

Fonctions de plusieurs variables

Notes du Cours LM216

Albert Cohen

Dans ce cours, on s'intéresse aux fonctions de plusieurs variables

$$(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n).$$

Celles-ci interviennent naturellement pour décrire la dépendance de grandeurs en fonction de plusieurs paramètres : par exemple la distribution de la température ou de la pression atmosphérique en fonction de la position dans l'espace et du temps. Elles peuvent aussi être définies par des expressions similaires à celles que l'on rencontre pour les fonctions d'une variable : par exemple $f(x, y, z) = \sin(xy^2 + 1) + \sqrt{z^4 + 3x^2}$ est une fonction des trois variables x , y et z .

L'objectif principal de ce cours est de généraliser aux fonctions de plusieurs variables les notions de dérivation et d'intégration qui ont été abordées dans le cadre des fonctions d'une variable.

Ces notes contiennent la totalité des résultats du cours sous une forme très condensée. En particulier, la plupart des démonstrations détaillées en amphi sont omises, et c'est un excellent exercice pour l'étudiant que d'essayer de les refaire uniquement à partir des indications données dans les notes. La présentation est orientée sur les fonctions de n variables, où n est un nombre arbitraire, mais l'étudiant est encouragé à traduire les résultats dans le cas de fonctions de 2 ou 3 variables afin de s'en forger une intuition visuelle.

Avertissement : ces notes sont régulièrement mises à jour et corrigées. Toutes les remarques permettant d'améliorer la rédaction peuvent être envoyées à l'adresse cohen@ann.jussieu.fr.

1 Notions de topologie dans \mathbb{R}^n

1.1 Vecteurs et normes dans \mathbb{R}^n

Rappelons qu'en dimension 2, on identifie un vecteur x de coordonnées (x_1, x_2) avec un point du plan de coordonnées (x_1, x_2) une fois fixée une origine. On généralisera ici cette identification en désignant le point ou le vecteur de coordonnées (x_1, \dots, x_n) par $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

La *norme euclidienne* d'un vecteur $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ est définie par

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

Le *produit scalaire* de x avec $y = (y_1, \dots, y_n)$ est défini par

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

on a en particulier $\|x\|^2 = x \cdot x$.

Inégalité de Cauchy-Schwarz : pour tout x et y dans \mathbb{R}^n

$$|x \cdot y| \leq \|x\| \|y\|,$$

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

Inégalité triangulaire : pour tout x et y dans \mathbb{R}^n

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

On rappelle aussi que si θ désigne l'angle entre deux vecteurs x et y non-nuls, on a

$$\cos(\theta) = \frac{x \cdot y}{\|x\| \|y\|}.$$

D'un point de vue géométrique, la norme euclidienne $\|x\|$ correspond à la longueur du vecteur x (ou encore à la distance du point x à l'origine). On définit ainsi la distance euclidienne entre deux points x et y par

$$d(x, y) = \|x - y\|,$$

c'est à dire la longueur du vecteur reliant le point x au point y .

La norme euclidienne n'est pas l'unique façon de mesurer la taille d'un vecteur x .

Définition 1.1 On appelle *norme sur \mathbb{R}^n* une application N de \mathbb{R}^n dans \mathbb{R}_+ vérifiant les trois propriétés suivantes

1. $N(x) = 0$ si et seulement si $x = 0$.
2. $N(\lambda x) = |\lambda| N(x)$ pour tout $x \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$.
3. $N(x + y) \leq N(x) + N(y)$ pour tout $x, y \in \mathbb{R}^n$.

On vérifie facilement que la norme euclidienne vérifie ces trois propriétés. Voici d'autres exemples de normes : la norme "sup" ou ℓ^∞ définie par

$$\|x\|_\infty := \sup_{i=1, \dots, n} |x_i|,$$

la norme ℓ^p définie pour $1 \leq p < +\infty$ par

$$\|x\|_p := (|x_1|^p + \dots + |x_n|^p)^{1/p}.$$

On voit que le cas $p = 2$ correspond à la norme euclidienne.

Etant donné une norme N , un point x et un nombre $r \geq 0$ on appelle *boule* de centre x et de rayon r pour la norme N l'ensemble des points à distance au plus r de x , c'est à dire

$$B(x, r) = \{y ; N(y - x) \leq r\}.$$

On appelle boule unité la boule $B(0, 1)$ de centre 0 et de rayon 1. En l'absence de précision sur N , il s'agira systématiquement de la norme euclidienne.

Définition 1.2 Deux normes N_1 et N_2 sont dites "équivalentes" si et seulement si il existe deux constantes $c, C > 0$ telles que pour tout $x \in \mathbb{R}^n$

$$cN_1(x) \leq N_2(x) \leq CN_1(x).$$

Cette propriété signifie intuitivement qu'un vecteur très petit dans une norme le sera aussi dans l'autre. On vérifie par exemple aisément (en utilisant l'inégalité de Cauchy-Schwarz) que

$$\|x\|_2 \leq \|x\|_1 \leq n^{1/2} \|x\|_2.$$

On a aussi

$$n^{-1/2} \|x\|_2 \leq \|x\|_\infty \leq \|x\|_2.$$

On a en fait la propriété fondamentale suivante.

Théorème 1.1 Toutes les normes sur \mathbb{R}^n sont équivalentes entre elles.

1.2 Quelques notions de topologie

Soit $(p_k)_{k>0}$ une suite de points (ou de vecteurs) de \mathbb{R}^n . On dit que cette suite converge vers une limite $p \in \mathbb{R}^n$ si et seulement si pour tout $\varepsilon > 0$, il existe k_0 tel que

$$k \geq k_0 \Rightarrow \|p_k - p\| \leq \varepsilon.$$

On note alors $\lim_{k \rightarrow +\infty} p_k = p$ ou encore parfois $p_k \rightarrow p$. Dans le cas contraire on dit que la suite diverge. Dans le cas $n = 1$, on a $\|p_k - p\| = |p_k - p|$ et on retrouve la définition usuelle de la convergence des suites réelles. On vérifie aisément que la limite d'une suite si elle existe est unique.

Grâce à l'équivalence des normes, il est possible de remplacer la norme euclidienne par n'importe quelle autre norme N sans altérer la définition de la convergence. *Il en est de même pour toutes les notions que nous allons introduire dans cette section.*

En particulier en utilisant la norme sup, on voit que si $p_k = (x_{k,1}, \dots, x_{k,n})$ et $p = (x_1, \dots, x_n)$, la convergence de la suite $(p_k)_{k \geq 0}$ vers p est équivalente à la convergence pour tout $i = 1, \dots, n$ de la suite des i -ème coordonnées $(x_{k,i})_{k \geq 0}$ vers la coordonnée x_i .

Une suite $(p_k)_{k>0}$ est dite "de Cauchy" si et seulement si pour tout $\varepsilon > 0$, il existe k_0 tel que

$$k, l \geq k_0 \Rightarrow \|p_k - p_l\| \leq \varepsilon.$$

On vérifie aisément que toute suite convergente est de Cauchy. La réciproque est aussi vraie : l'espace \mathbb{R}^n est *complet* ce qui signifie que toute suite de Cauchy est convergente. Cette propriété se déduit aisément de la propriété analogue connue pour \mathbb{R} en utilisant la norme sup. L'intérêt pratique de cette propriété est qu'il n'est ainsi pas nécessaire de connaître la limite d'une suite pour pouvoir prouver sa convergence.

Définition 1.3 Un ensemble $U \subset \mathbb{R}^n$ est dit "ouvert" si et seulement si pour tout $x \in U$ il existe $r > 0$ tel que $B(x, r) \subset U$.

Par convention l'ensemble vide \emptyset est ouvert.

Pour un ensemble E quelconque on dit que x est un point intérieur à E si il existe $r > 0$ tel que $B(x, r) \subset E$. L'ensemble des points intérieur à E est appelé l'intérieur de E et est noté $\overset{\circ}{E}$. On vérifie que $\overset{\circ}{E}$ est le plus grand ouvert contenu dans E et que E est ouvert si et seulement si il est égal à son intérieur.

Définition 1.4 Un ensemble $F \subset \mathbb{R}^n$ est dit “fermé” si et seulement si son complémentaire $F^c = \mathbb{R}^n \setminus F$ est ouvert.

Théorème 1.2 F est fermé si et seulement toute suite de point de F qui converge a sa limite contenue dans F .

Pour un ensemble E quelconque on dit que x est un point adhérent à E si il existe une suite $(x_k)_{k \geq 0}$ de points de E qui converge vers x . L’ensemble des points adhérents à E est noté \overline{E} et est appelé l’adhérence (ou la fermeture) de E . On vérifie que \overline{E} est le plus petit fermé contenant E et que E est fermé si et seulement si $E = \overline{E}$.

Quelques propriétés élémentaires :

1. Toute union finie ou infinie d’ouverts est un ouvert.
2. Toute union finie de fermés est un fermé
3. Toute intersection finie ou infinie de fermés est un fermé.
4. Toute intersection finie d’ouverts est un ouvert
5. Les seuls ensembles à la fois ouverts et fermés sont \emptyset et \mathbb{R}^n .
6. Un ensemble fini de points de \mathbb{R}^n est fermé.
7. On a $\overset{\circ}{E} \subset E \subset \overline{E}$.

Définition 1.5 On appelle frontière d’un ensemble E l’ensemble des points de son adhérence qui ne sont pas dans son intérieur, c’est à dire

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

Définition 1.6 Un ensemble $E \subset \mathbb{R}^n$ est dit borné si et seulement si il existe $R > 0$ tel que $E \subset B(0, R)$. Un ensemble fermé et borné de \mathbb{R}^n est dit “compact”.

Théorème 1.3 Pour $E \subset \mathbb{R}^n$, les trois propriétés suivantes sont équivalentes

1. E est compact
2. Propriété de Bolzano-Weierstrass : toute suite $(x_k)_{k \geq 0}$ de points de E admet une sous-suite qui converge vers une limite $x \in E$.
3. Propriété de Borel-Lebesgue : de toute famille d’ouverts $(U_i)_{i \in I}$ telle que $E \subset \cup_{i \in I} U_i$, on peut extraire une sous famille finie (U_1, \dots, U_m) telle que $E \subset U_1 \cup \dots \cup U_m$.

Définition 1.7 Un ensemble E est dit “connexe” si et seulement si il n’existe aucune paire d’ouverts (U_1, U_2) tels que $E \subset U_1 \cup U_2$ et $E \cap U_1$ et $E \cap U_2$ sont disjoints.

Autrement dit, E est connexe si on ne peut pas le séparer en deux partie disjointes en l’intersectant avec deux ouverts. Dans le cas d’un ensemble ouvert, cela se traduit par la propriété intuitive suivante.

Théorème 1.4 Un ensemble ouvert U est connexe si et seulement si pour tout x et y dans U , il existe une ligne brisée contenue dans U qui les relie, c’est à dire un ensemble fini de points $\{p_1, \dots, p_m\}$ tel que $p_1 = x$ et $p_m = y$, et tel que les segments d’extrémités p_i et p_{i+1} sont tous contenus dans U .

Définition 1.8 Un ensemble E est dit “convexe” si et seulement si pour tout $x, y \in E$ et $t \in [0, 1]$ on a $tx + (1 - t)y \in E$, i.e. le segment d’extrémité x et y est contenu dans E .

On peut vérifier qu’en dimension $n = 1$ les connexes et les convexes de \mathbb{R} sont exactement les intervalles (de taille finie ou infinie). En revanche, en dimension $n > 1$ tout convexe est connexe mais la réciproque est fausse.

On vérifie aussi qu’une intersection finie ou infinie de compacts est un compact et de même pour les convexes.

2 Fonctions continues

2.1 Définition et propriétés fondamentales

On va s'intéresser à des fonctions de plusieurs variables et à valeurs réelles $f : \mathbb{R}^n \rightarrow \mathbb{R}$: pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $f(x) = f(x_1, \dots, x_n) \in \mathbb{R}$.

