Chapitre IV. — INVERSION LOCALE, FONCTIONS IMPLICITES ET RANG CONSTANT

Comme d'habitude, tous nos espaces vectoriels sont de dimension finie. On note $(\mathrm{E}, ||.||_\mathrm{E})$ et $(\mathrm{F}, ||.||_\mathrm{F})$ des espaces vectoriels normés sur $\mathbb{R}$, U un ouvert non vide de E et $f\;:\mathrm{U}\to\mathrm{F}$ une application. Par un choix de bases pour E et F, on se ramène à $f : \mathrm{U}\,(\subset \mathbb{R}^p)\;\to \mathbb{R}^q$.

Ce chapitre est consacré à la démonstration de trois théorèmes fondamentaux: le théorème d'inversion locale, le théorème des fonctions implicites et le théorème du rang constant. On y trouvera aussi le lemme de Morse, et l'énoncé du théorème de Sard.

8 Difféomorphismes

Soient $\mathrm{U}\subset\mathrm{E}$ et $\mathrm{V}\subset\mathrm{F}$ des ouverts d'espaces vectoriels normés E et F respectivement. Un *difféomorphisme de classe* $\mathcal{C}^1$ $f : \mathrm{U}\to\mathrm{V}$ est un homéomorphisme $f : \mathrm{U}\to\mathrm{V}$ de classe $\mathcal{C}^1$ tel que son application réciproque $f^{-1} : \mathrm{V}\to\mathrm{U}$ est aussi de classe $\mathcal{C}^1$.

Si I et J sont des intervalles de $\mathbb{R}$, alors une application surjective $f : \mathrm{I}\to\mathrm{J}$ est un difféomorphisme si $f\in\mathcal{C}^1$ et si f' ne s'annule pas. Dans ce cas, f est strictement monotone, donc bijective sur son image qui est J. De plus, l'application inverse $g = f^{-1} : \mathrm{J}\to\mathrm{I}$ est dérivable de dérivée $g('x) = 1/f'(g(x))$ qui est bien continue.

Proposition 8.1. — *On a les propriétés suivantes.*

1. *Une application $f : \mathrm{U}\to\mathrm{V}$ est un difféomorphisme si et seulement si*
   - a) *f est de classe $\mathcal{C}^1$;*
   - b) *f est un homéomorphisme;*
   - c) *pour tout $x\in\mathrm{U}$, $\mathrm{D}_x f$ est un isomorphisme.*
2. *Soient $f : \mathrm{U}\to\mathrm{V}$ un difféomorphisme et $u : \mathrm{V}\to\mathrm{G}$ une application à valeurs dans un espace vectoriel normé G. Alors u est de classe $\mathcal{C}^1$ si et seulement si $u\circ f$ est de classe $\mathcal{C}^1$.*

Démonstration. — Supposons que f est un difféomorphisme de classe $\mathcal{C}^1$. Par définition, f est de classe $\mathcal{C}^1$ et un homéomorphisme. On écrit $g = f^{-1}$ de sorte que $g\circ f = \mathrm{Id}$. Du coup, pour tout $x\in\mathrm{U}$, $\mathrm{D}_{f(x)}g\circ\mathrm{D}_x f = \mathrm{Id}$, impliquant ainsi que $\mathrm{D}_x f$ est inversible. Réciproquement, si f est un homéomorphisme de classe $\mathcal{C}^1$ de sorte que $\mathrm{D}_x f$ est un isomorphisme pour tout $x\in\mathrm{U}$, alors il reste à montrer que $g = f^{-1}$ est de classe $\mathcal{C}^1$. La proposition 3.8 implique que g est différentiable en chaque point et $\mathrm{D}_y g = (\mathrm{D}_{g(y)}f)^{-1}$ pour tout $y\in\mathrm{V}$. Du coup, $\mathrm{D}g$ est la composée de $\varphi : u\in\mathrm{GL}(\mathrm{E},\mathrm{F})\mapsto u^{-1}\in\mathrm{GL}(\mathrm{F},\mathrm{E})$, $x\mapsto\mathrm{D}_x f$ et g qui sont toutes continues par hypothèses (on a $\mathrm{D}g = \varphi\circ\mathrm{D}f\circ g$). Donc $\mathrm{D}g$ est continue et on conclut que g est de classe $\mathcal{C}^1$.

Comme on peut écrire $u = (u\circ f)\circ f^{-1}$, on déduit facilement que u est de classe $\mathcal{C}^1$ si et seulement si $u\circ f$ est de classe $\mathcal{C}^1$ par stabilité de la propriété d'être $\mathcal{C}^1$. ∎

On remarque qu'être un homéomorphisme de classe $\mathcal{C}^1$ ne suffit pas à être un difféomorphisme. Prenons par exemple l'application $x\mapsto x^3$ sur $\mathbb{R}$ qui est bien un homéomorphisme, mais son application réciproque n'est pas différentiable en zéro. Cependant, il n'y a pas de problèmes pour passer à la régularité supérieure:

Proposition 8.2. — *Si un difféomorphisme $f : \mathrm{U}\to\mathrm{V}$ est de classe $\mathcal{C}^k$, $k\geq 2$, alors f^{-1} est aussi de classe $\mathcal{C}^k$.*

Démonstration. — On écrit $g = f^{-1}$ de sorte que $f\circ g = \mathrm{Id}$. Du coup, $\mathrm{D}_{g(y)}f\circ\mathrm{D}_y g = \mathrm{Id}$, impliquant ainsi $\mathrm{D}_y g = (\mathrm{D}_{g(y)}f)^{-1}$. En reprenant les notations de la démonstration précédente, on a donc $\mathrm{D}_y g = \varphi\circ\mathrm{D}f\circ g$.

Montrons par récurrence sur $k\geq 1$ que si f est de classe $\mathcal{C}^k$ alors g aussi. Pour $k = 1$, cela découle de la définition de difféomorphisme. Si c'est vrai au rang k et que f est de classe $\mathcal{C}^{k+1}$, alors $\mathrm{D}f$ est de classe $\mathcal{C}^k$ et g de classe $\mathcal{C}^k$. Or comme φ est de classe $\mathcal{C}^\infty$ (voir l'exemple 6.11), on en déduit que $\mathrm{D}g$ est de classe $\mathcal{C}^k$, donc g est de classe $\mathcal{C}^{k+1}$. ∎

On montre comment, dans la proposition 3.8, se passer de l'hypothèse d'être un homéomorphisme si on part d'une application de classe $\mathcal{C}^1$, comme en dimension 1.

