

Licence de Mathématiques 3

M62 : Fonctions réelles de plusieurs variables

Laurent Guillopé

www.math.sciences.univ-nantes.fr/~guillope/M62/

Ces notes de cours ont été utilisées dans une UE de Licence de l'année 2007-2008 sous la responsabilité de F. Wagemann, remercié chaleureusement pour sa lecture critique qui a suscité la révision 2008 de ces notes.

[9 juin 2008]

TABLE DES MATIÈRES

1. Applications différentiables	1
1.1. Différentiabilité.....	2
1.2. Fonctions composées.....	5
1.3. Fonctions de classe C^k	6
1.4. Dérivation et intégration.....	7
1.5. Accroissements finis.....	8
2. Le théorème d'inversion locale	11
2.1. L'énoncé du théorème.....	12
2.2. Le théorème du point fixe.....	13
2.3. La preuve du théorème d'inversion locale.....	14
2.4. Changement de variables.....	15
3. Le théorème des fonctions implicites	17
3.1. Le théorème et sa preuve.....	17
3.2. Racines d'équations polynomiales.....	18
4. Courbes et surfaces régulières	20
4.1. Équations régulières.....	20
4.2. Courbes régulières.....	20
4.3. Surfaces régulières.....	22
4.4. Gradient.....	23
5. Applications C^k	23
5.1. La version C^k des théorèmes classiques.....	25
5.2. Théorème de Schwarz.....	26
5.3. Formules de Taylor.....	27
5.4. Approximation de l'espace tangent.....	28
6. Points critiques et extrema	29
6.1. Extrema.....	29
6.2. Extrema liés.....	31
6.3. Appendice : rappels sur les formes quadratiques.....	33
7. Compléments	35
7.1. Fonctions convexes.....	35
7.2. La transformée de Legendre.....	36
7.3. Le lemme de Morse.....	37

CHAPITRE 1

APPLICATIONS DIFFÉRENTIABLES

Le calcul différentiel étudie des fonctions de plusieurs variables x_1, \dots, x_n à valeurs vectorielles, *i. e.* des fonctions définies sur une partie U (en général ouverte) de \mathbb{R}^n à valeurs dans \mathbb{R}^p . Pour quantifier certaines définitions (limite, continuité, différentiabilité), l'usage de normes sur \mathbb{R}^k est chose courante. Vu que deux normes N_1, N_2 quelconques sur \mathbb{R}^k sont équivalentes, *i. e.* il existe des constantes C, C' telles que

$$CN_1(x) \leq N_2(x) \leq C'N_1(x), \quad x \in \mathbb{R}^k,$$

le choix de normes n'influera en général pas sur les diverses définitions. Ainsi celle de voisinage : la partie V de \mathbb{R}^k est un voisinage de $m \in \mathbb{R}^k$ s'il existe $r > 0$ tel que la la boule $B(m, r) = \{\|x - m\| < r\}$ de centre m et de rayon r est incluse dans V . On vérifie bien sur cet exemple que peu importe la norme choisie, notamment parmi les normes classiques

$$\|x\|_1(x) = \sum_{i=1}^n |x_i|, \quad \|x\|_2(x) = \sqrt{\sum_{i=1}^n |x_i|^2}, \quad \|x\|_\infty(x) = \sup_{i=1}^n |x_i|, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

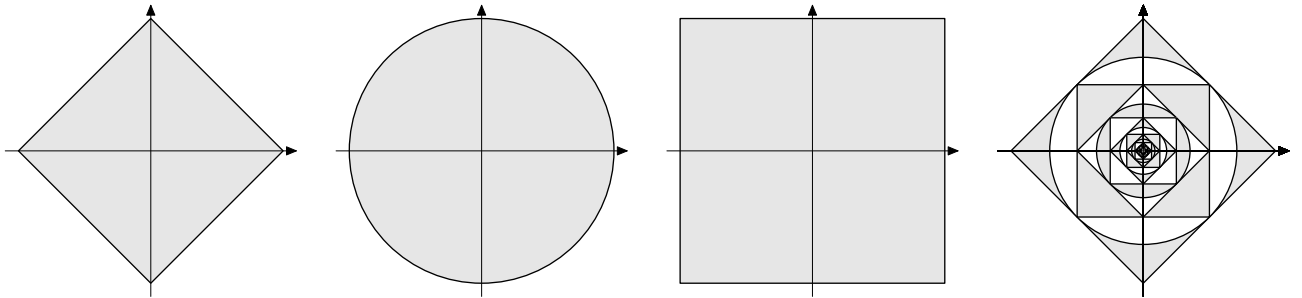


FIGURE 1. Dans \mathbb{R}^2 , les boules $B(1) = \{N(x, y) \leq 1\}$ pour les normes $\| \cdot \|_1, \| \cdot \|_2$ et $\| \cdot \|_\infty$, avec les inclusions de boules de types différents et rayons adéquats.

En général, on prendra comme norme sur des espaces de matrices $M_{n,m}(\mathbb{R})$ ou d'applications linéaires $L(E, F)$ des normes d'opérateurs : $\|A\| = \sup_{x \in \mathbb{R}^m \setminus \{0\}} \|Ax\|_n / \|x\|_m$ pour une matrice $A \in M_{n,m}(\mathbb{R})$ avec des normes $\| \cdot \|_m, \| \cdot \|_n$ sur \mathbb{R}^m et \mathbb{R}^n ou $\|\ell\| = \sup_{x \in \mathbb{E} \setminus \{0\}} \|\ell x\|_F / \|x\|_E$ pour un opérateur linéaire $\ell \in L(E, F)$ avec des normes $\| \cdot \|_E, \| \cdot \|_F$ sur E et F .

On note par $o(h)$ toute fonction q définie pour h dans un voisinage de 0 dans \mathbb{R}^n et à valeurs dans \mathbb{R}^p , telle que

$$\lim_{h \rightarrow 0} q(h) / \|h\| = 0.$$

Si h est dans \mathbb{R} , il suffit de considérer la limite de $q(h)/h$ lorsque $h \rightarrow 0$. De manière analogue on notera $\varepsilon(h)$ toute quantité telle que $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. Ainsi, on remplacera parfois un $o(h)$ par $\|h\|\varepsilon(h)$.

▷ **Exemple 1.1.** — $h\sqrt{|h|} = o(h)$, $\sin(h) = \varepsilon(h)$ pour $h \in \mathbb{R}$ a voisinage de 0, $HAH = o(H)$ pour $A, H \in M_n(\mathbb{R})$ et H au voisinage de 0. ◁

L'algèbre des $o(h)$ obéit aux « règles » suivantes

$$(1) \quad o(h) + o(h) = o(h), \quad b(h)o(h) = o(h), \quad o(o(h)) = o(h)$$

où on a noté b une fonction bornée au voisinage de $h = 0$. En effet, si q_1, q_2 sont des $o(h)$, on a

$$\begin{aligned} \frac{\|q_1(h) + q_2(h)\|}{\|h\|} &\leq \frac{\|q_1(h)\|}{\|h\|} + \frac{\|q_2(h)\|}{\|h\|}, \\ \frac{\|b(h)q_1(h)\|}{\|h\|} &\leq \sup \|b(h)\| \frac{\|q_1(h)\|}{\|h\|}, \\ \frac{\|q_1(q_2(h))\|}{\|h\|} &= \frac{\| \|q_2(h)\| \varepsilon_1(q_2(h)) \|}{\|h\|} = \frac{\|q_2(h)\|}{\|h\|} \|\varepsilon_1(q_2(h))\|. \end{aligned}$$

1.1. Différentiabilité

Une fonction d'une variable réelle $f : (x_0 - r, x_0 + r) \rightarrow \mathbb{R}$ est dite dérivable en x_0 , avec pour dérivée le nombre réel $f'(x_0)$ si

$$f(x_0 + h) = f(x_0) + f'(x_0)h + o(h).$$

La dérivabilité est équivalente à l'approximation locale par une fonction affine, autrement dit à l'existence d'un développement limité à l'ordre 1.

△ L'existence d'un développement limité à l'ordre 2

$$f(x_0 + h) = f(x_0) + Ah + Bh^2 + o(h^2)$$

au voisinage d'un point $x_0 = 0$ ne dit pas nécessairement l'existence d'une dérivée seconde pour f , comme la fonction $x \rightarrow x^3 \sin(1/x) = o(x^2)$ en $x_0 = 0$ le confirme. ◁

Définition 1.1. — Soit $f : B(m_0, r) \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Pour h vecteur non nul de \mathbb{R}^n , la *dérivée directionnelle* $D_h f(m_0)$ est, si elle existe, la dérivée en $t = 0$ de la fonction définie sur $(-r/\|h\|, r/\|h\|)$ qui à t associe $f(m_0 + th)$.

△ Pour une fonction de deux variables (on généralisera aisément pour f de n variables, avec $n \geq 3$) $f : m = (x, y) \in B(m_0, r) \subset \mathbb{R}^2 \rightarrow f(m) = f(x, y) \in \mathbb{R}$, les dérivées partielles $\frac{\partial f}{\partial x}(m_0)$ et $\frac{\partial f}{\partial y}(m_0)$ sont, si elles existent, les dérivées des fonctions F_1, F_2 définie au voisinage de $t = 0$ par

$$F_1(t) = f(x_0 + t, y_0), \quad F_2(t) = f(x_0, y_0 + t).$$

Les dérivées partielles sont les dérivées directionnelles suivant les axes de coordonnées. On notera aussi les dérivées partielles $\partial_x f(m_0)$, $\partial_y f(m_0)$, $\partial_{x_1} f(m_0)$ (voire $\partial_i f(m_0)$) et pour les dérivées successives (lorsque cela est bien défini) $\partial_x \partial_y f(m_0)$, $\partial_{xy}^2 f$, $\partial_{x_{i_1} \dots x_{i_k}}^k f$, $\partial_{i_1 \dots i_k}^k f$. ◁

On a donc $f(m_0 + th) = f(m_0) + D_h f(m_0)t + o(t)$ et on retrouve $\frac{\partial f}{\partial x}(m_0) = D_{e_1} f(m_0)$: par ailleurs, on a $D_{2e_1} f(m_0) = 2 \frac{\partial f}{\partial x}(m_0)$.

▷ **Exemples 1.2.** —

1. Si $f(x, y) = \sin(x + y)$, on a $D_{(u,v)} f(0, 0) = u + v$.
2. Si f est définie sur \mathbb{R}^2 par $f(x, y) = 2xy(x^2 + y^2)^{-1/2}$ si $(x, y) \neq 0$ et $f(0, 0) = 0$, on a $f(t(u, v)) = |t|f(u, v)$ pour $t \in \mathbb{R}$. Ainsi, la restriction de f à la droite $\mathbb{R}(u, v)$ n'est pas approchée par une expression linéaire au voisinage de $t = 0$: les dérivées directionnelles $D_{(u,v)} f(0, 0)$ n'existent pas. Si on utilise les coordonnées polaires $(x, y) = (r \cos \theta, r \sin \theta)$, on a $f(x, y) = r \sin 2\theta$. ◁

Définition 1.2. — Soit $f : B(m_0, r) \subset \mathbb{R}^n \rightarrow \mathbb{R}$. La fonction f est dite *différentiable* en m_0 s'il existe une application linéaire $L : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que

$$(2) \quad f(m_0 + h) = f(m_0) + L(h) + o(h).$$

L'application linéaire L est appelée *différentielle* de f en m_0 et notée $Df(m_0)$.

▷ **Exemples 1.3.** —

1. Soit $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ différentiable en t_0 . L'application linéaire $Df(t_0) : \mathbb{R} \rightarrow \mathbb{R}$ est habituellement identifiée à sa valeur sur le vecteur générateur $h = 1$ de l'espace unidimensionnel \mathbb{R} , autrement dit la dérivée de la fonction f en t_0 . On a

$$f(t_0 + h) = f(t_0) + Df(t_0)h + o(h).$$

2. Soit $A : \mathbb{R}^n \rightarrow \mathbb{R}$ linéaire et $b \in \mathbb{R}$. Alors l'application $f : x \in \mathbb{R}^n \rightarrow Ax + b \in \mathbb{R}$ est différentiable en tout $x \in \mathbb{R}^n$, avec $Df(x) = A$. En effet,

$$f(x + h) = A(x + h) + b = Ax + b + Ah = f(x) + Ah. \quad \triangleleft$$

Proposition 1.1. — Soit f une fonction différentiable en m_0 .

- La fonction f est continue en m_0 .
- L'application L de la définition 1.2 est unique.
- La fonction f admet en m_0 des dérivées directionnelles suivant toute direction et $D_v f(m_0) = Df(m_0)(v)$, $v \in \mathbb{R}^n$. En particulier, f admet des dérivées partielles.

Démonstration. — Soient L_1 et L_2 deux formes linéaires vérifiant (2). Alors pour, h assez petit,

$$f(x + h) - f(x) = L_1(h) + o(h) = L_2(h) + o(h)$$

ainsi

$$\lim_{\|h\| \rightarrow 0} (L_1 - L_2)(h/\|h\|) = \lim_{\|h\| \rightarrow 0} o(h)/\|h\| = 0,$$

soit la nullité de

$$\|L_1 - L_2\| = \sup_{\|v\|=1} \|(L_1 - L_2)(v)\| = \sup_{\|h\| \leq \alpha} \|(L_1 - L_2)(h/\|h\|)\|.$$

La continuité de f en m_0 est équivalente à $f(m_0 + h) = f(m_0) + \varepsilon(h)$: si f est différentiable, c'est bien le cas, vu que $f(m_0 + h) = f(m_0) + [Df(m_0)(h) + o(h)]$.

Enfin si f est différentiable, pour v fixé non nul, on a au voisinage de $t = 0$,

$$f(m_0 + tv) = f(m_0) + Df(m_0)(tv) + o(tv) = f(m_0) + Df(m_0)(v)t + o(t),$$

ce qui indique la dérivabilité de f dans la direction $v \in \mathbb{R}$, avec $D_v f(m_0) = Df(m_0)(v)$. □

△ La différentielle est notée parfois $df(m_0)$ ou $f'(m_0)$ (voire $\dot{f}(m_0)$), dernière notation qu'on évitera dans un premier temps, pour bien distinguer dérivée (d'une fonction d'une variable réelle) et différentielle. Si f à valeurs numériques est différentiable en $m = (x_0, y_0, z_0) \in \mathbb{R}^3$ (exemple qu'on généralisera au cas d'un nombre quelconque de variables), on écrit souvent sa différentielle df suivant Leibniz $df(m_0) = f'_x(m_0)dx + f'_y(m_0)dy + f'_z(m_0)dz$; il faut y interpréter dx comme la différentielle de la fonction coordonnée x . Si la fonction f dépend de variables groupées suivant (x, y) avec $x = (x_1, \dots, x_m)$ et $y = (y_1, \dots, y_n)$, la notation $D_x f$ signifiera que les variables y_1, \dots, y_n sont considérées comme paramètres, seules les variables x_1, \dots, x_m intervenant pour la différentiabilité. ▽

▷ **Exemple 1.4.** — Il se peut que f admette des dérivées directionnelles sans être différentiable. Ainsi de la fonction f définie sur \mathbb{R}^2 par $f(0, 0) = 0$ et $f(m) = 3y - 4y^3(x^2 + y^2)^{-1}$ sinon : en $m = (0, 0)$, les dérivées directionnelles sont $D_{(\cos \alpha, \sin \alpha)} f(m_0) = \sin 3\alpha = 3 \sin \alpha - 2 \sin^3 \alpha$, non linéaire en $\sin \alpha$. En coordonnées polaires, f prend la forme $f(r \cos \theta, r \sin \theta) = r \sin 3\theta$. ▽

La différentiabilité d'une fonction à valeurs vectorielles se définit en calquant la définition 1.2.

Définition 1.3. — Soit $f : B(m_0, r) \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. La fonction f est dite *différentiable* en m_0 s'il existe une application linéaire (unique!) $L : \mathbb{R}^n \rightarrow \mathbb{R}^p$, notée $Df(m_0)$ telle que

$$f(m_0 + h) = f(m_0) + Df(m_0)(h) + o(h).$$

▷ **Exemple 1.5.** — On peut reprendre les exemples 1.3 avec p entier quelconque. ◁

△ Si $f = (f_1, \dots, f_p)$, la différentiabilité de f en m_0 est équivalente à celle des applications $f_i, i = 1, \dots, p$ (exercice!). ▽

Définition 1.4. — Soit $f = (f_1, \dots, f_p) : B(m_0, r) \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ une application différentiable en m_0 . Sa *matrice jacobienne* est la matrice

$$\text{Jac}(f)(m_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(m_0) & \dots & \frac{\partial f_1}{\partial x_n}(m_0) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(m_0) & \dots & \frac{\partial f_p}{\partial x_n}(m_0) \end{pmatrix}.$$

C'est la représentation matricielle de la différentielle $Df(m_0)$ dans les bases canoniques de \mathbb{R}^n et \mathbb{R}^p .

▷ **Exemple 1.6.** — Les « coordonnées polaires » $(r, \theta) \in \mathbb{R}^{+*} \times \mathbb{R} \rightarrow (r \cos \theta, r \sin \theta) \in \mathbb{R}^2$ ont pour matrice jacobienne

$$\begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}. \quad \triangleleft$$

Proposition 1.2. — L'ensemble $\mathcal{D}(m_0)$ des fonctions numériques f définies sur $B(m_0, r_f)$ ($r_f > 0$ dépendant de f) et différentiables en m_0 est un espace vectoriel, l'application différentielle $f \in \mathcal{D}(m_0) \rightarrow Df(m_0) \in L(\mathbb{R}^n, \mathbb{R})$ étant linéaire, i. e.

$$D(\alpha_1 f_1 + \alpha_2 f_2)(m_0) = \alpha_1 Df_1(m_0) + \alpha_2 Df_2(m_0),$$

avec stabilité par le produit

$$D(f_1 f_2)(m_0) = f_1(m_0) Df_2(m_0) + f_2(m_0) Df_1(m_0).$$

Si $f(m_0) \neq 0$, alors f^{-1} est définie au voisinage de m_0 , différentiable en m_0 et $D(f^{-1})(m_0) = -(f(m_0))^{-2} Df(m_0)$.

Démonstration. — Que $\mathcal{D}(m_0)$ soit un espace vectoriel et l'application $f \in \mathcal{D}(m_0) \rightarrow Df(m_0) \in L(\mathbb{R}^n, \mathbb{R})$ est linéaire résulte de l'écriture

$$(\alpha f + \beta g)(m_0 + h) = (\alpha f + \beta g)(m_0) + (\alpha Df(m_0) + \beta Dg(m_0))(h) + o(h), \quad \|h\| < \inf(r_f, r_g)$$

obtenue par combinaison linéaire de

$$\begin{aligned} f(m_0 + h) &= f(m_0) + Df(m_0)(h) + o(h), & \|h\| < r_f, \\ g(m_0 + h) &= g(m_0) + Dg(m_0)(h) + o(h), & \|h\| < r_g. \end{aligned}$$

Pour le produit, on a

$$\begin{aligned} (f_1 f_2)(m_0 + h) &= (f_1(m_0) + Df_1(m_0)(h) + \|h\| \varepsilon_1(h))(f_2(m_0) + Df_2(m_0)(h) + \|h\| \varepsilon_2(h)) \\ &= f_1(m_0) f_2(m_0) + [f_2(m_0) Df_1(m_0)(h) + f_1(m_0) Df_2(m_0)(h)] \\ &\quad + \|h\| [\varepsilon_1(h)(f_2(m_0) + Df_2(m_0)(h) + \|h\| \varepsilon_2(h)) + \varepsilon_2(h)(f_1(m_0) + Df_1(m_0)(h))] \end{aligned}$$

où le dernier terme est un $o(h)$.

