

CALCUL DIFFÉRENTIEL ET OPTIMISATION

J. Féjoz
Université Paris-Dauphine
`jacques.fejoz@dauphine.fr`

2017

Cet ouvrage est sous licence *Creative Commons Attribution 4.0 International*. Pour accéder à une copie de cette licence, merci de se rendre à l'adresse

<http://creativecommons.org/licenses/by/4.0/>

ou d'envoyer un courrier à

Creative Commons
444 Castro Street, Suite 900
Mountain View, California, 94041, USA.

Avertissement

La découverte du calcul différentiel et des équations différentielles, notamment par I. Newton¹ et G. Leibniz,² est l'une des plus extraordinaires conquêtes de l'esprit humain. Ce cours en est une introduction. On y parle de dérivée d'une application, du théorème d'inversion locale et de certaines de ses variantes géométriques, des surfaces (sous-variétés) de \mathbb{R}^n , des multiplicateurs de Lagrange, de convexité — l'alphabet de bien des raisonnements en mathématiques et dans quasiment toutes les disciplines scientifiques quantitatives. C'est un vieux sujet, mais on trouvera dans ce cours le portrait de mathématiciens contemporains.

Sous les hypothèses de ce cours, les applications possèdent des formes normales linéaires (par exemple, le théorème d'inversion locale donne une condition nécessaire et suffisante pour qu'une application soit localement équivalente, à changement de coordonnées près, à l'identité). Il convient donc d'être à l'aise avec les rudiments de l'algèbre linéaire, dont nous avons fait des rappels au fil du cours. Nous avons pris le parti de nous concentrer sur le calcul différentiel en dimension finie. Les équations différentielles, ordinaires ou aux dérivées partielles, sont traitées dans d'autres cours à l'Université Paris-Dauphine.

Inévitablement, le cours paraîtra abstrait sans un important travail d'appropriation, qui passe par la lecture et la relecture du cours, ainsi que la résolution des exercices. Il faut savoir redire précisément les définitions et les énoncés des propositions et théorèmes, ainsi qu'avoir au moins une idée des démonstrations, sans quoi l'on ne peut pas prétendre comprendre les énoncés, ni aller plus loin dans des cours ultérieurs, même plus "concrets". Les exercices sont le plus souvent des applications immédiates du cours. Il faut aussi s'entraîner à faire *certaines des exercices plus ambitieux, comportant plusieurs questions interdépendantes qui ne sont pas directement reliées au chapitre en cours* ; c'est le seul test véritable pour voir si l'on a compris. Faire un dessin ou comprendre un cas particulier doivent permettre de ne jamais rester muet.

L'examen a pour objectif de vérifier la capacité à résoudre des exercices simples mais variés, et s'assurer que des cours plus poussés seront profitables. Essayer de s'y préparer spécifiquement en apprenant rapidement un petit nombre de "questions de cours" balisées, ou en apprenant à résoudre un petit nombre d'exercices type, serait une perversion de l'idée d'examen. Le jeu en est à la fois intéressant et difficile. J'espère que le lecteur y trouvera plaisir... Bon travail!

J. F.

1. Isaac NEWTON, physicien et mathématicien anglais (1642–1727), dont on peut légitimement soutenir qu'il marqua la naissance de la science moderne.

2. Gottfried Wilhelm LEIBNIZ, savant allemand (1646–1716).

Indications bibliographiques et remerciements

L'esprit de ce cours est proche des remarquables notes de cours de l'université Paris-Diderot (non publiées) d'A. Chenciner, du livres de M. Chaperon [Cha03] et de F. Laudenbach [Lau11], ainsi que du livre d'exercices de F. Rouvière [Rou03], livres que l'on consultera avec profit. Les appendices de topologie sont inspirés de cours de Choquet [Cho64] de T. Tao [Tao16].

Les chapitres portant une étoile peuvent être sautés en première lecture. Ceux sur la formule de Stokes sont une introduction heuristique, dont les trous pourront être comblés par la lecture d'ouvrages plus spécialisés.

Les livres suivants sont plus avancés et fourniront des références et des compléments sur divers aspects du cours :

- Arnold [Arn84] sur les équations différentielles,
- Arnold encore [Arn88, Arn89] sur les systèmes dynamiques et la mécanique classique,
- J.-P. Demailly [Dem06] sur l'analyse numérique,
- F. Pham [Pha92] sur le calcul et la géométrie différentiels,
- Rudin [Rud74] sur l'analyse complexe,
- Hirsch [Hir94] et Milnor [Mil97] sur la topologie différentielle,
- Hörmander [Hör07] sur l'analyse convexe,
- Kolmogorov-Fomin [KF57], H. Brezis [Bre99] ou Hirsch-Lacombe [HL99] sur l'analyse fonctionnelle.

Merci à ceux qui m'ont aidé à améliorer ce cours, dont A. Ben Amor, T. Castan, A. Chenciner, G. Flath, A. Frouvelle et D. Gontier, et qui continueront à le faire...

Table des matières

Avertissement	3
Indications bibliographiques	5
1 Dérivée d'un chemin	9
2 Dérivée d'une application	17
3 Classe C^1	27
4 Le théorème du point fixe	33
5 Inversion locale	39
6 Forme normale d'une application linéaire	47
7 Forme normale d'une application	51
8 Fonctions implicites	53
9 Submersions et immersions	55
10 Incursion en analyse complexe ★	59
11 Surfaces I – Espace tangent	63
12 Extrema de fonctions (semi-)continues	69
13 Factorisation des applications linéaires	75
14 Points critiques de fonctions	79

15 Surfaces II – Coordonnées	85
16 Déterminant	91
17 Formes différentielles ★	97
18 Surfaces III – Orientabilité ★	103
19 Intégration ★	107
20 Formule de Stokes ★	113
21 Dérivées d'ordre supérieur	117
22 Forme normale d'une forme quadratique	121
23 Fonctions de Morse	125
24 Convexité	131
25 Fonctions convexes	135
26 Et la dimension infinie ? ★	141
A Topologie	147
B Compacité	153
C Espaces métriques complets	157
Glossaire	161
Alphabet grec	165
Bibliographie	168
Index	169

Chapitre 1

Dérivée d'un chemin

Mots-clefs du chapitre Chemin, courbe paramétrée, vitesse, accélération, formule de Taylor, vecteur tangent, point régulier, point d'inflexion, point de rebroussement (cusp)

Soit $c : I$ intervalle ouvert $\subset \mathbb{R} \rightarrow \mathbb{R}^p$, $t \mapsto c(t)$. Une telle application, qui ne dépend que d'une variable, s'appelle un *chemin* ou une *courbe paramétrée* de \mathbb{R}^p : on peut penser à la variable t comme au temps et à $c(t)$ comme la position d'un point mobile, au temps t .

La *courbe* (non paramétrée) associée à c est l'image $c(I) \subset \mathbb{R}^p$ de l'application c . Réciproquement, c s'appelle un *paramétrage* de son image.

1.1 Exemple. Le *graphe* d'une fonction $f : I$ intervalle ouvert $\subset \mathbb{R} \rightarrow \mathbb{R}$ est la courbe

$$C = \{(x, f(x)), x \in I\} \subset \mathbb{R}^2.$$

Son paramétrage par l'abscisse est $c : I \rightarrow \mathbb{R}^2$, $x \mapsto (x, f(x))$.

1.2 Définition. La *dérivée* de c en $t \in I$ est la limite dans \mathbb{R}^p , si elle existe,

$$c'(t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} (c(t + \tau) - c(t)) \in \mathbb{R}^p.$$

Il est équivalent de dire que

$$c'(t) = \frac{1}{\tau} (c(t + \tau) - c(t)) + o(1),$$

ou que

$$c(t + \tau) = c(t) + c'(t)\tau + o(\tau), \tag{1.1}$$

quand τ tend vers 0, c'est-à-dire que c possède un développement limité du premier ordre en t , et que le coefficient dans le terme linéaire est $c'(t)$.

— La notation $o(\tau^k)$ désigne une fonction négligeable devant τ^k quand τ tend vers 0, c'est-à-dire telle que $\frac{o(\tau^k)}{\tau^k} \rightarrow 0$.

— En composantes, si l'on note¹ $c = \begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix}$, on obtient

$$\begin{pmatrix} c_1(t + \tau) \\ \vdots \\ c_p(t + \tau) \end{pmatrix} = \begin{pmatrix} c_1(t) \\ \vdots \\ c_p(t) \end{pmatrix} + \begin{pmatrix} c'_1(t) \\ \vdots \\ c'_p(t) \end{pmatrix} \tau + \begin{pmatrix} o(\tau) \\ \vdots \\ o(\tau) \end{pmatrix},$$

donc

$$c'(t) = \begin{pmatrix} c'_1(t) \\ \vdots \\ c'_p(t) \end{pmatrix}. \quad (1.2)$$

La caractérisation (1.1) a le mérite de garder un sens si τ est un vecteur, à condition que $c'(\tau)$ soit une application linéaire et que $o(\tau)$ soit pris au sens de $o(\|\tau\|)$; c'est celle-ci qui nous permettra de généraliser la notion de dérivée au cas de plusieurs variables.

1.a Exercice (Deux chemins de Lissajous). Tracer l'image des chemins $c = (x, y) : \mathbb{R} \rightarrow \mathbb{R}^2$ suivants :

$$\begin{cases} x(t) = \cos^3 t \sin t \\ y(t) = \cos^2 t, \end{cases} \quad \begin{cases} x(t) = 3 \cos t + 2 \cos(3t) \\ y(t) = 3 \sin t - 2 \sin(3t) \end{cases}$$

(ce sont deux exemples de *courbes de Lissajous*,² dont, par définition, les composantes sont des polynômes trigonométriques).

Indication. En exploitant la périodicité et la symétrie par rapport aux axes de coordonnées, on se ramène à étudier c sur l'intervalle $[0, \pi/2]$ dans le premier cas et $[0, \pi/4]$ dans le second. Pour tracer les tangentes verticales et horizontales, on pourra utiliser la notion de vecteur tangent en un point régulier, décrite ci-dessous; pour tracer les tangentes au point singulier $t = \pi/2$ de la première courbe, on utilisera la formule de Taylor rappelée ci-dessous ou, plus généralement, on pourra se référer à l'exercice 1.b. Voir la figure 1.1.

□

On note donc $c' : I \rightarrow \mathbb{R}^p$ la dérivée ou *vitesse* de c . C'est une application de même nature que c elle-même, à savoir un chemin de \mathbb{R}^p . En continuant à dériver, on note $c'' = (c')'$ l'*accélération*, $c''' = (c'')'$, puis $c^{(k)} = (c^{(k-1)})'$ pour les ordres de dérivation suivants. Par exemple, quand un homme politique affirme que la hausse du chômage décélère, c'est dire que, si c est le nombre de chômeurs, $c''' < 0$.

1. Noter les vecteurs verticalement est arbitraire. Cette convention a l'unique avantage de faciliter le calcul des produits matrice-vecteur. Nous ne serons pas toujours cohérents là-dessus et nous noterons parfois $c = (c_1, \dots, c_p)$ un chemin dans \mathbb{R}^p sans que cela doivent être compris différemment.

2. Jules Antoine LISSAJOUS (1822-1884), physicien français, célèbre pour son étude des phénomènes oscillatoires, ainsi que pour avoir expérimenté, pendant le siège de Paris en 1870, un système de communication optique via une montgolfière

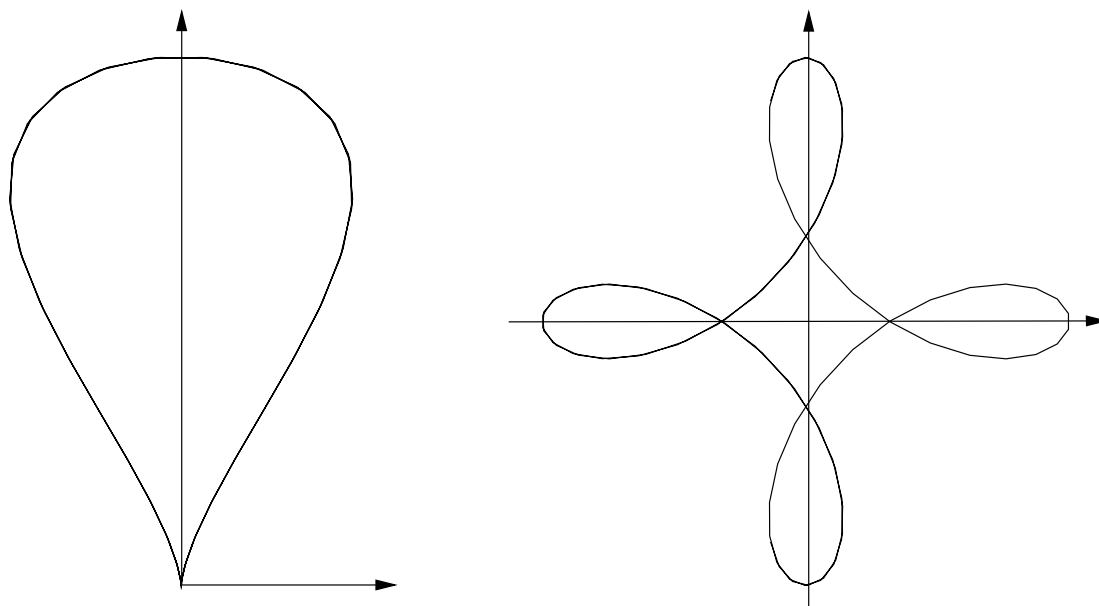


FIGURE 1.1 – Deux chemins de Lissajous

L'allure d'un chemin au voisinage d'un point (notamment un point où la vitesse s'annule) est déterminée par l'importante formule de Taylor, que nous rappelons maintenant : le développement de Taylor de c s'obtient simplement en développant chacune des coordonnées de c .

1.3 Rappel (Les deux formules fondamentales du calcul différentiel [Féj14]). Si $f : [a, b] \rightarrow \mathbb{R}$ est en escalier (constante sur chaque intervalle d'une subdivision finie de $[a, b]$), on définit de façon évidente l'intégrale de f . Si f est seulement continue (éventuellement par morceaux), il existe une suite (f_n) de fonctions en escalier qui converge uniformément vers f , et l'on montre que : 1) l'intégrale de f_n converge, 2) la limite ne dépend pas du choix de la suite. Si maintenant $c = (c_1, \dots, c_p)$ est un chemin de \mathbb{R}^p continu sur $[a, b]$, on pose $\int_a^b c(t) dt = \left(\int_a^b c_1(t) dt, \dots, \int_a^b c_p(t) dt \right)$. Avec ces définitions, on montre que,

1) Si $c : [a, b] \rightarrow \mathbb{R}^p$ est continue, la fonction $C(t) = \int_a^t c(s) ds$ est de classe C^1 et $C'(t) = c(t)$, i.e. C est une solution de l'équation différentielle $C' = c$ de donnée c .

2) Si $C : [a, b] \rightarrow \mathbb{R}^p$ est de classe C^1 , $C(b) - C(a) = \int_a^b C'(t) dt$, i.e. l'équation différentielle précédente possède au plus une solution une fois que $C(a)$ est fixé.

1.4 Notation. On note $c : (\mathbb{R}, t) \rightarrow \mathbb{R}^p$ un chemin défini sur un certain voisinage de t dans \mathbb{R} . Cette notation permet de ne pas donner de nom à ce voisinage. On peut aussi noter $c : (\mathbb{R}, t) \rightarrow (\mathbb{R}^p, a)$ si de plus $c(t) = a$.

1.5 Théorème (Formule de Taylor). *Si c est de classe C^{k+1} sur un voisinage de t (c 'est-à-dire que $c^{(k+1)}$ existe sur ce voisinage et y est continue), pour tout τ assez petit,*

$$c(t + \tau) = c(t) + c'(t)\tau + \dots + c^{(k)}(t) \frac{\tau^k}{k!} + \left(\int_0^1 \frac{(1-s)^k}{k!} c^{(k+1)}(t + s\tau) ds \right) \tau^{k+1}.$$

Démonstration. D'après la seconde formule fondamentale ci-dessus, appliquée au chemin $s \mapsto c(t + s\tau)$, $[0, 1] \rightarrow \mathbb{R}^p$,

$$\begin{aligned} c(t + \tau) &= c(t) + \int_0^1 \frac{d}{ds} c(t + s\tau) ds \\ &= c(t) + \left(\int_0^1 c'(t + s\tau) ds \right) \tau. \end{aligned}$$

Ensuite, une intégration par parties (en choisissant, comme primitive de 1, la fonction $s - 1$, qui a le mérite de s'annuler en $s = 1$ et donc de faire disparaître le terme en $c'(t + \tau)$) montre

$$\begin{aligned} c(t + \tau) &= c(t) - [(1 - s)c'(t + s\tau)]_0^1 \tau + \left(\int_0^1 (1 - s)c''(t + s\tau) ds \right) \tau^2 \\ &= c(t) + c'(t)\tau + \left(\int_0^1 (1 - s)c''(t + s\tau) ds \right) \tau^2, \end{aligned}$$

et, par récurrence, la formule voulue. \square

Voici maintenant une généralisation de la notion de vecteur vitesse. En effet, en un point t où la vitesse c' s'annule, celle-ci ne suffit pas pour déterminer la direction tangente au chemin; c'est qu'on arrive trop lentement au point $c(t)$ pour voir d'où l'on vient et où l'on va. On pallie cette difficulté en accélérant le temps.

1.6 Définition. La *droite tangente* à c en t est la droite limite, si elle existe, de la droite ("corde") joignant $c(t)$ à $c(t + \tau)$ quand τ tend vers 0. En prenant la limite quand τ tend vers 0^\pm , on obtient la notion de *droite tangente à gauche* ou à *droite*.

La droite $D_{t,\tau}$ joignant $c(t)$ à $c(t + \tau)$ est l'ensemble des points $p = (x, y)$ tels que les vecteurs $p - c(t)$ et $c(t + \tau) - c(t)$ soient colinéaires. Elle a donc pour équation

$$\det(p - c(t), c(t + \tau) - c(t)) = \begin{vmatrix} x - x(t) & x(t + \tau) - x(t) \\ y - y(t) & y(t + \tau) - y(t) \end{vmatrix} = 0.$$

Le cas le plus simple est celui où $c'(t) \neq 0$; on dit que t est un point *régulier* de c . D'après la formule de Taylor au premier ordre :

$$c(t + \tau) = c(t) + c'(t)\tau + o(\tau),$$

donc l'équation est de la forme

$$\tau(y'(t)(x - x(t)) - x'(t)(y - y(t)) + o(1)) = 0.$$

Après division de l'équation par τ (ce qui ne change pas la droite), quand τ tend vers 0 l'équation de $D_{t,\tau}$ tend vers l'équation

$$y'(t)(x - x(t)) - x'(t)(y - y(t)) = 0$$

de la droite tangente en t au chemin; c'est la droite passant par $c(t)$ et de pente $y'(t)/x'(t) \in \mathbb{R} \cup \{\infty\}$ (comme $x'(t)$ et $y'(t)$ ne sont pas tous deux nuls, ce rapport n'est pas indéterminé).

En un *point singulier*, où $c'(t) = 0$, on peut avoir besoin de calculer un équivalent plus précis de $c(t + \tau) - c(t)$ quand τ tend vers 0. Souvent, il suffit de calculer un développement limité de c jusqu'au premier terme non nul de la série de Taylor, et ce dernier indique alors la direction tangente. Si la série de Taylor de c en t est nulle, on parle de *point plat*, et il faut utiliser une autre échelle de comparaison.

1.b Exercice ★ (Vecteur tangent à un chemin en un point singulier).

Soient $c = (x, y) : (\mathbb{R}, t) \rightarrow \mathbb{R}^2$ et $k \geq 1$ tels que c soit de classe C^{k+1} et³

$$c'(t) = \dots = c^{(k-1)}(t) = 0, \quad c^{(k)}(t) \neq 0;$$

on suppose k minimal pour cette propriété.

En post-composant c par une transformation linéaire qui envoie le vecteur v sur le premier vecteur de la base canonique, on se ramène au cas où $x^{(k)}(t) \neq 0$. Soit l le plus petit entier $> k$ tel que $y^{(l)}(t) \neq 0$. Le point t est

- *régulier* si $k = 1$ (c'est le cas banal, les autres sont exceptionnels)
- un *point d'inflexion* si k et l sont impairs et $l > k$
- un *point de rebroussement* (ou *point cusp*, en français) si k est pair (de *première espèce* si l impair, de *seconde espèce* si l est pair).

1. Donner des exemples de points d'inflexion et de rebroussement, et les dessiner. Comment définir une *demi-tangente* en un point de rebroussement ?

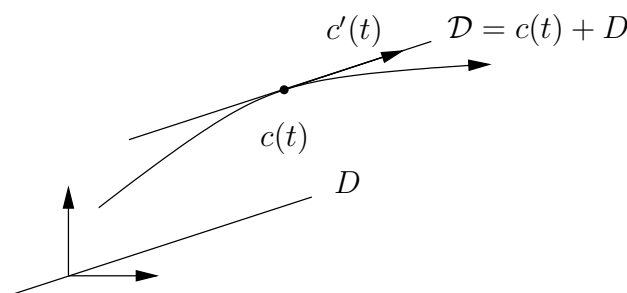


FIGURE 1.2 – Droite des vecteurs tangents en un point régulier de c ($k = 1$)

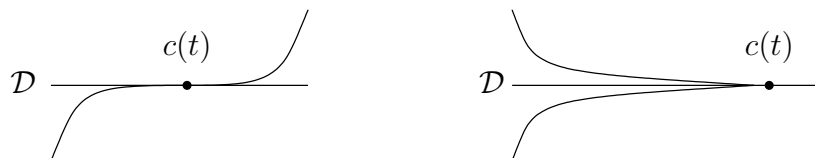


FIGURE 1.3 – Point d'inflexion et point de rebroussement

1.c Exercice (Cardioïde). Justifier le tracé figure 1.4 de la courbe définie en coordonnées polaires par l'équation $r = 1 + \cos \theta$. On déterminera la tangente en $\theta = 0$ et π .

3. Si $c'(t) = c''(t) = \dots = c^{(k-1)}(t) = 0$ et $c^{(k)}(t) \neq 0$, on dit que le chemin c est *plat* à l'ordre $k - 1$. Ici, on est en train de supposer que c n'est pas plat à tout ordre.

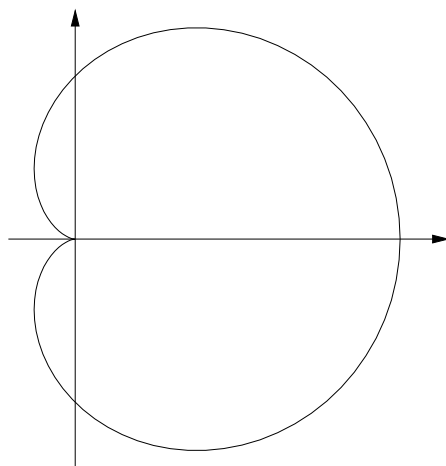


FIGURE 1.4 – Cardioïde

On montre qu'une courbe $c : (\mathbb{R}, a) \rightarrow (\mathbb{R}^p, b)$ est entièrement définie, localement au voisinage de a , par la donnée de son vecteur vitesse $c'(t)$ à chaque instant t proche de a . (Ça n'est pas évident, et l'on ne peut généralement pas calculer $c(t) = (x(t), y(t))$ explicitement, par de simples calculs d'intégrales.) C'est en fait le contenu essentiel du théorème de Cauchy-Lipschitz (voir l'exercice 4.d pour une première démonstration).

L'exercice suivant montre un cas où, à défaut de trouver explicitement le chemin c , on peut déterminer la courbe image.

1.d Exercice (Facteur intégrant).

Soient $c : \mathbb{R} \rightarrow \mathbb{R}^2$ un chemin dont la vitesse au point $(x, y) = c(t)$ est donnée par le champ de vecteurs

$$v : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (x + \sin x)(y, -x).$$

1. Montrer qu'il existe une fonction non triviale (= non constante localement au voisinage de c) f de classe C^∞ sur \mathbb{R}^2 telle f soit constante le long de c .
2. En déduire que c est bornée.

1.e Exercice (Courbure).

Soit $c : [t_0, t_1] \subset \mathbb{R} \rightarrow \mathbb{R}^2$ un chemin dans le plan, de classe C^2 . La *longueur* de c est

$$L(c) = \int_{t_0}^{t_1} \|c'(t)\| dt.$$

1. Montrer que $L(c)$ est invariant par changement de paramétrage (on fera des hypothèses plausibles sur le "changement de paramétrage").

Supposons que $t \in]t_0, t_1[$ soit un point régulier ($c'(t) \neq 0$). Le *vecteur tangent unitaire* est

$$\tau(t) = \frac{c'(t)}{\|c'(t)\|},$$

et le *vecteur normal unitaire* est le vecteur $\nu(t)$ obtenu à partir du précédent par une rotation de $\pi/2$. La *(droite) normale* est la droite passant par $c(t)$ et dirigée par le vecteur normal unitaire.

2. Montrer que généralement les normales en t et en s se coupent en un point unique, et que ce point a une limite quand s tend vers t . On appelle *centre de courbure* et l'on note $C(t)$ cette limite.

3. Déterminer le nombre $R(t) \geq 0$, appelé le *rayon de courbure*, tel que

$$C(t) = c(t) \pm R(t)\nu(t);$$

la *courbure* est $1/R(t)$.

Le chemin C est la *développée* de c et, inversement, c est la *développante* de C .

4. Trouver la développée de $c(t) = (3t - t^3, 3t^2)$.

Exercices complémentaires

1.f Exercice (Ellipse). Trouver un paramétrage rationnel de l'ellipse

$$c(t) = (a \cos t, b \sin t) \quad (t \in \mathbb{R})$$

(où a et b sont deux réels fixés), c'est-à-dire un chemin \tilde{c} dont l'image coïncide avec celle de c , et dont les composantes soient des fractions rationnelles.

1.g Exercice ★ (Étude des solutions d'une équation différentielle).

Étudier qualitativement les solutions de l'équation

$$t^2 x'' + x^2 x' - x^2 = 0$$

(sont-elles bornées?, quelle sont leurs limites au bord de leur domaine d'existence?); on pourra commencer par trouver un paramétrage rationnel de la *courbe cissoïde* d'équation

$$Y(X^2 + Y^2) - X^2 = 0,$$

et s'y ramener.

1.h Exercice (Longueur d'un chemin continu).

Soient $c : T = [a, b] \rightarrow \mathbb{R}^n$ un chemin continu et $\tau = \{t_0 = a \leq t_1 \leq \dots \leq t_{n-1} \leq t_n = b\}$ une subdivision de $[a, b]$. La *variation totale* de c relativement à τ est

$$V_\tau(c) = \sum_{i=1}^n \|c(t_i) - c(t_{i-1})\| \in [0, +\infty[,$$

où $\|\cdot\|$ est la norme euclidienne de \mathbb{R}^n . Autrement dit, $V_\tau(c)$ est la longueur du polygone de sommets successifs $c(t_0), \dots, c(t_n)$. Quand on choisit des subdivisions contenant de plus en plus de points de $[a, b]$, ce polygone approche la courbe c de plus en plus près. La *longueur* de c (voir la figure 1.5) est

$$L(c) = \sup_{\tau} V_\tau(c) \in [0, +\infty].$$

Le chemin c est *rectifiable* si il est de longueur finie.

1. Montrer que si c est lipschitzien, il est rectifiable.
2. Montrer que si c est de classe C^1 ,

$$L(c) = \int_a^b \|c'(t)\| dt;$$

on pourra appliquer le théorème des accroissements finis à la courbe $\varphi : t \mapsto c(t) - tc'(s)$, $s \in T$.

3. Montrer que

$$c : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto \begin{cases} t^2 \cos^2(\pi/t^2) & \text{si } t \neq 0 \\ 0 & \text{si } t = 0 \end{cases}$$

n'est pas rectifiable.

4. Montrer que si $T = [0, \pi]$, $n \in \mathbb{N}^*$ et $c_n : [0, \pi] \rightarrow \mathbb{R}^2$, $t \mapsto (t, n^{-1} \sin nt)$, la longueur de c_n ne dépend pas de $n \in \mathbb{N}_*$. La longueur est-elle une fonctionnelle continue?

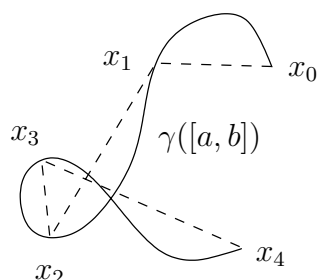


FIGURE 1.5 – La longueur de c est la limite supérieure des courbes polygonales inscrites

1.i Exercice (Limite d'une dérivée).

1. La fonction $f : x \mapsto x^2 \sin(1/x^2)$, prolongée par continuité en 0, est-elle dérivable sur $[0, 1]$? Sa dérivée est-elle continue?
2. Montrer que, si $f : [0, 1] \rightarrow \mathbb{R}^p$ est continue, dérivable sur $]0, 1]$, et telle que f' possède une limite $\ell \in \mathbb{R}^p$ en 0, f est dérivable en 0 et $f'(0) = \ell$.

Chapitre 2

Dérivée d'une application

Mots-clefs du chapitre Dérivée, dérivée partielle, gradient

On a défini la dérivée en t d'un chemin $c : (\mathbb{R}, t) \rightarrow \mathbb{R}^p$ comme le vecteur

$$c'(t) = \lim_{\tau \rightarrow 0} \frac{c(t + \tau) - c(t)}{\tau} \in \mathbb{R}^p.$$

Cette égalité équivaut au développement limité

$$c(t + \tau) = c(t) + c'(t)\tau + o(\tau),$$

où le vecteur $c'(t) = \begin{pmatrix} c'_1(t) \\ \vdots \\ c'_p(t) \end{pmatrix}$ peut aussi être vu comme une matrice-colonne,

soit une application linéaire de $L(\mathbb{R}, \mathbb{R}^p)$, et le scalaire τ comme un vecteur 1-dimensionnel, de sorte que $c'(t)\tau$ s'interprète comme un produit matrice-vecteur, à valeurs dans \mathbb{R}^p . Cette formulation se généralise directement au cas où l'espace de départ est multi-dimensionnel.

Soit donc $f : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$ une application.¹

2.1 Définition. L'application f est *dérivable* en x si il existe une application $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$ (linéaire) telle que

$$f(x + \xi) = f(x) + f'(x) \cdot \xi + o(\xi) \quad (2.1)$$

quand ξ tend vers 0. L'application $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$ s'appelle la *dérivée* de f en x ,² et, si $\xi \in \mathbb{R}^n$, le vecteur

$$f'(x) \cdot \xi = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + t\xi) - f(x)) \quad (2.2)$$

1. Rappelons pour la dernière fois que cette notation signifie que f est définie sur un certain voisinage (auquel on ne souhaite pas ici donner de nom) de x dans \mathbb{R}^n .

2. La dérivée s'appelle aussi la *différentielle*, le *jacobien*, l'*application linéaire tangente*, etc., et se note $df(x)$, $d_x f$, $Df(x)$, $Jf(x)$, $Tf(x)$, $f_*(x)$, etc. Le point de vue moderne est de ne pas changer de terminologie ni de notation, par rapport à la dérivée d'une fonction d'une variable, puisque celle-ci n'est qu'un cas particulier de la notion générale de dérivée. (Qui voudrait noter différemment, par exemple, l'addition des nombres rationnels et l'addition des nombres entiers ?) Dans le cas où $n = 1$, le produit matrice-vecteur $f'(x) \cdot \xi$ se réduit au produit d'un vecteur par un scalaire.

s'appelle la *dérivée directionnelle* de f en x dans la direction de ξ .

L'expression (2.2) montre que la dérivée, si elle existe, est unique. Mais attention, il ne suffit pas que la dérivée directionnelle de f existe dans toute direction pour que f soit dérivable (on pourra voir différentes pathologies possibles, propres à la dimension $n > 1$, dans l'exercice 2.h).

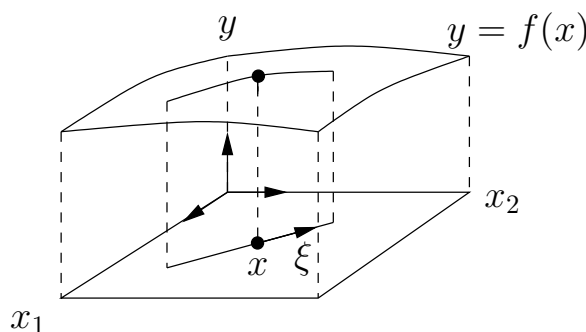


FIGURE 2.1 – Dérivée directionnelle : c'est la dérivée de la fonction le long du chemin $t \mapsto x + t\xi$

La notation $f'(x) \cdot \xi$ fait ressortir le fait que $f'(x)$ est une matrice, élément de $M_{p,n}(\mathbb{R})$, que l'on multiplie matriciellement par $\xi \in \mathbb{R}^n$, pour obtenir un vecteur de \mathbb{R}^p . La i -ième ligne de $f'(x)$ est la dérivée de f_i , la i -ième composante de f .

2.2 Exemple. Si f est une application linéaire,

$$f'(x) \cdot \xi = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + t\xi) - f(x)) = f(\xi).$$

La dérivée en un point x d'une fonction d'une variable $f : (\mathbb{R}, x) \rightarrow \mathbb{R}^p$ (chemin) était auparavant définie comme un *vecteur* de \mathbb{R}^p ,

$$f'(x) = \begin{pmatrix} f'_1(x) \\ \vdots \\ f'_p(x) \end{pmatrix}, \quad (2.3)$$

et non comme une application linéaire $\mathbb{R} \rightarrow \mathbb{R}^p$. Mais le vecteur (2.3) peut aussi bien être vu comme une matrice colonne $\in M_{n,1}(\mathbb{R})$, c'est-à-dire justement comme une application linéaire $\mathbb{R} \rightarrow \mathbb{R}^p$,

$$\tau \mapsto \begin{pmatrix} f'_1(x)\tau \\ \vdots \\ f'_p(x)\tau \end{pmatrix},$$

conformément à la définition 2.1.

2.3 Notation. Soient e_1, \dots, e_n les vecteurs de la base canonique de \mathbb{R}^n . La dérivée directionnelle de f dans la direction de e_i s'appelle la i -ième *dérivée partielle* de f en x et se note

$$\frac{\partial f}{\partial x_i}(x) = f'(x) \cdot e_i = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + te_i) - f(x));$$

c'est la dérivée du chemin de \mathbb{R}^p obtenu en figeant toutes les coordonnées de x sauf la i -ième. On la note aussi $\partial_{x_i} f(x)$ ou $\partial_i f(x)$.

Par exemple,

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \lim_{t \rightarrow 0} \frac{1}{t} (f(x_1 + t, x_2) - f(x_1, x_2)).$$

Attention, contrairement à ce que cette notation suggère, la i -ème dérivée partielle de f en x dépend de *toutes* les coordonnées utilisées, et non seulement de la i -ième. En effet, la courbe, paramétrée par x_i , le long de laquelle la i -ième dérivée partielle de f se calcule, a pour équation $x_j = cte_j$ ($j \neq i$). La forme de cette courbe dépend donc des coordonnées x_j (tandis que x_i ne détermine que son paramétrage). C'est une source d'erreur classique, qui a trompé même des mathématiciens tels que Lagrange ou Laplace au 18e siècle. Mais, au 21e siècle, cette erreur n'est plus tolérée dans une copie de L3!



2.4 Exemple. Considérons la fonction

$$f(x_1, x_2) = x_2$$

et le changement de coordonnées

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 + x_2 \end{pmatrix},$$

de sorte que

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \phi^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ -y_1 + y_2 \end{pmatrix}.$$

Posons $g = f \circ \phi^{-1}$, c'est-à-dire $g(y_1, y_2) = f(x_1, x_2) = x_2 = y_2 - y_1$ ³ :

$$\begin{array}{ccc} x & \xrightarrow{\phi} & y \\ & \searrow f & \swarrow g \\ & f(x) = g(y) & \end{array}$$

Alors

$$\partial_1 f(x) = 0 \neq \partial_1 g(y) = -1,$$

bien que $x_1 = y_1$. Géométriquement, on voit que ces deux dérivées n'ont rien à voir : elles sont les taux d'accroissement de la fonction f le long des chemin $x_2 = cte$ et $y_2 = x_1 + x_2 = cte$ respectivement. On peut faire remonter ce phénomène au fait qu'un élément e_i^* de la base duale (e_1^*, \dots, e_n^*) d'une base (e_1, \dots, e_n) d'un espace vectoriel dépend non seulement de e_i mais de tous les e_j .

La matrice de $f'(x)$ s'exprime de la façon suivante, dans les bases canoniques de \mathbb{R}^n et de \mathbb{R}^p .

3. Quand on écrit $g(y) = f(x)$, il est tacite que $y = \phi(x)$. Pour ainsi définir g , il faut que ϕ soit inversible.

2.5 Lemme. Si $f : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$ est dérivable en x ,

$$f'(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) = \begin{pmatrix} \partial_1 f_1(x) & \cdots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_p(x) & \cdots & \partial_n f_p(x) \end{pmatrix} \in M_{p,n}(\mathbb{R}).$$

On est ainsi ramené à des calculs de dérivées de fonctions réelles d'une variable réelle. Par exemple, si

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto \begin{pmatrix} \sin(x - y) \\ x + 2y \end{pmatrix},$$

sa dérivée en (x, y) est la matrice

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

Par ailleurs, il est commode de savoir écrire directement la formule développée :

$$f'(x) \cdot \xi = \sum_i \frac{\partial f}{\partial x_i}(x) \xi_i,$$

sans passer par la matrice. Cette formule est intuitive : la variation infinitésimale de f (« infinitésimal » signifie : la variation entre x et $x + \xi$, au premier ordre en ξ quand ξ tend vers 0) est la somme des ξ_i pondérés par les dérivées partielles de f .

Démonstration.

$$\begin{aligned} f'(x) \cdot \xi &= f'(x) \cdot \sum_i \xi_i e_i \\ &= \sum_i \xi_i f'(x) \cdot e_i && \text{(linéarité)} \\ &= \sum_i \xi_i \frac{\partial f}{\partial x_i}(x) && \text{(notation)} \\ &= \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} && \text{(produit matrice-vecteur).} \end{aligned}$$

La seconde égalité découle de (1.2). □

Par exemple, la dérivée en $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ de l'application

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_1^2 + x_2 \\ x_2^3 \end{pmatrix}$$

est

$$f'(x) = \begin{pmatrix} 1 & 0 \\ 2x_1 & 1 \\ 0 & 2x_2^2 \end{pmatrix} \in M_{3,2}(\mathbb{R}) = L(\mathbb{R}^2, \mathbb{R}^3).$$

Le lemme précédent connaît des généralisations évidentes. Par exemple, si $f : (\mathbb{R}^n \times \mathbb{R}^m, (x, y)) \rightarrow \mathbb{R}^p$ est dérivable, la dérivée de f se décompose naturellement en deux blocs horizontaux :

$$f'(x, y) = \left(\frac{\partial f}{\partial x}(x, y) \quad \frac{\partial f}{\partial y}(x, y) \right),$$

où par exemple le premier bloc $\frac{\partial f}{\partial x}(x, y) \in L(\mathbb{R}^n, \mathbb{R}^p)$ désigne la dérivée de f par rapport à la variable x , obtenue en figeant y .

2.6 Notation. Si f est une application linéaire $\mathbb{R}^n \rightarrow \mathbb{R}^p$, on a vu (exemple 2.2) qu'elle est égale à sa propre différentielle en tout point. Comme l'application linéaire $f'(x)$ ne dépend pas du point x où on la calcule, il est d'usage de noter alors df au lieu de $f'(x)$ ou de $df(x)$.⁴

En particulier avec les omniprésentes “applications coordonnées” $x_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $u \mapsto x_i(u) = u_i$,

$$dx_i : \mathbb{R}^n \rightarrow \mathbb{R}, \quad v \mapsto v_i$$

est la i -ième projection canonique $\mathbb{R}^n \rightarrow \mathbb{R}$,⁵ et

$$f'(x) = \partial_{x_1} f(x) dx_1 + \cdots + \partial_{x_n} f(x) dx_n.$$

Cette écriture est intuitive : la variation de f est la somme des contributions de dérivées partielles par rapport à chacune des variables.

2.7 Proposition (Règles de calcul). 1. La dérivation est linéaire : si $f, g : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$ sont dérivables en x et si $\alpha, \beta \in \mathbb{R}$,

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x).$$

2. Si $f : (\mathbb{R}^n, x) \rightarrow (\mathbb{R}^p, y)$ et $g : (\mathbb{R}^p, y) \rightarrow \mathbb{R}^q$ sont deux applications dérivables respectivement en x et en y , $g \circ f$ est dérivable en x et sa dérivée est le produit matriciel des dérivées de f et de g :

$$(g \circ f)'(x) = g'(y) \cdot f'(x). \tag{2.4}$$

3. Si $f : (\mathbb{R}^n, x) \rightarrow (\mathbb{R}^n, y)$ est localement inversible et si f et f^{-1} sont dérivables (respectivement en x et en y), $f'(x)$ est un automorphisme de \mathbb{R}^n et

$$(f^{-1})'(y) = f'(x)^{-1}.$$

Démonstration. 1. Trivial en utilisant l'unicité de la dérivée.

4. On aurait aussi pu penser à f' ...

5. La raison pour distinguer x_i de dx_i apparaîtra dans le chapitre sur les surfaces. En gros, l'argument de x_i est un “point” tandis que celui de dx_i est un “vecteur tangent”, choses qui ne seront alors plus les mêmes...

2. La formule découle du développement limité suivant :

$$\begin{aligned} g \circ f(x+h) &= g(f(x) + f'(x) \cdot h + o(h)) \\ &= g(y) + g'(y) \cdot (f'(x) \cdot h + o(h)) + o(f'(x) \cdot h + o(h)) \\ &= g(y) + g'(y) \cdot f'(x) \cdot h + o(h). \end{aligned}$$

3. La formule précédente appliquée à $f \circ f^{-1} = \text{id}$ et à $f^{-1} \circ f = \text{id}$ donne respectivement

$$f'(x) \cdot (f^{-1})'(y) = I \quad \text{et} \quad (f^{-1})'(y) \cdot f'(x) = I,$$

donc l'opérateur $(f^{-1})'(y)$ est inversible, d'inverse $f'(x)$.

□

Un cas particulier de la formule 2.4 est le suivant : Si $c : (\mathbb{R}, t) \rightarrow (\mathbb{R}^n, x)$ et $f : (\mathbb{R}^n, x) \rightarrow (\mathbb{R}^p, y)$ sont dérivables,

$$(f \circ c)'(t) = f'(x) \cdot c'(t).$$

Cette formule peut être utilisée pour définir de façon cinématique la dérivée de f dans la direction de $c'(t)$: $f'(x) \cdot c'(t)$ n'est autre que la vitesse du chemin $f \circ c$; il est ici important que cette quantité ne dépend de c qu'à travers $c'(t)$.

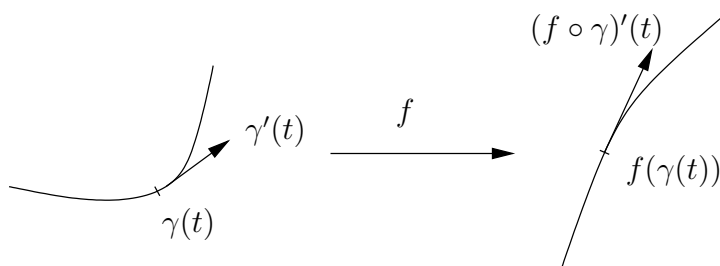


FIGURE 2.2 – Dérivée d'une application

2.a Exercice (Dérivée le long d'un chemin). Soient $c : (\mathbb{R}, 0) \rightarrow (\mathbb{R}^2, 0)$ et $f : (\mathbb{R}^2, 0) \rightarrow \mathbb{R}$ de classe C^∞ , tels que $c' = (\partial_2 f \circ c, -\partial_1 f \circ c)$. Montrer que la fonction $f \circ c$ est localement constante.