Théorème 8.3 (d'inversion locale). — *Soit $f : \mathrm{U} \to \mathrm{F}$ une application de classe $\mathcal{C}^1$. S'il existe $a \in \mathrm{U}$ tel que $\mathrm{D}_a f$ est inversible, alors il existe des voisinages V de a et W de $f(a)$ tels que $f : \mathrm{V} \to \mathrm{W}$ est un difféomorphisme.*

Remarque 8.4. — *Les hypothèses du théorème sont équivalentes à f de classe $\mathcal{C}^1$ et jacobien au point a non nul i.e,* $\det \mathrm{D}_a f \neq 0$.

Une application $f : \mathrm{U} \to \mathrm{F}$ est un *difféomorphisme local* si, pour tout $a \in \mathrm{U}$, il existe un voisinage V de a tel que $f : \mathrm{V} \to f(\mathrm{V})$ est un difféomorphisme. Le théorème d'inversion locale donne une condition nécessaire (et suffisante) pour qu'une application de classe $\mathcal{C}^1$ soit un difféomorphisme local.

Corollaire 8.5. — *Une application $f : \mathrm{U} \to \mathrm{F}$ de classe $\mathcal{C}^1$ est un difféomorphisme local si et seulement si, pour tout $x \in \mathrm{U}$, on a* $\det \mathrm{D}_x f \neq 0$.

Méthode de Newton.— La démonstration s'appuie sur une généralisation de la méthode de Newton. Rappelons le principe de cette méthode. On se donne une fonction numérique $u : \mathrm{I} \to \mathbb{R}$ de classe $\mathcal{C}^1$ dont on cherche un zéro i.e., un point $t \in \mathrm{I}$ tel que $u(t) = 0$. On part d'une condition initiale x_0 et on remplace le graphe de u par la tangente au point x_0 à laquelle on trouve le zéro correpondant que l'on note x_1. On recommence avec la condition initiale x_1 et on construit une suite (x_n) en espérant que celle-ci soit convergente. On montre alors que la limite, si elle existe, est le zéro recherché.

Si x_n est construit, alors l'équation de la tangente est $y - u(x_n) = u'(x_n)(x - x_n)$. Cette droite coupe l'axe des abscisses au point

$$x_{n+1} = x_n - \frac{u(x_n)}{u'(x_n)} .$$

Ce point est bien défini si $u'(x_n) \neq 0$, c'est-à-dire si sa tangente n'est pas horizontale!

Pour montrer que la suite est bien convergente, on peut essayer d'appliquer le théorème du point fixe suivant à la fonction $\mathrm{N}_u(x) = x - u(x)/u'(x)$:

Théorème 8.6 (du point fixe). — *Soit (X, d) un espace métrique complet et considérons une transformation $f : \mathrm{X} \to \mathrm{X}$. Si f est contractante i.e., il existe $0 < k < 1$ telle que, pour $x, y \in \mathrm{X}$, on a $d(f(x), f(y)) \leq kd(x, y)$. Alors f admet un unique point fixe p i.e., vérifiant $f(p) = p$. De plus, le point p s'obtient comme suit. Pour tout choix initial $x_0 \in \mathrm{X}$, la suite (x_n) définie par récurrence en posant $x_{n+1} = f(x_n)$ est convergente de limite p.*

Démonstration. — On montre par récurrence que $d(x_n, x_{n+1}) \leq k^n d(x_0, x_1)$. En effet, c'est vrai rang $n = 0$ et si c'est vrai au rang n, alors

$$d(x_{n+1}, x_{n+2}) = d(f(x_n), f(x_{n+1})) \leq kd(x_n, x_{n+1}) \leq k^{n+1} d(x_0, x_1) .$$

Par conséquent, pour tout $n, p \geq 0$, on a

$$d(x_n, x_{n+p}) \leq \sum_{j=0}^{p-1} d(x_{n+j}, x_{n+j+1}) \leq \sum_{j=0}^{p-1} k^{n+j} d(x_0, x_1) \leq \frac{k^n}{1-k} d(x_0, x_1)$$

donc (x_n) est une suite de Cauchy. Comme on a supposé X complet, cette suite est convergente; on note sa limite p. Comme f est continue, l'identité $x_{n+1} = f(x_n)$ donne à la limite $p = f(p)$. L'unicité du point fixe découle de la contraction de f car si q était un autre point fixe, on aurait $d(p, q) = d(f(p), f(q)) \leq kd(p, q) < d(p, q)$ qui est impossible. ■

Retournons à la méthode de Newton. Si $\mathrm{N}_u(p) = p$, cela implique

$$p = p - \frac{u(p)}{u'(p)}$$

soit $u(p) = 0$. Donc un point fixe pour la méthode de Newton fournit bien un zéro de u.

Démonstration du théorème d'inversion locale.— Nous allons mettre en œuvre la stratégie ci-dessus. Pour simplifier les notations, on se ramène au cas où $\mathrm{E} = \mathrm{F}$, $a = f(a) = 0$ et $\mathrm{D}_a f = \mathrm{Id}_\mathrm{E}$. En effet, f est un difféomorphisme local au point a si et seulement si l'application: $u \mapsto (\mathrm{D}_a f)^{-1}[f(a+u) - f(a)]$ est un difféomorphisme local au point 0. C'est-à-dire que moyennant une translation sur a et sur $f(a)$ et un isomorphisme linéaire, on peut se ramener à cette situation simplifiée.

Pour $y \in \mathrm{E}$, on considère la méthode de Newton $\mathrm{N}_y : x \mapsto x - (f(x) - y)$ de sorte qu'un point fixe de N_y nous donne un antécédent de y par f. Nous allons montrer qu'il existe $\varepsilon > 0$ assez petit de sorte que

1. $\mathrm{B}(0, \varepsilon) \subset \mathrm{U}$;

2. si $y \in \mathrm{B}(0, \varepsilon/2)$, alors $\mathrm{N}_y(\mathrm{B}(0, \varepsilon)) \subset \mathrm{B}(0, \varepsilon)$;

3. si $y \in \mathrm{B}(0, \varepsilon/2)$, alors N_y est contractante.

