Faculté des sciences et techniques
Département de mathématiques

Licence de Mathématiques

Filière Mathématiques classiques

Calcul différentiel

II. Fonctions différentiables

Laurent Guillopé

www.math.sciences.univ-nantes.fr/˜guillope/LC4

TABLE DES MATIÈRES

CHAPITRE 1

APPLICATIONS DIFFÉRENTIABLES

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

$$\lim_{h \to 0} o(h)/\|h\| = 0.$$

Si h est dans $\mathbb{R}$, il suffit de considérer la limite de $o(h)/h$ lorsque $h \to 0$.

De manière analogue on notera $\varepsilon(h)$ toute quantité telle que $\lim_{h\to 0} \varepsilon(h) = 0$. Ainsi, on remplacera parfois un $o(h)$ par $\|h\|\varepsilon(h)$.

Vu que toutes les normes de $\mathbb{R}^k$ sont équivalentes, ces définitions sont indépendantes des normes choisies dans $\mathbb{R}^n$ et $\mathbb{R}^p$.

1.1. Différentiabilité

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

$$f(x_0 + h) = f(x_0) + \dot{f}(x_0)h + o(h).$$

La dérivabilité est équivalente à l'approximation locale par une fonction affine, autrement dit à l'existence d'une 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 four f, comme la fonction $x \to x^3 \sin(1/x) = o(x^2)$ pour $x_0 = 0$ le confirme.

Pour une fonction de deux variables (on généralisera aisément pour f de n variables, avec $n \geq 3$) $f : m = (x, y) \in B(m_0, r) \to f(m) = f(x, y) \in \mathbb{R}$, les dérivées directionnelles s'obtiennent en considérant les restrictions de f à des droites passant par m_0 : les dérivées partielles $\frac{\partial f}{\partial x}(m_0)$ et $\frac{\partial f}{\partial y}(m_0)$ (notées parfois $\partial_x f(m_0), \partial_y f(m_0)$) sont, si elles existent, les dérivées des fonctions $t \to f(x_0 + t, y_0)$ et $s \to f(x_0, y_0 + s)$ en $t = 0$ et $s = 0$ resp. De manière plus générale, on a la notion de dérivée directionnelle.

Définition 1.1. — Soit $f : B(m_0, r) \subset \mathbb{R}^n \to \mathbb{R}$. 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 $f(m_0 + th)$.

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 1.1. —

1. Si $f(x, y) = \sin(x + y)$, on a $D_{(u,v)} f(0, 0) = u + v$.

2. Si $f(x,y) = r\sin 2\theta = 2xy(x^2+y^2)^{-1/2}$ où on a utilisé les coordonnées polaires $(x,y) = (r\cos\theta, r\sin\theta)$, on a $f(t(u,v)) = |t| f(u,v)$ pour $t \in \mathbb{R}$. Ainsi, la restriction de f, nulle en $t=0$, à une 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.

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

$$f(m_0 + h) = f(m_0) + L(h) + o(h). \tag{1}$$

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

△ La différentielle est notée parfois $f'(m_0)$ (voire $\dot{f}(m_0)$) : on l'évitera dans un premier temps, pour bien distinguer dérivée (d'une fonction d'une variable réelle) et différentielle. Si f est à valeurs numériques, on note souvent sa différentielle df suivant Leibniz, qui écrit $df = f'_x dx + f'_y dy + f'_z dz$; il faut y interpréter dx comme la différentielle de la fonction coordonnée x. ▽

Proposition 1.1. — *Soit f une fonction différentiable en m_0.*
- *L'application L de la définition 1.2 est unique.*
- *La fonction f est continue en m_0.*
- *La fonction f admet des dérivées directionnelles suivant toute direction en m_0 et $D_h f(m_0) = Df(m_0)(h)$. En particulier, f admet des dérivées partielles.*

Démonstration. — Soient L_1 et L_2 deux formes linéaires vérifiant (1). Alors pour, h assez petit,

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

ainsi $(L_1 - L_2)(h/\|h\|) = o(h)/\|h\|$ et donc 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 h fixé non nul, on a

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

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

Exemple 1.2. — Il se peut que f admette des dérivées directionnelles sans être différentiable. Ainsi de $f(m) = r\sin 3\theta = 3y - 4y^3(x^2+y^2)^{-1}$ en $m_0 = (0,0)$: les dérivées directionnelles sont $D_{(\cos\theta, \sin\theta)} f(m_0) = \sin 3\theta = 3\sin\theta - 2\sin^3\theta$, non linéaire en $\sin\theta$.

