

Analyse vectorielle

Béatrice de Tilière

13 janvier 2009

Table des matières

1	Espace euclidien \mathbb{E}^n	1
1.1	Espace vectoriel \mathbb{R}^n	1
1.2	Espace euclidien \mathbb{E}^n	3
1.2.1	Produit scalaire	4
1.2.2	Norme induite, distance	4
1.2.3	Angles	7
1.3	Limite et continuité des applications	8
1.4	Espace topologique \mathbb{E}^n	11
2	Les courbes paramétrées	15
2.1	Définitions et exemples	15
2.2	Dérivabilité d'une courbe	17
2.3	Longueur d'une courbe	18
2.3.1	Longueur d'une courbe	18
2.3.2	Courbes équivalentes	20
2.3.3	Paramétrisation par longueur d'arc	21
3	Fonctions de plusieurs variables	23
3.1	Définition et représentation	23
3.2	Dérivées partielles	24

3.3	Dérivabilité	25
3.4	Dérivée et gradient	27
3.5	Dérivation composées et interprétations géométriques du gradient	30
3.6	Dérivées d'ordre supérieur	33
3.7	Formule de Taylor	33
3.8	Maximums, minimums, points selles	37
3.8.1	Points critiques	38
3.8.2	Formes quadratiques	39
3.8.3	Classification des points critiques	40
3.8.4	Multiplicateurs de Lagrange	43
3.8.5	Maximums et minimums absolus	47
4	Applications de \mathbb{R}^m dans \mathbb{R}^n	49
4.1	Applications linéaires	49
4.2	Exemples	52
4.3	La matrice jacobienne	52
4.4	Dérivabilité	54
4.5	Dérivée et matrice jacobienne	55
4.6	Dérivation des applications composées	57
4.7	Théorèmes d'inversion locale et des applications implicites	59
4.7.1	Théorème d'inversion locale	59
4.7.2	Théorème des applications implicites	61
4.7.3	Lien entre les deux	62
5	Intégrales multiples	65
5.1	Intégration sur les pavés	65
5.2	Intégration sur des domaines plus généraux	69
5.3	Calculs explicites	71
5.4	Changement de variables	75

6	Intégrales curvilignes et Théorème de Green	79
6.1	Intégrale curviligne d'une fonction	79
6.2	Intégrale curviligne d'un champ de vecteurs	80
6.3	Théorème de Green	84
6.3.1	Le domaine d'intégration en question	84
6.3.2	Le Théorème de Green	85
7	Intégrales de surface - Théorèmes de la divergence et de Stokes	89
7.1	Surfaces paramétrées	89
7.1.1	Définition	89
7.1.2	Rappel sur le produit vectoriel	90
7.1.3	Dérivabilité et plan tangent	90
7.2	Intégrale de surface d'une fonction	92
7.3	Intégrale de surface d'un champ de vecteurs	95
7.4	Théorème de la divergence et Théorème de Stokes	97
7.4.1	Divergence et rotationnel	97
7.4.2	Les domaines d'intégration en question	99
7.4.3	Théorème de la divergence	100
7.4.4	Théorème de Stokes	103

Chapitre 1

Espace euclidien \mathbb{E}^n

Le but du cours d'analyse I est l'étude approfondie des fonctions de la droite réelle dans elle-même. Le cours d'analyse vectorielle s'inscrit dans la continuité de ce dernier, et a pour sujet l'étude des fonctions d'un espace euclidien dans un autre. Une partie des résultats et définitions seront des généralisations naturelles du cas réel, mais la structure plus riche de l'espace euclidien \mathbb{E}^n va également engendrer des phénomènes nouveaux et complexes.

Dans ce premier chapitre, nous définissons l'espace euclidien \mathbb{E}^n . Outre l'intérêt propre que porte une telle définition, il nous paraît important de passer un peu de temps à décrire l'espace qui sera l'ensemble de départ et d'arrivée des fonctions que nous allons étudier.

Résumé : l'espace euclidien \mathbb{E}^n est l'espace vectoriel \mathbb{R}^n muni du produit scalaire usuel. Le produit scalaire engendre les notions de *distance* et d'*angles*. L'espace euclidien possède aussi une *structure topologique*.

1.1 Espace vectoriel \mathbb{R}^n

L'espace \mathbb{R}^n est l'ensemble des n -tuples de nombres réels, c'est-à-dire :

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, \dots, x_n) \mid x_i \in \mathbb{R}\}.$$

Les points x_1, \dots, x_n sont appelés les *coordonnées* du point \mathbf{x} .

Notation

On utilisera les caractères gras, minuscules pour les points de \mathbb{R}^n , $n > 1$, et les lettres minuscules pour les points de \mathbb{R} .

Interprétation géométrique

On peut donner une interprétation géométrique de \mathbb{R} , \mathbb{R}^2 , \mathbb{R}^3 de la manière suivante. Supposons que l'on fixe une unité de longueur. Un point $x_1 \in \mathbb{R}$ peut être vu comme le point de coordonnée x_1 sur la droite réelle. Un point (x_1, x_2) (plus habituellement noté (x, y)) peut être vu comme le point de coordonnées (x_1, x_2) dans l'espace formé de deux axes perpendiculaires. De manière analogue un point (x_1, x_2, x_3) (souvent noté (x, y, z)) est le point de coordonnées (x_1, x_2, x_3) dans l'espace formé de trois axes mutuellement orthogonaux.

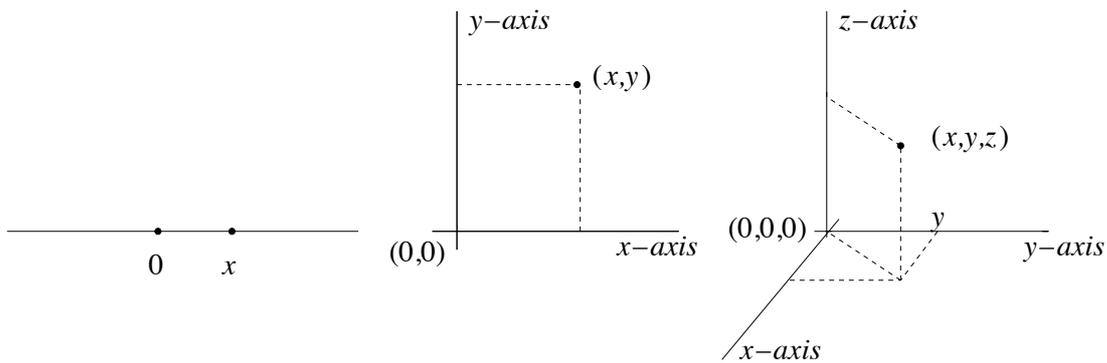


FIG. 1.1 – Interprétation géométrique des points de \mathbb{R} (gauche), \mathbb{R}^2 (milieu), \mathbb{R}^3 (droite).

Ces cas particuliers nous permettent d'imaginer une interprétation géométrique de \mathbb{R}^n , même si la visualisation devient plus difficile lorsque $n \geq 4$.

Espace vectoriel

On munit \mathbb{R}^n de deux opérations : l'addition et la multiplication par un scalaire. Si $\mathbf{x} = (x_1, \dots, x_n)$ et $\mathbf{y} = (y_1, \dots, y_n)$ sont deux points de \mathbb{R}^n , on définit la *somme de \mathbf{x} et \mathbf{y}* , notée $\mathbf{x} + \mathbf{y}$, par :

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n).$$

De plus, si a est un nombre réel, on définit la *multiplication par un scalaire de \mathbf{x} par a* , notée $a\mathbf{x}$ par :

$$a\mathbf{x} = (ax_1, \dots, ax_n).$$

On utilise la notation $\mathbf{0} = (0, \dots, 0)$, et $-\mathbf{x} = (-1)\mathbf{x} = (-x_1, \dots, -x_n)$.

Exemple 1.1 Soit $\mathbf{x} = (1, -2, 1.5)$ $\mathbf{y} = (\pi, 2, -\sqrt{7})$, $a = -3$, alors

$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (1 + \pi, 0, 1.5 - \sqrt{7}), \\ a\mathbf{x} &= (-3, 6, -4.5), \\ a\mathbf{y} &= (-3\pi, -6, 3\sqrt{7}).\end{aligned}$$

L'ensemble \mathbb{R}^n muni des opérations somme et multiplication par un scalaire satisfait les propriétés suivantes :

1. $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ (associativité).
2. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (commutativité).
3. $\mathbf{x} + \mathbf{0} = \mathbf{x}$ (existence identité pour addition).
4. $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ (existence inverse pour addition).
5. $(ab)\mathbf{x} = a(b\mathbf{x})$ (associativité).
6. $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$ (distributivité).
7. $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$ (distributivité).
8. $1\mathbf{x} = \mathbf{x}$ (existence identité pour multiplication par scalaire).

Ces propriétés sont des conséquences immédiates des propriétés des nombres réels \mathbb{R} . Elles confèrent à \mathbb{R}^n une structure d'*espace vectoriel*.

Rappel : un espace V muni d'une addition et d'une multiplication par un scalaire qui vérifie les conditions 1 à 8 est appelé un *espace vectoriel*.

L'ensemble des vecteurs

$$\{\mathbf{e}_1 = (1, 0, \dots, 0), \mathbf{e}_2 = (0, 1, 0, \dots, 0), \dots, \mathbf{e}_n = (0, \dots, 0, 1)\},$$

forment une base de \mathbb{R}^n , appelée *base canonique*. Un point $\mathbf{x} = (x_1, \dots, x_n)$ de \mathbb{R}^n s'écrit $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ dans la base canonique.

Un sous-ensemble V de \mathbb{R}^n est un *sous-espace vectoriel* s'il est contenu dans \mathbb{R}^n et est lui-même un espace vectoriel (en particulier il doit contenir $\mathbf{0}$ en vertu de la propriété 3).

Exemple 1.2

– Soient $\mathbf{x}_1, \dots, \mathbf{x}_k$, k points de \mathbb{R}^n . Alors l'ensemble

$$\{(a_1 \mathbf{x}_1 + \dots + a_k \mathbf{x}_k) \mid a_1, \dots, a_k \in \mathbb{R}\},$$

est un sous-espace vectoriel de \mathbb{R}^n , et s'appelle l'*espace engendré par $\mathbf{x}_1, \dots, \mathbf{x}_k$* .

– Les ensembles suivants sont des sous-espaces vectoriels de \mathbb{R}^2 :

$$V_1 = \{(x, y) \in \mathbb{R}^2 \mid y = 2x\}.$$

$$V_2 = \{a(1, 0) \mid a \in \mathbb{R}\}.$$

Remarquez que V_2 peut aussi s'écrire $V_2 = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$.

1.2 Espace euclidien \mathbb{E}^n

Afin d'obtenir la structure géométrique complète de \mathbb{R}^n , nous introduisons le produit scalaire usuel. Comme nous allons le voir, il définit de manière naturelle une notion de distance, une norme et les angles entre les points de \mathbb{R}^n . L'espace vectoriel \mathbb{R}^n muni de cette structure supplémentaire est appelé l'*espace euclidien*, noté \mathbb{E}^n .

1.2.1 Produit scalaire

Soit $\mathbf{x} = (x_1, \dots, x_n)$ et $\mathbf{y} = (y_1, \dots, y_n)$ deux points de \mathbb{R}^n . Le *produit scalaire usuel* de \mathbf{x} et \mathbf{y} , noté $\mathbf{x} \cdot \mathbf{y}$, ou encore $\langle \mathbf{x}, \mathbf{y} \rangle$ est défini de la manière suivante :

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \dots + x_n y_n.$$

Exemple 1.3 Soient $\mathbf{x} = (1, 0, 1)$, $\mathbf{y} = (0, 1, 0)$ $\mathbf{z} = (2, 1, 4)$, alors :

$$\mathbf{x} \cdot \mathbf{y} = 0, \quad \mathbf{x} \cdot \mathbf{z} = 6, \quad \mathbf{y} \cdot \mathbf{z} = 1.$$

Le produit scalaire usuel satisfait les propriétés suivantes :

1. $\mathbf{x} \cdot \mathbf{x} \geq 0$, et $\mathbf{x} \cdot \mathbf{x} = 0$ ssi $\mathbf{x} = \mathbf{0}$.
2. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.
3. $(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a\mathbf{x} \cdot \mathbf{z} + b\mathbf{y} \cdot \mathbf{z}$.

Ces propriétés découlent de celles de \mathbb{R} , et seront vérifiées en exercices.

Deux vecteurs \mathbf{x}, \mathbf{y} de \mathbb{R}^n sont dits *orthogonaux* si $\mathbf{x} \cdot \mathbf{y} = 0$.

Rappel : Soit un espace vectoriel V muni d'une opération \cdot de $V \times V$ dans \mathbb{R} . Si l'opération \cdot satisfait les propriétés 1 à 3, elle est appelée un *produit scalaire*.

1.2.2 Norme induite, distance

Norme induite

Le produit scalaire usuel induit la *norme euclidienne* sur \mathbb{R}^n : soit \mathbf{x} un point de \mathbb{R}^n , alors la norme euclidienne de \mathbf{x} , notée $\|\mathbf{x}\|$ est définie par,

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Dans le cas où $n = 2$ ou $n = 3$, cette définition coïncide avec notre notion intuitive de norme (dérivée du Théorème de Pythagore), voir les Figures 1.2 et 1.3 ci-dessous.

Exemple 1.4 Soit $\mathbf{x} = (2, 3)$, alors $\|\mathbf{x}\| = \sqrt{13}$. Soit $\mathbf{x} = (3, 1, 2)$, alors $\|\mathbf{x}\| = \sqrt{14}$.

La norme euclidienne satisfait les propriétés suivantes :

1. $\|\mathbf{x}\| \geq 0$, et $\|\mathbf{x}\| = 0$ ssi $\mathbf{x} = \mathbf{0}$.
2. $\|a\mathbf{x}\| = |a| \|\mathbf{x}\|$.
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (inégalité triangulaire).

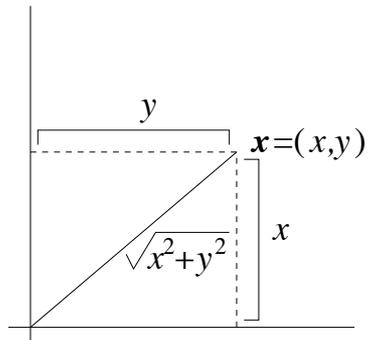


FIG. 1.2 – Théorème de Pythagore et norme, $n = 2$.

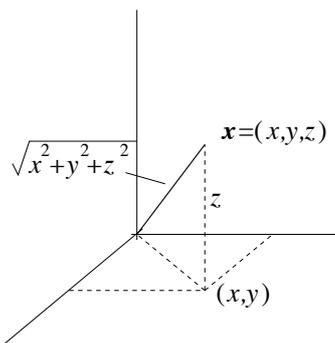


FIG. 1.3 – Théorème de Pythagore et norme, $n = 3$.

Les deux premières propriétés découlent de la définition de la norme et des propriétés du produit scalaire. L'inégalité triangulaire découle de l'inégalité de Cauchy-Schwarz, que nous énonçons et prouvons ci-dessous, étant donné sa grande importance.

Théorème 1 [Inégalité de Cauchy-Schwarz]

Si V est un espace vectoriel, \cdot est un produit scalaire sur V , et $\|\cdot\|$ est la norme correspondante, alors :

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

En particulier, lorsque $V = \mathbb{R}^n$, \cdot est le produit scalaire usuel, et $\|\cdot\|$ est la norme euclidienne, l'inégalité de Cauchy-Schwarz s'écrit :

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n y_i^2 \right)^{\frac{1}{2}}.$$

Preuve:

Si \mathbf{x} ou \mathbf{y} est $\mathbf{0}$, l'inégalité est triviale, donc supposons que les deux soient non nuls. Posons

$$\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|, \quad \mathbf{v} = \mathbf{y}/\|\mathbf{y}\|,$$

alors $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. Donc,

$$0 \leq \|\mathbf{u} - \mathbf{v}\|^2 = (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) = \|\mathbf{u}\|^2 - 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{v}\|^2 = 2 - 2\mathbf{u} \cdot \mathbf{v}.$$

Ainsi, $\mathbf{u} \cdot \mathbf{v} \leq 1$, ce qui est équivalent à $\frac{\mathbf{x}}{\|\mathbf{x}\|} \cdot \frac{\mathbf{y}}{\|\mathbf{y}\|} \leq 1$, ou encore

$$\mathbf{x} \cdot \mathbf{y} \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

En remplaçant \mathbf{x} par $-\mathbf{x}$ on obtient $-\mathbf{x} \cdot \mathbf{y} \leq \|\mathbf{x}\| \|\mathbf{y}\|$, d'où le résultat. \square

Preuve (inégalité triangulaire) :

Soit $\mathbf{x}, \mathbf{y} \in V$, alors

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \\ &= \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2, \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 \text{ (Cauchy-Schwarz) }, \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

\square

Remarquez que si \mathbf{x} et \mathbf{y} sont orthogonaux, l'inégalité triangulaire devient une égalité :

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

Ceci est le *Théorème de Pythagore* dans le cas général de \mathbb{E}^n .

Rappel : Soit V un espace vectoriel et $\|\cdot\|$ une application de V dans \mathbb{R} satisfaisant les propriétés 1 à 3. Alors $\|\cdot\|$ s'appelle une *norme*, et l'espace V muni de l'application $\|\cdot\|$ un *espace vectoriel normé*.

Distance

La norme euclidienne induit la notion de distance euclidienne entre deux points de \mathbb{R}^n . Si \mathbf{x}, \mathbf{y} sont deux points de \mathbb{R}^n , alors la *distance euclidienne* de \mathbf{x} à \mathbf{y} , notée $d(\mathbf{x}, \mathbf{y})$ est définie par :

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

La distance euclidienne satisfait les trois propriétés suivantes :

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$, $d(\mathbf{x}, \mathbf{y}) = 0$ ssi $\mathbf{x} = \mathbf{y}$.
2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$.
3. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (inégalité triangulaire).

Ces propriétés découlent de celles du produit scalaire.

Rappel : Soit V un espace vectoriel muni d'une application d de $V \times V$ dans \mathbb{R} . Si d satisfait les propriétés 1 à 3, alors d est appelée une *distance*, et V muni de d est appelé un *espace métrique*.

1.2.3 Angles

Le produit scalaire usuel et la norme nous permettent de définir la notion d'angle entre deux vecteurs. Remarquez que la définition est une généralisation naturelle du cas \mathbb{R}^2 . Si \mathbf{x}, \mathbf{y} sont deux vecteurs de \mathbb{R}^n , on définit l'*angle* entre \mathbf{x} et \mathbf{y} , noté $\angle(\mathbf{x}, \mathbf{y})$ par :

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \in [0, \pi].$$

Remarquez que l'angle est bien défini en vertu de l'inégalité de Cauchy-Schwarz qui garantit que le terme de droite est dans $[-1, 1]$.

Exemple 1.5 Soit $\mathbf{x} = (0, 1, 1)$, et $\mathbf{y} = (1, 1, 0)$, alors

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{1}{\sqrt{2}\sqrt{2}} = \arccos \frac{1}{2} = \frac{\pi}{3}.$$

On peut aussi interpréter cet angle géométriquement comme sur la figure ci-dessous.

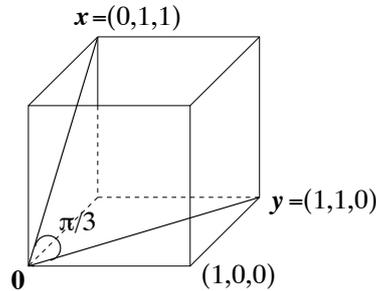


FIG. 1.4 – Interprétation géométrique de l'angle.

1.3 Limite et continuité des applications

Le but de cette section est de généraliser en dimensions supérieures la notion de limite et continuité des fonctions de \mathbb{R} dans \mathbb{R} .

Soit D un sous-ensemble de \mathbb{R}^m . Une *application* F de D dans \mathbb{R}^n , notée $F : D \rightarrow \mathbb{R}^n$, est une règle qui à tout point \mathbf{x} de D associe un point $F(\mathbf{x})$ dans \mathbb{R}^n . On peut représenter F par ses coordonnées, c'est-à-dire,

$$F(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x})),$$

où $F_i : \mathbb{R}^m \rightarrow \mathbb{R}$. On utilisera la terminologie *fonction* lorsque l'espace d'arrivée est \mathbb{R} , c'est-à-dire lorsque $n = 1$.

Dans les chapitres suivants, nous passerons du temps sur les cas particuliers $m = 1$, puis $n = 1$, de sorte que nous ne faisons pas maintenant un catalogue des différents genres d'applications de \mathbb{R}^m dans \mathbb{R}^n . Nous donnerons simplement quelques exemples pour illustrer les définitions de limite et de continuité.

Dans cette section, nous traitons directement le cas m et n entiers positifs quelconques. En effet, pour les notions de limite et continuité, la complexité supplémentaire par rapport au cas des fonctions réelles est minimale : il suffit de remplacer la valeur absolue par la norme euclidienne.

Limite d'une application

Avant de donner la définition de limite d'une application, nous devons encore mentionner le point technique suivant. On souhaite donner la définition de "l'application F admet \mathbf{b} comme limite au point \mathbf{a} ". On souhaite néanmoins autoriser que l'application F ne soit pas définie au point \mathbf{a} , mais seulement aux points qui sont très proches de \mathbf{a} . Afin de définir ceci de manière rigoureuse, on a besoin du concept fondamental suivant. Soit \mathbf{a} un point de \mathbb{R}^m , on appelle *boule ouverte de rayon r et de centre \mathbf{a}* , notée $B_r(\mathbf{a})$,

l'ensemble suivant :

$$B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n \mid d(\mathbf{a}, \mathbf{x}) < r\} = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \sqrt{\sum_{i=1}^n (x_i - a_i)^2} < r \right\}.$$

On dit qu'un point \mathbf{a} est un *point limite* de l'ensemble D , ssi toute boule ouverte centrée en \mathbf{a} contient des points de D autres que \mathbf{a} .

Exemple 1.6

1. Un ensemble fini de points n'a pas de point limite.
2. L'origine $\mathbf{0}$ est un point limite de $\mathbb{R}^m \setminus \{\mathbf{0}\}$.

Voici maintenant la définition de limite pour une application. Soit $F : D \rightarrow \mathbb{R}^n$ une application, soient \mathbf{a} un point limite de D , et \mathbf{b} un point de \mathbb{R}^n . On dit que F *admet* \mathbf{b} *comme limite au point* \mathbf{a} , écrit :

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} F(\mathbf{x}) = \mathbf{b},$$

ssi pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $\mathbf{x} \in B_\delta(\mathbf{a}) \cap D$, alors $F(\mathbf{x}) \in B_\varepsilon(\mathbf{b})$.

L'idée géométrique derrière la définition est que la valeur $F(\mathbf{x})$ peut être aussi proche de \mathbf{b} que l'on veut, si l'on prend \mathbf{x} dans une boule centrée en \mathbf{a} suffisamment petite.

Exemple 1.7 On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, définie par :

$$f(x, y) = x^2 + xy + y.$$

Montrons que $\lim_{\mathbf{x} \rightarrow (1,1)} f(x, y) = 3$. Dans la suite, nous utilisons plusieurs fois l'inégalité triangulaire.

$$\begin{aligned} |f(x, y) - 3| &= |x^2 + xy + y - 3| \leq |x^2 - 1| + |xy - 1| + |y - 1|, \\ &\leq |x + 1||x - 1| + |xy - y + y - 1| + |y - 1|, \\ &\leq (|x + 1| + |y|)|x - 1| + 2|y - 1|. \end{aligned}$$

D'autre part,

$$\|(x, y) - (1, 1)\| < \delta \Leftrightarrow (x - 1)^2 + (y - 1)^2 < \delta^2 \Rightarrow |x - 1| < \delta \text{ et } |y - 1| < \delta.$$

Ceci implique en particulier

$$\begin{aligned} |x + 1| &\leq |x - 1| + 2 < \delta + 2, \\ |y| &\leq |y - 1| + 1 < \delta + 1. \end{aligned}$$

En résumé, $\|(x, y) - (1, 1)\| < \delta$ implique $|f(x, y) - 3| < (2\delta + 5)\delta$. Ainsi, il suffit de prendre $\delta = \min\left(1, \frac{\varepsilon}{7}\right)$.

Remarque 1.1 Voici une conséquence importante de la définition. Si une application $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ admet une limite au point \mathbf{a} , alors la limite est la même quelque soit le chemin d'accès à \mathbf{a} . En particulier, la limite est la même si l'on accède à \mathbf{a} par des droites différentes. En conséquence, si l'on souhaite montrer qu'une droite n'admet pas de limite en un point \mathbf{a} , il suffit de trouver deux droites d'accès à \mathbf{a} , tel que l'application n'admette pas la même limite selon que l'on prenne une droite ou l'autre. Voici un exemple.

Exemple 1.8 Montrons que la fonction suivante n'admet pas de limite en $\mathbf{0} = (0, 0)$. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

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

Etudions les valeurs de $f(x, y)$ lorsque (x, y) approche $(0, 0)$ le long de la droite $(x, \alpha x)$, $\alpha \in \mathbb{R}$. Si $\alpha = \pm 1$, alors :

$$f(x, \pm x) = 0.$$

Si $\alpha \neq \pm 1$, alors

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

Ainsi f a différentes valeurs constantes sur différentes droites qui passent par $\mathbf{0}$, et donc ne peut avoir de limite en $\mathbf{0}$.

Le théorème suivant est utile pour étudier les limites des applications F de \mathbb{R}^m dans \mathbb{R}^n . Il sera démontré pendant les exercices.

Théorème 2 Soit une application $F : D \rightarrow \mathbb{R}^n$, où D est un sous-ensemble de \mathbb{R}^m , et soit \mathbf{a} un point limite de D . Alors, F admet $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$ comme limite au point \mathbf{a} , ssi pour tout $i = 1, \dots, n$, la fonction coordonnée $F_i : D \rightarrow \mathbb{R}$ admet b_i comme point limite au point \mathbf{a} :

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} F_i(\mathbf{x}) = b_i.$$

Continuité des applications

Soit $F : D \rightarrow \mathbb{R}^n$ une application, où D est un sous-ensemble de \mathbb{R}^m . On dit que F est continue au point $\mathbf{a} \in D$, ssi

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} F(\mathbf{x}) = F(\mathbf{a}).$$

On dit que F est continue sur D , si elle est continue en tout point de D .

Remarque 1.2 Si \mathbf{a} n'est pas un point limite de D , la définition que nous avons donnée de limite ne fait pas de sens. Pour lui en donner un, nous disons que F est toujours continue en de tels points.

Corollaire 3 Une application $F : D \rightarrow \mathbb{R}^n$, où D est un sous-ensemble de \mathbb{R}^m est continue au point \mathbf{a} , ssi toutes ses coordonnées sont continues au point \mathbf{a} .

Pour montrer la continuité des applications, on ne recourt que rarement à la définition, qui est relativement lourde à utiliser comme nous l'avons vu dans l'Exemple 1.7. On utilisera volontiers le Corollaire 3 ci-dessus, la continuité déjà établie des fonctions dans \mathbb{R} , et les théorèmes ci-dessous.

Soit les applications $F : D_1 \rightarrow \mathbb{R}^n$ et $G : D_2 \rightarrow \mathbb{R}^k$, où D_1 est un sous-ensemble de \mathbb{R}^m et D_2 est un sous-ensemble de \mathbb{R}^n tels que $F(D_1) \subset D_2$. On définit alors l'application composée : $G \circ F : D_1 \rightarrow \mathbb{R}^k$ par

$$G \circ F(\mathbf{x}) = G(F(\mathbf{x})).$$

Théorème 4 Si F est continue au point $\mathbf{a} \in D_1$ et G est continue au point $F(\mathbf{a})$, alors $G \circ F$ est continue au point \mathbf{a} .

Des résultats ci-dessus, on déduit,

Théorème 5 Soient f, g des fonctions de \mathbb{R}^m dans \mathbb{R} . Alors,

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

et

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} (f(\mathbf{x})g(\mathbf{x})) = \left(\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) \right) \left(\lim_{\mathbf{x} \rightarrow \mathbf{a}} g(\mathbf{x}) \right),$$

sous la condition que ces limites existent.

Exemple 1.9 Soit l'application $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ définie par :

$$F(x, y, z) = (\sin(x + y) \cos(y) + ze^y, \cos(z)).$$

Pour montrer que l'application F est continue sur tout \mathbb{R}^3 , par le Corollaire 3, il suffit de montrer que chacune de ses coordonnées est continue sur tout \mathbb{R}^3 . Or chacune des composantes est somme/produit/composition de fonctions usuelles qui sont continues sur les ensembles considérés, donc est continue.

1.4 Espace topologique \mathbb{E}^n

Dans cette section, nous introduisons la topologie naturelle sur \mathbb{E}^n induite par la distance euclidienne. En effet c'est le formalisme adéquat pour introduire certaines des

définitions et concept d'analyse vectorielle. Nous utiliserons essentiellement les notions d'ouverts, fermés, intérieur, bord, et continuité.

Dans la section précédente, nous avons déjà parlé de continuité des applications sans parler de topologie. En effet, nous voulions introduire la continuité d'une application après celle de limite d'une application, et pour ceci le formalisme topologique n'est pas le plus adapté. Mais, nous allons maintenant donner la définition en topologie d'une application continue, et vous montrerez aux exercices l'équivalence entre les deux.

Dernier avertissement avant de rentrer dans le coeur de ce chapitre. Nous travaillons dans le cas particulier de l'espace euclidien, et de la topologie induite par la distance euclidienne, mais il est aisé de voir que tout espace métrique a une topologie naturelle définie par sa distance.

Les concepts introduits ici seront traités de manière approfondie lors du cours de topologie, ainsi nous énonçons seulement quelques uns des théorèmes, et ne les démontrons pas.

Espace topologique \mathbb{E}^n

Un sous-ensemble U de \mathbb{R}^n est dit *ouvert*, ssi pour tout $\mathbf{x} \in U$, il existe $r > 0$ et une boule ouverte de rayon r , centrée en \mathbf{x} , qui est contenue dans U . On note \mathcal{T} l'ensemble des ouverts de \mathbb{R}^n .

Exemple 1.10 L'ensemble

$$B_1(\mathbf{0}) = \{(x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} < 1\},$$

est un ouvert de \mathbb{R}^2 .

L'ensemble \mathcal{T} vérifie les propriétés suivantes.

1. L'ensemble \emptyset et \mathbb{R}^n sont ouverts.
2. Toute réunion d'ouverts est un ouvert.
3. L'intersection de deux ouverts est un ouvert.

Ces propriétés seront vérifiées en exercices.

Rappel : un espace X muni d'un ensemble \mathcal{T} de parties de X qui satisfait les propriétés 1 à 3 est appelé un *espace topologique*. Les parties de \mathcal{T} sont appelées les *ouverts*.

Quelques notions sur l'espace topologique \mathbb{E}^n ...

Un sous-ensemble F de \mathbb{R}^n est dit *fermé* si son complémentaire est ouvert.

Soit S un sous-ensemble de \mathbb{R}^n . Un point \mathbf{x} de S est appelé un *point intérieur* de S , ssi il existe une boule ouverte centrée en \mathbf{x} de rayon strictement positif contenue dans

S . L'ensemble des points intérieurs de S s'appelle l'*intérieur* de S , et est noté $\text{int}(S)$. Remarquer que $\text{int}(S)$ est un ouvert de \mathbb{R}^n .

Un point \mathbf{x} de \mathbb{R}^n est appelé un point frontière de S , ssi toute boule ouverte centrée en \mathbf{x} contient un point de S et un point qui n'est pas dans S . L'ensemble des points frontières de S est appelé la *frontière* de S , et est noté ∂S . On peut montrer que ∂S est un fermé de \mathbb{R}^n .