Remarquons que N_y est différentiable sur U et $\mathrm{D}_x\mathrm{N}_y = \mathrm{Id} - \mathrm{D}_x f$. Puisque $\mathrm{D}f$ est continue et $\mathrm{D}_0 f = \mathrm{Id}$, il existe $\varepsilon > 0$ tel que $\mathrm{B}(0, \varepsilon) \subset \mathrm{U}$ et si $||x|| \leq \varepsilon$ alors $||\mathrm{D}_x f - \mathrm{Id}|| \leq 1/2$. D'après l'inégalité des accroissements finis, on a, pour tous x_1, x_2 dans la boule $\mathrm{B}(0, \varepsilon)$,

$$\|\mathrm{N}_y(x_1) - \mathrm{N}_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|\,. \tag{8.1}$$

Ceci montre en particulier que N_y est contractante. En remarquant que $\mathrm{N}_y(0) = y$, et en choisissant $y \in \mathrm{B}(0, \varepsilon/2)$, alors, pour $x \in \mathrm{B}(0, \varepsilon)$, on a

$$\|\mathrm{N}_y(x)\| \leq \|\mathrm{N}_y(x) - \mathrm{N}_y(0)\| + \|y\| \leq \frac{1}{2}\|x\| + \|y\| < \varepsilon$$

donc on obtient $\mathrm{N}_y(\mathrm{B}(0, \varepsilon)) \subset \mathrm{B}(0, \varepsilon)$ si $\|y\| < \varepsilon/2$. D'après le théorème du point fixe appliqué dans $\overline{\mathrm{B}(0, \varepsilon)}$, N_y admet un unique point fixe $g(y)$ dans $\overline{\mathrm{B}(0, \varepsilon)}$. Par conséquent, l'équation $f(x) = y$ admet une unique solution, $g(y)$, dans $\overline{\mathrm{B}(0, \varepsilon)}$. On pose $\mathrm{V} = \mathrm{B}(0, \varepsilon) \cap f^{-1}(\mathrm{B}(0, \varepsilon/2))$ qui est un ouvert non vide puisque $0 \in \mathrm{V}$; on note aussi $\mathrm{W} = f(\mathrm{V})$.

La restriction $f : \mathrm{V} \to \mathrm{W}$ est donc une bijection de classe $\mathcal{C}^1$ d'inverse g. De plus, $\mathrm{D}_x f$ est inversible sur V car $\mathrm{D}_x f = \mathrm{Id} - (\mathrm{Id} - \mathrm{D}_x f)$ et $\|\mathrm{Id} - \mathrm{D}_x f\| \leq 1/2 < 1$, cf. § 3.1 **5.**. Il reste à établir la continuité de g. D'après (8.1) et l'inégalité triangulaire inverse, on a, pour $y_1, y_2 \in \mathrm{W}$ et $x_1 = g(y_1)$, $x_2 = g(y_2)$,

$$\begin{aligned} \frac{1}{2}\|x_1 - x_2\| &\geq \|\mathrm{N}_0(x_1) - \mathrm{N}_0(x_2)\| \\ &= \|(x_1 - x_2) - (y_1 - y_2)\| \\ &\geq \|x_1 - x_2\| - \|y_1 - y_2\| \end{aligned}$$

ce qui implique $\|x_1 - x_2\| \leq 2\|y_1 - y_2\|$. Autrement dit, g est 2-lipschitzienne, donc continue. Le critère de la proposition 8.1 s'applique pour conclure que $f : \mathrm{V} \to \mathrm{W}$ est un difféomorphisme. ■

Corollaire 8.7 (inversion globale). — *Soient* E, F *des espaces vectoriels normés de dimension finie et soit* U *un ouvert de* E. *Si* $f : \mathrm{U} \to f(\mathrm{U})$ *est de classe* $\mathcal{C}^1$*, injective, et si* $\mathrm{D}_x f \in \mathrm{GL}(\mathrm{E}, \mathrm{F})$ *pour tout* $x \in \mathrm{U}$*, alors* f *est un difféomorphisme de classe* $\mathcal{C}^1$ *de* U *sur* $f(\mathrm{U})$.

Démonstration. — D'après le théorème d'inversion locale, f est un difféomorphisme local. Mais comme f est globalement injective, on en déduit que f est un difféomorphisme global sur son image. ■

Corollaire 8.8 (application ouverte). — *Soient* E, F *des espaces vectoriels normés de dimension finie et soit* U *un ouvert de* E. *Si* $f : \mathrm{U} \to \mathrm{F}$ *est de classe* $\mathcal{C}^1$ *et si* $\mathrm{D}_x f \in \mathrm{GL}(\mathrm{E}, \mathrm{F})$ *pour tout* $x \in \mathrm{U}$*, alors* f *est une application ouverte: pour tout ouvert* $\mathrm{V} \subset \mathrm{U}$, $f(\mathrm{V})$ *est ouvert.*

Démonstration. — Soit $\mathrm{V} \subset \mathrm{U}$ un ouvert et prenons $a \in \mathrm{V}$. Nous allons montrer que $f(a)$ admet un voisinage contenu dans $f(\mathrm{V})$. Cela découle du théorème d'inversion locale qui implique l'existence d'un voisinage ouvert W de a tel que $f : \mathrm{W} \to f(\mathrm{W})$ est un difféomorphisme; notons $g : f(\mathrm{W}) \to \mathrm{W}$ son inverse qui est donc continue. Du coup, $\mathrm{V} \cap \mathrm{W}$ est un ouvert non vide, et $f(\mathrm{V} \cap \mathrm{W}) = g^{-1}(\mathrm{V} \cap \mathrm{W})$ est ouvert par la continuité de g. ■

On conclut par un exemple important.

Coordonnées polaires dans le plan. — On considère l'application $f :]0, \infty[\times\mathbb{R} \to \mathbb{R}^2$ définie par $f(r, \theta) = (r\cos\theta, r\sin\theta)$. Cette application est de classe $\mathcal{C}^\infty$ et

$$\mathrm{D}f = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

donc $\det \mathrm{D}f = r\cos^2\theta + r\sin^2\theta = r \neq 0$ donc f est un difféomorphisme local de classe $\mathcal{C}^\infty$.

9 Fonctions implicites

On s'intéresse dans ce paragraphe aux solutions d'équations à paramètres. L'objet du théorème des fonctions implicites est de montrer comment de la connaissance d'une solution pour un paramètre fixé obtenir des solutions pour des paramètres voisins.