Pour une fonction $f : B(m_0, r) \subset \mathbb{R}^n \to \mathbb{R}^p$, on définit de manière analogue la différentiabilité de f par l'existence d'une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ telles que

$$f(m_0 + h) = f(m_0) + L(h) + o(h)$$

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

Exemples 1.3. —

1. Soit $f : I \subset \mathbb{R} \to \mathbb{R}^n$ différentiable en t_0. L'application linéaire $Df(t_0) : \mathbb{R} \to \mathbb{R}^n$ est habituellement identifiée à sa valeur sur le vecteur générateur $h = 1$, 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 $A : \mathbb{R}^n \to \mathbb{R}^p$ linéaire et $B \in \mathbb{R}^p$. Alors l'application $f : x \in \mathbb{R}^n \to Ax + B \in \mathbb{R}^p$ est différentiable en tout $x \in \mathbb{R}^n$, avec $Df(x) = A$. En effet,
$$f(x+h) = A(x+h) + B = Ax + B + Ah = f(x) + Ah.$$

Définition 1.3. — Soit $f = (f_1, \dots, f_p) : B(m_0, r) \subset \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable en m_0. Sa matrice jacobienne est la matrice
$$\operatorname{Jac}(f)(m_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(m_0) & \dots & \frac{\partial f_1}{\partial x_n}(m_0) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(m_0) & \dots & \frac{\partial f_p}{\partial x_n}(m_0) \end{pmatrix}.$$
C'est la représentation matricielle de la différentielle $Df(m_0)$ dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.

Exemple 1.4. — Les « coordonnées polaires » $(r, \theta) \in \mathbb{R}^{+*} \times \mathbb{R} \to (r\cos\theta, r\sin\theta)$ ont pour matrice jacobienne
$$\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Lemme 1.1. — *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 $f \in \mathcal{D}(m_0) \to Df(m_0) \in L(\mathbb{R}^n, \mathbb{R})$ 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
$$D(f_1 f_2)(m_0) = f_1(m_0) Df_2(m_0) + f_2(m_0) Df_1(m_0).$$

Démonstration. — Que $\mathcal{D}(m_0)$ soit un espace vectoriel et l'application $f \to Df(m_0)$ 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
$$\begin{aligned} f(m_0 + h) &= f(m_0) + Df(m_0)(h) + o(h), \quad \|h\| < r_f, \\ g(m_0 + h) &= g(m_0) + Dg(m_0)(h) + o(h), \quad \|h\| < r_g. \end{aligned}$$
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 $o(h)$. □

La fonction $\operatorname{Inv} : A \in \mathrm{GL}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}), A \text{ inversible}\} \to A^{-1}$ est différentiable de différentielle $D\operatorname{Inv}(A)(h) = -A^{-1} h A^{-1}, h \in M_n(\mathbb{R})$. En effet
$$\begin{aligned} (A + h)^{-1} &= A^{-1}(1 + hA^{-1})^{-1} = A^{-1} + A^{-1}((1 + hA^{-1})^{-1} - 1) \\ &= A^{-1} - A^{-1}(1 + hA^{-1})^{-1} h A^{-1} = A^{-1} - A^{-1} h A^{-1} + A^{-1}((1 + hA^{-1})^{-1} - 1) h A^{-1} \\ &= A^{-1} - A^{-1} h A^{-1} + A^{-1}(1 + hA^{-1})^{-1} h A^{-1} h A^{-1} \end{aligned}$$
où le dernier terme est borné par $\|h\|^2$ vu que le facteur $(1 + hA^{-1})^{-1} = \sum_{k=0}^{\infty}(-hA)^k$ est borné au voisinage de $h = 0$.

Théorème 1.1. — *Soit U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^p$, $f : U \to \mathbb{R}^p$ et $g : V \to \mathbb{R}^q$. 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*

$$D(g \circ f)(m_0) = Dg(f(m_0)) \circ Df(m_0),$$
$$\mathrm{Jac}(g \circ f)(m_0) = \mathrm{Jac}(g)(f(m_0))\,\mathrm{Jac}(f)(m_0).$$

Démonstration. — On compose les développements limités

$$f(m_0 + h) = f(m_0) + Df(m_0)(h) + o(h),$$
$$g(f(m_0) + k) = g(f(m_0)) + Dg(f(m_0))(k) + o(k),$$

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 $o(h)$. □

Exemples 1.5. —

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

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

2. Soit $h \in \mathbb{R}^n$ et $f : B(m_0, r) \to \mathbb{R}$. La dérivée directionnelle $D_h f(m_0)$ apparaît comme la dérivée de l'application composée $F = f \circ \gamma$ où $\gamma : t \in \mathbb{R} \to m_0 + th \in \mathbb{R}^n$. On a en effet
$$DF(t_0) = Df(m_0 + t_0 h) \circ D\gamma(t_0).$$
La dérivée $D\gamma(t_0)$ s'identifie au vecteur h.
3. Si f et g sont différentiables en m_0 et $f(m_0)$ resp. avec $g \circ f = \mathrm{Id}$, alors $Dg(f(m_0)) \circ Df(m_0) = \mathrm{Id}$. 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 $\mathcal{C}^1$

Définition 1.4. — Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^p$. Si f est différentiable en tout point $m \in U$ et l'application $Df : m \in U \to Df(m) \in L(\mathbb{R}^n, \mathbb{R}^p)$ est continue, l'application f est dite de classe C^1 sur U.

