

MATH IV, Analyse

(Automne 2009)

Résumé du cours

Avertissement : Les pages qui suivent ne contiennent PAS les notes du cours. Vous n'y trouverez que le plan détaillé du cours, un aide-mémoire pour ainsi dire. Une liste complète des définitions, des énoncés des résultats, et quelques exemples fréquemment rencontrés, le tout légèrement motivé par des remarques courtes. Pour découvrir le reste, sauf grippe, vous devez suivre le cours régulièrement.

En cours, certains des énoncés ci-dessous seront démontrés, ceux qui seront admis seront précisés et parfois certaines conclusions seront laissées aux travaux dirigés ou resteront tout juste de simples exercices d'entraînement. Si vous souhaitez passer au delà des limites de notre cours, vous pouvez feuilleter les pages des années précédentes qui contiennent des exercices, preuves, notions supplémentaires ou me contacter. Communiquer est une méthode efficace pour apprendre.

Au début, les notes ci-dessous seront plus chargées en raison des sujets qui nécessitent un plus grand nombre de notions nouvelles. Au fur et à mesure que les chapitres évolueront, elles deviendront plus légères. Néanmoins, les sections et la numérotation des énoncés que vous y verrez seront cohérentes avec celles adoptées en classe. A la moindre déviation que vous constaterez, n'hésitez pas à m'avertir.

Table des matières

1	Espaces vectoriels normés : généralités	5
1.1	Définitions, exemples ; distance	5
1.2	Topologie des espaces normés	6
1.3	Normes équivalentes	7
1.4	Construire de nouveaux ensembles	8
1.5	Suites et convergence dans les espaces normés	8
2	Fonctions de plusieurs variables : limites, continuité	11
2.1	Qu'est-ce qu'une fonction de plusieurs variables, comment est-elle représentée? .	11
2.2	Limites	12
2.3	Continuité	13
2.4	Opérations sur les limites	14
2.5	Opérations sur les fonctions continues	14
2.6	Fonctions continues et topologie	15
3	Calcul différentiel	19
3.1	Dérivabilité des fonctions d'une seule variable et à valeurs réelles ; rappels, une nouvelle conception	19
3.2	Dérivées partielles	19
3.3	Différentiabilité, différentielle d'une fonction de plusieurs variables	21
3.4	Liens avec dérivées partielles ; matrice jacobienne ; fonctions de classe \mathcal{C}^1	22
3.5	Opérations sur les fonctions différentiables ; opérations sur les fonctions de classe \mathcal{C}^1	23
3.6	Le théorème des accroissements finis	25
3.7	Applications géométriques ; le gradient	26
3.8	Dérivées directionnelles	27
4	Différentielles du second ordre	29
4.1	Dérivées partielles secondes ; fonctions de classe \mathcal{C}^2	29
4.2	Opérations sur les fonctions de classe \mathcal{C}^2	30
4.3	Le théorème de Schwarz	30
4.4	La formule de Taylor du second ordre et la matrice Hessienne	31
5	Extrema	33
5.1	Définitions de base ; rappels sur les fonctions d'une seule variable	33
5.2	Extrema locaux	34
5.3	Extrema globaux ; ensembles compacts	35
5.4	Quelques outils pratiques pour les extrema locaux	36
5.5	Extrema liés, multiplicateurs de Lagrange	37

6	Géométrie des courbes et des surfaces : le théorème des fonctions implicites	39
6.1	Le théorème des fonctions implicites : énoncé, un exemple simple, une application rencontrée	39
6.2	Courbes dans \mathbb{R}^2	40
6.3	Surfaces dans \mathbb{R}^3	42
7	Intégration ; intégrales multiples	45
7.1	Pavés dans \mathbb{R}^p ; parties pavables de \mathbb{R}^p	46
7.2	Fonctions intégrables	47
7.3	Propriétés des fonctions intégrables	49
7.4	Théorème de Fubini	49
7.5	Changement de variables et intégration	50
8	Différentiation, intégration : quelques liens	51
8.1	Longueur d'arc, intégration le long d'une courbe	51
8.2	Formes différentielles : recherche d'une primitive en dimension supérieure	53
8.3	Formes différentielles : intégration	54

Chapitre 1

Espaces vectoriels normés : généralités

Ce chapitre est consacré à l'étude des *espaces vectoriels normés*. Nous aurons tendance à les appeler des espaces normés, sans quand-même oublier que ce sont en particulier des espaces vectoriels.

Vous savez tous ce que c'est qu'un K -espace vectoriel où K est un corps commutatif, le corps des scalaires. En cas d'oubli ou de trou de mémoire, faites vos révisions. Dans ce cours, le corps K sera celui des réels, noté \mathbb{R} . Sauf mention contraire, nous dirons un espace vectoriel au lieu d'un \mathbb{R} -espace vectoriel.

Rappelons la notation pour les puissances cartésiennes de \mathbb{R} . Soit $p \in \mathbb{N}$, \mathbb{N} étant comme d'habitude l'ensemble des nombres naturels.

- Si $p = 0$, alors $\mathbb{R}^p = \{0\}$.
- Si $p \in \mathbb{N}^*$, alors $\mathbb{R}^p = \{(x_1, \dots, x_p) \mid \text{tout } x_i \text{ est un nombre réel}\}$. Dans le cas où $p = 1$, nous aurons tendance à omettre les parenthèses et écrire x au lieu de (x) .

Il y a divers choix de notation pour noter les éléments d'un espace vectoriel : \bar{u} , \vec{u} . Nous n'utiliserons aucune de ces notations ornementées. Nous nous contenterons des lettres, u , v , x ... toujours en précisant les ensembles d'appartenance des objets qu'elles notent.

1.1 Définitions, exemples ; distance

Définition 1.1.1 (Norme) Soit E un \mathbb{R} -espace vectoriel. Une fonction

$$\begin{array}{lcl} \|\cdot\| & : & E \longrightarrow \mathbb{R}_+ \\ u & \longmapsto & \|u\| \end{array}$$

est dite une norme si elle satisfait les trois conditions suivantes :

- (N1) pour tout $u \in E$, u est le vecteur nul si et seulement si $\|u\| = 0$;
- (N2) pour tout $r \in \mathbb{R}$ et $u \in E$, $\|ru\| = |r| \|u\|$ ($|\cdot|$ est la valeur absolue) ;
- (N3) pour tous $u, v \in E$, $\|u+v\| \leq \|u\| + \|v\|$ (la première somme est la loi interne de l'espace vectoriel E tandis que la deuxième est celle des nombres réels).

Voici des exemples de normes :

1. La valeur absolue $|\cdot|$ d'un nombre réel est une norme dans l'espace vectoriel des nombres réels. C'est la source intuitive qui a motivé la notion de norme.
2. (**La norme euclidienne**) Soit $p \in \mathbb{N}^*$. Pour tout $x = (x_1, \dots, x_p) \in \mathbb{R}^p$,

$$\|x\|_2 = \sqrt{\sum_{i=1}^p x_i^2}$$

définit une norme dans l'espace vectoriel \mathbb{R}^p .

3. Soit $p \in \mathbb{N}^*$. Pour tout $x = (x_1, \dots, x_p) \in \mathbb{R}^p$,

$$\|x\|_1 = \sum_{i=1}^p |x_i|$$

définit une norme dans l'espace vectoriel \mathbb{R}^p .

4. Soit $p \in \mathbb{N}^*$. Pour tout $x = (x_1, \dots, x_p) \in \mathbb{R}^p$,

$$\|x\|_\infty = \max_{i=1}^p |x_i|$$

définit une norme sur l'espace vectoriel \mathbb{R}^p .

5. Soit E l'ensemble des suites réelles bornées. Donc, tout élément $x \in E$ est une suite $(x_n)_{n \in \mathbb{N}}$ telle que $\sup_{n \in \mathbb{N}} |x_n| < +\infty$. C'est un espace vectoriel de dimension infinie dont la loi interne est la somme usuelle des suites et la loi externe est la multiplication par les scalaires, "coordonnée par coordonnée". Pour tout $x \in E$, $\|x\| = \sup_{n \in \mathbb{N}} |x_n|$ définit une norme sur E .

Les normes (2)-(4) sont définies sur des espaces vectoriels de dimensions finies. Dans ce cours, sauf mention contraire, nous travaillerons en dimension finie. Les exemples ci-dessus sont des normes naturelles fréquemment rencontrées, mais pas les seules. Vérifiez que la fonction

$$\begin{aligned} \|\cdot\| : \mathbb{R}^2 &\longrightarrow \mathbb{R}_+ \\ (x, y) &\longmapsto |x + y| + |x| \end{aligned}$$

définit une norme.

Définition 1.1.2 (Espace normé) *Un espace vectoriel E est un espace vectoriel normé s'il est muni d'une norme telle que celle-ci a été définie ci-dessus. La notation est $(E, \|\cdot\|)$.*

Tous les espaces vectoriels dans les exemples ci-dessus sont donc des espaces vectoriels normés. Nous utiliserons l'appellation courte "espace normé".

Proposition 1.1.1 *Soit $(E, \|\cdot\|)$ un espace normé. Alors, l'application*

$$\begin{aligned} d : E \times E &\longrightarrow \mathbb{R}_+ \\ (x, y) &\longmapsto \|x - y\| \end{aligned}$$

définit une notion de "distance" sur E . Plus précisément, d jouit des propriétés suivantes :

- (D1) pour tous $u, v \in E$, $u = v$ si et seulement si $d(u, v) = 0$;
- (D2) pour tous u et $v \in E$, $d(u, v) = d(v, u)$;
- (D3) (Inégalité triangulaire) pour tous $u, v, w \in E$, $d(u, w) \leq d(u, v) + d(v, w)$.

1.2 Topologie des espaces normés

Définition 1.2.1 (Boule ouverte, boule fermée, sphère) *Soient $(E, \|\cdot\|)$ un espace normé, $a \in E$, $r \in \mathbb{R}_+$.*

1. La boule ouverte de centre a , de rayon r est l'ensemble

$$B(a, r) = \{x \in E \mid \|x - a\| < r\} .$$

2. La boule fermée de centre a , de rayon r est l'ensemble

$$\overline{B}(a, r) = \{x \in E \mid \|x - a\| \leq r\} .$$

3. La sphère de centre a , de rayon r est l'ensemble

$$S(a, r) = \{x \in E \mid \|x - a\| = r\}.$$

Définition 1.2.2 (Partie fermée, partie ouverte) Soient $(E, \|\cdot\|)$ un espace normé, $A \subset E$.

(ouvert) L'ensemble A est un ouvert de E si pour tout $a \in A$, il existe $r \in \mathbb{R}_+^*$ tel que la boule ouverte $B(a, r)$ soit contenue dans A .

(fermé) L'ensemble A est dit un fermé de E si son complémentaire est un ouvert de E .

Remarquons immédiatement qu'il s'ensuit du premier point de cette définition que l'ensemble vide et E sont des ouverts. Ensuite, il découle du deuxième point qu'ils sont aussi des fermés. Donc, être ouvert et fermé ne sont pas deux propriétés qui s'excluent. Par ailleurs, un ensemble peut être ni ouvert ni fermé.

Les notions d'ensemble fermé et d'ensemble ouvert sont motivées par les notions d'intervalle fermé et d'intervalle ouvert dans $(\mathbb{R}, |\cdot|)$. Mais ce ne sont pas les seuls exemples.

Proposition 1.2.1 Soit $(E, \|\cdot\|)$ un espace normé. Alors les énoncés suivants sont vrais :

1. Toute boule ouverte dans un espace normé est un ouvert.
2. Toute boule fermée dans un espace normé est fermée.
3. L'union d'une famille arbitraire d'ensembles ouverts est un ouvert. L'intersection d'un nombre fini d'ensembles fermés est fermée.
4. L'intersection d'une famille arbitraire d'ensembles fermés est fermée. L'union d'un nombre fini d'ouverts est ouverte.

Définition 1.2.3 (Partie bornée) Une partie X d'un espace normé $(E, \|\cdot\|)$ est dite bornée s'il existe $a \in E$, $r \in \mathbb{R}_+$ tels que $X \subset \overline{B}(a, r)$.

1.3 Normes équivalentes

Définition 1.3.1 (Normes équivalentes) Soit E un espace vectoriel muni de deux normes, $\|\cdot\|$ et $\|\cdot\|'$. Les deux normes sont dites équivalentes s'il existe deux nombres réels r, s strictement positifs tels que pour tout $x \in E$

$$r\|x\|_1 \leq \|x\|_2 \leq s\|x\|_1.$$

Nous noterons $\|\cdot\| \sim \|\cdot\|'$. Notons aussi que r et s n'ont aucune raison d'être uniques.

Proposition 1.3.1 Soient E un espace vectoriel et $\|\cdot\|, \|\cdot\|', \|\cdot\|''$ trois normes définies dans E .

1. Toute norme définie dans E est équivalente à elle-même (relation réflexive).
2. Si $\|\cdot\|$ et $\|\cdot\|'$ sont deux normes définies dans E alors $\|\cdot\| \sim \|\cdot\|'$ si et seulement si $\|\cdot\|' \sim \|\cdot\|$ (relation symétrique).
3. Si $\|\cdot\|, \|\cdot\|'$ et $\|\cdot\|''$ sont trois normes définies dans E , $\|\cdot\| \sim \|\cdot\|'$ et que $\|\cdot\|' \sim \|\cdot\|''$, alors $\|\cdot\| \sim \|\cdot\|''$ (relation transitive).

En résumé, \sim est une relation d'équivalence.

Voici une raison pour laquelle il est important de savoir si deux normes sont équivalentes :

Proposition 1.3.2 Soit E un espace vectoriel muni de deux normes $\|\cdot\|$ et $\|\cdot\|'$. Si ces deux normes sont équivalentes, alors les ouverts de E par rapport à $\|\cdot\|$ sont les mêmes que ses ouverts par rapport à $\|\cdot\|'$. Par conséquent, il en est de même pour les fermés.

La preuve du théorème suivante, quoique abordable avec les outils de ce cours, nécessite plus de connaissances que n'en ont été fournies à ce stade. Néanmoins, nous l'énonçons dès maintenant et nous l'utiliserons librement.

Théorème 1.3.1 *Dans \mathbb{R}^p , plus généralement dans un espace vectoriel normé de dimension finie toutes les normes sont équivalentes.*

Nous montrerons le cas particulier suivant en utilisant des méthodes élémentaires :

Proposition 1.3.3 *Les trois normes $\| \cdot \|_1$, $\| \cdot \|_2$ et $\| \cdot \|_\infty$ dans \mathbb{R}^p ($p \in \mathbb{N}^*$) sont équivalentes.*

1.4 Construire de nouveaux ensembles

Définition 1.4.1 *Soient $(E, \| \cdot \|)$ un espace normé et $X \subset E$.*

1. L'intérieur de X , noté $\overset{\circ}{X}$ est l'ensemble suivant :

$$\{x \in X \mid \text{il existe } r \in \mathbb{R}_+^* \text{ tel que } B(x, r) \subset X\} .$$

En d'autres termes, il existe un ouvert de O de E qui est contenu dans X et auquel appartient x . Parfois, on dit que X est un voisinage de x si $x \in \overset{\circ}{X}$.

2. L'adhérence de X , notée \overline{X} est l'ensemble des points $x \in E$ tels que tout ouvert contenant x a une intersection non vide avec X (un tel ouvert "rencontre" X).

3. La frontière de X , notée $\text{Fr}(X)$ est l'ensemble $\overline{X} \cap \overline{(E \setminus X)}$.

Proposition 1.4.1 *Soient $(E, \| \cdot \|)$ un espace normé et $A \subset E$.*

1. L'intérieur $\overset{\circ}{A}$ de A est l'ouvert le plus large de E contenu dans A . En particulier, A est ouvert si et seulement si $A = \overset{\circ}{A}$.

2. L'adhérence \overline{A} de A est le fermé le plus petit de E qui contient A . En particulier, A est fermé si et seulement si $A = \overline{A}$.

3. Un point x de E appartient à $\text{Fr}(A)$ si et seulement tout voisinage de x rencontre A ou $E \setminus A$ en un point différent de x .

4. $\overline{A} = A \cup \text{Fr}(A)$.

1.5 Suites et convergence dans les espaces normés

La notion de norme généralise celle de valeur absolue et permet de parler de la distance entre deux points. Par conséquent, on s'attend à ce que cette notion fournisse une notion de *convergence* dans un espace normé arbitraire. C'est en effet le cas, la convergence d'une suite de points d'un espace normé nous servira en étudiant diverses propriétés liées à la topologie telles que l'adhérence, la continuité.

Définition 1.5.1 *Soient $(E, \| \cdot \|)$ un espace normé, $(x_n)_{n \in \mathbb{N}}$ une suite de points de E . La suite $(x_n)_{n \in \mathbb{N}}$ est dite de converger vers un point l dans E si la condition suivante est satisfaite :*

$$\text{pour tout } r \in \mathbb{R}_+^*, \text{ il existe } N \in \mathbb{N} \text{ tel que } n \geq N \text{ implique } \|x_n - l\| < r ;$$

de manière équivalente

$$\text{pour tout } r \in \mathbb{R}_+^*, \text{ il existe } N \in \mathbb{N} \text{ tel que } n \geq N \text{ implique } x_n \in B(l, r) .$$

Nous noterons ce phénomène par la notation usuelle :

$$\lim_{n \rightarrow \infty} x_n = l .$$

Une remarque importante sur la notion de convergence est qu'elle ne change pas si la norme dans sa définition est remplacée par une norme équivalente. Ceci est particulièrement utile dans \mathbb{R}^p .

Proposition 1.5.1 Soient $p \in \mathbb{N}^*$, $\| \cdot \|$ une norme définie dans \mathbb{R}^p et $(x_n)_{n \in \mathbb{N}}$ une suite de points dans \mathbb{R}^p . Notons $x_n = (x_{n,1}, \dots, x_{n,p})$ les coordonnées de chaque élément de la suite. Alors

$$\lim_{n \rightarrow \infty} x_n = (l_1, \dots, l_p)$$

si et seulement si pour chaque $i \in \{1, \dots, p\}$

$$\lim_{n \rightarrow \infty} x_{n,i} = l_i .$$

Notons que dans cette proposition le choix de la norme ne change rien à la conclusion puisque toutes les normes sont équivalentes dans \mathbb{R}^p .

Proposition 1.5.2 (Propriétés élémentaires de la convergence) Soient $(E, \| \cdot \|)$ un espace normé et $(x_n)_{n \in \mathbb{N}}$ et $(y_n)_{n \in \mathbb{N}}$ deux suites dans E qui convergent vers k et l respectivement. Alors

1. $\lim_{n \rightarrow \infty} (x_n + y_n) = k + l$;
2. pour tout scalaire $r \in \mathbb{R}$, $\lim_{n \in \mathbb{N}} (rx_n) = rk$;
3. toute suite extraite d'une suite convergente converge vers la même limite ;
4. toute suite convergente est bornée.

Voici un usage "topologique" de la convergence.

Proposition 1.5.3 Soient $(E, \| \cdot \|)$ un espace normé et $A \subset E$. Un point $x \in \overline{A}$ si et seulement si il existe une suite $(a_n)_{n \in \mathbb{N}}$ convergente vers x telle que tout $a_n \in A$.