Théorème 9.1 (des fonctions implicites). — *Soient* $\mathrm{E}, \mathrm{F}, \mathrm{G}$ *des evn de dimension finie,* U *un ouvert de* $\mathrm{E}\times\mathrm{F}$ *et* $f : \mathrm{U} \to \mathrm{G}$ *de classe* $\mathcal{C}^k$*,* $k \geq 1$*. On suppose qu'en un point* $(x_0, y_0) \in \mathrm{U}$*, la différentielle partielle* $\partial_{\mathrm{F}} f(x_0, y_0)$ *est un isomorphisme.*

Alors il existe des ouverts V *et* W *de* E *et* F *respectivement et une application* $g : \mathrm{V} \to \mathrm{W}$ *de classe* $\mathcal{C}^k$ *tels que*

1. *on a* $\mathrm{V}\times\mathrm{W} \subset \mathrm{U}$*;*
2. *on a* $g(x_0) = y_0$ *et* $\mathrm{D}_{x_0}g = -(\partial_{\mathrm{F}} f(x_0, y_0))^{-1} \circ \partial_{\mathrm{E}} f(x_0, y_0)$*;*
3. *pour tout* $x \in \mathrm{V}$*, l'équation* $f(x, y) = f(x_0, y_0)$ *en* y *admet une unique solution* $g(x)$ *dans* W*; autrement dit, pour tout* $x \in \mathrm{V}$*, on a* $f(x, g(x)) = f(x_0, y_0)$ *et si* $f(x, y) = f(x_0, y_0)$ *avec* $y \in \mathrm{W}$ *alors* $y = g(x)$*.*

On se donne donc une application $f : \mathrm{E}\times\mathrm{F} \to \mathrm{G}$ et on s'intéresse à l'équation $f(x, y) = z$ où $z \in \mathrm{G}$ est fixé et (x, y) sont des inconnues. On cherche à résoudre cette équation de manière à trouver y en fonction de x. Graphiquement, on s'intéresse à l'ensemble de niveau $\mathrm{E}_z = \{(x, y) \in \mathrm{E}\times\mathrm{F},\ f(x, y) = z\}$ et on veut déterminer si, au voisinage d'un point $(x_0, y_0) \in \mathrm{E}_z$, on peut exprimer E_z comme le graphe d'une fonction $y = g(x)$. Le théorème répond par l'affirmative si le plan tangent de E_z dans $\mathrm{E}\times\mathrm{F}$ n'est pas parallèle à F. Par exemple, si on considère $f(x, y) = x^2 + y^2 - 1 = 0$, alors $\partial_y f = 2y$ qui est non nul si $y \neq 0$. Dans ce cas, on peut exprimer le cercle unité comme le graphe de $x \mapsto \pm\sqrt{1 - x^2}$, $x \in]0, 1[$.

Démonstration. — Posons $\varphi : \mathrm{U} \to \mathrm{E}\times\mathrm{G}$ défini par $\varphi(x, y) = (x, f(x, y))$. Cette application est de classe $\mathcal{C}^k$ et

$$\mathrm{D}_{(x_0,y_0)}\varphi = \begin{pmatrix} \mathrm{Id}_{\mathrm{E}} & 0 \\ \partial_{\mathrm{E}} f(x_0, y_0) & \partial_{\mathrm{F}} f(x_0, y_0) \end{pmatrix}.$$

Comme $\partial_{\mathrm{F}} f(x_0, y_0)$ est inversible, $\det \mathrm{D}_{(x_0,y_0)}\varphi \neq 0$, donc φ est un difféomorphisme au voisinage du point (x_0, y_0) d'après le théorème d'inversion locale. Il existe donc des ouverts $\mathrm{V} \subset \mathrm{E}$, $\mathrm{W} \subset \mathrm{F}$ et $\mathrm{Y} \subset \mathrm{G}$ tels que $\varphi : \mathrm{V}\times\mathrm{W} \to \mathrm{V}\times\mathrm{Y}$ est un difféomorphisme de classe $\mathcal{C}^k$.

L'égalité $f(x, y) = f(x_0, y_0)$ est équivalente au fait que $\varphi(x, y) = (x, f(x_0, y_0))$, donc $\varphi^{-1}(x, f(x_0, y_0)) = (x, y)$. Notons $\pi_{\mathrm{F}} : \mathrm{E}\times\mathrm{F} \to \mathrm{F}$ la projection canonique et $g : \mathrm{V} \to \mathrm{W}$ définie par $g(x) = \pi_{\mathrm{F}} \circ \varphi^{-1}(x, f(x_0, y_0))$, de sorte que $(x, g(x)) = \varphi^{-1}(x, f(x_0, y_0))$. On vérifie que $g(x_0) = y_0$, $g(\mathrm{V}) \subset \mathrm{W}$, g est de classe $\mathcal{C}^k$ comme composée de fonctions de classe au moins $\mathcal{C}^k$ et $f(x, g(x)) = f(x_0, y_0)$.

Enfin en différentiant $x \mapsto f(x, g(x))$, on obtient $\partial_{\mathrm{E}} f + \partial_{\mathrm{F}} f \circ \mathrm{D}g = 0$, d'où l'expression de $\mathrm{D}g$ au point x_0. ■

Interprétation analytique.— On écrit $f(x, y) = f_x(y)$ pour voir x comme un paramètre. Supposons que l'on connaisse une solution pour le paramètre x_0 de l'équation $f_{x_0}(y) = 0$. On étudie le problème linéarisé induit par $\partial_{\mathrm{F}} f_{x_0}(y)$: si celui-ci peut être résolu au sens que c'est un isomorphisme, alors l'équation perturbée admet une unique solution au voisinage de la solution y_0, qui dépend de manière $\mathcal{C}^1$ de la perturbation.

Exemple: racines simples de polynômes.— On prend $\mathrm{E} = \mathbb{C}_d[\mathrm{X}]$, $\lambda \in \mathbb{C}$ et $\mathrm{P}_0 \in \mathrm{E}$ tels que $\mathrm{P}_0(\lambda) = 0$. On suppose que λ est une racine simple impliquant ainsi $\mathrm{P}_0'(\lambda) \neq 0$. On étudie l'équation $\mathrm{P}(z) = 0$ en posant $f(\mathrm{P}, z) = \mathrm{P}(z)$. On a $\partial_{\mathbb{C}} f(\mathrm{P}_0, \lambda) = \mathrm{P}_0'(\lambda) \neq 0$, donc est inversible. Du coup, il existe un voisinage de U de P_0 et V de λ et une transformation $g : \mathrm{U} \to \mathrm{V}$ de classe $\mathcal{C}^\infty$ tels que, pour tout polynôme Q assez proche de P_0, $g(\mathrm{Q})$ est l'unique racine de Q dans V, cette racine étant aussi simple.

Equivalence entre les théorèmes d'inversion locale et des fonctions implicites.— On a déduit le théorème des fonctions implicites du théorème d'invesion locale, mais on aurait pu procéder à l'inverse. En effet, calculer l'inverse d'une application $\varphi : \mathrm{U} \to \mathrm{F}$ revient à étudier l'équation $\varphi(x) - y = 0$. En posant $f(x, y) = \varphi(x) - y$, on obtient $\partial_{\mathrm{E}} f(x, y) = \mathrm{D}_x\varphi$. Donc si $\mathrm{D}_{x_0}\varphi$ est inversible, alors on trouve des voisinages V de x_0 et W de $\varphi(x_0)$ et une (unique) application $\psi : \mathrm{W} \to \mathrm{V}$ tels que $\varphi(\psi(y)) = y$ pour tout $y \in \mathrm{W}$. Ceci montre que φ est un difféomorphisme local au point x_0.