2.b Exercice (Équations de transport – méthode des caractéristiques (1)).

1. Trouver les fonctions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$, de classe C^1 vérifiant l'équation de transport

$$\partial_x f = \partial_t f;$$

on pourra utiliser le changement de variables $\varphi : (x, t) \mapsto (u, v) = (x + t, x - t)$.

2. Trouver les fonctions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$, de classe C^2 vérifiant l'équation de propagation des ondes

$$\partial_x^2 f = \partial_t^2 f;$$

on pourra utiliser le même changement de variables φ que dans la question précédente.

2.c Exercice (Une équation linéaire – méthode des caractéristiques (2)).

Soient $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction nulle en 0 et $f_0 : \mathbb{R} \rightarrow \mathbb{R}$ une fonction paire, toutes deux de classe C^1 .

1. Trouver l'expression de la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 telle que

$$-y \partial_x f(x, y) + x \partial_y f(x, y) = g(x, y) \quad \text{et} \quad f(x, 0) = f_0(x) \quad (\forall x, y),$$

en supposant qu'elle existe; on pourra chercher f en coordonnées polaires (r, θ) (avec $x = r \cos \theta$, $y = r \sin \theta$).

2. Cette solution existe-t-elle si $g(x, y) = 1$ (pour tous x, y) ?

3. Si $g(x, y) = xy$ (sans omettre de vérifier le caractère C^1 de f) ?

2.d Exercice (Application bilinéaire).

Soit $B : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ une application bilinéaire.

1. Montrer qu'il existe $C \geq 0$ tel que

$$\|B(x, y)\| \leq C \|x\| \|y\|;$$

la plus petite des constantes C pour cette inégalité est appelée la *norme* de B , i.e.

$$\|B\| = \sup_{\|x\|=1, \|y\|=1} \|B(x, y)\|.$$

2. En déduire la dérivée de B .

3. En déduire la dérivée de la composition des applications linéaires.

4. Soit $c : (\mathbb{R}, 0) \rightarrow (M_n(\mathbb{R}), I_n)$ un chemin dérivable de matrices isométriques, c'est-à-dire tracé sur

$$O_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}), {}^t M M = I\}.$$

Montrer que le vecteur vitesse $c'(0)$ est une matrice antisymétrique. Dans le cas $n = 3$, définir le *vecteur rotation* associé à $c'(0)$ (ceci utilise le *produit vectoriel*).

2.e Exercice. Calculer la dérivée des applications suivantes (en un point quelconque) :

1. $M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}), M \mapsto M^2$

2. $M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}), M \mapsto {}^t M M$.

3. $GL_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}), M \mapsto M^{-1}$ (on commencera par montrer que $GL_n(\mathbb{R})$ est ouvert dans $M_n(\mathbb{R})$)

4. $M_n(\mathbb{R}) \rightarrow \mathbb{R}, M \mapsto \det M$ (on pourra commencer par les matrices inversibles).

Un vecteur étant plus facile à représenter qu'une forme linéaire, on considère souvent le gradient, défini dans l'exercice suivant.

2.f Exercice (Gradient d'une fonction).

Soit $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}$ une fonction dérivable en a . Le *gradient* de f en a est le vecteur $\text{grad } f(a) \in \mathbb{R}^n$ défini par

$$\langle \text{grad } f(a) | \xi \rangle = f'(a) \cdot \xi \quad (\forall \xi \in \mathbb{R}^n),$$

où $\langle \cdot | \cdot \rangle$ est un produit scalaire (euclidien) de \mathbb{R}^n .

1. Justifier l'existence du gradient.

2. Montrer que $\text{grad } f(a)$ pointe dans la “direction de plus grande pente” de f , c'est-à-dire dans la direction dans laquelle il faut se déplacer pour que la valeur de f augmente le plus vite possible.

3. Dans le cas où $n = 2$, calculer les composantes de $\text{grad } f(a)$ pour le produit scalaire standard

$$\langle \xi | \eta \rangle = \xi_1 \eta_1 + \xi_2 \eta_2.$$

Soient

$$\rho : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad p = (r, \theta) \mapsto (x_1, x_2) = (r \cos \theta, r \sin \theta)$$

l'application “coordonnées polaires”, p tel que $\rho(p) = a$, et

$$F = f \circ \rho : (\mathbb{R}^2, p) \rightarrow \mathbb{R}, \quad (r, \theta) \mapsto f(r \cos \theta, r \sin \theta);$$

F est la fonction f “lue dans les coordonnées polaires”.

Soit $\langle \cdot | \cdot \rangle_p$ le produit scalaire induit en p par les coordonnées polaires, à savoir

$$\langle P | Q \rangle_p = \langle \rho'(p) \cdot P, \rho'(p) \cdot Q \rangle. \quad 6$$

4. Calculer les composantes du gradient de F en p relativement au produit scalaire $\langle \cdot | \cdot \rangle_p$, dans la base canonique, puis dans une base orthonormée.

2.g Exercice (\mathbb{C} -dérivabilité).

Soit $f : (\mathbb{C}, z) \rightarrow \mathbb{C}$. Montrer que les trois propriétés suivantes sont équivalentes :

1. f est \mathbb{C} -dérivable en z , c'est-à-dire que

$$f'(z) := \lim_{\zeta \rightarrow 0, \zeta \in \mathbb{C}_*} \frac{f(z + \zeta) - f(z)}{\zeta}$$

existe dans \mathbb{C}

2. f est différentiable en z (vue comme une fonction $(\mathbb{R}^2, z) \rightarrow \mathbb{R}^2$) et satisfait l'équation de Cauchy-Riemann :

$$\frac{\partial f}{\partial y}(z) = i \frac{\partial f}{\partial x}(z).$$

3. f est conforme, c'est-à-dire conserve les angles : si $f'(z) \neq 0$ et si $\zeta_1, \zeta_2 \in \mathbb{C}$, l'angle entre ζ_1 et ζ_2 est le même qu'entre $f'(z) \cdot \zeta_1$ et $f'(z) \cdot \zeta_2$.

En Analyse complexe, on montre que, si f est \mathbb{C} -dérivable, elle est localement développable en série entière, donc de classe C^∞ [Rud74].

6. La collection des produits scalaires ainsi obtenus en faisant varier le point p s'appelle une *métrique riemannienne*.

Exercices complémentaires

2.h Exercice (Dérivabilité directionnelle et dérivabilité).

1. Soit $f : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$. Se convaincre que les propriétés suivantes vont par force croissante :

1. f possède une dérivée dans toutes les directions en x , i.e. pour tout $\xi \in \mathbb{R}^n$ la limite

$$f'(x) \cdot \xi = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + t\xi) - f(x))$$

existe

2. f possède une dérivée dans toutes les directions en x et l'application $f'(x) : \xi \mapsto f'(x) \cdot \xi$ est linéaire⁷
3. f est dérivable en x ⁸

et que la dernière implique la continuité de f en x .

2. (Discontinuité) Trouver une fonction f discontinue en x mais possédant une dérivée directionnelle dans toutes les directions.

3. Montrer que si, pour une fonction $f : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$ et un vecteur $v \in \mathbb{R}^n$, $f'(x) \cdot v$ existe, pour tout $\lambda \in \mathbb{R}$ on a

$$f'(x) \cdot (\lambda v) = \lambda f'(x) \cdot v.$$

4. (Non-linéarité) Trouver une fonction telle que $f'(x)$ soit définie sur tout \mathbb{R}^n mais ne soit pas linéaire.

5. Montrer que, si f est différentiable en x , pour tout chemin $c : (\mathbb{R}, 0) \rightarrow (\mathbb{R}^n, x)$ dérivable,

$$(f \circ c)'(0) = f'(x) \cdot c'(0)$$

(ceci est une première instance de la formule de dérivation des fonctions composées).

6. (Non-uniformité) En déduire une fonction telle que $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$, mais telle que l'on n'ait pas

$$f(x + \xi) = f(x) + f'(x) \cdot \xi + o(\xi)$$

quand ξ tend vers 0.

Indication : À défaut d'imaginer soi-même les trois exemples demandés, on pourra donner des formules pour les fonctions dont les graphes sont tracés sur la figure 2.3.

7. On dit alors que f est *Gateaux-dérivable* (sans accents circonflexe, attention les francophones!). René Eugène GATEAUX (1889–1914) est un mathématicien français mort au combat à 25 ans le 3 octobre 1914. Ses travaux en analyse fonctionnelle, notamment sur l'intégration en dimension infinie, furent le point de départ de la construction la mesure de Wiener dans l'étude du mouvement brownien.

8. Parfois on précise alors *Fréchet-dérivable*, par opposition à *Gateaux-dérivable*. Maurice FRÉCHET est un mathématicien français (1878–1973). Remarquablement prolifique, il est célèbre notamment pour ses découvertes en topologie, en probabilités et en analyse fonctionnelle.

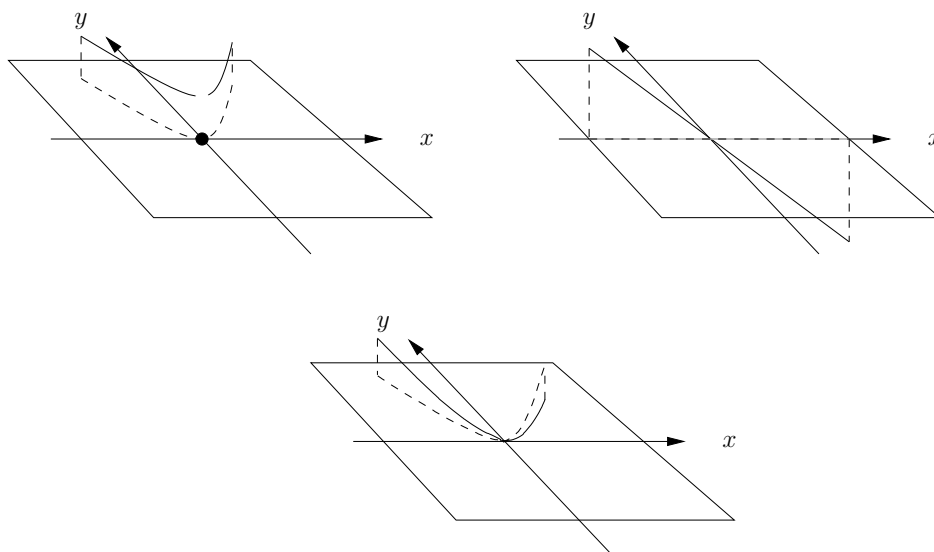


FIGURE 2.3 – Diverses raisons de ne pas être dérivable

2.i Exercice (Identité d'Euler).

Soient $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ une fonction dérivable et $k \geq 1$. Montrer que les deux propriétés suivantes sont équivalentes :

1. f est homogène de degré k : $f(tx) = t^k f(x)$ ($\forall x \in \mathbb{R}^n \setminus \{0\}, \forall t > 0$)
2. Identité d'Euler : $f'(x) \cdot x = kf(x)$ ($\forall x \in \mathbb{R}^n \setminus \{0\}$).

2.j Exercice (F. Laudenbach).

1. Soient (D_1, D_2, D_3) et (D'_1, D'_2, D'_3) deux familles de droites vectorielles de \mathbb{R}^2 . Montrer qu'il existe une application linéaire (donc dérivable) de \mathbb{R}^2 dans lui-même envoyant chaque D_i sur D'_i .

2. Soient maintenant (D_1, \dots, D_4) et (D'_1, \dots, D'_4) deux familles de quatre droites vectorielles de \mathbb{R}^2 . Montrer qu'il existe une application dérivable $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ inversible telle que $f(D_i) = D'_i$ si et seulement si le birapport des deux familles est le même; le *birapport* de quatre nombres complexes z_1, \dots, z_4 deux à deux distincts est

$$(z_1, \dots, z_4) = \frac{z_3 - z_1}{z_3 - z_2} \times \frac{z_4 - z_2}{z_4 - z_1},$$

et le birapport de quatre droites vectorielles de \mathbb{R}^2 est le birapport de leurs pentes.

Chapitre 3

Classe C^1

Mots-clefs du chapitre Classe C^k , formule de la moyenne, inégalité des accroissements finis, formule de Liouville

Soit $f : U \text{ ouvert } \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. Rappelons qu'on note $C^0(\mathbb{R}^n, \mathbb{R}^p)$ l'espace vectoriel des telles applications qui sont continues.

3.1 Définition. L'application f est *continûment dérivable* ou de *classe C^1* , ce qu'on note $f \in C^1(U, \mathbb{R}^p)$, si elle est dérivable en tout point de U et si sa dérivée $f' : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^p)$ est continue. Elle est de *classe C^k* si $f' : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^p)$ est de classe C^{k-1} . Elle est de classe C^∞ si elle est de classe C^k pour tout $k \in \mathbb{N}$.

3.2 Lemme. Si $f : U \text{ ouvert } \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $g : V \text{ ouvert } \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ sont de classe C^1 , $g \circ f$ est de classe C^1 sur $U \cap g^{-1}(V)$.

Démonstration. En tout point $x \in U_1 := U \cap g^{-1}(V)$, $g \circ f$ est dérivable de dérivée

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x).$$

Cette dérivée est donc la composée de l'application continue

$$U_1 \rightarrow L(\mathbb{R}^p, \mathbb{R}^q) \times L(\mathbb{R}^n, \mathbb{R}^p), \quad x \mapsto (g'(f(x)), f'(x))$$

et du produit matriciel (bilinéaire, donc continu, cf. l'exercice 2.d)

$$L(\mathbb{R}^p, \mathbb{R}^q) \times L(\mathbb{R}^n, \mathbb{R}^p), \quad (b, a) \mapsto b \cdot a,$$

donc est continue. □

3.3 Théorème (Formule de la moyenne). Si f est de classe C^1 sur U et si $c : [\alpha, \beta] \rightarrow U$ est un chemin de classe C^1 ,

$$f(c(\alpha)) - f(c(\beta)) = \int_{\alpha}^{\beta} f'(c(t)) \cdot c'(t) dt.$$

Démonstration. D'après la seconde formule fondamentale du calcul différentiel à une variable,

$$f(c(\alpha)) - f(c(\beta)) = \int_{\alpha}^{\beta} \frac{d}{dt} f(c(t)) dt,$$

où

$$\frac{d}{dt}f(c(t)) = f'(c(t)) \cdot c'(t).$$

□

Rappelons que, si $c : [\alpha, \beta] \rightarrow \mathbb{R}^n$ est un chemin de classe C^1 , sa *longueur* est

$$\ell(c) = \int_{\alpha}^{\beta} \|c'(t)\| dt.$$

Aussi, pour $M \in L(\mathbb{R}^n, \mathbb{R}^p)$, on note $\|M\|$ la norme d'opérateur de M , subordonnée aux normes euclidiennes sur \mathbb{R}^n et \mathbb{R}^p :

$$\|M\| = \max_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Mx\|_{\mathbb{R}^p}}{\|x\|_{\mathbb{R}^n}}.$$

(Exercice : Montrer que la “norme subordonnée” est bien une norme sur $L(\mathbb{R}^n, \mathbb{R}^p)$, et définir de même une norme subordonnée pour les applications bilinéaires.)

3.4 Corollaire (Inégalité des accroissements finis). *Si $f : U$ ouvert $\subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ est de classe C^1 ,*

$$\|f(y) - f(x)\| \leq \ell(c) \sup_{[\alpha, \beta]} \|f' \circ c\|,$$

pour tout chemin c de classe C^1 joignant x à y dans U .

En particulier, pour tous $x, y \in U$ tels que $[x, y] \in U$,

$$\|f(y) - f(x)\| \leq \|y - x\| \sup_{[x, y]} \|f'\|.$$

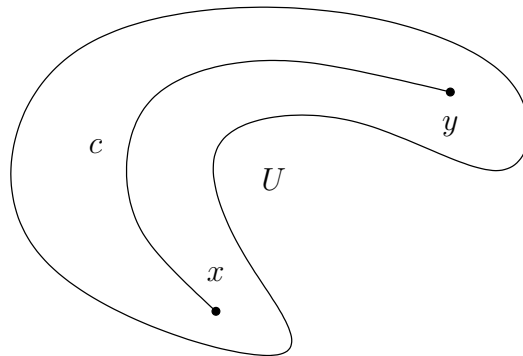


FIGURE 3.1 – L’inégalité des accroissements finis majore la variation de f entre x et y

Démonstration. On a

$$\begin{aligned} \|f(y) - f(x)\| &= \left\| \int_0^1 f'(c(t)) \cdot c'(t) dt \right\| \\ &\leq \int_0^1 \|f'(c(t))\| \|c'(t)\| dt, \end{aligned}$$

d’où la conclusion. Le cas particulier s’obtient en choisissant $c(t) = (1 - t)x + ty$, $0 \leq t \leq 1$. □

L'inégalité des accroissements finis est en fait vérifiée sous des hypothèses légèrement plus larges.

3.a Exercice (Inégalité des accroissements finis).

Soient $F : [\alpha, \beta] \rightarrow \mathbb{R}^p$ et $\varphi : [\alpha, \beta] \rightarrow \mathbb{R}$ deux applications continues, dérivables sur $] \alpha, \beta [$, telles que

$$\|F'(t)\| \leq \varphi'(t)$$

sur $] \alpha, \beta [$. Montrer que

$$\|F(\beta) - F(\alpha)\| \leq \varphi(\beta) - \varphi(\alpha).$$

3.b Exercice. Montrer qu'une fonction $f : B(a, r) = \{x \in \mathbb{R}^n, \|x - a\| < r\} \rightarrow \mathbb{R}^p$ est constante si et seulement si $f' \equiv 0$.

3.5 Corollaire. Soit $f : U$ ouvert $\subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. Les propriétés suivantes sont équivalentes :

1. f est de classe C^1
2. les dérivées partielles $\partial_{x_i} f(x) = \lim_{t \rightarrow 0} \frac{1}{t}(f(x + te_i) - f(x))$ existent et sont continues.

Démonstration. L'implication directe est triviale, puisque, pour tout $i = 1, \dots, n$,

$$\partial_{x_i} f(x) = f'(x) \cdot e_i.$$

Réciproquement, soit $x \in U$. Notons $L(x)$ l'application linéaire de \mathbb{R}^n dans \mathbb{R}^p définie par

$$L(x) \cdot \xi = \partial_{x_1} f(x) \xi_1 + \dots + \partial_{x_n} f(x) \xi_n.$$

Il n'y a qu'à montrer que

$$f(x + \xi) = f(x) + L(x) \cdot \xi + o(\xi),$$

puisque $L(x)$ dépend continûment de x par hypothèse. Pour un vecteur ξ donné assez petit, soit $c : [0, 1] \rightarrow U$, $t \mapsto f(x + t\xi)$.

Or, c est dérivable, de dérivée

$$c'(t) = L(x + t\xi) \cdot \xi.$$

En effet, la quantité

$$\frac{1}{\tau} (f(x + (t + \tau)\xi) - f(x + t\xi))$$

peut se décomposer comme

$$\sum_{1 \leq i \leq n} \frac{1}{\tau} (f(x + t\xi + \tau(\xi_1, \dots, \xi_i, 0, \dots, 0)) - f(x + t\xi + \tau(\xi_1, \dots, \xi_{i-1}, 0, \dots, 0)))$$

(de façon à ne faire apparaître que des variations parallèles aux axes de coordonnées), qui, par hypothèse quand τ tend vers 0, tend vers

$$c'(t) = \sum_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(x + t\xi) \xi_i = L(x + t\xi) \cdot \xi.$$

D'après cette expression de $c'(t)$, c' est continue, donc c est de classe C^1 .

D'après la formule de la moyenne,

$$f(x + \xi) - f(x) = \int_0^1 L(x + t\xi) \cdot \xi dt.$$

Donc

$$\|f(x + \xi) - f(x) - L(x) \cdot \xi\| \leq \int_0^1 \|L(x + t\xi) - L(x)\| dt \|\xi\|.$$

Or, cette dernière intégrale tend vers 0 avec t , puisque $t \mapsto L(x + t\xi)$ est continue sur l'intervalle compact $[0, 1]$, donc uniformément continue. Donc

$$\|f(x + \xi) - f(x) - L(x) \cdot \xi\| = o(\xi)$$

comme voulu. □

Dans l'exercice suivant, on démontre le théorème de fonction implicite par récurrence sur la dimension de l'équation. Une démonstration plus géométrique sera donnée aux chapitres 5–9 (cette dernière passera facilement à la dimension infinie, chapitre 26).

3.c Exercice (Fonction implicite).

Soit

$$f : (\mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^p, (0, 0)) \rightarrow (\mathbb{R}, 0), \quad (x, y) \mapsto f(x, y)$$

une fonction de classe C^1 telle que

$$\det \frac{\partial f}{\partial y}(0, 0) \neq 0.$$

On veut montrer qu'il existe une unique fonction

$$\varphi : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^p, 0)$$

de classe C^1 telle que, localement au voisinage de $(0, 0)$ dans $\mathbb{R}^n \times \mathbb{R}^p$,

$$f(x, y) = 0 \Leftrightarrow y = \varphi(x). \tag{3.1}$$

1. Le montrer dans le cas où $p = 1$.

Rappel. Soit $M \in M_p(\mathbb{R})$. La méthode du pivot de Gauss montre qu'il existe une suite d'opérations sur les lignes (permutations $\ell'_i = \ell_{\sigma_i} \forall i$, dilatations $\ell'_i = \lambda \ell_i$ et transvections $\ell'_i = \ell_i + \lambda \ell_j$) qui mette M sous forme triangulaire supérieure S . Ces opérations peuvent être vues comme autant de multiplications à gauche de M par des matrices (resp. de permutation, de dilatation ou de transvection) :

$$A_1 \cdots A_k M = U,$$

soit

$$M = (A_k^{-1} \cdots A_1^{-1}) U.$$

2. Conclure par récurrence sur p .

3.d Exercice (Formule de Liouville).

1. Soient $A : \mathbb{R} \rightarrow M_n(\mathbb{R})$ une fonction continue, et $x_1, \dots, x_n : \mathbb{R} \rightarrow \mathbb{R}^n$ des solutions de l'équation différentielle linéaire

$$x'(t) = A(t)x(t).$$

Montrer que le (*déterminant*) *wronskien* (aussi appelé (*déterminant*) *jacobien*)

$$w(t) = \det(x_1(t), \dots, x_n(t))$$

vérifie l'équation différentielle linéaire

$$w'(t) = \operatorname{tr} A(t) w(t);$$

on rappelle que la dérivée de l'application déterminant (exercice 2.e) est

$$(\det)'(A) \cdot H = \operatorname{tr} ({}^t \bar{A} H).$$

2. Soient maintenant $v : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 et (φ_t) son flot, c'est-à-dire que, pour tout x proche de a , $t \mapsto \varphi_t(x)$ est l'unique solution maximale du problème de Cauchy

$$\frac{d}{dt}(\varphi_t(x)) = v(\varphi_t(x)), \quad \varphi_0(x) = x.$$

Montrer que

$$\left. \frac{d}{dt} \det(\varphi'_t)(x) \right|_{t=0} = \operatorname{div} v(x),$$

avec

$$\operatorname{div} v(x) = \operatorname{tr} v'(x) = \sum_{1 \leq i \leq n} \frac{\partial v_i}{\partial x_i}(x)$$

(*divergence* de v en x).

3. À quelle condition sur son champ de vecteurs vitesse un fluide est-il incompressible?

Chapitre 4

Le théorème du point fixe

Mots-clefs du chapitre Application lipschitzienne, application contractante, point fixe

Soient (X, d_X) et (Y, d_Y) deux espaces métriques. Par exemple, X et Y peuvent être des ouverts de \mathbb{R}^n , munis de la distance euclidienne $\|x - x'\|$.

Soit $v : X \rightarrow Y$. Dire que v est continue en $x \in X$, c'est dire qu'il existe une fonction $\eta(\epsilon) > 0$ qui tend vers 0 quand ϵ tend vers 0, telle que que

$$\forall \epsilon > 0 \forall x, x' \quad d_X(x, x') < \eta(\epsilon) \Rightarrow d_Y(v(x), v(x')) < \epsilon.$$

Si v est « peu » continue, η doit tendre rapidement vers 0 en fonction de ϵ , par exemple comme ϵ^{36} ou $e^{-1/\epsilon}$. Si on peut choisir $\eta = C\epsilon$, cela correspond à une forme de continuité forte, qui s'appelle la continuité lipschitzienne.

4.1 Définition. Le rapport de Lipschitz¹ v est

$$\text{lip } v = \sup_{x, x' \in X, x \neq x'} \frac{d_Y(v(x), v(x'))}{d_X(x, x')} \in [0, +\infty].$$

L'application v est

- *lipschitzienne* si $\text{lip } v < \infty$
- *contractante* si $\text{lip } v \leq 1$
- *strictement contractante* si $\text{lip } v < 1$.

(Le champ de vecteurs v est *localement lipschitzien* si, pour tout $a \in X$, il existe un voisinage V de a dans X (ou, de façon équivalente, une boule $B(a, r)$ $r > 0$) tel que $\text{lip } v|_V < +\infty$. Etc.)

4.2 Exemples. – Une application linéaire $v : \mathbb{R}^n \hookrightarrow \mathbb{R}^m$ est lipschitzienne, et

$$\text{lip } v = \|v\|$$

(le montrer, en utilisant la compacité de la sphère unité en dimension finie).

1. Rudolph Otto Sigismund LIPSCHITZ, mathématicien allemand (1832–1903). Il travailla dans beaucoup de domaines mathématiques, et découvrit en particulier la version moderne du théorème d'existence et d'unicité locales de solutions d'équations différentielles ordinaires.

– Une isométrie de \mathbb{R}^n (soit une application $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dans lui-même qui préserve la distance euclidienne : $d(v(x), v(x')) = d(x, x')$ pour tous x, x') est une contraction.

4.3 Remarques. – Si $f : X \rightarrow Y$ et $g : Y \rightarrow Z$,

$$\frac{d_Z(g \circ f(x), g \circ f(x'))}{d_X(x, x')} = \frac{d_Z(g \circ f(x), g \circ f(x'))}{d_Y(f(x), f(x'))} \frac{d_Y(f(x), f(x'))}{d_X(x, x')},$$

donc

$$\text{lip}(g \circ f) \leq \text{lip } g \text{ lip } f.$$

– Si $f, g : X \rightarrow \mathbb{R}^n$,

$$\text{lip}(f + g) \leq \text{lip } f + \text{lip } g.$$

Le corollaire 3.4 se reformule dans les termes suivants.

4.4 Proposition (Inégalité des accroissements finis). *Si une application $f : U$ ouvert convexe de $\mathbb{R}^n \rightarrow \mathbb{R}^p$ est de classe C^1 ,*

$$\text{lip } f \leq \sup_{x \in U} \|f'(x)\|.$$

Comme précédemment, $\|f'(x)\|$ désigne la norme d'opérateur de $f'(x)$, soit aussi sa constante de Lipschitz.

4.a Exercice (Théorème du point fixe, Peano² [Pea88], Picard³ [Pic90]).

Soient (X, d) un espace métrique complet et $f : X \rightarrow X$ une contraction stricte.

1. Montrer que f possède un unique point fixe $a \in X$ et que, pour tout $x \in X$, la suite des images itérées de x par f converge géométriquement vers a .

2. Montrer que la convergence est localement uniforme et, si X est borné, uniforme.

Supposons maintenant que $f = f_\lambda$ dépende d'un paramètre $\lambda \in \Lambda$, où Λ est un ouvert de \mathbb{R}^n , et que la famille (f_λ) soit uniformément strictement contractante (il existe α tel que $\text{lip } f_\lambda \leq \alpha < 1$ quel que soit λ). D'après ce qui précède il existe une application $a : \Lambda \rightarrow X$, $\lambda \mapsto a_\lambda =$ l'unique point fixe de f_λ .

3. Montrer que, si $\lambda \mapsto f_\lambda(x)$ est continue pour tout $x \in X$, a est continue.

4. Montrer que, si $\lambda \mapsto f_\lambda(x)$ est localement lipschitzienne pour tout x , $\lambda \mapsto a_\lambda$ est localement lipschitzienne.

5. Montrer que, si $X = \mathbb{R}^p$, si Λ est un ouvert d'espace de Banach et si f est de classe C^r , $\lambda \mapsto a_\lambda$ est de classe C^r .

Il existe beaucoup d'autres théorèmes de point fixe. Par exemple, le *théorème de Brouwer* affirme qu'une application continue de la boule fermée dans elle-même

2. Giuseppe PEANO, mathématicien et linguiste italien (1858–1932, qui axiomatisa l'arithmétique)

3. Charles Émile PICARD, mathématicien français (1856–1941), célèbre pour ses travaux sur les singularités de fonctions holomorphes et sur les équations intégro-différentielles.

possède un point fixe. Il généralise le théorème des valeurs intermédiaires (qui est équivalent au fait qu'une fonction continue $f : [0, 1] \hookrightarrow [0, 1]$ possède un point fixe). Nous n'utiliserons pas ces théorèmes de nature topologique, ni a fortiori leurs généralisations en dimension infinie (*théorème de Schauder*), malgré leur importance en analyse fonctionnelle.

4.b Exercice (Points fixes attractifs ou répulsifs).

Soit $f : (\mathbb{R}, a) \rightarrow (\mathbb{R}, a)$ une application de classe C^2 ayant a comme point fixe. Comme en Théorie des Systèmes dynamiques, on notera ici f^n la composée de f n fois avec elle-même : $f^0(x) = x$ et, pour tout $n \in \mathbb{N}$, $f^{n+1}(x) = f(f^n(x))$.

1. Supposons que a est *attractif* : $|f'(a)| < 1$. Montrer que, si x est assez proche de a , la suite $(f^n(x))$ est définie sur \mathbb{N} et, si $f'(a) \neq 0$, converge vers a géométriquement :

$$f^{n+1}(x) - a \sim f'(a)(f^n(x) - a) \quad \text{quand } n \rightarrow +\infty.$$

2. Supposons que a est même *surattractif* : $|f'(a)| = 0$. La question précédente montre que, si x est assez proche de a , la suite $(f^n(x))$ est définie sur \mathbb{N} et converge vers a . Montrer que, si $f''(a) \neq 0$, la convergence est quadratique :

$$f^{n+1}(x) - a \sim \frac{f''(a)}{2}(f^n(x) - a)^2 \quad \text{quand } n \rightarrow +\infty.$$

3. Supposons au contraire que a est *répulsif* : $|f'(a)| > 1$. Montrer qu'il existe un voisinage V de a tel que quel que soit $x \in V \setminus \{a\}$, il existe $n \in \mathbb{N}$ tel que $f^n(x) \notin V$.

4. Comment calculer les décimales du *nombre d'or* $a = \frac{1+\sqrt{5}}{2}$ à ϵ près ? *Indication* : Le nombre a est un point fixe attractif de l'application $f : [0, +\infty[$, $x \mapsto \sqrt{x+1}$.

Les choses se compliquent si l'on passe d'une récurrence $x_{n+1} = f(x_n)$ à une récurrence « dépendant du temps n » $x_{n+1} = f(x_n, n)$. On est alors obligé de passer en dimension infinie, en raisonnant dans un espace de suites. L'exercice suivant en donne un exemple.

4.c Exercice (J.-C. Yoccoz).

Soit (x_n) une suite réelle définie par $x_0 \in \mathbb{R}$ et

$$x_{n+1} = x_n^2 - 100 + \sin n \quad (\forall n \in \mathbb{N}).$$

On veut montrer qu'il existe un unique $x_0 \in \mathbb{R}$ tel que (x_n) soit bornée à valeurs ≥ 0 .

1. Montrer que, si (x_n) est bornée, pour tout n on a $10 \leq x_n \leq 11$.

Soient X l'ensemble des suites réelles (y_n) de $[10, 11]$, d la distance

$$d((y_n), (z_n)) = \sup_n |y_n - z_n|$$

et F l'opérateur

$$F : X \hookrightarrow X, \quad (y_n) \mapsto (z_n), \quad z_n = \sqrt{100 - \sin n + y_{n+1}}.$$

2. Vérifier que F est bien défini. Quelle est son rapport de Lipschitz ?

3. Conclure. On donnera une estimation de l'unique point fixe (a_n) de F .

L'une des applications les plus célèbres du théorème du point fixe est la démonstration (par Lipschitz) de l'existence et de l'unicité des solutions d'équations différentielles, donnée ci-dessous en exercice. L'autre application que nous en ferons est la démonstration du théorème d'inversion, dans le chapitre suivant.

4.d Exercice (Théorème de Cauchy-Lipschitz).

1. Montrer que le problème de Cauchy

$$\begin{cases} \gamma'(t) = \gamma^{1/3}(t) \\ \gamma(0) = 0 \end{cases}$$

possède plusieurs solutions sur \mathbb{R} (une infinité, même).

Soient $E = \mathbb{R}^d$, U un ouvert de E et v un champ de vecteurs⁴ sur U localement lipschitzien.

Considérons, pour $\chi > 0$ et $k > 0$, la norme $\|\cdot\|_\chi$ de $C^0([-T, T], E)$ définie par

$$\|x\|_\chi = \max_{t \in [-T, T]} e^{-\chi k |t|} \|x(t)\|$$

(cette norme écrase la croissance exponentielle des trajectoires en fonction du temps); le paramètre k sera choisi plus tard. En particulier, $\|\cdot\|_0$ est la norme C^0 .

2. Montrer que toutes les χ -normes sont équivalentes.

3. Montrer que $\|\cdot\|_\chi$ fait de $C^0([-T, T], E)$ un espace de Banach.

Supposons d'abord que $U = E$ et que v est lipschitzien sur E tout entier, et notons $k = \text{lip } v$. Soit $T > 0$ et soit $\phi : C^0([-T, T], E) \rightarrow C^0([-T, T], E)$ l'application qui, à un chemin continu $x \in C^0([-T, T], E)$, associe le chemin $\phi x \in C^0([-T, T], E)$ défini par

$$(\phi x)(t) = a + \int_0^t v(x(s)) ds.$$

Les courbes intégrales de v passant par a à $t = 0$ sont les points fixes de ϕ .

4. Montrer que, pour tout point $a \in U$ il existe un réel $T > 0$ tel qu'il existe une unique courbe $\gamma_a : [-T, T] \rightarrow U$ dérivable, solution du problème de Cauchy

$$\begin{cases} \gamma'(t) = v(\gamma(t)) \\ \gamma(0) = a. \end{cases} \quad (4.1)$$

5. En déduire le cas général, où U est un ouvert quelconque et v est seulement localement lipschitzien.

6. Si x est la solution cherchée et si x_0 est la valeur approchée initiale, montrer que, pour tout $n \in \mathbb{N}$ on a

$$\|x - \phi^n x_0\|_0 \leq \left(\frac{ekT}{n} \right)^n \frac{n}{n - kT} \|x_0 - \phi x_0\|_0.$$

4. Une *champ de vecteurs* sur U est une application $U \rightarrow E$.

7. Calculer $\phi^n 0$ dans le cas où $d = 1$, $v(x) = x$ et $a = 1$.
8. Si v est (globalement) lipschitzien sur E , montrer que la solution maximale du problème de Cauchy est définie sur \mathbb{R} .
9. Si le champ de vecteurs v dépend du temps, quelle régularité de v suffit-elle?
10. Que peut-on dire dans le cas où E est un espace de Banach?

Chapitre 5

Inversion locale

Mots-clefs du chapitre Lipéomorphisme, difféomorphisme

Le théorème d'inversion locale affirme, sous certaines hypothèses sur une application f , qu'il existe une unique application f^{-1} telle que, localement,

$$y = f(x) \Leftrightarrow x = f^{-1}(y).$$

Le théorème des fonctions implicites, son frère jumeau, affirme que, sous certaines hypothèses sur une application $F(\cdot, \cdot)$, il existe une unique application φ telle que, localement,

$$F(x, y) = 0 \Leftrightarrow y = \varphi(x)$$

c'est-à-dire que, localement,

$$F(x, \varphi(x)) \equiv 0$$

et que $\varphi(x)$ est l'unique solution de cette équation. Ces deux théorèmes se déduisent facilement l'un de l'autre (dans l'un des deux sens, poser $F(x, y) = (x, f(x))$). Ils émergent des travaux : d'I. Newton et G. Leibniz, pour qui la question était moins de démontrer ce théorème que de calculer la fonction inverse ; de J.-L. Lagrange,¹ qui démontra une version particulière du théorème utile dans la détermination de la position de la Terre en fonction du temps, dans le Système solaire ; d'A. Cauchy, qui démontra rigoureusement le théorème pour la première fois, pour les fonctions analytiques ; d'U. Dini,² qui démontra la version 5.2 ci-dessous, etc.

5.1 Définition. Soit $f : U$ ouvert de $\mathbb{R}^n \rightarrow V$ ouvert de \mathbb{R}^p une bijection. Elle est

- un *homéomorphisme* si f et f^{-1} sont continues
- un *lipéomorphisme* si f et f^{-1} sont lipschitziennes
- un *difféomorphisme* si f et f^{-1} sont de classe C^1
- un C^k -*difféomorphisme* si f et f^{-1} sont de classe C^k ($k \in \mathbb{N}_* \cup \{\infty\}$)
- un *isomorphisme analytique* si f et f^{-1} sont analytiques réelles.

5.2 Théorème (Inversion locale). *Une application $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^p$ de classe C^1 est un difféomorphisme local si et seulement si $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est inversible (donc $n = p$).*

1. Joseph L. LAGRANGE, mathématicien et astronome italo-français (1736–1813). Voir <http://www.breves-de-maths.fr/joseph-louis-lagrange/>

2. Ulisse DINI, mathématicien et homme politique italien (1845–1918)

Rappelons qu'une application $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^p$ de classe C^1 est un difféomorphisme local, par définition, si et seulement si il existe un voisinage ouvert U de a , assez petit, tel que la restriction $f : U \rightarrow f(U)$ soit un difféomorphisme.

Démonstration du sens facile. Si $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ est un C^1 -difféomorphisme local, $f^{-1} \circ f = \text{id}$ et $f \circ f^{-1} = \text{id}$ au voisinage respectivement de a et de b , donc, par dérivation,

$$(f^{-1})'(b) \cdot f'(a) = I_n \quad \text{et} \quad f'(a) \cdot (f^{-1})'(b) = I_n.$$

Donc $f'(a)$ est un isomorphisme, et, notamment, $p = n$. Il se trouve que la réciproque est vraie, et constitue le résultat principal de ce chapitre. \square

5.3 Remarque \star . Soit $f : U$ ouvert de $\mathbb{R}^n \rightarrow V$ ouvert de \mathbb{R}^p une bijection. On vient de voir que, si f est un C^1 -difféomorphisme local, forcément $p = n$. La même conclusion est-elle automatique pour une bijection f plus générale? Si f est un homéomorphisme, le *théorème de l'invariance du domaine* de Brouwer³ (1912) affirme que oui. Si f est une simple bijection, la conclusion est fautive en général, comme le démontra G. Cantor.⁴

Le reste du chapitre consiste en la démonstration du sens non trivial du théorème.

1. Théorie lipschitzienne

5.4 Théorème (Inversion globale). *Soient $h : \mathbb{R}^n \hookrightarrow \mathbb{R}^n$ un lipéomorphisme et $u : \mathbb{R}^n \hookrightarrow \mathbb{R}^n$ une application lipschitzienne. Si*

$$\text{lip } u \text{ lip}(h^{-1}) < 1,$$

$h + u$ est un lipéomorphisme et

$$\text{lip}((h + u)^{-1}) \leq \frac{\text{lip}(h^{-1})}{1 - \text{lip } u \text{ lip}(h^{-1})}.$$

Trivialement, $h + u$ est lipschitzienne, puisque $\text{lip}(h + u) \leq \text{lip } h + \text{lip } u$. Donc la nouveauté est que d'une part $h + u$ est bijective et que d'autre part le rapport de Lipschitz de son inverse satisfait l'inégalité donnée.

Démonstration. L'application $v = u \circ h^{-1}$ est une contraction stricte, puisque $\text{lip } v \leq \text{lip } u \text{ lip}(h^{-1}) < 1$. Comme

$$h + u = (\text{id} + v) \circ h,$$

il suffit de montrer que $\text{id} + v$ est un lipéomorphisme et que

$$\text{lip}((\text{id} + v)^{-1}) \leq \frac{1}{1 - \text{lip } v}.$$

3. Luitzen BROUWER, mathématicien néerlandais (1881–1966), auteur de découvertes fondamentales en topologie, et initiateur de la logique intuitionniste.

4. Georg CANTOR, mathématicien allemand (1845–1918). Pour répondre à certaines questions sur la convergence des séries de Fourier, il donna un fondement à la théorie des ensembles.

Soit $y \in \mathbb{R}^n$. On veut montrer que y possède un antécédant unique par $\text{id} + v$. Ses antécédants éventuels sont les points fixes de $v_y : \mathbb{R}^n \rightarrow \mathbb{R}^n, x \mapsto y - v(x)$:

$$x + v(x) = y \Leftrightarrow x = y - v(x).$$

Or v_y a la même constante de Lipschitz que v . Elle aussi, donc, est une contraction stricte. D'après le théorème de Picard, v_y possède un unique point fixe. Donc $\text{id} + v$ est bijective.

Si $x = (\text{id} + v)^{-1}(y)$ et $x' = (\text{id} + v)^{-1}(y')$, on a

$$|x - x'| = |v_y(x) - v_{y'}(x')| = |y - v(x) - y' + v(x')| \leq |y - y'| + (\text{Lip } v) |x - x'|,$$

donc

$$|x - x'| \leq \frac{|y - y'|}{1 - \text{Lip } v},$$

ce qui montre l'estimation de la constante de Lipschitz de l'inverse. \square

5.5 Exemple. L'équation de Kepler $y = x - \epsilon \sin x$ définit un lipéomorphisme de \mathbb{R} dans \mathbb{R} si $|\epsilon| < 1$. En mécanique céleste, y est le temps et x est un angle (l'anomalie excentrique) qui repère la position de la Terre sur l'ellipse qu'elle décrit autour du Soleil.

5.6 Remarque. Si de plus v est linéaire, l'inverse de $\text{id} + v$ est la somme de la série normalement convergente (pour la norme des opérateurs linéaires) $\sum (-1)^n v^n$. Pourquoi l'adaptation de cette idée au cas non linéaire n'est-elle pas directe? Réponse : la constante de Lipschitz n'est pas une norme pour les opérateurs non linéaires.

5.a Exercice (Inversion globale).

Soit $f : \mathbb{R}^n \hookrightarrow \mathbb{R}^n$ une application continue.

1. Montrer que, si f est injective et propre (au sens que, si (x_n) tend vers l'infini, il en est de même de $(f(x_n))$), f est surjective ; on pourra commencer par montrer que $f(\mathbb{R}^n)$ est fermée.

Supposons dorénavant que,

$$\text{lip } f = \sup_{x \neq x'} \frac{\|f(x) - f(x')\|}{\|x - x'\|} < +\infty \quad \text{et} \quad \text{lip inf } f := \inf_{x \neq x'} \frac{\|f(x) - f(x')\|}{\|x - x'\|} > 0.$$

2. Montrer que f est surjective. En déduire que f est un lipéomorphisme.

3. En déduire une autre démonstration du théorème 5.4.

5.7 Théorème (Inversion locale). Soient $h : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^n$ un lipéomorphisme et $u : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^n$ une application lipschitzienne. Si, localement au voisinage de a ,

$$\text{lip } u \text{ lip}(h^{-1}) < 1,$$

$h + u$ est un lipéomorphisme localement au voisinage de a .

Début de la démonstration du théorème 5.7. Comme précédemment, on note $v = u \circ h^{-1}$, de sorte que

$$h + u = (\text{id} + v) \circ h,$$

et il s'agit donc de montrer que $\text{id} + v$ est un lipéomorphisme local. Cela découlera immédiatement du théorème 5.4, après qu'on aura prolongé $\text{id} + v$ convenablement.