Théorème 1.2. — *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^p$. 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

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

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 $\varepsilon_{ij}(h) \to 0$ lorsque $h \to 0$. On a donc

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

et par suite

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

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 jacobienne $\mathrm{Jac}(f)$, dont les coordonnées sont les dérivées partielles de f, est continue ; il en est de même pour la dérivée $m \in U \to Df(m) \in L(\mathbb{R}^n, \mathbb{R}^p)$. □

Exemple 1.6. — L'expression de l'inverse d'une matrice $A \in \mathrm{GL}(n, \mathbb{R}))$ en terme de déterminants

$$A^{-1} = (\text{dét}\, A)^{-1} \,\mathrm{Comat}\, A$$

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

Une fonction continue est dite de classe C^0. La définition des fonctions de classe C^1 est complétée par celle des fonctions de classe C^k de manière récursive.

Définition 1.5. — Pour $k \geq 1$, une fonction est de dite de classe C^k si elle est différentiable et sa différentielle Df est de classe C^{k-1}. Elle est dite de classe C^∞ si elle est C^k pour tout entier k.

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 1.7. — Les fonctions polynômes sont de classe C^∞.
La fonction exponentielle sur $\mathbb{C}$ est de classe C^∞.

1.4. 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 $f : t \in \mathbb{R} \to (\cos t, \sin t) \in \mathbb{R}^2$, 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 égalité des accroissements finis pour les fonctions différentiables sur des ouverts de $\mathbb{R}^n$ et à valeurs vectorielles, source de nombreuses applications.

***Théorème 1.3* (Inégalité des accroissements finis).** — *Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^p$ 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*

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

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

$$g(1) - g(0) = \int_0^1 \dot{g}(s) ds \tag{2}$$

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. □

Seconde 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. □

Troisième 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

$$\|g(t) - g(0)\| \leq \left(\sup_{s \in [0,t]} \|\dot{g}(s)\| + \varepsilon \right) t. \tag{3}$$

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

$$A_\varepsilon = \{T \in [0,1] : \text{l'inégalité (3) est vérifiée pour } t \text{ dans } [0,T]\}.$$

[(1)] Pour montrer qu'une fonction d'une variable, à valeurs numériques et continûment dérivable, est l'intégrale de sa dérivée, on utilise l'identité des dérivées des deux membres de cette égalité et l'identité des accroissements finis $f(b) - f(a) = (b-a)f'(c)$ en une variable, prouvée comme corollaire du théorème de Rolle. L'identité intégrale (2) pour une fonction d'une variable à valeurs vectorielles dans un espace de dimension finie en résulte, et donc par suite l'inégalité des accroissements finis : en dimension 1 et pour une fonction à valeurs numériques, on évitera l'argument de calcul intégral pour utiliser directement l'égalité des accroissements finis ; pour des fonctions à valeurs vectorielles, on vérifie bien qu'on démontre l'inégalité des accroissements finis sans l'utiliser subrepticement (même si on peut utiliser l'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 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

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

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

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

inégalité qui exprime l'appartenance de $T_\varepsilon + \eta$ à A_ε : l'hypothèse $T_\varepsilon < 1$ est donc erronée. Ainsi, faisant tendre ε vers 0 dans (3) avec $t = 1$, on obtient l'inégalité du théorème, en remarquant la majoration $|\dot g(t)| \le \|Df(m_0 + t(m_1 - m_0))\| \, \|m_1 - m_0\|$. □

Corollaire 1.1. — *Soit $U \subset \mathbb{R}^n$ un ouvert connexe. Si $f : U \to \mathbb{R}^p$ différentiable 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 $\widetilde m$ de $B(m,r)$, le segment $[m,\widetilde m]$ est inclus dans la boule $B(m,r)$ et d'après l'inégalité des accroissements finis $f(\widetilde m) = f(m)$, ainsi $B(m_0,r) \subset U$. Par connexité de U, $V_0 = U$ et la fonction est constante sur U. □

Corollaire 1.2. — *Si $f : U \to \mathbb{R}^p$ est de classe C^1, la fonction f est localement lipschitzienne.*

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

$$\|f(x) - f(y)\| \le 2M_{m,r}r, \quad x, y \in B(m,r),$$

ce qui est le caractère localement lipschitzien de f. □

CHAPITRE 2

LE THÉORÈME D'INVERSION LOCALE

2.1. L'énoncé du théorème

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

Exemples 2.1. —

1. La fonction $\mathrm{tg} : (-\pi/2, \pi/2) \to \mathbb{R}$ est une application bijective différentiable, de dérivée $1 + \mathrm{tg}^2$ et d'application inverse arctg dérivable : $\mathrm{arctg}'(y) = (1+y^2)^{-1}$.
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 $t \to t^{1/3}$ 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
$$\dot{f}(1/(2k\pi + u)) = 1 - \cos u + 2(2k\pi + u)^{-1}\sin u \sim_{k\to+\infty, u\to 0} (k\pi)^{-1}u$$
change de signe au voisinage de $u = 0$: la fonction f n'est pas monotone au voisinage de $t = 0$.
4. La fonction exponentielle de $\mathbb{C}$ est différentiable,
$$\mathrm{e}^{z+h} = \mathrm{e}^z\mathrm{e}^h = \mathrm{e}^z(1 + h + o(h))$$
avec pour différentielle $D\exp(z)$ l'application linéaire induite par la multiplication par e^z : bien que la différentielle $D\exp(z)$ soit inversible pour tout z, l'application $\exp : \mathbb{C} \to \mathbb{C}$ n'est pas injective ($\mathrm{e}^{2\mathrm{i}\pi} = \mathrm{e}^0$).

