SOUS-VARIÉTÉS DE $\mathbb{R}^n$, THÉORÈME D'INVERSION LOCALE ET CONSÉQUENCES

AGRÉGATION DE MATHÉMATIQUES, RENNES 1
ISMAËL BAILLEUL

Je ne saurai trop recommander pour les sujets abordés dans ces notes le livre "*Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*", de Rouvière. Le livre *Catastrophes et bifurcations*, de Demazure, est aussi très bien pour visualiser les choses et par la qualité de ses explications et démonstrations.

1. Définitions et exemples

Soit V un sous-ensemble de $\mathbb{R}^n$ et $1 \leqslant d \leqslant n$.

Définition **1.** *On dit que V est une sous-variété de $\mathbb{R}^n$ de dimension d si chaque point de V admet un voisinage $\mathcal{U}$ dans $\mathbb{R}^n$ sur lequel est défini un $\mathcal{C}^1$-difféomorphisme $\varphi : \mathcal{U} \to \mathbb{R}^n$ envoyant $V \cap \mathcal{U}$ sur un ouvert de $\mathbb{R}^d$.*

Théorème **2.** *Les assertions suivantes sont équivalentes.*

- **i) Définition.** *V est une sous-variété de $\mathbb{R}^n$.*
- **ii) Graphe.** *Tout point de V admet un voisinage $\mathcal{U}$ dans $\mathbb{R}^n$ sur lequel sont définies des coordonnées*[1] *$(x_1, \dots, x_n)$ et une application $f = (f_{d+1}, \dots, f_n)$ définie sur un ouvert Ω de $\mathbb{R}^d$ et à valeurs dans $\mathbb{R}^{n-d}$, tels que $\mathcal{U} \cap V$ est le graphe de f : $m \in \mathcal{U}$, de coordonnées $(x_1, \dots, x_n)$, appartient à $\mathcal{U} \cap V$ ssi $x_{d+1} = f_{d+1}(x_1, \dots, x_d), \dots, x_n = f_n(x_1, \dots, x_d)$.*
- **iii) Paramétrage.** *Tout point a de V admet un voisinage $\mathcal{U}$ dans $\mathbb{R}^n$ auquel on peut associer un homéomorphisme $\psi : \Omega \to V \cap \mathcal{U}$ de classe $\mathcal{C}^1$ d'un voisinage ouvert connexe Ω de 0 dans $\mathbb{R}^d$ sur $V \cap \mathcal{U}$, tel que $\psi(0) = a$ et $D_0\psi$ est injective.*
- **iv) Description implicite.** *Tout point a de V admet un voisinage $\mathcal{U}$ dans $\mathbb{R}^n$ sur lequel est défini une fonction $g : \mathcal{U} \to \mathbb{R}^{n-d}$ telle que $V \cap \mathcal{U} = g^{-1}\big(\{0\}\big)$ et $D_a g : \mathbb{R}^n \to \mathbb{R}^{n-d}$ est surjective.*

Noter que dans la caractérisation à l'aide d'un paramétrage local, on demande à ce paramétrage d'être un homéomorphisme : deux points de V voisins dans $\mathbb{R}^n$ doivent donc avoir des paramètres voisins. La démonstration de ce théorème repose entièrement sur le théorème d'inversion locale ; l'un et l'autre sont démontrés dans la section 3.2 dédiée aux démonstrations.

Exemples.

(1) Le graphe d'une fonction $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}^{n-d}$ est une sous-variété. Etant donné la description d'une sous-variété donnée par l'item ii), posons $\varphi(m) = \big(x_1, \dots, x_d, x_{d+1} - f_{d+1}(m), \dots, x_n - f_n(m)\big)$, pour $m \in \mathcal{U}$. On vérifie directement

[1] C'est-à-dire un $\mathcal{C}^1$-difféomorphisme $\mathfrak{x} : \mathcal{U} \to \mathbb{R}^n$ qui à m associe $\mathfrak{x}(m) = (x_1, \dots, x_n)$.

sans mal que φ est un $\mathcal{C}^1$-difféomorphisme envoyant $V \cap \mathcal{U}$ sur Ω. Cela prouve l'implication **ii)** $\Rightarrow$ **i)**.

(2) La courbe du plan décrite par le paramétrage $x(t) = \frac{3t}{1+t^3}$, $y(t) = \frac{3t^2}{1+t^3}$, pour $t > -1$, n'est pas une sous-variété du plan comme il y a sur cette courbe des points arbitrairement proches avec des paramètres non proches, correspondant à $t = 0$ et $t \to +\infty$.

(3) Voir l'exercice 86 du livre "*Petit guide de calcul différentiel (...)*" de Rouvière pour d'autres exemples et contre-exemples instructifs.