On étend à ce cadre la notion de *domaine de définition* connue pour les fonctions d'une variable. Par exemple le domaine de définition de la fonction $f(x, y) = \log(\sqrt{3 - x^2 - y^2})$ est l'ensemble $D = \{(x, y) ; x^2 + y^2 < 3\}$ c'est à dire l'intérieur de la boule de centre 0 et de rayon $\sqrt{3}$.

On s'intéressera aussi parfois à la notion plus générale de fonctions de n variables et à valeurs vectorielles $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$: pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ $f(x) = (f_1(x), \dots, f_m(x)) \in \mathbb{R}^m$. Notons que chacune des fonctions f_i est une fonction de n variables et à valeur réelle. On dit parfois que f est un *champ de vecteurs* à m composantes définis sur \mathbb{R}^n .

Si $E \subset \mathbb{R}^n$ on appelle *image* de E par f l'ensemble $f(E)$ des images par f des points de E . On a donc

$$f(E) = \{f(x) ; x \in E\}.$$

On dit que la fonction f est bornée sur E si $f(E)$ est un ensemble borné, autrement dit, il existe $R > 0$ tel que $\|f(x)\| \leq R$ pour tout $x \in E$.

Si $F \subset \mathbb{R}^m$, on appelle *image réciproque* de F par f l'ensemble $f^{-1}(F)$ des antécédents par f des points de F . On a donc

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

Une fonction d'une variable et à valeur réelle est décrite par son graphe, qui est le sous-ensemble de \mathbb{R}^2 défini par

$$G_f = \{(x, f(x)) ; x \in \mathbb{R}\},$$

et que l'on représente par la courbe du plan d'équation $y = f(x)$. Dans le cas d'une fonction de deux variables à valeur réelle on définit de même le graphe

$$G_f = \{(x, y, f(x, y)) ; (x, y) \in \mathbb{R}^2\},$$

que l'on peut représenter comme une surface d'équation $z = f(x, y)$ dans l'espace à 3 dimensions. Pour $\lambda \in \mathbb{R}$, on appelle courbe de niveau λ de la fonction f l'ensemble C_λ des points du plan dont l'image par f vaut λ c'est à dire

$$C_\lambda := \{(x, y) ; f(x, y) = \lambda\} = f^{-1}(\{\lambda\}).$$

Les notions de graphe et de courbes de niveau s'étendent de manière évidente au cas des fonctions de n variables.

Définition 2.1 Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est continue au point $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ si et seulement si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tout $y = (y_1, \dots, y_n) \in \mathbb{R}^n$,

$$\|y - x\| \leq \alpha \Rightarrow |f(y) - f(x)| \leq \varepsilon.$$

Une fonction f est continue sur un ensemble E si et seulement si elle est continue en tout point de cet ensemble. L'ensemble des fonctions continues sur E est noté $\mathcal{C}(E)$ ou $C^0(E)$.

Dans le cas de fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ à valeur vectorielle, il faut remplacer $|f(y) - f(x)|$ par $\|f(y) - f(x)\|$ dans la définition ci-dessus. Grâce à l'équivalence des normes, il est possible de remplacer la norme euclidienne par n'importe quelle autre norme N sans altérer la définition de la continuité. On remarque que si f est continue au point x alors elle est toujours bornée sur une boule $B(x, r)$ pour r suffisamment petit.

Une manière équivalente d'exprimer la continuité au point x est de dire que pour tout suite (x_k) telle que $x_k \rightarrow x$, on a $f(x_k) \rightarrow f(x)$. On peut aussi démontrer que f est continue sur \mathbb{R}^n si et seulement si l'image réciproque de tout ensemble ouvert est un ensemble ouvert.

On démontre enfin que l'image d'un connexe par une fonction continue est aussi connexe, et que l'image d'un compact par une fonction continue est aussi compact. Cette dernière propriété entraîne qu'une fonction continue est toujours bornée sur un compact.

Définition 2.2 Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est lipschitzienne si et seulement si pour tout $x, y \in \mathbb{R}^n$ on a

$$\|f(x) - f(y)\| \leq C\|x - y\|,$$

où C est une constante indépendante de x et y .

On dit parfois que f est C -lipschitzienne. On voit immédiatement que toute fonction lipschitzienne est continue. En revanche la réciproque est fautive. Un exemple de fonction 1-lipschitzienne est l'application $x \mapsto \|x\|$ (cela se déduit aisément de l'inégalité triangulaire).

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est une application linéaire, alors en introduisant la base canonique (e_1, \dots, e_n) de \mathbb{R}^n et en notant $M = \max \|f(e_i)\|$, on voit que pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ on a

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

avec $B = M\sqrt{n}$. En remplaçant x par $x - y$ on voit ainsi que toute application linéaire est lipschitzienne et donc continue. On appelle *norme* de l'application linéaire f la quantité

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

On vérifie que $\|f\|$ est la plus petite constante C telle que $\|f(x)\| \leq C\|x\|$ pour tout x .

2.2 Opérations sur les fonctions continues

Voici trois propriétés dont les démonstrations sont élémentaires :

1. Une combinaison linéaire ou une multiplication de deux (ou d'un nombre fini de) fonctions continues en un point x est continue en x .
2. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continue au point x et $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ est continue au point $f(x)$, alors $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est continue au point x .
3. La fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ qui à $x = (x_1, \dots, x_n)$ associe la i -ème coordonnée $f(x) = x_i$ est continue. On l'appelle la i -ème application coordonnée.

Ces propriétés permettent de construire de nombreuses fonctions continues en utilisant les résultats connus sur la continuité des fonctions d'une seule variable. Par exemple, la fonction $f(x, y) = \log(\sqrt{3 - x^2 - y^2})$ est continue sur son domaine de définition, par composition des applications coordonnées, carré, racine, logarithme. De même, on voit que si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est continue au point x et $f(x) \neq 0$, alors $1/f$ est aussi continue au point x .

Une classe importante de fonctions continues est celle des *polynômes*. On rappelle qu'un polynôme d'une variable est une fonction de la forme

$$p(x) = a_0 + a_1x + \dots + a_mx^m,$$

c'est à dire une combinaison linéaire d'un nombre fini de fonctions puissance x^k pour k entier positif. L'entier m est appelé "degré" du polynôme p . Le terme a_kx^k est appelé monôme de degré k . Dans le cas de fonctions à n variable, un polynôme est une combinaison linéaire d'un nombre fini de fonctions du type $x_1^{k_1}x_2^{k_2}\dots x_n^{k_n}$ où k_1, \dots, k_n sont des entiers positifs. Le degré de $x_1^{k_1}\dots x_n^{k_n}$ est par définition $k_1 + \dots + k_n$. Le degré de p est celui du monôme de plus haut degré présent dans p . En notant $k = (k_1, \dots, k_n) \in \mathbb{N}^n$, on voit ainsi qu'un polynôme de degré m est de la forme

$$p(x) = p(x_1, \dots, x_n) = \sum_{k_1 + \dots + k_n \leq m} a_k x_1^{k_1} \dots x_n^{k_n}.$$

Une manière naturelle d'étudier les variations d'une fonction f de n variables consiste à fixer $n - 1$ d'entre elles et à étudier alors la fonction par rapport à l'unique variable restante. Ainsi si $x = (x_1, \dots, x_n)$ est un point fixé, on définit la i -ème application partielle au point x par

$$f_{i,x}(t) = f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n).$$

Il est facile de vérifier que la continuité de f en x entraîne celle de l'application partielle $f_{i,x}$ au point $t = x_i$. En revanche, la réciproque est fautive : par exemple la fonction définie par $f(x, y) = \frac{xy}{x^4+y^4}$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$ admet des applications partielles continues en tout point, mais cependant cette fonction n'est pas continue au point $(0, 0)$.

2.3 Fonctions continues et ensembles compacts

Une fonction est dite uniformément continue sur un ensemble E si et seulement si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tout $x, y \in E$

$$\|x - y\| \leq \alpha \Rightarrow |f(x) - f(y)| \leq \varepsilon,$$

($\|f(x) - f(y)\|$ dans le cas des fonctions à valeurs vectorielles). Le point important est que α dépend seulement de ε , pas de x et y . Ainsi une fonction uniformément continue est continue mais la réciproque est fautive en générale. On a cependant le résultat suivant que l'on peut démontrer à l'aide de la propriété de Borel-Lebesgue.

Théorème 2.1 *Si E est un ensemble compact, toute fonction continue sur E est aussi uniformément continue.*

Une fonction f à valeur réelle est dite minorée si et seulement si il existe M tel que $f(x) \geq M$ pour tout $x \in E$. Dans ce cas, l'ensemble $f(E)$ admet un infimum qui est aussi le plus grand M ayant la propriété précédente. Cette valeur est notée $\inf_{x \in E} f(x)$ et est caractérisée par les propriétés

1. Pour tout $x \in E$ on a $f(x) \geq \inf_{x \in E} f(x)$
2. Il existe une suite (x_k) de points de E telle que $f(x_k) \rightarrow \inf_{x \in E} f(x)$.

Dans le cas où f n'est pas minorée, on pose $\inf_{x \in E} f(x) = -\infty$. Dans le cas où l'infimum est fini, il peut-être atteint, c'est à dire qu'il existe $x_0 \in E$ tel que $f(x_0) = \inf_{x \in E} f(x)$, ou ne pas être atteint. Si il est atteint on parle de minimum de f sur E et on note alors $f(x_0) = \min_{x \in E} f(x)$. On définit de manière similaire le suprémum $\sup_{x \in E} f(x)$ sur E (fini ou égal à $+\infty$) que l'on appelle maximum $\max_{x \in E} f(x)$ dans le cas ou il est atteint. Le théorème suivant peut se démontrer à l'aide de la propriété de Bolzano-Weierstrass.

Théorème 2.2 *Si E est un ensemble compact, toute fonction continue sur E admet un minimum et un maximum.*

On peut utiliser ce résultat afin de prouver l'équivalence de toutes les normes définies sur \mathbb{R}^n , en montrant leur équivalence avec la norme sup.

3 Différentiabilité et dérivées partielles

3.1 Définitions et propriétés élémentaires

Définition 3.1 Soit f une fonction à valeur réelle définie sur un ouvert $U \subset \mathbb{R}^n$. On dit que f admet une dérivée partielle au point $x = (x_1, \dots, x_n) \in U$ par rapport à la i -ème coordonnée si la i -ème application partielle $f_{i,x}(t)$ associée au point x est dérivable au point $t = x_i$. On note

$$\frac{\partial f}{\partial x_i}(x) = f'_{i,x}(x_i) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x)}{h}.$$

Le calcul de $\frac{\partial f}{\partial x_i}$ consiste donc à ne dériver l'expression de f que par rapport à la variable x_i . Notons que les dérivées partielles $\frac{\partial f}{\partial x_i}$ sont aussi des fonctions de n variables à valeurs dans \mathbb{R} . Par exemple si $f(x, y, z) = -2x \cos(y)$, on a $\frac{\partial f}{\partial x}(x, y, z) = -2 \cos(y)$, $\frac{\partial f}{\partial y}(x, y, z) = 2x \sin(y)$ et $\frac{\partial f}{\partial z}(x, y, z) = 0$.

Définition 3.2 Soit f une fonction à valeur réelle définie sur un ouvert $U \subset \mathbb{R}^n$ et soit $v \in \mathbb{R}^n$ un vecteur fixé. La dérivée de f en $x \in U$ suivant le vecteur v est la valeur de la dérivée de la fonction d'une variable $f_v(t) = f(x + tv)$ en $t = 0$ si elle existe, c'est à dire $\lim_{t \rightarrow 0} \frac{f(x+tv) - f(x)}{t}$.

Intuitivement cette dernière quantité mesure les variations de f lorsqu'on se déplace autour de x dans la direction du vecteur v à la "vitesse" $\|v\|$. On voit en particulier que $\frac{\partial f}{\partial x_i}(x)$ correspond à la dérivée en x suivant le vecteur de base e_i (avec $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, ..., $e_n = (0, \dots, 0, 1)$).