10 Structure locale des applications différentiables

Dans ce paragraphe, on étudie la forme d'une application de classe $\mathcal{C}^k$ à difféomorphisme près. On étudie plusieurs cas de figure selon la régularité de f au voisinage d'un point a donné.

10.1 Rang constant

Si E et F sont des espaces vectoriels normés, le rang d'une application linéaire $\mathrm{L} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ est la dimension de $\mathrm{Im}\,\mathrm{L}$. Pour une application $\varphi : \mathrm{U} \to \mathrm{F}$ différentiable au point $a \in \mathrm{U}$, on définit

$$\mathrm{rg}_a\,\varphi \stackrel{\text{def.}}{=} \mathrm{rg}\,\mathrm{D}_a\varphi\,.$$

On rappelle le lemme suivant d'algèbre linéaire.

Lemme 10.1. — *Soient $p, q \geq 1$ et $r \geq 0$. Si une matrice $\mathrm{A} \in \mathcal{M}_{q,p}(\mathbb{R})$ est de rang r alors il existe $\mathrm{P} \in \mathrm{GL}_p(\mathbb{R})$ et $\mathrm{Q} \in \mathrm{GL}_q(\mathbb{R})$ telles que*

$$\mathrm{QAP}^{-1} = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Démonstration. — On considère l'application linéaire $u_{\mathrm{A}} : \mathbb{R}^p \to \mathbb{R}^q$ définie par $u_{\mathrm{A}}(\mathrm{X}) = \mathrm{AX}$. Soit $(\mathrm{e}_1, \ldots \mathrm{e}_n)$ une base de $\mathrm{Ker}\, u_{\mathrm{A}}$ que l'on complète en une base $(\mathrm{e}_1, \ldots, \mathrm{e}_n, \varepsilon_1, \ldots, \varepsilon_m)$ de $\mathbb{R}^p$.

Montrons que $(u_{\mathrm{A}}(\varepsilon_1), \ldots, u_{\mathrm{A}}(\varepsilon_m))$ est une base de $\mathrm{Im}\, u_{\mathrm{A}}$. Tout d'abord, cette famille est génératrice par définition: si $y \in \mathrm{Im}\, u_{\mathrm{A}}$, alors il existe $x \in \mathbb{R}^p$ tel que $u_{\mathrm{A}}(x) = y$. Or, on peut trouver des scalaires $x_1, \ldots, x_n$ et $y_1, \ldots, y_m$ tels que

$$x = \sum_{k=1}^{n} x_k \mathrm{e}_k + \sum_{j=1}^{m} y_j \varepsilon_j\,.$$

Du coup, comme $\sum x_k \mathrm{e}_k \in \mathrm{Ker}\, u_{\mathrm{A}}$, on obtient

$$y = u_{\mathrm{A}}(x) = \sum_{j=1}^{m} y_j u_{\mathrm{A}}(\varepsilon_j)\,.$$

Il reste à vérifier que cette famille est aussi libre. Pour cela, on choisit $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ tels que $\sum \lambda_j u_{\mathrm{A}}(\varepsilon_j) = 0$. Cela s'écrit aussi $u_{\mathrm{A}}(\sum \lambda_j \varepsilon_j) = 0$ ce qui implique que $\sum \lambda_j \varepsilon_j \in \mathrm{Ker}\, u_{\mathrm{A}}$. Comme $(\mathrm{e}_1, \ldots, \mathrm{e}_n)$ est une base de $\mathrm{Ker}\, u_{\mathrm{A}}$ et $(\mathrm{e}_1, \ldots, \mathrm{e}_n, \varepsilon_1, \ldots, \varepsilon_m)$ est libre, on en déduit que $\sum \lambda_j \varepsilon_j = 0$ donc $\lambda_1 = \ldots = \lambda_m = 0$ et cette famille est bien une base de $\mathrm{Im}\, u_{\mathrm{A}}$.

Comme $\mathrm{rg}\,\mathrm{A} = r$, on en déduit que $m = r$ et $p = n + r$. Enfin, on note $f_j = u_{\mathrm{A}}(\varepsilon_j)$ et on complète $(f_1, \ldots, f_r)$ en une base $\mathcal{F} = (f_1, \ldots, f_q)$ de $\mathbb{R}^q$. On réordonne la base de $\mathbb{R}^p$ pour l'écrire sous la forme $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_r, \mathrm{e}_1, \ldots, \mathrm{e}_n)$. La matrice de u_{A} dans ces bases est donc

$$\mathrm{Mat}(u_{\mathrm{A}}, \mathcal{E}, \mathcal{F}) = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

En notant P la matrice de passage de la base canonique de $\mathbb{R}^p$ vers $\mathcal{E}$ et Q la matrice de passage de la base canonique de $\mathbb{R}^q$ vers $\mathcal{F}$, on conclut la démonstration du lemme. ■

On montre une version « à géométrie variable » de ce résultat.

Théorème 10.2 (du rang constant). — *Soit $f : \mathrm{U} \to \mathrm{F}$ une application de classe $\mathcal{C}^k$, $k \geq 1$, définie sur U et supposons qu'il existe $r \geq 1$ tel que $\mathrm{rg}_x f = r$ pour tout x dans un voisinage d'un point $a \in \mathrm{U}$. Alors il existe un voisinage $\mathrm{V} \subset \mathrm{U}$ de a, un voisinage $\mathrm{W} \subset \mathrm{F}$ de $f(a)$, des difféomorphismes $\varphi : \mathrm{V} \to \mathrm{V}' (\subset \mathbb{R}^p)$ et $\psi : \mathrm{W} \to \mathrm{W}' (\subset \mathbb{R}^q)$ tels que $\psi \circ f \circ \varphi^{-1} = \mathrm{D}_a f : \mathrm{V}' \to \mathrm{W}'$. Dans des bases appropriées, on obtient*

$$(x_1, \ldots, x_p) \mapsto (x_1, \ldots, x_r, \underbrace{0, \ldots, 0}_{q-r})$$

Ce théorème affirme qu'une application de classe au moins $\mathcal{C}^1$ de rang constant est localement linéaire à changement de coordonnées près par des difféomorphismes. Autrement dit, une application de classe au moins $\mathcal{C}^1$ de rang constant n'est pas plus compliquée qu'une application linéaire.