Soit $r > 0$. Notons ρ la « rétraction »⁵

$$\rho : \mathbb{R}^n \longrightarrow \bar{B}(a, r), \quad x \longmapsto \begin{cases} a + r \frac{x-a}{\|x-a\|} & \text{pour } \|x-a\| \geq r \\ x & \text{pour } \|x-a\| \leq r \end{cases}$$

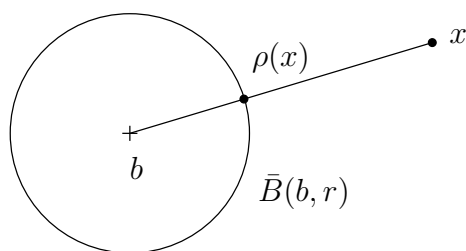


FIGURE 5.1 – La rétraction ρ

Comme la restriction de ρ à la boule est l'identité, si u est initialement définie sur $\bar{B}(a, r)$, l'application $v \circ \rho$ prolonge v à \mathbb{R}^n entier, sans la modifier sur la boule. \square

5.8 Lemme. *On a $\text{lip } \rho = 1$, et donc $\text{lip}(v \circ \rho) \leq \text{lip } v$.*

Démonstration du lemme. Le lemme découle du fait que $\text{lip } \rho = 1$. Par translation on peut se ramener au cas où $a = 0$.

- Si $\|x\|, \|y\| \leq r$, $\|\rho(y) - \rho(x)\| = \|y - x\|$.
- Si $\|x\| \leq r \leq \|y\|$, on a

$$\begin{aligned} \|\rho(y) - \rho(x)\|^2 &= (ry/\|y\| - x)^2 = r^2 - 2rx \cdot y/\|y\| + \|x\|^2 \\ &\leq \|y\|^2 - 2x \cdot y + \|x\|^2 = (y - x)^2 = \|y - x\|^2. \end{aligned}$$

- Si enfin $r \leq \|x\| \leq \|y\|$, on a

$$\begin{aligned} \|\rho(y) - \rho(x)\| &= \left\| \rho\left(\frac{ry}{\|y\|}\right) - \rho\left(\frac{rx}{\|x\|}\right) \right\| \\ &\leq \left\| \frac{ry}{\|y\|} - \frac{rx}{\|x\|} \right\| \quad (\text{d'après le cas précédent}) \\ &\leq \frac{r}{\|x\|} \|y - x\| \leq \|y - x\| \end{aligned}$$

\square

5. Si A est une partie d'un espace métrique ou topologique X , une *rétraction* de X sur A est une application $X \rightarrow A$ dont la restriction à A est l'identité.

Fin de la démonstration du théorème 5.7. Comme $\text{lip } v < 1$, le théorème d'inversion globale, montre que l'application $\text{id} + (v \circ \rho)$ est un lipéomorphisme. Donc il en est de même de sa restriction $\text{id} + v$, donc de $h + u = (\text{id} + v) \circ h$. \square

5.9 Remarque \star . Dans un espace de Banach qui ne serait pas de Hilbert, la constante de Lipschitz de ρ se détériore d'un facteur 2 : par exemple si $\|x\| \leq r \leq \|y\|$, on a

$$\begin{aligned} \|\rho(y) - \rho(x)\| &= \|ry/\|y\| - rx/\|x\|\| \leq \|ry/\|y\| - y\| + \|y - x\| \\ &\leq \|y\| - r - \|y - x\| \leq \|y\| - \|x\| + \|y - x\| \\ &\leq 2\|y - x\|. \end{aligned}$$

Dans le théorème, il faut alors supposer que $2 \text{lip } u \text{lip}(h^{-1}) < 1$.

2. Régularité

Fin de la démonstration du théorème 5.2. Soient $h : x \mapsto b + f'(a) \cdot (x - a)$ et $u = f - h$. Par hypothèse, $f'(a)$ est inversible, donc $\text{lip}(h^{-1}) < +\infty$. De plus, localement

$$u(x + \xi) - u(x) = \left(\int_0^1 u'(x + t\xi) dt \right) \xi,$$

donc

$$\text{lip } u|_{B(a,\epsilon)} \leq \sup_{y \in \bar{B}(a,\epsilon)} \|u'(y)\|.$$

Or, par continuité de f' , $u'(y) \rightarrow 0$ quand $y \rightarrow 0$. Donc, si ϵ est assez petit,

$$\text{lip}(u|_{B(a,\epsilon)}) \text{lip}(h^{-1}) < 1.$$

D'après le théorème d'inversion locale lipschitzien, f est un lipéomorphisme local. La régularité de l'inverse de f s'en déduit grâce à la proposition suivante. \square

5.10 Proposition (Régularité). *Soit $h : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$ un lipéomorphisme.*

- *Si h est dérivable en x , $h'(x)$ est un isomorphisme, h^{-1} est dérivable en $y = h(x)$ et $(h^{-1})'(y) = h'(x)^{-1}$.*
- *De plus, si h est de classe C^k , $1 \leq k \leq \infty$, il en va de même de h^{-1} .*

Démonstration. — Supposons h dérivable en x . Montrons que l'application $h'(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est un isomorphisme. Quels que soient $\xi \in \mathbb{R}^n$ et $\epsilon > 0$ assez petit, on a

$$\|\xi\| = \frac{1}{\epsilon} \|h^{-1}(h(x + \epsilon\xi)) - h^{-1}(h(x))\| \leq \text{Lip}(h^{-1}) \frac{1}{\epsilon} \|h(x + \epsilon\xi) - h(x)\|,$$

donc, à la limite quand $\epsilon \rightarrow 0$,

$$\|\xi\| \leq \text{Lip}(h^{-1}) \|h'(x) \cdot \xi\|;$$

donc $h'(x)$ est injective, donc inversible.

Pour tout $\eta \in \mathbb{R}^p$,

$$\begin{aligned} \eta &= h(h^{-1}(y + \eta)) - h(h^{-1}(y)) \\ &= h'(x) \cdot (h^{-1}(y + \eta) - x) + o(\|h^{-1}(y + \eta) - x\|) \\ &= h'(x) \cdot (h^{-1}(y + \eta) - x) + o(\eta) \end{aligned}$$

(la dernière égalité résulte du fait que h^{-1} est lipschitzienne). En appliquant $h'(x)^{-1}$ aux deux membres, on voit que

$$h'(x)^{-1} \cdot \eta = h^{-1}(y + \eta) - h^{-1}(y) + o(\eta).$$

Donc h^{-1} est différentiable en y et $(h^{-1})'(y) = h'(x)^{-1}$.

— Si de plus h est de classe C^k , on a encore

$$(h^{-1})'(y) = h'(h^{-1}(y))^{-1}$$

pour tout $y \in V$. Par ailleurs, pour $m \in \mathbb{N}$, cette même formule montre que si h^{-1} et h' sont de classe C^m , $(h^{-1})'$ l'est aussi, comme composée de h^{-1} , de h' et de l'application analytique $\text{GL}(E, F) \rightarrow \text{GL}(F, E)$, $A \mapsto A^{-1}$; il en résulte que h^{-1} est de classe C^{m+1} . Ceci montre par récurrence que h^{-1} est de classe C^k .

□

5.11 Remarque ★. L'argument fonctionne aussi avec $k = \omega$, i.e. dans la catégorie des applications analytiques réelles. De plus, comme l'inverse d'une bijection \mathbb{C} -linéaire est \mathbb{C} -linéaire, l'argument fonctionne, *mutatis mutandis*, en classe holomorphe.

5.b Exercice (Algorithme de Newton).

Soit $f : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^n, 0)$ de classe C^2 telle que $f'(0)$ soit inversible. Sous ces hypothèses, on va retrouver le théorème d'inversion locale. Soit, pour $y \in \mathbb{R}^n$,

$$\phi : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}^n, \quad x \mapsto x + f'(x)^{-1}(y - f(x)).$$

On définit $x_0 = 0$, et, aussi longtemps que la récurrence a un sens, $x_{n+1} = \phi(x_n) = \phi^n(0)$.

1. Montrer que, si y est assez proche de 0, x_n existe pour tout n et converge vers un $x \in \mathbb{R}^n$ tel que $f(x) = y$, quadratiquement (c'est-à-dire que $\ln \ln \|x_n - x\|$ est linéaire en n).

2. Comment calculer $\sqrt{2}$ à ϵ près ?

5.c Exercice (Surface stable d'une application).

Soient $X = \mathbb{R}^n$ et $Y = \mathbb{R}^p$, et

$$f : X \times Y \hookrightarrow, \quad (x, y) \mapsto (a(x) + r(x, y), h(x, y))$$

avec

$$a : (X, 0) \hookrightarrow, \quad r : (X \times Y, 0) \rightarrow (X, 0), \quad h : (X \times Y, 0) \rightarrow (Y, 0).$$

On suppose que a est un lipéomorphisme et que

$$\text{lip } r, \text{lip } h \ll 1.$$

On veut montrer qu'il existe une unique application contractante $\alpha : (X, 0) \rightarrow (Y, 0)$ dont le graphe soit invariant par f :

$$f(\text{graphe } \alpha) = \text{graphe } \alpha$$

(théorème de Perron). Le graphe de α s'appelle une *surface* ou *variété stable* (lip-schitzienne) de f .

1. Montrer que, pour toute application contractante $\alpha : (X, 0) \rightarrow (Y, 0)$, il existe une unique application $\beta = G(\alpha) : (X, 0) \rightarrow (Y, 0)$ telle que

$$\text{graphe } \beta = f(\text{graphe } \alpha);$$

$\beta = G(\alpha)$ est la *transformée de graphe* de α .

Soit \mathcal{A} l'espace de toutes les contractions $\alpha : (X, 0) \rightarrow (Y, 0)$ définies sur un voisinage fixe de 0 dans X .

2. Montrer que \mathcal{A} est stable par G .

3. ★ Montrer qu'il existe $0 \leq k < 1$ tel que, pour tous $\alpha \neq \beta \in \mathcal{A}$,

$$\text{lip}(G(\alpha) - G(\beta)) \leq k \text{lip}(\alpha - \beta);$$

on admettra que cela implique l'existence et l'unicité d'un point fixe de G .

4. L'énoncé ci-dessous persiste-t-il si X et Y sont des espaces de Banach ?

Chapitre 6

Forme normale d'une application linéaire

Mots-clefs du chapitre Applications linéaires équivalentes, forme normale, rang

Soient E et F deux espaces vectoriels de dimension finie, respectivement n et p , et φ une application linéaire de E dans F .

6.1 Définition. Une application linéaire $\tilde{\varphi} : \tilde{E} \rightarrow \tilde{F}$ est *équivalente* à φ s'il existe des isomorphismes $P : E \rightarrow \tilde{E}$ et $Q : F \rightarrow \tilde{F}$ tels que

$$\tilde{\varphi} = Q \circ \varphi \circ P^{-1}.$$

Une *forme normale* de φ est une application linéaire $\mathbb{R}^n \rightarrow \mathbb{R}^p$ équivalente à φ et de la forme

$$J_r = \left(\begin{array}{c|c} \text{id}_{\mathbb{R}^r} & 0 \\ \hline 0 & 0 \end{array} \right) : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

pour un certain entier $r \leq \min(n, p)$.

6.2 Proposition (théorème du rang). *La forme normale de φ est J_r , où $r = \dim \text{im } \varphi$ (rang de φ). En particulier, $\dim \ker \varphi + \text{rang } \varphi = n$.*

Démonstration. Soit (e_{s+1}, \dots, e_n) une base de $\ker \varphi$ (où l'on a noté s est la codimension de $\ker \varphi$), que l'on complète en une base (e_1, \dots, e_n) de E . La restriction de φ à $\text{Vect}(e_1, \dots, e_s)$ est injective, et réalise donc un isomorphisme sur l'image de φ . Donc $r = s$ (c'est-à-dire que $\dim \ker \varphi + \text{rang } \varphi = n$). Soit maintenant (f_1, \dots, f_r) la base de $\text{im } \varphi$ définie par $f_i = \varphi(e_i)$ pour $i = 1, \dots, r$; on la complète en une base (f_1, \dots, f_p) de F . Alors J_r est la matrice de φ dans les bases (e_1, \dots, e_n) et (f_1, \dots, f_p) . \square

Autrement dit, toutes les matrices de rang r ont pour matrice J_r dans des bases bien choisies.

6.3 *Cas particuliers.* Si φ est surjective, $r = p \leq n$ et la forme normale de φ est la surjection

$$J_p = \left(\text{id}_{\mathbb{R}^p} \mid 0 \right).$$

Si φ est injective, $r = n \leq p$ et la forme normale de φ est l'injection

$$J_n = \left(\frac{\text{id}_{\mathbb{R}^n}}{0} \right).$$

6.a Exercice (Critère de contrôlabilité de Kalman).

Soient $A \in M_{d,d}(\mathbb{R})$ et $B \in M_{d,r}(\mathbb{R})$. On considère le problème de Cauchy

$$\begin{cases} f(0) = x \\ f'(t) = Af(t) + Bg(t), \end{cases}$$

sur l'intervalle de temps $[0, T]$, où $g : [0, T] \rightarrow \mathbb{R}^r$ est une fonction bornée continue. Soit \mathcal{A}_g l'ensemble des points de la forme $f(T)$ obtenus en faisant varier g . La fonction g s'appelle un *contrôle* et les points $y \in \mathcal{A}_x$ sont qualifiés d'*atteignables*. Montrer que \mathcal{A}_x est \mathbb{R}^d entier si et seulement si

$$\text{rang}[B, AB, \dots, A^{d-1}B] = d$$

(la notation $[B, AB, \dots, A^{d-1}B]$ désigne la matrice en ligne par blocs, dont les blocs successifs sont les $A^k B$, $k = 0, \dots, d-1$).

L'action du groupe $GL(E)$ sur l'espace $L(E)$ des endomorphismes de E (par conjugaison : $P \cdot \varphi \mapsto P \circ \varphi \circ P^{-1}$) est significativement plus compliquée et possède une infinité d'orbites (étant donnée déjà que les homothéties λI , $\lambda \in \mathbb{R}$, ne sont pas équivalentes entre elles). Il existe plusieurs formes normales relativement à cette action de groupe.

6.b Exercice (Forme normale de Jordan¹).

Soient E un espace vectoriel complexe de dimension finie et φ un endomorphisme de E de polynôme minimal

$$\mu(X) = \prod_{k=1}^r (X - \lambda_k)^{\alpha_k},$$

où les $\lambda_k \in \mathbb{C}$ sont deux à deux distincts.

Écrivons la décomposition en éléments simples de $1/\mu$ de la façon suivante :

$$\frac{1}{\mu(X)} = \sum_{1 \leq k \leq r} \frac{Q_k(X)}{(X - \lambda_k)^{\alpha_k}}.$$

1. Montrer que les endomorphismes

$$P_i = Q_i(\varphi) \prod_{j \neq i} (\varphi - \lambda_j I)^{\alpha_j}$$

1. Camille JORDAN, mathématicien français, spécialiste de théorie des groupes

sont des projecteurs algébriquement orthogonaux ($P_i P_j = 0$ si $i \neq j$), tels que $\sum_i P_i = \text{id}$.

2. Montrer le *lemme des noyaux* :

$$E = \bigoplus_{k=1}^r \ker(\varphi - \lambda_r \text{id})^{\alpha_k}.$$

3. En déduire que φ se décompose comme la somme d'un endomorphisme diagonalisable

$$D = \sum_i \alpha_i P_i$$

et d'un endomorphisme nilpotent $N = \varphi - D$ (*décomposition de Dunford*).

4. En déduire qu'il existe une base de E dans laquelle la matrice d'un opérateur donné $\varphi \in L(E)$ est une matrice diagonale par blocs, dont chaque bloc est un bloc de Jordan

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

6.c Exercice (Diagonalisabilité sur \mathbb{R}).

Montrer que, si A est une matrice réelle diagonalisable dans \mathbb{C} dont les valeurs propres sont réelles, elle est diagonalisable dans \mathbb{R} .

La *forme normale de Frobenius*² a la mérite d'être "algorithmique" (c'est-à-dire que son calcul ne passe pas par la détermination des racines du polynôme caractéristique, et est valable quel que soit le corps de base).

6.d Exercice ★ (Forme normale de Frobenius).

Soient E un espace vectoriel de dimension finie sur $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , et $\varphi \in L(E)$.

1. Montrer qu'il existe un entier $r \geq 1$, des sous-espaces vectoriels E_1, \dots, E_r de E , des vecteurs $x_1 \in E_1, \dots, x_r \in E_r$ et des entiers $d_1, \dots, d_r \geq 1$ tels que

- $E = E_1 \oplus \cdots \oplus E_r$
- E_i a pour base $(x_i, \varphi(x_i), \dots, \varphi^{d_i-1}(x_i))$ ($i = 1, \dots, r$)
- π_{i+1} divise π_i ($i = 1, \dots, r-1$), où l'on a noté π_i le polynôme minimal de la restriction de φ à E_i .

Indication : On pourra commencer par montrer qu'il existe un vecteur $x_1 \in E$ tel que le polynôme minimal de φ soit le polynôme minimal ponctuel de φ en x_1 (c'est-à-dire le polynôme unitaire qui engendre l'idéal des polynômes $P \in \mathbb{K}[X]$ tels que $P(\varphi)(x_1) = 0$).

2. En déduire qu'il existe une forme bilinéaire symétrique non-dégénérée b (produit scalaire pseudo-euclidien) telle que φ soit autoadjoint pour cette forme :

$$b(\varphi(x), y) = b(x, \varphi(y)).$$

2. Ferdinand Georg FROBENIUS, mathématicien allemand (1849–1917), spécialiste notamment de la théorie des représentations de groupes.

Indication : malgré les efforts pour rendre cet énoncé intrinsèque, la démonstration est matricielle, et il s'agit de montrer que, si $M \in M_n(\mathbb{K})$, M est semblable à sa transposée.

Une application en est donnée dans l'exercice 22.f.

Chapitre 7

Forme normale d'une application

Mots-clefs du chapitre Applications équivalentes, forme normale

La notion d'applications équivalentes que nous introduisons ici permet de simplifier l'énoncé de diverses variantes géométriques du théorème d'inversion locale.

Soit une application $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ de classe C^1 . Réinterprétons la conclusion du théorème d'inversion locale de la façon suivante. Si $f'(x)$ est inversible, f est un difféomorphisme local donc on peut voir l'application

$$f : x \mapsto y = f(x)$$

comme un "changement de coordonnées" ; au lieu de repérer un point $x \in \mathbb{R}^n$ par la collection de ses coordonnées (x_1, \dots, x_n) , on le repère par une autre collection de nombres (y_1, \dots, y_n) . Alors, tautologiquement, l'application f vue dans les coordonnées y n'est autre que l'identité :

$$\begin{array}{ccc} x \in (\mathbb{R}^n, a) & \xrightarrow{f} & y \in (\mathbb{R}^n, b). \\ \downarrow f & \nearrow \text{id} = f \circ f^{-1} & \\ y \in (\mathbb{R}^n, b = f(a)) & & \end{array}$$

On a trouvé des coordonnées locales y dans lesquelles f prend la forme de l'identité, qu'on appelle sa *forme normale*.

7.1 Définition. Une application $\tilde{f} : (\mathbb{R}^n, \tilde{a}) \rightarrow (\mathbb{R}^p, \tilde{b})$ de classe C^1 est *équivalente* à f s'il existe des difféomorphismes $\alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^n, \tilde{a})$ et $\beta : (\mathbb{R}^p, b) \rightarrow (\mathbb{R}^p, \tilde{b})$ tels que

$$\tilde{f} = \beta \circ f \circ \alpha^{-1}.$$

Une *forme normale* de f est une application linéaire $J_r : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^p, 0)$ équivalente à f .¹

1. Quand aucune application J_r n'est équivalente à f , on peut chercher des formes normales de f plus compliquées, par exemple polynomiales. On en verra des exemples plus loin avec le lemme de Morse.

Autrement dit, le diagramme d'applications suivant commute :

$$\begin{array}{ccc} (\mathbb{R}^n, a) & \xrightarrow{f} & (\mathbb{R}^p, b) \\ \downarrow \alpha & & \downarrow \beta \\ (\mathbb{R}^n, \tilde{a}) & \xrightarrow{\tilde{f}} & (\mathbb{R}^p, \tilde{b}). \end{array} \quad (7.1)$$

Les difféomorphismes « verticaux » sont vus comme des changements de variables, tandis que l'application qui nous intéresse est horizontale : c'est f dans les coordonnées initiales, et sa réincarnation \tilde{f} dans les nouvelles coordonnées.

Par exemple, f est toujours équivalente à une application $\tilde{f} : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^p, 0)$, à laquelle on se ramène par de simples translations :

$$\begin{array}{ccc} (\mathbb{R}^n, a) & \xrightarrow{f} & (\mathbb{R}^p, b) \\ \downarrow \cdot -a & & \downarrow \cdot -b \\ (\mathbb{R}^n, \tilde{a}) & \xrightarrow{F} & (\mathbb{R}^p, \tilde{b}), \end{array}$$

soit

$$\tilde{f}(\tilde{x}) = f(\tilde{x} + a) - b.$$

7.a Exercice.

Montrer que si une application est équivalente à une application localement surjective (respectivement injective), la première elle-même est localement surjective (resp. injective).

Chapitre 8

Fonctions implicites

Mots-clefs du chapitre Fonction implicite

Dans ce chapitre, nous allons voir une variante de l'inversion locale, correspondant au cas où la forme normale de l'application est non l'identité, mais une projection linéaire d'un espace produit sur un facteur. Dans le chapitre suivant, nous verrons des variantes correspondant aux formes normales suivantes : projection linéaire, injection linéaire ou, plus généralement, la matrice J_r (subimmersion linéaire) du chapitre 6.

8.1 Théorème (des fonctions implicites). *Soit*

$$f : (\mathbb{R}^n \times \mathbb{R}^p, (a, b)) \rightarrow (\mathbb{R}^p, c), \quad (x, y) \mapsto f(x, y)$$

une application de classe C^1 . Si $\partial_y f(a, b) \in L(\mathbb{R}^p)$ est un isomorphisme, f est équivalente à la projection canonique

$$J_p : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^p, \quad (x, z) \mapsto z.$$

Démonstration. L'application

$$\alpha : (\mathbb{R}^n \times \mathbb{R}^p, (a, b)) \rightarrow (\mathbb{R}^p, c), \quad (x, y) \mapsto (x, z) = (x, f(x, y))$$

vérifie

$$\alpha'(a, b) = \left(\begin{array}{c|c} \text{id}_{\mathbb{R}^n} & 0 \\ \hline \partial_x f(a, b) & \partial_y f(a, b) \end{array} \right),$$

où cette dernière matrice carrée est de rang maximum $n + p$. Donc, d'après le théorème d'inversion locale, α est un difféomorphisme local, qu'on peut utiliser comme changement de coordonnées à la source de f . Donc f est équivalente à la projection canonique

$$J_p = f \circ \alpha^{-1} : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^p, \quad (x, z) \mapsto z.$$

□

La forme la plus opérationnelle du théorème est donnée par l'exercice suivant.

8.a Exercice (Théorème des fonctions implicites).

Sous les hypothèses du théorème des fonctions implicites, montrer qu'il existe une unique fonction $Y : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ telle que, localement,

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

8.b Exercice (Relation entre dérivées partielles).

Soit $f : (\mathbb{R}^3, (a, b, c)) \rightarrow (\mathbb{R}, 0)$ une fonction de classe C^1 telle que

$$\partial_x f(a, b, c) \neq 0, \partial_y f(a, b, c) \neq 0, \partial_z f(a, b, c) \neq 0.$$

On note X, Y, Z les fonctions, définies implicitement et localement, telles que

$$f(x, y, z) = 0 \Leftrightarrow x = X(y, z) \Leftrightarrow y = Y(x, z) \Leftrightarrow z = Z(x, y).$$

Que vaut le produit $\partial_y X \partial_z Y \partial_x Z$ au point (a, b, c) ?

8.c Exercice (Résolution approchée d'une équation).

1. Donner une valeur approchée de la solution, proche de 1, de l'équation

$$x^7 + 0,99x - 2,03 = 0,$$

en calculant le développement limité au premier ordre de la fonction $x(p, q)$ définie implicitement par l'équation $x^7 + px + q = 0$ au voisinage de $(x, p, q) = (1, 1, -2)$.

2. Donner une méthode pour majorer l'erreur ainsi commise ; on pourra déjà montrer que la solution exacte vérifie $1 \leq x \leq 2$.

8.d Exercice (L'équation de degré trois). Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}, (x, p, q) \mapsto x^3 + px + q$. Dessiner, pour p fixé respectivement $< 0, = 0$ et > 0 , les courbes d'équation $f(x, p, q) = 0$ dans le plan des (x, q) . Puis tracer dans le plan des (p, q) le lieu des points où l'équation $f = 0$ ne détermine pas de fonction implicite $x(p, q)$.

Chapitre 9

Submersions et immersions

Mots-clefs du chapitre immersion, subimmersion

Soit une application $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ de classe C^1 .

Submersions Le théorème de submersion est une version géométrique du théorème des fonctions implicites, dans laquelle l'hypothèse de non-dégénérescence porte sur un mineur quelconque de la dérivée, pas forcément le premier (ou le dernier) mineur carré.

9.1 Théorème (de submersion). *Si $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est surjective (donc $n \geq p$), une forme normale de f est la surjection canonique*

$$J_p = (I_p \mid 0) : \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p} \rightarrow \mathbb{R}^p, \quad (x, y) \mapsto x;$$

on dit alors que f est une submersion en a .

Comme $f'(a)$ est elle-même équivalente (comme application linéaire) à J_p , le théorème dit que f est équivalente à sa partie linéaire $f'(a)$, autrement dit que f est *linéarisable*.

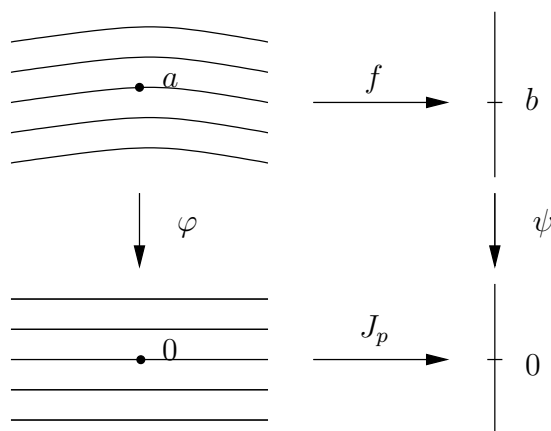


FIGURE 9.1 – Submersion f et sa forme normale J_p

9.a Exercice (Théorème de submersion).

1. Démontrer le théorème de submersion. *Indication* : S'inspirer du cas où f est une application linéaire pour trouver un isomorphisme quelque part... (Si $\mathbb{R}^n \equiv \ker f'(a) \oplus V$, appliquer le théorème d'inversion locale au « déploiement » $F : (V \times \ker f'(a), a) \rightarrow \mathbb{R}^p \times \mathbb{R}^{n-p}$, $(x, y) \mapsto (f(x, y), y)$.)

2. En déduire que f possède un inverse à droite (appelé *section* de f), c'est-à-dire une fonction $g : (\mathbb{R}^p, b) \rightarrow (\mathbb{R}^n, a)$ telle que

$$f \circ g = \text{id}_{(\mathbb{R}^p, b)}.$$

Le fait que $f'(a)$ soit surjective dit que, dans la matrice

$$f'(a) = \begin{pmatrix} \partial_1 f_1(a) & \cdots & \partial_n f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_p(a) & \cdots & \partial_n f_p(a) \end{pmatrix}$$

on peut trouver une sous-matrice carrée inversible d'ordre p . Le théorème des fonctions implicites est donc un cas particulier du théorème de submersion, où la sous-matrice carrée inversible est simplement constituée des p dernières colonnes.

Les applications du théorème de submersion viendront dans les chapitres suivants, notamment dans la définition d'une surface.

Immersion

9.2 Théorème (d'immersion). *Si $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est injective (donc $n \leq p$), une forme normale de f est l'injection canonique*

$$J_n = \begin{pmatrix} I_n \\ 0 \end{pmatrix} : \mathbb{R}^n \rightarrow \mathbb{R}^p = \mathbb{R}^n \times \mathbb{R}^{p-n}, \quad x \mapsto (x, 0_{\mathbb{R}^{p-n}});$$

on dit alors que f est une immersion.

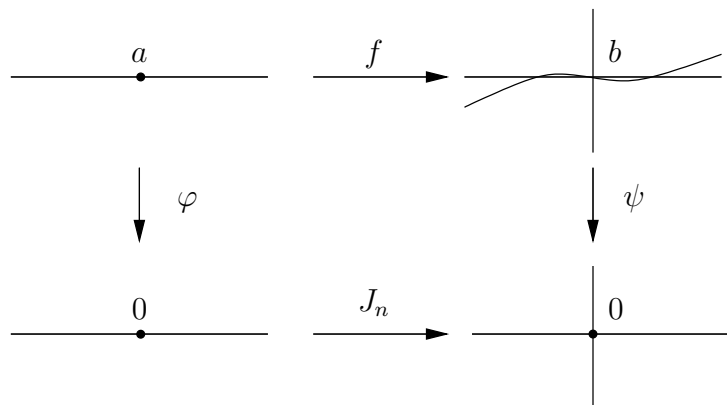


FIGURE 9.2 – Immersion f et sa forme normale J_n

9.b Exercice. 1. Démontrer le théorème d'immersion. *Indication* : Si E est une supplémentaire de $\text{im } f'(a) : \mathbb{R}^p \equiv \text{im } f'(a) \oplus E$, considérer l'application $F : (\mathbb{R}^n \times E, a) \rightarrow \mathbb{R}^p$, $(x, y) \mapsto f(x) + y$.

2. En déduire que f possède un inverse local à gauche, c'est-à-dire une fonction $g : (\mathbb{R}^p, b) \rightarrow (\mathbb{R}^n, a)$ telle que

$$g \circ f = \text{id}_{(\mathbb{R}^n, a)}.$$

Subimmersions ★ Le théorème du rang constant ci-dessous englobe les théorèmes de submersion et d'immersion. La complication est que, si le rang n'est pas maximal en un point, il peut ne pas être constant.

9.3 Définition. Le rang d'une application $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^p$ en a est le rang de $f'(a)$ ($\leq \min(n, p)$). L'application est *de rang maximal* en a si son rang en a vaut $\min(n, p)$. Le point a est *régulier* ou *critique* selon que le rang est maximal ou pas.

9.c Exercice (Rang d'une application).

1. Donner un exemple de discontinuité du rang d'une application.

2. Montrer que le rang d'une application en un point est semi-continu inférieurement, c'est-à-dire que, si (x_n) tend vers a , la limite du rang de f en les x_n est supérieure au rang de f en a .

9.4 Théorème ★ (du rang constant). *Si le rang de $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est localement constant égal à r au voisinage de a , une forme normale de f est*

$$J_r = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right) : \mathbb{R}^n = \mathbb{R}^r \times \mathbb{R}^{n-r} \rightarrow \mathbb{R}^p = \mathbb{R}^r \times \mathbb{R}^{p-r}, \quad (x, y) \mapsto (x, 0_{\mathbb{R}^{p-r}}).$$

on dit alors que f est une subimmersion.

9.d Exercice. Déduire le théorème du rang constant du théorème d'inversion locale.

9.5 Remarque. Dans les théorèmes, 9.1, 8.1, 9.2 et 9.4, si f est de classe C^k ($k \in \mathbb{N} \cup \{+\infty\}$) il en est de même des fonctions construites à l'aide du théorème d'inversion locale, d'après le théorème de régularité du chapitre précédent.

9.e Exercice. Soit $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^n$ une application localement injective.

1. L'application f est-elle automatiquement un difféomorphisme local ?

2. Montrer qu'il existe des points b arbitrairement proches de a tels que f soit un difféomorphisme local au voisinage de b .

Chapitre 10

Incursion en analyse complexe ★

La rigidité de la classe analytique permet d'obtenir, pour les fonctions analytiques, des théorèmes de type inversion locale plus précis. Nous effleurons ici le sujet à travers quelques exemples, et renvoyons à un cours spécialisé pour des démonstrations détaillées.

10.a Exercice (Méthode des séries majorantes [Car61]).

Montrer le théorème d'inversion locale directement en classe analytique. On pourra utiliser la méthode suivante, dite des séries majorantes. Commencer par montrer le théorème dans la classe des séries formelles¹. Puis utiliser une *série majorante* pour montrer la convergence de la série formelle réciproque trouvée ; la série majorante d'une série $\sum a_n z^n$ est une série $\sum_n \alpha_n z^n$ à coefficients réels positifs telle que $|a_n| \leq \alpha_n$ pour tout n .

10.1 Théorème (de Rouché). Soient f et g deux fonctions holomorphes sur un ouvert borné D de \mathbb{C} bordé par des chemins continus γ_i , $i \in I$. Si

$$|f(z) - g(z)| < |g(z)| \quad \text{le long des } \gamma_i,$$

f et $f + g$ ont le même nombre de zéros dans D .

Idée de démonstration. La fonction méromorphe $h = f/g$ vérifie

$$|h(z) - 1| = \frac{|f(z) - g(z)|}{|g(z)|} < 1.$$

Lui appliquer le théorème de l'argument. □

10.b Exercice (Théorème fondamental de l'algèbre).

Montrer que tout polynôme complexe de degré n possède exactement n racines complexes.

10.c Exercice (Localisation des zéros d'un polynôme).

Soient $a \in \mathbb{C}$, $|a| \geq 1$ et $n \in \mathbb{N}$, $n \geq 2$. Montrer que les racines du polynôme $1 + z + az^n$ ont un module ≤ 2 .

1. Une *série formelle* est une série entière, ici à plusieurs variables, dont le domaine de convergence est éventuellement nul.

10.d Exercice (Polygone de Newton).

Soit $\Sigma \subset \mathbb{R}^2$ l'ensemble des zéros du polynôme réel

$$P(x, y) = y^3 + x^2y^2 - xy + x^4.$$

On s'intéresse à la structure de Σ au voisinage de l'origine. Trouver les réels $\alpha > 0$ tels qu'il existe une unique fonction analytique $Y : (\mathbb{R}, 0) \rightarrow \mathbb{R}$ telle que localement

$$P(x, x^\alpha Y(x)) \equiv 0.$$

On commencera poura procéder en trois temps : (1) déterminer les valeurs de α pour lesquelles l'équation $P(x, x^\alpha Y(x)) = 0$ possède une unique solution formelle, (2) déterminer lesquelles de ces solutions formelles correspondent à des solutions analytiques, et (3) montrer que les autres valeurs de α ne donnent pas d'autres solutions.

10.e Exercice ★ (Théorème de division de Malgrange).

Soient $f : (\mathbb{C}^n, 0) \rightarrow (\mathbb{C}^n, 0)$, $x \mapsto y = f(x)$, une application holomorphe, et μ sa *multiplicité*, définie par

$$\mu = \lim_{r \rightarrow 0^+} \max_y \#(f^{-1}(y) \cap B(0, r)),$$

où $B(0, r)$ est la boule de rayon r centrée en l'origine.

On définit de plus l'*algèbre locale* de f comme

$$Q = H_x / (f^* \mathfrak{M}_y) H_x,$$

où H_x est l'algèbre des germes en 0 de fonctions holomorphes dans \mathbb{R}^n , \mathfrak{M}_y est l'idéal maximal des germes en 0 de fonctions holomorphes dans \mathbb{R}^n (c'est-à-dire l'idéal des germes s'annulant en 0) et $f^* \mathfrak{M}_y$ est l'ensemble des fonctions de la forme $\varphi(f(x))$, où $\varphi \in \mathfrak{M}_y$.

1. ★ Montrer que la multiplicité μ est bien définie, que, si r est assez petit > 0 , presque tout point y proche de 0 possède exactement μ préimages dans $B(0, r)$, et que Q est de dimension μ (voir [AVZK86, paragraphe 5]).

2. Montrer que l'*application de Viète*, définie par le fait que les y_i sont les polynômes symétriques élémentaires

$$y_1 = \sum_i x_i, \quad y_2 = \sum_{i < j} x_i x_j, \quad \dots, \quad y_n = \prod_i x_i,$$

a pour multiplicité $\mu = n!$.

Soit $e = (e_1, \dots, e_\mu) \in H_x^\mu$ une famille de représentants d'une base de Q .

3. Montrer que, pour toute série formelle $\varphi \in \mathbb{C}[[x_1, \dots, x_n]]$, il existe un unique champ formel $a \in \mathbb{C}[[y_1, \dots, y_n]]^\mu$ telle que

$$\varphi = f^* a \cdot e,$$

soit

$$\varphi(x) = a_1(f(x))e_1(x) + \dots + a_\mu(f(x))e_\mu(x).$$

Autrement dit, en “divisant” (au sens du produit scalaire) la fonction φ par e on obtient une fonction qui se factorise par f , de la forme $a(f(x))$, d’où le nom de *théorème de division*.

4. ★ Montrer que, si φ est convergente, il en est de même de a ; on pourra utiliser la méthode des séries majorantes [Hou61].

5. En déduire que, si φ est une fonction symétrique, elle se factorise (de façon analytique) par l’application de Viète.

6. Montrer que, si f dépend de paramètres, on peut remplacer les fonctions $e_k(x)$ par des fonctions arbitraires $E_k(x, t)$ telles que $E_k(x, 0) = e_k(x)$. *Indication* : Les $E_k(x, t)$ forment un ensemble de générateurs de l’algèbre locale de l’application $F : (x, t) \mapsto (f(x, t), t)$ parce que $F(x, t) - F(x, 0)$ et les $E_k(x, t) - E_k(x, 0)$ sont dans l’idéal engendré par t dans l’espace $\mathbb{C}\{x, t\}$ des fonctions holomorphes en 0 de x et de t .

7. En déduire le *théorème de préparation de Weierstrass* : Soit

$$F(x, t) = x^\mu + u_1(t)x^{\mu-1} + \cdots + u_\mu(t), \quad u_s(0) = 0.$$

Alors toute fonction $\varphi(x, t)$ se laisse mettre sous la forme

$$\varphi(x, t) = a(x, t)F(x, t) + \sum_{r=0}^{\mu-1} h_r(t)x^r;$$

a est le *quotient* de φ par F , et le polynôme de degré $\mu - 1$ est le *reste*.

Chapitre 11

Surfaces I – Espace tangent

Tout le monde sait ce que c'est qu'une courbe, jusqu'à ce qu'il ait étudié suffisamment de mathématiques pour que ses idées soient brouillées par le nombre infini d'exceptions possibles.

F. Klein

Mots-clefs du chapitre Surface, dimension, codimension, redressement, équation, espace tangent

Deux moyens pour définir une surface de \mathbb{R}^3 sont : par une équation

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

ou par un paramétrage

$$x = X(s, t), \quad y = Y(s, t), \quad z = Z(s, t)$$

(où s et t sont des paramètres).¹ Par exemple, le plan d'équation $x + y + z = 0$ peut aussi être paramétré de la façon suivante

$$x = s, \quad y = t, \quad z = -s - t \quad (s, t \in \mathbb{R}).$$

Mais, en fonction de f ou de (X, Y, Z) , on peut obtenir un ensemble "singulier", qu'on n'a pas envie d'appeler une surface (lisse). C'est le cas par exemple de l'équation $x^2 + y^2 + z^2 = 0$ et du paramétrage constant $x = y = z = 0$. Dans ce chapitre, on verra des conditions suffisantes pour obtenir une surface : en l'occurrence il suffit que f soit une submersion, ou (X, Y, Z) une immersion.

1. On pourrait aussi prendre l'exemple d'une courbe de \mathbb{R}^3 , qui peut être défini par un système de deux équations réelles $f(x, y, z) = 0, g(x, y, z) = 0$ ou par un paramétrage $x = X(t), y = Y(t), z = Z(t)$.

11.1 Définition. Une *surface (lisse)*² de dimension d de \mathbb{R}^n est une partie S de \mathbb{R}^n telle que, pour tout $a \in S$, il existe un difféomorphisme

$$\alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^d \times \mathbb{R}^{n-d}, 0)$$

tel que, localement au voisinage de 0,

$$(\tilde{x}, \tilde{y}) \in \alpha(S) \Leftrightarrow \tilde{y} = 0.$$

Le difféomorphisme α est un *redressement* local de S (voir la figure 11.1), l'entier d est la *dimension* de S , et l'entier $p = n - d$ est sa *codimension*.

Quand $d = 1$, on parle plus spécifiquement de *courbe* et, quand $p = 1$, d'*hyper-surface*.

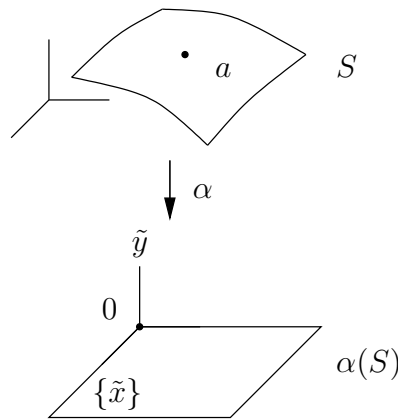


FIGURE 11.1 – Surface S et son difféomorphisme de redressement α

11.a Exercice (Graphe d'une application).

Soit $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ une application de classe C^1 . Montrer que son graphe

$$G(f) = \{(x, y), y = f(x)\}$$

est une surface de $\mathbb{R}^n \times \mathbb{R}^p$ de dimension n .

11.2 Proposition (Surface définie par une équation). Soient $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ une *submersion* et S la partie de \mathbb{R}^n définie par l'équation $f = b$, i.e. $S = f^{-1}(b)$; par abus, on appelle parfois f elle-même une équation de S . Alors S est une surface de \mathbb{R}^n de codimension p , et donc de dimension $n - p$.

Démonstration. En effet, en notant $d = n - p$, le théorème de submersion montre qu'il existe des difféomorphismes

$$\begin{cases} \alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^d \times \mathbb{R}^p, 0) \\ \beta : (\mathbb{R}^p, b) \rightarrow (\mathbb{R}^p, 0) \end{cases}$$

2. Nous utiliserons ce vocable pour désigner les sous-variétés de \mathbb{R}^n . Parfois on réserve le terme de surface aux sous-variétés de dimension 2.

tels que le diagramme suivant commute :

$$\begin{array}{ccc} (\mathbb{R}^n, a) & \xrightarrow{f} & (\mathbb{R}^p, b) \\ \alpha \downarrow & & \downarrow \beta \\ (\mathbb{R}^d \times \mathbb{R}^p, 0) & \xrightarrow{J_p} & (\mathbb{R}^p, 0). \end{array}$$

Donc, localement, si $\alpha(x) = (\tilde{x}, \tilde{y})$, et puisque $\beta \circ f = J_p \circ \alpha$,

$$(\tilde{x}, \tilde{y}) \in \alpha(S) \Leftrightarrow f(x) = b \Leftrightarrow (\beta \circ f)(x) = 0 \Leftrightarrow (J_p \circ \alpha)(x) = 0 \Leftrightarrow \tilde{y} = 0.$$

Donc le difféomorphisme α redresse S , qui est bien une surface de \mathbb{R}^n . □

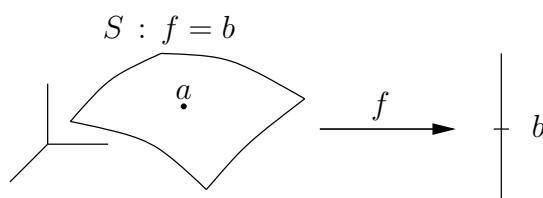


FIGURE 11.2 – Surface S d'équation $f = 0$

11.3 Exemple. La sphère

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1}, \|x\| = 1\}$$

est une surface de dimension n , définie globalement par l'équation

$$f(x) = 0, \quad f(x) = \|x\|^2 - 1;$$

f est bien un submersion au voisinage de tout point de \mathbb{S}^n (mais pas en $0 \notin \mathbb{S}^n$) puisque

$$f'(x) \cdot \xi = 2x \cdot \xi = 0 \quad (\forall \xi) \Leftrightarrow x = 0.$$

11.4 Exemple. Notons $\mathbb{T} = \mathbb{S}^1$ le cercle. C'est une courbe de \mathbb{R}^2 . Le tore de dimension n , \mathbb{T}^n , est alors une surface de \mathbb{R}^{2n} , de dimension n .