Un ensemble V est appelé un *voisinage* d'un point \mathbf{x} , s'il contient un ouvert contenant \mathbf{x} .

Un sous-ensemble K de \mathbb{R}^n est dit *compact* ssi toute suite admet une sous-suite qui est convergente. Le théorème suivant donne une caractérisation utile des compacts.

Théorème 6 *Un sous-ensemble K de \mathbb{R}^n est compact, ssi il est fermé et borné.*

Théorème 7 *Soit $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ une application. Alors l'application F est continue sur \mathbb{R}^m , ssi pour tout ouvert $U \subset \mathbb{R}^n$, l'image inverse $F^{-1}(U)$ est un ouvert de \mathbb{R}^m .*

Ce théorème sera démontré en séance d'exercices, en effet il est important que l'étudiant comprenne le lien entre les deux approches de continuité.

Théorème 8 *Soit K un sous-ensemble compact de \mathbb{R}^m , et $F : K \rightarrow \mathbb{R}^n$ une application continue. Alors $F(K)$ est un sous-ensemble compact de \mathbb{R}^n .*

Théorème 9 *Soit K un sous-ensemble compact de \mathbb{R}^m , et $F : K \rightarrow \mathbb{R}$ une application continue. Alors F atteint un maximum et un minimum en des points de K .*

Théorème 10 *Soit K un sous-ensemble compact de \mathbb{R}^n , et $F : K \rightarrow \mathbb{R}$ une fonction continue. Alors F est uniformément continue.*

Chapitre 2

Les courbes paramétrées

Le deuxième chapitre a pour sujet les courbes paramétrées. C'est un moyen doux de commencer à travailler dans \mathbb{R}^n . En effet, il est d'habitude possible de s'aider d'une interprétation géométrique, et la théorie des fonctions réelles s'applique souvent assez directement.

Par ailleurs, les courbes font partie des mathématiques depuis l'antiquité. Les premiers à les étudier sont sans doute les grecs, elles eurent ensuite beaucoup d'importance dans la physique de Newton. Leur étude a été formalisée et approfondie avec l'apparition de la géométrie différentielle d'une part et de la géométrie algébrique de l'autre. Une famille de courbes amusantes récemment découvertes sont les courbes de Peano, qui remplissent tout le plan.

Mots clés : courbes paramétrées, vecteur vitesse, accélération, longueur, courbes équivalentes, paramétrisation par longueur d'arc.

2.1 Définitions et exemples

Soit $I = [a, b]$ un intervalle de \mathbb{R} . Une *courbe paramétrée* γ est une application continue $\gamma : I \rightarrow \mathbb{R}^n$. En coordonnées, γ s'écrit :

$$\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t)).$$

L'ensemble $\gamma(I)$ s'appelle la *trace de la courbe* γ .

Une courbe est dite *simple*, ssi l'application γ est injective sur (a, b) . Une courbe est dite *fermée*, ssi $\gamma(a) = \gamma(b)$. Une courbe fermée et simple est appelée une *courbe de Jordan*.

Exemple 2.1 On considère le segment reliant deux points \mathbf{a}, \mathbf{b} de \mathbb{R}^n . Une paramétrisation du segment est donnée par :

$$\begin{aligned} \gamma : [0, 1] &\rightarrow \mathbb{R}^n \\ t &\mapsto \gamma(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) = (a_1 + t(b_1 - a_1), \dots, a_n + t(b_n - a_n)). \end{aligned}$$

Exemple 2.2 La courbe paramétrée suivante représente une hélice dans \mathbb{R}^3 :

$$\begin{aligned} \gamma : \mathbb{R} &\rightarrow \mathbb{R}^3 \\ t &\mapsto \gamma(t) = (\cos t, \sin t, t). \end{aligned}$$

Une courbe de \mathbb{R}^2 est dite *paramétrée en coordonnées polaires*, si elle est de la forme

$$\begin{aligned} \gamma : I &\rightarrow \mathbb{R}^2 \\ \theta &\mapsto \gamma(\theta) = (r(\theta) \cos \theta, r(\theta) \sin \theta), \end{aligned}$$

où r est une fonction continue, positive (représentant le rayon).

Exemple 2.3 Dans cet exemple, nous allons construire les coniques comme des courbes paramétrées en coordonnées polaires.

Une *conique* est par définition l'intersection d'un cône de révolution et d'un plan. Cette intersection est classée en différents types : ellipse, hyperbole, parabole etc...

Il existe une manière équivalente de définir les coniques, qui est la suivante. Soit D une droite du plan, \mathbf{f} un point qui n'est pas situé sur D , et $\varepsilon > 0$. Alors, on appelle conique de droite directrice D , de foyer \mathbf{f} et d'excentricité ε , l'ensemble des points \mathbf{x} du plan vérifiant :

$$d(\mathbf{f}, \mathbf{x}) = \varepsilon d(D, \mathbf{x}), \quad (2.1)$$

où $d(\mathbf{f}, \mathbf{x})$ est la distance du point \mathbf{x} au point \mathbf{f} , et $d(D, \mathbf{x})$ est la distance du point \mathbf{x} à la droite D .

On considère le repère de centre \mathbf{f} , d'axe des abscisses perpendiculaire à la droite D , et d'axe des ordonnées parallèle à D . On écrit un point (x, y) dans ce repère en coordonnées polaires, c'est à dire $(x, y) = (r \cos \theta, r \sin \theta)$. Alors l'équation (2.1) est équivalente à

$$r = \varepsilon(m + r \cos \theta),$$

où m est la distance du foyer \mathbf{f} à la droite D . Cette équation a une solution positive $r = r(\theta) = \frac{\varepsilon m}{1 - \varepsilon \cos \theta}$, ssi $\theta \in I_\varepsilon = \{\theta \in [0, 2\pi) \mid 1 - \varepsilon \cos \theta > 0\}$. On vérifie aussi que sur I_ε , r est continue.

Donc les coniques s'écrivent comme des courbes paramétrées en coordonnées polaires sous la forme :

$$\begin{aligned} \gamma : I_\varepsilon &\rightarrow \mathbb{R}^2 \\ \theta &\mapsto \gamma(\theta) = \left(\frac{k \cos \theta}{1 - \varepsilon \cos \theta}, \frac{k \sin \theta}{1 - \varepsilon \cos \theta} \right). \end{aligned}$$

A partir, de ces équations en coordonnées polaires, on peut retrouver les équations algébriques et la classification connue, voir les exercices.

2.2 Dérivabilité d'une courbe

L'intervalle I sur lequel est définie une courbe peut être ouvert, fermé, ou ouvert à gauche (droite) / fermé à droite (gauche). Pour définir la dérivabilité, nous supposons que l'intervalle I contient plus d'un point. De plus, lorsque l'on considère la limite ci-dessous en un point extrême de I et que I est fermé en ce point, on considèrera uniquement la limite à gauche, ou à droite selon le cas, de sorte que $t+h$ soit également dans I .

Soit $\gamma : I \rightarrow \mathbb{R}^n$ une courbe, et soit $t \in I$. Alors, la courbe γ est dite *dérivable* au point t si la limite suivante existe :

$$\lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

Dans ce cas, on appelle cette limite la *dérivée au point t* et on la note $\gamma'(t)$, $\frac{d\gamma}{dt}$, ou encore $\dot{\gamma}(t)$. Le vecteur $\gamma'(t)$ est aussi connu sous le nom de *vecteur vitesse au point t* .

La courbe est dite *dérivable sur I* si elle est dérivable en tout point de I . En utilisant le Théorème 2, on déduit que la courbe γ est dérivable, ssi toutes ses coordonnées le sont, et dans ce cas :

$$\gamma'(t) = (\gamma'_1(t), \dots, \gamma'_n(t)).$$

En utilisant les propriétés des dérivées des fonctions réelles pour chacune des coordonnées, on obtient le théorème suivant :

Théorème 11 Soient $\gamma_1, \gamma_2 : \mathbb{R} \rightarrow \mathbb{R}^n$ deux courbes, et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction, toutes les trois dérivables, alors :

1. $(\gamma_1 + \gamma_2)' = \gamma'_1 + \gamma'_2$.
2. $(\varphi \gamma_1)' = \varphi' \gamma_1 + \varphi \gamma'_1$.
3. $(\gamma_1 \cdot \gamma_2)' = \gamma'_1 \cdot \gamma_2 + \gamma_1 \cdot \gamma'_2$.
4. $(\gamma \circ \varphi)'(t) = \varphi'(t) \gamma'(\varphi(t))$.

Une courbe γ est dite de *classe \mathcal{C}^1* sur I , si elle est dérivable et que sa dérivée est continue en tout point de I . Une courbe de classe \mathcal{C}^1 est dite *régulière*, si $\gamma'(t) \neq \mathbf{0}$ pour tout $t \in I$.

Lorsque la courbe γ est régulière, on peut visualiser $\gamma'(t)$ comme étant le vecteur tangent à la courbe au point $\gamma(t)$, voir Figure 2.1. Dans ce cas, on définit *la tangente* à γ au point $\gamma(t)$ comme étant la droite qui passe par $\gamma(t)$, et qui est parallèle au vecteur $\gamma'(t)$.

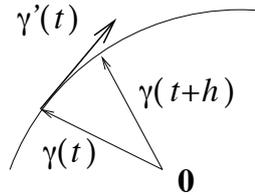


FIG. 2.1 – Le vecteur tangent

La *vitesse au point t* , notée $v(t)$, d'une courbe γ supposée dérivable, est la norme du vecteur vitesse au point t :

$$v(t) = \|\gamma'(t)\|.$$

Le *vecteur d'accélération au point t* d'une courbe γ deux fois dérivable, est le vecteur $\gamma''(t)$. L'*accélération au point t* , notée $a(t)$, est la norme du vecteur d'accélération au point t :

$$a(t) = \|\gamma''(t)\|.$$

2.3 Longueur d'une courbe

2.3.1 Longueur d'une courbe

Soit $\gamma : I = [a, b] \rightarrow \mathbb{R}^n$ une courbe paramétrée de classe \mathcal{C}^1 . La *longueur* de γ notée $\ell(\gamma)$, est par définition l'intégrale de la vitesse entre a et b , i.e.

$$\ell(\gamma) = \int_a^b v(t) dt = \int_a^b \|\gamma'(t)\| dt.$$

En utilisant la définition de la norme euclidienne, cette définition peut se réécrire ainsi,

$$\ell(\gamma) = \int_a^b \sqrt{(\gamma'_1(t))^2 + \dots + (\gamma'_n(t))^2} dt. \quad (2.2)$$

Remarque 2.1 *Justification de la définition de la longueur.*

Cette définition peut être comprise intuitivement de la manière suivante. Dans l'esprit de l'intégrale de Riemann, il est naturel de définir la longueur de la courbe en

considérant une partition de l'intervalle $[a, b]$, de sommer les "petits bouts" de longueur, et de prendre la limite lorsque la partition devient de plus en plus petite. Soit $\mathcal{P} = \{a = t_0 < t_1 < \dots < t_k = b\}$ une partition de l'intervalle $[a, b]$. Alors, une approximation de la longueur de la courbe correspondant à cette partition est :

$$\begin{aligned} s(\gamma, \mathcal{P}) &= \sum_{i=1}^k \|\gamma(t_i) - \gamma(t_{i-1})\| \\ &= \sum_{i=1}^k \left(\sum_{\ell=1}^n (\gamma_\ell(t_i) - \gamma_\ell(t_{i-1}))^2 \right)^{\frac{1}{2}} \quad (\text{où } \gamma_1, \dots, \gamma_n \text{ sont les coordonnées de } \gamma). \end{aligned}$$

En utilisant le théorème de la valeur intermédiaire pour chacune des composantes, on sait que pour tout ℓ , il existe t_i^ℓ dans l'intervalle (t_{i-1}, t_i) tel que :

$$\gamma_\ell(t_i) - \gamma_\ell(t_{i-1}) = \gamma'_\ell(t_i^\ell)(t_i - t_{i-1}).$$

Ainsi,

$$s(\gamma, \mathcal{P}) = \sum_{i=1}^k \left(\sum_{\ell=1}^n \gamma'^2_\ell(t_i^\ell) \right)^{\frac{1}{2}} (t_i - t_{i-1}). \quad (2.3)$$

Nous ne justifions pas l'étape suivante, mais elle peut l'être rigoureusement. On suppose que pour tous les ℓ , le point t_i^ℓ est un même point t_i^* dans l'intervalle (t_{i-1}, t_i) , alors la somme peut s'écrire :

$$s(\gamma, \mathcal{P}) = \sum_{i=1}^k \|\gamma'(t_i^*)\| (t_i - t_{i-1}),$$

qui est une somme de Riemann pour l'intégrale (2.2). Ainsi,

$$\ell(\gamma) = \lim_{|\mathcal{P}| \rightarrow 0} s(\gamma, \mathcal{P}).$$

Ça marche pas, ça reviendrait à dire que l'égalité des acc. finis vectorielle est valide, ce qui est faux. Voir "Étude métrique des courbes planes"

Exemple 2.4 On considère une courbe $\gamma : I = [a, b] \rightarrow \mathbb{R}^2$ paramétrée en coordonnées polaires :

$$\gamma(\theta) = (r(\theta) \cos \theta, r(\theta) \sin \theta).$$

Alors, $\gamma'(\theta) = (r'(\theta) \cos \theta - r(\theta) \sin \theta, r'(\theta) \sin \theta + r(\theta) \cos \theta)$, et

$$\|\gamma'(\theta)\| = \sqrt{r'(\theta)^2 + r(\theta)^2}.$$

Ainsi,

$$\ell(\gamma) = \int_a^b \sqrt{r'(\theta)^2 + r(\theta)^2} d\theta.$$

En particulier, si on considère le cercle de rayon r paramétré en coordonnées polaire, on a $I = [0, 2\pi]$, $r(\theta) \equiv r$, de sorte que :

$$\ell(\gamma) = 2\pi r.$$

Exemple 2.5 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . Alors le graphe de f est la trace de la courbe $\gamma : [a, b] \rightarrow \mathbb{R}^2$ définie par :

$$\gamma(x) = (x, f(x)).$$

Ainsi, la longueur du graphe de f est :

$$\ell(\gamma) = \int_a^b \sqrt{1 + f'(x)^2} dx.$$

2.3.2 Courbes équivalentes

La question naturelle suivante se pose. Si deux courbes paramétrées de classe \mathcal{C}^1 ont la même trace, est-ce qu'elles ont même longueur ? La réponse est positive dans le cas suivant.

Soit $\gamma_1 : I_1 = [a_1, b_1] \rightarrow \mathbb{R}^n$, $\gamma_2 : I_2 = [a_2, b_2] \rightarrow \mathbb{R}^n$ deux courbes paramétrées. On dit que γ_1 et γ_2 sont équivalentes, ssi il existe un difféomorphisme- \mathcal{C}^1 , $\varphi : I_1 \rightarrow I_2$ tel que :

$$\gamma_1 = \gamma_2 \circ \varphi.$$

On rappelle qu'un difféomorphisme- \mathcal{C}^1 , $\varphi : I_1 \rightarrow I_2$ est une fonction bijective, qui est de classe \mathcal{C}^1 sur $\text{int}(I_1)$, et dont l'inverse est de classe \mathcal{C}^1 sur $\text{int}(I_2)$.

Remarque 2.2 La relation ci-dessus est une *relation d'équivalence*, c'est-à-dire qu'elle est *réflexive* (γ_1 est équivalente à γ_1), *symétrique* (si γ_1 et γ_2 sont équivalentes, alors γ_2 et γ_1 le sont aussi) et *transitive* (si γ_1 est équivalente à γ_2 , et γ_2 est équivalente à γ_3 , alors γ_1 est équivalente à γ_3).

Théorème 12 Soient $\gamma_1 : I_1 \rightarrow \mathbb{R}^n$, $\gamma_2 : I_2 \rightarrow \mathbb{R}^n$ deux courbes paramétrées de classe \mathcal{C}^1 . Si elles sont équivalentes, alors elles ont même longueur.

Preuve:

Comme γ_1 et γ_2 sont équivalentes, il existe un difféomorphisme- \mathcal{C}^1 , $\varphi : I_1 \rightarrow I_2$ tel que :

$$\gamma_1 = \gamma_2 \circ \varphi.$$

Comme φ est un difféomorphisme- \mathcal{C}^1 , $\varphi' \neq 0$ sur I_1 . Ainsi, il y a deux cas possible, soit $\varphi' > 0$ et φ est strictement croissante, soit $\varphi' < 0$ et φ est strictement décroissante.

Nous traitons les deux cas en même temps dans la preuve.

$$\begin{aligned}
 \ell(\gamma_1) &= \int_{a_1}^{b_1} \|\gamma_1'(t)\| dt, \\
 &= \int_{a_1}^{b_1} \|\gamma_2'(\varphi(t))\| |\varphi'(t)| dt, \quad (\text{dérivation fonction composée}), \\
 &= \operatorname{sgn} \varphi' \int_{a_1}^{b_1} \|\gamma_2'(\varphi(t))\| |\varphi'(t)| dt, \\
 &= \operatorname{sgn} \varphi' \int_{\varphi(a_1)}^{\varphi(b_1)} \|\gamma_2'(s)\| ds, \quad (\text{changement de variable pour fonctions réelles}), \\
 &= \int_{a_2}^{b_2} \|\gamma_2'(s)\| ds = \ell(\gamma_2).
 \end{aligned}$$

□

2.3.3 Paramétrisation par longueur d'arc

Nous avons défini une relation d'équivalence pour les courbes paramétrées. Il est intéressant d'observer que certaines propriétés physiques se reflètent mieux sous certaines paramétrisations. Par exemple, si la vitesse de la courbe est constante, alors le vecteur vitesse (autrement dit le vecteur tangent) est orthogonal au vecteur accélération (le vecteur deuxième dérivée), ce qui est cohérent avec notre intuition physique. Le Théorème 13 ci-dessous donne une condition nécessaire pour qu'une courbe puisse être reparamétrisée avec une vitesse constante. La preuve fournit la reparamétrisation explicite, appelée paramétrisation par *longueur d'arc*, ou par *abscisse curviligne*.

Théorème 13 Soit $\gamma : I = [a, b] \rightarrow \mathbb{R}^n$ une courbe régulière. Alors γ est équivalente à une courbe de vitesse égale à 1.

Preuve:

Soit $L = \ell(\gamma)$. On définit $\sigma : [a, b] \rightarrow [0, L]$ par :

$$\sigma(t) = \int_a^t \|\gamma'(u)\| du.$$

Ainsi, $\sigma(t)$ est la longueur de la courbe jusqu'à l'instant t , d'où la terminologie *longueur d'arc*.

En utilisant le théorème fondamental du calcul intégral, on obtient $\sigma'(t) = \|\gamma'(t)\| > 0$, de sorte que σ est une fonction de classe \mathcal{C}^1 dont la dérivée est strictement positive. Ainsi, la fonction inverse $\varphi : [0, L] \rightarrow [a, b]$ existe, et sa dérivée est donnée par :

$$\varphi'(s) = \frac{1}{\sigma'(\varphi(s))} > 0.$$

On déduit que φ' est aussi continue, de sorte que φ est un difféomorphisme- \mathcal{C}^1 . Soit $\hat{\gamma} : [0, L] \rightarrow \mathbb{R}^n$, la courbe équivalente à γ définie par $\hat{\gamma} = \gamma \circ \varphi$. Alors,

$$\|\hat{\gamma}'(s)\| = \|\gamma'(\varphi(s))\varphi'(s)\| = \frac{\|\gamma'(\varphi(s))\|}{\sigma'(\varphi(s))} = 1,$$

car $\sigma'(t) = \|\gamma'(t)\|$. □

Lemme 14 *Soit $\gamma : I \rightarrow \mathbb{R}^n$ une courbe paramétrée de classe \mathcal{C}^2 , qui a vitesse constante. Alors, le vecteur vitesse et le vecteur accélération sont orthogonaux, c'est-à-dire,*

$$\forall t \in I, \quad \gamma'(t) \cdot \gamma''(t) = 0.$$

Preuve:

Ce lemme sera démontré aux exercices. □

Chapitre 3

Fonctions de plusieurs variables

3.1 Définition et représentation

Soit S un sous-ensemble de \mathbb{R}^m . Une *fonction* $f : S \rightarrow \mathbb{R}$ est une application qui à tout point de S associe un élément de \mathbb{R} . Souvenez-vous qu'un point \mathbf{x} de S est en fait un m -tuple (x_1, \dots, x_m) de points de \mathbb{R} .

Exemple 3.1 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$f(x, y) = x^2 + y^2.$$

On définit le *graphe* de f comme étant l'ensemble des points de \mathbb{R}^{m+1} de la forme :

$$\{(x_1, \dots, x_m, f(x_1, \dots, x_m)) : (x_1, \dots, x_m) \in S\}.$$

Exemple 3.2 Graphe de la fonction de l'Exemple 3.1.

Remarquez que l'on peut seulement représenter graphiquement les graphes des fonctions de \mathbb{R} ou \mathbb{R}^2 dans \mathbb{R} . Nous allons néanmoins décrire une autre manière de représenter les fonctions qui permet de visualiser d'une certaine manière les fonctions de \mathbb{R}^3 dans \mathbb{R} .

Pour tout nombre $a \in \mathbb{R}$, l'équation $f(x, y) = a$ est l'équation d'une courbe dans \mathbb{R}^2 , appelée la *courbe de niveau* a de f . Elle décrit l'ensemble des points où f prend la valeur a . En dessinant un nombre suffisant de courbes de niveau, on obtient une idée de l'allure de la courbe, cf l'exemple ci-dessous.

Exemple 3.3 Courbes de niveau de la fonction de l'Exemple 3.1.

L'idée des courbes de niveau est celle des lignes de contour sur les cartes géographiques. Chaque ligne décrit l'"altitude" de la fonction. Si le graphe est interprété comme une région montagneuse, chaque ligne de niveau donne l'ensemble des points à altitude constante. En physique, une telle fonction f peut donner par exemple la distribution de la température sur une certaine région. Les courbes de niveau sont alors appelées *lignes isothermes*.

Dans le cas où $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, on peut définir de manière analogue l'équation $f(x, y, z) = a$, qui est maintenant l'équation d'une surface dans \mathbb{R}^3 , appelée *surface de niveau*. La représentation de ces surfaces pour différentes valeurs de a peut aider à visualiser la fonction f .

3.2 Dérivées partielles

Soit $f : U \subset \mathbb{R}^m \rightarrow \mathbb{R}$ une fonction, où U est un ouvert de \mathbb{R}^m .

Etant donné que U est un ouvert de \mathbb{R}^m , on sait que si f est définie au point $\mathbf{x} = (x_1, \dots, x_m)$, alors f est aussi définie en tout point d'une petite boule ouverte autour de \mathbf{x} , cf Section 1.4. En particulier, f est définie au point $(x_1, \dots, x_i + h, \dots, x_m)$, lorsque h est suffisamment petit. L'expression suivante a donc un sens :

$$\frac{f(x_1, \dots, x_i + h, \dots, x_m) - f(x_1, \dots, x_m)}{h}.$$

Si la limite lorsque h tend vers zéro de la quantité ci-dessus existe, on l'appelle la *i -ème dérivée partielle* de f , que l'on note $D_i f(x_1, \dots, x_m)$ ou $D_i f(\mathbf{x})$ ou encore $\frac{\partial f}{\partial x_i}$.

Une dérivée partielle est donc obtenue comme une dérivée ordinaire dans une des directions, gardant toutes les autres directions fixées. Lorsque $m = 2$ on notera habituellement x_1, x_2 par x, y , et lorsque $m = 3$, on utilisera la notation x, y, z pour x_1, x_2, x_3 .

Exemple 3.4 Soit $f(x, y) = \sin xy$. Alors

$$D_1 f(x, y) = y \cos xy, \quad D_2 f(x, y) = x \cos xy.$$

En particulier $D_1 f(0, \pi) = \pi$, $D_2 f(0, \pi) = 0$.

Soit f une fonction définie sur un ouvert U de \mathbb{R}^m . Supposons que toutes les dérivées partielles existent en tous points de U . Alors on appelle *gradient*, noté $(\text{grad} f)$ ou encore ∇f , le vecteur formé de toutes les dérivées partielles, c'est-à-dire :

$$(\nabla f)(\mathbf{x}) = (\text{grad} f)(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_m} \right) = (D_1 f(\mathbf{x}), \dots, D_m f(\mathbf{x})).$$

Exemple 3.5 Reprenant l'exemple précédent où $f(x, y) = \sin xy$, le gradient est :

$$(\nabla f)(x, y) = (y \cos xy, x \cos xy).$$

En utilisant les propriétés des dérivées des fonctions réelles, on obtient le théorème suivant :

Théorème 15 Soient f, g deux fonctions définies sur un ouvert U de \mathbb{R}^m . Supposons que toutes les dérivées partielles existent en tout point de U , et soit $a \in \mathbb{R}$. Alors

1. $\nabla(f + g) = \nabla f + \nabla g$.
2. $\nabla(af) = a\nabla f$.

Remarque 3.1 Dès que $m \geq 2$, l'existence des dérivées partielles d'une fonction en un point n'implique pas la continuité de la fonction en ce point. Ceci est illustré par l'exemple suivant, qui sera traité en exercice. La fonction f ci-dessous admet toutes ses dérivées partielles en $(0, 0)$, mais n'est pas continue en ce point.

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

Noter que l'existence des dérivées partielles lorsque $m \geq 2$ est plus faible que l'existence de la dérivée pour des fonctions réelles, qui elle impliquait la continuité de la fonction.

Le but de la prochaine section est de définir la notion de *dérivabilité* pour les fonctions de m variables, puis de la comparer avec l'existence des dérivées partielles.

3.3 Dérivabilité

Dans cette section, nous définissons la notion de dérivabilité des fonctions de \mathbb{R}^m dans \mathbb{R} . C'est un concept plus fort que celui d'existence des dérivées partielles introduit à la Section 3.2. En effet, pour calculer des dérivées partielles, on considère les dérivées dans une direction après l'autre (en l'espace de départ). Pour définir la dérivée, nous allons prendre toutes les directions en même temps.

Regardons d'abord ce qu'il serait naturel de faire au vu de nos connaissances des fonctions réelles. Regardons ensuite pourquoi cela n'est pas possible, puis trouvons un moyen de contourner le problème.

Soit f une fonction définie sur un ouvert U de \mathbb{R}^m , et soit $\mathbf{x} \in U$. Comme U est ouvert, pour tout vecteur $\mathbf{h} = (h_1, \dots, h_m)$ suffisamment petit (i.e. la norme $\|\mathbf{h}\|$ est suffisamment

petite), $\mathbf{x} + \mathbf{h} = (x_1 + h_1, \dots, x_m + h_m)$ est encore dans U . Comme dans le cas des fonctions d'une variable, on souhaiterait étudier la limite du quotient :

$$\frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})}{\mathbf{h}},$$

mais ceci n'a AUCUN SENS, car on divise par un vecteur !!! Donc, il faut trouver autre chose.

Considérons à nouveau le cas d'une fonction d'une variable réelle. Supposons qu'elle soit dérivable au point x . La dérivée de f au point x est alors définie comme étant la limite :

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}.$$

Définissons

$$g(h) = \frac{f(x+h) - f(x)}{h} - f'(x).$$

Alors g n'est pas définie en 0, mais $\lim_{h \rightarrow 0} g(h) = 0$. On peut donc écrire pour tout $h \neq 0$, suffisamment petit,

$$f(x+h) - f(x) = f'(x)h + hg(h),$$

où $\lim_{h \rightarrow 0} g(h) = 0$. Au lieu d'avoir h dans le terme de droite, on souhaite avoir $|h|$, ainsi lorsque h est négatif, on écrit :

$$f(x+h) - f(x) = f'(x)h - h(-g(h)),$$

où $\lim_{h \rightarrow 0} -g(h) = 0$. Nous avons donc montré que si f est dérivable, il existe une fonction g telle que :

$$\begin{aligned} f(x+h) - f(x) &= f'(x)h + |h|g(h), \quad \text{lorsque } h \neq 0 \\ \lim_{h \rightarrow 0} g(h) &= 0. \end{aligned}$$

Réciproquement, supposons qu'il existe un nombre a_x et une fonction g telle que :

$$\begin{aligned} f(x+h) - f(x) &= a_x h + |h|g(h), \quad \text{lorsque } h \neq 0 \\ \lim_{h \rightarrow 0} g(h) &= 0. \end{aligned}$$

Alors, pour $h \neq 0$, on a :

$$\frac{f(x+h) - f(x)}{h} = a_x + \frac{|h|}{h}g(h)$$

Prenant la limite lorsque $h \rightarrow 0$, on obtient a_x . De sorte que f est dérivable et sa dérivée est égale à a_x . Ainsi, on aurait pu utiliser cette définition pour la dérivabilité. Cette dernière définition a le grand avantage de ne pas faire intervenir h au dénominateur. Elle se prête donc bien à une généralisation au cas des fonctions de \mathbb{R}^m dans \mathbb{R} .

Soit f une fonction définie sur un ouvert U de \mathbb{R}^m . Soit $\mathbf{h} = (h_1, \dots, h_m)$ un vecteur tel que $\|\mathbf{h}\|$ est suffisamment petit pour que $\mathbf{x} + \mathbf{h} \in U$, lorsque $\mathbf{x} \in U$. On dit que f est *dérivable au point \mathbf{x}* s'il existe un vecteur \mathbf{a}_x de \mathbb{R}^m et une fonction g définie pour \mathbf{h} petit, tels que :

$$\begin{aligned} f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) &= \mathbf{a}_x \cdot \mathbf{h} + \|\mathbf{h}\|g(\mathbf{h}) \\ \lim_{\mathbf{h} \rightarrow \mathbf{0}} g(\mathbf{h}) &= 0. \end{aligned}$$

Le vecteur \mathbf{a}_x s'appelle alors la *dérivée de f au point \mathbf{x}* . On dit que f est dérivable en U , ssi elle est dérivable en tout point de U .

Théorème 16 *Soit $f : U \rightarrow \mathbb{R}$, où U est un ouvert de \mathbb{R}^m . Si f est dérivable au point \mathbf{x} , alors la dérivée est unique.*

Preuve:

Supposons qu'il existe $\mathbf{a}_x, \mathbf{b}_x$ deux vecteurs de \mathbb{R}^m , et deux fonctions g_1, g_2 telles que :

$$\begin{aligned} f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) &= \mathbf{a}_x \cdot \mathbf{h} + \|\mathbf{h}\|g_1(\mathbf{h}) \\ f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) &= \mathbf{b}_x \cdot \mathbf{h} + \|\mathbf{h}\|g_2(\mathbf{h}), \end{aligned}$$

et $\lim_{\|\mathbf{h}\| \rightarrow 0} g_1(\mathbf{h}) = \lim_{\|\mathbf{h}\| \rightarrow 0} g_2(\mathbf{h}) = 0$. Prenant la différence des deux équations, on obtient :

$$(\mathbf{a}_x - \mathbf{b}_x) \cdot \mathbf{h} = \|\mathbf{h}\|(g_1(\mathbf{h}) - g_2(\mathbf{h})), \quad (3.1)$$

pour tout \mathbf{h} suffisamment petit. On considère maintenant $\mathbf{y} \in \mathbb{R}^m$, et $t > 0$ petit, de sorte que $t\mathbf{y}$ soit suffisamment petit, alors remplaçant \mathbf{h} par $t\mathbf{y}$ dans (3.1), et simplifiant par $t > 0$, on obtient :

$$(\mathbf{a}_x - \mathbf{b}_x) \cdot \mathbf{y} = \|\mathbf{y}\|(g_1(t\mathbf{y}) - g_2(t\mathbf{y})).$$

Prenant la limite lorsque $t \rightarrow 0$ implique que pour tout $\mathbf{y} \in \mathbb{R}^m$,

$$(\mathbf{a}_x - \mathbf{b}_x) \cdot \mathbf{y} = 0,$$

ce qui est équivalent à $\mathbf{a}_x - \mathbf{b}_x = \mathbf{0}$. □

3.4 Dérivée et gradient

Les deux théorèmes suivants permettent de relier la notion de dérivée et celle de dérivées partielles.