Donnons les grandes étapes de la démonstration. Dans un premier temps, on construit φ. Pour cela, on part du fait que la fonction f se comporte en première approximation comme l'application affine $x \mapsto f(a) + \mathrm{D}_a f$, donc envoie U proche du sous-espace $f(a) + \mathrm{Im}\,\mathrm{D}_a f$. Cela nous conduit à considérer une projection $\pi : \mathrm{F} \to \mathrm{Im}\,\mathrm{D}_a f$ et d'étudier dans un premier temps l'application $\pi \circ f : \mathrm{U} \to \mathrm{Im}\,\mathrm{D}_a f$. On se place dans des repères adaptés de E et F qui tiennent compte de $\mathrm{D}_a f$ c'est-à-dire pour que $\mathrm{D}_a f$ soit la matrice donnée par le lemme 10.1. On décompose donc E et F sous la forme $\mathrm{E} = \mathrm{E}' \oplus \mathrm{Ker}\,\mathrm{D}_a f$ et $\mathrm{F} = \mathrm{Im}\,\mathrm{D}_a f \oplus \mathrm{Ker}\,\pi$. Dans cette situation, on montre l'existence d'un difféomorphisme local φ tel que $\pi \circ f \circ \varphi^{-1} = \mathrm{D}_a f$ est linéaire.

L'étape suivante consiste à montrer que $f \circ \varphi^{-1}$ est de la forme $(x, y) \mapsto (x, g(x))$ au voisinage de 0; on utilise ici que le rang est localement constant. De là, on construit ψ au voisinage de $f(a)$ en posant $\psi(x, z) = (x, z - g(x))$.

Démonstration. — Quitte à remplacer f par $x \mapsto (f(a+x) - f(a))$, on peut supposer que $a = f(a) = 0$. Nous donnons deux présentations en parallèle de la preuve: l'une intrinsèque et l'autre en coordonnées.

On choisit une base $\mathcal{B}_\mathrm{E} = (\mathrm{e}_1, \ldots, \mathrm{e}_p)$ de E telle que $(\mathrm{e}_{r+1}, \ldots, \mathrm{e}_p)$ soit une base de $\mathrm{Ker}\,\mathrm{D}_0 f$; notons E' l'espace engendré par $(\mathrm{e}_1, \ldots, \mathrm{e}_r)$. Du coup, si on pose $f_j = \mathrm{D}_0 f(\mathrm{e}_j)$ pour $j = 1, \ldots, r$, alors $(f_1, \ldots, f_r)$ est une famille libre que l'on peut compléter en une base de $\mathcal{B}_\mathrm{F} = (f_1, \ldots, f_q)$ de F. Cela nous permet d'identifier E avec $\mathbb{R}^p$ et F avec $\mathbb{R}^q$. On travaille maintenant dans ces repères.

Dans cette base, on a l'écriture en matrice par blocs

$$\mathrm{D}_0 f = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}$$

Du coup, la restriction $\mathrm{D}_0 f|_{\mathrm{E}'} : \mathrm{E}' \to \mathrm{Im}\,\mathrm{D}_0 f$ est un isomorphisme dont on notera l'inverse

$$(\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} : \mathrm{Im}\,\mathrm{D}_0 f \to \mathrm{E}'.$$

Notez que $\mathrm{D}_0 f \circ (\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} = \mathrm{Id}$ et $(\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \mathrm{D}_0 f = \pi_1$ où $\pi_1 : \mathrm{E} \to \mathrm{E}$ est la projection de E sur E' parallèlement à $\mathrm{Ker}\,\mathrm{D}_0 f$.

Un développement limité de f montre que f envoie un voisinage de 0 dans $\mathbb{R}^p$ très proche de $\mathrm{Im}\,\mathrm{D}_0 f = \mathbb{R}^r \times \{0\}_{q-r} \subset \mathrm{F}$. On considère la projection $\pi : \mathrm{F} \to \mathrm{F}$ sur $\mathrm{Im}\,\mathrm{D}_0 f$ parallèlement à $\mathrm{F}' = \mathrm{vec}(f_{r+1}, \ldots, f_q) = \{0\} \times \mathbb{R}^{q-r}$ définie par $\pi(x, z) = x$, pour $(x, z) \in \mathbb{R}^r \times \mathbb{R}^{q-r}$.

On définit $\varphi : \mathrm{E} \to \mathrm{E}$ par $\varphi = ((\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi \circ f) + \pi_2$, où $\pi_2 : \mathrm{E} \to \mathrm{E}$ désigne la projection de E sur $\mathrm{Ker}\,\mathrm{D}_0 f$ parallèlement à E'. L'application φ s'écrit, dans les repères choisis, $\varphi : \mathbb{R}^r \times \mathbb{R}^{p-r} \to \mathbb{R}^r \times \mathbb{R}^{p-r}$ en posant $\varphi(x, y) = (((\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi \circ f)(x, y), y)$. C'est une application de classe $\mathcal{C}^k$. On a

$$\mathrm{D}_0 \varphi = ((\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi \circ \mathrm{D}_0 f) + \pi_2 = \pi_1 + \pi_2 = \mathrm{Id}$$

ou, plus prosaïquement, on calcule la matrice jacobienne. On obtient la matrice par blocs

$$\mathrm{D}_0 \varphi = \begin{pmatrix} \pi \circ \partial_1 f(0) & \pi \circ \partial_2 f(0) \\ 0 & \mathrm{I}_{p-r} \end{pmatrix} = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & \mathrm{I}_{p-r} \end{pmatrix}$$

car $\pi = \mathrm{Id}$ sur $\mathbb{R}^r \times \{0\}_{q-r}$ et $\{0\}_r \times \mathbb{R}^{q-r} = \mathrm{Ker}\,\pi$.

Par conséquent $\mathrm{D}_0\varphi$ est inversible et le théorème d'inversion locale affirme que φ est un difféomorphisme $\mathcal{C}^k$ au voisinage de 0; notons V, V′ ces voisinages de sorte que $\varphi : \mathrm{V} \to \mathrm{V}'$; on choisit V′ convexe. On a

$$\pi \circ (f \circ \varphi^{-1}) = \mathrm{D}_0 f - (\mathrm{D}_0 f \circ \pi_2 \circ \varphi^{-1}) = \mathrm{D}_0 f$$

car $\mathrm{Im}\,\pi_2 \subset \mathrm{Ker}\,\mathrm{D}_0 f$. Autrement dit, on a $(\pi \circ f \circ \varphi^{-1})(x,y) = x$ par construction.