11.b Exercice (Groupes classiques).

Montrer que les groupes suivants sont des surfaces de $M_n(\mathbb{R}) \equiv \mathbb{R}^{n^2}$, dont on calculera la dimension :

- le groupe linéaire $GL_n(\mathbb{R})$
- le groupe linéaire spécial

$$SL_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}), \det M = 1\}$$

- le groupe orthogonal

$$O_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}), {}^t M M = I\}.$$

11.c Exercice (Un lemme de division ★). Soient $f : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}$ une fonction de classe C^∞ et $g : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}^p$ une submersion de classe C^∞ . Montrer que

$$g(x) = 0 \Rightarrow f(x) = 0 \quad (\forall x)$$

si et seulement si il existe une fonction $a : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^p, 0)$ telle que $f(x) = a(x) \cdot g(x)$ (le produit scalaire euclidien de $a(x)$ et $g(x)$).

Un chemin γ est *tracé sur* S (d'équation $f = 0$) si $f \circ \gamma \equiv 0$.³

11.5 Définition. L'*espace tangent* à une surface $S \subset \mathbb{R}^n$ en un point $x \in S$ est l'ensemble $T_x S$ des vecteurs vitesses $\gamma'(0)$ de chemins $\gamma : (\mathbb{R}, 0) \rightarrow (S, x)$ (on dit que les tels chemins sont *tracés* sur S) dérivables (vus comme chemins $(\mathbb{R}, 0) \rightarrow \mathbb{R}^n$).

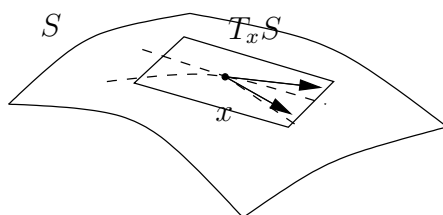


FIGURE 11.3 – Plan tangent à une surface. On représente généralement le *plan affine tangent* $x + T_x S$, passant par x et dirigé par $T_x S$.

11.6 Proposition. Si $f : (\mathbb{R}^n, x) \rightarrow (\mathbb{R}^p, 0)$ est une équation locale de S ,

$$T_x S = \ker f'(x).$$

En particulier, $T_x S$ est un sous-espace vectoriel de même dimension que S .

Démonstration. L'inclusion $T_x S \subset \ker f'(x)$ est triviale, puisque si $\gamma : (\mathbb{R}, 0) \rightarrow (S, x)$ est un chemin dérivable tracé sur S , $f \circ \gamma \equiv 0$, donc, par dérivation en 0, $f'(x) \cdot \gamma'(0) = 0$, donc $\gamma'(0) \in \ker f'(x)$.

Montrons l'inclusion réciproque. Soit $\xi \in \ker f'(x)$. On veut réaliser ξ comme le vecteur vitesse $\gamma'(0)$ d'un chemin $\gamma : (\mathbb{R}, 0) \rightarrow (\mathbb{R}^n, x)$ tracé sur S ($f \circ \gamma \equiv 0$). D'après le théorème de submersion, il existe deux difféomorphismes α et β faisant commuter le diagramme suivant :

$$\begin{array}{ccc} (\mathbb{R}^n, x) & \xrightarrow{f} & (\mathbb{R}^n, 0) \\ \downarrow \alpha & & \downarrow \beta \\ (\mathbb{R}^n, 0) & \xrightarrow{J_p} & (\mathbb{R}^p, 0). \end{array} \quad (11.1)$$

Notons $\tilde{\xi} = \alpha'(x) \cdot \xi$. Comme $J_p \circ \alpha = \beta \circ f$, on a

$$J_p \cdot \tilde{\xi} = J_p \cdot \alpha'(x) \cdot \xi = \beta'(0) f'(x) \cdot \xi = 0.$$

3. Rappelons que cette notation signifie que la fonction $f \circ \gamma$ est identiquement nulle, alors que la notation $f \circ \gamma = 0$ pourrait signifier qu'elle s'annule juste en certains points.

Donc $\tilde{\xi} \in \ker J_p$ (traduction du côté de la forme normale, de l'hypothèse selon laquelle $\xi \in \ker f'(x)$). Cette fois, il est facile de réaliser $\tilde{\xi}$ comme le vecteur vitesse d'un chemin $\tilde{\gamma}$ tracé sur $\alpha(S) = \ker J_p$: à cause de la forme très particulière de $J_p : (x, y) \mapsto (x, 0)$, l'hypothèse dit que $\tilde{\xi}$ est de la forme $\tilde{\xi} = (0, \tilde{\xi}_2)$, et il suffit donc de poser $\tilde{\gamma}(t) = (0, t\tilde{\xi}_2)$. Il ne reste qu'à poser $\gamma(t) = \alpha^{-1} \circ \tilde{\gamma}(t)$. On vérifie que

$$f \circ \gamma = f \circ \alpha^{-1} \circ \tilde{\gamma} = \beta^{-1} \circ J_p \circ \tilde{\gamma} = 0$$

et que

$$\gamma'(0) = \alpha'(x)^{-1} \cdot \tilde{\gamma}'(0) = \alpha'(x)^{-1} \cdot \tilde{\xi} = \xi.$$

□

11.d Exercice. Calculer le plan tangent à \mathbb{S}^n en un point x .

11.e Exercice. Calculer la droite tangente à l'ellipse d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

11.7 Définition. Soient S et T des surfaces respectivement de \mathbb{R}^n et de \mathbb{R}^p . Soit $f : (S, a) \rightarrow (T, b)$ une application. La *dérivée* de f en a , si elle existe, est l'application $f'(a) : T_a S \rightarrow T_b T$ définie par

$$f'(a) \cdot \gamma'(0) = (f \circ \gamma)'(0),$$

pour tout chemin $\gamma : (\mathbb{R}, 0) \rightarrow (S, a)$ dérivable tracé sur S .

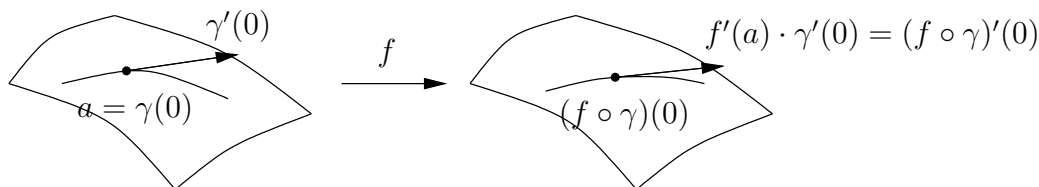


FIGURE 11.4 – Définition cinématique de la dérivée $f'(a)$: c'est l'application qui envoie la vitesse d'un chemin sur la vitesse du chemin image

Démonstration (de quoi ?) Il faut démontrer que, si $\tilde{\gamma} : (\mathbb{R}, 0) \rightarrow (S, x)$ est un chemin tel que $\tilde{\gamma}'(0) = \gamma'(0)$, $(f \circ \tilde{\gamma})'(0) = (f \circ \gamma)'(0)$.

Il existe un difféomorphisme de redressement $\alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^d \times \mathbb{R}^{n-d}, 0)$ tel que $\alpha(S) = \{(x, 0), x \in \mathbb{R}^d\}$. Donc $\alpha \circ \gamma(t)$ est de la forme $(x(t), 0)$, et $\alpha \circ \tilde{\gamma}(t)$ de la forme $(\tilde{x}(t), 0)$. D'après la formule habituelle de dérivation des fonctions composées (définies sur des ouverts d'espaces euclidiens, et non sur des surfaces),

$$\begin{aligned} (\alpha \circ \gamma)'(0) &= \alpha'(a) \cdot \gamma'(0) \\ &= \alpha'(a) \cdot \tilde{\gamma}'(0) \\ &= (\alpha \circ \tilde{\gamma})'(0), \end{aligned}$$

donc $x'(0) = \tilde{x}'(0)$. Posons $F = f \circ \alpha^{-1}$. Alors

$$\begin{aligned}
 (f \circ \gamma)'(0) &= (F \circ \alpha \circ \gamma)'(0) \\
 &= (F \circ (x, 0))'(0) \\
 &= F'(\alpha(a)) \cdot (x'(0), 0) & (*) \\
 &= F'(\alpha(a)) \cdot (\tilde{x}'(0), 0) \\
 &= (F \circ (\tilde{x}, 0))'(0) & (*) \\
 &= (F \circ \alpha \circ \tilde{\gamma})'(0) \\
 &= (f \circ \tilde{\gamma})'(0),
 \end{aligned}$$

où les égalités (*) découlent de la formule habituelle de dérivation des fonctions composées. \square

11.8 Proposition. *Si S est une surface de \mathbb{R}^n de dimension d , au voisinage de tout point a de S il existe une base $e(x) = (e_1(x), \dots, e_d(x))$ de $T_x S$ telle que $e : (S, x) \rightarrow M_{n,d}(\mathbb{R})$ soit de classe C^1 .*

Démonstration. Soit $\alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^d \times \mathbb{R}^{n-d}, 0)$, $x \mapsto (X, Y)$, un difféomorphisme de redressement de S , de sorte que $\alpha(S)$ ait pour équation $Y = 0$. Soit (e_1, \dots, e_d) la base canonique de \mathbb{R}^d . C'est une base de l'espace tangent à $\alpha(S)$ en tout point voisin de 0. Alors $(\alpha'(x)^{-1}(e_i))_{i=1, \dots, d}$ est une base de $T_x S$ (parce que $\alpha'(x)$ est un isomorphisme), ce qui démontre la proposition, au changement de notation près. \square

11.f Exercice (Transversalité).

Soient $f : (\mathbb{R}^n, x) \rightarrow (\mathbb{R}^p, y)$ C^1 et S une surface de \mathbb{R}^p contenant y , de dimension d . On suppose que f est *transverse* à S en x , au sens que

$$\text{im } f'(x) + T_y S = \mathbb{R}^p.$$

Montrer que $f^{-1}(S)$ est localement une surface de \mathbb{R}^n de dimension $n - p + d$.

11.g Exercice (Contour apparent d'une surface).

Soient Q une forme quadratique définie positive sur \mathbb{R}^3 et S l'*ellipsoïde* d'équation $Q = 1$. Soit a un point à l'extérieur de S , c'est-à-dire tel que $Q(a) > 1$ (pourquoi est-ce bien l'extérieur de S ?). Le *contour apparent* de S vu de a est l'ensemble C des points x de S où l'espace affine tangent $x + T_x S$ contient a . Montrer que C est une ellipse (intersection de S et d'un plan affine de \mathbb{R}^3).

Chapitre 12

Extrema de fonctions (semi-)continues

Mots-clefs du chapitre Suite minimisante, minimum, continuité, semi-continuité, compacité

Soit $f : X \rightarrow \mathbb{R}$ une fonction sur un espace métrique. Soit

$$m = \inf_X f \in [-\infty, +\infty[$$

la borne inférieure de f . Une *suite minimisante* de f est une suite (x_n) de X telle que

$$\lim f(x_n) = m;$$

une telle suite existe forcément, par définition de la borne inférieure. Mais on ne sait pas, en général, si elle converge dans X .

12.1 Définition. Un *minimum* est un point $x \in X$ tel que $f(x) = m$ (forcément $\in \mathbb{R}$ par définition de f). Le nombre m s'appelle alors la *valeur minimum*.

On définit de même les notions de borne supérieure, de suite maximisante, de maximum et de valeur maximale. Un *extremum* est un point x qui est soit un minimum soit un maximum.

Par exemple, la fonction x^2 sur $X = \mathbb{R}$ possède un minimum en $x = 0$, mais pas de maximum puisque sa borne supérieure est $+\infty$.

12.2 Théorème. *Si X est compact, une fonction continue $f : X \rightarrow \mathbb{R}$ possède (au moins) un minimum et un maximum.*

Démonstration. Soit (x_n) une suite minimisante. Comme X est supposé compact, il existe une suite extraite (x_{n_k}) convergente dans X , disons vers un point x . Par continuité de f , $f(x) = m$, donc x est un minimum. \square

12.3 Exemple. Soient C une courbe fermée de \mathbb{R}^n (*fermée* signifie qu'elle possède un paramétrage $c : \mathbb{T} \rightarrow C$) et $f : C^3 \rightarrow \mathbb{R}$ définie par

$$f(x, x', x'') = \|x - x'\| + \|x' - x''\| + \|x'' - x\|.$$

Alors f possède un minimum et un maximum, puisque f est continue et que, comme \mathbb{T} est compact, C l'est aussi. Les minima sont en nombre infinis, puisqu'ils contiennent déjà tous les triplets de la forme (x, x, x) , $x \in C$. À supposer que C ne contienne pas de triplets de points alignés (par exemple si C est strictement convexe), les maxima, eux, sont constitués de points deux à deux distincts, puisque, si $x \neq x'$,

$$f(x, x', x'') > f(x, x, x'').$$

12.a Exercice. Soient X l'ensemble des fonctions f croissantes de classe C^1 sur \mathbb{R} telles que $f(-1) = -1$ et $f(1) = 1$, et $J : X \rightarrow \mathbb{R}$ telle que

$$J(f) = \int_{-1}^{+1} x^2 f'(x) dx.$$

Montrer que $\inf_X J = 0$ et que J n'atteint pas sa borne inférieure.

12.4 Remarque. En fonction de la "forme" de X , on peut souvent inférer l'existence, non seulement d'extrema, mais plus généralement de points critiques des fonctions dérivables sur X . Leur nombre minimal s'appelle la *catégorie de Lyusternik-Schnirelmann* k et, s'ils sont non dégénérés (c'est-à-dire si la fonction est de Morse), le nombre de points critiques minimum est le *nombre de Betti* $\geq k$. Un point critique qui n'est pas un extremum local s'appelle un *point col* (voir la figure 12.1).

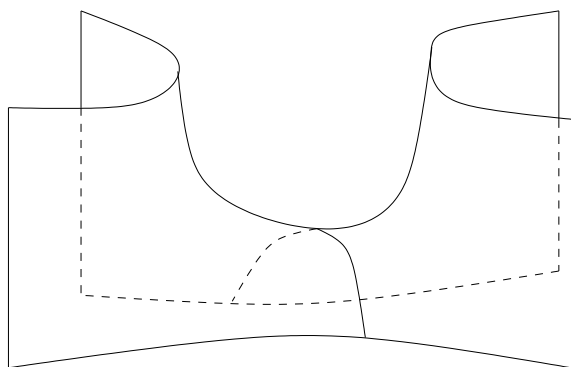


FIGURE 12.1 – Graphe d'une fonction (telle que $x^2 - y^2$) possédant un point col.

Dans le théorème 12.2, si l'on cherche seulement un minimum de f , toute la force de la continuité de f n'est pas requise.

Rappel : la limite inférieure d'une fonction f en un point a est

$$\liminf_{x \rightarrow a} f(x) = \lim_{\eta \rightarrow 0} \inf f(B_\eta(a));$$

cette limite existe parce que la fonction $\eta \mapsto \inf f(B_\eta(a))$ est croissante.

12.5 Définition. Soient $f : X \rightarrow \bar{\mathbb{R}}$ une fonction et $a \in X$ un point. La fonction f est *semi-continue inférieurement* (sci) en a si l'une de ces deux conditions équivalentes suivantes est satisfaite :

$$- f(a) = \liminf_{x \rightarrow a} f(x).$$

— pour tout $\epsilon > 0$ il existe un voisinage V de a tel que $f(a) - \epsilon < f(V)$.
Elle est *semi-continue supérieurement* en a si $-f$ est semi-continue inférieurement en a .

Remarquons que, si f est sci en a , on a donc

$$f(a) \leq \limsup_{x \rightarrow a} f(x),$$

avec égalité si et seulement si f est continue en a .

12.6 Exemple. La fonction indicatrice de $]0, 1[$ est sci, tandis que celle de $[0, 1]$ est scs.

12.7 Exemple. La fonction indicatrice de \mathbb{Q} sur \mathbb{R} est sci en tout point x irrationnel, et scs en tout point x rationnel.

12.8 Théorème. *Si X est compact, une fonction semi-continue inférieurement $f : X \rightarrow \bar{\mathbb{R}}$ possède un minimum.*

Démonstration. Soit $m = \inf_X f$. Soit (x_n) une suite minimisante de f : pour tout $\epsilon > 0$, il existe N tel que, pour tout $n \geq N$, $f(x_n) \leq m + \epsilon$.

Par compacité de X , il existe une suite extraite (x_{n_k}) convergente dans X . Notons a sa limite. Par semi-continuité de f en a , il existe K tel que, pour tout $k \geq K$, $f(a) - \epsilon \leq f(x_{n_k})$.

Donc, pour $k \geq K$ tel que $n_k \geq N$, $f(a) \leq m + 2\epsilon$. Donc $f(a) = m$. □

Terminons ce chapitre par quelques propriétés des fonctions semi-continues, qui seront utiles dans l'exercice 12.b.

12.9 Proposition. *L'application f est semi-continue inférieurement sur X si et seulement si l'une quelconque de ces propriétés est satisfaite :*

1. pour tout $\lambda \in \bar{\mathbb{R}}$ l'ensemble $f^{-1}(] \lambda, +\infty])$ est ouvert.
2. pour tout $\lambda \in \bar{\mathbb{R}}$ l'ensemble $f^{-1}([-\infty, \lambda])$ est fermé.
3. l'épigraphe

$$E_f = \{(x, y) \in X \times \bar{\mathbb{R}} ; y \geq f(x)\}$$

est fermé dans $X \times \bar{\mathbb{R}}$.

Démonstration. 1. La semi-continuité inférieure de f équivaut à ce que, pour tout point $a \in X$ et pour tout λ tels que $\lambda < f(a)$, l'ensemble des x tels que $\lambda < f(x)$ soit un voisinage de a ; autrement dit, à ce que, pour tout λ , l'ensemble des a tels que $\lambda < f(a)$ soit un voisinage de chacun de ses points, donc soit ouvert.

2. Comme $\complement f^{-1}([-\infty, \lambda]) = f^{-1}(] \lambda, +\infty])$, la seconde propriété est équivalente à la première.

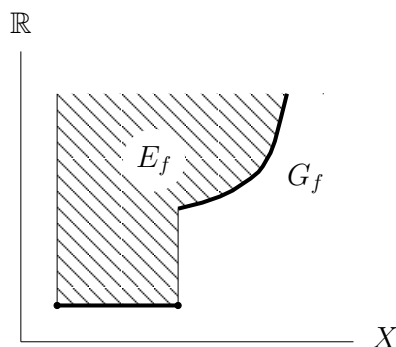


FIGURE 12.2 – Épigraphe d'une fonction semi-continue inférieurement

3. E_f est fermé si et seulement si son complémentaire est voisinage de chacun de ses points. Or, f est semi-continue inférieurement sur X si et seulement si, pour tout couple $(x, \lambda) \in \complement E_f$ ($\lambda < f(x)$) et pour tout μ tel que $\lambda < \mu < f(x)$, on a $\mu < f(y)$ pour tout y dans un voisinage V de x ; c'est-à-dire s'il existe un voisinage $V \times]-\infty, \mu[\subset \complement E_f$ de (x, λ) . Donc la troisième propriété énoncée dans la proposition est bien équivalente à la semi-continuité inférieure de f sur X .

□

12.10 Exemple. La fonction caractéristique $\mathbb{1}_A$ d'une partie A de X est sci sur X si et seulement si A est ouverte. En effet,

$$\mathbb{1}_A^{-1}(] \lambda, +\infty]) = \begin{cases} \emptyset & \text{si } 1 \leq \lambda \\ A & \text{si } 0 \leq \lambda < 1 \\ X & \text{si } \lambda < 0. \end{cases}$$

La borne supérieure d'une famille de fonctions continues n'est pas, en général, continue, mais seulement semi-continue inférieurement. Ceci découle du corollaire suivante.

12.11 Corollaire. *L'enveloppe supérieure $\sup\{f_j\}_{j \in J}$ d'une famille $(f_j)_{j \in J}$ de fonctions numériques semi-continues inférieurement est aussi semi-continue inférieurement.*

Démonstration. L'épigraphe de l'enveloppe supérieure est l'intersection des épigraphes :

$$E_{\sup f_j} = \bigcap E_{f_j}.$$

D'après la proposition 12.9, les E_{f_j} sont fermés. Donc leur intersection est fermée. Donc l'épigraphe de $\sup f_j$ est fermé, d'où l'affirmation. □

12.b Exercice (Géodésiques dans un espace métrique).

1. Montrer que, si un espace métrique X est compact, une fonction semi-continue inférieurement sur X atteint son minimum.

2. Montrer que la longueur d'un chemin continu (cf. exercice 1.h) dépend du chemin

de façon semi-continue inférieure; on pourra utiliser que la borne supérieure de fonctions continues est semi-continue inférieurement.

3. Montrer que, pour toute courbe $\gamma \in C(T, X)$, $T = [a, b]$, $a < b$, telle que $L(\gamma) < l \in \mathbb{R}^+$ il existe un homéomorphisme croissant $\phi : [0, 1] \rightarrow [a, b]$ tel que la courbe $\gamma \circ \phi$ soit l -lipschitzienne.

4. En déduire que, pour toute famille $(\gamma_j)_{j \in J}$ de courbes paramétrées de X de longueur $< l$, il existe une famille $(\hat{\gamma}_j)_{j \in J}$ de courbes paramétrées appartenant à $C([0, 1], X)$, telle que pour tout $j \in J$ les courbes γ_j et $\hat{\gamma}_j$ soient équivalentes par homéomorphisme lipschitzien et que $\hat{\gamma}_j$ soit lipschitzienne de rapport l .

5. ★ Supposons compact l'espace métrique X . Soient A et B deux parties fermées disjointes de X . Montrer que, s'il existe dans X des courbes rectifiables d'extrémités dans A et B respectivement, et si k désigne la borne inférieure de leurs longueurs, il existe aussi un arc simple γ de longueur k et d'extrémités dans A et B respectivement (γ est une *géodésique*).

Chapitre 13

Factorisation des applications linéaires

Mots-clefs du chapitre Espace vectoriel quotient, factorisation d'une application linéaire par une autre

On commence par rappeler le concept de quotient d'espaces vectoriels. Soient E un espace vectoriel et V un sous-espace vectoriel de E . On cherche à construire un espace vectoriel ressemblant le plus possible à E mais dans lequel on ait la règle de calcul supplémentaire selon laquelle tout vecteur de V peut être identifié à 0.

13.1 Définition. Le *quotient* de E par V est l'ensemble, noté E/V , des classes d'équivalences de E pour la relation d'équivalence

$$x \sim x' \Leftrightarrow x - x' \in V \Leftrightarrow x' \in x + V.$$

Autrement dit, E/V est l'ensemble des sous-espaces affines de E dirigés par V :

$$E/V = \{x + V, x \in E\}.$$

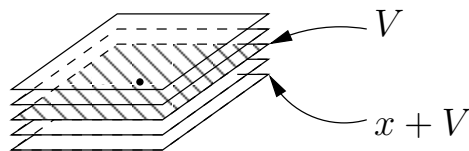


FIGURE 13.1 – La partition de E en les sous-espaces affines $x + V$, $x \in E$, dirigés par V

Les deux cas extrêmes sont

$$E/\{0\} = \{x + \{0\}, x \in E\} = \{\{x\}, x \in E\} \equiv E$$

et

$$E/E = \{x + E, x \in E\} = \{E\} \equiv \{0\}.$$

13.a Exercice. Montrer que le quotient E/V d'un espace vectoriel E par un sous-espace V hérite d'une structure d'espace vectoriel définie par les relations

$$\begin{cases} \bar{x} + \bar{x}' = \overline{x + x'} \\ \lambda \cdot \bar{x} = \overline{\lambda \cdot x}, \end{cases}$$

et que, si E est de dimension finie,

$$\dim E/V = \text{codim } V.$$

13.b Exercice (Réalisation du quotient). Montrer que, si $E = V \oplus W$, la restriction à W de la projection canonique $p : E \rightarrow E/V$ est un isomorphisme.

13.c Exercice (Propriété universelle du quotient). Montrer que, pour toute application linéaire $f : E \rightarrow F$ s'annulant sur V , il existe une unique application linéaire $\varphi : E/V \rightarrow F$ telle que

$$f = \varphi \circ p,$$

où p est la surjection canonique de E sur E/V . Comme E/V ne dépend pas de f , la propriété précédente est qualifiée de *propriété universelle du quotient*. Montrer de plus que p est unique à isomorphisme près parmi les applications linéaires sur E s'annulant sur V et vérifiant la propriété universelle.

13.2 Lemme. Une application linéaire $f : E \rightarrow F$ induit une application linéaire $\varphi : E/\ker f \rightarrow \text{im } \varphi$, qui est un isomorphisme.

Démonstration. φ est définie par la formule

$$\varphi(x + \ker f) = f(x),$$

qui n'est pas ambiguë. Elle a même image que f , et est injective parce que, si $f(x) = 0$, $x + \ker f = \ker f$. \square

13.3 Proposition. Soient E , V et W trois espaces vectoriels et $\alpha \in L(E, V)$ et $\beta \in L(E, W)$ deux applications linéaires. Les deux propriétés suivantes sont équivalentes :

— Il existe une application linéaire $\lambda \in L(W, V)$ telle que $\alpha = \lambda \circ \beta$:

$$\begin{array}{ccc} E & \xrightarrow{\alpha} & V \\ & \searrow \beta & \uparrow \lambda? \\ & & W \end{array} .$$

— $\ker \beta \subset \ker \alpha$.

Alors la restriction de λ à $\text{im } \beta$ est unique.

Si β est inversible, il suffit de poser $\lambda = \alpha \circ \beta^{-1}$. Mais si ce n'est pas le cas, il n'y a pas de formule aussi simple donnant λ . La démonstration consiste à se ramener au cas inversible.

Démonstration du lemme. L'implication directe est évidente. Montrons l'implication réciproque. Les applications α et β induisent des isomorphismes

$$\begin{cases} \hat{\alpha} : E/\ker \alpha \rightarrow \operatorname{im} \alpha \subset V \\ \hat{\beta} : E/\ker \beta \rightarrow \operatorname{im} \beta \subset W. \end{cases}$$

Par ailleurs, la projection canonique $E \rightarrow E/\ker \alpha$ induit une projection $E/\ker \beta \xrightarrow{p} E/\ker \alpha$ puisque que par hypothèse $\ker \beta \subset \ker \alpha$:

$$\begin{array}{ccccc} E & \xrightarrow{\alpha} & \operatorname{im} \alpha & \xleftarrow{\hat{\alpha}} & E/\ker \alpha . \\ & \searrow \beta & \uparrow \lambda|_{\operatorname{im} \beta} & & \uparrow p \\ & & \operatorname{im} \beta & \xleftarrow{\hat{\beta}} & E/\ker \beta \end{array}$$

Il suffit donc de poser $\lambda = \hat{\alpha} \circ p \circ \hat{\beta}^{-1}$ et de prolonger λ , sur un supplémentaire de $\operatorname{im} \beta$, par exemple par 0 . □

13.d Exercice.

Démontrer la proposition 13.3 sans utiliser de quotients d'espaces vectoriels. Indication : choisir des supplémentaires de $\ker \alpha$ et de $\ker \beta$.

Chapitre 14

Points critiques de fonctions

Mots-clefs du chapitre Extremum, point critique, multiplicateur de Lagrange
Soient D une partie de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$ une fonction de classe C^1 au sens où f possède un prolongement C^1 sur un voisinage ouvert de D .¹ Trouver les extrema (minima et maxima) de f s'appelle *optimiser f sous la contrainte D* .

Dans le chapitre précédent, on a vu que si D est compact, f possède toujours un maximum et un minimum. Mais comment les trouver ? La réponse n'est pas toujours immédiate, et il faut alors décomposer D en plusieurs sous-parties, chercher les extrema (voire les extrema locaux) de f en restriction à chacune de ces sous-parties, pour enfin comparer les valeurs prises par f en ces différents points pour décider lesquels de ces points sont les extrema sur D entier.

Les sous-parties que nous allons considérer ici sont : les ouverts de \mathbb{R}^n et les surfaces de \mathbb{R}^n .

14.1 Exemple. Si D est le carré fermé $[0, 1]^2$, avec les outils de ce cours on décomposera D en l'ouvert $]0, 1[^2$, les quatre côtés ouverts ($]0, 1[\times \{0\}$, etc.) et les quatre sommets $((0, 0), \text{etc.})$.

Optimisation sur un ouvert de \mathbb{R}^n

Commençons par quelques rappels sur l'optimisation locale sur un ouvert de \mathbb{R}^n . Soit $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}$ une fonction de classe C^1 .

14.2 Définition. Le point a est *régulier* si $f'(a) \neq 0$ et *critique* si $f'(a) = 0$.²

14.3 Lemme. *Si a est un extremum local³ de f , il en est un point critique.*

1. Comme on n'a pas supposé que D est ouvert, f n'est peut-être pas définie au voisinage de chacun des points de D . Donc la propriété d'être de classe C^1 requiert bien une définition, qui prolonge celle où D est ouvert.

2. Plus généralement, pour une application $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^p$, a est *critique* si $f'(a) < \min(n, p)$. Cette définition est cohérente avec la notion de point critique introduite au premier chapitre pour les chemins ($n = 1$).

3. C'est-à-dire un minimum local ou un maximum local

Démonstration. Comme f est dérivable,

$$f(a + \xi) = f(a) + f'(a) \cdot \xi + o(\xi) = f(a) + \sum_{1 \leq j \leq n} \partial_j f(x) \xi_j + o(\xi).$$

Si au moins l'une des dérivées partielles $\partial_j f(x)$ est non nulle, avec $\xi = (0, \dots, 0, 1, 0, \dots, 0)$ (avec un 1 en position j) la formule précédente devient

$$f(a + \xi) = f(a) + \partial_j f(x) \xi_j (1 + o(1)),$$

ce qui montre que a n'est pas un maximum local de f . □

14.4 Corollaire. *Si f est localement constante, $f'(x) \equiv 0$ au voisinage de a (puisque tous les points sont des extrema locaux).*

14.5 Remarque. Un point critique n'est pas forcément un extremum, même localement. Sur la figure 14.1, on voit les exemples les plus simples à avoir en tête, avec deux extrema locaux en haut, et, en bas, un point critique dégénéré (au sens défini dans le chapitre sur les fonctions de Morse : la dérivée seconde s'annule) et un *point selle* (où la partie quadratique n'est ni positive ni négative). Le lemme de Morse permettra de ramener à de l'algèbre bilinéaire la question de savoir si un point critique non dégénéré est un extremum local ou non.

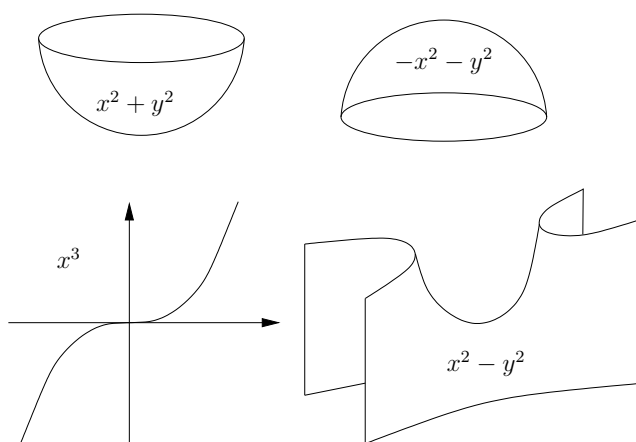


FIGURE 14.1 – Les points critiques ne sont pas tous des extrema locaux.

Optimisation sur une surface de \mathbb{R}^n

Nous allons trouver une condition nécessaire analogue pour qu'un point soit un extremum local d'une fonction *en restriction à une surface*. Soit S une surface de \mathbb{R}^n contenant a et d'équation locale $g = 0$, où $g : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^p$ est une submersion. Comment détecter les extrema locaux de la restriction de f à S ?

14.6 Définition. Le point a est un *point critique* de $f|_S$ si $f'(a)|_{T_a S} = 0$.

Pourquoi cette définition ? Parce ce que $f'(a)|_{T_a S} = 0$ s'interprète comme la dérivée de $f|_S$, puisque, pour tout chemin $C^1 \gamma : (\mathbb{R}, 0) \rightarrow (\mathbb{R}^n, a)$ tracé sur S ,

$$(f \circ \gamma)'(0) = f'(a) \cdot \gamma'(0)$$

avec $\gamma'(0) \in T_a S$.

14.7 Lemme. *Si a est un extremum local de $f|_S$, a en est un point critique.*

Démonstration. Si a est un extremum local de $f|_S$, c'est en particulier un extremum local de la fonction

$$f|_S \circ \gamma = f \circ \gamma : (\mathbb{R}, 0) \rightarrow \mathbb{R}$$

pour tout chemin $C^1 \gamma : (\mathbb{R}, 0) \rightarrow (S, a)$. Donc (on peut ici invoquer le lemme 14.3),

$$(f \circ \gamma)'(0) = f'(a) \cdot \gamma'(0) = 0.$$

Donc

$$f'(a)|_{T_a S} = 0.$$

□

Une première façon de s'attaquer au problème d'optimisation consiste à se ramener au paragraphe précédent en choisissant des coordonnées sur S (ce concept sera formalisé au chapitre suivant), en exprimant $f|_S$ dans ces coordonnées, et en cherchant les extrema locaux de $f|_S$ avec la condition nécessaire du lemme 14.3.

14.8 Exemple. Un point (a, b) est un extremum local de la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ restreinte à la courbe $S : \sin x + y = 0$ si et seulement si $b = -\sin a$ et a est un extremum local de la fonction d'une variable $\varphi : x \mapsto f(x, -\sin x)$ (et donc $\varphi'(a) = 0$).

Mais cette méthode conduit parfois à des calculs compliqués pouvant être évités. Avec le petit ingrédient supplémentaire d'algèbre linéaire vu dans le chapitre précédent, on obtient en effet le théorème suivant, très utile pour détecter les points critiques sur une surface.

14.9 Théorème. *Si a est un point critique de $f|_S$, il existe une forme linéaire $\lambda \in (\mathbb{R}^p)^*$ telle que*

$$f'(a) = \lambda \cdot g'(a).$$

La forme linéaire λ est alors unique et s'appelle le multiplicateur de Lagrange.

Par définition, l'espace tangent en a à la surface $S = \{x, g(x) = 0\}$ est l'ensemble des vecteurs vitesses au point a de chemins tracés sur S . On a montré que $T_a S = \ker g'(a)$.

Démonstration du théorème. Supposons que a soit un point critique de $f|_S$. D'après le lemme 13.3, pour conclure il suffit de voir que $\ker g'(a) \subset \ker f'(a)$, ce qui n'est autre que l'hypothèse selon laquelle a est un point critique de $f|_S$. D'après le même lemme, le multiplicateur de Lagrange est unique sur l'image de $g'(a)$, c'est-à-dire, puisque g est une submersion, sur \mathbb{R}^p entier. □

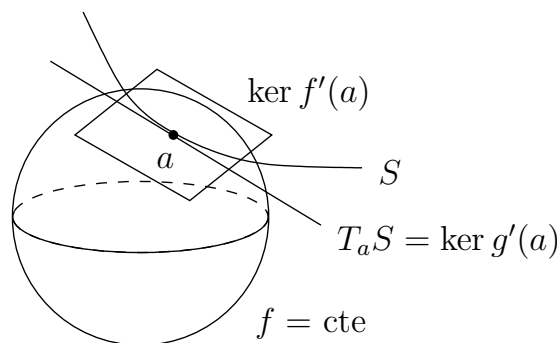


FIGURE 14.2 – Extremum lié

14.a Exercice (Parallélépipèdes rectangles de volume maximal).

Quel est le parallélépipède rectangle (engendré par trois vecteurs orthogonaux, donc) d'aire $A > 0$ donnée et de volume maximal? On commencera par justifier l'existence de ce maximum, puis on donnera deux calculs : un direct en utilisant des coordonnées de l'espace des parallélépipèdes rectangles d'aire donnée, et un utilisant un multiplicateur de Lagrange.

14.b Exercice (Maximisation d'un volume).

Maximiser la fonction xyz sur la courbe obtenue par intersection de la sphère $\mathbb{S}^2 : x^2 + y^2 + z^2 = 1$ et du plan $x + y + z = 1$.

14.c Exercice (Formule de Héron). Application : Montrer que l'aire d'un triangle dont le périmètre est fixé est maximale lorsque le triangle est équilatérale; on pourra commencer par démontrer que l'aire A d'un triangle de côtés a , b et c vérifie

$$A^2 = p(p-a)(p-b)(p-c), \quad p = \frac{a+b+c}{2}.$$

14.d Exercice (Axes principaux d'un ellipsoïde). Soit q une forme quadratique positive définie sur \mathbb{R}^n . Quels sont les extrema de $\|x\|$ sur l'hypersurface d'équation $q = 1$ en fonction des sous-espaces propres de q ?

14.e Exercice (Diagonalisation d'un opérateur symétrique).

Soit q une forme quadratique sur \mathbb{R}^n associées à $B \in L_s^2(\mathbb{R}^n, \mathbb{K})$. Montrer qu'un maximum x de la restriction de q à la sphère \mathbb{S}^{n-1} est une valeur propre de B . En déduire par récurrence que B est diagonalisable sur \mathbb{R} , dans une base orthonormée.

14.f Exercice (Inégalité de Minkowski).

Soit $p \geq 1$ un réel. Notons

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

pour tout $x \in \mathbb{R}^n$. Pour tous vecteurs $x, y \in \mathbb{R}^n$ on a

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

(ce qui montre que $\|\cdot\|_p$ est une norme).

14.g Exercice (Inégalité de Hadamard).

Montrer que, pour tous vecteurs $x_1, \dots, x_n \in \mathbb{R}^n$,

$$|\det(x_1, \dots, x_n)| \leq \|x_1\| \cdots \|x_n\|,$$

où $\|\cdot\|$ est la norme euclidienne de \mathbb{R}^n ; on pourra commencer par montrer que, si x_2, \dots, x_n est une famille libre donnée et si x_1 est un maximum de la fonction $\xi^{n-1} \rightarrow \mathbb{R}, \xi \mapsto \det(\xi, x_2, \dots, x_n)$, x_1 est orthogonal à x_2, \dots, x_n , puis en déduire les maxima de $\det(x_1, \dots, x_n)$ lorsque $x_1, \dots, x_n \in \xi^{n-1}$.

14.h Exercice (Inégalité de Hölder).

Quels que soient les entiers naturels non nuls p, q, r tels que $1/p + 1/q = 1/r$ et quels que soient les vecteurs $x, y \in \mathbb{R}^n$ on a

$$\|(x_1 \cdot y_1, \dots, x_n \cdot y_n)\|_r \leq \|x\|_p \|y\|_q$$

où, rappelons-le, la norme ℓ^p est définie par

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

Chapitre 15

Surfaces II – Coordonnées

Mots-clefs du chapitre Système de coordonnées locales, application lue dans des coordonnées, dérivabilité des applications entre surfaces, difféomorphisme

15.1 Rappel. Une partie S de \mathbb{R}^n est une *surface* de dimension d si, au voisinage de chacun de ses points a , il existe un difféomorphisme de redressement de S , soit un difféomorphisme

$$\alpha = (\alpha_X, \alpha_Y) : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^d \times \mathbb{R}^{n-d}, (b, c)), \quad x \mapsto (X, Y)$$

tel que $\alpha(S)$ ait pour équation $Y = c$ (figure 11.1). On dit que S est de classe C^k si on peut choisir α de classe C^k ($k \geq 1$).

Avec les notations ci-dessus, on remarque que, S ayant pour équation $Y = \alpha_Y(x) = c$, les points de S se distinguent les uns des autres par leur “coordonnée” $X = \alpha_X(x)$.

(On commet parfois l’abus consistant à noter X à la place de α_X .)

15.2 Définition. Une bijection $(S, a) \rightarrow (\mathbb{R}^d, b)$ de la forme $x \mapsto X = \alpha_X(x)$ s’appelle un *système de coordonnées (locales)* de S .

15.3 Exemple. Considérons la droite D d’équation $x + y = 1$ dans \mathbb{R}^2 . L’application

$$\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (X, Y) = (x, 1 - x - y)$$

est un redressement (global) de D ($\alpha(D)$ ayant pour équation $Y = 0$). Donc une coordonnée (ou un système de coordonnées) est l’application $D \mapsto \mathbb{R}$, $(x, y) \mapsto x$. Une autre coordonnée est l’application $(x, y) \mapsto y$. En fait, n’importe quelle forme linéaire λ de \mathbb{R}^2 non constante sur D (c’est-à-dire non proportionnelle à $(x, y) \mapsto -x - y$) convient, d’après le théorème d’inversion locale.

15.4 Exemple. L’application “coordonnées sphériques” est

$$\sigma : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} r \\ \theta \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \end{pmatrix}.$$

Cette application n’est pas inversible, pour deux “raisons” :

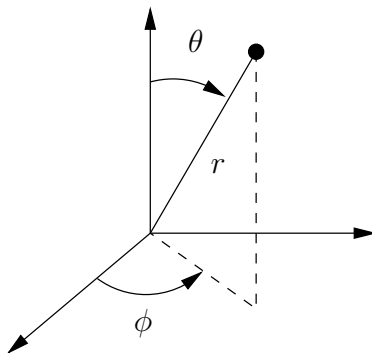


FIGURE 15.1 – Coordonnées sphériques

— Sa dérivée a pour déterminant

$$r^2 \begin{vmatrix} \sin \theta \cos \phi & \cos \theta \cos \phi & -\sin \theta \sin \phi \\ \sin \theta \sin \phi & \cos \theta \sin \phi & \sin \theta \cos \phi \\ \cos \theta & -\sin \theta & 0 \end{vmatrix} = r^2 \sin \theta,$$

qui s'annule en $\theta = 0 \pmod{\pi}$ et en $r = 0$. Donc σ n'est pas un difféomorphisme local en ces points.

— Les angles θ et ϕ ne sont pas définis globalement, puisque σ est 2π -périodique par rapport à θ et à ϕ . Tout au plus peut-on choisir une détermination continue de θ et de ϕ sur des ouverts suffisamment petit de \mathbb{R}^3 privé de la droite $0z$.

Localement, en dehors de $\theta = 0 \pmod{\pi}$ et de $r = 0$, l'application σ définit un difféomorphisme local. Une branche inverse quelconque de σ redresse la sphère \mathbb{S}^2 , puisque l'équation de \mathbb{S}^2 dans ces coordonnées est

$$r = 1.$$

Donc les variables (θ, ϕ) forment un système de coordonnées locales de la sphère \mathbb{S}^2 , localement au voisinage de n'importe quel point différent des pôles. C'est ce qu'on appelle les *coordonnées sphériques*.

15.a Exercice (Graphe d'une application).

Montrer que le graphe d'une application $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$, soit

$$G = \{(x, f(x)), x \in \mathbb{R}^n \text{ proche de } a\}$$

est une surface de $\mathbb{R}^n \times \mathbb{R}^p$ de dimension n , admettant la première projection $p_1 : (x, f(x)) \mapsto x$ comme système de coordonnées (globales).

15.b Exercice (Surfaces comme graphes).

Montrer réciproquement que, si S est une surface de dimension d de \mathbb{R}^n , il existe une décomposition en somme directe $\mathbb{R}^n = E \oplus F$ avec $\dim E = d$, et une application $\varphi : (E, a_1) \rightarrow (F, a_2)$ de classe C^1 (avec $a = a_1 + a_2$) telle que localement S soit le graphe de φ , i.e. la surface d'équation $y = \varphi(x)$ et

$$T_a S = \{(\xi, \varphi'(a_1) \cdot \xi); \xi \in E\}.$$

Soit $f : (S, a) \rightarrow (U, b)$ une application entre deux surfaces. Le calcul différentiel classique peut-être étendu aux telles applications de la façon suivante.