Théorème 17 Soit $f : U \rightarrow \mathbb{R}$, où U est un ouvert de \mathbb{R}^m . Si f est dérivable au point \mathbf{x} , alors toutes les dérivées partielles de f existent au point \mathbf{x} , et la dérivée est le gradient au point \mathbf{x} , i.e.

$$\mathbf{a}_{\mathbf{x}} = (\nabla f)(\mathbf{x}).$$

Preuve:

Comme f est dérivable au point \mathbf{x} , il existe un unique vecteur $\mathbf{a}_{\mathbf{x}}$ et une fonction g tels que $\lim_{\|\mathbf{h}\| \rightarrow 0} g(\mathbf{h}) = 0$, et pour tout \mathbf{h} suffisamment petit :

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \mathbf{a}_{\mathbf{x}} \cdot \mathbf{h} + \|\mathbf{h}\|g(\mathbf{h}).$$

En particulier, si on définit $\mathbf{h}_i = (0, \dots, h, \dots, 0)$, où $h \in \mathbb{R}$ est la $i^{\text{ème}}$ coordonnée, et si on utilise la définition du produit scalaire et de la norme, on obtient :

$$f(\mathbf{x} + \mathbf{h}_i) - f(\mathbf{x}) = \mathbf{a}_{\mathbf{x},i} h + |h|g(\mathbf{h}_i),$$

où $\mathbf{a}_{\mathbf{x},i}$ est la $i^{\text{ème}}$ coordonnée du vecteur dérivée $\mathbf{a}_{\mathbf{x}}$. Supposons que $h \neq 0$. On peut alors diviser par h , et prendre la limite lorsque $h \rightarrow 0$ pour obtenir :

$$\lim_{h \rightarrow 0} \frac{f(\mathbf{x} + \mathbf{h}_i) - f(\mathbf{x})}{h} = \mathbf{a}_{\mathbf{x},i}.$$

Ceci implique que la $i^{\text{ème}}$ dérivée partielle de f existe et est égale à $\mathbf{a}_{\mathbf{x},i}$ de sorte que :

$$\mathbf{a}_{\mathbf{x}} = (\nabla f)(\mathbf{x}).$$

□

Théorème 18 Soit $f : U \rightarrow \mathbb{R}$, où U est un ouvert de \mathbb{R}^m . Si toutes les dérivées partielles de f existent dans U , et qu'elles sont de plus continues, alors f est dérivable en tout point de U .

Preuve:

Pour simplifier les notations, on va supposer que $m = 2$. La preuve dans le cas d'un m quelconque est tout à fait similaire. On considère $\mathbf{x} = (x, y) \in U$, et $\mathbf{h} = (h, k)$. Pour \mathbf{h} suffisamment petit, nous devons étudier la différence $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = f(x + h, y + k) - f(x, y)$, qui peut aussi s'écrire :

$$f(x + h, y + k) - f(x, y + k) + f(x, y + k) - f(x, y).$$

Utilisant le théorème de la valeur intermédiaire pour les fonctions d'une variable, ainsi que la définition des dérivées partielles, nous savons qu'il existe $s \in (x, x+h)$ et $t \in (y, y+k)$ tels que :

$$\begin{aligned} f(x+h, y+k) - f(x, y+k) &= D_1 f(s, y+k)h \\ f(x, y+k) - f(x, y) &= D_2 f(x, t)k. \end{aligned}$$

Ainsi, $f(x+h, y+k) - f(x, y) = D_1 f(s, y+k)h + D_2 f(x, t)k$. En posant $g_1(\mathbf{h}) = D_1 f(s, y+k) - D_1 f(x, y)$ et $g_2(\mathbf{h}) = D_2 f(x, t) - D_2 f(x, y)$, on peut réécrire ceci de la manière suivante :

$$\begin{aligned} f(x+h, y+k) - f(x, y) &= D_1 f(x, y)h + D_2 f(x, y)k + g_1(\mathbf{h})h + g_2(\mathbf{h})k \\ f(x+h, y+k) - f(x, y) &= \nabla f(\mathbf{x}) \cdot \mathbf{h} + \|\mathbf{h}\|g(\mathbf{h}), \end{aligned}$$

où $g(\mathbf{h}) = \frac{h}{\|\mathbf{h}\|}g_1(\mathbf{h}) + \frac{k}{\|\mathbf{h}\|}g_2(\mathbf{h})$. Ainsi, il suffit de démontrer que

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} g_1(\mathbf{h}) = \lim_{\mathbf{h} \rightarrow \mathbf{0}} g_2(\mathbf{h}) = 0.$$

Mais ceci est effectivement vrai par continuité des dérivées partielles, et parce que :

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} (s, y+k) = (x, y) = \lim_{\mathbf{h} \rightarrow \mathbf{0}} (x, t).$$

□

Remarque 3.2 Les Théorèmes 17 et 18 sont optimaux comme illustré par ces deux exemples qui seront traités aux exercices.

Exemple 3.6 Si f est dérivable au point x , alors les dérivées partielles ne sont pas forcément continues. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$f(x, y) = \begin{cases} x^3 \sin \frac{1}{x} + y^2 & \text{lorsque } x \neq 0, \\ y^2 & \text{lorsque } x = 0. \end{cases}$$

Alors f est dérivable en $(0, 0)$, mais $D_1 f$ n'est pas continue en $(0, 0)$.

Exemple 3.7 Si les dérivées partielles existent au point \mathbf{x} alors f n'est pas forcément dérivable. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

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

Alors les dérivées partielles existent en $(0, 0)$, mais f n'est pas dérivable en $(0, 0)$.

3.5 Dérivation composées et interprétations géométriques du gradient

Dans cette section, nous démontrons la formule donnant la dérivée de la composition d'une fonction de plusieurs variables et d'une courbe. C'est une généralisation de la formule pour la dérivée des fonctions composées, que vous avez étudiée au cours d'analyse réelle. Cette formule plus générale permet d'interpréter le gradient géométriquement ; elle est également utilisée dans la preuve de la formule de Taylor en dimensions supérieures.

Théorème 19 *Soit f une fonction qui est dérivable sur un ouvert U de \mathbb{R}^n , et soit $\gamma : I \rightarrow \mathbb{R}^n$ une courbe dérivable, telle que $\gamma(I) \subset U$. Alors, la fonction*

$$g = f \circ \gamma : I \rightarrow \mathbb{R},$$

est dérivable sur I , et :

$$g'(t) = (\nabla f)(\gamma(t)) \cdot \gamma'(t).$$

En développant le produit scalaire, on obtient :

$$g'(t) = \frac{\partial f}{\partial x_1} \frac{d\gamma_1}{dt} + \cdots + \frac{\partial f}{\partial x_n} \frac{d\gamma_n}{dt} = D_1 f(\gamma(t)) \gamma'_1(t) + \cdots + D_n f(\gamma(t)) \gamma'_n(t).$$

Preuve:

Comme g est une fonction d'une variable réelle, pour établir sa dérivabilité, nous devons étudier :

$$\frac{g(t+h) - g(t)}{h} = \frac{f(\gamma(t+h)) - f(\gamma(t))}{h}.$$

Soit $\mathbf{k} = \gamma(t+h) - \gamma(t)$. Alors le quotient ci-dessus s'écrit :

$$\frac{f(\gamma(t) + \mathbf{k}) - f(\gamma(t))}{h}.$$

Par définition de la dérivabilité de la fonction f , pour \mathbf{k} suffisamment petit, il existe une fonction φ telle que :

$$f(\gamma(t) + \mathbf{k}) - f(\gamma(t)) = (\nabla f)(\gamma(t)) \cdot \mathbf{k} + \|\mathbf{k}\| \varphi(\mathbf{k}),$$

et $\lim_{\|\mathbf{k}\| \rightarrow 0} \varphi(\mathbf{k}) = 0$. En remplaçant \mathbf{k} par sa définition, et en divisant par h , on obtient :

$$\frac{f(\gamma(t+h)) - f(\gamma(t))}{h} = (\nabla f)(\gamma(t)) \cdot \frac{\gamma(t+h) - \gamma(t)}{h} + \left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \varphi(\mathbf{k}).$$

De plus, si $h \rightarrow 0$, alors $\mathbf{k} \rightarrow 0$, de sorte que le premier terme tend vers $(\nabla f)(\gamma(t)) \cdot \gamma'(t)$ (en utilisant que γ est dérivable en t), et le deuxième tend vers 0. \square

Exemple 3.8 Soient $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, et $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$ définies par :

$$f(x, y, z) = e^{xy} \cos z, \quad \gamma(t) = (t, \sin t, t^2).$$

Alors la dérivée au point t de la fonction $g = f \circ \gamma$ est :

$$\begin{aligned} g'(t) &= D_1 f(\gamma(t))\gamma'_1(t) + D_2 f(\gamma(t))\gamma'_2(t) + D_3 f(\gamma(t))\gamma'_3(t), \\ &= \sin(t)e^{t \sin(t)} \cos(t^2) + te^{t \sin(t)} \cos(t^2) \cos(t) - e^{t \sin t} \sin(t^2)2t, \\ &= e^{t \sin(t)} [\sin(t) \cos(t^2) + t \cos(t^2) \cos(t) - \sin(t^2)2t]. \end{aligned}$$

Interprétation 1 : Direction de croissance maximale

Soit f une fonction dérivable définie sur un ouvert U de \mathbb{R}^m . Si \mathbf{x} est un point de U tel que le gradient n'est pas nul, alors il donne la direction de croissance maximale. En effet, soit la fonction réelle $g(t) = f(\mathbf{x} + t\mathbf{h})$, où \mathbf{h} est suffisamment petit pour que $\mathbf{x} + t\mathbf{h}$ soit encore dans U . Alors pour un \mathbf{h} fixé, la fonction g représente l'évolution de la fonction f dans la direction \mathbf{h} . Utilisant le Théorème 19, on obtient :

$$g'(0) = (\nabla f)(\mathbf{x}) \cdot \mathbf{h},$$

ainsi la croissance est maximale au point \mathbf{x} lorsque cette dérivée est maximale. Fixons la norme du vecteur \mathbf{h} à 1, le produit scalaire est maximal lorsque les vecteurs sont colinéaires, c'est à dire lorsque $\mathbf{h} = \frac{(\nabla f)(\mathbf{x})}{\|(\nabla f)(\mathbf{x})\|}$. Ainsi la direction de croissance maximale est donnée par le gradient.

Interprétation 2 : Plan tangent à une surface

Nous supposons $m \geq 2$. Soit $f : U \rightarrow \mathbb{R}$, une fonction dérivable, où U est un ouvert de \mathbb{R}^m , et soit $c \in \mathbb{R}$. Le but est de définir le *plan tangent à la surface* S_c , au point $\mathbf{x}_0 \in S_c$, où S_c est donnée par :

$$S_c = \{\mathbf{x} \in U : f(\mathbf{x}) = c\}.$$

La surface S_c est aussi appelée *surface de niveau* c de la fonction f . Remarquer que S_c est un objet " $m - 1$ -dimensionnel". En particulier, lorsque $m = 2$, on parlera de *courbe de niveau* c de la fonction f .

Soit $\gamma : I \rightarrow \mathbb{R}^m$ une courbe paramétrée simple et dérivable. On dit que γ est *sur la surface* S_c , ssi $\gamma(I) \subset S_c$, autrement dit

$$\forall t \in I, f(\gamma(t)) = c.$$

En utilisant le Théorème 19, on obtient :

$$\forall t \in I, \nabla f(\gamma(t)) \cdot \gamma'(t) = 0.$$

On considère un point $\mathbf{x}_0 \in S_c$, et une courbe $\gamma : I \rightarrow \mathbb{R}^m$ sur la surface S_c passant par \mathbf{x}_0 , qui est simple et régulière. Alors, il existe $t_0 \in I$ tel que $\mathbf{x}_0 = \gamma(t_0)$, et donc :

$$\nabla f(\mathbf{x}_0) \cdot \gamma'(t_0) = 0.$$

Ainsi, si $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$, le gradient de f au point \mathbf{x}_0 est orthogonal au vecteur tangent à la courbe en ce point. Ceci reste vrai pour toute courbe sur la surface, régulière et passant par \mathbf{x}_0 .

Lorsque $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$, il est donc naturel de définir le *plan tangent à la surface S_c au point \mathbf{x}_0* , noté $\mathcal{T}_{\mathbf{x}_0}$, comme étant le plan passant par \mathbf{x}_0 , et orthogonal au gradient de f en ce point, i.e.

$$\mathcal{T}_{\mathbf{x}_0} = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x} \cdot \nabla f(\mathbf{x}_0) = \mathbf{x}_0 \cdot \nabla f(\mathbf{x}_0)\}. \quad (3.2)$$

Si $\nabla f(\mathbf{x}_0) = \mathbf{0}$, on dit que le plan tangent au point \mathbf{x}_0 n'est pas bien défini.

Lorsque $m = 2$, on parlera de *droite tangente à la courbe de niveau S_c au point \mathbf{x}_0* . Ainsi, si $f : U \rightarrow \mathbb{R}$, où U est un ouvert de \mathbb{R}^2 , alors le gradient $\nabla f(\mathbf{x})$ en un point $\mathbf{x} \in U$ est orthogonal à la courbe de niveau passant par ce point.

Exemple 3.9 Trouver la droite tangente à la courbe,

$$x^2y + y^3 = 10,$$

au point $\mathbf{x}_0 = (1, 2)$.

Soit $f(x, y) = x^2y + y^3$, alors le gradient de f au point \mathbf{x}_0 est $\nabla f(\mathbf{x}_0) = (4, 13)$. On cherche donc l'équation de la droite orthogonale à ce vecteur, passant par \mathbf{x}_0 . Par (3.2) l'équation de cette droite est donnée par :

$$4x + 13y = 30.$$

Exemple 3.10 Soit $f : U \rightarrow \mathbb{R}$ une fonction dérivable, où U est un ouvert de \mathbb{R}^{m-1} . Le graphe de f :

$$\text{Graphe}(f) = \{(x_1, \dots, x_{m-1}, x_m) \in U \times \mathbb{R} : x_m - f(x_1, \dots, x_{m-1}) = 0\},$$

est une surface de \mathbb{R}^m . C'est la surface de niveau 0 de la fonction $g : U \times \mathbb{R} \rightarrow \mathbb{R}$ définie par $g(x_1, \dots, x_m) = x_m - f(x_1, \dots, x_{m-1})$.

Vous montrerez aux exercices que le plan tangent au graphe de f , au point $(\mathbf{x}_0, f(\mathbf{x}_0))$ est le plan passant par $(\mathbf{x}_0, f(\mathbf{x}_0))$, engendré par les vecteurs $(\mathbf{e}_i, D_i f(\mathbf{x}_0))_{i=1, \dots, m-1}$, où $\mathbf{e}_i = (0, \dots, 1, \dots, 0)$ est le i -ème vecteur de base canonique de \mathbb{R}^{m-1} . Ainsi, dans ce cas on montre que le plan tangent est de dimension $m - 1$.

3.6 Dérivées d'ordre supérieur

Il est maintenant naturel d'étudier les dérivées d'ordre supérieur. L'enjeu principal est de comprendre quand est-ce que l'on a le droit d'échanger l'ordre dans lequel on exécute ces dérivées.

Soit $f : U \rightarrow \mathbb{R}$ une fonction, où U est un ouvert de \mathbb{R}^m . La fonction f est dite de classe \mathcal{C}^k sur U , si toutes les dérivées partielles de f d'ordre au plus k existent et sont continues. Plus précisément, cela signifie que, pour toute suite i_1, \dots, i_q où $q \leq k$, et $i_j \in \{1, \dots, m\}$ pour tout j , la dérivée partielle itérée,

$$D_{i_1} D_{i_2} \cdots D_{i_q} f = \frac{\partial}{\partial x_{i_1}} \frac{\partial}{\partial x_{i_2}} \cdots \frac{\partial}{\partial x_{i_q}} f,$$

existe, et est continue sur U .

Dans ce cas, l'ordre dans lequel on prend les dérivées partielles n'a pas d'importance, comme le dit le théorème suivant.

Théorème 20 Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^k , où U est un ouvert de \mathbb{R}^m . Soit i_1, \dots, i_q une suite d'indices comme ci-dessus, et soit i'_1, \dots, i'_q une permutation de cette suite, alors :

$$D_{i'_1} D_{i'_2} \cdots D_{i'_q} f = D_{i_1} D_{i_2} \cdots D_{i_q} f.$$

Nous ne donnons pas la démonstration de ce théorème. L'idée de la preuve est de procéder par induction sur k .

Remarque 3.3 Il est important pour que ce théorème soit vrai que les dérivées partielles soient continues. Ceci est illustré par l'exemple suivant qui sera traité aux exercices.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

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

Alors, $D_1 D_2 f$ et $D_2 D_1 f$ existent en $(0, 0)$, mais ne sont pas égales.

3.7 Formule de Taylor

Nous souhaitons maintenant établir la formule de Taylor pour les fonctions de plusieurs variables. Comme dans le cas des fonctions réelles, le but est d'approcher une fonction

par un polynôme. Ceci est très utile en pratique étant donné que les polynômes sont des fonctions au comportement bien compris. Dans la majorité des cas, cette approximation est locale, c'est-à-dire qu'elle est valable au voisinage d'un point.

Dans la suite, nous étudions la formule de Taylor dans le cas où le reste est sous la forme de Lagrange. Comme vous vous rappelez peut-être, il y a différentes manières d'exprimer le reste, selon la régularité de la fonction, mais nous ne souhaitons pas rentrer dans ces considérations ici.

La formule de Taylor dans ce contexte plus général est très similaire au cas des fonctions réelles. Elle est cependant compliquée au niveau des notations, de sorte qu'il nous paraît utile de d'abord rappeler le Théorème de Taylor-Lagrange dans le cas des fonctions réelles.

Théorème 21 (Formule de Taylor-Lagrange, cas $m = 1$)

Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^{k+1} sur un ouvert U de \mathbb{R} . Soient $x \in \mathbb{R}$, et $h \neq 0$, tels que U contienne le segment $[x, x+h]$, si $h > 0$, $[x+h, x]$, si $h < 0$. Alors, il existe un nombre η strictement compris dans cet intervalle, tel que :

$$f(x+h) = f(x) + f'(x)h + \dots + \frac{f^{(k)}(x)}{k!}h^k + \frac{f^{(k+1)}(\eta)}{(k+1)!}h^{k+1}.$$

Nous énonçons maintenant le théorème dans le cas général. Ensuite, nous expliciterons les notations, et donnerons un exemple d'application. Finalement, nous démontrerons le théorème, dont la preuve est une jolie application de la dérivation des fonctions composées.

Théorème 22 (Formule de Taylor-Lagrange, cas m quelconque)

Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^{k+1} sur un ouvert U de \mathbb{R}^m . Soit $\mathbf{x}, \mathbf{h} \in \mathbb{R}^m$ tels que U contienne tous les points du segment $L = \{\mathbf{x} + t\mathbf{h} : t \in [0, 1]\}$. Alors il existe un point $\boldsymbol{\eta} \in \text{int}(L)$, tel que :

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + (\mathbf{h} \cdot \nabla)f(\mathbf{x}) + \dots + \frac{(\mathbf{h} \cdot \nabla)^k f(\mathbf{x})}{k!} + \frac{(\mathbf{h} \cdot \nabla)^{k+1} f(\boldsymbol{\eta})}{(k+1)!}. \quad (3.3)$$

La quantité $P_k(\mathbf{h}) = f(\mathbf{x}) + (\mathbf{h} \cdot \nabla)f(\mathbf{x}) + \dots + \frac{(\mathbf{h} \cdot \nabla)^k f(\mathbf{x})}{k!}$ est un polynôme de degré k en les variables h_1, \dots, h_m , que l'on appelle *polynôme de Taylor de degré k de f au point \mathbf{x}* .

Remarque 3.4 Essayons d'abord de comprendre l'équation (3.3). Elle a le grand avantage d'être compacte, de ressembler à la formule dans le cas $m = 1$, mais les notations méritent un éclaircissement.

En utilisant la définition du produit scalaire, et du gradient, on obtient

$$\mathbf{h} \cdot \nabla = h_1 D_1 + \cdots + h_m D_m = \sum_{i=1}^m h_i D_i,$$

où souvenez-vous que $D_i = \frac{\partial}{\partial x_i}$. Afin de comprendre le terme $(\mathbf{h} \cdot \nabla)^r$ pour $1 \leq r \leq k+1$, il est utile de rappeler la *formule du multinôme*. C'est une généralisation de la formule du binôme de Newton lorsque le nombre de termes est plus grand que 2.

$$(x_1 + \cdots + x_m)^r = \sum_{j_1 + \cdots + j_m = r} \binom{r}{j_1 \cdots j_m} x_1^{j_1} \cdots x_m^{j_m}.$$

La somme s'effectue sur tous les m-tuples d'indices entiers (j_1, \dots, j_m) compris entre 0 et r , dont la somme vaut r . Les coefficients du polynôme s'appellent les *coefficients multinômiaux* et sont donnés par la formule suivante, démontrée aux exercices :

$$\binom{r}{j_1 \cdots j_m} = \frac{r!}{j_1! \cdots j_m!}.$$

En appliquant cette formule à $\mathbf{h} \cdot \nabla$, et en utilisant le fait que les dérivées partielles d'ordre au plus $k+1$ commutent lorsque f est de classe \mathcal{C}^{k+1} , on obtient :

$$(\mathbf{h} \cdot \nabla)^r f = \sum_{j_1 + \cdots + j_m = r} \binom{r}{j_1 \cdots j_m} h_1^{j_1} \cdots h_m^{j_m} D_1^{j_1} \cdots D_m^{j_m} f.$$

Même si l'équation (3.3) est plus claire, la lourdeur des indices reste néanmoins non-négligeable, de sorte qu'il nous semble utile d'écrire explicitement les premiers termes lorsque $m = 2$. Dans ce cas, on remplacera $\mathbf{x} = (x_1, x_2)$ par $\mathbf{x} = (x, y)$, et $\mathbf{h} = (h_1, h_2)$ par $\mathbf{h} = (h, k)$. Alors,

$$\begin{aligned} (\mathbf{h} \cdot \nabla) f &= h D_1 f + k D_2 f = h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y}. \\ (\mathbf{h} \cdot \nabla)^2 f &= (h D_1 f + k D_2)^2 f = h^2 D_1^2 f + 2hk D_1 D_2 f + k^2 D_2^2 f, \\ &= h^2 \frac{\partial^2 f}{\partial x^2} + 2hk \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} + k^2 \frac{\partial^2 f}{\partial y^2}. \end{aligned}$$

Ainsi, si f est au moins de classe \mathcal{C}^3 , les premiers termes du développement de Taylor-Lagrange sont :

$$f(x+h, y+k) = f(x, y) + h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} + \frac{1}{2} h^2 \frac{\partial^2 f}{\partial x^2} + hk \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} + \frac{1}{2} k^2 \frac{\partial^2 f}{\partial y^2} + \cdots$$

Preuve:

La preuve est une jolie application du Théorème 19. On définit la fonction composée

$g(t) = f(\gamma(t))$, où $\gamma(t) = \mathbf{x} + t\mathbf{h}$ pour $t \in [0, 1]$. Alors, $g(0) = f(\mathbf{x})$ et $g(1) = f(\mathbf{x} + \mathbf{h})$, de sorte qu'en appliquant la formule de Taylor à la fonction réelle g autour de 0, on obtient :

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \sum_{\ell=1}^k \frac{g^{(\ell)}(0)}{\ell!} + \frac{g^{(k+1)}(s)}{(k+1)!},$$

où $s \in (0, 1)$. Il suffit donc de montrer que pour tout $\ell \in \{1, \dots, k+1\}$,

$$g^{(\ell)}(t) = (\mathbf{h} \cdot \nabla)^\ell f(\mathbf{x} + t\mathbf{h}).$$

Nous allons procéder par induction sur ℓ . Lorsque $\ell = 1$, on a :

$$\begin{aligned} g'(t) &= (\nabla f)(\gamma(t)) \cdot \gamma'(t) \quad (\text{par le Théorème 19}), \\ &= (\nabla f)(\mathbf{x} + t\mathbf{h}) \cdot \mathbf{h} \quad (\text{en utilisant la définition de } \gamma), \\ &= (\mathbf{h} \cdot \nabla)f(\mathbf{x} + t\mathbf{h}) \quad (\text{autre notation}). \end{aligned}$$

Supposons maintenant que ce soit démontré pour ℓ , alors $g^{(\ell)}(t) = (\mathbf{h} \cdot \nabla)^\ell f(\mathbf{x} + t\mathbf{h})$, et :

$$\begin{aligned} g^{(\ell+1)}(t) &= \frac{d}{dt} g^{(\ell)}(t), \\ &= (\mathbf{h} \cdot \nabla)(\mathbf{h} \cdot \nabla)^\ell f(\mathbf{x} + t\mathbf{h}) \quad (\text{par hypothèse de récurrence}), \\ &= (\mathbf{h} \cdot \nabla)^{(\ell+1)} f(\mathbf{x} + t\mathbf{h}). \end{aligned}$$

□

Remarque 3.5 Dans la pratique, il est utile de poser $\mathbf{x} = \mathbf{a} + \mathbf{h}$, de sorte que la formule de Taylor Lagrange devient (sous les mêmes hypothèses) :

$$f(\mathbf{x}) = f(\mathbf{a}) + ((\mathbf{x} - \mathbf{a}) \cdot \nabla)f(\mathbf{a}) + \dots + \frac{((\mathbf{x} - \mathbf{a}) \cdot \nabla)^k f(\mathbf{a})}{k!} + \frac{((\mathbf{x} - \mathbf{a}) \cdot \nabla)^{k+1} f(\boldsymbol{\eta})}{(k+1)!}.$$

Le polynôme de Taylor $P_k(\mathbf{x} - \mathbf{a})$ de f au point \mathbf{a} , est un polynôme de degré k en les variables $x_1 - a_1, \dots, x_m - a_m$.

Exemple 3.11 Calculons le polynôme de Taylor de degré 2 aux points $(0, 0)$ et $(1, 1)$ de la fonction $f(x, y) = \log(1 + x + 2y)$. Remarquons que f est de classe \mathcal{C}^3 au voisinage de $(0, 0)$ et $(1, 1)$, de sorte que f admet un polynôme de Taylor de degré 2 au voisinage de ces points. Les dérivées partielles d'ordre 1 et 2 sont :

$$\begin{aligned} D_1 f(x, y) &= \frac{1}{1 + x + 2y}, \quad D_2 f(x, y) = \frac{2}{1 + x + 2y}, \\ D_1 D_2 f(x, y) &= -\frac{2}{(1 + x + 2y)^2}, \quad D_1^2 f(x, y) = -\frac{1}{(1 + x + 2y)^2}, \quad D_2^2 f(x, y) = -\frac{4}{(1 + x + 2y)^2}. \end{aligned}$$

Ainsi on a les polynômes de Taylor suivants aux points $(0, 0)$ et $(1, 1)$:

$$P_2(x, y) = x + 2y - \frac{1}{2}(x^2 + 4xy + 4y^2),$$

$$P_2(x - 1, y - 1) = \log(4) + \frac{x - 1}{4} + \frac{y - 1}{2} - \frac{1}{2} \left(\frac{(x - 1)^2}{16} + \frac{(x - 1)(y - 1)}{4} + \frac{(y - 1)^2}{4} \right).$$

Le théorème suivant donne l'unicité du polynôme de Taylor. Il est utile pour trouver ce polynôme sans devoir calculer explicitement toutes les dérivées partielles. Nous le donnons sans démonstration.

Théorème 23 Soit $f : U \rightarrow \mathbb{R}$, une fonction de classe \mathcal{C}^{k+1} sur un ouvert U de \mathbb{R}^m . Si Q est un polynôme de degré k qui vérifie :

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - Q(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|^k} = 0,$$

alors Q est le polynôme de Taylor de degré k de f au point \mathbf{a} .

Exemple 3.12 Reprenons l'exemple précédent. Du cours d'analyse réelle, vous vous souvenez que le développement de Taylor de la fonction $\log(1 + u)$ à l'ordre 2 en $u = 0$ est :

$$\log(1 + u) = u - \frac{u^2}{2} + R(u),$$

où $\lim_{u \rightarrow 0} \frac{R(u)}{u^2} = 0$. Remplaçant u par $x + 2y$, on obtient :

$$f(x, y) = \log(1 + x + 2y) = x + 2y - \frac{1}{2}(x^2 + 4xy + 4y^2) + R(x + 2y).$$

Pour montrer que $P_2(x, y) = x + 2y - \frac{1}{2}(x^2 + 4xy + 4y^2)$ est bien le polynôme de Taylor de degré 2 de f autour de $(0, 0)$, il suffit de vérifier que :

$$\lim_{\mathbf{x} \rightarrow \mathbf{0}} \frac{R(x + 2y)}{x^2 + y^2} = 0.$$

$$\text{Or } \lim_{\mathbf{x} \rightarrow \mathbf{0}} \frac{R(x + 2y)}{x^2 + y^2} = \lim_{\mathbf{x} \rightarrow \mathbf{0}} \left(\frac{R(x + 2y)}{(x + 2y)^2} \frac{(x + 2y)^2}{x^2 + y^2} \right) = 0.$$

3.8 Maximums, minimums, points selles

Pour les fonctions d'une variable sur un intervalle ouvert, vous avez étudié les points critiques, et appris à déterminer si ces derniers sont des maximums, minimums, ou points selles en fonction du comportement de la dérivée seconde. Dans cette section, nous allons faire une étude similaire pour les fonctions de plusieurs variables.

3.8.1 Points critiques

Soit f une fonction dérivable sur un ensemble ouvert U de \mathbb{R}^m , et soit $\mathbf{x} \in U$. Le point \mathbf{x} est appelé *point critique* de f si :

$$(\nabla f)(\mathbf{x}) = \mathbf{0}.$$

Exemple 3.13 On considère la fonction $f(x, y) = e^{-(x^2+y^2)}$. Les dérivées partielles de f sont,

$$D_1 f(x, y) = -2xe^{-(x^2+y^2)}, \quad \text{et } D_2 f(x, y) = -2ye^{-(x^2+y^2)}.$$

Il n'y a qu'un seul point pour lequel les deux dérivées partielles s'annulent : $(0, 0)$. Ce dernier est donc l'unique point critique de f .

De manière analogue au cas des fonctions réelles, si une fonction est dérivable sur un ouvert U , et si elle admet un extremum en un point $\mathbf{x} \in U$, alors \mathbf{x} est un point critique de f . Le réciproque n'est cependant pas vraie. Ecrivons maintenant ceci de manière rigoureuse.

Soit f une fonction, dérivable ou non, définie sur un ouvert U de \mathbb{R}^m . La fonction f admet un *maximum local* (resp. *minimum local*) au point $\mathbf{x} \in U$, s'il existe une boule ouverte $B \subset U$, centrée en \mathbf{x} , telle que :

$$f(\mathbf{y}) \leq f(\mathbf{x}), \quad \forall \mathbf{y} \in B, \quad (\text{resp. } f(\mathbf{y}) \geq f(\mathbf{x}), \quad \forall \mathbf{y} \in B).$$

Le maximum local (minimum local) est dit *strict* si l'inégalité ci-dessus est stricte. La fonction f admet un *extremum local* au point $\mathbf{x} \in U$, si elle admet soit un maximum, soit un minimum local en ce point.

