Analyse dans $\mathbf{R}^n$

Chapitre no. 2, Calcul différentiel

On suppose $U \subset \mathbf{R}^n$ un ouvert. Soient $f, g : U \to \mathbf{R}^p$ deux applications continues. On dit que f et g sont *tangentes* en $a \in U$ si

$$\lim_{x \to a,\, x \neq a} \frac{||f(x) - g(x)||}{||x - a||} = 0\,,$$

ce qui implique bien evidemment que $f(a) = g(a)$. D'après l'inégalité triangulaire : $||f(x) - h(x)|| \leq ||f(x) - g(x)|| + ||g(x) - h(x)||$, si f, g sont tangentes en a, et g, h sont tangentes en a, alors f, h sont tangentes en a.

Parmi les applications tangentes à f en a, il exite au plus une de la forme $x \to f(a) + u(x-a)$ avec u linéaire :

Preuve : Supposons que existe deux telles applications $x \to f(a) + u_1(x - a)$ et $x \to f(a) + u_2(x - a)$. Il s'ensuit que l'application linéare $v = u_1 - u_2$ vérifie

$$\lim_{y \to 0,\, y \neq 0} \frac{||v(y)||}{||y||} = 0.$$

Mais dans ce cas $v = 0$, car, donné $\varepsilon > 0$, il existe $r > 0$ tel que $||y|| \leq r \Rightarrow ||v(y)|| \leq \varepsilon ||y||$. Puisque v est linéaire, on voit que cette inégalité est vérifié pour tout $x \neq 0$ (on pose $y = rx/||x||$). Puisque ε est arbitraire, on voit que $v(x) = 0$ pour tout x (il suffit de supposer que $||x|| = 1$). $\square$

On dit qu'une application $f : U \to \mathbf{R}^p$ est *différentiable en* $a \in U$ s'il existe une application linéaire $u : \mathbf{R}^n \to \mathbf{R}^p$ telle que $x \to f(a) + u(x - a)$ est tangente à f en a. On vient de voir qu'une telle application est unique ; elle s'appelle la *dérivée de* f *en* a, et s'écrit $Df(a)$ (ou $f'(a)$). La dérivée de f est caractérisée par la propriété :

$$f(x) = f(a) + Df(a)(x-a) + ||x-a||\varphi(x-a) \quad \text{ou} \quad f(a+h) = f(a) + Df(a)(h) + ||h||\varphi(h)$$

pour x assez proche de a (h assez petit), où $\lim_{h \to 0} \varphi(h) = 0$.

On remarque que si la dérivée existe en a, alors $f(x) \to f(a)$ lorsque $x \to a$, d'où, f est continue en a.

Cas particuliers : 1. Soit $f :]a, b[\to \mathbf{R}$ et soit $x \in]a, b[$. Si la dérivée existe en x, alors $Df(x) = f'(x)$ où

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

2. Si $f : \mathbf{R}^n \to \mathbf{R}^p$ est constante, alors $Df(a) = 0$ pour tout $a \in \mathbf{R}^n$.

3. Soit $f : U \to \mathbf{R}$, avec U un ouvert dans $\mathbf{R}^n$ et soit $a = (a_1, a_2, \ldots, a_n) \in U$. On suppose $\mathbf{R}^n$ muni des coordonnées cartesiennes $(x_1, x_2, \ldots, x_n)$. Si elles existent, les dérivées partielles de f en a, sont définies par

$$\frac{\partial f}{\partial x_j}(a) = \lim_{h_j \to 0} \frac{f(a_1, \ldots a_{j-1}, a_j + h_j, a_{j+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h_j}$$

pour chaque $j = 1, \ldots, n$. (On fixe $a_1, a_2, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n$ et on prend la dérivée da la fonction réelle $t \to f(a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_n)$ en $t = a_j$.)