Si $f(m_0) \neq 0$, par continuité il en est de même pour $f(m)$ sur un voisinage de m_0 et

$$\begin{aligned} f(m_0 + h)^{-1} &= \frac{1}{f(m_0) + Df(m_0)h + o(h)} = \frac{1}{f(m_0)} \frac{1}{1 + f(m_0)^{-1}Df(m_0)h + o(h)} \\ &= f(m_0)^{-1}(1 - f(m_0)^{-1}Df(m_0)h - o(h) + o(f(m_0)^{-1}Df(m_0)h + o(h))) \\ &= f(m_0)^{-1} - f(m_0)^{-2}Df(m_0)h + o(h), \end{aligned}$$

où on a utilisé $(1 + u)^{-1} = 1 - u + o(u)$. □

▷ **Exemple 1.7.** — La fonction

$$\text{Inv} : A \in \text{GL}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}), A \text{ inversible}\} \rightarrow A^{-1} \in M_n(\mathbb{R})$$

est différentiable de différentielle

$$D \text{Inv}(A)(h) = -A^{-1}hA^{-1}, \quad h \in M_n(\mathbb{R}).$$

En effet

$$\begin{aligned} (A + h)^{-1} &= A^{-1}(1 + hA^{-1})^{-1} = A^{-1} + A^{-1}((1 + hA^{-1})^{-1} - 1) \\ &= A^{-1} - A^{-1}(1 + hA^{-1})^{-1}hA^{-1} = A^{-1} - A^{-1}hA^{-1} + A^{-1}((1 + hA^{-1})^{-1} - 1)hA^{-1} \\ &= A^{-1} - A^{-1}hA^{-1} + A^{-1}(1 + hA^{-1})^{-1}hA^{-1}hA^{-1} \end{aligned}$$

où le dernier terme est en norme uniformément borné au voisinage de $h = 0$ par $\|h\|^2$ vu que le facteur $(1 + hA^{-1})^{-1} = \sum_{k=0}^{\infty} (-hA^{-1})^k = 1 + o(h)$ est borné au voisinage de $h = 0$. ◁

1.2. Fonctions composées

Théorème 1.1. — Soit U un ouvert de \mathbb{R}^n , V un ouvert de \mathbb{R}^p , $f : U \rightarrow \mathbb{R}^p$ et $g : V \rightarrow \mathbb{R}^q$. On suppose $f(U) \subset V$, f différentiable en $m_0 \in U$ et g différentiable en $f(m_0)$. Alors $g \circ f$ est différentiable en m_0 et

$$\begin{aligned} D(g \circ f)(m_0) &= Dg(f(m_0)) \circ Df(m_0), \\ \text{Jac}(g \circ f)(m_0) &= \text{Jac}(g)(f(m_0)) \text{Jac}(f)(m_0). \end{aligned}$$

Démonstration. — On compose les développements limités

$$\begin{aligned} f(m_0 + h) &= f(m_0) + Df(m_0)(h) + o(h), \\ g(f(m_0) + k) &= g(f(m_0)) + Dg(f(m_0))(k) + o(k), \end{aligned}$$

de telle sorte que

$$\begin{aligned} (g \circ f)(m_0 + h) &= g(f(m_0 + h)) = g(f(m_0) + Df(m_0)(h) + o(h)) \\ &= g(f(m_0)) + Dg(f(m_0))(Df(m_0)(h) + o(h)) + o(Df(m_0)(h) + o(h)) \\ &= (g \circ f)(m_0) + (Dg(f(m_0)) \circ Df(m_0))(h) \\ &\quad + (Dg(f(m_0))(o(h)) + o(Df(m_0)(h) + o(h))). \end{aligned}$$

Le terme $Dg(f(m_0))(o(h)) = Dg(f(m_0))(\varepsilon(h))\|h\|$ est un $o(h)$, alors que le dernier terme, noté $\delta(h)$, est majoré en norme par

$$\begin{aligned} \|\delta(h)\| &\leq \|Df(m_0)(h) + o(h)\| \|\varepsilon(Df(m_0)(h) + o(h))\| \\ &\leq [\|Df(m_0)\| + \|\varepsilon(h)\|] \|h\| \|\varepsilon(Df(m_0)(h) + o(h))\| \end{aligned}$$

et est donc un $o(h)$ (c'est en fait un $o(o(h))$ comme introduit dans l'arithmétique (1)). □

▷ **Exemples 1.8.** —

1. Soit $f : (r, \theta) \in \mathbb{R}^{+*} \times \mathbb{R} \rightarrow (x, y) = (r \cos \theta, r \sin \theta) \in \mathbb{R}^2$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ différentiable. Alors $G = g \circ f$ l'est aussi et, avec $m_0 = (r_0 \cos \theta_0, r_0 \sin \theta_0)$

$$\begin{aligned} \text{Jac}(G)(r_0, \theta_0) &= \text{Jac}(g)(m_0) \text{Jac}(f)(r_0, \theta_0) \\ &= \begin{pmatrix} \frac{\partial g}{\partial x}(m_0) & \frac{\partial g}{\partial y}(m_0) \end{pmatrix} \begin{pmatrix} \cos \theta_0 & -r_0 \sin \theta_0 \\ \sin \theta_0 & r_0 \cos \theta_0 \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta_0 \frac{\partial g}{\partial x}(m_0) + \sin \theta_0 \frac{\partial g}{\partial y}(m_0) & -r_0 \sin \theta_0 \frac{\partial g}{\partial x}(m_0) + r_0 \cos \theta_0 \frac{\partial g}{\partial y}(m_0) \end{pmatrix} \end{aligned}$$

2. Soit $v \in \mathbb{R}^n$ et $f : B(m_0, r) \rightarrow \mathbb{R}$. La dérivée directionnelle $D_v f(m_0)$ apparaît comme la dérivée de l'application composée $F = f \circ \gamma_v$ où $\gamma_v : t \in \mathbb{R} \rightarrow m_0 + tv \in \mathbb{R}^n$. On a en effet

$$DF(t_0) = Df(m_0 + t_0 v) \circ D\gamma_v(t_0).$$

La dérivée $D\gamma_v(t_0)$ s'identifie au vecteur v .

3. Si f et g sont différentiables en m_0 et $f(m_0)$ resp., avec $g \circ f = \text{Id}$, alors $Dg(f(m_0)) \circ Df(m_0) = \text{Id}$. Ainsi, si les dimensions des espaces d'arrivée et de départ de f et g sont égales, $Df(m_0)$ et $Dg(f(m_0))$ sont inversibles. ◁

1.3. Fonctions de classe C^k

Définition 1.5. — Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^p$. Si f est différentiable en tout point $m \in U$ et l'application $Df : m \in U \rightarrow Df(m) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est continue, l'application f est dite *fonction de classe C^1* sur U .

Théorème 1.2. — Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^p$. Si les dérivées partielles $\partial_{x_1} f, \dots, \partial_{x_n} f$ existent en tout point de U et définissent des fonctions continues sur U , alors f est de classe C^1 .

Démonstration. — Pour $h = (h_1, \dots, h_n)$, on écrit

$$\begin{aligned} f(m_0 + h) - f(m_0) &= [f(m_0 + h) - f(m_0 + (h_1, \dots, h_{n-1}, 0))] + \dots \\ &\quad + [f(m_0 + (h_1, \dots, h_j, h_{j+1}, \dots)) - f(m_0 + (h_1, \dots, h_j, 0, \dots))] \\ &\quad + \dots + [f(m_0 + (h_1, 0, \dots, 0)) - f(m_0)] \end{aligned}$$

Dans chaque terme, on applique l'inégalité des accroissements finis pour des fonctions d'une variable à chaque fonction coordonnée : il existe $\theta_{ij}(h) \in [0, 1]$

$$\begin{aligned} f_i(m_0 + (h_1, \dots, h_j, 0, \dots, 0)) - f_i(m_0 + (h_1, \dots, h_{j-1}, 0, \dots, 0)) \\ &= \partial_{x_j} f_i(m_0 + (h_1, \dots, h_{j-1}, \theta_{ij}(h)h_j, 0, \dots, 0))h_j \\ &= [\partial_{x_j} f_i(m_0) + \varepsilon_{ij}(h)]h_j \end{aligned}$$

où $\varepsilon_{ij}(h) \rightarrow 0$ lorsque $h \rightarrow 0$ vu la continuité de $\partial_{x_j} f_i$ au voisinage de m_0 . On a donc

$$f(m_0 + (h_1, \dots, h_j, 0, \dots, 0)) - f(m_0 + (h_1, \dots, h_{j-1}, 0, \dots, 0)) = [\partial_{x_j} f(m_0) + \varepsilon_j(h)]h_j$$

où on a noté $\varepsilon_j(h)$ le vecteur $\varepsilon_j(h) = (\varepsilon_{1j}(h), \dots, \varepsilon_{pj}(h))$, et par suite

$$f(m_0 + h) = f(m_0) + \sum_{j=1}^n \partial_{x_j} f(m_0)h_j + \sum_{j=1}^n \varepsilon_j(h)h_j,$$

où la somme est du type $o(h)$ vu que

$$\left\| \sum_{j=1}^n \varepsilon_j(h)h_j \right\| \leq \left[\sum_{j=1}^n \|\varepsilon_j(h)\| \right] \sup_{j=1}^n |h_j|.$$

La fonction f est différentiable en m_0 : sa jacobienne $\text{Jac}(f)$, dont les vecteurs colonnes sont les dérivées partielles $\partial_{x_j} f$, est continue ; il en est de même pour la différentielle $m \in U \rightarrow Df(m) \in L(\mathbb{R}^n, \mathbb{R}^p)$. \square

▷ **Exemple 1.9.** — L'expression de l'inverse d'une matrice $A \in \text{GL}(n, \mathbb{R})$ en terme de déterminants

$$A^{-1} = (\det A)^{-1} \text{Comat } A$$

où $\text{Comat } A$ est la comatrice de A , matrice transposée des déterminants des cofacteurs de A , indique que les fonctions coordonnées $a_{ij}^1, 1 \leq i, j \leq n$, de A^{-1} sont des fractions rationnelles des coordonnées $(a_{k\ell}), 1 \leq k, \ell \leq n$, admettant des dérivées partielles continues sur l'ouvert $\text{GL}(n, \mathbb{R}) = \{\det[a_{k\ell}] \neq 0\}$: on retrouve ainsi la différentiabilité de l'application Inv de l'exemple 1.3. \triangleleft

La définition des fonctions de classe C^1 est complétée par celle des fonctions de classe C^k pour $k \in \mathbb{N}$ de manière récursive.

Définition 1.6. — Soit U un ouvert de \mathbb{R}^n et f une fonction définie sur U et à valeurs dans \mathbb{R}^p . La fonction f est dite de *fonction de classe C^0* sur U si elle est continue sur U . Pour $k \geq 1$, la fonction f est dite de *fonction de classe C^k* sur U si elle est différentiable et sa différentielle Df est de classe C^{k-1} sur U . La fonction f est dite *fonction de classe C^∞* sur U si elle est C^k sur U pour tout entier k .

On note par $D^k f$ la *dérivée d'ordre k* de f , i. e. la différentielle de $D^{k-1} f$.

△ Si f est de classe C^k , l'application $D^k f$ est une application à valeurs dans l'espace d'applications linéaires $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \dots, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)) \dots)$. ∇

Les fonctions de classe C^k seront étudiées au chapitre 5 : on verra notamment qu'une fonction est de classe C^k si elle admet des dérivées partielles continues jusqu'à l'ordre k .

▷ **Exemple 1.10.** — Les fonctions polynômes, \exp , \log , \cos, \dots sont de classe C^∞ . La fonction exponentielle sur $\mathbb{C} \simeq \mathbb{R}^2$ est de classe C^∞ . On peut montrer qu'il existe une fonction C^∞ sur \mathbb{R}^n et à valeurs numériques telle que $f(x) = 1$ si $\|x\| \leq 1$ et $f(x) = 0$ si $\|x\| \geq 2$ (fonction qui n'est pas obtenue par des opérations algébriques en les fonctions usuelles précédentes...). \triangleleft

1.4. Dérivation et intégration

En une variable, dérivation et intégration (au sens de Riemann dans ce cours) sont des opérations inverses l'une de l'autre.

Théorème 1.3. — Soit I un intervalle ouvert de \mathbb{R} , $m_0 \in I$ et f une fonction définie sur I , à valeurs numériques et de classe C^1 . On a alors

$$(3) \quad f(m) = f(m_0) + \int_{m_0}^m f'(t) dt, \quad m \in I.$$

Démonstration. — Soit F définie par $F(m) = f(m_0) + \int_{m_0}^m f'(t) dt, m \in I$. La continuité de la dérivée f' au voisinage de m_0

$$f'(t) = f'(m_0) + \varepsilon(t - m_0)$$

donne

$$\int_{m_0}^{m_0+T} f'(t) dt = \int_{m_0}^{m_0+T} [f'(m_0) + \varepsilon(t - m_0)] dt = T f'(m_0) + o(T)$$

vu que

$$\left| \int_{m_0}^{m_0+T} \varepsilon(t - m_0) dt \right| \leq T \sup_{\tau \in [0, T]} |\varepsilon(\tau)|,$$

ce qui exprime la différentiabilité de la fonction F en m_0 , avec dérivée $f'(m_0)$. La différentiabilité de F en $m_1 \in I$ résulte de l'écriture

$$F(m) = f(m_0) + \int_{m_0}^{m_1} f'(t)dt + \int_{m_1}^m f'(t)dt, m \in I.$$

Ainsi les deux membres de (3) ont même dérivée, et par suite sont égales. En effet une fonction dérivable et à dérivée nulle, est constante d'après le théorème des accroissements finis en une variable. \square

\triangle Si l'intégrale d'une fonction f d'une variable à valeurs dans \mathbb{R}^p est définie comme le vecteur de \mathbb{R}^p obtenu en intégrant les p fonctions coordonnées f_1, \dots, f_p de f , le théorème précédent est aussi valable pour les fonctions de I dans \mathbb{R}^p et de classe C^1 . Le théorème est aussi valable avec des hypothèses de régularité de la dérivée f' plus faibles, par exemple son caractère réglé. ∇

1.5. Accroissements finis

La formule des accroissements finis pour une fonction numérique dérivable sur $[a, b]$ donne l'existence d'un $c \in [a, b]$ tel que $f(b) - f(a) = f'(c)(b - a)$. Pour une fonction à valeurs vectorielles, il n'y a pas de telle formule des accroissements finis : ainsi de la fonction $f : t \in \mathbb{R} \rightarrow (\cos t, \sin t) \in \mathbb{R}^2$, dont la dérivée $Df(t) = (-\sin t, \cos t)$ ne s'annule jamais, alors que $f(2\pi) - f(0) = 0$ ne peut être de la forme $Df(\theta)2\pi$. Néanmoins, il existe une inégalité des accroissements finis pour les fonctions différentiables sur des ouverts de \mathbb{R}^n et à valeurs vectorielles, source de nombreuses applications.

Théorème 1.4 (Inégalité des accroissements finis). — Soit U un ouvert de \mathbb{R}^n , $f : U \rightarrow \mathbb{R}^p$ de classe C^1 et m_0, m_1 deux points de U tel que le segment $[m_0, m_1] = \{m_0 + t(m_1 - m_0), t \in [0, 1]\}$ soit inclus dans U . Alors ⁽¹⁾

$$(4) \quad \|f(m_1) - f(m_0)\| \leq \left(\sup_{m \in [m_0, m_1]} \|Df(m)\| \right) \|m_1 - m_0\|.$$

L'inégalité des accroissements finis est un résultat tout à fait central du calcul différentiel, si important qu'on vient à l'oublier. Pour insister sur cette importante, on en donnera trois démonstrations ⁽²⁾.

Première démonstration. — La fonction $g : t \in [0, 1] \rightarrow f(m_0 + t(m_1 - m_0))$ est dérivable, de dérivée $\dot{g}(t) = Df(m_0 + t(m_1 - m_0))(m_1 - m_0)$ continue. On a

$$(5) \quad g(1) - g(0) = \int_0^1 \dot{g}(s)ds$$

1. La norme de la différentielle $\|Df(m)\|$ est bien sûr la norme matricielle induite sur $L(\mathbb{R}^n, \mathbb{R}^p)$ par les normes sur \mathbb{R}^n et \mathbb{R}^p utilisées pour estimer les variations respectives $\|m_1 - m_0\|$ et $\|f(m_1) - f(m_0)\|$.

2. La première preuve repose sur le théorème 1.3, qui implique l'identité intégrale (5) pour une fonction d'une variable à valeurs vectorielles dans un espace de dimension finie et donc par suite l'inégalité des accroissements finis : pour montrer le théorème 1.3, on a invoqué l'égalité des accroissements finis valable pour des fonctions d'une variable réelle à valeurs réelles. On s'assure bien qu'on démontre l'inégalité des accroissements finis (4) sans l'utiliser subrepticement, même si on peut utiliser cette inégalité des accroissements finis pour montrer qu'une fonction est l'intégrale de sa dérivée. Les démonstrations suivantes de l'inégalité des accroissements finis n'utilisent pas cette égalité des accroissements finis en dimension 1.

La troisième preuve repose sur l'identité de Grönwall classique d'un cours d'équations différentielles : sa preuve est similaire à celle de notre seconde démonstration de l'inégalité des accroissements finis.

et par suite

$$\|g(1) - g(0)\| = \left\| \int_0^1 \dot{g}(s) ds \right\| \leq \int_0^1 \|\dot{g}(s)\| ds \leq \int_0^1 \|Df(m_0 + s(m_1 - m_0))\| \|m_1 - m_0\| ds,$$

ce qui donne l'inégalité annoncée. \square

Seconde démonstration. — Soit g la fonction définie par $g(t) = f(m_0 + t(m_1 - m_0))$ pour $t \in [0, 1]$, qui est dérivable de dérivée $\dot{g}(t) = Df(m_0 + t(m_1 - m_0))(m_1 - m_0)$. Considérons l'inégalité

$$(6) \quad \|g(t) - g(0)\| \leq \left(\sup_{s \in [0, t]} \|\dot{g}(s)\| + \varepsilon \right) t.$$

et, pour $\varepsilon > 0$, la partie A_ε définie par

$$A_\varepsilon = \{T \in [0, 1] : \text{l'inégalité (6) est vérifiée pour } t \text{ dans } [0, T]\}.$$

La partie A_ε contient $T = 0$ et est un intervalle fermé : les inégalités avec des fonctions continues sont stables par passage à la limite. Soit T_ε la borne supérieure de A_ε et supposons $T_\varepsilon < 1$. La fonction f étant différentiable en T_ε , il existe $\eta > 0$ tel que $T_\varepsilon + \eta < 1$ et

$$\|g(T_\varepsilon + h) - g(T_\varepsilon) - \dot{g}(T_\varepsilon)h\| \leq \varepsilon h, \quad h \in [0, \eta].$$

Alors, pour $h \in [0, \eta]$,

$$\begin{aligned} \|g(T_\varepsilon + h) - g(0)\| &\leq \|g(T_\varepsilon + h) - g(T_\varepsilon) - \dot{g}(T_\varepsilon)h\| + \|\dot{g}(T_\varepsilon)\|h + \|g(T_\varepsilon) - g(0)\| \\ &\leq \varepsilon h + \|\dot{g}(T_\varepsilon)\|h + \left[\sup_{s \in [0, T_\varepsilon]} \|\dot{g}(s)\| + \varepsilon \right] T_\varepsilon \\ &\leq \left[\sup_{s \in [0, T_\varepsilon + h]} \|\dot{g}(s)\| + \varepsilon \right] (T_\varepsilon + h), \end{aligned}$$

inégalité qui exprime l'appartenance de $T_\varepsilon + h$ à A_ε : l'hypothèse $T_\varepsilon < 1$ est donc erronée. Ainsi, faisant tendre ε vers 0 dans (6) avec $t = 1$, on obtient l'inégalité du théorème, en remarquant la majoration $|\dot{g}(t)| \leq \|Df(m_0 + t(m_1 - m_0))\| \|m_1 - m_0\|$. \square

Troisième démonstration. — Soit x la fonction définie sur $[0, 1]$ par

$$x(t) = f(m_0 + t(m_1 - m_0)) - f(m_0), \quad t \in [0, 1].$$

Elle vérifie l'équation différentielle $\dot{x}(t) = Df(m_0 + t(m_1 - m_0))(m_1 - m_0)$, avec membre de droite majorée par la constante M donnée par le majorant de l'inégalité du théorème. L'inégalité $\|x(1)\| \leq \|x(0)\| + M$ résulte de la remarque suivant le corollaire I.4.2 de l'inégalité de Grönwall, c'est l'inégalité annoncée. \square

Corollaire 1.1. — Soit $U \subset \mathbb{R}^n$ un ouvert connexe. Si $f : U \rightarrow \mathbb{R}^p$ de classe C^1 a une différentielle nulle, la fonction f est constante.

Démonstration. — Soit $m_0 \in U$ et $V_0 = \{m \in U : f(m) = f(m_0)\}$. Cette partie V_0 de U est non vide, fermée, comme image réciproque d'un fermé par une fonction continue. Elle est ouverte, puisque si $m \in V_0$, il existe une boule $B(m, r)$ incluse dans U sur laquelle f est constante : pour tout point de \tilde{m} de $B(m, r)$, le segment $[m, \tilde{m}]$ est inclus dans la boule $B(m, r)$ et d'après l'inégalité des accroissements finis $f(\tilde{m}) = f(m)$, ainsi $B(m, r) \subset V_0$. Par connexité de U , $V_0 = U$ et la fonction est constante sur U . \square

Rappelons qu'une fonction $f : U \rightarrow V$ est K -lipschitzienne si $\|f(x) - f(y)\| \leq K\|x - y\|$, $x, y \in U$. Une norme est un exemple de fonction 1-lipschitzienne, comme l'indique une des formes de l'inégalité triangulaire : $\| \|x\| - \|y\| \| \leq \|x - y\|$; cet exemple montre que la différentiabilité n'est pas une condition nécessaire pour le caractère lipschitzien, même si l'inégalité des accroissements finis implique toute fonction de classe C^1 est lipschitzienne.