(4) **Sous-variétés de GL($\mathbb{R}^n$).** L'ensemble $\mathrm{SL}(\mathbb{R}^n)$ des matrices de déterminant 1 est une sous-variété de $\mathrm{GL}(\mathbb{R}^n)$. Posons $g : \mathrm{GL}(\mathbb{R}^n) \to \mathbb{R}, A \mapsto \mathrm{dét} A - 1$; l'ensemble $\mathrm{SL}(\mathbb{R}^n) = g^{-1}\big(\{0\}\big)$. Calculons sa différentielle en un point A quelconque A de $\mathrm{SL}(\mathbb{R}^n)$. On a $\mathrm{dét}(A + M) = \big(\mathrm{dét} A\big)\mathrm{dét}(\mathrm{Id} + A^{-1}M)$. Pour obtenir un développement limité de $\mathrm{dét}(\mathrm{Id} + A^{-1}M)$ au premier ordre en M, notez que l'ensemble des applications linéaires M telles que $A^{-1}M$ est diagonalisable est dense dans $\mathrm{L}(\mathbb{R}^n)$; voyez-vous pourquoi ? Pour de telles applications le calcul de $\mathrm{dét}(\mathrm{Id} + A^{-1}M)$ se fait trivialement dans une base où la matrice de $A^{-1}M$ est diagonale, d'où l'on tire : $\mathrm{dét}(\mathrm{Id} + A^{-1}M) = \mathrm{tr}(A^{-1}M) + o(|M|)$, ce qui donne $(D_A g)(M) = (\mathrm{dét} A)\,\mathrm{tr}(A^{-1}M) = \mathrm{tr}(co(A)^* M)$, où $co(A)^*$ désigne la transposée de la comatrice de A. La continuité de l'application déterminant permet d'étendre cette relation à toute $M \in \mathrm{L}(\mathbb{R}^n)$. On lit sur cette formule la surjectivité de l'application $D_A g : \mathrm{L}(\mathbb{R}^n) \to \mathbb{R}$. Le groupe spécial linéaire $\mathrm{SL}(\mathbb{R}^n)$ est donc une sous-variété de $\mathrm{GL}(\mathbb{R}^n)$ de dimension $n^2 - 1$.

Voici un autre exemple. Le groupe orthogonal $\mathrm{O}(\mathbb{R}^n) = \{A \in \mathrm{L}(\mathbb{R}^n)\,;\, A^*A = \mathrm{Id}\}$. Notons $\mathcal{S}$ l'ensemble des matrices symétriques et définissons $g : \mathrm{L}(\mathbb{R}^n) \to \mathcal{S}$ par la formule $g(A) = A^*A - \mathrm{Id}$. On voit sans mal sur l'expression $(D_A g)(M) = A^{-1}M + \big(A^{-1}M\big)^*$ de $D_A g$ qu'elle est surjective, comme $\big(D_A g\big)\big(\frac{1}{2}AS\big) = S$ pour toute matrice symétrique S. Le groupe orthogonal est donc une sous-variété de $\mathrm{GL}(\mathbb{R}^n)$ de dimension $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

(5) On démontre le théorème suivant dans la section 3.2 consacrée aux démonstrations.

Théorème **3** (von Neumann)**.** *Tout sous-groupe fermé non discret de* $\mathrm{GL}(\mathbb{R}^n)$ *est une sous-variété de* $\mathrm{L}(\mathbb{R}^n)$.

2. Plan tangent

2.1. Définition et conséquences.

Définition **4.** *Soit V une sous-variété de $\mathbb{R}^n$ et $a \in V$. On appelle* ***plan tangent en** a **à** V *l'ensemble des vecteurs $\gamma'(0)$ où $\gamma : (-1, 1) \to V$ est une courbe $\mathcal{C}^1$ à valeurs dans V et $\gamma(0) = a$. On le note T_aV.*

Proposition **5.** *L'ensemble T_aV est un espace vectoriel de dimension d.*

Démonstration – L'assertion est claire si $V = \mathbb{R}^d \subset \mathbb{R}^n$. Dans le cas générique, il existe d'après la définition un $\mathcal{C}^1$-difféomorphisme φ^{-1} d'un voisinage ouvert Ω de 0 dans $\mathbb{R}^n$ sur un voisinage $\mathcal{U}$ de a dans $\mathbb{R}^n$ tel que $V \cap \mathcal{U} = \varphi^{-1}(\Omega)$. Comme toute courbe γ à valeurs dans $V \cap \mathcal{U}$ est de la forme $\varphi^{-1} \circ \rho$ pour une courbe ρ à valeurs dans $\mathbb{R}^d \subset \mathbb{R}^n$, il s'ensuit que $T_aV = \big(D_0\varphi^{-1}\big)(\mathbb{R}^d)$, qui est bien un espace vectoriel de dimension d puisque $D_0\varphi^{-1}$ est inversible. ▷

Exemples.

(6) Décrivez le plan tangent au graphe d'une fonction.

(7) Supposons V donné sous la forme d'un ensemble de niveau $V = g^{-1}(\{0\})$ d'une fonction $g : \mathcal{U}(\subset \mathbb{R}^n) \to \mathbb{R}^{n-d}$; alors $T_aV = \ker(D_a g)$. En effet, comme $t \mapsto g(\gamma(t))$ est constant pour toute courbe γ à valeurs dans V, on a $T_aV \subset \ker(D_a g)$. C'est une égalité comme ces deux espaces ont même dimension. Notez que si $g = (g_1, \dots, g_{n-d})$ alors $\ker(D_a g) = \bigcap_{i=1..n-d} \ker(D_a g_i)$.

(8) Trouver les espaces tangents en l'identité à $\mathrm{SL}(\mathbb{R}^n)$ et $\mathrm{O}(\mathbb{R}^n)$.

La proposition suivante donne une version précisée de la description d'une sous-variété de $\mathbb{R}^n$ en tant que graphe d'une fonction, et justifie l'impication **iii)** $\Rightarrow$ **ii)** de la caractérisation des sous-variétés donnée au théorème 2.

