UNIVERSITE PARIS 7- DENIS DIDEROT

CALCUL DIFFERENTIEL

M. Joseph le Potier

Ce polycopie à été effectué à la reprographie du second cycle par Mme GOEMINNE
TOUR 56/66 1er ETAGE -2 Place Jussieu- 75251 PARIS Cedex 5

Calcul différentiel

Joseph Le Potier

TABLE DES MATIÈRES

1. Applications différentiables 2
1.1. Différentielles et différentielles partielles 2
1.2. Théorème des accroissements finis 11
2. Différentielles d'ordre supérieur 15
2.1. Applications bilinéaires 15
2.2. La différentielle seconde 16
2.3. Formules de Taylor 18
2.4. Le théorème de symétrie de Schwarz 20
2.5. Applications de classe C^r, avec $r \geq 3$. 22
3. Difféomorphismes 25
3.1. Le théorème d'inversion locale 25
3.2. Théorème des fonctions implicites 30
4. Théorème du rang 31
4.1. Le rang 31
4.2. Théorème de submersion 32
4.3. Théorème d'immersion 34
4.4. Subimmersion 34
5. Sous-variétés de $\mathbb{R}^n$ 36
5.1. La notion de sous-variété 36
5.2. Espaces tangents 38
5.3. Sous-variétés définies par des équations 39
5.4. Sous-variétés définies par un paramétrage 42
5.5. Morphismes 45
6. Fonctions de Morse 50

6.1. Lemme de Morse 51
6.2. Fonctions définies sur des sous-variétés de $\mathbb{R}^n$ 54

Le but principal de ce cours est de démontrer le théorème d'inversion locale, qui donne des conditions pour qu'une application différentiable soit au voisinage d'un point, un difféomorphisme. Nous introduirons à partir de là les sous-variétés de dimension m de $\mathbb{R}^n$, objets qui ressemblent localement à des ouverts de $\mathbb{R}^m$, et sur lesquelles la notion de fonction différentiable a encore un sens. Nous pourrons alors examiner sur ces sous-variétés les problèmes d'extrema des fonctions différentiables, dits extremas avec contrainte, et plus généralement, nous expliquerons comment trouver les points critiques des fonctions différentiables sur ces sous-variétés, et l'intérêt topologique que peut avoir l'étude des points critiques, avec notamment les fonctions de Morse. En premier lieu, nous rappelons les notions de base du calcul différentiel.

Tous les espaces vectoriels considérés seront supposés de dimension finie. On les suppose munis de normes. Rappelons que toutes les normes sont équivalentes. Si E et F sont deux espaces vectoriels normés, on désigne par $\mathrm{L}(\mathrm{E},\mathrm{F})$ l'espace vectoriel des applications linéaires de E dans F. Cet espace vectoriel est équipé de la norme suivante : $\| u \| = \sup_{\|x\|=1} \| u(x) \|$ de sorte que pour tout vecteur $x \in \mathrm{E}$, on a la majoration

$$\| u(x) \| \leq \| u \| \| x \| .$$

Etant donnée une application linéaire $u : \mathrm{E} \to \mathrm{F}$, et $x \in \mathrm{E}$, on écrit souvent $u.x$ pour $u(x)$.

1. Applications différentiables

1.1. Différentielles et différentielles partielles

Soit U un ouvert d'un espace vectoriel E, et $f : \mathrm{U} \to \mathrm{F}$ une application définie sur U.

DÉFINITION 1. — *Soit $a \in \mathrm{U}$. On dit que f est négligeable devant $\| x - a \|$ quand $x \to a$ si pour tout réel $\epsilon > 0$, il existe un voisinage ouvert $\mathrm{V} \subset \mathrm{U}$ de a tel que pour tout $x \in \mathrm{V}$ on ait*

$$\| f(x) \| \leq \epsilon \| x - a \|$$

Cette définition est évidemment indépendante du choix des normes choisies sur E et F. Si f est négligeable devant $\| x - a \|$ quand $x \to a$, on a $f(a) = 0$, et la fonction f est continue en a. L'énoncé est équivalent au suivant : il existe une fonction $x \mapsto \epsilon(x)$ définie au voisinage de a telle que $\lim_{x \to a} \epsilon(x) = 0$ et telle que

$$\| f(x) \| \leq \epsilon(x) \| x - a \| .$$

DÉFINITION 2. — *Soient f et $g : \mathrm{U} \to \mathrm{F}$ deux applications définies sur un voisinage ouvert U d'un point a, à valeurs dans un espace vectoriel F. On dit que f et g sont tangentes en a si $f - g$ est négligeable devant $\| x - a \|$ quand $x \to a$.*

Dans l'ensemble des applications définies au voisinage de a, la relation «f et g sont tangentes en a » est une relation d'équivalence.

DÉFINITION 3. — *Une application $f : \mathrm{E} \to \mathrm{F}$ est dite affine l'application $u : x \to f(x) - f(0)$ est linéaire.*

Une telle application affine s'écrit donc de manière unique $f(x) = a + u.x$, avec $u \in \mathrm{L}(\mathrm{E}, \mathrm{F})$. L'application linéaire u s'appelle la différentielle de f; elle est notée df.

Exemples

Les translations $x \mapsto x + a$ sont des applications affines. Le composé $g \circ f$ de deux applications affines $f : \mathrm{E} \to \mathrm{F}$ et $g : \mathrm{F} \to \mathrm{G}$ est une application affine, et on a $d(g \circ f) = dg \circ df$.

PROPOSITION 4. — *Considérons la relation d'équivalence définie ci-dessus dans l'ensemble des applications définies au voisinage de a. Dans chaque classe d'équivalence, il existe au plus une application affine.*

Démonstration. On peut supposer, quitte à faire une translation, que $a = 0$. Il s'agit de montrer que si f est une application affine, négligeable devant $\| x \|$ quand $x \to 0$, alors $f = 0$. On a évidemment $f(0) = 0$. Donc f est une application linéaire. Mais alors il découle de la définition pour tout $\epsilon > 0$, il existe une boule $\mathrm{B}(0, r)$ dans E telle que pour $\| x \| \leq r$ on ait

$$\| f(x) \| \leq \epsilon \| x \| .$$

Mais cette inégalité reste conservée quand on change x en λx, et par conséquent, elle est vraie pour tout vecteur de E. Par suite, $\| f \| \leq \epsilon$. Puisque ceci est vrai pour tout $\epsilon > 0$, ceci implique que $\| f \| = 0$ et par suite $f = 0$. □

DÉFINITION 5. — *Soit $f : \mathrm{U} \to \mathrm{F}$ une application définie au voisinage* U *de $a \in \mathrm{E}$. On dit que f est différentiable en a si f est tangente en a à une application affine.*

Si f est différentiable au point a, il y a une seule application affine g tangente à f au point a. Sa différentielle dg s'appelle la *différentielle de f au point a*, et est notée $d_a f$ ou $df(a)$. Dans la pratique, il est souvent commode de se ramener au cas $a = 0$, ce qu'on peut faire en remarquant que f est différentiable au point a si et seulement si il existe une application linéaire $u : \mathrm{E} \to \mathrm{F}$ telles que les fonctions, définies au voisinage de a, $h \mapsto f(a + h)$ et $h \mapsto f(a) + u(h)$ soient tangentes en 0. Ceci signifie donc qu'étant donné un nombre réel $\epsilon > 0$, il existe un voisinage ouvert W de 0 dans E tel que pour $h \in \mathrm{W}$, on ait $a + h \in \mathrm{U}$ et

$$\| f(a + h) - f(a) - u(h) \| \leq \epsilon \| h \| .$$

Bien sûr, l'application linéaire u est la différentielle $d_a f$.

Il résulte de la définition qu'une application affine f est différentiable en tout point, et que la différentielle en un point a est en fait indépendante de a.

REMARQUE 6. — Si f est différentiable en a, f est évidemment continue en a. En effet, une application affine g est continue, et on a déjà vu que si f et g sont tangentes en a, la différence $f - g$ est continue en a.

REMARQUE 7. — Si $\mathrm{E} = \mathbb{R}$, la différentiabilité de $f : \mathrm{U} \to \mathrm{F}$ en a est équivalente à l'existence d'une dérivée $f'(a)$ pour la fonction f; la différentielle

$d_a f \in \mathrm{L}(\mathbb{R}, \mathrm{F})$ est alors donnée pour $h \in \mathbb{R}$ par la formule

$$d_a f(h) = hf'(a).$$

En effet, si f est différentiable en a, sa différentielle s'écrit $d_a f.h = hv$, où v est un vecteur de F; on a alors

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = v.$$

Réciproquement, si f est dérivable en a de dérivée v, la fonction $f(a+h) - f(a) - hv$ est négligeable devant $\| h \|$ quand $h \to 0$, et donc f est différentiable en a de différentielle $h \mapsto hv$.

DÉFINITION 8. — *Soit* $f : \mathrm{U} \to \mathrm{F}$ *une application définie sur un ouvert* U *d'un espace vectoriel* E. *On dit que* f *est différentiable sur* U *si* f *est différentiable en tout point de* U. *On dit que* f *est de classe* C^1 *si l'application* f *est différentiable, et si l'application* $df : \mathrm{U} \to \mathrm{L}(\mathrm{E}, \mathrm{F})$ *qui associe à un point* x *sa différentielle* $d_x f$ *est continue.*

Applications à valeurs dans un produit

PROPOSITION 9. — *Soit*

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$$

une application $\mathrm{U} \to \mathrm{F} = \mathrm{F}_1 \times \ldots \times \mathrm{F}_m$ *une application définie au voisinage* U *d'un point* $a \in \mathrm{E}$. *Alors* f *est différentiable en* a *si et seulement si chacune des applications* f_i *est différentiable en* a, *et sa différentielle* $d_a f$ *est l'application linéaire*

$$h \mapsto \begin{pmatrix} d_a f_1.h \\ \vdots \\ d_a f_m.h \end{pmatrix}$$

Démonstration. On choisit une norme sur chacun des espaces vectoriels F_i, et on équipe l'espace vectoriel F de la norme $\| y \| = \sup_{i=1,\dots,m} \| y_i \|$ pour

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in F.$$

Considérons une application affine $g : E \to F$. Les composantes g_i de g sont encore des applications affines, et on a pour tout $x \in U$

$$\| f(x) - g(x) \| = \sup_{i=1,\dots,m} \| f_i(x) - g_i(x) \| .$$

Par conséquent, si f et g sont tangentes en a alors seulement si f_i et g_i sont tangentes en a. Il en découle que si f est différentiable en a alors f_i l'est aussi, et que sa différentielle $d_a f_i$ est la i-ème composante de $d_a f$. Réciproquement, supposons que pour tout i l'application f_i soit différentiable en a, et considérons l'application linéaire u de composantes $d_a f_i$. Alors l'application affine $x \to f(a) + u(x - a)$ est tangente à f en a, et par conséquent f est différentiable en a, de différentielle $d_a f = u$.

Différentielles partielles

Soit $f : U \to F$ une application définie au voisinage U d'un point a. Soit $S \subset E$ un sous-espace vectoriel de E. On dit que f a au point a une différentielle partielle dans la direction S si l'application, définie sur le voisinage ouvert de 0 des points $h \in S$ tels que $a + h \in U$ par $h \to f(a + h)$ est différentiable en 0. Si c'est le cas, la différentielle de cette fonction en 0 est notée $\partial_S f(a)$.

PROPOSITION 10. — *Si f est différentiable en a, alors f a une différentielle partielle au point a dans la direction* S *et on a* $\partial_S f(a) = d_a f|_S$.

Démonstration. Par définition, les applications $h \to f(a + h)$ et $h \to f(a) + d_a f.h$ sont tangentes en 0. Quand on les compose par l'inclusion $F \hookrightarrow E$, les applications obtenues sont encore tangentes. L'énoncé en découle. □

Ceci s'applique bien sûr quand

$$E = \oplus_{i=1}^{n} E_i$$

est une somme directe de sous-espaces E_i, à chacun des sous-espaces vectoriels E_i. Dans ce cas, la différentielle partielle $\partial_{\mathrm{E}_i} f(a)$ est plutôt notée $\partial_i f(a)$. Rappelons qu'une application linéaire $u : \mathrm{E} \to \mathrm{F}$ est déterminée par ses restrictions $u_i = u|_{\mathrm{E}_i}$: l'image d'un vecteur x s'écrivant dans cette somme directe $x = \sum_{i=1}^n x_i$, avec $x_i \in \mathrm{E}_i$, est le vecteur $\sum_{i=1}^n u_i(x_i)$. Une telle application sera notée comme une matrice ligne

$$(u_1, \ldots, u_n)$$

Les éléments de cette matrice ne sont plus des scalaires, mais des applications linéaires.

PROPOSITION 11. — *On considère une application $f : \mathrm{U} \subset \mathrm{E} \to \mathrm{F}$ définie sur un ouvert d'un espace vectoriel* E *s'écrivant comme somme directe*

$$\mathrm{E} = \oplus_{i=1}^n \mathrm{E}_i.$$

Soit $a \in \mathrm{U}$. Supposons que f soit différentiable en a. Alors f a des dérivée partielles en a dans la direction E_i, notée $\partial_i f(a)$, et la différentielle $d_a f$ est donnée par la matrice

$$(\partial_1 f(a), \ldots, \partial_p f(a))$$

Quand les espaces vectoriels E_i sont de dimension 1, chacun d'entre eux étant muni d'un vecteur e_i non nul, les vecteurs $(e_1, \ldots, e_n)$ constitue une base de E. On note $(x_1, ..., x_n)$ les coordonnées d'un point dans cette base. La différentiabilité partielle de f dans la direction E_i équivaut à la dérivabilité de $t \mapsto f(a + te_i)$ à l'origine. La différentielle partielle $\partial_i f(a)$ est alors donnée par sa valeur en e_i, et coïncide avec cette dérivée à l'origine ; on l'appelle la dérivée partielle au point a dans la direction e_i, et on la note

$$\frac{\partial f}{\partial x_i}(a).$$

En général l'existence de différentielles partielles de f en un point a ne permet pas de conclure à la différentiabilité de f en ce point. Il faudra donc rajouter des hypothèses supplémentaires si on veut obtenir, à partir de l'existence de dérivées partielles, la différentiabilité de f (*cf.* théorème 17).

Exercice 1.1

Considérons sur $\mathbb{R}^2$ l'application f définie par

$$\begin{cases} f(x,y) = & \dfrac{xy}{x^2+y^2} \quad \text{si } (x,y) \neq 0 \\ f(0) = & 0 \end{cases}$$

Cette application a des dérivées partielles à l'origine, mais elle n'est pas continue en 0.

Exercice 1.2

On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$\begin{cases} f(x,y) = & \dfrac{xy^2}{x^2+y^4} \quad \text{si } (x,y) \neq 0 \\ f(0) = & 0 \end{cases}$$

Soit $e \in \mathbb{R}^2$ un vecteur non nul. Montrer que pour tout e, cette fonction f a une dérivée partielle dans la direction e. Vérifier que cette fonction n'est pas continue à l'origine.

On peut combiner les deux énoncés ci-dessus. Supposons données deux décompositions en somme directe

$$\mathrm{E} = \oplus_{i=1}^{n} \mathrm{E}_i \ ; \ \mathrm{F} = \oplus_{j=1}^{m} \mathrm{F}_j$$

et considérons une application $f = \mathrm{U} \to \mathrm{F}$ définie sur un voisinage ouvert U d'un point a, à valeurs dans F, différentiable en a. On désigne par $(f_1, \ldots, f_m)$ les composantes de f. Alors f_j a des différentielles partielles $\partial_i f_j(a)$ et la matrice de la différentielle $df(a)$ de f en a dans les sommes directes ci-dessus est la matrice

$$(\partial_i f_j(a))$$

(où i est l'indice de colonne et j l'indice de ligne) qu'on appelle la matrice jacobienne de f en a. Ceci signifie que cette différentielle est donnée pour $h = \sum_i h_i$, où $h_i \in \mathrm{E}_i$ par

$$h \mapsto \sum_{i,j} \partial_i f_j(a) h_i.$$

Composition des applications différentiables

PROPOSITION 12. — *Soient* E, F *et* G *trois espaces vectoriels, et* $\mathrm{U} \subset \mathrm{E}$ *et* $\mathrm{V} \subset \mathrm{F}$ *deux ouverts. On considère deux applications* $f : \mathrm{U} \to \mathrm{V}$ *et* $g : \mathrm{V} \to \mathrm{G}$, *et on suppose que* f *est différentiable en* a *et que* g *est différentiable en* $b = f(a)$. *Alors* $g \circ f$ *est différentiable en* a *et on a*

$$d_a(g \circ f) = d_b g \circ d_a f.$$

Démonstration. On choisit des normes sur les trois espaces vectoriels. Soient $u = d_a f$ et $v = d_b g$. On écrit, pour $a + h \in \mathrm{U}$ et $b + k \in \mathrm{V}$

$$\begin{aligned} f(a+h) =& f(a) + u(h) + \alpha(h) \\ g(b+k) =& g(b) + v(k) + \beta(k) \end{aligned}$$

Par hypothèse, $\alpha(h)$ est négligeable devant $\| h \|$ quand $h \to 0$, et $\beta(k)$ est négligeable devant $\| k \|$ quand $k \to 0$. Alors pour $a + h \in \mathrm{U}$, on a

$$g(f(a+h)) = g(b) + v(u(h)) + v(\alpha(h)) + \beta(u(h) + \alpha(h))$$

On va montrer que les deux derniers termes sont négligeables devant $\| h \|$ quand $h \to 0$, ce qui entraînera que $g \circ f$ est différentiable en a, de différentielle $v \circ u$. Puisque v est une application linéaire, la majoration

$$\| v(\alpha(h)) \| \leq \| u \| \| \alpha(h) \|$$

montre que $v(\alpha(h))$ est négligeable devant $\| h \|$ quand $h \to 0$. On peut écrire d'autre part $\beta(k) = \epsilon(k) \| k \|$, où ϵ est une fonction à valeurs dans $\mathbb{R}_+^*$ définie sur le voisinage ouvert de 0 dans F des points k tels que $b + k \in \mathrm{V}$, telle que $\lim_{k \to 0} \epsilon(k) = 0$. Alors

$$\| \beta(u(h) + \alpha(h)) \| \leq \epsilon(u(h) + \alpha(h)) \| u(h) + \alpha(h) \|$$

La fonction $h \mapsto \epsilon(u(h) + \alpha(h))$ tend vers 0 quand $h \to 0$. D'autre part, on a la majoration $\| u(h) + \alpha(h) \| \leq \| u \| \| h \| + \| \alpha(h) \|$. Soit $\mathrm{B}(a, r)$ une boule ouverte contenue dans U et telle que pour $\| h \| \leq r$ on ait $\| \alpha(h) \| \leq \| h \|$. On a alors pour $\| h \| \leq \rho$

$$\| \beta(u(h) + \alpha(h)) \| \leq \epsilon(u(h) + \alpha(h))(\| u \| + 1) \| h \|$$

ce qui implique que la fonction $h \mapsto \beta(u(h)+\alpha(h))$ est négligeable devant 0 quand $h \to 0$. □

Opérations sur les fonctions différentiables

Soient f et g deux applications définie sur un ouvert U d'un espace vectoriel E, à valeurs dans un espace vectoriel F, différentiables en a, et $\alpha : \mathrm{U} \to \mathbb{R}$ une fonction différentiable en a.