Définition 2.1. — $f : U \to V$ 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 $f^{-1} : V \to U$.

Théorème 2.1. — *Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^n$ 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 $f : (s,t) \to (s + st, t + st)$. Sa jacobienne est $\begin{pmatrix} 1+t & t \\ s & 1+s \end{pmatrix}$, inversible au voisinage de $(s,t) = (0,0)$.

2.2. Le théorème du point fixe

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

Définition 2.2. — Une contraction de X est une application $T : X \to X$ telle qu'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 $x \to Ax + b$ est une contraction de $(\mathbb{R}^n, \|\ \|)$ dès que $\|A\| < 1$. Le choix de la norme $\|\ \|$ 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 $\|\ \|_\infty$ sur $\mathbb{R}^n$ et $\sup(1/2, \varepsilon + 1/3)$ pour celle dérivée de $\|\ \|_1$: 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 une version avec paramètre.

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

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

Sa limite $x_\infty(y)$ est un point fixe de T_y :

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

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

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

et la continuité de l'application $y \to a(y)$. □

△ 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

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 $\widetilde{f} = Df(0)^{-1} \circ \tau_{-f(m_0)} \circ f \circ \tau_{m_0}$ qui permet de 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 $T_y(x) = y - f(x) + x$, soit $T_y = y - \varphi$ avec $\varphi(x) = f(x) - x, x \in U$.

Vu sur $D\varphi(0) = 0$ et 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, 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 $y \in B(0, r/2) \to g(y) \in B(0, r)$ 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 $f : U' \to V'$, d'inverse $g : V' \to U'$ continue, est un homéomorphisme.

Pour montrer que g est dérivable sur V', soit $y \in V'$, $x = g(y)$, k tel que $y + k \in V'$ et $h = g(y+k) - g(y)$. Tout d'abord, 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

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

on a

$$h = Df(x)^{-1}(k) + Df(x)^{-1}(o(h)) = Df(x)^{-1}(k) + o(h),$$

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 $k \to 0$, un $o(h)$ est ainsi un $o(k)$ et

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

c'est la différentiabilité de g en y, avec comme différentielle $(Df(x))^{-1} = (Df(g(y)))^{-1}$.

Pour conclure, remarquons que la différentielle $Dg(y) = [Df(g(y))]^{-1}$ est la composée d'applications continues et est donc est continue.

2.4. Changement de variables

L'exemple le plus simple de difféomorphisme est donné par une matrice inversible A. L'interprétation géométrique d'une 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 $\mathbf{b} = (b_1, \ldots, b_n)$ une base et $X = (x_1, \ldots, x_n)$ les coordonnées de $m \in \mathbb{R}^n$; en fait si $(x_1, \ldots, x_n)$ note la base de formes linéaires sur $\mathbb{R}^n$ duale de la base $\mathbf{b}$ (*i. e.* telle $x_i(b_j) = \delta_{ij}$), on a la i-ème coordonnée de m donnée par l'évaluation de x_i sur m. Si $\mathbf{c} = (c_1, \ldots, c_n)$ est la base image de $\mathbf{b}$ par P, *i. e.* $c_i = Pb_i$, 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 $X = PY$. En dimension $n = 2$, les droites $y_i = y_i(0), i = 1, 2$ correspondent aux droites $P(y_1(0), u), u \in \mathbb{R}$ (parallèles au vecteur c_2) et $P(v, y_2(0)), v \in \mathbb{R}$ (parallèles au vecteur c_1).

Un difféomorphisme $f : U \to V$ 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 $a \in U$ est repéré par les coordonnées $(y_1, \ldots, y_n)$ de $f(m)$ au voisinage de $f(a)$. Autrement dit, en dimension $n = 2$, le tissu des droites $y_i = y_i(0), i = 1, 2$ au voisinage de $f(a)$ correspond au tissu de courbes $(x_1(y_1(0), u), x_2(y_1(0), u)), u \sim y_2(0)$ et $(x_1(v, y_2(0)), x_2(v, y_2(0))), v \sim y_1(0)$.

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

Exemple 2.4. — Les coordonnées polaires $(x, y) \to (r, \theta)$ avec $(x, y) = (r\cos\theta, r\sin\theta)$, soit $r = qrtx^2 + y^2, \theta = \operatorname{arctg}(y/x)$ ou $\theta = \pi/2 - \operatorname{arctg}(x/y)$, donnent un changement de variable.