15.5 Définition. Soient $x : (S, a) \rightarrow (\mathbb{R}^d, 0)$ et $y : (U, b) \rightarrow (\mathbb{R}^p, 0)$ des systèmes de coordonnées locales. L'application f lue dans ces systèmes de coordonnées est l'application \tilde{f} faisant commuter le diagramme suivant :

$$\begin{array}{ccc} (S, a) & \xrightarrow{f} & (U, b) \\ x \downarrow & & \downarrow y \\ (\mathbb{R}^d, 0) & \xrightarrow{\tilde{f}} & (\mathbb{R}^p, 0). \end{array}$$

Contrairement à f , \tilde{f} est définie sur un ouvert d'espace euclidien, sur lequel les notions habituelles de calcul différentiel s'appliquent.

On avait défini cinématiquement la dérivée d'une application entre surfaces (définition 11.7) :

$$f'(a) \cdot \gamma'(0) = (f \circ \gamma)'(0) \quad (\gamma'(0) \in T_a S)$$

quand le membre de droite existe. Le lemme suivant découle de la formule

$$f \circ \gamma = y^{-1} \circ \tilde{f} \circ (x \circ \gamma).$$

15.6 Lemme. L'application f est dérivable en a si et seulement si \tilde{f} est dérivable en 0 .

15.7 Définition. Supposons S de classe C^k ($k \geq 1$). L'application f est de classe C^k si \tilde{f} est de classe C^k (au sens habituel). Elle est un C^k -difféomorphisme si elle est inversible et si f et f^{-1} sont de classe C^k , c'est-à-dire si \tilde{f} est un C^k -difféomorphisme.

15.c Exercice (Classe C^k).

Montrer que le fait d'être de classe C^k ne dépend pas des systèmes de coordonnées C^k x et y choisis.

15.8 Théorème (Inversion locale). Une application $f : (S, a) \rightarrow (U, b)$ de classe $C^{k \geq 1}$ est un C^k -difféomorphisme local si et seulement si l'application linéaire $f'(a) : T_a S \rightarrow T_b U$ est un isomorphisme.

Démonstration. f est un C^k -difféomorphisme local si et seulement si \tilde{f} est un C^k -difféomorphisme local (avec la notation de la définition 15.7), donc, d'après le théorème d'inversion locale, si et seulement si l'application linéaire $\tilde{f}'(0) : \mathbb{R}^d \rightarrow \mathbb{R}^p$ est un isomorphisme (et donc $d = p$). Comme

$$\tilde{f}'(0) = (y \circ f \circ x^{-1})'(0) = y'(b) \cdot f'(a) \cdot (x^{-1})'(0)$$

et comme $x'(a)$ et $y'(b)$ sont eux-mêmes des isomorphismes, le théorème est démontré. \square

15.d Exercice ★ (R. Krikorian).

1. Montrer que, si $H \in M_n(\mathbb{R})$ est antisymétrique, son exponentielle est dans $SO_n(\mathbb{R})$.

2. Soit $M \in M_n(\mathbb{R})$ telle que $\text{tr}(HM) = 0$ pour toute matrice H antisymétrique. Montrer que M est symétrique ; on pourra considérer le produit scalaire $H \cdot M = \text{tr}(H^t M)$.

3. Soit A une matrice symétrique définie positive dont les valeurs propres sont distinctes deux à deux. Calculer

$$\max\{\det(A + RAR^{-1}), R \in SO_n\}.$$

Solution succincte. 1. Dériver $(e^{tH}x) \cdot y$ par rapport à t .

2. Les matrices symétriques sont orthogonales aux matrices antisymétriques. Un argument de dimension permet de conclure.

3. En un maximum R , pour toute matrice antisymétrique H ,

$$\det(A + RAR^{-1} + [H, RAR^{-1}]) = O(H^2)$$

(remplacer R par $e^{\epsilon H}R$ et dériver par rapport à ϵ). Donc

$$\text{tr}((A + RAR^{-1})^{-1}[H, RAR^{-1}]) = 0,$$

soit

$$\text{tr}(H [(A + RAR^{-1})^{-1}, RAR^{-1}]) = 0.$$

D'après la question précédente,

$$S = [(A + RAR^{-1})^{-1}, RAR^{-1}]$$

est antisymétrique. Mais comme le crochet de deux matrices antisymétriques est symétrique, $S = 0$. Donc R et $(A + RAR^{-1})^{-1}$ et RAR^{-1} commutent, donc $A + RAR^{-1}$ et RAR^{-1} commutent, donc A et RAR^{-1} commutent.

Donc ces dernières sont simultanément diagonalisables dans une base orthonormée : il existe P orthogonale, et D et Δ diagonales telles que

$$A = PDP^{-1} \quad \text{et} \quad RAR^{-1} = P\Delta P^{-1}.$$

Posons $U = P^{-1}RP$. Alors $U\Delta U^{-1} = D$. Comme D a ses valeurs propres distinctes, P est une matrice de permutation. On est donc ramené à calculer le maximum de $\det(D + U^{-1}DU)$ lorsque U parcourt l'ensemble des matrices de permutation.

Si l'on note $\lambda_1 < \dots < \lambda_n$ les valeurs propres de A , et si U est la matrice d'une permutation σ ,

$$\det(D + U^{-1}DU) = \prod_{1 \leq i \leq n} (\lambda_i + \lambda_{\sigma(i)}).$$

Le maximum est atteint pour la permutation décroissante. En effet, si $a \leq b$ et $c \leq d$, on a

$$(a + c)(b + d) \leq (a + d)(b + c),$$

puisque

$$0 \leq (b - a)(d - c) = ac + bd - ad - bc.$$

Donc par récurrence on voit qu'on augmente le produit en ordonnant les seconds termes $\lambda_{\sigma(i)}$, dans chaque facteur, de façon décroissante. \square

Chapitre 16

Déterminant

Mots-clefs du chapitre Volume orienté, orientation

Si (x_1, \dots, x_n) est une famille de vecteurs de \mathbb{R}^n , on peut définir leur *déterminant* comme le volume orienté du parallélotope¹ engendré,

$$\{a_1x_1 + \dots + a_nx_n, a_1, \dots, a_n \in [0, 1]\}.$$

L'exemple suivant montre comment calculer ce volume orienté. Considérons deux vecteurs u et v dans \mathbb{R}^2 . Pour calculer l'aire orientée du parallélogramme qu'ils engendrent (voir la figure 16.1), on découpe le parallélogramme et l'on se ramène en deux étapes à un rectangle de même aire :

$$\text{aire}(u, v) = \text{aire}(u', v) = \text{aire}(u', v').$$

L'aire du parallélogramme initial est alors simplement le produit des longueurs de u' et de v' . Mais ces opérations géométriques correspondent exactement aux opérations sur les lignes et les colonnes d'une matrice 2×2 destinées à la ramener à une matrice diagonale, et dont chaque étape préserve le déterminant :

$$\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} = \begin{vmatrix} u'_1 & v_1 \\ 0 & v_2 \end{vmatrix} = \begin{vmatrix} u'_1 & 0 \\ 0 & v'_2 \end{vmatrix} = u'_1 v'_2.$$

Comme $u' = u - \frac{u_2}{v_2}v$ et $v' = v - \frac{v_1}{u_1}u'$, on a $u'_1 = u_1 - \frac{u_2}{v_2}v_1$ et $v'_2 = v_2$, donc l'aire orientée vaut

$$u'_1 v'_2 = u_1 v_2 - u_2 v_1.$$

Ces considérations conduisent à *définir* le déterminant

$$\det : (\mathbb{R}^n)^n \equiv M_n(\mathbb{R}) \rightarrow \mathbb{R}$$

comme une forme n -linéaire antisymétrique. On peut montrer que, si on la normalise de façon que

$$\det I_n = 1,$$

elle vaut

$$\det(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon(\sigma) x_{1, \sigma(1)} \cdots x_{n, \sigma(n)}.$$

1. Généralisation en dimension quelconque du parallélogramme en dimension $n = 2$ et du parallélépipède en dimension 3.

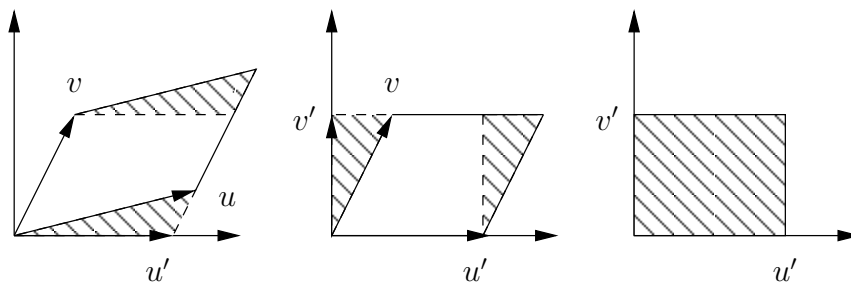


FIGURE 16.1 – Le calcul de l’aire orientée de 2 vecteurs dans \mathbb{R}^2 est le même que le calcul de leur déterminant.

16.1 Rappel (Signature). Tout permutation σ de $\{1, \dots, n\}$ se décompose comme composée de transpositions (i.e. de permutations qui, individuellement, ne font que permuter *deux* éléments). Le nombre de transpositions nécessaires n’est pas unique, mais sa parité l’est. La *signature* de σ est $\epsilon(\sigma) = \pm 1$ selon que σ se décompose en un nombre pair ou impair de transpositions. Par exemple, la signature d’une transposition est -1 , tandis que celle d’une bitransposition est $+1$.

Par exemple,

$$\det \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} = x_{1,1}x_{2,2} - x_{2,1}x_{1,2}.$$

16.a Exercice (Aire d’un triangle). **1.** Calculer l’aire orientée du triangle de sommets

$$A = (1, 2), \quad B = (3, 3), \quad C = (2, 1).$$

2. Calculer l’aire (non orientée) du triangle dans \mathbb{R}^3 de sommets

$$A = (1, 2, 3), \quad B = (4, 4, 4), \quad C = (3, 2, 1)$$

(une solution est d’utiliser le produit vectoriel, mais on essaiera plutôt de faire un calcul qui se généralise en dimension quelconque).

16.2 Remarque. La multilinéarité et l’antisymétrie permettent de calculer certains déterminants. Mais par exemple son caractère intrinsèque : $\det(P^{-1}AP) = \det A$, ni a fortiori son caractère multiplicatif : $\det(AB) = \det A \det B$, n’en découlent simplement. Ces propriétés découleront de la suite et sont provisoirement admises.

16.b Exercice (Matrices de Gram).

La *matrice de Gram*² de p vecteurs $x_1, \dots, x_p \in \mathbb{R}^n$ est la matrice $p \times p$ des produits scalaires,

$$G(x_1, \dots, x_p) = ((x_i | x_j))_{1 \leq i, j \leq p} \in M_p(\mathbb{R}),$$

2. Jørgen Pedersen GRAM, danois, n’étudia pas à l’Université Paris-Dauphine et fut pourtant mathématicien et actuaire (1850–1916).

où $(\cdot|\cdot)$ désigne le produit scalaire standard de \mathbb{R}^n (question complémentaire : refaire l'exercice avec un espace euclidien quelconque).

1. Montrer qu'une matrice $H \in M_p(\mathbb{R})$ est une matrice de Gram si et seulement si il existe $A \in M_{n,p}(\mathbb{R})$ telle que $H = {}^tAA$.

2. Montrer que la matrice $G(x_1, \dots, x_p)$ et la famille (x_1, \dots, x_p) ont même rang.

3. Si (f_1, \dots, f_p) est une famille libre de \mathbb{R}^n et $x \in \mathbb{R}^n$, montrer que la distance de x au sous-espace $F = \text{Vect}(f_1, \dots, f_p)$ est

$$d(x, F) = \sqrt{\frac{\det G(x, f_1, \dots, f_p)}{\det G(f_1, \dots, f_p)}}.$$

4. En déduire que $G(x_1, \dots, x_p)$ est le carré du volume du paralléloèdre

$$P(x_1, \dots, x_p) = \{\alpha_1 x_1 + \dots + \alpha_p x_p, 0 \leq \alpha_i \leq 1\}$$

engendré par les vecteurs $x_1, \dots, x_p \in \mathbb{R}^n$.

5. En déduire une interprétation géométrique de la formule de Binet-Cauchy³ : pour toute matrice $A \in M_{n,p}(\mathbb{R})$ (avec $p \leq n$),

$$\det({}^tAA) = \sum_{J \in \mathcal{J}_{n,d}} \det(A_J)^2,$$

où $\mathcal{J}_{n,d}$ est l'ensemble des parties à d éléments de $\{1, \dots, n\}$ (d étant un entier fixé) et où $A_J \in M_d(\mathbb{R})$ est la sous-matrice de A obtenue en ne gardant que les lignes et les colonnes dont les indices sont dans J . Que dit cette formule pour $d = 1$?

6. Question subsidiaire : Montrer la formule de Binet-Cauchy ; on pourra commencer par montrer que, si $M \in M_n(\mathbb{R})$,

$$\det(I_n + xM) = \sum_{d=1}^n T_d(M)x^d, \quad T_d(M) = \sum_{J \in \mathcal{J}_{n,d}} \det M_J.$$

16.c Exercice (Polynômes de Legendre).

Soient $I = [-1, 1]$ et $E = L^2(I)$ l'espace des fonctions réelles de carré intégrable sur I . Le produit scalaire

$$(f|g) = \int_I f(t) g(t) dt$$

fait de E un espace de Hilbert.

Soit $G_n = ((x^i|x^j)_{0 \leq i,j \leq n})$ la matrice de Gram des monômes x^i . Comme le coefficient (i, j) de G_n ne dépend que de $i + j$, on le notera γ_{i+j} .

1. Montrer que G_n est définie positive.

Soit $(p_k)_{k=0, \dots, n}$ la famille obtenue par orthonormalisation de la famille de monômes $(1, \dots, x^n)$.

2. Montrer que

$$p_n(x) = \frac{1}{\sqrt{\Delta_{n-1}\Delta_n}} q_n(x),$$

3. Jacques Philippe Marie BINET, mathématicien et astronome français (1786–1756)

où

$$q_n(x) = \det \begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & & \vdots \\ \gamma_{n-1} & \cdots & \gamma_{2n-1} \\ 1 & \cdots & x^n \end{pmatrix} \quad \text{et} \quad \Delta_n = \det G_n.$$

3. Montrons d'abord qu'il existe $c_n \in \mathbb{R}$ tel que $p_n = c_n q_n(x)$. Comme q_n est de degré n , il suffit de montrer que

$$(q_n(x)|x^i) = O \quad \forall i = 0, \dots, n-1.$$

Or

$$(q_n(x)|x^i) = \int_I \det \begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & & \vdots \\ \gamma_{n-1} & \cdots & \gamma_{2n-1} \\ x^i & \cdots & x^{n+i} \end{pmatrix} = \det \begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & & \vdots \\ \gamma_{n-1} & \cdots & \gamma_{2n-1} \\ \gamma_i & \cdots & \gamma_{n+i} \end{pmatrix}.$$

Dans ce dernier déterminant, la ligne i et la dernière sont les mêmes, donc le déterminant est nul.

De plus $\|p_n\|^1 = 1$. Or, le coefficient dominant de p_n est $c_n \Delta_{n-1}$.

Soit E un espace vectoriel réel de dimension finie n .

Soient $e = (e_1, \dots, e_n)$ et $f = (f_1, \dots, f_n)$ deux bases de E . Les applications de \mathbb{R}^n dans E définies par

$$\alpha : (x_1, \dots, x_n) \mapsto \sum_{1 \leq j \leq n} x_j e_j \quad \text{et} \quad \beta : (y_1, \dots, y_n) \mapsto \sum_{1 \leq j \leq n} y_j f_j$$

sont des isomorphismes. Le diagramme d'isomorphismes suivant définit une matrice carrée $\alpha \circ \beta^{-1} \in M_n(\mathbb{R})$:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\alpha} & E \\ \alpha \circ \beta^{-1} \downarrow & \nearrow \beta & \\ \mathbb{R}^n & & \end{array}$$

16.3 Définition. Les deux bases ont même *orientation* si le déterminant de la matrice $\alpha \circ \beta^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a un déterminant > 0 .

16.d Exercice. Montrer que le fait de définir une même orientation est une relation d'équivalence ; on utilisera la multiplicativité du déterminant.

16.4 Définition. Chacune des deux classes d'équivalences de bases de même orientation s'appelle une *orientation* de E .

Autres exercices

16.e Exercice (Exemples de déterminants).

Calculer les déterminants suivants (le dernier étant plus difficile) :

$$A_n = \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{i < j} (x_j - x_i)$$

$$B_n = \det \begin{pmatrix} 1 + x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & 1 + x_2 y_2 & \cdots & x_2 y_n \\ \vdots & & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & 1 + x_n y_n \end{pmatrix} = 1 + x_1 y_1 + \cdots + x_n y_n$$

$$C_n = \det(PGCD(i, j)_{1 \leq i, j \leq n}) = \prod_{1 \leq k \leq n} \varphi(k),$$

où φ est l'indicatrice d'Euler.

Chapitre 17

Formes différentielles ★

Mots-clefs du chapitre Forme extérieure, produit extérieur, produit vectoriel, forme différentielle

Un champ de vecteurs sur une surface est une application qui, à chaque point, associe un vecteur tangent en ce point. Une forme différentielle, par définition, est un champ d'applications multilinéaires antisymétriques. Pourquoi considérer ce type d'objet ? Parce que, comme É. Cartan¹ l'a découvert, les formes différentielles jouent un rôle crucial en géométrie et peuvent, notamment, être intégrées.

Commençons par l'aspect algébrique linéaire. Soit E un espace vectoriel réel de dimension n . On connaît le *dual* de E ; c'est l'espace vectoriel $E^* = L(E, \mathbb{R})$, de même dimension n , des formes linéaires sur E . Nous allons considérer d'autres espaces vectoriels associés à E .

17.1 Définition. Une p -forme extérieure (ou forme extérieure de degré p) α sur E est une forme $E^p \rightarrow \mathbb{R}$

- p -linéaire, c'est-à-dire linéaire par rapport à chacun de ses p arguments
- et antisymétrique : pour toute permutation $\sigma \in \mathfrak{S}_p$,

$$\alpha(\xi_{\sigma(1)}, \dots, \xi_{\sigma(p)}) = \epsilon(\sigma)\alpha(\xi_1, \dots, \xi_p).$$

On note $\wedge^p E^*$ l'espace des p -formes extérieures sur E ($p \geq 1$).² Conventionnellement, $\wedge^0 E^* = \mathbb{R}$.

17.2 Exemple. Le déterminant $\det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ est une n -forme extérieure (figure 17.1).

17.3 Exemple. Si v est un vecteur de \mathbb{R}^3 , la 2-forme extérieure (appartenant à $\wedge^2(\mathbb{R}^3)^*$)

$$(\xi, \eta) \mapsto \det(v, \xi, \eta)$$

1. Élie CARTAN, géomètre français (1869–1951), auteur de contributions fondamentales en physique mathématique, en géométrie différentielle et en théorie des groupes. Il fut aussi le père d'Henri CARTAN, autre mathématicien français qui a exercé une influence considérable.

2. Pourquoi cette notation bizarre ? On peut plus fondamentalement définir la puissance extérieure $\wedge^p E$ de E lui-même, mais nous ne le ferons pas ici.

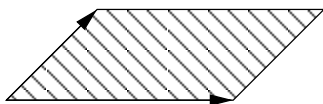


FIGURE 17.1 – Aire orientée de deux vecteurs

s'interprète, en mécanique des fluides, comme le flux d'un fluide de vitesse v à travers le parallélogramme engendré par ξ et η (figure 17.2). Exercice : se convaincre que le mot « flux » n'est pas usurpé.

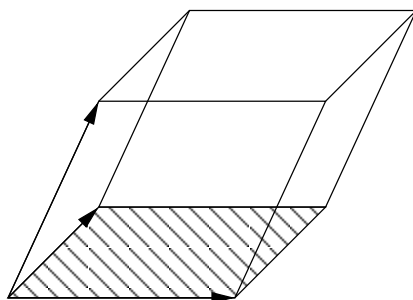


FIGURE 17.2 – Flux d'un fluide à travers un parallélogramme

Les formes extérieures ont un produit naturel, qualifié d'extérieur parce que par exemple le produit de deux 1-formes est une 2-forme, et non une 1-forme.

17.4 Définition. Le *produit extérieur* de p formes $\alpha_1, \dots, \alpha_p$ de degré 1 est la p -forme définie par

$$\alpha_1 \wedge \dots \wedge \alpha_p(\xi_1, \dots, \xi_p) = \det((\alpha_i(\xi_j))) = \begin{vmatrix} \alpha_1(\xi_1) & \dots & \alpha_1(\xi_p) \\ \vdots & & \vdots \\ \alpha_p(\xi_1) & \dots & \alpha_p(\xi_p) \end{vmatrix}.$$

17.5 Exemple. Soient (e_1, \dots, e_n) une base de E et (e_1^*, \dots, e_n^*) la base duale de $E^* = L(E, \mathbb{R})$ (i.e. $e_i^*(e_j) = \delta_{ij}$).

- Chaque e_i^* est une 1-forme.
- Chaque $e_i^* \wedge e_j^*$ est une 2-forme, et $e_i^* \wedge e_j^*(\xi_1, \xi_2)$ est l'aire orientée du parallélogramme engendré par les projections de ξ_1 et ξ_2 dans le plan (e_i, e_j) .
- ...
- $e_1^* \wedge \dots \wedge e_n^*$ est une n -forme ; c'est le *déterminant*.

17.6 Lemme. Si $\{e_1, \dots, e_n\}$ est une base de E , une base de $\wedge^p E^*$ est

$$\{e_{i_1}^* \wedge \dots \wedge e_{i_p}^*, 1 \leq i_1 < i_2 < \dots < i_p \leq n\}.$$

En particulier, $\wedge^p E^*$ est de dimension $\binom{n}{p} = \frac{n!}{p!(n-p)!}$.

Démonstration. Soit α une p -forme. Pour tout p -uplet (i_1, \dots, i_p) tel que $i_1 < \dots < i_p$, notons

$$a_{i_1 \dots i_p} = \alpha(e_{i_1}, \dots, e_{i_p}).$$

Alors

$$\alpha = \sum_{1 \leq i_1 < \dots < i_p \leq n} a_{i_1 \dots i_p} e_{i_1}^* \wedge \dots \wedge e_{i_p}^*. \quad (17.1)$$

En effet, soit maintenant un p -uplet de vecteurs de base $(e_{i_1}, \dots, e_{i_p})$ général (pas forcément ordonné). Si deux des indices sont égaux,

$$\alpha(e_{i_1}, \dots, e_{i_p}) = 0$$

par antisymétrie. Sinon il existe une unique permutation σ de $\{i_1, \dots, i_p\}$ telle que

$$\sigma(i_1) < \dots < \sigma(i_p).$$

Par antisymétrie,

$$\alpha(e_{i_1}, \dots, e_{i_p}) = \epsilon(\sigma) a_{i_1 \dots i_p} = \left(\sum_{1 \leq j_1 < \dots < j_p \leq n} a_{i_1 \dots i_p} e_{j_1}^* \wedge \dots \wedge e_{j_p}^* \right) (e_{i_1}, \dots, e_{i_p}).$$

La décomposition (17.1) est unique. \square

En particulier, si E est de dimension n , seules les $n + 1$ premières puissances extérieures de E^* sont non triviales (en partant de 0) :

$$\left\{ \begin{array}{ll} \wedge^0 E^* = \mathbb{R} & \dim = 1 \\ \wedge^1 E^* = E^* & \dim = n \\ \vdots & \\ \wedge^{n-1} E^* & \dim = n \\ \wedge^n E^* & \dim = 1 \\ \wedge^{n+1} E^* = 0 & \dim = 0 \\ \vdots & \end{array} \right.$$

Le produit vectoriel de deux vecteurs de \mathbb{R}^3 est le vecteur de \mathbb{R}^3 noté³ par

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} yz' - zy' \\ zx' - xz' \\ xy' - yx' \end{pmatrix}.$$

Le produit extérieur généralise le produit vectoriel en dimension quelconque. (Mais la propriété que le produit de deux vecteurs est encore un vecteur, est propre à la dimension 3.)

3. Parfois on utilise aussi la notation \wedge pour le produit vectoriel, ce que nous éviterons de faire ici par risque d'ambiguïté.

17.a Exercice (Produit vectoriel).

Dans cet exercice, on identifie \mathbb{R}^3 et son dual grâce au produit scalaire euclidien de \mathbb{R}^3 (via l'isomorphisme $\mathbb{R}^3 \rightarrow (\mathbb{R}^3)^*$, $\xi \mapsto (\eta \mapsto \xi \cdot \eta)$).

1. Vérifier que l'application

$$f : \mathbb{R}^3 \mapsto A_3(\mathbb{R}) = \wedge^2(\mathbb{R}^3)^*, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$$

est un isomorphisme (on a noté $A_3(\mathbb{R})$ l'espace des matrices 3×3 antisymétriques) tel que

$$f(\xi) \cdot \eta = \xi \times \eta.$$

2. Montrer que le diagramme suivant commute :

$$\begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \xrightarrow{\times} & \mathbb{R}^3 \\ \cong \downarrow & & \downarrow f \\ (\mathbb{R}^3)^* \times (\mathbb{R}^3)^* & \xrightarrow{\wedge} & \wedge^2(\mathbb{R}^3)^*. \end{array}$$

17.7 Définition. Une p -forme différentielle ω sur une surface S de \mathbb{R}^n est une application de classe C^1 qui, à tout point x de S , associe une p -forme extérieure de $T_x S \subset \mathbb{R}^n$. On note $\Lambda^p(S)$ l'espace de ces formes différentielles.

17.8 Notation. La dérivée d'une fonction $f : S \rightarrow \mathbb{R}$ (par opposition à une application plus générale) en un point x , soit $f'(x) \in L(T_x S, \mathbb{R})$, se note aussi $df(x)$.

Notamment, si $x : (S, a) \rightarrow \mathbb{R}^p$ est un système de coordonnées locales,⁴ pour tout i , $dx_i(a)$ (abusivement noté dx_i) est une forme linéaire sur $T_a S$, et $\{dx_1, \dots, dx_p\}$ est une base de $T_a S^*$.

17.9 Notation. On note $\partial_{x_1}, \dots, \partial_{x_p}$ (ou $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_p}$, ou $\partial_1, \dots, \partial_p$) la base duale de (dx_1, \dots, dx_p) , sur $T_a S$.

(La raison de cette notation provient du fait que l'on peut identifier un vecteur $v \in T_x S$ et l'opérateur de dérivation⁵

$$f \mapsto f'(x) \cdot v$$

agissant sur la fonction $f : (S, x) \rightarrow \mathbb{R}$. Exercice : le montrer.)

17.10 Exemple. Si $S = \mathbb{R}^d$ et $x = \text{id}$, $dx_i = x_i : \mathbb{R}^d \rightarrow \mathbb{R}$ est la projection sur le i -ième facteur, soit la i -ième forme de la base canonique duale de \mathbb{R}^d , et ∂_{x_i} est le i -ième vecteur de la base canonique de \mathbb{R}^d .

4. Cela est en particulier le cas si $S = \mathbb{R}^p$ et $x = \text{id}$.

5. Un *opérateur de dérivation* (du premier ordre) est un opérateur \mathbb{R} -linéaire agissant sur les fonctions dérivables et satisfaisant l'identité de Leibnitz $D(fg) = (Df)g + f(Dg)$.

Par suite, localement au voisinage de a , une p -forme différentielle ω a une expression de la forme

$$\omega(x) = \sum_{i_1 < \dots < i_p} \omega_{i_1 \dots i_p}(x) dx_{i_1} \wedge \dots \wedge dx_{i_p},$$

où chaque composante $\omega_{i_1 \dots i_p}$ est une fonction de classe C^1 . Rappelons que, si ξ_1, \dots, ξ_p sont des vecteurs tangents à S en x , chaque ξ_i est de la forme

$$\xi_i = \sum_{1 \leq j \leq d} \xi_{i,j} \partial_j,$$

et

$$\omega(x)(\xi_1, \dots, \xi_p) = \sum_{i_1 < \dots < i_p} \omega_{i_1, \dots, i_p}(x) \begin{vmatrix} \xi_{1,i_1} & \dots & \xi_{p,i_1} \\ \vdots & & \vdots \\ \xi_{1,i_p} & \dots & \xi_{p,i_p} \end{vmatrix}.$$

Chapitre 18

Surfaces III – Orientabilité ★

Mots-clefs du chapitre Forme volume, orientabilité

Soit S une surface de \mathbb{R}^n de dimension d et $x : (S, a) \rightarrow \mathbb{R}^d$ un système de coordonnées locales.¹ Rappelons que, pour tout point $a \in S$, $\wedge(T_a S)$ est une droite vectorielle engendrée par

$$dx_1 \wedge \cdots \wedge dx_d.$$

donc toute forme de degré d de classe C^k est du type

$$f(a) dx_1 \wedge \cdots \wedge dx_d$$

où $f : (S, a) \rightarrow \mathbb{R}$ est une fonction de classe C^k .

18.1 Définition. Une *forme volume* sur S est une d -forme différentielle sur S qui ne s'annule en aucun point. La surface S est *orientable* si une forme volume existe sur S .

18.2 Exemple (L'espace euclidien). La *forme volume standard* de \mathbb{R}^d est

$$\Omega = dx_1 \wedge \cdots \wedge dx_d.$$

(On verra bientôt que le *volume* d'une partie A de \mathbb{R}^d est l'intégrale, dans un sens à définir et si elle existe, de Ω sur A : $\text{Vol}(A) = \int_A \Omega$.)

18.3 Proposition (Orientabilité des hypersurfaces). *Soit S une hypersurface de \mathbb{R}^n définie par une équation globale, c'est-à-dire qu'il existe une submersion $f : (\mathbb{R}^n, S) \rightarrow \mathbb{R}$ (définie sur un voisinage de S) telle que $S = f^{-1}(0)$. Alors S est orientable.*

Démonstration. Soit ν le champ de vecteurs normaux à S défini par

$$\nu(x) = \text{grad } f(x).$$

1. Il faut bien sûr ne pas confondre la dimension d de S et l'opérateur d de dérivation des fonctions.

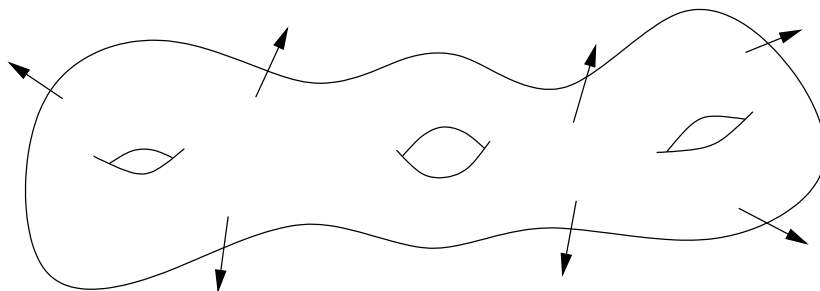


FIGURE 18.1 – Une hypersurface et un champ de vecteurs normaux

Comme f est une submersion, la 1-forme $f'(x)$ ne s'annule pas, donc le vecteur $\text{grad } f(x)$ ne s'annule pas. Ce dernier est normal à S , puisque, pour tout vecteur tangent $\xi \in T_x S$, par définition du gradient,

$$\xi \cdot \nu(x) = df(x) \cdot \xi = 0$$

(le premier point est le produit scalaire, le second le produit d'une matrice par un vecteur). Alors la $(d-1)$ -forme

$$i_{\nu(x)}\Omega := \Omega(\nu(x), \dots),$$

où Ω est la forme volume de \mathbb{R}^{d+1} , est une forme volume sur S . En effet, un point x de S étant fixé, soit e_1, \dots, e_d une base de $T_x S$. Alors $(e_1, \dots, e_d, \nu(x))$ est une base de \mathbb{R}^n , puisque $\nu(x)$ n'est pas dans $T_x S$. Donc $i_{\nu(x)}\Omega$ est non nulle. \square

On a utilisé la notation suivante.

18.4 Définition. Si $\omega \in \wedge^p(T_x S)^*$ (p -forme sur $T_x S$) et $\xi \in T_x S$,

$$i_\xi \omega = \omega(\xi, \dots)$$

est le *produit intérieur* de ω par ξ .

18.5 Exemple (La sphère). La sphère $\mathbb{S}^d : x_1^2 + \dots + x_{d+1}^2 = 1$ est orientable.

18.6 Proposition. Si une surface S de dimension d est orientable, il existe des ouverts U_i de \mathbb{R}^n qui recouvrent S et des systèmes de coordonnées locales $x_i : U_i \cap S \rightarrow \mathbb{R}^d$ tels que les applications de changement de coordonnées $x_i \circ x_j^{-1}$ aient un jacobien > 0 .

Démonstration. Admise. \square

18.7 Exemple (Ruban de Möbius). Le ruban de Möbius n'est pas orientable; la figure 18.3 en est presque une démonstration.

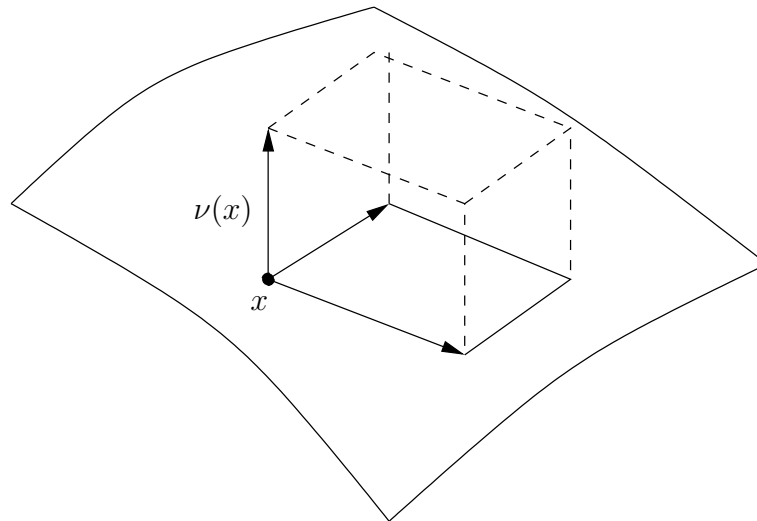
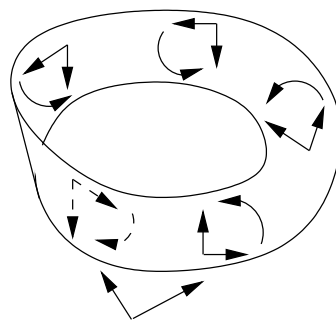


FIGURE 18.2 – Pour la forme volume induite sur une hypersurface, on mesure le volume en ajoutant une direction normale à l'hypersurface.



Deux orientations opposées

FIGURE 18.3 – Le ruban de Möbius n'est pas orientable : il n'existe pas de choix d'orientation continu sur tout le ruban.

Chapitre 19

Intégration ★

Mots-clefs du chapitre Image réciproque d'une forme différentielle, intégrale, forme d'aire

Chaque type d'objet mathématique se transforme par changement de coordonnées d'une façon qui lui est propre.

19.1 Exemple (Image direct d'un champ de vecteurs). Soit $\gamma : (\mathbb{R}, 0) \rightarrow (U, a)$ un chemin de classe C^1 . L'*image directe* de γ par un difféomorphisme $\varphi : U \rightarrow V$ entre deux ouverts de \mathbb{R}^d est naturellement le chemin

$$\varphi \circ \gamma : (\mathbb{R}, 0) \rightarrow (V, \varphi(a)).$$

Soit maintenant X un champ de vecteurs sur U , c'est-à-dire une application $X : U \rightarrow \mathbb{R}^d$ de classe C^1 . L'*image directe* de X par φ est l'unique champ de vecteur $\varphi_* X$ sur V dont les courbes intégrales soient les $\varphi \circ \gamma$, où γ est une courbe intégrale de X .¹ Autrement dit, si $\gamma'(t) = X(\gamma(t))$, on veut

$$(\varphi \circ \gamma)'(t) = \varphi'(\gamma(t)) \cdot X(\gamma(t)) = (\varphi_* X)(\varphi \circ \gamma(t)),$$

de quoi on déduit la formule suivante (en posant $y = \varphi(\gamma(t))$) pour l'image directe de X :

$$(\varphi_* X)(y) = \varphi'(\varphi^{-1}(y)) \cdot X(\varphi^{-1}(y)),$$

soit

$$\varphi_* X = (\varphi' \cdot X) \circ \varphi^{-1}.$$

Considérons maintenant une fonction $g : V \text{ ouvert } \subset \mathbb{R}^d \rightarrow \mathbb{R}$ de classe C^1 . Soit

$$\varphi : U \text{ ouvert } \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^d, \quad x \mapsto y = \varphi(x)$$

une application (éventuellement un difféomorphisme, mais ce n'est pas nécessaire ici), la fonction g "lue dans les coordonnées x " est la fonction $g \circ \varphi$ (définition 15.5),

1. Rappelons que γ est une *courbe intégrale* du champ de vecteurs X si à chaque instant t sa vitesse est le vecteur prescrit par X en ce point : $\gamma'(t) = X(\gamma(t))$, c'est-à-dire que γ est une solution de l'équation différentielle associée à X .

qui fait commuter le diagramme suivant :

$$\begin{array}{ccc} U & \xrightarrow{\varphi} & V; \\ & \searrow g \circ \varphi & \swarrow g \\ & \mathbb{R} & \end{array}$$

c'est la fonction qui donne le même résultat quand on change la façon de désigner les éléments à la source de g (x ou y).

Plus généralement, si β est une forme différentielle de degré k sur V (le cas d'une fonction g correspondant à $k = 0$), on lui associe la forme différentielle suivante.

19.2 Définition. L'image réciproque de β par φ est la k -forme différentielle sur U définie par

$$(\varphi^* \beta)(x)(\xi_1, \dots, \xi_k) = \beta(\varphi(x))(\varphi'(x) \cdot \xi_1, \dots, \varphi'(x) \cdot \xi_k).$$

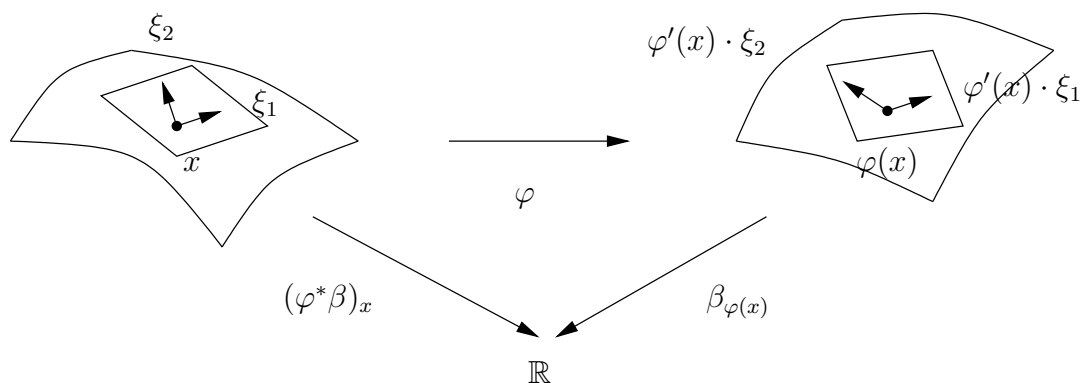


FIGURE 19.1 – Image réciproque d'une forme différentielle

C'est l'unique forme telle que, quand φ est un difféomorphisme,

$$(\varphi^* \beta)(x) \cdot X_x = \beta(\varphi(x)) \cdot (\varphi_* X)(\varphi(x)),$$

où $\varphi_* X$ est l'image directe du champ de vecteurs X (exemple 19.1).

Plus explicitement, si

$$\beta(y) = \sum_{i_1 < \dots < i_k} \beta_{i_1, \dots, i_k}(y) dy_{i_1} \wedge \dots \wedge dy_{i_k}$$

et si l'on note $\varphi = (\varphi_1, \dots, \varphi_k)$, on a

$$\begin{aligned} (\varphi^* \beta)(x) &= \sum_{i_1 < \dots < i_k} \beta_{i_1, \dots, i_k}(\varphi(x)) (dy_{i_1} \circ \varphi'(x)) \wedge \dots \wedge (dy_{i_k} \circ \varphi'(x)) \\ &= \sum_{i_1 < \dots < i_k} \beta_{i_1, \dots, i_k}(\varphi(x)) d\varphi_{i_1}(x) \wedge \dots \wedge d\varphi_{i_k}(x); \end{aligned}$$

autrement dit, l'expression de l'image réciproque d'une forme différentielle s'obtient simplement en exprimant les y_i en fonction des x_i (et φ n'a même pas besoin d'être un difféomorphisme).

19.3 Exemple. L'image réciproque de la forme d'aire $\beta = dx \wedge dy$ sur \mathbb{R}^2 par l'application "coordonnées polaires" $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (r, \theta) \mapsto (x, y) = (r \cos \theta, r \sin \theta)$ est

$$\begin{aligned}\varphi^* \beta(r, \theta) &= d(r \cos \theta) \wedge d(r \sin \theta) \\ &= (\cos \theta dr - r \sin \theta d\theta) \wedge (\sin \theta dr + r \cos \theta d\theta) \\ &= r(\cos^2 \theta + \sin^2 \theta) dr \wedge d\theta \\ &= r dr \wedge d\theta.\end{aligned}$$

19.a Exercice (Forme surface de \mathbb{R}^2).

1. Calculer l'expression de la forme surface standard $\Omega = dx \wedge dy$ de \mathbb{R}^2 en coordonnées polaires, c'est-à-dire l'image réciproque de Ω par l'application « coordonnées polaires » $\rho : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (r, \theta) \mapsto (x, y) = (r \cos \theta, r \sin \theta)$.
2. Faire de même avec la forme volume standard de \mathbb{R}^3 et les coordonnées sphériques.

La remarque fondamentale est la suivante. *L'intégrale de la fonction g dépend des systèmes de coordonnées choisis.* En effet, la formule du changement de variable démontrée dans le cours d'intégration est :

$$\int_V g(y) dy_1 \cdots dy_d = \int_U g(\varphi(x)) |\det \varphi'(x)| dx_1 \cdots dx_d.$$

Autrement dit, si l'on veut calculer l'intégrale de g dans les coordonnées x , il ne suffit pas d'intégrer la fonction g lue dans les coordonnées x , c'est-à-dire d'intégrer $(g \circ \varphi)(x)$; il faut multiplier cette dernière par la valeur absolue du déterminant jacobien du changement de variables. Mais comme, sur une surface, il n'existe aucun système de coordonnées locales privilégié, on ne sait tout simplement pas d'où partir pour intégrer f .

Il en est de même pour les formes de tout degré de 0 à $d - 1$, mais, en degré maximal d , un miracle se produit ! Soit β une d -forme sur $V \subset \mathbb{R}^d$, i.e.

$$\beta(y) = g(y) dy_1 \wedge \cdots \wedge dy_d$$

où $g : V \rightarrow \mathbb{R}$ est une fonction de classe C^1 .