Or $\mathrm{D}_0 f = \pi \circ \mathrm{D}_0 f$ donc $\pi[(f \circ \varphi^{-1}) - \mathrm{D}_0 f] = 0$ car $\pi = \mathrm{Id}$ sur $\mathrm{Im}\,\mathrm{D}_0 f$; du coup, on a $((f \circ \varphi^{-1}) - \mathrm{D}_0 f)(\mathrm{V}') \subset \mathrm{Ker}\,\pi$. Par conséquent, il existe une application $g : \mathrm{V}' \to \mathrm{Ker}\,\pi = \mathrm{F}'$ de classe $\mathcal{C}^k$ telle que $(f \circ \varphi^{-1}) = \mathrm{D}_0 f + g$, soit $g : \mathbb{R}^p \to \mathbb{R}^{q-r}$ telle que $(f \circ \varphi^{-1})(x,y) = (x, g(x,y))$.

Nous allons montrer qu'en fait g ne dépend que de $x \in \mathrm{E}' = \mathbb{R}^r$. C'est là que l'hypothèse de rang constant va jouer. On a, pour $x \in \mathrm{V}'$,

$$r = \mathrm{rg}_x(f \circ \varphi^{-1}) \geq \dim \mathrm{D}_x(f \circ \varphi^{-1})(\mathrm{E}') \geq \dim \pi \circ \mathrm{D}_x(f \circ \varphi^{-1})(\mathrm{E}') = \dim \mathrm{D}_0 f(\mathrm{E}') = r$$

donc $\dim \mathrm{D}_x(f\circ\varphi^{-1})(\mathrm{E}') = r$ et on déduit d'une part que $\dim \mathrm{D}_x(f\circ\varphi^{-1})(\mathrm{Ker}\,\mathrm{D}_0 f) \subset \dim \mathrm{D}_x(f\circ\varphi^{-1})(\mathrm{E}')$ car le rang vaut r et d'autre part que la restriction $\pi : \dim \mathrm{D}_x(f\circ\varphi^{-1})(\mathrm{E}') \to \mathrm{Im}\,\mathrm{D}_0 f$ est un isomorphisme car elle est surjective et les dimensions coïncident. Comme $\pi(\mathrm{D}_x(f \circ \varphi^{-1})(\mathrm{Ker}\,\mathrm{D}_0 f)) = 0$, il vient $\mathrm{D}_x(f \circ \varphi^{-1})(\mathrm{Ker}\,\mathrm{D}_0 f) = 0$. Alternativement, on considère la matrice jacobienne de $(f \circ \varphi^{-1})$ en un point $(x,y) \in \mathrm{V}'$:

$$\mathrm{D}_{(x,y)} g = \begin{pmatrix} \mathrm{I}_r & 0 \\ \partial_1 g & \partial_2 g \end{pmatrix}$$

Or, pour tout $(x,y) \in \mathrm{V}'$, $\mathrm{rg}_{(x,y)}(f \circ \varphi^{-1}) = r$, donc $\partial_2 g \equiv 0$ d'après le lemme 10.3 ci-dessous.

Autrement dit, comme V′ est convexe, pour $z, w \in \mathrm{V}'$ tel que $z - w \in \mathrm{Ker}\,\mathrm{D}_0 f$, on obtient $g(z) = g(w)$. Ceci implique que g ne dépend que de x, ce que l'on peut encore écrire sous la forme $g = g \circ \pi_1$. Il vient $(f \circ \varphi^{-1})(x,y) = (x, g(x))$ pour $(x,y) \in (\mathbb{R}^r \times \mathbb{R}^{q-r}) \cap \mathrm{V}'$.

Enfin, on note $\psi = \mathrm{Id} - g \circ (\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi$ c'est-à-dire $\psi : (x,z) \mapsto (x, z - g(x))$, bien définie au voisinage de $0 \in \mathrm{F}$. Cette application est injective, de classe $\mathcal{C}^k$ et

$$\mathrm{D}\psi = \mathrm{Id} - \mathrm{D}_{(\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi} g \circ (\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi\,.$$

Montrons que c'est un isomorphisme: si $\mathrm{D}\psi(h) = 0$ alors on déduit d'abord que $h \in \mathrm{Ker}\,\pi$ car g y prend ses valeurs donc

$$h = \mathrm{D}_{(\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi} g \circ (\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi(h) = 0\,.$$

Ceci montre que $\mathrm{D}\psi$ est inversible. En coordonnées, on obtient

$$\mathrm{D}\psi = \begin{pmatrix} \mathrm{I}_r & 0 \\ -\partial_1 g & \mathrm{I}_{q-r} \end{pmatrix}$$

sur son ensemble de définition, donc on retrouve que $\mathrm{D}\psi$ est inversible en tout point où ψ est défini. Le théorème d'inversion globale implique que ψ est un difféomorphisme sur son image. On obtient ainsi

$$\psi \circ f \circ \varphi^{-1} = [\mathrm{D}_0 f + g] - g \circ (\mathrm{D}_0 f|_{\mathrm{E}'})^{-1} \circ \pi \circ [\mathrm{D}_0 f + g] = \mathrm{D}_0 f + g - g \circ \pi_1 = \mathrm{D}_0 f$$

car g prend ses valeurs dans $\mathrm{F}' = \mathrm{Ker}\,\pi$. On peut aussi conduire le calcul en coordonnées:

$$(\psi \circ f \circ \varphi^{-1})(x,y) = \psi(x, g(x)) = (x, 0)\,.$$

Cela termine la démonstration. ■

Lemme 10.3. — *On considère une matrice par blocs* $\mathrm{A} = \begin{pmatrix} \mathrm{I}_r & 0 \\ \mathrm{B} & \mathrm{C} \end{pmatrix}$. *Si* $\mathrm{rg}\,\mathrm{A} = r$ *alors* $\mathrm{C} = 0$.

Démonstration. — On note $v_1, \ldots, v_n$ les vecteurs colonnes de A. On constate que les r premières colonnes $(v_1, \ldots, v_r)$ forment une famille libre car elles sont échelonnées. Comme on a supposé que A est de rang r, tout autre vecteur colonne v_j, $r < j \leq n$, est combinaison linéaire des r premières. On a donc $v_j = \sum_{k=1}^r \lambda_k v_k$. En regardant la $k$ième ligne de A, on constate que l'on doit avoir $\lambda_k = 0$. Donc $v_j = 0$ pour tout $r < j \leq n$. Cela signifie $\mathrm{C} = 0$. ■

10.2 Points critiques

On s'intéresse ici au cas où f admet un point critique isolé, de différentielle seconde non dégénérée. On note $\mathcal{S}_n(\mathbb{R})$ l'espace des matrices symétriques. On rappelle le théorème de Gauss concernant les formes quadratiques non dégénérées définies sur $\mathbb{R}$.