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

Définition 2.4. — Soit $h : U \to \mathbb{R}$ une fonction différentiable. Le point $m_0 \in U$ est dit *critique* 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 $(v_1, \ldots, v_n)$ une base orthonormée de vecteurs propres de A, avec valeurs propres associées $\lambda_1, \ldots, \lambda_n$ ordonnées de manière croissante : $\lambda_1 \leq \ldots \leq \lambda_n$. Si $v = \sum_{i=1}^n x_i v_i$ est un point critique de h_A et I le premier indice 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 $h : U(\subset \mathbb{R}^n) \to \mathbb{R}$ 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 $f = (f_1, \ldots, f_n) : U' \to \mathbb{R}$ avec $f_1 = h$.*

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 $L_2, \ldots, L_n$ telles que $(Dh(a), L_2, \ldots, L_n)$ soit une base de ce dual. Alors l'application $F = (h, L_2, \ldots, L_n)$ 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 $Dh(m_0), L_2, \ldots, L_n$ sont linéairement indépendantes. Ainsi $(h, L_2, \ldots, L_n)$ 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 $Dh(x, y) = (3x^2, -2y)$ 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 $\Phi : (x, y) \to (h(x, y), y)$ valant

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

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

Le théorème des fonctions implicites a pour objet d'exprimer la relation entre x et y imposée 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 que c'est toujours le cas.

Exemples 3.1. —

1. Le cercle C d'équation $x^2+y^2=1$ est l'union d'arcs paramétrés $x=\pm\sqrt{1-y^2}$ et $y=\pm\sqrt{1-x^2}$. 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 $S=(A\,B)$ 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. — *Soit U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^p$ et $f:U\times V\to\mathbb{R}^p$ 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 $D_yf(m_0)$ 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 $\varphi:U_0\to V_0$ telle que $D_yf(m)$ 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 $D\varphi(x)=-[D_yf(x,\varphi(x))]^{-1}D_xf(x,\varphi(x)), x\in U_0$.*

Exemple 3.2. — La fonction $f(x,y)=y^3/3-y-x$ est C^1 sur le plan $\mathbb{R}^2$, avec pour différentielle $Df(x,y)=(-1\ y^2-1)$. 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 $(x=2/3, y=\pm1)$. Pour $x\neq 2/3$, on a la différentielle $D\varphi$ donnée par $1/(\varphi^2(x)-1)$.

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

$$\begin{pmatrix}\mathrm{Id} & 0\\ J_xf & J_yf\end{pmatrix},$$

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

B avec l'inverse $h = g^{-1}$ vérifiant $h(X,Z) = (X, \varphi(X,Z))$. On a donc

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

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 $R_1 = U_0 \times V_1$ 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. Différentiant $f(x, \varphi(x)) = 0$, 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. $\square$

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 $a_i, i = 1,2,3$ 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 $P(\lambda, Y)$, *i. e.* $\partial_y p(\lambda_0, y_0) \neq 0$. Le théorème des fonctions implicites assure l'existence de $r > 0, \delta > 0$, 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 $P(Z) = p_n Z^n + p_{n-1} Z^{n-1} + \ldots + p_1 Z + p_0$ 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

$$p(z) = z^n (1 + p_{n-1} z^{-1} + \ldots + p_1 z^{1-n} + p_0 z^{-n}),$$

où le facteur de droite tend vers 1 lorsque $|z| \to \infty$. 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)

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

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

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

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', on reprend l'expression précédente, où N est le minimum des k tels que $p^{(k)}(z) \neq 0$,

$$p(z+h) = p(z) + \frac{p^{(N)}(z)}{N!}h^N \left[1 + \frac{N!p^{(N+1)}(z)}{(N+1)!p^{(N)}(z)}h + \ldots + \frac{N!p(z)}{p^{(N)}(z)}h^{n-N}\right]$$

Notant $q_{z,N}(h)$ le dernier facteur, soit $f_{z,N}$ la fonction définie sur $\mathbb{C}^2 \simeq \mathbb{R}^4$ par $f_{z,N}(h,y) = y^N - q_{z,N}(h), (h,y) \in \mathbb{C}^2$. La 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 Nième 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 $h \to r_{z,N}\varphi_{z,N}(h)h$ est inversible au voisinage de $h = 0$ puisque sa différentielle 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 $\zeta \to \zeta^N$: 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

Définition 4.1. — Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ 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.

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

$m_0 = (a_1, a_2, \ldots, a_n)$ il existe une boule B centrée en $(a_1, a_2, \ldots, a_{n-1})$, un $\delta > 0$ et une application $\varphi : B \to (a_n - \delta, a_n + \delta)$ tel que $\frac{\partial f}{\partial x_n}(m)$ soit non nul sur $B \times (a_n - \delta, a_n + \delta)$ et que, pour $m = (x_1, x_2, \ldots, x_n) \in B \times (a_n - \delta, a_n + \delta)$, l'équation $f(m) = 0$ soit équivalente à $x_n = \varphi(x_1, x_2, \ldots, x_{n-1})$.

