Feuille d'exercices n^{o}12

Corrigé

Exercice 1

1. a) On peut prendre $\phi(a,b) = (a,b)$ et $\psi(a,b) = (a, b-a^2)$.
b) Pour tout $(x,y) \in \mathbb{R}^2$, $df(x,y).(h,l) = (h, 2yl)$. On a donc $\text{rang}(df(x,y)) = 2$ si $y \neq 0$ et $\text{rang}(df(x,y)) = 1$ si $y = 0$.
Supposons qu'il existe ψ, ϕ des difféomorphismes au voisinage de $(0,0)$ tels que $\psi \circ f \circ \phi(a,b) = (a,0)$ pour tous (a,b) assez proches de $(0,0)$. Puisque l'application $(a,b) \to (a,0)$ a une différentielle de rang 1 en tout point, f doit aussi avoir une différentielle de rang 1 en tout point assez proche de 0. Or on a vu que ce n'était pas le cas.
c) Soit $U = \{x \in V \text{ tq } \text{rang}(df) \text{ admet un maximum local en } x\}$.

Lemme 1.1. *U est dense dans V.*

Démonstration. Soit W un ouvert de V quelconque. L'application $x \to \text{rang}(df(x))$ ne prend qu'un nombre fini de valeurs sur W. Elle admet donc un maximum sur W. Si x_0 est un point où le maximum est atteint, $x_0 \in U$. □

Lemme 1.2. *U est ouvert et df est de rang localement constant sur U.*

Démonstration. Soit x_0 un point où $\text{rang}(df)$ est localement maximal. Soit r le rang en question. Alors, si, pour tout x, on note $M(x)$ la matrice associée à $df(x)$ dans les bases canoniques, il existe une sous-matrice de $M(x_0)$, de taille $r \times r$, dont le déterminant n'est pas nul. Notons $M_{I,J}(x_0)$ cette sous-matrice.
Pour x assez proche de x_0, $\det(M_{I,J}(x)) \neq 0$ (par continuité du déterminant) donc $\text{rang}(M(x)) = \text{rang}(df(x)) \geq r$. Puisque r est le rang maximal de df au voisinage de x_0, on a exactement, pour x assez proche de x_0, $\text{rang}(df(x)) = r$.
Il existe donc un voisinage $W \subset V$ de x_0 sur lequel le rang de df est constant égal à r. Ce voisinage W est inclus dans U.
Donc, pour tout x_0, U contient un voisinage de x_0 et on peut choisir le voisinage de telle sorte que df y soit de rang constant. □

Montrons maintenant que, sur U, df est injective. Soit $x \in U$ quelconque.
Puisque df est de rang localement constant au voisinage de x, il existe (d'après le théorème de forme normale), des $\mathcal{C}^1$-difféomorphismes ϕ et ψ, définis au voisinage de x et $f(x)$, tels que $\phi(x) = x, \psi(f(x)) = f(x)$ et, pour tout $y = (y_1, ..., y_n)$ assez proche de x :

$$\psi \circ f \circ \phi(y_1, ..., y_n) = (y_1, ..., y_r, 0, ..., 0) \in \mathbb{R}^m$$

où r est le rang de $df(x)$.

On a nécessairement $r = n$, sinon f n'est pas injective : $f \circ \phi(x + (0, ..., 0, t)) = f \circ \phi(x)$ pour tout t assez petit.
Donc $df(x)$ est une application de rang n de $\mathbb{R}^n$ vers $\mathbb{R}^m$. Cela implique que $df(x)$ est injective et que $m \geq n$.

2. a) $dg(x_0, y_0).(h, l) = (h, \partial_x f(x_0, y_0).h + \partial_y f(x_0, y_0).l)$.
Cette application admet pour réciproque $(h, l) \to (h, (\partial_y f(x_0, y_0))^{-1}.(l - \partial_x f.h))$. C'est donc un isomorphisme.
D'après le théorème d'inversion locale, g est donc un $\mathcal{C}^1$-difféomorphisme au voisinage de (x_0, y_0).
b) Pour tout (x, z) dans l'ensemble de définition de γ, on doit avoir $(x, z) = g(\gamma(x, z)) = (\gamma_1(x, z), f(\gamma_1(x, z), \gamma_2(x, z)))$. On doit donc avoir $\gamma_1(x, z) = x$ et $z = f(\gamma_1(x, z), \gamma_2(x, z)) = g(x, \gamma_2(x, z))$.
c) Pour tout (x, y) assez proche de 0, il existe un unique couple $(x', z) \in E \times G$ tel que $(x, y) = \gamma(x', z) = (x', \gamma_2(x', z))$. On a alors $x' = x$ et $f(x, y) = f(x, \gamma_2(x, z)) = z$. Donc $f(x, y) = 0$ si et seulement si $z = 0$, ce qui implique $y = \gamma_2(x, 0)$.
Réciproquement, si $y = \gamma_2(x, 0)$, on a bien $f(x, y) = f(x, \gamma_2(x, 0)) = 0$.