Proposition **6.** *Soit E un supplémentaire de T_aV dans $\mathbb{R}^n$. Il existe un voisinage U de 0 dans T_aV et un voisinage $\mathcal{U}$ de a dans $\mathbb{R}^n$ tels que $V \cap \mathcal{U}$ est le graphe d'une application de classe $\mathcal{C}^1$ de U dans E.*

Démonstration – Notons $\pi_1 : \mathbb{R}^n \to T_aV$ la projection sur T_aV parallèlement à E et π_2 la projection sur E parallèlement à T_aV. Soit $\psi : \Omega \to \mathbb{R}^n$ un paramétrage local d'un voisinage $U = \mathcal{U} \cap V$ de a dans V, et $(\psi_1, \psi_2) = (\pi_1 \circ \psi, \pi_2 \circ \psi)$ ses composantes sur T_aV et E respectivement. Comme $D_0\psi$ est injective elle a T_aV pour image (même dimensions !) et $\mathrm{Im}(D_0\psi_1) = \pi_1\big(\mathrm{Im}(D_0\psi)\big) = \pi_1(T_aV) = T_aV$, le théorème d'inversion locale nous dit que ψ_1 est un difféomorphisme. L'ensemble $V \cap \mathcal{U}$ est le graphe de l'application $\psi_2 \circ \psi_1^{-1}$. (Faites un dessin !) $\triangleright$

Corollaire **7** (Théorème des fonctions implicites)**.** *Soit V une sous-variété de $\mathbb{R}^n \times \mathbb{R}^d$ et $a = (a', a'')$ un point de V. Si la restriction à T_aV de la projection sur $\mathbb{R}^n$ parallelement à $\mathbb{R}^d$ est une bijection, c'est-à-dire si $T_aV \oplus \mathbb{R}^d = \mathbb{R}^n \times \mathbb{R}^d$, il existe un voisinage ouvert U' de a' dans $\mathbb{R}^n$, un voisinage ouvert U'' de a'' dans $\mathbb{R}^d$ et une application $f : U' \to U''$ de classe $\mathcal{C}^1$ telle que $V \cap (U' \times U'')$ soit le graphe $\big\{(x'.f(x'))\big\}_{x' \in U'}$ de f.*

Démonstration – On utilise la proposition 6 et la fait que la projection π de T_aV sur U' est une bijection linéaire. Avec les notations précédentes : $f = \psi_2 \circ \psi_1^{-1} \circ \pi$. $\triangleright$

Dans une situation pratique, V est donné sous la forme $h^{-1}(\{0\})$ pour une fonction $h : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^d$ telle qu'en $a = (a', a'')$ la différentielle $\partial_2 h$ de h par rapport à sa seconde variable est un isomorphisme.

2.2. **Extrema liés.**

Proposition **8.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe $\mathcal{C}^1$.*

- *Si $f_{|V}$ a un extremum en un point a de V alors $D_a f_{|T_aV} = 0$.*
- *En particulierm si V est définie par les équations $g_1 = \cdots = g_{n-d} = 0$, où les g_i sont comme dans le théorème 2, il existe des constantes $\lambda_1, \dots, \lambda_{n-d}$ telles que*

$$D_a f = \lambda_1 D_a g_1 + \cdots + \lambda_{n-d} D_a g_{n-d},$$

appelées multiplicateurs de Lagrange.

Démonstration – Comme on a $f \circ \gamma = 0$ pour toute courbe γ à valeurs dans V, le premier point se lit sur la définition du plan tangent. Pour le second, on commence par compléter la famille *libre* des formes linéaires $D_a g_i$ en une base $(\epsilon_i^*)_{i=1..n}$ de $(\mathbb{R}^n)^*$

dont on note $(\epsilon_i)_{i=1..n}$ la base duale dans $\mathbb{R}^n$. Si l'on avait $D_af = \sum_{i=1}^{n-d}\lambda_i D_a g_i + \sum_{j=n-d+1}^{n}\lambda_j\epsilon_j^*$, avec un $\lambda_{j_0} \neq 0$ pour un indice $n-d+1 \leqslant j_0 \leqslant n$, on aurait $(D_af)(\epsilon_{j_0}) = \lambda_{j_0} \neq 0$ alors que $\epsilon_{j_0} \in \cap_{i=1}^{n-d}\ker(D_a g_i)$, contredisant le premier point de la proposition. ▷

Corollaire **9.** *Etant donnés une forme quadratique q sur $\mathbb{R}^n$ définie positive et une forme quadratique quelconque q' sur $\mathbb{R}^n$, il existe une base q orthonormée dans laquelle la matrice de q' est diagonale.*

Démonstration – Remarquons d'abord que puisque q est définie positive, sa sphère unité $V = \{q = 1\}$ est une sous-variété compacte de $\mathbb{R}^n$. (Voyez-vous pourquoi ?) En un point a où la restriction de q' à V est maximale on a $D_aq' = \lambda D_a q$ pour un certain multiplicateur $\lambda \in \mathbb{R}$. Choississons une base q-orthonormée quelconque et notons M la matrice de q' dans cette base. L'identité précédente se lit $a^*M = \lambda a^*\mathrm{Id}$, c'est-à-dire, $Ma = \lambda a$; aussi le vecteur a est-il propre pour M. Notons que si $q(a,b) = 0$ alors $q'(a,b) = \lambda q(a,b) = 0$, si bien que M préserve le q-orthogonal de la droite engendré par a. Le résultat s'ensuit par récurrence sur la dimension en considérant les restrictions de q et q' à l'orthogonal de la droite engendré par a. ▷