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 $f : U_0 \to \mathbb{R}$ 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.

Exemple 4.2. — Le cercle $x^2 + y^2 - 1 = 0$, l'hyperbole $x^2 - y^2 - 1 = 0$ sont des courbes régulières, mais pas le lieu $x^2 - y^2 = 0$.

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*

$$\frac{\partial f}{\partial x}(m)(X - x) + \frac{\partial f}{\partial y}(m)(Y - y) = 0, \quad (X, Y) \in \mathbb{R}^2. \tag{4}$$

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 (4) 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 (4) 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 $m = (\psi(y), y)$. □

4.3. Surfaces régulières

Définition 4.3. — 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 $f : U_0 \to \mathbb{R}$ 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 $g : V(\subset \mathbb{R}^2) \to \mathbb{R}$ 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 $x^2 - y^2 - z^2 = 0$.

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 $\Phi(S \cap U_0) = (\mathbb{R}^2 \times \{0\}) \cap \Phi(U_0)$.*

△ 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 $(\{0\}\times\mathbb{R}^2)\cap\Phi(U)$.

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. □

Pour S une partie de $\mathbb{R}^3$ et $m_0\in S$, on note par $\mathcal{C}_m(S)$ 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 $\gamma(-\varepsilon,\varepsilon)\subset S$.

Définition 4.4. — L'espace tangent $T_{m_0}S$ à S est l'ensemble des vecteurs dérivés $\dot\gamma(0)$ où γ est un arc de $\mathcal{C}_{m_0}(S)$.

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* $\operatorname{Ker} Df(m_0)$. *Le plan $T_{m_0}S$ a pour équation*

$$(5)\qquad \frac{\partial f}{\partial x}(m_0)X+\frac{\partial f}{\partial y}(m_0)Y+\frac{\partial f}{\partial z}(m_0)Z=0,\quad (X,Y,Z)\in\mathbb{R}^3.$$

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 $v_\pm=(0,\pm1,1\}$, vecteurs dérivés en $t=0$ des droites $v(t)=(0,\pm t,t)$ incluses dans C. Mais le vecteur v_++v_- 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 $t\neq0$ 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 $t\to(t,0,t)$ contenue dans C, les vecteurs v_+,v_-,w, linéairement indépendants, engendreraient un espace vectoriel tangent de dimension $3\ldots$

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 $t=0$, $Df(m_0)\dot\gamma(0)$ et donc l'inclusion $T_{m_0}S\subset\operatorname{Ker}Df(m_0)$.

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 $\mathcal{C}_{m_0}S$: $\gamma:t\in(-\varepsilon,\varepsilon)\to\Phi^{-1}(\Phi(m_0)+tv)$ 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 $\operatorname{Ker}Df(m_0)$ 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 $\operatorname{Ker}Df(m_0)$, dont l'équation est bien (5). □

Terminons par une définition, qui a une version en dimension $n=2$ (ou n quelconque bien sûr).

Définition 4.5. — Soit f différentiable sur U et m un point de U. Le gradient de f en m, noté $\operatorname{grad} f(m)$ ou $\nabla f(m)$, est le vecteur défini par

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

Ainsi, si $\langle\ ,\ \rangle$ désigne le produit scalaire standard sur $\mathbb{R}^3$, l'équation du plan tangent $T_{m_0}S$ d'une surface d'équation régulière $f = 0$ est

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

L'espace tangent $\tau_{m_0}S$ passant par m_0 est l'espace affine contenant m_0 de direction $T_{m_0}S$, il a pour équation

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

CHAPITRE 5

APPLICATIONS $\mathcal{C}^k$

Pour k entier, la définition 1.5 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 $f : U \to \mathbb{R}^p$ est de classe C^k (resp. $\mathcal{C}^\infty$) si et seulement si ses fonctions coordonnées f_i, $i = 1, \dots, p$ le sont.*

Démonstration. — On le démontre par récurrence. Pour $k = 1$, c'est la remarque précédant l'exemple 1.3. Supposons le lemme montré pour la régularité C^k et soit f de classe C^{k+1}. Ainsi f est différentiable et sa différentielle Df est de classe C^k. Les fonctions coordonnées de Df sont $\frac{\partial f_i}{\partial x_j}, i = 1, \dots, p, j = 1, \dots, n$, elles sont de classe C^k d'après l'hypothèse de récurrence. Par suite, pour $i = 1, \dots, p$, la fonction f_i est différentiable et, à nouveau d'après l'hypothèse de récurrence, Df_i, avec fonctions de coordonnées $\frac{\partial f_i}{\partial x_j}, j = 1, \dots, n$ est de classe C^k. Ainsi f_i est de classe C^{k+1}. Pour la réciproque, il suffit de remarquer que tous les arguments précédents sont en fait des équivalences. □

Exemple 5.1. — L'application $P_{npq} : (A, B) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q) \to B \circ A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^q)$ est de classe C^∞, vu que ses fonctions coefficients sont polynomiales. On calcule facilement sa différentielle au point (A, B)

$$dP_{npq}(A, B)(H, K) = B \circ H + K \circ A, \quad (H, K) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q),$$

et l'application dP_{npq} est une application linéaire (qui est C^∞).