3. Posons $A = P^{-1}(P(S))$.
Sur $\mathbb{C} - A$, P est un $\mathcal{C}^\infty$-difféomorphisme local, d'après le théorème d'inversion locale. En effet, sa différentielle est inversible en tous les points de $\mathbb{C} - A$ car $\mathbb{C} - A \subset \mathbb{C} - S$. L'image par P de $\mathbb{C} - A$ est donc un ouvert de $\mathbb{C} - P(S)$.
De plus, l'image de $\mathbb{C} - A$ par P est aussi un fermé de $\mathbb{C} - P(S)$: si $P(x_n) \to z \in \mathbb{C} - P(S)$, la suite $(x_n)_{n \in \mathbb{N}}$ est bornée (car $|P(x)| \to +\infty$ quand $|x| \to +\infty$) donc, quitte à extraire, on peut supposer qu'elle converge vers une limite x_∞. Alors $z = P(x_\infty)$ et, puisque $z \in \mathbb{C} - P(S)$, $x_\infty \in \mathbb{C} - A$ donc $z \in P(\mathbb{C} - A)$.
Puisque $\mathbb{C} - P(S)$ est connexe, l'image de $\mathbb{C} - A$ par P, qui est un ouvert et fermé non-vide, est égale à $\mathbb{C} - P(S)$. Donc $\mathbb{C} - P(S) \subset P(\mathbb{C})$. Puisque $P(S)$ est aussi inclus dans $P(\mathbb{C})$, $\mathbb{C} \subset P(\mathbb{C})$. Donc $P(\mathbb{C}) = \mathbb{C}$.

4. On munit $\mathbb{R}^n$ de la distance euclidienne. Pour tout $\eta > 0$ et tout fermé $F \subset \mathbb{R}^n$, on note $F_\eta = \{x \in \mathbb{R}^n \text{ tq } d(x, F) \leq \eta\}$.

Lemme 1.3. *Soit $F \subset \mathbb{R}^n$ un fermé. Pour tout $\epsilon > 0$, il existe $f_\epsilon \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ telle que $F \subset f^{-1}(\{0\}) \subset F_\epsilon$.*

Démonstration. Soit $\chi : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathcal{C}^\infty$ non-nulle, à support compact inclus dans $[-1; 1]^n$. Une telle fonction existe : dans le TD précédent, par exemple, on en a construit une pour $n = 1$. Si on note ψ cette fonction correspondant à $n = 1$, la fonction $\chi(x_1, ..., x_n) = \psi(x_1)...\psi(x_n)$ convient.
Quitte à élever χ au carré, on peut supposer que χ est positive. Quitte à la diviser par sa norme 1, on peut aussi supposer que $\int_{\mathbb{R}^n} \chi = 1$.
Posons $\phi(x) = \frac{2^n n^{n/2}}{\epsilon^n} \chi\left(\frac{2\sqrt{n}x}{\epsilon}\right)$. C'est une fonction de classe $\mathcal{C}^\infty$, d'intégrale 1 (cela se vérifie par le calcul, en faisant un changement de variables) et à support dans $\left[-\frac{\epsilon}{2\sqrt{n}}; \frac{\epsilon}{2\sqrt{n}}\right]^n \subset \overline{B}(0, \epsilon/2)$.
Notons maintenant g la fonction qui vaut 0 sur $F_{\epsilon/2}$ et 1 sur le complémentaire. Posons $f_\epsilon = g \star \phi$. C'est une fonction de classe $\mathcal{C}^\infty$ car ϕ l'est.

De plus, pour tout $x \in F$, $\overline{B}(x,\epsilon/2) \subset F_{\epsilon/2}$ donc $f_\epsilon(x) = \int_{\mathbb{R}^n} g(y)\phi(x-y)d^ny = 0$ (en effet, soit $x-y \notin \overline{B}(0,\epsilon/2)$ et alors $\phi(x-y)=0$, soit $y \in \overline{B}(x,\epsilon/2)$ et alors $g(y)=0$).
Pour tout $x \notin F_\epsilon$, $g(y)=1$ sur $\overline{B}(x,\epsilon/2)$ (car $\overline{B}(x,\epsilon/2) \cap F_{\epsilon/2} = \emptyset$) donc $f_\epsilon(x) = \int_{\mathbb{R}^n} \phi(x-y)d^ny = 1$. En particulier, $f_\epsilon(x) \neq 0$.
Donc $F \subset f_\epsilon^{-1}(\{0\}) \subset F_\epsilon$. □