Remarquer que dans le cas de dimension 1, une boule ouverte est un intervalle ouvert, de sorte que vous retrouvez la définition du cours d'analyse réelle.

Théorème 24 Soit f une fonction dérivable sur un ensemble ouvert U de \mathbb{R}^m . Si f admet un extremum local en un point $\mathbf{x} \in U$, alors \mathbf{x} est un point critique de f .

Preuve:

Supposons que f admette un maximum local en \mathbf{x} (le cas d'un minimum est symétrique). Soit la fonction d'une variable $g(t) = f(\mathbf{x} + t\mathbf{h})$. Alors pour \mathbf{h} suffisamment petit, $\mathbf{x} + t\mathbf{h}$ est dans la boule ouverte B , de sorte que :

$$g(t) \leq g(0).$$

Ainsi la fonction d'une variable g , admet un maximum local en 0, donc sa dérivée en 0 est nulle. En utilisant la dérivation des fonctions composées pour calculer $g'(0)$, on obtient :

$$g'(0) = (\nabla f)(\mathbf{x}) \cdot \mathbf{h} = 0, \quad \text{pour tout } \mathbf{h} \text{ suffisamment petit,}$$

ce qui est équivalent à dire que $(\nabla f)(\mathbf{x}) = 0$. \square

3.8.2 Formes quadratiques

Avant d'établir une classification des points critiques, il nous semble utile de faire un rappel sur les formes quadratiques.

Par définition, une *forme quadratique* sur \mathbb{R}^m est un polynôme homogène de degré 2 en les variables x_1, \dots, x_m . Un polynôme est dit *homogène* si tous les monômes ont le même degré.

Exemple 3.14 Le polynôme suivant est une forme quadratique sur \mathbb{R}^3 .

$$q(x, y, z) = z^2 + 3y^2 + 4xy + xz.$$

Le polynôme $x^3 + yx + 4xyz$ n'est pas un polynôme homogène de degré 3.

Soit

$$q(\mathbf{x}) = \sum_{i=1}^m a_{ii}x_i^2 + \sum_{1 \leq i < j \leq m} a_{ij}x_i x_j$$

une forme quadratique sur \mathbb{R}^m . Alors, on peut écrire q de la manière suivante :

$$q(\mathbf{x}) = {}^t \mathbf{x} Q \mathbf{x}$$

où ${}^t \mathbf{x}$ est la transposée du vecteur \mathbf{x} , et Q est la matrice $m \times m$, symétrique, définie par :

$$Q_{ij} = \begin{cases} a_{ii} & \text{si } i = j \\ \frac{1}{2}a_{ij} & \text{si } i < j. \end{cases}$$

On appelle Q la *matrice de la forme quadratique* q . Voici quelques définitions qui seront utiles pour déterminer la nature des points critiques.

Une forme quadratique est dite *définie positive* (resp. *définie négative*) si :

$$q(\mathbf{x}) > 0, \quad \forall \mathbf{x} \neq \mathbf{0}, \quad (\text{resp. } q(\mathbf{x}) < 0, \quad \forall \mathbf{x} \neq \mathbf{0}).$$

Elle est dite *semi-définie positive* ou *semi-définie négative* si l'inégalité ci-dessus n'est pas stricte. Une forme quadratique est dite *non-semi-définie*, si il existe $\mathbf{x}_1 \neq \mathbf{0}$ et

$\mathbf{x}_2 \neq 0$, tels que $q(\mathbf{x}_1) < 0$ et $q(\mathbf{x}_2) > 0$, c'est à dire, si elle n'est ni semi-définie positive, ni semi-définie négative.

Les critères suivants sont utiles pour déterminer si une forme quadratique est définie positive. Etant donné qu'ils sont classiques, et appartiennent plutôt à un cours d'algèbre linéaire, nous les donnons sans démonstration.

Lemme 25 *Les assertions suivantes sont équivalentes :*

1. *La forme quadratique q est définie positive.*
2. *La matrice Q de la forme quadratique q satisfait :*

$${}^t\mathbf{x}Q\mathbf{x} > 0, \forall \mathbf{x} \neq 0.$$

3. *Les valeurs propres de Q sont toutes strictement positives.*
4. *Pour $k = 1, \dots, m$, les déterminants des $k \times k$ coins supérieurs gauches de Q sont strictement positifs.*

Pour les formes quadratiques semi-définies et non semi-définies, on a la correspondance suivante avec les valeurs propres de la matrice Q .

Lemme 26

1. *La forme quadratique q est semi-définie positive (négative), ssi les valeurs propres de Q sont ≥ 0 (≤ 0).*
2. *La forme quadratique q est non-semi-définie, ssi Q a au moins une valeur propre positive et une négative.*

3.8.3 Classification des points critiques

Dans cette section, nous établissons la généralisation du critère de la deuxième dérivée pour déterminer si en un point critique une fonction de plusieurs variables admet un maximum local, minimum local ou un point selle.

On considère une fonction f définie sur un ouvert U de \mathbb{R}^m , qui est de classe \mathcal{C}^3 dans un voisinage d'un point critique \mathbf{x} de U . On peut écrire le développement de Taylor à l'ordre 2 :

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \frac{(\mathbf{h} \cdot \nabla)^2 f(\mathbf{x})}{2} + R_2(\mathbf{h}),$$

où $\lim_{\mathbf{h} \rightarrow 0} \frac{R_2(\mathbf{h})}{\|\mathbf{h}\|^2} = 0$. Noter qu'en vertu du Théorème 20, $D_i D_j = D_j D_i$. On appelle

$$q_{\mathbf{x}}(\mathbf{h}) = \frac{(\mathbf{h} \cdot \nabla)^2 f(\mathbf{x})}{2} = \frac{1}{2} \sum_{i=1}^m h_i^2 D_i^2 f(\mathbf{x}) + \sum_{1 \leq i < j \leq m} h_i h_j D_i D_j f(\mathbf{x}),$$

la forme quadratique de la fonction f au point critique \mathbf{x} . Dans ce cas, la matrice symétrique associée à q où les coefficients sont multipliés par 2, revêt une importance particulière, et est appelée la *matrice hessienne de f au point \mathbf{x}* , notée $Hf(\mathbf{x})$. En vertu de la section précédente la matrice hessienne de f au point \mathbf{x} est une matrice de taille $m \times m$, définie par :

$$Hf(\mathbf{x})_{ij} = \begin{cases} D_i^2 f(\mathbf{x}) & \text{si } i = j \\ D_i D_j f(\mathbf{x}) & \text{si } i < j. \end{cases}$$

Voici le théorème qui permet la classification des points critiques.

Théorème 27 *Soit f une fonction définie sur un ouvert U de \mathbb{R}^m , qui est de classe C^3 dans le voisinage d'un point critique \mathbf{x} de U .*

1. *Si la forme quadratique $q_{\mathbf{x}}$ est définie positive, f admet un minimum local strict en \mathbf{x} .*
2. *Si la forme quadratique $q_{\mathbf{x}}$ est définie négative, f admet un maximum local strict en \mathbf{x} .*
3. *Si la forme quadratique $q_{\mathbf{x}}$ est non-semi-définie, alors f n'admet ni maximum, ni minimum en \mathbf{x} .*

Preuve:

L'idée de la preuve pour les deux premiers points est la suivante (il manque des détails techniques). Si q est définie positive, alors pour \mathbf{h} suffisamment petit, $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) > 0$, de sorte que \mathbf{x} est un minimum local strict. En effet, on peut montrer que le reste est négligeable par rapport à la forme quadratique, ainsi localement, le comportement de la fonction est le même que celui de la forme quadratique.

Pour le troisième point, on montre la contraposée : si f admet un minimum (maximum) local (pas forcément strict), alors la forme quadratique $q_{\mathbf{x}}$ est semi-définie positive (négative). \square

Remarque 3.6

1. Il est IMPORTANT de noter que si la forme quadratique est semi-définie positive, ou semi-définie négative, le Théorème ci-dessus ne permet pas de conclure. En effet, tout peut arriver comme l'illustre l'Exemple 3.17 ci-dessous. Remarquer qu'au niveau de la matrice hessienne, ce cas se présente lorsque le déterminant est 0, et que les valeurs propres sont soit toutes ≥ 0 , soit toutes ≤ 0 .
2. Un point critique qui n'est ni un minimum, ni un maximum est appelé un *point selle*, pour des raisons géométriques expliquées ci-dessous, et illustrées dans l'Exemple 3.16.

3. Voici une description géométrique de ce qui se passe. Comme la matrice hessienne est symétrique, elle peut être diagonalisée dans une base orthogonale. De sorte que l'on peut écrire

$$q_{\mathbf{x}}(\mathbf{h}) = \lambda_1 a_1^2 + \cdots + \lambda_m a_m^2,$$

où $\mathbf{h} = \sum_{i=1}^m a_i \mathbf{v}_i$ est décomposé dans une base orthogonale, et $\lambda_1, \dots, \lambda_m$ sont les valeurs propres de $Hf(\mathbf{x})$. De plus, dans les cas sus-mentionnés la forme quadratique approche bien la fonction f , de sorte que q décrit le comportement de f au voisinage de \mathbf{x} . Dans le premier cas les valeurs propres sont positives, ainsi on voit bien que q admet un minimum en $\mathbf{0}$. On raisonne de manière analogue dans le deuxième cas. Dans le troisième cas, si on considère le sous-espace vectoriel engendré par les vecteurs propres des valeurs propres positives, alors f a un minimum local sur ce sous-espace. De manière analogue, f a un maximum local sur le sous-espace engendré par les vecteurs propres des valeurs propres négatives. Cela donne une idée géométrique de ce qui se passe pour les points selles.

Exemple 3.15

Lorsque $m = 1$, la matrice hessienne est simplement la dérivée seconde, de sorte que l'on retrouve le critère déjà connu pour les fonctions d'une variable.

Lorsque $m = 2$, la matrice hessienne $Hf(\mathbf{x})$ est donnée par :

$$\begin{pmatrix} D_1^2 f(\mathbf{x}) & D_1 D_2 f(\mathbf{x}) \\ D_1 D_2 f(\mathbf{x}) & D_2^2 f(\mathbf{x}) \end{pmatrix}$$

Donc en utilisant le critère 4 du Lemme 25, et le fait qu'une forme quadratique q est définie négative si $-q$ est définie positive, on déduit le critère suivant. Une fonction f de deux variables qui est de classe \mathcal{C}^3 dans un voisinage d'un point critique \mathbf{x} , admet

1. un minimum local si $|Hf(\mathbf{x})| > 0$ et $D_1^2 f(\mathbf{x}) > 0$,
2. un maximum local si $|Hf(\mathbf{x})| > 0$, et $D_1^2 f(\mathbf{x}) < 0$,
3. un point selle si $|Hf(\mathbf{x})| < 0$,
4. si $|Hf(\mathbf{x})| = 0$, on ne peut pas conclure.

Exemple 3.16 On suppose que $m = 2$, et que f est une fonction de classe \mathcal{C}^3 au voisinage d'un point critique \mathbf{x} . Soit $q_{\mathbf{x}}$ la forme quadratique associée à f au point \mathbf{x} . Voici quelques exemples pour $q_{\mathbf{x}}$.

1. Si $q_{\mathbf{x}}(h_1, h_2) = h_1^2 + h_2^2$, alors la forme quadratique est définie positive, et f admet un minimum local au point \mathbf{x} . Si $q_{\mathbf{x}}(h_1, h_2) = -h_1^2 - h_2^2$, alors la forme quadratique est définie négative, et f admet un maximum local au point \mathbf{x} .
2. Si $q_{\mathbf{x}}(h_1, h_2) = h_1^2 - h_2^2$, alors la forme quadratique est non-semi-définie, car $q(2, 1) > 0$ et $q(1, 2) < 0$, donc f admet un point selle en \mathbf{x} .

3. Si $q_{\mathbf{x}}(h_1, h_2) = h_2^2$ ou $q(h_1, h_2) = 0$, le Théorème 27 ne s'applique pas, donc on ne peut rien dire sur le point critique de f .

Exemple 3.17 Soit $f(x, y) = x^2 - 2xy + y^2 + x^4 + y^4$, et $g(x, y) = x^2 - 2xy + y^2 - x^4 - y^4$. Pour ces deux fonctions la forme quadratique au point critique $(0, 0)$ est :

$$q(h_1, h_2) = (h_1 - h_2)^2,$$

qui est semi-définie positive, donc le Théorème 27 ne permet pas de conclure. En effet, f peut s'écrire $f(x, y) = (x - y)^2 + x^4 + y^4$, de sorte que $(0, 0)$ est un minimum local pour f . Par contre pour g , on a :

$$g(t, t) = -2t^4, \quad g(t, -t) = 2t^2(2 - t^2),$$

de sorte que $(0, 0)$ est un point selle de g . Ceci illustre le fait que lorsque la forme quadratique est semi-définie positive, tout peut arriver.

3.8.4 Multiplicateurs de Lagrange

Dans cette section nous allons étudier les points critiques d'un problème d'optimisation sous contrainte.

La contrainte est déterminée par une surface S définie de la manière suivante.

• **La surface en question**

Soit $g : U \rightarrow \mathbb{R}$, une fonction de classe \mathcal{C}^1 sur un ouvert U de \mathbb{R}^m , et soit S la surface définie par :

$$S = \{\mathbf{x} \in U : g(\mathbf{x}) = 0\}.$$

Dans la suite, on suppose que $\nabla g(\mathbf{x}) \neq \mathbf{0}$, pour tout $\mathbf{x} \in U \cap S$.

Exemple 3.18 Soit $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ la fonction définie par $g(x, y, z) = x^2 + y^2 + z^2 - 1$, alors la surface S correspondante est la sphère de rayon 1 dans \mathbb{R}^3 .

Remarque 3.7 Dans la Section 3.5, on a défini le plan tangent en un point \mathbf{x} de S , noté $\mathcal{T}_{\mathbf{x}}$, comme le plan passant par \mathbf{x} , et orthogonal au gradient en ce point, i.e.

$$\mathcal{T}_{\mathbf{x}} = \{\mathbf{y} \in \mathbb{R}^m : (\mathbf{y} - \mathbf{x}) \cdot \nabla g(\mathbf{x}) = 0\}.$$

Sous les hypothèses ci-dessus, on peut montrer que le plan tangent $\mathcal{T}_{\mathbf{x}}$ est aussi le plan engendré par les vecteurs tangents aux courbes sur la surface S passant par \mathbf{x} . Nous avons démontré une des inclusions à la Section 3.5; nous ne donnons pas la démonstration de l'autre inclusion.

• **Le problème d'optimisation**

Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 sur un ouvert U de \mathbb{R}^m (le même que celui de la fonction g). On souhaite trouver l'ensemble des points $\mathbf{x} \in S$ tels que $f(\mathbf{x})$ soit un extremum local (maximum local ou minimum local) de f sur la surface S , c'est à dire, on cherche tous les points $\mathbf{x} \in U$ tels que $g(\mathbf{x}) = 0$, et tel qu'il existe une boule ouverte $B \subset U$ vérifiant :

$$f(\mathbf{x}) \geq f(\mathbf{y}), \text{ pour tout } \mathbf{y} \in B \cap S, \text{ ou,}$$

$$f(\mathbf{x}) \leq f(\mathbf{y}), \text{ pour tout } \mathbf{y} \in B \cap S.$$

Si \mathbf{x} satisfait une des deux conditions, on dit que la fonction f soumise à la contrainte g admet un *extremum local au point* \mathbf{x} . Elle admet un *extremum strict* si les inégalités ci-dessus sont strictes.

Théorème 28 Soit $g : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 sur un ouvert U de \mathbb{R}^m , et soit S la surface correspondante. On suppose que $\nabla g(\mathbf{x}) \neq \mathbf{0}, \forall \mathbf{x} \in U \cap S$. Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 sur le même ouvert U . Si la fonction f soumise à la contrainte g admet un *extremum local au point* \mathbf{x} , alors il existe un nombre λ , tel que :

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}). \quad (3.4)$$

Avant de donner une esquisse de démonstration, voici quelques définitions. Un point $\mathbf{x} \in U \cap S$ pour lequel il existe λ tel que $\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x})$ s'appelle un *point critique* pour le problème d'optimisation de la fonction f sous la contrainte g . Le nombre λ correspondant s'appelle le *multiplicateur de Lagrange* en \mathbf{x} .

Preuve:

Supposons que f admette un extremum local au point \mathbf{x} , et supposons que ce soit un maximum local. Soit $\gamma : I \rightarrow S$ une courbe sur la surface S qui passe par \mathbf{x} , i.e. il existe $t_0 \in I$ tel que $\gamma(t_0) = \mathbf{x}$. Alors la fonction composée $f \circ \gamma : I \rightarrow \mathbb{R}$ admet un maximum en t_0 , de sorte que sa dérivée en t_0 s'annule, i.e.

$$\nabla f(\mathbf{x}) \cdot \gamma'(t_0) = 0.$$

Ainsi $\nabla f(\mathbf{x})$ est orthogonal à toute courbe sur la surface S passant par \mathbf{x} . En utilisant la Remarque 3.7, cela signifie que $\nabla f(\mathbf{x})$ est orthogonal au plan tangent $\mathcal{T}_{\mathbf{x}}$ à S au point \mathbf{x} . D'autre part, on peut montrer que l'espace orthogonal au plan tangent est de dimension 1, de sorte qu'il existe λ tel que,

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}).$$

□

Nous souhaitons maintenant classifier les points critiques du problème d'optimisation sous contrainte. Pour cela, nous introduisons la *fonction auxiliaire* $\mathcal{L} : U \rightarrow \mathbb{R}$ définie par :

$$\mathcal{L}(\mathbf{x}) = f(\mathbf{x}) - \lambda g(\mathbf{x}).$$

Remarquer que si \mathbf{x} est un point critique du problème d'optimisation de la fonction f sous la contrainte g , alors \mathbf{x} est un point critique de la fonction \mathcal{L} .

Supposons de plus que f et g soient de classe \mathcal{C}^3 . La *forme quadratique* du problème d'optimisation de la fonction f soumise à la contrainte g au point \mathbf{x} est, par définition, la forme quadratique de la fonction \mathcal{L} au point critique \mathbf{x} , et est notée $q_{\mathbf{x}}^{\mathcal{L}}$:

$$q_{\mathbf{x}}^{\mathcal{L}}(\mathbf{h}) = \frac{(\mathbf{h} \cdot \nabla)^2 \mathcal{L}(\mathbf{x})}{2} = \frac{1}{2} \sum_{i=1}^m h_i^2 D_i^2 \mathcal{L}(\mathbf{x}) + \sum_{1 \leq i < j \leq m} h_i h_j D_i D_j \mathcal{L}(\mathbf{x}),$$

On a alors le théorème suivant :

Théorème 29 Soit $g : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^3 sur un ouvert U de \mathbb{R}^m , et soit S la surface correspondante. Supposons que $\nabla g(\mathbf{x}) \neq \mathbf{0}$, pour tout $\mathbf{x} \in S \cap U$. Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^3 , et \mathbf{x} un point critique du problème d'optimisation de la fonction f soumise à la contrainte g . Soit \mathcal{L} la fonction auxiliaire ci dessus, et $q_{\mathbf{x}}^{\mathcal{L}}$ la forme quadratique associée. Alors f soumise à la contrainte g admet :

1. un minimum local si $q_{\mathbf{x}}^{\mathcal{L}}$ est définie positive sur le plan tangent $\mathcal{T}_{\mathbf{x}}$ à S au point \mathbf{x} .
2. un maximum local si $q_{\mathbf{x}}^{\mathcal{L}}$ est définie négative sur le plan tangent $\mathcal{T}_{\mathbf{x}}$ à S au point \mathbf{x} .
3. un point selle si $q_{\mathbf{x}}^{\mathcal{L}}$ est non semi-définie sur le plan tangent $\mathcal{T}_{\mathbf{x}}$ à S au point \mathbf{x} .

Preuve:

Soit \mathbf{x} un point critique du problème d'optimisation sous contrainte. Nous souhaitons déterminer le signe de $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$, lorsque $\mathbf{x}, \mathbf{x} + \mathbf{h} \in S$. Par hypothèse, la fonction auxiliaire \mathcal{L} est de classe \mathcal{C}^3 . Effectuons son développement de Taylor à l'ordre 2 au point \mathbf{x} :

$$\mathcal{L}(\mathbf{x} + \mathbf{h}) = \mathcal{L}(\mathbf{x}) + q_{\mathbf{x}}^{\mathcal{L}}(\mathbf{h}) + R_2(\mathbf{h}).$$

Si \mathbf{x} et $\mathbf{x} + \mathbf{h} \in S$, alors $\mathcal{L}(\mathbf{x} + \mathbf{h}) = f(\mathbf{x} + \mathbf{h})$ et $\mathcal{L}(\mathbf{x}) = f(\mathbf{x})$, de sorte que :

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + q_{\mathbf{x}}^{\mathcal{L}}(\mathbf{h}) + R_2(\mathbf{h}).$$

De plus, si $\mathbf{x} + \mathbf{h} \in S$, et \mathbf{h} est suffisamment petit, le point $\mathbf{x} + \mathbf{h}$ est proche du plan tangent $\mathcal{T}_{\mathbf{x}}$ de S au point \mathbf{x} . On peut montrer que le comportement de $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$ est décrit par le comportement de la forme quadratique restreinte au plan tangent $\mathcal{T}_{\mathbf{x}}$, dans les limite données dans le théorème.

□

Exemple 3.19 Voici un exemple pour illustrer la méthode. On souhaite minimiser la surface totale d'un parallélépipède qui a pour volume 1000. Soit $f : U = (\mathbb{R}^+)^3 \rightarrow \mathbb{R}$ définie par

$$f(x, y, z) = 2(xy + xz + yz),$$

et soit $g : U \rightarrow \mathbb{R}$ définie par $g(x, y, z) = xyz - 1000$. Alors résoudre notre problème revient à trouver le minimum de la fonction f sous la contrainte g . Remarquons que U est un ouvert, et que f et g sont de classe \mathcal{C}^3 sur U . De plus $\nabla g(x, y, z) = (yz, xz, xy) \neq \mathbf{0}$ si $\mathbf{x} \in S$. Ainsi, si f admet un extremum au point \mathbf{x} , alors \mathbf{x} est un point critique du problème d'optimisation. Cherchons ces points critiques. On veut résoudre :

$$\begin{aligned} \nabla f(\mathbf{x}) &= \lambda \nabla g(\mathbf{x}) \\ g(\mathbf{x}) &= 0. \end{aligned}$$

Ceci est équivalent à résoudre :

$$\begin{aligned} 2y + 2z &= \lambda yz \\ 2x + 2z &= \lambda xz \\ 2x + 2y &= \lambda xy \\ xyz &= 1000. \end{aligned}$$

En multipliant la première équation par x , la deuxième par y et la troisième par z , on obtient :

$$xy + xz = xy + yz = xz + yz = 500\lambda.$$

De plus, comme $\mathbf{x} \in U$, on a $x, y, z \neq 0$. De là on déduit que $x = y = z$. En utilisant la dernière équation, on conclut qu'il n'y a qu'un seul point critique pour le problème d'optimisation : $(x, y, z) = (10, 10, 10)$, et que le multiplicateur de Lagrange correspondant est $\lambda = \frac{2}{5}$. La fonction auxiliaire est :

$$\mathcal{L}(x, y, z) = 2xy + 2xz + 2yz - \frac{2}{5}xyz + 400.$$

Après calculs, on montre que la forme quadratique de \mathcal{L} au point $(10, 10, 10)$ est :

$$q(x, y, z) = -2xy - 2xz - 2yz.$$

Cette forme est non-semi-définie, mais ce qui nous intéresse est son comportement sur le plan tangent $\mathcal{T}_{\mathbf{x}}$ à S au point critique $(x, y, z) = (10, 10, 10)$. Le plan tangent $\mathcal{T}_{\mathbf{x}}$ est par définition le plan orthogonal au gradient en ce point, i.e. c'est le plan orthogonal au point $\nabla g(\mathbf{x}) = (100, 100, 100)$. Il est engendré par les vecteurs :

$$\mathbf{v}_1 = (1, -1, 0), \quad \mathbf{v}_2 = (1, 0, -1).$$

Ainsi, si $\mathbf{v} \in \mathcal{T}_{\mathbf{x}}$, on peut écrire $\mathbf{v} = s\mathbf{v}_1 + t\mathbf{v}_2 = (s + t, -s, -t)$. En remplaçant, on déduit :

$$q(\mathbf{v}) = 2s^2 + st + 2t^2.$$

Cette forme est définie positive, de sorte que l'on a un minimum local en ce point.

3.8.5 Maximums et minimums absolus

Toute la théorie développée jusqu'à maintenant ne donne que des extremas locaux. Quels outils avons-nous pour trouver des extremas globaux ?

On rappelle les faits suivants du chapitre 1. Un sous-ensemble K de \mathbb{R}^m est compact, ssi il est fermé et borné. De plus, si $f : K \rightarrow \mathbb{R}$ est une fonction continue sur un compact K de \mathbb{R}^m , alors f admet un maximum et un minimum sur K . Si f est suffisamment régulière, on utilisera la stratégie suivante pour les trouver.

1. Chercher les extremums locaux dans $\text{int}(K)$ qui est ouvert avec la théorie élaborée à la Section 3.8.4.
2. Chercher les extremums locaux sur le bord ∂K . Ce problème peut parfois se ramener à une question de multiplicateurs de Lagrange.
3. On sait que les extremums globaux se trouvent parmi ces extremums locaux. Comparer la valeur de la fonction en ces différents points, et en déduire quels sont les extremums globaux.

Chapitre 4

Applications de \mathbb{R}^m dans \mathbb{R}^n

Rappelons la définition d'une application (voir aussi Chapitre 1, Section 1.3). Soit D un sous-ensemble de \mathbb{R}^m . Une *application* F de D dans \mathbb{R}^n , notée $F : D \rightarrow \mathbb{R}^n$, est une règle qui à tout point \mathbf{x} de D associe un point $F(\mathbf{x})$ dans \mathbb{R}^n . On peut représenter F par ses coordonnées, c'est-à-dire,

$$F(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{pmatrix}$$

où $F_i : \mathbb{R}^m \rightarrow \mathbb{R}$. On utilisera la terminologie *fonction* lorsque l'espace d'arrivée est \mathbb{R} , c'est-à-dire lorsque $n = 1$.

Remarque : Dans ce chapitre, il est important de noter que les points \mathbf{x} sont des vecteurs colonnes de \mathbb{R}^m , et les images par $F(\mathbf{x})$ de ces points sont des vecteurs colonnes de \mathbb{R}^n . Néanmoins, afin d'économiser de l'espace, nous les écrirons sous forme de vecteurs lignes dans les exemples.

Un des buts principaux de ce chapitre est d'approcher une telle application, qui peut être compliquée, par une application linéaire. Ainsi, la première section consiste en un rappel sur les applications linéaires.

4.1 Applications linéaires

Soit L une application de \mathbb{R}^m dans \mathbb{R}^n . Alors L est dite *linéaire* si elle vérifie les propriétés suivantes :

1. Pour tout $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$,

$$L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}).$$

2. Pour tout $a \in \mathbb{R}$,

$$L(a\mathbf{x}) = aL(\mathbf{x}).$$

De manière équivalente, l'application L est linéaire, ssi pour tout $a, b \in \mathbb{R}$, et tout $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$:

$$L(a\mathbf{x} + b\mathbf{y}) = aL(\mathbf{x}) + bL(\mathbf{y}).$$

Exemple 4.1 Soit $a \in \mathbb{R}$, alors la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, définie par $f(x) = ax$ est linéaire. Réciproquement, soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction linéaire, alors $f(x) = f(x.1) = xf(1)$. Ainsi $f(x) = ax$, où $a = f(1)$.

Exemple 4.2

- L'application identité $\text{Id} : \mathbb{R}^m \rightarrow \mathbb{R}^m$, définie par $\text{Id}(\mathbf{x}) = \mathbf{x}$ est linéaire. En coordonnées, l'application identité s'écrit : $\text{Id}_i(\mathbf{x}) = x_i$.
- Soit $\mathbf{a} \in \mathbb{R}^n$. Alors l'application $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$, définie par $F(\mathbf{x}) = \mathbf{a}$, est linéaire ssi $\mathbf{a} = \mathbf{0}$.

Exemple 4.3 Soit $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ l'application projection orthogonale définie par $F(x, y, z) = (x, y)$. Alors, F est linéaire. En coordonnées, l'application F s'écrit :

$$\begin{aligned} F_1(x, y, z) &= 1.x + 0.y + 0.z \\ F_2(x, y, z) &= 0.x + 1.y + 0.z. \end{aligned}$$

Voici le Théorème fondamental qui relie les applications linéaires aux matrices.

Théorème 30 Soit $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ une application. Alors F est linéaire, ssi il existe une matrice A de taille $n \times m$, telle que $F(\mathbf{x}) = A\mathbf{x}$ pour tout $\mathbf{x} \in \mathbb{R}^m$, i.e.

$$\begin{aligned} F_1(\mathbf{x}) &= a_{11}x_1 + \cdots + a_{1m}x_m, \\ &\vdots \\ F_n(\mathbf{x}) &= a_{n1}x_1 + \cdots + a_{nm}x_m. \end{aligned}$$

Alors A est la matrice dont la j -ème colonne est donnée par $F(\mathbf{e}_j)$, où $\mathbf{e}_j^t = (0, \dots, 1, \dots, 0)$ est le j -ème vecteur unité.

Preuve:

Soit A une matrice de taille $n \times m$, et soit $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ l'application définie par $F(\mathbf{x}) = A\mathbf{x}$. Soit $a, b \in \mathbb{R}$, et $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$. Alors, par les propriétés des opérations sur les matrices, on a :

$$F(a\mathbf{x} + b\mathbf{y}) = A(a\mathbf{x} + b\mathbf{y}) = aA\mathbf{x} + bA\mathbf{y} = aF(\mathbf{x}) + bF(\mathbf{y}),$$

ainsi F est linéaire, et $F(\mathbf{e}_j) = A\mathbf{e}_j$ est bien la j -ème colonne de la matrice A .

Réciproquement, soit F une application linéaire. On définit pour tout $j = 1, \dots, m$:

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} := F(\mathbf{e}_j),$$

et la matrice $A = (a_{ij})$. Soit maintenant $\mathbf{x} = (x_1, \dots, x_m)^t \in \mathbb{R}^m$. En décomposant \mathbf{x} dans la base canonique, on obtient :

$$\mathbf{x} = \sum_{j=1}^m x_j \mathbf{e}_j.$$

Ainsi, en utilisant la linéarité de l'application F , on a :

$$\begin{aligned} F(\mathbf{x}) &= F\left(\sum_{j=1}^m x_j \mathbf{e}_j\right), \\ &= \sum_{j=1}^m x_j F(\mathbf{e}_j), \\ &= \sum_{j=1}^m x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}, \\ &= \begin{pmatrix} \sum_{j=1}^m a_{1j} x_j \\ \vdots \\ \sum_{j=1}^m a_{nj} x_j \end{pmatrix}, \\ &= A\mathbf{x}. \end{aligned}$$

□

Exemple 4.4 Soit $f : \mathbb{R}^m \rightarrow \mathbb{R}$ une fonction linéaire. Alors la matrice A donnée par le Théorème 30 est de la forme :

$$A = (a_{11} \ a_{12} \ \cdots \ a_{1m}),$$

$$\text{ainsi } f(x_1, \dots, x_m) = (a_{11} \ \cdots \ a_{1m}) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = a_{11}x_1 + \cdots + a_{1m}x_m.$$

Exemple 4.5 Soit $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ une courbe linéaire. Alors la matrice A donnée par le Théorème 30 est de la forme :

$$A = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix},$$

de sorte que $\gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{pmatrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} t = \begin{pmatrix} a_{11}t \\ \vdots \\ a_{n1}t \end{pmatrix}$. Ainsi, la trace de la courbe est la droite dans \mathbb{R}^n , passant par les points $\mathbf{0} = (0, \dots, 0)$, et $\mathbf{a} = (a_{11}, \dots, a_{n1})$.