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

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

Exemple 5.2. — L'application $A \in \mathrm{GL}(n,\mathbb{R}) \to A^{-1} \in \mathcal{M}_n$ 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 $\mathcal{C}^\infty$.

5.1. La version C^k des théorèmes classiques

Théorème 5.1. — *Si f et g sont de classe C^k (resp. $\mathcal{C}^\infty$), la composée $g \circ f$ 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, et aussi $D(g \circ f)$ vu comme composée des applications précédentes et de l'application $(A,B) \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^p,\mathbb{R}^q) \to B \circ A \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^q)$. Ainsi $D(g \circ f)$ est de classe C^k, et $g \circ f$ de classe C^{k+1}. □

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

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 $A \in Gl(n,\mathbb{R}) \to A^{-1}$. □

5.2. Théorème de Schwarz

***Théorème 5.3* (Schwarz)**. — *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ de classe C^2. Alors $\partial_{x_i}(\partial_{x_j} f) = \partial_{x_j}(\partial_{x_i} f)$. 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.*

Démonstration. — Il suffit de le montrer pour une fonction f de deux variables x,y différentiable à l'origine $(0,0)$. Soient F et H_1 définies au voisinage de l'origine par

$$F(x,y) = f(x,y) - f(x,0) - f(0,y) + f(0,0), \quad H_1(x,y) = F(x,y) - xy\partial_y\partial_x F(0,0).$$

On a $\partial_x H_1(x,y) = \partial_x F(x,y) - y\partial_y\partial_x F(0,0)$. Par ailleurs, la différentiabilité de $\partial_x F$ à l'origine permet d'écrire

$$\partial_x F(s,t) = \partial_x F(0,0) + \partial_x\partial_x F(0,0)s + \partial_y\partial_x F(0,0)t + (|s|+|t|)\varepsilon(s,t)$$

On a $\partial_x F(x,y) = \partial_x f(x,y) - \partial_x f(x,0)$, $\partial_x F(x,0) = 0$ et $\partial_x\partial_x F(0,0) = 0$. Ainsi la formule précédente se réduit à

$$\partial_x F(s,t) = t\partial_y\partial_x F(0,0) + (|s|+|t|)\varepsilon(s,t)$$

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

$$|H_1(x,y)| \leq \sup_{s\in[0,x]} |\partial_x F(s,y) - y\partial_y\partial_x F(0,0)|\,|x| \leq (|x|+|y|)\varepsilon_1(x,y)|x|$$

soit

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

On a une inégalité analogue en échangeant x et y

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

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

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

soit

$$|\partial_x\partial_y F(0,0) - \partial_y\partial_x F(0,0)| \leq 2(\varepsilon_1(x,y) + \varepsilon_2(x,y)),$$

et donc l'annulation du membre de gauche en faisant tendre (x, y) vers $(0, 0)$, ce qu'il fallait démontrer. □

5.3. Formules de Taylor

***Théorème 5.4* (Formule de Taylor à l'ordre 2).** — *Soit f de classe C^2. Alors*

$$f(m_0 + h) = f(m_0) + Df(m_0)h + D^2 f(m_0)(h, h) + o(\|h\|^2)$$

Démonstration. — On considère la fonction g d'une variable définie par $g(t) = f(m_0 + th)$ pour t dans un voisinage de $[0, 1]$.

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

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

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

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

$$D^n f(h, \ldots, h) = \sum_{(i_1, i_2, \ldots, i_n) \in \mathbb{N}_*^n} \frac{\partial^n f}{\partial x_{ci_1} \ldots \partial x_{i_n}} h_{i_1} \ldots h_{i_n}.$$

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

CHAPITRE 6

POINTS CRITIQUES ET EXTREMA

6.1. Extrema

***Définition 6.1* (maximum/minimum local).** — Soit $f : U(\subset \mathbb{R}^n) \to \mathbb{R}$. Le point $m_0 \in U$ est un *maximum* (resp. *minimum*) *local* pour f s'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. $\triangledown$

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 $Df(m_0)$ non nulle. Il existe une direction h telle que $Df(m_0)h$ soit non nulle. Alors

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

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

Théorème 6.1. — *Soit f de classe C^2, m_0 un point critique de f avec hessienne* $\operatorname{Hess} f(m_0)$ *de rang r et d'indice i.*

Si m_0 est un maximum (resp. minimum) local de f, alors $\operatorname{Hess}(f)(m_0)$ *est une forme quadratique négative (resp. positive).*

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.