Pour tout $n \in \mathbb{N}^*$, soit donc $f_{1/n}$ une fonction comme dans le lemme précédent. Quitte à l'élever au carré, on peut supposer qu'elle est positive. Nous allons choisir convenablement une suite $(\alpha_n)_{n\in\mathbb{N}}$ de réels positifs et définir $g = \sum\limits_n \alpha_n f_{1/n}$.
Si la suite a été choisie de telle sorte que la série définissant g converge uniformément sur tout compact (en fait, pour l'affirmation qui suit, la convergence simple suffit), alors g s'annule exactement sur l'ensemble $F' = \bigcap\limits_n f_{1/n}^{-1}(\{0\})$.
D'après la définition des $f_{1/n}$, $F \subset F' \subset \bigcap\limits_n F_{1/n} = F$ donc $F' = F$ et $F = g^{-1}(\{0\})$.
Il faut donc montrer que, si on choisit bien la suite $(\alpha_n)_{n\in\mathbb{N}}$, la série définissant g converge uniformément sur tout compact et définit une fonction de classe $\mathcal{C}^\infty$.
Pour tout n, notons $A_n = \sup\limits_{k,l\leq n, ||x||\leq n} |f_{1/k}^{(l)}(x)|$ et prenons $\alpha_n = \frac{1}{2^n A_n}$.
Montrons que, avec cette définition, $\sum\limits_n \alpha_n f_{1/n}^{(l)}$ converge uniformément sur tout compact, pour tout $l \in \mathbb{N}$. Fixons donc $l \in \mathbb{N}$ et K un compact quelconques.
Soit $n_0 \in \mathbb{N}^*$ tel que $K \subset B(0,n_0)$. Pour tout $n \geq \max(n_0,l)$, $\sup\limits_{x\in K} |\alpha_n f_{1/n}^{(l)}(x)| \leq \alpha_n A_n = 2^{-n}$.
La série converge donc normalement sur K. Elle y converge donc uniformément.
La série définissant g converge donc uniformément sur tout compact, ainsi que les séries des dérivées. La fonction g est donc bien définie et de classe $\mathcal{C}^\infty$.

Exercice 2

1. a) Si on identifie $\mathcal{M}_n(\mathbb{R})$ et $\mathbb{R}^{n^2}$, chacune des coordonnées de ϕ est une application polynomiale. L'application ϕ est donc $\mathcal{C}^\infty$.
$\phi(\mathrm{Id}+H) = A_0 + {}^tHA_0 + A_0H + {}^tHA_0H = \phi(\mathrm{Id}) + {}^tHA_0 + A_0H + o(H)$
La différentielle de ϕ en Id vaut donc $d\phi(\mathrm{Id}).H = {}^tHA_0 + A_0H$.
b) Pour tout $S \in \mathcal{S}_n(\mathbb{R})$, $d\phi(\mathrm{Id}).(A_0^{-1}S/2) = S$ donc $d\phi(\mathrm{Id})$ est surjective.
c) L'application ϕ est une submersion en Id.
D'après le théorème de forme normale des submersions, il existe $\chi : \mathrm{Ker}\,(d\phi(\mathrm{Id})) \times \mathcal{S}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$ un difféomorphisme local, défini au voisinage de $(0,0)$, tel que :

$$\chi(0,0) = \mathrm{Id}$$

et :

$$\forall (X,A) \text{ assez proche de } (0,0), \quad \phi \circ \chi(X,A) = \phi(\mathrm{Id}) + A = A_0 + A$$

Posons $P(A) = \chi(0, A-A_0)$ pour tout A assez proche de A_0. Cette application est bien de classe $\mathcal{C}^\infty$ et, pour tout A :

$$A = A_0 + (A - A_0) = \phi \circ \chi(0, A-A_0) = \phi(P(A)) = {}^tP(A)A_0P(A)$$

2. a) D'après la formule de Taylor, pour tout $x \in U$, si le segment $[0; x]$ est inclus dans U (ce qui est vrai sur un voisinage V assez petit de 0) :

$$f(x) = f(0) + df(0).x + \frac{1}{2}\int_0^1 d^{(2)}f(tx).(x,x)dt = \frac{1}{2}\int_0^1 d^{(2)}f(tx).(x,x)dt$$

Pour tout $y \in U$, on note $M(y) = (M_{i,j}(y))_{i,j\leq n}$ la matrice qui représente la forme bilinéaire symétrique $d^{(2)}f(y)$ dans la base canonique. Avec cette notation :

$$f(x) = \frac{1}{2}\int_0^1 \left(\sum_{i,j} M_{i,j}(tx)x_i x_j\right) dt = \sum_{i,j}\left(\frac{1}{2}\int_0^1 M_{i,j}(tx)dt\right) x_i x_j$$