Corollaire 1.2. — Si $f : U \rightarrow \mathbb{R}^p$ est de classe C^1 , la fonction f est localement lipschitzienne, i. e. pour $m \in U$, il existe une constante $r > 0$ une constante $K_{m,r}$ telle que $\|f(x) - f(y)\| \leq K_{m,r}\|x - y\|$ pour x, y dans la boule $B(m, r)$.

Démonstration. — Soit $m \in U$ et $r > 0$ telle que la boule fermée $\overline{B(m, r)}$ soit incluse dans U . Alors si $K_{m,r} = \sup_{x \in \overline{B(m, r)}} \|Df(x)\|$, d'après le théorème des accroissements finis, on a

$$\|f(x) - f(y)\| \leq K_{m,r}\|x - y\|, \quad x, y \in B(m, r),$$

ce qui est le caractère localement lipschitzien de f . \square

\triangle Bien que très pratique, la condition de différentiabilité n'est pas une condition nécessaire du caractère lipschitzien : d'après l'inégalité triangulaire écrite sous la forme

$$|N(x) - N(y)| \leq N(x - y), \quad x, y \in \mathbb{R}^n,$$

toute norme N est lipschitzienne sur \mathbb{R}^n . ∇

Le théorème 1.3 de dérivabilité de l'intégrale par rapport à une de ses bornes est complété par un théorème de dérivabilité d'une intégrale dépendant d'un paramètre.

Théorème 1.5. — Soit I un intervalle de \mathbb{R} , U un ouvert de \mathbb{R}^m , $y_0 \in U$. Soit f une fonction de $I \times U$ dans \mathbb{R}^p , $[a, b]$ un segment inclus dans I et $g : (a, b) \rightarrow \mathbb{R}$ continue telle que $\int_a^b |g(t)| dt < +\infty$. Si f est de classe C^1 , alors la fonction F définie par

$$F(x) = \int_a^b f(t, x)g(t)dt, \quad x \in U,$$

est différentiable sur U , de différentielle $DF(x) = \int_a^b D_x f(t, x)g(t)dt, x \in U$.

Démonstration. — Soit $x_0 \in U$ et $B(x_0, r_0)$ une boule telle que son adhérence $\{\|x - x_0\| \leq r_0\}$ soit incluse dans U . Pour $t \in I$, la fonction

$$\Phi_t : x \in U \rightarrow f(t, x) - f(t, x_0) - D_x f(t, x_0)(x - x_0)$$

est différentiable avec différentielle $D\Phi_t(x) = D_x f(t, x) - D_x f(t, x_0)$ et $\Phi_t(x_0) = 0$. Ainsi d'après l'inégalité des accroissements finis,

$$(7) \quad \|f(t, x_0 + h) - f(t, x_0) - D_x f(t, x_0)h\| \leq \sup_{z \in B(x_0, \|h\|)} \|D_x f(t, z) - D_x f(t, x_0)\| \|h\|.$$

Le premier facteur du second membre est majoré par $M_0(\|h\|)$ avec

$$M_0(r) = \sup_{z \in B(x_0, r), t \in [a, b]} \|D_x f(t, z) - D_x f(t, x_0)\|,$$

qui tend vers 0 lorsque $r \rightarrow 0$: sinon, il existerait $\alpha > 0$ et une suite $(t_n, x_n) \in [a, b] \times \overline{B(x_0, 1/n)}$ telles que $\|D_x f(t_n, x_n) - D_x f(t_n, x_0)\| \geq \alpha$ et la continuité de $D_x f$ en un point d'adhérence (t_∞, x_0) de la suite (t_n, x_n) amènerait une contradiction. Ainsi

$$\left\| F(x_0 + h) - F(x_0) - \int_a^b D_x f(t, x_0)hg(t)dt \right\| \leq M_0(\|h\|)\|h\| \int_a^b |g(t)|dt,$$

soit

$$F(x_0 + h) = F(x_0) + \int_a^b D_x f(t, x_0)hg(t)dt + o(\|h\|),$$

ce qui est bien l'expression de la différentiabilité de F en x_0 avec comme différentielle $DF(x_0) = \int_a^b D_x f(t, x_0)g(t)dt$. \square

▷ **Exemples 1.11.** —

1. Pour n entier, la fonction L_n définie par

$$L_n(x) = \frac{1}{\pi} \int_0^\pi \left[x + i\sqrt{1-x^2} \cos t \right]^n dt, \quad x \in (-1, 1)$$

est dérivable. En fait, en développant l'intégrale, on s'assure que la fonction L_n est polynomiale, et donc de classe C^∞ . La fonction L_n est connue sous le nom de polynôme de Legendre.

2. La fonction Γ d'Euler définie sur \mathbb{R}_*^+ par

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt, \quad x > 0,$$

est de classe C^∞ : le théorème précédent est insuffisant pour le démontrer, vu que l'intervalle d'intégration $[0, +\infty)$ est non compact. Néanmoins, en remarquant que $\partial_x t^x = \log t t^x$, $\partial_x^2 t^x = (\log t)^2 t^x$, la formule de Taylor-Lagrange

$$t^x = t^{x_0} + \log t t^{x_0} + \int_{x_0}^x (x - \sigma)(\log t)^2 t^\sigma d\sigma$$

amène la majoration

$$|t^x - t^{x_0} - \log t t^{x_0}| = \left| \int_{x_0}^x (x - \sigma)(\log t)^2 t^\sigma d\sigma \right| \leq |\log t|^2 \frac{|x - x_0|^2}{2} \sup_{\sigma \in [x, x_0]} t^\sigma$$

analogue de la majoration (7), étape essentielle de la démonstration générale précédente. On conclut à la dérivabilité pour $x > 0$, et même indéfiniment, avec l'équation fonctionnelle

$$\Gamma(x+1) = x\Gamma(x), \quad x > 0$$

obtenue par une simple intégration par parties.

3. Pour ν réel au moins égal à $-1/2$, la fonction J_ν définie par

$$J_\nu(x) = \frac{(x/2)^\nu}{\sqrt{\pi}\Gamma(\nu+1/2)} \int_0^\pi \cos(x \cos t) \sin^{2\nu} t dt, \quad x > 0,$$

est de classe C^∞ . Elle est connue sous le nom de fonction de Bessel d'indice ν . Pour n entier positif, on montre que $J_n(x) = \pi^{-1} \int_0^\pi \cos(x \sin \theta - n\theta) d\theta$. ◁

CHAPITRE 2

LE THÉORÈME D'INVERSION LOCALE

Le théorème d'inversion locale est au centre de ce cours et de multiples développements du calcul différentiel. Après avoir énoncé le théorème et discuté sa validité, notamment dans ses rapports d'extension des propriétés connues des fonctions d'une variable réelle, on revient au théorème du point fixe en en présentant une version avec paramètre. La démonstration du théorème d'inversion locale en fait usage et on termine par la notion de changement de variable ou fonctions coordonnées que justifie le théorème d'inversion locale.

2.1. L'énoncé du théorème

Pour les fonctions d'une variable réelle, les homéomorphismes entre intervalles sont caractérisées comme les applications continues monotones. D'autre part, un homéomorphisme $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^p$ avec f, f^{-1} différentiables, a une différentielle inversible.

▷ **Exemples 2.1.** —

1. La fonction $\operatorname{tg} : (-\pi/2, \pi/2) \rightarrow \mathbb{R}$ est une application bijective différentiable, de dérivée $1 + \operatorname{tg}^2$ et d'application inverse arctg dérivable : $\operatorname{arctg}'(y) = (1 + y^2)^{-1}$.
2. La fonction $f(t) = t^3$ est une bijection de \mathbb{R} sur lui-même, de dérivée $\dot{f}(t) = 3t^2$ inversible si et seulement si t non nul. La bijection réciproque $t \rightarrow t^{1/3}$ n'est pas différentiable en $t = 0$.
3. La fonction $f(t) = t + t^2 \sin(1/t)$ est dérivable sur \mathbb{R} : si t est non nul, sa dérivée est $\dot{f}(t) = 1 - \cos(1/t) + 2t \sin(1/t)$ et en $t = 0$ elle vaut 1. La dérivée \dot{f} n'est pas continue en $t = 0$ et, pour k entier non nul, la dérivée

$$\dot{f}(1/(2k\pi + u)) = 1 - \cos u + 2(2k\pi + u)^{-1} \sin u \sim_{k \rightarrow +\infty, u \rightarrow 0} (k\pi)^{-1} u$$

change de signe au voisinage de $u = 0$: la fonction f n'est pas monotone au voisinage de $t = 0$.

4. L'élevation au carré $C : z \in \mathbb{C} \rightarrow z^2 \in \mathbb{C}$, soit $(x, y) \in \mathbb{R}^2 \rightarrow (x^2 - y^2, 2xy) \in \mathbb{R}^2$ via l'identification $z = x + iy \in \mathbb{C} \simeq (x, y) \in \mathbb{R}^2$, est de classe C^∞ , avec différentielle $DC(z)$, multiplication par $2z$ dans \mathbb{C} (c'est l'application $(h, k) \in \mathbb{R}^2 \rightarrow 2(xh - yk, yh + xk) \in \mathbb{R}^2$ sur le plan \mathbb{R}^2), inversible pour tout z non nul. L'application $C : \mathbb{C}^* \rightarrow \mathbb{C}^*$ n'est pas injective.
5. La fonction exponentielle $E_{\mathbb{C}}$ de \mathbb{C} est différentiable,

$$E_{\mathbb{C}}(z + h) = e^{z+h} = e^z e^h = e^z(1 + h + o(h)) = E_{\mathbb{C}}(z) + e^z h + o(h)$$

avec pour différentielle $DE_{\mathbb{C}}(z)$ l'endomorphisme induit par la multiplication par e^z : bien que la différentielle $DE_{\mathbb{C}}(z)$ soit inversible pour tout z , l'application $E_{\mathbb{C}} : \mathbb{C} \rightarrow \mathbb{C}$ n'est pas injective ($e^{2i\pi} = e^0$). ◁

Définition 2.1. — $f : U \rightarrow V$ est un *difféomorphisme* de classe C^1 si f est une bijection de U sur V , f est de classe C^1 , de différentielle $Df(m)$ inversible pour tout $m \in U$ et il en est de même pour $f^{-1} : V \rightarrow U$.

Théorème 2.1 (Théorème d'inversion locale). — Soit U un ouvert de \mathbb{R}^n , $f : U \rightarrow \mathbb{R}^n$ une fonction de classe C^1 , $m_0 \in U$ tel que $Df(m_0)$ soit inversible. Alors il existe un voisinage ouvert U' de m_0 dans U , V' de $f(m_0)$ dans \mathbb{R}^n tels que $f|_{U'}$ soit un difféomorphisme de classe C^1 de U' sur V' .

▷ **Exemple 2.2.** — Soit $f : (s, t) \rightarrow (s + st, t + st)$. Sa jacobienne est $\begin{pmatrix} 1+t & t \\ s & 1+s \end{pmatrix}$, inversible au voisinage de $(s, t) = (0, 0)$. ◁

△ Dire que f est un difféomorphisme local en m_0 , c'est affirmer l'existence d'un voisinage ouvert de m_0 tel que $f : U \rightarrow f(U)$ soit un difféomorphisme. Le théorème d'inversion locale énonce simplement qu'une application f de classe C^1 et avec $Df(m_0)$ inversible est un difféomorphisme local (en m_0). ▽

2.2. Le théorème du point fixe

On se place dans un espace métrique (X, d) .

Définition 2.2. — L'application $T : X \rightarrow X$ est une *contraction* de X s'il existe $k \in [0, 1)$ avec

$$d(Tx, Ty) \leq kd(x, y), \quad x, y \in X.$$

T est dit *k-contractante*.

▷ **Exemple 2.3.** — Une application affine $x \rightarrow Ax + b$ est une contraction de $(\mathbb{R}^n, \| \cdot \|)$ dès que $\|A\| < 1$. Le choix de la norme $\| \cdot \|$ importe. Par exemple, $A_\varepsilon = \begin{pmatrix} 1/2 & \varepsilon \\ 0 & 1/3 \end{pmatrix}$ est de norme $\sup(1/2 + \varepsilon, 1/3)$ pour la norme matricielle dérivée de la norme $\| \cdot \|_\infty$ sur \mathbb{R}^n et $\sup(1/2, \varepsilon + 1/3)$ pour celle dérivée de $\| \cdot \|_1$: si $\varepsilon = 1/2$, A n'est pas contractante relativement à la première, mais l'est pour la seconde. ◁

Les contractions sont très utiles à travers le théorème du point fixe dit de Picard.

Théorème 2.2 (Picard). — Si (X, d) est complet, toute contraction T de X a un unique point fixe.

Le théorème de Picard 2.2 a sa preuve incluse dans celle de sa version avec paramètres comme suit.

Théorème 2.3 (Picard avec paramètres). — Soit X, Y des espaces métriques, avec X complet et $T : X \times Y \rightarrow X$ continue, avec les applications partielles $T_y : x \in X \rightarrow T(x, y) \in X$ k -contractantes de même constante $k \in [0, 1)$ pour tout $y \in Y$. L'application T_y a un point fixe unique a_y et l'application $y \in Y \rightarrow a_y \in X$ est continue.

Démonstration. — Si T_y a un point fixe, il est unique : sinon, avec deux points fixes x_1, x_2 on aurait

$$d(x_1, x_2) = d(T_y(x_1), T_y(x_2)) \leq kd(x_1, x_2) < d(x_1, x_2)$$

ce qui est absurde.

Soit $x_0 \in X$. La suite $(x_n(y))_{n \geq 0}$ définie par $x_0(y) = x_0$ et $x_n(y) = T_y(x_{n-1}(y))$ si $n \geq 1$ est de Cauchy :

$$\begin{aligned} d(x_{p+\ell}(y), x_p(y)) &\leq d(x_{p+\ell}(y), x_{p+\ell-1}(y)) + \dots + d(x_{p+1}(y), x_p(y)) \\ &\leq d(T_y^{p+\ell-1}(x_1(y)), T_y^{p+\ell-1}(x_0(y))) + \dots + d(T_y^p(x_1(y)), T_y^p(x_0(y))) \\ &\leq (k^{p+\ell-1} + \dots + k^{p-1})d(x_1(y), x_0(y)) \leq \frac{k^{p-1}}{1-k}d(T_y(x_0), x_0) \end{aligned}$$

Vu la complétude de l'espace X , la suite $(x_n(y))_{n \geq 0}$ est convergente et sa limite $x_\infty(y)$ est un point fixe de T_y :

$$x_\infty(y) = \lim_{p \rightarrow \infty} T_y^p(x) = T_y(\lim_{p \rightarrow \infty} T_y^{p-1}(x)) = T_y(x_\infty(y))$$

c'est l'unique point fixe, noté $a(y)$ et indépendant du point x_0 , de T_y .

On a alors

$$\begin{aligned} d(a(y), a(y')) &\leq d(a(y), T_{y'}(a(y))) + d(T_{y'}(a(y)), a(y')) \\ &\leq d(a(y), T_{y'}(a(y))) + d(T_{y'}(a(y)), T_{y'}(a(y'))) \\ &\leq d(a(y), T_{y'}(a(y))) + kd(a(y), a(y')) \end{aligned}$$

d'où

$$(1 - k)d(a(y), a(y')) \leq d(a(y), T(a(y), y')),$$

où, le point y étant fixé, le membre de droite de cette inégalité tend vers 0 lorsque y' tend vers y : c'est la continuité de l'application $z \rightarrow a(z)$ au point y . ◻

△ La preuve d'existence du point fixe avec paramètre est celle du théorème de Picard : la convergence uniforme de la suite $(x_n(y))_{n \geq 0}$, qui apparaît dans la démonstration de la convergence, aurait pu être invoquée pour la continuité du point fixe. ▽

2.3. La preuve du théorème d'inversion locale

Par composition par des translations τ_v ($\tau_v(m) = m + v, m \in \mathbb{R}^n$) et d'une application linéaire, on remplace f par l'application $\tilde{f} = Df(m_0)^{-1} \circ \tau_{-f(m_0)} \circ f \circ \tau_{m_0} : f$ et \tilde{f} sont, ou ne sont pas, pareillement des difféomorphismes locaux. on peut donc se ramener au cas où $m_0 = 0, f(m_0) = 0$ et $Df(m_0) = Id$, ce qui sera supposé désormais.

L'équation $f(x) = y$ est remplacée par la recherche du point fixe pour T_y définie par $T_y(x) = y + x - f(x)$, soit $T_y = \tau_y \circ \varphi$ avec $\varphi(x) = x - f(x), x \in U$.

Vu que $D\varphi(0) = 0$ et que φ est C^1 , il est possible de choisir $r > 0$ tel que $\overline{B(0, r)} \subset U$ et $\|D\varphi(x)\| \leq 1/2$ pour $x \in \overline{B(0, r)}$. Par l'inégalité des accroissements finis, l'application φ , et par suite T_y , est $1/2$ -contractante sur $\overline{B(0, r)}$. Par ailleurs, en prenant y avec $\|y\| < r/2$, on a

$$\|T_y(x)\| = \|y + \varphi(x)\| \leq \|y\| + \|\varphi(x)\| < r/2 + r/2 = r$$

si $\|x\| \leq r$, *i. e.* T_y laisse invariante la boule fermée $\overline{B(0, r)}$, qui est un espace complet. D'après le théorème du point fixe avec paramètre, T_y admet un point fixe unique, noté $g(y)$, dans $\overline{B(0, r)}$ et l'application $y \in B(0, r/2) \rightarrow g(y) \in B(0, r)$ est continue. On pose alors

$$V' = B(0, r/2), \quad U' = \{x, \|f(x)\| < r/2 \text{ et } \|x\| < r\} = f^{-1}(V') \cap B(0, r).$$

L'application $f : U' \rightarrow V'$, d'inverse $g : V' \rightarrow U'$ continue, est un homéomorphisme.

Pour montrer que g est de classe C^1 sur V' , soit $y \in V', x = g(y), k$ tel que $y + k \in V'$ et $h = g(y + k) - g(y)$. La continuité de g affirme que $h \rightarrow 0$ lorsque $k \rightarrow 0$. Tout d'abord, pour $x \in U'$, la différentielle $Df(x) = 1 - D\varphi(x)$ est inversible, avec inverse donné par la série convergente $\sum_{k=0}^{\infty} (D\varphi(x))^k$ puisque $\|D\varphi(x)\| \leq 1/2$. Alors, vu $x + h = g(y + k)$, soit

$$k = f(x + h) - y = f(x + h) - f(x) = Df(x)(h) + o(h),$$

on a

$$(8) \quad h = [Df(x)]^{-1}[k - o(h)] = [Df(x)]^{-1}(k) + o(h),$$

et donc, pour h assez petit

$$\|h\| \leq \|[Df(x)]^{-1}\| \|k\| + o(\|h\|) \leq \|[Df(x)]^{-1}\| \|k\| + \|h\|/2$$

et par suite $\|h\| \leq 2\|[Df(x)]^{-1}\| \|k\|$. Si $k \rightarrow 0$, un $o(h) = \|h\|\varepsilon(h)$, est ainsi un $o(k)$ et, d'après (8),

$$g(y + k) - g(y) = h = [Df(x)]^{-1}(k) + o(k).$$

C'est exactement la différentiabilité de g en $y = f(x)$, avec comme différentielle

$$Dg(y) = [Df(x)]^{-1} = [Df(g(y))]^{-1}.$$

Remarquons enfin que la différentielle $y \in V' \rightarrow Dg(y) = [Df(g(y))]^{-1}$ est la composée d'applications continues et est donc continue, ce qui achève la preuve du théorème d'inversion locale.