On voit alors que si la dérivée de f existe en a, pour h assez petit,

$$f(a+h) = f(a) + \frac{\partial f}{\partial x_1}(a)h_1 + \cdots \frac{\partial f}{\partial x_n}(a)h_n + ||h||\varphi(h)$$

où $h = (h_1, \ldots, h_n) \in \mathbf{R}^n$ et $\lim_{h\to 0} \varphi(h) = 0$. Il s'ensuit que la dérivée de f est l'application linéaire $Df(a) : \mathbf{R}^n \to \mathbf{R}$ donnée par la $1 \times n$–matrice (agissant sur les vecteurs colonnes $h \in \mathbf{R}^n$)

$$\left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right)$$

4. Si $f : \mathbf{R}^n \to \mathbf{R}^p$ est linéaire, alors $Df(a) = f$ pour tout $a \in \mathbf{R}^n$. En effet $f(a+h) - f(a) = f(h)$ et f est déjà linéaire.

5. Si $f : U \to \mathbf{R}^p \times \mathbf{R}^q$ ($U \subset \mathbf{R}^n$ ouvert) et on écrit $f = (f_1, f_2)$ où $f_1 : U \to \mathbf{R}^p$ et $f_2 : U \to \mathbf{R}^q$, alors, pour $a \in U$, on a $Df(a) : \mathbf{R}^n \to \mathbf{R}^p \times \mathbf{R}^q$ l'application linéaire donnée par $Df(a) = (Df_1(a), Df_2(a)$. Si on écrit les vecteurs comme vecteurs colonnes, on en déduit du cas 3 que, pour une application $f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_p \end{pmatrix} : U \to \mathbf{R}^p$ (U ouvert dans $\mathbf{R}^n$), la dérivée est l'application linéaire de matrice

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

Cette matrice s'appelle la *matrice jacobienne* de f en a. Elle représente l'application linéaire $Df(a)$ dans les coordonnées données, dans ce cas, les coordonnées cartesiennes. On la note par $J_f(a)$ ou par $\dfrac{D(f_1, f_2, \ldots, f_p)}{D(x_1, x_2, \ldots, x_n)}$ si on veut indiquer explicitment les coordonnées par rapport auxquelles on prend les dérivées. Si $n = p$, la quantité $\det J_f(a)$ s'appelle le *jacobien* de f en a. Il intervient dans la formule de changement de variables dans un intégrale à plusieurs variables.

<u>Dérivée suivant le vecteur X</u>. Soit $f : U \to \mathbf{R}$ ($U \subset \mathbf{R}^n$ ouvert) une fonction différentiable en $a \in U$ et soit $X \in \mathbf{R}^n$. La fonction $\varphi(t) = f(a+tX)$ est définie pour t assez petit et vérifie :

$$\lim_{t\to 0} \frac{\varphi(t) - \varphi(0) - Df(a)(tX)}{||tX||} = \lim_{t\to 0} \frac{f(a+tX) - f(a) - Df(a)(tX)}{||tX||} = 0$$

(on pose $h = tX$ dans la définition de dérvée), donc

$$\lim_{t\to 0} \frac{\varphi(t) - \varphi(0) - tDf(a)(X)}{|t|} = 0$$

et φ est dérivable en $t = 0$ avec dérivée $\varphi'(0) = Df(a)(X)$. On dit que $Df(a)(X)$ *est la dérivée de f en a suivant le vecteur X*. Il s'ensuit que les dérivées partielles $\partial f/\partial x_i(a) = Df(a)(e_i)$ sont les dérivées de f suivant les vecteurs $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ (avec 1 dans la i'ième place).