Posons, pour tous $i, j \leq n$, $a_{i,j}(x) = \frac{1}{2}\int_0^1 M_{i,j}(tx)dt$. Puisque $M_{i,j}$ est de classe $\mathcal{C}^1$ pour tous i, j, l'application $a_{i,j}$ est également de classe $\mathcal{C}^1$ (on peut différencier sous le signe intégrale).
b) Avec les définitions de la question précédente, $A(x) = (a_{i,j}(x))_{i,j\leq n}$ est une matrice symétrique pour tout $x \in V$ (et si on avait définis autrement les $a_{i,j}$, on pourrait remplacer $a_{i,j}$ par $\frac{a_{i,j}+a_{j,i}}{2}$ et la matrice serait symétrique).
De plus, par continuité des $a_{i,j}$ en 0, on a $f(x) = {}^t x A(0) x + o(||x||^2)$ lorsque $x \to 0$. Donc $A(0)$ est la matrice associée à la forme bilinéaire $d^{(2)}f(0)$ dans la base canonique.
D'après la première question, il existe un voisinage $\mathcal{V}$ de $A(0)$ dans $\mathcal{S}_n(\mathbb{R})$ et une application $P : \mathcal{V} \to \mathcal{M}_n(\mathbb{R})$ de classe $\mathcal{C}^\infty$ telles que $P(A(0)) = \mathrm{Id}$ et, pour tout $A \in \mathcal{V}$, $A = {}^tP(A)A(0)P(A)$.
Quitte à diminuer la taille de V, on peut supposer que $A(x) \in \mathcal{V}$ pour tout $x \in V$. On a alors, pour tout $x \in V$:

$$f(x) = {}^t x A(x) x = {}^t x {}^t P(A(x))A(0)P(A(x))x$$

Puisque $A(0)$ est de signature $(p, n-p)$, il existe $G \in \mathcal{M}_n(\mathbb{R})$ telle que :

$$A(0) = {}^t G \begin{pmatrix} 1 & 0 & \dots & & 0 \\ & \ddots & & & \\ \vdots & & 1 & & \vdots \\ & & & -1 & \\ 0 & & \dots & \ddots & -1 \end{pmatrix} G$$

En notant D la matrice diagonale de l'équation précédente, on a, pour tout $x \in V$:

$$f(x) = {}^t(GP(A(x))x)D(GP(A(x))x)$$

Posons $\psi(x) = GP(A(x))x$. On a $d\psi(0) = GP(A(0))$ donc, d'après le théorème d'inversion locale, il existe deux voisinages V_1, V_2 de 0 tels que ψ réalise un $\mathcal{C}^1$-difféomorphisme de V_1 sur V_2.
Si on pose $\phi = \psi^{-1}$, on a :

$$f(\phi(x)) = {}^t(\psi(\phi(x)))D(\psi(\phi(x))) = {}^t x D x = x_1^1 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_n^2$$

Exercice 3

1. Le cours dit que, lorsqu'une fonction est strictement convexe, le minimum est atteint en un unique point s'il est atteint. Montrons qu'il est atteint.
Soit $x_0 \in \mathbb{R}^n$ quelconque. Soit $R > 0$ tel que, pour tout x tel que $||x|| > R$, $f(x) > f(x_0)$. Soit m le minimum de f sur $\overline{B}(0, R)$. Il est atteint car la convexité de f implique sa continuité. C'est aussi le minimum sur tout $\mathbb{R}^n$.
2. a)

$$\begin{aligned}
f(x_{n+1}) - f(x_n) &= \int_0^1 \langle \nabla f(x_n + t(x_{n+1} - x_n)), x_{n+1} - x_n \rangle dt \\
&= \int_0^1 \langle \nabla f(x_n), x_{n+1} - x_n \rangle dt \\
&\quad + \int_0^1 \langle \nabla f(x_n + t(x_{n+1} - x_n)) - \nabla f(x_n), x_{n+1} - x_n \rangle dt \\
&\leq \int_0^1 \langle \nabla f(x_n), x_{n+1} - x_n \rangle dt + \int_0^1 tL||x_{n+1} - x_n||^2 dt \\
&= \langle \nabla f(x_n), x_{n+1} - x_n \rangle + \frac{L}{2}||x_{n+1} - x_n||^2 \\
&= -\alpha(1 - \frac{L\alpha}{2})||\nabla f(x_n)||^2
\end{aligned}$$

b) La fonction f étant convexe, on a, pour tous y_1, y_2, $f(y_1) - f(y_2) \geq \langle \nabla f(y_2), y_1 - y_2 \rangle$. Pour $y_1 = x^*$ et $y_2 = x_n$, on obtient $f(x_n) \leq f(x^*) + \langle \nabla f(x_n), x_n - x^* \rangle$.