Chapitre 2

Fonctions de plusieurs variables : limites, continuité

2.1 Qu'est-ce qu'une fonction de plusieurs variables, comment est-elle représentée ?

Une fonction est en général une loi qui associe à chaque élément d'une partie d'un *ensemble de départ* un élément et un seul membre d'un *ensemble d'arrivée*. En voici une :

$$\begin{array}{lll} \# & : & \text{Etres humains} \longrightarrow \mathbb{N} \\ & & x \longmapsto \text{numéro d'étudiant à l'UCBL en 2009-10} \end{array}$$

La loi $\#$, sauf erreur administrative, définit une fonction. Son *domaine* est le sous-ensemble des étudiants à l'UCBL, une partie bien plus petite que l'ensemble de départ affiché. Bien sûr, l'ensemble d'arrivée est aussi infiniment plus large que l'*image* (ou l'*image directe*) de $\#$. Tout ça, c'est bien loin des espaces normés, n'est-ce pas ?

Précisons ce que nous entendons par une fonction de plusieurs variables. Soient $p, q \in \mathbb{N}^*$ et $D \subset \mathbb{R}^p$. Une fonction f définie sur D et d'ensemble d'arrivée est une fonction de plusieurs variables. Bien évidemment, quand $p = 1$, la fonction est d'une seule variable, un cas particulier qui fera néanmoins partie de la discussion générale des fonctions de plusieurs variables. Si ce cas particulier présente des aspects particuliers non vérifiés en général, ceci sera précisé.

Le cas des fonctions $p = q = 1$ était systématiquement étudié dans des cours précédents d'analyse. Dans ce cours, nous essayerons de développer une étude aussi systématique que possible des fonctions de plusieurs variables en utilisant la notion de norme.

Pour certaines valeurs de p et de q , il existe une terminologie spécifique. Introduisons cette terminologie aussi quoique ceci ne représente rien d'indispensable pour ce que nous ferons. Ainsi, une fonction de plusieurs variables avec $q = 1$ est dite une fonction *scalaire* et les fonctions avec $q > 1$ sont parfois dites *vectérielles*.

Comment représenter une fonction de plusieurs variables ? En outre de sa définition en tant que loi entre deux ensembles fixés, il existe des représentations des fonction de plusieurs variables d'intérêt géométrique.

Définition 2.1.1 (Le graphe d'une fonction) Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$ et $f : D \longrightarrow \mathbb{R}^q$ une fonction. Le graphe de f est l'ensemble suivant :

$$\mathcal{G}(f) = \{(x_1, \dots, x_p, y_1, \dots, y_q) \in \mathbb{R}^{p+q} \mid f(x_1, \dots, x_p) = (y_1, \dots, y_q)\}.$$

Cette notion si bien connue a une valeur géométrique importante pour ce cours. Illustrons ceci par deux exemples simples et non moins connus.

$$\begin{array}{lll} f & : & \mathbb{R} \longrightarrow \mathbb{R} \\ & & x \longmapsto x^2 \end{array}$$

Le graphe de cette fonction

$$\mathcal{G}(f) = \{(x, x^2) \mid x \in \mathbb{R}\}$$

est une *courbe* dans \mathbb{R}^2 , une parabole. Nous pouvons élaborer notre exemple un peu pour définir

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x_1, x_2) \longmapsto x_1^2 + x_2^2$$

Le graphe de cette fonction

$$\mathcal{G}(f) = \{(x_1, x_2, x_1^2 + x_2^2) \mid (x_1, x_2) \in \mathbb{R}^2\}$$

est une *surface* dans \mathbb{R}^3 , un *paraboloïde*.

Afin d'être étudiés, ces objets géométriques sont parfois entièrement dessinés dans les limites du possible. Parfois, on se restreint à des présentations partielles. Des exemples de celles-ci seront présentés en cours. Ici, nous nous contentons de donner une définition importante qui sera utile tout au long du cours et qui concerne le cas des fonctions scalaires de deux variables.

Définition 2.1.2 (Lignes de niveau) Soient $p = 2$, $q = 1$, $D \subset \mathbb{R}^2$ et $f : D \longrightarrow \mathbb{R}$. Fixons un réel k . La ligne de niveau k est l'ensemble

$$L_k = \{(x_1, x_2) \in D \mid f(x_1, x_2) = k\},$$

en d'autres c'est l'image inverse du singleton $\{k\} : f^{-1}(\{k\})$.

Intuitivement, il s'agit de la "projection" sur \mathbb{R}^2 de l'intersection du graphe $\mathcal{G}(f)$ avec le "plan" $\{(x_1, x_2, k) \in \mathbb{R}^3 \mid (x_1, x_2) \in \mathbb{R}^2\}$. Cette notion naturelle sera utile pour l'étude de certaines fonctions.

2.2 Limites

Comment étendre la notion de limite aux fonctions de plusieurs variables ? Nous utiliserons la notion de norme. Notons que les normes sur \mathbb{R}^p ni sur \mathbb{R}^q ne seront précisées puisque c'est inutile. En effet, la définition utilise des boules ouvertes autour de certains points fixés. Comme les normes sont équivalentes en dimension finie, les boules ouvertes (resp. fermées) gardent leurs natures topologiques quand la norme change.

La notion d'adhérence joue un rôle important dans la définition. En particulier, un point est dit *adhérent* à une partie de \mathbb{R}^p s'il appartient à l'adhérence de cette partie.

Définition 2.2.1 (Limite d'une fonction de plusieurs variables en un point adhérent à son domaine) Soient $p, q \in \mathbb{N}^*$, $f : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ une fonction de plusieurs variables définie sur $D \subset \mathbb{R}^p$. Soit $a \in \mathbb{R}^p$ un point adhérent à D . La fonction f est dite d'avoir la limite b au point a ou de tendre vers b quand $x \in D$ tend vers a si pour tout $\epsilon \in \mathbb{R}_+^*$ il existe $\delta \in \mathbb{R}_+^*$ tel que

$$\|x - a\| < \delta \Rightarrow \|f(x) - b\| < \epsilon.$$

Ce fait sera noté

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

Il est très important de noter que a peut ne pas appartenir à D . Voici un exemple où ceci joue un rôle :

$$f : \mathbb{R}_+ \longrightarrow \mathbb{R} \\ x \longmapsto \begin{cases} x^2 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \end{cases}$$

Si nous posons $D = \mathbb{R}_+^*$, alors $\lim_{x \rightarrow 0} f(x) = 0$.

Voici deux propriétés de la notion de limite, telle que nous l'avons introduite en plusieurs variables, qui nous disent que la notion est robuste :

Proposition 2.2.1 Nous gardons la notation de la définition 2.2.1. Soit $f : D \subset \mathbb{R}^p \longrightarrow \mathbb{R}^q$ une fonction de plusieurs variables.

1. La limite de f en un point $a \in \overline{D}$ est unique si elle existe.
2. La limite de f en un point est indépendante du choix de normes dans \mathbb{R}^p et dans \mathbb{R}^q (Plus généralement, des normes équivalentes dans l'ensemble de départ et dans l'ensemble d'arrivée donnent la même limite).

Comme dans le cas des fonctions d'une seule variable, il est possible de caractériser la notion de limite en utilisant la notion de convergence :

Proposition 2.2.2 Nous gardons la notation de la définition 2.2.1. Soit $D \subset \mathbb{R}^p$ ($p \in \mathbb{N}^*$). On considère $f : D \longrightarrow \mathbb{R}^q$ ($q \in \mathbb{N}^*$). Si $a \in \overline{D}$, alors

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

si et seulement si pour toute suite $(x_n)_{n \in \mathbb{N}}$ dans D qui converge à a , $\lim_{n \rightarrow +\infty} f(x_n) = l$.

2.3 Continuité

Définition 2.3.1 (Continuité d'une fonction de plusieurs variables en un point de son domaine) Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$ et $f : D \longrightarrow \mathbb{R}^q$ une application définie sur D . L'application f est dite continue en $a \in D$ si $\lim_{x \rightarrow a} f(x) = f(a)$.

Cette définition implique les trois points suivants :

1. le point a est dans le domaine de f ;
2. la fonction f a une limite en a ;
3. la limite de f en a est $f(a)$.

Définition 2.3.2 (Continuité d'une fonction de plusieurs variables sur une partie de son domaine) Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$ et $f : D \longrightarrow \mathbb{R}^q$ une application définie sur D . L'application f est dite continue sur A si elle est continue à chaque point de A .

Proposition 2.3.1 (Caractérisations de la continuité) Soient f, D, p, q, a comme dans la définition 2.3.1. Alors les conditions suivantes sont équivalentes :

1. f est continue en a ;
2. pour tout $\epsilon \in \mathbb{R}_+^*$, il existe $\delta \in \mathbb{R}_+^*$ tel que

$$\|x - a\| < \delta \text{ entraîne } \|f(x) - f(a)\| < \epsilon ;$$

3. pour toute suite $(x_n)_{n \in \mathbb{N}}$ dans D ,

$$\text{si } \lim_{n \rightarrow +\infty} x_n = a \text{ alors } \lim_{n \rightarrow +\infty} f(x_n) = f(a)$$

$$\left(f\left(\lim_{n \rightarrow +\infty} x_n\right) = \lim_{n \rightarrow +\infty} f(x_n) \right).$$

Citons quelques exemples aussi simples qu'utiles de fonctions continues. La proposition précédente ainsi que les propriétés de la convergence (la proposition 1.5.2) sont clés à la vérification de leur continuité. Ces exemples et d'autres seront abordés en détail en cours et en travaux dirigés. Notons que, sauf mention contraire, p, q sont arbitrairement fixés dans \mathbb{N}^* .

1. La fonction identité :

$$f : \mathbb{R}^p \longrightarrow \mathbb{R}^p \\ x \longmapsto x$$

2. La fonction constante, c étant un point fixé dans \mathbb{R}^q

$$\begin{aligned} f &: \mathbb{R}^p \longrightarrow \mathbb{R}^q \\ x &\longmapsto c \end{aligned}$$

3. La projection sur la i ème coordonnée ($1 \leq i \leq p$)

$$\begin{aligned} \pi_i &: \mathbb{R}^p \longrightarrow \mathbb{R}^q \\ x = (x_1, \dots, x_i, \dots, x_p) &\longmapsto x_i \end{aligned}$$

4. La somme dans \mathbb{R}^p

$$\begin{aligned} + &: \mathbb{R}^p \times \mathbb{R}^p \longrightarrow \mathbb{R}^p \\ (x, y) &\longmapsto x + y \end{aligned}$$

5. Le produit dans \mathbb{R}

$$\begin{aligned} \cdot &: \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) &\longmapsto xy \end{aligned}$$

2.4 Opérations sur les limites

Les diverses opérations telles que la somme, le produit, le quotientement, la composition des fonctions ont des effets sur les limites qui sont réminiscentes de ce qui se passe dans le cas des fonctions d'une seule variable... quitte à prendre soin de certaines subtilités. Comme c'est souvent fait en mathématiques, elles permettent d'étudier des propriétés des fonctions "compliquées", qui sont déjà connues pour celles qui sont plus "simples". La section suivante suit la même approche dans l'étude de la continuité.

Proposition 2.4.1 Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$, g, f deux fonctions de \mathbb{R}^p vers \mathbb{R}^q définies sur D et a un point adhérent à D . S'il existe deux points l_f et l_g dans \mathbb{R}^q telles que

$$\lim_{x \rightarrow a} f(x) = l_f \text{ et } \lim_{x \rightarrow a} g(x) = l_g ,$$

alors

1. $\lim_{x \rightarrow a} (f + g)(x) = \lim_{x \rightarrow a} (f(x) + g(x)) = l_f + l_g ;$
2. pour $q = 1$, $\lim_{x \rightarrow a} (fg)(x) = \lim_{x \rightarrow a} (f(x) \cdot g(x)) = l_f \cdot l_g ;$
3. pour $q = 1$ et à condition qu'il existe un voisinage de l_g où $g(x) \neq 0$,
 $\lim_{x \rightarrow a} (f/g)(x) = \lim_{x \rightarrow a} (f(x)/g(x)) = l_f/l_g.$

Proposition 2.4.2 Soient $m, p, q \in \mathbb{N}^*$, $D_f \subset \mathbb{R}^m$, $D_g \subset \mathbb{R}^p$, f et g deux fonctions de \mathbb{R}^m vers \mathbb{R}^p et de \mathbb{R}^p vers \mathbb{R}^q , définies sur D_f et D_g respectivement. Si $a \in \overline{D_f}$, $b \in \overline{D_g}$, $l \in \mathbb{R}^q$, $\lim_{x \rightarrow a} f(x) = b$, $\lim_{y \rightarrow b} g(y) = l$ et $f(D_f) \subset D_g$ (en d'autres termes, la composition $g \circ f$ est définie sur D_f), alors

$$\lim_{x \rightarrow a} (g \circ f)(x) = l .$$

Les propositions 2.4.1 et 2.4.2 fournissent des règles de calcul pratiques pour le calcul des limites qui ont des conséquences sur la continuité aussi.

2.5 Opérations sur les fonctions continues

Les propositions 2.4.1 et 2.4.2 s'adaptent à la discussion de la continuité et permettent de conclure la continuité des fonctions "compliquées" à partir des fonctions "simples".

Proposition 2.5.1 Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$, g, f deux fonctions de \mathbb{R}^p vers \mathbb{R}^q définies sur D , et $a \in D$. Si f et g sont continues en a , alors

1. $f + g$ est continue en a ;
2. pour $q = 1$, fg est continue en a ;
3. pour $q = 1$ et à condition qu'il existe un voisinage de $g(a)$ où $g(x) \neq 0$, f/g est continue en a .

Proposition 2.5.2 Soient $m, p, q \in \mathbb{N}^*$, $D_f \subset \mathbb{R}^m$, $D_g \subset \mathbb{R}^p$, f et g deux fonctions de \mathbb{R}^m vers \mathbb{R}^p et de \mathbb{R}^p vers \mathbb{R}^q , définies sur D_f et D_g respectivement. Si $a \in D_f$, $b \in D_g$, $f(a) = b$, f est continue en a , g est continue en $f(a)$, et $f(D_f) \subset D_g$ (en d'autres termes, la composition $g \circ f$ est définie sur D_f), alors $g \circ f$ est continue en a .

Proposition 2.5.3 Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$, f_1, \dots, f_q des fonctions de \mathbb{R}^p dans \mathbb{R} définies respectivement sur D_i ($1 \leq i \leq q$), et $a \in D = D_1 \cap \dots \cap D_q$. Alors toutes les fonctions f_1, \dots, f_q sont continues en a si et seulement si la fonction

$$\begin{aligned} f &: D \longrightarrow \mathbb{R}^q \\ x &\longmapsto (f_1(x), \dots, f_q(x)) \end{aligned}$$

est continue en a .

Des applications des trois propositions précédentes seront données en cours et en travaux dirigés. Notons dès maintenant que les polynômes de p variables ($p \in \mathbb{N}^*$) sont continus sur \mathbb{R}^p .

2.6 Fonctions continues et topologie

Dans tout domaine des mathématiques, les notions principales sont liées à des familles particulières de fonctions. En analyse, les notions topologiques, donc les notions qui sont définies à partir des notions d'ensemble fermé et d'ensemble ouvert sont étroitement liées aux fonctions continues. Le théorème suivant très important illustre bien ce lien.

Théorème 2.6.1 (Caractérisation topologique des applications continues) Soient $p, q \in \mathbb{N}^*$, et $f : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ une application. Les conditions suivantes sont équivalentes :

1. f est une application continue ;
2. si O est une partie ouverte de \mathbb{R}^q , alors il en est de même pour $f^{-1}(O)$;
3. si F est une partie fermée de \mathbb{R}^q , alors il en est de même pour $f^{-1}(F)$.

Nous admettrons ce théorème bien que sa preuve soit à la portée de nos connaissances. Ceux qui s'intéressent à la connaître peuvent s'adresser aux pages des années précédentes. Ce qui est important est de comprendre bien son énoncé et de pouvoir l'appliquer correctement. Donnons un exemple simple pour souligner son énorme valeur pratique.

Nous nous posons la question suivante. Nous avons tendance à croire que la sphère unité de centre l'origine dans \mathbb{R}^p est un fermé de \mathbb{R}^p :

$$S((0, \dots, 0), 1) = \{x \in \mathbb{R}^p \mid \|x\| = 1\}.$$

La vérification directe de cette conclusion est faisable mais nécessite un certain calcul. Or, nous pouvons utiliser le théorème 2.6.1 très efficacement pour répondre à la question. En effet, l'application

$$\begin{aligned} f &: \mathbb{R}^p && \longrightarrow \mathbb{R} \\ (x_1, \dots, x_p) &&& \longmapsto x_1^2 + \dots + x_p^2 \end{aligned}$$

est continue partout dans \mathbb{R}^p . En plus, la sphère unité $S((0, \dots, 0), 1)$ est exactement $f^{-1}(\{1\})$, l'image inverse du singleton $\{1\}$ dans \mathbb{R} . Or, ce dernier ensemble est un fermé de \mathbb{R} . Par conséquent, le théorème 2.6.1 montre que son image inverse est fermée dans \mathbb{R}^p . Pour arriver à cette conclusion, la fonction que nous avons utilisée n'est pas le seul choix, mais il suffit d'en trouver une.

Le théorème 2.6.1 a beaucoup d'autres applications que nous verrons tout au long de ce cours. Néanmoins, il faut bien comprendre ses hypothèses et en savoir la portée. Par exemple, l'énoncé ne dit aucunement que l'image directe d'un ouvert (resp. fermé) est ouverte (resp. fermée). Donnons un exemple très simple. Soient $p, q \in \mathbb{N}^*$ et $c \in \mathbb{R}^q$ un point fixé. Considérons l'application constante

$$\begin{aligned} f : \mathbb{R}^p &\longrightarrow \mathbb{R}^q \\ x &\longmapsto c \end{aligned}$$

L'espace \mathbb{R}^p est un ouvert de \mathbb{R}^p . Son image directe sous f est le singleton $\{c\}$ qui n'est pas un ouvert de \mathbb{R}^q .

Pour continuer notre étude, nous introduisons une notion fondamentale, celle d'une partie *compacte* de \mathbb{R}^p (plus généralement, un espace normé de dimension finie).

Définition 2.6.1 (Parties compactes de \mathbb{R}^p) *Un ensemble K dans \mathbb{R}^p est dit compact si K est fermé et borné.*

Cette définition est un mensonge! Non, pas tout à fait, c'est un cas particulier de la notion générale d'ensemble compact. Celle-ci sera donnée au cours de topologie, et pour mieux l'apprécier il est indispensable de comprendre et de retenir toutes les propriétés des ensembles compacts de notre cas particulier.

Commençons par une caractérisation très utile dont nous admettrons la preuve :

Théorème 2.6.2 *Soient $p \in \mathbb{N}^*$ et $K \subset \mathbb{R}^p$. Alors K est compact si et seulement si, de toute suite $(x_n)_{n \in \mathbb{N}^*}$ dans K (en d'autres termes, pour tout $n \in \mathbb{N}$, $x_n \in K$), on peut extraire une suite convergente.*

Illustrons ce théorème par un exemple très simple. L'intervalle $[0, 1]$ est à la fois fermé et borné dans \mathbb{R} . Il est donc compact. Voici une suite très simple

$$x_n = 0 \text{ si } n \text{ est pair} \quad ; \quad x_n = 1 \text{ si } n \text{ est impair} \quad .$$

Les suites constantes de valeur 0 et 1 respectivement sont deux sous-suites convergentes. Dans ce cas, c'est presque évident. Ce qui est remarquable est que, d'après le théorème 2.6.2, nous pouvons extraire de telles sous-suites de chaque suite dans $[0, 1]$.

Ce qui n'était pas en général vrai pour les ensembles fermés ou bornés l'est pour les ensembles compacts :

Proposition 2.6.1 (Parties compactes et fonctions continues) *Soient $p, q \in \mathbb{N}^*$ et $f : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ une application continue sur une partie D de \mathbb{R}^p . Si K est un compact de \mathbb{R}^p contenu dans D , alors $f(K)$ est aussi compact.*

La preuve de cette proposition sera admise. Pour mieux apprécier son importance, il suffit de se rappeler que l'image par rapport à une fonction continue d'un fermé n'est pas nécessairement un fermé, et que l'image d'un ensemble borné n'est pas nécessairement bornée. Voici deux exemples :

1.