(1) La somme $f+g$ est différentiable en a, de différentielle

$$d_a(f+g) = d_a f + d_a g.$$

(2) Le produit αf est différentiable en a, de différentielle

$$d_a(\alpha f) = (d_a\alpha)f + \alpha(a)d_a f.$$

(3) Si $\alpha(a) \neq 0$, la fonction $\dfrac{f}{\alpha}$ est différentiable en a, de différentielle

$$d_a(\frac{f}{\alpha}) = \frac{\alpha(a)d_a f - (d_a\alpha)f(a)}{\alpha(a)^2}.$$

Exercice 1.3

Soient $\mathrm{E}, \mathrm{F}, \mathrm{G}$ trois espaces vectoriels (de dimension finie), et $\mathrm{B} : \mathrm{E} \times \mathrm{F} \to \mathrm{G}$ une application bilinéaire. Alors l'application B est différentiable en tout point $(a,b) \in \mathrm{E} \times \mathrm{F}$. La différentielle $d_{(a,b)}\mathrm{B}$ est l'application linéaire $\mathrm{E} \times \mathrm{F} \to \mathrm{G}$ définie par

$$(h,k) \mapsto \mathrm{B}(h,b) + \mathrm{B}(a,k).$$

Généraliser cet énoncé au cas d'une application multilinéaire. Démontrer que tout polynôme $\mathrm{P} : \mathrm{E} \to \mathbb{R}$ est différentiable en tout point.

Exercice 1.4

Soit End(E) l'espace vectoriel des endomorphismes d'un espace vectoriel E. L'application End(E) $\to$ End(E) définie par $u \mapsto u^n$ est différentiable de différentielle en $a \in \mathrm{End(E)}$ l'application linéaire

$$h \mapsto ha^{n-1} + aha^{n-2} + \ldots + a^{n-1}h$$

Exercice 1.5

Soit GL(E) l'ensemble des applications linéaires inversibles E $\to$ E. Vérifier que GL(E) est un ouvert, et l'application GL(E) $\to$ GL(E) définie par $u \mapsto u^{-1}$ est différentiable, de différentielle $h \mapsto -a^{-1}ha^{-1}$ au point a.

Exercice 1.6

1) Soit $f(x) = \sum_n a_n x^n$ une série entière à coefficients réels, de rayon de convergence ρ. Soit E un espace vectoriel normé de dimension finie ; on munit l'espace vectoriel End(E) de la norme associée. On considère l'application, définie sur la boule $B(0, \rho)$ de End(E), à valeurs dans End(E), par

$$u \mapsto f(u) = \sum_{n \geq 0} a_n u^n$$

Démontrer que f est différentiable.

2) Démontrer que l'application End(E) $\to$ End(E) définie par

$$u \mapsto \exp(u) = \sum_n \frac{u^n}{n!}$$

est différentiable, et que son image contient la boule ouverte $B(1, 1)$.

Exercice 1.7

Soit E un espace vectoriel normé de dimension finie ; on munit l'espace vectoriel End(E) de la norme associée. Soit $u \in$ End(E) tel que $\| u \| < 1$; montrer que l'endomorphisme $1 + u$ est inversible.

1.2. Théorème des accroissements finis

Soient E et F deux espaces vectoriels normés de dimension finie. On considère une fonction différentiable f : U $\to$ F sur un ouvert convexe U de E. Etant donnés deux points a et $b \in$ U, on se pose le problème est de majorer la distance $\| f(b) - f(a) \|$. Il s'agit donc d'étendre le théorème de Rolle. On commence par le cas d'une fonction dérivable sur un intervalle de $\mathbb{R}$.

LEMME 13. — *Soit* $\varphi : [0, 1] \to$ F *une fonction dérivable, dont la dérivée* $\varphi'(t)$ *est bornée sur* $[0, 1]$. *Alors*

$$\| \varphi(1) - \varphi(0) \| \leq \sup_{0 \leq t \leq 1} \| \varphi'(t) \|$$

Démonstration. Soit m un nombre tel que $\| \varphi'(t) \| < m$ pour tout $t \in [0,1]$. Considérons l'ensemble A_m des points $t \in [0,1]$ tels que $\| \varphi(t) - \varphi(0) \| \leq mt$. Comme $t \mapsto \| \varphi(t) - \varphi(0) \|$ est continue, A_m est un fermé, et il est non vide, car il contient 0. Il a donc une un plus grand élément c. On va montrer que $c = 1$, ce qui entraînera évidemment le résultat attendu. Supposons $c < 1$. On a par définition

$$\lim_{h \to 0} \| \frac{\varphi(c+h) - \varphi(c)}{h} \| = \| \varphi'(c) \|$$

et donc pour $h > 0$ assez petit on aura

$$\frac{\| \varphi(c+h) - \varphi(c) \|}{h} \leq m$$

ce qui implique

$$\| \varphi(c+h) - \varphi(0) \| \leq m(c+h).$$

Alors $c + h \in \mathrm{A}_m$; mais ceci contredit la définition de c. Donc $c = 1$, et par suite $\| \varphi(1) - \varphi(0) \| \leq m$. Ceci devant être vrai pour tout m, on obtient le résultat attendu. □

REMARQUE 14. — On sait d'après le théorème de Rolle que si $\varphi : [0,1] \to \mathbb{R}$ est une fonction dérivable, il existe un nombre réel $0 < \theta < 1$ tel que

$$\varphi(1) - \varphi(0) = \varphi'(\theta).$$

Cependant, cet énoncé ne s'étend pas aux fonctions dérivables à valeurs vectorielles : par exemple, la fonction $t \to e^{2i\pi t} : [0,1] \to \mathbb{C}$ est dérivable, de dérivée non nulle en tout point, et pourtant $\varphi(0) = \varphi(1)$.

THÉORÈME 15. — *(Théorème des accroissements finis)*

Soit U *un ouvert convexe de l'espace vectoriel* E, *et* $f : \mathrm{U} \to \mathrm{F}$ *une application différentiable telle que* $df : x \mapsto d_x f$ *soit bornée sur* U. *On pose*

$$\| df \| = \sup_{x \in \mathrm{U}} \| d_x f \| .$$

Alors pour tout a *et* $b \in \mathrm{U}$, *on a*

$$\| f(a) - f(b) \| \leq \| df \| \, \| a - b \|$$

Démonstration. On applique le lemme qui précède à la fonction $\varphi : [0,1] \to \mathrm{F}$ définie par $\varphi(t) = f((ta + (1-t)b)$. Cette application est dérivable, de dérivée

$$\varphi'(t) = d_{ta+(1-t)b}f.(a-b)$$

au point t. Il en résulte que pour tout $t \in [0,1]$, on a $\| \varphi'(t) \| \leq \| df \| \| a - b \|$. Comme $\varphi(1) = f(a)$ et $\varphi(0) = f(b)$, ceci conduit à l'énoncé. □

COROLLAIRE 16. — *Sur un ouvert connexe, une application différentiable dont la différentielle est nulle est constante.*

Démonstration. Sur un ouvert convexe, une fonction différentiable dont la différentielle est nulle est évidemment constante d'après l'énoncé ci-dessus. Il en résulte que sur un ouvert quelconque U, une fonction différentiable f dont la différentielle est nulle est localement constante. Ceci implique que si $a \in \mathrm{U}$ l'ensemble des points $x \in \mathrm{U}$ tels que $f(x) = f(a)$ est à la fois un ouvert et un fermé. Si l'ouvert U est connexe, cet ensemble est exactement l'ouvert U. □

Retour aux différentielles partielles

THÉORÈME 17. — *On suppose que* $\mathrm{E} = \oplus_{i=1}^{n} \mathrm{E}_i$, *et on considère une fonction* $f : \mathrm{U} \to \mathrm{F}$ *définie que un ouvert* U *de* E. *Alors* f *est de classe* C^1 *si et seulement si* f *a sur* U *des différentielles partielles* $x \mapsto \partial_i f(x)$ *continues sur* U.

Démonstration. Si f est de classe C^1, on a vu que f a des dérivées partielles $\partial_i f(x)$ en tout point $x \in \mathrm{U}$, et on a $\partial_i f(x) = d_x f|_{\mathrm{E}_i}$. L'application restriction $\mathrm{L}(\mathrm{E},\mathrm{F}) \to \mathrm{L}(\mathrm{E}_i,\mathrm{F})$ qui associe à u la restriction $u|_{\mathrm{E}_i}$ est évidement linéaire donc continue. Il en résulte que l'application $\mathrm{U} \to \mathrm{L}(\mathrm{E}_i,\mathrm{F})$ définie par $x \mapsto \partial_i f(x)$ est continue.

C'est dans l'autre sens que la question est plus intéressante. Supposons que f admette des différentielles partielles continues $x \mapsto \partial_i f(x)$. Il s'agit de montrer que f est de classe C^1. On se ramène immédiatement par récurrence à démontrer l'énoncé au cas de deux facteurs : $\mathrm{E} = \mathrm{E}_1 \oplus \mathrm{E}_2$. On peut supposer, quitte à changer au besoin la norme sur E, que l'on a pour $h \in \mathrm{E}_1$ et $k \in \mathrm{E}_2$

$$\| h + k \| = \| h \| + \| k \|$$

Montrons d'abord l'existence d'une différentielle en un point a. La différentielle en un point a, si elle existe, est obligatoirement l'application linéaire

$$h + k \mapsto \partial_1 f(a).h + \partial_2 f(a).k$$

(où $h \in \mathrm{E}_1$ et $k \in \mathrm{E}_2$). On se fixe une boule $\mathrm{B}(a, \rho) \subset \mathrm{U}$ centrée en a. On doit alors majorer, pour $\| h + k \|$ assez petit, la quantité

$$\| f(a + h + k) - f(a) - (\partial_1 f(a).h + \partial_2 f(a).k) \|$$

On majore cette quantité par

$$\| f(a + h + k) - f(a + h) - \partial_2 f(a).k \| + \| f(a + h) - f(a) - \partial_1 f(a).h \| .$$

Le second terme ne pose aucun problème : soit en effet ϵ un réel > 0. Par définition de la différentielle partielle de $\partial_1 f(a)$, on peut trouver un réel $0 < r < \rho$ tel que pour $h \in \mathrm{E}_1$ et $\| h \| \leq r$ on ait

$$\| f(a + h) - f(a) - \partial_1 f(a).h \| \leq \frac{\epsilon}{2} \| h \|$$

Pour majorer le premier terme, on considère, h étant fixé de norme $\| h \| < \rho$, la fonction $g(v) = f(a+h+v) - \partial_2 f(a).v$, définie pour $v \in \mathrm{E}_2$ dans la boule $\| v \| < \rho$. Cette application est différentiable, de différentielle en v

$$d_v g = \partial_2 f(a + h + v) - \partial_2 f(a)$$

Il en résulte de l'hypothèse de continuité de $\partial_2 f$ en a qu'il existe un réel $0 < \eta \leq r$ tel que pour $\| u \| \leq \eta$ on ait

$$\| \partial_2 f(a + u) - \partial_2 f(a) \| \leq \frac{\epsilon}{2}$$

Maintenant, on peut appliquer le théorème des accroissements finis à la fonction g : il entraîne pour $\| h + k \| \leq \eta$

$$\| f(a + h + k) - f(a + h) - \partial_2 f(a) k \| \leq \frac{\epsilon}{2} \| k \|$$

Alors pour $\| h + k \| \leq \eta$

$$\| f(a + h + k) - f(a) - (\partial_1 f(a)(h) + \partial_2 f(a)(k)) \| \leq \epsilon \| h + k \| .$$

Ainsi, f est différentiable en tout point a et sa différentielle $d_a f$ est donnée dans la somme directe $\mathrm{E} = \mathrm{E}_1 \oplus \mathrm{E}_2$ par la matrice

$$(\partial_1 f(a) \quad \partial_2 f(a))$$

Puisque les différentielles partielles $x \mapsto \partial_i f(x)$ sont des fonctions continues, il résulte de l'isomorphisme d'espaces vectoriels $\mathrm{L}(\mathrm{E}, \mathrm{F}) \simeq \mathrm{L}(\mathrm{E}_1, \mathrm{F}) \times \mathrm{L}(\mathrm{E}_2, \mathrm{F})$ que l'application $x \mapsto d_x f$ est elle aussi continue. Ainsi f est de classe C^1. Ceci achève la démonstration. □

2. Différentielles d'ordre supérieur

2.1. Applications bilinéaires

Soient E_1, E_2 et F trois espaces vectoriels de dimension finie, et normés. On désigne par $\mathrm{L}(\mathrm{E}_1, \mathrm{E}_2; \mathrm{F})$ l'espace vectoriel des applications bilinéaires $b : \mathrm{E}_1 \times \mathrm{E}_2 \to \mathrm{F}$. Puisque l'on est en dimension finie, une telle application est évidemment continue, et donc bornée sur le compact constitué du produit des boules unité. En particulier

$$\| b \| = \sup_{\substack{\|x\|=1 \\ \|y\|=1}} (\| b(x, y) \|$$

a un sens. Ceci définit une norme sur l'espace vectoriel $\mathrm{L}(\mathrm{E}_1, \mathrm{E}_2; \mathrm{F})$. Pour tout $h \in \mathrm{E}_1$ et $k \in \mathrm{E}_2$, on a l'inégalité

$$\| b(h, k) \| \leq \| b \| \| h \| \| k \|$$

Lemme 1. — *Il existe une isométrie canonique*

$$\mathrm{L}(\mathrm{E}_1, \mathrm{E}_2; \mathrm{F}) \simeq \mathrm{L}(\mathrm{E}_1, \mathrm{L}(\mathrm{E}_2, \mathrm{F}))$$

Démonstration. Soit $f : \mathrm{E}_1 \to \mathrm{L}(\mathrm{E}_2, \mathrm{F}))$ une application linéaire. A cette application linéaire, on associe l'application bilinéaire $b : \mathrm{E}_1 \times \mathrm{E}_2 \to \mathrm{F}$ définie par

$$b(h, k) = f(h)(k).$$

Réciproquement, à une application bilinéaire $b : \mathrm{E}_1 \times \mathrm{E}_2 \to \mathrm{F}$ on associe l'application linéaire f définie de la manière suivante : pour $h \in \mathrm{E}_1$, l'application linéaire $f(h) : \mathrm{E}_2 \to \mathrm{F}$ est $k \mapsto b(h,k)$. On a alors $b(h,k) = f(h)(k)$ pour tout $(h,k) \in \mathrm{E}_1 \times \mathrm{E}_2$. Il est clair que les applications $f \mapsto b$ et $b \mapsto f$ sont linéaires et inverses l'une de l'autre.

Il reste à vérifier que cette correspondance est une isométrie. Soit $f \in \mathrm{L}(\mathrm{E}_1, \mathrm{L}(\mathrm{E}_2, \mathrm{F}))$, et b l'application bilinéaire associée. Pour $h \in \mathrm{E}_1$ et $k \in \mathrm{E}_2$ de norme 1 on a $\| f(h)(k) \| \leq \| f(h) \| \leq \| f \|$, et par suite $\| b \| \leq \| f \|$. Pour obtenir l'égalité, on remarque qu'il existe $h \in \mathrm{E}_1$ de norme 1 tel que $\| f(h) \| = \| f \|$ et pour ce vecteur h un vecteur $k \in \mathrm{E}_2$ de norme 1 tel que $\| f(h)(k) \| = \| f(h) \|$. Alors $\| f(h)(k) \| = \| f \|$, et par suite $\| f \| \leq \| b \|$. Ainsi, $\| f \| = \| b \|$ ce qui termine la preuve du lemme. □

2.2. La différentielle seconde

Dans tout cette section, on se donne un ouvert U d'un espace vectoriel E, et $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^1 à valeurs dans un espace vectoriel F.

DÉFINITION 2. — *On dit que f est deux fois différentiable en un point $a \in \mathrm{U}$ si la différentielle $df : \mathrm{U} \to \mathrm{L}(\mathrm{E}, \mathrm{F})$ est différentiable en a.*

Si f est deux fois différentiable en a, la différentielle $d_a(df) \in \mathrm{L}(\mathrm{E}, \mathrm{L}(\mathrm{E}, \mathrm{F}))$ peut être identifiée à une application bilinéaire $\mathrm{E} \times \mathrm{E} \to \mathrm{F}$ qui sera notée $d_a^2 f$, ou $d^2 f(a)$. Cette application bilinéaire est la différentielle seconde au point a.. On a donc par définition, pour tout $(h,k) \in \mathrm{E} \times \mathrm{E}$

$$d_a^2 f(h,k) = d_a(df)(h)(k).$$

LEMME 3. — *Si f est deux fois différentiable en a, pour tout $k \in \mathrm{E}$ l'application $\mathrm{U} \to \mathrm{F}$ définie par $x \mapsto d_x f(k)$ est différentiable en a, et a pour différentielle $h \mapsto d_a^2 f(h,k)$.*

Démonstration. Il suffit de remarquer que l'application $x \mapsto d_x f(k)$ est la composée de $df : \mathrm{U} \to \mathrm{L}(\mathrm{E}, \mathrm{F})$ et de l'application linéaire $\mathrm{L}(\mathrm{E}, \mathrm{F}) \to \mathrm{F}$ définie par $u \mapsto u(k)$. □

DÉFINITION 4. — *On dit que $f : \mathrm{U} \to \mathrm{F}$ est de classe C^2 si f est deux fois différentiable en tout point, et si la différentielle seconde $x \mapsto \mathrm{L}(\mathrm{E}, \mathrm{E}; \mathrm{F})$ est continue.*

Supposons que U soit un ouvert d'une somme directe

$$\mathrm{E} = \oplus_{i=1}^{n} \mathrm{E}_i.$$

Alors f est différentiable en a si et seulement si les différentielles partielles $x \mapsto \partial_j f(x) \in \mathrm{L}(\mathrm{E}_j, \mathrm{F})$ sont différentiables en a. La différentielle $\partial_i(\partial_j f)(a)$ définit est une application linéaire $\mathrm{E}_i \to \mathrm{L}(\mathrm{E}_j, \mathrm{F}))$, qui s'identifie à une application bilinéaire $\partial^2_{i,j} f : \mathrm{E}_i \times \mathrm{E}_j \to \mathrm{F}$ qu'on appelle la différentielle partielle d'ordre 2 relative aux sous-espaces E_i et E_j.