$$\begin{aligned}
f(x_{n+1}) &\leq f(x_n) - ||\nabla f(x_n)||^2 \alpha(1 - \frac{L\alpha}{2}) \\
&\leq f(x_n) - \frac{\alpha}{2}||\nabla f(x_n)||^2 \\
&\leq f(x^*) + \langle \nabla f(x_n), x_n - x^* \rangle - \frac{\alpha}{2}||\nabla f(x_n)||^2 \\
&= f(x^*) + \frac{1}{2\alpha}(2\alpha\langle \nabla f(x_n), x_n - x^* \rangle - \alpha^2||\nabla f(x_n)||^2) \\
&= f(x^*) + \frac{1}{2\alpha}(2\langle x_n - x_{n+1}, x_n - x^* \rangle - ||x_n - x_{n+1}||^2) \\
&= f(x^*) + \frac{1}{2\alpha}(||x_n - x^*||^2 - ||x_{n+1} - x^*||^2)
\end{aligned}$$

c) Puisque, pour tout n, $f(x_{n+1}) - f(x^*) \leq \frac{1}{2\alpha}(||x_n - x^*||^2 - ||x_{n+1} - x^*||^2)$:

$$\sum_{n=1}^{N}(f(x_n) - f(x^*)) \leq \frac{1}{2\alpha}(||x_0 - x^*||^2 - ||x_N - x^*||^2) \leq \frac{||x_0 - x^*||^2}{2\alpha}$$

Puisque $(f(x_n))_n$ est une suite décroissante (d'après la question a)), cela donne, pour tout $N \geq 1$:

$$f(x_N) - f(x^*) \leq \frac{1}{N}\sum_{n=1}^{N}(f(x_n) - f(x^*)) \leq \frac{1}{2N\alpha}||x_0 - x^*||^2$$

3. a) Notons $g(x) = f(x) - \frac{m}{2}||x||^2$. Cette fonction est de classe $\mathcal{C}^1$ (puisque f l'est) et vérifie $\nabla g(x) = \nabla f(x) - mx$.
Puisque g est convexe, on a, pour tous x, y :

$$g(y) - g(x) \geq \langle \nabla f(x) - mx, y - x \rangle = \langle \nabla f(x), y - x \rangle - m\langle x, y - x \rangle$$

On obtient donc :

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle + \frac{m}{2}||y||^2 - \frac{m}{2}||x||^2 - m\langle x, y - x \rangle = \langle \nabla f(x), y - x \rangle + \frac{m}{2}||x - y||^2$$

b) D'après la question a), $\langle \nabla f(x_n), x_n - x^* \rangle \geq f(x_n) - f(x^*) + \frac{m}{2}||x_n - x^*||^2 \geq \frac{m}{2}||x_n - x^*||^2$.
Donc :

$$\begin{aligned} ||x_{n+1} - x^*||^2 &= ||x_n - x^*||^2 - 2\alpha\langle \nabla f(x_n), x_n - x^* \rangle + \alpha^2 ||\nabla f(x_n)||^2 \\ &\leq ||x_n - x^*||^2 - m\alpha||x_n - x^*||^2 + \alpha^2||\nabla f(x_n)||^2 \end{aligned}$$

Puisque ∇f est L-lipschitzienne et $\nabla f(x^*) = 0$, on a, pour tout n, $||\nabla f(x_n)|| \leq L||x_n - x^*||$.
On obtient donc l'inégalité :

$$||x_{n+1} - x^*||^2 \leq (1 - m\alpha + L^2\alpha^2)||x_n - x^*||^2$$

Si α est suffisamment petit pour que $1 - m\alpha + L^2\alpha^2 < 1$, on a, pour tout n :

$$||x_{n+1} - x^*|| \leq c||x_n - x^*||$$

avec $c = \sqrt{1 - m\alpha + L^2\alpha^2}$.
Par récurrence, cela implique bien l'inégalité demandée pour tout n.

Exercice 4

1. a) C'est une conséquence du théorème des fonctions implicites.
La différentielle de g en a est $dg(a) = (dg_1(a), ..., dg_k(a))$. Puisque les formes linéaires $dg_1(a), ..., dg_k(a)$ sont linéairement indépendantes, $\dim \mathrm{Ker}\,(dg(a)) = n - k$.
Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de $\mathbb{R}^n$ telle que $\mathrm{Ker}\,(dg(a)) = \mathrm{Vect}\,\{e_{k+1}, ..., e_n\}$. Alors $dg(a)$ réalise une bijection de $\mathrm{Vect}\,\{e_1, ..., e_k\}$ vers $\mathbb{R}^k$. D'après le théorème des fonctions implicites, il existe donc un voisinage de $(a_{k+1}, ..., a_n)$ et une application ϕ de ce voisinage vers $\mathbb{R}^k$, de classe $\mathcal{C}^1$, telle que, pour tout $x = (x_1, ..., x_n)$ assez proche de a :