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x^2+1} \end{aligned}$$

Notez que f est continue en tout point de \mathbb{R} mais que $f(\mathbb{R}) =]0, 1]$.

2.

$$\begin{aligned} g :]0, 1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x} \end{aligned}$$

La fonction g est continue sur son domaine qui est une partie bornée de \mathbb{R} . Néanmoins, son image ne l'est pas.

Avant de continuer, soulignons que la proposition 2.6.1 ne dit rien sur les images inverses des ensembles compacts. En effet, il n'est pas nécessaire que l'image inverse d'un compact soit compact. L'application constante ci-dessus fournit un exemple.

La proposition 2.6.1 a une conséquence importante et utile en pratique :

Corollaire 2.6.1 *Soient $p \in \mathbb{N}^*$, $f : \mathbb{R}^p \rightarrow \mathbb{R}$ une application continue sur une partie D de \mathbb{R} . Si K est une partie compacte de \mathbb{R}^p contenue dans D , alors f est bornée sur K et elle atteint ses bornes sur K . En d'autres termes, il existe $a, b \in K$ tels que pour tout $x \in K$,*

$$f(a) \leq f(x) \leq f(b) .$$

Pour mieux apprécier la notion suivante nous retournons au deuxième exemple ci-dessus. Nous pouvons *prolonger* l'application g ci-dessus à l'intervalle $[0, 1]$ qui est compact.

$$\begin{aligned} \bar{g} :]0, 1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} \frac{1}{x} & \text{si } x \in]0, 1] \\ c & \text{si } x = 0 \end{cases} \end{aligned}$$

(c est un réel arbitrairement fixé). Néanmoins, ce prolongement n'est plus continu.

Définition 2.6.2 (Prolongement par continuité) *Soient $p, q \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$, f une fonction de \mathbb{R}^p vers \mathbb{R}^q définie sur D . Si a est adhérent à D et que nous pouvons définir*

$$\begin{aligned} \bar{f} : D \cup \{a\} &\longrightarrow \mathbb{R}^q \\ x &\longmapsto \begin{cases} f(x) & \text{si } x \in D \\ c & \text{si } x = a \end{cases} \end{aligned}$$

de telle façon que \bar{f} soit continue sur $D \cup \{a\}$, alors la fonction \bar{f} est dite le prolongement par continuité de f au point a .

La dernière notion topologique de ce chapitre sera importante dans le calcul différentiel et dans l'intégration. Il s'agit de la *connexité par arcs*. C'est un cas particulier de la notion de connexité qui est plus géométrique et suffisante pour notre cours.

Définition 2.6.3 (Arcs dans \mathbb{R}^q) *Soit $[a, b]$ un intervalle fermé et borné de \mathbb{R} . On appelle arc une fonction $\gamma : [a, b] \rightarrow \mathbb{R}^q$ ($q \in \mathbb{N}^*$) continue sur $[a, b]$. Les points $\gamma(a)$ et $\gamma(b)$ sont dits les extrémités de l'arc. L'arc γ est dit de joindre $\gamma(a)$ à $\gamma(b)$.*

La notion d'arc ne vous est pas inconnue :

1.

$$\begin{aligned} \gamma : [0, 1] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (\cos(2\pi t), \sin(2\pi t)) \end{aligned}$$

Dans cet exemple, les extrémités sont le même point : $\gamma(0) = \gamma(1) = (1, 0)$.

2. Le segment de droite joignant deux points $x = (x_1, \dots, x_q)$ et $y = (y_1, \dots, y_q)$ dans \mathbb{R}^q ($q \in \mathbb{N}^*$).

$$\begin{aligned} \gamma : [0, 1] &\longrightarrow \mathbb{R}^q \\ t &\longmapsto ty + (1 - t)x \end{aligned}$$

Les extrémités sont $\gamma(0) = x$ et $\gamma(1) = y$.

Définition 2.6.4 (Parties connexes par arcs de \mathbb{R}^q) *Soit $C \subset \mathbb{R}^q$. L'ensemble C est dit connexe par arcs si toute paire de points sont joignables par un arc dans C . En d'autres termes, l'ensemble d'arrivée de la fonction qui définit l'arc en question est contenue dans C .*

Le deuxième exemple ci-dessus nous fournit immédiatement un exemple d'ensemble connexe par arcs dans $\mathbb{R}^q : \mathbb{R}^q$. Voici quelques (non) exemples :

1. Soit P un point fixé de \mathbb{R}^2 , alors l'ensemble $\mathbb{R}^2 \setminus \{P\}$ est connexe par arcs.
2. L'ensemble $\mathbb{R}^2 \setminus B((0, 0), 1)$ est connexe par arcs.
3. L'ensemble \mathbb{R}^* n'est pas connexe par arcs.
4. L'ensemble $\{(0, 0), (1, 1)\}$ n'est pas connexe par arcs.

Comment vérifier tout cela? Les énoncés suivants sont indispensables dans l'étude de la connexité par arcs :

Proposition 2.6.2 (Connexité par arcs et fonctions continues) *Soient $p, q \in \mathbb{N}^*$, f une application de \mathbb{R}^p vers \mathbb{R}^q continue sur une partie D de \mathbb{R}^p . Si D est connexe par arcs, alors $f(D)$ est connexe par arcs aussi.*

Théorème 2.6.3 (Connexité par arcs dans \mathbb{R}) *Une partie de \mathbb{R} est connexe par arcs si et seulement si c'est un intervalle.*

Théorème 2.6.4 (Le théorème e des valeurs intermédiaires) *Soient $p \in \mathbb{N}^*$, D une partie de connexe par arcs de \mathbb{R}^p . Si $f : \mathbb{R}^p \rightarrow \mathbb{R}$ est une fonction continue sur D et $x_1, x_2 \in D$, alors pour tout nombre y entre $f(x_1)$ et $f(x_2)$, il existe $x \in D$ tel que $f(x) = y$.*

La preuve du théorème des valeurs intermédiaires découle des deux énoncés précédents. Celle du théorème 2.6.3 nécessite plus que nous en verrons dans ce cours.

Chapitre 3

Calcul différentiel

3.1 Dérivabilité des fonctions d'une seule variable et à valeurs réelles ; rappels, une nouvelle conception

Voici une fonction très simple de variables réelles,

$$\begin{array}{ccc} f & : & \mathbb{R} \longrightarrow \mathbb{R} \\ & & x \longmapsto x^3 \end{array}$$

... si simple que nous savons tous la dériver :

$$f'(x) = 3x^2 .$$

Mais qu'est-ce que cela veut dire ? Comment l'interpréter ?

Tout d'abord, il s'agit d'une nouvelle fonction avec sa propre loi, son propre domaine, sa propre arrivée et ses propriétés. Elle associe à chaque nombre réel a la valeur réelle $3a^2$. Pour obtenir cette définition, nous calculons une limite :

$$\lim_{x \rightarrow a ; x \neq a} \frac{x^3 - a^3}{x - a} = 3a^2 .$$

Ceci équivaut à

$$\lim_{x \rightarrow a ; x \neq a} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = 0 ,$$

soit encore

$$f(x) = f(a) + f'(a)(x - a) + o(x - a) .$$

Les deux premiers termes de l'expression forment l'équation de la tangente au graphe de f au point $(a, f(a))$. La pente de cette droite est $f'(a)$, et à chaque réel a correspond une application linéaire :

$$\begin{array}{ccc} a \longmapsto f'(a) & : & \mathbb{R} \longrightarrow \mathbb{R} \\ & & t \longmapsto f'(a).t \end{array}$$

Cette observation souligne l'aspect fondamental de la différentiation. Il s'agit d'associer à chaque point du domaine, une application linéaire dont les entrées de la matrice représentante seront déterminées en utilisant les *dérivées partielles*.

3.2 Dérivées partielles

Les dérivées partielles, de simple définition, essaient d'étendre la notion de différentiation aux fonctions de plusieurs variables. Quoique trop faible pour caractériser la différentiabilité en plusieurs variables, elles fournissent un outil indispensable dans la détermination de la matrice jacobienne.

Définition 3.2.1 (Dérivées partielles en un point) Soient $p \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p et f une fonction de \mathbb{R}^p vers \mathbb{R} définie sur U . La fonction f est dite d'avoir une dérivée partielle en $a = (a_1, \dots, a_p)$ à par rapport à la i ème variable si la fonction

$$t \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_p)$$

est dérivable en a_i , en d'autres termes, si la limite

$$\lim_{t \rightarrow a_i} \frac{f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_p) - f(a)}{t - a_i}$$

existe. On note cette limite

$$\frac{\partial f}{\partial x_i}(a) \quad \text{ou} \quad \partial_i f(a) \quad \text{ou} \quad D_i f(a).$$

Cette notion de dérivée utilise de façon essentielle la notion de dérivée des fonctions d'une seule variable. Elle permet de définir une nouvelle fonction de \mathbb{R}^p vers \mathbb{R} quand elle existe sur un ouvert :

$$\begin{aligned} \partial_i f &: U \longrightarrow \mathbb{R} \\ a &\longmapsto \partial_i f(a) \end{aligned} .$$

Voici quelques exemples :

1.

$$\begin{aligned} f &: \mathbb{R}^3 \longrightarrow \mathbb{R} \\ (x_1, x_2, x_3) &\longmapsto 3x_1^4 + x_2^3 x_3 + x_1^2 x_3^5 \quad . \\ \partial_1 f &: \mathbb{R}^3 \longrightarrow \mathbb{R} \\ (x_1, x_2, x_3) &\longmapsto 12x_1^3 + 2x_1 x_3^5 \quad ; \\ \partial_2 f &: \mathbb{R}^3 \longrightarrow \mathbb{R} \quad ; \quad \partial_3 f : \mathbb{R}^3 \longrightarrow \mathbb{R} \\ (x_1, x_2, x_3) &\longmapsto 3x_2^2 x_3 \quad ; \quad (x_1, x_2, x_3) \longmapsto x_2^3 + 5x_1^2 x_3^4 \quad . \end{aligned}$$

2.

$$\begin{aligned} g &: \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases} \quad . \\ \partial_1 g &: \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \begin{cases} y \frac{y^2-x^2}{(x^2+y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases} \quad . \\ \partial_2 g &: \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \begin{cases} x \frac{x^2-y^2}{(x^2+y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases} \quad . \end{aligned}$$

Les dérivées partielles de g sont définies en tout point de \mathbb{R}^2 , mais la fonction g ne doit pas être "dérivable" en $(0, 0)$ quelle que soit la notion de dérivée qui attend d'être introduite puisqu'elle n'est pas continue à l'origine. Cette conclusion est aussi liée à l'observation suivante : les dérivées partielles de f ne sont pas continues en $(0, 0)$.

La conclusion que nous tirons du deuxième exemple est que pour être "dérivable" au voisinage d'un point il ne doit pas suffire pas d'avoir toutes les dérivées partielles sur ce voisinage. La raison principale est que les dérivées partielles dépendent des directions, donc des chemins, en l'occurrence les axes de coordonnées. Ces remarques seront précisées au long des sections qui viennent.

3.3 Différentiabilité, différentielle d'une fonction de plusieurs variables

Nous commençons immédiatement avec la “bonne” définition de la dérivabilité. Il convient de noter que cette notion correspond aussi à la nouvelle approche à la dérivabilité des fonctions d'une seule variable présentée dans la section 3.1

Définition 3.3.1 (Différentiabilité en un point) Soient $p, q \in \mathbb{N}^*$, $U \subset \mathbb{R}^p$ un ouvert, f une fonction de \mathbb{R}^p vers \mathbb{R}^q définie sur U , et $a \in U$. La fonction f est dite différentiable en a s'il existe une application linéaire notée $L \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$ telle que

$$\lim_{x \rightarrow a} \frac{\|f(x) - f(a) - L(x - a)\|_{\mathbb{R}^q}}{\|x - a\|_{\mathbb{R}^p}} = 0 ;$$

En d'autres termes, s'il existe $L \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$ telle que

$$\|f(x) - f(a) - L(x - a)\| = o(\|x - a\|) .$$

La fonction f est dite différentiable sur U si elle est différentiable en tout point de U .

Deux remarques méritent d'être faites dès le début :

1. la différentiabilité ne dépend pas de la norme choisie ;
2. si $p = q = 1$, alors la notion de différentiabilité équivaut à la dérivabilité usuelle.

Proposition 3.3.1 Si f est différentiable en a , alors elle y est continue.

Les exemples suivants seront détaillés en cours :

1. Soit $p \in \mathbb{N}^*$,

$$f : \mathbb{R}^p \longrightarrow \mathbb{R}^p \\ x \longmapsto x .$$

La fonction f est différentiable en tout point de \mathbb{R}^p . L'application linéaire qui en témoigne est la matrice identité.

- 2.

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x_1, x_2) \longmapsto x_1^2 + x_2^2 .$$

La fonction g est différentiable en tout point a de \mathbb{R}^2 . Si $a = (a_1, a_2)$ est un tel point, alors l'application linéaire qui correspond est représentée dans la base canonique par la matrice 1×2 $(2a_1 \ 2a_2) = (\partial_1 f(a) \ \partial_2 f(a))$:

$$L : \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (h_1, h_2) \longmapsto (2a_1 \ 2a_2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} .$$

Proposition 3.3.2 (Unicité de la différentielle) Nous gardons la même notation. Si la fonction $f : U \longrightarrow \mathbb{R}^q$ est différentiable en a , alors l'application linéaire associée à ce point est unique.

Cette unicité permet de définir sans ambiguïté la *fonction différentielle*.

Définition 3.3.2 (Différentielle d'une fonction en un point) Soient $p, q \in \mathbb{N}^*$, $U \subset \mathbb{R}^p$ un ouvert, f une fonction de \mathbb{R}^p vers \mathbb{R}^q définie et différentiable sur U . L'application linéaire uniquement associée à chaque point a est dite la différentielle de f en a ; l'application qui associe à $a \in U$ la différentielle de f en a est dite la différentielle de f .

La notation qui correspond à cette définition est la suivante :

$$df : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ a \longmapsto (df)_a$$

3.4 Liens avec dérivées partielles ; matrice jacobienne ; fonctions de classe \mathcal{C}^1

La proposition 3.3.1 a solidement vérifié que les dérivées partielles sont trop faibles pour conclure la différentiabilité. Néanmoins, les liens sont forts :

Proposition 3.4.1 *Soient $p \in \mathbb{N}^*$, U un ouvert dans \mathbb{R}^p , $a \in U$ et f une fonction de \mathbb{R}^p vers \mathbb{R} définie sur U . Si f est différentiable au point a , alors ses dérivées partielles en ce point sont définies. En plus, la matrice représentant l'application linéaire $(df)_a$ par rapport à la base canonique est*

$$(\partial_1 f(a) \quad \dots \quad \partial_p f(a)) \quad .$$

Résumons : si une fonction f est différentiable en un point a , alors sa différentielle est déterminée par les valeurs des dérivées partielles en ce point. Celles-ci y sont définies. Par contre, la seule existence des dérivées partielles en un point n'entraîne pas la différentiabilité en ce point. D'autres conditions seront nécessaires. Avant d'introduire ces conditions supplémentaires, nous présenterons le cas général de ce que nous avons vérifié pour les fonctions scalaires.

Nous travaillons avec la même notation, le seul changement étant le suivant : la fonction f a \mathbb{R}^q comme ensemble d'arrivée où q est un naturel non nul arbitraire. Elle est toujours supposée être différentiable en a . Notre fonction est donc de la forme :

$$f : U \longrightarrow \mathbb{R}^q \\ a \longmapsto (f_1(a), \dots, f_q(a)) \quad .$$

Alors, la différentielle de

$$df : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ a \longmapsto (df)_a$$

Alors l'application linéaire qui représente la différentielle de f en a est la matrice suivante

$$\begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_p f_1(a) \\ \vdots & & \vdots \\ \vdots & \dots & \partial_j f_i(a) \quad \dots & \vdots \\ \vdots & & \vdots \\ \partial_1 f_q(a) & \dots & \partial_p f_q(a) \end{pmatrix}_{1 \leq i \leq q ; 1 \leq j \leq p}$$

soit encore

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_p}(a) \\ \vdots & & \vdots \\ \vdots & \dots & \frac{\partial f_i}{\partial x_j}(a) \quad \dots & \vdots \\ \vdots & & \vdots \\ \frac{\partial f_q}{\partial x_1}(a) & \dots & \frac{\partial f_q}{\partial x_p}(a) \end{pmatrix}_{1 \leq i \leq q ; 1 \leq j \leq p}$$

Cette matrice est dite *la matrice jacobienne* de f au point a .

Il est temps d'introduire une condition suffisante de différentiabilité qui utilise les dérivées partielles. Une définition d'abord :

Définition 3.4.1 (Fonctions de classe \mathcal{C}^1 sur un ouvert) *Soient $p, q \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p et f une fonction de \mathbb{R}^p vers \mathbb{R}^q définie sur U . On dira que f est continûment différentiable sur U , ou de classe \mathcal{C}^1 sur U , si*

$$df : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ a \longmapsto (df)_a$$

est une application continue sur U . La notation sera $f \in \mathcal{C}^1(U)$.

3.5. OPÉRATIONS SUR LES FONCTIONS DIFFÉRENTIABLES ; OPÉRATIONS SUR LES FONCTIONS DE CLASSE \mathcal{C}^1

La continuité de df a un sens puisqu'il s'agit d'une application de \mathbb{R}^p vers \mathbb{R}^{pq} , deux espaces normés.

Muni de cet arsenal, nous pouvons énoncer un théorème qui sera de grande valeur pratique :

Théorème 3.4.1 *Soient $p, q \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une fonction dont toutes les dérivées partielles sont définies sur U . Si ces dérivées partielles sont continues sur U , alors f est de classe \mathcal{C}^1 sur U .*

Certaines remarques doivent être immédiatement faites :

1. La conclusion du théorème est strictement plus forte que la différentiabilité sur U . Des exemples de fonctions différentiables sans l'être continûment seront étudiés aux travaux dirigés.
2. Plus tard, nous introduirons la notion de fonction de classe \mathcal{C}^k pour tout $k \in \mathbb{N}^*$ en ayant recours aux dérivées partielles. Celles-ci, quoique faibles, sont des outils dont on se sert fréquemment.

3.5 Opérations sur les fonctions différentiables ; opérations sur les fonctions de classe \mathcal{C}^1

Cette section a une valeur pratique importante pour calculer des différentielles. Elle illustre aussi les liens forts entre notre cours et l'algèbre linéaire, un phénomène que nous continuerons de rencontrer.