2.4. Changement de variables

L'exemple le plus simple de difféomorphisme entre ouverts de \mathbb{R}^n est donné par une matrice inversible A , difféomorphisme de \mathbb{R}^n sur lui-même. L'interprétation géométrique d'une telle matrice A est double : soit c'est une transformation bijective de \mathbb{R}^n , soit elle permet un changement de coordonnées de l'espace \mathbb{R}^n en apparaissant comme la matrice d'un changement de bases. Pour préciser ce deuxième point de vue, soit E un espace vectoriel avec une base $\mathbf{b} = (b_1, \dots, b_n)$ et $X = (x_1, \dots, x_n)$ le vecteur des coordonnées de $m \in E$ relativement à la base \mathbf{b} . Si $\mathbf{c} = (c_1, \dots, c_n)$ est une autre base de E , avec P matrice de changement de base de \mathbf{b} à \mathbf{c} , i. e. $c_i = P b_i$, le vecteur des coordonnées $Y = (y_i)$ de m relativement à la base \mathbf{c} est donné par $Y = AX$ avec $A = P^{-1}$. En effet,

$$m = \sum_i x_i b_i = \sum_j y_j c_j = \sum_j y_j P b_j = \sum_j y_j \sum_i P_{ij} b_i = \sum_i \left[\sum_j P_{ij} y_j \right] b_i$$

soit $x_i = \sum_j P_{ij} y_j$ ou encore $X = PY$. Si E est de dimension 2, les droites parallèles au vecteur c_2 ont pour équation $y_1 = \alpha_1$ dans le système de coordonnées (y_i) , alors que dans le système de coordonnées (x_i) elles ont pour représentation paramétrique $u \in \mathbb{R} \rightarrow P(\alpha_1, u)$: le changement de variable induit par A permet un *redressement* de droites en position quelconque dans le système de coordonnées initial (x_i) en un système de droites parallèles à un des axes de coordonnées dans le système de coordonnées (y_i) ; en dimension n , on a les mêmes remarques pour des systèmes d'hyperplans ou de sous-espaces de codimension supérieures parallèles.

Un difféomorphisme $f : U \rightarrow V$ est simplement une application bijective, différentiable ainsi que son inverse. Il peut être vu comme un changement de coordonnées : un point m au voisinage de $a \in U$, repéré par ses coordonnées (x_i) dans l'ouvert $U \subset \mathbb{R}^n$, l'est tout autant par les coordonnées $(y_i) = (y_i(x_1, \dots, x_n))$ de son image $f(m)$ dans $V = f(U)$. En particulier, en dimension $n = 2$, le système de segments de droites d'équation $y_1 = \alpha_1$ dans $f(V)$, avec α_1 dans un voisinage $I_{y_1(f(a))}$ de $y_1(f(a))$ correspond au tissu de courbes paramétrées $u \in I_{y_1(f(a))} \rightarrow (x_1(\alpha_1, u), x_2(\alpha_1, u))$ où $(x_1(y_1, y_2), x_2(y_1, y_2)) = f^{-1}(y_1, y_2)$. Ce système d'arcs courbés dans U est *redressé* par le difféomorphisme f en un système de segments de droites parallèles dans l'ouvert V :

Définition 2.3. — Soient f_1, \dots, f_n des fonctions numériques sur $U \subset \mathbb{R}^n$. On dit que (f_1, \dots, f_n) est un *changement de coordonnées* ou *changement de variables* au voisinage de a si l'application $F = (f_1, \dots, f_n)$ est un difféomorphisme local au voisinage de a . La famille (f_1, \dots, f_n) induit un *système de coordonnées*.

▷ **Exemple 2.4.** — Les coordonnées polaires sont liées au difféomorphisme local

$$\Psi : (r, \theta) \in \mathbb{R}^{+*} \times \mathbb{R} \rightarrow (x, y) = (r \cos \theta, r \sin \theta) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

L'application Ψ n'est pas bijective : on peut définir des inverses partiels, tels

$$\Phi_1 : (x, y) \in \mathbb{R}^2 \setminus \{(x, 0), x \leq 0\} \rightarrow (r, \theta) = (r, \theta_1(x, y))$$

avec

$$\theta_1(x, y) = \begin{cases} \pi/2 - \operatorname{arctg}(x/y) & \text{si } x > 0, \\ \pi/2 - \operatorname{arctg}(y/x) & \text{si } y > 0, \\ -\pi/2 - \operatorname{arctg}(y/x) & \text{si } y < 0. \end{cases}$$

ou plus simplement $\theta_1(x, y) = 2 \operatorname{arctg}(y/(x + \sqrt{x^2 + y^2}))$, qui donne un changement de variable communément noté (r, θ) où on ne précise pas la formule exacte pour θ , ni son domaine de validité. ◁

Avec ce concept de changement de variables, on va pouvoir considérer localement toute fonction non critique comme une forme linéaire.

Définition 2.4. — Soit $h : U \rightarrow \mathbb{R}$ une fonction différentiable. Le point $m_0 \in U$ est dit *point critique* de h si la différentielle $Dh(m_0)$ est nulle.

▷ **Exemple 2.5.** — Soit A une matrice symétrique et h_A la fonction définie sur $\mathbb{R}^n \setminus \{0\}$ par

$$h_A(x) = \frac{\langle Ax, x \rangle}{\langle x, x \rangle}, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

La différentielle de h_A est donnée par

$$dh_A(x)(X) = \frac{2\langle Ax, X \rangle}{\langle x, x \rangle} - \frac{2\langle Ax, x \rangle \langle x, X \rangle}{\langle x, x \rangle^2}, \quad X \in \mathbb{R}^n.$$

Soit (v_1, \dots, v_n) une base orthonormée de vecteurs propres de A , avec valeurs propres associées $\lambda_1, \dots, \lambda_n$ ordonnées de manière croissante : $\lambda_1 \leq \dots \leq \lambda_n$. Si $v = \sum_{i=1}^n x_i v_i$ est un point critique de h_A et I le premier indice i tel que x_i soit non nul, alors l'annulation de $dh_A(x)(v_I)$ donne

$$\lambda_I x_I \sum_{i=I}^n x_i^2 = \left[\sum_{i=I}^1 \lambda_i x_i^2 \right] x_I,$$

ce qui implique $x_i = 0$ pour tous les i tels que $\lambda_i > \lambda_I$: le vecteur v est donc dans l'espace propre de la valeur propre λ_I . Les points critiques de h_A sont les vecteurs propres de la matrice A . ◁

Théorème 2.4. — Soit $h : U(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$ une fonction de classe C^1 . Alors, pour tout $m_0 \in U$ non critique pour h , il existe un voisinage U' de m_0 et des fonctions coordonnées $f = (f_1, \dots, f_n) : U' \rightarrow \mathbb{R}$ avec $f_1 = h$.

△ Ce théorème peut-être vu comme un théorème de redressement des lignes de niveau $L_{F_0} = \{m \in U, f(m) = F_0\}$: si on change de coordonnées en prenant les coordonnées (h_1, \dots, h_n) , la ligne de niveau L_{F_0} apparaît comme la ligne de niveau de la forme linéaire f_1 (la projection sur la première coordonnée), *i. e.* une portion d'hyperplan. ▽

Démonstration. — La forme linéaire $Dh(m_0)$ étant non nulle, le théorème de la base incomplète appliquée dans le dual de \mathbb{R}^n affirme l'existence de $n - 1$ formes linéaire L_2, \dots, L_n telles que $(Dh(m_0), L_2, \dots, L_n)$ soit une base de ce dual. Alors l'application $F = (h, L_2, \dots, L_n)$ de U dans \mathbb{R}^n a pour matrice jacobienne en m_0

$$JF(m_0) = \begin{pmatrix} Dh(m_0) \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$$

qui est inversible puisque les formes linéaires $Dh(m_0), L_2, \dots, L_n$ sont linéairement indépendantes. Ainsi (h, L_2, \dots, L_n) est un système de coordonnées au voisinage de m_0 . □

▷ **Exemple 2.6.** — Soit h définie sur \mathbb{R}^2 par $h(x, y) = x^3 - y^2$. Sa différentielle est $Dh(x, y) = 3x^2 dx - 2y dy$ et son seul point critique est l'origine $(0, 0)$. Au point (x_0, y_0) avec $x_0 \neq 0$, on peut prendre comme seconde fonction coordonnée L_2 la forme linéaire y , vu que la matrice jacobienne de $\Phi : (x, y) \rightarrow (h(x, y), y)$ valent

$$J\Phi(x_0, y_0) = \begin{pmatrix} 3x_0^2 & -2y_0 \\ 0 & 1 \end{pmatrix}$$

est inversible. Pour un point $(0, y_0)$ distinct de l'origine, on peut prendre la forme linéaire L_2 définie par $L_2(x, y) = x$. ◁

CHAPITRE 3

LE THÉORÈME DES FONCTIONS IMPLICITES

Dans le plan ou pour des fonctions f de deux variables (x, y) , le théorème des fonctions implicites a pour objet d'exprimer la relation entre x et y le long du lieu géométrique donné par l'équation $f(x, y) = 0$ sous la forme d'une relation $y = \varphi(x)$ ou $x = \psi(y)$: toute courbe est un graphe. Les exemples suivants sont particuliers puisqu'on peut y donner ces fonctions φ ou ψ par des formules explicites utilisant des fonctions usuelles. Le théorème des fonctions implicites, avec ses hypothèses minimales qui sont exactement les conditions linéarisées telles qu'elles apparaissent dans le troisième exemple de l'algèbre linéaire ci-dessous, affirme la validité de cette description du lieu $\{f(x, y) = 0\}$ comme graphe.

▷ **Exemples 3.1.** —

1. Le cercle C d'équation $x^2 + y^2 = 1$ est l'union d'arcs paramétrés $x = \pm\sqrt{1 - y^2}, y \in (-1, 1)$ et $y = \pm\sqrt{1 - x^2}, x \in (-1, 1)$. Le cercle C n'est pas un graphe globalement.
2. L'écriture de l'équation $y - \varphi(x) = 0$ sous la forme $x = \psi(y)$ revient exactement au problème de l'inversion de la fonction φ . On a vu que le théorème d'inversion locale donnait une solution satisfaisante à ce problème d'inversion.
3. Soit $SV = 0$ un système linéaire à p équations et $p + n$ inconnues $V = (x_1, \dots, x_{n+p})$. On suppose S de rang p avec les p dernières colonnes indépendantes et on écrit $S = (AB)$ avec $A \in \mathcal{M}_{pn}$ et $B \in \mathcal{M}_p$. À cette décomposition correspond celle de V en $V = (x, y)$ avec les variables libres $x \in \mathbb{R}^n$ et les variables principales $y \in \mathbb{R}^p$. Alors $SV = 0$ s'écrit $Ax + By = 0$, résolu suivant $y = -B^{-1}Ax$. ◁

3.1. Le théorème et sa preuve

Théorème 3.1 (Théorème des fonctions implicites). — Soit U un ouvert de \mathbb{R}^n , V un ouvert de \mathbb{R}^p et $f : U \times V \rightarrow \mathbb{R}^p$ de classe C^1 . On considère un point $m_0 = (x_0, y_0)$ de $U \times V$ tel que $f(m_0) = 0$ avec la différentielle partielle $D_y f(m_0)$ inversible. Alors il existe un voisinage U_0 de x_0 inclus dans U , V_0 de y_0 inclus dans V et une application $\varphi : U_0 \rightarrow V_0$ tels que $D_y f(m)$ soit inversible pour $m \in U_0 \times V_0$ et que dans $U_0 \times V_0$ l'équation $f(x, y) = 0$ soit équivalente à $y = \varphi(x)$. De plus, la différentielle $D\varphi$ est donnée par

$$D\varphi(x) = -[D_y f(x, \varphi(x))]^{-1} D_x f(x, \varphi(x)), \quad x \in U_0.$$

Démonstration. — Soit $g : U \times V \rightarrow \mathbb{R}^n \times \mathbb{R}^p$ définie par $g(x, y) = (x, f(x, y))$. La fonction g est C^1 , de matrice jacobienne

$$\begin{pmatrix} \text{Id}_{\mathbb{R}^n} & 0 \\ J_x f & J_y f \end{pmatrix},$$

inversible en m_0 . Ainsi, d'après le théorème d'inversion locale, il existe une boule B contenant $(x_0, 0) = g(x_0, y_0)$ et W un voisinage de (x_0, y_0) tels que g soit un difféomorphisme de W

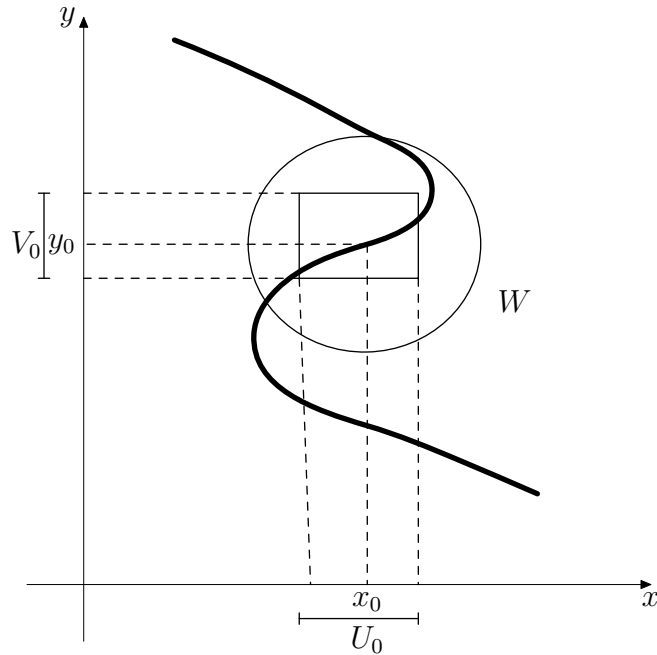


FIGURE 2. Une courbe est localement un graphe.

sur B avec l'inverse $h = g^{-1}$ vérifiant $h(X, Z) = (X, \Phi(X, Z))$. On a donc pour $(x, y) \in W$,

$$\begin{aligned} f(x, y) = 0 &\iff g(x, y) \in B \cap \{z = 0\} \iff (x, y) \in g^{-1}(B \cap \{z = 0\}) \\ &\iff (x, y) = (x, \Phi(x, 0)) \iff y = \varphi(x) \end{aligned}$$

où on a posé $\varphi(x) = \Phi(x, 0)$.

Soient U_1 et V_1 des ouverts contenant x_0 et y_0 resp. et tels que $R_1 = U_1 \times V_1$ soit inclus dans W . Il n'est pas sûr que $R_1 \cap \{f(x, y) = 0\}$ soit un graphe : il se peut que la fibre $\{x\} \times V_1$ ne contienne pas de points de $f = 0$. Néanmoins, du fait que φ est continue, il existe un ouvert U_0 tel que $\varphi(U_0) \subset V_1$. Alors, posant $V_0 = V_1$, on a un rectangle $R_0 = U_0 \times V_0$ tel que $R_0 \cap \{f = 0\}$ soit un graphe.

La fonction Φ étant C^1 , la fonction φ l'est aussi. Différenciant $f(x, \varphi(x)) = 0$, on obtient

$$D_x f(x, \varphi(x)) + D_y f(x, \varphi(x)) D_x \varphi(x) = 0, \quad x \in U_0,$$

d'où la formule pour la différentielle de φ au point x . □

▷ **Exemple 3.2.** — La fonction $f(x, y) = y^3/3 - y - x$ est de classe C^1 sur le plan \mathbb{R}^2 , avec pour différentielle $Df(x, y) = -dx + (y^2 - 1)dy$. Alors, le lieu C d'équation $f = 0$ est au voisinage de tout point $m \in C$ localement un graphe de la forme $y = \varphi(x)$, sauf les deux points $(x = \mp 2/3, y = \pm 1)$. Pour $x \neq \pm 2/3$, on a la différentielle $D\varphi$ donnée par $1/(\varphi^2(x) - 1)$. ◁

3.2. Racines d'équations polynomiales

3.2.1. Variation de racines simples. — Soit $P_\lambda(Y) = Y^3 + a_2(\lambda)Y^2 + a_1(\lambda)Y + a_0(\lambda)$ un polynôme dépendant d'un paramètre λ parcourant un ouvert $\Lambda \subset \mathbb{R}^\ell$, avec les coefficients a_0, a_1, a_2 des fonctions de classe C^1 sur Λ . On note $p(\lambda, y)$ la fonction définie sur $\Lambda \times \mathbb{R}$ associée : le polynôme dérivé $P'_\lambda(Y)$ est identifié pareillement avec la dérivée partielle $\partial_y p(\lambda, y)$. On suppose $y_0 \in \mathbb{R}$ racine simple du polynôme $P_{\lambda_0}(Y)$, i. e. $\partial_y p(\lambda_0, y_0) \neq 0$. Le théorème des fonctions implicites assure l'existence de $r > 0$ et $\delta > 0$ tels que pour tout $\lambda \in B(\lambda_0, r)$, le polynôme $P_\lambda(Y)$ a une unique racine simple $y = \varphi(\lambda)$ dans $(y_0 - \delta, y_0 + \delta)$, avec dépendance C^1 en λ .

3.2.2. Le théorème fondamental de l'algèbre. —

Théorème 3.2. — *Tout polynôme complexe P de degré $n > 0$ admet au moins une racine dans \mathbb{C} .*

Démonstration. — Soit $P(Z) = p_n Z^n + p_{n-1} Z^{n-1} + \dots + p_1 Z + p_0$ un polynôme de degré n ; on peut supposer sans restriction $p_n = 1$. On raisonne par l'absurde, *i. e.* on suppose que P n'a pas de zéro dans \mathbb{C} . On notera p la fonction polynomiale associée au polynôme P : c'est une fonction de $\mathbb{C} \simeq \mathbb{R}^2$ dans lui-même.

Lorsque $|z|$ tend vers ∞ , il en est de même de $|p(z)|$, vu que

$$p(z) = z^n(1 + p_{n-1}z^{-1} + \dots + p_1z^{1-n} + p_0z^{-n}),$$

où le facteur de droite tend vers 1 lorsque $|z| \rightarrow \infty$. Ainsi, il existe $R > 0$ tel que pour $|z| > R$, on a $|p(z)| > |p(0)| + 1$ et la borne inférieure $m_p = \inf_{z \in \mathbb{C}} |p(z)|$, infimum de $|p|$ sur la boule compacte $\{|z| \leq R\}$, est atteinte en un z_0 : d'après l'hypothèse $p(z_0) \neq 0$ et donc $m_p > 0$.

Lemme 3.1. — *L'application p est ouverte.*

Admettons un instant le lemme : l'image $p(\mathbb{C})$ est incluse dans $\{|z| \geq m_p\}$ et $p(z_0)$ avec $|p(z_0)| = m_p$ ne peut avoir de voisinage ouvert inclus dans $p(\mathbb{C})$, ce qui est contradictoire avec le caractère ouvert de $p(\mathbb{C})$. \square

Preuve du lemme. — Il s'agit de montrer que pour tout complexe z , l'image d'un voisinage de z par p est un voisinage de $p(z)$. Remarquons tout d'abord que la formule de Taylor (algébrique)

$$P(Z + H) = P(Z) + P'(Z)H + \frac{P''(Z)}{2!}H^2 + \dots + \frac{P^{(n)}(Z)}{n!}H^n$$

permet d'écrire, pour $z, h \in \mathbb{C}$ et en notant $p^{(k)}$ la fonction polynomiale sur \mathbb{C} associé au polynôme dérivé $P^{(k)}$,

$$(9) \quad \begin{aligned} p(z + h) &= p(z) + p'(z)h + \frac{p''(z)}{2!}h^2 + \dots + \frac{p^{(n)}(z)}{n!}h^n \\ &= p(z) + p'(z)h + o(h), \quad h \rightarrow 0, \end{aligned}$$

Ainsi l'application p est une application différentiable avec comme différentielle la multiplication par $p'(z)$. Ainsi, si $p'(z)$ est non nul, l'application linéaire de multiplication par $p'(z)$ est un isomorphisme et donc p est un difféomorphisme local d'un voisinage de z sur un voisinage de $p(z)$: l'application est donc ouverte sur le complémentaire des zéros de P' .

En un zéro z de P' , il existe un entier n_z non nul tel que $p^{(n_z)}(z)$ soit non nul (sinon p serait l'application constante). Pour N le plus petit de ces entiers, on a une expression analogue à (9)

$$p(z + h) = p(z) + \frac{p^{(N)}(z)}{N!}h^N q_{z,N}(h)$$

avec

$$q_{z,N}(h) = 1 + \frac{N!p^{(N+1)}(z)}{(N+1)!p^{(N)}(z)}h + \dots + \frac{N!p^{(n)}(z)}{n!p^{(N)}(z)}h^{n-N}.$$

Alors, si $f_{z,N}$ est la fonction définie par

$$f_{z,N}(h, y) = y^N - q_{z,N}(h) \in \mathbb{C} \simeq \mathbb{R}^2, \quad (h, y) \in \mathbb{C}^2 \simeq \mathbb{R}^4,$$

sa différentielle partielle $D_y f_{z,N}(h, y)$ est, comme précédemment, l'application linéaire induite par la multiplication par Ny^{N-1} , non nulle en $(h_0, y_0) = (0, 1)$. D'après le théorème des fonctions implicites, il existe une fonction différentiable $\varphi_{z,N}$, définie sur un voisinage de

$h_0 = 0$ dans \mathbb{C} et à valeurs dans \mathbb{C} , telle que $\varphi_{z,N}(h)^N = q_{z,N}(h)$ et $\varphi_{z,N}(0) = 1$. Ainsi, si $r_{z,N}$ est une racine N ième de $P^{(N)}(z)/N!$, on a pour h petit

$$p(z+h) = p(z) + [r_{z,N}\varphi_{z,N}(h)h]^N.$$

L'application $h \rightarrow r_{z,N}\varphi_{z,N}(h)h$ est inversible au voisinage de $h_0 = 0$ puisque sa différentielle en $h_0 = 0$ est la multiplication par le complexe non nul $r_{z,N}$. Ainsi, au voisinage d'un zéro de P' , l'application p est la composée d'une application inversible et de la fonction puissance $\zeta \rightarrow \zeta^N$, qui est ouverte : l'application p envoie un voisinage du zéro z de P' sur un voisinage de $p(z)$. On a donc terminé de montrer que l'application p est ouverte sur \mathbb{C} . \square

CHAPITRE 4

COURBES ET SURFACES RÉGULIÈRES

4.1. Équations régulières

La notion d'équation régulière est la transcription géométrique de la notion de point critique introduite dans la définition 2.4.