Définition 3.3 Soit f une fonction définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R}^m . On dit que f est différentiable en un point $x \in U$ si et seulement si il existe une application linéaire de \mathbb{R}^n dans \mathbb{R}^m notée df_x telle que pour tout $h \in \mathbb{R}^n$ suffisamment petit

$$f(x + h) = f(x) + df_x(h) + \|h\|\varepsilon(h),$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. L'application df_x est appelée différentielle de f au point x ou encore application linéaire tangente de f au point x .

La formule ci-dessus n'est rien d'autre qu'un développement limité de f à l'ordre 1 au voisinage de x . Dans le cas d'une fonction d'une variable à valeur réelle, cette définition redonne la notion classique de dérivabilité, et on a alors $df_x(h) = f'(x)h$. Clairement, si f est différentiable au point x , elle est continue en ce point.

Dans le cas d'une fonction de n variables et à valeur dans \mathbb{R} , df_x est une forme linéaire de \mathbb{R}^n dans \mathbb{R} . Pour $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ on a donc $df_x(h) = a_1 h_1 + \dots + a_n h_n$. On vérifie aisément que $a_i = \frac{\partial f}{\partial x_i}(x)$. Le vecteur

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

est appelé le *gradient* de f au point x . On a donc

$$f(x + h) = f(x) + \nabla f(x) \cdot h + \|h\|\varepsilon(h) = f(x) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i + \|h\|\varepsilon(h).$$

On déduit en particulier que la dérivée de f en x suivant le vecteur v est donnée par le produit scalaire $\nabla f(x) \cdot v$.

Ainsi une fonction différentiable en un point x admet des dérivées partielles et plus généralement des dérivées suivant tout vecteur en ce point. La réciproque est fautive : il existe des fonctions qui admettent des dérivées suivant tout vecteur en un point mais ne sont pas même continues en ce point. Un exemple est le comportement au point $(0, 0)$ de la fonction définie par $f(x, y) = y^2/x$ si $x \neq 0$ et $f(0, y) = 0$.

Dans le cas d'une fonction f de n variables à valeur dans \mathbb{R}^m , on vérifie que df_x est la matrice $m \times n$ dont l'élément de ligne i et de colonne j est donné par $\frac{\partial f_i}{\partial x_j}(x)$. Cette matrice est appelée *matrice jacobienne* de f au point x . On peut écrire pour chaque composante de f

$$f_i(x + h) = f_i(x) + \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x) h_j + \|h\|\varepsilon_i(h).$$

3.2 Opérations sur les dérivées, fonctions \mathcal{C}^1

Les propriétés suivantes se prouvent de façon élémentaire :

1. Si f est constante, alors df_x est l'application nulle en tout point x . Si f est linéaire alors $df_x = f$ en tout point.
2. Une combinaison linéaire de deux (ou d'un nombre fini de) fonctions différentiables en un point est différentiable, et on a $d(\alpha f + \beta g)_x = \alpha df_x + \beta dg_x$, ce qui s'écrit aussi $\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g$ ou encore $\frac{\partial(\alpha f + \beta g)}{\partial x_i} = \alpha \frac{\partial f}{\partial x_i} + \beta \frac{\partial g}{\partial x_i}$.
3. Si f et g sont deux fonctions à valeur réelle différentiables au point x , alors fg l'est aussi et on a la formule de Leibniz $\nabla(fg) = f\nabla g + g\nabla f$.
4. Si f est différentiable au point x et g est différentiable au point $f(x)$ alors $g \circ f$ est différentiable en x avec $d(g \circ f)_x = dg_{f(x)} \circ df_x$.

La dernière propriété signifie que la matrice jacobienne de $g \circ f$ en x est obtenu en effectuant le produit de la matrice jacobienne de g en $f(x)$ par la matrice jacobienne de f en x (on rappelle que le produit des matrices n'est pas commutatif). Dans le cas où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$, on obtient $\nabla(g \circ f)(x) = g'(f(x))\nabla f(x)$. Un exemple d'application est le suivant : si f est différentiable et non nulle au point x , $g = 1/f$ l'est aussi et

$$\nabla g(x) = -\frac{\nabla f(x)}{f^2(x)}.$$

Un cas plus délicat est celui où $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}$. Le théorème de composition donne alors $\nabla(g \circ f)(x) = (df_x)^T \nabla g(f(x))$. Ceci se traduit plus intuitivement par la règle

$$\frac{\partial g \circ f}{\partial x_i}(x) = \sum_{j=1}^n \frac{\partial f_j}{\partial x_i}(x) \frac{\partial g}{\partial x_j}(f(x)).$$

Une application de la propriété de composition est la suivante : si f est une fonction de \mathbb{R}^n dans \mathbb{R} qui est homogène de degré α , c'est à dire telle que $f(\lambda x) = \lambda^\alpha f(x)$ pour tout $x \in \mathbb{R}^n$ et $\lambda \geq 0$, et si f est différentiable au point $x = (x_1, \dots, x_n)$, on a alors l'identité d'Euler

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(x) = \alpha f(x).$$

Les résultats qui suivent généralisent des propriétés bien connues pour les fonctions d'une variable.

Théorème 3.1 *Si une fonction f à valeur réelle est différentiable en tout point du segment S d'extrémités x et y , on a*

$$|f(x) - f(y)| \leq C \|x - y\|,$$

avec $C = \sup_{z \in S} \|\nabla f(z)\|$. Par conséquent si f est différentiable sur un ouvert convexe U et si $C = \sup_{z \in U} \|\nabla f(z)\| < +\infty$, alors f est C -lipschitzienne sur U .

Dans le cas des fonctions à valeur vectorielles, on a le même résultat en remplaçant $|f(x) - f(y)|$ par $\|f(x) - f(y)\|$ et $\|\nabla f(z)\|$ par $\|df_z\|$.

Théorème 3.2 *Si U est un ouvert connexe et $\nabla f(x)$ est nul en tout point $x \in U$ alors f est constante sur U .*

Dans le cas des fonctions à valeur vectorielle, on a le même résultat en remplaçant $\nabla f(x)$ par df_x . La nécessité de supposer U connexe dans ce dernier théorème vient du fait que sinon, on peut le séparer en deux parties disjointes $U \cap U_1$ et $U \cap U_2$ et considérer une fonction f qui prend deux valeurs constantes différentes sur chaque partie.

Théorème 3.3 *Si f est une fonction définie sur \mathbb{R}^n ou plus généralement sur un convexe et à valeur dans \mathbb{R} , et si la i -ème dérivée partielle $\frac{\partial f}{\partial x_i}$ est partout nulle, alors $f(x)$ est une fonction des variables x_j pour $j \neq i$, i.e. ne dépend pas de i .*

On a vu qu'une fonction dont les dérivées partielles sont partout définies n'est pas nécessairement différentiable. La situation est différente lorsque ces dérivées partielles sont continues.

Définition 3.4 Une fonction f définie sur un ouvert U de \mathbb{R}^n et à valeur dans \mathbb{R} est dite de classe \mathcal{C}^1 sur U si et seulement ses dérivées partielles $\frac{\partial f}{\partial x_i}$ sont des fonctions continues sur U . On note $\mathcal{C}^1(U)$ l'ensemble des fonctions de classe \mathcal{C}^1 sur U .

Cette définition se généralise aux fonctions vectorielles en écrivant que toutes les composantes f_1, \dots, f_m ont cette propriété.

Théorème 3.4 Une fonction \mathcal{C}^1 sur un ouvert U de \mathbb{R}^n est différentiable en tout point de U et sa différentielle est une fonction continue.

Définition 3.5 Une fonction f de classe \mathcal{C}^1 définie sur un ouvert U de \mathbb{R}^n à valeur dans \mathbb{R}^m est un \mathcal{C}^1 -difféomorphisme si et seulement si il existe un ouvert V de \mathbb{R}^m tel que f est bijective de U dans V , et si la bijection réciproque f^{-1} est aussi de classe \mathcal{C}^1 .

Un \mathcal{C}^1 -difféomorphisme est parfois appelé changement de variable. Le théorème de composition montre que pour tout $x \in U$ on a alors

$$df_x \circ df_{f(x)}^{-1} = I \quad \text{et} \quad df_{f(x)}^{-1} \circ df_x = I,$$

où I est l'application identité. Conséquence : df_x est inversible en tout point x et

$$df_{f(x)}^{-1} = (df_x)^{-1}.$$

Ceci nous montre aussi qu'on a nécessairement $m = n$: un \mathcal{C}^1 -difféomorphisme ne peut exister qu'entre espaces de même dimension.

Définition 3.6 Pour une fonction f de classe \mathcal{C}^1 définie sur un ouvert U de \mathbb{R}^n à valeur dans \mathbb{R}^n , on appelle "jacobien de f au point x ", la valeur absolue du déterminant de la matrice jacobienne df_x , et on note cette quantité $J_f(x)$.

Dans le cas d'un \mathcal{C}^1 -difféomorphisme, $J_f(x)$ ne s'annule pas puisque df_x est inversible en tout point x

3.3 Dérivées partielles d'ordre supérieur, fonctions \mathcal{C}^k

Définition 3.7 Une fonction f définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R} est de classe \mathcal{C}^2 si et seulement si elle est de classe \mathcal{C}^1 et si ses dérivées partielles sont aussi de classe \mathcal{C}^1 . On note

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}$$

c'est à dire la i -ème dérivée partielle de la j -ème dérivée partielle de f . Lorsque $i = j$ on utilise la notation simplifiée $\frac{\partial^2 f}{\partial x_i^2}$.

Un résultat fondamental est la relation de Schwarz qui affirme que l'ordre des variables peut-être échangé dans le calcul des dérivées partielles sans altérer le résultat.

Théorème 3.5 Si f est de classe \mathcal{C}^2 sur un ouvert $U \subset \mathbb{R}^n$, on a en tout point de U

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

Ce résultat nous montre que la matrice des dérivées partielles d'ordre 2 au point x

$$d^2 f_x = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} (x) \right)_{i,j=1,\dots,n}$$

est une matrice symétrique. On l'appelle la *matrice hessienne* de f au point x .

Plus généralement, si les propriétés de f le permettent, on introduit par récurrence les dérivées partielles d'ordre k ,

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}} = \frac{\partial \frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \cdots \partial x_{i_1}}}{\partial x_{i_k}},$$

où i_1, \dots, i_k sont des indices choisis dans $\{1, \dots, n\}$ qui correspondent aux variables par rapports auxquelles on dérive successivement.

Définition 3.8 Une fonction f définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R} est de classe \mathcal{C}^k si et seulement si toutes ses dérivées partielles d'ordre au plus k sont définies et continues sur U . On dit que f est \mathcal{C}^∞ si et seulement si elle est de classe \mathcal{C}^k pour tout $k > 0$.

Cette définition se généralise aux fonctions vectorielles en écrivant que toutes les composantes f_1, \dots, f_m ont cette propriété.

En appliquant itérativement la relation de Schwarz, on voit que l'ordre d'apparition des variables x_i dans la dérivation peut être permuté. On emploie donc la notation

$$\frac{\partial^k f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}},$$

pour désigner la dérivée partielle d'ordre k faisant intervenir k_i fois la variable x_i avec $k_1 + \dots + k_n = k$.

Les propriétés suivantes se démontrent par récurrence sur k :

1. Une combinaison linéaire de deux (ou d'un nombre fini de) fonctions \mathcal{C}^k est aussi de classe \mathcal{C}^k .
2. Si f et g sont deux fonctions à valeur réelle et de classe \mathcal{C}^k , alors fg l'est aussi.
3. Si f est de classe \mathcal{C}^k d'un ouvert $U \subset \mathbb{R}^n$ dans \mathbb{R}^m et si g est de classe \mathcal{C}^k d'un ouvert $V \subset \mathbb{R}^m$ contenant $f(U)$ dans \mathbb{R}^p , alors $g \circ f$ est de classe \mathcal{C}^k de U dans \mathbb{R}^p .
4. Si f est un \mathcal{C}^1 -difféomorphisme et est de classe \mathcal{C}^k , la réciproque f^{-1} est aussi de classe \mathcal{C}^k .

3.4 Points critiques et extremas

Définition 3.9 Soit f une fonction \mathcal{C}^1 définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur réelle. On dit que $x \in U$ est un point critique de f si et seulement si $\nabla f(x) = 0$.