LEMME 5. — *Pour $h \in \mathrm{E}_i$ et $k \in \mathrm{E}_j$, on a*

$$\partial^2_{i,j} f(a)(h,k) = d^2_a f(h,k)$$

Démonstration. Soit $h \in \mathrm{E}_i$. On a dans $\mathrm{L}(\mathrm{E}_j, \mathrm{F})$

$$\partial^2_{i,j} f(a)(h,-) = (d_a(\partial_j f))(h)$$

Comme dans le lemme précédent, le second membre appliqué au vecteur $k \in \mathrm{E}_j$ est aussi la valeur en h de la différentielle en a de l'application $\mathrm{U} \to \mathrm{F}$ définie par $x \mapsto \partial_j f(k) = d_x f(k)$, c'est-à-dire en appliquant le lemme ci-dessus $d^2_a f(h,k)$. □

Il résulte de ce lemme que l'on a, si on décompose $h \in \mathrm{E}$ sous la forme $h = \sum_i h_i$ dans la somme directe ci-dessus on a

$$d^2_a f(h,h) = \sum_{i,j} \partial^2_{i,j} f(h_i, h_j) \tag{9.1}$$

Supposons donnée une base $(e_1, \ldots, e_n)$ de E; on désigne par $(x_1, \ldots, x_n)$ les coordonnées dans cette base. On applique ce qui précède à la somme directe $\mathrm{E} = \oplus_{i=1}^{n} \mathbb{R} e_i$. La différentielle partielle $\partial_j f(x)$ est alors donnée par la dérivée partielle

$$\frac{\partial f}{\partial x_j}(x) = \partial_j f(x).e_j.$$

Alors f est de classe C^2 si et seulement les dérivées partielles de ces fonctions existent sur U et sont de classe C^1. On note

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

On a donc compte-tenu du lemme 5

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = d_a^2 f(e_i, e_j),$$

et la formule (9.2) devient, pour $h = \sum_i h_i e_i$, avec $h_i \in \mathbb{R}$

$$d^2 f(a)(h,h) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j.$$

2.3. Formules de Taylor

Il existe plusieurs versions possibles suivant le but à atteindre. La première version est l'existence d'un développement limité en un point. Elle ne nécessite que l'existence d'une différentielle seconde en un point a, mais les renseignements apportés ne sont valable qu'au voisinage de ce point. La seconde formule apporte des renseignements globaux : c'est l'analogue du théorème des accroissements finis ; elle nécessite l'existence de la différentielle seconde $d^2 f$ sur l'ouvert U, et une majoration $\| d^2 f \|$. La troisième, dite formule de Taylor avec reste sous forme d'intégrale, nécessite une hypothèse de continuité sur cette différentielle seconde.

PROPOSITION 6. — *(Développement limité à l'ordre 2) Soit $f : \mathrm{U} \to \mathrm{F}$ une fonction de classe C^1, admettant une différentielle seconde en $a \in \mathrm{U}$. Alors*

$$f(a+h) = f(a) + d_a f(h) + \frac{1}{2} d_a^2 f(h,h) + \epsilon(h) \| h \|^2$$

où $h \mapsto \epsilon(h)$ est une fonction, définie au voisinage de 0, telle que $\lim_{h \to 0} \epsilon(h) = 0$.

Démonstration. Soit $\mathrm{B}(a,r) \subset \mathrm{U}$ une boule centrée en a. Soit h un vecteur tel que $\| h \| < r$. Considérons la fonction, définie sur $[0,1]$,

$$t \mapsto f(a+th) - f(a) - t d_a f(h) - \frac{t^2}{2} d_a^2 f(h,h).$$

Cette fonction est de classe C^1 sur $[0,1]$ est a pour dérivée au point t

$$\varphi'(t) = (d_{a+th} f.h - d_a f(h) - t d_a^2 f(h,h).$$

Soit ϵ un réel > 0. De la définition de la différentielle $d_a^2 f$, il découle que qu'il existe un réel $0 < \eta \leq r$ tel que pour $\| h \| \leq \eta$ on ait $\| (d_{a+h}f - d_a f - d_a^2 f(h, -) \| \leq \epsilon \| h \|$. Par suite, pour $\| h \| \leq \eta$, on aura $\| \varphi'(t) \| \leq \epsilon \| h \|^2$. Du théorème des accroissements finis appliqué à φ, il découle que $\| \varphi(1) - \varphi(0) \| \leq \epsilon \| h \|^2$, ce qui démontre le résultat annoncé.

La seconde formule est à rapprocher du théorème des accroissements finis, dont il est d'ailleurs une conséquence.

PROPOSITION 7. — *(Taylor-Young) On suppose que* $f : \mathrm{U} \to \mathrm{F}$ *a une différentielle seconde sur* U, *et qu'il existe un nombre* M *tel que* $\| d_x^2 f \| \leq \mathrm{M}$ *pour tout* $x \in \mathrm{U}$. *Soient* a *et* $a + h$ *deux points de* U, *tels que le segment* $[a, a + h]$ *soit contenu dans* U. *Alors*

$$\| f(a + h) - f(a) - d_a f.h \| \leq \frac{1}{2}\mathrm{M} \| h \|^2$$

Avant de démontrer cet énoncé, commençons par un rappel qui sera aussi utile pour la troisième formule. Soit $\varphi : [0, 1] \to \mathrm{F}$ une application continue à valeurs dans F. Alors l'intégrale $\int_0^1 \varphi(t)dt$ a un sens, et

$$\| \int_0^1 \varphi(t)dt \| \leq \int_0^1 \| \varphi(t) \| \, dt.$$

Ceci s'obtient facilement en approchant φ, au sens de la topologie de la convergence uniforme sur $[0, 1]$ par une suite φ_n de fonctions en escalier ; dans le cas des fonctions en escaliers, cette inégalité est évidente.

Démonstration. La fonction $h \mapsto g(h) = f(a+h) - f(a) - d_a f.h$ est de classe C^1 sur l'ouvert translaté de U par la translation $x \mapsto x - a$, et a pour différentielle $d_{a+th}f - d_a f$ au point $a+th$. D'après le théorème des accroissements finis appliqué à cette fonction,

$$\| d_{a+th}f - d_a f \| \leq \mathrm{M}t \| h \|$$

Maintenant, on écrit la formule intégrale $\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt$ pour la fonction de classe C^1 définie sur $[0, 1]$ par $\varphi(t) = g(th)$

$$f(a + h) - f(a) - d_a f.h = \int_0^1 (d_{a+th}f - d_a f).hdt$$

ce qui fournit la majoration

$$\| f(a+h) - f(a) - d_a f.h \| \leq \frac{1}{2} \mathrm{M} \| h \|^2$$

et achève la démontration. □

PROPOSITION 8. — *(Reste sous forme d'intégrale) Soit* $f : \mathrm{U} \to \mathrm{F}$ *une application de classe* C^2. *Soient* a *et* $a+h$ *deux points de* U, *tels que le segment* $[a, a+h]$ *soit contenu dans* U. *Alors*

$$f(a+h) - f(a) - d_a f.h = \int_0^1 (1-t) d^2_{a+th} f(h,h) dt$$

Il est clair que si f est de classe C^2, et si on a la majoration $\| d^2_x f \| \leq \mathrm{M}$, alors ceci fournit la majoration de Taylor-Young, en majorant l'intégrale. Ainsi, la plupart du temps, la formule de Taylor avec reste intégrale sera la plus utile et la plus performante, puisqu'elle permet une évaluation précise de l'erreur que l'on fait en remplaçant $f(a+h)$ par l'application affine $f(a) + d_a f.h$. La seconde formule donne seulement une majoration de cette erreur, mais avec une hypothèse un peu plus faible.

Démonstration. On considère une fonction $\varphi : [0,1] \to \mathbb{R}$ de classe C^2. On a

$$\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 (1-t)\varphi''(t) dt.$$

Cette formule s'obtient de façon évidente en calculant l'intégrale par intégration par parties. On applique cette formule à la fonction $t \mapsto f(a+th)$ qui est par hypothèse de classe C^2. Cette fonction a pour dérivée $t \mapsto d_{a+th} f.h$ et pour dérivée seconde en t, d'après le lemme 3, $d^2_{a+th} f(h,h)$. D'où l'énoncé. □

2.4. Le théorème de symétrie de Schwarz

THÉORÈME 9. — *Soit* $f : \mathrm{U} \to \mathrm{F}$ *une application de classe* C^1 *deux fois différentiable en* a. *La différentielle seconde* $d^2_a f$ *est bilinéaire symétrique.*

On munit dans le lemme qui suit le produit $\mathrm{E} \times \mathrm{E}$ de la norme

$$\| (h,k) \| = \sup(\| h \|, \| k \|).$$

Soit $\mathrm{B}(a, 2r)$ une boule ouverte centrée en a contenue dans U.

LEMME 10. — *Soit ϵ un réel > 0. Alors il existe un réel $0 < \eta \leq r$ tel que pour $\| h \|$ et $\| k \| \leq \eta$ on ait*

$$\| f(a+h+k) - f(a+h) - f(a+k) + f(a) - d_a^2 f(h,k) \| \leq \epsilon \| (h,k) \|^2$$

Démonstration. On considère pour $\| h \| < r$ la fonction φ sur la boule centrée en 0 et de rayon r dans E définie par

$$\varphi(k) = f(a+h+k) - f(a+h) - f(a+k) + f(a) - d_a^2 f(h,k)$$

Cette fonction est de classe C^1, et sa différentielle au point k est donnée par $d_{a+h+k}f - d_{a+k}f - d_a^2 f(h,-)$. Soit ϵ un réel > 0. D'après la définition de la différentielle seconde, il existe un réel $0 < \eta \leq r$ tel que pour $\| u \| \leq \eta$ on ait

$$\| d_{a+u}f - d_a f - d_a^2 f(u,-) \| \leq \frac{1}{3}\epsilon \| h \| .$$

En remplaçant succesivement u par $h+k$ et et par u, ceci conduit à la majoration suivante, pour $\| h \|$ et $\| k \| \leq \frac{\eta}{2}$,

$$\| d_{a+h+k}f - d_{a+k}f - d_a^2 f(h,-) \| \leq \epsilon \| (h,k) \| .$$

Le théorème des accroissements finis appliqué à la fonction φ montre que pour $\| h \| \leq \frac{\eta}{2}$ et $\| h \| \leq \frac{\eta}{2}$ on a $\| \varphi(k) - \varphi(0) \| \leq \epsilon \| (h,k) \|^2$ ce qui démontre le lemme. □

Démonstration du théorème 9

On voit que dans l'inégalité ci-dessus, si on échange les rôles de h et k, on obtient que la différence $d_a^2 f(h,k) - d_a^2 f(h,k)$ est négligeable devant $\| (h,k) \|^2$ quand $(h,k) \to 0$. Ceci entraîne que pour tout $\epsilon > 0$, l'application bilinéaire $(h,k) \mapsto d_a^2 f(h,k) - d_a^2 f(h,k)$ est de norme $\leq \epsilon$. Elle est donc de norme nulle, et par suite nulle. Ceci signifie que la différentielle $d_a^2 f$ est symétrique. □

COROLLAIRE 11. — *Supposons que* $E = \oplus_{i=1}^n E_i$. *Si* $f : U \to F$ *est de classe* C^1 *sur un ouvert* U *de* E *et admet une différentielle seconde* $d_a^2 f$ *en un point* $a \in U$, *la différentielle partielle* $\partial_{i,j}^2 f(a)$ *est la transposée de* $\partial_{j,i}^2 f(a)$.

COROLLAIRE 12. — *Supposons donnée une base de* E, *et désignons par* $(x_1, \ldots, x_n)$ *les coordonnées dans cette base. Si* $f : \mathrm{U} \to \mathrm{F}$ *est de classe* C^2 *sur un ouvert* U *de* E *on a pour tout* $a \in \mathrm{U}$

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

2.5. Applications de classe C^r, avec $r \geq 3$.

Désignons par $\mathrm{L}^r(\mathrm{E}, \mathrm{F})$ l'espace vectoriel des applications $r-$multilinéaires

$$\underbrace{\mathrm{E} \times \ldots \times \mathrm{E}}_{r} \to \mathrm{F}.$$

LEMME 13. — *Pour* $1 \leq p < r$, *on a un isomorphisme canonique*

$$\mathrm{L}^p(\mathrm{E}, \mathrm{L}^{r-p}(\mathrm{E}, \mathrm{F})) \simeq \mathrm{L}^r(\mathrm{E}, \mathrm{F})$$

Démonstration. Cet isomorphisme s'obtient en associant à l'application $p-$multilinéaire $f : \underbrace{\mathrm{E} \times \ldots \times \mathrm{E}}_{p} \to \mathrm{L}^{r-p}(\mathrm{E}, \mathrm{F})$ l'application $r-$linéaire g définie par

$$g(x_1, \ldots, x_r) = f(x_1, \ldots, x_p)(x_{p+1}, \ldots, x_n).$$

□

Exercice 2.1
On pose pour $f \in \mathrm{L}^r(\mathrm{E}, \mathrm{F})$

$$\| f \| = \sup_{\|x_1\|=1, \ldots, \|x_r\|=1} \| f(x_1, \ldots, x_r) \|$$

Montrer que ceci définit une norme sur $\mathrm{L}^r(\mathrm{E}, \mathrm{F})$, et que l'isomorphisme ci-dessus est une isométrie.

On considère une application $f : \mathrm{U} \to \mathrm{F}$ de classe C^2 sur un ouvert U de E. Il est alors équivalent de demander que $df : \mathrm{U} \to \mathrm{L}(\mathrm{E}, \mathrm{F})$ est de classe C^2, ou que $d^2 f$ est de classe C^1. On dira alors que f est de classe C^3. En un point $a \in \mathrm{U}$ la différentielle $d_a(d(df))) \in \mathrm{L}(\mathrm{E}, \mathrm{L}(\mathrm{E}, \mathrm{L}(\mathrm{E}, \mathrm{F})))$ définit alors une application

3−linéaire symétrique, qui s'identifie, modulo les isomorphismes ci-dessus à $d_a^2(df)$ et à $d_a(d^2f)$.

On définit plus généralement par récurrence, la notion d'application de classe C^r. Supposons en effet définie la notion d'application de classe C^p pour $p \leq r$ sur un ouvert $U \subset E$; en chaque point $x \in U$, la notion de différentielle $d_x^r f \in L^r(E, F)$ a donc un sens. Soit f une telle application de classe C^r.

Exercice 2.2

Les assertions suivantes sont équivalentes

(i) la différentielle $df : U \to L(E, F)$ est de classe C^r,

(ii) Pour $1 \leq p \leq r$, la différentielle d'ordre p

$$d^{r-p}f : U \to L^{r-p}(E, F))$$

est de classe C^{p+1}.

(iii) L'application $d^r f : U \to L^r(E, F)$ est de classe C^1.

et si l'une de ces conditions est satisfaite, pour tout $x \in U$, on a, via les identifications fournies dans le lemme 13,

$$d_x(d^r f) = d_x^p(d^{r+1-p}f) = d_x^r(df)$$

dans $L^{r+1}(E, F)$. (On pourra mettre de telles identités à l'ordre $\leq r$ dans l'hypothèse de récurrence).

DÉFINITION 14. — *On dit que $f : U \to F$ est de classe C^∞ si f est de classe C^r pour tout r.*

Pour une application de classe C^r, la différentielle d'ordre r en un point $d_x^r f$; est une application n−linéaire $E \times \ldots \times E \to F$. C'est, en fait, d'après le théorème de symétrie de Schwarz, une application r−linéaire symétrique. La composée de deux applications de classe C^r est de classe C^r. Les applications de classe $C^r : U \to \mathbb{R}$ constituent une algèbre, et les applications de classe C^r définie sur U et à valeurs dans un espace vectoriel F constituent un module sur cette algèbre.

La formule de Taylor se généralise aux applications de classe C^r. Comme dans le cas $r = 2$, il y a trois versions utiles, dont deux donnent des renseignements précis sur le reste. Nous les démontrerons à partir de la version qui permet d'écrire le reste sous forme intégrale. On désigne pour $h \in E$, par $h^{(i)}$ l'élément de E^i défini par $\underbrace{(h, \ldots, h)}_{i}$.

PROPOSITION 15. — *(Développement limité à l'ordre r) Soit* $f : \mathrm{U} \to \mathrm{F}$ *une fonction de classe* C^r *sur un ouvert* U *de* E, *et* a *un point de* U. *Alors*

$$f(a+h) = f(a) + \sum_{1 \leq i \leq r} \frac{1}{i!} d_a{}^i f(h^{(i)}) + \epsilon(h) \parallel h \parallel^p$$

où $\epsilon(h)$ *est une fonction définie au voisinage de* 0 *dans* E, *à valeurs dans* F, *qui tend vers* 0 *quand* $h \to 0$.

Pour l'énoncé qui suit, on suppose que E et F sont normés, et on rappelle que l'espace vectoriel $\mathrm{L}^r(\mathrm{E}, \mathrm{F})$ se trouve alors équipé d'une norme naturelle.