$$g(x) = 0 \Leftrightarrow (x_1, ..., x_k) = \phi(x_{k+1}, ..., x_n)$$

b) Puisque $0 = g(\phi(x), x)$ pour tout x assez proche de $(a_{k+1}, ..., a_n)$, on a :

$$0 = dg(\phi(x), x) \circ \begin{pmatrix} d\phi(x) \\ \mathrm{Id} \end{pmatrix}$$

c'est-à-dire, pour tout x et tout $l \in \mathbb{R}^{n-k}$: $0 = dg(\phi(x), x).(d\phi(x).l, l)$.
Pour tout $h \in \mathbb{R}^n$ tel que $(h_1, ..., h_k) = d\phi(\alpha)(h_{k+1}, ..., h_n)$, on a donc, en posant $l = (h_{k+1}, ..., h_n)$, que $dg(a).h = dg(\phi(\alpha), \alpha).(d\phi(l), l) = 0$.

Donc $\{(h_1, ..., h_n) \in \mathbb{R}^n \text{ tq } (h_1, ..., h_k) = d\phi(\alpha)(h_{k+1}, ..., h_n)\}$ est inclus dans $\operatorname{Ker} dg(a)$. Puisque $\operatorname{Ker} dg(a)$ est de dimension $n-k$, on a égalité.

c) Pour tout $x = (x_{k+1}, ..., x_n)$ assez proche de α, $(\phi(x), x)$ appartient à Z. Le point α est donc un extremum local de $x \to F(\phi(x), x)$, puisque a est un extremum local de F sur Z.

La différentielle de cette fonction en α doit donc être nulle : pour tout $h \in \mathbb{R}^{n-k}$, $dF(a).(d\phi(\alpha).h, h) = 0$. Donc $\operatorname{Ker} dg(a) \subset \operatorname{Ker} dF(a)$.

D'après le théorème fondamental de l'algèbre linéaire (?), $\operatorname{Vect}\{dg_1(a), ..., dg_k(a)\} = \{\phi \in (\mathbb{R}^n)' \text{ tq } \bigcap_{i \le k} dg_i(a) \subset \operatorname{Ker} \phi\}$.

D'après la définition de g, $\operatorname{Ker} dg(a) = \bigcap_{i \le k} dg_i(a)$. D'après ce qu'on vient de voir, on a donc $dF(a) \in \operatorname{Vect}\{dg_1(a), ..., dg_k(a)\}$ donc $dF(a)$ est une combinaison linéaire des $dg_i(a)$.

2. Puisque a est aussi un extremum local de F sur $Z' = \{x \in U \text{ tq } g_1(x) = ... = g_k(x) = 0\} \subset Z$, il existe, d'après la première partie, des réels $\lambda_1, ..., \lambda_k$ tels que :

$$dF(a) = \sum_i \lambda_i dg_i(a)$$

Il faut montrer que les λ_i sont positifs.

Pour cela, nous allons imiter le raisonnement de la première question. Dans la première question, nous avions localement paramétrisé Z en écrivant :

$$Z \cap V = \{(\phi(x_{k+1}, ..., x_n), x_{k+1}, ..., x_n) \text{ tq } x_{k+1}, ..., x_n \text{ dans un voisinage de } \alpha\}$$

Pour la nouvelle définition de Z, nous allons obtenir une paramétrisation de la forme :

$$Z \cap W = \{(\zeta(y_1, ..., y_k, x_{k+1}, ..., x_n), x_{k+1}, ..., x_n) \text{ tq } \\ x_{k+1}, ..., x_n \in \text{Vois}(\alpha) \text{ et } y_1, ..., y_k \in \text{Vois}(0) \cap (\mathbb{R}^+)^k\}$$

où ζ est une application de classe $\mathcal{C}^1$ telle que $g(\zeta(y_1, ..., y_k, x_{k+1}, ..., x_n), x_{k+1}, ..., x_n) = (y_1, ..., y_k)$.

Considérons, comme dans la question 1.a), une base $\mathcal{B} = (e_1, ..., e_n)$ de $\mathbb{R}^n$ telle que $\operatorname{Ker}(dg(a)) = \operatorname{Vect}\{e_{k+1}, ..., e_n\}$ et donc telle que $dg(a)$ réalise une bijection de $\operatorname{Vect}\{e_1, ..., e_k\}$ vers $\mathbb{R}^k$.

L'application $\psi : x \to (g(x), x_{k+1}, ..., x_n)$ est bien définie au voisinage de a et de différentielle inversible. D'après le théorème d'inversion locale, il existe donc un voisinage W de a et un voisinage W' de $(0, ..., 0, a_{k+1}, ..., a_n)$ tels que ψ réalise un $\mathcal{C}^1$-difféomorphisme de W vers W'.