19.4 Lemme. *L'image réciproque de β par φ est*

$$(\varphi^* \beta)(x) = (g \circ \varphi)(x) (\det \varphi'(x)) dx_1 \wedge \cdots \wedge dx_d. \quad (19.1)$$

Démonstration. L'image réciproque vaut

$$\begin{aligned}(\varphi^* \beta)(x) &= (g \circ \varphi)(x) d\varphi_1(x) \wedge \cdots \wedge d\varphi_d(x) \\ &= (g \circ \varphi)(x) \left(\sum_{i_1} \partial_{x_{i_1}} \varphi_1(x) dx_{i_1} \right) \wedge \cdots \wedge \left(\sum_{i_d} \partial_{x_{i_d}} \varphi_d(x) dx_{i_d} \right) \\ &= (g \circ \varphi)(x) \sum_{i_1, \dots, i_d} \partial_{x_{i_1}} \varphi_1(x) \cdots \partial_{x_{i_d}} \varphi_d(x) dx_{i_1} \wedge \cdots \wedge dx_{i_d}.\end{aligned}$$

Dans la majorité des cas, à savoir dès que l'application $(1, 2, \dots, d) \mapsto (i_1, \dots, i_d)$ n'est pas bijective, le terme en

$$dx_{i_1} \wedge \dots \wedge dx_{i_d}$$

est nul (parce qu'au moins deux facteurs sont identiques). Les seuls termes ayant une contribution non nulle sont donc ceux pour lesquels l'application $\sigma : (1, 2, \dots, d) \mapsto (i_1, \dots, i_d)$ est une permutation, de sorte que

$$\begin{aligned} (\varphi^* \beta)(x) &= (g \circ \varphi)(x) \sum_{\sigma \in \mathfrak{S}_d} \partial_{x_{\sigma 1}} \varphi_1(x) \cdots \partial_{x_{\sigma d}} \varphi_d(x) dx_{i_1} \wedge \dots \wedge dx_{i_d} \\ &= (g \circ \varphi)(x) \left(\sum_{\sigma \in \mathfrak{S}_d} \epsilon(\sigma) \partial_{x_{\sigma 1}} \varphi_1(x) \cdots \partial_{x_{\sigma d}} \varphi_d(x) \right) dx_1 \wedge \dots \wedge dx_d, \end{aligned}$$

d'où la formule voulue. □

L'apparition du déterminant jacobien de φ dans la formule (19.1) est précisément le facteur qu'il faut pour rendre intrinsèque l'intégrale d'une d -forme, définie de la façon suivante.

19.5 Définition. L'intégrale d'une d -forme différentielle

$$\beta(y) = g(y) dy_1 \wedge \dots \wedge dy_d$$

définie sur un ouvert V de \mathbb{R}^d est

$$\int_V \beta = \int_V g(y) dy_1 \cdots dy_d,$$

où le membre de droite est l'intégrale de la fonction g (g étant supposée de classe C^1 , ses intégrales de Riemann et de Lebesgue coïncident).

L'équation (19.1) montre que cette définition ne change pas si l'on change de coordonnées en préservant l'orientation :

$$\int_V \beta = \int_U \varphi^* \beta.$$

Sur une surface orientée de dimension d , on se ramène bien sûr localement à un ouvert de \mathbb{R}^d en choisissant des coordonnées (le résultat ne dépendant pas des coordonnées choisies).

19.6 Exemple (Aire d'un paraboloidé). Soit S le graphe de la fonction

$$f : [-1, 1]^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2.$$

Comme l'équation du graphe s'écrit

$$F(x, y, z) = z - x^2 - y^2 = 0,$$

le champ de vecteurs normaux sur S (orienté, disons, vers le haut) est

$$\nu : S \rightarrow \mathbb{R}^3, \quad (x, y, z) \mapsto \frac{\text{grad } F(x, y, z)}{\|\text{grad } F(x, y, z)\|} = \frac{1}{\sqrt{1 + 4x^2 + 4y^2}} \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix}.$$

Donc la forme d'aire sur S est

$$\begin{aligned} \sigma &= i_\nu(dx \wedge dy \wedge dz) \\ &= \frac{1}{\sqrt{1 + 4x^2 + 4y^2}} (-2x \, dy \wedge dz + 2y \, dx \wedge dz + dx \wedge dy) \\ &= \sqrt{1 + 4x^2 + 4y^2} \, dx \wedge dy \quad (z = x^2 + y^2). \end{aligned}$$

Donc l'aire de S vaut

$$\text{aire}(S) = \int_S \sigma = \int_{[-1,1]^2} \sqrt{1 + 4x^2 + 4y^2} \, dx \, dy,$$

intégrale qui, à défaut de se calculer rigoureusement, peut être évaluée numériquement.

19.7 Exemple (Aire du disque \mathbb{D}^2). L'aire du disque $\mathbb{D}^2 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$ est l'intégrale sur \mathbb{D}^2 de la forme d'aire $dx \wedge dy$ de \mathbb{R}^2 , soit

$$\begin{aligned} \text{aire}(\mathbb{D}^2) &= \int_{\mathbb{D}^2} dx \wedge dy \\ &= \int_{\mathbb{D}^2} dx \, dy \\ &= 4 \int_{\mathbb{R}_*^{+2} \cap \mathbb{D}^2 \setminus [0,1]} dx \, dy \\ &\quad (\text{par soustraction d'une partie de mesure nulle}) \\ &= 4 \int_{]0,1[\times]0,\pi/2[} r \, dr \, d\theta \quad (\text{passage en polaires}) \\ &= \pi. \end{aligned}$$

19.8 Exemple (Longueur du cercle \mathbb{S}^1). La longueur est l'intégrale sur \mathbb{S}^1 de la « forme longueur » induite par la forme d'aire standard $\Omega = dx \wedge dy$ de \mathbb{R}^2 sur \mathbb{S}^1 , soit $\omega = i_\nu \Omega$ si ν est le vecteur normal unitaire (disons sortant) du cercle, soit encore

$$\omega = x \, dy - y \, dx.$$

En coordonnées polaires, ω a pour expression

$$\omega = \cos \theta \, d(\sin \theta) - \sin \theta \, d(\cos \theta) = d\theta.$$

Remarquons qu'au voisinage de tout point du cercle une détermination de θ est une coordonnée locale, mais que θ n'est pas une coordonnée globale. En coupant le cercle au moins en deux, on se convaincra (malgré ceci) que l'intégrale de ω vaut

$$\text{longueur}(\mathbb{S}^1) = \int_{\mathbb{S}^1} \omega = 2\pi.$$

19.9 Exemple (Aire de la sphère \mathbb{S}^2). On se convaincra que l'aire du méridien $\mathbb{S}^2 \cap \{y = 0\}$ est nulle. Donc l'aire de la sphère est l'intégrale sur $\mathbb{S}^2 \setminus (\{\pm p\} \cup \{y = 0\})$ de la forme d'aire

$$i_\nu(dx \wedge dy \wedge dz).$$

Notons (r, θ, φ) les coordonnées sphériques (voir la figure 15.1). Celles-ci redressent la sphère et les angles θ et ϕ forment un système de coordonnées globales sur l'hémisphère $\mathbb{S}^2 \cap \{y > 0\}$. Or $\nu = \partial_r$ et $dx \wedge dy \wedge dz = r^2 \sin \theta dr \wedge d\theta \wedge d\varphi$, donc

$$i_\nu(dx \wedge dy \wedge dz) = r^2 \sin \theta d\theta \wedge d\varphi,$$

de sorte que

$$\text{aire}(\mathbb{S}^2) = 2 \int_{]0, \pi[\times]0, \pi[} \sin \theta d\theta d\phi = 4\pi.$$

Plus généralement, la sphère de rayon R a pour aire $4\pi R^2$.

19.b Exercice (Revêtements).

Une application $f : S \rightarrow T$ entre deux surfaces orientables est un *revêtement* de degré k si

- f est un difféomorphisme local au voisinage de tout point de S
- tout point de T possède exactement k antécédents par f .

1. Donner des exemples de revêtements de degré 2 et infini.
2. Montrer que, si β est une forme différentielle de degré maximal sur T ,

$$\int_S f^* \beta = k \int_T \beta.$$

Chapitre 20

Formule de Stokes ★

Mots-clefs du chapitre Surface à coins, bord, orientation, dérivation extérieure, gradient, rotationnel, divergence

Nous démontrons ici la formule de Stokes¹ (20.1), d'ailleurs plutôt vue ici comme une définition de l'opérateur de dérivation extérieure des formes différentielles, et qui peut légitimement passer pour la plus importante formule de l'analyse mathématique.

20.1 Définition. Une *surface à coins* de dimension d de \mathbb{R}^n est une partie S de \mathbb{R}^n telle que, pour tout $a \in S$, il existe un difféomorphisme

$$\alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^+)^d \times (\mathbb{R}^+)^{n-d}$$

tel que, localement au voisinage de 0,

$$(\tilde{x}, \tilde{y}) \in \alpha(S) \Leftrightarrow \tilde{y} = 0.$$

Comme pour les surfaces sans coin, le difféomorphisme α est un *redressement* local de S (voir la figure 11.1), l'entier d est la *dimension* de S , et l'entier $n - d$ est sa codimension.

Le bord de $(\mathbb{R}_+)^k$ est

$$\partial(\mathbb{R}_+)^k = \{x \in (\mathbb{R}_+)^k, \exists i x_i = 0\}.$$

Les redressements locaux permettent de transposer immédiatement cette définition aux surfaces à coins.

20.2 Définition. La *bord* de S est l'ensemble des points de S appartenant à un ensemble

$$\alpha^{-1}(\partial(\mathbb{R}_+)^n)$$

où $\alpha : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^+)^d \times (\mathbb{R}^+)^{n-d}$ est un redressement local de S .

1. George Gabriel STOKES, mathématicien britannique (1819–1903), célèbre pour ses travaux notamment en mécanique des fluides et en optique

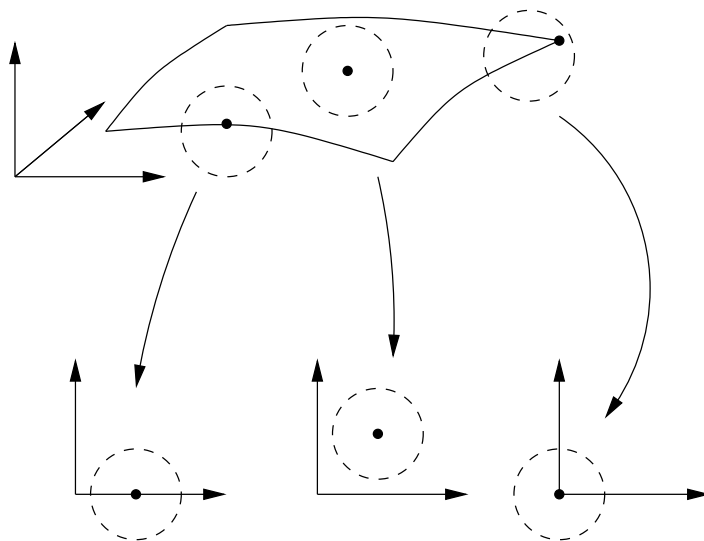


FIGURE 20.1 – Surface à coin, avec plusieurs systèmes de coordonnées locales

20.a Exercice (Bord).

1. Montrer que le bord ∂S de S est une surface de \mathbb{R}^n de dimension $d - 1$.
2. En déduire que, si S est orientée, il en est de même de ∂S .

20.3 Théorème. *Pour toute surface à coins S de dimension n , il existe un unique opérateur $d = d_S : \Lambda^k(S) \rightarrow \Lambda^{k+1}(S)$ ($k = 0, \dots, n$), appelé la dérivation extérieure, tel que*

1. *d est naturel vis-à-vis des images réciproques :*

$$\varphi^* \circ d_S = d_T \circ \varphi^*$$

pour toute application $\varphi : T \rightarrow S$ entre surfaces ;

2. *pour toute $(n - 1)$ -forme α sur S*

$$\int_S d\alpha = \int_{\partial S} \alpha. \quad (20.1)$$

De plus, pour toute fonction f ,

$$d(f dx_{i_1} \wedge \cdots \wedge dx_{i_k}) = df \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \quad (20.2)$$

Idée de démonstration. Unicité. Comme d est naturelle vis-à-vis des images réciproques (ce qui rend cet opérateur insensible aux changements de coordonnées), en prenant des coordonnées locales on peut supposer que S est un ouvert de \mathbb{R}^d . On applique la formule (20.1) sur des cubes $C_\epsilon = [0, \epsilon]^d$, et l'on trouve la formule (20.2).

Pour l'existence, il faut montrer que la formule (20.1) est vérifiée non seulement quand S est (à de petits bouts de cubes près, au bord) un cube mais pour une surface à coins orientable quelconque. Ceci se fait en découpant une telle surface S en petits cubes d -dimensionnels, dont on fait tendre la dimension vers 0 (pour faire tendre les erreurs vers 0).

On trouvera les détails par exemple dans l'excellent livre de F. Pham [Pha92]. \square

20.b Exercice.

1. Montrer que, pour toute k -forme différentielle α , $d^2\alpha = 0$ (on pourra utiliser le lemme de Schwarz).

2. En déduire que, pour toute surface à bord orientable S , $\partial^2 S = 0$.

20.4 Exemple (\mathbb{R}^3). Dans l'espace \mathbb{R}^3 euclidien, on peut transcrire les opérateurs de dérivation extérieure, pour les différents degrés, en termes d'opérateurs agissant sur les champs de vecteurs, c'est-à-dire des applications $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ de classe C^1 . Notons $\chi(\mathbb{R}^3)$ l'espace de ces champs de vecteurs.

On définit préalablement deux isomorphismes \mathbb{R} -linéaires

$$\sharp : \Lambda^1(\mathbb{R}^3) \rightarrow \chi(\mathbb{R}^3) \quad \text{et} \quad * : \Lambda^2(\mathbb{R}^3) \rightarrow \Lambda^1(\mathbb{R}^3)$$

par les formules suivantes :

$$\left\{ \begin{array}{l} \sharp dx = \partial_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ \sharp dy = \partial_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ \sharp dz = \partial_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} *(dy \wedge dz) = dx \\ *(dz \wedge dx) = dy \\ *(dx \wedge dy) = dz. \end{array} \right.$$

(On peut donner une définition intrinsèque de ces opérateurs, en termes uniquement du produit scalaire euclidien.) Alors les opérateurs *gradient*, *rotationnel* et *divergence* sont définis par le diagramme suivant :

$$\begin{array}{ccccccc} \Lambda^0(\mathbb{R}^3) & \xrightarrow{d} & \Lambda^1(\mathbb{R}^3) & \xrightarrow{d} & \Lambda^2(\mathbb{R}^3) & \xrightarrow{d} & \Lambda^3(\mathbb{R}^3) \\ \parallel & & \downarrow \sharp & & \downarrow \sharp* & & \downarrow * \\ \mathcal{F}(\mathbb{R}^3) & \xrightarrow{\text{grad}} & \chi(\mathbb{R}^3) & \xrightarrow{\text{rot}} & \chi(\mathbb{R}^3) & \xrightarrow{\text{div}} & \mathcal{F}(\mathbb{R}^3). \end{array}$$

Du fait que $d^2=0$ découle immédiatement que le rotationnel d'un gradient ou la divergence d'un rotationnel sont nuls.

20.c Exercice (Formule de Cauchy).

Soit f une fonction complexe définie sur un voisinage de $D = D(0, 1)$, de classe C^1 .

1. Montrer que, pour tout $\zeta \in D$,

$$i2\pi f(\zeta) = \int_{\partial D} \frac{f(z)}{z - \zeta} dz + \int_D \frac{(\partial f / \partial \bar{z})(z)}{z - \zeta} dz d\bar{z}.$$

2. En déduire que, si f est holomorphe,

$$i2\pi f(\zeta) = \int_{\partial D} \frac{f(z)}{z - \zeta} dz,$$

donc développable en série entière.

Chapitre 21

Dérivées d'ordre supérieur

Mots-clefs du chapitre Classe C^k , hessienne, formule de Taylor, applications tangentes

Soit $f : U$ ouvert $\subset \mathbb{R}^n \rightarrow \mathbb{R}^p$.

21.1 Définition. L'application f est de classe C^k ($k \in \mathbb{N}_*$) si f est différentiable sur U et si $f' : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^p) \equiv \mathbb{R}^{np}$ est elle-même de classe C^{k-1} ; par récurrence, on définit alors $f^{(k)} = (f^{(k-1)})'$. L'application est de classe C^∞ si elle est de classe C^k sur U pour tout k .

21.2 Exemple ($k = 2$). La hessienne de f en x est

$$f''(x) \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p)) \equiv L_2(\mathbb{R}^n, \mathbb{R}^p),$$

$L_2(\mathbb{R}^n, \mathbb{R}^p)$ désignant l'espace des applications bilinéaires de $\mathbb{R}^n \times \mathbb{R}^n$ dans \mathbb{R}^p . (Si $p = 1$, la hessienne est une matrice.) En dérivant

$$f'(x) \cdot \xi = \sum_{1 \leq j \leq n} (\partial_{x_j} f)(x) \xi_j$$

par rapport à x , on voit que

$$f''(x) \cdot (\xi, \eta) = \sum_{1 \leq j, k \leq n} (\partial_{x_k} \partial_{x_j} f)(x) \xi_j \xi_k;$$

si l'on donne deux fois le même argument ξ à $f'(x)$, on note

$$f''(x) \cdot \xi^2 = f''(x) \cdot (\xi, \xi).$$

Plus généralement, $f^{(k)}$ peut être vue comme une application k -linéaire $(\mathbb{R}^n)^k \rightarrow \mathbb{R}^p$.

21.3 Lemme (Schwarz). Si $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}^p$ est de classe C^2 , la hessienne de f est symétrique : pour tous i, j ,

$$\partial_{x_i} \partial_{x_j} f = \partial_{x_j} \partial_{x_i} f.$$

Comme c'est un résultat démontré dans les cours antérieurs, nous en donnons ici une démonstration plus sophistiquée (et plus élégante ?), qui utilise la formule de Stokes.

Démonstration. Comme on ne dérive pas dans les autres directions que celles de x_i et x_j , on peut supposer que $n = 2$, $i = 1$ et $j = 2$, le cas général s'en déduisant directement. Soit $D = [0, \epsilon]^2$, muni de l'orientation habituelle du plan. Son bord est

$$\partial D = \ell_1 + \ell_2 + \ell_3 + \ell_4,$$

union de quatre segments de droites, orientés comme sur la figure 21.1. Le bord de son bord est

$$\partial^2 D = \partial(\ell_1 + \ell_2 + \ell_3 + \ell_4) = (B - A) + (C - B) + (D - C) + (A - D) = \emptyset$$

(c'est un fait fondamental que le bord du bord d'un domaine, en général, est vide).

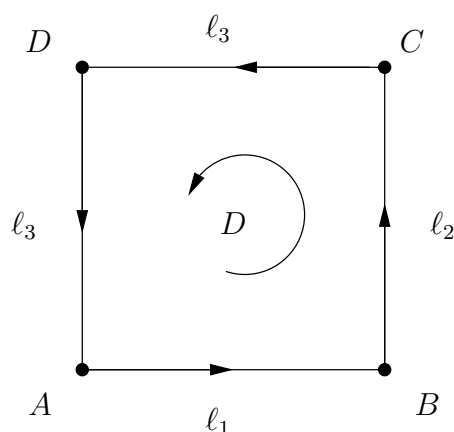


FIGURE 21.1 – Le carré D , son bord et le bord de son bord

Donc la formule de Stokes (appliquée deux fois de suite) montre que

$$0 = \int_{\partial^2 D} f = \int_{\partial D} df = \int_D d^2 f.$$

Comme cela est vrai pour tout $\epsilon > 0$ assez petit, $d^2 f = 0$. Or,

$$df = \partial_x f dx + \partial_y f dy,$$

donc

$$0 = d^2 f(0, 0) = (\partial_x \partial_y f(0, 0) - \partial_y \partial_x f(0, 0)) dx \wedge dy.$$

□

21.4 Proposition (Formule de Taylor d'ordre k). *Si $f : (\mathbb{R}^n, x) \rightarrow \mathbb{R}^p$ est de classe C^{k+1} , pour tout $\xi \in \mathbb{R}^n$ assez petit,*

$$f(x + \xi) = f(x) + f'(x) \cdot \xi + \cdots + \frac{1}{k!} f^{(k)}(x) \cdot \xi^k + \left(\int_0^1 f^{(k+1)}(x + t\xi) \frac{(1-t)^k}{k!} dt \right) \cdot \xi^{k+1}.$$

Démonstration. Des intégrations par parties successives montrent que

$$\begin{aligned}
 f(x + \xi) &= f(x) + \int_0^1 \frac{d}{dt} f(x + t\xi) dt \\
 &\quad \text{(formule fondamentale du calcul différentiel, cf. rappel 1.3)} \\
 &= f(x) + \left(\int_0^1 f'(x + t\xi) dt \right) \cdot \xi \\
 &\quad \text{(formule voulue à l'ordre zéro)} \\
 &= f(x) - [(1-t)f'(x + t\xi)]_0^1 \cdot \xi + \left(\int_0^1 (1-t)f''(x + t\xi) dt \right) \cdot \xi^2 \\
 &= f(x) + f'(x) \cdot \xi + \left(\int_0^1 (1-t)f''(x + t\xi) dt \right) \cdot \xi^2 \\
 &\quad \text{(formule voulue au premier ordre)} \\
 &= \dots,
 \end{aligned}$$

d'où la formule. □

21.5 Remarque. $f^{(k)}(x) \cdot \xi^k$ est la dérivée directionnelle k -ième de f dans la direction de ξ . Donc, pour expliciter ce terme, il suffit d'appliquer k fois l'opérateur $\xi_1 \partial_{x_1} + \dots + \xi_n \partial_{x_n}$.

21.a Exercice (Cas particuliers de la formule de Taylor).

Écrire la formule de Taylor à l'ordre trois pour une fonction de deux variables, en fonction des dérivées partielles de la fonction.

21.b Exercice (Tangence d'ordre k).

Deux applications $f : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ et $g : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^p, b)$ sont *tangentes d'ordre k* ou *k -tangentes*, ce qu'on note $f \sim_k g$, si

$$(f - g)(a + \xi) = o(\|\xi\|^k).$$

1. Montrer que la k -tangence est une relation d'équivalence.

2. Montrer que f est continue en x si et seulement si f et $x \mapsto b$ sont 0-tangentes, et que f est dérivable en x avec pour dérivée $f'(x)$ si et seulement si f et $x \mapsto b + f'(x) \cdot (x - a)$ sont 1-tangentes.

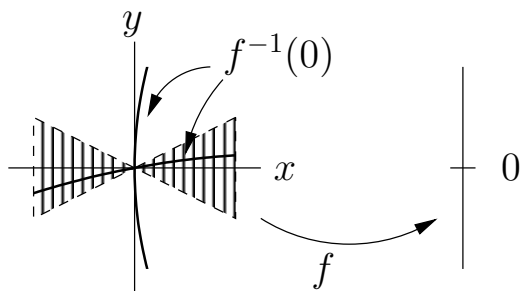
3. Montrer que, si $\varphi : (\mathbb{R}^n, a') \rightarrow (\mathbb{R}^n, a)$ est un difféomorphisme local de classe C^k , $f \sim_k g$ (en a) si et seulement si $f \circ \varphi \sim_k g \circ \varphi$ (en a').

21.c Exercice (Éclatement d'une singularité).

Soit $f : (\mathbb{R}^2, 0) \rightarrow (\mathbb{R}, 0)$ une fonction locale de classe C^∞ dont 0 soit un point critique non dégénéré d'indice 1 :

$$f(x, y) = xy + O_3.$$

Montrer que, localement, l'ensemble $\Sigma_0 = f^{-1}(0)$ de niveau critique de f est l'union de deux courbes transverses, dont chacune est tangente à l'un des axes de coordonnées.



Chapitre 22

Forme normale d'une forme quadratique

Mots-clefs du chapitre Orthogonalité, groupe orthogonal, diagonalisation, forme normale, index

Soit b une *forme bilinéaire symétrique* (éventuellement dégénérée) sur un \mathbb{K} -espace vectoriel E ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}) :

$$b(x, y) = b(y, x) \quad (\forall x, y).$$

On dit que x et y sont *b -orthogonaux* si $b(x, y) = 0$.

22.1 Définition. La *forme quadratique* associée à b est la fonction

$$q : E \rightarrow \mathbb{K}, \quad x \mapsto q(x) = b(x, x).$$

La forme q détermine b , comme le montre la formule de restitution

$$b(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

(cette formule s'appelle aussi la *formule du parallélogramme*, parce qu'elle restitue la valeur $b(x, y)$ en fonction de la valeur de b aux autres sommets du parallélogramme engendré par x et y).

22.a Exercice (Forme quadratique abstraite).

Montrer que, si une fonction $q : E \rightarrow \mathbb{K}$ est une forme quadratique “abstraite”, au sens où q est homogène de degré 2 : $q(\lambda x) = \lambda^2 q(x)$ ($x \in E$, $\lambda \in \mathbb{K}$) et l'application

$$b : E^2 \rightarrow \mathbb{K}, \quad b_q(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

est bilinéaire, alors q est la forme quadratique associée à b , et que donc l'espace des formes quadratiques s'identifie à celui des formes bilinéaires symétriques.¹

1. Il est essentiel ici que la caractéristique du corps \mathbb{K} ne soit pas 2.

22.b Exercice. La fonction $q : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto \sqrt{x^4 + y^4}$ est-elle une forme quadratique ?

22.2 Définition. Le *groupe orthogonal* $O(q)$ est l'ensemble des automorphismes a de E qui préservent $q : q(a(x)) = q(x)$ pour tout x . Si B est la matrice de b dans une base, et A celle de a , A préserve B si et seulement si

$${}^tABA = B.$$

22.3 Théorème (Diagonalisation des formes bilinéaires symétriques). *Soit b une forme bilinéaire symétrique. Il existe une base de E formée de vecteurs deux à deux b -orthogonaux.*

Démonstration. Si $b = 0$, n'importe quelle base de E convient. Sinon, il existe un vecteur e_1 tel que $b(e_1, e_1) \neq 0$. Soit alors $E_1 = e_1^\perp$. Comme E_1 est le noyau de la forme linéaire non nulle $b(e_1, \cdot)$, $\dim E_1 = n - 1$. De plus, $e_1 \notin E_1$. Donc $E = \mathbb{K}e_1 \oplus E_1$. Le théorème s'en déduit par récurrence, en répétant le raisonnement. \square

22.4 Corollaire (Forme normale des formes quadratiques réelles). *Si q est une forme quadratique sur un espace vectoriel réel de dimension n , il existe des entiers j et k , et une base q -orthonormée dans laquelle q est la forme quadratique sur \mathbb{R}^n*

$$-x_1^2 - \cdots - x_j^2 + x_{j+1}^2 + \cdots + x_{j+k}^2.$$

Les entiers j et k sont des invariants.

Démonstration. Dans la démonstration précédente, il suffit de remplacer e_1 par $e'_1 = e_1/\sqrt{|q(e_1)|}$ (de sorte que $q(e'_1) = \pm 1$) et ainsi de suite.

Le fait que la signature ne dépend pas de la base choisie provient du fait que j est la dimension du plus grand sous-espace sur lequel b est définie négative. De même, k est la dimension du plus grand sous-espace sur lequel b est définie positive. \square

22.5 Définition. La paire (j, k) est la *signature*, l'entier j (soit le nombre de carrés négatifs) est l'indice, et l'entier $r = j + k$ est le rang. Une forme quadratique réelle est

- *positive* si $j = 0$
- *négative* si $k = 0$
- *non-dégénérée* si $r = n$
- *définie* si j ou $k = n$ (donc non-dégénérée).

22.c Exercice (Tracé d'une conique).

Tracer la courbe conique d'équation $3x^2 - 2xy + 3y^2 = 1$.

Démonstration. Soit $Q = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$. Considérons donc la courbe C a pour équation ${}^tXQX = 1$, $X = (x, y)$.

On voit que

$$Q \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{et} \quad Q \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Ces deux vecteurs forment une base orthogonale. En les normalisant on obtient une base orthonormale (e_1, e_2) , avec

$$b_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad b_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Soit $P = (b_1, b_2)$, vue comme une matrice de passage. Comme P est un opérateur orthogonal,

$$D := P^{-1}QP = {}^tPQP = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

En posant $X' = P^{-1}X$, on voit que l'équation de $C' = P^{-1}(C)$ est ${}^tX'DX' = 1$. Si l'on note $X' = (x', y')$, l'équation est, plus explicitement,

$$2x'^2 + 4y'^2 = 1.$$

C'est une ellipse dont les axes de symétrie sont les axes Ox' et Oy' , et de demi grands axes $1/\sqrt{2}$ et $1/2$. La courbe C est l'image de cette ellipse par P , elle est donc encore une ellipse de demi grands axes $1/\sqrt{2}$ et $1/2$. \square

22.d Exercice (Étude d'un point critique).

Déterminer la nature du point critique 0 pour la fonction $f(x, y) = 3x^2 - 2xy + 3y^2$.

Démonstration. Première solution : D'après l'exercice 22.c, $f \circ P(x', y') = 2x'^2 + 4y'^2$, donc l'origine est un minimum strict de f .

Deuxième solution : en complétant le premier carré, on voit que

$$f(x, y) = \left(\sqrt{3}x - \frac{y}{\sqrt{3}} \right)^2 + \frac{8}{3}y^2,$$

donc l'origine est un minimum strict de f . \square

22.6 Lemme (Réduction lisse). *Quelle que soit la matrice symétrique inversible H_0 , il existe une application $P : (S_n(\mathbb{R}), H_0) \rightarrow (M_n(\mathbb{R}), I)$ de classe C^∞ telle que, pour toute matrice symétrique H proche de H_0 ,*

$$H = {}^tP(H)H_0P(H).$$

Démonstration. Il s'agit de trouver un inverse à droite de l'application

$$h : (M_n(\mathbb{R}), I) \rightarrow (S_n(\mathbb{R}), H_0), \quad P \mapsto {}^tPH_0P.$$

L'application h est polynomiale, donc C^∞ . Sa dérivée en l'identité est

$$h'(I) : M \mapsto H_0M + {}^tMH_0,$$

donc est surjective (si S est une matrice symétrique donnée, il suffit de poser $M = \frac{1}{2}H_0^{-1}S$ pour avoir $S = h'(I) \cdot M$). Donc h est une submersion. D'après le théorème de submersion, h possède un inverse à droite local, qui est l'application P recherchée. \square

22.e Exercice (Forme quadratique complexe).

Montrer que, si q est une forme quadratique sur un espace vectoriel complexe de dimension finie, il existe une base q -orthonormée dans laquelle q est la forme quadratique sur \mathbb{C}^n

$$x_1^2 + \cdots + x_r^2;$$

r est le rang, qui est un invariant.

En termes de forme bilinéaire, ce qui précède se reformule de la façon suivante.

22.7 Corollaire (Forme normale d'une forme bilinéaire symétrique). Si $\mathbb{K} = \mathbb{R}$, il existe une base dans laquelle la matrice de b est de la forme

$$I_{j,k} = \begin{pmatrix} -I_j & & \\ & I_k & \\ & & 0 \end{pmatrix}.$$

Si $\mathbb{K} = \mathbb{C}$, il existe une base dans laquelle la matrice de b est de la forme I_r .

Les formes bilinéaires symétriques non dégénérées ($r = n$) et les groupes d'automorphismes associés $O(q)$ portent des noms particuliers selon les cas :

- Si $\mathbb{K} = \mathbb{R}$, b s'appelle un *produit scalaire pseudo-euclidien* ou *lorentzien*. La forme quadratique pseudo-euclidienne standard d'indice j est

$$q_j(x) = -x_1^2 - \cdots - x_j^2 + x_{j+1}^2 + \cdots + x_{j+k=n}^2$$

et le groupe orthogonal correspondant $O(q)$ se note $O_{j,k}$.

- Si de plus b est positive ($j = 0$), b s'appelle un *produit scalaire euclidien*. La forme quadratique euclidienne standard sur \mathbb{R}^n est

$$q(x) = x_1^2 + \cdots + x_n^2$$

et le groupe associé se note $O(q) = O_n$.

- Si $\mathbb{K} = \mathbb{C}$, b s'appelle un *produit scalaire euclidien complexe*. La forme quadratique euclidienne complexe standard sur \mathbb{C}^n est

$$q(x) = x_1^2 + \cdots + x_n^2$$

(attention : elle n'est pas à valeurs réelles ; le véritable analogue complexe d'un produit scalaire euclidien n'est pas bilinéaire mais sesquilinéaire, et l'on parle alors de forme *hermitienne*) et $O(q)$ se note alors $O_n(\mathbb{C})$.

22.f Exercice (Spectre d'un opérateur isométrique).

Soient E un espace vectoriel de dimension finie sur $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , et $\varphi \in L(E)$ un opérateur préservant une forme bilinéaire non dégénérée (ce qui s'appelle un *produit scalaire*²). Montrer que φ est inversible et semblable à son inverse. *Indication* : on pourra utiliser la forme normale de Frobenius.

2. Par exemple, un produit scalaire antisymétrique s'appelle un *produit scalaire symplectique*

Chapitre 23

Fonctions de Morse

Mots-clefs du chapitre Point critique non dégénéré, forme normale

Soit $f : (\mathbb{R}^n, a) \rightarrow \mathbb{R}$ une fonction de classe C^3 . Si $f'(a)$ est non nulle (donc de rang 1), le théorème de submersion affirme que f est équivalente à sa partie linéaire (une projection de \mathbb{R}^n sur \mathbb{R}). Si $f'(a) = 0$, il n'en est rien en général, comme le montre l'exemple de la fonction $f(x) = x^2$ au voisinage de 0 dans \mathbb{R} (dont la partie linéaire en 0 est nulle, et à laquelle f n'est donc pas équivalente). Si toutefois la partie quadratique $f''(a)$ est non-dégénérée, f lui est équivalente; c'est l'objet du lemme de Morse, exposé ci-dessous. Si l'on continue dans la hausse du niveau de dégénérescence, le théorème des jets suffisants donne une condition de non-dégénérescence suffisante pour que f soit équivalente à son polynôme de Taylor d'un ordre assez élevé. C'est le début de la théorie des singularités.

23.1 Définition. Le point a est *critique* si $f'(a)$ n'est pas de rang maximal, i.e., ici, $f'(a) = 0$.

23.2 Lemme. Si a est un point critique de f , $f''(a)$ définit une forme bilinéaire sur \mathbb{R}^n qui est indépendante des coordonnées choisies, dans le sens où, si $\varphi : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^n, a)$ est un C^2 -difféomorphisme local,

$$(f \circ \varphi)''(a)(\xi, \xi) = f''(a) \cdot (\varphi'(a) \cdot \xi, \varphi'(a) \cdot \xi).$$

Démonstration. Cela résulte directement de la formule de dérivation des fonctions composées : $(f \circ \varphi)' = f' \circ \varphi \cdot \varphi'$ donc

$$(f \circ \varphi)'' = f'' \circ \varphi \cdot (\varphi', \varphi') + f' \circ \varphi \cdot \varphi'',$$

d'où, si a est critique, la formule voulue. □

23.3 Définition. Si a est un point critique, la *hessienne* de f en a est la forme bilinéaire $f''(a)$. Le point a est alors *non-dégénéré* si $f''(a)$ est non-dégénérée, comme forme bilinéaire (c'est-à-dire par exemple si le déterminant de sa matrice, dans une base quelconque, est non nul).

23.4 Lemme (de Morse). *Si a est un point critique non-dégénéré de f ,¹ il existe un difféomorphisme local $\varphi : (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^n, a)$ tel que*

$$(f \circ \varphi)(u) = f(a) + \frac{1}{2}f''(a) \cdot (u - a)^2$$

et φ est tangent à l'identité : $\varphi'(a) = I$.

Dans le chapitre 22, on a vu que, puisque d'après l'hypothèse de non-dégénérescence le rang est maximal, égal à n , il existe un entier j tels que la forme quadratique $f''(a)$ soit, dans une base convenable,

$$-v_1^2 - \dots - v_j^2 + v_{j+1}^2 + \dots + v_n^2 = {}^t v J_{j,n-j} v,$$

avec $J_{j,n-k} = \begin{pmatrix} -I_j & 0 \\ 0 & I_{n-j} \end{pmatrix}$; on dit que $f''(a)$ lui est ainsi équivalente. Donc la conclusion du lemme de Morse s'exprime par l'une quelconque des trois équivalences suivantes :

$$f(x) \sim f(0) + \frac{1}{2}f''(0) \cdot (x - a)^2 \sim {}^t v \cdot J_{j,n-j} \cdot v.$$

Dans le cas $n = 2$, il existe trois uniques formes normales possibles, représentées figure 23.1.

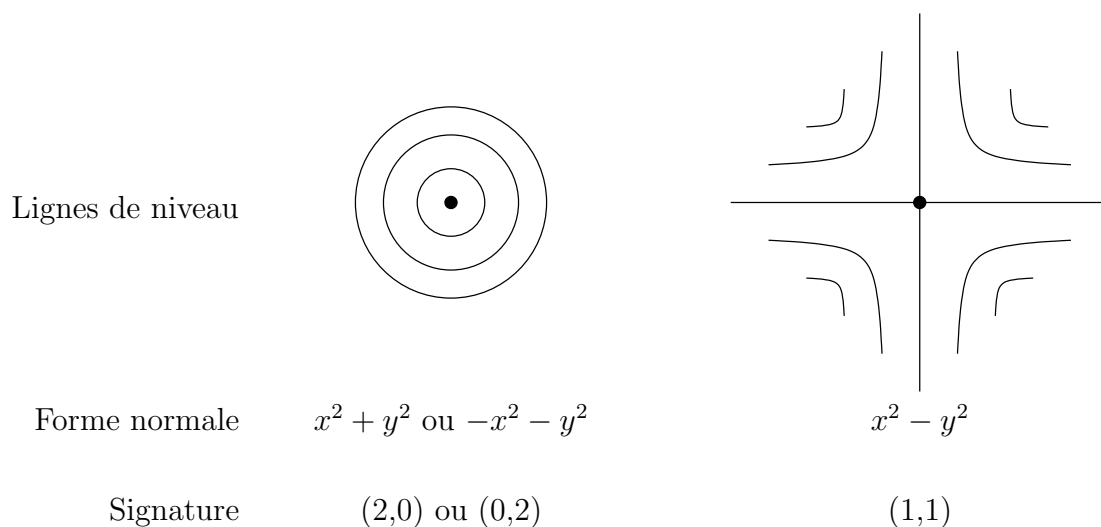


FIGURE 23.1 – Les trois formes normales possibles en dimension 2, avec le dessin de leurs lignes de niveau et leur signature.

Démonstration. Quitte à remplacer f par $\xi \mapsto f(a + \xi) - f(a)$, on peut supposer que $f : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}, 0)$. La formule de Taylor à l'ordre 1 s'écrit

$$f(x) = {}^t x \cdot H(x) \cdot x, \quad H(x) = \int_0^1 (1-t)f''(tx) dt \in M_{nn}(\mathbb{R}).$$

1. La conclusion du lemme de Morse reste vraie si f est seulement de classe C^2 , ce que nous ne démontrerons pas ici.

L'application H ainsi définie est de classe C^1 , puisque f est supposée de classe C^3 . On voit de plus que $H(0) = \frac{1}{2}f''(0)$, donc que $H(0)$ est symétrique et non dégénérée.

D'après l'exercice 22.6, il existe une application $P : (S_n(\mathbb{R}), H(0)) \rightarrow (M_n(\mathbb{R}), I)$ de classe C^∞ telle que, localement, ${}^tP(H)H(0)P(H) = H$. On a

$$f(x) = {}^t(P(H(x)) \cdot x) \cdot H(0) \cdot (P(H(x)) \cdot x).$$

L'application $\psi : (\mathbb{R}^n, 0) \hookrightarrow, x \mapsto y = P(H(x)) \cdot x$ est un difféomorphisme, puisque sa dérivée $\psi'(0) = P(H(0)) = I$ est inversible. Il ne reste qu'à poser $\varphi = \psi^{-1}$ pour avoir la formule voulue, à savoir

$$f \circ \varphi(y) = {}^ty \cdot H(0) \cdot y = \frac{1}{2}{}^ty \cdot f''(0) \cdot y.$$

Enfin on voit que

$$\varphi'(0) = \psi'(0)^{-1} = I.$$

□

23.a Exercice.

Tracer, au voisinage de l'origine, l'ensemble d'équation

$$x^2 - y^2 + 2x^3 + y^3 = 0.$$

Préciser les tangentes au point double et la position de la courbe par rapport aux tangentes.

23.5 Corollaire. *Un point critique non dégénéré a de f est un minimum local si et seulement si la forme quadratique $f''(a)$ est positive.*

23.b Exercice (Distance à une hypersurface).

Soient $\varphi : (\mathbb{R}^n, t) \rightarrow (\mathbb{R}^{n+1}, a)$ une immersion de classe C^∞ , et $\theta : (\mathbb{R}^n, t) \rightarrow \mathbb{R}^{n+1}$ une application C^∞ telle que, pour tout u , $\theta(u)$ soit un vecteur unitaire orthogonal à $\text{Im } \varphi'(u)$.

1. Montrer que l'application $F : (\mathbb{R}^n \times \mathbb{R}, (t, 0)) \rightarrow \mathbb{R}^{n+1}$, $(u, v) \mapsto \varphi(u) + v\theta(u)$ est un difféomorphisme local.

2. Montrer que la fonction $v(x) = \text{pr}_2 \circ F^{-1}(x)$ satisfait l'équation eikonale

$$\|v'(x)\|^2 = (\partial_{x_1}v)^2 + \cdots + (\partial_{x_{n+1}}v)^2 = 1.$$

3. Soit de plus z est un point fixé de \mathbb{R}^{n+1} . On suppose que t est un point critique de la fonction

$$f : (\mathbb{R}^n, t) \rightarrow \mathbb{R}, \quad u \mapsto \|\varphi(u) - z\|^2.$$

Soient Q_1 et Q_2 les deux formes quadratiques sur \mathbb{R}^n définies par

$$Q_1(\tau) = \|\varphi'(t) \cdot \tau\|^2 \quad \text{et} \quad Q_2(\tau) = (\varphi''(t) \cdot \tau^2) \cdot \theta(t).$$

Montrer que $z - a$ et $\theta(t)$ sont colinéaires, et qu'il existe un réel α tel que

$$f(t + \tau) - f(t) = Q_1(\tau) - \alpha Q_2(\tau) + o(\|\tau\|^2).$$

En déduire une condition nécessaire et suffisante sur les valeurs propres de $Q_1^{-1}Q_2$ et sur α pour que t soit un minimum local strict de f .

23.c Exercice ★ (Théorème des jets suffisants).

Soient $f : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}$ une fonction et $r \geq 1$ un entier satisfaisant la propriété suivante :

Quelle que soit la fonction $\varphi : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}$ telle que $\varphi(x) = o(\|x\|^r)$, il existe un champ de vecteurs $\xi : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}^n$ tel que $\varphi(x) = Df(x) \cdot \xi(x)$ et $\|\xi(x)\| = o(\|x\|)$.

On veut montrer que toute fonction $g : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}$ telle que $g(x) - f(x) = o(\|x\|^r)$, g est équivalente à f :

Il existe un difféomorphisme local $\theta : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^n, 0)$ tel que $g \circ \theta = f$; en particulier, f est équivalente à son polynôme de Taylor d'ordre r .

Remarques :

1. Le théorème d'inversion locale pour les fonctions (qui est évident) se déduit de ce théorème dans le cas $r = 1$. En effet, supposons que $f : (\mathbb{R}, 0) \rightarrow \mathbb{R}$ est une fonction telle que $f'(0) \neq 0$. Par continuité, $f'(x) \neq 0$ si x est proche de 0. Donc l'hypothèse du théorème est satisfaite avec $\xi(x) = \varphi(x)/f'(x)$. La conclusion appliquée à la partie affine $g(x) = f(0) + f'(0)x$ de f affirme que f est *linéarisable* (affinisable, devrait-on dire) : il existe un difféomorphisme local $\theta : (\mathbb{R}, 0) \rightarrow (\mathbb{R}, 0)$ tel que

$$f \circ \theta^{-1}(y) = f(0) + f'(0)y,$$

donc tel que

$$f \left(\theta^{-1} \left(\frac{z - f(0)}{f'(0)} \right) \right) = z ;$$

donc f est un difféomorphisme local.