PROPOSITION 16. — *(Formule de Taylor-Young) Soit* $f : \mathrm{U} \to \mathrm{F}$ *une fonction de classe* C^r *sur un ouvert* U *de* E, *et* a *un point de* U. *On suppose que pour tout* $x \in \mathrm{U}$ *on a la majoration*

$$\parallel d_x^r f \parallel \leq \mathrm{M}.$$

Alors si $[a, a+h] \subset \mathrm{U}$

$$\parallel f(a+h) - f(a) - \sum_{1 \leq i \leq r-1} \frac{1}{i!} d_a^i f(h^{(i)}) \parallel \leq \frac{\mathrm{M} \parallel h \parallel^r}{r!}$$

PROPOSITION 17. — *(Formule de Taylor avec reste sous forme intégrale) Soit* $f : \mathrm{U} \to \mathrm{F}$ *une fonction de classe* C^r *sur un ouvert* U *de* E, *et* a *un point de* U. *Alors si* $[a, a+h] \subset \mathrm{U}$

$$f(a+h) = \sum_{0 \leq i \leq r-1} \frac{1}{i!} d_a^i f(h^{(i)}) + \frac{1}{(r-1)!} \int_0^1 (1-t)^{r-1} d_{a+th}^r f(h^{(r)}) dt.$$

Démonstration. Soit $\varphi : [0,1] \to \mathrm{F}$ une application de classe C^r à valeurs dans F. On désigne par $\varphi^{(i)}$ la i-ème dérivée de φ. On a alors

$$\varphi(1) = \varphi(0) + \ldots + \frac{\varphi^{(r-1)}(0)}{(r-1)!} + \frac{1}{(r-1)!} \int_0^1 (1-t)^{r-1} \varphi^{(r)}(t) dt.$$

Cette formule se vérifie par récurrence sur r, en intégrant par parties l'intégrale

$$\frac{1}{(r-1)!} \int_0^1 (1-t)^{r-1} \varphi^{(r)}(t) dt.$$

On considère maintenant, si le segment $[a, a+h]$ est contenu dans U, la fonction $\varphi : [0,1] \to \mathrm{F}$ définie par $\varphi(t) = f(a+th)$. Cette fonction est de classe C^r et on a $\varphi^{(i)}(t) = d^i_{a+th} f(h, \ldots, h)$. D'où l'énoncé 17.

L'énoncé 16 en résulte aussitôt, en majorant le reste dans la formule ci-dessus. Enfin, pour obtenir 15, on utilise la continuité de $d^r f$ en a : étant donné un nombre réel $\epsilon > 0$ il existe un nombre réel $\eta > 0$ tel que pour $\| h \| \leq \eta$ on ait $\| d^r_{a+th} f - d^r_a f \| \leq \epsilon$. Il en résulte que si $\| h \| \leq \eta$ on a

$$\| \int_0^1 (1-t)^{r-1} (d^r_{a+th} f - d^r_a f)(h^{(r)}) dt \| \leq \frac{\epsilon \| h \|^r}{r}.$$

Il en résulte l'énoncé 15. □

Naturellement, on peut comme dans la section 2.3 obtenir les mêmes conclusions que dans les énoncés 15 et 16 sous des hypothèses légèrement moins restrictives : ceci permet de les comprendre comme des généralisations de la définition de la différentielle, ou de la formule des accroissements finis.

3. Difféomorphismes

3.1. Le théorème d'inversion locale

Dans ce qui suit, r est un entier ≥ 1, ou $r = \infty$.

DÉFINITION 1. — *Soient* $\mathrm{U} \subset \mathrm{E}$ *et* $\mathrm{V} \subset \mathrm{F}$ *deux ouverts, et* $f : \mathrm{U} \to \mathrm{V}$ *une application. On dit que* f *est un difféoméorphisme de classe* C^r *si* f *est inversible et si* f *et* f^{-1} *sont de classe* C^r.

PROPOSITION 2. — *Soient* $\mathrm{U} \subset \mathrm{E}$ *et* $\mathrm{V} \subset \mathrm{F}$ *deux ouverts des espaces vectoriels* E *et* F, *et* $f : \mathrm{U} \to \mathrm{V}$ *un difféomorphisme de classe* C^r. *Alors* E *et* F *sont de même dimension, et* $d_a f$ *est inversible en tout point de* U.

Démonstration. Soit $b = f(a)$. D'après la règle de composition, on a $d_{fb} f^{-1} \circ d_a f = \mathrm{id}_\mathrm{E}$, et $d_a f \circ d_b f^{-1} = \mathrm{id}_\mathrm{F}$. Donc $d_a f$ est inversible, d'inverse $d_b f^{-1}$. □

Exemples

(1) Soit $f : \mathrm{E} \to \mathrm{F}$ une application de classe C^r d'un ouvert $\mathrm{U} \subset \mathrm{E}$ à valeurs dans F. Alors l'application $\phi : \mathrm{U} \times \mathrm{F} \to \mathrm{U} \times \mathrm{F}$ définie par $(x, y) \mapsto (x, y - f(x))$ est un difféomorphisme.

Exercice 3.1

Ecrire la matrice des différentielles partielles de ϕ.

(2) Considérons l'application $\phi :]0, \infty[\times]-\pi, \pi[\to \mathbb{R}^2 - \{(x, y), y = 0, x \leq 0\}$ définie par

$$(r, \theta) \mapsto (x = r\cos\theta, y = r\sin\theta)$$

C'est un difféomorphisme de classe C^∞; son inverse est donné par

$$\phi^{-1}(x, y) = (\sqrt{x^2 + y^2}, 2\mathrm{Arctg}\frac{y}{\sqrt{x^2 + y^2} + x})$$

Même si la différentielle d'une fonction $f : \mathrm{U} \to \mathrm{F}$ est inversible en tout point, ceci n'entraîne pas que f est injective : par exemple l'application $\mathbb{R}^2 \to \mathbb{R}^2 \setminus \{0\}$ définie par $(r, \theta) \mapsto (x = r\cos\theta, y = r\sin\theta)$ est de différentielle inversible en tout point. Pourtant, cette application n'est pas injective. Le théorème d'inversion locale va nous dire que localement cette application est injective, pourvu que f soit au moins de classe C^1. Plus précisément :

THÉORÈME 3. —*(Théorème d'inversion locale) Soit r un entier ≥ 1. Soit $f : \mathrm{U} \subset \mathrm{E} \to \mathrm{F}$ une application de classe C^r et a un point de U. Si $d_a f$ est inversible, il existe un voisinage ouvert $\mathrm{V} \subset \mathrm{U}$ de a satisfaisant aux conditions suivantes :*

(i) *l'image $\mathrm{W} = f(\mathrm{V})$ est un ouvert de F ;*

(ii) *l'application $f_\mathrm{V} : \mathrm{V} \to \mathrm{W}$ induite par f est un difféomorphisme de classe C^r.*

Avant de démontrer cet énoncé, donnons quelques conséquences immédiates.

COROLLAIRE 4. — *Soit $f : \mathrm{U} \subset \mathrm{E} \to \mathrm{F}$ une application de classe C^r.*

(i) *Si la différentielle $d_a f$ est inversible en tout point de U, f est une application ouverte.*

(ii) *Si en outre f est injective, $f : \mathrm{U} \to \mathrm{V} = f(\mathrm{U})$ est un difféomorphisme de classe C^r.*

Exercice 3.2

Montrer que l'application $]-1,1[\times\mathbb{R} \to]-1,1[\times\mathbb{R}$ définie par $(t,x) \mapsto (t, x - \sin tx)$ est un difféomorphisme de classe C^∞.

Exercice 3.3

On propose dans cet exercice une démonstration du théorème de d'Alembert. Soit $f(z)$ un polynôme complexe de degré n, et Z l'ensemble des zéros du polynôme dérivé. Soit $V = \mathbb{C} \setminus f(Z)$, et U l'ouvert $f^{-1}(V)$.

1) Démontrer que l'application $f : U \to V$ induite par f est ouverte.

2) Démontrer qu'il existe des constantes $C > 0$ et $R > 0$ telles que pour $|z| \geq R$ on ait $|P(z)| \geq C|z|^n$. En déduire que f est propre, et par suite que $f : U \to V$ est propre, donc fermée (*cf.* proposition 3.13).

3) En déduire que f est surjective. En particulier, f a au moins un zéro.

La démonstration du théorème d'inversion locale repose sur le théorème des points fixes que nous commençons par rappeler.

THÉORÈME 5. — *Soit* X *un espace métrique complet,* $\phi : X \to X$ *une application contractante de rapport* $\lambda < 1$, *c'est-à-dire que pour tout* x *et* $y \in X$ *on a* $d(\phi(x), \phi(y)) \leq \lambda d(x,y)$. *Alors* ϕ *a un point fixe et un seul.*

Démonstration.

Soit x_0 un point de X. Considérons la suite x_n de points de X définie par récurrence par $x_{n+1} = \phi(x_n)$. Alors $d(x_n, x_{n+1}) \leq \lambda^n d(x_0, x_1)$ et par conséquent la série $\sum_n d(x_n, x_{n+1})$ est convergente. Alors $d(x_n, x_{n+p}) \leq \sum_{n \leq i \leq n+p-1} d(x_i, x_{i+1})$ et par conséquent la suite x_n est de Cauchy. Puisque X est un espace métrique complet, cette suite converge vers une limite ℓ. Puisque ϕ est continue, l'égalité $\phi(x_n) = x_{n+1}$ donne par passage à la limite $\phi(\ell) = \ell$. Ceci montre l'existene d'un point fixe. L'unicité est évidente. □

Démonstration du théorème d'inversion locale

Nous utiliserons sur l'espace $L(E,F)$ des applications linéaires $E \to F$ la norme définie par

$$\| u \| = \sup_{\|x\| \leq 1} \| u(x) \|$$

En particulier, pour $u \in L(E,E)$, $\| u^k \| \leq \| u \|^k$ de sorte que toute application linéaire $E \to E$ de la forme $\mathrm{id}_E + v$, avec $\| v \| < 1$ est inversible.

a) On peut bien sûr supposer, en faisant éventuellement des translations que $a = 0$, $f(a) = 0$. D'autre part, en remplaçant f par $(d_a f)^{-1} \circ f$, on peut supposer que $\mathrm{E} = \mathrm{F}$ et que $d_a f = \mathrm{id}_{\mathrm{E}}$.

b) On écrit dans ce cas

$$f(x) = x + u(x)$$

avec $u(x)$ négligeable devant x quand $x \to 0$. Bien sûr u est encore de classe C^r, et $d_0 u = 0$. Montrons d'abord qu'on peut trouver un voisinage ouvert de 0 contenu dans U sur lequel f est injective. Par continuité de la différentielle, on voit qu'il existe une boule $\mathrm{B}(0, r)$ telle que pour $x \in \mathrm{B}(0, r)$ on ait

$$\| d_x u \| \leq \frac{1}{2}$$

Il résulte du théorème des accroissements finis que pour x' et $x'' \in \mathrm{B}(0, r)$ on a

$$\| u(x') - u(x'') \| \leq \frac{1}{2} \| x' - x'' \| \tag{1.1}$$

Il en résulte que

$$\| f(x') - f(x'') \| \geq \frac{1}{2} \| x' - x'' \| \tag{1.2}$$

ce qui prouve l'injectivité de f sur $\mathrm{B}(0, r)$.

c) Montrons maintenant que $f(\mathrm{B}(0, r))$ est un voisinage de 0. Ainsi, il s'agit de montrer que si y est suffisamment voisin de 0, que l'équation

$$x + u(x) = y \tag{1.3}$$

a une solution $x \in \mathrm{B}(0, r)$. Nous écrivons cette équation sous la forme

$$x = y - u(x)$$

Remarquons que si $\| x \| \leq \frac{r}{2}$ et $\| y \| \leq \frac{r}{4}$, on a d'après (1.1) $\| u(x) \| \leq \frac{1}{2} \| x \|$ et par conséquent $\| \phi(x) \| \leq \frac{r}{2}$. Considérons l'application $\phi : \overline{\mathrm{B}}(0, \frac{r}{2}) \to \overline{\mathrm{B}}(0, \frac{r}{2})$ définie par $\phi(x) = y - u(x)$. L'application ϕ est contractante de rapport $\frac{1}{2}$ d'après (1.1), et le fermé $\overline{\mathrm{B}}(0, \frac{r}{2})$ est un espace métrique complet. Il en résulte que l'application ϕ a un point fixe, et donc que l'équation (1.3) a une solution $x \in \mathrm{B}(0, r)$ si $\| y \| \leq \frac{r}{4}$.

d) Soit V l'ouvert défini par $\| x \| < r$, et $\| f(x) \| < \frac{r}{4}$. L'application induite par f

$$f_{\mathrm{V}} : \mathrm{V} \to \mathrm{B}(0, \frac{r}{4})$$

est alors bijective. On va montrer que $g = (f_{\mathrm{V}})^{-1}$ est encore de classe C^r, ce qui prouvera le théorème. En vertu de l'inégalité (1.2), elle est Lipchitzienne de rapport 2, et par suite continue. Donc, f_{V} est un homéomorphisme. Montrons maintenant que g est différentiable.

En tout point $x \in \mathrm{V}$, la différentielle $d_x f$ est donnée par $d_x f = \mathrm{id}_{\mathrm{E}} + d_x u$. La majoration $\| d_x u \| \leq \frac{1}{2}$ montre que $d_x f$ est inversible pour tout $x \in \mathrm{V}$. On est ramené à prouver le lemme suivant :

LEMME 6. — *Soit $f : \mathrm{U} \to \mathrm{V}$ un homéomorphisme d'un ouvert* U *d'un espace vectoriel* E *dans un ouvert* V *d'un espace vectoriel* F. *On suppose que f est de classe C^r et que pour tout $x \in \mathrm{U}$, la différentielle $d_x f$ est inversible. Alors f est un difféomorphisme de classe C^r*

Démonstration. a) Soit $a \in \mathrm{U}$. Montrons que g est différentiable au point $b = f(a)$. Comme dans la démontration ci-dessus, on peut supposer que $a = b = 0$, et que $d_0 f = \mathrm{id}_{\mathrm{E}}$. On écrit comme ci-dessus, $f(x) = x + u(x)$ ce qui s'écrit encore pour $y \in \mathrm{V}$

$$g(y) = y - u(g(y))$$

avec $u(x)$ négligeable devant $\| x \|$ quand $x \to 0$. Comme ci-dessus, g est lipchitzienne de rapport 2 au voisinage de 0 et donc $\| g(y) \| \leq 2 \| y \|$ au voisinage de 0. Maintenant, $\epsilon > 0$ étant donné, il existe un voisinage ouvert $\mathrm{V} \subset \mathrm{U}$ de 0 sur lequel on a $\| u(x) \| \leq \epsilon \| x \|$. Il existe alors un voisinage ouvert W de 0 tel que pour $y \in \mathrm{W}$ on ait

$$\| u(g(y)) \| \leq \epsilon \| g(y) \| \leq 2\epsilon \| y \|$$

ce qui prouve que $u(g(y))$ est négligeable devant $\| y \|$ quand $y \to 0$. Ainsi, l'application g est différentiable en 0.

b) Ceci montre que g est différentiable en tout point, et sa différentielle est donnée par

$$d_y g = (d_x f)^{-1}$$

Ainsi, dg s'obtient en composant df avec l'application $u \mapsto u^{-1}$. Compte-tenu de l'exercice 1.1 et du fait que df est de classe C^{r-1}, on obtient que dg est de classe C^{r-1}, et par suite que g est de classe C^r. Ceci achève la démonstration. □

Exercice 3.4

On considère la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = x + 2x^2 \sin \frac{1}{x}.$$

Montrer que f est dérivable, et que $f'(0) = 1$. Vérifier qu'il n'existe pas de voisinage de 0 sur lequel la fonction f est injective. Expliquer pourquoi ceci ne contredit pas le théorème s'inversion locale.

3.2. Théorème des fonctions implicites

Grossièrement parlant, le problème des fonctions implicites consiste étant donnée une équation $f(x, y) = c$ à pouvoir exprimer la variable y en fonctions de x. Les données précises sont les suivantes : on se donne E et F deux espaces vectoriels, et $f : U \to F$ une application de classe C^r sur un ouvert U de $E \times F$. On considère dans U l'équation $f(x, y) = Cte$. A condition de faire une hypothèse sur la différentielle partielle $\partial_2 f$ dans la direction F, on va pouvoir exprimer y en fonction de x au voisinage d'un point de U.

THÉORÈME 7. — *Soit (a, b) un point de* U, *et $f(a, b) = c$. On suppose que la différentielle partielle $\partial_2 f(a, b)$ dans la direction* F *est inversible. Alors il existe un voisinage ouvert* U′ *de (a, b) tel que la trace dans* U′ *du fermé $f(x, y) = c$ soit le graphe d'une application $u : V \to F$ de classe C^r sur un voisinage ouvert* V *de a à valeurs dans* F.

Autrement dit, on a l'équivalence

$$(x, y) \in U', \text{ et } f(x, y) = c \Longleftrightarrow x \in V \text{ et } y = u(x).$$

Démonstration. Considérons l'application $\phi : U \to E \times F$ définie par

$$(x, y) \mapsto (x, f(x, y))$$

L'application ϕ est de classe C^r, et sa différentielle $d_{(a,b)}\phi$ au point (a,b), est donnée par la matrice

$$\begin{pmatrix} \mathrm{id}_E & 0 \\ \partial_1 f(a,b) & \partial_2 f(a,b) \end{pmatrix}$$

Puisque $\partial_2 f(a,b)$ est inversible, cette application linéaire est aussi inversible. Donc, il existe un voisinage ouvert U' de (a,b) tel que ϕ induise un difféomorphisme $\phi_{U'}$ de U' sur un voisinage ouvert V' de (a,c). Soit $\Psi : V' \to U' \subset E \times F$ l'inverse de $\phi_{U'}$; cette application inverse est de la forme $(X,Y) \mapsto (X, \psi(X,Y))$, avec ψ de classe C^r. La trace G' du fermé d'équation $f(x,y) = c$ sur U' est l'image par le difféomorphisme Ψ du fermé de V' défini par les couples (X,Y) tels que $Y = c$. Soit V l'ensemble des $x \in E$ tels que $(x,c) \in V'$. C'est un ouvert de E, et l'application $u : V \to F$ définie par $u(x) = \psi(x,c)$ a pour graphe G'. L'application u est de classe C^r, ce qui termine donc la démonstration. □

Exemple

Considérons l'équation $x^3 - x + y^2 = 0$. Au voisinage du point (1,0), cette équation peut se résoudre sous la forme $x = u(y)$, avec u de classe C^∞.

4. Théorème du rang

4.1. Le rang

Soient E et F deux espaces vectoriels, et $f : U \to F$ une application de classe C^r d'un ouvert U de E à valeurs dans F. On appelle rang de f en a, et on note $\mathrm{rg}_a f$ le rang de la différentielle $d_a f$.

LEMME 1. — *La fonction* $U \to \mathbb{N}$ *définie par*

$$x \mapsto \mathrm{rg}_x f$$

est semi-continue inférieurement.

Démonstration. Soit $a \in U$ et ρ le rang de $d_a f$. Désignons par N le noyau de $d_a f$ et par I son image. Considérons des supplémentaires S et T de N et I, de sorte que l'on a des sommes directes $E = N \oplus S$ et $F = I \oplus T$. Alors f s'écrit dans la dernière somme directe $f = (f_1, f_2)$ et la différentielle partielle $\partial_2 f_1(x)$

est inversible au point $x = a$, et de rang ρ. Par continuité, elle reste inversible au voisinage de a, de sorte qu'au voisinage de a le rang de $d_x f$ est $\geq \rho$. □

Exercice 4.1

Retrouver le résultat en choisissant des bases de E et F et en considérant les mineurs de la matrice jacobienne de f au point a.