Cette définition s'étend aux fonctions à valeur vectorielle en remplaçant $\nabla f(x)$ par df_x .

Définition 3.10 Soit f une fonction définie sur un ensemble $E \subset \mathbb{R}^n$ et à valeur réelle. On dit que f admet un minimum global au point $x \in E$ si et seulement si $f(x) \leq f(y)$ pour tout $y \in E$. On dit que f admet un minimum local au point $x \in E$ si et seulement si il existe un ouvert U contenant x tel que $f(x) \leq f(y)$ pour tout $y \in U \cap E$.

On définit de la même manière la notion de maximum global ou local. On utilise la dénomination d'*extremum* pour désigner sans distinction un maximum ou minimum. On a déjà signalé l'existence d'un maximum et minimum global sur un compact. Par ailleurs si f admet un extremum local en x , alors pour tout $u \in \mathbb{R}^n$ la fonction d'une variable $g_u(t) = f(x + tu)$ admet un extremum local en $t = 0$ et par conséquent $g'(0) = \nabla f(x) \cdot u = 0$. Comme u est arbitraire, ceci entraîne le résultat suivant.

Théorème 3.6 Soit f une fonction \mathcal{C}^1 définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R} . Si f admet un extremum local en un point $x \in U$ alors x est un point critique de f .

Être un point critique est donc une condition nécessaire pour être un extremum, elle n'est cependant pas suffisante : considérer par exemple la fonction $f(x, y) = x^2 - y^2$ au point $(0, 0)$.

Afin d'identifier plus précisément les extrema parmi les points critiques, on est amené à introduire un développement limité au deuxième ordre qui fait intervenir la matrice hessienne $d^2 f_x$.

Théorème 3.7 *Soit f une fonction \mathcal{C}^2 définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur réelle. Pour $x \in U$ et pour tout $h \in \mathbb{R}^n$ suffisamment petit, on a*

$$f(x+h) = f(x) + \nabla f(x) \cdot h + \frac{1}{2} d^2 f_x h \cdot h + \|h\|^2 \varepsilon(h),$$

où $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

Le terme du deuxième ordre s'écrit plus précisément en fonction des dérivées partielles

$$d^2 f_x h \cdot h = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j.$$

C'est la *forme quadratique* associée à la matrice $d^2 f_x$.

On rappelle qu'une matrice symétrique se diagonalise dans une base orthonormée et que toutes ses valeurs propres sont réelles. En notant (f_1, \dots, f_n) et $(\lambda_1, \dots, \lambda_n)$ la base orthonormée et les valeurs propres pour la matrice $d^2 f_x$, et en décomposant h suivant cette base selon $h = \sum_{i=1}^n a_i f_i$ on obtient l'expression plus simple

$$d^2 f_x h \cdot h = \sum_{i=1}^n \lambda_i a_i^2.$$

On peut à partir de cette remarque obtenir une condition nécessaire pour avoir un extremum au point x qui précise celle de point critique énoncée précédemment.

Théorème 3.8 *Soit f une fonction \mathcal{C}^2 définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur réelle. Si f admet un minimum local au point x alors les λ_i sont tous positifs ou nuls. Si f admet un maximum local en x alors les λ_i sont tous négatifs ou nuls.*

On peut aussi obtenir une condition suffisante pour avoir un extremum au point x .

Théorème 3.9 *Soit f une fonction \mathcal{C}^2 définie sur un ouvert $U \subset \mathbb{R}^n$ et à valeur réelle. Soit $x \in U$ un point critique de f . Si les λ_i sont tous strictement positifs alors f admet un minimum local au point x . Si les λ_i sont tous strictement négatifs alors f admet un maximum local au point x .*

Dans le cas $n = 2$ la matrice hessienne est de taille 2×2 et il est particulièrement facile de vérifier le signe des λ_i en remarquant que le déterminant $\det(d^2 f_x)$ vaut $\lambda_1 \lambda_2$ et que la trace $\text{tr}(d^2 f_x)$ vaut $\lambda_1 + \lambda_2$. On voit donc que si x est un point critique et si $\det(d^2 f_x) > 0$, alors f admet un extremum au point x . On voit aussi que cet extremum est un maximum si $\text{tr}(d^2 f_x) < 0$ et un minimum si $\text{tr}(d^2 f_x) > 0$.

3.5 Opérateurs du calcul différentiel

Il est fréquent d'identifier le symbole ∇ à un "vecteur" de coordonnées $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ que l'on "applique" à f pour obtenir ∇f . Notons bien que cette identification est purement symbolique et n'a rien de rigoureux, mais elle peut être utile d'un point de vue mnémotechnique.

De même si $g = (g_1, \dots, g_n)$ est une fonction vectorielle de classe \mathcal{C}^1 allant de \mathbb{R}^n dans \mathbb{R}^n , c'est à dire un champs de vecteurs sur \mathbb{R}^n , on définit la *divergence* de g comme le "produit scalaire" de ∇ avec g , c'est à dire

$$\text{div}(g) := \nabla \cdot g = \sum_{i=1}^n \frac{\partial g_i}{\partial x_i}.$$

Si f est une fonction de classe \mathcal{C}^2 allant de \mathbb{R}^n dans \mathbb{R} , alors ∇f est un champs de vecteurs de classe \mathcal{C}^1 sur \mathbb{R}^n , et on appelle *laplacien* de f sa divergence, qui est notée

$$\Delta f := \operatorname{div}(\nabla f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}.$$

Enfin rappelons que le produit vectoriel de deux vecteurs $x = (x_1, x_2, x_3)$ et $y = (y_1, y_2, y_3)$ de \mathbb{R}^3 est donné par

$$x \wedge y = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1).$$

Si $g = (g_1, g_2, g_3)$ est un champs de vecteur de classe \mathcal{C}^1 sur \mathbb{R}^3 , on définit ainsi son *rotationnel* comme le champs de vecteur

$$\operatorname{rot}(g) := \nabla \wedge g = \left(\frac{\partial g_3}{\partial x_2} - \frac{\partial g_2}{\partial x_3}, \frac{\partial g_1}{\partial x_3} - \frac{\partial g_3}{\partial x_1}, \frac{\partial g_2}{\partial x_1} - \frac{\partial g_1}{\partial x_2} \right).$$

En dimension 2, le produit vectoriel de $x = (x_1, x_2)$ et $y = (y_1, y_2)$ est le déterminant $x_1 y_2 - x_2 y_1$ qui est la seule composante non-nulle du produit vectoriel des vecteurs $(x_1, x_2, 0)$ et $y = (y_1, y_2, 0)$. On définit donc le rotationnel d'un champs de vecteur $g = (g_1, g_2)$ de classe \mathcal{C}^1 sur \mathbb{R}^2 comme la fonction de \mathbb{R}^2 dans \mathbb{R}

$$\operatorname{rot}(g) := \nabla \wedge g = \frac{\partial g_2}{\partial x_1} - \frac{\partial g_1}{\partial x_2}.$$

Les opérateurs ∇ , div , Δ et rot sont utilisés en permanence dans les équations issues de la physique et de la mécanique. Voici quelques formules qu'il est facile de démontrer :

1. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$, on a $\operatorname{div}(fg) = f \operatorname{div}(g) + \nabla f \cdot g$.
2. Si $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ et $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, on a $\operatorname{rot}(fg) = f \operatorname{rot}(g) + \nabla f \wedge g$.
3. Si $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, on a $\operatorname{rot}(\nabla f) = 0$.
4. Si $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, on a $\operatorname{div}(\operatorname{rot}(g)) = 0$.
5. Si $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, on a $\operatorname{rot}(\operatorname{rot}(f)) = \nabla(\operatorname{div}(f)) - \Delta f$, où $\Delta f := (\Delta f_1, \Delta f_2, \Delta f_3)$.

4 Résultats plus avancés

4.1 Equations aux dérivées partielles

De nombreux phénomènes physiques se modélisent par des équations faisant intervenir les dérivées partielles de la fonction solution de l'équation. De telles équations sont appelées *équations aux dérivées partielles* et souvent notées EDP.

L'exemple le plus simple d'EDP est l'équation de transport : étant donné un réel $c \neq 0$, on cherche $u(x, t)$ de classe \mathcal{C}^1 de \mathbb{R}^2 dans \mathbb{R} qui vérifie

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0.$$

On peut faire le changement de variable $y = x - ct$ et $z = x + ct$ qui est un \mathcal{C}^1 difféomorphisme de \mathbb{R}^2 et on écrit donc $u(x, t) = v(y, z)$. On voit alors que l'équation se re-écrit

$$\frac{\partial v}{\partial z} = 0,$$

Ce qui montre que u est uniquement fonction de $y = x - ct$. On dit aussi que u est constante le long des *droites caractéristiques* d'équation $x = x_0 + ct$, pour tout $x_0 \in \mathbb{R}$.

En imposant une *condition initiale* au temps $t = 0$,

$$u(x, 0) = u_0(x),$$

où u_0 est une fonction fixée, on voit ainsi que la solution de l'équation transport s'écrit

$$u(x, t) = u_0(x - ct),$$

ce qui correspond au "déplacement" de la fonction u_0 à la vitesse c le long de l'axe des x .

Voici d'autres exemples d'équations aux dérivées partielles :

1. L'équation de la chaleur : étant donné $a > 0$ et u_0 , on cherche $u(x, t)$ de classe \mathcal{C}^2 de $\mathbb{R} \times \mathbb{R}_+$ dans \mathbb{R} qui vérifie

$$\frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} = 0, \quad t > 0 \quad \text{et} \quad u(x, 0) = u_0(x).$$

On pourra essayer de montrer que la solution est donnée par la formule

$$u(x, t) = (2\pi at)^{-1/2} \int_{\mathbb{R}} u_0(x - y) e^{-\frac{y^2}{2at}} dy.$$

2. L'équation des ondes : étant donné $c \in \mathbb{R}$ on cherche $u(x, t)$ de classe \mathcal{C}^2 de \mathbb{R}^2 dans \mathbb{R} qui vérifie

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0.$$

On pourra à nouveau utiliser le changement de variable $(x, t) \rightarrow (y, z)$ pour décrire la solution à partir des conditions initiales $u_0(x) = u(x, 0)$ et $u_1(x) = \frac{\partial u}{\partial t}(x, 0)$.

3. L'équation de Burgers : on cherche $u(x, t)$ de classe \mathcal{C}^1 de $\mathbb{R} \times \mathbb{R}_+$ dans \mathbb{R} qui vérifie

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0, \quad t > 0 \quad \text{et} \quad u(x, 0) = u_0(x).$$

Il s'agit d'une équation *non-linéaire* à cause de la présence du facteur u qui signifie que la vitesse est proportionnelle à l'amplitude de la solution. On pourra essayer de montrer que u est constante le long des droites caractéristiques d'équation $x = x_0 + u_0(x_0)t$ pour tout $x_0 \in \mathbb{R}$, et de comprendre pourquoi les solutions de cette équation ne peuvent pas rester de classe \mathcal{C}^1 (ni même continues) au bout d'un temps critique $T > 0$.

4.2 Inversion locale et fonctions implicites

Un résultat classique pour les fonctions d'une variable affirme que si f est \mathcal{C}^1 et $f'(x) \neq 0$, alors f est strictement monotone sur un intervalle I contenant x et est une bijection entre les intervalles I et $f(I)$. De plus la bijection réciproque est aussi de classe \mathcal{C}^1 .

Afin d'établir un résultat du même type pour les fonctions de plusieurs variables, on utilise un résultat simple mais fondamental : le théorème du point fixe de Picard. Rappelons que x est un point fixe d'une fonction f si et seulement si $f(x) = x$.

Théorème 4.1 *Soit $E \subset \mathbb{R}^n$ et $f : E \rightarrow E$ une fonction C -Lipschitzienne avec $C < 1$. Alors f admet un unique point fixe $x \in E$.*

Ce théorème se démontre en remarquant que toute suite $(x_n)_{n \geq 0}$ définie par $x_n = f(x_{n-1})$ avec $x_0 \in E$ vérifie $\|x_{n+1} - x_n\| \leq C^n \|x_1 - x_0\|$ ce qui entraîne qu'elle est de Cauchy. Sa limite est nécessairement un point fixe de f et son unicité découle de la propriété $\|f(x) - f(y)\| \leq C\|x - y\|$.