Proposition 3.5.1 (Sommes et produits par les scalaires) *Soient $p, q \in \mathbb{N}^*$, U, V deux ouverts dans \mathbb{R}^p , f, g deux fonctions différentiables (de classe \mathcal{C}^1) sur U et V respectivement. Alors $f + g$, définie sur $U \cap V$ par la loi $x \mapsto (f + g)(x) = f(x) + g(x)$ est différentiable (resp. de classe \mathcal{C}^1) sur $U \cap V$. Sa différentielle est définie comme suit :*

$$d(f + g) : \begin{array}{l} U \cap V \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ x \longmapsto (df + dg)_x = (df)_x + (dg)_x \end{array} .$$

En pratique il s'agit de la somme de deux matrices jacobiniennes :

$$\begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_p f_1(a) \\ \vdots & & \vdots \\ \vdots & \dots & \partial_j f_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 f_q(a) & \dots & \partial_p f_q(a) \end{pmatrix} + \begin{pmatrix} \partial_1 g_1(a) & \dots & \partial_p g_1(a) \\ \vdots & & \vdots \\ \vdots & \dots & \partial_j g_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 g_q(a) & \dots & \partial_p g_q(a) \end{pmatrix} = \begin{pmatrix} \partial_1 (f_1 + g_1)(a) & \dots & \partial_p (f_1 + g_1)(a) \\ \vdots & & \vdots \\ \vdots & \dots & \partial_j (f_i + g_i)(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 (f_q + g_q)(a) & \dots & \partial_p (f_q + g_q)(a) \end{pmatrix}$$

Si $\lambda \in \mathbb{R}$, alors la fonction λf définie sur U par $x \mapsto \lambda \cdot f(x)$ est différentiable (resp. de classe \mathcal{C}^1) sur U de différentielle

$$d(\lambda f) : \begin{array}{l} U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ x \longmapsto \lambda(df)_x \end{array} .$$

En pratique il s'agit de multiplier la matrice jacobienne par une matrice scalaire :

$$\begin{aligned}
& \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ & \ddots & & & \\ & 0 & \lambda & 0 & \\ & & & \ddots & \\ 0 & & \dots & 0 & \lambda \end{pmatrix}_{q \times q} \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_p f_1(a) \\ \vdots & \ddots & \vdots \\ \vdots & \dots & \partial_j f_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 f_q(a) & \dots & \partial_p f_q(a) \end{pmatrix}_{q \times p} = \\
& \begin{pmatrix} \lambda \partial_1 f_1(a) & \dots & \lambda \partial_p f_1(a) \\ \vdots & \ddots & \vdots \\ \vdots & \dots & \lambda \partial_j f_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \lambda \partial_1 f_q(a) & \dots & \lambda \partial_p f_q(a) \end{pmatrix}_{q \times p} = \\
& \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_p f_1(a) \\ \vdots & \ddots & \vdots \\ \vdots & \dots & \partial_j f_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 f_q(a) & \dots & \partial_p f_q(a) \end{pmatrix}_{q \times p} \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ & \ddots & & & \\ & 0 & \lambda & 0 & \\ & & & \ddots & \\ 0 & & \dots & 0 & \lambda \end{pmatrix}_{p \times p}
\end{aligned}$$

Proposition 3.5.2 (Composition) Soient $p, q, r \in \mathbb{N}^*$, U et V des ouverts de \mathbb{R}^p et \mathbb{R}^q respectivement. Si f et g sont deux fonctions différentiables (de classe \mathcal{C}^1) sur U et V respectivement, et que $f(U) \subset V$, alors $g \circ f$ est différentiable (resp. \mathcal{C}^1) sur U de différentielle

$$\begin{aligned}
d(g \circ f) & : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^r) \\
x & \longmapsto (dg)_{f(x)} \circ (df)_x \quad .
\end{aligned}$$

La deuxième composition est celle de deux applications linéaires, et en pratique, correspond au produit des matrices jacobiniennes :

$$\begin{aligned}
& \begin{pmatrix} \partial_1(g \circ f)_1(a) & \dots & \partial_p(g \circ f)_1(a) \\ \vdots & \ddots & \vdots \\ \vdots & \dots & \partial_j(g \circ f)_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1(g \circ f)_r(a) & \dots & \partial_p(g \circ f)_r(a) \end{pmatrix}_{r \times p} = \\
& \begin{pmatrix} \partial_1 g_1(f(a)) & \dots & \partial_q g_1(f(a)) \\ \vdots & \ddots & \vdots \\ \vdots & \dots & \partial_j g_i(f(a)) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 g_r(f(a)) & \dots & \partial_q g_r(f(a)) \end{pmatrix}_{r \times q} \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_p f_1(a) \\ \vdots & \ddots & \vdots \\ \vdots & \dots & \partial_j f_i(a) & \dots & \vdots \\ \vdots & & \vdots & & \vdots \\ \partial_1 f_q(a) & \dots & \partial_p f_q(a) \end{pmatrix}_{q \times p}
\end{aligned}$$

Nous aborderons deux applications en détail :

1. une formule qui lie la différentielle d'une fonction à valeurs réelles et différentiable sur un ouvert à ses dérivées partielles ;
2. changement de coordonnées.

3.6 Le théorème des accroissements finis

Rappelons d'abord le théorème des accroissements finis dans le cas particulier des fonctions d'une seule variable et à valeurs réelles :

Fait 3.6.1 Soit f une fonction de \mathbb{R} vers \mathbb{R} , définie et continue sur un intervalle $[a, b] \subset \mathbb{R}$ ($a, b \in \mathbb{R}$ et $a < b$), et dérivable sur l'intérieur $]a, b[$ du même intervalle. Alors il existe $c \in]a, b[$ tel que

$$f'(c) = \frac{f(b) - f(a)}{b - a} .$$

C'est un résultat simple et fondamental dont une conséquence (parmi d'autres) est souvent utilisée : une fonction d'une seule variable et à valeurs réelles est constante sur un intervalle ouvert si et seulement si sa dérivée est nulle sur cet intervalle. La généralisation que nous étudierons est aussi importante et a des conséquences similaires... quitte à trouver des conditions suffisantes convenables.

Théorème 3.6.1 (Le théorème des accroissements finis pour les fonctions de plusieurs variables et à valeurs réelles) Soient $p \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une application différentiable sur U . Supposons aussi que U contient deux points P et Q liés par un arc γ dans U , $\gamma : [a, b] \rightarrow U$ avec γ différentiable sur l'intervalle ouvert $]a, b[$. Alors il existe $t_0 \in]a, b[$ tel que

$$f(\gamma(b)) - f(\gamma(a)) = (df)_{\gamma(t_0)} \gamma'(t_0) (b - a) ,$$

où

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

Pour mieux apprécier cet énoncé, c'est un bon exercice de retrouver le cas du fait 3.6.1 comme cas particulier de la conclusion générale du théorème 3.6.1. D'autres exemples seront étudiés en cours non seulement pour illustrer le théorème des accroissements finis mais aussi en vue des conséquences géométriques des notions de ce chapitre.

Maintenant nous citerons deux conséquences de ce théorème liant l'invariance des valeurs d'une fonction de plusieurs variables à sa différentielle. Nous le ferons en deux étapes dont la première est l'occasion d'introduire la notion importante de *convexité*

Définition 3.6.1 (Parties convexes de \mathbb{R}^p) Une partie C de \mathbb{R}^p est dite convexe si pour toute paire de points x_1, x_2 dans C , le segment de droite $\{(1-t)x_1 + tx_2 | t \in [0, 1]\}$ qui joint x_1 à x_2 est contenu dans C .

Un exemple de partie convexe est la boule ouverte (resp. fermée) de centre x et de rayon r . En effet, pour toute paire de points a et b dans $B(x, r)$ (resp. $\overline{B}(x, r)$) et toute valeur de $t \in [0, 1]$

$$\|a(1-t) + bt - x\| < r \quad (\text{resp. } \leq r) .$$

Corollaire 3.6.1 Soient $p \in \mathbb{N}^*$, U un ensemble convexe et ouvert de \mathbb{R}^p , f une application de différentielle nulle sur U . Alors, f est constante sur U .

Le deuxième corollaire est une généralisation du premier. Néanmoins, sa preuve se réduit en utilisant la notion générale de compacité, au cas particulier du corollaire 3.6.1

Corollaire 3.6.2 Soient $p \in \mathbb{N}^*$, U un ensemble connexe par arcs et ouvert de \mathbb{R}^p , f une application de différentielle nulle sur U . Alors, f est constante sur \overline{U} .

3.7 Applications géométriques ; le gradient

La différentielle d'une fonction de plusieurs variables et à valeur réelles, disons de \mathbb{R}^p vers \mathbb{R} , est une fonction qui associe une matrice ligne $1 \times p$ à chaque point de \mathbb{R}^p où elle est différentiable. Nous pouvons traiter une telle matrice comme un point ou vecteur, ce qui permet de l'interpréter et de l'utiliser de façon plus géométrique.

Définition 3.7.1 (Le vecteur gradient) Soient $p \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R} différentiable U . Pour tout $a \in U$, le vecteur $(\partial_1 f(a), \dots, \partial_p f(a))$ est dit le gradient de f . Il est noté $\nabla f(a)$ ou $\text{grad} f(a)$.

De façon similaire à la différentielle, le vecteur gradient permet de définir une fonction :

$$\begin{aligned} \text{grad} f &: U \longrightarrow \mathbb{R}^p \\ a &\longmapsto \nabla f(a) = (\partial_1 f(a), \dots, \partial_p f(a)) \end{aligned}$$

Avec cette notation, nous pouvons compactifier la définition de la différentielle d'une fonction à valeurs réelles :

$$\begin{aligned} df &: U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}) \\ a &\longmapsto \nabla f(a) \end{aligned}$$

Pour tout élément (h_1, \dots, h_p) de \mathbb{R}^p , nous obtenons l'égalité suivante :

$$\nabla f(a) \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix} = \sum_{j=1}^p \partial_j f(a) h_j .$$

Remarquablement, l'expression à droite est un *produit scalaire*, une opération de haute valeur géométrique qui permet de définir l'angle entre deux éléments de \mathbb{R}^p .

La remarque précédente, quoique applicable dans tout \mathbb{R}^p , devient observable quand $p = 2$. La proposition suivante montre son utilité dans l'étude des lignes de niveau :

Proposition 3.7.1 Soient $k \in \mathbb{R}$, U un ouvert de \mathbb{R}^2 et $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1(U)$. Si $a \in U$ appartient aussi à la ligne de niveau $L_k = \{(x_1, x_2) \in \mathbb{R}^2 \mid f(x_1, x_2) = k\}$, alors $\nabla f(a)$ est perpendiculaire à la droite tangente à la "courbe" $f(x_1, x_2) = k$ au point a .

La perpendicularité (l'orthogonalité) de deux éléments de \mathbb{R}^p (deux vecteurs) équivaut à ce que leur produit scalaire soit nul.

Plus généralement dans \mathbb{R}^p , l'étude des liens entre les angles (donc, choix de direction) et le gradient donne les conclusions suivantes : comme à un point a de différentiabilité sur un ouvert U d'une fonction f ,

$$\lim_{\|h\| \rightarrow 0} \frac{|f(a+h) - f(a) - (df)_a h|}{\|h\|} = 0 ,$$

$$\begin{aligned} f(a+h) &= f(a) + (df)_a h + o(\|h\|) \\ &= f(a) + \nabla f(a) \cdot h + o(\|h\|) \\ &= f(a) + \|\nabla f(a)\| \|h\| \cos \theta + o(\|h\|) . \end{aligned}$$

Cette valeur est maximisée quand $\cos \theta = 1$. De façon équivalente, si on ne permet à θ que des valeurs entre 0 et π , ce qui est le cas du produit scalaire, alors $\theta = 0$; soit encore que la direction de la plus grande augmentation est exactement celle du vecteur gradient. Ce phénomène se résume souvent par le slogan suivant : le "gradient indique la direction de la plus grande pente".

3.8 Dérivées directionnelles

Les dérivées partielles sont des dérivées calculées en suivant une direction particulière le long d'une droite (ou d'un vecteur à un point). La même idée s'étend à toutes les directions :

Définition 3.8.1 (Dérivées directionnelles à un point) Soient $p \in \mathbb{N}^*$, $U \subset$ un ouvert de \mathbb{R}^p , $f : \mathbb{R}^p \rightarrow \mathbb{R}$ définie sur U , et $v \in \mathbb{R}^p \setminus \{(0, \dots, 0)\}$. La fonction f est dite d'avoir une dérivée directionnelle au point $a \in U$ dans la direction de v si la limite suivante existe :

$$\lim_{h \rightarrow 0; h \in \mathbb{R}^*} \frac{1}{h} [f(a + hv) - f(a)] .$$

Des notations fréquentes sont : $f'_v(a)$, $D_v f(a)$.

Les dérivées directionnelles, comme leur cas particulier qui est celui des dérivées partielles, sont très utiles dans l'étude des fonctions et de leurs aspects géométriques. Néanmoins, elles dépendent du choix de direction. En conséquence, leur seule existence est trop faible pour entraîner la différentiabilité.

Voici une proposition qui facilite le calcul des dérivées directionnelles sous des hypothèses plus fortes :

Proposition 3.8.1 Soient $p \in \mathbb{N}^*$, $U \subset \mathbb{R}^p$ un ouvert de \mathbb{R}^p , $f : \mathbb{R}^p \rightarrow \mathbb{R}$ différentiable sur U . Alors

$$D_v f(a) = \nabla f(a) \cdot v$$

Notons que c'est le deuxième terme du membre de droite de l'expansion utilisée à la fin de la section précédente.

Chapitre 4

Différentielles du second ordre

Dans le chapitre 3, nous avons introduit une bonne notion de différentiation pour les fonctions de plusieurs variables qui ne dépend pas de choix de chemin. Remarquablement, quitte à raisonnablement renforcer les hypothèses, il était possible de trouver des caractérisations de la différentiabilité en fonction des outils pratiques mais faibles que sont les dérivées partielles : le théorème 3.4.1.

Une question naturelle qui se pose est si on ne peut pas continuer ce procédé afin de l'étendre aux "dérivées d'ordres supérieurs", et à la rigueur, introduire une notion de "dérivée partielle d'ordre supérieur". C'est ce que nous ferons dans ce chapitre. Encore une fois, les dérivées partielles seront très utiles.

Il convient de rappeler une subtilité du théorème 3.4.1 et de la notion de fonction de classe \mathcal{C}^1 . C'est une notion définie *sur un ouvert*. Sinon, nous avons vu dans le cas de la fonction $(x, y) \mapsto \sin(|xy|)$ de la fiche 4, que tout peut ne pas marcher comme on s'y attend.

4.1 Dérivées partielles secondes ; fonctions de classe \mathcal{C}^2

Définition 4.1.1 (Fonctions de classe \mathcal{C}^2 sur un ouvert) Soient $p, q \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , Une f fonction de \mathbb{R}^p vers \mathbb{R}^q définie sur U est dite de classe \mathcal{C}^2 sur U , noté $f \in \mathcal{C}^2(U)$, si la différentielle

$$\begin{aligned} df & : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \\ a & \longmapsto (df)_a \end{aligned}$$

est définie sur U et de classe \mathcal{C}^1 sur U ; en d'autres termes, si la différentielle seconde

$$d(df) : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q))$$

est continue en tout point de U .

Comme dans les remarques faites immédiatement après la définition 3.4.1, cette définition a un sens puisque df est une application de \mathbb{R}^p vers \mathbb{R}^{pq} , deux espaces normés.

Telle quelle, la définition 4.1.1 est difficile à appliquer directement dans des cas concrets. Pour surmonter cette difficulté, comme dans l'étude des fonctions de classe \mathcal{C}^1 , les dérivées partielles s'avèrent très utiles quitte à leur imposer des hypothèses de continuité. Cette fois-ci nous utiliserons les dérivées partielles secondes.

Définition 4.1.2 (Dérivées partielles secondes) Soient $p, q \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p et f une fonction de \mathbb{R}^p vers \mathbb{R}^q définie sur U dont toutes les dérivées partielles

$$\partial_j f_i : U \longrightarrow \mathbb{R} \quad (1 \leq j \leq p, 1 \leq i \leq q)$$

sont définies sur U .

Si une telle dérivée partielle admet une dérivée partielle seconde par rapport à sa k ème variable en un point $a \in U$, alors cette dérivée seconde est notée

$$\partial_k(\partial_j f_i(a)) = \partial_{kj} f_i(a).$$

La fonction f est dite d'admettre des dérivées partielles secondes sur U si toutes les dérivées partielles $\partial_{kj} f_i$ sont définies sur U .

Introduisons une autre notation très souvent utilisée pour les dérivées partielles secondes :

$$\frac{\partial^2 f_i}{\partial x_j \partial x_k}(a) = \partial_{jk} f_i(a) \quad \text{si } j \neq k ;$$

$$\frac{\partial^2 f_i}{\partial x_j^2}(a) = \partial_{jj} f_i(a) \quad \text{si } j = k ;$$

Le théorème suivant dont nous admettrons la preuve est analogue au théorème 3.4.1.

Théorème 4.1.1 Soient $p, q \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R}^q définie sur U . Alors $f \in \mathcal{C}^2(U)$ si et seulement si f a des dérivées partielles secondes en tout point de U et que celles-ci sont continues.

Certaines remarques sont nécessaires :

1. En suivant la même ligne, il est possible d'introduire sur un ouvert U de \mathbb{R}^p , les classes supérieures $\mathcal{C}^k(U)$ et les caractériser par les k èmes dérivées partielles. Dans le cas très particulier des fonctions qui appartiennent à toutes les classes $\mathcal{C}^k(U)$, nous parlerons des fonctions de classe \mathcal{C}^∞ sur U , noté $\mathcal{C}^\infty(U)$.
2. Notons que $\mathcal{C}^1(U) \supsetneq \mathcal{C}^2(U)$. Nous étudierons dans les travaux dirigés la fonction

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

qui appartient à $\mathcal{C}^1(\mathbb{R}^2) \setminus \mathcal{C}^2(\mathbb{R}^2)$. La vérification de ce phénomène est liée à la différence

$$\partial_{12} f((0, 0)) \neq \partial_{21} f((0, 0))$$

C'est en fait le sujet du *théorème de Schwarz* que nous aborderons sous peu.

4.2 Opérations sur les fonctions de classe \mathcal{C}^2

Tout se passe comme prévu comme le montre l'énoncé suivant :

Proposition 4.2.1 Soient $p, q, r \in \mathbb{N}^*$, U et V des ouverts de \mathbb{R}^p et de \mathbb{R}^q respectivement.

1. Si $f, g \in \mathcal{C}^2(U)$, alors $f + g \in \mathcal{C}^2(U)$.
2. Si $q = 1$ et $f, g \in \mathcal{C}^2(U)$, alors $f \cdot g \in \mathcal{C}^2(U)$.
3. Si $f \in \mathcal{C}^2(U)$, $g \in \mathcal{C}^2(V)$ et $f(U) \subset V$, alors $g \circ f \in \mathcal{C}^2(U)$.

4.3 Le théorème de Schwarz

Dans les remarques après le théorème 4.1.1, nous avons vu que la classe \mathcal{C}^1 est strictement plus large que la classe \mathcal{C}^2 . L'outil principal pour détecter des fonctions qui causent cette différence est le théorème suivant dont la preuve relativement compliquée sera abordée dans le cours de calcul différentiel en troisième année.

Théorème 4.3.1 (Le théorème de Schwarz) Soient $p, q \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R}^q de classe \mathcal{C}^2 sur U . Alors en tout $a \in U$ et pour tous $1 \leq i \leq q$, $1 \leq j, k \leq p$

$$\partial_{jk}f_i(a) = \partial_{kj}f_i(a) \quad .$$

Quelques remarques seront utiles pour apprécier mieux ce théorème fondamental qui nous sera indispensable dans le développement de la formule de Taylor de plusieurs variables et donc dans l'étude des extrema.

1. Il existe des fonctions deux fois différentiables mais pas de classe \mathcal{C}^2 sur un ouvert. L'exemple de la section 4.1 nous permettra d'observer cette différence aussi.
2. Quand $q = 1$, la symétrie qui découle du théorème de Schwarz entraîne la symétrie (au sens des formes bilinéaires) de la matrice suivante où U est un ouvert de \mathbb{R}^p , $f \in \mathcal{C}^2(U)$ et $a \in U$:

$$H_f(a) = \begin{pmatrix} \partial_{11}f(a) & \partial_{21}f(a) & \dots & \partial_{p1}f(a) \\ \partial_{12}f(a) & \partial_{22}f(a) & \dots & \partial_{p2}f(a) \\ & & \vdots & \\ \vdots & \dots & \partial_{jk}f(a) & \dots & \vdots \\ & & \vdots & & \\ \partial_{p1}f(a) & & \dots & & \partial_{pp}f(a) \end{pmatrix}_{p \times p} \quad .$$

C'est la *matrice Hessienne* qui jouera un rôle important dans la formule de Taylor et tout ce qui en découle.