DÉFINITION 2. — *Soient* U *et* U′ *deux ouverts de* E, *et* V *et* V′ *deux ouverts de* F; *enfin, soient* $f : \mathrm{U} \to \mathrm{V} \subset \mathrm{F}$ *et* $f' : \mathrm{U}' \subset \mathrm{V}'$ *deux applications de classe* C^r.

On dit que f *et* f' *sont* C^r*−conjuguées s'il existe des difféomorphismes* $\phi : \mathrm{U} \to \mathrm{U}'$ *et* $\psi : \mathrm{V} \to \mathrm{V}'$ *tels que*

$$f' = \psi f \phi^{-1}.$$

LEMME 3. — *Si* f *est conjuguée à une application* $g : \mathrm{U}' \to \mathrm{V}'$ *induite par une application linéaire, le rang de* f *est constant.*

Démonstration. Soient comme ci-dessus des difféomorphismes ϕ et ψ tels que $g = \psi f \phi^{-1}$. Soient $a \in \mathrm{E}$, $b = f(a)$ et $a' = \phi(a)$ et $b' = \psi(b)$ les points correspondants de U′ et V′. La différentielle de g en $a' \in \mathrm{U}'$ est donnée par

$$d_{a'} g = d_b \psi \circ d_a f \circ d_a \phi^{-1}$$

Par suite, $d_{a'} g$ et $d_a f$ on même rang. Puisque g est induite par une application linéaire, sa différentielle est constante. D'où le lemme. □

Le but de la section est de montrer la réciproque ; cependant, le résultat ne sera vrai que localement, au voisinage d'un point : autrement dit, il faudra éventuellement diminuer la taille des ouverts U et V. On commence par deux cas particuliers, qui en fait seront suffisants pour la plupart des exemples. Le théorème du rang est une combinaison de ces deux cas particuliers.

4.2. Théorème de submersion

Une application $f : \mathrm{U} \to \mathrm{F}$ de classe C^r de différentielle $d_a f$ surjective en un point a reste de rang maximum au voisinage de a. On dit que c'est une submersion au voisinage de a.

THÉORÈME 4. — *Soit* $f : \mathrm{U} \to \mathrm{F}$ *une application de classe* C^r *sur un ouvert* U *de* E *et* a *un point de* U.

On suppose que la différentielle $d_a f$ *est surjective.*

Alors il existe un voisinage ouvert $\mathrm{V} \subset \mathrm{U}$ *de* a *et un difféomorphisme* $\phi : \mathrm{V} \to \mathrm{W} \subset \mathrm{E}$ *de classe* C^r *tel que l'application* $f \circ \phi^{-1}$ *soit induite par une application linéaire* $\mathrm{E} \to \mathrm{F}$.

Démonstration. On pose $b = f(a)$. Soit $\mathrm{N} = \ker d_a f$, et S un supplémentaire de N, de sorte que dans la somme directe $\mathrm{E} = \mathrm{N} \oplus \mathrm{S}$ la différentielle partielle $\partial_2 f(a) = d_a f|_{\mathrm{S}}$ est, en raison de l'hypothèse, un isomorphisme. On considère l'application $\phi : \mathrm{E} \to \mathrm{N} \oplus \mathrm{F}$ de classe C^r définie par

$$\phi(x, y) = (x, f(x, y))$$

La matrice jacobienne de f au point a dans les sommes directes ci-dessus est donnée par

$$d_a f = \begin{pmatrix} \mathrm{id}_{\mathrm{N}} & 0 \\ \partial_1 f(a) & \partial_2 f(a) \end{pmatrix}$$

Cette application linéaire est inversible. D'après le théorème d'inversion locale, il existe un voisinage ouvert $\mathrm{V} \subset \mathrm{U}$ de a, un voisinage ouvert W de $(0, b)$ tels que ϕ induise un difféomorphisme, noté encore $\phi : \mathrm{V} \to \mathrm{W}$. Le diagramme

$$\begin{array}{ccc} \mathrm{V} & \xrightarrow{\phi} & \mathrm{W} \\ & f\searrow & \downarrow p_2 \\ & & \mathrm{F} \end{array}$$

où p_2 est la seconde projection, est commutatif, de sorte que $f \circ \phi^{-1} : \mathrm{W} \to \mathrm{F}$ est induit par la seconde projection, qui est évidemment linéaire. Puisque $\mathrm{E} \simeq \mathrm{N} \oplus \mathrm{F}$, ceci conduit au résultat. □

DÉFINITION 5. — *Une application* $f : \mathrm{U} \to \mathrm{F}$ *de classe* C^r *sur un ouvert* U *de* E *est appelée une submersion si la différentielle* $d_a f$ *est surjective en tout point de* U.

4.3. Théorème d'immersion

Une application $f : \mathrm{U} \to \mathrm{F}$ de classe C^r sur un ouvert U d'un espace vectoriel de dimension n dont la différentielle est injective en un point a reste de rang n au voisinage de a. On dit que c'est une immersion au voisinage de a. Cette notion est justifiée par le théorème suivant :

THÉORÈME 6. — *Soit $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^r sur un ouvert U de E. Soit $a \in \mathrm{E}$ et $b = f(a)$. On suppose que la différentielle $d_a f$ est injective.*

Alors il existe un voisinage ouvert $\mathrm{R} \subset \mathrm{U}$ de a, un voisinage ouvert V de b tels que $f(\mathrm{R}) \subset \mathrm{V}$ et un difféomorphisme $\phi : \mathrm{V} \to \mathrm{W} \subset \mathrm{F}$ de classe C^r tels que l'application $\phi \circ f|_{\mathrm{R}}$ soit induite par une application linéaire $\mathrm{E} \to \mathrm{F}$.

Démonstration. Soit I l'image de $d_a f$, et S un supplémentaire de I. Considérons l'application $\Phi : \mathrm{U} \times \mathrm{S} \to \mathrm{F}$ définie sur l'ouvert $\mathrm{U} \times \mathrm{S}$ de $\mathrm{E} \oplus \mathrm{S} = \mathrm{E} \times \mathrm{S}$ par

$$(x, y) \mapsto f(x) + y$$

La différentielle de ϕ en $(a, 0)$ est la matrice

$$d_a\Phi = \begin{pmatrix} d_a f & \mathrm{id}_{\mathrm{S}} \end{pmatrix}$$

Puisque la différentielle $d_a f$ est injective, cette application linéaire est inversible. Donc il existe un voisinage ouvert $\mathrm{W} \subset \mathrm{U} \times \mathrm{S}$ de $(a, 0)$ et un voisinage ouvert V de b tel que Φ induise un difféomorphisme $\mathrm{W} \to \mathrm{V}$. Soit ι l'inclusion $\mathrm{E} \hookrightarrow \mathrm{E} \times \mathrm{S}$, et $\mathrm{R} = \iota^{-1}(\mathrm{W})$. On a alors $f(\mathrm{R}) \subset \mathrm{W}$, et le diagramme

$$\begin{tikzcd} \mathrm{R} \arrow[d, "\iota"'] \arrow[dr, "f"] & \\ \mathrm{W} \arrow[r, "\Phi"] & \mathrm{V} \end{tikzcd}$$

est commutatif. Considérons le difféomorphisme $\phi = \Phi^{-1} : \mathrm{V} \to \mathrm{W}$. Alors $\phi \circ f|_{\mathrm{R}}$ est induit par l'application linéaire ι. □

DÉFINITION 7. — *Une application $f : \mathrm{U} \to \mathrm{F}$ de classe C^r sur un ouvert U de E est appelée une immersion si la différentielle $d_a f$ est injective en tout point de U.*

4.4. Subimmersion

THÉORÈME 8. — *(Théorème du rang) Soit $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^r ; soit a un point de U, et $b = f(a)$. On suppose que le rang de f est constant au voisinage de a. Alors il existe des difféomorphismes ϕ et ψ définies au voisinage de a et b respectivement, de classe C^r, tels que l'application composée*

$$g = \psi \circ f \circ \phi^{-1},$$

soit, au voisinage de a, induite par une application linéaire.

Bien entendu, l'application composée ci-dessus n'a de sens qu'au voisinage de a. Il s'agit uniquement d'un théorème local. Ce théorème va s'obtenir en combinant le théorème de submersion et le théorème d'immersion.

Démonstration. Soit I l'image de $d_a f$, et S un supplémentaire de I. Dans la somme directe $\mathrm{F} = \mathrm{I} \oplus \mathrm{S}$ on écrit $f = (f_1, f_2)$ de sorte que $d_a f_1$ est surjective. D'après le théorème de submersion, il existe un difféomorphisme ϕ de classe C^r défini sur un voisinage ouvert de a tel que $f_1 \circ \phi^{-1}$ soit, au voisinage de a, induite par une application linéaire (obligatoirement surjective).

On peut donc supposer que $\mathrm{E} = \mathrm{N} \oplus \mathrm{I}$ et que f_1 est donné au voisinage de a par la projection $(x, y) \mapsto y$; l'application f s'écrit, sur un voisinage ouvert de $a = (a', a'')$ qu'on peut supposer de la forme $\mathrm{U}' \times \mathrm{U}'' \subset \mathrm{U}$

$$(x, y) \mapsto (y, f_2(x, y))$$

On peut en outre supposer que U' est connexe, et que le rang de f est constant sur $\mathrm{U}' \times \mathrm{U}''$. La matrice jacobienne de f en (x, y) s'écrit dans les sommes directes ci-dessus

$$d_a f = \begin{pmatrix} 0 & \mathrm{id}_{\mathrm{I}} \\ \partial_1 f(x, y) & \partial_2 f(x, y) \end{pmatrix}$$

L'hypothèse sur le rang montre que $\partial_1 f(x, y) = 0$ ce qui implique que l'application f_1 ne dépend pas de x. Posons pour $y \in \mathrm{U}'' \subset \mathrm{I}$, $g(y) = f(a', y)$. Alors g est de classe C^r. Sur l'ouvert $\mathrm{U}' \times \mathrm{U}''$, l'application f est la composée de la projection $(x, y) \mapsto y$ et de l'application $\mathrm{G} : \mathrm{U}'' \to \mathrm{F}$ définie par

$$y \mapsto (y, g(y))$$

D'après le théorème d'immersion, il existe un difféomorphisme ψ de classe C^r défini au voisinage de $b = G(a'')$ dans F tel que $\psi \circ G$ soit, au voisinage de a'', induit par une application linéaire v. Il en résulte qu'au voisinage $\psi \circ f$ est au voisinage de a, induit par l'application linéaire $(x, y) \mapsto v(y)$. □

5. Sous-variétés de $\mathbb{R}^n$

Soit E un espace vectoriel de dimension n. Le concept de sous-variété de E généralise celui de courbe ou de surface. Il s'agit de parties de E qu'on peut décrire localement par de bonnes équations.

5.1. La notion de sous-variété

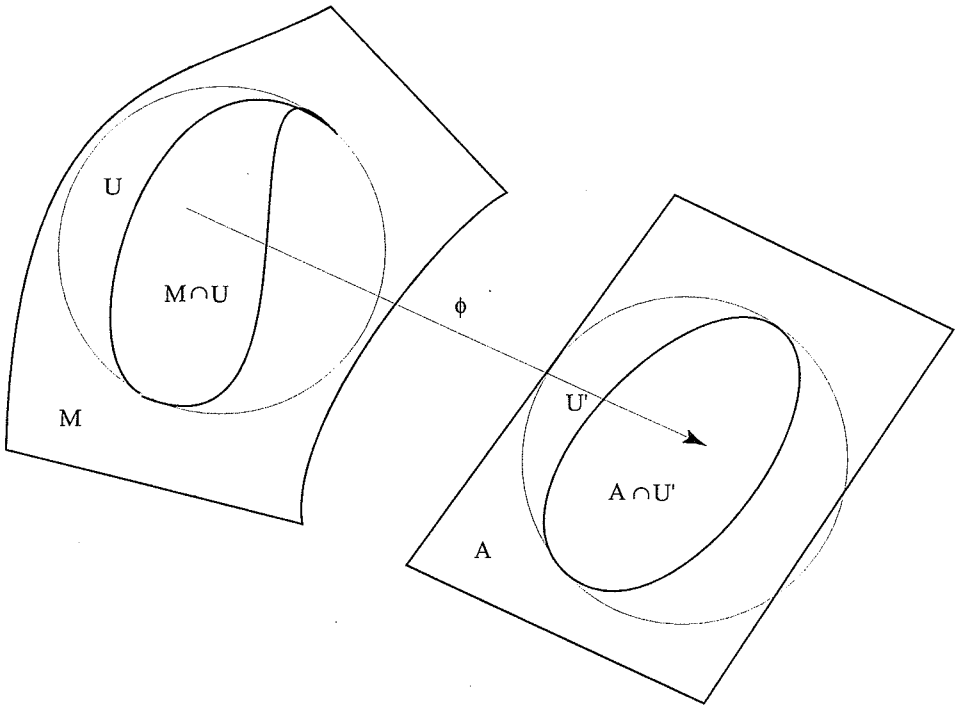
DÉFINITION 1. — *Une partie* $A \subset E$ *s'appelle sous-espace affine de* E *si elle est l'image d'un sous-espace vectoriel* F *de* E *par une translation.*

Soit A un sous-espace affine de E. Si a est un point de A, le sous-espace vectoriel F est alors l'image réciproque de A par la translation $x \to a + x$. On l'appelle espace vectoriel tangent à A. La dimension de cet espace tangent est par définition la dimension de A.

Exemple

Soit G un autre espace vectoriel de dimension p. Soit $f : E \to G$ une application linéaire, et b un point de G; alors l'ensemble $A = f^{-1}(b)$ est, s'il n'est pas vide, un sous-espace affine de E d'espace tangent $\ker f$.

DÉFINITION 2. — *Soit* M *une partie de* E. *On dit que* M *est une sous-variété de classe* C^r *de* E *si pour tout point* $a \in M$ *il existe un voisinage ouvert* U *de* a *dans* E *et un difféomorphisme* $\phi : U \to U' \subset E$ *tel que* $\phi(M \cap U)$ *soit la trace sur* U' *d'un sous-espace affine* A *de* E.



On peut bien sûr, quitte à composer ϕ par une translation, supposer que A contient 0, ce qui signifie alors que A un sous-espace vectoriel de E. On peut alors trouver une base $(e_i)_{i=1\ldots,m}$ de A qu'on complète pour obtenir une base $(e_i)_{i=1,\ldots,n}$ de E. On désigne par $(\phi_i)_{i=1,\ldots,n}$ les composantes dans cette base. Alors, dans l'ouvert U de E, l'intersection $\mathrm{M} \cap \mathrm{U}$ est définie par les équations

$$\phi_{m+1} = \ldots = \phi_n = 0.$$

Remarquons que la différentielle de l'application $(\phi_{m+1}, \ldots, \phi_n)$ est surjective en tout point de a. Réciproquement, on verra au paragraphe 3.3 que si M est une partie de E qui peut être définie localement par de telles équations, M est une sous-variété de E.

Il résulte de la définition une sous-variété de E est localement fermée dans E, c'est-à-dire intersection d'un ouvert et d'un fermé. Les ouverts de E sont des cas particuliers de sous-variétés de E.

Exemple

Considérons le graphe $\mathrm{G} \subset \mathrm{E} \times \mathrm{F}$ d'une application $f : \mathrm{U} \to \mathrm{F}$ de classe C^r d'un ouvert U de E à valeurs dans un espace vectoriel F. Alors le difféomorphisme

$\phi : \mathrm{U} \times \mathrm{F} \to \mathrm{U} \times \mathrm{F}$ défini par

$$\phi(x, y) = (x, y - u(x))$$

transforme le graphe de u en $\mathrm{U} \times \{0\}$. C'est la trace sur $\mathrm{U} \times \mathrm{F}$ d'un sous-espace vectoriel de $\mathrm{E} \times \mathrm{F}$. Donc le graphe G est une sous-variété de $\mathrm{E} \times \mathrm{F}$.

5.2. Espaces tangents

DÉFINITION 3. — *Soit* M *une partie de* E, *et* a *un point de* M. *Un vecteur* $t \in \mathrm{E}$ *est dit tangent à* M *en* a *s'il existe un chemin* $\alpha :]-\epsilon, \epsilon[\to \mathrm{E}$ *défini sur un petit intervalle ouvert contenant* 0, *dérivable en* 0, *tel que*

(i) *l'image de* α *est contenu dans* M

(ii) *le vecteur vitesse* $\alpha'(0)$ *est égal à* t.

On désigne par $\mathrm{T}_a\mathrm{M}$ l'ensemble des vecteurs tangents à M en a.

Exercice 5.1

Soit $\mathrm{M} \subset \mathbb{R}^2$ le fermé défini par les points (x, y) définis par l'équation $xy = 0$. Déterminer l'ensemble $\mathrm{T}_0\mathrm{M}$ et vérifier que ce n'est pas un sous-espace vectoriel de $\mathbb{R}^2$.

Exercice 5.2

Soit $\phi = \mathrm{U} \to \mathrm{V}$ un difféomorphisme de classe C^1 d'un ouvert U de E sur un ouvert V de E. Soit M une sous-variété de E. Montrer que $\phi(\mathrm{M} \cap \mathrm{U})$ est une sous-variété de E et que l'on a

$$d_a\phi(\mathrm{T}_a\mathrm{M}) = \mathrm{T}_{\phi(a)}(\phi(\mathrm{M})).$$

PROPOSITION 4. — *Soit* M *une sous-variété de classe* C^1 *de* E. *Alors l'ensemble des vecteurs tangents* $\mathrm{T}_a\mathrm{M}$ *est un sous-espace vectoriel de* E.

Démonstration. Au voisinage U de a, il existe un difféomorphisme $\phi : \mathrm{U} \to \mathrm{V}$ de classe C^1 tel que $\phi(\mathrm{U} \cap \mathrm{M})$ soit la trace sur V d'un sous-espace vectoriel $\mathrm{F} \subset \mathrm{E}$. En vertu de l'exercice 3.2 qui précède, on est ramené à vérifier qu'en un point quelconque de F on a $\mathrm{T}_b\mathrm{F} = \mathrm{F}$, ce qui est évident. □

DÉFINITION 5. — *Soit* M *une sous-variété de* E. *On appelle dimension de* M *en* $a \in \mathrm{M}$ *la dimension de l'espace tangent* $\mathrm{T}_a\mathrm{M}$.