Définition 4.1. — Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^1 . L'équation $f = 0$ est dite *régulière*, et le lieu $\{f = 0\}$ *régulier*, si la fonction f est non critique en tout point m vérifiant $f(m) = 0$.

▷ **Exemple 4.1.** — L'équation $x^2 + y^2 + z^2 - 1 = 0$ est régulière, même si la fonction $f(x, y, z) = x^2 + y^2 + z^2 - 1$ a l'origine comme point critique. \triangleleft

Soit f une fonction à valeurs numériques définie sur un ouvert de \mathbb{R}^n telle que $f = 0$ soit régulière et m_0 tel que $f(m_0) = 0$. En numérotant convenablement les coordonnées, on peut supposer $\frac{\partial f}{\partial x_n}(m_0)$ non nul. Ainsi, d'après le théorème des fonctions implicites, si $m_0 = (a_1, a_2, \dots, a_n)$ il existe une boule B centrée en $(a_1, a_2, \dots, a_{n-1})$, un $\delta > 0$ et une application $\varphi : B \rightarrow (a_n - \delta, a_n + \delta)$ tel que $\frac{\partial f}{\partial x_n}(m)$ soit non nul sur $B \times (a_n - \delta, a_n + \delta)$ et que, pour $m = (x_1, x_2, \dots, x_n) \in B \times (a_n - \delta, a_n + \delta)$, l'équation $f(m) = 0$ soit équivalente à $x_n = \varphi(x_1, x_2, \dots, x_{n-1})$.

4.2. Courbes régulières

Définition 4.2. — Une partie C du plan \mathbb{R}^2 est dite *courbe régulière* si pour tout point m_0 de C , il existe un voisinage ouvert U_0 de m_0 et une fonction $f : U_0 \rightarrow \mathbb{R}$ de classe C^1 telle que $C \cap U_0 = \{m \in U_0, f(m) = 0\}$ et l'équation $f(m) = 0, m \in U_0$ soit régulière.

Le théorème des fonctions implicites énonce que toute courbe régulière est localement un graphe $\{(x, \varphi(x))\}$ ou $\{(\psi(y), y)\}$.

▷ **Exemples 4.2.** —

1. On a étudié dans l'exemple 3.2 la cubique d'équation $y^3/3 - y - x = 0$.

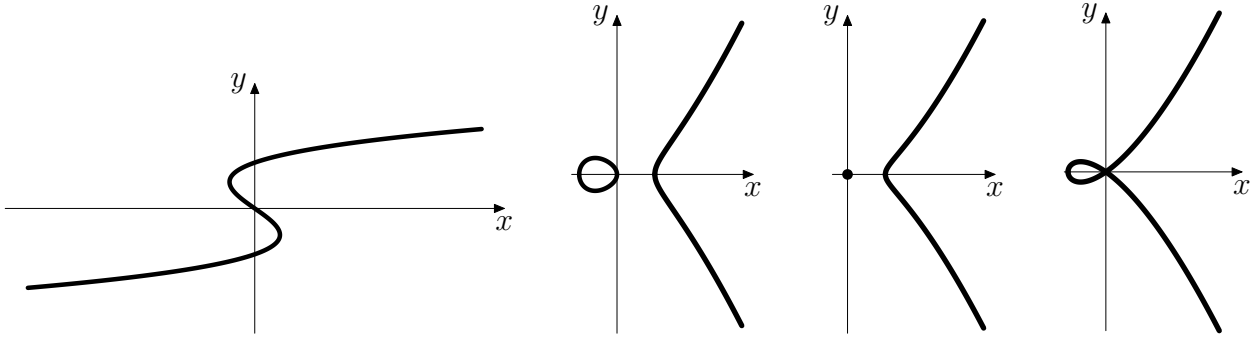


FIGURE 3. Les cubiques des exemples 4.2.

2. La cubique d'équation $y^2 = x(x^2 - 1)$ est une courbe régulière : elle a deux composantes connexes, une bornée correspondant à $x \in [-1, 0]$, l'autre non bornée où $x \in [1, \infty)$.
3. La cubique d'équation $y^2 = x^2(x - 1)$ est une courbe régulière, sauf au point $(0, 0)$ qui est un point isolé de la courbe.
4. La cubique d'équation $y^2 = x^2(x + 1)$ est une courbe régulière, sauf au point $(0, 0)$ qui est un point *double* de la courbe : par le difféomorphisme $(x, y) \rightarrow (X, Y) = (x\sqrt{1+x}, y)$ défini au voisinage de $(0, 0)$, la courbe au voisinage de $(0, 0)$ est homéomorphe à l'union de deux segments des deux droites $X = \pm Y$. \triangleleft

Définition 4.3. — Soit C une courbe. La droite Δ passant par m_0 est dite tangente à C en $m_0 \in C$ si elle est obtenue comme limite de droites $\Delta_{m_0 n}$ passant par m_0 et $n \in C$ lorsque $n \rightarrow m_0$.

\triangle Si $\gamma : t \in (a, b) \rightarrow \gamma(t) = (x(t), y(t)) \in \mathbb{R}^2$ est un arc paramétré de classe C^1 avec $\gamma'(t)$ ne s'annulant pas sur (a, b) , la courbe $\gamma(a, b)$ a comme tangente en $m_0 = \gamma(t_0)$ la droite de vecteur directeur $\gamma'(t_0)$ passant par m_0 . En effet, la droite $\Delta_{\gamma(t)\gamma(t_0)}$ a pour équation

$$\begin{vmatrix} X - x(t_0) & x(t) - x(t_0) \\ Y - y(t_0) & y(t) - y(t_0) \end{vmatrix} = 0$$

soit, pour $t \neq t_0$

$$\begin{vmatrix} X - x(t_0) & \frac{x(t) - x(t_0)}{t - t_0} \\ Y - y(t_0) & \frac{y(t) - y(t_0)}{t - t_0} \end{vmatrix} = 0.$$

En passant à la limite $t \rightarrow t_0$, on obtient donc

$$\begin{vmatrix} X - x(t_0) & x'(t_0) \\ Y - y(t_0) & y'(t_0) \end{vmatrix} = 0,$$

ce qui est l'équation de la droite passant par m_0 et de direction $\gamma'(t_0)$. ∇

Proposition 4.1. — Soit une courbe régulière C d'équation $f = 0$ au voisinage de $m = (x, y)$. La tangente à C au point m a pour équation

$$(10) \quad \frac{\partial f}{\partial x}(m)(X - x) + \frac{\partial f}{\partial y}(m)(Y - y) = 0, \quad (X, Y) \in \mathbb{R}^2.$$

Démonstration. — Si $\frac{\partial f}{\partial y}(m)$ est non nul, le lieu C est localement le graphe $\{(x, \varphi(x)), x \in I\}$ pour une fonction φ . La tangente est la droite passant par $(x, y) = (x, \varphi(x))$ de pente $\varphi'(x)$, elle a donc pour équation $(Y - y)/(X - x) = \varphi'(x)$, soit (10) après avoir remplacé $\varphi'(x)$ par son expression $\varphi'(x) = -\frac{\partial f}{\partial x}(x, \varphi(x))/\frac{\partial f}{\partial y}(x, \varphi(x))$ au point $m = (x, \varphi(x))$.

Si $\frac{\partial f}{\partial x}(m)$ est non nul, le lieu C est localement la courbe paramétrée $\{(\psi(y), y), y \in I\}$ pour une fonction ψ . La tangente $\tau_m C$ au point $m = (x, y)$ de C est la droite passant

par $(x, y) = (\psi(y), y)$ de direction $(\psi'(y), 1)$, elle a donc pour équation $\begin{vmatrix} X - x & \psi'(y) \\ Y - y & 1 \end{vmatrix} = 0$ ou $(X - x) - \psi'(y)(Y - y) = 0$, soit (10) après avoir remplacé $\psi'(y)$ par son expression $\psi'(y) = -\frac{\partial f}{\partial y}(\psi(y), y) / \frac{\partial f}{\partial x}(\psi(y), y)$ au point $m = (\psi(y), y)$. \square

4.3. Surfaces régulières

Définition 4.4. — Une partie S de l'espace \mathbb{R}^3 est dite *surface régulière* si pour tout point m_0 de S , il existe un voisinage ouvert U_0 de m_0 et une fonction $f : U_0 \rightarrow \mathbb{R}$ de classe C^1 telle que $S \cap U_0 = \{m \in U_0, f(m) = 0\}$ et l'équation $f(m) = 0, m \in U_0$ soit régulière.

▷ **Exemples 4.3.** —

1. Soit $g : V(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ une fonction de classe C^1 . Le graphe $\Gamma_g = \{(x, y, g(x, y)), (x, y) \in V\}$ est une surface régulière. Une équation régulière de Γ_g est $z - g(x, y) = 0, (x, y, z) \in V \times \mathbb{R}$. Si g est donnée par $g(x, y) = x^2 - y^2, (x, y) \in \mathbb{R}^2$, le graphe Γ_g est une surface régulière, même si son intersection $\Gamma_g \cap \{z = 0\}$ avec le plan $z = 0$ n'est pas une courbe régulière.
2. La sphère $x^2 + y^2 + z^2 - 1 = 0$, l'hyperboloïde $x^2 - y^2 - z^2 - \varepsilon = 0$ avec $\varepsilon \neq 0$ sont des surfaces régulières, mais pas le cône $x^2 - y^2 - z^2 = 0$. \triangleleft

Théorème 4.1. — Soit S une partie de \mathbb{R}^3 . Sont équivalentes

1. S est une surface régulière.
2. Pour tout $m_0 \in S$, il existe un voisinage U_0 de m_0 tel que $U_0 \cap S$ soit un graphe au-dessus d'un ouvert d'un des plans de coordonnées.
3. Pour tout m_0 de S , il existe un voisinage U_0 et un difféomorphisme C^1 Φ de U_0 sur $\Phi(U_0) \subset \mathbb{R}^3$ tel que $\Phi(S \cap U_0) = (\mathbb{R}^2 \times \{0\}) \cap \Phi(U_0)$.

\triangle Le difféomorphisme local Φ de la troisième description est dit *difféomorphisme de linéarisation* de la surface S au point m_0 . ∇

▷ **Exemple 4.4.** — Soit S la sphère $x^2 + y^2 + z^2 = 1$. Sur le voisinage $U_- = \{x < 0\}$ du point $(-1, 0, 0)$, la sphère S est le graphe de la fonction $g(y, z) = -\sqrt{1 - y^2 - z^2}$ définie sur $D_1 = \{(y, z), y^2 + z^2 < 1\}$. Le difféomorphisme Φ défini par $\Phi(x, y, z) = (x + \sqrt{1 - y^2 - z^2}, y, z)$ défini sur $U = \mathbb{R} \times D_1$ applique $S \cap U$ sur $(\{0\} \times \mathbb{R}^2) \cap \Phi(U)$. \triangleleft

Démonstration. — La définition de surface régulière et le théorème des fonctions implicites donnent immédiatement l'implication de (2) à partir de (1). Si $\Gamma_g = \{(x, y, g(x, y)), (x, y) \in V\}$ est une présentation locale de S sous forme de graphe, alors le difféomorphisme Φ défini sur $U = V \times \mathbb{R}$ par $\Phi(x, y, z) = (x, y, z - g(x, y))$ répond aux spécifications de (3). Si l'assertion (3) est vérifiée pour Φ d'applications coordonnées Φ_1, Φ_2, Φ_3 , alors $\Phi_3 = 0$ est une équation de S , régulière puisque Φ étant un difféomorphisme, $D\Phi$ est inversible et $D\Phi_3$ est non nulle, i. e. Φ_3 est non critique sur S . \square

Pour Σ une partie de \mathbb{R}^3 et $m_0 \in \Sigma$, on note par $\mathcal{C}_{m_0}(\Sigma)$ l'ensemble des arcs de courbe γ définis comme application de $(-\varepsilon, \varepsilon)$ (pour un $\varepsilon > 0$ dépendant de γ) dans \mathbb{R}^3 de classe C^1 telle que $\gamma(0) = m_0$ et $\gamma(-\varepsilon, \varepsilon) \subset \Sigma$.

Définition 4.5. — Soit Σ une partie de \mathbb{R}^3 . L'espace tangent $T_{m_0}\Sigma$ à Σ est l'ensemble des vecteurs dérivés $\dot{\gamma}(0)$ où γ est un arc de $\mathcal{C}_{m_0}(\Sigma)$.

\triangle L'ensemble $T_{m_0}\Sigma$ est une réunion de droites limite $\lim_{m \rightarrow m_0} \Delta_{m_0 m}$. ∇

Théorème 4.2. — Si S est une surface régulière d'équation régulière $f = 0$ et m_0 un point de S , l'espace tangent $T_{m_0}S$ est le plan vectoriel $\text{Ker } Df(m_0)$. Le plan $T_{m_0}S$ a pour équation

$$(11) \quad \frac{\partial f}{\partial x}(m_0)X + \frac{\partial f}{\partial y}(m_0)Y + \frac{\partial f}{\partial z}(m_0)Z = 0, \quad (X, Y, Z) \in \mathbb{R}^3.$$

Démonstration. — Si $\gamma \in \mathcal{C}_{m_0}(S)$, alors $f(\gamma(t)) = 0$ pour t voisin de 0. La dérivée $\frac{d}{dt}f(\gamma(t)) = Df(\gamma(t))\dot{\gamma}(t)$ est nulle, soit pour $t = 0$, $Df(m_0)\dot{\gamma}(0) = 0$ et donc l'inclusion $T_{m_0}(S) \subset \text{Ker } Df(m_0)$.

Soit par ailleurs Φ un difféomorphisme de linéarisation sur un voisinage U_0 de m_0 : $\Phi(U_0 \cap S) = P \cap \Phi(U_0)$ pour un plan P . Pour $v \in P$ et ε assez petit, l'image de la portion de droite $\Phi(m_0) + tv, t \in (-\varepsilon, \varepsilon)$ par Φ^{-1} est une courbe de $\mathcal{C}_{m_0}(S)$: $\gamma : t \in (-\varepsilon, \varepsilon) \rightarrow \Phi^{-1}(\Phi(m_0) + tv)$ et donc le vecteur $\dot{\gamma}(0) = D\Phi^{-1}(\Phi(m_0))(v)$ est dans $T_{m_0}S$. Vu que Φ^{-1} est un difféomorphisme, $D\Phi^{-1}(\Phi(m_0))$ est un isomorphisme linéaire et $D\Phi^{-1}(\Phi(m_0))(P)$ est un plan vectoriel. Ainsi, l'espace tangent $T_{m_0}S$ est inclus dans le plan vectoriel $\text{Ker } Df(m_0)$ en même temps qu'il contient le plan $D\Phi^{-1}(\Phi(m_0))(P)$: ces deux plans sont égaux et $T_{m_0}S$ est le sous-espace vectoriel $\text{Ker } Df(m_0)$, dont l'équation est bien (11). \square

\triangleright **Exemple 4.5.** — Le cône $C = \{x^2 + y^2 - z^2\}$ est une surface régulière en dehors de l'origine $(0, 0, 0)$. L'espace tangent à ce cône au point $m = (x, y, z) \neq (0, 0, 0)$ est le plan d'équation $xX + yY - zZ = 0$.

La partie $T_{(0,0,0)}C$ contient les vecteurs $v_{\pm} = (0, \pm 1, 1)$, vecteurs dérivés en $t = 0$ des droites $v(t) = (0, \pm t, t)$ incluses dans C . Mais le vecteur $v_+ + v_- = (0, 0, 2)$ n'est pas dans $T_{(0,0,0)}C$. En effet, si c'était le cas, il existerait une courbe γ telle que $\gamma(t) = t(v_+ + v_-) + t\varepsilon(t)$ au voisinage de $t = 0$ incluse dans C : $t^2\varepsilon_x(t)^2 + t^2\varepsilon_y(t)^2 - (2t + t\varepsilon_z(t))^2 = 0$ soit $\varepsilon_x(t)^2 + \varepsilon_y(t)^2 - (2 + \varepsilon_z(t))^2 = 0$ pour $t \neq 0$ et en passant à la limite $-4 = 0$, ce qui n'est pas. On peut aussi remarquer que, w étant le vecteur $w = (1, 0, 1)$ qui est dans $T_{(0,0,0)}C$ comme dérivée en $t = 0$ de la droite paramétrée $t \rightarrow (t, 0, t)$ contenue dans C , les vecteurs v_+, v_-, w , linéairement indépendants, engendreraient un espace vectoriel tangent de dimension 3... \triangleleft

4.4. Gradient

On considère \mathbb{R}^n muni de sa structure euclidienne standard : le produit scalaire $\langle \cdot, \cdot \rangle$ est donné par $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ et la base canonique est orthonormée.

Définition 4.6. — Soit f différentiable sur $U \subset \mathbb{R}^n$ et m un point de U . Le *gradient* de f en m , noté $\text{grad } f(m)$ ou $\nabla f(m)$, est le vecteur représentant la différentielle $Df(m)$, *i. e.* tel que

$$\langle \text{grad } f(m), \xi \rangle = \langle \nabla f(m), \xi \rangle = Df(m)(\xi), \quad \xi \in \mathbb{R}^n$$

Ainsi, en dimension $n = 3$, le gradient $\nabla f(m)$ est donné par

$$\text{grad } f(m) = \nabla f(m) = \left(\frac{\partial f}{\partial x}(m), \frac{\partial f}{\partial y}(m), \frac{\partial f}{\partial z}(m) \right).$$

L'équation du plan tangent $T_{m_0}S$ d'une surface d'équation régulière $f = 0$ est

$$\langle \text{grad } f(m_0), v \rangle = 0, \quad v \in \mathbb{R}^3.$$

et l'espace tangent $\tau_{m_0}S$ passant par m_0 défini comme l'espace affine contenant m_0 de direction $T_{m_0}S$ a pour équation

$$\langle \text{grad } f(m_0), m - m_0 \rangle = 0, \quad m \in \mathbb{R}^3.$$

Ces espaces tangents ont pour vecteur orthogonal le gradient $\nabla f(m)$, de même que l'espace tangent $T_m C$ à la courbe plane C d'équation locale $f = 0$ est l'orthogonal de $\nabla f(m)$.

CHAPITRE 5

APPLICATIONS C^k

Pour k entier, la définition 1.6 d'une fonction de classe C^k se fait par récurrence sur l'entier k , comme se feront les démonstrations des versions C^k des théorèmes établis précédemment en régularité C^1 . Commençons par un lemme.

Lemme 5.1. — Soit U un ouvert de \mathbb{R}^n . Une fonction $f : U \rightarrow \mathbb{R}^p$ est de classe C^k (resp. C^∞) si et seulement si ses fonctions coordonnées f_i , $i = 1, \dots, p$ le sont.

Démonstration. — On le démontre par récurrence avec la propriété de récurrence

\mathcal{P}_k : une fonction $f : x \in U \subset \mathbb{R}^n \rightarrow f(x) = (f_1(x), \dots, f_p(x)) \in \mathbb{R}^p$ est de classe C^k si et seulement si ses fonctions coordonnées $f_1, \dots, f_p : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ le sont.