3. Le théorème d'inversion locale, applications

3.1. **Le théorème d'inversion locale.** Soient E et F deux espaces de Banach, $\mathcal{U}$ un voisinage ouvert d'un point x_0 de E et $f : \mathcal{U} \to F$ une application de classe $\mathcal{C}^1$. Rappelons qu'en dimension infinie, une application linéaire n'est pas forcément continue. C'est cependant un conséquence immédiate du théorème de l'application ouverte de Banach que l'inverse d'une application linéaire *bijective continue* de E dans F est automatiquement continue.[2]

Théorème **10** (d'inversion locale). *Si $D_{x_0}f$ est une bijection continue de E dans F alors f est un $\mathcal{C}^1$-difféomorphisme d'un voisinage $\mathcal{V}$ de x_0 sur $f(\mathcal{V})$.*

La démonstration de ce théorème a mauvaise presse bien qu'il s'agisse d'une application élémentaire du théorème du point fixe. Il suffit en effet de trouver un voisinage $\mathcal{V}$ de x_0 sur lequel f définit une *bijection* de $\mathcal{V}$ sur $f(\mathcal{V})$ pour conclure. La formule de Taylor à l'ordre 1 nous assure alors que puisqu'on a

$$f(y) - f(x) = (D_xf)(y-x) + o(y-x)$$

pour tous $x, y \in \mathcal{V}$, on a d'une part $o(y-x) = o\big(f(y) - f(x)\big)$ et que puisque $\big(D_{f^{-1}(a)}f\big)^{-1}$ est continue on a d'autre part

$$\big(D_{f^{-1}(a)}f\big)^{-1}(b-a) = f^{-1}(b) - f^{-1}(a) + o(b-a)$$

[2]Notons L cette application linéaire continue et bijective de E dans F. La continuité de L^{-1} est équivalente au fait que L envoit un ouvert quelconque de E sur un ouvert de F. Il suffit pour démontrer cela de voir que l'image par L de la boule de E de rayon 1 contient un voisinage ouvert de 0 dans F. Comme F d'intérieur non vide (!), s'écrit $F = \bigcup_{n\geqslant 1}\overline{L\big(B_E(0,n)\big)}$ comme un réunion de fermés, l'un d'entre eux (donc chacun d'entre eux par homogénéité) est d'intérieur non vide d'après la *propriété de Baire*. On a donc un certain n_0 pour lequel $B_F(0,1) \subset \overline{L\big(B_E(0,n_0)\big)}$, soit encore $B_F(0,2^{-k}) \subset \overline{L\big(B_E(0,2^{-k}n_0)\big)}$ pour tout $k \geqslant 1$. Etant donné un point a de $B_F(0,1)$, on construit alors par récurrence une suite de points x_k de $B_E(0,2^{-k}n_0)$ telle que $a - L(x_1 + \cdots + x_k)$ soit de norme inférieure à $2^{-(k+1)}$. La série $\sum x_k$ converge vers un point y de $B_E(0,2n_0)$ qui satisfait $L(y) = a$ puisque L est continue.

pour tous $a, b \in f(\mathcal{V})$. Cette formule signifie que f^{-1} est différentiable en tout point $a \in f(\mathcal{V})$, de différentielle $\big(D_{f^{-1}(a)}f\big)^{-1}$ en a ; elle est à fortiori continue et sa différentielle est continue en tant que composée de fonctions continues.

Démonstration – Comme $(D_{x_0}f)^{-1}$ est une application linéaire continue, quitte à considérer
$(D_{x_0}f)^{-1}\big(f(x_0+\cdot) - f(x_0)\big)$ plutôt que $f(\cdot)$, on se ramène à considérer le cas où $E = F, x_0 = f(x_0) = 0$ et $D_0 f = \mathrm{Id}$. L'identité $f(x) = x + u(x)$ définit alors une fonction u de classe $\mathcal{C}^1$ dont la différentielle est nulle en 0. Comme $D_z u$ dépend continûment de z il existe une boule $B(0,r)$ sur laquelle $\|D_z u\| \leqslant \frac{1}{2}$; l'identité élémentaire $u(y) - u(x) = \int_0^1 \big(D_{x+t(y-x)}f\big)(y-x)dt$, nous assure alors que la restriction de u à $B(0,r)$ est $\frac{1}{2}$-lipschitzienne. Pour $x, y \in B(0,r)$ on a donc l'inégalité $\big|f(x) - f(y)\big| = \big|x + u(x) - y - u(y)\big| \geqslant \frac{1}{2}|x-y|$, sur laquelle on lit l'injectivité de la restriction de f à $B(0,r)$. On montre que $B\big(0,\frac{r}{2}\big) \subset f\big(B(0,r)\big)$ en remarquant que pour $a \in B\big(0,\frac{r}{2}\big)$ et $x \in B(0,r)$ on a $\big|a - u(x)\big| \leqslant r$. L'application $\frac{1}{2}$-lipschitzienne $x \mapsto a - u(x)$ envoit donc la boule $B(0,r)$ dans elle même, et y a un unique point fixe y qui vérifie $a - u(y) = y$, c'est-à-dire $a = y + u(y) = f(y)$. ▷

Notez qu'en dimension finie $f = (f_1, \ldots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$, la différentielle $D_{x_0}f$ de f en x_0 est inversible ssi les formes linéaires $\big(D_{x_0}f_i\big)_{i=1..n}$ forment une base de $(\mathbb{R}^n)^*$.