Il s'ensuit qu'une condition nécessaire pour que f soit différentiable en a est que ses dérivées partielles existent. Pourtant l'existence des n dérivées partielles $\dfrac{\partial f}{\partial x_i}(a)$ ne suffit pas pour assurer la différentiabilité de f. Par exemple, soit $f : \mathbf{R}^2 \to \mathbf{R}$ définie par

$$f(x,y) = \frac{xy}{\sqrt{x^2+y^2}} \quad \text{pour } (x,y) \neq (0,0) \quad \text{et } f(0,0) = 0$$

Puisque f se réduit à la constante nulle sur les axes des coordonnées, elle vérifie donc

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0\,.$$

Si cette fonction était différentiable à l'origine, sa différentielle en ce points serait nulle, et par définition de dérivée, le rapport

$$\frac{f(x,y)}{\sqrt{x^2+y^2}} = \frac{xy}{x^2+y^2}$$

tendrait vers zéro à l'origine, ce qui n'est manifestement pas le cas. Donc f n'est pas différentiable à l'origine. Cependant elle admet non seulement des dérivées partielles, mais aussi les dérivées suivant toutes les directions.

<u>Exercice</u> : (a) Soit $f : \mathbf{R}^2 \to \mathbf{R}$ définie par

$$f(x,y) = \frac{xy^2}{x^2+y^2} \text{ pour } (x,y) \neq (0,0) \quad \text{et } f(0,0) = 0\,.$$

Montrer que pour tout $x, y \in \mathbf{R}^2$, la limite

$$\lim_{t\to 0, t\neq 0} \frac{f(x+ty) - f(x)}{t} := g(x,y)$$

existe, mais que $y \to g(0,y)$ n'est pas linéaire et donc f n'est pas différentiable au point $(0,0)$.

(b) Soit f définie par

$$f(x,y) = \frac{x^3 y}{x^4+y^2} \text{ pour } (x,y) \neq (0,0) \quad \text{et} f(0,0) = 0\,.$$

Montrer que la limite $g(x,y)$ existe pour tout x et y, et que $y \to g(x,y)$ est linéaire pour tout $x \in \mathbf{R}^2$, mais que f n'est pas différentiable en $(0,0)$ (on considère les points (x,y) tels que $y = x^2$).

Loi de la dérivée d'une application composée. On suppose que $f : U \to \mathbf{R}^p$ (U ouvert dans $\mathbf{R}^n$) et $g : V \to \mathbf{R}^q$ (V ouvert dans $\mathbf{R}^p$) avec $f(U) \subset V$. On peut définir l'application composée $h = g \circ f$ définie par $h(x) = g(f(x))$ pour tout $x \in U$. Alors, si f est dérivable en $a \in U$ et g est dérivable en $b = f(a)$, il s'ensuit que h est dérivable en a avec dérivée

$$Dh(a) = Dg(f(a)) \circ Df(a)$$

où la partie droite est la composée des application linéaire (en termes des matrices jacobiennes, il s'agit de multiplication des matrices).

Preuve : Par hypothèse, donné ε avec $0 < \varepsilon < 1$, il existe $\delta > 0$ tel que pour $||h|| < \delta$ et $||k|| < \delta$ on a

$$\begin{aligned} f(a+h) &= f(a) + Df(a)(h) + o_1(h) \\ g(b+k) &= g(b) + Dg(b)(k) + o_2(k) \end{aligned}$$

avec $||o_1(h)|| < \varepsilon ||h||$ et $||o_2(k)|| < \varepsilon ||k||$. D'autre part, puisque la dérivée est une application linéaire, il existe des constantes A, B tels que, quel que soit h et k, on a

$$||Df(a)(h)|| \leq A||h|| \quad \text{et} \quad ||Dg(b)(k)|| \leq B||k||$$

(on peut prendre $A = \max\{||Df(a)(h)|| : ||h|| = 1\}$, par exemple). Il s'ensuit que

$$||Df(a)(h) + o_1(h)|| \leq (A+1)||h||$$

pour tout $||h|| < \delta$, d'où, pour $||h|| \leq \delta/(A+1)$, on a

$$||o_2(Df(a)(h) + o_1(h)|| < (A+1)\varepsilon||h|| \quad \text{et} \quad ||Dg(b)(o_1(h))|| \leq B\varepsilon||h||$$

On peut alors écrire

$$h(a+h) = g(b + Df(a)(h) + o_1(h)) = g(b) + Dg(b)(Df(a)(h)) + o_3(h)$$

avec

$$||o_3(h) \leq (A+B+1)\varepsilon||h||$$

ce qui démontre la formule. □

<u>Exemples</u> : 1. Soit $f :]a,b[\to]c,d[$ et $g :]c,d[\to \mathbf{R}$ deux fonctions différentiables ; alors $(g \circ f)'(x) = g'(f(x))f'(x)$.