2. Le lemme de Morse se déduit du cas $r = 2$ de ce théorème.² En effet, supposons que $f : (\mathbb{R}^n, 0) \rightarrow \mathbb{R}$ est une fonction dont l'origine dans \mathbb{R}^n soit un point critique ($Df(0) = 0$) non dégénéré ($D^2f(0)$ inversible). D'après la formule de Taylor, il existe une application F localement à valeurs dans les formes bilinéaires symétriques non dégénérées telles que

$$f(x) = f(0) + F(x) \cdot x^2 = f(0) + \sum_{i,j} F_{i,j}(x) x_i x_j$$

et telle que $F(0) = D^2f(0)/2$. La différentielle de f satisfait

$$Df(x) \cdot \xi = 2F(x) \cdot (x, \xi) + DF(x) \cdot (x^2, \xi), \quad DF(x) \cdot (x^2, \xi) = \sum_{i,j,k} x_i x_j DF_{i,j}(x) \cdot \xi.$$

Soit $\varphi : (\mathbb{R}^n, 0) \rightarrow 0$ une fonction nulle à l'ordre 2. D'après la formule de Taylor, de même on a $\varphi(x) = \Phi(x) \cdot x^3$. Pour vérifier l'hypothèse du

2. Le cas $r = 2$ est parfois appelé « lemme $\mathcal{M}^2 \subset \mathcal{M}^2 \mathcal{J}$ », parce que l'hypothèse revient à supposer que le carré de l'idéal maximal \mathcal{M} de l'anneau des germes de fonctions en 0 est inclus dans le produit de \mathcal{M}^2 par l'idéal jacobien \mathcal{J} de f , idéal engendré par les dérivées partielles de f .

théorème, il faut trouver un champ de vecteurs $\xi : x \mapsto \xi(x) = \bar{\xi}(x) \cdot x$ s'annulant au premier ordre tel que $Df(x) \cdot \xi(x) = \varphi(x)$, c'est-à-dire tel que

$$(2F(x) \cdot x + DF(x) \cdot x^2) \cdot \bar{\xi}(x) \cdot x = \Phi(x) \cdot x^2.$$

Or

$$2F(x) + DF(x) \cdot x = D^2f(0) + O(\|x\|)$$

est inversible si x est suffisamment proche de 0. Donc $\xi(x)$ existe bien au voisinage de 0 et s'annule au premier ordre. Le théorème appliqué à la partie quadratique

$$g(x) = f(0) + F(0) \cdot x^2$$

de f qu'il existe un difféomorphisme local θ tel que

$$f \circ \theta^{-1}(y) = f(0) + F(0) \cdot y^2.$$

En composant encore à droite par un isomorphisme $z \mapsto y = Az$ qui mette la forme bilinéaire $F(0)$ sous la forme canonique

$$F(0) \cdot (Az)^2 = - (z_1^2 + \dots + z_k^2) + (z_{k+1}^2 + \dots + z_n^2),$$

où k est l'indice de $F(0)$, on obtient l'expression voulue :

$$f \circ (\theta^{-1} \circ A)(z) = - (z_1^2 + \dots + z_k^2) + (z_{k+1}^2 + \dots + z_n^2).$$

3. Quand $r \geq 3$, le théorème généralise le lemme de Morse à des cas où la hessienne n'est pas inversible. La condition de non dégénérescence implique le développement limité (ou jet) de f à un ordre supérieur r . Le résultat est que f est équivalente à son développement limité (ou jet) d'ordre r : le jet d'ordre fini r de f est suffisant pour déterminer la classe d'équivalence de f sous l'action du groupe des difféomorphismes locaux. C'est un fait remarquable que l'étude des fonctions C^∞ qui tombent dans le champ d'application du théorème soit ramenée, modulo un changement de coordonnées à la source, à l'étude de polynômes. (Mais, comme le montre la fonction $e^{-1/x}$, dont la série de Taylor est nulle, il n'existe pas toujours un entier r permettant cette réduction.)

23.d Exercice ★ (B-A-BA de théorie de Morse³).

1. Soient S une surface de \mathbb{R}^n et $f : S \rightarrow \mathbb{R}$ une fonction. Soient $a < b$ tels que l'ensemble $f^{-1}([a, b])$ soit compact et ne contienne aucun point critique de f . Montrer que $f^{-1}(]-\infty, a])$ et $f^{-1}(]-\infty, b])$ sont difféomorphes [Mil63, Chap. 1, section 3].

2. Dessiner les courbes de niveau de la fonction

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto -\cos x + y^2;$$

3. Harold Calvin Marston MORSE, mathématicien américain (1892–1977), spécialiste de calcul des variations et l'un des fondateurs de la topologie différentielle.

on pourra chercher les points critiques de f , utiliser le lemme de Morse pour déterminer les deux niveaux de f au voisinage de ces points critiques, puis utiliser la question précédente pour interpoler entre les niveaux critiques. (Cette fonction s'interprète en Mécanique comme l'énergie d'un pendule pesant, où x est l'angle du pendule avec la verticale et y la vitesse angulaire.)

La théorie de Morse décrit comment les sous-niveau de f sont modifiés au travers d'une valeur critique, mais cela dépasse le cadre du cours.

Chapitre 24

Convexité

Mots-clefs du chapitre Enveloppe convexe, centre de masse d'une mesure

Soit E un espace vectoriel réel.

24.1 Définition. Une partie $X \subset E$ est *convexe* si, pour tous $x, y \in X$, le segment

$$[x, y] = \{(1 - t)x + ty, 0 \leq t \leq 1\}$$

est inclus dans X . Elle est *strictement convexe* si de plus, pour tous $x, y \in \partial X = \bar{X} - \overset{\circ}{X}$, le segment ouvert $]x, y[$ ne rencontre pas ∂X .

Convexe implique connexe par arc, mais pas le contraire.

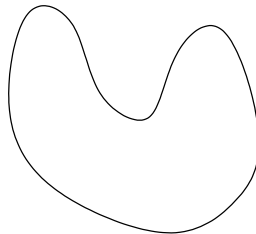


FIGURE 24.1 – Une partie connexe par arcs de \mathbb{R}^2 , non convexe

24.2 Proposition. 1) *L'intersection de toute famille de parties convexes de E est convexe.*

2) *L'image d'un convexe par une fonction affine est convexe.*

3) *L'image réciproque d'un convexe par une fonction affine est convexe.*

4) *Le produit d'une famille quelconque de parties non vides est convexe si et seulement si chaque partie est convexe.*

Démonstration. 1) Immédiat.

2) Cela découle du fait que l'image de $[x, x']$ par une fonction affine f est le segment $[f(x), f(x')]$.

3) $[x, x'] \subset f^{-1}([f(x), f(x')])$.

4) Si $A = \prod_i A_i$ est convexe, pour tout i A_i est la projection de A par la i -ième projection, qui est affine, donc A_i est convexe. Réciproquement, si A_i est convexe pour tout i , $\prod_i A_i = \cap_i \text{pr}_i^{-1}(A_i)$ est convexe. \square

24.a Exercice (Théorème de Minkowski).

Supposons E de dimension finie.

1. Soit C un convexe fermé non vide de E . Montrer que tout point $p \in E$ n'appartenant pas à C est strictement séparé de C par un hyperplan affine.

(Ce lemme est un cas particulier en dimension finie du théorème de Hahn-Banach.)

2. En déduire qu'un convexe fermé C de E est l'intersection des demi-espaces affines (quelconques, ou tous fermés, ou tous ouverts) qui le contiennent.

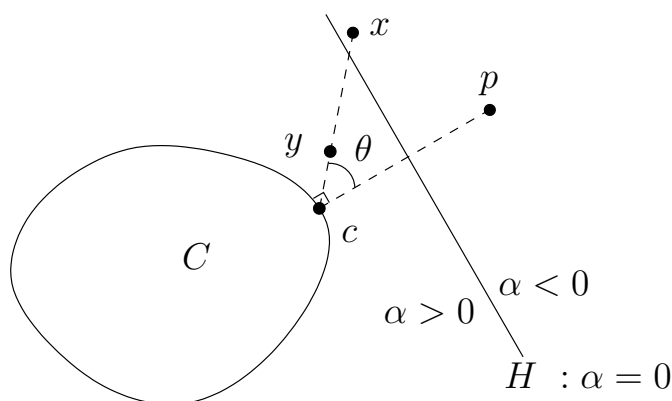


FIGURE 24.2 – Une propriété de séparation des convexes (première question de l'exercice 24.a)

24.3 Définition. L'enveloppe convexe d'une partie X de E est le plus petit convexe contenant X , i.e. l'intersection des convexes contenant X , i.e. l'ensemble des combinaisons linéaires barycentriques de points de X :

$$\text{conv}(X) = \{\alpha_0 x_0 + \dots + \alpha_n x_n, n \geq 1, x_i \in X, \alpha_i \geq 0, \alpha_0 + \dots + \alpha_n = 1\}.$$

24.b Exercice (Inégalité de Jensen).

Supposons E est de dimension finie n .

1. Montrer que l'enveloppe convexe d'une partie X de E est

$$\text{conv}(X) = \{\alpha_0 x_0 + \dots + \alpha_n x_n, x_j \in X, \alpha_j \geq 0, \alpha_0 + \dots + \alpha_n = 1\}$$

(théorème de Carathéodory).

2. Si X est fermée, $\text{conv}(X)$ est-elle automatiquement fermée ?

On suppose dorénavant X compacte.

3. Montrer que $\text{conv}(X)$ est compacte.

4. Montrer que $\text{conv}(X)$ est l'intersection des demi-espaces affines (quelconques, ou tous ouverts, ou tous fermés) contenant X .

Soient m une mesure de E borélienne, finie et à support compact. Le *centre de masse* de m est le point

$$c(m) = \frac{1}{m(E)} \int_E x \, dm \in E.$$

5. Montrer que $c(m)$ appartient à l'enveloppe convexe du support de m .

6. En déduire que, si $f : I$ intervalle ouvert borné $\subset \mathbb{R} \rightarrow \mathbb{R}$ est une fonction convexe et si μ est une probabilité sur I ,

$$f\left(\int_I x \, d\mu\right) \leq \int_I f \, d\mu.$$

Remarque : Comme la formule d'intégration par rapport à une mesure image le montre, le membre de droite de l'inégalité est le centre de masse

$$\int_I f(x) \, d\mu(x) = \int_{f(I)} y \, d(f_*\mu)(y)$$

de la loi $f_*\mu$ de f . Donc l'inégalité de Jensen affirme que :

$$f(c(\mu)) \leq c(f_*\mu).$$

24.c Exercice (Théorème de Gauss-Lucas). Soit P un polynôme complexe. Montrer que les points critiques de P sont dans l'enveloppe convexe de $P^{-1}(a)$ pour tout $a \in \mathbb{C}$ (théorème de Gauss-Lucas).

(Ce théorème est un analogue complexe du théorème de Rolle. Il est élémentaire si P est de degré 2 (l'unique point critique de P est alors au milieu du segment des deux racines), ou quand $|a|$ est grand.)

Chapitre 25

Fonctions convexes

25.1 Définition. Une fonction $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est *convexe* si l'épigraphe

$$E_f = \{(x, y) \in X \times \mathbb{R}, y \geq f(x)\}$$

de f est convexe (ce qui implique que X le soit), i.e. X est convexe et

$$f((1-t)x + tx') \leq (1-t)f(x) + tf(x') \quad (\forall x, x' \in X, t \in [0, 1]).$$

Elle est *strictement convexe* si

$$f((1-t)x + tx') < (1-t)f(x) + tf(x') \quad (\forall x, x' \in X, t \in]0, 1[).$$

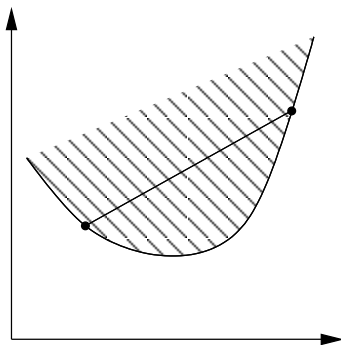


FIGURE 25.1 – Une fonction est convexe si son épigraphe est convexe

25.2 Proposition. Soient a un point intérieur à un convexe X de E et $f : X \rightarrow \mathbb{R}$ une fonction convexe et majorée sur un voisinage de a . Alors f est continue en a . Plus précisément, si $f(x) - f(a) \leq M$ sur une petite boule fermée $B(a, r)$ ($r > 0$),

$$|f(x) - f(a)| \leq \frac{M}{r} \|x - a\| \quad (\forall x \in B(a, r)).$$

Démonstration. On peut supposer que $a = 0$ et $f(a) = 0$, après quoi la démonstration se lit sur la figure 25.2 □

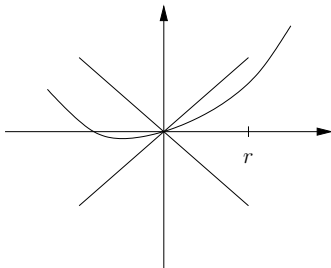


FIGURE 25.2 – Continuité d'une fonction localement convexe

25.3 Proposition. Soit X un ouvert convexe de \mathbb{R}^n . Une fonction dérivable $f : X \rightarrow \mathbb{R}$ est convexe si et seulement si

$$f(\hat{x}) - f(x) \geq f'(x) \cdot (\hat{x} - x) \quad (\forall x, \hat{x} \in X).$$

Géométriquement, ce critère de convexité revient à dire que le graphe de f est situé au-dessus de ses tangentes. Cela découle du théorème de Minkowski, mais nous en donnons maintenant une démonstration directe.

Démonstration. Supposons f convexe. Soient $x, \hat{x} \in X$. Notons $\varphi(t) = f((1-t)x + t\hat{x})$. Par définition de la convexité,

$$\frac{\varphi(t) - \varphi(0)}{t} \leq \varphi(1) - \varphi(0) = f(\hat{x}) - f(x) \quad (\forall t \in [0, 1]).$$

En faisant tendre t vers 0, on voit que

$$\varphi'(0) = f'(x) \cdot (\hat{x} - x) \leq f(\hat{x}) - f(x).$$

Réciproquement, appliquons l'inégalité de l'énoncé entre x et $z = (1-t)x + t\hat{x}$ et entre \hat{x} et z ($t \in [0, 1]$) :

$$f'(x) \cdot (x - z) \leq f(x) - f(z) \quad \text{et} \quad f'(z) \cdot (\hat{x} - z) \leq f(\hat{x}) - f(z).$$

En faisant une combinaison linéaire de ces deux inégalités, on obtient

$$0 = f'(z) \cdot ((1-t)x + t\hat{x} - z) \leq (1-t)f(x) + tf(\hat{x}) - f(z),$$

donc f est convexe. □

25.4 Corollaire. Soient X est un ouvert convexe de \mathbb{R}^n et $f : X \rightarrow \mathbb{R}$ convexe. Si f dérivable en $a \in X$ et si a est un point critique de f , a est un minimum global de f .

Démonstration. La démonstration de la proposition 25.3 montre que, pour tout $x \in X$,

$$f(x) - f(a) \geq f'(a) \cdot (x - a)$$

(l'argument est encore valide, même si f n'est pas dérivable sur U entier). Comme on suppose que le membre de droite est nul, la conclusion en découle. □

25.a Exercice (Fonctions convexes sur un intervalle de \mathbb{R}).

Soit $f : I$ intervalle de $\mathbb{R} \rightarrow \mathbb{R}$ une fonction convexe.

1. Montrer que les propriétés suivantes sont équivalentes :

1. f est convexe
2. pour tous $x_0 < x_1$ dans I ,

$$f\left(\frac{x_0 + x_1}{2}\right) \leq \frac{f(x_0) + f(x_1)}{2}.$$

3. pour tous $x_0 < x < x_1$ dans I ,

$$\det \begin{pmatrix} 1 & x_0 & f(x_0) \\ 1 & x & f(x) \\ 1 & x_1 & f(x_1) \end{pmatrix} \geq 0.$$

4. le taux d'accroissement

$$\frac{f(y) - f(x)}{y - x} \quad (x < y, x, y \in I)$$

est une fonction croissante de chacune de ses variables.

2. Retrouver la propriété que, si f est convexe, f est continue sur l'intérieur de I .

3. Montrer que, si f est de classe C^1 , f est convexe si et seulement si f' est croissante.

4. Montrer que, si f est de classe C^2 , f est convexe si et seulement si $f'' \geq 0$.

5. Montrer l'*inégalité de Jensen* : si f est convexe et si $x_1, \dots, x_n \in I$, $p_1, \dots, p_n \in [0, 1]$ et $\sum p_k = 1$,

$$f(p_1x_1 + \dots + p_nx_n) \leq p_1f(x_1) + \dots + p_nf(x_n).$$

6. Montrer que, pour un triangle dont on note α, β, γ les mesures d'angles dans $]0, \pi[$,

$$\frac{1}{\sin \alpha} + \frac{1}{\sin \beta} \geq \frac{8}{3 + 2 \cos \gamma}.$$

25.b Exercice (Géodésiques du plan euclidien).**25.c Exercice.**

1. Montrer que si $f : E \rightarrow \mathbb{R}$ est convexe et majorée, elle est constante.

2. Montrer que, si E est de dimension finie et $f : X$ convexe $\subset E \rightarrow \mathbb{R}$ est convexe, f est continue en tout point intérieur de X .

25.5 Théorème. Si f est de classe C^2 sur un ouvert convexe X de \mathbb{R}^n , f est convexe si et seulement si sa hessienne est positive en tout point.

Si la hessienne est positive définie en tout point, f est strictement convexe.

La réciproque de la seconde affirmation est fautive, comme le montre l'exemple de x^4 .

Démonstration. Supposons f convexe. Fixons $x \in X$ et $\xi \in \mathbb{R}^n$. D'après la proposition 25.3,

$$f(x + t\xi) - f(x) \geq tf'(x) \cdot \xi$$

pour tout t assez petit. Or, d'après la formule de Taylor,

$$0 \leq f(x + t\xi) - f(x) - tf'(x) \cdot \xi = \frac{t^2}{2} f''(x) \cdot \xi^2 + o(t^2).$$

Donc

$$f''(x) \cdot \xi^2 + o(1) \geq 0.$$

En faisant tendre t vers 0, on voit que la forme bilinéaire $f''(x)$ est positive.

Réciproquement, supposons que $f''(x)$ soit positive pour tout $x \in X$. D'après la formule de Taylor, pour tout $\xi \in \mathbb{R}^n$ tel que $x + \xi \in X$,

$$f(x + \xi) - f(x) - f'(x) \cdot \xi = \int_0^1 (1-t) f''(x + t\xi) \cdot \xi^2 dt \geq 0,$$

ce qui dit, d'après la proposition 25.3, que f est convexe. □

25.6 Théorème. *L'ensemble des minima d'une fonction convexe est convexe, et réduit à un singleton si la fonction est strictement convexe.*

Tout minimum local d'une fonction convexe est global.

Démonstration. Exercice. □

25.d Exercice (Fonction fortement convexe).

Soit $f : X$ convexe fermé $\subset \mathbb{R}^n$ continue. On dit que f est α -convexe s'il existe $\alpha > 0$ tel que

$$f(x) - \frac{\alpha}{2} \|x\|^2$$

soit convexe. (On dit aussi *fortement convexe*.)

1. Montrer que, si f est α -convexe, f est minorée, puis qu'elle possède un unique minimum.

On suppose de plus que f est de classe C^2 .

2. Montrer que f est α -convexe si et seulement si

$$f''(x) \cdot (v, v) \geq \alpha \|v\|^2 \quad (\forall x \in X, v \in \mathbb{R}^n).$$

3. Montrer que, si f est α -convexe, elle possède un minimum unique a , que de plus toute suite minimisante converge vers a , et que l'on a

$$\|x - a\|^2 \leq \frac{4}{\alpha} (f(x) - f(a)).$$

25.e Exercice (Algorithme de relaxation).

Soit $f : X$ convexe fermé $\subset \mathbb{R}^n$ continue et α -convexe. Construisons la suite (x^k) de \mathbb{R}^n par l'*algorithme de relaxation* suivant, qui consiste à minimiser f par rapport à chacune de ses variables, prises dans un ordre cyclique fixe :

- Commençons avec un point $x^0 = (x_1^0, \dots, x_n^0) \in X$ quelconque.
- La fonction $t \mapsto f(t, x_2^0, \dots, x_n^0)$ possède un unique minimum t_1^0 .
- Puis on appelle t_2^0 l'unique minimum de la fonction $t \mapsto f(t_1^0, t, x_3^0, \dots, x_n^0)$.
- ...
- Puis on appelle t_n^0 l'unique minimum de la fonction $t \mapsto f(t_1^0, t_2^0, \dots, x_{n-1}^0, t)$.
- On pose $x^{0,i} = (t_1^0, \dots, t_i^0, x_{i+1}^0, \dots, x_n^0)$ pour tout i et $x^1 = (t_1^0, \dots, t_n^0)$, et l'on recommence en remplaçant x^0 par x^1 , et ainsi de suite.

Montrer que (x^k) converge vers le minimum de f .

25.7 Exemple (Mesures de Gibbs). Soit $X = \{1, \dots, N\}$ l'ensemble fini des états d'un système physique. À chaque état x est associée une énergie $E_x \in \mathbb{R}$. Un *état statistique* du système est une probabilité $p = (p_x)$ sur X . À chaque état statistique on associe deux grandeurs fondamentales :

- l'énergie $\mathcal{E}(p)$, soit l'espérance de E :

$$\mathcal{E}(p) = \sum_{x \in X} p_x E_x.$$

- l'entropie

$$\mathcal{S}(p) = \sum_{x \in X} p_x \ln p_x.$$

Soient E_- et E_+ les bornes inférieure et supérieure de E , et $I = [E_-, E_+]$, de sorte que, pour tout x , $E_x \in I$. Un état statistique est un *état d'équilibre* s'il maximise l'entropie parmi les états de son énergie. Pour toute énergie $\mathcal{E} \in I$, un état d'équilibre existe par compacité du polyèdre

$$\mathcal{P}_{\mathcal{E}} = \left\{ p, \sum_x p_x = 1, p_x \geq 0, \sum_x p_x E_x = \mathcal{E} \right\}.$$

Il est unique si I n'est pas réduit à un point, par convexité de l'entropie, et situé dans l'intérieur de $\mathcal{P}_{\mathcal{E}}$. D'après le théorème de Lagrange, il existe deux nombres $\lambda, \mu \in \mathbb{R}$ tels que, pour tout $x \in X$,

$$\ln p_x = \lambda - \mu E_x,$$

donc

$$p_x = \frac{1}{Z} e^{-\mu E_x}, \quad Z := \sum_x e^{-\mu E_x}.$$

En Thermodynamique, on note k la *constante de Boltzmann* et la quantité $T = 1/(k\mu)$ s'appelle la *température*. On vérifie que \mathcal{S} est une fonction dérivable sur $\mathcal{P}_{\mathcal{E}}$ et que

$$\frac{\partial \mathcal{S}}{\partial \mathcal{E}} = \frac{1}{kT}.$$

Autres exercices

25.f Exercice.

1. Soit $f : [a, b] \rightarrow \mathbb{R}$ continue convexe. Montrer que

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(t) dt \leq \frac{f(a) + f(b)}{2}$$

(inégalité de Hermite-Hadamard). Quels sont les cas d'égalité ?

Nous allons appliquer cette inégalité dans plusieurs cas.

2. Soient $g : [a, b] \rightarrow \mathbb{R}$ de classe C^2 , $m = \min g''$ et $M = \max g''$. Montrer que

$$\begin{cases} \frac{m(b-a)^2}{24} \leq \frac{1}{b-a} \int_a^b g(t) dt \leq \frac{M(b-a)^2}{24} \\ \frac{m(b-a)^2}{12} \leq \frac{g(a)+g(b)}{2} \leq \frac{M(b-a)^2}{12}. \end{cases}$$

3. Montrer que, pour tous $x, y > 0$,

$$\sqrt{xy} \leq \frac{x-y}{\ln x - \ln y} \leq \frac{x+y}{2}$$

(inégalité arithmético-logarithmico-géométrique).

4. Montrer que, si $x \geq 0$,

$$x - \frac{x^2}{2+x} \leq \ln(1+x) \leq x - \frac{x^2}{2(1+x)},$$

en déduire que, pour tout $n \in \mathbb{N}_*$,

$$\frac{1}{n + \frac{1}{2}} \leq \ln(n+1) - \ln n \leq \frac{1}{2} \left(\frac{1}{n} + \frac{1}{n+1} \right).$$

En déduire la formule de Stirling :

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

25.g Exercice (Méthode du trapèze).

1. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ convexe dérivable. Montrer que, pour tout $n \in \mathbb{N}_*$,

$$0 \leq \frac{f(0)}{2} + f(1) + f(2) + \cdots + f(n-1) + \frac{f(n)}{2} - \int_0^n f \leq \frac{f'(n) - f'(0)}{8}.$$

2. Interpréter cette inégalité en terme d'erreur dans le calcul approché d'une intégrale par la méthode du trapèze.

Chapitre 26

Et la dimension infinie ? ★

Mots-clefs du chapitre Espace de Banach, facteur direct

26.1 Lemme. Soit $f : E \rightarrow F$ une application linéaire entre espaces vectoriels normés. Les propriétés suivantes sont équivalentes :

1. f est continue en 0
2. f est continue sur E
3. f est lipschitzienne
4. la norme d'opérateur de f (subordonnée à celles de E et de F) est finie.

Démonstration. (1) \Rightarrow (2) Si f est continue en 0, pour tout $x \in E$, quand h tend vers 0, on a

$$f(x+h) = f(x) + f(h) \rightarrow f(x)$$

donc f est continue en x .

(2) \Rightarrow (3) Supposons f continue. En particulier, il existe $\eta > 0$ tel que $\|f(k)\| \leq 1$ pour tout k tel que $\|k\| \leq \eta$. Donc

$$\|f(x+h) - f(x)\| = \|f(h)\| = \frac{\|h\|}{\eta} \left\| f\left(\eta \frac{h}{\|h\|}\right) \right\| \leq \frac{1}{\eta} \|h\|,$$

donc f est lipschitzienne.

(3) \Rightarrow (4) Si f est lipschitzienne, pour tout x tel que $\|x\| \leq 1$,

$$f(x) \leq \text{lip } f \|x\| \leq \text{lip } f < \infty.$$

(4) \Rightarrow (1) Trivial. □

En dimension finie, en prenant une base on voit que f est automatiquement continue. L'une des difficultés est que ce n'est plus le cas en dimension infinie.

26.2 Exemple (Application linéaire discontinue). L'application $f : \mathbb{R}[X] \hookrightarrow P, P \mapsto P'$, n'est pas continue quand on munit $\mathbb{R}[X]$ de la norme $\|\sum_j a_j X^j\|_\infty = \max_j |a_j|$. En effet, on a $f(X^k) = kX^{k-1}$, donc

$$\|f(X^k)\|_\infty = k \rightarrow_{k \rightarrow \infty} \infty$$

tandis que $\|X^k\|_\infty = 1$ pour tout k . Donc $\|f\| = \infty$.

Dans les espaces vectoriels normés, quand on parle d'application linéaire, on sous-entend généralement qu'elle doit être continue ; pour nous, $L(E, F)$ désignera donc l'espace des applications linéaires continues de E vers F . On note donc E' le *dual topologique* de E , soit l'espace des formes linéaires continues sur E . De même, un "isomorphisme" doit avoir un inverse continu.

Notre définition de la dérivée se généralise mot pour mot aux fonctions entre ouverts d'espaces normés.

26.3 Exemple. Pour une application linéaire discontinue $f : E \rightarrow F$, $f'(x) \cdot v = f(v)$ pour tout v , donc $f'(x)$ n'est pas continue.

26.a Exercice. La norme uniforme sur l'espace des fonctions continues sur un ouvert U de \mathbb{R}^n est-elle différentiable en dehors de la fonction nulle ?

Pour généraliser facilement le théorème d'inversion locale en dimension infinie, il faut que la suite de Cauchy construite dans la démonstration du théorème du point fixe converge (exercice 4.a). Le plus simple est de se supposer que l'espace normé considéré est complet.

26.4 Définition. Un *espace de Banach* est un espace vectoriel normé complet.

26.b Exercice (Inversion locale entre espaces de Banach).

Soient E et F deux espaces de Banach et $f : (E, a) \rightarrow F$ une application de classe C^1 . Montrer que f est un difféomorphisme local si et seulement si $f'(a)$ est un isomorphisme.

26.5 Définition. Un sous-espace F d'un espace de Banach est un *facteur direct* s'il existe un isomorphisme φ de E sur le produit $V \times W$ de deux espaces de Banach, tel que $F = \varphi^{-1}(V \times \{0\})$; autrement dit, F est fermé et possède un supplémentaire fermé.

La définition ci-dessus est justifiée par le fait qu'il existe des sous-espaces qui ne sont pas facteurs directs, ce que nous admettrons.¹

26.c Exercice (Exemples de facteurs directs).

1. Montrer qu'un sous-espace de dimension finie ou de codimension finie est facteur direct.

2. Montrer qu'un sous-espace fermé d'un espace de Hilbert est facteur direct.

26.6 Définition. Soient E et F deux espaces de Banach et $f : (E, a) \rightarrow (F, b)$ une application de classe C^1 . Cette application est une *submersion* (resp. une *immersion*) en a si $f'(a)$ est surjective (resp. injective) et si le noyau (resp. l'image) de $f'(a)$ est facteur direct.

1. Par exemple, on peut montrer que l'espace c_0 des suites de limite nulle est fermé mais pas facteur direct dans l'espace ℓ^∞ . En effet, ce dernier est faiblement séparable (i.e. il existe une suite bornée (α_n) de $(\ell^\infty)'$ telle que, pour tout $u \in \ell^\infty$, $\|u\|_\infty \leq \sup_n |\alpha_n(u)|$). Si c_0 était facteur direct, le quotient ℓ^∞/c_0 aussi serait faiblement séparable, ce qui n'est pas le cas.

26.d Exercice. Montrer les théorèmes de submersion, d'immersion et de fonction implicite pour des applications entre espaces de Banach. Montrer les théorèmes du rang et du corang constants.

26.e Exercice (Théorème de Cauchy-Lipschitz).

Soient E un espace de Banach et v un champ de vecteurs lipschitzien sur un voisinage ouvert U d'un point $a \in E$.

1. Il existe une unique application lipschitzienne locale $\phi : (U \times \mathbb{R}, (a, 0)) \rightarrow (U, a)$, $(x, t) \mapsto \phi_t(x)$ telle que pour tout x l'application partielle $t \mapsto \phi_t(x)$ soient de classe C^1 et qui satisfasse le problème de Cauchy suivant :

$$\begin{cases} \frac{d}{dt}(\phi_t(x)) = v(\phi_t(x)) & (\forall x, t) \\ \phi_0(x) = x & (\forall x). \end{cases}$$

2. Si v est de classe C^k ($0 \leq k \leq \infty, \Omega$), il en est de même pour ϕ et, si $k < \infty$, pour tout x l'application partielle $t \mapsto \phi_t(x)$ est de classe C^{k+1} .

L'application ϕ est le *flot local* de v , et réciproquement $v = \frac{d\phi}{dt} \circ \phi_t^{-1}$ est le *générateur infinitésimal* de ϕ .

Avec les *fonctionnelles*, c'est-à-dire des applications entre espaces de fonctions, le calcul infinitésimal prend le nom de *calcul des variations*.

26.f Exercice (Hamilton [Ham82, p. 75]).

On considère une bulle de savon s'appuyant sur deux cercles parallèles dans \mathbb{R}^3 , situés dans les plans $x = \pm 1$, et de même rayon $r > 0$. En Statique des fluides, on montre que la bulle de savon prend la forme d'une surface de révolution obtenue en faisant tourner la courbe $y = f(x)$ autour de l'axe des x , avec la contrainte $f(\pm 1) = r$, et que cette surface est d'aire $A(f)$ minimale.

1. Montrer que

$$A(f) = \int_{-1}^1 2\pi f \sqrt{1 + f'^2} dx.$$

2. En déduire que

$$f f'' = f'^2 + 1;$$

on pourra montrer que la fonction

$$A : E = \{f \in C^\infty([-1, 1]), f(\pm 1) = r\} \rightarrow \mathbb{R}, \quad f \mapsto A(f)$$

a toutes ses dérivées directionnelles nulles.

3. Montrer qu'il existe des solutions si r est suffisamment grand; on pourra remarquer que ces solutions sont de la forme $f(x) = \frac{1}{s} \cosh(sx)$, pour un $s \in \mathbb{R}_*$ que l'on caractérisera en fonction de r .

Terminons ce chapitre par l'étude des "géodésiques d'une surface de révolution". Soit

$$g : U \text{ ouvert } \subset \mathbb{R}^2 \rightarrow \text{Bil}(\mathbb{R}^2 \times \mathbb{R}^2, \mathbb{R})$$

une application (infiniment différentiable) à valeur dans le cône des produits scalaires euclidiens ; g est une *métrique riemannienne* sur U , et permet de mesurer la longueur des vecteurs vitesses de chemins tracés sur U d'une façon qui dépend du point de U .

26.7 Exemple. Soit S une surface de révolution, d'équation $\rho = \rho(z)$ en coordonnées cylindriques (ρ, θ, z) , où $\rho : I$ intervalle $\subset \mathbb{R} \rightarrow]0, +\infty[$ est une fonction de classe C^∞ . L'application

$$r : \mathbb{T} \times I \rightarrow \mathbb{R}^3, \quad (\theta, z) \mapsto (\rho(z) \cos \theta, \rho(z) \sin \theta, z)$$

est un paramétrage de S (immersion injective et homéomorphisme). La *première forme fondamentale* de S est la restriction du produit scalaire euclidien de \mathbb{R}^3 aux vecteurs tangents à S , et l'expression de cette forme fondamentale dans les coordonnées (θ, z) est une métrique riemannienne g sur $\mathbb{T} \times I$ qui, identifiée à un application à valeurs dans les formes quadratiques, vérifie

$$g_{(\theta,z)}(\dot{\theta}, \dot{z}) = \left\| dr_{(\theta,z)} \cdot (\dot{\theta}, \dot{z}) \right\|^2 = \rho(z)^2 \dot{\theta}^2 + (1 + \rho'(z)^2) \dot{z}^2 ;$$

autrement dit, la matrice de $g_{\theta,z}$ dans la base canonique de \mathbb{R}^2 est

$$G_{(\theta,z)} = \begin{pmatrix} \rho(z)^2 & 0 \\ 0 & 1 + \rho'(z)^2 \end{pmatrix}.$$

Soient maintenant

$$L : TU = U \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (u, v) \mapsto L_u(v)$$

une fonction sur l'espace des positions et des vitesses et, pour tout chemin $\gamma : I = [0, 1] \rightarrow U$ de classe C^1 , \mathcal{A}_L la fonctionnelle

$$\mathcal{A}_L(\gamma) = \int_0^1 L_{\gamma(t)}(\dot{\gamma}(t)) dt.$$

26.8 Exemple. Si L est la vitesse scalaire :

$$L_{(\theta,z)}(\dot{\theta}, \dot{z}) = \sqrt{g_{(\theta,z)}(\dot{\theta}, \dot{z})},$$

$\mathcal{A}_L(\gamma)$ est la longueur du chemin $r \circ \gamma$ tracé sur S .

Si L est le carré scalaire de la vitesse : $L_{(\theta,z)}(\dot{\theta}, \dot{z}) = g_{(\theta,z)}(\dot{\theta}, \dot{z})$ (énergie cinétique, à un facteur $1/2$ près, pour un point matériel de masse unité), $\mathcal{A}_L(\gamma)$ est l'*intégrale d'énergie* de γ .

Notons $\mathcal{C} = C_{a,b}^1(I, U)$ l'espace des chemins γ de classe C^1 tracés sur U tels que $\gamma(0) = a$ et $\gamma(1) = b$, $a, b \in U$; c'est un espace affine modelé sur un espace de Banach.

26.9 Remarque. Un vecteur tangent à \mathcal{C} en γ est un champ de vecteurs $\delta : I \rightarrow \mathbb{R}^2$ de classe C^1 tel que $\delta(0) = \delta(1) = 0$; autrement dit,

$$T_\gamma \mathcal{C} = C_{(0,0)}^1(I, \mathbb{R}^2).$$

Démonstration. En effet, \mathcal{C} est la sous-variété (banachique) de l'ouvert $C^1(I, U)$ de l'espace de Banach $C^1(I, \mathbb{R}^2)$ définie par l'équation

$$\gamma(0) = a, \quad \gamma(1) = b.$$

Autrement dit,

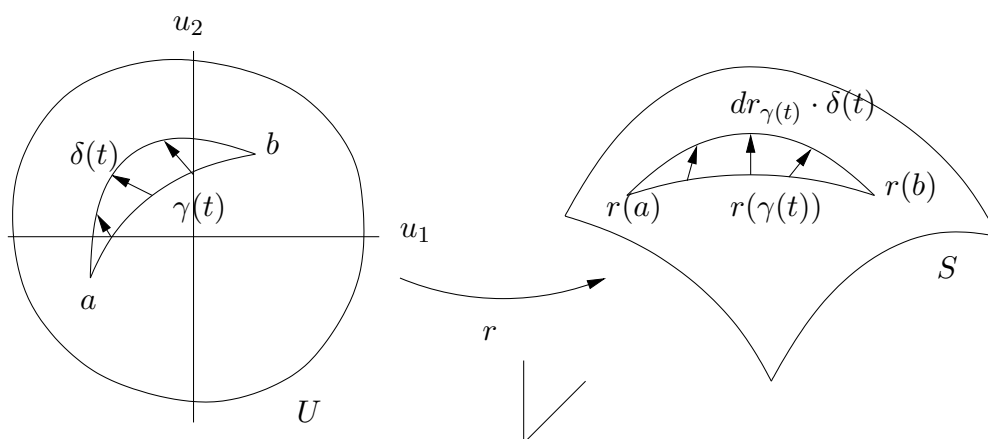
$$\mathcal{C} = \Phi^{-1}(a, b),$$

où $\Phi : C^1(I, U) \rightarrow U \times U$ est la submersion linéaire définie par $\Phi(\gamma) = (\gamma(0), \gamma(1))$. Donc $T_\gamma \mathcal{C} = \ker d\Phi_\gamma$, d'où le résultat voulu. \square

Nous allons nous intéresser aux points critiques de la fonctionnelle $\mathcal{A}_L : \mathcal{C} \rightarrow \mathbb{R}$. Nous nous contenterons de calculer la dérivée radiale de \mathcal{A}_L , ce qui ramène le problème à la dimension un.

26.10 Lemme. *Quelle que soit $\delta \in C^1_{(0,0)}(I, \mathbb{R}^2)$, la dérivée radiale $\left. \frac{\mathcal{A}_L(\gamma + \epsilon \delta)}{d\epsilon} \right|_{\epsilon=0}$ dans la direction de δ existe et vaut*

$$d\mathcal{A}_L(\gamma) \cdot \delta = \int_0^1 \left(\frac{\partial}{\partial v} L_{\gamma(t)}(\gamma'(t)) - \frac{d}{dt} \left(\frac{\partial}{\partial v} L_{\gamma(t)}(\gamma'(t)) \right) \right) \delta(t) dt.$$



26.11 Corollaire (Équation d'Euler-Lagrange). *Si γ est un extremum de \mathcal{A}_L ,*

$$\frac{\partial}{\partial u} L_{\gamma(t)}(\gamma'(t)) - \frac{d}{dt} \left(\frac{\partial}{\partial v} L_{\gamma(t)}(\gamma'(t)) \right) = 0 \quad (\forall t \in [0, 1]).$$

Démonstration. Supposons que l'équation d'Euler-Lagrange ne soit pas satisfaite : il existe un temps $t_0 \in [0, 1]$ et un indice $j \in \{1, 2\}$ tel que la quantité

$$[L]_j^\gamma(t) = \frac{\partial}{\partial u_j} L_{\gamma(t)}(\gamma'(t)) - \frac{d}{dt} \left(\frac{\partial}{\partial v_j} L_{\gamma(t)}(\gamma'(t)) \right)$$

soit non nulle à $t = t_0$. Par continuité, il existe un intervalle ouvert $J \subset [0, 1]$ contenant t_0 sur lequel $[L]_j^\gamma(t) \neq 0$. Quitte à restreindre J , on peut supposer que $\bar{J} \subset]0, 1[$ (alors J ne contient plus forcément le point t_0 initial, mais peu importe), et que $[L]_j^\gamma$ soit de signe constant, disons positif, sur J . Soit δ un champ de vecteur tel que $\delta(t) = 0$ en dehors de J , $\delta_k \equiv 0$ pour $k \neq j$, $\delta_j(t) \geq 0$ pour tout t , et $\delta(t) > 0$ sur un sous-intervalle de J . Alors $d\mathcal{A}_L(\gamma) \cdot \delta > 0$. D'après le théorème des accroissements finis, γ n'est donc pas un extremum. \square

26.12 *Exemple* (Équations de la Mécanique). Si $L : TU \rightarrow \mathbb{R}$ est de la forme

$$L_u(v) = \frac{1}{2} \|v\|^2 + V(u),$$

où $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ est une fonction de classe C^1 , les équations d'Euler-Lagrange s'écrivent

$$\gamma''(t) = -\frac{\partial V}{\partial u}(\gamma(t)).$$

Ces équations sont les équations de Newton, en Mécanique classique, d'un point matériel soumis à un champ de force dérivant du potentiel V . Ceci signifie que les trajectoires de ce système matériel sont les points critiques de l'action : le point matériel choisit, en un sens, parmi toutes les trajectoires possibles, celle qui est un point critique de \mathcal{A}_L .

26.13 *Remarque*. La longueur d'une courbe est invariante par changement de paramétrage de la courbe. Donc, si γ est un extremum, n'importe quel paramétrage de la courbe $\gamma(I)$ doit être solution de l'équation d'Euler-Lagrange associée à la fonction $L = \sqrt{g}$; dans ce sens, cette équation est dégénérée.

26.14 Définition. Une courbe paramétrée $\gamma \in \mathcal{C}$ est une *géodésique* si elle est une solution de l'équation d'Euler-Lagrange pour la fonction $L = g$ (intégrale d'énergie).

26.15 *Exemple*. Pour la surface de révolution, l'équation d'Euler-Lagrange est

$$\begin{cases} \frac{d}{dt}(\rho(z)^2\theta'(t)) = 0 \\ \rho(z(t))\rho'(z(t))\theta'(t)^2 - \rho'(z(t))\rho''(z(t))z'(t)^2 - (1 + \rho'(z(t))^2)z''(t) = 0 \end{cases}$$

L'intégrale de Clairaut (ou moment cinétique)

$$C : TU = U \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad C_{(\theta,z)}(\dot{\theta}, \dot{z}) = \rho(z)^2\dot{\theta}$$

est donc une fonction constante le long des trajectoires (i.e. $C_{\gamma(t)}(\gamma'(t)) \equiv \text{cte}$). Par ailleurs, un calcul direct montre qu'il en est de même de l'énergie

$$H : TU = U \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad H_{(\theta,z)}(\dot{\theta}, \dot{z}) = \frac{1}{2}(\rho'(z)^2 + 1)\dot{z}^2 + \frac{C_{(\theta,z)}(\dot{\theta}, \dot{z})^2}{2\rho^2}.$$

L'intégrale de Clairaut permet de déterminer comment la vitesse angulaire $\dot{\theta}$ varie en fonction de la distance ρ à l'axe de rotation de S . L'intégrale de l'énergie permet elle de déterminer comment la vitesse verticale \dot{z} varie en fonction de ρ , ρ' et C . Par exemple, si $\rho(z) = z^2 + 1$, quelles que soient les conditions initiales, le point matériel va osciller une infinité de fois entre deux valeurs z_0 et z_1 de l'altitude.

Annexe A

Topologie

Dans un ensemble X nu, il n'existe aucune notion de proximité entre les points et, par exemple, la question de savoir si une suite de X converge dans X n'a pas de sens.