Pour tous les théorèmes classiques sur les applications linéaires, on se réfère à un livre d'algèbre linéaire.

4.2 Exemples

Les applications linéaires de la section précédente sont des exemples d'applications. En voici quelques autres.

Exemple 4.6

1. $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par : $F(x, y) = (xy, \sin x, x^2y)$.
2. Soit U un ouvert de \mathbb{R}^m . Alors un *champ de vecteurs sur U* est une application $F : U \rightarrow \mathbb{R}^m$. Un champ de vecteurs est interprété en physique comme un champ de forces.
3. $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par : $F(r, \theta) = (r \cos \theta, r \sin \theta)$, appelée application *coordonnées polaires*.
4. $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $F(r, \theta, h) = (r \cos \theta, r \sin \theta, h)$, appelée application *coordonnées cylindriques*.
5. $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $F(r, \varphi, \theta) = (r \sin \varphi \cos \theta, r \sin \varphi \sin \theta, r \cos \varphi)$, appelée application *coordonnées sphériques*.

4.3 La matrice jacobienne

Dans cette section, nous définissons les dérivées partielles des applications et la matrice jacobienne. C'est une généralisation de ce que nous avons fait pour les fonctions de m variables au Chapitre 3.

Soit une application $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$, et soit $F_1, \dots, F_n : \mathbb{R}^m \rightarrow \mathbb{R}$ ses fonctions coordonnées, c'est à dire :

$$F(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{pmatrix}.$$

Remarquer que pour tout $i = 1, \dots, n$, F_i est une fonction de m variables comme au Chapitre 3. Ainsi nous connaissons déjà la notion de dérivée partielle pour chacune des fonctions coordonnées F_i .

Supposons que pour chacune des fonctions F_i , toutes les dérivées partielles existent. Alors, l'ensemble de ces dérivées partielles est représenté sous forme d'une matrice de taille $n \times m$, appelée *matrice jacobienne*, notée $J_F(\mathbf{x})$, de la manière suivante :

$$J_F(\mathbf{x}) = (D_j F_i(\mathbf{x})) = \left(\frac{\partial F_i}{\partial x_j} \right) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}.$$

Exemple 4.7 Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ l'application définie par :

$$F(x, y) = (xy, \sin x, x^2y).$$

Alors, toutes les coordonnées de F admettent toutes leurs dérivées partielles en tout point $(x, y) \in \mathbb{R}^2$. Ainsi, la matrice jacobienne de F au point (x, y) existe, et est donnée par :

$$J_F(x, y) = \begin{pmatrix} y & x \\ \cos x & 0 \\ 2xy & x^2 \end{pmatrix}$$

Voici une définition importante qui va intervenir dans le chapitre sur l'intégration. Soit U un ouvert de \mathbb{R}^m , et $F : U \rightarrow \mathbb{R}^m$ une application vers un espace de même dimension. On suppose que toutes les dérivées partielles de toutes les coordonnées de l'application F existent. Alors la matrice jacobienne est de taille $m \times m$. Le déterminant de la matrice jacobienne $\det J_F(\mathbf{x})$ est appelé *déterminant jacobien* ou plus simplement *jacobien*.

Exemple 4.8 Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'application définie par :

$$F(x, y) = (x^2 + y^2, e^{xy}).$$

Alors la matrice jacobienne de F au point (x, y) est :

$$J_F(x, y) = \begin{pmatrix} 2x & 2y \\ ye^{xy} & xe^{xy} \end{pmatrix},$$

et le jacobien est, $\det(J_F(x, y)) = 2x^2e^{xy} - 2y^2e^{xy}$.

4.4 Dérivabilité

Après s'être échauffé avec le cas des fonctions de \mathbb{R}^m dans \mathbb{R} au Chapitre 3, nous sommes maintenant prêts à définir la notion de dérivabilité pour les applications de \mathbb{R}^m dans \mathbb{R}^n .

Soit F une application d'un ouvert $U \subset \mathbb{R}^m$ dans \mathbb{R}^n . On dit que F est *dérivable au point* $\mathbf{x} \in U$, s'il existe une application linéaire $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$ et une application $G : \mathbb{R}^m \rightarrow \mathbb{R}^n$, définie pour tous les vecteurs \mathbf{h} suffisamment petits, telles que :

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}) + L(\mathbf{h}) + \|\mathbf{h}\|G(\mathbf{h}), \quad (4.1)$$

$$\lim_{\|\mathbf{h}\| \rightarrow 0} G(\mathbf{h}) = \mathbf{0}. \quad (4.2)$$

Lorsqu'elle existe, l'application linéaire L ci-dessus est appelée la *différentielle* de F au point \mathbf{x} , et est notée $dF_{\mathbf{x}}$. La matrice qui la représente, notée $F'(\mathbf{x})$ ou $DF(\mathbf{x})$, est appelée la *dérivée* de F au point \mathbf{x} . C'est-à-dire que $F'(\mathbf{x})$ est l'unique matrice donnée par le Théorème 30, telle que :

$$dF_{\mathbf{x}}(\mathbf{y}) = F'(\mathbf{x})\mathbf{y}, \quad \forall \mathbf{y} \in \mathbb{R}^m.$$

La condition (4.1) signifie que L approche F pour des points proches de \mathbf{x} , à un facteur de petite amplitude près. Autrement dit, nous avons localement approché l'application F , qui peut-être très compliquée, par une application linéaire.

Remarque 4.1 On dit qu'une application $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ définie pour des petites valeurs de \mathbf{h} , est $o(\mathbf{h})$ si :

$$\lim_{\|\mathbf{h}\| \rightarrow 0} \frac{\varphi(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}.$$

En utilisant cette nouvelle définition, on peut écrire l'équation (4.1) sous la forme

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}) + L(\mathbf{h}) + o(\mathbf{h}).$$

Exemple 4.9 Si $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ est constante, c'est à dire qu'il existe $\mathbf{a} \in \mathbb{R}^n$ tel que $F(\mathbf{x}) = \mathbf{a}$ pour tout $\mathbf{x} \in \mathbb{R}^m$, alors F est dérivable partout, et la différentielle est l'application nulle, i.e. pour tout $\mathbf{x} \in \mathbb{R}^m$, $dF_{\mathbf{x}} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ satisfait,

$$dF_{\mathbf{x}}(\mathbf{y}) = \mathbf{0}, \quad \forall \mathbf{y} \in \mathbb{R}^m.$$

En effet, pour tout $\mathbf{h} \in \mathbb{R}^m$, on a :

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}).$$

Dans ce cas, l'application G est identiquement nulle.

Exemple 4.10 Si $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ est une application linéaire, alors F est dérivable partout et la différentielle est l'application F , c'est-à-dire $dF_{\mathbf{x}} = F$ pour tout $\mathbf{x} \in \mathbb{R}^m$. En effet, pour tout $\mathbf{x} \in \mathbb{R}^m$, et pour tout $\mathbf{h} \in \mathbb{R}^m$, on a par linéarité de F :

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}) + F(\mathbf{h}).$$

Dans ce cas, l'application G est à nouveau identiquement nulle. On aurait pu deviner ce résultat intuitivement : on veut approcher F par une application linéaire, mais comme F est linéaire, on peut l'approcher par elle-même, et il n'y aura pas de terme d'erreur (l'application G est nulle).

Le théorème suivant relie la dérivabilité de l'application F et celle de ses coordonnées.

Théorème 31 Soit U un ouvert de \mathbb{R}^m . L'application $F : U \rightarrow \mathbb{R}^n$ est dérivable au point \mathbf{x} , ssi chacune de ses coordonnées F_1, \dots, F_n l'est.

Preuve:

Utiliser les définitions. □

Les Théorèmes 32, 33, 34 ci-dessous, se déduisent du Théorème 31 qui permet de se restreindre à l'étude des coordonnées (qui sont des fonctions de plusieurs variables), et des Théorèmes 16, 17, 18 qui sont les résultats analogues pour les fonctions de plusieurs variables.

Théorème 32 Soit U un ouvert de \mathbb{R}^m , et $F : U \rightarrow \mathbb{R}^n$ une application. Si F est dérivable au point \mathbf{x} , alors la différentielle de F au point \mathbf{x} est unique.

4.5 Dérivée et matrice jacobienne

Cette section est l'analogue de la Section 3.4. Elle permet de relier la dérivabilité et la matrice jacobienne.

Théorème 33 Soit une application $F : U \rightarrow \mathbb{R}^n$, où U est un ouvert de \mathbb{R}^m . Si F est dérivable au point \mathbf{x} , alors toutes les dérivées partielles de toutes les composantes de F existent au point \mathbf{x} , et la dérivée est la matrice jacobienne au point \mathbf{x} , i.e.

$$F'(\mathbf{x}) = J_F(\mathbf{x}).$$

Exemple 4.11

- Soit $f : U \rightarrow \mathbb{R}$ une fonction définie sur un ouvert U de \mathbb{R}^m . Si f est dérivable au point $\mathbf{x} \in U$, alors la dérivée au point \mathbf{x} est :

$$f'(\mathbf{x}) = (D_1f(\mathbf{x}), \dots, D_mf(\mathbf{x})) = \nabla f(\mathbf{x})^t.$$

Remarque : Maintenant que nous différencions les vecteurs lignes et les vecteurs colonnes, nous introduisons la convention usuelle qui consiste à dire que le gradient est un vecteur *colonne*, ainsi $f'(\mathbf{x})$ est la transposée du gradient $\nabla f(\mathbf{x})$.

- Soit $\gamma : I \rightarrow \mathbb{R}^n$ une courbe définie sur un intervalle ouvert. Si γ est dérivable au point $t \in I$, alors la dérivée au point t est :

$$\gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_n(t) \end{pmatrix}.$$

Théorème 34 Soit une application $F : U \rightarrow \mathbb{R}^n$, où U est un ouvert de \mathbb{R}^m . Si toutes les dérivées partielles de toutes les coordonnées de l'application F existent, et qu'elles sont de plus continues, alors F est dérivable en tout point de U .

Soit $F : U \rightarrow \mathbb{R}^n$ une application, où U est un ouvert de \mathbb{R}^m . L'application F est dite de *classe \mathcal{C}^k* sur U , si toutes les dérivées partielles d'ordre au plus k de toutes les coordonnées existent et sont continues.

En particulier, le Théorème 34 dit que si l'application F est de classe \mathcal{C}^1 sur U , alors F est dérivable en tout point de U .

Remarque 4.2 Soit $f : \mathbb{R}^m \rightarrow \mathbb{R}$ une fonction dérivable. Maintenant que nous avons une définition précise de différentielle et de dérivées partielles, nous pouvons donner un sens précis à la formule connue suivante :

$$df = \frac{\partial f}{\partial x_1} dx_1 + \dots + \frac{\partial f}{\partial x_m} dx_m.$$

Soit $\text{Id} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ l'application identité. Soient x_1, \dots, x_m ses coordonnées, i.e.

$$\forall \mathbf{p} = (p_1, \dots, p_m) \in \mathbb{R}^m, \quad \text{Id}(\mathbf{p}) = (x_1(\mathbf{p}), \dots, x_m(\mathbf{p})) = (p_1, \dots, p_m).$$

Comme x_i est une fonction linéaire, sa différentielle est elle-même, autrement dit pour tout $\mathbf{a} \in \mathbb{R}^m$, $\mathbf{p} \in \mathbb{R}^m$:

$$dx_{i,\mathbf{a}}(\mathbf{p}) = x_i(\mathbf{p}) = p_i. \quad (4.3)$$

De plus, comme f est une fonction dérivable de m variables, sa dérivée au point \mathbf{a} est le gradient de f au point \mathbf{a} , c'est-à-dire :

$$\begin{aligned} df_{\mathbf{a}}(\mathbf{p}) &= f'(\mathbf{a})\mathbf{p}, \\ &= \nabla f(\mathbf{a})^t \mathbf{p} = \nabla f(\mathbf{a}) \cdot \mathbf{p}, \\ &= \sum_{i=1}^m \frac{\partial f}{\partial x_i} p_i, \\ &= \sum_{i=1}^m \frac{\partial f}{\partial x_i} dx_{i,\mathbf{a}}, \quad \text{par (4.3).} \end{aligned}$$

Ce qui est exactement la formule ci-dessus où l'on a omis d'écrire l'indice \mathbf{a} .

4.6 Dérivation des applications composées

Voici la généralisation "ultime" de la dérivation des fonctions composées.

Théorème 35 *Soient U, V des ouverts de \mathbb{R}^m et \mathbb{R}^n respectivement. Soient $F : U \rightarrow \mathbb{R}^n$ et $G : V \rightarrow \mathbb{R}^k$ deux applications, telles que $F(U) \subset V$. Soit $\mathbf{x} \in U$. Si l'application F est dérivable au point \mathbf{x} , et l'application G est dérivable au point $F(\mathbf{x})$, alors l'application composée $H = G \circ F : U \rightarrow \mathbb{R}^k$ est dérivable au point \mathbf{x} et on a :*

$$\begin{aligned} dH_{\mathbf{x}} &= dG_{F(\mathbf{x})} \circ dF_{\mathbf{x}} \quad (\text{composition des applications linéaires}), \\ H'(\mathbf{x}) &= G'(F(\mathbf{x}))F'(\mathbf{x}) \quad (\text{multiplication de matrices}). \end{aligned}$$

Preuve:

Comme F est une application dérivable au point \mathbf{x} , il existe une application $\varphi_1 : \mathbb{R}^m \rightarrow \mathbb{R}^n$, telle que :

$$\begin{aligned} F(\mathbf{x} + \mathbf{h}) &= F(\mathbf{x}) + F'(\mathbf{x})\mathbf{h} + \|\mathbf{h}\|\varphi_1(\mathbf{h}), \\ \lim_{\|\mathbf{h}\| \rightarrow 0} \varphi_1(\mathbf{h}) &= \mathbf{0}. \end{aligned} \tag{4.4}$$

De manière analogue, comme G est dérivable au point $F(\mathbf{x})$, il existe une application $\varphi_2 : \mathbb{R}^n \rightarrow \mathbb{R}^k$, telle que :

$$\begin{aligned} G(F(\mathbf{x}) + \mathbf{k}) &= G(F(\mathbf{x})) + G'(F(\mathbf{x}))\mathbf{k} + \|\mathbf{k}\|\varphi_2(\mathbf{k}), \\ \lim_{\|\mathbf{k}\| \rightarrow 0} \varphi_2(\mathbf{k}) &= \mathbf{0}. \end{aligned} \tag{4.5}$$

Posons $\mathbf{k} = F(\mathbf{x} + \mathbf{h}) - F(\mathbf{x})$, ainsi $F(\mathbf{x}) + \mathbf{k} = F(\mathbf{x} + \mathbf{h})$. De plus, par (4.4), on a aussi $\mathbf{k} = F'(\mathbf{x})\mathbf{h} + \|\mathbf{h}\|\varphi_1(\mathbf{h})$. Remplaçant \mathbf{k} dans (4.5), on obtient :

$$G(F(\mathbf{x} + \mathbf{h})) = G(F(\mathbf{x})) + G'(F(\mathbf{x}))F'(\mathbf{x})\mathbf{h} + \|\mathbf{h}\|G'(F(\mathbf{x}))\varphi_1(\mathbf{h}) + \|\mathbf{k}\|\varphi_2(\mathbf{k}).$$

On définit $\varphi_3 : \mathbb{R}^m \rightarrow \mathbb{R}^k$ par :

$$\varphi_3(\mathbf{h}) = G'(F(\mathbf{x}))\varphi_1(\mathbf{h}) + \frac{\|F(\mathbf{x} + \mathbf{h}) - F(\mathbf{x})\|\varphi_2(F(\mathbf{x} + \mathbf{h}) - F(\mathbf{x}))}{\|\mathbf{h}\|},$$

alors on peut montrer (il suffit de réfléchir un peu et d'écrire bien les choses) que $\lim_{\|\mathbf{h}\| \rightarrow 0} \varphi_3(\mathbf{h}) = \mathbf{0}$. Ainsi, on a :

$$\begin{aligned} G(F(\mathbf{x} + \mathbf{h})) &= G(F(\mathbf{x})) + G'(F(\mathbf{x}))F'(\mathbf{x})\mathbf{h} + \|\mathbf{h}\|\varphi_3(\mathbf{h}), \\ \lim_{\|\mathbf{h}\| \rightarrow 0} \varphi_3(\mathbf{h}) &= \mathbf{0}. \end{aligned}$$

On conclut que $G \circ F$ est dérivable au point \mathbf{x} , et que sa dérivée est $G'(F(\mathbf{x}))F'(\mathbf{x})$. \square

Exemple 4.12 Comme corollaire on retrouve le Théorème 19. En effet, soit une fonction $f : U \rightarrow \mathbb{R}$ sur un ouvert U de \mathbb{R}^m , et soit $\gamma : I \rightarrow \mathbb{R}^m$ une courbe, telle que $\gamma(I) \subset U$. Par le Théorème 35, si γ est dérivable au point t , et f est dérivable au point $\gamma(t)$, alors $h = f \circ \gamma$ est dérivable au point t , et

$$h'(t) = f'(\gamma(t))\gamma'(t).$$

Nous devons donc calculer ces dérivées. En utilisant le Théorème 33, on a :

$$f'(\gamma(t)) = \nabla f(\gamma(t))^t, \quad \text{et} \quad \gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_m(t) \end{pmatrix},$$

ainsi

$$h'(t) = \nabla f(\gamma(t))^t \gamma'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t), \quad (\text{produit scalaire}).$$

et on retrouve le Théorème 19.

Exemple 4.13 Soit $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'application coordonnées polaires, définie par

$$T(r, \theta) = (r \cos \theta, r \sin \theta),$$

alors T est dérivable en tout point $(r, \theta) \in \mathbb{R}^2$. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction dérivable. On définit $g = f \circ T$, c'est-à-dire $g(r, \theta) = f(r \cos \theta, r \sin \theta)$. Les hypothèses du Théorème 35 sont vérifiées, donc la dérivée de g au point (r, θ) est :

$$g'(r, \theta) = f'(T(r, \theta))T'(r, \theta), \quad (\text{notations matricielles}).$$

Calculons ces différentes dérivées en utilisant le Théorème 33, dont les hypothèses sont aussi vérifiées.

$$\begin{aligned} g'(r, \theta) &= \nabla g(r, \theta)^t = \left(\frac{\partial g}{\partial r} \quad \frac{\partial g}{\partial \theta} \right) \\ f'(x, y) &= \nabla f(x, y)^t = \left(\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right) \\ T'(r, \theta) &= \begin{pmatrix} \frac{\partial T_1}{\partial r} & \frac{\partial T_1}{\partial \theta} \\ \frac{\partial T_2}{\partial r} & \frac{\partial T_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}. \end{aligned}$$

Alors la dérivée de g est donnée par :

$$\left(\frac{\partial g}{\partial r} \quad \frac{\partial g}{\partial \theta} \right) = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix},$$

en coordonnées, cela donne :

$$\frac{\partial g}{\partial r} = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta, \quad \frac{\partial g}{\partial \theta} = -\frac{\partial f}{\partial x} r \sin \theta + \frac{\partial f}{\partial y} r \cos \theta.$$

4.7 Théorèmes d'inversion locale et des applications implicites

Dans cette section, nous étudions deux théorèmes classiques d'analyse vectorielle. Même si ils semblent de nature différente, ils sont en fait équivalents, comme nous allons le montrer. Par manque de temps, nous n'allons pas donner la preuve du Théorème d'inversion locale. Le lecteur intéressé pourra trouver une preuve dans l'ouvrage de référence [3]. Ces deux théorèmes constituent le socle sur lequel est bâti la théorie des variétés différentiables.

4.7.1 Théorème d'inversion locale

Voici d'abord quelques définitions.

Soit $F : U \rightarrow \mathbb{R}^n$ une application, où U est un ouvert de \mathbb{R}^n . Alors l'application F est un *difféomorphisme*, si F est bijective de U sur $F(U)$, F est dérivable sur U , et son inverse est dérivable sur $F(U)$. L'application F est un *difféomorphisme- \mathcal{C}^k* , si F est bijective de U sur $F(U)$, F est de classe \mathcal{C}^k sur U et son inverse est de classe \mathcal{C}^k sur $F(U)$.

L'application F est une *difféomorphisme- \mathcal{C}^k local* au point $\mathbf{a} \in U$, s'il existe un voisinage ouvert $W \subset U$ de \mathbf{a} , tel que F soit un difféomorphisme- \mathcal{C}^k sur W .

Soit $F : U \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^k . On dit que F est \mathcal{C}^k -*inversible sur* U , si F est un difféomorphisme- \mathcal{C}^k . On dit que F est *localement \mathcal{C}^k -inversible au voisinage du point $\mathbf{a} \in U$* , si F est un difféomorphisme- \mathcal{C}^k local au point \mathbf{a} .

Motivons le Théorème d'inversion locale par les remarques suivantes.

Remarque 4.3 Soit $F : U \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^1 , où U est un ouvert de \mathbb{R}^n .

1. Si l'application F est bijective, l'application inverse est clairement bijective, mais sera-t-elle de classe \mathcal{C}^1 ? Non, pas forcément. Voici un contre-exemple tout simple. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$, définie par $f(x) = x^3$. Alors f est bijective, de classe \mathcal{C}^1 , mais la fonction inverse $g(y) = y^{\frac{1}{3}}$ n'est pas dérivable en 0. Remarquer que dans ce cas $f'(0) = 0$, en particulier f' n'est pas inversible en 0. Ainsi, f est un exemple d'application bijective de classe \mathcal{C}^1 qui n'est pas un difféomorphisme.
2. Si l'application F est bijective, et que l'on suppose de plus que sa dérivée est inversible sur U , est-ce que l'application inverse est de classe \mathcal{C}^1 ? Oui, comme nous allons le voir. C'est à dire que sous cette hypothèse l'application F est un difféomorphisme.
3. Réciproquement, si F est un difféomorphisme, est-ce que sa dérivée est inversible sur U ? Oui. Ceci est une conséquence du Théorème des applications composées. En effet, soit $G : F(U) \rightarrow U$ l'application réciproque. Alors G vérifie :

$$G \circ F = \text{Id}.$$

Par hypothèse, F et G sont de classe \mathcal{C}^1 . De plus l'application Id est linéaire, de sorte que sa différentielle est elle-même en tout point de U , et la matrice qui la représente est la matrice identité. Ainsi, par le théorème des applications composées :

$$\forall \mathbf{x} \in U, G'(F(\mathbf{x}))F'(\mathbf{x}) = \text{I}_n,$$

et la dérivée F' est donc inversible en tout point $\mathbf{x} \in U$. En particulier, soit $\mathbf{y} \in F(U)$, alors il existe $\mathbf{x} \in U$ tel que $F(\mathbf{x}) = \mathbf{y}$, et donc la dérivée de l'application G au point \mathbf{y} est :

$$G'(\mathbf{y}) = F'(\mathbf{x})^{-1}, \quad (\text{inverse matriciel!})$$

4. Finalement, on peut se poser la question suivante. On se souvient que la différentielle en un point est l'application linéaire qui approche l'application F en ce point. Si la différentielle est inversible en un point, est-ce que cette propriété reste vraie pour l'application elle-même? La réponse est positive, localement, et est justement le contenu du Théorème d'inversion locale. Il donne également la régularité de l'inverse local.

Théorème 36 (Inversion locale) Soit $F : U \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^1 , où U est un ouvert de \mathbb{R}^n . Supposons que la dérivée au point $\mathbf{a} \in U$ est inversible. Alors, l'application F est localement \mathcal{C}^1 -inversible au voisinage de \mathbf{a} , i.e. c'est un difféomorphisme- \mathcal{C}^1 local au point \mathbf{a} .

Corollaire 37 Soit $F : U \rightarrow \mathbb{R}^n$ une application bijective de classe \mathcal{C}^1 , où U est un ouvert de \mathbb{R}^n . Supposons que la dérivée est inversible en tout point de U . Alors F est \mathcal{C}^1 -inversible, i.e. F est un difféomorphisme- \mathcal{C}^1 .

Exemple 4.14 Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, définie par $F(x, y) = (x + (y + 2)^2 + 1, (x - 1)^2 + y + 1)$, alors F est de classe \mathcal{C}^1 sur tout \mathbb{R}^2 . Soit $\mathbf{a} = (1, -2)$, donc $F(\mathbf{a}) = (2, -1)$. On veut déterminer si F est localement \mathcal{C}^1 -inversible au voisinage du point \mathbf{a} . Comme l'application est de classe \mathcal{C}^1 , la dérivée est donnée par la matrice jacobienne :

$$J_F(x, y) = \begin{pmatrix} 1 & 2(y + 2) \\ 2(x - 1) & 1 \end{pmatrix}.$$

Donc $F'(\mathbf{a}) = I$, et $\det(F'(\mathbf{a})) = 1$. Ainsi par le théorème d'inversion locale, on sait que F est localement \mathcal{C}^1 -inversible au voisinage de \mathbf{a} . Autrement dit, les équations :

$$\begin{aligned} u &= x + (y + 2)^2 + 1, \\ v &= (x - 1)^2 + y + 1, \end{aligned}$$

peuvent être résolues pour x, y lorsque (u, v) est dans un voisinage de $(2, -1)$. On sait de plus que la dérivée de l'application inverse locale au point $(2, -1)$ est aussi la matrice identité :

$$G'(2, -1) = F'(1, -2)^{-1} = I.$$

4.7.2 Théorème des applications implicites

Le théorème d'inversion locale permet de résoudre l'équation $\mathbf{x} = F(\mathbf{y})$, ou de manière équivalente $\mathbf{x} - F(\mathbf{y}) = \mathbf{0}$, pour \mathbf{y} en fonction de \mathbf{x} . De manière analogue, le théorème des fonctions implicites permet de résoudre l'équation

$$F(\mathbf{x}, \mathbf{y}) = \mathbf{0}, \tag{4.6}$$

pour \mathbf{y} en fonction de \mathbf{x} . Si une telle solution existe, on dit que F définit \mathbf{y} comme une *application implicite de \mathbf{x}* .

De sorte à pouvoir espérer trouver une solution, il faut autant d'équations que d'inconnues, ainsi si $\mathbf{x} \in \mathbb{R}^m$, et $\mathbf{y} \in \mathbb{R}^n$, F est une application de \mathbb{R}^{m+n} dans \mathbb{R}^n .

Remarquer que l'équation (4.6) ne définit pas forcément une application implicite, en effet, à un point \mathbf{x} peut correspondre deux valeurs de \mathbf{y} .

Exemple 4.15 Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, l'application définie par, $F(x, y) = x^2 + y^2 - 2$. On suppose que les valeurs de y sont données au moyen de l'équation implicite : $F(x, y) = 0$. Alors pour le point $x = 0$, correspondent les valeurs $y = \pm\sqrt{2}$. Ainsi, l'ensemble des points :

$$\{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}, \quad (4.7)$$

ne peut être le graphe d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$.

Nous avons besoin de la notation suivante. Soit $F : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^1 au voisinage d'un point (\mathbf{a}, \mathbf{b}) . Soit $J_F(\mathbf{a}, \mathbf{b})$ la matrice jacobienne en ce point. On notera $J_F^2(\mathbf{a}, \mathbf{b})$ la sous-matrice de $J_F(\mathbf{a}, \mathbf{b})$ de taille $n \times n$ constituée des n dernières colonnes de $J_F(\mathbf{a}, \mathbf{b})$.

Théorème 38 (Applications implicites) Soit $F : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^1 au voisinage du point (\mathbf{a}, \mathbf{b}) , où $F(\mathbf{a}, \mathbf{b}) = \mathbf{0}$. Supposons que $\det(J_F^2(\mathbf{a}, \mathbf{b})) \neq 0$. Alors, il existe un voisinage U de \mathbf{a} dans \mathbb{R}^m et un voisinage V de \mathbf{b} dans \mathbb{R}^n , et une application $\varphi : U \rightarrow V$ de classe \mathcal{C}^1 , telle que pour tout $(\mathbf{x}, \mathbf{y}) \in U \times V$, on ait :

$$F(\mathbf{x}, \mathbf{y}) = \mathbf{0} \Leftrightarrow \mathbf{y} = \varphi(\mathbf{x}).$$

Exemple 4.16 Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par, $F(x, y) = x^2 + y^2 - 2$, et $(a, b) = (1, 1)$. Alors $F(1, 1) = 0$. De plus, $J_F^2(x, y) = 2y$, de sorte que $\det(J_F^2(a, b)) = 2 \neq 0$. Ainsi, on peut exprimer y en fonction de x localement autour du point $x = 1$. Remarquer que dans ce cas, on aurait aussi pu résoudre explicitement l'équation $x^2 + y^2 - 2 = 0$, pour obtenir

$$y = \sqrt{2 - x^2}.$$

Si on choisit $(a, b) = (-1, -1)$, alors la résolution explicite donne :

$$y = -\sqrt{2 - x^2}.$$

Si on choisit $(a, b) = (\sqrt{2}, 0)$, ou $(a, b) = (-\sqrt{2}, 0)$, alors le Théorème ne s'applique pas, comme on peut le voir graphiquement.

4.7.3 Lien entre les deux

• Montrons que le Théorème d'inversion locale implique le théorème des applications implicites.

Soit $F : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^1 au voisinage d'un point (\mathbf{a}, \mathbf{b}) , telle que $F(\mathbf{a}, \mathbf{b}) = \mathbf{0}$. On suppose de plus que $\det(J_F^2(\mathbf{a}, \mathbf{b})) \neq 0$.

Définissons l'application $G : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^{m+n}$ de la manière suivante,

$$G(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y})).$$

4.7. THÉORÈMES D'INVERSION LOCALE ET DES APPLICATIONS IMPLICITES 63

Alors G est de classe \mathcal{C}^1 au voisinage de (\mathbf{a}, \mathbf{b}) , et $G(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{0})$. Calculons $J_G(\mathbf{a}, \mathbf{b})$:

$$\left(\begin{array}{c|c} I & \mathbf{0} \\ \hline J_F^1(\mathbf{a}, \mathbf{b}) & J_F^2(\mathbf{a}, \mathbf{b}) \end{array} \right)$$

Ainsi, $\det(J_G(\mathbf{a}, \mathbf{b})) = \det(J_F^2(\mathbf{a}, \mathbf{b})) \neq 0$ par hypothèse. On peut donc appliquer le théorème d'inversion locale. Il existe un voisinage de (\mathbf{a}, \mathbf{b}) que l'on peut supposer de la forme $U \times V$, et un voisinage W de $(\mathbf{a}, \mathbf{0})$, tels que G est \mathcal{C}^1 -inversible de $U \times V$ sur W . Soit H l'application inverse.

Soit $(\mathbf{x}, \mathbf{y}) \in U \times V$ tel que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. Alors,

$$(\mathbf{x}, \mathbf{y}) = H \circ G(\mathbf{x}, \mathbf{y}) = H(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = H(\mathbf{x}, \mathbf{0}).$$

En écrivant H en coordonnées, nous obtenons :

$$(\mathbf{x}, \mathbf{y}) = (H_1(\mathbf{x}, \mathbf{0}), H_2(\mathbf{x}, \mathbf{0})),$$

donc $\mathbf{y} = H_2(\mathbf{x}, \mathbf{0})$. Posons, $\varphi(\mathbf{x}) = H_2(\mathbf{x}, \mathbf{0})$.