4.4 La formule de Taylor du second ordre et la matrice Hessienne

Dans cette section, nous travaillons avec les données suivantes :

$p \in \mathbb{N}^*$, U est un ouvert de \mathbb{R}^p et f est une fonction de \mathbb{R}^p vers \mathbb{R} de classe \mathcal{C}^2 sur U .

L'hypothèse d'appartenir à la classe $\mathcal{C}^2(U)$ s'exprime pour f de deux manières différentes mais équivalentes :

1. la fonction différentielle

$$\begin{aligned} df & : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}) \\ a & \longmapsto (df)_a = (\partial_1 f(a) \dots \partial_p f(a)) \end{aligned}$$

est de classe \mathcal{C}^1 sur U ;

2. les dérivées partielles secondes de f sont toutes définies et continues sur U .

La représentation matricielle du point (1) qui permet de considérer la différentielle comme une fonction de U vers \mathbb{R}^p et l'usage des dérivées partielles permettent de trouver la matrice Hessienne comme nous l'avons constaté dans la section précédente :

$$\begin{aligned} d(df) & : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p) \\ a & \longmapsto (d(df))_a = H_f(a) = i \begin{pmatrix} & & & j \\ & & & \vdots \\ \dots & \partial_{ji}f(a) & \dots & \\ & & & \vdots \\ & & & \end{pmatrix}_{p \times p} \quad . \end{aligned}$$

Un calcul qui sera détaillé en cours et qui utilise la *symétrie de la matrice Hessienne* permet alors de vérifier les égalités suivantes :

$$\begin{aligned}
 f(a+h) &= f(a) + (df)_a h + \frac{1}{2!} h (d(df))_a h + o(\|h\|^2) \\
 &= f(a) + (df)_a h + \frac{1}{2} (h_1, \dots, h_p) H_f(a) \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix} + o(\|h\|^2) \\
 &= f(a) + \sum_{i=1}^p \partial_i f(a) h_i + \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^p \partial_{ji} f(a) h_i h_j + o(\|h\|^2)
 \end{aligned}$$

où $h = (h_1, \dots, h_p)$. Les deuxième et troisième termes à droite de chaque égalité correspondent à des produits matriciels. Le premier représente la différentielle tandis que le deuxième représente une forme bilinéaire et symétrique correspondant à la matrice Hessienne.

C'est le point de départ d'une nouvelle et fructueuse interaction entre notre cours et l'algèbre linéaire dont nous cuillerons les fruits dans le chapitre suivant.

Chapitre 5

Extrema

Ce chapitre est consacré à l'étude des valeurs extrémales d'une fonction de plusieurs variables et à valeurs réelles. Trois directions principales se distinguent : les extrema locaux, les extrema globaux et les extrema liés. Comme dans les chapitres précédents, la motivation pour l'étude qui sera menée sera les résultats sur les fonctions d'une seule variable. Pour mettre en place un développement rigoureux à partir de cette motivation initiale, les différentielles du second ordre du chapitre précédent seront indispensables.

5.1 Définitions de base ; rappels sur les fonctions d'une seule variable

Avant tout, il faut définir précisément les notions de base.

Définition 5.1.1 Soient $p \in \mathbb{N}^*$, $D \subset \mathbb{R}^p$ et f une fonction de \mathbb{R}^p vers \mathbb{R} définie sur D . La fonction f est dite d'admettre un maximum (resp. minimum) local en un point $a \in D$ s'il existe un ouvert U de \mathbb{R}^p tel que $a \in U \subset D$ et que pour tout $x \in U$, $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$).

La fonction f est dite d'admettre un maximum (resp. minimum) global en $a \in D$ si pour tout $x \in D$, $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$).

Un extremum local (resp. global) est un maximum ou minimum local (resp. global).

Remarquons immédiatement que bien que le point de départ de l'étude des extrema locaux ou globaux soient le même, les chemins suivis pour aboutir aux conclusions sont susceptibles de diverger suivant la topologie de D . L'étude locale utilise les techniques du calcul différentiel développés depuis le troisième chapitre tandis que l'étude globale peut ne pas aboutir du tout si D ne possède pas certaines propriétés topologiques assurant l'existence d'extrema globaux. Le cas le plus fréquemment rencontré qui assure cette existence est celui où D est compact grâce au corollaire 2.6.1.

Afin de motiver la discussion qui suivra dans les sections suivantes, nous rappelons le cas des fonctions d'une seule variable réelle. Etudions donc la fonction suivante assez simple qui est par ailleurs de classe \mathcal{C}^2 sur la totalité de \mathbb{R} :

$$\begin{aligned} f &: \mathbb{R} \longrightarrow \mathbb{R} \\ x &\longmapsto x(x-1)(x+1) \end{aligned}$$

En particulier, en tout point $a \in \mathbb{R}$, la formule de Taylor du second ordre est définie :

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{1}{2}f''(a)h^2 + o(h^2) \\ &= a(a-1)(a+1) + (3a^2-1)h + 3ah^2 + o(h^2). \end{aligned}$$

Les extrema locaux sont susceptibles d'être atteints aux points satisfaisant l'équation différentielle

$$f'(a) = 0.$$

Dans notre exemple, les deux candidats sont $a = \pm \frac{1}{\sqrt{3}}$.

Il n'est en général pas nécessaire que chaque point candidat (*point critique*) fournisse une valeur extrême. Dans le cas des fonctions d'une seule variable, l'étude rigoureuse consiste à étudier la *concavité* de f au voisinage des points candidats, en d'autres termes le signe de $f''(a)$. Dans notre cas, cette valeur est définie par la formule 6a qui est strictement positive quand $a = \frac{1}{\sqrt{3}}$ et strictement négative quand $a = -\frac{1}{\sqrt{3}}$. La première valeur signifie un minimum local, $f\left(\frac{1}{\sqrt{3}}\right)$, tandis que la deuxième correspond à un maximum local, $f\left(-\frac{1}{\sqrt{3}}\right)$. Ces conclusions sont naturelles. En effet, dans le premier cas, $f''\left(\frac{1}{\sqrt{3}}\right) > 0$, ce qui signifie qu'autour du point $a = \frac{1}{\sqrt{3}}$ il y a un intervalle I sur lequel $f(a+h) \geq f(a)$; dans le deuxième cas, $f''\left(-\frac{1}{\sqrt{3}}\right) < 0$, ce qui signifie l'existence d'un intervalle contenant ce point et sur lequel $f(a+h) \leq f(a)$.

Si $f''(a) = 0$, alors l'étude résumée ci-dessus est inconclusive. En effet, il suffit de comparer les fonctions $x \mapsto x^2$ et $x \mapsto x^3$ en 0. La première fonction a un minimum en 0 tandis que la deuxième fonction n'a ni maximum ni minimum en 0. Ces observations nous guideront dans le cas des fonctions de plusieurs variables aussi.

Il convient de souligner l'importance de la formule de Taylor : a condition que la fonction soit suffisamment dérivable, tout est décrit dans la formule de Taylor.

5.2 Extrema locaux

La première étape de l'étude des extrema consiste à dresser la liste des points candidats où une fonction donnée peut atteindre une valeur extrême. Soulignons que cette liste n'est qu'une liste de candidats, il est nécessaire d'y être mais pas suffisant comme nous l'avons déjà constaté dans le cas de la fonction $x \mapsto x^3$ définie sur \mathbb{R} . La proposition suivante fournit la liste des candidats pour les extrema locaux.

Proposition 5.2.1 Soient $p \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R} différentiable sur U . Si f atteint un extremum local en $a \in U$, alors $(df)_a = (0 \dots 0)$

La définition suivante découle de cette proposition.

Définition 5.2.1 Soient $p \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R} différentiable sur U . Un point $a \in U$ est dit critique si $(df)_a = 0$.

Comme conséquence du point de départ précis que nous venons d'établir, nous concluons que toute recherche d'extrema locaux commence par la détermination des solutions du système suivant d'équations différentielles :

$$\begin{cases} \partial_1 f(a) = 0 \\ \vdots \\ \partial_p f(a) = 0 \end{cases}$$

Pour pouvoir aller plus loin, nous essayerons de généraliser les méthodes du cas des fonctions d'une seule variable rappelées dans la section précédente en utilisant la formule de Taylor ainsi que la matrice hessienne et les symétries de celle-ci. Pour ce faire, nous aurons besoin de supposer que nos fonctions sont de classe \mathcal{C}^2 sur les ouverts concernés.

Soit donc f une fonction de classe \mathcal{C}^2 sur un ouvert U . En un point critique $a \in U$, nous observons la situation suivante en utilisant la formule de Taylor :

$$\begin{aligned} f(a+h) &= f(a) + (df)_a h + \frac{1}{2} h H_f(a) h + o(\|h\|^2) \\ &= f(a) + 0 + \frac{1}{2} (h_1, \dots, h_p) H_f(a) (h_1, \dots, h_p) + o(\|h\|^2). \end{aligned}$$

D'après le théorème de Schwarz (le théorème 4.3.1), $H_f(a)$ est une matrice symétrique. Si, par miracle, elle était une matrice diagonale, alors elle aurait la forme suivante :

$$H_f(a) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & \lambda_p \end{pmatrix}$$

Les λ_i étant les valeurs propres de la matrice. Dans ce cas particulier, la formule de Taylor a la forme suivante très réminiscente du cas des fonctions d'une seule variable :

$$f(a+h) = f(a) + \frac{1}{2} \sum_{i=1}^p h_i^2 \lambda_i + o(\|h\|^2).$$

Les cas suivants se présentent :

Minimum local Chaque $\lambda_i \in \mathbb{R}_+$. Le point a appartient à une boule ouverte B dans U qui a la propriété suivante : pour tout $h \in \mathbb{R}^p$ tel que $a+h \in B$, $f(a+h) \geq f(a)$. La fonction atteint un minimum local, $f(a)$, en a .

Maximum local Chaque $\lambda_i \in \mathbb{R}_-$. Le point a appartient à une boule ouverte B dans U qui a la propriété suivante : pour tout $h \in \mathbb{R}^p$ tel que $a+h \in B$, $f(a+h) \leq f(a)$. La fonction atteint un maximum local, $f(a)$, en a .

Point selle Chaque $\lambda_i \in \mathbb{R}^*$ mais il existe des valeurs propres négatives ainsi que d'autres positives. C'est un point selle.

Inconclusif Certaines des valeurs propres sont nulles, il n'est pas possible de conclure.

Dans le cas général, $H_f(a)$ n'est pas nécessairement une matrice diagonale. Néanmoins, le miracle n'est pas hors de portée. En effet, le théorème suivant fait un nouveau lien avec MathIV algèbre :

Théorème 5.2.1 *Toute matrice symétrique à entrées réelles se diagonalise avec des valeurs propres réelles.*

Alors, il suffira de faire un changement de base dans \mathbb{R}^p pour diagonaliser la matrice hessienne. On peut se demander pourquoi la nouvelle base est aussi légitime que l'ancienne. Un changement de base correspond à de nouvelles directions pour la détermination des dérivées partielles. Or, f est de classe \mathcal{C}^2 au voisinage de a . Par conséquent, les dérivées directionnelles sont définies et continues dans toutes les directions possibles.

5.3 Extrema globaux ; ensembles compacts

Un extremum local n'est pas nécessairement un extremum global. En général, il n'est pas clair si une fonction, même quand elle admet des extrema locaux, admettrait des extrema globaux. Néanmoins, dans un cas particulier mais très important, il est possible d'arriver à des conclusions satisfaisantes en suivant une recette bien établie. Nous supposons que K soit un compact de \mathbb{R}^p et que $f \in \mathcal{C}^2(\overset{\circ}{K})$. Alors, le corollaire 2.6.1 assure qu'il existe un point où f atteint son maximum global et un autre point où elle atteint son minimum global.

Pour déterminer les points où les extrema sont atteints, on dresse la liste suivante de candidats :

1. les points $a \in \overset{\circ}{K}$ où $(df)_a = (0 \dots 0)$;
2. les points $a \in \text{Fr}(K)$; plus précisément, dans ce deuxième cas, nous étudions la restriction de f à $\text{Fr}(K)$. Ceci revient à dire que nous étudierons une nouvelle fonction dont le nombre de variables libres aura diminué.

Une fois la liste déterminée, la tâche est simple : on évalue f à chacun des points candidats et retient ceux qui donnent les valeurs extrémales.

Il faut bien noter que la notion d'extremum global a un aspect relatif. Si K varie, la même loi de fonction est susceptible de fournir de nouveaux extrema atteints à de nouveaux points. Pensez à la fonction $x \mapsto x^2$ restreinte aux intervalles $[-1, 1]$, $[0, 3]$ et $[1, 4]$ par exemple. Des exemples seront étudiés en détail pendant les travaux dirigés.

5.4 Quelques outils pratiques pour les extrema locaux

Appuyée par la force de l'algèbre linéaire, la discussion de la section 5.2 est robuste. Néanmoins, la détermination des signes des valeurs propres peut s'avérer compliquée. Dans le cas particulier $p = 2$, la tâche est plus simple et une méthode particulière peut être développée.

Une matrice 2×2 symétrique a la forme générale suivante :

$$H = \begin{pmatrix} R & S \\ S & T \end{pmatrix}.$$

La détermination de ses valeurs propres équivaut à la détermination des racines du déterminant de la matrice suivante :

$$\lambda I - H = \begin{pmatrix} \lambda - R & -S \\ -S & \lambda - T \end{pmatrix}.$$

Ce déterminant est le polynôme caractéristique

$$\lambda^2 - (R + T)\lambda + RT - S^2.$$

Ce n'est pas une coïncidence, et vous devez savoir pourquoi, que le coefficient de λ est la trace de H (la somme des valeurs propres) tandis que $RT - S^2$ est son déterminant (le produit des valeurs propres).

Le théorème 5.2.1 se vérifie rapidement pour H . En effet, le discriminant Δ du polynôme caractéristique de H est $(R - T)^2 + 4S^2$, un nombre positif. Les racines sont donc réelles, et leurs valeurs sont

$$\lambda_1 = \frac{R + T + \sqrt{(R - T)^2 + 4S^2}}{2} ; \quad \lambda_2 = \frac{R + T - \sqrt{(R - T)^2 + 4S^2}}{2}.$$

Par conséquent, si $\Delta > 0$, alors les deux racines sont réelles et distinctes, ce qui implique que H est diagonalisable avec des valeurs propres réelles. Si $\Delta = 0$, alors $R = T$ et $S = 0$: une seule valeur propre, toujours réelle, et la matrice est déjà diagonale.

En utilisant aussi le fait que le déterminant d'une matrice reste invariant quand on change de base, nous constatons alors que la division en quatre cas de la section précédente prend alors une nouvelle forme :

Minimum local Chaque $\lambda_i \in \mathbb{R}_+^*$; de manière équivalente $RT - S^2 > 0$ et $R + T > 0$.

Maximum local Chaque $\lambda_i \in \mathbb{R}_-^*$; de manière équivalente $RT - S^2 > 0$ et $R + T < 0$.

Point selle Chaque $\lambda_i \in \mathbb{R}^*$ mais il existe des valeurs propres négatives ainsi que d'autres positives ; de manière équivalente $RT - S^2 < 0$.

Inconclusif Certaines des valeurs propres sont nulles ; ceci équivaut à $RT - S^2 = 0$.

Si $p > 2$, alors le théorème suivant de l'algèbre linéaire est pratique :

Théorème 5.4.1 *Soit*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{12} & a_{22} & & \\ \vdots & & \ddots & \\ a_{1p} & \dots & \dots & a_{pp} \end{pmatrix}$$

une matrice symétrique à entrées réelles.

1. Les valeurs propres de A sont toutes strictement positives (resp. négatives) s'il en est de même pour toute matrice

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{12} & a_{22} & & \\ \vdots & & \ddots & \\ a_{1k} & \cdots & \cdots & a_{kk} \end{pmatrix} \quad (1 \leq k \leq p) .$$

2. Les valeurs propres incluent 0 si une des matrices A_k est de déterminant 0.

Notons qu'aucune des connaissances de cette section n'est indispensable pour l'étude des extrema.

5.5 Extrema liés, multiplicateurs de Lagrange

Parfois, il faut étudier les extrema avec "contraintes". Voici un exemple simple mais concret : maximiser le produit de deux nombres réels dont la somme est 2. En fait, nous avons deux fonctions, ou encore une fonction produit à maximiser contrainte par une équation qui utilise la fonction somme :

$$\begin{array}{llll} f : \mathbb{R}^2 & \longrightarrow & \mathbb{R} & \phi : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & xy & (x, y) & \longmapsto & x + y - 2 . \end{array}$$

La condition est $\phi(x, y) = 0$, équivalentement, $x + y = 2$. Dans ce cas simple, nous pouvons résoudre le problème à la main : si $\phi(x, y) = 0$, alors $f(x, y) = f(x, 2 - x) = (2 - x)x$. Par conséquent l'étude de f se réduit à celle d'une seule variable $f_\phi(x) = (2 - x)x$ sur l'ensemble $\{(x, y) \in \mathbb{R}^2 \mid \phi(x, y) = 0\}$. La maximisation de cette fonction d'une seule variable, certes très bien connue par tous nos lecteurs, donne $x = 1$.

Néanmoins, un peu d'attention permet de trouver une solution générale qui peut résoudre des problèmes de même forme mais plus compliquée. En effet, au point $(1, 1)$ les gradient $\nabla f(1, 1)$ et $\nabla \phi(1, 1)$ sont parallèles. Tracer les lignes de niveau correspondantes montre que ce n'est pas une coïncidence puisque les graphes de f et de ϕ sont tangents au point $(1, 1)$. La proposition suivante montre que sous certaines hypothèses, cela est un phénomène tout à fait naturel.

Proposition 5.5.1 Soient $p \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^p , f et ϕ deux fonctions de \mathbb{R}^p vers \mathbb{R} de classes \mathcal{C}^1 sur U . Notons f_ϕ , la restriction de f aux points de l'ensemble $\{x \in U \mid \phi(x) = 0\}$. Si $a \in U$ est tel que

1. $\phi(a) = 0$;
2. la restriction f_ϕ admette un extremum en a ;
3. $\nabla \phi(a) \neq 0$;

alors il existe $\lambda \in \mathbb{R}$ tel que $(df)_a = \lambda(d\phi)_a$.

Le coefficient λ de la proposition 5.5.1 est un *multiplicateur de Lagrange*. En fait, l'existence d'un tel multiplicateur équivaut au parallélisme des gradients de f et de ϕ aux points concernés. Les "surfaces de niveau" (donc les lignes de niveau si $p = 2$), sont tangents à la surface déterminée par la contrainte ϕ . Par conséquent, dresser la liste des candidats d'extrema, revient à résoudre le système d'équations différentielles :

$$\nabla f(a) = \lambda \nabla \phi(a) .$$

Des exemples de ce phénomène de nature géométrique ainsi que leurs conséquences seront étudiés aux travaux dirigés.

Chapitre 6

Géométrie des courbes et des surfaces : le théorème des fonctions implicites

Ce chapitre commence avec un cas particulier d'un théorème fondamental : le théorème des fonctions implicites. Malgré son énoncé relativement compliqué, le TFI décrit des phénomènes que nous avons déjà rencontrés, et il forme la clé de beaucoup d'applications géométriques.