2. Soit $t \to (x(t), y(t))$ une courbe dans $\mathbf{R}^2$ et soit $f : \mathbf{R}^2 \to \mathbf{R}$ une fonction (les deux différentiables). On peut former la fonction composée : $t \to f(x(t), y(t))$. En termes de matrices jacobiennes, on a

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x(t), y(t))|_{t=t_0} = \begin{pmatrix} \frac{\partial f}{\partial x}(x(t_0), y(t_0)) & \frac{\partial f}{\partial y}(x(t_0), y(t_0)) \end{pmatrix} \begin{pmatrix} x'(t_0) \\ y'(t_0) \end{pmatrix} = \frac{\partial f}{\partial x} x' + \frac{\partial f}{\partial y} y'$$

Dans le cas où $t \to (x(t), y(t))$ est un niveau de f : $f(x(t), y(t)) = c$ constante, alors $(\mathrm{d}/\mathrm{d}t) f(x(t), y(t)) = 0$, ce qui traduit au fait que

$$\begin{pmatrix} \frac{\partial f}{\partial x}(x(t), y(t)) \\ \frac{\partial f}{\partial y}(x(t), y(t)) \end{pmatrix} \cdot \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = 0$$

<u>Gradient</u>. Un champ de vecteurs est une correspondence qui à tout point x d'un domaine $U \subset \mathbf{R}^n$ associe un vecteur : $x \to \vec{v}(x)$. Le gradient est un cas particulier d'un champ de vecteurs : si $f : U \to \mathbf{R}$ est différentiable en tout point $x \in U \subset \mathbf{R}^n$, alors son *gradient* est défini par

$$\operatorname{grad} f(x) := \vec{\nabla} f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$$

On a vu ci-dessus que le gradient est orthogonal aux niveaux de f ; il est dirigé dans le sens de *plus grande croissance de* f.

<u>Exercice</u> : (gradient et dérivée suivant la direction X) : Montre que pour une fonction différentiable $f : U \to \mathbf{R}$ ($U \subset \mathbf{R}^n$ ouvert), on a la relation suivante entre la dérivée suivant la direction X et le gradient :

$$Df(a)(X) = \operatorname{grad} f(a) \cdot X$$

où, dans la partie droite, le produit s'agit du produit scalaire entre vecteurs.

3. Soit $f = (f_1, f_2, f_3) : U \to \mathbf{R}^3$ ($U \subset \mathbf{R}^2$) et $g = (g_1, g_2) : V \to \mathbf{R}^2$ ($V \subset \mathbf{R}^3$ avec $f(U) \subset V$). En termes de matrices jacobiennes, on a

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) \\ \frac{\partial f_3}{\partial x_1}(a) & \frac{\partial f_3}{\partial x_2}(a) \end{pmatrix} \quad \text{et} \quad J_g(a) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(b) & \frac{\partial g_1}{\partial y_2}(b) & \frac{\partial g_1}{\partial y_3}(b) \\ \frac{\partial g_2}{\partial y_1}(b) & \frac{\partial g_2}{\partial y_2}(b) & \frac{\partial g_2}{\partial y_n}(b) \end{pmatrix}$$

où (x_1, x_2) sont les coordonnées du domaine de f, (y_1, y_2, y_3) les coordonnées du domaine de g et $b = f(a)$. L'application composée $h = g \circ f : U \to \mathbf{R}^2$ est alors représentée par la matrice jacobienne :

$$J_h(a) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(b) & \frac{\partial g_1}{\partial y_2}(b) & \frac{\partial g_1}{\partial y_3}(b) \\ \frac{\partial g_2}{\partial y_1}(b) & \frac{\partial g_2}{\partial y_2}(b) & \frac{\partial g_2}{\partial y_n}(b) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) \\ \frac{\partial f_3}{\partial x_1}(a) & \frac{\partial f_3}{\partial x_2}(a) \end{pmatrix}$$