Soit à présent f de classe \mathcal{C}^1 sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R}^n et telle que df_x est inversible pour un point $x \in U$. On peut étudier l'inversibilité de f au voisinage de x en considérant $z \in \mathbb{R}^n$ tel que $\|z - f(x)\| \leq \varepsilon$, et en cherchant une solution $y = x + h$ à l'équation $z = f(y)$. En utilisant la différentiabilité de f en x , on peut ré-écrire cette équation suivant

$$z = f(x) + df_x(h) + \phi(h),$$

où ϕ est une fonction \mathcal{C}^1 telle que $\phi(0) = 0$ et $d\phi_0 = 0$. La solution h est solution du problème de point fixe

$$h = df_x^{-1}(f(x) - z) + df_x^{-1}(\phi(h)) = F(h).$$

On vérifie alors que pour ε suffisamment petit, il existe $\alpha > 0$ tel que F est C -Lipschitzienne de $B(0, \alpha)$ dans $B(0, \alpha)$ avec $C < 1$. En appliquant théorème du point fixe, on en déduit l'existence d'un unique antécédent y de z dans $B(x, \alpha)$. Ce raisonnement nous conduit au *théorème d'inversion locale* énoncé ci-dessous.

Théorème 4.2 *Soit f de classe \mathcal{C}^1 sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R}^n et soit $x \in U$ telle que df_x est inversible. Alors il existe deux ouverts, $U' \subset U$ et V tels que $x \in U'$ et $f(x) \in V$ et tels que f est un \mathcal{C}^1 difféomorphisme de U' dans V .*

Il existe une version de ce théorème qui permet d'établir la propriété de \mathcal{C}^1 difféomorphisme sur la totalité du domaine U . C'est le *théorème d'inversion globale*.

Théorème 4.3 *Soit f de classe \mathcal{C}^1 sur un ouvert $U \subset \mathbb{R}^n$ et à valeur dans \mathbb{R}^n , et telle que f est injective sur U . Alors $f(U)$ est un ouvert V de \mathbb{R}^n . Dans ces conditions f est un \mathcal{C}^1 difféomorphisme de U dans V si et seulement si df_x est inversible en tout point x de U .*

Une application du théorème d'inversion locale concerne le problème suivant : pour f fonction \mathcal{C}^1 de deux variables, on considère l'équation $f(x, y) = 0$ et on cherche à comprendre si cette équation est en un certain sens équivalente à $y = g(x)$ où g est une fonction d'une variable. L'exemple de la fonction $f(x, y) = x^2 + y^2 - 1$ nous montre que ceci n'est possible que localement : certaines portions du cercle unité s'identifient au graphe de la fonction $y = (1 - x^2)^{1/2}$, d'autres à celui de la fonction $y = -(1 - x^2)^{1/2}$. D'autres portions, telles qu'au voisinage du point $(1, 0)$ ne peuvent s'identifier à un graphe. Le *théorème des fonctions implicites* donne un résultat général allant dans ce sens.

Théorème 4.4 *Soit f une fonction \mathcal{C}^1 de \mathbb{R}^2 dans \mathbb{R} et (x_0, y_0) un couple tel que $f(x_0, y_0) = 0$ et tel que $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Alors il existe deux intervalles ouverts I et J tels que $x_0 \in I$ et $y_0 \in J$, et une fonction g de classe \mathcal{C}^1 de I dans J telle que pour tout $(x, y) \in I \times J$ on ait*

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

De plus on a

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

La preuve de ce théorème peut s'effectuer en considérant la fonction $h(x, y) = (x, f(x, y))$ et en lui appliquant le théorème d'inversion locale qui permet de définir $(x, g(x))$ comme l'antécédent de $(x, 0)$.

Le théorème des fonctions implicites se généralise aux fonctions de plus de deux variables : si f est \mathcal{C}^1 de $\mathbb{R}^n \times \mathbb{R}^m$ dans \mathbb{R}^m avec $f(x_0, y_0) = 0$ et si sa différentielle *par rapport à la variable y* est inversible en (x_0, y_0) , alors il existe deux ouverts U et V tels que $x_0 \in U$ et $y_0 \in V$, et une fonction g de classe \mathcal{C}^1 de U dans V telle que pour tout $(x, y) \in U \times V$ on ait $f(x, y) = 0$ si et seulement si $y = g(x)$.

4.3 Extremas liés

On s'est déjà intéressé à l'étude des extremas d'une fonction f définie sur un ouvert de \mathbb{R}^n . Dans de nombreuses applications, il est fréquent que l'on cherche à minimiser ou maximiser f sur un sous-ensemble restreint, ce qui correspond à imposer une *contrainte* sur la variable x . Par exemple, minimiser une fonction f sur la sphère unité signifie que l'on impose sur x la contrainte $x_1^2 + \dots + x_n^2 - 1 = 0$.

De façon générale, étant données deux fonctions f et g de classe \mathcal{C}^1 définies sur un ouvert U de \mathbb{R}^n et à valeur dans \mathbb{R} , on considère l'ensemble

$$E := \{x \in U ; g(x) = 0\},$$

et on s'intéresse aux extrémums de f sur E . Le *théorème des extrémums liés* permet de préciser les propriétés de tels extrémums et d'en faciliter la recherche.

Théorème 4.5 *Soit $x \in E$ tel que $\nabla g(x) \neq 0$ et tel que f restreinte à E admet un extremum local en x . Alors, il existe un réel λ appelé multiplicateur de Lagrange tel que*

$$\nabla f(x) = \lambda \nabla g(x).$$

Esquignons la preuve : pour $u \in \mathbb{R}^n$ et $v \in \mathbb{R}^n$ tel que $\nabla g(x) \cdot v \neq 0$, on définit une fonction de \mathbb{R}^2 dans \mathbb{R}

$$F(s, t) = g(x + su + tv).$$

On s'intéresse aux (s, t) tels que $F(s, t) = 0$ c'est à dire $x + su + tv \in E$. Le théorème des fonctions implicites montre qu'il existe une fonction h telle que $t = h(s)$. On voit que si x est un extremum local de f sur E , alors la fonction d'une variable $G(s) = f(x + su + h(s)v)$ admet un extremum local en $s = 0$. On a donc

$$0 = G'(0) = \nabla f(x) \cdot u + h'(0) \nabla f(x) \cdot v = \nabla f(x) \cdot u - \frac{\nabla g(x) \cdot u}{\nabla g(x) \cdot v} \nabla f(x) \cdot v,$$

Ce qui montre que $\nabla f(x) \cdot u = \lambda \nabla g(x) \cdot u$ pour tout u avec $\lambda = \frac{\nabla f(x) \cdot v}{\nabla g(x) \cdot v}$, et donc $\nabla f(x) = \lambda \nabla g(x)$.

Voici finalement une généralisation dans le cas où l'ensemble E est défini par plusieurs contraintes associées à des fonctions (g_1, \dots, g_m) , c'est à dire

$$E := \{x \in U ; g_1(x) = \dots = g_m(x) = 0\},$$

Théorème 4.6 *Soit $x \in E$ tel que les vecteurs $(\nabla g_1(x), \dots, \nabla g_m(x))$ sont indépendants et tel que f restreinte à E admet un extremum local en x . Alors, il existe des réels λ_i appelés multiplicateurs de Lagrange tel que*

$$\nabla f(x) = \lambda_1 \nabla g_1(x) + \dots + \lambda_m \nabla g_m(x).$$

5 Construction de l'intégrale multiple

5.1 Intégrale sur un rectangle

Pour une fonction f d'une variable définie sur un intervalle $[a, b]$, rappelons que l'intégrale $\int_a^b f(t)dt$ est définie par un passage à la limite des *sommes de Riemann*

$$\Sigma = \sum_{k=0}^{m-1} (x_{k+1} - x_k) f(y_k),$$

(avec $a = x_0 < x_1 < \dots < x_{m-1} < x_m = b$ et $y_k \in [x_k, x_{k+1}]$) lorsque $\max_k |x_{k+1} - x_k|$ tend vers 0.

Il est aisé de reproduire cette construction pour une fonction f de n variables à valeur dans \mathbb{R} lorsque le domaine d'intégration est de forme rectangulaire, c'est à dire

$$R = [a_1, b_1] \times \dots \times [a_n, b_n].$$

Par abus de notation, on appellera "rectangle" un tel domaine, quelque soit la dimension n considérée. On notera

$$|R| = \text{vol}(R) = (b_1 - a_1) \dots (b_n - a_n),$$

le volume de R .

On peut dans ce cas définir une subdivision de R à partir de subdivisions de chaque intervalle $[a_i, b_i]$: si pour tout i , on a la subdivision

$$a_i = x_{0,i} < x_{1,i} \dots x_{m_i-1,i} < x_{m_i,i} = b_i,$$

alors l'ensemble des rectangles

$$[x_{k_1,1}, x_{k_1+1,1}] \times \dots \times [x_{k_n,n}, x_{k_n+1,n}],$$

lorsque (k_1, \dots, k_n) parcourt $\{0, \dots, m_1 - 1\} \times \dots \times \{0, \dots, m_n - 1\}$ forme une partition de R (à ceci près que certains de ces rectangles ont des faces communes). Afin de simplifier, on notera

$$\mathcal{S} = (R_j)_{j=1, \dots, M},$$

où $M = m_1 m_2 \dots m_n$ est le nombre total de rectangles présent dans \mathcal{S} .

Appelons \mathcal{S} la subdivision de R constituée par cet ensemble de rectangle et désignons par

$$\delta(\mathcal{S}) := \max\{|x_{k_i+1,i} - x_{k_i,i}| ; k_i = 0, \dots, m_i - 1, i = 1, \dots, n\}$$

le *pas* de la subdivision qui correspond au plus grand côté sur l'ensemble des rectangles R_j de \mathcal{S} .

Comme dans le cas unidimensionnel, on dit que \mathcal{S}_2 est plus fine que \mathcal{S}_1 si tout rectangle de \mathcal{S}_1 est l'union de plusieurs rectangles de \mathcal{S}_2 . Si \mathcal{S}_1 et \mathcal{S}_2 sont deux subdivisions de R alors on peut construire une subdivision $\mathcal{S}_3 = \mathcal{S}_3(\mathcal{S}_1, \mathcal{S}_2)$ plus fine que \mathcal{S}_1 et \mathcal{S}_2 en considérant pour chaque $i = 1, \dots, n$ l'union des subdivisions d'une variable associées à \mathcal{S}_1 et \mathcal{S}_2 .

Pour $\mathcal{S} = (R_j)_{j=1, \dots, M}$ une subdivision de R et $y_j \in R_j$, on définit la somme de Riemann

$$\Sigma = \Sigma(\mathcal{S}) = \sum_{j=1}^M f(y_j) |R_j|.$$

Remarquons que Σ dépend aussi du choix des y_j dans chaque rectangle R_j . On peut par exemple décider de prendre pour y_j le centre de R_j .

Définition 5.1 Une fonction f définie sur R est dite *intégrable* (au sens de Riemann) sur R si et seulement si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour toutes subdivisions \mathcal{S}_1 et \mathcal{S}_2 de pas inférieur à α , on a

$$|\Sigma(\mathcal{S}_1) - \Sigma(\mathcal{S}_2)| \leq \varepsilon,$$

pour toutes les sommes de Riemann associées à \mathcal{S}_1 et \mathcal{S}_2 . En conséquence il existe un nombre I , tel que

$$|I - \Sigma(\mathcal{S})| \leq \varepsilon.$$

pour tout \mathcal{S} tel que $\delta(\mathcal{S}) \leq \alpha$, c'est à dire $I = \lim_{\delta(\mathcal{S})=0} \Sigma(\mathcal{S})$. On dit que I est l'intégrale de f sur R et on la note

$$I = \int_R f(x) dx.$$

Théorème 5.1 Si f est une fonction continue sur R , alors elle est intégrable sur R .