6.1 Le théorème des fonctions implicites : énoncé, un exemple simple, une application rencontrée

Commençons par la description de l'exemple simple. Notre objet d'étude est le cercle euclidien dans \mathbb{R}^2 (la sphère euclidienne)

$$S((0,0),1) = \{ (x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \} .$$

Cet ensemble est la ligne de niveau L_1 de la fonction $f : (x,y) \mapsto x^2 + y^2$, ou de manière équivalente, la ligne de niveau L_0 de la fonction $(x,y) \mapsto x^2 + y^2 - 1$. Il s'agit donc d'une "courbe" dans \mathbb{R}^2 . Alors, une question naturelle qui se pose est si $S((0,0),1)$ est le graphe d'une fonction. Il ne l'est pas puisqu'à chaque valeur entre -1 et 1 de la première ou de la deuxième coordonnée correspondent deux valeurs distinctes de l'autre sauf pour ± 1 .

Néanmoins, "localement" nous pouvons trouver des fonctions dont les graphes couvrent $S((0,0),1)$. Par exemple, nous pouvons exprimer la deuxième coordonnée en fonction de la première après avoir fixé un intervalle ouvert : $y = \phi_+(x) = \sqrt{1-x^2}$. L'ensemble d'arrivée de ϕ_+ est l'intervalle $[0,1]$, la moitié supérieure du cercle. Cette fonction est même dérivable, quitte à éviter les points $(1,0)$ et $(-1,0)$, les points où $\partial_2 f = 0$. Éviter ces deux points équivaut à fixer des intervalles ouverts pour les valeurs de la première et de la deuxième coordonnée : $] -1, 1[$ et $]0, 1[$ respectivement.

Il existe bien sûr un deuxième choix auquel s'applique les raisonnements analogues : $y = \phi_-(x) = -\sqrt{1-x^2}$. Dans ce cas, les ensembles de départ et d'arrivée de la fonction différentiable seront $] -1, 1[$ et $] -1, 0[$ respectivement. L'ensemble d'arrivée correspond à la moitié négative du cercle.

Remarquablement, ce que nous venons de faire peut se faire pour trouver des fonctions qui expriment la première coordonnée en fonction de la deuxième :

$$x = \psi_+(y) = \sqrt{1-y^2} \quad ; \quad x = \psi_-(y) = -\sqrt{1-y^2} .$$

Les deux fonctions ψ_+ et ψ_- deviennent différentiables si nous évitons les valeurs qui annulent $\partial_1 f$.

En conclusion, nous avons explicité des liens a priori implicites entre les deux variables. Ces liens existaient parce que nous avons considéré contraint les valeurs de la fonction f . Le théorème des fonctions implicites énonce que, quitte à satisfaire certaines hypothèses, nous pouvons conclure l'existence d'un tel lien même quand c'est pratiquement impossible à expliciter.

Théorème 6.1.1 (Le théorème des fonctions implicites) Soient $p \in \mathbb{N}^*$, U un ouvert \mathbb{R}^{p+1} , f une fonction de \mathbb{R}^{p+1} vers \mathbb{R} de classe \mathcal{C}^1 sur U , $S = \{x \in U \mid f(x) = 0\}$ et $a = (a_1, \dots, a_{p+1}) \in S$. Nous fixons une coordonnée $i \in \{1, \dots, p+1\}$. Si $\partial_i f(a) \neq 0$, alors il existe un ouvert V de \mathbb{R}^p contenant $(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_{p+1})$ et une fonction $g \in \mathcal{C}^1(V)$ telle que pour tout $(x_1, \dots, x_{p+1}) \in U$

$$f(x_1, \dots, x_{p+1}) = 0 \text{ si et seulement si } x_i = g(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{p+1}).$$

En plus, pour tout $j \in \{1, \dots, p+1\} \setminus \{i\}$,

$$\partial_j g(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{p+1}) = -\frac{\partial_j f(x_1, \dots, x_{p+1})}{\partial_i f(x_1, \dots, x_{p+1})}.$$

Dans ce qui précède, on peut remplacer \mathcal{C}^1 par \mathcal{C}^k .

Illustrons immédiatement ce théorème dans l'exemple que nous avons étudié au début de cette section. La fonction f est $(x, y) \mapsto x^2 + y^2 - 1$ tandis que S est le cercle $S((0, 0), 1)$. En tout point sur S , sauf en $(1, 0)$, $(-1, 0)$, $(0, 1)$ et $(0, -1)$, les deux dérivées partielles sont non nulles et une coordonnée s'exprime en fonction de l'autre continûment différentiablement. En suivant la notation du théorème 6.1.1, si $i = 1$, alors g correspond à ϕ_+ ou ϕ_- . Effectivement, il y a un choix, le TFI ne nous dit pas que g est unique. De façon similaire, si $i = 2$, alors g est ψ_+ ou ψ_- . Quant aux ensembles U et V , on peut respectivement poser $U = \mathbb{R}^2$ et $V =]-1, 1[$, ces choix n'étant pas les seuls.

Nous avons déjà utilisé le TFI de façon plus fondamentale que notre exemple simple à propos du cercle. La proposition 3.7.1 qui énonce la perpendicularité du vecteur gradient à la droite tangente à une "courbe" dans \mathbb{R}^2 ne serait pas possible sans savoir l'existence d'un lien entre les coordonnées. En effet, les hypothèses de la proposition 3.7.1 permettent de conclure que nous pouvons trouver localement une fonction ϕ de classe \mathcal{C}^1 telle que $x_2 = \phi(x_1)$ sur la région décrite par $f(x_1, x_2) = 0$. Alors,

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\phi} & \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R} \\ x_1 & \mapsto & (x_1, \phi(x_1)) & \mapsto & f(x_1, \phi(x_1)) \end{array}.$$

Comme nous sommes sur une région où f est constante (la courbe $f(x_1, x_2) = 0$),

$$(\partial_1 f(x_1, \phi(x_1)) \quad \partial_2 f(x_1, \phi(x_1))) \begin{pmatrix} 1 \\ \phi'(x_1) \end{pmatrix} = 0.$$

Tout ce que cette dernière équation dit est ce que nous savons déjà : $\nabla f(a)$ est perpendiculaire à la droite tangente de la courbe $f(x_1, x_2) = 0$ au point $a = (x_1, \phi(x_1))$.

6.2 Courbes dans \mathbb{R}^2

Cette section est une introduction plus rigoureuse à la géométrie des courbes. Le TFI sera un outil principal.

Définition 6.2.1 (Courbes dans \mathbb{R}^2) Une partie C de \mathbb{R}^2 est dite une courbe si elle est décrite sous l'une des trois formes suivantes :

1. si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction de deux variables et $k \in \mathbb{R}$, alors C est la ligne de niveau

$$L_k(f) = \{ (x_1, x_2) \in \mathbb{R}^2 \mid f(x_1, x_2) = k \};$$

2. si $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction d'une seule variable (resp. $\psi : \mathbb{R} \longrightarrow \mathbb{R}$), alors C est le graphe

$$\mathcal{G}(\phi) = \{ (x_1, x_2) \in \mathbb{R}^2 \mid x_2 = \phi(x_1) \}$$

(resp. $\mathcal{G}(\psi) = \{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 = \psi(x_2) \}$);

3. si $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^2$ est une fonction de la forme $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ (“arc généralisé”), alors C est l'image de γ :

$$\{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 = \gamma_1(t), x_2 = \gamma_2(t), t \in \mathbb{R} \}.$$

Cette dernière définition, la définition paramétrique, est celle qui se généralise à tout \mathbb{R}^p : $\gamma(t) = (\gamma_1(t), \dots, \gamma_p(t))$.

Une courbe est donc un objet géométrique de “dimension 1” : sa définition dépend d'une seule variable libre, x_1 ou x_2 dans (1), x_1 pour ϕ et x_2 pour ψ dans (2), et t dans (3) dans la définition 6.2.1.

Voici deux exemples simples mais fondamentaux :

1. Une droite dans \mathbb{R}^2 est une courbe. Suivant la division de cas dans la définition 6.2.1 elle est définie sous l'une des formes suivantes que vous connaissez déjà très bien :

(Ligne de niveau) Si f est la fonction $(x, y) \mapsto ax + by$ à condition que $(a, b) \neq (0, 0)$, alors la “droite $ax + by = c$ ” est la ligne de niveau $L_c(f)$. On peut bien sûr remplacer f par $g : (x, y) \mapsto ax + by - c$ dans lequel cas la même droite est la ligne de niveau $L_0(g)$.

(Graphe) Si ϕ est la fonction $x \mapsto mx + n$ à condition que $(m, n) \neq (0, 0)$, alors la “droite $y = mx + n$ ” est le graphe $G(\phi)$.

(Version paramétrique) Une droite dans \mathbb{R}^p est la donnée d'un point $a = (a_1, \dots, a_p)$ et d'un vecteur directeur non nul $v = (v_1, \dots, v_p)$

$$\{ a + tv \mid t \in \mathbb{R} \} = \begin{cases} x_1 & = a_1 + tv_1 \\ & \vdots \\ x_p & = a_p + tv_p \end{cases}$$

2. Les coordonnées polaires nous permettent de définir des courbes paramétrées. Voici un exemple connu :

$$S((0, 0), 1) = \{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 = \sin(t), x_2 = \cos(t), t \in [0, 2\pi[\}.$$

La discussion sur l'application du TFI à la perpendicularité du vecteur gradient à une courbe, nous permet de trouver l'équation de la droite tangente à une courbe en un point donné, quand la courbe est donnée comme une ligne de niveau ou un graphe. Si la courbe est décrite par l'équation $f(x_1, x_2) = k$ avec f de classe \mathcal{C}^1 au voisinage d'un point $a = (a_1, a_2) \in \mathbb{R}^2$ tel que $f(a_1, a_2) = 0$ et que $\partial_2 f(a) \neq 0$, alors l'équation de la droite tangente à la courbe au point a est

$$x_2 - a_2 = -\frac{\partial_1 f(a)}{\partial_2 f(a)} (x_1 - a_1),$$

ce qui équivaut à

$$\partial_2 f(a)(x_2 - a_2) + \partial_1 f(a)(x_1 - a_1) = 0,$$

soit encore

$$\nabla f(a) \cdot (x_1 - a_1, x_2 - a_2) = 0.$$

Aux travaux dirigés, nous développerons cette approche et nous étudierons l'équation de la droite tangente quand une courbe est donnée paramétriquement.

6.3 Surfaces dans \mathbb{R}^3

Dans cette section, nous poursuivons l'approche rigoureuse à la géométrie dans le contexte des surfaces dans \mathbb{R}^3 .

Définition 6.3.1 (Surfaces dans \mathbb{R}^3) Une partie S de \mathbb{R}^3 est dite une surface si elle est décrite sous l'une des formes suivantes :

1. si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction de deux variables, alors $S = \mathcal{G}(f)$;
2. si $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ est une fonction de trois variables et $k \in \mathbb{R}$, alors S est la "surface de niveau"

$$\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid f(x_1, x_2, x_3) = k \} .$$

Une surface est un objet géométrique de "dimension 2" : sa définition fait intervenir exactement deux variables libres. La forme explicite de (1) et la forme implicite de (2) de la définition 6.3.1 sont liées par le TFI.

Comme dans le cas des courbes, donnons des exemples simples mais importants :

1. Un plan dans \mathbb{R}^3 est une surface. Suivant la division de cas de la définition 6.3.1, elle est définie sous l'une des formes suivantes :

(Graphe) Si f est la fonction $(x, y) \mapsto ax + by$ à condition $(a, b) \neq (0, 0)$, alors le "plan $ax + by = z$ " est le graphe $\mathcal{G}(f)$.

(Surface de niveau) Si f est la fonction $(x, y, z) \mapsto ax + by + cz$ et $(a, b, c) \neq (0, 0, 0)$, alors le "plan $ax + by + cz = d$ " est la "surface de niveau"

$$\{ (x, y, z) \in \mathbb{R}^3 \mid f(x, y, z) = d \} .$$

On peut bien sûr remplacer f par $g(x, y, z) = f(x, y, z) - d$ dans lequel cas le niveau est 0.

(Vecteur perpendiculaire) Un plan est la donnée d'un point $a = (a_1, a_2, a_3) \in \mathbb{R}^3$ et d'un vecteur $u = (u_1, u_2, u_3) \in \mathbb{R}^3 \setminus \{(0, 0, 0)\}$ perpendiculaire au plan :

$$\begin{aligned} S &= \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid (u_1, u_2, u_3) \cdot (x_1 - a_1, x_2 - a_2, x_3 - a_3) = 0 \} \\ &= \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid u_1(x_1 - a_1) + u_2(x_2 - a_2) + u_3(x_3 - a_3) = 0 \} \\ &= \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid u_1x_1 + u_2x_2 + u_3x_3 = u_1a_1 + u_2a_2 + u_3a_3 \} \end{aligned}$$

2. La sphère euclidienne $S((0, 0, 0), 1)$ est la "surface de niveau 1" pour la fonction

$$f : \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto x^2 + y^2 + z^2 .$$

3. Si

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto x^2 + y^2 ,$$

alors le graphe

$$\begin{aligned} \mathcal{G}(f) &= \{ (x, y, z) \in \mathbb{R}^3 \mid z = f(x, y) \} \\ &= \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z = 0 \} \end{aligned}$$

est une surface. Notons que la deuxième description transforme une représentation par un graphe en une représentation par une "surface de niveau".

Le rôle similaire à celui d'une droite tangente à une courbe est joué par le plan tangent à une surface en un point donné. Ceci est déjà visible dans la troisième forme que nous avons donnée de la représentation d'un plan. Plus généralement si S est une surface déterminée par l'équation

$f(x_1, x_2, x_3) = 0$ où f est une fonction de classe \mathcal{C}^1 si $a = (a_1, a_2, a_3) \in S$, alors l'équation du plan tangent à S au point a est

$$\nabla f(a) \cdot (x_1 - a_1, x_2 - a_2, x_3 - a_3) = 0 ,$$

ou de façon plus détaillée

$$\partial_1 f(a)(x_1 - a_1) + \partial_2 f(a)(x_2 - a_2) + \partial_3 f(a)(x_3 - a_3) = 0 .$$

Nous détaillerons la vérification de cette conclusion et étudierons certains exemples aux travaux dirigés.

Chapitre 7

Intégration ; intégrales multiples

Voici quelques questions et réponses dans lesquelles l'usage du mot "volume" peut paraître un peu déroutant dans un premier temps.

0. Quel est le volume d'un point dans \mathbb{R}^p ?

Pauvre singleton, il n'est que de dimension 0. Son volume est donc négligeable, disons 0. Il en est de même pour toute partie finie de \mathbb{R}^p .

1. Quel est le volume d'un segment compact dans \mathbb{R} ?

La tendance serait de parler de la "longueur" du segment $[a, b]$ ($a, b \in \mathbb{R}, a \leq b$), mais il est clair que l'objectif est de *mesurer* la "quantité de matière" contenue dans un certain endroit de \mathbb{R} . Dans ce cas particulier, cette quantité est effectivement $b - a$.

2. Quel est le volume d'un rectangle

$$[a_1, b_1] \times [a_2, b_2] = \{ (x_1, x_2) \in \mathbb{R}^2 \mid a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2 \}$$

dans \mathbb{R}^2 ?

Nous semblons parler de l'aire plutôt que du volume, mais encore une fois, il s'agit de mesurer la quantité de matière, contenue cette fois-ci dans un certain endroit de \mathbb{R}^2 . Et cette quantité est $(b_1 - a_1) \times (b_2 - a_2)$.

3. Quel est le volume d'un parallélépipède rectangle

$$[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] = \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, a_3 \leq x_3 \leq b_3 \}$$

dans \mathbb{R}^3 ?

Ça, on le connaît : $(b_1 - a_1) \times (b_2 - a_2) \times (b_3 - a_3)$.

p. Quel est le volume d'un "rectangle généralisé"

$$[a_1, b_1] \times \dots \times [a_p, b_p] = \{ (x_1, \dots, x_p) \in \mathbb{R}^p \mid a_i \leq x_i \leq b_i \ i \in \{1, \dots, p\} \}$$

dans \mathbb{R}^p ?

Les rectangles généralisés ont pour volume $\prod_{i=1}^p (b_i - a_i)$.

Ce que nous avons appelé un "volume" est donc une longueur dans \mathbb{R} , une aire \mathbb{R}^2 , un volume dans \mathbb{R}^3 , ... une mesure de quantité de matière contenue dans une région de \mathbb{R}^p . L'intégration est l'activité de mesurer cette quantité sur des régions qui ont suffisamment de propriétés communes avec les régions rectangulaires dont nous savons déterminer les volumes.

7.1 Pavés dans \mathbb{R}^p ; parties pavables de \mathbb{R}^p

Définition 7.1.1 (Pavés dans \mathbb{R}^p ; parties pavables de \mathbb{R}^p)

1. Un pavé de \mathbb{R}^p est une partie de \mathbb{R}^p de la forme

$$\prod_{i=1}^p [a_i, b_i] = \{ (x_1, \dots, x_p) \in \mathbb{R}^p \mid a_i \leq x_i \leq b_i \ i \in \{1, \dots, p\} \} .$$

2. Une partie de \mathbb{R}^p est dite pavable si elle est union finie de pavés.

Les remarques suivantes sur ces notions clés sont simples, intuitives mais importantes pour mieux comprendre la discussion en cours :

1. Toute partie pavable de \mathbb{R}^p est fermée et bornée, donc compacte.
2. Une partie pavable P de \mathbb{R}^p a un volume, ou une “mesure”, que l’on peut déterminer facilement. Si P est un pavé $\prod_{i=1}^p [a_i, b_i]$, alors sa mesure est

$$\mu(P) = \prod_{i=1}^p (b_i - a_i) .$$

Si $P = \bigcup_{i=1}^m P_i$ et que $\overset{\circ}{P}_i \cap \overset{\circ}{P}_j = \emptyset$ si et seulement si $i \neq j$, alors

$$\mu(P) = \sum_{i=1}^m \mu(P_i) .$$

Plus généralement, il faut enlever de cette somme la mesure de l’adhérence des intersections des intérieurs.

Intuitivement, toute partie de \mathbb{R}^p a une mesure “calculable”, ou plus facilement calculable, si et seulement si elle est “approximable” par des pavés. La définition suivante clarifie ce procédé d’approximation de manière rigoureuse :

Définition 7.1.2 (Parties quarrables de \mathbb{R}^p) Soit A une partie bornée de \mathbb{R}^p . Alors A est dite quarrable si pour tout $\epsilon \in \mathbb{R}_+^*$, il existe deux parties pavables de \mathbb{R}^p , R et R' , telles que $R \subset A \subset R'$ et que $\mu(R' \setminus R) < \epsilon$.

Afin d’illustrer comment la définition 7.1.2 marche en pratique, nous étudierons le cas d’une partie bornée de \mathbb{R}^2 . Nous ne précisons pas la forme géométrique de A puisque la seule condition est celle que A soit bornée. Néanmoins, tout lecteur qui préfère un exemple concret peut prendre A comme la boule euclidienne (fermée ou ouverte) de rayon 1.