Réciproquement, soit $(\mathbf{x}, \mathbf{y}) \in U \times V$ tel que $\mathbf{y} = \varphi(\mathbf{x}) = H_2(\mathbf{x}, \mathbf{0})$. Alors

$$G(\mathbf{x}, \mathbf{y}) = G(\mathbf{x}, H_2(\mathbf{x}, \mathbf{0})) = G(H_1(\mathbf{x}, \mathbf{0}), \mathbf{x}, H_2(\mathbf{x}, \mathbf{0})) = (\mathbf{x}, \mathbf{0}).$$

De plus, par définition, $G(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$. Donc, $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$.

• Montrons que le Théorème des applications implicites implique le Théorème d'inversion locale.

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^1 dans un voisinage d'un point $\mathbf{a} \in \mathbb{R}^n$. On définit l'application $G : \mathbb{R}^{n+n} \rightarrow \mathbb{R}^n$ par :

$$G(\mathbf{x}, \mathbf{y}) = F(\mathbf{y}) - \mathbf{x}.$$

Alors G est de classe \mathcal{C}^1 au voisinage de $(F(\mathbf{a}), \mathbf{a})$. De plus, $G(F(\mathbf{a}), \mathbf{a}) = \mathbf{0}$, et $\det(J_G^2(F(\mathbf{a}), \mathbf{a})) = \det(J_F(\mathbf{a})) \neq 0$, par hypothèse. Ainsi, on peut appliquer le Théorème des applications implicites. Il existe un voisinage $U \subset \mathbb{R}^n$ de $F(\mathbf{a})$, et $V \subset \mathbb{R}^n$ de \mathbf{a} , et une application $\varphi : U \rightarrow V$, telle que pour tout $(\mathbf{x}, \mathbf{y}) \in U \times V$, $G(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, ssi $\mathbf{y} = \varphi(\mathbf{x})$. Ainsi, pour tout $\mathbf{x} \in U$, $F \circ \varphi(\mathbf{x}) = \mathbf{x}$. De manière analogue, on peut montrer que $\varphi \circ F(\mathbf{y}) = \mathbf{y}$, pour tout $\mathbf{y} \in V$.

Chapitre 5

Intégrales multiples

Cette première section d'intégration a pour sujet l'intégration de Riemann des fonctions de \mathbb{R}^m dans \mathbb{R} . C'est une généralisation naturelle de l'intégration de Riemann que vous avez rencontrée dans le premier cours d'Analyse. Il y aura néanmoins quelques subtilités supplémentaires dues au fait que l'on travaille en dimensions supérieures.

Souvenez-vous que l'intégrale de Riemann d'une fonction réelle calcule l'aire contenue sous le graphe d'une fonction. De manière analogue, l'intégrale de Riemann des fonctions de plusieurs variables, représente le volume contenu sous le graphe de la fonction.

La plupart des théorèmes de cette section sont donnés sans démonstration.

5.1 Intégration sur les pavés

Souvenez-vous que l'intégrale de Riemann en dimension 1 est définie pour des fonctions dont le domaine de définition est un intervalle borné (ou réunion finie de tels intervalles). La généralisation naturelle des intervalles est la notion de *pavé* introduite ci-dessous. Nous commençons donc par définir l'intégrale de Riemann sur des pavés. Ensuite, nous étendons cette définition à des domaines plus généraux. En effet, si nous souhaitons intégrer une fonction définie sur un disque, cela n'entrerait même pas dans cette catégorie. Dans tout ce qui suit, il est utile de se représenter le cas de \mathbb{R}^2 , où la visualisation géométrique est plus aisée.

Un *pavé* P de \mathbb{R}^m est le produit cartésien de m intervalles bornés non-vides de \mathbb{R} . Il peut être écrit explicitement sous la forme,

$$P = \prod_{j=1}^m I_j, \quad \text{où } I_j = [a_j, b_j] \subset \mathbb{R}.$$

Le *volume* $v(P)$ du pavé P est par définition :

$$v(P) = \prod_{j=1}^m (b_j - a_j).$$

Exemple 5.1

En dimension 1, les pavés sont les intervalles fermés, et le volume est habituellement appelé *longueur*.

En dimension 2, les pavés sont les rectangles, et le volume est habituellement appelé *aire*.

En dimension 3, les pavés sont les parallélépipèdes, et le volume est habituellement appelé *volume*.

Avant de définir rigoureusement l'intégrale de Riemann d'une fonction f définie sur un pavé P , voici une description en mots de ce que l'on va faire.

- Découper le pavé P en pavés plus petits P_i .
- Pour chacun des petits pavés P_i , définir l'intégrale d'une fonction constante de valeur K sur un petit pavé P_i par $Kv(P_i)$. Ainsi, pour les fonctions constantes l'intégrale représente bien le volume sous le graphe de la fonction.
- Approcher la fonction définie sur P par des fonctions constantes sur les petits pavés P_i .
- Définir l'intégrale de f sur P comme la limite, si elle existe, des sommes des volumes des fonctions constantes sur des pavés de plus en plus petits.

Une *partition* du pavé P , est un ensemble fini de pavés $\mathcal{P} = \{P_i\}_{1 \leq i \leq r}$, qui satisfont :

$$P = \bigcup_{i=1}^r P_i, \quad \text{et } \text{int}(P_i) \cap \text{int}(P_j) = \emptyset, \quad \forall i \neq j.$$

La *taille* d'une partition $\mathcal{P} = \{P_i\}_{1 \leq i \leq r}$, notée $\delta(\mathcal{P})$, est définie de la manière suivante. Pour chacun des sous-pavés P_i , on note δ_i la taille du plus grand intervalle dans le produit cartésien de P_i , alors

$$\delta(\mathcal{P}) = \max_{\{1 \leq i \leq r\}} \delta_i.$$

Soit f une fonction bornée sur un pavé P , soit $\mathcal{P} = \{P_i\}_{1 \leq i \leq r}$ une partition de P . On définit les *sommes de Riemann inférieures* et *supérieures* :

$$S_1(f, \mathcal{P}) = \sum_{i=1}^r \inf_{P_i} (f) v(P_i),$$

$$S_2(f, \mathcal{P}) = \sum_{i=1}^r \sup_{P_i} (f) v(P_i).$$

La quantité que nous souhaitons définir est clairement supérieure à $S_1(f, \mathcal{P})$ et inférieure à $S_2(f, \mathcal{P})$. Il serait naturel de définir l'intégrale comme le sup sur toutes les partitions de $S_1(f, \mathcal{P})$, et l'inf sur toutes les partitions de $S_2(f, \mathcal{P})$, mais il se pourrait que ces deux valeurs ne coïncident pas. Ceci nous amène naturellement à la définition suivante.

Une fonction f bornée définie sur un pavé P est dit *intégrable au sens de Riemann sur P* , ssi

$$\sup_{\mathcal{P}} S_1(f, \mathcal{P}) = \inf_{\mathcal{P}} S_2(f, \mathcal{P}).$$

L'*intégrale de f sur P* , notée $\int_P f$, est alors la valeur commune de ces deux quantités.

Le théorème suivant donne une condition nécessaire et suffisante pour l'intégrabilité au sens de Riemann, qui peut être utile dans la pratique.

Théorème 39 *Soit f une fonction bornée définie sur un pavé P . La fonction f est intégrable au sens de Riemann sur P , et l'intégrale vaut I , ssi :*

Pour tout $\varepsilon > 0$, il existe $\delta > 0$, tel que :

$$\left| I - \sum_{i=1}^r f(\mathbf{x}_i)v(P_i) \right| \leq \varepsilon,$$

pour toute partition $\mathcal{P} = \{P_i\}_{i=1}^r$ de taille inférieure à δ , et pour tout choix $\{\mathbf{x}_i\}_{i=1}^r$ tels que $\mathbf{x}_i \in P_i$. On écrit aussi :

$$\lim_{\delta(\mathcal{P}) \rightarrow 0} \sum_{i=1}^r f(\mathbf{x}_i)v(P_i) = \int_P f.$$

Si f est bornée et intégrable au sens de Riemann sur un pavé P , on appelle la quantité $\sum_{i=1}^r f(\mathbf{x}_i)v(P_i)$, une *somme de Riemann* pour f .

Voici les propriétés vérifiées par l'intégrale de Riemann.

Proposition 40

1. *L'ensemble des fonctions bornées sur un pavé P , qui sont intégrables au sens de Riemann, forment un espace vectoriel.*
2. *L'application $f \mapsto \int_P f$ est linéaire. C'est à dire que si f et g sont deux fonctions bornées sur un pavé P , intégrables au sens de Riemann, et si $\alpha, \beta \in \mathbb{R}$, alors :*

$$\int_P \alpha f + \beta g = \alpha \int_P f + \beta \int_P g.$$

3. Si $f \geq 0$ est une fonction bornée sur un pavé P , et intégrable au sens de Riemann, alors :

$$\int_P f \geq 0.$$

Corollaire 41 Soient f et g deux fonctions bornées sur un pavé P , intégrables au sens de Riemann.

1. Si $f \leq g$, alors :

$$\int_P f \leq \int_P g.$$

2. Si $|f|$ est intégrable au sens de Riemann, alors :

$$\left| \int_P f \right| \leq \int_P |f|.$$

Preuve:

1. Il suffit de poser $h = g - f$, et d'utiliser 1, 2 et 3.
2. Remarquer que $|f| - f \geq 0$ et $|f| + f \geq 0$, et utiliser 1, 2 et 3.

□

Avec la définition ci-dessus, il est difficile en pratique de déterminer si une fonction est intégrable au sens de Riemann. Le critère ci-dessous donne une condition suffisante. Nous avons besoin de la définition suivante.

Un sous-ensemble A de \mathbb{R}^m est *négligeable*, ssi pour tout $\varepsilon > 0$, il existe des pavés P_1, \dots, P_r de \mathbb{R}^m tels que :

$$A \subset \bigcup_{i=1}^r P_i, \quad \text{et} \quad \sum_{i=1}^r v(P_i) < \varepsilon.$$

Une fonction f bornée définie sur un pavé P est dite *admissible*, ssi elle est continue sauf sur un sous-ensemble négligeable de P .

Remarque 5.1 On a les propriétés immédiates suivantes :

1. Une réunion (intersection) finie d'ensembles négligeables est négligeable.
2. L'ensemble des fonctions admissibles forme un espace vectoriel.
3. Si f et g sont deux fonctions admissibles, alors $\max\{f, g\}$, $\min\{f, g\}$, $|f|$ sont admissibles.

Théorème 42 *Toute fonction admissible sur un pavé P est intégrable au sens de Riemann sur P . De plus, si f et g sont deux fonctions admissibles sur P , et $f(\mathbf{x}) = g(\mathbf{x})$ sauf sur un ensemble négligeable, alors*

$$\int_P f = \int_P g.$$

5.2 Intégration sur des domaines plus généraux

Nous établissons ici l'intégrale de Riemann sur une famille de sous-ensembles de \mathbb{R}^m .

Un sous-ensemble A de \mathbb{R}^m est dit *admissible*, ssi il est borné, et son bord (ou frontière) ∂A , est négligeable.

Remarque 5.2 On a les propriétés suivantes :

1. Une réunion (intersection) finie d'ensembles admissibles est admissible.
2. Si A et B sont deux sous-ensembles admissibles de \mathbb{R}^m , et $A \subset B$, alors $B \setminus A$ est un sous-ensemble admissible de \mathbb{R}^m .
3. Si A est un sous-ensemble admissible de \mathbb{R}^m , et B est un sous-ensemble admissible de \mathbb{R}^n , alors $A \times B$ est un sous-ensemble admissible de \mathbb{R}^{m+n} .

Soit A un sous-ensemble admissible de \mathbb{R}^m , et soit f une fonction bornée définie sur A . Comme A est borné, on peut l'inclure dans un pavé P de \mathbb{R}^m . Définissons $f_A : P \rightarrow \mathbb{R}$ par :

$$f_A(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{lorsque } \mathbf{x} \in A, \\ 0 & \text{lorsque } \mathbf{x} \in P \setminus A. \end{cases}$$

On dit qu'une fonction f bornée définie sur A , est *intégrable au sens de Riemann sur A* , ssi la fonction f_A est intégrable au sens de Riemann sur P . La quantité $\int_P f_A$ s'appelle alors l'*intégrale de Riemann de f sur A* , et est notée $\int_A f$. Il est facile de voir que cette définition ne dépend pas du pavé P choisi.

Les propriétés de l'intégrale de Riemann sur un domaine admissible A sont les mêmes que celle de la Proposition 40 et du Corollaire 41. On a en plus les propriétés suivantes :

Proposition 43

1. Soit A, B deux sous-ensembles admissibles de \mathbb{R}^m . Soient f une fonction bornée définie sur $A \cup B$, intégrable au sens de Riemann sur $A \cup B$, alors :

$$\int_{A \cup B} f = \int_A f + \int_B f - \int_{A \cap B} f.$$

2. Soit A un sous-ensemble admissible de \mathbb{R}^m , et f une fonction bornée sur A , intégrable au sens de Riemann sur A . Si A est négligeable, alors :

$$\int_A f = 0.$$

Etablissons maintenant un critère suffisant d'intégrabilité au sens de Riemann pour une fonction bornée, définie sur un sous-ensemble admissible A de \mathbb{R}^m .

Une fonction f bornée définie sur un sous-ensemble admissible A de \mathbb{R}^m est dite *admissible*, ssi elle est continue sauf sur un sous-ensemble négligeable de A .

Théorème 44 Soit A un sous-ensemble admissible de \mathbb{R}^m , et f une fonction admissible sur A . Alors f est intégrable au sens de Riemann sur A .

Preuve:

Soit P un pavé qui contient A . Nous devons montrer que f_A est intégrable au sens de Riemann sur P . Les points de discontinuité de f_A sont les points de discontinuité de f , et les points de ∂A . Comme A est admissible, on sait que ∂A est négligeable. Ainsi l'ensemble des points de discontinuité de f est réunion de deux ensembles négligeables, et est donc négligeable. On en déduit que f_A est admissible sur P , et donc intégrable sur P par le Théorème 42. \square

Soit A un sous-ensemble admissible de \mathbb{R}^m , alors le *volume de A* , noté $v(A)$, est défini par :

$$v(A) = \int_A 1.$$

Remarquer que le volume est bien défini par le théorème ci-dessus. Lorsque A est un sous-ensemble admissible de \mathbb{R} , on appelle le volume de A , la *longueur*, et lorsque A est un sous-ensemble admissible de \mathbb{R}^2 , on appelle le volume de A , l'*aire*.

On a encore la propriété suivante, qui sera démontrée aux exercices.

Proposition 45 Soit A un sous-ensemble admissible de \mathbb{R}^m , et soit f une fonction bornée, intégrable au sens de Riemann sur A , alors :

$$\left| \int_A f \right| \leq \left(\sup_{x \in A} |f(x)| \right) v(A).$$

5.3 Calculs explicites

Nous avons maintenant des critères utiles qui permettent de déterminer si une fonction est intégrable au sens de Riemann ou non. Mais, nous n'avons pas de méthode efficace pour calculer ces intégrales dans le cas où elles existent. L'idée est de se ramener aux intégrales de fonctions réelles, que vous savez calculer grâce au Théorème fondamental du calcul intégral. Le théorème clé qui permet de faire ceci, est le Théorème de Fubini ci-dessous. Rappelons d'abord le Théorème fondamental du calcul intégral.

Soit $f : I \rightarrow \mathbb{R}$ une fonction définie sur un intervalle I . Une fonction F est appelée une *primitive* de f , ssi $F'(x) = f(x)$, $\forall x \in I$.

Théorème 46 (Fondamental du calcul intégral) Soit $f : I \rightarrow \mathbb{R}$ une fonction continue sur l'intervalle $I = [a, b]$. Alors f admet une primitive F , et

$$\int_I f = F(b) - F(a).$$

Notation : on écrit aussi $\int_I f = \int_a^b f(x)dx$.

Théorème 47 (Fubini) Soit P un pavé de \mathbb{R}^m et Q un pavé de \mathbb{R}^n . Soit $f : P \times Q \rightarrow \mathbb{R}$ une fonction intégrable au sens de Riemann sur $P \times Q$. On suppose que pour tout \mathbf{x} sauf ceux dans un sous-ensemble négligeable de P , la fonction $f_{\mathbf{x}}(\mathbf{y}) := f(\mathbf{x}, \mathbf{y})$ est intégrable au sens de Riemann sur Q . Alors, $\int_Q f_{\mathbf{x}}$ est intégrable sur P , et :

$$\int_{P \times Q} f = \int_P \left(\int_Q f_{\mathbf{x}} \right).$$

Notation : on écrit aussi $\int_{P \times Q} f = \int_P \left(\int_Q f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) d\mathbf{x}$.

Exemple 5.2 Soit $f : [1, 2] \times [-3, 4] \rightarrow \mathbb{R}$ définie par $f(x, y) = x^2y$. Alors, f est continue sur $[1, 2] \times [-3, 4]$, et pour tout $x \in [1, 2]$, f_x est continue sur $[-3, 4]$. Ainsi, f est intégrable sur $[1, 2] \times [-3, 4]$, et pour tout $x \in [1, 2]$, f_x est intégrable sur $[-3, 4]$

(Théorème 42). Donc, le Théorème de Fubini s'applique, et

$$\begin{aligned}
 \int_{[1,2] \times [-3,4]} f &= \int_1^2 \left(\int_{-3}^4 f(x,y) dy \right) dx, \\
 &= \int_1^2 x^2 \left(\frac{y^2}{2} \Big|_{-3}^4 \right) dx, \text{ (Théorème fondamental du calcul intégral)} \\
 &= \frac{7}{2} \int_1^2 x^2 dx, \\
 &= \frac{7}{2} \frac{x^3}{3} \Big|_1^2, \text{ (Théorème fondamental du calcul intégral)} \\
 &= \frac{49}{6}.
 \end{aligned}$$

Remarquer que l'on aurait aussi pu calculer les intégrales dans l'“autre sens”. En effet, pour tout $y \in [-3,4]$, f_y est aussi continue sur $[1,2]$, donc intégrable sur $[1,2]$. Ainsi, le Théorème de Fubini s'applique, et

$$\begin{aligned}
 \int_{[1,2] \times [-3,4]} f &= \int_{-3}^4 \left(\int_1^2 f(x,y) dx \right) dy, \\
 &= \int_{-3}^4 y \left(\frac{x^3}{3} \Big|_1^2 \right) dy, \text{ (Théorème fondamental du calcul intégral)} \\
 &= \frac{7}{3} \int_{-3}^4 y dy, \\
 &= \frac{7}{3} \frac{y^2}{2} \Big|_{-3}^4, \text{ (Théorème fondamental du calcul intégral)} \\
 &= \frac{49}{6}.
 \end{aligned}$$

En répétant de manière récursive le raisonnement de l'exemple ci-dessus, on obtient le corollaire suivant :

Corollaire 48 Soit $f : P \rightarrow \mathbb{R}$ une fonction continue sur un pavé $P = [a_1, b_1] \times \cdots \times [a_m, b_m]$ de \mathbb{R}^m . Alors pour toute permutation σ de \mathcal{S}_m , on a :

$$\int_P f = \int_{a_{\sigma(1)}}^{b_{\sigma(1)}} \cdots \left(\int_{a_{\sigma(m)}}^{b_{\sigma(m)}} f(x_1, \cdots, x_m) dx_{\sigma(m)} \right) \cdots dx_{\sigma(1)}.$$

Voici deux conséquences du Théorème de Fubini qui sont utiles en pratique.

Théorème 49 Soit $g_1, g_2 : \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions continues, telles que $g_1(x) \leq g_2(x)$, $\forall x \in \mathbb{R}$. Soient deux réels $a < b$, et

$$A = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}.$$

Soit $f : A \rightarrow \mathbb{R}$ une fonction continue sur A , alors f est intégrable sur A , et :

$$\int_A f = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx.$$

Preuve:

Par l'exercice 5 de la Série 10, on sait que A est un sous-ensemble admissible de \mathbb{R}^2 . Ainsi par le Théorème 44, f est intégrable sur A . Soit $[c, d]$ un intervalle qui contient $[\min_{x \in [a, b]} g_1(x), \max_{x \in [a, b]} g_2(x)]$ (existe car une fonction continue sur un compact est bornée et atteint son minimum et son maximum). Alors par définition de l'intégrale sur un domaine admissible, on a :

$$\int_A f = \int_{[a, b] \times [c, d]} f_A.$$

De plus pour tout $x \in [a, b]$, la fonction $(f_A)_x(y) = f_A(x, y)$ est continue sur $[c, d]$ sauf en $\{g_1(x), g_2(x)\}$ qui est négligeable, ainsi $(f_A)_x$ est intégrable sur $[c, d]$. Le Théorème de Fubini s'applique donc, et donne :

$$\begin{aligned} \int_{[a, b] \times [c, d]} f_A &= \int_a^b \left(\int_c^d f_A(x, y) dy \right) dx, \\ &= \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx. \end{aligned}$$

□

Exemple 5.3 On veut calculer l'aire de la région bornée, comprise entre la fonction $g_1(x) = x^2$ et la droite $g_2(x) = x$. Soit

$$A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, g_1(x) \leq y \leq g_2(x)\}.$$

Alors, par l'exercice 5 de la Série 10, A est admissible, l'aire de A , notée $v(A)$, est bien définie, et

$$v(A) = \int_A 1.$$

Comme la fonction 1 est continue sur A , les hypothèses du Théorème 49 sont vérifiées, et donc :

$$\begin{aligned} v(A) &= \int_A 1 = \int_0^1 \left(\int_{x^2}^x 1 \, dy \right) dx, \\ &= \int_0^1 (x - x^2) dx = \left. \frac{x^2}{2} - \frac{x^3}{3} \right|_0^1 = \frac{1}{6}. \end{aligned}$$

La deuxième application du Théorème de Fubini que nous étudions permet en particulier de calculer des volumes de révolution.

Théorème 50 (Principe de Cavalieri) *Soit A un sous-ensemble admissible de \mathbb{R}^3 , tel que $A \subset [a, b] \times P$, où $a < b$, et P est un pavé de \mathbb{R}^2 . On suppose que pour tout $t \in [a, b]$, l'ensemble*

$$A_t := \{(x, y) \in \mathbb{R}^2 \mid (t, x, y) \in A\} \subset P,$$

est admissible dans \mathbb{R}^2 , et on note $A(t) = v(A_t)$. Alors,

$$v(A) = \int_a^b A(t) \, dt.$$

Preuve:

Comme A est admissible le volume de A est bien défini, et

$$v(A) = \int_A 1 = \int_{[a,b] \times P} 1_A.$$

De plus, pour tout $t \in [a, b]$, la fonction $(1_A)_t(x, y) = 1_A(t, x, y) = 1_{A_t}(x, y)$ est continue sur P , sauf sur ∂A_t qui est négligeable par hypothèse, ainsi elle est intégrable sur P . Le Théorème de Fubini s'applique donc, et on déduit :

$$\begin{aligned} v(A) &= \int_A 1 = \int_{[a,b] \times P} 1_A, \\ &= \int_a^b \left(\int_P 1_{A_t} \right) dt, \\ &= \int_a^b \left(\int_{A_t} 1 \right) dt, \text{ (définition de l'intégrale sur un domaine admissible)} \\ &= \int_a^b A(t) \, dt. \end{aligned}$$

□

Exemple 5.4 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue, positive. Soit A le sous-ensemble de \mathbb{R}^3 obtenu par révolution du graphe de la fonction autour de l'axe $(0x)$. Alors, vous allez montrer en exercice que le Principe de Cavalieri s'applique, et donne :

$$v(A) = \pi \int_a^b f(t)^2 dt.$$

5.4 Changement de variables

La formule du changement de variables permet dans certains cas de calculer plus facilement des intégrales. C'est une généralisation en dimensions supérieures de la formule que vous avez étudiée pour l'intégration des fonctions réelles ; mais sous des hypothèses plus fortes.

Nous ne démontrons pas entièrement la formule, mais donnons assez d'éléments afin que vous vous en fassiez une idée intuitive.

La première chose à comprendre est l'évolution du volume d'un pavé lorsqu'il est transformé par une application linéaire. Ceci est donné par le théorème suivant, que nous démontrons seulement dans le cas $m = 2$. Si L est une application linéaire, on note aussi L la matrice associée, donnée par le Théorème 30.

Théorème 51 Soit P un pavé de \mathbb{R}^m , et soit $L : \mathbb{R}^m \rightarrow \mathbb{R}^m$ une application linéaire, alors :

$$v(L(P)) = |\det(L)|v(P).$$

Preuve:

Soit P un pavé de \mathbb{R}^2 , c'est à dire que P est un rectangle. On admet (ceci est intuitivement clair) que le volume d'un pavé est invariant par translation. Ainsi, nous supposons que $P = [0, a] \times [0, b]$, de sorte que :

$$v(P) = ab.$$

On peut réécrire P sous la forme,

$$P = \{\alpha \mathbf{e}_1 + \beta \mathbf{e}_2 \mid 0 \leq \alpha \leq a, 0 \leq \beta \leq b\},$$

ainsi

$$L(P) = \{\alpha L(\mathbf{e}_1) + \beta L(\mathbf{e}_2) \mid 0 \leq \alpha \leq a, 0 \leq \beta \leq b\}.$$

On suppose connue la formule pour le volume d'un parallélogramme :

$$v(L(P)) = ab \|L(\mathbf{e}_1)\| \cdot \|L(\mathbf{e}_2)\| \cdot |\sin \theta|,$$

où θ est l'angle entre $L(\mathbf{e}_1)$ et $L(\mathbf{e}_2)$. Notons $\mathbf{l}_1, \mathbf{l}_2$ les vecteurs $L(\mathbf{e}_1)$ et $L(\mathbf{e}_2)$. De plus,

$$|\sin \theta| = \sqrt{1 - \cos^2 \theta} = \sqrt{1 - \frac{(\mathbf{l}_1 \cdot \mathbf{l}_2)^2}{(\|\mathbf{l}_1\| \cdot \|\mathbf{l}_2\|)^2}} = \frac{1}{\|\mathbf{l}_1\| \cdot \|\mathbf{l}_2\|} \sqrt{(\|\mathbf{l}_1\| \cdot \|\mathbf{l}_2\|)^2 - (\mathbf{l}_1 \cdot \mathbf{l}_2)^2}.$$

Donc,

$$v(L(P)) = ab\sqrt{(\|\mathbf{l}_1\| \cdot \|\mathbf{l}_2\|)^2 - (\mathbf{l}_1 \cdot \mathbf{l}_2)^2}.$$

La dernière étape, qui est laissée en exercice, est la suivante. En écrivant, $\mathbf{l}_1 = (l_{11}, l_{21}), \mathbf{l}_2 = (l_{12}, l_{22})$ en coordonnées, montrer que :

$$(\|\mathbf{l}_1\| \cdot \|\mathbf{l}_2\|)^2 - (\mathbf{l}_1 \cdot \mathbf{l}_2)^2 = (l_{11}l_{22} - l_{12}l_{21})^2.$$

De là on déduit :

$$v(L(P)) = ab|\det(L)| = |\det(L)|v(P).$$

□

Soit $T : U \rightarrow \mathbb{R}^m$ une application, où U est un ouvert de \mathbb{R}^m . Le théorème suivant dit que si l'application T est suffisamment régulière, alors on peut localement remplacer T par sa différentielle, qui est une application linéaire. Nous l'énonçons de manière imprécise (sans les ε).

Théorème 52 *Soit P un pavé suffisamment petit, soit \mathbf{x} son centre. Soit $T : U \rightarrow \mathbb{R}^m$ un difféomorphisme- \mathcal{C}^1 , où U est un ouvert de \mathbb{R}^m qui contient P . Alors, on peut approcher T sur P par sa différentielle au point \mathbf{x} , $dT_{\mathbf{x}}$.*

Énonçons maintenant la formule du changement de variables, et donnons un argument intuitif dans un cas particulier.

Théorème 53 (Changement de variables) *Soient A un sous-ensemble admissible de \mathbb{R}^m , et $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ une application de classe \mathcal{C}^1 , qui est un difféomorphisme- \mathcal{C}^1 sur $\text{int}(A)$. Soit f une fonction admissible sur $T(A)$, alors $f \circ T$ est admissible sur A , et :*

$$\int_{T(A)} f = \int_A (f \circ T) |\det J_T|.$$

Remarque 5.3 On peut montrer que :

1. Si A est un sous-ensemble admissible de \mathbb{R}^m et T est une application de classe \mathcal{C}^1 sur \mathbb{R}^m , alors $T(A)$ est un sous-ensemble admissible de \mathbb{R}^m .
2. Sous les hypothèses du Théorème, l'application $f \circ T$ est admissible.

Preuve:

Donnons une heuristique lorsque A est un pavé P , ceci n'est pas une preuve.

Soit $\mathcal{P} = \{P_1, \dots, P_r\}$ une partition de P en tout petit pavés, et soit \mathbf{x}_i le centre du pavé P_i . Alors, en utilisant une approximation de l'intégrale dans l'idée des sommes de Riemann, on obtient

$$\begin{aligned} \int_{T(P)} f &= \sum_{i=1}^r \int_{T(P_i)} f \\ &\sim \sum_{i=1}^r f(T(\mathbf{x}_i))v(T(P_i)), \quad (\text{puisque les } P_i \text{ sont petits}) \\ &\sim \sum_{i=1}^r f(T(\mathbf{x}_i))v(dT_{\mathbf{x}_i}(P_i)), \quad (\text{par le Théorème 52}) \\ &= \sum_{i=1}^r f(T(\mathbf{x}_i))|\det T'(\mathbf{x}_i)|v(P_i), \quad (\text{par le Théorème 51}) \\ &\sim \int_P (f \circ T)|\det T'| = \int_P (f \circ T)|\det J_T|, \end{aligned}$$

(T étant de classe \mathcal{C}^1 , la dérivée T' est donnée par la matrice jacobienne J_T).

□

Exemple 5.5 Soit $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'application coordonnées polaires, donnée par :

$$T(r, \theta) = (r \cos \theta, r \sin \theta).$$

Alors T est de classe \mathcal{C}^1 sur \mathbb{R}^2 , donc dérivable sur \mathbb{R}^2 , et :

$$|\det(J_T(r, \theta))| = |r|.$$

Soit $U = (0, \infty) \times (0, 2\pi)$, alors l'application T est un difféomorphisme- \mathcal{C}^1 sur U .

Soit A un sous-ensemble admissible de $[0, \infty) \times [0, 2\pi]$, alors $\text{int}(A) \subset U$, de sorte que T est un difféomorphisme- \mathcal{C}^1 sur $\text{int}(A)$. Ainsi, si f est une fonction admissible sur $T(A)$, on a :

$$\int_{T(A)} f = \int_A f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Prenons un cas particulier. Soit $A = [0, R] \times [0, 2\pi]$ qui est un sous-ensemble admissible de $[0, \infty) \times [0, 2\pi]$, alors $T(A)$ est le disque de centre $\mathbf{0}$, et de rayon R , noté D . Soit f une fonction continue sur D , alors en utilisant aussi le Théorème de Fubini, on a :

$$\int_D f = \int_0^{2\pi} \int_0^R f(r \cos \theta, r \sin \theta) dr d\theta.$$

Prenons par exemple $f(x, y) = e^{-(x^2+y^2)}$, alors :

$$\int_D e^{-(x^2+y^2)} dx dy = \int_0^{2\pi} \int_0^R e^{-r^2} r dr d\theta = \pi(1 - e^{-R^2}).$$

Une autre application est de calculer l'aire du disque D :

$$v(D) = \int_D 1 = \int_0^{2\pi} \int_0^R r dr d\theta = \pi R^2.$$

Chapitre 6

Intégrales curvilignes et Théorème de Green

Dans un premier temps, nous allons définir l'intégration le long d'une courbe paramétrée (objet uni-dimensionnel). Il existe deux sortes d'objets que l'on peut intégrer : les fonctions, et les champs de vecteurs définis le long de cette courbe. Dans le premier cas, c'est une généralisation de l'intégrale de Riemann sur un intervalle de \mathbb{R} (objet uni-dimensionnel particulier).