3.2. **Caractérisation des sous-variétés de $\mathbb{R}^n$.** On démontre dans ce paragraphe le théorème 2.

• **i) ⇒ ii), iii), iv).** Soit V une sous-variété de $\mathbb{R}^n$ et a un point de V et $\mathcal{U}$ un voisinage de a dans $\mathbb{R}^n$ sur lequel est défini un $\mathcal{C}^1$-difféomorphisme $\varphi : \mathcal{U} \to \mathbb{R}^n$ envoyant $V \cap \mathcal{U}$ sur un ouvert de $\mathbb{R}^d$. Notons $\pi_i : \mathbb{R}^n \to \mathbb{R}$ la projection sur la i$^{\text{ème}}$ coordonnée. Le difféomorphisme φ définit des coordonnées $(x_1, \ldots, x_n) = (\pi_1 \circ \varphi, \ldots, \pi_n \circ \varphi)$ sur $\mathcal{U}$ et un point de coordonnées $(x_1, \ldots, x_n)$ appartient à $\mathcal{U} \cap V$ ssi $x_{d+1} = 0, \ldots, x_n = 0$. L'ensemble V est donc bien décrit localement par le graphe d'une fonction (constante !). La restriction de φ^{-1} à $\mathbb{R}^d$ fournit par ailleurs un paramétrage de V au sens du iii), et $V \cap \mathcal{U} = g^{-1}\big(\{0\}\big)$, où $g = (\pi_{d+1} \circ \varphi, \ldots, \pi_n \circ \varphi)$ a bien sa différentielle en a surjective.

iii) ⇒ ii) ⇒ i). Ces implications ont été montrées à la proposition 6 et l'exemple (1).

iv) ⇒ i) Puisque $D_a g : \mathbb{R}^n \to \mathbb{R}^{n-d}$ est surjective, on peut choisir une famille libre $\big(\epsilon^*_{n-d+1}, \ldots, \epsilon^*_n\big)$ de $(\mathbb{R}^n)^*$ telle que $\big(D_a g_1, \ldots, D_a g_{n-d}, \epsilon^*_{n-d+1}, \ldots, \epsilon^*_n\big)$ est une base de $(\mathbb{R}^n)^*$; notons $(\epsilon_i)_{i=1..n}$ la base duale dans $\mathbb{R}^n$ et repérons un point m de $\mathcal{U}$ par ses coordonnées $m = a + \sum_{i=1}^n \lambda_i \epsilon_i$ dans le repère affine attaché à a. On définit alors une application $\varphi : \mathcal{U} \to \mathbb{R}^n$ de classe $\mathcal{C}^1$ par la formule

$$\begin{aligned}\varphi(m) &= \big(g_1(m), \ldots, g_{n-d}(m), \lambda_{n-d+1}, \ldots, \lambda_n\big)\\ &= \big(g_1(m), \ldots, g_{n-d}(m), \epsilon^*_{n-d+1}(m), \ldots, \epsilon^*_n(m)\big).\end{aligned}$$

Le théorème d'inversion locale s'applique comme les formes linéaires $D_a g_1, \ldots, D_a g_{n-d}, \epsilon^*_{n-d+1}, \ldots, \epsilon^*_n$ forment une base de $(\mathbb{R}^n)^*$.

3.3. **Illustrations du théorème d'inversion locale.** On donne dans cette section quelques illustrations du théorème d'inversion locale.

a) Application exponentielle et petits sous-groupes de GL($\mathbb{R}^n$). Rappelons que l'exponentielle d'une matrice est définie par la série absolument convergente $\exp(A) =$

$\sum_{k\geqslant 0}\frac{A^k}{k!}$, pour tous $A \in \mathrm{L}(\mathbb{R}^n)$. On lit sur le terme linéaire de cette série que la différentielle en 0 de l'application exponentielle est l'identité. L'application exponentielle est donc un difféomorphisme d'un voisinage $\mathcal{V}$ de 0 dans $\mathrm{L}(\mathbb{R}^n)$ sur un voisinage de l'identité dans $\mathrm{GL}(\mathbb{R}^n)$.

PROPOSITION **11.** *Il n'existe pas dans* $\mathrm{GL}(\mathbb{R}^n)$ *de sous-groupes arbitrairement petit, au sens où il existe un voisinage* $\mathcal{U}$ *de l'identité tel que tout sous-groupe de* $\mathrm{GL}(\mathbb{R}^n)$ *contenu dans* $\mathcal{U}$ *est trivial, égal à* $\{\mathrm{Id}\}$.