Pour $k = 1$, c'est la remarque précédant l'exemple 1.3. Supposons la propriété de récurrence montrée au rang k . Si $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ est de classe C^{k+1} , alors f est différentiable et sa différentielle $Df : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^p) \simeq \mathbb{R}^{np}$ est de classe C^k . Les fonctions coordonnées $\frac{\partial f_i}{\partial x_j}$, $i = 1, \dots, p$, $j = 1, \dots, n$ de Df sont de classe C^k d'après l'hypothèse de récurrence au rang k . Par suite, pour $i = 1, \dots, p$, la différentielle Df_i , dont les fonctions coordonnées $\frac{\partial f_i}{\partial x_j}$, $j = 1, \dots, n$ sont de classe C^k , est de classe C^k d'après l'hypothèse de récurrence au rang k (appliquée cette fois-ci de droite à gauche). Ainsi les fonctions coordonnées f_i , $i = 1, \dots, p$ de f sont de classe C^{k+1} . Pour la réciproque, on reprend l'argumentation précédente en partant de la fin, en remarquant que tous ses arguments sont en fait des équivalences. \square

▷ **Exemples 5.1.** —

1. Soit M_{pq} l'espace des matrices à coefficients réels à p colonnes et q lignes. Pour des entiers n, p, q , la multiplication matricielle

$$P_{npq} : (A, B) \in M_{np} \times M_{pq} \rightarrow BA \in M_{nq}.$$

est de classe C^∞ vu que ses fonctions coordonnées sont polynomiales (de degré 2!), avec différentielle

$$DP_{npq}(A, B)(H, K) = B \circ H + K \circ A, \quad (H, K) \in M_{np} \times M_{pq}.$$

La différentielle $(A, B) \in M_{np} \times M_{pq} \rightarrow DP_{npq}(A, B) \in \mathcal{L}(M_{np} \times M_{pq}, M_{nq})$ est linéaire : la différentielle seconde $D(DP_{npq})$ est donc constante

$$D(DP_{npq})(A, B) : (H, K) \in M_{np} \times M_{pq} \rightarrow DP_{npq}(H, K) \in \mathcal{L}(M_{np} \times M_{pq}, M_{nq})$$

avec

$$DP_{npq}(H, K) : (H', K') \in M_{np} \times M_{pq} \rightarrow K \circ H' + K' \circ H \in M_{pq}$$

et les différentielles d'ordre au moins 3 nulles.

2. Soient E, F, G espaces vectoriels de dimension finie. La composition d'applications linéaires

$$P_{EFG} : (A, B) \in \mathcal{L}(E, F) \times \mathcal{L}(F, G) \rightarrow B \circ A \in \mathcal{L}(E, G)$$

est différentiable, avec différentielle au point (A, B)

$$DP_{EFG}(A, B)(H, K) = B \circ H + K \circ A, \quad (H, K) \in \mathcal{L}(E, F) \times \mathcal{L}(F, G).$$

Il suffit de reprendre la définition 1.2 de la différentiabilité en écrivant pour $(H, K) \in \mathcal{L}(E, F) \times \mathcal{L}(F, G)$

$$P_{EFG}(A + H, B + K) = P_{EFG}(A, B) + B \circ H + K \circ A + K \circ H,$$

où $K \circ H = o((H, K))$. La différentielle DP_{EFG} est une application linéaire qui est C^∞ : il en est de même pour P_{EFG} , pareillement à la multiplication matricielle P_{npq} de l'exemple précédent.

En fait, si \mathbf{b} est une base d'un espace vectoriel V et \mathbf{b}' une base d'un espace vectoriel V' , notant par $M_{\mathbf{b}}^{\mathbf{b}'} : \mathcal{L}(V, V') \rightarrow M_{\dim V \dim V'}$ l'isomorphisme de représentation matricielle des morphismes linéaires de V dans V' relativement aux bases \mathbf{b}, \mathbf{b}' on a, avec $\mathbf{b}_E, \mathbf{b}_F, \mathbf{b}_G$ des bases de E, F, G de dimensions respectives n, p, q ,

$$P_{npq} = M_{\mathbf{b}_E}^{\mathbf{b}_G} \circ P_{EFG} \circ (M_{\mathbf{b}_E}^{\mathbf{b}_F} \times M_{\mathbf{b}_F}^{\mathbf{b}_G})$$

soit au niveau des différentielles, d'après la règle de composition des fonctions différentiables,

$$DP_{npq} = M_{\mathbf{b}_E}^{\mathbf{b}_G} \circ DP_{EFG} \circ (M_{\mathbf{b}_E}^{\mathbf{b}_F} \times M_{\mathbf{b}_F}^{\mathbf{b}_G})^{-1}$$

Les remarques de l'exemple précédent sur les différentielles d'ordre supérieur de P_{npq} valent *mutatis mutandis* pour P_{EFG} , soit par examen directe soit en utilisant les relations précédentes. \triangleleft

Corollaire 5.1. — Soient P, Q polynômes à n variables et U_Q l'ouvert $U_Q = Q^{-1}(\mathbb{R}^*)$. Alors l'application $R : (x_1, \dots, x_n) \in U_Q \rightarrow P(x_1, \dots, x_n)/Q(x_1, \dots, x_n)$ est de classe C^∞ .

Démonstration. — La fraction rationnelle $R = P/Q$ est différentiable et ses dérivées partielles $\partial_{x_i} R = (\partial_{x_i} P Q - \partial_{x_i} Q P)/Q^2$, fonctions coordonnées de la différentielle DR , sont des fractions rationnelles bien définies sur U_Q . On fait alors une récurrence sur k . \square

\triangleright **Exemple 5.2.** — L'application $A \in \text{GL}(n, \mathbb{R}) \rightarrow A^{-1} \in M_n$ a des fonctions coordonnées rationnelles comme il résulte de l'expression de l'inverse d'une matrice par sa comatrice et son déterminant : elle est donc C^∞ . \triangleleft

5.1. La version C^k des théorèmes classiques

Théorème 5.1. — Si $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $g : V \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$, avec $f(U) \subset V$, sont de classe C^k (resp. C^∞), la composée $g \circ f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^q$ l'est aussi.

Démonstration. — Une récurrence démarre par le cas C^1 : c'est le théorème 1.1 de composition de fonctions différentiables. Supposons la proposition vraie en rang k et soient f et g de classe C^{k+1} . La différentielle de $g \circ f$ est $[Dg \circ f] \circ Df$: d'après l'hypothèse de récurrence, $Dg \circ f$ est de classe C^k , comme composée de deux fonctions de classe C^k , ainsi que

$$D(g \circ f) = DP_{\mathbb{R}^n \mathbb{R}^p \mathbb{R}^q}(Df, Dg \circ f)$$

vu comme composée des applications précédentes et de l'application de composition $P_{\mathbb{R}^n \mathbb{R}^p \mathbb{R}^q} : (A, B) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \rightarrow B \circ A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^q)$. Ainsi $D(g \circ f)$ est de classe C^k , et $g \circ f$ de classe C^{k+1} . \square

Théorème 5.2. — Si f est C^k (resp. C^∞) et f est un difféomorphisme C^1 , alors f^{-1} est de classe C^k (resp. C^∞).

Démonstration. — Pour la régularité C^1 , c'est le théorème 2.1 d'inversion locale. On passe de C^k à C^{k+1} via la formule de l'inverse $D(f^{-1}) = [Df \circ f^{-1}]^{-1}$, le théorème de composition C^k et le caractère C^∞ de l'application $A \in \text{GL}(n, \mathbb{R}) \rightarrow A^{-1} \in M_n$ de l'exemple 5.2. \square

5.2. Théorème de Schwarz

Théorème 5.3 (Schwarz). — Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 . Alors, pour $i \neq j$, on a l'égalité

$$\partial_{x_i}(\partial_{x_j} f)(m) = \partial_{x_j}(\partial_{x_i} f)(m), \quad m \in U.$$

Plus généralement, si f est de classe C^k , l'ordre des dérivées partielles n'importe pas pour les dérivées d'ordre au plus k : $\partial_{x_{i_1}} \partial_{x_{i_2}} \dots \partial_{x_{i_k}} f(m) = \partial_{x_{i_{\sigma(1)}}} \partial_{x_{i_{\sigma(2)}}} \dots \partial_{x_{i_{\sigma(k)}}} f(m)$ où σ est une permutation de $\{1, \dots, k\}$ et i une fonction de $\{1, \dots, k\}$ dans $\{1, \dots, n\}$.

Démonstration. — Il suffit de le montrer pour une fonction f de deux variables x, y différentiable à l'origine $(0, 0)$. Soit F définie au voisinage de l'origine par

$$F(x, y) = f(x, y) - f(x, 0) - f(0, y) + f(0, 0).$$

La différentiabilité de $\partial_x F$ à l'origine permet d'écrire

$$(12) \quad \partial_x F(x, y) = \partial_x F(0, 0) + \partial_x \partial_x F(0, 0)x + \partial_y \partial_x F(0, 0)y + (|x| + |y|)\varepsilon(x, y).$$

On a

$$\begin{aligned} \partial_x F(x, y) &= \partial_x f(x, y) - \partial_x f(x, 0), \\ \partial_{x^2}^2 F(x, y) &= \partial_{x^2}^2 f(x, y) - \partial_{x^2}^2 f(x, 0), \\ \partial_y \partial_x F(x, y) &= \partial_y \partial_x f(x, y). \end{aligned}$$

Ainsi la formule (12) se réduit à

$$\partial_x F(x, y) = y \partial_y \partial_x f(0, 0) + (|x| + |y|)\varepsilon(x, y)$$

On en déduit par l'inégalité des accroissements finis avec y fixé, et $\varepsilon_1(x, y) = \sup_{s \in [0, x]} |\varepsilon(s, y)|$,

$$|F(x, y) - xy \partial_y \partial_x f(0, 0)| \leq \sup_{s \in [0, x]} |\partial_x F(s, y) - y \partial_y \partial_x F(0, 0)| |x| \leq (|x| + |y|)\varepsilon_1(x, y)|x|$$

soit

$$|F(x, y) - xy \partial_y \partial_x f(0, 0)| \leq (|x| + |y|)|x|\varepsilon_1(x, y).$$

On a une inégalité analogue en échangeant x et y

$$|F(x, y) - yx \partial_x \partial_y f(0, 0)| \leq (|x| + |y|)|y|\varepsilon_2(x, y),$$

d'où en additionnant, et en prenant $x = y = t$,

$$|t^2 \partial_x \partial_y f(0, 0) - t^2 \partial_y \partial_x f(0, 0)| \leq 2t^2(\varepsilon_1(t, t) + \varepsilon_2(t, t))$$

soit, pour $t \neq 0$,

$$|\partial_x \partial_y f(0, 0) - \partial_y \partial_x f(0, 0)| \leq 2(\varepsilon_1(t, t) + \varepsilon_2(t, t)),$$

et donc l'annulation du membre de gauche en faisant tendre t vers 0, ce qu'il fallait démontrer. \square

\triangleright **Exemple 5.3.** — Soit f la fonction continue sur \mathbb{R}^2 telle que

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}, \quad (x, y) \neq (0, 0).$$

La fonction f admet des dérivées partielles continues sur \mathbb{R}^2

$$\frac{\partial f}{\partial x}(0, y) = -y, \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1, \quad \frac{\partial f}{\partial y}(x, 0) = x, \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1.$$

Ainsi la différentielle Df n'est pas différentiable en $(0, 0)$. \triangleleft

△ Le théorème de Schwarz vaut évidemment pour les fonctions à valeurs dans \mathbb{R}^p . D'après la définition 1.6, l'application $D^2f = D(Df)$ est une application définie sur U dont la valeur $D^2f(m)$, $m \in U$ est une application linéaire de $\mathbb{R}^n \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$, qu'on peut identifier à une application bilinéaire $\widetilde{D^2f}(m)$ de $\mathbb{R}^n \times \mathbb{R}^n$ dans \mathbb{R}^p , via

$$D^2f(m)(u)(v) = \widetilde{D^2f}(m)(u, v) \quad m \in U, u, v \in \mathbb{R}^n.$$

On conviendra de noter pareillement la différentielle D^2f et la fonction $\widetilde{D^2f}$. Le théorème de Schwarz 5.3 énonce simplement le fait que l'application bilinéaire $\widetilde{D^2f}(m)$ est symétrique.

▽

Relativement à la base canonique de \mathbb{R}^n et pour f une fonction numérique, la matrice la forme bilinéaire $D^2f(m)$ s'identifie à la matrice, dite *hessienne* de la fonction f et notée $\text{Hess } f(m)$, des dérivées partielles de f d'ordre 2 relativement à la base canonique (e_1, \dots, e_n)

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

avec

$$D^2f(m)(u, v) = {}^tU \text{Hess } f(m)V$$

où U et V sont les vecteurs des coordonnées de u et v relativement à la base canonique. Le théorème de Schwarz 5.3 n'a d'autre conclusion que le caractère symétrique de la matrice $\text{Hess } f(m)$.

Définition 5.1. — Soit f définie sur un ouvert U de \mathbb{R}^n , à valeurs réelles et de classe C^2 . La *hessienne* $\text{Hess } f$ de f est la fonction sur U à valeurs forme bilinéaire symétrique définie par

$$\text{Hess } f(m)(u, v) = D^2f(m)(u)(v), \quad m \in U, u, v \in \mathbb{R}^n.$$

On désignera parfois par la même notation la forme quadratique associée : $\text{Hess } f(m)(u) = D^2f(m)(u)(u)$.

La hessienne $\text{Hess } f(m)$ est une forme bilinéaire symétrique, dont la matrice relativement à la base canonique (e_1, \dots, e_n) de \mathbb{R}^n est $(\partial_{x_i x_j}^2 f(m))$, vu que $\text{Hess } f(m)(e_i)(e_j) = \partial_{x_i x_j}^2 f(m)$: la dérivée partielle $\partial_{x_i} f(m)$ est exactement la dérivée directionnelle $Df(m)(e_i)$.

5.3. Formules de Taylor

Théorème 5.4 (Formule de Taylor à l'ordre 2). — Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 . Alors

$$f(m_0 + h) = f(m_0) + Df(m_0)h + \frac{\text{Hess } f(m_0)(h)}{2} + o(\|h\|^2)$$

Démonstration. — La fonction g d'une variable définie par $g(t) = f(m_0 + th)$ pour t dans un voisinage de $[0, 1]$ a pour dérivées d'ordre au plus 2

$$g'(t) = Df(m_0 + th)(h), \quad g''(t) = D^2f(m_0 + th)(h)(h) = \text{Hess } f(m_0 + th)(h).$$

Ainsi

$$\begin{aligned}
 f(m_0 + h) &= g(1) = g(0) + \int_0^1 g'(s) ds = g(0) + g'(0) + \int_0^1 (1-s)g''(s) ds \\
 &= g(0) + g'(0) + \frac{g''(0)}{2} + \int_0^1 (1-s)[g''(s) - g''(0)] ds \\
 &= f(m_0) + Df(m_0)h + \frac{\text{Hess } f(m_0)(h)}{2} + \int_0^1 (1-s)[g''(s) - g''(0)] ds \\
 &= f(m_0) + Df(m_0)h + \frac{\text{Hess } f(m_0)(h)}{2} + o(\|h\|^2).
 \end{aligned}$$

On aura utilisé dans la dernière égalité

$$\begin{aligned}
 |g''(s) - g''(0)| &= |\text{Hess } f(m_0 + sh)(h) - \text{Hess } f(m_0)(h)| \\
 &\leq \sup_{s \in [0,1]} \|\text{Hess } f(m_0 + sh) - \text{Hess } f(m_0)\| \|h\|^2. \quad \square
 \end{aligned}$$

\triangle On a des formules de Taylor à l'ordre k : si f est de classe C^k au voisinage de m_0

$$f(m_0 + h) = f(m_0) + Df(m_0)h + \frac{D^2 f(m_0)(h, h)}{2} + \dots + \frac{D^k f(m_0)(h, h, \dots, h)}{k!} + o(\|h\|^k).$$

On aura identifié $D^k f$ à une forme k -linéaire sur \mathbb{R}^n , symétrique d'après le théorème de Schwarz, avec en coordonnées locales pour $h = (h_1, \dots, h_n)$

$$D^k f(h, \dots, h) = \sum_{(i_1, i_2, \dots, i_k) \in \mathbb{N}_*^k} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} h_{i_1} \dots h_{i_k}.$$

Par la suite, seul le développement limité à l'ordre 2 sera utilisé, c'est pourquoi les dérivées $D^k f$ pour $k \geq 3$ n'ont pas été étudiées. ∇

5.4. Approximation de l'espace tangent

Commençons par considérer le cas particulier d'une surface $\Sigma = \{(x, y, z(x, y)), (x, y) \in U\}$ où U est un voisinage de l'origine $(0, 0)$ dans \mathbb{R}^2 , $z : U \rightarrow \mathbb{R}$ de classe $C^{k \geq 2}$ avec $Dz(0, 0) = 0$. Le plan tangent à Σ en $m_0 = (0, 0, z(0, 0))$ est le plan horizontal $Z = 0$. Le développement de Taylor à l'ordre 2 au voisinage de $(0, 0)$

$$(13) \quad z(x, y) = z(0, 0) + \frac{1}{2} \text{Hess } z(0, 0)(x, y) + o(x^2 + y^2),$$

permet de préciser la position de Σ relativement au plan tangent $\tau_{m_0} \Sigma$:

- si la hessienne $\text{Hess } z(0, 0)$ est définie positive (resp. négative), alors, au voisinage de m_0 , la surface $\Sigma \setminus \{m_0\}$ est au-dessus (resp. au-dessous) du plan tangent, *i. e.* toute entière contenue dans le demi-espace $Z > 0$ (resp. $Z < 0$). En effet, dans ce cas, il existe $C > 0$ tel que $|\text{Hess } z(0, 0)(x, y)| > C(x^2 + y^2)$ pour (x, y) non nul, ainsi peut-on écrire le dernier terme du développement (13) sous la forme $o(x^2 + y^2) = \frac{1}{2} \text{Hess } z(0, 0)(x, y)\varepsilon(x, y)$ avec $\varepsilon(x, y) \rightarrow 0$ lorsque $(x, y) \rightarrow (0, 0)$ et par suite

$$z(x, y) - z(0, 0) = \frac{1}{2} \text{Hess } z(0, 0)(x, y)(1 + \varepsilon(x, y))$$

- si la hessienne $\text{Hess } z(0, 0)$ est non dégénérée et non définie, alors la surface Σ_0 d'équation $z = z(0, 0) + \text{Hess } z(0, 0)(x, y)$ intersecte le plan tangent $\tau_{m_0} \Sigma$ suivant deux droites $D_{\pm}(Z_0)$ et est de part et d'autre du plan tangent ; on démontre que la surface Σ intersecte son plan tangent suivant deux courbes tangentes aux droites, $D_{\pm}(Z_0)$.
- si la hessienne est dégénérée, on ne peut rien dire.

En général, soit Σ donnée par un paramétrage $(u, v) \rightarrow P(u, v)$ de classe $C^{k \geq 2}$ et $m_0 = P(u_0, v_0)$. Si N_0 est un vecteur normal au plan des directions tangentes $T_{m_0}\Sigma$, alors le développement à l'ordre 2

$$P(u, v) = m_0 + DP(u_0, v_0)(u - u_0, v - v_0) + \frac{1}{2}D^2P(m_0)(u - u_0, v - v_0) + o((u - u_0)^2 + (v - v_0)^2)$$

donne

$$\langle P(u, v) - m_0, N_0 \rangle = \frac{1}{2} \langle D^2P(m_0)(u - u_0, v - v_0), N_0 \rangle + o((u - u_0)^2 + (v - v_0)^2)$$

et amène à la même discussion que précédemment concernant la position de la surface Σ par rapport à son plan tangent $T_{m_0}\Sigma$ d'équation $\langle M - m_0, N_0 \rangle = 0$. Que les conclusions soient exprimées en terme de la signature de la forme quadratique $\langle D^2P(m_0)(X, Y), N_0 \rangle$, et ce indépendamment du choix du paramétrage, résulte du lemme suivant

Lemme 5.2. — Soit $\tilde{P} : (\tilde{u}, \tilde{v}) \in \tilde{U}(\subset \mathbb{R}^2) \rightarrow \tilde{P}(\tilde{u}, \tilde{v})$ un autre paramétrage de la surface au voisinage de $m_0 = \tilde{P}(\tilde{u}_0, \tilde{v}_0)$. Alors $\langle D^2P(u_0, v_0), N_0 \rangle$ et $\langle D^2\tilde{P}(\tilde{u}_0, \tilde{v}_0), N_0 \rangle$ ont même signature.

Démonstration. — Soit $\varphi : U_0 \rightarrow \tilde{U}_0$ le difféomorphisme tel que $P = \tilde{P} \circ \varphi$. On a alors

$$\begin{aligned} DP &= D\tilde{P} \circ D\varphi, \\ D^2P &= D^2\tilde{P}(D\varphi, D\varphi) + D\tilde{P} \circ D^2\varphi \end{aligned}$$

et donc

$$\langle D^2P(u_0, v_0), N_0 \rangle = \langle D^2\tilde{P}(\tilde{u}_0, \tilde{v}_0)(D\varphi(u_0, v_0), D\varphi(u_0, v_0)), N_0 \rangle$$

ce qui conclut, vu que les formes quadratiques Q et $Q(A, A)$ ont mêmes signature et indice si A est un isomorphisme. \square

CHAPITRE 6

POINTS CRITIQUES ET EXTREMA

6.1. Extrema

Définition 6.1 (maximum/minimum local). — Soit $f : U(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$. Si le point $m_0 \in U$ est un *maximum* (resp. *minimum*) *local* pour f , alors il existe $r > 0$ tel que $f(m_0) \geq f(m)$ (resp. $f(m_0) \leq f(m)$) si $\|m - m_0\| \leq r$. On dira que m_0 est un *extremum* (local) de f si c'est un minimum ou un maximum (local) pour f . L'extremum est dit *strict* si il y a inégalité stricte pour $m \neq m_0$; il est *global* sur U si les inégalités valent sur U .