Démonstration. — Soient $(v_i, \alpha_i), i = 1, \ldots, n$ les éléments propres de la hessienne $\operatorname{Hess}(f)(m_0)$ avec $(v_i)_{i=1,\ldots,n}$ une base orthonormée de $\mathbb{R}^n$: $\operatorname{Hess}(f)(m_0)(v_i, v_i) = \alpha_i$. On écrit alors le développement limité au voisinage de $v = 0$ pour $f(m_o + v)$ avec $v = \sum_{i=1}^n y_i v_i$

$$f\left(m_0 + \sum_{i=1}^n y_i v_i\right) = f(m_0) + \operatorname{Hess} f(m_0)\left(\sum_{i=1}^n y_i v_i\right) + o(\|v\|^2) = f(m_0) + \sum_{i=1}^n \alpha_i y_i^2 + o(\|v\|^2).$$

Si m_0 est un maximum local, on a en prenant $v = tv_j$ dans l'expression précédente,

$$0 \geq f(m_0 + tv_j) - f(m_0) = \alpha_j t^2 + o(t^2) = t^2(\alpha_j + \varepsilon(t))$$

au voisinage de $t = 0$, d'où en divisant par t et passant à la limite $t \to 0$, l'inégalité $0 \geq \alpha_j$: la hessienne $\operatorname{Hess} f(m_0)$ est négative.

Supposons la hessienne $\operatorname{Hess} f(m_0)$ non dégénérée. Si elle est définie négative (*i. e.* $i = n$), soit α le maximum des $\alpha_i, i = 1, \ldots, n$, qui est donc strictement négatif. On a, pour $v = \sum_{i=1}^n y_i v_i$,

$$f(m_0 + v) - f(m_0) = \sum_{i=1}^n \alpha_i y_i^2 + o(\|v\|^2) \leq \alpha \sum_{i=1}^n y_i^2 + o(\|v\|^2) = \|v\|^2(\alpha + \varepsilon(v))$$

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 $\operatorname{Hess} f(m_0)$ 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 f 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 pour $n = 2$ 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 $(Dh(m_0))_{|\operatorname{Ker}\ell} = 0$ est l'inclusion $V = \operatorname{Ker}\ell \subset \operatorname{Ker} Dh(m_0)$: il existe $\lambda \in \mathbb{R}$ tel que $Dh(m_0) = \lambda\ell$. En fait, λ est non nul si $\operatorname{Ker}\ell = \operatorname{Ker} Dh(m_0)$, nul si m_0 est un point critique de h.

Si L est l'intersection de deux hyperplans affines distincts d'équations $\ell_1(m) = \ell_1(m_0)$ et $\ell_2(m) = \ell_2(m_0)$ (*i. e.* une droite dans $\mathbb{R}^3$), alors $V = \operatorname{Ker}\ell_1 \cap \operatorname{Ker}\ell_2$ 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 $(\ell_i)_{i=1,\dots,n}$ est une base de formes linéaires complétant les deux formes indépendantes ℓ_1, ℓ_2, de base duale $(v_i)_{i=1,\dots,n}$, on a $Dh(m_0) = \sum_{i=1}^n \lambda_i\ell_i$ et les $\lambda_i = Dh(m_0)(v_i)$ sont tous nuls pour $i \geq 3$ puisque $v_i \in V$.

Le scalaire λ, ou les scalaires λ_1, λ_2, sont appelés *multiplicateur(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 6.2. — *Soit $S = \{f = 0\}$ une surface régulière dans $\mathbb{R}^3$. Le point $m_0 \in S$ est un extrema local de $h_{|S}$ s'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

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

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

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

Proposition 6.3. — *Soit $C = \{f = 0\}$ une courbe régulière dans $\mathbb{R}^2$. Le point $m_0 \in S$ est un extrema local de $h_{|S}$ 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_+. 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 $f \to +\infty$ au bord de Q_+, minimum strict vu que $\operatorname{Hess}(f)(m_-) = \begin{pmatrix} -c & 0 \\ 0 & -b \end{pmatrix}$. 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 proies/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 6.4. — *Soit $C = \{f_1 = 0, f_2 = 0\}$ une courbe régulière dans $\mathbb{R}^3$. Le point $m_0 \in S$ est une extrema local de $h_{|S}$ s'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 telle que C soit localement donnée par $\{(x, \varphi_1(x), \varphi_2(x))\}$. L'espace tangent à la courbe C est de direction $(1, \varphi_1'(x_0), \varphi_2'(x_0))$, égale à l'intersection $\operatorname{Ker} Df_1(m_0) \cap \operatorname{Ker} Df_2(m_0)$. 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

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

ainsi $\operatorname{Ker} Df_1(m_0) \cap \operatorname{Ker} Df_2(m_0)$ est inclus dans $\operatorname{Ker} Dh(m_0)$ 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 $h : (x, y, z) \to z$ 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. 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 $\langle\ ,\ \rangle$, la donnée de la 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 $(v_i)_{i=1,\ldots,n}$. Ainsi, si α_i est la valeur propre du vecteur propre v_i de la matrice A, associée à la forme bilinéaire déterminée par q, on a

$$q\left(\sum_{i=1}^n y_i v_i\right) = \sum_{i=1}^n \alpha_i y_i^2.$$

Le nombre de α_i 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

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

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