Pour tout $x \in W$, $\psi(x)$ et x ont les mêmes $n-k$ dernières coordonnées. L'application ψ^{-1} est donc de la forme $z = (y_1, ..., y_k, x_{k+1}, ..., x_n) \to (\zeta(z), x_{k+1}, ..., x_n)$. Soit ζ l'application qui vérifie cette relation.

Puisque $\psi(\zeta(y_1, ..., x_n), x_{k+1}, ..., x_n) = \psi \circ \psi^{-1}(y_1, ..., x_n) = (y_1, ..., x_n)$, on a, par définition de ψ, $g(\zeta(y_1, ..., x_n), x_{k+1}, ..., x_n) = (y_1, ..., y_k)$.

Donc, pour tout $x \in W$, $x \in Z$ si et seulement si x est de la forme $x = (\zeta(y_1, ..., x_n), x_{k+1}, ..., x_n)$ avec $y_1, ..., y_k \ge 0$. On a donc obtenu la paramétrisation de Z annoncée.

Pour tout $t > 0$ assez petit, notons $z(t) = (\zeta(t, 0, ..., 0, a_{k+1}, ..., a_n), a_{k+1}, ..., a_n)$. D'après ce qui précède, c'est un élément de Z pour t assez petit. Donc $F(\zeta(t, 0, ..., 0, a_{k+1}, ..., a_n), a_{k+1}, ..., a_n) \ge F(a)$. En dérivant, on obtient donc $dF(a).(d_1\zeta(0, ..., a_n), 0, ..., 0) \ge 0$ (où d_1 désigne la différentielle par rapport à la première variable).

En outre, puisque $g(\zeta(t,0,...,0,a_{k+1},...,a_n),a_{k+1},...,a_n) = (t,0,...,0)$ pour tout t assez petit, $dg(a).(d_1\zeta(0,...,a_n),0,...,0) = (1,0,...,0)$, c'est-à-dire que $dg_i(a).(d_1\zeta(0,...,a_n),0,...,0) = \delta_{1i}$ pour tout $i = 1,...,k$.
Comme $dF(a) = \sum_i \lambda_i dg_i(a)$, on a $0 \leq dF(a).(d_1\zeta(0,...,a_n),0,...,0) = \lambda_1$. Donc λ_1 est positif.
De la même façon, $\lambda_i \geq 0$ pour tout $i = 2,...,k$.

Exercice 5

1. Si (1) est vraie, alors $df(x)$ est inversible pour tout x car $df^{-1}(f(x)) \circ df(x) = \mathrm{Id}$ pour tout x.
De plus, pour tout K compact, $f^{-1}(K)$ est l'image du compact K par l'application continue f^{-1} donc est compacte. Donc f est propre.
On a donc montré que (1) impliquait (2).
Supposons maintenant que (2) est vérifiée. Alors son image est un ouvert de $\mathbb{R}^n$. En effet, puisque $df(x)$ est inversible pour tout x, f est un difféomorphisme local, d'après le théorème d'inversion locale.
Le fait que f est propre implique que l'image de f est fermée. En effet, si $(f(x_n))_{n\in\mathbb{N}}$ converge vers y, la suite $(x_n)_{n\in\mathbb{N}}$ est bornée : $\{y\} \cup \{f(x_n)\}_{n\in\mathbb{N}}$ est un compact donc son antécédant par f est également un compact, donc un ensemble borné. On peut donc, quitte à extraire, supposer que $(x_n)_{n\in\mathbb{N}}$ converge vers une limite x_∞. Par continuité de f, $y = f(x_\infty)$.

2. Il s'agit d'un ensemble compact, puisque f est propre et $\{z\}$ est compact.
De plus, tous les points de S sont isolés car f est un difféomorphisme local (d'après le théorème d'inversion locale).
Cela implique que S est fini. En effet, si ce n'est pas le cas, on peut trouver une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de S tous différents. Cette suite n'admet pas de sous-suite convergente dans S (puisqu'une suite non-stationnaire ne peut pas converger vers un point isolé d'un ensemble), ce qui est en contradiction avec la compacité de S.

3. Le système admet une unique solution maximale puisque $x \to -(df(x))^{-1}.(f(x) - f(z))$ est de classe $\mathcal{C}^1$. Soit $]a;b[$ son intervalle de définition.
Montrons que $b = +\infty$.
Remarquons que $f(x)' = df(x).\dot{x} = f(z) - f(x)$. Donc $t \to f(x(t))$ est solution d'une équation différentielle du premier ordre qu'on sait résoudre.
Puisque $f(x(0)) = f(x_0)$, on doit avoir, pour tout $t \in [0;b[$:

$$f(x(t)) = f(z) + (f(x_0) - f(z))e^{-t}$$

Comme $f(x)$ est bornée sur $[0;b[$, x est bornée aussi (puisque f est propre) et $\dot{x}$ aussi. Donc si $b \neq +\infty$, on peut prolonger x en b (en posant $x(b) = x(0) + \int_0^b \dot{x}(t)dt$), ce qui est en contradiction avec le fait que la solution maximale est définie sur un ouvert.
Donc $b = +\infty$.