\triangle Si m_0 est un maximum local de f , m_0 est un minimum local de $-f$. On se limitera au possible à donner les preuves pour m maximum. ∇

Proposition 6.1. — Si f est différentiable en m_0 et m_0 est un extremum local pour f , alors m_0 est un point critique de f .

Démonstration. — Supposons la différentielle $Df(m_0)$ non nulle. Il existe une direction h telle que $Df(m_0)h$ soit non nul. Alors, pour t suffisamment petit,

$$f(m_0 + th) = f(m_0) + tDf(m_0)h + o(t) = f(m_0) + t(Df(m_0)h + \varepsilon(t)),$$

où le dernier terme, équivalent à $tDf(m_0)h$ au voisinage de $t = 0$, change de signe en $t = 0$: m_0 ne peut être un extremum local. \square

Proposition 6.2. — Si f est de classe C^2 et m_0 est un minimum (resp. maximum) local pour f , alors $\text{Hess } f(m_0)$ est positive (resp. négative).

Démonstration. — Soit g_ξ la fonction d'une variable définie au voisinage de $t = 0$ par $g_\xi(t) = f(m_0 + t\xi)$. La formule de Taylor à l'ordre 2 donne l'existence de $\theta_t \in [0, 1]$ telle que

$$g_\xi(t) = g_\xi(0) + tg'_\xi(0) + \frac{t^2}{2}g''_\xi(\theta_t t) = f(m_0) + \frac{t^2}{2}\text{Hess } f(m_0 + \theta_t t)(\xi)$$

d'où en faisant $t \rightarrow 0^+$, $\text{Hess } f(m_0)(\xi) \geq 0$. \square

Les conditions nécessaires d'extrema des deux propositions précédentes ne sont pas suffisantes : une condition suffisante est donnée par la proposition suivante.

Proposition 6.3. — Soit f de classe C^2 , m_0 un point critique de f avec hessienne $\text{Hess } f(m_0)$ de rang r et d'indice i .

Si $r = n$ et $i = n$ (resp. $i = 0$), m_0 est un maximum (resp. minimum) local strict.

Si $0 < i < r$, m_0 n'est pas un extremum.

Si $r < n$ et $i \in \{0, r\}$, on ne peut rien dire.

▷ **Exemple 6.1.** — Les fonctions définies sur \mathbb{R}^n

$$F_0(x, y) = \alpha x^2 + \beta y^2, \quad f_1(x, y) = xy, \quad f_{2,k,\pm}(x, y) = x^2 \pm y^k$$

avec α, β non nuls et $k = 3, 4$, ont $m_0 = (0, 0)$ comme point critique et illustrent les différentes assertions du théorème. En particulier, la fonction $f_{2,k,\pm}$ montre que dans le cas de hessienne dégénérée ($r < n$), on ne peut conclure en général si m_0 est (ou n'est pas) un extremum. \triangleleft

Démonstration. — Soient (v_j, α_j) , $j = 1, \dots, n$ les éléments propres de la hessienne $\text{Hess } f(m_0)$ avec $(v_j)_{j=1, \dots, n}$ une base orthonormée de \mathbb{R}^n : $\text{Hess } f(m_0)(v_j, v_j) = \alpha_j$. On écrit alors le développement limité au voisinage de 0 pour $f(m_0 + v)$ avec $v = \sum_{j=1}^n y_j v_j$

$$\begin{aligned} f\left(m_0 + \sum_{j=1}^n y_j v_j\right) &= f(m_0) + \frac{1}{2}\text{Hess } f(m_0)\left(\sum_{j=1}^n y_j v_j\right) + o(\|v\|^2) \\ &= f(m_0) + \frac{1}{2}\sum_{j=1}^n \alpha_j y_j^2 + o(\|v\|^2). \end{aligned}$$

Supposons la hessienne $\text{Hess } f(m_0)$ non dégénérée. Si elle est définie négative (i. e. $i = n$), soit α le maximum des α_j , $j = 1, \dots, n$, qui est donc strictement négatif. On a, pour $v = \sum_{j=1}^n y_j v_j$,

$$f(m_0 + v) - f(m_0) = \frac{1}{2}\sum_{j=1}^n \alpha_j y_j^2 + o(\|v\|^2) \leq \frac{1}{2}\alpha \sum_{j=1}^n y_j^2 + o(\|v\|^2) = \|v\|^2\left(\frac{\alpha}{2} + \varepsilon(v)\right)$$

et donc $f(m_0 + v) < f(m_0)$ pour v voisin de zéro non nul : m_0 est un maximum local strict de f .

Si $\text{Hess } f(m_0)$ est non dégénérée, de signe non constant, à ordre près on peut supposer $\alpha_1 < 0$ et $\alpha_2 > 0$: f restreinte à la droite passant par m_0 et de direction v_1 (resp. v_2) a un maximum (resp. minimum) local strict en m_0 : m_0 n'est ni maximum, ni minimum local de la fonction f . \square

6.2. Extrema liés

Soit h une fonction définie sur un ouvert U de \mathbb{R}^n et L un lieu géométrique inclus dans U . On s'intéresse à l'étude des extrema (locaux ou globaux) de la restriction de h à L . Si L est régulier (surface dans \mathbb{R}^3 , courbe dans \mathbb{R}^2 ou \mathbb{R}^3) et la fonction h différentiable (ce qui sera supposé dans la suite), le théorème des fonctions implicites permet d'établir une condition nécessaire pour que $m_0 \in C$ soit un extrema local : c'est cette condition qui sera établie dans cette section pour des cas particuliers en dimension $n = 2$ ou 3 .

▷ **Exemple 6.2.** — On considère le cas où L est un sous-espace affine, de direction vectorielle V (sous-espace vectoriel de \mathbb{R}^n). La restriction de h à L est différentiable, avec différentielle $D(h|_L)$ donnée par la restriction $(Dh)|_V$ de Dh à V . Ainsi, si $m_0 \in L$ est un extremum local de $h|_L$, alors on a $(Dh(m_0))|_V = 0$.

Si L est un hyperplan affine d'équation $\ell(m) = \ell(m_0)$ avec ℓ une forme linéaire sur \mathbb{R}^n , la condition précédente $(Dh(m_0))|_V = 0$ est l'inclusion $V = \text{Ker } \ell \subset \text{Ker } Dh(m_0)$: il existe $\lambda \in \mathbb{R}$ tel que $Dh(m_0) = \lambda \ell$. En fait, λ est non nul si $\text{Ker } \ell = \text{Ker } Dh(m_0)$, nul si m_0 est un point critique de h .

Si L est une droite dans \mathbb{R}^3 , intersection de deux hyperplans affines distincts d'équations $\ell_1(m) = \ell_1(m_0)$ et $\ell_2(m) = \ell_2(m_0)$, alors $V = \text{Ker } \ell_1 \cap \text{Ker } \ell_2$ et la condition énonce que tout v annulé par ℓ_1 et ℓ_2 annule $Dh(m_0)$. On en déduit l'existence de réels λ_1, λ_2 tels que $Dh(m_0) = \lambda_1 \ell_1 + \lambda_2 \ell_2$. En effet, si (ℓ_1, ℓ_2, ℓ_3) est une base de formes linéaires complétant les deux formes indépendantes ℓ_1, ℓ_2 , de base duale (v_1, v_2, v_3) , on a $Dh(m_0) = \sum_{j=1}^3 \lambda_j \ell_j$ et $\lambda_3 = Dh(m_0)(v_3)$ est nul puisque $v_3 \in V$.

Définition 6.2. — Le scalaire λ , ou les scalaires λ_1, λ_2 , sont appelés *multiplieur(s) de Lagrange*.

Si on applique la condition précédente à la fonction $h(x, y, z) = x^2 + y^2 + z^2$ sur l'hyperplan $\mathcal{H}_{n,C}$ d'équation $\langle n, m \rangle = C$ avec $n \in \mathbb{R}^3, C \in \mathbb{R}$, on a, en exprimant la condition $(Dh(m_0))|_V = 0$ sur le gradient de h , l'existence de λ tel que $2m_0 = \lambda n$: la distance $\sqrt{h} = d(0, m)$ d'un point m de \mathcal{H} à l'origine 0 est minimale au projeté m_0 de l'origine sur $\mathcal{H}_{n,C}$. ◁

Commençons par traiter le cas des extrema de fonctions restreintes à des surfaces régulières.

Proposition 6.4. — Soit $S = \{f = 0\}$ une surface régulière dans \mathbb{R}^3 . Si le point $m_0 \in S$ est un extrema local de $h|_S$, alors il existe λ tel que $Dh(m_0) = \lambda Df(m_0)$.

Démonstration. — Supposons $D_x f(m_0)$ non nul. Alors, d'après le théorème des fonctions implicites, il existe une fonction φ définie sur un voisinage V_0 de (y_0, z_0) telle que dans un voisinage de m_0 la surface S soit le graphe $\{(\varphi(y, z), y, z), (y, z) \in V_0\}$. La fonction H définie par $H(y, z) = h(\varphi(y, z), y, z), (y, z) \in V_0$ a donc (y_0, z_0) comme point critique : la nullité de la différentielle $DH(y_0, z_0)$ en terme de dérivées partielles donne

$$\frac{\partial h}{\partial x}(m_0) \frac{\partial \varphi}{\partial y}(y_0, z_0) + \frac{\partial h}{\partial y}(m_0) = 0, \quad \frac{\partial h}{\partial x}(m_0) \frac{\partial \varphi}{\partial z}(y_0, z_0) + \frac{\partial h}{\partial z}(m_0) = 0$$

On a par ailleurs, vu $f(\varphi(y, z), y, z) = 0$,

$$\frac{\partial f}{\partial x}(m_0) \frac{\partial \varphi}{\partial y}(y_0, z_0) + \frac{\partial f}{\partial y}(m_0) = 0, \quad \frac{\partial f}{\partial x}(m_0) \frac{\partial \varphi}{\partial z}(y_0, z_0) + \frac{\partial f}{\partial z}(m_0) = 0,$$

et donc

$$\frac{\partial h}{\partial y}(m_0) = \frac{\frac{\partial h}{\partial x}(m_0) \frac{\partial f}{\partial y}(m_0)}{\frac{\partial f}{\partial x}(m_0)}, \quad \frac{\partial h}{\partial z}(m_0) = \frac{\frac{\partial h}{\partial x}(m_0) \frac{\partial f}{\partial z}(m_0)}{\frac{\partial f}{\partial x}(m_0)}.$$

On en déduit $\text{grad } h = \lambda \text{grad } f$ ou $Dh(m_0) = \lambda Df(m_0)$ avec $\lambda = \frac{\partial h}{\partial x}(m_0) / \frac{\partial f}{\partial x}(m_0)$. ◻

On a évidemment une proposition analogue dans le cas d'une courbe dans le plan, dont on laisse la démonstration au lecteur.

Proposition 6.5. — Soit $C = \{f = 0\}$ une courbe régulière dans \mathbb{R}^2 . Le point $m_0 \in C$ est un extrema local de $h|_C$ s'il existe λ tel que $Dh(m_0) = \lambda Df(m_0)$.

▷ **Exemple 6.3.** — Soient a, b, c, d des constantes positives non nulles et f la fonction définie sur $Q_+ = \{(x, y) \in \mathbb{R}^2, x > 0, y > 0\}$ par $f(x, y) = cx - d \log x + by - a \log y$. Vu que les constantes a, b, c, d sont positives, la valeur $f(m)$ tend vers $+\infty$ lorsque m tend vers le bord $\partial Q_+ = \{(x, y), x = 0 \text{ ou } y = 0\}$ de Q_+ ou vers l'infini. La fonction f est indéfiniment différentiable sur Q_+ , de différentielle $Df(x, y) = (c - d/x, b - a/y)$: son seul point critique y est $m_- = (d/c, a/b)$, qui est nécessairement un minimum pour f , puisque $f(m) \rightarrow +\infty$ lorsque m tend vers le bord de Q_+ ou l'infini, minimum strict vu que $\text{Hess } f(m_-) = \begin{pmatrix} c^2/d & 0 \\ 0 & b^2/a \end{pmatrix}$. L'image $f(Q_+)$ est l'intervalle $[v_-, +\infty)$ avec $v_- = f(m_-)$.

Pour $v \in (v_-, +\infty)$, l'ensemble de niveau $C_v = \{m \in Q_+, f(m) = v\}$ est une courbe régulière, bornée dans Q_+ , donc compacte, puisque f est infinie au bord de Q_+ .

La hauteur $h(x, y) = y$ est extremum aux points où la condition d'extrema liés $Dh = \lambda Df$ est réalisée, i. e. $x = d/c$: au voisinage d'un tel point, C_v est le graphe d'une fonction $y(x)$ qui vérifie $c - d/x + by'(x) - ay'(x)/y = 0$, soit $y'(x) = \frac{c-d/x}{a/y-b}$. Évaluant la dérivée de la relation précédente en $x = d/c$ où $y'(d/c) = 0$, on a $y''(d/c) = \frac{c^2/d}{a/y(d/c)-b}$: les extrema sont stricts. De plus sur la droite $x = d/c$ privée du point m_- , la fonction f est strictement monotone, ainsi C_v intersecte cette droite en au plus deux points : il n'y a qu'un maximum et un minimum possibles, et ils existent bien sur le compact C_v . Il en est de même pour l'abscisse x qui est extremum le long de C_v en deux points de la droite $y = a/b$, de part et d'autre du point m_- . Les courbes C_v sont donc des courbes fermées qui entourent le point critique m_- .

On remarquera que la fonction f est intégrale première du système proie/prédateurs

$$\begin{cases} \dot{x} = ax - bxy, \\ \dot{y} = cxy - dy. \end{cases}$$

puisque le long d'une trajectoire

$$\frac{d}{dt} f(x(t), y(t)) = c\dot{x} - d\dot{x}/x + b\dot{y} - a\dot{y}/y = (c - d/x)(ax - bxy) + (b - a/y)(cxy - dy) = 0$$

Les trajectoires de ce système sont donc toutes bornées et périodiques. ◁

Proposition 6.6. — Soit $C = \{f_1 = 0, f_2 = 0\}$ une courbe régulière dans \mathbb{R}^3 . Si le point $m_0 \in S$ est un extrema local de $h|_S$, alors il existe λ_1, λ_2 tels que $Dh(m_0) = \lambda_1 Df_1(m_0) + \lambda_2 Df_2(m_0)$.

Démonstration. — On peut supposer que $D_{y,z}(f_1, f_2)(m_0)$ est inversible. Il existe des fonctions φ_1, φ_2 définies au voisinage de x_0 telles que C soit localement donnée par $\{(x, \varphi_1(x), \varphi_2(x))\}$. La tangente à la courbe C est de direction $(1, \varphi_1'(x_0), \varphi_2'(x_0))$, égale à l'intersection $\text{Ker } Df_1(m_0) \cap \text{Ker } Df_2(m_0)$. L'extrémalité de m_0 dit que x_0 est un point critique de la fonction $H(x) = h(x, \varphi_1(x), \varphi_2(x))$. L'annulation de la dérivée $H'(x_0)$ donne

$$\frac{\partial h}{\partial x}(m_0) + \frac{\partial h}{\partial y}(m_0)\varphi_1'(x_0) + \frac{\partial h}{\partial z}(m_0)\varphi_2'(x_0) = 0,$$

ainsi $\text{Ker } Df_1(m_0) \cap \text{Ker } Df_2(m_0)$ est inclus dans $\text{Ker } Dh(m_0)$ et donc il existe λ_1, λ_2 tels que $Dh(m_0) = \lambda_1 Df_1(m_0) + \lambda_2 Df_2(m_0)$. ◻

▷ **Exemple 6.4.** — Soit f_1 définie sur \mathbb{R}^3 par $f_1(x, y, z) = 3x^2 + 2y^2 - 1$ et f_2 par $f_2(x, y, z) = x^2 + y^2 + z^2 - 1$. La surface cylindrique $S_1 = \{3x^2 + 2y^2 = 1\}$ de base l'ellipse dans le plan $\{z = 0\}$ et la sphère $S_2 = \{x^2 + y^2 + z^2 = 1\}$ s'intersectent suivant la courbe C : la courbe C ne rencontre pas le plan $z = 0$ (les sections par le plan $z = 0$ de S_1 et S_2 sont une ellipse et un cercle resp. qui ne s'intersectent pas) et est symétrique par rapport au plan $\{z = 0\}$: on notera C_{\pm} les parties $C_{\pm} = C \cap \{\pm z > 0\}$.

Les courbes C_{\pm} sont régulières. En effet, la différentielle $D(f_1, f_2)(x, y, z) = \begin{pmatrix} 6x & 4y & 0 \\ 2x & 2y & 2z \end{pmatrix}$ est de rang 2 le long de C_{\pm} : si $y \neq 0$, la différentielle partielle $D_{y,z}(f_1, f_2)$ est inversible et le théorème des fonctions implicites assure l'existence d'un paramétrage local $(x, y(x), z(x))$ de C , alors que si $x \neq 0$, il existe un paramétrage local $(x(y), y, z(y))$ de C_{\pm} .

Considérons la fonction hauteur $h : (x, y, z) \rightarrow z$ sur C_+ . La fonction h atteint ses extrema en des points $m = (x, y, z)$ où les formes linéaires Dh, Df_1, Df_2 sont linéairement dépendantes : la matrice

$$\begin{pmatrix} Dh \\ Df_1 \\ Df_2 \end{pmatrix} (x, y, z) = \begin{pmatrix} 0 & 0 & 1 \\ 6x & 4y & 0 \\ 2x & 2y & 2z \end{pmatrix}$$

est non régulière si et seulement si $xy = 0$. Les points critiques de h sur C sont donc $m_{\pm} = (0, \pm\sqrt{1/2}, \sqrt{1/2})$ (minimum de $h|_{C_+}$) et $M_{\pm} = (\pm\sqrt{1/3}, 0, \sqrt{2/3})$ (maximum de $h|_{C_+}$). Au voisinage de m_{\pm} , en utilisant le paramétrage en x , on a le long de C

$$6x + 4y'(x)y(x) = 0, \quad 2x + 2y'(x)y(x) + 2z'(x)z(x) = 0$$

d'où $z'(x) = x/(2z(x))$ et $z''(x(m_{\pm})) = 1/(2z(x(m_{\pm})))$: m_{\pm} est un minimum. Au voisinage de M_{\pm} on utilisera un paramétrage en y et on aura $z'(y) = -y/(3z(y))$, d'où $z''(M_{\pm}) = -1/(3z(M_{\pm}))$: M_{\pm} est un maximum local de $z|_{C_+}$. ◁

6.3. Appendice : rappels sur les formes quadratiques

Soit E un espace vectoriel réel. Une forme quadratique q sur E est associée à une forme bilinéaire symétrique S suivant $q(v) = S(v, v), v \in E$. La forme quadratique q détermine la forme bilinéaire symétrique S

$$2S(v, w) = q(v + w) - q(v) - q(w), \quad v, w \in E.$$

Une forme quadratique q est dite positive (resp. négative) si $q(v) \geq 0$ (resp. ≤ 0) pour tout vecteur v ; elle est dite définie si l'inégalité est stricte pour tout vecteur v non nul.

Dans l'espace euclidien \mathbb{R}^n muni du produit scalaire canonique $\langle \cdot, \cdot \rangle$ tel que

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j, \quad x = (x_1, \dots, x_n), y = (y_1, \dots, y_n),$$

la donnée d'une forme bilinéaire symétrique S est équivalente à celle d'un opérateur (ou d'une matrice) symétrique $A : \langle Av, w \rangle = S(v, w), v, w \in \mathbb{R}^n$. Une matrice symétrique A est diagonalisable dans une base orthonormée $(v_j)_{j=1, \dots, n}$. Ainsi, si α_j est la valeur propre du vecteur propre v_j de la matrice A , associée à la forme bilinéaire déterminée par q , on a

$$q\left(\sum_{j=1}^n y_j v_j\right) = \sum_{j=1}^n \alpha_j y_j^2.$$

Le nombre de α_j non nuls est égal au rang de A : c'est le *rang* de la forme quadratique q . Le nombre de valeurs propres positives (resp. négatives) est noté n_+ (resp. n_-) : le couple (n_+, n_-) est appelé la *signature* de la forme quadratique et l'entier n_- son *indice*.