Comme A est borné, il existe $a_1, a_2, b_1, b_2 \in \mathbb{R}$ tels que $a_1 \leq b_1$, $a_2 \leq b_2$ et que $A \subset [a_1, b_1] \times [a_2, b_2]$. En d’autres termes, A est contenu dans un pavé. Si $a_1 = b_1$ ou $a_2 = b_2$, alors il n’y a rien à faire, le volume de A sera 0. Sinon, pour chaque choix de $m, n \in \mathbb{N}$, on effectue des subdivisions des intervalles $[a_1, b_1]$ et $[a_2, b_2]$:

$$x_0 = a_1 < x_1 \dots < x_m = b_1$$

et

$$y_0 = a_2 < y_1 \dots < y_n = b_2 .$$

Il s’agit donc des segments sur les intervalles $[a_1, b_1]$ et $[a_2, b_2]$. Ensuite, sont définies deux sommes :

$$s(\sigma) = \sum_{[x_i, x_{i+1}] \times [y_j, y_{j+1}] \subset A} (y_{j+1} - y_j)(x_{i+1} - x_i)$$

et

$$S(\sigma) = \sum_{[x_i, x_{i+1}] \times [y_j, y_{j+1}] \cap A \neq \emptyset} (y_{j+1} - y_j)(x_{i+1} - x_i) .$$

Comme l'union des pavés qui contribuent à $s(\sigma)$ est contenue dans celle des pavés qui contribuent à $S(\sigma)$, $s(\sigma) \leq S(\sigma)$. Notons que ces sommes sont toujours des nombres positifs.

Il faut bien constater que la construction ci-dessus peut se faire avec n'importe quel pavé contenant l'ensemble A et avec n'importe quel subdivision d'un choix particulier de pavé. Par ailleurs, des choix de subdivisions plus fines augmentent la valeur de $s(\sigma)$ et diminuent celle de $S(\sigma)$. L'ensemble A est quarrable précisément quand ces augmentations et diminutions convergent vers une même mesure :

Proposition 7.1.1 (Caractérisation d'une partie quarrable de \mathbb{R}^p) *Soit A une partie bornée de \mathbb{R}^p . Alors les conditions suivantes sont équivalentes :*

1. A est quarrable ;
2. $\sup_{\sigma}(s(\sigma)) = \inf_{\sigma}(S(\sigma))$;
3. $Fr(A)$ est de mesure 0.

A ce stade, une question naturelle se pose : comment déterminer si une partie bornée de \mathbb{R}^p est quarrable ? A priori, c'est difficile à déterminer. Intuitivement, toute partie dont nous savons calculer le volume est quarrable, et il en est de même de sa frontière. Voici quelques exemples :

- les pavés ;
- les frontières des pavés (de mesure 0) ;
- un disque dans \mathbb{R}^2 de rayon r : la mesure est πr^2 ;
- la frontière d'un disque dans \mathbb{R}^2 : la mesure est 0.

Evidemment, ces réponses sont loin d'être satisfaisantes. Non seulement, quoique intuitivement claires, elles ne sont pas justifiées, mais elles n'offrent pas de méthode générale. L'intégration fournira la méthode

7.2 Fonctions intégrables

Dans cette section, nous introduirons la notion de *fonction intégrable*. Cette notion n'est pas moins facile à remanier que celle d'un ensemble quarrable. Néanmoins, le théorème de Fubini qui sera introduit à travers des exemples changera le paysage pratique complètement et fera le lien entre l'approche utilisant les sommes, les volumes, les mesures, et l'approche analytique des techniques d'intégration des fonctions d'une seule variable.

Nous commençons en fixant une partie quarrable A de \mathbb{R}^p . Soit f une fonction de \mathbb{R}^p vers \mathbb{R} , définie et *bornée* sur A . Cette deuxième hypothèse équivaut à dire qu'il existe $M \in \mathbb{R}_+$ tel que pour tout $x \in A$, $|f(x)| \leq M$. Un exemple fréquent est le cas d'une fonction continue sur un ensemble compact.

Comme A est quarrable, nous pouvons définir des subdivisions comme dans la section précédente. Comme A est en particulier borné, il existe $a_1, \dots, a_p, b_1, \dots, b_p \in \mathbb{R}$ tels que $a_1 \leq b_1, a_2 \leq b_2, \dots, a_p \leq b_p$ et que

$$A \subset \prod_{i=1}^p [a_i, b_i] .$$

Pour chaque coordonnée, les subdivisions auront la forme suivante :

$$x_{i,0} = a_i \leq x_{i,1} \leq \dots \leq x_{i,k_i} = b_i \quad (1 \leq i \leq p, k_i \in \mathbb{N}) .$$

Notons que les seules contributions seront apportées par les subdivisions où chaque intervalle est de longueur non nulle. Alors, un pavé est de la forme suivante :

$$R_{(j_1, \dots, j_p)} = \prod_{i=1}^p [x_{i,j_i}, x_{i,j_i+1}] \quad (\text{pour tout } i, 0 \leq j_i \leq k_i) .$$

Cette fois-ci, les subdivisions serviront à calculer des sommes pondérées par la fonction f . En utilisant cette notation, nous définissons

$$s_f(\sigma) = \sum_{R_{(j_1, \dots, j_p)} \cap A \neq \emptyset} \inf_{R_{(j_1, \dots, j_p)} \cap A} f(x) \prod_{i=1}^p (x_{i, j_{i+1}} - x_{i, j_i})$$

$$S_f(\sigma) = \sum_{R_{(j_1, \dots, j_p)} \cap A \neq \emptyset} \sup_{R_{(j_1, \dots, j_p)} \cap A} f(x) \prod_{i=1}^p (x_{i, j_{i+1}} - x_{i, j_i}) .$$

Ces deux sommes sont réminiscentes des sommes de Riemann. Ce n'est pas une coïncidence, en effet la somme suivante est une somme de Riemann généralisée :

$$R_f^{(\sigma)} = \sum_{R_{(j_1, \dots, j_p)} \cap A \neq \emptyset} \inf_{R_{(j_1, \dots, j_p)} \cap A} f(a_{(j_1, \dots, j_p)}) \prod_{i=1}^p (x_{i, j_{i+1}} - x_{i, j_i}) ,$$

le point $a_{(j_1, \dots, j_p)}$ étant arbitrairement choisi dans $R_{(j_1, \dots, j_p)} \cap A$.

Après toute cette préparation, nous pouvons définir l'intégrabilité d'une fonction.

Définition 7.2.1 (Fonctions intégrables) *La fonction f est dite intégrable sur A si et seulement si l'une des deux conditions équivalentes suivantes est vraie :*

1. $\sup_{\sigma} s_f(\sigma) = \inf_{\sigma} S_f(\sigma)$;
2. $\lim_{\mu(R_{(j_1, \dots, j_p)}) \rightarrow 0} R_f^{(\sigma)}$ existe.

Remarquons immédiatement que si les conditions de la définition 7.2.1 sont satisfaites, alors

$$\sup_{\sigma} (s_f) = \inf_{\sigma} S_f = \lim_{\sigma} R_f .$$

Cette valeur, l'intégrale de f sur A est notée

$$\int \dots \int_A f(x_1, \dots, x_p) dx_1 \dots dx_p .$$

Dans le cas particulier d'une fonction f dont la restriction à A est la fonction constante de valeur 1, l'intégrale est exactement $\mu(A)$.

ATTENTION!!! Avertissons nos lecteurs. A ce stade du développement de l'intégration des fonctions de plusieurs variables, l'écriture

$$\int \dots \int_A f(x_1, \dots, x_p) dx_1 \dots dx_p .$$

n'est effectivement qu'un choix particulier de notation. Il ne faut pas penser qu'elle veut dire qu'on a droit à faire "intégration partielle", c'est à dire, intégrer par rapport à une variable, puis à une autre, ainsi de suite. Ceci sera possible suite au théorème de Fubini. Avant d'aborder ce théorème indispensable dans le calcul d'une intégrale sur un ensemble quarrable, nous précisons des conditions suffisantes afin de décider de l'intégrabilité d'une fonction.

Théorème 7.2.1 *Soient A une partie quarrable de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R} définie et bornée sur A . Si l'ensemble de points de discontinuité de f sur A est de mesure 0, alors f est intégrable sur A .*

Vous étudierez la preuve de ce théorème fondamental en L3. Le corollaire suivant est très utile :

Corollaire 7.2.1 *Soient A une partie quarrable de \mathbb{R}^p , f une fonction de \mathbb{R}^p vers \mathbb{R} définie et bornée sur A . Si f a un nombre fini de discontinuités sur A , alors f est intégrable sur A . En particulier, une fonction continue sur A est intégrable sur A .*

Ce corollaire nous permet de décider dans les cas les plus fréquents dans ce cours si une certaine partie bornée de \mathbb{R}^p est quarrable. Il suffit que sa frontière soit décrite par des fonctions continues.

7.3 Propriétés des fonctions intégrables

Toute notion de fonction intégrable digne de cette appellation vérifie certaines conditions. La proposition suivante en donne l'aperçu nécessaire pour notre cours.

Notation efficace pour les discussions théoriques : Afin d'éviter l'usage pour le moment inutile d'une multitude de symboles d'intégration, nous écrivons

$$\int_A f(x) dx$$

au lieu de

$$\int \dots \int_A f(x_1, \dots, x_p) dx_1 \dots dx_p .$$

pour noter l'intégrale d'une fonction f de p variables sur une partie quarrable A de \mathbb{R}^p .

Proposition 7.3.1 Soient A une partie quarrable de \mathbb{R}^p et f, g deux fonctions de \mathbb{R}^p vers \mathbb{R} définies et intégrables sur A .

Linéarité Si $\lambda_f, \lambda_g \in \mathbb{R}$ sont deux scalaires, alors

$$\int_A (\lambda_f f + \lambda_g g)(x) dx = \lambda_f \int_A f(x) dx + \lambda_g \int_A g(x) dx .$$

Croissance Si pour tout $a \in A$, $f(a) \leq g(a)$, alors

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

Additivité Si $A = A_1 \cup \dots \cup A_m$ tel que $\overset{\circ}{A}_i \cap \overset{\circ}{A}_j = \emptyset$ si et seulement si $i \neq j$, alors

$$\int_A f(x) dx = \sum_{i=1}^m \int_{A_i} f(x) dx$$

Preuve. Vous pouvez démontrer ces résultats vous-même en appliquant directement les sommes généralisées de Riemann introduites dans la section précédente pour définir l'intégrabilité d'une fonction. \square

7.4 Théorème de Fubini

Le théorème de Fubini est réminiscent de la proposition 3.4.1 qui énoncent que sous certaines conditions la détermination de la différentielle se réduit à celle des dérivées partielles. Ci-dessous nous en donnons un énoncé, son étude approfondie sera développée sur des exemples détaillés en cours et aux travaux dirigés.

Théorème 7.4.1 (Théorème de Fubini) Soient A une partie quarrable de \mathbb{R}^p et f une fonction de \mathbb{R}^p vers \mathbb{R} intégrable sur A . Alors,

$$\int_A f(x) dx = \int_a^b \left(\dots \left(\int_{x_{i_2}=\phi_2(x_1, \dots, \hat{x}_{i_1}, \dots, x_{i_2}, \dots, x_p)} \left(\int_{x_{i_1}=\phi_1(x_1, \dots, \hat{x}_{i_1}, \dots, x_p)} f(x_1, \dots, x_p) dx_{i_1} \right) dx_{i_2} \right) \dots \right) dx_{i_p}$$

où $a, b \in \mathbb{R}$ et les $\hat{}$ symbolisent les variables omises.

7.5 Changement de variables et intégration

Une question d'intégration comme la suivante est conceptuellement claire et simple, donc facile à aborder, sans pour autant être aussi simple au niveau calculatoire qui est indispensable pour aboutir à une solution finale :

Calculer le volume de la boule unité fermée par rapport à la norme euclidienne dans \mathbb{R}^3 .

Il s'agit du calcul du volume de la région

$$A = \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq 1 \},$$

le théorème de Fubini permet d'écrire

$$\int_A 1 \, dx = \int_{-1}^1 \left(\int_{-\sqrt{1-x_3^2}}^{\sqrt{1-x_3^2}} \left(\int_{-\sqrt{1-x_3^2-x_2^2}}^{\sqrt{1-x_3^2-x_2^2}} dx_1 \right) dx_2 \right) dx_3.$$

La complication est causée par les bornes de l'intégration, donc par la frontière de A , qui sont plutôt sphériques et qui s'expriment avec des équations faisant intervenir des polynômes du second degré en coordonnées cartésiennes. Or, dans les coordonnées sphériques, A s'exprime de façon très simple en fonction des trois variables (ρ, ϕ, θ) : pour un point P , ces variables expriment respectivement, la distance euclidienne à $(0, 0, 0)$, l'angle entre l'axe des x_3 et le segment liant $(0, 0, 0)$ à P , l'angle entre l'axe des x_1 et la projection dans le plan x_1x_2 du segment liant $(0, 0, 0)$ à P . Dans ce système de coordonnées A est plutôt un parallélépipède rectangle dont les bornes fournissent une intégrale simple à déterminer

$$A = \{ (\rho, \phi, \theta) \mid 0 \leq \rho \leq 1, 0 \leq \phi \leq \pi, 0 \leq \theta < 2\pi \}.$$

Le théorème suivant nous permet de profiter de cette simplification

Théorème 7.5.1 Soient $D \subset \mathbb{R}^p$ un ensemble quarrable, $f : D \rightarrow \mathbb{R}$ une fonction intégrable sur D , U et V deux ouverts dans \mathbb{R}^p . On suppose $D \subset V$ et qu'il existe un difféomorphisme de $\Phi : U \rightarrow V$: la notation sera $\Phi(u) = \Phi(u_1, \dots, u_p) = (v_1, \dots, v_p) = \Phi(v)$ avec $v_i = \Phi_i(u_1, \dots, u_p)$ pour tout $i \in \{1, \dots, p\}$. Alors,

$$\int_D f(v) \, dv = \int_{\Phi^{-1}(D)} (f \circ \Phi)(u_1, \dots, u_p) |(d\Phi)_{(u_1, \dots, u_p)}| \, du.$$

$|(d\Phi)_{(u_1, \dots, u_p)}|$ est la valeur absolue du déterminant du jacobien.

Une fonction bijective $\Phi : U \rightarrow V$ entre deux ouverts respectivement de \mathbb{R}^p et \mathbb{R}^p est dite un *difféomorphisme* si elle est de classe \mathcal{C}^1 et son inverse aussi est de classe \mathcal{C}^1 . En utilisant vos connaissances sur le calcul différentiel, vous pouvez vérifier que ceci équivaut à ce que le Jacobien de Φ se représente par des matrices inversibles aux points de U .

Illustrons ce changement dans l'exemple que nous venons de donner. Nous pouvons supposer $D = V = \overset{\circ}{A}$, $U = \mathbb{R}_+^* \times]0, \pi[\times]0, 2\pi[$ et

$$\begin{aligned} \Phi : \mathbb{R}_+^* \times]0, \pi[\times]0, 2\pi[&\longrightarrow \overset{\circ}{A} \\ (\rho, \phi, \theta) &\longmapsto (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi). \end{aligned}$$

La condition sur U et V d'être ouverts nous fait perdre la frontière de la boule unité fermée. Néanmoins, cette perte ne cause aucune variation dans la valeur de l'intégrale. Nous savons que les frontières d'un ensemble quarrable sont de mesure négligeable.

Dans un premier temps, la formule de changement de variables peut paraître difficile à saisir, même à motiver. Il est très utile de penser au cas particulier où $p = 1$ que vous connaissez déjà à une valeur absolue près :

$$\int_a^b f(x) \, dx = \int_{t_1}^{t_2} (f \circ \gamma)(t) \gamma'(t) \, dt$$

si $x = \gamma(t)$ et $\gamma(t) \in \mathcal{C}^1(I)$ avec I un intervalle ouvert contenant $[t_1, t_2]$.

Chapitre 8

Différentiation, intégration : quelques liens

Dans ce chapitre nous étudierons certains liens entre la différentiation et l'intégration en plusieurs variables. Cette étude généralise les liens entre les primitives et les dérivées dans le cas des fonctions d'une seule variable. Elle se lie aux idées géométriques du chapitre 6.

8.1 Longueur d'arc, intégration le long d'une courbe

Nous nous posons la question suivante : nous avons introduit une notion d'intégration bien établie qui marche, sous des conditions raisonnables, dans tout \mathbb{R}^p , nous pouvons par conséquent intégrer des fonctions sur diverses régions dans \mathbb{R}^p et calculer des volumes ; ne pouvons-nous pas mesurer la longueur des arcs qui ne sont pas nécessairement des segments de droite dans \mathbb{R}^p ? Cette question qui peut être généralisée est pertinente en mathématique qu'en ses applications. Nous l'étudierons dans le cas des courbes dans \mathbb{R}^p en utilisant et développant certaines données du chapitre 6.

Pour nos objectifs, le cadre le mieux établi est fourni par la *version paramétrique* de la définition d'un courbe :

si $\gamma : \mathbb{R} \rightarrow \mathbb{R}^p$ est une fonction de la forme $\gamma(t) = (\gamma_1(t), \dots, \gamma_p(t))$ ("arc généralisé"), alors la courbe C est l'image de γ :

$$C = \{ (x_1, \dots, x_p) \in \mathbb{R}^p \mid x_i = \gamma_i(t), i \in \{1, \dots, p\}, t \in \mathbb{R} \} .$$

Pour ne pas compliquer la notation trop, nous préférons ne mentionner que γ comme courbe, en n'oubliant pas qu'il s'agit de son image.

Voici une version rallongée d'une courbe dans \mathbb{R}^3 que vous connaissez bien :

$$\begin{array}{lcl} \gamma & : & \mathbb{R} \longrightarrow \mathbb{R}^3 \\ & & t \longmapsto (\cos t, \sin t, t) \end{array}$$

Maintenant, retournons à la discussion générale. Nous introduirons une méthode pour mesurer la longueur d'une courbe γ dans \mathbb{R}^p qui sera supposée une fonction différentiable. Evoquons d'abord certains aspects pertinents du chapitre 6. En tout point $a = (a_1, \dots, a_p) \in \mathbb{R}^p$ qui se trouve sur γ , en d'autres termes, qui satisfait pour au moins une valeur t_a de t

$$\left\{ \begin{array}{l} a_1 = \gamma_1(t_a) \\ \vdots \\ a_p = \gamma_p(t_a) \end{array} \right.$$

nous pouvons associer une droite tangente en utilisant la différentielle :

$$\{ u \in \mathbb{R}^p \mid u = a + (\gamma'_1(t_a)t, \dots, \gamma'_p(t_a)t), t \in \mathbb{R} \},$$

soit encore

$$\begin{cases} u_1 &= a_1 + \gamma'_1(t_a)t \\ &\vdots \\ u_p &= a_p + \gamma'_p(t_a)t \end{cases} \quad (t \in \mathbb{R})$$

Evidemment, cette droite est une autre courbe dans \mathbb{R}^p .

Nous utiliserons ces données pour calculer la longueur de l'arc γ pour $t \in [a, b]$, à condition que les valeurs de a et b conviennent. Si nous étions dans le cas de l'intégration comme dans le chapitre précédent, alors on s'attendrait à une intégrale de la forme

$$\int_{\gamma} 1 dt$$

avec une expression pour qui s'obtient en ajoutant les "valeurs" de γ sur des subdivisions de $[a, b]$. Or, cette notion de valeur a peu de sens. Néanmoins, une courbe est un objet de dimension 1 dont les morceaux ressemblent à des segments de droites d'autant plus que les subdivisions de $[a, b]$ deviennent fines. En effet, sur un intervalle d'une telle subdivision γ est presque une droite tangente en un point arbitraire de l'intervalle en question.