Ainsi, la dimension d'une sous-variété est une fonction localement constante. Si elle est constante, on dit que la sous-variété est équidimensionnelle ; cette constante est la dimension de M. Chacune des composantes connexes d'une sous-variété est équidimensionnelle. Dans la pratique, on se limitera souvent aux sous-variétés équidimensionnelles de dimension m : ceci revient à imposer que dans la définition 2, le sous-espace affine A est de dimension m. Une sous-variété de dimension 1 (resp. 2) s'appelle une courbe (resp. une surface) de E ; une sous-variété de dimension $n-1$ s'appelle une hypersurface de E. Une sous-variété M de E de dimension 0 de E est un ensemble discret, c'est-à-dire que la topologie induite sur M est la topologie discrète.

Exercice 5.3

Soient M et N deux sous-variétés d'espaces vectoriels E et F respectivement. Soit U un ouvert contenant M et $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^r telle que $f(\mathrm{M}) \subset \mathrm{N}$. Démontrer que

$$d_a f(\mathrm{T}_a\mathrm{M}) \subset \mathrm{T}_{f(a)}\mathrm{N}.$$

5.3. Sous-variétés définies par des équations

Soient E et F deux espaces vectoriels, et $\Omega \subset \mathrm{E}$ un ouvert de E.

THÉORÈME 6. — *Soit $f : \Omega \to \mathrm{F}$ une application de classe C^r et de rang constant.*

Soit $b \in \mathrm{F}$ un point de F. L'ensemble $\mathrm{M} = \{x \in \Omega, f(x) = b\}$ est une sous-variété de Ω de classe C^r, dont l'espace tangent en un point $a \in \mathrm{M}$ est

$$\mathrm{T}_a\mathrm{M} = \ker d_a f$$

Démonstration. Soit $a \in \mathrm{M}$. D'après le théorème du rang, on peut trouver des voisinages $\mathrm{U} \subset \Omega$ et V de a et b respectivement, et des difféomorphismes $\phi : \mathrm{U} \to \mathrm{U}'$ et $\psi : \mathrm{V} \to \mathrm{V}'$ tels que $f(\mathrm{U}) \subset \mathrm{V}$ et tels que $g = \psi \circ f \circ \phi^{-1}$ soit induite par une application linéaire. Soit $a' = \phi(a)$, et $b' = \psi(b)$. Considérons le diagramme commutatif

$$\begin{array}{ccc} \mathrm{U} & \xrightarrow{f} & \mathrm{V} \\ \phi \downarrow & & \downarrow \psi \\ \mathrm{U}' & \xrightarrow{g} & \mathrm{V}' \end{array}$$

On voit que $x \in \mathrm{M} \cap \mathrm{U}$ si et seulement si $x' = \phi(x)$ satisfait à l'équation $g(x') = b'$. Puisque g est la restriction à U' d'une application linéaire, $\phi(\mathrm{M} \cap \mathrm{U})$ est la trace sur U' d'un sous-espace affine : ainsi, M est une sous-variété de classe C^r de E. L'espace tangent à $\phi(\mathrm{M} \cap \mathrm{U})$ en a' est le noyau de dg. En vertu du diagramme commutatif

$$\begin{array}{ccc} \mathrm{E} & \xrightarrow{d_a f} & \mathrm{F} \\ {\scriptstyle d_a\phi}\downarrow & & \downarrow{\scriptstyle d_b\psi} \\ \mathrm{E} & \xrightarrow{dg} & \mathrm{F} \end{array}$$

on obtient que $\mathrm{T}_a\mathrm{M} = \ker d_a f$.

DÉFINITION 7. — *Soit $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^r sur un ouvert* U *de* E. *Un point $a \in \mathrm{U}$ est appelé un point critique de f si*

$$\mathrm{rg}_a f < \dim \mathrm{F}$$

La valeur de f en un point critique s'appelle valeur critique.

COROLLAIRE 8. — *Soit $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^r. Si b est un point de* F *qui n'est pas valeur critique, le fermé de* U *défini par* $\mathrm{M} = \{x \in \mathrm{U}, f(x) = b\}$ *est une sous-variété de* E, *dont l'espace tangent en un point $a \in \mathrm{M}$ est* $\ker d_a f$.

Exercice 5.4

Soit $f : \mathbb{R}^n \to \mathbb{R}$ l'application définie par

$$(x_1, \ldots, x_n) \mapsto x_1^d + \ldots + x_n^d$$

Vérifier que la seule valeur critique de f est 0.

Il en résulte que l'équation

$$x_1^d + \ldots + x_n^d = 1$$

définit une hypersurface de $\mathbb{R}^n$.

Exercice 5.5

Soit E un espace vectoriel euclidien. Soit End(E) l'espace vectoriel des endomorphismes de E. Pour $u \in \mathrm{End(E)}$, on désigne par u^* l'adjointe de u. On considère l'application $f : \mathrm{End(E)} \to \mathrm{End(E)}$

$$f : u \mapsto uu^*$$

1) Vérifier que f est de rang constant sur GL(E).

2) En déduire que le groupe O(E) des applications orthogonales est une sous-variété de End(E). Préciser la dimension et l'espace tangent en un point.

Exercice 5.6

1) Trouver les valeurs critiques de l'application $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par

$$f(x, y, z) = \begin{pmatrix} x^3 + y^3 + z^3 \\ x^2 + y^2 + z^2 \end{pmatrix}$$

Vérifier que l'ensemble des valeurs critiques est de mesure nulle.

2) En déduire que les équations

$$\begin{cases} x^3 + y^3 + z^3 = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

définissent une courbe compacte de $\mathbb{R}^3$.

Exercice 5.7

On considère deux espaces vectoriels E et F de dimensions respectives n et m. Soit $\mathrm{M} \subset \mathrm{L(E, F)}$ l'ensemble des applications linéaires de rang ρ.

1) Soit $u \in \mathrm{M}$. On désigne par N le noyau de u et par I son image. On choisit des supplémentaires S et T de N et I respectivement, de sorte que l'on dispose de sommes directes $\mathrm{E} = \mathrm{N} \oplus \mathrm{S}$ et $\mathrm{F} = \mathrm{I} \oplus \mathrm{T}$. On écrit dans ces sommes directes, pour $v \in \mathrm{End(E)}$

$$v = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Démontrer que dans l'ouvert U défini par les v tels que b soit inversible, $\mathrm{M} \cap \mathrm{U}$ est défini par l'équation à valeurs dans $\mathrm{L(N, T)}$

$$c - db^{-1}a = 0$$

2) En déduire que M est une sous-variété localement fermée de dimension $(n-\rho)(m-\rho)$ de l'espace vectoriel End(E), et que son espace tangent au point u est l'espace vectoriel noyau de l'application linéaire canonique

$$\mathrm{End}(\mathrm{E}) \to \mathrm{L}(\ker u, \mathrm{coker}\ u)$$

définie par composition avec l'inclusion $\ker u \hookrightarrow \mathrm{E}$ et la projection $\mathrm{F} \to \mathrm{coker}\ u$.

5.4. Sous-variétés définies par un paramétrage

On considère une application $f : \Omega \to \mathrm{F}$ de classe C^r définie sur un ouvert Ω d'un espace vectoriel E. Alors qu'il n'y a pas de précaution particulière, en dehors de conditions portant sur le rang, pour vérifier que les fibres de f sont des sous-variétés, il faut être beaucoup plus prudent pour affirmer que l'image de f une sous-variété. Par exemple, en général, même si f est une immersion injective, l'image de f n'est pas toujours une sous-variété de F.

THÉORÈME 9. — *Soit Ω un ouvert de* E, *et $f : \Omega \to \mathrm{F}$ une application de classe C^r de rang constant. On suppose que l'application induite $f : \Omega \to \mathrm{Im}\, f$ est ouverte. Alors*

(i) *l'image* $\mathrm{Im}\, f$ *est une sous-variété de* F *de classe* C^r;

(ii) *pour $a \in \Omega$, l'espace tangent à* $\mathrm{Im}\, f$ *au point $b = f(a)$ est donné par*

$$\mathrm{T}_b \mathrm{Im}\, f = \mathrm{Im}\, d_a f$$

L'hypothèse demandant que l'application $f : \Omega \to \mathrm{Im}\, f$ est ouverte est très restrictive : elle impose que l'image de $d_a f$ est indépendante de a quand a parcourt la fibre $f^{-1}(b)$.

DÉFINITION 10. — *Soit* U *un ouvert de* E, *et $f : \mathrm{U} \to \mathrm{F}$ une application de classe C^r. On dit que f est un plongement de classe C^r si c'est une immersion et si $f : \mathrm{U} \to f(\mathrm{U})$ est un homéomorphisme.*

COROLLAIRE 11. — *Soit $f : \mathrm{U} \to \mathrm{F}$ un plongement de classe C^r. L'image* $\mathrm{Im}\, f$ *est une sous-variété de* F *de classe* C^r.

Vérifier qu'une application injective $f : \mathrm{U} \to \mathrm{F}$ est un homéomorphisme sur son image revient à vérifier qu'elle induit une application fermée : $\mathrm{U} \to \mathrm{Im}\, f$. Pour ceci, il est souvent utile d'introduire la notion de propreté.

DÉFINITION 12. — *Soient* X *et* Y *deux espaces topologiques localement compacts ; une application continue* $f : \mathrm{X} \to \mathrm{Y}$ *est dite propre si l'image réciproque d'un compact de* Y *est un compact.*

Si $f : \mathrm{X} \to \mathrm{Y}$ est une application propre, les fibres de f sont évidemment compactes.

PROPOSITION 13. — *Soient* X *et* Y *deux espaces topologiques localement compacts et* $f : \mathrm{X} \to \mathrm{Y}$ *une application propre. Alors* f *est fermée.*

Démonstration. Soit F un fermé de X ; il s'agit de vérifier que $f(\mathrm{F})$ est fermé. Soit $y \in \mathrm{Y}$ appartenant à l'adhérence de $f(\mathrm{F})$. Considérons un voisinage compact K de y. Alors $f^{-1}(\mathrm{K})$ est un compact, qui rencontre F. Pour tout voisinage compact $\mathrm{W} \subset \mathrm{K}$ le fermé $\mathrm{F} \cap f^{-1}(\mathrm{W})$ est non vide ; ces parties sont fermées dans le compact $\mathrm{F} \cap f^{-1}(\mathrm{K})$, et cette famille de fermés est stable par intersection finie. Il en résulte que

$$\cap_{\mathrm{W}} f^{-1}(\mathrm{W}) \cap \mathrm{F} \neq \emptyset$$

Si x est un point de cette intersection, le point $f(x)$ appartient à tous les voisinages W et par suite, puisque Y est séparé, $f(x) = y$. Donc y appartient à $f(\mathrm{F})$. □

Exercice 5.8

Démontrer que l'application $f : \mathbb{R} \to \mathbb{C}$ définie par $t \mapsto e^{(1+i)t}$ est un plongement.

Exercice 5.9

On considère l'application $f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^4 \setminus \{0\}$ définie par

$$f(x, y) = (x^3, x^2y, xy^2, y^3)$$

1) Montrer que f est une immersion propre.

2) Démontrer que f est ouverte sur son image. En déduire que l'image M de f est une sous-variété de classe C^∞ de $\mathbb{R}^4$, de dimension 2.

3) Trouver des équations locales pour M au voisinage du point $(1, 0, 0, 0)$.

Exercice 5.10

1) Montrer que l'application $\mathbb{R} \to \mathbb{C}$ définie par

$$f : t \mapsto e^{it} \cos 2t$$

est une immersion et que l'image n'est pas une sous-variété de $\mathbb{C} = \mathbb{R}^2$.

2) Montrer que la restriction g de f à l'intervalle $-]\frac{\pi}{2}, \frac{\pi}{4}[$ est une immersion injective. Vérifier que l'image de g n'est pas une sous-variété de $\mathbb{C} = \mathbb{R}^2$.

Exercice 5.11

Soit E un espace vectoriel de dimension n, et $\mathrm{N} \subset \mathrm{End(E)}$ l'ensemble des endomorphismes nilpotents de degré exactement n. On considère l'action naturelle de GL(E) sur End(E) par conjugaison.

1) Soit $v \in \mathrm{N}$. Démontrer qu'il existe une base $e_1, \ldots, e_n$ de E telle que $v(e_i) = e_{i+1}$ pour $1 \leq i \leq n-1$ et $v(e_n) = 0$. En déduire que N est une orbite de GL(E).

2) Soit $v \in \mathrm{N}$. Démontrer que l'application $\phi : \mathrm{GL(E)} \to \mathrm{N}$ définie par $\phi(g) = gvg^{-1}$ est de classe C^∞, et qu'elle a des sections locales continues. En déduire que N est une sous-variété de End(E).

3) Préciser l'espace tangent en un point et la dimension de N.

Démonstration du théorème 9.

Soit $a \in \Omega$ et $b = f(a)$. D'après le théorème du rang, on peut trouver des voisinages $\mathrm{U} \subset \Omega$ et V de a et b respectivement, et des difféomorphismes $\phi : \mathrm{U} \to \mathrm{U}'$ et $\psi : \mathrm{V} \to \mathrm{V}'$ tels que $f(\mathrm{U}) \subset \mathrm{V}$ et tels que $g = \psi \circ f \circ \phi^{-1}$ soit induite par une application linéaire. Soit $a' = \phi(a)$, et $b' = \psi(b)$. Considérons le diagramme commutatif

$$\begin{array}{ccc} \mathrm{U} & \xrightarrow{f} & \mathrm{V} \\ \phi \downarrow & & \downarrow \psi \\ \mathrm{U}' & \xrightarrow{g} & \mathrm{V}' \end{array}$$

Par hypothèse, f est ouverte sur son image : ainsi, il existe un ouvert W de F tel que $f(\mathrm{U}) = \mathrm{Im}\, f \cap \mathrm{W}$. Remarquons que $g(\mathrm{U}') \subset \psi(\mathrm{V} \cap \mathrm{W})$. Quitte à remplacer au besoin V par $\mathrm{V} \cap \mathrm{W}$, et ψ par sa restriction : $\mathrm{V} \cap \mathrm{W} \to \psi(\mathrm{V} \cap \mathrm{W})$, on peut supposer que $f(\mathrm{U}) = \mathrm{Im}\, f \cap \mathrm{V}$. L'image de $\mathrm{Im}\, f \cap \mathrm{V}$ par ψ est alors $g(\mathrm{U}')$, et, d'après l'exercice 3.2, il s'agit de vérifier que $g(\mathrm{U}')$ est une sous-variété de F.

On est donc ramené à vérifier que l'image d'un ouvert de E par une application linéaire $G : E \to F$ est une sous-variété. Mais $\operatorname{Im} G$ est un sous-espace vectoriel, et l'application linéaire surjective $G : E \to \operatorname{Im} G$ est ouverte. L'image d'un ouvert de E est donc un ouvert d'un espace vectoriel.

Pour trouver l'espace tangent à $\operatorname{Im} f$ en $b = f(a)$, il suffit de contempler le diagramme commutatif

$$\begin{array}{ccc} E & \xrightarrow{d_a f} & F \\ {\scriptstyle d_a\phi}\downarrow & & \downarrow{\scriptstyle d_b\psi} \\ E & \xrightarrow{dg} & F \end{array}$$

et de remarquer que l'espace tangent à $\operatorname{Im} g$ en un point est l'image de dg. Ceci achève la démonstration. □

5.5. Morphismes

Soit ν un entier ≥ 1, ou $\nu = \infty$.

DÉFINITION 14. — *Soit* M *une sous-variété de classe* C^ν *de* $\mathbb{R}^n$. *Un homéomorphisme* $\varphi : V \to V'$ *d'un ouvert* V *de* M *sur un ouvert* V′ *d'un sous-espace vectoriel* F *de* $\mathbb{R}^n$ *s'appelle une carte locale s'il existe un difféomorphisme* $\phi : U \to U'$ *de classe* C^ν *d'un ouvert* U *de* $\mathbb{R}^n$ *sur un ouvert* U′ *de* $\mathbb{R}^n$ *tel que* $V = M \cap U$, *et* $\varphi = \phi|_V$.

L'ouvert V s'appelle la source de la carte locale, et l'ouvert V′ de F le but. Quitte à diminuer U et U′, on peut au besoin imposer en outre que $U' \cap F = V'$. Cette condition n'est pas automatique avec la définition ci-dessus. Si φ est une carte locale, sa restriction $\varphi|_W : W \to \varphi(W)$ à un ouvert $W \subset V$ est encore une carte locale.

DÉFINITION 15. — *Soit* $f : M \to \mathbb{R}^p$ *une application. On dit que* f *est de classe* C^ν *si pour tout* $a \in M$, *il existe une carte locale* $\varphi : V \to V'$ *dont la source contient* a, *et telle que* $f \circ \varphi^{-1}$ *soit de classe* C^ν *sur* V′.

L'application $f \circ \varphi^{-1}$ s'appelle l'expression de f dans la carte ϕ. Si $f : M \to \mathbb{R}^p$ prend ses valeurs dans une sous-variété $N \subset \mathbb{R}^p$ on dira que $f : M \to N$ est de classe C^ν si l'application composée $M \to N \hookrightarrow \mathbb{R}^p$ est de classe C^ν.

PROPOSITION 16. — *Soient* M *et* N *deux sous-variétés de classe* C^ν *de* $\mathbb{R}^n$ *et* $\mathbb{R}^p$ *respectivement et* $f : M \to N$ *une application de classe* C^ν. *Soit* $v \in T_aM$, *et* $\gamma :]-\epsilon, \epsilon[\to M$ *un chemin dérivable en 0 tel que* $\gamma(0) = a, \gamma'(0) = v$.

1. *Le chemin* $t \mapsto f(\gamma(t))$ *est dérivable en 0.*
2. *Sa dérivée en* 0 *ne dépend que de* v. *On la note* $T_af.v$
3. *L'application* $T_af : T_aM \to T_{f(a)}N$ *est linéaire.*

DÉFINITION 17. — *L'application linéaire* $T_af : T_aM \to T_{f(a)}N$ *ainsi définie s'appelle l'application linéaire tangente en* a *à* f.

Démonstration. Remarquons que si M est un ouvert de $\mathbb{R}^n$, l'énoncé est évident : l'application T_af n'est autre que la différentielle de $d_af : \mathbb{R}^n \to \mathbb{R}^p$, qui prend ses valeurs dans $T_{f(a)}N$.