La preuve de ce théorème s'effectue comme dans le cas des fonctions d'une variable : on utilise la subdivision $\mathcal{S}_3(\mathcal{S}_1, \mathcal{S}_2)$ pour re-exprimer les sommes de Riemann $\Sigma(\mathcal{S}_1)$ et $\Sigma(\mathcal{S}_2)$, et on contrôle ensuite la différence $|\Sigma(\mathcal{S}_1) - \Sigma(\mathcal{S}_2)|$ grâce à la continuité uniforme de f sur le compact R .

On utilise parfois une autre notation pour l'intégrale multiple :

$$I = \int \int_R f(x_1, x_2) dx_1 dx_2,$$

dans le cas des intégrales à 2 variables,

$$I = \int \int \int_R f(x_1, x_2, x_3) dx_1 dx_2 dx_3,$$

dans le cas des intégrales à 3 variables, etc. Cette notation sera justifiée par une technique de calcul des intégrales multiples donnée en section 6.

5.2 Intégrale sur un domaine quarrable

On souhaite à présent étendre la notion d'intégrale multiple à des domaines plus généraux que des rectangles. Afin de réaliser cet objectif, nous allons nous limiter à une classe restreinte, néanmoins suffisamment riche, de domaines.

Soit E un ensemble borné de \mathbb{R}^n . Cela entraîne en particulier que E est contenu dans un rectangle suffisamment grand :

$$E \subset R = [a_1, b_1] \times \cdots \times [a_n, b_n].$$

Considérons à présent une subdivision \mathcal{S} de R . On lui associe le pavage extérieur et intérieur de E par \mathcal{S} qui sont les sous-ensembles de \mathcal{S} définis par

$$\mathcal{S}_{ext}(E) = \{R_j \in \mathcal{S} ; R_j \cap \overline{E} \neq \emptyset\},$$

et

$$\mathcal{S}_{int}(E) = \{R_j \in \mathcal{S} ; R_j \subset \overset{\circ}{E}\}.$$

On a clairement

$$\cup_{R_j \in \mathcal{S}_{int}(E)} R_j \subset E \subset \cup_{R_j \in \mathcal{S}_{ext}(E)} R_j$$

On remarque aussi que la différence entre le pavage extérieur et intérieur correspond aux rectangles qui touchent la frontière ∂E de E :

$$\mathcal{S}_{ext}(E) \setminus \mathcal{S}_{int}(E) = \{R_j \in \mathcal{S} ; R_j \cap \partial E \neq \emptyset\}.$$

Définition 5.2 Un ensemble E est dit quarrable si et seulement si la propriété suivante est vérifiée : pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tout \mathcal{S} tel que $\delta(\mathcal{S}) \leq \alpha$ on a

$$\sum_{R_j \in \mathcal{S}_{ext}(E)} |R_j| - \sum_{R_j \in \mathcal{S}_{int}(E)} |R_j| \leq \varepsilon.$$

Ceci équivaut à l'existence d'un un nombre A tel que

$$\lim_{\delta(\mathcal{S}) \rightarrow 0} \sum_{R_j \in \mathcal{S}_{ext}(E)} |R_j| = \lim_{\delta(\mathcal{S}) \rightarrow 0} \sum_{R_j \in \mathcal{S}_{int}(E)} |R_j| = A.$$

Le nombre A est appelé *volume* ou *mesure* de E (ou aire en dimension $n = 2$) et noté

$$A = |E| = \text{vol}(E).$$

Cette définition nous montre en particulier que E est quarrable si et seulement si sa frontière est quarrable et est de volume nul :

$$|\partial E| = 0.$$

Les propriétés suivantes se démontrent assez facilement :

1. Un rectangle R est quarrable et son aire définie ci-dessus est bien égale au produit de ses côtés.
2. Si E et F sont quarrables alors $E \cap F$ et $E \cup F$ le sont aussi et on a

$$|E \cup F| + |E \cap F| = |E| + |F|.$$

3. Si E et F sont quarrables alors $E \setminus F$ l'est aussi et on a

$$|E \setminus F| + |E \cap F| = |E|.$$

4. L'aire est invariante par translation et par rotation (on le démontre d'abord sur un rectangle puis sur un ensemble général).

On peut à présent définir la notion de somme de Riemman pour une fonction f définie sur E : étant donné une subdivision \mathcal{S} du rectangle R , on définit

$$\Sigma = \Sigma(\mathcal{S}) = \sum_{R_j \in \mathcal{S}(E)} f(y_j) |R_j|.$$

où $y_j \in R_j$ et $\mathcal{S}(E)$ est un ensemble vérifiant $\mathcal{S}_{int}(E) \subset \mathcal{S}(E) \subset \mathcal{S}_{ext}(E)$.

Définition 5.3 Soit E un ensemble quarrable fermé. Une fonction f définie sur E est dite *intégrable* sur E si et seulement si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour toutes subdivisions \mathcal{S}_1 et \mathcal{S}_2 de pas inférieur à α , on a

$$|\Sigma(\mathcal{S}_1) - \Sigma(\mathcal{S}_2)| \leq \varepsilon,$$

pour toutes les sommes de Riemann associées à \mathcal{S}_1 et \mathcal{S}_2 . En conséquence il existe un nombre I , tel que

$$|I - \Sigma(\mathcal{S})| \leq \varepsilon.$$

pour tout \mathcal{S} tel que $\delta(\mathcal{S}) \leq \alpha$, c'est à dire $I = \lim_{\delta(\mathcal{S}) \rightarrow 0} \Sigma(\mathcal{S})$. On dit que I est l'*intégrale* de f sur E et on la note

$$I = \int_E f(x) dx.$$

Théorème 5.2 Si f est une fonction continue sur E , alors elle est intégrable sur E .

La preuve de ce théorème ressemble à celle dans le cas du rectangle, mais il faut en plus utiliser le fait que l'ensemble est quarrable. Les propriétés suivantes se démontrent facilement :

1. L'intégrale de la constante 1 donne le volume : $\int_E 1 dx = |E|$
2. L'intégrale d'une fonction sur un domaine de volume nul est toujours nulle.
3. Linéarité de l'intégrale : $\int_E (\alpha f + \beta g)(x) dx = \alpha \int_E f(x) dx + \beta \int_E g(x) dx$.
4. Additivité par rapport aux ensembles : si E et F sont quarrables et $|E \cap F| = 0$ alors $\int_{E \cup F} f(x) dx = \int_E f(x) dx + \int_F f(x) dx$.
5. Si f et $|f|$ sont intégrables, on a $|\int_E f(x) dx| \leq \int_E |f(x)| dx$.

6. Cauchy-Schwarz : $|\int_E f(x)g(x)dx|^2 \leq (\int_E |f(x)|^2 dx)(\int_E |g(x)|^2 dx)$.
 7. Monotonie : si $f \leq g$ sur E on a $\int_E f(x)dx \leq \int_E g(x)dx$.

De la dernière propriété, on déduit que si f est continue sur E on a

$$|E| \min_{x \in E} f(x) \leq \int_E f(x)dx \leq |E| \max_{x \in E} f(x),$$

ce qui entraîne la formule de la moyenne en supposant l'ensemble E connexe et d'aire non-nulle : il existe $x \in E$ tel que

$$f(x) = \frac{1}{|E|} \int_E f(x)dx.$$

Comme dans le cas des fonctions d'une variable il est possible d'étendre la notion d'intégrale à des classes de fonctions plus générales que les fonctions continues, en particulier les fonctions continues par morceaux.

Théorème 5.3 *Si un ensemble quarrable E peut s'écrire*

$$E = \cup_{i=1}^m E_i,$$

ou chaque E_i est quarrable et $E_i \cap E_j = \emptyset$, et si f est une fonction uniformément continue sur chaque partie E_i , alors f est intégrable sur E .

On voit en particulier que si f est continue sur un fermé quarrable E et que l'on définit l'extension de f sur le rectangle R par

$$\tilde{f}(x) = f(x) \text{ si } x \in E, \quad \tilde{f}(x) = 0, \text{ si } x \notin E,$$

alors on a

$$\int_R \tilde{f}(x)dx = \int_E f(x)dx.$$

En particulier, si on définit la fonction indicatrice 1_E de E par

$$1_E(x) = 1 \text{ si } x \in E, \quad 1_E(x) = 0, \text{ si } x \notin E,$$

on a

$$\int_R 1_E(x)dx = |E|.$$

On peut remarquer au passage qu'un ensemble est quarrable si et seulement si son indicatrice est intégrable.

Il est temps de donner une classe importante d'ensembles quarrables.

Théorème 5.4 *Soit $E \subset \mathbb{R}^n$ un ensemble dont la frontière ∂E est contenue dans une union finie de boules $B_i = B(x_i, r_i)$, $i = 1, \dots, m$, avec la propriété suivante : pour chaque i , il existe un repère orthonormé de coordonnées $(\alpha_1, \dots, \alpha_n)$ centré autour du point x_i et une fonction continue f_i de \mathbb{R}^{n-1} dans \mathbb{R} , tels que $\partial E \cap B_i$ s'identifie au graphe de f_i :*

$$\partial E \cap B_i = \{(\alpha_1, \dots, \alpha_n) ; \alpha_n = f_i(\alpha_1, \dots, \alpha_{n-1}) \text{ et } \alpha_1^2 + \dots + \alpha_{n-1}^2 \leq r_i^2\}.$$

Alors l'ensemble E est quarrable.

Ce théorème se démontre en prouvant que ∂E est quarrable avec $|\partial E| = 0$. Entre autres conséquences, on voit qu'un polygone (ensemble dont la frontière est une ligne brisée fermée) est quarrable, que l'ensemble des points compris entre les graphes de deux fonctions continues $f \leq g$ restreintes à un ensemble quarrable est quarrable, etc.

Il est aussi facile d'exhiber des ensembles non-quarrables : par exemple l'ensemble E des points d'un rectangle R dont les coordonnées sont des nombres rationnels. On voit aussi que la fonction indicatrice de E n'est pas intégrable au sens de Riemann sur R (elle l'est avec une définition plus fine de l'intégrale, due à Lebesgue).

6 Calcul des intégrales multiples

6.1 Le théorème de Fubini

Il est bien connu que le calcul des intégrales des fonctions d'une variable est lié au calcul des primitives : si f est telle que $f = F'$, on a

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

Le calcul pratique des intégrales multiples se heurte au fait que la notion de primitive est uniquement bien définie pour les fonctions d'une variable. Le théorème de Fubini offre un moyen de ramener le calcul des intégrales multiples à celui des intégrales des fonctions d'une variable. Une version élémentaire de ce théorème s'énonce pour les fonctions de deux variables définies sur un rectangle.

Théorème 6.1 Soit f une fonction continue de $R = [a, b] \times [c, d]$ dans \mathbb{R} . Si $F(x) = \int_c^d f(x, y)dy$ alors F est continue et on a

$$\int \int_R f(x, y)dx dy = \int_a^b F(x)dx.$$

Ce résultat peut aussi s'écrire sous la forme

$$\int \int_R f(x, y)dx dy = \int_a^b \left(\int_c^d f(x, y)dy \right) dx.$$

En échangeant le rôle des variables x et y , on voit que l'on a aussi

$$\int \int_R f(x, y)dx dy = \int_c^d \left(\int_a^b f(x, y)dx \right) dy,$$

autrement dit l'ordre des variables d'intégration n'importe pas.