Théorème : Soit f un homéomorphisme (bijection continue) d'un ouvert U de $\mathbf{R}^n$ dans un ouvert V de $\mathbf{R}^n$, et soit g l'homéomorphisme inverse. On suppose que f est différentiable au point a et que $Df(a) : \mathbf{R}^n \to \mathbf{R}^n$ est un isomorphisme linéaire ; alors g est différentiable au point $b = f(a)$ et $Dg(b)$ est l'application inverse de $Df(a)$: $Dg(f(a)) = (Df(a))^{-1}$ (voir ci-dessus : Théorème de l'application inverse).

Théorème des accroissements finis. On se rappelle, que pour une fonction $f : [a, b] \to \mathbf{R}$, continue sur l'intervalle fermé $[a, b]$ et dérivable sur l'intervalle ouvert $]a, b[$, il existe $c \in]a, b[$ vérifiant

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

ce qui implique, bien evidemment, que si $|f'(x)| \leq M$ pour tout $x \in]a, b[$, on a l'inégalité

$$|f(b) - f(a)| \leq M|b - a|$$

Soit $f : U \to \mathbf{R}$ différentiable ($U \subset \mathbf{R}^n$ ouvert) et supposons que $a, b \in U$ sont deux points tels que le segment droit $[a, b] = \{(1 - t)a + tb : 0 \leq t \leq\}$ soit contenu dans U. On peut alors appliquer la formule des accroissments finis à la fonction réelle $g(t) = f((1-t)a+tb)$ définie sur l'intervalle $[0, 1]$, pour en déduire que $g(1)-g(0) = g'(s)$ pour $0 < s < 1$. Mais $g(1) = f(y)$, $g(0) = f(x)$ et $g'(s) = Df((1 - s)a + sb)(b - a)$, d'où

$$f(y) - f(x) = Df((1 - s)a + sb)(b - a) = \operatorname{grad} f((1 - s)a + sb) \cdot (b - a)$$

En appliquant l'inégalité de Cauchy-Schwarz à la partie droite on en déduit que

$$|f(y) - f(y)| \leq \max_{c \in [x,y]} ||Df(c)|| \times ||x - y||$$

où la norme sur la dérivée est donnée par $||Df(c)|| = \sup_{||v||=1} ||Df(c)(v)||$. Cette formule se généralise aux applications $f : \mathbf{R}^n \to \mathbf{R}^p$:

$$||f(y) - f(y)|| \leq \max_{c \in [x,y]} ||Df(c)|| \times ||x - y||$$

Théorème de l'application inverse. Ce théorème affirme que si la dérivée en un point est un isomorphisme (linéaire) alors l'application est un homéomorphsme dans un voisinage de ce point. Ce résultat est important plus tard pour montrer l'existence d'une application implicite. Pour le montrer il nous faut d'abord un résultat prèliminaire.

Théorème du point fixe : Soit (X, d) un espace métrique complet (on peut penser à la boule unitée fermée munie de sa norme euclidienne) et soit $\varphi : X \to X$ une application contractante non-constante, c'est à dire il existe un nombre $k \in]0, 1[$ tel que pour tout $x, y \in X$ on a :

$$d(\varphi(x), \varphi(y)) \leq kd(x, y) \, .$$

Alors, il existe un point unique $x^* \in X$ tel que $\varphi(x^*) = x^*$ (un point fixe).