Dans le cas général, on choisit au voisinage de a une carte locale $\varphi : V \to V' \subset F$ dont la source contient a induite par un difféomorphisme ϕ d'un ouvert U sur un ouvert U′ de l'espace vectoriel $\mathbb{R}^n$ ambiant, et telle que $g = f \circ \phi^{-1} : V' \to \mathbb{R}^p$ soit de classe C^ν. Alors $t \mapsto \phi(\gamma(t))$ est dérivable en 0, et donc $f(\gamma(t)) = g(\phi(\gamma(t)))$ est dérivable en 0; sa dérivée est donnée par $d_{\phi(a)}g(d_a\phi(v))$. Ainsi, cette dérivée ne dépend pas de γ, (ni de φ) mais seulement de v, et est évidemment linéaire en v. D'où l'énoncé. □

PROPOSITION 18. — *Soient* M *et* N *deux sous-variétés de classe* C^ν *de* $\mathbb{R}^n$ *et* $\mathbb{R}^p$ *respectivement, et* $f : M \to N$ *une application. Alors* f *est de classe* C^ν *si et seulement si pour tout point* $a \in M$ *il existe un voisinage ouvert* U *de* a *dans* $\mathbb{R}^n$ *et une application* $G : U \to \mathbb{R}^p$ *de classe* C^ν *telle que*

$$G|_{U\cap M} = f|_{U\cap M}.$$

On a alors $T_af = d_aG|_{T_aM}$.

Démonstration. Supposons que f soit de classe C^ν. Choisissons une carte locale $\varphi : V \to V' \subset F$ au voisinage de a, induite par un difféomorphisme ϕ d'ouverts de l'espace vectoriel ambiant et telle que $g = f \circ \phi^{-1} : V' \to \mathbb{R}^p$ soit de classe C^ν. Choisissons un facteur direct S de F dans $\mathbb{R}^n$ et désignons par π la projection $\mathbb{R}^n \to F$. Alors $g \circ \pi$ est de classe C^ν sur l'ouvert $\pi^{-1}(V')$. Considérons l'application $G = g \circ \pi \circ \phi$: cette application est définie et de classe C^ν sur un ouvert U de $\mathbb{R}^n$ contenant V et coïncide avec f sur $U \cap M$.

Réciproquement, supposons que f s'étende au voisinage U de $a \in \mathrm{M}$ dans $\mathbb{R}^n$ en une application G de classe C^ν. Soient $\varphi : \mathrm{V} \to \mathrm{V}'$ une carte locale arbitraire induite par un difféomorphisme ϕ d'un ouvert $\mathrm{W} \subset \mathbb{R}^n$ sur un ouvert $\mathrm{W}' \subset \mathbb{R}^n$, et a un point de V; l'application $f \circ \varphi^{-1}$ est alors induite par $\mathrm{G} \circ \phi^{-1}$ au voisinage de $\varphi(a)$ et est donc de classe C^ν au voisinage de $\varphi(a)$. Elle est donc de classe C^ν sur V'.

Soit $\gamma :]-\epsilon, \epsilon[\to \mathrm{M}$ un chemin dérivable tel que $\gamma'(0) = v$. Puisque $\mathrm{G}|_{\mathrm{U}\cap\mathrm{M}} = f|_{\mathrm{U}\cap\mathrm{M}}$, on a $f(\gamma(t)) = \mathrm{G}(\gamma(t))$ au voisinage de 0, et donc en prenant la dérivée à l'origine, on obtient $\mathrm{T}_a f(v) = d_a\mathrm{G}(v)$. Ceci démontre la formule $\mathrm{T}_a f = d_a\mathrm{G}|_{\mathrm{T}_a\mathrm{M}}$.

Dans la pratique, les applications de classe C^ν seront toujours obtenues comme restrictions d'applications de classe C^ν définies sur des ouverts de l'espace ambiant.

COROLLAIRE 19. — *Si $f : \mathrm{M} \to \mathrm{N} \subset \mathbb{R}^p$ est de classe C^ν, l'expression de f dans toute carte locale φ de M est de classe C^ν.*

Composition

Soient M, N et P trois sous-variétés de classe C^ν plongées dans des espaces affines et $f : \mathrm{M} \to \mathrm{N}$ et $g : \mathrm{N} \to \mathrm{P}$ des applications de classe C^ν. Alors $g \circ f : \mathrm{M} \to \mathrm{P}$ est de classe C^ν, et on a pour $a \in \mathrm{M}$ et $b = f(a)$

$$\mathrm{T}_a(g \circ f) = \mathrm{T}_{f(a)}g \circ \mathrm{T}_a f$$

Cette propriété est immédiate en choisissant des applications de classe C^ν définies au voisinage de a et $f(a)$ respectivement dans les espaces vectoriels ambiants, qui induisent f et g.

Difféomorphismes

DÉFINITION 20. — *Soient* M *et* N *deux sous-variétés de classe* C^ν *de* $\mathbb{R}^n$ *et* $\mathbb{R}^p$ *respectivement. Une application $f : \mathrm{M} \to \mathrm{N}$ est un difféomorphisme si elle est inversible, et si f et f^{-1} sont de classe* C^ν.

Exemples

1. Soit M une sous-variété de $\mathbb{R}^n$ et $\varphi : \mathrm{V} \to \mathrm{V}'$ une carte locale. Alors φ est un difféomorphisme de classe C^ν.

2. Considérons l'hyperboloïde $\mathrm{M} \subset \mathbb{R}^3$ défini par l'équation $x^2 + y^2 - z^2 = 1$. L'application $\mathrm{M} \to \mathrm{S}_1 \times \mathbb{R}$ définie par

$$(x, y, z) \mapsto (\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}, z)$$

est un difféomorphisme de classe C^∞.

Si f est un difféomorphisme de classe C^1, l'application linéaire tangente $\mathrm{T}_a f$ en un point $a \in \mathrm{M}$ est inversible, et on a,

$$(\mathrm{T}_a f)^{-1} = \mathrm{T}_{f(a)}\mathrm{M}.$$

Réciproquement, si cette application linéaire tangente est inversible, on peut en déduire que f est localement au voisinage de a, un difféomorphisme :

THÉORÈME 21. — *Soient* M *et* N *deux sous-variétés de classe* C^ν *de* $\mathbb{R}^n$ *et* $\mathbb{R}^p$ *respectivement, et* $f : \mathrm{M} \to \mathrm{N}$ *un application de classe* C^ν. *On suppose que l'application linéaire tangente* $\mathrm{T}_a f : \mathrm{T}_a\mathrm{M} \to \mathrm{T}_{f(a)}\mathrm{N}$ *est inversible en un point* $a \in \mathrm{M}$. *Alors il existe un voisinage ouvert* U *de* a *dans* M *tel que*

(i) *l'image* $\mathrm{V} = f(\mathrm{U})$ *est un ouvert de* N;

(ii) *l'application* $f|_{\mathrm{U}} : \mathrm{U} \to \mathrm{V}$ *est un difféomorphisme de classe* C^ν.

Démonstration. On prend des cartes locales $\varphi : \mathrm{V} \to \mathrm{V}'$ pour M au voisinage de a, et $\psi : \mathrm{W} \to \mathrm{W}'$ pour N au voisinage de $b = f(a)$ telles que $f(\mathrm{V}) \subset \mathrm{W}$ et on considère l'application

$$g = \psi \circ f \circ \varphi^{-1} : \mathrm{V}' \to \mathrm{W}'$$

Cette application est de classe C^ν, et sa différentielle $d_{a'}g$ est inversible au point $a' = f(a)$. Il résulte du théorème d'inversion locale que g induit un difféomorphisme de classe C^ν d'un voisinage ouvert $\mathrm{V}'' \subset \mathrm{V}'$ de a' sur un voisinage ouvert W'' de $g(a')$. Alors $f = \psi^{-1} \circ g \circ \varphi$ est un difféomorphisme de classe C^ν de l'ouvert $\varphi^{-1}(\mathrm{V}'')$ sur l'ouvert $\psi^{-1}(\mathrm{W}'')$. □

COROLLAIRE 22. — *Soient* M *et* N *deux sous-variétés de* $\mathbb{R}^n$ *et* $\mathbb{R}^p$ *respectivement, et* $f : \mathrm{M} \to \mathrm{N}$ *une application de classe* C^ν, *bijective, telle que l'application linéaire tangente* $\mathrm{T}_a f$ *soit inversible en tout point* a *de* M. *Alors* f *est un difféomorphisme de classe* C^ν.

On peut aussi transposer aux sous-variétés de $\mathbb{R}^n$ tous les énoncés obtenus à partir du théorème du rang constant.

DÉFINITION 23. — *Soit* $f : \mathrm{M} \to \mathrm{N}$ *une application de classe* C^ν. *On appelle rang de* f *en un point* a, *et on note* $\mathrm{rg}_a(f)$, *le rang de l'application linéaire tangente* $\mathrm{T}_a f$.

Exercice 5.12
Montrer que la fonction $a \mapsto \mathrm{rg}_a f$ est semi-continue inférieurement.

PROPOSITION 24. — *Soient* M *et* N *deux sous-variétés de* $\mathbb{R}^n$ *et* $\mathbb{R}^p$ *respectivement, et* $f : \mathrm{M} \to \mathrm{N}$ *une application de classe* C^ν *de rang constant. Alors*

(i) *Les fibres* $\mathrm{M}_b = f^{-1}(b)$ *sont des sous-variétés de* $\mathbb{R}^n$, *d'espace vectoriel tangent en un point* a

$$\mathrm{T}_a\mathrm{M}_b = \ker \mathrm{T}_a f$$

(ii) *Si* f *est ouverte sur son image, l'image de* f *est une sous-variété de* $\mathbb{R}^p$, *et on a en un point* $f(a) \in \mathrm{Im}\, f$

$$\mathrm{T}_{f(a)}\mathrm{Im}\, f = \mathrm{Im}\, \mathrm{T}_a f.$$

On utilise le plus couramment la partie (i) de cet énoncé quand l'application linéaire tangente à f est surjective en tout point : on dit alors que f est une submersion.

DÉFINITION 25. — *Un point critique de* f *est un point* $a \in \mathrm{M}$ *tel que* $\mathrm{rg}_a(f) < \dim \mathrm{N}$; *une valeur critique de* f *une valeur de* f *en un point critique.*

Si b n'est pas valeur critique, l'application f est une submersion sur un voisinage ouvert de la fibre $\mathrm{M}_b = f^{-1}(b)$. On obtient :

COROLLAIRE 26. — *Si* b *n'est pas valeur critique de* $f : \mathrm{M} \to \mathrm{N}$, *la fibre* $\mathrm{M}_b = f^{-1}(b)$ *est une sous-variété de* $\mathbb{R}^n$.

La partie (ii) de l'énoncé s'utilise souvent quand l'application linéaire $\mathrm{T}_a f$ est injective en tout point : on dit alors que f est une immersion. Ainsi,

COROLLAIRE 27. — *Si* $f : \mathrm{M} \to \mathrm{N}$ *est une immersion injective et propre, l'image* $f(\mathrm{M})$ *est une sous-variété de* $\mathbb{R}^p$*, et l'application* $\mathrm{M} \to \mathrm{Im}\, f$ *induite par* f *un difféomorphisme de classe* C^ν.

Exercice 5.13

Montrer que la surface de $\mathbb{R}^3$ d'équation $x^4 + y^4 + z^4 = 1$ est difféomorphe à la sphère S_2.

Exercice 5.14

La cubique de $\mathbb{R}^2$ d'équation $x^3 - x + y^2 = 0$ a deux composantes connexes, dont l'une est difféomorphe au cercle S_1, et l'autre à $\mathbb{R}$.

Exercice 5.15

Le tore de $\mathbb{R}^3$ d'équation

$$(\sqrt{x^2 + y^2} - 2)^2 + z^2 = 1$$

est difféomorphe à $\mathrm{S}_1 \times \mathrm{S}_1$.

Exercice 5.16

Soit S_2 la sphère de centre 0 et de rayon 1 dans $\mathbb{R}^3$.

1. Montrer que l'application $f : \mathrm{S}_2 \to \mathbb{R}^4$ définie par

$$(x, y, z) \mapsto (x^2 - y^2, xy, yz, zx)$$

est une immersion et que l'image de f est une sous-variété de dimension 2 de $\mathbb{R}^4$.

2. Trouver des équations locales pour $\mathrm{Im}\, f$ au voisinage du point $(1, 0, 0, 0)$.

6. Fonctions de Morse

On suppose dans ce qui suit que ν est un entier ≥ 2. Soit $f : \mathrm{U} \to \mathbb{R}$ une application de classe C^ν définie sur un ouvert U d'un espace vectoriel E. Les points critiques de f sont les points $a \in \mathrm{U}$ tels que $d_a f = 0$. On sait que si $a \in \mathrm{U}$ n'est pas un point critique, une telle application est conjuguée (au voisinage de a) à une application linéaire $\mathrm{E} \to \mathbb{R}$, modulo l'action d'un difféomorphisme défini au voisinage de a. On se préoccupe dans cette section d'examiner ce qui se passe au voisinage des points critiques raisonnables.

6.1. Lemme de Morse

Rappelons que la différentielle seconde $d_a^2 f$ au point $a \in \mathrm{U}$ est par définition la forme bilinéaire sur $\mathrm{E} \times \mathrm{E} \to \mathbb{R}$ définie par $d_a(df) \in \mathrm{L}(\mathrm{E}, \mathrm{L}(\mathrm{E}, \mathbb{R}))$; autrement dit, pour u et $v \in \mathrm{E}$ on a $d_a^2 f(u, v) = d_a(df)(u)(v)$. Le lemme de Schwarz dit que cette forme bilinéaire est symétrique. En termes de dérivées partielles, si $(x_1, \ldots, x_n)$ sont les coordonnées dans une base e_i on a

$$d_a^2 f(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

On désigne encore par $\mathrm{H}_a f$ la forme quadratique associée à cette forme bilinéaire symétrique : c'est le Hessien de f en a. On a donc par définition pour $x \in \mathrm{E}$

$$\mathrm{H}_a f(x) = d_a^2 f(x, x)$$

DÉFINITION 1. — *Soit a un point critique de f. On dit que f est un point critique non dégénéré si le Hessien $\mathrm{H}_a f$ est une forme quadratique non dégénérée.*

DÉFINITION 2. — *Un point critique non dégénéré a est dit d'indice p si la forme quadratique $\mathrm{H}_a f$ est de signature $(n-p, p)$.*

Exercice 6.1

Soit $\phi : \mathrm{U} \to \mathrm{U}'$ un difféomorphisme et $b = \phi(a)$ et $g = f \circ \phi^{-1}$.

1. Démontrer que a est un point critique de f si et seulement si b est un point critique de g et on a alors

$$\mathrm{H}_b(b) \circ \phi) = \mathrm{H}_a f \circ d_b \phi.$$

2. En déduire que si a est un point critique non dégénéré d'indice p pour f, le point b est un point critique non dégénéré d'indice p pour la fonction g.

Exercice 6.2

Démontrer que si $f : \mathrm{U} \to \mathbb{R}$ n'a que des points critiques non dégénérés, l'ensemble des points critiques de f est un ensemble discret.

Au voisinage d'un point critique $a \in \mathrm{U}$ le développement de Taylor à l'ordre 2 s'écrit

$$f(a+h) = f(a) + \frac{1}{2}\mathrm{H}_a f(h) + \epsilon(h) \parallel h \parallel^2$$

où $\epsilon(h)$ est une fonction réelle définie au voisinage de 0 telle que $\lim_{h\to 0} \epsilon(h) = 0$. On va montrer que si a est un point critique non dénégénéré, il existe un difféomorphisme qui transforme f en son développement de Taylor à l'ordre 2 : c'est le lemme de Morse.

LEMME 3. — *(Lemme de Morse)*

Soit U *ouvert de* E *et* $f : \mathrm{U} \to \mathbb{R}$ *une fonction de classe* C^∞. *On suppose que* $a \in \mathrm{U}$ *est un point critique non dégénéré pour* f. *Alors il existe un difféomorphisme* $\phi : \mathrm{V} \to \mathrm{W}$ *défini sur voisinage ouvert de* 0 *satisfaisant aux conditions suivantes :*

(i) $\phi(0) = 0$ *et* $d_0\phi = \mathrm{id}_\mathrm{E}$ *;*

(ii) *au voisinage de* 0 *on a*

$$f(a+\phi(x)) = f(a) + \frac{1}{2}\mathrm{H}_a f(x).$$

Démonstration. Commençons par écrire la petite formule de Taylor avec reste sous forme intégrale, sur une boule ouverte $\mathrm{B}(a,r) \subset \mathrm{U}$. Du fait que a est un point critique, on obtient

$$f(a+x) = f(a) + \int_0^1 (1-t)\mathrm{H}_{a+tx} f(x)dt.$$

Considérons la forme quadratique $q_x = \int_0^1 (1-t)\mathrm{H}_{a+tx} f dt$. Du fait f est de classe C^∞, il résulte que l'application définie sur $\mathrm{B}(0,r)$ et à valeurs dans l'espace vectoriel des formes quadratiques $x \mapsto q_x$ est de classe C^∞. La formule ci-dessus s'écrit

$$f(a+x) = f(a) + q_x(x)$$

pour $x \in \mathrm{B}(0,r)$.

On munit E de la forme quadratique non dégénérée $q_0 = \frac{1}{2}\mathrm{H}_a f$, et on désigne par $(u,v) \mapsto \langle u,v\rangle$ la forme polaire de q_0. On désigne par $\mathrm{Sym(E)} \subset \mathrm{End(E)}$ l'espace vectoriel des applications linéaires $\mathrm{E} \to \mathrm{E}$, symétriques pour cette forme quadratique. On désigne d'autre part par $\mathrm{S}^2\mathrm{E}^*$ l'espace vectoriel des formes

quadratiques sur E. L'application linéaire $\mathrm{Sym(E)} \to \mathrm{S^2E^*}$ qui associe à un endomorphisme symétrique α la forme quadratique Q_α définie par

$$\mathrm{Q}_\alpha(x) = \langle \alpha(x), x \rangle$$

est un isomorphisme d'espaces vectoriels.