Théorème 10.4. — *Si q est une forme quadratique non dégénérée définie sur un espace vectoriel de dimension finie, il existe un unique couple d'entiers naturels (s,t), appelé la signature de q tel que $s+t = \dim \mathrm{E} \stackrel{def.}{=} n$, et tel qu'il existe une base dans laquelle*

$$q(x_1,\dots,x_n) = \sum_{1\leq j\leq s} x_j^2 - \sum_{s+1\leq j\leq n} x_j^2.$$

On montre une version « à géométrie variable » du résultat précédent.

Lemme de Morse. — *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe $\mathcal{C}^k$, $k\geq 2$ telle que $f(0) = \mathrm{D}_0 f = 0$ et telle que $\mathrm{D}_0^2 f$ soit inversible. Alors il existe un voisinage V de l'origine et des applications de classe $\mathcal{C}^{k-2}$ $y_1,\dots,y_r$ et $z_1,\dots,z_s$ définies sur V telles que $r+s=n$ et, pour $x\in\mathrm{V}$, on ait*

$$f(x) = \sum_{1\leq j\leq r} y_j^2 - \sum_{1\leq j\leq s} z_j^2\,.$$

Démonstration. — On considère le développement limité avec reste intégrale de f au voisinage de l'origine. On a

$$f(x) = \int_0^1 (1-t)\mathrm{D}^2_{tx} f(x,x)dt = {}^t\mathrm{X}\cdot \int_0^1 (1-t)\mathrm{D}^2_{tx} f dt\cdot \mathrm{X}\,.$$

On note

$$\mathrm{A}(x) = \left(\int_0^1 (1-t)\frac{\partial^2 f}{\partial x_i\partial x_j}(tx)dt\right)_{i,j} \qquad \text{et} \qquad \mathrm{A}_0 = \mathrm{A}(0)\,.$$

On utilise alors le lemme suivant.

Lemme 10.5. — *Si $\mathrm{A}_0\in \mathcal{S}_n(\mathbb{R})\cap \mathrm{GL}_n(\mathbb{R})$, il existe un voisinage U de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application infiniement différentiable $\psi : \mathrm{U}\to \mathcal{M}_n(\mathbb{R})$ tels que $\psi(\mathrm{A}_0) = \mathrm{I}$ et, pour tout $\mathrm{A}\in\mathrm{U}$,*

$${}^t\psi(\mathrm{A})\cdot \mathrm{A}_0\cdot \psi(\mathrm{A}) = \mathrm{A}\,.$$

Du coup, si x est assez proche de l'origine, alors $f(x) = {}^tx{}^t\psi(\mathrm{A}(x))\mathrm{A}_0\psi(\mathrm{A}(x))x$. On pose $\psi_1(x) = \psi(\mathrm{A}(x))x$, et on obtient

$$f(x) = {}^t\psi_1(x)\mathrm{A}_0\psi_1(x)\,.$$

Or il existe une base P dans laquelle A_0 soit diagonale avec r valeurs sur la diagonale égales à 1 et s égales à -1. On note J cette matrice et on pose $\psi_2 = \mathrm{P}\cdot\psi_1$. Il vient $f(x) ={}^t\psi_2(x)\mathrm{J}\psi_2(x)$. Si on appelle $y_1,\dots,y_r$ et $z_1,\dots,z_s$ les coordonnées de ψ_2, on obtient la forme recherchée.

Ceci établit le lemme de Morse modulo le Lemme 10.5. ∎

Démonstration du Lemme 10.5. — On considère l'application $h : \mathcal{M}_n(\mathbb{R}) \to \mathcal{S}_n(\mathbb{R})$ définie par $h(\mathrm{M}) = {}^t\mathrm{M}\mathrm{A}_0\mathrm{M}$. On calcule la différentielle à l'identité de h.

$$h(\mathrm{I}+\mathrm{M}) = {}^t(\mathrm{I}+\mathrm{M})\mathrm{A}_0(\mathrm{I}+\mathrm{M}) = \mathrm{A}_0 + ({}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}) + {}^t\mathrm{M}\mathrm{A}_0\mathrm{M} = \mathrm{A}_0 + ({}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}) + \mathrm{O}(\|\mathrm{M}\|^2)\,.$$

Donc $\mathrm{D_I}h(\mathrm{M}) = {}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}$. Cette application n'est pas inversible puisque $\dim\mathcal{S}_n(\mathbb{R}) < \dim\mathcal{M}_n(\mathbb{R})$. En revanche, le noyau consiste en les matrices M telles que $\mathrm{A}_0\mathrm{M}$ soit antisymétrique. On considère l'espace E des matrices M telles que $\mathrm{A}_0\mathrm{M}$ soit symétrique. Cet espace est supplémentaire à $\mathrm{Ker}\,\mathrm{D_I}h$, et la restriction de $\mathrm{D_I}h$ à E devient maintenant inversible (injective car A_0 est inversible et espaces source et but de même dimension).

Le théorème d'inversion locale appliqué à $h|_\mathrm{E}$ montre qu'il existe des voisinages U de A_0 et V de I et un difféomorphisme infiniment différentiable $\psi : \mathrm{U}\to\mathrm{V}$ qui inverse $h|_\mathrm{E}$. ∎

On conclut par un dernier énoncé important. On considère une fonction f définie sur un ouvert U de $\mathbb{R}^p$, à valeurs dans $\mathbb{R}^q$, et de classe $\mathcal{C}^k$.

On appelle *points critiques* les points en lesquels l'application différentielle de f est non surjective, et *valeurs critiques* les images des points critiques. Les valeurs non critiques sont dites *régulières* (qu'elles soient des valeurs effectivement prises par f ou non). On rassemble deux résultats en un, dûs à Sard et Brown respectivement.

Théorème 10.6 (Sard, Brown). — *Soit* $f : \mathrm{U} \to \mathbb{R}^q$ *de classe* $\mathcal{C}^k$*, où* U *est un ouvert de* $\mathbb{R}^p$*.*

1. *Si* $k > \max(0, p - q)$*, alors l'ensemble des valeurs critiques a une mesure de Lebesgue nulle.*

2. *L'ensemble des valeurs régulières est dense dans* $\mathbb{R}^q$*.*

En revanche, l'ensemble des points critiques peut être très important, par exemple si $p < q$, tous les points sont critiques, mais l'ensemble image de f sera quand même de mesure nulle.