Preuve : Soit $x_0 \in X$ quelconque. On définit une suite (x_n) en posant $x_n = \varphi(x_{n-1})$. Par récurrence on voit que

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

Il s'ensuit que (x_n) est une suite de Cauchy. En effet, soit $m, n \in \mathbf{N}$ tels que $m > n$:

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq (k^{m-1} + k^{m-2} + \cdots + k^n) d(x_1, x_0) \\ &= k^n d(x_1, x_0) \sum_{j=0}^{m-n-1} k^j \leq k^n d(x_1, x_0) \sum_{j=0}^{\infty} k^j = k^n d(x_1, x_0) \left(\frac{1}{1-k} \right) \end{aligned}$$

Quel que soit $\varepsilon > 0$, on peut trouver N tel que

$$k^N < \frac{\varepsilon(1-k)}{d(x_1, x_0)}$$

Pour $m, n > N$ on a

$$d(x_m, x_n) \leq k^n d(x_1, x_0) \left(\frac{1}{1-k} \right) < \varepsilon$$

Puisque la suite est de Cauchy, par la complétitude de X il existe une limite $x^* \in X$ ce qui nous donne le point fixe :

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} \varphi(x_{n-1}) = \varphi(\lim_{n\to\infty} x_{n-1}) = \varphi(x^*)$$

où on a inverser la limite et φ par la continuité de φ, qui s'ensuit du fait qu'elle est une application contractante.

Le point x^* est unique, sinon, si y^* est un autre point fixe, on aurait

$$0 < d(\varphi(x^*), \varphi(y^*)) = d(x^*, y^*) > kd(x^*, y^*),$$

une contradiction. □

Application de classe C^1. Soit $f : U \to \mathbf{R}^p$ ($U \subset \mathbf{R}^n$ ouvert) différentiable en tout point de U. Alors pour tout $x \in U$ sa dérivée $Df(x)$ est un élément de $L(\mathbf{R}^n, \mathbf{R}^p)$, l'espace des applications linéaires de $\mathbf{R}^n$ dans $\mathbf{R}^p$. Cet espace admet des normes définies dans le chapitre 1 (toutes équivalentes). On peut alors parler de la continuité de $Df : U \to L(\mathbf{R}^n, \mathbf{R}^p)$. Si cette application est continue, on dit que f est de classe C^1. En termes des coordonnées canoniques, si on représente la dérivée par sa matrice jacobienne, f est de classe C^1 si et seulement si ses dérivées partielles sont continues.

Théorème (de l'application inverse). Soit $U \subset \mathbf{R}^n$ ouvert et soit $x_0 \in U$. On suppose que $f : U \to \mathbf{R}^n$ est de classe C^1 et que $Df(x_0)$ est inversible. Alors il existe des voisinages $V \subset U$ et $W \subset \mathbf{R}^n$ de x_0 et $f(x_0)$ respectivement, et une application $g : W \to V$ de classe C^1, tels que :

$$f(g(y)) = y \quad \text{et} \quad g(f(x)) = x, \qquad \forall x \in V, \forall y \in W\,.$$

En plus $Dg(f(x)) = (Df(x))^{-1}$.

Preuve : Quitte à remplacer x par $x - x_0$, $f(x)$ par $f(x) - f(x_0)$ et $f(x)$ par $(Df(0))^{-1} f(x)$, on peut supposer que $x_0 = 0$, $f(0) = 0$ et $Df(0) = I$ (l'application identité). Soit $\psi(x) = x - f(x)$. Clairement ψ est de classe C^1 et $D\psi(0) = 0$ (l'application nulle). Il s'ensuite qu'il existe $\delta > 0$ tel que

$$(1) \qquad ||x|| \le \delta \ \Rightarrow\ ||D\psi(x)|| \le \frac{1}{2}$$

(on applique le fait que ψ est de classe C^1).