Plus précisément, on fixe d'abord une subdivision de $[a, b]$:

$$t_0 = a \leq t_1 \dots \leq t_m = b.$$

Plus cette division est fine, plus le vecteur tangent en t_i ($i \in \{1, \dots, m\}$)

$$(\gamma'_1(t_i), \dots, \gamma'_p(t_i))$$

est identique au vecteur qui lie $\gamma(t_i)$ au point $\gamma(t_{i+1})$. Ainsi les longueurs $\|\gamma'(t_i)\|_2$ et $\|\gamma(t_{i+1}) - \gamma(t_i)\|_2$ se convergent quand les longueurs des intervalles des subdivisions convergent vers 0. Ceci motive la définition de la longueur d'un *arc de courbe* de $t = a$ à $t = b$ qui est la suivante :

$$L_{\gamma}(a, b) = \int_a^b \|(d\gamma)_t\|_2 dt.$$

Nous utilisons la métrique euclidienne pour rester conforme aux calculs de distance usuelles dans \mathbb{R}^p . Ainsi,

$$L_{\gamma}(a, b) = \int_a^b \sqrt{\gamma'_1(t)^2 + \dots + \gamma'_p(t)^2} dt.$$

Une question naturelle se pose. Est-ce bien défini ? En effet, s'il était possible de déterminer la même courbe en utilisant une fonction θ avec les mêmes propriétés de différentiation, comment saurait-on que la même longueur est obtenue ? Un tel changement correspond à un nouveau choix de paramétrage $\Phi : [c, d] \rightarrow [a, b]$ qui est forcément différentiable pour que le tout reste différentiable. Il s'agit d'un changement de variables du type $\theta = \gamma \circ \Phi$ avec Φ un difféomorphisme. On obtient alors en tout $u \in [c, d]$

$$\begin{aligned} \int_c^d \|(d\theta)_u\|_2 du &= \int_c^d \|(d\gamma)_{\phi(u)}\phi'(u)\|_2 du \\ &= \int_a^b \|(d\gamma)_t\|_2 dt \end{aligned}$$

Une fois que nous avons la définition de la longueur d'arc dans \mathbb{R}^p , l'intégration d'une fonction le long de cet arc n'est que le calcul d'une somme pondérée de façon réminiscente de l'intégration d'une fonction dans le chapitre précédent. Soit donc une fonction f de \mathbb{R}^p vers \mathbb{R} définie sur $\gamma([a, b])$. Son *intégrale curviligne* est

$$\int_a^b f(\gamma_1(t), \dots, \gamma_p(t)) \sqrt{\gamma'_1(t)^2 + \dots + \gamma'_p(t)^2} dt.$$

8.2 Formes différentielles : recherche d'une primitive en dimension supérieure

Cette section est consacrée à un autre lien entre l'intégration et la différentiation des fonctions de plusieurs variables : la recherche d'une primitive. Cette étude nécessite l'introduction des formes différentielles qui est un vaste sujet. Nous nous contenterons de très peu ce qui néanmoins suffira de découvrir un univers géométrique très riche.

Soient $U \subset \mathbb{R}^p$ un ouvert et $f : \mathbb{R}^p \rightarrow \mathbb{R}$ une fonction différentiable sur U . Nous savons bien que la différentielle est la fonction

$$\begin{aligned} df & : U \longrightarrow \mathcal{L}(\mathbb{R}^p, \mathbb{R}) \\ a & \longmapsto (df)_a = (\partial_1 f(a) \dots \partial_p f(a)) \end{aligned}$$

Comme $\mathcal{L}(\mathbb{R}^p, \mathbb{R})$ est un espace vectoriel, nous pouvons considérer de fixer une base pour les fonctions différentiables sur U et exprimer une df comme une combinaison linéaire des éléments de cette base. La base canonique (e_1, \dots, e_p) de \mathbb{R}^p nous guidera dans la détermination d'une base canonique pour $\mathcal{L}(\mathbb{R}^p, \mathbb{R})$.

Si nous voyons $(df)_a$ plutôt comme un point dans \mathbb{R}^p (un "vecteur"), en d'autres termes, si nous considérons $\nabla f(a)$, alors celui-ci s'écrit

$$\nabla f(a) = (\partial_1 f(a), \dots, \partial_p f(a)) = \partial_1 f(a)e_1 + \dots + \partial_p f(a)e_p.$$

Chaque élément e_i de la base canonique projetée sur $\nabla f(a)$ sur la coordonnée correspondante i . Les différentielles de ces projections, que nous avons notées π_i au deuxième chapitre, forment la base canonique de $\mathcal{L}(\mathbb{R}^p, \mathbb{R})$. Elles sont notées dx_i et nous permettent d'écrire la combinaison linéaire suivante :

$$(df)_a = \partial_1 f(a)dx_1 + \dots + \partial_p f(a)dx_p.$$

On rencontre souvent l'une des deux écritures suivantes utilisées pour noter la fonction df :

$$df = \sum_{i=1}^p \partial_i f dx_i \quad ; \quad df = \sum_{i=1}^p \frac{\partial_i f}{\partial x_i} dx_i.$$

Cette écriture motive la forme générale suivante :

$$\sum_{i=1}^p Q_i(x) dx_i \quad (\text{pour tout } i \in \{1, \dots, p\}, Q_i : \mathbb{R}^p \rightarrow \mathbb{R}).$$

Il s'agit d'une fonction de \mathbb{R}^p vers $\mathcal{L}(\mathbb{R}^p, \mathbb{R})$ définie sur et peut-être plus sur l'ouvert concerné.

Définition 8.2.1 (Forme différentielle de degré 1) Une fonction de la forme

$$\omega(x) = \sum_{i=1}^p Q_i(x) dx_i \quad (\text{pour tout } i \in \{1, \dots, p\}, Q_i : \mathbb{R}^p \rightarrow \mathbb{R})$$

est dite une forme différentielle de degré 1. Elle est dite de classe $\mathcal{C}^k(U)$ s'il existe un ouvert U de \mathbb{R}^p sur lequel chaque "coefficient" Q_i est de classe $\mathcal{C}^k(U)$.

Il est clair que chaque fonction f différentiable sur un ouvert U de \mathbb{R}^p y définit une forme différentielle de degré 1, qui n'est rien d'autre que sa différentielle df . Mais est-ce que ce sont les seules possibilités ? Ou encore, si ω est une forme différentielle de degré 1 arbitrairement définie, existe-t-il une fonction $f : \mathbb{R}^p \rightarrow \mathbb{R}$ telle que $df = \omega$ sur un ouvert de \mathbb{R}^p ? Cette question est la recherche de primitive évoquée au début de cette section.

Comme dans le cas des fonctions d'une seule variable, cette recherche n'est ni évidente ni facile. La réponse générale est négative, et quand elle est positive, elle est très liée à l'intégration.

Notons que si une “primitive” f telle que $df = \omega$ existe, alors $Q_i = \partial_i f$ pour tout $i \in \{1, \dots, p\}$. Les dérivées partielles seront des outils importantes donc.

Jusqu’à mention contraire, nous nous restreignons à $p = 2$. Dans ce contexte très particulier, nous noterons les formes en suivant une notation traditionnelle :

$$\omega = Pdx + Qdy .$$

Bien évidemment, ce n’est qu’une notation. De manière plus détaillée, ω est une fonction de \mathbb{R}^2 vers $\mathcal{L}(\mathbb{R}^2, \mathbb{R})$ (nous pouvons voir ce dernier espace comme \mathbb{R}^2), qui associe à chaque point $a \in \mathbb{R}^2$, le point (ou encore le vecteur, ou l’application linéaire de \mathbb{R}^2 vers \mathbb{R}) dont les coordonnées par rapport à la base canonique sont

$$(P(a), Q(a)) .$$

Ce type de fonction porte un autre nom que nous avons déjà cité parmi les exemples de la première section du deuxième chapitre au sujet des représentations de fonctions de plusieurs variables : *un champ de vecteurs*.

Un champ de vecteurs, d’usage courant en physique, ne nous est pas étranger. Pendant l’étude des courbes, nous avons souvent parlé de la pente de la droite tangente à une courbe à un point donné. Cette association était décrite par une fonction associant à un point $a \in \mathbb{R}^2$ un vecteur perpendiculaire au vecteur $\nabla f(a)$ (voir la discussion à la fin de la section 6.1).

Proposition 8.2.1 *Soit $\omega = Pdx + Qdy$ une forme différentielle de degré 1 de classe $\mathcal{C}^1(U)$ sur un ouvert de U . S’il existe $f \in \mathcal{C}^2(U)$ telle que $df = \omega$ sur U , alors $\partial_2 P = \partial_1 Q$.*

L’idée de la preuve, réminiscente du théorème de Schwarz, est simple. Si une telle f existe, alors $\partial_1 f = P$ et $\partial_2 f = Q$. Il s’ensuit sous les hypothèses données que

$$\partial_2 P = \partial_2 \partial_1 f = \partial_1 \partial_2 f = \partial_1 Q .$$

Cette proposition relativement simple définit une condition nécessaire pour l’existence d’une “primitive”. Pour exprimer nos notions de manière plus claire, nous introduisons la terminologie suivante très importante sans aucune restriction sur la valeur de p :

Définition 8.2.2 (Formes exactes/fermées) *Soit U un ouvert de \mathbb{R}^p .*

1. *Une forme ω de classe $\mathcal{C}^1(U)$ est dite exacte sur U s’il existe une fonction f de classe $\mathcal{C}^2(U)$ telle que $df = \omega$. La fonction f est dite une primitive de ω .*
2. *Pour toute forme différentielle $\omega = Pdx + Qdy$ de classe $\mathcal{C}^1(U)$, nous noterons $d\omega$ la fonction $\partial_1 Q - \partial_2 P$. Une forme ω est dite fermée si $d\omega = 0$.*

Nous nous sentons obligés de faire la remarque suivante, qui n’aura pas de conséquences pratiques dans ce cours où les formes différentielles ne sont étudiées qu’au niveau d’introduction : la définition donnée pour “la différentielle de ω ”, $d\omega$, n’est pas très rigoureuse ni complète. Elle sera suffisante pour nous.

En utilisant la définition 8.2.2, nous pouvons énoncer la conclusion de la proposition 8.2.1 de la manière suivante : si une forme différentielle est exacte, alors elle est fermée. L’inverse est faux, et pour déterminer des cas où il est vrai il nous faudra utiliser certains théorèmes. Ces résultats forment les fondements de la recherche des primitives pour les fonctions de plusieurs variables. C’est la théorie de l’intégration des formes différentielles.

8.3 Formes différentielles : intégration

La définition 8.2.2 a fourni la terminologie et les notions pour développer une théorie de l’intégration des formes différentielles dont les rudiments seront introduits dans cette section. Ce qui reste à faire est d’introduire une notion solide d’intégration d’une forme différentielle de

degré 1 qui permettra en outre de déterminer les primitives quand elles existent. L'intégration d'une forme différentielle suit une ligne différente de celle des intégrales curvilignes malgré des points communs. Jusqu'à mention contraire, il n'y aura pas de restriction sur la valeur de $p = 2$.

Soient U un ouvert de \mathbb{R}^p , pour tout $x = (x_1, \dots, x_p) \in U$

$$\omega(x) = Q_1 dx_1 + \dots + Q_p dx_p$$

une forme continue sur U (les coefficients Q_i sont des fonctions continues sur U), γ une courbe donnée paramétriquement

$$\begin{aligned} \gamma &: \mathbb{R} \longrightarrow \mathbb{R}^p \\ t &\longmapsto (\gamma_1(t), \dots, \gamma_p(t)) \end{aligned}$$

telle que l'image d'un intervalle fermé $[a, b]$ tombe dans U où la courbe est de classe \mathcal{C}^1 . Nous posons alors

$$\begin{aligned} \int_{\gamma} \omega &= \int_{\gamma} Q_1(x) dx_1 + \dots + Q_p(x) dx_p \\ &= \int_{\gamma} [Q_1(x(t))x'_1(t) + \dots + Q_p(x(t))x'_p(t)] dt \end{aligned}$$

Plus généralement, γ peut être l'union d'un nombre fini de courbes de classe \mathcal{C}^1 qui sont les images des parties d'une subdivision de γ , disons $\gamma_1, \dots, \gamma_k$. Nous appellerons une telle courbe \mathcal{C}^1 par morceaux. Dans le cas d'une courbe \mathcal{C}^1 par morceaux, l'intégrale est la somme des intégrales sur les morceaux \mathcal{C}^1 de γ :

$$\int_{\gamma} \omega = \sum_{i=1}^k \int_{\gamma_i} \omega.$$

Les intégrales curvilignes peuvent être vues comme un cas très particulier de cette définition définies par "une forme de longueur". Nous ne détaillerons pas ces liens. Voici un premier théorème qui illustre la force de l'hypothèse d'exactitude pour les formes différentielles :

Théorème 8.3.1 Soient U un ouvert de \mathbb{R}^p , ω une forme exacte de classe \mathcal{C}^2 sur U , γ une courbe paramétrée de classe \mathcal{C}^1 par morceaux qui joint un point $A \in U$ à un point $B \in U$ sur l'intervalle fermé $[a, b]$ dont l'image $\gamma([a, b])$ est incluse dans U . Alors

$$\int_{\gamma} \omega = f(B) - f(A),$$

où f est une primitive de ω : $df = \omega$.

Vous pouvez comparer ce théorème à celui sur l'intégration des fonctions d'une seule variable qui énonce que pour toute fonction g

$$\int_a^b g(x) dx = G(b) - G(a)$$

quand $G' = g$ sur un intervalle ouvert contenant $[a, b]$; G est donc une primitive de g . Ceci est en fait un phénomène qui se produit naturellement : l'intégration se fait sur l'intervalle $[a, b]$ tandis que le calcul de différence se fait sur la frontière de cet intervalle.

Le théorème 8.3.1 montre que le même phénomène de "passage à la frontière" a lieu quand nous intégrons sur une courbe puisque $\text{Fr}(\gamma([a, b])) = \{A, B\}$. En particulier, si le point de départ coïncide avec le point d'arrivée, il n'y a pas de contribution à la valeur de l'intégrale. Comparer par exemple les valeurs de l'intégrale

$$\int_{\gamma} y dx + x dy$$

où $\gamma(t) = (\cos t, \sin t)$ sur les intervalles $[0, \frac{\pi}{4}]$, $[0, 2\pi]$, $[0, \frac{9\pi}{4}]$, $[0, 4\pi]$ pour les valeurs de t .

Les dernières remarques illustrent qu'éventuellement les conditions topologiques et géométriques liées aux courbes qui entrent dans les calculs peuvent avoir des effets dans le développement de l'intégration des formes différentielles. Nous aurons besoin de la définition suivante :

Définition 8.3.1 (Ensemble étoilé dans \mathbb{R}^p) Une partie E de \mathbb{R}^p est dite étoilée si elle contient un point C tel que si P est un point arbitrairement choisi dans E , alors le segment de droite qui joint C à P soit contenu dans E .

Donnons quelques exemples pour motiver cette définition.

1. Un ensemble convexe est étoilé. Ceci découle immédiatement de la définition d'un ensemble convexe. En effet, tout point peut jouer le rôle du point C de la définition 8.3.1. Comme le montre l'exemple suivant, un ensemble étoilé n'est pas nécessairement convexe.
2. Dans \mathbb{R}^2 , l'union des deux axes $\{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ est étoilé mais pas convexe. Pour le point C il y a un candidat et un seul : $(0, 0)$.
3. La boule euclidienne ouverte $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ est convexe et donc étoilée, mais quand privée d'un point, elle n'est ni convexe ni étoilée. Essayez de voir pourquoi, vous constaterez vite qu'il s'agit d'un "trou" dans la boule qui l'empêche d'être étoilée.

La notion d'ensemble étoilé joue un rôle fondamental dans la détermination des formes exactes. Nous avons dans la proposition 8.2.1 qu'une forme exacte est nécessairement fermée. Comme nous le verrons à travers des exemples en cours et aux travaux dirigés, l'inverse est faux en général. Or, tant il est facile de vérifier qu'une forme est fermée, on dérive, tant il est éventuellement difficile de vérifier qu'une forme fermée a aussi une primitive, il faut intégrer. Le théorème suivant montre que sous certaines hypothèses nous pouvons profiter de la notion d'ensemble étoilé.

Théorème 8.3.2 (Théorème de Poincaré) Soient U un ouvert étoilé de \mathbb{R}^p et ω une forme de classe C^1 sur U . Alors ω est exacte si et seulement si elle est fermée.

Comme nous l'avons indiqué dans le point (3) des exemples ci-dessus, le fait d'être étoilé est lié à l'absence de "trous" dans une partie de \mathbb{R}^p . En effet, le théorème de Poincaré est valable sur tout ouvert "sans trous". Néanmoins, rendre rigoureux notre intuition de "avec/sans trous" nécessite une préparation qui sera faite dans des cours plus avancés.

Voici un exemple pertinent des limites de l'applicabilité du théorème de Poincaré :

$$\omega = P(x, y)dx + Q(x, y)dy$$

avec

$$P(x, y) = \frac{-y}{x^2 + y^2} \quad \text{et} \quad Q(x, y) = \frac{x}{x^2 + y^2}$$

et $U = \mathbb{R} \setminus \{(0, 0)\}$. Nous pouvons remplacer \mathbb{R}^2 par tout ouvert contenant $(0, 0)$ sans changer la conclusion : la forme ω , quoique fermée, n'est exacte sur aucun ouvert contenant $(0, 0)$. Néanmoins, sur tout ouvert ne contenant pas $(0, 0)$, le théorème de Poincaré s'applique.

Le théorème de Poincaré règle le problème de l'existence d'une primitive efficacement. Nous finirons ce chapitre avec un théorème qui permet de faire de l'intégration en utilisant les formes différentielles. C'est un cas particulier pour $p = 2$ d'un résultat général. En effet, il est possible de développer la théorie générale des formes différentielles et leur intégration de telle façon à obtenir une version plus différentielle de ce que nous avons fait au septième chapitre.

Théorème 8.3.3 (Formule de Green-Riemann) Soit U un ouvert de \mathbb{R}^2 sur lequel est définie une forme différentielle $\omega = Pdx + Qdy$ de classe C^1 . Soit E une région dans \mathbb{R}^2 limitée par une courbe que l'on notera ∂E et que l'on supposera C^1 par morceaux. Alors

$$\int_{\partial E} \omega = \int_E d\omega \, d(x, y) .$$

Nous ferons plusieurs remarques sur ce théorème important.

1. Le membre droit de la formule de Green-Riemann est l'intégrale double

$$\int \int_E (\partial_1 Q - \partial_2 P) d(x, y) .$$

Tout cela signale que l'on doit pouvoir utiliser cette égalité pour déterminer des aires en profitant de la marge de manoeuvre offerte par deux intégrales différentes ayant la même valeur. En effet, l'égalité suivante est vraie pour toute partie $E \subset \mathbb{R}^2$ satisfaisant les hypothèses de la formule de Green-Riemann :

$$\int_{\partial E} x dy = \iint_E dx dy .$$

Il suffit de poser $Q(x, y) = x$ et $P(x, y) = 0$. Cette égalité permet de déterminer l'aire de E si sa frontière est décrite par une courbe dont nous connaissons les équations paramétriques. Une autre égalité de ce genre est la suivante :

$$\int_{\partial E} \frac{1}{2}(x dy - y dx) = \iint_E dx dy .$$

Le membre gauche devient en coordonnées polaires (r, t)

$$\frac{1}{2} \int_{\partial E} (x(t)y'(t) - y(t)x'(t)) dt = \frac{1}{2} \int_a^b r^2 dt$$

où a et b sont des bornes à déterminer suivant la forme de la frontière. Cette dernière forme est utile pour les frontières données par des équations polaires où il y a un lien entre r et t , par exemple si r est une fonction de t .

2. L'endroit déterminé par cette courbe est la frontière $Fr(E)$. Néanmoins, comme les paramétrisations peuvent varier, $Fr(E)$ peut être l'image directe de plusieurs courbes paramétrées.