Ensuite, nous énonçons le Théorème de Green, et le démontrons dans un cas particulier. Ce théorème relie l'intégrale d'un champ de vecteurs sur le bord d'un domaine suffisamment régulier (intégrale curviligne), à une intégrale sur le domaine tout entier (intégrale multiple en dimension 2).

6.1 Intégrale curviligne d'une fonction

Dans cette section, nous intégrons une fonction le long d'une courbe paramétrée.

Soit $\gamma : I = [a, b] \rightarrow \mathbb{R}^n$ une courbe paramétrée de classe \mathcal{C}^1 , et soit $f : \gamma[a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors l'intégrale de la fonction f le long de la courbe γ , notée $\int_{\gamma} f d\ell$, est par définition :

$$\int_{\gamma} f d\ell = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt.$$

Cette intégrale est aussi appelée *intégrale curviligne de la fonction f le long de la courbe γ* .

Exemple 6.1 Soit $\gamma : I = [a, b] \rightarrow \mathbb{R}^n$ une courbe paramétrée de classe \mathcal{C}^1 . Alors le centre de gravité de la courbe, noté \mathbf{x}_G , est par définition :

$$\mathbf{x}_G = \frac{1}{\int_{\gamma} dl} \left(\int_{\gamma} x_1 dl, \dots, \int_{\gamma} x_n dl \right).$$

Remarque 6.1

- La longueur de la courbe γ , définie à la Section 2.3, est l'intégrale de la fonction identiquement égale à 1 le long de la courbe γ , c'est-à-dire :

$$\ell(\gamma) = \int_{\gamma} dl.$$

- La justification de cette définition est exactement dans l'esprit de celle de la définition de la longueur d'une courbe de la Section 2.3, de sorte que nous ne la répétons pas ici.
- Soit $I = [a, b]$ un intervalle de \mathbb{R} (considéré comme la trace d'une courbe dans \mathbb{R}), et soit f une fonction continue sur I . Alors, une paramétrisation de classe \mathcal{C}^1 de I est donnée par $\gamma : I \rightarrow \mathbb{R}$, avec $\gamma(t) = t$. Ainsi,

$$\int_{\gamma} f dl = \int_a^b f(t) dt.$$

La question naturelle suivante se pose. Si deux courbes paramétrées ont la même trace, est-ce que l'intégrale d'une fonction f le long de ces deux courbes est la même? On s'était déjà posé cette question pour la longueur d'une courbe. La réponse est la même, et la preuve est tout à fait similaire, de sorte qu'on ne la répète pas ici.

Théorème 54 Soient $\gamma_1 : I_1 \rightarrow \mathbb{R}^n$ et $\gamma_2 : I_2 \rightarrow \mathbb{R}^n$ deux courbes paramétrées de classe \mathcal{C}^1 , telles que $\gamma_1(I_1) = \gamma_2(I_2)$. Soit $f : \gamma_1(I_1) \rightarrow \mathbb{R}$ une fonction continue. Alors si les deux courbes sont équivalentes, on a :

$$\int_{\gamma_1} f dl = \int_{\gamma_2} f dl.$$

6.2 Intégrale curviligne d'un champ de vecteurs

Dans cette section, nous intégrons un champ de vecteurs le long d'une courbe paramétrée.

Soit $\gamma : I = [a, b] \rightarrow \mathbb{R}^n$ une courbe paramétrée de classe \mathcal{C}^1 , et soit $F : \gamma[a, b] \rightarrow \mathbb{R}^n$ un champ de vecteurs continu. D'un point de vue physique, il faut se représenter $\gamma(t)$

comme étant la position d'une particule à l'instant t , qui se déplace sous l'influence du champ de forces F (ainsi $F(\gamma(t))$ est la force agissant sur la particule à l'instant t). L'intégrale du champ de vecteurs le long de la courbe γ , notée $\int_{\gamma} F$, est par définition le travail effectué par ce champ de forces pour déplacer la particule le long de la courbe de $\gamma(a)$ à $\gamma(b)$. Mathématiquement, cela se traduit de la manière suivante :

$$\int_{\gamma} F = \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt.$$

Cette intégrale est appelée *intégrale curviligne du champ de vecteurs F le long de la courbe γ* .

Remarque 6.2 *Justification de la définition*

Comme pour la définition de la longueur d'une courbe, c'est une justification infinitésimale. Soit $\mathcal{P} = \{a = t_0 < t_1 < \dots < t_k = b\}$ une partition de l'intervalle $[a, b]$. On approche la trace de la courbe par la réunion des petits segments de $\gamma(t_{i-1})$ à $\gamma(t_i)$. Sur chacun des petits segments, on considère que le champ de forces est constant, et égal à $F(\gamma(t_i))$, de sorte qu'en accord avec les lois de la physique, le travail effectué par ce champ de forces le long du petit segment de $\gamma(t_{i-1})$ à $\gamma(t_i)$ est :

$$F(\gamma(t_i)) \cdot (\gamma(t_i) - \gamma(t_{i-1})).$$

Ainsi, on considère l'approximation suivante pour le travail effectué par le champ de forces F le long de la courbe γ :

$$W(\gamma, F, \mathcal{P}) = \sum_{i=1}^k F(\gamma(t_i)) \cdot ((\gamma(t_i) - \gamma(t_{i-1}))).$$

Les étapes suivantes sont tout à fait similaire à la justification de la définition de la longueur (utilisation du Théorème de la valeur intermédiaire, etc...) de sorte qu'on ne les répète pas. Finalement, on trouve :

$$\int_{\gamma} F = \lim_{|\mathcal{P}| \rightarrow 0} W(\gamma, F, \mathcal{P}).$$

Remarque 6.3 Supposons que la courbe paramétrée $\gamma : I \rightarrow \mathbb{R}^n$ soit régulière en plus d'être de classe \mathcal{C}^1 . Alors, on définit le *vecteur tangent normalisé* au point $t \in I$, noté, $T(t)$, par

$$T(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|}.$$

Ainsi, on obtient :

$$\begin{aligned} \int_{\gamma} F &= \int_a^b F(\gamma(t)) \cdot T(t) \|\gamma'(t)\| dt, \\ &= \int_{\gamma} F \cdot T dl, \quad (\text{utilisant les notations de l'intégrale curviligne}). \end{aligned}$$

On se pose de nouveau la question naturelle suivante. Si deux courbes paramétrées ont la même trace, est-ce que l'intégrale du champ de vecteurs F le long de ces courbes est la même ? Dans ce cas, la réponse dépend du "sens" du difféomorphisme- \mathcal{C}^1 .

Théorème 55 Soient $\gamma_1 : I_1 = [a_1, b_1] \rightarrow \mathbb{R}^n$ et $\gamma_2 : I_2 = [a_2, b_2] \rightarrow \mathbb{R}^n$ deux courbes paramétrées de classe \mathcal{C}^1 , telles que $\gamma_1(I_1) = \gamma_2(I_2)$. Soit $F : \gamma_1[a_1, b_1] \rightarrow \mathbb{R}^n$ un champ de vecteurs continu. Supposons que les deux courbes sont équivalentes, et soit $\varphi : I_1 \rightarrow I_2$ le difféomorphisme- \mathcal{C}^1 associé. Alors

$$\int_{\gamma_1} F = \operatorname{sgn}(\varphi') \int_{\gamma_2} F.$$

Preuve:

Comme γ_1 et γ_2 sont équivalentes, et que φ est le difféomorphisme- \mathcal{C}^1 associé, pour tout $t \in I_1$, on a :

$$\gamma_1 = \gamma_2 \circ \varphi.$$

$$\begin{aligned} \int_{\gamma_1} F &= \int_{a_1}^{b_1} F(\gamma_1(t)) \cdot \gamma_1'(t) dt, \text{ (définition)} \\ &= \int_{a_1}^{b_1} F(\gamma_2(\varphi(t))) \cdot \gamma_2'(\varphi(t)) \varphi'(t) dt, \text{ (remplaçant } \gamma_1 \text{ par } \gamma_2 \circ \varphi) \\ &= \int_{\varphi(a_1)}^{\varphi(b_1)} F(\gamma_2(s)) \cdot \gamma_2'(s) ds, \text{ (changement de variables pour fonctions réelles)} \\ &= \operatorname{sgn}(\varphi') \int_{a_2}^{b_2} F(\gamma_2(s)) \cdot \gamma_2'(s) ds = \operatorname{sgn}(\varphi') \int_{\gamma_2} F. \end{aligned}$$

□

Ce théorème nous permet de définir de manière rigoureuse l'intégration d'un champ de vecteurs le long de la trace d'une courbe, sous certaines hypothèses, de la manière suivante.

Dans la suite, nous supposons que les paramétrisations considérées sont simples (i.e. injective sur l'intérieur de leur domaine de définition, autrement dit, elles ne parcourent la trace de la courbe qu'une seule fois). Remarquons que la trace d'une courbe paramétrée $\gamma : [a, b] \rightarrow \mathbb{R}^n$ a une orientation naturelle de $\gamma(a)$ vers $\gamma(b)$.

Soient maintenant $\gamma_1 : I_1 = [a_1, b_1] \rightarrow \mathbb{R}^n$, et $\gamma_2 : I_2 = [a_2, b_2] \rightarrow \mathbb{R}^n$ deux courbes paramétrées régulières et simples qui ont la même trace ($\gamma_1(I_1) = \gamma_2(I_2)$), et qui engendrent la même orientation sur la trace commune. Comme les courbes sont régulières

et simples, par le Théorème 13, chacune des paramétrisations est équivalente à la même paramétrisation par longueur d'arc, de sorte que les deux courbes sont équivalentes. De plus, comme les deux courbes engendrent la même orientation sur leur trace commune, on peut montrer que le difféomorphisme reliant les deux courbes a une dérivée de signe positif. Ainsi, par le Théorème 55, si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est un champ de vecteurs continu, on a :

$$\int_{\gamma_1} F = \int_{\gamma_2} F.$$

Un sous-ensemble C de \mathbb{R}^n est appelé une *courbe régulière orientée*, ssi C est la trace d'une courbe paramétrée $\gamma : [a, b] \rightarrow C$, qui est régulière, munie de l'orientation engendrée par la paramétrisation γ . Par ce que l'on a vu ci-dessus, l'intégrale curviligne d'un champ de vecteurs continu le long de cette courbe ne dépend pas de la paramétrisation γ vérifiant les hypothèses ci-dessus, ainsi il fait sens de définir l'*intégrale curviligne du champ de vecteurs continu F le long de la courbe C* par :

$$\int_C F = \int_{\gamma} F.$$

Exemple 6.2 On souhaite intégrer le champ de vecteurs $F(x, y) = (x^2y, y^3)$ le long du segment I qui va de $(0, 0)$ à $(1, 1)$. Le champ de vecteurs F est continu sur tout \mathbb{R}^2 , donc en particulier sur le segment. De plus le segment I est une courbe régulière orientée, étant donné que c'est la trace de la paramétrisation $\gamma : [0, 1] \rightarrow I$, définie par :

$$\gamma(t) = (t, t),$$

qui est régulière. Ainsi, on a :

$$\int_I F = \int_{\gamma} F = \int_0^1 (t^3 + t^3) dt = \frac{1}{2}.$$

Exemple 6.3 On souhaite intégrer le champ de vecteurs $F(x, y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$ le long du cercle unité C orienté dans le sens trigonométrique. Le champ de vecteurs F est continu sur $\mathbb{R}^2 \setminus \{(0, 0)\}$, donc en particulier sur le cercle unité. De plus le cercle unité orienté dans le sens trigonométrique est une courbe régulière orientée, étant donné que c'est la trace de la paramétrisation $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, définie par :

$$\gamma(\theta) = (\cos \theta, \sin \theta),$$

qui est régulière. Ainsi, on a :

$$\int_C F = \int_{\gamma} F = \int_0^{2\pi} d\theta = 2\pi.$$

Remarque 6.4 Soit C une courbe régulière orientée. Alors la courbe C^- , qui est la courbe C orientée dans le sens opposé, est aussi régulière, et :

$$\int_C F = - \int_{C^-} F.$$

En effet, comme C est une courbe régulière orientée, il existe une paramétrisation régulière $\gamma : [a, b] \rightarrow C$ qui engendre l'orientation de C . Définissons $\gamma^- : [b, a] \rightarrow C$ par $\gamma^-(t) = \gamma(a + b - t)$. Alors γ^- est aussi régulière, et la trace C est orientée de $\gamma(b)$ vers $\gamma(a)$, donc C^- est une courbe régulière orientée dans le sens opposé. De plus, les paramétrisations γ et γ^- sont équivalentes, avec $\varphi(t) = a + b - t$. En utilisant le Théorème 55, on déduit :

$$\int_C F = - \int_{C^-} F.$$

6.3 Théorème de Green

Dans cette section, on se restreint à \mathbb{R}^2 .

6.3.1 Le domaine d'intégration en question

Un sous-ensemble C de \mathbb{R}^2 est appelé un *chemin de \mathbb{R}^2* , ssi $C = \{C_1, \dots, C_m\}$, et les C_i sont des courbes régulières orientées "injectives" deux à deux disjointes, comme sur la figure ci-dessous (il est plus facile de donner la définition sur un dessin, car l'écrire exactement créerait plus de confusion qu'autre chose).

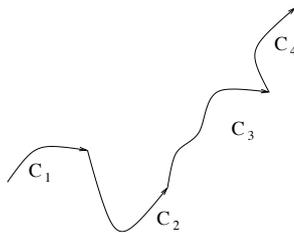


FIG. 6.1 – Un chemin de \mathbb{R}^2 .

Soit $F : C \rightarrow \mathbb{R}^2$ un champ de vecteurs continu. Alors, on définit l'*intégrale curviligne du champ de vecteurs le long du chemin C* , notée $\int_C F$, par :

$$\int_C F = \int_{C_1} F + \dots + \int_{C_m} F.$$

Un sous-ensemble de \mathbb{R}^2 est appelé un *chemin fermé*, ssi c'est un chemin, et que le point initial de C_1 et le point final de C_m sont identifiés.

Les domaines que nous considérons pour l'énoncé du Théorème de Green sont de la forme suivante : D est un sous-ensemble de \mathbb{R}^2 , connexe, compact (fermé et borné), dont le bord ∂D , est la réunion d'un nombre fini de chemins fermés.

En utilisant le Théorème de Jordan ("Toute courbe de Jordan sépare le plan en deux composantes connexes, une bornée et une non-bornée". C'est un théorème difficile dont la première preuve rigoureuse date de 1905), on déduit que parmi toutes les courbes fermées formant ∂D , il en existe une unique qui contient toutes les autres dans son intérieur. On appelle cette courbe le *bord extérieur*, et les autres (si elles existent), le *bord intérieur*. Des exemples de tels domaines sont illustrés sur la Figure 6.2.

Pour un tel domaine, on demande encore que les chemins fermés de ∂D soient orientés de sorte que $\text{int}(D)$ soit toujours à gauche, voir Figure 6.2.

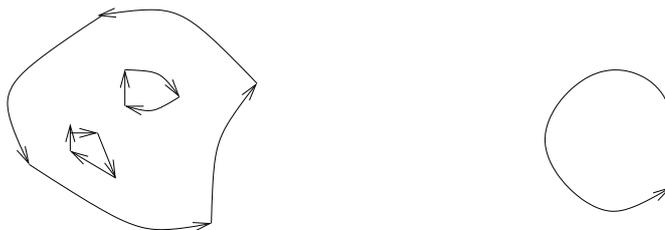


FIG. 6.2 – Exemples de domaines admissibles pour le Théorème de Green.

Nous appelons un domaine vérifiant les propriétés ci-dessus, un *domaine admissible pour le Théorème de Green*.

6.3.2 Le Théorème de Green

Théorème 56 (Green) Soit D un domaine admissible pour le Théorème de Green, et soit $F : U \rightarrow \mathbb{R}^2$ un champ de vecteurs de classe \mathcal{C}^1 , où U est un ouvert de \mathbb{R}^2 contenant D . Alors :

$$\int_{\partial D} F = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right), \quad (6.1)$$

où F_1, F_2 sont les deux composantes du champ de vecteurs F .

Remarque 6.5

1. Un domaine admissible pour le Théorème de Green est aussi un domaine admissible pour l'intégrale de Riemann, de sorte que le membre de droite de (6.1) est bien défini.

2. Si F est un champ de vecteurs comme dans l'énoncé du théorème, qui satisfait

$$\frac{\partial F_2}{\partial x} = \frac{\partial F_1}{\partial y},$$

alors son intégrale sur ∂D est nulle.

Exemple 6.4 Trouver l'intégrale du champ de vecteurs $F(x, y) = (3x + y, -x + 2y)$ le long de l'ellipse $4x^2 + y^2 = 4$, orientée dans le sens trigonométrique.

Soit E l'ensemble formé de l'ellipse et de son intérieur. Le champ de vecteurs F est de classe C^1 sur \mathbb{R}^2 , et E est un domaine admissible pour le Théorème de Green, donc on peut l'appliquer :

$$\int_{\partial E} F = \int_E (-2) = -2A(E) = -4\pi.$$

Preuve:

Démontrons ce théorème dans un cas particulier. On se place sous les hypothèses du théorème, et on suppose que :

- $F(x, y) = (F_1(x, y), 0)$,
- $D = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$, où $g_1 \leq g_2$ sont de classe C^1 .

Appelons C_1, C_2, C_3, C_4 les courbes comme sur la Figure 6.3 ci-dessous.

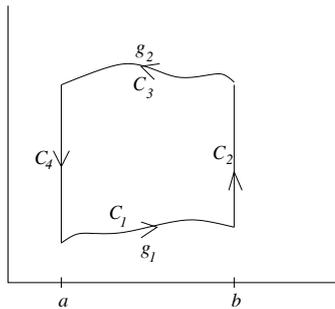


FIG. 6.3 – Les courbes C_1, C_2, C_3, C_4 .

Soit γ_i une paramétrisation régulière correspondant à C_i :

$$\begin{aligned} \gamma_1 : [a, b] &\rightarrow \mathbb{R}^2, \quad \gamma_1(t) = (t, g_1(t)), \\ \gamma_2 : [g_1(b), g_2(b)] &\rightarrow \mathbb{R}^2, \quad \gamma_2(t) = (b, t). \end{aligned}$$

Alors,

$$\int_{C_1} F = \int_a^b F_1(t, g_1(t)) dt$$

$$\int_{C_2} F = \int_{g_1(b)}^{g_2(b)} 0 dt = 0.$$

Et de manière similaire :

$$\int_{C_3} F = - \int_a^b F_1(t, g_2(t)) dt$$

$$\int_{C_4} F = 0.$$

De plus,

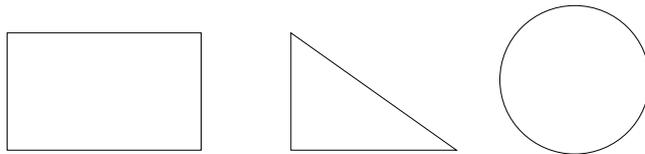
$$\int_D \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} - \frac{\partial F_1}{\partial y} dy \right) dx, \text{ (Théorème de Fubini),}$$

$$= \int_a^b F_1(x, g_1(x)) - F_1(x, g_2(x)) dx.$$

De manière analogue, on peut montrer le théorème dans le cas où :

- $F(x, y) = (0, F_2(x, y))$,
- $D = \{(x, y) \in \mathbb{R}^2 \mid g_1(y) \leq x \leq g_2(y), c \leq y \leq d\}$, où $g_1 \leq g_2$ sont de classe \mathcal{C}^1 .

La combinaison de ces deux cas particuliers donne une preuve du Théorème de Green, lorsque le domaine est à la fois du premier type, et du deuxième. Par exemple :



□

Chapitre 7

Intégrales de surface - Théorèmes de la divergence et de Stokes

Dans une première partie, nous introduisons les surfaces paramétrées. C'est l'analogue des courbes paramétrées, lorsque le domaine de départ est un sous-ensemble de \mathbb{R}^2 , et non un intervalle de \mathbb{R} . Ensuite, nous définissons l'intégration sur les surfaces paramétrées. Comme dans le cas des courbes paramétrées, il existe deux sortes d'objets que l'on peut intégrer : les fonctions et les champs de vecteurs définis sur la surface. Finalement, nous étudions le Théorème de la divergence et le Théorème de Stokes.

7.1 Surfaces paramétrées

7.1.1 Définition

Soit D un sous-ensemble admissible de \mathbb{R}^2 . Souvenez-vous que "admissible" signifie que D est un sous-ensemble borné de \mathbb{R}^2 , dont le bord est négligeable (voir le chapitre "Intégrales multiples"). Une *surface paramétrée* Σ est une application continue $\Sigma : D \rightarrow \mathbb{R}^n$. En coordonnées, Σ s'écrit :

$$\Sigma(u, v) = (\Sigma_1(u, v), \dots, \Sigma_n(u, v)).$$

L'ensemble $\Sigma(D)$ s'appelle la *trace de la surface* Σ . Une surface est dite *simple*, ssi l'application Σ est injective.

Exemple 7.1 La surface paramétrée suivante est la sphère de rayon r dans \mathbb{R}^3 :

$$\begin{aligned} \Sigma : D = [0, \pi] \times [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ (\varphi, \theta) &\mapsto \Sigma(\varphi, \theta) = (r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi). \end{aligned}$$

Exemple 7.2 Soient $0 < b < a$. La surface paramétrée suivante est le tore de paramètres a, b dans \mathbb{R}^3 :

$$\begin{aligned} \Sigma : D = [0, 2\pi] \times [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ (\theta, \varphi) &\mapsto \Sigma(\theta, \varphi) = ((a + b \cos \varphi) \cos \theta, (a + b \cos \varphi) \sin \theta, b \sin \varphi). \end{aligned}$$

A partir de maintenant, nous nous restreignons à l'étude du cas $n = 3$, c'est-à-dire aux surfaces paramétrées dans \mathbb{R}^3 , ceci pour deux raisons. D'une part, un traitement plus général nécessiterait plus de temps ; et d'autre part, le but est d'étudier le Théorème de la divergence de Gauss-Ostrogradsky et le Théorème de Stokes, pour lesquels l'étude des surfaces paramétrées de \mathbb{R}^3 est suffisante.

7.1.2 Rappel sur le produit vectoriel

Dans la suite, nous définissons le vecteur normal à une surface. Pour ce faire, nous avons besoin du produit vectoriel. Voici donc un bref rappel.

Soit $\mathbf{x} = (x_1, x_2, x_3)$ et $\mathbf{y} = (y_1, y_2, y_3)$ deux vecteurs de \mathbb{R}^3 . Alors, le *produit vectoriel* de \mathbf{x} et \mathbf{y} , noté $\mathbf{x} \wedge \mathbf{y}$, est par définition :

$$\mathbf{x} \wedge \mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

Le produit vectoriel satisfait les propriétés suivantes, dont les preuves sont supposées connues, ou laissées en exercice.

Proposition 57

1. $\mathbf{x} \wedge \mathbf{y} = -\mathbf{y} \wedge \mathbf{x}$.
2. $\mathbf{x} \wedge (\mathbf{y} + \mathbf{z}) = \mathbf{x} \wedge \mathbf{y} + \mathbf{x} \wedge \mathbf{z}$.
3. Pour tout $a \in \mathbb{R}$, $(a\mathbf{x}) \wedge \mathbf{y} = a(\mathbf{x} \wedge \mathbf{y}) = \mathbf{x} \wedge (a\mathbf{y})$.
4. $(\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z} = (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{y} \cdot \mathbf{z})\mathbf{x}$.
5. $\mathbf{x} \wedge \mathbf{y}$ est perpendiculaire à \mathbf{x} et \mathbf{y} .
6. $\|\mathbf{x} \wedge \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| |\sin \theta|$, où θ est l'angle entre les vecteurs \mathbf{x} et \mathbf{y} .

7.1.3 Dérivabilité et plan tangent

Soient D un sous-ensemble admissible de \mathbb{R}^2 , et $\Sigma : D \rightarrow \mathbb{R}^3$ une surface paramétrée. La surface est dite *de classe \mathcal{C}^1* , ssi toutes les dérivées partielles de toutes les composantes de Σ existent en tout point d'un ouvert contenant D et sont continues (on suppose donc implicitement que Σ est définie sur un ouvert contenant D). Une surface de classe \mathcal{C}^1 est dite *lisse*, ssi la matrice jacobienne de Σ au point (u, v) , $J_\Sigma(u, v)$, est de rang 2

pour tout $(u, v) \in \text{int}(D)$. La notion de surface lisse est l'analogie de celle de courbe régulière.

Soit $\Sigma : D \rightarrow \mathbb{R}^3$ une surface lisse, et soit $(u, v) \in D$. Alors, on peut visualiser les vecteurs

$$D_1\Sigma(u, v) = \frac{\partial \Sigma}{\partial u} = \begin{pmatrix} \frac{\partial \Sigma_1}{\partial u} \\ \frac{\partial \Sigma_2}{\partial u} \\ \frac{\partial \Sigma_3}{\partial u} \end{pmatrix}, \text{ et } D_2\Sigma(u, v) = \frac{\partial \Sigma}{\partial v} = \begin{pmatrix} \frac{\partial \Sigma_1}{\partial v} \\ \frac{\partial \Sigma_2}{\partial v} \\ \frac{\partial \Sigma_3}{\partial v} \end{pmatrix},$$

comme étant les vecteurs tangents à la surface au point $\Sigma(u, v)$. On définit le *plan tangent à la surface Σ au point $\Sigma(u, v)$* , comme étant le plan engendré par les deux vecteurs tangents. Comme la surface est lisse, pour tout point $(u, v) \in \text{int}(D)$, ces deux vecteurs ne sont pas colinéaires, et donc le plan tangent est de dimension 2.

Exemple 7.3 Les vecteurs tangents de la sphère de rayon r de l'Exemple 7.1, sont :

$$D_1\Sigma(\varphi, \theta) = \begin{pmatrix} r \cos \theta \cos \varphi \\ r \sin \theta \cos \varphi \\ -r \sin \varphi \end{pmatrix}, \text{ et } D_2\Sigma(\varphi, \theta) = \begin{pmatrix} -r \sin \theta \sin \varphi \\ r \cos \theta \sin \varphi \\ 0 \end{pmatrix}.$$

Soit $\Sigma : D \rightarrow \mathbb{R}^3$ une surface de classe \mathcal{C}^1 , et soit $(u, v) \in D$, alors on définit le *vecteur normal à la surface Σ au point $\Sigma(u, v)$* , noté $N(u, v)$, par :

$$N(u, v) = D_1\Sigma(u, v) \wedge D_2\Sigma(u, v).$$

Si la surface Σ est de plus supposée lisse, alors le vecteur normal est non nul en tout point $(u, v) \in \text{int}(D)$, et on définit le *vecteur normal unitaire*, noté $N^1(u, v)$, par

$$N^1(u, v) = \frac{N(u, v)}{\|N(u, v)\|}.$$

Exemple 7.4 Le vecteur normal de la sphère de rayon r de l'Exemple 7.1, est :

$$N(\varphi, \theta) = r \sin \varphi \Sigma(\varphi, \theta).$$

Comme $r > 0$ et $\sin \varphi \geq 0$, le vecteur normal a la même direction que le vecteur $\Sigma(\varphi, \theta)$, et ainsi pointe en dehors de la sphère.

En tout point (φ, θ) de $\text{int}(D)$, le vecteur normal est non nul, et le vecteur normal unitaire est :

$$N^1(\varphi, \theta) = \frac{1}{r} \Sigma(\varphi, \theta).$$

Remarquer que le plan tangent au point $\Sigma(u, v)$, est aussi le plan orthogonal au vecteur $N(u, v)$, passant par ce point.

Remarque 7.1 Soit $\Sigma : D \rightarrow \mathbb{R}^3$ une surface paramétrée lisse. Supposons qu'il existe une fonction $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 telle que $\Sigma(D) = \{(x, y, z) \in \mathbb{R}^3 \mid f(x, y, z) = 0\}$. Alors, le vecteur normal $N(u_0, v_0)$ est colinéaire à $\nabla f(\mathbf{x}_0)$, le gradient de f au point $\mathbf{x}_0 = \Sigma(u_0, v_0)$, lorsqu'il ne s'annule pas. De plus, à la Section ??, nous avons défini le plan tangent au point \mathbf{x}_0 d'une surface donnée sous cette forme, comme étant le sous-espace passant par \mathbf{x}_0 et orthogonal au gradient en ce point. Ainsi, les deux définitions de plan tangent sont cohérentes.

7.2 Intégrale de surface d'une fonction

Dans cette section, nous intégrons une fonction sur une surface paramétrée. Lorsque la fonction est identiquement égale à 1, l'intégrale en question représente l'aire de la surface. Cette section est l'analogue de l'intégrale curviligne d'une fonction définie le long d'une courbe.

Soit $\Sigma : D \rightarrow \mathbb{R}^3$ une surface paramétrée de classe \mathcal{C}^1 , où D est un sous-ensemble admissible de \mathbb{R}^2 , et soit $f : \Sigma(D) \rightarrow \mathbb{R}$ une fonction continue. Alors l'intégrale de la fonction f sur la surface Σ , notée $\int_{\Sigma} f \, d\sigma$, est par définition :

$$\int_{\Sigma} f \, d\sigma = \int_D f(\Sigma(u, v)) \|N(u, v)\| \, du \, dv. \quad (7.1)$$

Cette intégrale est appelée *intégrale de surface de la fonction f sur la surface Σ* . Remarquer que comme la surface Σ est de classe \mathcal{C}^1 sur un ouvert U contenant D , et que D est un domaine admissible par définition, l'intégrale de Riemann ci-dessus est bien définie.

L'aire de la surface Σ , notée $A(\Sigma)$, est par définition :

$$A(\Sigma) = \int_{\Sigma} d\sigma.$$

Remarque 7.2 *Justification de la définition de l'aire.*

Donnons une heuristique justifiant la définition de l'aire lorsque le domaine est un pavé P de \mathbb{R}^2 . La justification de la définition avec une fonction f est tout à fait similaire. Soit $\mathcal{P} = \{P_1, \dots, P_r\}$ une partition de P en petits pavés, et soit \mathbf{u}_i le centre du pavé P_i . Alors, l'aire de la surface Σ est donnée par :

$$A(\Sigma, \mathcal{P}) = \sum_{i=1}^r A(\Sigma(P_i)).$$

Comme les pavés P_i sont petits, et que Σ est de classe \mathcal{C}^1 , on peut approcher Σ sur le pavé P_i par sa différentielle au point \mathbf{u}_i , $d\Sigma_{\mathbf{u}_i}$:

$$A(\Sigma, \mathcal{P}) \sim \sum_{i=1}^r A(d\Sigma_{\mathbf{u}_i}(P_i)).$$

A translation près, le pavé P_i peut s'écrire $P_i = [0, a_i] \times [0, b_i]$, de sorte que $A(P_i) = a_i b_i$. Ainsi,

$$\begin{aligned} d\Sigma_{\mathbf{u}_i}(P_i) &= \{\alpha d\Sigma_{\mathbf{u}_i}(\mathbf{e}_1) + \beta d\Sigma_{\mathbf{u}_i}(\mathbf{e}_2) \mid 0 \leq \alpha \leq a_i, 0 \leq \beta \leq b_i\}, \\ &= \{\alpha D_1\Sigma(\mathbf{u}_i) + \beta D_2\Sigma(\mathbf{u}_i) \mid 0 \leq \alpha \leq a_i, 0 \leq \beta \leq b_i\}. \end{aligned}$$

Dans la deuxième égalité, on a utilisé le fait que sous ces hypothèses, la matrice correspondante à la différentielle est la jacobienne.