Il existe de multiples formes plus générales du théorème de Fubini. On peut tout d'abord considérer des fonctions de n variables continues sur un rectangle $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$. On a alors

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

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

On peut ensuite considérer des fonctions de n variables continues sur un ensemble quarrable E contenu dans $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Divisons l'ensemble des variables en deux parties $\{1, \dots, m\}$ et $\{m+1, \dots, n\}$. et notons $x = (x_1, \dots, x_n) := (y, z)$ avec $y = (x_1, \dots, x_m)$ et $z = (x_{m+1}, \dots, x_n)$, et $f(y, z) := f(x)$. On note $T = [a_1, b_1] \times \cdots \times [a_m, b_m]$, et

$$E_y := \{z \in \mathbb{R}^{n-m} \mid (y, z) \in E\},$$

la *section* de l'ensemble E dont les m premières coordonnées sont données par y . En supposant que les ensembles E_y sont quarrables, on peut alors montrer que la fonction F définie par

$$F(y) = \int_{E_y} f(y, z)dz$$

est intégrable au sens de Riemann, et que l'on a

$$\int_E f(x)dx = \int_T \left(\int_{E_y} f(y, z)dz \right) dy.$$

Par exemple, dans le cas des fonctions de 3 variables et avec $m = 1$, cela revient à "couper le domaine en tranches" et à calculer sur chaque tranche une intégrale de deux variables que l'on intègre ensuite par

rapport à la variable restante. A nouveau, il est possible de changer l'ordre des variables dans l'intégration. Il est intéressant de noter que pour un domaine général, la continuité de f n'entraîne pas nécessairement celle de $F(y)$: considérer par exemple une fonction constante sur un domaine en forme de L constitué de 3 carrés adjacents.

Enfin, il est possible d'appliquer le théorème de Fubini à des classes de fonctions plus générales que les fonctions continues, par exemples des fonctions continues par morceaux du type considéré dans la section précédente.

6.2 Changements de variables

Soit ϕ un C^1 difféomorphisme envoyant un ensemble F sur un $E = \phi(F)$. On vérifie aisément que F est fermé (resp. ouvert) si et seulement si E est fermé (resp. ouvert). On peut aussi montrer que E est quarrable si et seulement si F est quarrable. Si f est une fonction intégrable sur E , on va chercher à relier l'intégrale de f sur E à celle de $f \circ \phi$ sur F .

Considérons tout d'abord le cas particulier où ϕ est une application affine, c'est à dire

$$\phi(x) = Ax + b,$$

avec $b \in \mathbb{R}^n$ et A une matrice $n \times n$ inversible.

Théorème 6.2 *Lorsque ϕ est une application affine du type ci-dessus, on a*

$$\int_E f(x)dx = |\det(A)| \int_F (f \circ \phi)(y)dy.$$

La preuve de ce théorème s'effectue en remarquant que si \mathcal{S} est une subdivision en rectangles utilisée pour approcher l'intégrale sur E par la somme de Riemann

$$\Sigma = \sum_{R_j \in \mathcal{S}(E)} f(y_j)|R_j|,$$

alors on obtient un pavage $\mathcal{S}(F)$ par les parallélépipèdes $T_j = \phi^{-1}(R_j)$ tels que $|R_j| = |\det(A)||T_j|$, et on peut écrire

$$\Sigma = |\det(A)| \sum_{T_j \in \mathcal{S}(F)} (f \circ \phi)(x_j)|T_j|,$$

avec $x_j = \phi^{-1}(y_j) \in T_j$.

Intuitivement l'élément infinitésimal rectangulaire dy est transformé par ϕ en un élément infinitésimal dx avec $|dx| = |\det(A)||dy|$.

Dans le cas où ϕ est un C^1 -difféomorphisme plus général, lorsque l'on se place au voisinage de y , ϕ est proche de l'application affine

$$\tilde{\phi}(z) = \phi(y) + d\phi_y(z - y),$$

ce qui signifie (toujours intuitivement) que l'élément infinitésimal rectangulaire dy au voisinage de y est transformé par ϕ en un élément infinitésimal dx avec $|dx| = J_\phi(y)|dy|$, où $J_\phi(y) = |\det(d\phi_y)|$ est le Jacobien de ϕ en y . Ceci justifie le résultat suivant dont la preuve est difficile.

Théorème 6.3 *Lorsque ϕ est un C^1 difféomorphisme général on a*

$$\int_E f(x)dx = \int_F (f \circ \phi)(y)J_\phi(y)dy.$$

Voici l'application de ce théorème pour trois changements de variables très classiques :

1. Coordonnées polaires (r, θ) dans \mathbb{R}^2 : $(x, y) = \phi(r, \theta) = (r \cos \theta, r \sin \theta)$ et $J_\phi(r, \theta) = r$. Si $f(x, y) = g(r, \theta)$ avec $g = f \circ \phi$, on peut écrire

$$\int_E f(x, y)dx dy = \int_F g(r, \theta)rdr d\theta$$

2. Coordonnées cylindriques (r, θ, z) dans \mathbb{R}^3 : $(x, y, z) = \phi(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$ et $J_\phi(r, \theta) = r$. Si $f(x, y, z) = g(r, \theta, z)$ avec $g = f \circ \phi$, on peut écrire

$$\int_E f(x, y, z) dx dy dz = \int_F g(r, \theta, z) r dr d\theta dz$$

3. Coordonnées sphériques (r, θ, φ) dans \mathbb{R}^3 : $(x, y, z) = \phi(r, \theta, \varphi) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta)$ et $J_\phi(r, \theta) = r^2 |\sin \theta|$. Si $f(x, y, z) = g(r, \theta, \varphi)$ avec $g = f \circ \phi$, on peut écrire

$$\int_E f(x, y, z) dx dy dz = \int_F g(r, \theta, \varphi) r^2 |\sin \theta| dr d\theta d\varphi$$

6.3 Intégrales généralisées

Il est aussi possible d'étendre aux fonctions de plusieurs variables la notion d'intégrale généralisée.

Définition 6.1 Soit E un ensemble borné ou non. On dit que f est localement intégrable sur E si et seulement si elle est intégrable sur tout fermé borné quarrable contenu dans E .

Définition 6.2 Soit E un ensemble borné ou non. Une fonction f est dite "intégrable sur E " si et seulement si f et $|f|$ sont localement intégrables sur E et si pour tout $\varepsilon > 0$, il existe un fermé borné $F \subset E$ tel que pour tout fermé borné quarrable $G \subset E \setminus F$ on a

$$\int_G |f(x)| dx \leq \varepsilon.$$

On voit que cette propriété entraîne l'existence d'un nombre I tel que pour tout $\varepsilon > 0$, il existe un fermé borné $F \subset E$ tel que pour tout fermé borné quarrable G vérifiant $F \subset G \subset E$ on a

$$|I - \int_G f(x) dx| \leq \varepsilon.$$

Le nombre I est appelé l'intégrale de f sur E et est noté $I = \int_E f(x) dx$. On peut ainsi définir l'intégrale généralisée sur un domaine non-borné un ensemble fermé privé d'un point, de sa frontière, etc. La définition ci-dessus équivaut à la propriété qu'il existe une constante $C < +\infty$ telle que

$$\int_G |f(x)| dx \leq C,$$

pour tout ensemble G fermé borné quarrable contenu dans E . Cette dernière propriété est souvent plus facile à vérifier car il suffit qu'elle soit valable pour une famille \mathcal{F} de fermés bornés telle que tout fermé borné quarrable contenu dans E soit contenu dans un $G \in \mathcal{F}$. Par exemple, pour que f admette une intégrale généralisée sur \mathbb{R}^n , il suffit qu'il existe $C < +\infty$ telle que

$$\int_{B(0,R)} |f(x)| dx \leq C,$$

pour tout $R > 0$. Les propriétés classiques des intégrales énoncées dans les sections précédentes s'étendent aux intégrales généralisées. Il est en particulier possible d'appliquer les techniques de calcul basées sur le changement de variable et le théorème de Fubini pour prouver l'existence d'intégrales généralisées et les calculer.

A titre d'exemple, on peut considérer le calcul de $I = \int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy$. En se restreignant sur $B(0, R)$ et en effectuant le changement de variable en coordonnées polaires, on obtient

$$\int_{B(0,R)} e^{-(x^2+y^2)} dx dy = \int_0^{2\pi} \int_0^R e^{-r^2} r dr d\theta = 2\pi \int_0^R e^{-r^2} r dr = \pi[-e^{-r^2}]_0^R,$$

d'où $I = \pi$. D'autre part le théorème de Fubini nous montre que

$$I = \int_{\mathbb{R}} e^{-y^2} \left(\int_{\mathbb{R}} e^{-x^2} dx \right) dy = \left(\int_{\mathbb{R}} e^{-x^2} dx \right)^2.$$

Ce calcul permet ainsi d'obtenir l'intégrale de la gaussienne pour laquelle on ne dispose pas de primitive explicite : $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$.

Terminons par un résultat simple mais important : le théorème de majoration.

Théorème 6.4 *Soient f et g définies sur un ensemble E tel que f et $|f|$ sont localement intégrables sur E et $|f(x)| \leq g(x)$ pour tout $x \in E$. Si g est intégrable sur E , alors f l'est aussi.*

7 Intégrales curvilignes

7.1 Arcs orientés et courbes fermées orientées

De façon générale, une courbe du plan est définie par une fonction continue

$$\gamma = (\gamma_x, \gamma_y) : \mathbb{R} \rightarrow \mathbb{R}^2.$$

Les coordonnées γ_x et γ_y sont donc des fonctions de \mathbb{R} dans \mathbb{R} . Afin de définir la notion d'intégrale curviligne, nous allons d'abord nous restreindre à certains types de courbes.

Définition 7.1 *Un arc orienté Γ est défini par la donnée d'une fonction γ continue et injective d'un intervalle $[a, b] \subset \mathbb{R}$ dans \mathbb{R}^2 . On note*

$$\Gamma = \{\gamma(t) ; t \in [a, b]\},$$

On dit que l'arc est \mathcal{C}^1 si γ est \mathcal{C}^1 sur $]a, b[$.

On dit que $\gamma(t)$ est un *paramétrage* de Γ . Dans le cas d'un arc \mathcal{C}^1 on définit la *vitesse* du paramétrage γ au point $\gamma(t)$ de la courbe par le vecteur

$$\gamma'(t) = (\gamma'_x(t), \gamma'_y(t)).$$

Lorsqu'il est non nul, le vecteur $\gamma'(t)$ est un vecteur tangent à l'arc, de longueur $\|\gamma'(t)\|$. Il faut cependant remarquer qu'aux points où γ' s'annule, la courbe Γ n'admet pas nécessairement une tangente bien définie en tout point même si γ est \mathcal{C}^1 : considérer par exemple les courbes $\gamma(t) = (t^3, t^2)$ ou $\gamma(t) = (t^3, |t|^3)$ au point $(0, 0)$.

L'injectivité entraîne que la courbe Γ ne se croise pas elle-même. Notons que plusieurs paramétrages différents peuvent donner la même courbe. L'orientation de l'arc Γ précise le sens dans lequel on le parcourt, c'est à dire de $x_0 = \gamma(a)$ vers $x_1 = \gamma(b)$: on voit ainsi que si on définit $\mu(t) = \gamma(b + t(a - b))$ pour $t \in [0, 1]$ on obtient un paramétrage du même arc mais orienté dans le sens inverse. Le résultat facile suivant montre comment obtenir différents paramétrages du même arc avec la même orientation.

Théorème 7.1 *Deux paramétrages γ sur $[a, b]$ et μ sur $[c, d]$ définissent le même arc Γ avec la même orientation si et seulement si il existe φ une fonction continue et strictement croissante de $[c, d]$ dans $[a, b]$, telle que $\varphi(c) = a$ et $\varphi(d) = b$ et $\mu = \gamma \circ \varphi$.*

Définition 7.2 *La longueur de l'arc orienté Γ est définie par l'intégrale de la norme du vecteur vitesse*

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

Il est facile de voir que cette définition est indépendante du paramétrage : si $\mu = \gamma \circ \varphi$, on a

$$\int_a^b \|\gamma'(t)\| dt = \int_c^d \|\gamma'(\varphi(s))\| \varphi'(s) ds = \int_c^d \|\mu'(s)\| ds.$$

On vérifie aussi que la longueur est invariante par changement d'orientation du paramétrage.