Lemme 5.1. *Soit $z_0 \in S$. Il existe $\epsilon > 0$ tel que, si $||x_0 - z_0|| < \epsilon$, alors la solution x de l'équation différentielle converge vers z_0 en $+\infty$.*

Démonstration. Soit $r > 0$ tel que $\overline{B}(z_0, r)$ ne contienne pas d'autre élément de S que z_0.
Soit m le minimum de $||f(x) - f(z_0)||$ sur l'ensemble des x tels que $||x - z_0|| = r$.
Soit $\epsilon \in]0; r[$ tel que, pour tout $x_0 \in B(z_0, \epsilon)$, $||f(x_0) - f(z_0)|| < m$. Alors, pour tout $x_0 \in B(z_0, \epsilon)$, la solution x vérifie, pour tout $t \geq 0$:

$$||f(x(t)) - f(z)|| = ||f(x_0) - f(z)||e^{-t} \leq ||f(x_0) - f(z)|| < m$$

Donc on n'a jamais $||x(t) - z_0|| = r$. Cela implique que $x(t) \in B(z_0, r)$ pour tout $t \geq 0$.
Comme $f(x(t)) \to f(z)$ lorsque $t \to +\infty$ et comme z_0 est le seul point de $\overline{B}(z_0, r)$ où $f(z_0) = f(z)$, cela implique que $x(t) \to z_0$. En effet, si V est un voisinage ouvert de z_0, il existe $\alpha > 0$ tel que, pour tout $x \in \overline{B}(z_0, r) - V$, $||f(z_0) - x|| \geq \alpha$. Puisque $||f(z_0) - x(t)|| < \alpha$ pour tout t assez grand, $x(t) \in V$ pour tout t assez grand. □

Corollaire 5.2. *Quel que soit $x_0 \in \mathbb{R}^n$, la solution x de l'équation différentielle pour la condition initiale x_0 converge vers un élément de S en $+\infty$.*

Démonstration. Puisque x est bornée en $+\infty$, $\{x(t) \text{ tq } t \geq 0\}$ admet une valeur d'adhérence. Comme $f(x(t)) \to f(z)$, cette valeur d'adhérence appartient à S.
Notons z_0 cette valeur d'adhérence et soit $\epsilon > 0$ comme au lemme précédent. Soit $t_0 > 0$ tel que $x(t_0) \in B(z_0, \epsilon)$. Alors, d'après le lemme précédent, l'application $t \to x(t_0 + t)$, qui est solution du système différentiel considéré pour la condition initiale $x(t_0 + 0) = x(t_0)$, converge vers z_0. Donc x converge vers z_0. □

Pour tout $z_0 \in S$, on note $A(z_0)$ l'ensemble des x_0 tels que la solution x de l'équation différentielle considérée avec condition initiale x_0 converge vers z_0.
D'après le corollaire précédent, $\mathbb{R}^n = \bigcup_{z_0 \in S} A(z_0)$.
Or, pour tout $z_0 \in S$, $A(z_0)$ est ouvert. En effet, si $x_0 \in A(z_0)$, soit $\epsilon > 0$ comme dans le lemme. Il existe $t_0 > 0$ tel que $x(t_0) \in B(z_0, \epsilon)$ (où x est la solution de l'équation pour la condition initiale x_0). Puisque les solutions de l'équation différentielles varient continument en fonction des conditions initiales, si x_1 est assez proche de x_0, alors $x^1(t_0)$ appartient aussi à $B(z_0, \epsilon)$ (où x^1 est la solution pour condition initiale x_1 au lieu de x_0). Donc, d'après le lemme, x^1 converge vers z_0 en $+\infty$.
Donc $\mathbb{R}^n$ est l'union disjointe des ouverts $A(z_0)$, où l'ensemble des z_0 varie dans S et où chaque $A(z_0)$ est non-vide (car il contient un voisinage de z_0). Puisque $\mathbb{R}^n$ est connexe, cela implique que S ne contient qu'un seul élément.

On a montré que, pour tout z, $\{x \in \mathbb{R}^n \text{ tq } f(x) = f(z)\}$ est un singleton. Cela revient à dire que f est injective.
Comme on a vu que f était aussi surjective, f est bijective. De plus, f est un $\mathcal{C}^2$-difféomorphisme local, d'après le théorème d'inversion locale, donc f^{-1} est aussi un $\mathcal{C}^2$-difféomorphisme local. Donc f est un $\mathcal{C}^2$-difféomorphisme global.