DÉMONSTRATION – Travaillons dans le voisinage $\mathcal{V}$ de 0 dans $\mathrm{L}(\mathbb{R}^n)$ mis en avant ci-dessus, sur lequel l'application exponentielle est un difféomorphisme ; on ne perd rien à le supposer borné. Si $\exp\left(\frac{\mathcal{V}}{2}\right)$ contenait un sous-groupe non trivial, un quelconque de ses éléments $M \neq \mathrm{Id}$ s'écrirait $M = \exp(A)$, pour une matrice $A \in \frac{\mathcal{V}}{2}$ non nulle. On aurait donc d'une part $M^k = \exp(kA)$ pour tous $k \geqslant 1$, et d'autre part $M^k = \exp(B_k)$, pour un unique élément B_k de $\frac{\mathcal{V}}{2}$. L'injectivité de l'application exponentielle sur $\mathcal{V}$ forcerait alors l'égalité $B_k = kA$ pour tous les k tels que $kA \in \mathcal{V}$. Si cependant k_0 désignait le plus petit entier k tel que $kA \in \frac{\mathcal{V}}{2}$ et $(k+1)A \in \mathcal{V}\backslash\frac{\mathcal{V}}{2}$, le caractère borné de $\mathcal{V}$ nous assurerait que k_0 est fini, amenant la contradiction $B_{k_0+1} = (k_0+1)A \in \mathcal{V}\backslash\frac{\mathcal{V}}{2}$. $\triangleright$

On trouve cette démonstration p.59 du livre de Mneimé et Testard.

b) Etude locale d'une fonction. Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^n$ et $f : \mathcal{U} \to \mathbb{R}^p$ une fonction de classe $\mathcal{C}^1$. Etant donné $1 \leqslant r \leqslant p$, on note $A_r : \mathbb{R}^n \to \mathbb{R}^p$ l'application linéaire $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0) \in \mathbb{R}^p$. (Il n'y a de 0 dans cette définition que si $r < p$.)

THÉORÈME **12** (du rang constant.)**.** *Si la différentielle de* f *est de rang constant sur tout* $\mathcal{U}$*, disons* r*, alors on peut associer à chaque point* a *de* $\mathcal{U}$ *un voisinage* $\mathcal{V}_1$ *de* a *dans* $\mathcal{U}$ *et un voisinage* $\mathcal{V}_2$ *de* $f(a)$ *dans* $\mathbb{R}^p$ *sur lequels sont définies deux difféomorphismes* $h_1 : \mathcal{V}_1 \to \mathcal{V}_1$ *et* $h_2 : \mathcal{V}_2 \to \mathcal{V}_2$ *tels que* $f = h_2 \circ A_r \circ h_1$.

L'assertion signifie que quitte à changer de coordonnées au départ et à l'arrivée, l'application f n'est localement rien d'autre que l'application linéaire A_r. La démonstration classique suivante suit celle de Demazure.

DÉMONSTRATION – On peut supposer que l'on a $a = 0$ et $f(a) = 0$. On commence par choisir des coordonnées linéaires $(u_1, \dots, u_n)$ sur $\mathbb{R}^n$ et $(v_1, \dots, v_p)$ sur $\mathbb{R}^p$ dans lesquelles la matrice $D_a f$ est A_r. Le théorème d'inversion locale nous assure alors que les fonctions

$$x_1 = f_1, \dots, x_r = f_r, x_{r+1} = u_{r+1}, \dots, x_n = u_n$$

forment un système de coordonnées locales en a. Dans le système (x) et (v) l'application f s'exprime par $f(x) = \big(x_1, \dots, x_r, \phi_{r+1}(x), \dots, \phi_p(x)\big)$. Comme le rang de sa différentielle est exactement r, cela impose aux fonctions $\phi_{r+1}, \dots, \phi_p$ de ne pas dépendre de $x_{r+1}, \dots, x_n$ (voyez-vous pourquoi ?). La formule

$$y_1 = v_1, \dots, y_r = v_r, y_{r+1} = v_{r+1} - \phi_{r+1}(v_1, \dots, v_r), \dots, y_p = v_p - \phi_p(v_1, \dots, v_r)$$

définit un système de coordonnées locales au voisinag ede 0 (on calcule explicitement v en fonction de y) dans lequel f s'exprime sous la forme voulue. $\triangleright$

Ce théorème est démontré dans le livre de Rouvière, exercice 73.

c) Réduction des formes quadratiques et lemme de Morse. Soit $\mathcal{S}$ l'ensemble des matrices symétriques réelles et A_0 une matrice symétrique inversible. L'énoncé suivant signifie que tout un voisinage de A_0 dans $\mathcal{S}$ est forméde matrices similaires à A_0. Notons $E = \{M \in \mathrm{L}(\mathbb{R}^n)\,;\, A_0M \in \mathcal{S}\}$.

LEMME **13.** *L'application $\phi : E \to \mathcal{S}, M \mapsto M^*A_0M$, est un difféomorphisme local entre voisinage de l'identité dans E et un voisinage de A_0 dans $\mathcal{S}$.*

DÉMONSTRATION – Il suffit de voir que la différentielle $\big(D_{\mathrm{Id}}\phi\big)(H) = (A_0H)^* + (A_0H) = 2A_0H$ de ϕ est bijective. Elle est injective du fait de l'inversibilité de A_0, et partant bijective comme E et $\mathcal{S}$ sont deux espaces vectoriels de même dimension. ▷

L'énoncé suivant du à Morse exprime le fait qu'au voisinage d'un point critique a d'une fonction f non seulement la fonction s'exprime-t-elle dans les coordonnées données sous la forme $f(x) = f(a) + H(x) + o(x^2)$, où H est une forme quadratique, mais elle s'écrit exactement $f(x) = f(a) + H(x)$ dans un système de coordonnées bien choisi.