Le rang r est égal à la somme des entiers de signature ($r = n_+ + n_-$). L'indice n_- de la forme quadratique q sur \mathbb{R}^n est caractérisé comme étant le maximum de la dimension

d'un sous-espace E de \mathbb{R}^n tel que $q|_E$ soit définie négative. C'est aussi le nombre de carrés négatifs dans toute écriture

$$q(v) = \pm \ell_1(v)^2 \pm \ell_2(v)^2 \pm \dots \pm \ell_m(v)^2,$$

où les $\ell_k, k = 1, \dots, m$ sont des formes linéaires indépendantes : m est alors le rang de q .

CHAPITRE 7

COMPLÉMENTS

7.1. Fonctions convexes

Définition 7.1. — Une partie C de l'espace vectoriel E est dite *convexe* si pour tout $x, y \in C$ le segment $[x, y] = \{x + \lambda(y - x), \lambda \in [0, 1]\}$ est inclus dans C .

▷ **Exemple 7.1.** — Un sous-espace linéaire, une boule pour une norme sont des parties convexes. Une intersection de convexes reste convexe, mais pas une réunion. ◁

Définition 7.2. — Soit C une partie convexe de l'espace vectoriel E . La fonction f définie sur C et à valeurs numériques est dite *convexe* si

$$(14) \quad f(x + \lambda(y - x)) \leq f(x) + \lambda(f(y) - f(x)), \quad x, y \in C, \lambda \in [0, 1].$$

▷ **Exemple 7.2.** — Une fonction affine est convexe, de même qu'une forme quadratique définie positive. La somme de deux fonctions convexes est convexe, de même que la fonction obtenue par multiplication par un $\Lambda > 0$. Ainsi est convexe toute fonction polynomiales P de degré au plus 2, *i. e.* T définie par $T(t) = at^2 + bt + c, t \in \mathbb{R}$ pour des réels $a > 0, b, c$, est convexe. La fonction valeur absolue $t \in \mathbb{R} \rightarrow |t|$ est convexe. ◁

Proposition 7.1. — Soit C un convexe et f une fonction définie sur C à valeurs réelle. Sont équivalents

1. f est convexe,
2. pour tout $x, y \in C$, la restriction $f_{[x,y]}$ de f au segment $[x, y]$ est convexe,
3. pour tout $x_1, \dots, x_n \in C, \lambda_1, \dots, \lambda_n \in [0, 1]$ avec $\lambda_1 + \dots + \lambda_n = 1$, on a

$$f(\lambda_1 x_1 + \dots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \dots + \lambda_n f(x_n).$$

Si f est continue, f est convexe si et seulement si pour tout $x, y \in C$,

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}.$$

Démonstration. — Pour montrer la dernière équivalence, on remarque que l'inégalité de convexité pour les milieux entraîne l'inégalité de convexité (14) pour les diadiques $\lambda = \sum_{i=1}^N \varepsilon_i 2^{-i}$ avec $\varepsilon_i = 0$ ou 1, ce qui l'entraîne pour tout $\lambda \in [0, 1]$ par continuité. ◻

Théorème 7.1. — Soit f une fonction définie sur un ouvert convexe C et de classe C^2 . La fonction f est convexe si et seulement si $\text{Hess } f(x)$ est positive pour tout $x \in C$.

Démonstration. — Supposons f convexe. Alors, pour tout $x \in C, v \in E$ et $\varepsilon > 0$ assez petit, $x \pm \varepsilon v$ est dans C et

$$f\left(\frac{x - \varepsilon v}{2} + \frac{x + \varepsilon v}{2}\right) = f(x) \leq \frac{f(x - \varepsilon v) + f(x + \varepsilon v)}{2}$$

et par suite

$$\text{Hess } f(x)(v) = \lim_{\varepsilon \rightarrow 0^+} \frac{f(x - \varepsilon v) + f(x + \varepsilon v) - 2f(x)}{\varepsilon} \geq 0$$

Réciproquement, soit $f_{|[x,y]}$ la restriction de f au segment $[x, y]$ définie comme fonction sur $[0, 1]$ par

$$f_{|[x,y]}(t) = f(x + t(y - x)), \quad t \in [0, 1].$$

En appliquant successivement la formule des accroissements finis

$$\begin{aligned} f(x) + f(y) - 2f((x+y)/2) &= f_{|[x,y]}(0) + f_{|[x,y]}(1) - 2f_{|[x,y]}(1/2) = \frac{f'_{|[x,y]}(\theta_1) - f'_{|[x,y]}(\theta_0)}{2} \\ &= \frac{\theta_1 - \theta_0}{2} f''_{|[x,y]}(\theta) \geq 0 \end{aligned}$$

où $\theta_0 \in (0, 1/2)$, $\theta_1 \in (1/2, 1)$ et $\theta \in (0, 1)$, ce qui implique la convexité de f d'après la proposition précédente. \square

7.2. La transformée de Legendre

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^k avec $k \geq 2$ telle que $\Phi = \text{grad } f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ soit un difféomorphisme global d'inverse Ψ et $\text{Hess } f(m)$ définie positive pour tout $m \in \mathbb{R}^n$ (f est en particulier convexe). Pour $\xi \in \mathbb{R}^n$, la fonction g_ξ définie par

$$g_\xi(m) = \langle \xi, m \rangle - f(m), \quad m \in \mathbb{R}^n$$

a pour gradient $\text{grad } g_\xi = \xi - \text{grad } f$ et un unique point critique $m_* = \Psi(\xi)$ qui est un maximum global strict de g_ξ puisque $\text{Hess } g_\xi = -\text{Hess } f$.

Définition 7.3. — La transformée de Legendre de f est la fonction f^* définie sur \mathbb{R}^n par

$$f^*(\xi) = \sup_{m \in \mathbb{R}^n} [\langle \xi, m \rangle - f(m)] = [\langle \xi, m_* \rangle - f(m_*)]_{m_* = \Psi(\xi)}$$

▷ **Exemple 7.3.** — Si f_2 est la norme euclidienne au carré $f_2(m) = \|m\|^2$, on a $f^*(\xi) = \|\xi\|^2/4$. Si φ est une fonction lisse à support compact, alors pour ε assez petit, $f_2 + \varepsilon\varphi$ admet une transformée de Legendre. ◁

Proposition 7.2. — La transformée de Legendre f^* de f de classe C^k est de classe C^k . La dérivée de f^* est inverse de celle de f : $Df^* \circ Df = \text{Id}$ et f^* a une hessienne définie positive.

La transformée de Legendre de f^* est f .

Démonstration. — La transformée de Legendre f^* définie par $f^*(\xi) = \langle \xi, \Psi(\xi) \rangle - f \circ \Psi(\xi)$ est de classe C^{k-1} d'après le théorème d'inversion locale C^k . Sa différentielle appliquée à $h \in \mathbb{R}^n$ vaut

$$\begin{aligned} Df^*(\xi)(h) &= \langle h, \Psi(\xi) \rangle + \langle \xi, D\Psi(\xi)h \rangle - Df(\Psi(\xi)) \circ D\Psi(\xi)h \\ &= \langle h, \Psi(\xi) \rangle + \langle \xi, D\Psi(\xi)h \rangle - \langle \text{grad } f(\Psi(\xi)), D\Psi(\xi)h \rangle \end{aligned}$$

où les deux derniers termes sont de somme nulle puisque $\Phi(\Psi(\xi)) = \xi$. Ainsi $\text{grad } f^* = \Psi$ est de classe C^{k-1} et f^* est de classe C^k .

La hessienne de f^* en ξ est $D\Psi(\xi) = (D\Phi(\Psi(\xi)))^{-1}$, qui est définie positive comme l'est $D\Phi = \text{Hess } f$.

La transformée de Legendre f^{**} de f^* vérifie

$$f^{**}(x) = [\langle x, \xi \rangle - f^*(\xi)]_{x = \Psi(\xi)} = \langle \Psi(\xi), \xi \rangle - [\langle \Psi(\xi), \xi \rangle - f(x)] = f(x),$$

soit $f^{**} = f$. \square

7.3. Le lemme de Morse

Lemme 7.1. — Soit f une fonction numérique définie sur une boule $B(m_0, r)$ centrée en m_0 de classe $C^{k \geq 1}$. Il existe des fonctions $f_i, i = 1, \dots, n$ de classe C^{k-1} sur $B(0, r)$ telles que

$$f(m_0 + x) = f(m_0) + \sum_{i=1}^n x_i f_i(x), \quad x = (x_1, \dots, x_n) \in B(0, r)$$

et $f_i(0) = \partial_{x_i} f(m_0)$. Si $k \geq 2$, on a $\partial_{x_j} f_i(0) = \frac{1}{2} \partial_{x_j x_i}^2 f(m_0)$.

Démonstration. — Si on considère la fonction $t \in [0, 1] \rightarrow f(m_0 + tx)$, on a

$$f(m_0 + x) = f(m_0) + \int_0^1 \frac{d}{ds} [f(m_0 + sx)] ds = f(m_0) + \sum_{i=1}^n x_i \int_0^1 \partial_{x_i} f(m_0 + sx) ds.$$

Les fonctions f_i définies sur $B(0, r)$ par $f_i(x) = \int_0^1 \partial_{x_i} f(m_0 + sx) ds$ sont de classe C^{k-1} avec $f_i(0) = \partial_{x_i} f(m_0)$ et, si $k \geq 2$

$$\partial_{x_j} f_i(0) = \int_0^1 s [\partial_{x_j} \partial_{x_i} f(m_0 + sx)]_{|x=0} ds = \frac{1}{2} \partial_{x_j x_i}^2 f(m_0).$$

□

Lemme 7.2 (Lemme de Morse en dimension 1). — Soit f une fonction numérique définie sur un voisinage I de t_0 de classe $C^{k \geq 3}$ avec $f'(t_0) = 0$ et $f''(t_0) \neq 0$. Il existe un difféomorphisme φ d'un voisinage U de $u = 0$ sur un voisinage de t_0 de classe C^{k-2} tel que

$$f \circ \varphi(u) = f(t_0) + \varepsilon u^2, \quad u \in U,$$

où $\varepsilon = \pm 1$ suivant le signe de $f''(t_0)$.

Démonstration. — D'après le lemme précédent, il existe une fonction g définie au voisinage de $\tau = 0$, de classe C^{k-1} avec $g(0) = f'(t_0) = 0, g'(0) = f''(t_0)/2$ telle que

$$f(t) = f(t_0) + (t - t_0)g(t - t_0), \quad t \sim t_0$$

et une fonction h au voisinage de $\tau = 0$ de classe C^{k-2} avec $g(\tau) = \tau h(\tau)$ et $h(0) = g'(0)$ soit

$$f(t) = f(t_0) + (t - t_0)^2 h(t - t_0), \quad h(0) = f''(t_0)/2.$$

Ainsi, au voisinage de $\tau = 0$ et si ε est ± 1 suivant le signe de $h(0) = f''(t_0)$, on a $h(\tau) = \varepsilon h_1(\tau)^2$ avec $h_1 = \sqrt{\varepsilon h}$ de classe C^{k-2} . Alors, l'application ψ définie au voisinage de t_0

$$\psi(t) = (t - t_0)h_1(t - t_0), \quad t \sim t_0$$

est de classe C^{k-2} avec $\psi(t_0) = 0$ et $\psi'(t_0) = h_1(0) = \sqrt{|f''(t_0)|/2}$ non nul : ψ est donc inversible au voisinage de t_0 . Son inverse local φ , de classe C^{k-2} vérifie bien $f \circ \varphi(u) = f(t_0) + \varepsilon [\psi \circ \varphi(u)]^2 = f(t_0) + \varepsilon u^2$ sur un voisinage de $u = 0$. □

Théorème 7.2 (Lemme de Morse en dimension 2). — Soit f une fonction numérique définie sur un voisinage $U \subset \mathbb{R}^2$ de m_0 de classe $C^{k \geq 3}$ avec $Df(m_0) = 0$ et $\text{Hess } f(m_0)$ non dégénérée. Il existe ε, η de carrés 1 et un difféomorphisme Φ d'un voisinage $V \subset \mathbb{R}^2$ de $(u, v) = 0$ sur un voisinage de m_0 de classe C^{k-2} tel que

$$f \circ \Phi(u, v) = f(m_0) + \varepsilon u^2 + \eta v^2, \quad (u, v) \in V.$$

Démonstration. — La hessienne $\text{Hess } f(m_0)$ étant non dégénérée, il existe un isomorphisme linéaire A_0 et des réels ε, η de carrés 1 tels que $\text{Hess } f(m_0)(A_0(X, Y)) = \varepsilon X^2 + \eta Y^2$ si $(X, Y) \in \mathbb{R}^2$. Vu que, pour un endomorphisme linéaire A , $\text{Hess}(f \circ \tilde{A})(v) = \text{Hess}(f)(Av)$ si

\tilde{A} est défini par $m \rightarrow m_0 + A(m - m_0)$, quitte à composer f avec \tilde{A}_0 , on peut supposer que $\text{Hess } f(m_0)$ est la matrice diagonale $\text{Diag}(\varepsilon, \eta)$. D'après le lemme 7.1,

$$f(m_0 + (x, y)) = f(m_0) + xf_1(x, y) + yf_2(x, y)$$

avec

$$f_1(x, y) = \int_0^1 \partial_x f(m_0 + s(x, y)) ds, \quad f_2(x, y) = \int_0^1 \partial_y f(m_0 + s(x, y)) ds$$

et

$$f_1(0, 0) = \partial_x f(m_0) = 0, \quad f_2(0, 0) = \partial_y f(m_0) = 0,$$

puis

$$f(m_0 + (x, y)) = f(m_0) + x^2 f_{11}(x, y) + yx f_{12}(x, y) + xy f_{21}(x, y) + y^2 f_{22}(x, y)$$

avec

$$\begin{aligned} f_{11}(0, 0) &= \partial_x f_1(0, 0) = \partial_{xx}^2 f(m_0)/2 = \varepsilon/2, & f_{12}(0, 0) &= \partial_y f_1(0, 0) = \partial_{xy}^2 f(m_0)/2 = 0, \\ f_{21}(0, 0) &= \partial_x f_2(0, 0) = \partial_{yx}^2 f(m_0)/2 = 0, & f_{22}(0, 0) &= \partial_y f_2(0, 0) = \partial_{yy}^2 f(m_0) = \eta/2. \end{aligned}$$

Ainsi on a

$$f(m_0 + (x, y)) = f(m_0) + a(x, y)x^2 + 2b(x, y)xy + c(x, y)y^2$$

avec a, b, c de classe C^{k-2} et

$$a(0, 0) = \partial_{xx}^2 f(m_0)/2 = \varepsilon/2, \quad b(0, 0) = \partial_{xy}^2 f(m_0)/2 = 0, \quad c(0, 0) = \partial_{yy}^2 f(m_0)/2 = \eta/2.$$

On peut écrire donc au voisinage de $(x, y) = (0, 0)$

$$f(m_0 + (x, y)) = f(m_0) + \varepsilon \left(\sqrt{\varepsilon a(x, y)} x + \frac{\varepsilon b(x, y)}{\sqrt{\varepsilon a(x, y)}} y \right)^2 + \eta \left(\sqrt{\eta c(x, y) - \eta \frac{b(x, y)^2}{a(x, y)}} y \right)^2.$$

L'application Ψ définie au voisinage de m_0 par

$$\Psi(m_0 + (x, y)) = \left(\sqrt{\varepsilon a(x, y)} x + \frac{\varepsilon b(x, y)}{\sqrt{\varepsilon a(x, y)}} y, \sqrt{\eta c(x, y) - \eta \frac{b(x, y)^2}{a(x, y)}} y \right), \quad (x, y) \sim (0, 0),$$

est de classe C^{k-2} avec jacobienne $\text{Jac } \Psi(m_0) = \text{Id}/\sqrt{2}$ et $\Psi(m_0) = (0, 0)$: c'est un difféomorphisme local dont l'inverse Φ définie au voisinage de $(u, v) = (0, 0)$ vérifie

$$f \circ \Phi(u, v) = f(m_0) + \varepsilon u^2 + \eta v^2. \quad \square$$

\triangle Le couple (ε, η) est déterminé, à l'ordre près, par la signature de la hessienne $\text{Hess } f(m_0)$. Par ailleurs, le lemme de Morse vaut aussi pour des fonctions à n variables, avec un énoncé très similaire à celui pour deux variables. . . ∇

On peut appliquer le lemme de Morse à l'étude de l'intersection d'une surface S dans \mathbb{R}^3 et de ses plans tangents $\tau_M S$. Soit $M_0 \in S$ et, en choisissant convenablement les axes de coordonnées, une représentation de S au voisinage de M_0 par le graphe

$$G_f = \{(x, y, f(x, y)), (x, y) \sim (0, 0)\}$$

avec $M_0 = (0, 0, f(0, 0))$ et le plan tangent

$$\tau_{M_0} S = \{(\alpha, \beta, f(0, 0) + \alpha \partial_x f(0, 0) + \beta \partial_y f(0, 0)), \alpha, \beta \in \mathbb{R}\}.$$

On suppose la hessienne $\text{Hess } f(0, 0)$ non dégénérée. Alors, en utilisant le changement de variable $\Phi : (u, v) \rightarrow (x(u, v), y(u, v))$ donné par le lemme de Morse pour la fonction

$$h(x, y) = f(x, y) - f(0, 0) - x \partial_x f(0, 0) - y \partial_y f(0, 0)$$

qui a la même hessienne que f en $(0, 0)$, l'intersection $S \cap \tau_{M_0} S = \{(x, y, h(x, y)) : h(x, y) = 0\}$ au voisinage de M_0 est le lieu d'équation $\varepsilon u^2 + \eta v^2 = 0$ dans les coordonnées (u, v) données par le lemme de Morse. Si la hessienne $\text{Hess } f(0, 0)$ est définie, on a $\varepsilon \eta = 1$, le seul point d'intersection est M_0 et la surface est localement d'un seul côté du plan tangent,

dessus ou en dessous suivant que $\text{Hess } f(0, 0)$ est définie positive ou négative. Si la hessienne est non définie, soit $\varepsilon\eta = -1$, l'intersection $S \cap \tau_{M_0}S$ est donnée en coordonnées (u, v) par $u = \pm v$ au voisinage de $(0, 0)$: ce sont les portions de courbes $u \rightarrow (\Phi(u, \pm u), f(0, 0)) = (x(u, \pm u), y(u, \pm u), f(0, 0))$ dont les tangentes $(\partial_u \Phi(0, 0) \pm \partial_v \Phi(0, 0), 0)$ en M_0 sont suivant les directions annulant $\text{Hess } f(0, 0)$.

▷ **Exemple 7.4.** — La sphère $S^2 = \{x^2 + y^2 + z^2 = 1\}$ au voisinage de $M_0 = (0, 0, 1)$ est donnée par le graphe de la fonction $f(x, y) = \sqrt{1 - x^2 - y^2} \simeq 1 - (x^2 + y^2)/2 + o(x^2 + y^2)$: la sphère est toute entière, à l'exception du point de tangence, à l'intérieur d'un des demi-plans déterminés par chacun de ses espaces tangents.

Pour l'hyperboloïde à une nappe $H_1 = \{x^2 - y^2 + z^2 = 1\}$, au voisinage de $M_0 = (0, 0, 1)$, la surface H_1 coïncide avec le graphe de la fonction $h(x, y) = \sqrt{1 - x^2 + y^2} \simeq 1 - (x^2 - y^2)/2 + o(x^2 + y^2)$ et le plan tangent $\{Z = 1\}$ intersecte H_1 suivant les deux droites $\{x = \pm y, z = 1\}$. ◁

INDEX

- équation régulière, 20
- W. BESSEL, 11
- changement de coordonnées, 15
- changement de variables, 15
- contractante, 13
- contraction, 13
- courbe régulière, 20
- dérivée d'ordre k , 7
- dérivée directionnelle, 2
- difféomorphisme, 12
- différentielle, 3
- espace tangent, 22
- L. EULER, 11
- extremum, 29
 - global, 29
 - local, 29
 - strict, 29
- fonction
 - dérivable, 2
 - de classe C^0 , 7
 - de classe C^1 , 6
 - de classe C^∞ , 7
 - de classe C^k , 7
 - différentiable, 3, 4
- T. GRÖNWALL, 9
- gradient, 23
- hessienne, 27
- J.-L. LAGRANGE, 11, 31
- A.-M. LEGENDRE, 11, 36
- G. W. LEIBNIZ, 3
- lieu régulier, 20
- matrice jacobienne, 4
- maximum, 29
- minimum, 29
- M. MORSE, 37
- multiplicateur(s) de Lagrange, 31
- point critique, 16
- G. RIEMANN, 7
- H. SCHWARZ, 27, 28
- surface régulière, 22
- système de coordonnées, 15
- B. TAYLOR, 11, 19, 27