Par le théorème des accroissements finis :

$$(2) \qquad ||x|| \le \delta \ \Rightarrow\ ||\psi(x)|| \le \max_{c \in [0,x]} ||D\psi(c)||\, ||x|| \le \frac{\delta}{2}$$

Soit $y \in \mathbf{R}^n$ avec $||y|| < \delta/2$ et soit

$$\varphi_y(x) = \psi(x) + y = x - f(x) + y\,.$$

Par l'équation (2)

$$||x|| \le \delta \ \Rightarrow\ ||\varphi_y(x)|| = ||x - f(x) + y|| \le ||x - f(x)|| + ||y|| < \delta\,.$$

Par la continuité de φ_y on peut affirmer que $\varphi_y : B(0,\delta) \to B(0,\delta)$. En appliquant encore l'inégalité des accroissements finis, pour tout $x_1, x_2 \in B(0,\delta)$

$$||\varphi_y(x_1) - \varphi_y(x_2)|| = ||\psi(x_1) - \psi(x_2)|| \le \frac{1}{2}||x_1 - x_2||$$

Il s'ensuit que φ_y est une application contractante et par le théorème du point fixe, il présente un point fixe unique, disons $x_y \in B(0,\delta)$, d'où, il existe un unique point $x_y \in B(0,\delta)$ tel que

$$f(x_y) = y\,.$$

Puisque $||y|| < \delta/2$, on a $x_y \in B(0,\delta)$.

Définissons $W = B(0,\frac{\delta}{2})$, $V = f^{-1}(W) \cap B(0,\delta)$ et $g : W \to V$ par $g(y) = x_y$. Clairement $(f \circ g)(y) = y$ et $(g \circ f)(x) = x$ $\forall x \in V$ $\forall y \in W$. On montre que $g \in C^1$.

$\underline{g \text{ est continue}}$: Donné $\varepsilon >$, soient $y, z \in W$ tels que $||y - z|| < \varepsilon/2$. On pose $u = g(y)$ et $v = g(z)$. Alors

$$||\psi(u) - \psi(v)|| \leq \frac{1}{2}||u - v|| \quad \Rightarrow \quad ||f(u) - f(v) - (u - v)|| \leq \frac{1}{2}||u - v||$$
$$\Rightarrow ||u - v|| \leq 2||f(u) - f(v)|| \quad \Rightarrow \quad ||g(y) - g(z)|| \leq 2||y - z)) < \varepsilon$$

$\underline{g \text{ différentiable}}$: Soit $y \in W$ et soit $x = g(y)$. On montre que $Dg(y) = (Df(x))^{-1}$. Ecrivons A pour $Df(x)$. Puique W ouvert, on peut supposer k assez petit que $y + k \in W$. Posons $g(y + k) = x + h \Rightarrow h = g(y + k) - g(y)$. Il nous faut

$$\begin{aligned}
& \frac{g(y+k) - g(y) - A^{-1}k}{||k||} \to 0 \text{ lorsque } k \to 0 \\
\Leftrightarrow \quad & \frac{A(g(y+k) - g(y)) - k}{||k||} \to 0 \text{ lorsque } k \to 0 \\
\Leftrightarrow \quad & \frac{Ah - k}{||k||} = \frac{Ah - (f(x+h) - f(x))}{||k||} \to 0 \text{ lorsque } k \to 0 \\
\Leftrightarrow \quad & -\frac{f(x+h) - f(x) - Ah}{||h||} \times \frac{||h||}{||k||} \to 0 \text{ lorsque } k \to 0
\end{aligned}$$

Puisque $||h|| \leq 2||k||$, la limite de la dernière ligne est bien zéro, d'où g est différentiable et $Dg(y) = (Df(g(y)))^{-1}$. En plus Dg est la composée $Dg = i \circ Df \circ g$ où i est inversion. Il s'ensuit que Dg est continue et que $g \in C^1$. $\square$

Références 1. J. Dieudonné, Foundations of Modern Analysis, Academic Press 1969.

2. J. Lelong-Ferrand et J. M. Arnaudiès, Cours de Mathématiques Tome 2 : Analyse, 4ème édition, Dunod Université 1977.