THÉORÈME **14** (Lemme de Morse)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathcal{C}^3$ de différentielle nulle en 0 et de différentielle seconde $D_0^2 f$ non dégénérée. Notons q_0 la forme quadratique associée à $D_0^2 f$. Il existe un $\mathcal{C}^1$-difféomorphisme φ d'un voisinage U de 0 sur $\varphi(U)$ tel que*

$$f(x) = f(0) + q_0\big(\varphi(x)\big)$$

pour tout $x \in U$.

DÉMONSTRATION – On peut écrire $f(x) = f(0) + \int_0^1 (1-t) D^2_{tx} f(x,x)dt$, d'après la formule de Taylor avec reste intégral ; la forme quadratique $q_x := \int_0^1 (1-t) D^2_{tx} f(\cdot,\cdot)dt$ dépend de manière $\mathcal{C}^1$ de x. Notons A_x sa matrice et A_0 la matrice de la forme quadratique $D_0^2 f = q_0$. La matrice A_x s'écrit d'après le lemme 13 sous la forme $A_x = \big(\phi^{-1}(A_x)\big)^* A_0 \big(\phi^{-1}(A_x)\big)$. On définit une application φ de classe $\mathcal{C}^1$ en posant $\varphi(x) = \big(\phi^{-1}(A_x)\big)(x)$; elle vérifie $f(x) = f(0) + q_0\big(\varphi(x)\big)$. On calcule la différentielle de φ pour voir qu'il s'agit d'un difféomorphisme local. Comme $\varphi(x) = (D_0\varphi)(x) + o(x)$ on a d'une part $f(x) = f(0) + q_0(x) + o(x^2)$ et d'autre part $f(x) = f(0) + q_0\big((D_0\varphi)(x)\big) + o(x^2)$, dont on déduit que $D_0\varphi$ définit une isométrie de la forme quadratique définie positive q_0 ; elle est inversible à ce titre, comme son déterminant ne peut être nul. ▷

Pour une référence, voir les exercices 66 et 105-110 du livre de Rouvière.

d) Théorème de von Neumann sur les sous-groupes continus de $\mathrm{GL}(\mathbb{R}^n)$. On démontre dans ce paragraphe le théorème 3 selon lequel tout sous-groupe fermé non discret de $\mathrm{GL}(\mathbb{R}^n)$ est une sous-variété de $\mathrm{L}(\mathbb{R}^n)$.

On exhibe pour cela un certain sous-espace vectoriel $\mathrm{LIE}(G)$ de $\mathrm{L}(\mathbb{R}^n)$ dont on note E un supplémentaire quelconque, et tel que l'application $\phi : \mathrm{LIE}(G) \times E \to \mathrm{GL}(\mathbb{R}^n), (A, X) \mapsto \exp(A)\exp(X)$, définit un difféomorphisme d'un voisinage $\mathcal{V}$ de 0 dans $\mathrm{LIE}(G) \times E$ sur un voisinage $\mathcal{U}$ de l'identité dans $\mathrm{GL}(\mathbb{R}^n)$, avec $G \cap \mathcal{U} = \phi\big(\mathrm{LIE}(G) \cap \mathcal{V}\big)$. Cela montre que G est une variété au voisinage de l'identité ; on voit que c'est aussi le cas au voisinage de chacun de ses points $M \in G$ en considérant l'application $M\phi(\cdot)$. Rappelons qu'on note $[A, B] = AB - BA$ le crochet de A et B.

LEMME **15.** *Quels que soient $A, B \in L(\mathbb{R}^n)$ on a les limites suivantes lorsque n tend vers ∞*

$$\big(\exp(A/n)\exp(B/n)\big)^n \to \exp(A+B),$$

$$\big(\exp(A/n)\exp(B/n)\exp(-A/n)\exp(-B/n)\big)^{n^2} \to \exp\big([A,B]\big).$$

DÉMONSTRATION – Il suffit de prendre le logarithmes de chaque côté et de faire un développement limité. $\triangleright$

Posons

$$\mathrm{LIE}(G) = \{A \in \mathrm{L}(\mathbb{R}^n)\,;\, \exp(tA) \in G, \text{ pour tous } t \in \mathbb{R}\}.$$

Les formules du lemme montrent que $\mathrm{LIE}(G)$ est un sous-espace vectoriel de $\mathrm{L}(\mathbb{R}^n)$ stable par l'opération de crochet $[\cdot,\cdot]$. On voit que l'application ϕ définie plus haut est un difféomorphisme local en remarquant que sa diffé]rentielle en 0 est l'identité. On montre que $\phi\big(\mathrm{LIE}(G) \cap \mathcal{V}\big)$ est un voisinage de l'identité dans G en raisonnant par l'absurde. Dans le cas contraire, il existerait une suite $(g_n)_{n\geqslant 1}$ d'éléments de G dont chacun s'écrirait $\exp(A_n)\exp(X_n)$, avec A_n et X_n tendant vers 0 et X_n non nul ; le fait que G n'est pas discret sert ici. Comme $\exp(A_n) \in G$ par définition, on devrait avoir $\exp(X_n) \in G$ pour tout $n \geqslant 1$. Quitte à extraire une sous-suite, notons $X \in E$ la limite non nulle de la suite $\frac{X_n}{|X_n|}$. On va voir que cet élement non nul de E devrait aussi appartenir à $\mathrm{LIE}(G)$, ce qui constituerait une contradiction puisque ces deux espaces sont supplémentaires.