A.1 Définition. Une *structure de voisinages* sur X est la donnée, pour tout point $x \in X$, d'un ensemble non vide $\mathcal{V}(x)$ de parties V de X appelées les *voisinages* de x , vérifiant les quatre axiomes suivants :

- (V₁) Réflexivité : $x \in V$ pour tout $V \in \mathcal{V}(x)$
- (V₂) Stabilité par croissance : $V \in \mathcal{V}(x)$ et $V \subset W$ entraînent $W \in \mathcal{V}(x)$
- (V₃) Stabilité par intersection finie : $V, W \in \mathcal{V}(x)$ entraîne $V \cap W \in \mathcal{V}(x)$
- (V₄) Existence de l'intérieur : si $V \in \mathcal{V}(x)$, il existe $W \subset V$ tel que $W \in \mathcal{V}(y)$ pour tout $y \in W$; W s'appelle l'*intérieur* de V , et se note $\overset{\circ}{V}$.

A.a Exercice (Voisinages dans un espace métrique).

Soit (X, d) un espace métrique. Pour $x \in X$, soit $\mathcal{V}(x)$ l'ensemble des parties $V \subset X$ telles qu'il existe $r > 0$ telles que $B(x, r) \subset V$. Montrer que ces $\mathcal{V}(x)$ définissent une structure de voisinages sur X .

Il s'ensuit immédiatement une notion de convergence des suites¹ : une suite $(x_n)_{n \in \mathbb{N}}$ converge vers $\ell \in X$ si

$$\forall V \in \mathcal{V}(\ell) \quad \exists N \quad \forall n \geq N \quad x_n \in V.$$

A.2 Exemple (Droite à deux origines). Soit $X = (\mathbb{R} \setminus \{0\}) \cup \{O_1, O_2\}$ (où O_1 et O_2 sont deux points n'appartenant pas à \mathbb{R} , par exemple $\pm i \in \mathbb{C}$). On prolonge à X la relation d'ordre de \mathbb{R}_* en posant

$$\begin{cases} x \leq O_1 \text{ et } x \leq O_2 & \text{si } x \leq 0 \\ O_1 \leq x \text{ et } O_2 \leq x & \text{si } 0 \leq x. \end{cases}$$

On n'obtient pas un ordre total parce que l'on n'a ni $O_1 \leq O_2$ ni $O_2 \leq O_1$. On peut tout de même définir, pour $a \leq b$, les intervalles ouverts

$$]a, b[= \{x \in X, a < x < b\}$$

1. La notion générale de convergence concerne les *filtres*.

($x < y$ signifie $x \leq y$ et $x \neq y$). Si $x \in X$, on définit alors $\mathcal{V}(x)$ comme l'ensemble des parties V de X qui contiennent un intervalle de la forme $]x - r, x + r[$, avec $r > 0$.

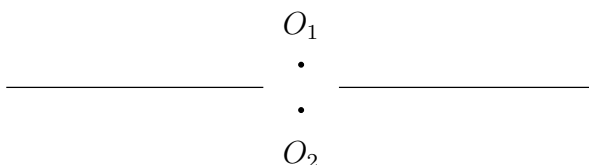


FIGURE A.1 – La droite à deux origines

La structure de voisinage ainsi définie ne provient d'aucune métrique puisque l'espace obtenu n'est pas *séparé*, au sens où l'intersection d'un voisinage de O_1 et d'un voisinage de O_2 n'est jamais vide. Une conséquence est que la suite $(1/n)_{n \geq 1}$ converge simultanément vers O_1 et vers O_2 (tandis que la limite d'une suite convergente serait unique si l'espace était séparé).

Soient X et Y deux ensembles munis des structures de voisinages \mathcal{V}_X et \mathcal{V}_Y . Rappelons que la notation

$$f : (X, x) \rightarrow Y$$

désigne une application d'un voisinage de x dans X à valeurs dans Y . On peut préciser $f : (X, x) \rightarrow (Y, y)$ si $y = f(x)$.

A.3 Définition (Continuité en un point). Une application $f : (X, x) \rightarrow (Y, y)$ est *continue* en x si

$$\forall W \in \mathcal{V}_Y(y) \quad f^{-1}(W) \in \mathcal{V}_X(x).$$

De façon équivalente, f est continue si et seulement si

$$\forall W \in \mathcal{V}_Y(y) \quad \exists V \in \mathcal{V}_X(x) \quad f(V) \subset W.$$

Dans le cas particulier où X et Y sont des espaces métriques, on retrouve la définition classique :

$$\forall \epsilon > 0 \quad \exists \eta > 0 \quad f(B(x, \eta)) \subset B(y, \epsilon).$$

A.4 Définition. Un *ouvert* de X est une partie de X qui est un voisinage de chacun de ses points.

Par exemple, un intervalle ouvert de \mathbb{R} , ou l'union de deux intervalles disjoints de \mathbb{R} , sont des ouverts de \mathbb{R} .

A.b Exercice (Continuité).

Montrer que $f : X \rightarrow Y$ est continue (sous-entendu : en tout point de X) si et seulement si, pour tout ouvert W de Y , $f^{-1}(W)$ est un ouvert de X .

Il s'avère que la donnée des voisinages est équivalente à celle des ouverts, et que la seconde est souvent plus commode à manipuler. Voici les axiomes auxquels les ouverts satisfont.

A.5 Définition. Une *topologie* sur X est un ensemble \mathcal{O} de parties de X , appelées *ouverts*, satisfaisant aux axiomes suivants :

- (O_1) Stabilité par union : $O_i \in \mathcal{O}$ entraîne $\cup_i O_i \in \mathcal{O}$
- (O_2) Stabilité par intersection finie : $O, O' \in \mathcal{O}$ entraîne $O \cap O' \in \mathcal{O}$
- (O_3) $\emptyset, X \in \mathcal{O}$.

L'ensemble X ainsi muni d'une topologie prend le nom d'*espace topologique*. Un *fermé* est alors une partie de X dont le complémentaire est ouvert.

A.c Exercice (Correspondance entre ouverts et voisinages).

Une topologie étant donnée, une partie V de X contenant un point x est un voisinage si et seulement si V contient un ouvert contenant x . Montrer qu'il est équivalent de se donner une structure de voisinage ou une topologie sur X .

A.6 Exemples. La *topologie grossière* sur X est $\{\emptyset, X\}$. La *topologie discrète* est $\mathcal{P}(X)$. Pour ces topologies extrêmes les notions de limite et de continuité sont caricaturales :

	Topologie grossière	Topologie discrète
Suite de X	toujours convergente	convergente \Leftrightarrow constante
Chemin sur X	toujours continu	continu \Leftrightarrow constant
Fonction sur X	continue \Leftrightarrow constant	toujours continue

A.d Exercice (Topologie initiale).

Soient X un ensemble, Y un espace topologique et $f : X \rightarrow Y$ une application. La *topologie initiale* de f est la topologie de X la plus petite pour l'inclusion (on dit aussi la plus *grossière*) qui rende f continue. Si f est injective, on parle de *topologie induite*.

1. Montrer que la topologie initiale existe.
2. Montrer qu'elle est caractérisée par la propriété universelle suivante : pour tout espace topologique Z , une application $g : Z \rightarrow X$ est continue si et seulement si $f \circ g$ est continue.

On généralise cela de façon évidente à une famille quelconque d'applications. En particulier, la *topologie faible* d'un espace vectoriel topologique E est la topologie initiale de ses formes linéaires continues $\alpha \in E'$.

A.e Exercice (Topologie finale).

Soient X un espace topologique, Y un ensemble, et $f : X \rightarrow Y$ une application. La *topologie finale* de f est la topologie de Y la plus grande pour l'inclusion (on dit aussi la plus *fine*) qui rende f continue. Si f est surjective, on parle de *topologie quotient* ; si f est injective (et pour toute topologie de Y plus petite que la topologie finale), f s'appelle un *plongement*.

1. Montrer que la topologie finale existe.
2. Montrer qu'elle est caractérisée par la propriété universelle suivante : pour tout espace topologique Z , une application $g : Y \rightarrow Z$ est continue si et seulement si $g \circ f$ est continue.
3. On munit $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ de la topologie quotient associée à la projection canonique

$p : \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$, et $\mathbb{U} = \{z \in \mathbb{C}; |z| = 1\}$ de la topologie initiale de l'inclusion $j : \mathbb{U} \hookrightarrow \mathbb{C}$. Montrer que \mathbb{T} et \mathbb{U} sont homéomorphes.

A.f Exercice (Topologie de l'ordre).

Soit (X, \leq) un ensemble non vide totalement ordonné.² On note $<$ la relation d'ordre strict associée à \leq ($a < b$ si et seulement si $a \leq b$ et $a \neq b$). Un *intervalle ouvert* est une partie de l'une des quatre formes suivantes :

$$]a, b[:= \{x, a < x < b\}, \quad]\alpha, a[:= \{x, x < a\}, \quad]a, \omega[:= \{x, a < x\}, \quad]\alpha, \omega[:= X,$$

où $a, b \in X$ (et où α et ω ne sont que deux symboles qui n'ont aucun sens hérité d'ailleurs).

1. Montrer que l'ensemble des parties O de X telles que quel que soit $a \in O$ la partie O contient un intervalle ouvert contenant a est une topologie de X ; c'est la *topologie de l'ordre* de (X, \leq) .

2. Dans le cas de (\mathbb{R}, \leq) , montrer que la topologie de l'ordre coïncide avec la topologie de la métrique euclidienne.

Notons $-\infty$ et $+\infty$ deux éléments quelconques n'appartenant pas à \mathbb{R} , et $X = \mathbb{R} \cup \{-\infty, +\infty\}$. Soit \leq l'unique relation d'ordre sur X qui prolonge celle de \mathbb{R} et telle que

$$-\infty \leq x \leq +\infty \quad (\forall x \in \mathbb{R}).$$

On munit X de la topologie de la relation d'ordre.

3. Donner une base de voisinages de $\pm\infty$.

4. Montrer que \mathbb{R} est dense dans X , c'est-à-dire que X est le plus petit fermé contenant \mathbb{R} . (Pour cette raison, on note souvent $X = \overline{\mathbb{R}}$.)

5. Montrer que la fonction $\varphi : \mathbb{R} \rightarrow]0, 1[$ définie par

$$\varphi(x) = \frac{x}{1 + |x|}$$

se prolonge par continuité en une application $\overline{\mathbb{R}} \rightarrow [0, 1]$ et que ce prolongement est un homéomorphisme, c'est-à-dire une bijection bicontinue.

6. En déduire une distance d sur X dont la topologie métrique associée soit la topologie de l'ordre de X .

(Le lecteur rigoureux remarquera que $]a, \omega[= [-\infty, +\infty]$, ce qui montre que l'on ne peut généralement pas identifier α et ω aux plus petit et plus grand éléments de X , même à supposer que ceux-ci existent !)

7. Montrer que la tribu borélienne de X (c'est-à-dire engendrée par les ouverts de X) est la "tribu borélienne" de $\overline{\mathbb{R}}$, c'est-à-dire la tribu engendrée par les ouverts de $\overline{\mathbb{R}}$.

2. On rappelle qu'une *relation d'ordre* est une relation binaire \leq

- réflexive : $x \leq x$
- antisymétrique : $x \leq y$ et $y \leq x \Rightarrow x = y$
- transitive : $x \leq y \leq z \Rightarrow x \leq z$.

A.7 Définition. Si A est une partie de X , l'*adhérence* de A est le plus petit fermé \bar{A} contenant A ; l'*intérieur* de A est le plus gros ouvert $\overset{\circ}{A}$ contenu dans X ; le *bord* (topologique) de A est $\partial A = \bar{A} \setminus \overset{\circ}{A}$.

A.g Exercice (Connexité).

Un espace topologique X est *connexe* s'il n'est pas l'union disjointe de deux fermés non vides.

1. Montrer que l'union de parties connexes dont l'intersection est non vide est connexe.
2. Montrer que les parties connexes de \mathbb{R} sont les intervalles
3. Montrer que l'image d'un connexe par une application continue est connexe.

La *composante connexe* d'un point $x \in X$ est la plus grande partie connexe contenant x .

4. Montrer que les composantes connexes forment une partition de X .
5. Montrer que, si entre deux points quelconques de X il existe un chemin continu (on dit que X est *connexe par arcs*), X est connexe.
6. Montrer que le groupe orthogonal O_n possède deux composantes connexes. On note SO_n la composante connexe de l'identité (et O_n^- l'autre composante, qui n'est pas un groupe).

Annexe B

Compacité

Les affirmations suivantes sont fausses pour un ensemble X quelconque, mais vraies si X est fini :

- Une fonction sur X est bornée
- Une fonction sur X atteint son maximum
- De toute suite de X on peut extraire une suite constante
- De tout recouvrement de X (une famille (U_i) de parties de X dont l'union soit X entier) on peut extraire un recouvrement fini.

Elles sont fausses par exemple si $X =]0, 1[$. Mais elles deviennent vraies pour $X = [0, 1]$ si on modifie les hypothèses et les conclusions en ajoutant un ingrédient topologique adéquat :

- Une fonction *continue* sur X est bornée
- Une fonction *continue* sur X atteint son maximum
- De toute suite de X on peut extraire une suite *convergente* (théorème de Bolzano-Weierstrass)
- De tout recouvrement *ouvert* de X (une famille (U_i) de parties ouvertes de X dont l'union soit X entier) on peut extraire un recouvrement fini (théorème de Heine-Borel).

Ces propriétés restent vraies si X est une partie fermée bornée de \mathbb{R}^n . Par ailleurs, elles sont étroitement corrélées. Par exemple, si l'on sait que toute suite possède une sous-suite convergente, alors on peut en déduire que toute fonction continue possède atteint son maximum (prendre une suite (x_n) maximisante, i.e. telle que $f(x_n)$ converge vers la borne supérieure de f). En gros, X est compact si ces propriétés sont toutes vérifiées. En réalité, ces propriétés ne sont pas exactement équivalentes, et il est commode de ne retenir que la dernière propriété.

Dorénavant, on suppose que les espaces topologiques sont séparés. (Les espaces non séparés, par exemple ceux munis de la topologie grossière, ont souvent peu d'intérêt.)

B.1 Définition. Un espace topologique séparé X est *compact* s'il vérifie la *propriété de Borel-Lebesgue* : de tout recouvrement ouvert, on peut extraire un recouvrement fini.

B.2 Proposition. *Si X est compact, toute suite décroissante de parties fermées*

non vides est non vide. En particulier, toute suite de points possède un point d'accumulation et, si ce point d'accumulation est unique, la suite converge.

Démonstration. La première affirmation s'obtient par passage au complémentaire. La seconde s'obtient en remarquant que l'ensemble des points d'accumulation d'une suite (x_n) est $\bigcap_n \bar{A}_n$, où $A_n = \{x_m, m \geq n\}$. \square

Par exemple, \mathbb{R} n'est pas compact, puisque la suite des intervalles $[n, +\infty[$ contredirait la proposition.

B.a Exercice. Montrer que tout intervalle fermé borné de \mathbb{R} est compact.

B.3 Théorème (admis). 1) *L'image d'un compact par une application continue est compacte.*

2) *Le produit d'espaces compacts est compact (théorème de Tychonov).*

3) *Les boules fermées d'un espace vectoriel normé E sont compactes si et seulement si E est de dimension finie (théorème de Riesz).*

La compacité implique d'importantes propriétés d'existence d'extrema de fonctions. Deux exercices sont donnés à ce sujet dans le chapitre 12.

Beaucoup d'espaces ne sont bien sûr pas compacts. On peut alors récupérer de la compacité en ajoutant un ou plusieurs points à l'espace, qui jouent le rôle de l'infini.

B.4 Définition. Une *compactification* de X est une application $j : X \rightarrow \bar{X}$ injective et d'image dense dans un espace compact \bar{X} ; \bar{X} est le *compactifié* de X .

Par exemple, on peut compactifier \mathbb{R} en ajoutant un point à chaque extrémité de la droite, pour obtenir la *droite réelle étendue* $[-\infty, +\infty]$, que l'on peut munir de la topologie de l'ordre (exercice A.f). On généralise ainsi la notion de limite d'une suite réelle, qui peut désormais être $\pm\infty$ en plus d'un nombre réel, rendant convergentes certaines suites qui étaient divergentes dans \mathbb{R} . Remarquons toutefois que le compactifié obtenu n'hérite pas de toute la structure de \mathbb{R} . Notamment, $[-\infty, +\infty]$ ne possède pas de structure d'espace vectoriel qui prolonge celle de \mathbb{R} : il n'existe pas de façon cohérente de définir la somme $-\infty + \infty$. On peut aussi compactifier \mathbb{R} en ne lui ajoutant qu'un seul point ; c'est la compactification minimale $\mathbb{R} \cup \{\infty\}$. À l'opposé, la compactification maximale d'un espace X s'appelle la *compactification de Stone-Cech* βX (par exemple $\beta\mathbb{N}$ est l'espace des *ultrafiltres*). On peut utiliser différentes compactifications pour distinguer plusieurs types de divergence de suites.

B.5 Exemple (Quelques compactifications de \mathbb{R}^n). 1. Adjonction d'un point à l'infini (par *projection stéréographique*) : l'injection

$$j : \mathbb{R}^n \rightarrow \mathbb{S}^n$$

$$x \mapsto y = (y_0, \hat{y}) = \left(\frac{\|x\|^2 - 1}{\|x\|^2 + 1}, \frac{2x}{\|x\|^2 + 1} \right)$$

envoie \mathbb{R}^n sur la sphère moins son pôle nord $(1, 0, \dots, 0)$.

2. Adjonction d'une sphère à l'infini : l'injection

$$j : \mathbb{R}^n \rightarrow \bar{\mathbb{B}}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n, \|x\| \leq 1\}, \quad x \mapsto \frac{x}{1 + \|x\|}$$

envoie \mathbb{R}^n sur la boule ouverte.

3. Adjonction d'un projectif à l'infini : l'injection

$$j : \mathbb{R}^n \rightarrow \mathbb{R}\mathbb{P}^n, \quad (x_1, \dots, x_n) \mapsto [1, x_1, \dots, x_n]$$

envoie \mathbb{R}^n sur le complémentaire du sous-espace homéomorphe à $\mathbb{R}\mathbb{P}^{n-1}$ des points de la forme $[0, x_1, \dots, x_n]$.

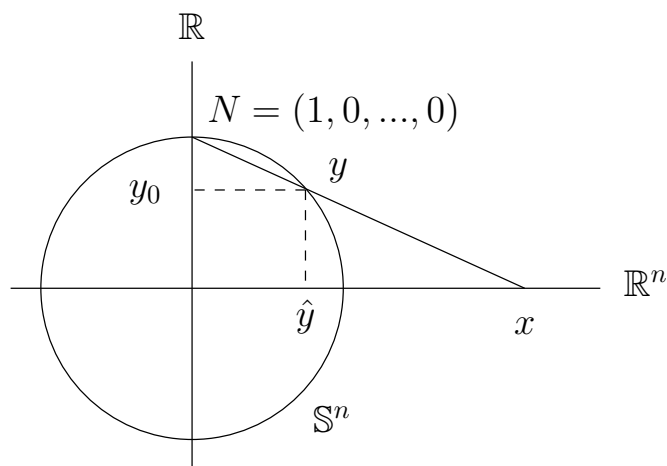


FIGURE B.1 – La projection stéréographique $j : \mathbb{R}^n \rightarrow S^n, x \mapsto y = (y_0, \hat{y})$

Un autre usage des compactifications est de voir certains objets mathématiques comme des objets limites d'autres objets. Par exemple, la mesure de Dirac, qui n'est pas une fonction, existe dans certains espaces de compactification (locale) d'espaces de fonctions, tels que les espaces de mesures ou de distributions. De même, la suite des groupes cycliques $\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/3\mathbb{Z}, \dots$, peut être compactifiée en lui ajoutant le tore $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ à l'infini. Pour un panorama des usages modernes de la compacité, voir le livre de T. TAO [Tao13].

Annexe C

Espaces métriques complets

La notion d'*espace métrique* fut introduit par M. Fréchet en 1906. Elle généralise l'espace euclidien \mathbb{R}^n et la notion d'espace de Hilbert. La *complétude* est une propriété magique de nombreux espaces, permettant de montrer que certaines suites convergent sans connaître leur limite *a priori*, et donc de montrer l'existence de solutions d'équations dont on ne connaît pas de solution explicite.

C.1 Définition. Une *distance* sur un ensemble X est une fonction $d : X \times X \rightarrow [0, +\infty[$ vérifiant les axiomes de

- symétrie : $d(x, y) = d(y, x)$
- séparation : $d(x, y) = 0 \Rightarrow x = y$
- inégalité triangulaire : $d(x, z) \leq d(x, y) + d(y, z)$.

Le couple (X, d) s'appelle alors un *espace métrique*.

Le premier exemple est \mathbb{R}^n muni de la *distance euclidienne*

$$d(x, y) = ((y_1 - x_1)^2 + \cdots + (y_n - x_n)^2)^{1/2}.$$

Beaucoup de notions classiquement utilisées dans \mathbb{R}^n se généralisent naturellement aux espaces métriques. Par exemple,

– la *boule (ouverte)* de centre $x \in X$ et de rayon $r > 0$ est

$$B(x, r) = \{y \in X \mid d(x, y) < r\}.$$

– une partie $V \subset X$ est un *voisinage* de x si il existe $\epsilon > 0$ tel que

$$B(x, \epsilon) \subset V.$$

C.2 Définition. Une suite (x_n) de X *converge* vers une limite ℓ dans X si $d(\ell, x_n)$ tend vers 0, i.e. pour tout $\epsilon > 0$ il existe N tel que, pour tout $n \geq N$, $d(\ell, x_n) < \epsilon$.

Elle est *de Cauchy*^{1, 2} si $d(x_p, x_q)$ tend vers 0 quand p et q tendent vers l'infini, i.e.

1. Augustin L. CAUCHY, mathématicien français (1789–1857), considéré comme le fondateur de l'analyse moderne. On lui doit en particulier, pour les passages à la limite, le formalisme $\epsilon - \eta$.

2. La notion de suite de Cauchy n'est pas purement topologique puisque par exemple la suite $(1/n)_{n \geq 1}$ est une suite de Cauchy de $]0, +\infty[$, qui est transformée, par l'homéomorphisme $x \mapsto 1/x$, en la suite $(n)_{n \geq 1}$, qui n'est pas de Cauchy. Dans ce chapitre on se place dans le cadre simple et important des espaces métriques.

pour tout $\epsilon > 0$ il existe N tel que, pour tous $p, q \geq N$, $d(x_p, x_q) < \epsilon$.

Si (x_n) converge, elle est de Cauchy, puisque

$$d(x_p, x_q) \leq d(x_p, \ell) + d(\ell, x_q).$$

La réciproque est fautive en général, comme le montre l'exemple de la suite $(1/n)$ dans $]0, +\infty[$.

C.3 Définition. X est *complet* si toute suite de Cauchy converge (dans X).

C.4 Théorème. \mathbb{R}^n est complet.

Démonstration. Une suite réelle (x_k) est de Cauchy si et seulement si ses décimales se stabilisent (c'est-à-dire que, pour tout p , après un certain rang K la p -ième décimale de x_k ne dépend plus de k), ce qui implique le théorème pour $n = 1$. La généralisation à n quelconque est immédiate (les décimales de toutes les composantes d'une suite de Cauchy se stabilisent, donc convergent). \square

C.a Exercice. Soient (X, d) est métrique complet, $\Lambda(X)$ l'espace de ses chemins $\gamma : [0, 1] \rightarrow X$ continus. Notons

$$d_\Lambda(\gamma, \gamma') = \sup_{[0,1]} d(\gamma, \gamma').$$

Montrer que $(\Lambda(X), d_\Lambda)$ est un espace métrique complet ; on pourra utiliser le fait que la limite uniforme d'une suite d'applications continues est continue.

C.b Exercice.

Montrer que, si X est complet, l'intersection d'une suite décroissante (A_n) de fermés non vides dont le diamètre tend vers 0 est un singleton.

C.c Exercice.

Montrer qu'un espace métrique compact est complet.

C.5 Théorème (admis). 1) *Le produit fini d'espaces métriques complets est complet.*

2) *L'image d'un espace métrique complet par une application uniformément continue est complète.*

C.6 Définition. Une *complétion* d'un espace métrique X est une application injective $j : X \rightarrow \bar{X}$ à valeurs dans un espace métrique complet, d'image dense.

C.7 Proposition. *Tout espace métrique possède une complétion, qui est unique à isométrie fixant les points de X près. Si de plus X est vectoriel normé, son complété admet une unique structure banachique. Si de plus X est préhilbertien, son complété est hilbertien.*

C.8 Exemple (espaces de Lebesgue). Si Ω est un espace mesuré et si $p \in [1, +\infty[$, l'espace $L^p(\Omega)$ est le complété de l'espace des fonctions étagées pour la norme L^p .³ C'est une façon rapide de définir $L^p(\Omega)$. La définition habituelle, consistant à construire explicitement $L^p(\Omega)$ comme le quotient (par l'égalité presque partout) de l'espace des fonctions dont le module à la puissance p est intégrable, présente l'immense avantage de réaliser les points ajoutés comme des (classes de) fonctions sur Ω .

C.9 Proposition. *Soient X et Y deux espaces métriques, D une partie dense de X et $f : D \rightarrow Y$ une application uniformément continue. Si Y est complet, il existe un unique prolongement continu de f sur X . En outre, ce prolongement est uniformément continu.*

C.10 Exemple (Transformation de Fourier-Plancherel). L'espace $L^2(\mathbb{R}^n)$ des fonctions de carré sommable sur \mathbb{R}^n est un espace de Hilbert. La transformation de Fourier des fonctions $f \in L^1 \cap L^2$ est une isométrie de L^2 , qui est donc uniformément continue et qui se prolonge en une isométrie de L^2 , la *transformation de Fourier-Plancherel*.

3. Ce résultat est faux pour $L^\infty(\mathbb{R}^n)$.

Glossaire

Fonction Application (définie sur un ouvert de \mathbb{R}^n ou une surface) à valeurs dans \mathbb{R} (ou, parfois, dans un espace vectoriel par opposition à une surface).

Chemin Application définie sur un intervalle de \mathbb{R} , à valeurs dans \mathbb{R}^n ou une surface. On dit aussi *courbe paramétrée*.

Courbe Image d'un chemin (c'est donc un ensemble de points).

Courbe intégrale d'un champ de vecteurs Chemin γ dont la vitesse $\gamma'(t)$ égale, à chaque instant t , le champ de vecteurs au point $\gamma(t)$.

Champ de vecteurs Application qui, à tout point, associe un vecteur (tangent en ce point). Sur un ouvert U de \mathbb{R}^n , c'est une application $v : U \rightarrow \mathbb{R}^n$. Si v est de classe C^1 (Lipschitz suffit), le théorème de Cauchy-Lipschitz affirme que, pour tout point $u \in U$, il existe un unique chemin $\gamma_u : (\mathbb{R}, 0) \rightarrow U$ défini sur un intervalle maximal $] -t_0(u), t_1(u)[$ et tel que

$$\gamma_u(0) = u \quad \text{et} \quad \gamma'_u(t) = v(\gamma_u(t)) \quad (\forall t).$$

Les fonctions $t_0, t_1 : U \rightarrow]0, +\infty]$ sont semi-continues inférieurement (c'est-à-dire que, par petite perturbation de u , $t_i(u)$ ne peut subir de discontinuité que vers le haut). Le *flot* de v est alors l'application

$$\phi : \mathcal{D}_v = \{(t, u) \in \mathbb{R} \times U, -t_0(u) \leq t < t_1(u)\} \rightarrow U, \quad (t, u) \mapsto \phi_t(u) = \gamma_u(t).$$

Réciproquement, le champ de vecteur

$$v = \frac{\partial \phi_t}{\partial t} \circ \phi_t^{-1}$$

est le *générateur infinitésimal* du flot ϕ .

Espace de Banach Espace vectoriel normé complet. Par exemple, l'espace de Lebesgue $L^1(\mathbb{R})$, muni de la norme $\|f\| = \int_{\mathbb{R}} |f(x)| dx$, est de Banach.

Espace de Hilbert Espace vectoriel E sur $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} muni d'un *produit scalaire hilbertien*, c'est-à-dire d'une forme $g : E^2 \rightarrow \mathbb{K}$

- sesquilinéaire⁴ : g est linéaire par rapport au premier argument, et *semi-linéaire* par rapport au second argument :

$$g(x, \lambda y + \lambda' y') = \bar{\lambda} g(x, y) + \bar{\lambda}' g(x, y') \quad (x, y, y' \in E, \lambda \in \mathbb{K})$$

- hermitienne : $g(x, y) = \overline{g(y, x)}$ ($x, y \in E$)
- positive : $g(x, x) \in \mathbb{R}_+$ ($x \in E$)
- définie / non dégénérée : $g(x, x) = 0 \Rightarrow x = 0$
- complète : les suites qui sont de Cauchy pour la norme induite $\|x\| = \sqrt{g(x, x)}$ convergent dans E .

Par exemple, l'espace de Lebesgue $L^2_{\mathbb{C}}(\mathbb{R})$, muni du produit scalaire

$$(f|g) = \int_{\mathbb{R}} f(x) \overline{g(x)} dx,$$

est de Hilbert. Les espaces hilbertiens sont les espaces de dimension éventuellement infinie dont les propriétés les rapprochent le plus des espaces euclidiens.

Norme d'opérateur subordonnée Si $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont deux espaces vectoriels normés, par exemple \mathbb{R}^n muni de la norme euclidienne, la norme d'opérateur subordonnée est la norme définie sur $L(E, F)$ par

$$\|a\|_{E,F} = \sup_{x \in E \setminus \{0\}} \frac{\|a(x)\|_F}{\|x\|_E}.$$

D'après la linéarité de a et l'homogénéité des normes,

$$\|a\|_{E,F} = \sup_{x \in E, \|x\|_E=1} \|a(x)\|_F.$$

Si E est de dimension finie, sa sphère unité est compacte (lemme de Riesz), donc la borne supérieure est atteinte :

$$\|a\|_{E,F} = \max_{x \in E, \|x\|_E=1} \|a(x)\|_F.$$

Dans tous les cas, les normes d'opérateurs sont sous-multiplicatives : si $(G, \|\cdot\|_G)$ est un troisième espace vectoriel normé et si $b \in L(F, G)$,

$$\|b \circ a\|_{E,G} \leq \|b\|_{F,G} \|a\|_{E,F}.$$

Théorème des accroissements finis Soient $\gamma : [a, b] \rightarrow \mathbb{R}$ continue et dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que

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

4. En latin, *sesqui* veut dire "un et demi". Par exemple, l'anniversaire sesquicentenaire est le cent cinquantième.

(Démonstration : appliquer le théorème de Rolle à la fonction $t \mapsto \gamma(t) - \frac{\gamma(b) - \gamma(a)}{b-a}(t-a)$.) En particulier,

$$\left| \frac{\gamma(b) - \gamma(a)}{b-a} \right| \leq \sup_{]a,b[} |\gamma'|.$$

Alphabet grec

L'usage conjoint des alphabets latin et grec offre plus de flexibilité dans les notations. Ne pas les connaître peut être sources de difficultés, que le concepteur d'un sujet d'examen ne soupçonne même pas !

Lettre capitale	Lettre minuscule	Nom en français	Usage dans ce cours
A	α	alpha	*
B	β	beta	*
Γ	γ	gamma	*
Δ	δ	delta	*
E	ϵ, ε	epsilon	*
Z	ζ	zêta	*
H	η	êta	*
Θ	θ, ϑ	thêta	*
I	ι	iota	
K	κ, \varkappa	kappa	*
Λ	λ	lambda	
M	μ	mu	*
N	ν	nu	*
Ξ	ξ	xi	*
O	o	omicron	
Π	π, ϖ	pi	*
P	ρ, ϱ	rhô	*
Σ	σ, ς	sigma	*
T	τ	tau	*
Y	υ	upsilon	
Φ	ϕ, φ	phi	*
X	χ	khi	
Ψ	ψ	psi	*
Ω	ω	oméga	*

D'autres alphabets sont utilisés en mathématique. Par exemple, depuis Georg CANTOR on note \aleph (aleph, de l'alphabet hébreu) les cardinaux des ensembles infinis bien ordonnés. Des symboles sont parfois créés à partir d'autres, comme pour la constante de PLANCK $\hbar = \frac{h}{2\pi}$, ou le signe de l'intégrale \int , allongement italique, inventé par LEIBNIZ, du caractère typographique ancien « s long » (comme abbréviation de « summa », qui veut dire « somme » en latin), lui-même disparu depuis l'uniformisation de l'écriture par CHARLEMAGNE.

Bibliographie

- [Arn84] V. I. Arnold. *Chapitres supplémentaires de la théorie des équations différentielles ordinaires*. Éditions Mir, Moscou, 1984.
- [Arn88] V. I. Arnold. *Équations différentielles ordinaires*. Éditions Mir, Moscou, 1988.
- [Arn89] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.
- [AVZK86] V.I. Arnold, A.N. Varchenko, S.G. Zadé, and V. Kotliar. *Singularités des applications différentiables*. Number ptie. 1 in TR. : Mathématiques. Mir, 1986.
- [Bre99] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maî trise. 2e édition chez Dunod, 1999. Théorie et applications.
- [Car61] H. Cartan. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Avec le concours de Reiji Takahashi. Hermann, Paris, 1961.
- [Cha03] M. Chaperon. *Calcul différentiel et intégral de troisième année*. Dunod, 2003.
- [Cho64] G. Choquet. *Cours de topologie : espaces topologiques et espaces métriques, fonctions numériques, espaces vectoriels topologiques*. Masson (Dunod), 1964.
- [Dem06] J.-P. Demailly. *Analyse numérique et équations différentielles*. EDP Sciences, 2006.
- [Féj14] J. Féjóz. Chapitres d'intégration et de probabilités. Cours polycopié de l'Université Paris-Dauphine, 2014. <https://www.ceremade.dauphine.fr/~fejoz/Integration/integration-probabilites.pdf>.
- [Ham82] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1) :65–222, 1982.
- [Hir94] M. W. Hirsch. *Differential topology*, volume 33 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994. Corrected reprint of the 1976 original.
- [HL99] F. Hirsch and G. Lacombe. *Éléments d'analyse fonctionnelle*. Dunod, 1999.
- [Hör07] L. Hörmander. *Notions of convexity*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2007. Reprint of the 1994 edition.

- [Hou61] C. Houzel. Géométrie analytique locale i. *Sém. Henri Cartan*, 18, 1960/61.
- [KF57] A. N. Kolmogorov and S. V. Fomin. *Elements of the theory of functions and functional analysis. Vol. 1. Metric and normed spaces*. Graylock Press, Rochester, N. Y., 1957.
- [Lau11] F. Laudenbach. *Calcul différentiel et intégral*. Éditions de l'École Polytechnique, 2011.
- [Mil63] J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.
- [Mil97] J. W. Milnor. *Topology from the differentiable viewpoint*. Princeton University Press, Princeton, NJ, 1997.
- [Pea88] G. Peano. Intégration par séries des équations différentielles linéaires. *Math. Ann.*, 32 :450–456, 1888.
- [Pha92] F. Pham. *Géométrie et calcul différentiel sur les variétés*. Interéditions, 1992.
- [Pic90] É. Picard. Mémoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives. *J. Math.*, 6(4) :145–210, 1890.
- [Rou03] F. Rouvière. *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*. Cassini, 2003.
- [Rud74] W. Rudin. *Real and complex analysis*. McGraw-Hill, 1974.
- [Tao13] T. Tao. *Compactness and Contradiction*. American Mathematical Society, 2013.
- [Tao16] T. Tao. Compactness and compactification. Preprint, 2016.

Index

- Adhérence, 195
- Algorithme
 - de Newton, 61
 - de relaxation, 177
- Algèbre locale d’une application, 80
- Application
 - bilinéaire, 30
 - conforme, 33
 - contractante, 47
 - de Viète, 81
 - lipschitzienne, 47
 - lue, 113
 - équivalente, 65, 69
- BINET, J. P. M., 119
- Bord, 141, 195
- Boule, 201
- BROUWER, L., 56
- Calcul des variations, 185
- CANTOR, G., 56
- Cardioïde, 14
- CARTAN, É., 125
- CAUCHY, A., 55, 201
- Centre de courbure, 15
- Centre de masse, 169
- Champ de vecteurs, 46, 205
- Chemin, 9, 205
 - rectifiable, 18
- Classe C^k , 113, 145
- Codimension, 83
- Compactification, 198
- Complétion, 202
- Complétude, 202
- Composante connexe, 195
- Connexité, 195
- Continuité, 192
 - en un point, 192
- Contour apparent, 74, 89
- Contrôle, 66
- Convergence
 - géométrique, 48, 50
 - normale, 57
 - quadratique, 50, 61
- Convexe, 167
- Coordonnées sphériques, 111
- Courbe, 9, 83, 205
 - cissoïde, 17
 - intégrale, 183, 205
- Courbe discriminante, 74
- Courbe paramétrée, 9
- Courbure, 15
- Critère de contrôlabilité de Kalman, 66
- \mathbb{C} -dérivable, 33
- Dérivée, 21, 88
- Déterminant, 117
 - jacobien, 46
 - wronskien, 46
- Difféomorphisme, 55
 - de redressement, 83
- Différentielle, 21
- Dilatation, 43
- Dimension, 83
- DINI, U., 55
- Distance, 201
- Divergence, 46, 143
- Droite
 - affine tangente, 13
 - tangente, 12, 13
- Dual topologique, 182
- Décomposition de Dunford, 66
- Dérivée
 - extérieure, 142
- Développante, 15
- Développée, 15
- Éclatement, 148
- Énergie, 178
- Entropie, 178

- Enveloppe supérieure, 94
- Épigraphe, 173
- Équation
 - d'Euler-Lagrange, 188
 - d'une surface, 84
 - de Cauchy-Riemann, 33
 - de Kepler, 57
 - de propagation, 27
 - de transport, 27
 - eikonale, 158
- Espace
 - complet, 202
 - de Banach, 205
 - de Hilbert, 206
 - de Lebesgue, 203
 - métrique, 201
 - séparé, 191
 - tangent, 86
 - topologique, 193
- État statistique, 178
- Extremum, 101
- Facteur direct, 183
- Facteur intégrant, 15
- Flot, 46, 205
- Fonction, 205
 - α -convexe, 177
 - convexe, 173
 - de Morse, 148
 - fortement convexe, 177
 - homogène, 36
 - plate, 13
- Forme
 - bilinéaire symétrique, 151
 - extérieure, 125
 - quadratique, 89, 151
 - volume, 131
- Forme normale, 65, 69
 - de Frobenius, 68
 - de Jordan, 66
- Formule
 - de Binet-Cauchy, 119
 - de Cauchy, 144
 - de Héron, 105
 - de Liouville, 46
 - de Stirling, 179
 - de Stokes, 142
 - de Taylor, 11, 146
 - de la moyenne, 39
- Formules
 - du calcul différentiel, 11
- FRÉCHET, M., 34, 201
- GATEAUX, R. E., 34
- Gradient, 32, 143
- GRAM, J. P., 119
- Graphe, 9, 84, 112
- Groupe
 - linéaire, 85
 - orthogonal, 85, 152
- Groupes classiques, 85
- Générateur infinitésimal, 183, 205
- Géodésique, 94, 189
- Hessienne, 145
- Homéomorphisme, 55, 194
- Hypersurface, 83
- Identité d'Euler, 36
- Image
 - directe, 135
 - réciproque, 136
- Immersion, 77
- Indice, 152
- Inégalité
 - arithmético..., 179
 - de Hermite-Hadamard, 178
 - de Hölder, 109
 - de Jensen, 169, 175
 - de Minkowski, 107
 - des accroissements finis, 40, 48
 - du trapèze, 180
- Intervalle
 - ouvert, 194
- Intégrale de Clairaut, 189
- Intérieur, 195
- JORDAN, C., 66
- LAGRANGE, J.-L., 55
- LEIBNIZ, G. W., 3
- Lemme
 - de Morse, 155
 - des noyaux, 66
- Linéarisabilité, 75

- LIPSCHITZ, R., 47
 Lipéomorphisme, 55
 LISSAJOUS, Jules A., 10
 Longueur, 15, 18, 40

 MATHER, John N., 163
 Matrice de Gram, 119
 Maximum, 101
 Mesure de Gibbs, 178
 Minimum, 91, 101
 MORSE, H. C. M., 164
 Multiplicateur de Lagrange, 103
 Multiplicité d'une application, 80
 Méthode des caractéristiques, 27, 28
 Méthode des séries majorantes, 79
 Métrique riemannienne, 32, 186

 NEWTON, I., 3
 Nombre d'or, 50
 Norme d'opérateur subordonnée, 206

 Opérateur de dérivation, 128
 Orientabilité, 131
 Orientation, 123
 Orthogonalité, 151
 Ouvert, 192, 193

 Paramétrage, 9
 PEANO, G., 48
 Permutation, 43
 PICARD, É., 48
 Pivot de Gauss, 43
 Plongement, 193
 Point
 - col, 92
 - critique, 77, 155
 - cusp, 13
 - d'inflexion, 13
 - de rebroussement, 13
 - régulier, 13, 77
 Point atteignable, 66
 Point critique, 101
 Point fixe
 - attractif, 50
 - répulsif, 50
 Point régulier, 12
 Polygone de Newton, 80
 Produit
 - extérieur, 126
 - intérieur, 132
 - scalaire euclidien, 154
 - scalaire lorentzien, 154
 - vectoriel, 127
 Propriété universelle
 - de la topologie finale, 193
 - de la topologie initiale, 193
 - du quotient, 98
 Quotient d'espace vectoriel, 97
 Quotient topologique, 193

 Rang, 65
 - d'une application, 77
 - maximal, 77
 Rapport de Lipschitz, 47
 Rayon de courbure, 15
 Redressement, 83
 Revêtement, 140
 Rotationnel, 143
 Rétraction, 58

 \mathbb{S}^n , 85
 Section, 76
 Segment, 167
 Série
 - formelle, 79
 - majorante, 79
 Signature, 118, 152
 Sous-variété, 83
 Sphère, 85
 STOKES, G. G., 141
 Strictement convexe, 167
 Structure de voisinage, 191
 Subimmersion, 78
 Submersion, 75
 Suite
 - minimisante, 91
 Suite de Cauchy, 201
 Surface, 83
 - à coins, 141
 Surface stable, 63
 Système de coordonnées, 111

 \mathbb{T} , 193
 Tangence, 147
 Température, 178

- Théorème
- d’immersion, 77
 - d’inversion globale, 56
 - d’inversion locale, 55
 - d’inversion locale lipschitzienne, 58
 - de Bolzano-Weierstrass, 197
 - de Carathéodory, 169
 - de Cauchy-Lipschitz, 51, 183
 - de Gauss-Lucas, 171
 - de Heine-Borel, 197
 - de Minkowski, 168
 - de Perron, 63
 - de Riesz, 198
 - de Rouché, 79
 - de Tychonov, 198
 - de division, 80
 - de l’invariance du domaine, 56
 - de préparation, 80
 - de subimmersion, 78
 - de submersion, 75
 - des accroissements finis, 206
 - des extrema liés, 103
 - des fonctions implicites, 43, 71
 - des jets suffisants, 161
 - du point fixe, 48
 - du rang, 65
 - du rang constant, 78
 - fondamental de l’algèbre, 79
- Théorie de Morse, 164
- THOM, René, 163
- Topologie, 193
- de l’ordre, 194
 - discrète, 193
 - faible, 193
 - finale, 193
 - grossière, 193
 - induite, 193
 - initiale, 193
 - quotient, 193
- Tore, 85
- Transformation
- de Fourier-Plancherel, 203
- Transformée de graphe, 63
- Transvection, 43
- Transversalité, 89
- Tribu borélienne, 194
- Variation totale, 18
- Variété stable, 63, 148
- Vecteur
- rotation, 30
- Voisinage, 191, 201
- Wronskien, 46