Tout endomorphisme $\alpha \in \mathrm{End(E)}$ définit à partir de q_0 une nouvelle forme quadratique $\alpha^*(q_0) = q_0 \circ \alpha$. Considérons l'équation à valeurs dans l'espace vectoriel des formes quadratiques $\mathrm{S^2E^*}$

$$\alpha^*(q_0) = q_x$$

où $\alpha \in \mathrm{Sym(E)}$ est l'inconnue. L'application $\mathrm{F} : \mathrm{B}(0,r) \times \mathrm{Sym(E)} \to \mathrm{S^2E^*}$ définie par $(x, \alpha) \mapsto \alpha^*(q_0) - q_x$ est de classe C^∞. La différentielle partielle $\partial_2\mathrm{F}(0, \mathrm{id_E})$ est donnée par $\alpha \mapsto 2\mathrm{Q}_\alpha$. Cette différentielle partielle est donc inversible. Il résulte du théorème des fonctions implicites qu'il existe une application $x \mapsto \alpha(x)$ de classe C^∞, définie au voisinage de 0 dans E et à valeurs dans Sym(E) telle que $\alpha(0) = \mathrm{id_E}$ et telle que

$$\alpha(x)^*(q_0) = q_x$$

Cette égalité entraîne en particulier $q_x(x) = q_0(\alpha(x)x)$, et donc

$$f(a+x) = f(a) + \frac{1}{2}\mathrm{H}_a f(\alpha(x)x).$$

Considérons l'application $\beta : x \mapsto \alpha(x)x$, définie au voisinage de 0 dans E et à valeurs dans E. Cette application est de classe C^∞, et sa différentielle en $x = 0$ est $d_0\beta = \mathrm{id_E}$. Ainsi, β induit un difféomorphisme d'un voisinage ouvert de 0 sur un voisinage ouvert de 0. Le difféomorphisme inverse ϕ satisfait aux conditions demandées. □

On peut formuler le lemme de Morse de manière plus concrète :

COROLLAIRE 4. — *Soit $f : \mathrm{U} \to \mathbb{R}$ ayant en a un point critique non dégénéré d'indice p. Il existe au voisinage de a des fonctions scalaires $u_1, \ldots, u_n$ de classe C^∞ s'annulant en a, telles que*

(i) $f(x) = f(a) + u_1^2 + \ldots + u_{n-p}^2 - u_{n-p+1}^2 - \ldots - u_n^2$

(ii) *L'application $x \mapsto (u_1(x), \ldots, u_n(x))$ induit un difféomorphisme d'un voisinage ouvert de a sur un voisinage ouvert de 0 dans $\mathbb{R}^n$.*

Démonstration. Il suffit de choisir une base $e_1, \ldots, e_n$ orthogonale pour la forme quadratique $q_0 = \frac{1}{2}\mathrm{H}_a f$ et telle $q_0(e_i) = 1$ si $1 \leq i \leq n-p$ et $q_0(e_i) = -1$ si $n-p+1 \leq i \leq n$. Une telle base orthogonale existe d'après le théorème de réduction des formes quadratiques réelles de signature $(n-p, p)$. Si β est l'application introduite dans la démonstration ci-dessus, on peut prendre pour fonctions u_i les composantes de l'application $a + x \mapsto \beta(x)$ dans cette base. □

Exercice 6.3

Soit $f : \mathrm{U} \to \mathbb{R}$ une fonction de classe C^∞ n'ayant qu'un nombre fini de points critiques non dégénérés. On considère un point critique a d'indice p, et on désigne par c la valeur critique correspondante. Etudier, au voisinage de a, l'allure de l'hypersurface de niveau définie par l'équation $f(x) = y$ quand y est voisin de la valeur critique c.

6.2. Fonctions définies sur des sous-variétés de $\mathbb{R}^n$

On considère une sous-variété M de classe C^ν, de dimension m d'un espace vectoriel $\mathrm{E} = \mathbb{R}^n$ et une fonction $\mathrm{M} \to \mathbb{R}$ de classe C^ν. Parmi les points critiques, il y a bien les points où f a un maximum ou un minimum local, mais il y en a en général beaucoup d'autres.

On sait que localement, f est induite au voisinage d'un point $a \in \mathrm{M}$ par une fonction de classe C^ν sur un voisinage ouvert de a dans l'espace ambiant ; on peut en déduire, en faisant l'usage de partitions de l'unité, qu'il existe une fonction $\mathrm{F} : \mathrm{U} \to \mathbb{R}$ de classe C^ν sur un ouvert U contenant M telle que $\mathrm{F}|_\mathrm{M} = f$. Dans la pratique, les fonctions f qui seront considérées seront toujours présentées de cette façon, de sorte qu'il n'est pas très utile de vérifier cet énoncé de nature globale.

PROPOSITION 5. — *Soit* U *un ouvert de* $\mathrm{E} = \mathbb{R}^n$ *et* $g = (g_1, \ldots, g_r) : \mathrm{U} \to \mathbb{R}^r$ *une submersion. On considère la sous-variété* $\mathrm{M} \subset \mathrm{U}$ *définie par les équations* $g_1 = \ldots = g_r = 0$.

Soit $\mathrm{F} : \mathrm{U} \to \mathbb{R}$ *une fonction de classe* C^ν, *et* $f = \mathrm{F}|_\mathrm{M}$. *Un point* a *de* M *est un point critique pour* f *si et seulement si il existe* $(\lambda_1, \ldots, \lambda_r) \in \mathbb{R}^r$ *tels que* a *soit un point critique pour la fonction*

$$\mathrm{F} - \sum_{i=1}^{r} \lambda_i g_i$$

Localement, on est toujours dans la situation de la proposition. Commençons, avant de démontrer cet énoncé, par un lemme d'algèbre linéaire :

LEMME 6. — *Soit* T *un sous-espace vectoriel de* E, *de codimension* r, *défini par les équations linéaires* $\varphi_1(x) = \ldots = \varphi_r(x) = 0$. *Alors l'orthogonal de* T *dans* E^* *est le sous-espace engendré par* $(\varphi_1, \ldots, \varphi_r)$.

Démonstration. Soit S un supplémentaire de T; alors $\varphi = (\varphi_1, \ldots, \varphi_r)$ induit un isomorphisme $\mathrm{S} \to \mathbb{R}^r$. Toute forme linéaire ψ sur $\mathrm{E} = \mathrm{T} \oplus \mathrm{S}$ s'écrit dans cette somme directe $\psi = (\psi' \ \psi'')$. On peut factoriser $\psi'' = \lambda \circ \varphi|_{\mathrm{S}}$ où $\lambda : \mathbb{R}^r \to \mathbb{R}$ est une forme linéaire. Si $\lambda = (\lambda_1, \ldots, \lambda_r)$ cette égalité s'écrit aussi

$$\psi'' = \sum_i \lambda_i \varphi_i|_{\mathrm{S}}$$

et par suite $\psi - \sum_{i=1}^r \lambda_i \varphi_i$ s'annule sur S. Dire que ψ appartient à l'orthogonal de S signifie que ψ s'annule sur T ; ceci implique donc qu'il existe $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ tels que

$$\psi = \sum_{i=1}^r \lambda_i \varphi_i.$$

Réciproquement, une telle égalité entraîne évidemment que ψ appartient à l'orthogonal de T. □

Démonstration de la proposition 5

Par définition, un point $a \in \mathrm{M}$ est un point critique pour f si et seulement si $\mathrm{T}_a f = 0$, autrement dit d'après la proposition 13.18

$$d_a \mathrm{F}|_{\mathrm{T}_a \mathrm{M}} = 0.$$

On sait que $\mathrm{T}_a\mathrm{M}$ est le sous-espace vectoriel de E défini par les vecteurs $v \in \mathrm{E}$ satisfaisant aux équations $d_a g_1(v) = \ldots = d_a g_r(v) = 0$. Ainsi, si a est un point critique de f, la différentielle $d_a\mathrm{F}$ appartient à l'orthogonal de $\mathrm{T}_a\mathrm{M}$, et d'après le lemme précédent, il existe des scalaires $(\lambda_1, \ldots, \lambda_r)$ tels que

$$d_a \mathrm{F} = \sum_{i=1}^r \lambda_i d_a g_i.$$

Ainsi, la différentielle de $\mathrm{F} - \sum_{i=1}^r \lambda_i g_i$ est nulle en a. Réciproquement, si c'est le cas, on voit que la restriction de $d_a\mathrm{F}$ à $\mathrm{T}_a\mathrm{M}$ est nulle, et par conséquent a est un point critique pour f. □

Si a est un point critique pour f, les nombres $(\lambda_1, \ldots, \lambda_r)$ sont déterminés de manière unique, puisque les différentielles $(d_a g_1, \ldots, d_a g_r)$ sont indépendantes. On les appelle les *multiplicateurs de Lagrange.*

Exemple

Considérons sur la sphère S_n de $\mathbb{R}^{n+1}$ d'équation

$$g(x) = x_0^2 + \ldots + x_n^2 - 1 = 0$$

la fonction $f = x_n|_{\mathrm{S}_n}$. Les points critiques de f sont les points a de la sphère S_n où les différentielles $(dx_n \ d_a g)$ sont liées. On trouve ainsi 2 points critiques $(0, \ldots, 0, \pm 1)$. Puisqu'il n'y qu'une équation pour S_n, il y a pour chaque valeur critique un seul multiplicateur de Lagrange, qui est $\lambda = \frac{1}{2}$ pour le pôle nord $(0, \ldots, 0, 1)$, et $-\frac{1}{2}$ pour le pôle sud $(0, \ldots, 0, -1)$. Ces deux points critiques correspondent au maximum et au minimum de la fonction f.

Nature des points critiques

Soit M une sous-variété de dimension m, de classe C^ν de $\mathrm{E} = \mathbb{R}^n$. Considérons une fonction $f : \mathrm{M} \to \mathbb{R}$ de classe C^ν, ayant au point a un point critique. Pour déterminer la nature du point critique, il convient d'étudier le comportement de f le long des chemins dérivables $\gamma :]-\epsilon, \epsilon[\to \mathrm{M}$ qui passent par a à l'instant $t = 0$. Considérons une carte locale $\varphi : \mathrm{V} \to \mathrm{V}'$ pour f au voisinage de a, et l'expression $g = f \circ \varphi^{-1}$ de f; on a alors, si $v = \gamma'(0)$ et $b = \varphi(a)$

$$f(\gamma(t)) = g(\varphi(\gamma(t)) = f(a) + \frac{t^2}{2} \mathrm{H}_b g \circ d_a \varphi(v) + \epsilon(t)|t|^2$$

où $\lim_{t \to 0} \epsilon(t) = 0$.

LEMME 7. — *La forme quadratique sur* $\mathrm{T}_a\mathrm{M}$ *définie par*

$$v \mapsto \mathrm{H}_b g \circ d_a \varphi$$

est indépendante du choix de la carte locale.

Démonstration. Il suffit d'utiliser le développement limité à l'ordre 2 de la fonction $t \mapsto f(\gamma(t)) :]-\epsilon, \epsilon[\to \mathbb{R}$: ce développement limité ne dépend pas du choix de la carte φ. □

Exercice 6.4

On considère un autre carte locale φ_1 telle que $\varphi_1(a) = b_1$ et on écrit $\varphi_1 = \theta\varphi$, où θ est un difféomorphisme de classe C^ν au voisinage de b. Alors l'expression g_1 de f au voisinage de b_1 est donnée par $g_1 = g \circ \theta^{-1}$. Démontrer directement que

$$H_b g \circ d_a\varphi = H_{b_1} g_1 \circ d_a\varphi_1$$

en utilisant l'exercice 14.1.

DÉFINITION 8. — *La forme quadratique ainsi définie sur* $T_a M$ *s'appelle le Hessien de* f *au point critique* a *et en notée* $H_a f$.

DÉFINITION 9. — *Soit* $m = \dim M$. *Si le Hessien* $H_a f$ *est une forme quadratique non dégénérée, on dit que* a *est un point critique non dégénéré. Si cette forme quadratique a pour signature* $(m-p, p)$ *on dit que* a *est un point critique d'indice* p.

Si a est un point critique d'indice m, alors f a en a un maximum local ; si a est un point critique d'indice 0, alors f a en a un minimum local.

DÉFINITION 10. — *Soit* M *une sous-variété compacte de* $E = \mathbb{R}^n$ *de classe* C^∞. *Une fonction de Morse sur* M *est une fonction qui n'a que des points critiques non dégénérés.*

On démontre qu'il existe toujours des fonctions de Morse. Les points critiques d'une telle fonction f sont donc en nombre fini. Soit $C(f)$ l'ensemble des points critiques d'une fonction de Morse f ; désignons par $I(f, a)$ l'indice d'un point critique. On peut alors former la somme

$$\sum_{a \in C(f)} (-1)^{I(f,a)}.$$

On démontre que cette somme est un invariant qui ne dépend pas du choix de la fonction f. En fait, cet invariant ne dépend que de l'espace topologique sous-jacent à M : c'est la caractéristique d'Euler-Poincaré, notée $\chi(M)$. Ceci signifie que si M' est une autre sous-variété d'un espace vectoriel $\mathbb{R}^{n'}$ telle que M et M' soient homéomorphes, alors $\chi(M) = \chi(M')$. La démonstration d'un tel résultat utilise de façon essentielle le lemme de Morse qui permet de dire exactement, en termes de

coordonnées, le comportement de f au voisinage des points critiques : par exemple les hypersurfaces de niveau $f = c$ restent difféomorphes entre elles tant que c ne traverse pas une valeur critique. Si c traverse une valeur critique correspondant à un seul point critique a, il faut étudier comment on peut construire le fermé $\{f(x) \leq c+\epsilon\}$ à partir du fermé $\{f(x) \leq c-\epsilon\}$, et la description de f au voisinage du point critique a permet de décrire de façon très précise ce passage critique. C'est là le point essentiel de la théorie de Morse, qui dépasse le niveau de ce cours.

Voyons comment on peut dans la pratique trouver l'indice des points critiques. On peut bien sûr, si on dispose d'une carte locale φ au voisinage du point critique a, calculer l'expression g de f dans cette carte, et trouver un développement limité de g au voisinage de $b = \varphi(a)$.

Exemple

Reprenons sur la sphère S_n l'exemple de la fonction $f = x_n|_{\mathrm{S}_n}$. On prend pour carte locale au voisinage du pôle nord l'application définie par $x \mapsto (x_0, \ldots, x_{n-1})$. Alors

$$g(x_0, \ldots, x_{n-1}) = \sqrt{1 - (x_0^2 + \ldots + x_{n-1}^2)}$$

On a alors pour $v \in \mathrm{T}_a\mathrm{S}_n$

$$\mathrm{H}_0 f(v) = -(v_0^2 + \ldots + v_{n-1}^2)$$

ce qui signifie que le pôle nord est un point critique d'indice n. On voit de même que le pôle sud est un point critique d'indice 0. Ainsi, la fonction f est une fonction de Morse sur S_n. La caractéristique d'Euler-Poincaré de S_n est donc

$$\chi(\mathrm{S}_n) = 1 + (-1)^n = \begin{cases} 0 \text{ si } n \text{ est impair} \\ 2 \text{ si } n \quad \text{ est pair} \end{cases}$$

Une autre méthode peut s'imposer à partir des multiplicateurs de Lagrange, quand on ne dispose pas d'une carte locale.

PROPOSITION 11. — *Soit* U *un ouvert de* $\mathrm{E} = \mathbb{R}^n$ *et* $g = (g_1, \ldots, g_r) : \mathrm{U} \to \mathbb{R}^r$ *une submersion. On considère la sous-variété* $\mathrm{M} \subset \mathrm{U}$ *définie par les équations* $g_1 = \ldots = g_r = 0$.

Soit $\mathrm{F} : \mathrm{U} \to \mathbb{R}$ *une fonction de classe* C^ν, *et* $f = \mathrm{F}|_{\mathrm{M}}$. *Soit* $a \in \mathrm{M}$ *un point critique de* f, *et* $(\lambda_1, \ldots, \lambda_r) \in \mathbb{R}^r$ *les multiplicateurs de Lagrange correspondants.*

Alors le Hessien de f en a est donné par

$$\mathrm{H}_a f = \mathrm{H}_a(\mathrm{F} - \sum_{i=1}^{r} \lambda_i g_i)|_{\mathrm{T}_a\mathrm{M}}$$

Démonstration. Par définition des multiplicateurs de Lagrange, a est un point critique de $\mathrm{F} - \sum_{i=1}^{r} \lambda_i g_i$. Soit $\gamma :]-\epsilon, \epsilon[\to \mathrm{M}$ un chemin dérivable tracé sur M tel que $\gamma(0) = a$ et $\gamma'(0) = v$. Le développement limité de $t \mapsto (\mathrm{F} - \sum_{i=1}^{r} \lambda_i g_i)(\gamma(t))$ s'écrit à l'ordre 2

$$f(\gamma(t)) = f(a) + \frac{t^2}{2}\mathrm{H}_a(\mathrm{F} - \sum_{i=1}^{r} \lambda_i g_i)(v) + \epsilon(t)|t|^2$$

Ainsi, l'énoncé découle de l'unicité des développements limités à l'ordre 2. □

Exemple

Reprenons une nouvelle fois l'exemple de la fonction $f = x_n|_{\mathrm{S}_n}$ sur la sphère S_n. Alors le multiplicateur de Lagrange pour le pôle nord $a = (0, \ldots, 0, 1)$ est $\frac{1}{2}$. Donc

$$\mathrm{H}_a f = \mathrm{H}_a(x_n - \frac{1}{2}(x_0^2 + \ldots + x_n^2 - 1))|_{\mathrm{T}_a\mathrm{S}_n}$$

Ce calcul se fait maintenant dans $\mathbb{R}^{n+1}$. L'espace tangent $\mathrm{T}_a\mathrm{S}_n$ est défini par les vecteurs $h = (h_0, \ldots, h_n)$ tels que $h_n = 0$. Si on pose $\mathrm{F} = x_n - \frac{1}{2}(x_0^2 + \ldots + x_n^2 - 1)$, on doit alors calculer le développement limité à l'ordre 2 de $h \mapsto \mathrm{F}(a + h)$ au voisinage de 0 dans $\mathbb{R}^{n+1}$ et prendre la restriction à l'espace tangent $h_n = 0$. Ici, comme il s'agit d'un polynôme de degré 2, il n'y a pas de reste ; on a ainsi

$$\begin{aligned}\mathrm{F}(h_0, \ldots, h_{n-1}, 1) =& \frac{3}{2} - \frac{1}{2}(1 + h_0^2 + \ldots + h_{n-1}^2)\\ =& 1 - \frac{1}{2}(h_0^2 + \ldots + h_{n-1}^2)\end{aligned}$$

Par suite $\mathrm{H}_a f(h) = -(h_0^2 + \ldots + h_{n-1}^2)$.

Exercice 6.5

On considère dans $\mathbb{R}^3$ le tore T d'équation $g(x, y, z) = (\sqrt{x^2 + z^2} - 2)^2 + y^2 = 1$ et la fonction $f : \mathrm{T} \to \mathbb{R}$ définie par $f(x, y, z) = z$. Montrer que f est une fonction de Morse, et calculer la caractéristique d'Euler-Poincaré de T.