Définition 7.3 *Le paramétrage γ sur $[a, b]$ est dit "normal" pour l'arc Γ si et seulement si $\|\gamma'(t)\| = 1$ pour tout $t \in]a, b[$. On a alors $L = b - a$.*

Définition 7.4 *Une courbe fermée orientée Γ est définie par la donnée d'une fonction γ continue d'un intervalle $[a, b] \subset \mathbb{R}$ dans \mathbb{R}^2 , telle que γ est injective sur $[a, b[$ et $\gamma(a) = \gamma(b)$. On note*

$$\Gamma = \{\gamma(t) ; t \in [a, b]\},$$

La fonction γ peut être étendue par périodicité à une fonction continue définie sur tout \mathbb{R} de période $b - a$. On dit que la courbe est \mathcal{C}^1 si γ ainsi prolongée est \mathcal{C}^1 sur \mathbb{R} .

A nouveau, plusieurs paramétrages différents peuvent aboutir à une même courbe fermée. On note d'autre part que puisque la courbe est fermée, on peut la parcourir à partir de n'importe quel point : par exemple les fonctions $\gamma(t) = (\sin(t), \cos(t))$ et $\mu(t) = (\sin(t + \pi/3), \cos(t + \pi/3))$ pour $t \in [0, 2\pi]$ sont deux paramétrages du cercle unité parcouru dans le même sens. Remarquons aussi que selon notre définition $\nu(t) = (\sin(2t), \cos(2t))$ ne donne pas un paramétrage du cercle sur $[0, 2\pi]$ car l'injectivité n'est pas vérifiée, mais donne un paramétrage sur $[0, \pi]$. Par analogie avec le cas des arcs orientés, on introduit la définition suivante pour l'orientation d'une courbe fermée.

Définition 7.5 *Un paramétrage μ de Γ défini sur $[c, d]$ a la même orientation que γ si et seulement si il existe une fonction φ continue et strictement croissante sur $[c, d]$ telle que $\varphi(d) - \varphi(c) = b - a$ et telle que $\mu = \gamma \circ \varphi$.*

Deux sens de parcours sont uniquement possibles : celui "des aiguilles d'une montre" dit *sens indirect* (ou négatif), et le sens inverse dit *sens direct* (ou positif). Une façon de déterminer l'orientation du paramétrage est la suivante : pour tout $y \in \mathbb{R}$ tel qu'il existe des points d'ordonnée y contenus dans la courbe Γ on considère

$$\alpha(y) = \min\{x ; (x, y) \in \Gamma\} \text{ et } \beta(y) = \max\{x ; (x, y) \in \Gamma\},$$

c'est à dire l'abscisse minimale et maximale des points de la courbe ayant l'ordonnée y . On regarde alors la deuxième coordonnée $\gamma'_y(t)$ du vecteur vitesse aux points $(\alpha(y), y)$ et $(\beta(y), y)$. On peut montrer que deux cas sont alors possibles :

1. Pour tout y , on a $\gamma'_y(t) \leq 0$ aux points $(\alpha(y), y)$ et $\gamma'_y(t) \geq 0$ aux points $(\beta(y), y)$: la courbe est orientée dans le sens direct.
2. Pour tout y , on a $\gamma'_y(t) \geq 0$ aux points $(\alpha(y), y)$ et $\gamma'_y(t) \leq 0$ aux points $(\beta(y), y)$: la courbe est orientée dans le sens indirect.

7.2 Intégrale curviligne et formule de Green-Riemann

Définition 7.6 *Soit Γ un arc orienté ou une courbe fermée orientée de paramétrage γ sur un intervalle $[a, b]$, et soit g une fonction continue de \mathbb{R}^2 dans \mathbb{R} . On pose*

$$\int_{\Gamma} g(x, y) dx = \int_a^b g(\gamma(t)) \gamma'_x(t) dt \text{ et } \int_{\Gamma} g(x, y) dy = \int_a^b g(\gamma(t)) \gamma'_y(t) dt.$$

En utilisant la formule de changement de variable, on vérifie aisément que les quantités $\int_{\Gamma} g(x, y) dx$ et $\int_{\Gamma} g(x, y) dy$ sont invariantes par changement de paramétrage préservant l'orientation. En revanche, elles changent de signe si on change d'orientation. Intuitivement $dx = \gamma'_x(t) dt$ et $dy = \gamma'_y(t) dt$ correspondent à la mesure des déplacements infinitésimaux dans les directions x et y lorsque l'on se déplace sur la courbe en faisant varier le paramètre t de dt .

Il est aussi possible de définir une intégrale où le déplacement infinitésimal est mesuré le long de la courbe.

Définition 7.7 *Soit Γ un arc orienté ou une courbe fermée orientée de paramétrage γ sur un intervalle $[a, b]$, et soit g une fonction continue de \mathbb{R}^2 dans \mathbb{R} . On appelle intégrale de g sur Γ la quantité*

$$\int_{\Gamma} g(x, y) ds = \int_a^b g(\gamma(t)) \|\gamma'(t)\| dt.$$

On peut vérifier que cette dernière quantité ne dépend pas de l'orientation du paramétrage. On retrouve les propriétés classiques des intégrales :

1. Linéarité : $\int_{\Gamma} \alpha f + \beta g = \alpha \int_{\Gamma} f + \beta \int_{\Gamma} g$ pour les intégrales par rapport à dx , dy ou ds .
2. Monotonie : $f \geq 0 \Rightarrow \int_{\Gamma} f \geq 0$ seulement pour l'intégrale par rapport à ds . En particulier $\int_{\Gamma} 1 ds$ donne la longueur de Γ .
3. Additivité : si on coupe Γ en deux parties (Γ_1, Γ_2) , on a $\int_{\Gamma} f = \int_{\Gamma_1} f + \int_{\Gamma_2} f$. Dans le cas des intégrales par rapport à dx ou dy , il faut supposer que Γ_1 et Γ_2 ont la même orientation que Γ .

Définition 7.8 Soit Γ un arc orienté ou une courbe fermée orientée, de paramétrage γ sur un intervalle $[a, b]$, et soit $g = (g_1, g_2)$ un champ de vecteur continu de \mathbb{R}^2 dans \mathbb{R}^2 . On appelle "circulation de g " le long de Γ la quantité

$$\int_{\Gamma} g_1(x, y)dx + g_2(x, y)dy = \int_a^b g(\gamma(t)) \cdot \gamma'(t)dt.$$

On vérifie que cette quantité est aussi l'intégrale sur Γ de la fonction $g(x, y) \cdot \tau(x, y)$ où τ est le vecteur unitaire tangent à la courbe orientée dans la direction choisie.

Nous allons à présent établir un lien entre les intégrales curvilignes et les intégrales multiples.

Théorème 7.2 Soit Ω un ensemble compact dont la frontière est une courbe Γ fermée et orientée dans le sens direct. Soit g une fonction de classe C^1 sur un ouvert contenant Ω . On a alors

$$\int_{\Omega} \frac{\partial g}{\partial x}(x, y)dx dy = \int_{\Gamma} g(x, y)dy$$

et

$$\int_{\Omega} \frac{\partial g}{\partial y}(x, y)dx dy = - \int_{\Gamma} g(x, y)dx$$

Il est important de noter ici la dissymétrie des rôles des variables x et y par la présence du signe $-$ dans la deuxième formule. Dans le cas d'une courbe orientée dans le sens indirect le signe $-$ est alors dans la première formule. On peut facilement démontrer ces formules dans le cas particulier où les sections horizontales et verticales de Ω sont des intervalles fermés. Pour la première formule par exemple, on considère les intervalles

$$[\alpha(y), \beta(y)] = \{x ; (x, y) \in \Omega\}, \quad (7.1)$$

où y varie dans l'intervalle $[y_0, y_1]$ en dehors duquel les sections horizontales de Γ sont vides. On applique alors le théorème de Fubini qui donne

$$\begin{aligned} \int_{\Omega} \frac{\partial g}{\partial x}(x, y)dx dy &= \int_{y_0}^{y_1} \left(\int_{\alpha(y)}^{\beta(y)} \frac{\partial g}{\partial x}(x, y)dx \right) dy \\ &= \int_{y_0}^{y_1} g(\beta(y), y)dy - \int_{y_0}^{y_1} g(\alpha(y), y)dy \\ &= \int_{\Gamma_{\beta}} g(x, y)dy - \int_{\Gamma_{\alpha}} g(x, y)dy, \end{aligned}$$

où Γ_{α} et Γ_{β} sont les deux parties de Γ définies respectivement par les paramétrage $\gamma_{\alpha}(y) = (\alpha(y), y)$ et $\gamma_{\beta}(y) = (\beta(y), y)$ pour $y \in [y_0, y_1]$. En remarquant que Γ_{β} est orientée dans le même sens que Γ et Γ_{α} dans le sens inverse, et en appliquant la propriété d'additivité, on a

$$\int_{\Gamma_{\beta}} g(x, y)dy - \int_{\Gamma_{\alpha}} g(x, y)dy = \int_{\Gamma} g(x, y)dy,$$

ce qui prouve la première formule. La preuve de la deuxième est similaire.

La *formule de Green-Riemann* synthétise les deux formules de la façon suivante : si $g = (g_1, g_2)$ est une fonction de classe C^1 sur un ouvert contenant Ω dans \mathbb{R}^2 , on a

$$\int_{\Gamma} g_1(x, y)dx + g_2(x, y)dy = \int_{\Omega} \left(\frac{\partial g_2}{\partial x}(x, y) - \frac{\partial g_1}{\partial y}(x, y) \right) dx dy.$$

On remarque que la quantité intégrée dans membre de droite est le rotationnel de g introduit dans la section §3.5. On obtient ainsi la formule de Stokes

$$\int_{\Gamma} g(x, y) \cdot \tau(x, y)ds = \int_{\Omega} \text{rot}(g)(x, y)dx dy.$$

Si on applique la formule de Green-Riemann à la fonction $\tilde{g} = (-g_2, g_1)$, on obtient

$$\int_{\Gamma} g_1(x, y)dx - g_2(x, y)dy = \int_{\Omega} \left(\frac{\partial g_1}{\partial x}(x, y) + \frac{\partial g_2}{\partial y}(x, y) \right) dx dy.$$

Le membre de gauche correspond à l'intégrale sur Γ du produit scalaire $\tilde{g}(x, y) \cdot \tau(x, y)$ qui est égal à $g(x, y) \cdot n(x, y)$ où $n(x, y)$ désigne le vecteur unitaire normal à Γ au point (x, y) et pointant vers l'extérieur de Ω . Cette quantité est appelée *flux* du champ de vecteur g sur Γ . La quantité intégrée dans le membre de droite est la divergence de g , introduite dans la section §3.5. On a ainsi établi la *formule de la divergence* (ou formule d'Ostrogradski) qui exprime que le flux de g sur Γ est égal à l'intégrale de sa divergence sur Ω :

$$\int_{\Gamma} g(x, y) \cdot n(x, y) ds = \int_{\Omega} \operatorname{div}(g)(x, y) dx dy.$$

Terminons ce cours en indiquant qu'il est possible d'étendre les notions d'intégrales curvilignes à des courbes $\Gamma = \{\gamma(t) ; t \in [a, b]\}$ définies dans \mathbb{R}^3 . Il est aussi possible de définir des intégrales sur des surfaces $S \subset \mathbb{R}^3$ du type

$$\int_S f(x, y, z) ds,$$

où ds est l'*élément d'aire* qui joue un rôle analogue à l'élément de longueur dans $\int_{\Gamma} f(x, y) ds$, et dont la définition rigoureuse dépasse le cadre de ce cours.

Le formule de Stokes s'étend à ce cadre de la façon suivante : si S est une surface de \mathbb{R}^3 délimitée par une courbe Γ on a

$$\int_{\Gamma} g(x, y, z) \cdot \tau(x, y, z) ds = \int_S \operatorname{rot}(g)(x, y, z) \cdot n(x, y, z) ds,$$

où n est le vecteur unitaire normal à la surface orienté dans le sens direct (ce qui signifie intuitivement vers le haut quand on assimile S au plan horizontal et que Γ est orientée dans le sens direct).

On peut aussi étendre la formule de la divergence : si Ω est un domaine de \mathbb{R}^3 délimité par une surface S , on a

$$\int_S g(x, y, z) \cdot n(x, y, z) ds = \int_{\Omega} \operatorname{div}(g)(x, y, z) dx dy dz.$$

où n est le vecteur unitaire normal à S pointant vers l'extérieur de Ω .