Ainsi l'aire du pavé $d\Sigma_{\mathbf{u}_i}(P_i)$ est donnée par :

$$A(d\Sigma_{\mathbf{u}_i}(P_i)) = a_i b_i \|D_1\Sigma(\mathbf{u}_i)\| \|D_2\Sigma(\mathbf{u}_i)\| |\sin \theta|,$$

où θ est l'angle entre les vecteurs $D_1\Sigma(\mathbf{u}_i)$ et $D_2\Sigma(\mathbf{u}_i)$. Par la propriété 6 du produit vectoriel, on a :

$$\|D_1\Sigma(\mathbf{u}_i)\| \|D_2\Sigma(\mathbf{u}_i)\| |\sin \theta| = \|D_1\Sigma(\mathbf{u}_i) \wedge D_2\Sigma(\mathbf{u}_i)\|.$$

Ainsi,

$$A(\Sigma, \mathcal{P}) = \sum_{i=1}^r \|D_1\Sigma(\mathbf{u}_i) \wedge D_2\Sigma(\mathbf{u}_i)\| a_i b_i = \sum_{i=1}^r \|N(\mathbf{u}_i)\| A(P_i),$$

qui est une somme de Riemann pour (7.1).

Exemple 7.5 Calculons l'aire de la sphère S^2 de l'Exemple 7.1. En utilisant la définition on a :

$$\begin{aligned} A(S^2) &= \int_S \|N(\varphi, \theta)\| d\varphi d\theta, \\ &= \int_0^\pi \left(\int_0^{2\pi} r^2 \sin \varphi d\theta \right) d\varphi, \quad (\text{Théorème de Fubini}) \\ &= 2\pi r^2 \int_0^\pi \sin \varphi d\varphi = 4\pi r^2. \end{aligned}$$

Exemple 7.6 Soit D un sous-ensemble admissible de \mathbb{R}^2 , et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 , où U est un ouvert contenant D . Alors le graphe de f est la trace de la surface paramétrée $\Sigma : D \rightarrow \mathbb{R}^3$, définie par :

$$\Sigma(x, y) = (x, y, f(x, y)).$$

Le vecteur normal $N(x, y)$ est :

$$N(x, y) = \begin{pmatrix} 1 \\ 0 \\ D_1 f(x, y) \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ D_2 f(x, y) \end{pmatrix} = \begin{pmatrix} -D_1 f(x, y) \\ -D_2 f(x, y) \\ 1 \end{pmatrix}.$$

Ainsi, l'aire du graphe de la fonction est donnée par :

$$A(\Sigma) = \int_D \sqrt{1 + D_1 f(x, y)^2 + D_2 f(x, y)^2} dx dy.$$

On peut maintenant se poser la même question que pour les courbes paramétrées. Si deux surfaces paramétrées de classe \mathcal{C}^1 ont la même trace, est-ce qu'elles ont même aire ? La réponse est positive dans le cas suivant.

Soit $\Sigma_1 : D_1 \rightarrow \mathbb{R}^3$, $\Sigma_2 : D_2 \rightarrow \mathbb{R}^3$ deux surfaces paramétrées. On dit que Σ_1 et Σ_2 sont *équivalentes*, ssi il existe une application $T : D_1 \rightarrow D_2$, bijective sur D_1 , qui est un difféomorphisme- \mathcal{C}^1 de $\text{int}(D_1)$ sur $\text{int}(D_2)$, telle que :

$$\Sigma_1 = \Sigma_2 \circ T.$$

Théorème 58 Soient $\Sigma_1 : D_1 \rightarrow \mathbb{R}^3$, $\Sigma_2 : D_2 \rightarrow \mathbb{R}^3$ deux surfaces paramétrées de classe \mathcal{C}^1 supposées équivalentes. Soit $f : \Sigma_1(D_1) \rightarrow \mathbb{R}$ une fonction continue. Alors,

$$\int_{\Sigma_1} f d\sigma = \int_{\Sigma_2} f d\sigma.$$

Preuve:

Comme Σ_1 et Σ_2 sont équivalentes, il existe une application $T : D_1 \rightarrow D_2$, bijective sur D_1 , qui est un difféomorphisme- \mathcal{C}^1 sur $\text{int}(D_1)$, telle que :

$$\Sigma_1 = \Sigma_2 \circ T.$$

Par définition, on a :

$$\int_{\Sigma_1} f d\sigma = \int_{D_1} f(\Sigma_1(u, v)) \|D_1 \Sigma_1(u, v) \wedge D_2 \Sigma_1(u, v)\| du dv.$$

Aux exercices, vous démontrerez que :

$$D_1 \Sigma_1(u, v) \wedge D_2 \Sigma_1(u, v) = D_1 \Sigma_2(T(u, v)) \wedge D_2 \Sigma_2(T(u, v)) \det(J_T(u, v)).$$

Ainsi :

$$\int_{\Sigma_1} f d\sigma = \int_{D_1} f \circ \Sigma_2(T(u, v)) \|D_1 \Sigma_2(T(u, v)) \wedge D_2 \Sigma_2(T(u, v))\| |\det(J_T(u, v))| du dv.$$

Comme l'application T est un difféomorphisme- \mathcal{C}^1 sur $\text{int}(D_1)$ qui est admissible, et que f est supposée continue, toutes les hypothèses du Théorème 53 (changement de variables) sont vérifiées, de sorte que :

$$\int_{\Sigma_1} f \, d\sigma = \int_{D_2} f(\Sigma_2(s, t)) \|D_1 \Sigma_2(s, t) \wedge D_2 \Sigma_2(s, t)\| \, ds \, dt = \int_{\Sigma_2} f \, d\sigma.$$

□

7.3 Intégrale de surface d'un champ de vecteurs

Dans cette section, nous intégrons un champ de vecteurs sur une surface paramétrée.

Soit $\Sigma : D \rightarrow \mathbb{R}^3$ une surface paramétrée de classe \mathcal{C}^1 , et soit $F : \Sigma(D) \rightarrow \mathbb{R}^3$ un champ de vecteurs continu. D'un point de vue physique il faut se représenter la surface comme étant une membrane, et le champ de vecteurs comme donnant les vecteurs vitesse de l'écoulement d'un fluide à travers la membrane. L'intégrale du champ de vecteurs F sur la surface Σ , notée $\int_{\Sigma} F$, est par définition le flux de liquide à travers la membrane. Mathématiquement, cela se traduit de la manière suivante :

$$\int_{\Sigma} F = \int_D F(\Sigma(u, v)) \cdot N(u, v) \, du \, dv.$$

Cette intégrale est appelée *intégrale de surface du champ de vecteurs F sur la surface Σ* .

La justification de cette définition est dans l'esprit de la justification de l'intégrale d'un champ de vecteurs le long d'une courbe, de sorte que nous ne la répétons pas ici.

Le théorème suivant donne la relation entre les intégrales de surface d'un champ de vecteurs sur deux surfaces équivalentes.

Remarque 7.3 Soient D_1 et D_2 deux domaines admissibles de \mathbb{R}^2 , et soit $T : \text{int}(D_1) \rightarrow \text{int}(D_2)$ un difféomorphisme- \mathcal{C}^1 . Alors la matrice jacobienne $J_T(u, v)$ de T est inversible en tout point (u, v) de $\text{int}(D_1)$, de sorte que le jacobien $\det(J_T(u, v))$ est non nul en tout point (u, v) de $\text{int}(D_1)$. De plus, comme T est un difféomorphisme- \mathcal{C}^1 , le jacobien $\det(J_T(u, v))$ est une application continue sur $\text{int}(D_1)$, ainsi il est soit strictement positif, soit strictement négatif sur $\text{int}(D_1)$. On définit $\text{sgn}(\det(J_T))$ comme étant $+1$ si le jacobien est strictement positif, et -1 si il est strictement négatif.

Théorème 59 Soient $\Sigma_1 : D_1 \rightarrow \mathbb{R}^3$, $\Sigma_2 : D_2 \rightarrow \mathbb{R}^3$ deux surfaces paramétrées de classe \mathcal{C}^1 , supposées équivalentes, et soit $T : D_1 \rightarrow D_2$ le difféomorphisme- \mathcal{C}^1 associé.

Soit $F : \Sigma_1(D_1) \rightarrow \mathbb{R}^3$ un champ de vecteurs continu, alors :

$$\int_{\Sigma_1} F = \operatorname{sgn}(\det(J_T)) \int_{\Sigma_2} F.$$

Preuve:

Comme Σ_1 et Σ_2 sont équivalentes, il existe une application $T : D_1 \rightarrow D_2$, bijective sur D_1 , qui est un difféomorphisme- \mathcal{C}^1 sur $\operatorname{int}(D_1)$, telle que :

$$\Sigma_1 = \Sigma_2 \circ T.$$

Par définition, on a :

$$\int_{\Sigma_1} F = \int_{D_1} F(\Sigma_1(u, v)) \cdot (D_1 \Sigma_1(u, v) \wedge D_2 \Sigma_1(u, v)) \, du \, dv$$

De manière analogue à la preuve du Théorème 58, on obtient :

$$\begin{aligned} \int_{\Sigma_1} F &= \int_{D_1} F \circ \Sigma_2(T(u, v)) \cdot [D_1 \Sigma_2(T(u, v)) \wedge D_2 \Sigma_2(T(u, v))] \det(J_T(u, v)) \, du \, dv, \\ &= \operatorname{sgn}(\det(J_T)) \int_{D_1} F \circ \Sigma_2(T(u, v)) \cdot [D_1 \Sigma_2(T(u, v)) \wedge D_2 \Sigma_2(T(u, v))] |\det(J_T(u, v))| \, du \, dv. \end{aligned}$$

Le Théorème du changement de variables permet à nouveau de conclure. \square

Comme dans le cas des courbes, ce théorème nous permet de définir l'intégration d'un champ de vecteurs sur la trace d'une surface, sous certaines hypothèses, de la manière suivante.

Nous introduisons d'abord le concept de *surface orientable*. Ceci est fait de manière non rigoureuse pour deux raisons : tout d'abord c'est une notion qui est intuitive, et deuxièmement la rendre rigoureuse demande d'introduire les variétés différentiables, une notion qui est en dehors de l'objectif de ce cours, et qui appartient à un cours de géométrie différentielle.

Une surface S dans \mathbb{R}^3 est dite *orientable*, ssi elle a "deux côtés". Un des deux côtés est nommé *extérieur*, l'autre *intérieur*. Lorsqu'il n'y a pas d'ambiguïté, l'extérieur (resp. intérieur) est le côté qui correspond à la notion intuitive d'extérieur (resp. intérieur). Lorsque ce n'est pas clair, on précisera quel côté est quoi.

Exemple 7.7 La sphère, le tore dans \mathbb{R}^3 sont des surfaces orientables. Le ruban de Möbius dans \mathbb{R}^3 est l'exemple type de surface non-orientable.

Dans la suite nous supposons que les paramétrisations de surface que nous considérons sont injectives sur l'intérieur de leur domaine de définition.

Soient maintenant $\Sigma_1 : D_1 \rightarrow \Sigma(D_1)$ et $\Sigma_2 : D_2 \rightarrow \Sigma(D_2)$ deux paramétrisations lisses d'une même surface orientable. On suppose de plus que le vecteur normal des deux paramétrisations pointe vers l'extérieur de la surface. Alors, on peut montrer que ces deux paramétrisations sont équivalentes, et que le jacobien du difféomorphisme- \mathcal{C}^1 les reliant est de signe positif. Ainsi, si $F : \Sigma_1(D_1) \rightarrow \mathbb{R}^3$ est un champ de vecteurs continu, on a :

$$\int_{\Sigma_1} F = \int_{\Sigma_2} F.$$

Un sous-ensemble S de \mathbb{R}^3 est appelée une *surface lisse orientée*, ssi S est une surface orientable, et S est la trace d'une surface paramétrée $\Sigma : D \rightarrow \mathbb{R}^3$ lisse, dont le vecteur normal pointe vers l'extérieur de S en tout point de $\text{int}(D)$. Par ce que l'on a vu ci-dessus, l'intégrale de surface d'un champ de vecteurs F ne dépend pas de la paramétrisation Σ vérifiant les hypothèses ci-dessus, ainsi il fait sens de définir l'*intégrale de surface du champ de vecteurs F sur la surface S* par :

$$\int_S F = \int_{\Sigma} F.$$

Exemple 7.8 On souhaite intégrer le champ de vecteurs $F(x, y, z) = (x, y, 0)$ sur la sphère de rayon r ($r > 0$) dans \mathbb{R}^3 , notée S_r . Le champ de vecteurs F est continu sur tout \mathbb{R}^3 , donc en particulier sur S_r . De plus, la sphère S_r est une surface lisse orientée, étant donné que c'est la trace de la paramétrisation coordonnées sphériques $\Sigma : D = [0, 2\pi] \times [0, \pi] \rightarrow \mathbb{R}^3$ qui est lisse, et dont le vecteur normal pointe bien vers l'extérieur de la sphère S_r . Ainsi,

$$\int_{S_r} F = \int_{\Sigma} F = r^3 \int_0^{2\pi} \int_0^{\pi} \sin^3 \varphi \, d\varphi = \frac{8\pi r^3}{3}.$$

7.4 Théorème de la divergence et Théorème de Stokes

7.4.1 Divergence et rotationnel

Soit $F : U \rightarrow \mathbb{R}^3$ un champ de vecteurs dérivable, où U est un ouvert de \mathbb{R}^3 . En coordonnées F s'écrit :

$$F(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)).$$

La *divergence* de F , notée $\text{div}(F)$, est par définition la fonction :

$$\text{div}(F) = D_1 F_1 + D_2 F_2 + D_3 F_3 = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}.$$

Le *rotationnel* de F , noté $\text{rot}(F)$, est par définition le champ de vecteurs :

$$\begin{aligned}\text{rot}(F) &= (D_2F_3 - D_3F_2, D_3F_1 - D_1F_3, D_1F_2 - D_2F_1) \\ &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right).\end{aligned}$$

Remarque 7.4 On utilise souvent la notation suivante :

$$\nabla = (D_1, D_2, D_3) = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Ainsi, la divergence et le rotationnel peuvent aussi s'écrire :

$$\begin{aligned}\text{div}(F) &= \nabla \cdot F \\ \text{rot}(F) &= \nabla \wedge F\end{aligned}$$

Exemple 7.9 Soit $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ le champ de vecteurs défini par

$$F(x, y, z) = (\sin(xy), e^{xz}, 2x + yz^4),$$

alors :

$$\begin{aligned}\text{div}(F) &= y \cos(xy) + 4yz^3 \\ \text{rot}(F) &= (z^4 - xe^{xz}, -2, ze^{xz} - x \cos(xy)).\end{aligned}$$

Ces deux opérateurs satisfont les propriétés suivantes que vous démontrerez en exercice.

Proposition 60 Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 , où U est un ouvert de \mathbb{R}^3 , alors :

1. $\text{rot}(\text{grad}(f)) = \nabla \wedge \nabla f = 0$.
2. $\text{div}(\text{grad}(f)) = \nabla \cdot \nabla f = \Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$.

Soient $F, G : U \rightarrow \mathbb{R}^3$ deux champs de vecteurs de classe \mathcal{C}^2 , où U est un ouvert de \mathbb{R}^3 , alors :

1. $\text{div}(\text{rot}(F)) = \nabla \cdot (\nabla \wedge F) = 0$.
2. $\text{rot}(\text{rot}(F)) = \nabla \wedge (\nabla \wedge F) = \text{grad}(\text{div}(F)) - \Delta F$.
3. $\text{div}(F \wedge G) = \nabla \cdot (F \wedge G) = G \cdot (\nabla \wedge F) - F \cdot (\nabla \wedge G)$.

Remarque 7.5 D'un point de vue physique, la divergence d'un champ de vecteurs mesure le flux du champ en un point donné. Le rotationnel d'un champ de vecteurs décrit la rotation engendrée par ce champ.

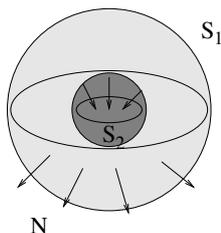
7.4.2 Les domaines d'intégration en question

• Volume admissible pour le Théorème de la divergence

Un *volume admissible pour le Théorème de la divergence*, est un sous-ensemble V de \mathbb{R}^3 , connexe, compact, dont le bord ∂V est une surface orientable dans \mathbb{R}^3 , qui est réunion d'un nombre fini de surfaces lisses orientées.

Exemple 7.10

1. V est un pavé de \mathbb{R}^3 .
2. V est la boule de rayon r dans \mathbb{R}^3 (la boule est la sphère et son intérieur).
3. $V = \{(x, y, z) \in \mathbb{R}^3 \mid z \geq x^2 + y^2, 0 \leq z \leq 2\}$, c'est-à-dire que V est un parabolôïde avec son intérieur.
4. V est le volume compris entre les deux boules comme sur la figure ci-dessous.



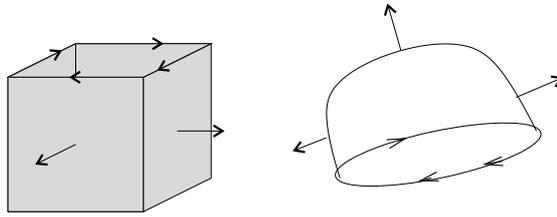
• Surface admissible pour le Théorème de Stokes

Pour le Théorème de Green, on a défini ce qu'est un chemin de \mathbb{R}^2 , et un chemin fermé de \mathbb{R}^2 . La définition d'un chemin de \mathbb{R}^3 , et d'un chemin fermé de \mathbb{R}^3 est tout à fait similaire, on remplace simplement \mathbb{R}^2 par \mathbb{R}^3 .

Une *surface admissible pour le Théorème de Stokes*, est une surface orientable S de \mathbb{R}^3 , connexe, compacte, qui est réunion d'un nombre fini de surfaces lisses orientées de \mathbb{R}^3 . On suppose de plus que le bord ∂S de S , est réunion finie de chemins fermés de \mathbb{R}^3 . On oriente le bord de sorte à ce que la surface S soit à gauche.

Exemple 7.11

1. S est l'hémisphère nord de la sphère de rayon r dans \mathbb{R}^3 .
2. La figure ci-dessous donne deux exemples de surfaces admissibles pour le Théorème de Stokes.



7.4.3 Théorème de la divergence

Théorème 61 (divergence) Soit V un volume admissible pour le Théorème de la divergence. Soit $F : U \rightarrow \mathbb{R}^3$ un champ de vecteurs de classe C^1 , défini sur un ouvert U contenant V . Alors :

$$\int_{\partial V} F = \int_V \operatorname{div}(F).$$

Remarque 7.6

1. Ce Théorème s'appelle aussi Théorème de la divergence de Gauss, ou Théorème de la divergence de Gauss-Ostrogradsky.
2. Par définition, ∂V est réunion d'un nombre fini de surfaces lisses orientées, disons S_1, \dots, S_m . On définit donc naturellement :

$$\int_{\partial V} F = \int_{S_1} F + \dots + \int_{S_m} F.$$

Chacun des termes est alors une intégrale de surface du champ de vecteurs F , définie à la fin de la Section 7.3.

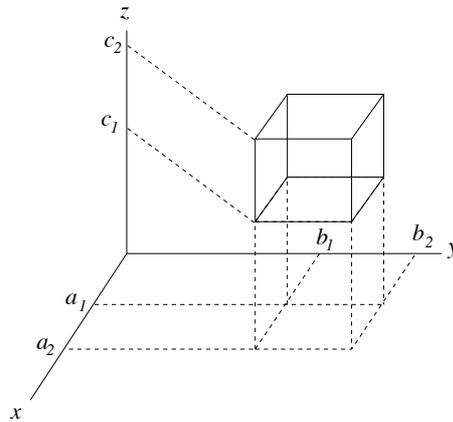
3. Le terme de droite est une intégrale de Riemann multiple en dimension 3. Elle est bien définie car l'intégrand est continu sur son domaine de définition, et le domaine d'intégration est admissible au sens de Riemann.
4. Souvent, le Théorème de la divergence est utilisé pour calculer l'intégrale de surface d'un champ de vecteurs. En effet, l'intégrale de Riemann en dimension 3 est usuellement plus facile à évaluer.

Preuve:

Nous démontrons le Théorème de la divergence dans le cas particulier où V est le pavé $V = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2]$, voir figure ci-dessous.

Soit S_1 la face "avant" du pavé. Alors, S_1 admet la paramétrisation suivante qui est lisse, et dont le vecteur normal pointe bien vers l'extérieur du pavé :

$$\begin{aligned} \Sigma : D = [b_1, b_2] \times [c_1, c_2] &\rightarrow \mathbb{R}^3 \\ (y, z) &\mapsto \Sigma(y, z) = (a_2, y, z). \end{aligned}$$



Le vecteur normal est : $N(y, z) = (0, 1, 0) \wedge (0, 0, 1) = (1, 0, 0)$, et :

$$\begin{aligned} \int_{S_1} F &= \int_{\Sigma} F = \int_D F(\Sigma(y, z)) \cdot N(y, z) dy dz, \\ &= \int_D F_1(a_2, y, z) dy dz, \\ &= \int_{c_1}^{c_2} \left(\int_{b_1}^{b_2} F_1(a_2, y, z) dy \right) dz, \quad (\text{Théorème de Fubini}). \end{aligned}$$

Soit S_2 la face “arrière” du pavé. Alors, de manière similaire, on trouve :

$$\int_{S_2} F = - \int_{c_1}^{c_2} \left(\int_{b_1}^{b_2} F_1(a_1, y, z) dy \right) dz.$$

Ainsi,

$$\begin{aligned} \int_{S_1} F + \int_{S_2} F &= \int_{c_1}^{c_2} \left(\int_{b_1}^{b_2} [F_1(a_2, y, z) - F_1(a_1, y, z)] dy \right) dz, \\ &= \int_{c_1}^{c_2} \left(\int_{b_1}^{b_2} \left(\int_{a_1}^{a_2} D_1 F_1(x, y, z) dx \right) dy \right) dz, \\ &= \int_V D_1 F_1. \end{aligned}$$

En répétant cet argument pour les faces latérales S_3, S_4 du pavé, on trouve :

$$\int_{S_3} F + \int_{S_4} F = \int_V D_2 F_2,$$

et pour les faces supérieures, et inférieures S_5, S_6 :

$$\int_{S_5} F + \int_{S_6} F = \int_V D_3 F_3.$$

En additionnant le tout, on obtient :

$$\int_{\partial V} F = \int_V \operatorname{div}(F).$$

□

Exemple 7.12 Calculer l'intégrale du champ de vecteurs $F(x, y, z) = (x, y, z)$ sur le bord du cube centré à l'origine de côté de longueur 2.

Soit V le cube avec son intérieur, alors V est un volume admissible pour le Théorème de la divergence, et le champ de vecteurs F est de classe \mathcal{C}^1 sur tout \mathbb{R}^3 . Ainsi, par le Théorème de la divergence, on a

$$\int_{\partial V} F = \int_V \operatorname{div}(F).$$

Calculons,

$$\operatorname{div}(F) = 1 + 1 + 1 = 3.$$

Ainsi, en utilisant le Théorème de Fubini, on a :

$$\int_{\partial V} F = \int_V 3 = \int_{-1}^1 \left(\int_{-1}^1 \left(\int_{-1}^1 3 \, dx \right) dy \right) dz = 24.$$

Corollaire 62 Soient S_1, S_2 deux surfaces orientables dans \mathbb{R}^3 , telles que S_2 est contenue dans l'intérieur de S_1 , et soit V la région entre les deux (voir Figure 7.1). Soit $F : U \rightarrow \mathbb{R}^3$ un champ de vecteurs de classe \mathcal{C}^1 défini sur un ouvert U contenant V , qui satisfait $\operatorname{div}(F) \equiv 0$, sur U . Alors :

$$\int_{S_1} F = - \int_{S_2} F.$$

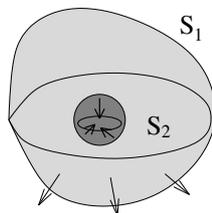


FIG. 7.1 – Exemple de volume admissible pour le corollaire.

7.4.4 Théorème de Stokes

Théorème 63 (Stokes) Soit S une surface admissible pour le Théorème de Stokes. Soit $F : U \rightarrow \mathbb{R}^3$ un champ de vecteurs de classe \mathcal{C}^1 , défini sur un ouvert U contenant S . Alors :

$$\int_S \text{rot}(F) = \int_{\partial S} F.$$

Remarque 7.7

1. Souvenez-vous que le rotationnel d'un champ de vecteurs est un champ de vecteurs, ainsi le terme de gauche du Théorème de Stokes est une intégrale de surface d'un champ de vecteurs, comme définie à la fin de la Section 7.3. Le terme de droite est une intégrale curviligne d'un champ de vecteurs, comme définie à la fin de la Section 6.2.
2. Comme corollaire du Théorème de Stokes, on obtient à nouveau le Théorème de Green.

Soit $D \subset \mathbb{R}^2$ un domaine admissible pour le Théorème de Green, alors $S = D \times \{0\}$ est une surface admissible pour le Théorème de Stokes. Soit $F : U \rightarrow \mathbb{R}^2$ un champ de vecteurs de classe \mathcal{C}^1 , où U est un ouvert contenant D . Alors $\tilde{F} : \tilde{U} = U \times \mathbb{R} \rightarrow \mathbb{R}^3$, défini par :

$$\tilde{F}(x, y, z) = (F(x, y), 0) = (F_1(x, y), F_2(x, y), 0).$$

est un champ de vecteurs de classe \mathcal{C}^1 sur \tilde{U} , un ouvert qui contient S . Ainsi, par le Théorème de Stokes :

$$\int_{\partial S} \tilde{F} = \int_S \text{rot}(\tilde{F}).$$

Calculons le membre de droite.

$$\text{rot}(\tilde{F}) = (0, 0, D_1 F_2(x, y) - D_2 F_1(x, y)).$$

Une paramétrisation lisse Σ de la surface S est donnée par :

$$\begin{aligned} \Sigma : \quad D &\rightarrow S \\ (x, y) &\mapsto (x, y, 0). \end{aligned}$$

Le vecteur normal $N(x, y, z) = (0, 0, 1)$ pointe bien vers l'extérieur de S . Ainsi, par définition de l'intégrale de surface des champs de vecteurs, on a :

$$\int_S \text{rot}(\tilde{F}) = \int_D D_1 F_2(x, y) - D_2 F_1(x, y).$$

De plus, il est facile de se convaincre que $\int_{\partial S} \tilde{F} = \int_{\partial D} F$.

3. Une manière de prouver le Théorème de Stokes est de se ramener au contexte du Théorème de Green au moyen de paramétrisations, et d'utiliser ce dernier pour prouver le Théorème de Stokes. Étant donné que c'est là l'idée principale, et que les notations sont lourdes, nous décidons d'omettre cette preuve.

Exemple 7.13 Vérifions le Théorème de Stokes sur un exemple. On considère le champ de vecteurs $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, défini par :

$$F(x, y, z) = (z - y, x + z, -(x + y)),$$

et le parabolôide $S = \{(x, y, z) \in \mathbb{R}^3 \mid z = 4 - x^2 - y^2, 0 \leq z \leq 4\}$.

Remarquons d'abord que S est une surface admissible pour le Théorème de Stokes, et que F est un champ de vecteurs de classe \mathcal{C}^1 sur tout \mathbb{R}^3 , de sorte que le Théorème de Stokes s'applique bien.

Calculons d'abord le membre de droite de la formule de Stokes. Le bord ∂S de S est le cercle C de rayon 2 dans le plan (xOy) . Ainsi, une paramétrisation régulière de C , telle que la surface S soit à gauche, est :

$$\begin{aligned} \gamma : [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ \theta &\mapsto \gamma(\theta) = (2 \cos \theta, 2 \sin \theta, 0). \end{aligned}$$

Par définition de l'intégrale curviligne d'un champ de vecteurs, on a :

$$\begin{aligned} \int_{\partial S} F &= \int_0^{2\pi} F(\gamma(\theta)) \cdot \gamma'(\theta) d\theta, \\ &= \int_0^{2\pi} (4 \sin^2 \theta + 4 \cos^2 \theta) d\theta, \\ &= 8\pi. \end{aligned}$$

Calculons maintenant le membre de gauche de la formule de Stokes. Une paramétrisation lisse du parabolôide, telle que le vecteur normal pointe vers l'extérieur, est donnée par :

$$\begin{aligned} \Sigma : [0, 2] \times [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ (r, \theta) &\mapsto \Sigma(r, \theta) = (r \cos \theta, r \sin \theta, 4 - r^2). \end{aligned}$$

Le vecteur normal $N(r, \theta)$ de cette paramétrisation est :

$$N(r, \theta) = \begin{pmatrix} \cos \theta \\ \sin \theta \\ -2r \end{pmatrix} \wedge \begin{pmatrix} -r \sin \theta \\ r \cos \theta \\ 0 \end{pmatrix} = \begin{pmatrix} 2r^2 \cos \theta \\ 2r^2 \sin \theta \\ r \end{pmatrix},$$

qui pointe bien vers l'extérieur de S . Calculons maintenant $\text{rot}(F)$:

$$\text{rot}(F) = (-1 - 1, 1 + 1, 1 + 1) = (-2, 2, 2).$$

Ainsi, par définition de l'intégrale de surface d'un champ de vecteurs, on a :

$$\begin{aligned} \int_S \operatorname{rot}(F) &= \int_{[0,2] \times [0,2\pi]} \operatorname{rot}(F)(\Sigma(r, \theta)) \cdot N(r, \theta) \, dr \, d\theta, \\ &= \int_0^2 \left(\int_0^{2\pi} (4r^2(-\cos \theta + \sin \theta) + 2r) \, d\theta \right) \, dr, \quad (\text{Théorème de Fubini}) \\ &= 4\pi \int_0^2 r \, dr, \\ &= 8\pi. \end{aligned}$$

Corollaire 64 Soient S_1 et S_2 deux surfaces admissibles pour le Théorème de Stokes, dont le bord est un même chemin fermé, qui se situent de part et d'autre du bord (voir Figure 7.2). Soit $F : U \rightarrow \mathbb{R}^3$ un champ de vecteurs de classe \mathcal{C}^1 défini sur un ouvert U contenant $S_1 \cup S_2$, alors :

$$\int_{S_1} \operatorname{rot}(F) = - \int_{S_2} \operatorname{rot}(F).$$

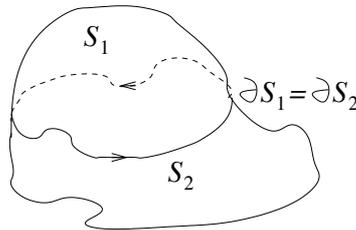


FIG. 7.2 – Exemple de surface admissible pour le corollaire.

Bibliographie

- [1] R. C. Buck. Advanced calculus. Third edition. *McGraw-Hill Book Company*. (1978).
- [2] P. Buser. Cours de Géométrie. *EPFL*. (1995).
- [3] C.H. Edwards, Jr. Advanced calculus of several variables. *Academic Press Inc. New York and London..* (1973).
- [4] S. Lang, Calculus of several variables. Third edition. *Undergraduate Texts in Mathematics, Springer*. (1987).