Il suffit pour cela de remarquer que puisque X_n tend vers 0 on peut associer à tout $t > 0$ une suite $(m_n)_{n\geqslant 1}$ d'entiers telle que $m_n|X_n|$ tend vers t. Le groupe G étant fermé, la limite $\exp(tX)$ des éléments $\exp(m_nX_n) = \big(\exp(X_n)\big)^{m_n}$ de G devrait aussi appartenir à G ; son inverse $\exp(-tX)$ aussi, ce qui signifie que X appartient à $\mathrm{LIE}(G)$.

Vous trouverez ce théorème dans le livre de Mneimé et Testard, ainsi que dans le livre de Gonnord et Tosel sur le calcul différentiel, chez Ellipses.

e) Diffeomorphismes de $\mathbb{R}^n$. On donne dans ce paragraphe un énoncé caractérisant les difféomorphismes globaux de $\mathbb{R}^n$ de classe $\mathcal{C}^2$. Rappelons qu'une application de $\mathbb{R}^n$ dans lui-même est dit propre si f envoie l'infini à l'infini : $\big|f(x)\big| \to \infty$ lorsque $|x| \to \infty$. C'est en particulier le cas s'il existe une constante positive k telle que $\big|f(x) - f(y)\big| \geqslant k|x-y|$, pour tous x, y en dehors d'un compact.

THÉORÈME **16** (Hadamard)**.** *Une application $f : \mathbb{R}^n \to \mathbb{R}^n$ de classe $\mathcal{C}^2$ est un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}^n$ si, et seulement si, f est propre et sa différentielle est inversible en chaque point.*

DÉMONSTRATION – Seule l'implication $\Leftarrow$ pose problème. On va pour la démontrer construire une application *surjective* $s : \mathbb{R}^n \to \mathbb{R}^n$ telle que

$$f \circ s = \mathrm{Id}.$$

L'existence d'une telle fonction prouve à la fois que f est surjective (parce que s l'est !) et injective, parce que si $f(x) = f(y)$ avec $x = s(a)$ et $y = s(b)$ alors $a = f\big(s(a)\big) = f\big(s(b)\big) = b$ et $x = y$. En fait, on va chercher une famille $(s_t)_{0\leqslant t\leqslant 1}$ d'applications continues telles que

$$f \circ s_t = t\mathrm{Id}.$$

Que gagne-t-on à faire cela ? Etant donné $a \in \mathbb{R}^n$, dériver la relation $f\big(s_t(a)\big) = ta$ donne $\big(D_{s_t(a)}f\big)\big(s_t'(a)\big) = a$. Comme D_xf est partout inversible, cette relation impose à $s_t(a)$ d'être solution de l'équation différentielle

$$s_t'(a) = \big(D_{s_t(a)}f\big)^{-1}(a), \quad s_0(a) = 0.$$

Puisqu'on a supposé f de classe $\mathcal{C}^2$, sa différentielle est de classe $\mathcal{C}^1$ et l'équation dépendant linéairement de a, elle possède une unique solution maximale $(t\,;\,a) \mapsto s_t(a)$, définie sur un ouvert $\bigcup_{a\in\mathbb{R}^n}\big(T_*(a), T^*(a)\big) \times \{a\} \subset \mathbb{R} \times \mathbb{R}^n$, où elle est de classe $\mathcal{C}^1$. Tant que cela a un sens, on a $f\big(s_t(a)\big) = ta$. Un point a étant fixé, si

$s_t(a)$ avait un temps de vie futur $T^*(a)$ fini, $s_t(a)$ devrait tendre vers l'infini lorsque t tend vers $T^*(a)$, impliquant la convergence de $f\big(s_t(a)\big)$ vers l'infini comme f est propre. Cela contredirait la relation $f\big(s_t(a)\big) = ta$, dans laquelle ta reste bornée pour $t \leqslant T^*(a) < \infty$. Chaque fonction $s_1(a)$ est donc bien définie, et continue par rappport à a. Mais est-elle surjective ? L'application Puisque f est un difféomorphisme local d'après le théorème d'inversion locale, et s_1 l'inverse de f, l'application s_1 est aussi un difféomorphisme local ; elle est en particulier ouverte. On montre qu'elle est surjective en montrant que son image est fermée dans $\mathbb{R}^n$; la connexité de $\mathbb{R}^n$ permet de conclure. Mais supposant $s_1(a_n) \to x$, on a $a_n = f\big(s_1(a_n)\big) \to f(x)$ et $s_1(a_n) \to s_1\big(f(x)\big)$, puisque s_1 est continue, ce dont on conclut que $x = s_1\big(f(x)\big)$ est dans l'image de s_1. $\triangleright$

Vous trouverez une référence pour cet énoncé dans le Zuily-Queffelec et probablement dans le livre de Gonnord et Tosel. La méthode utilisée s'appelle méthode du chemin. Vous trouverez dans le livre de Chaperon, "*Calcul différentiel et calcul intégral*", chez Dunod, p. 314, d'autres illustrations de cette méthode.

Pour d'autres applications, on pourra aussi penser au théorème fondamental de l'algèbre, affirmant l'existence dans le plan complexe de zéros à tout polynôme (Chambert-Loire), au théorème de redressement des champs de vecteurs, ou à la régularité des racines simples d'un polynôme par rapport à ses coefficients (Chambert-Loire).