Calcul Différentiel

par Alain Prouté
Université Denis Diderot — Paris 7

Table des matières

1 Dérivée d'une fonction.

1.1 L'application linéaire tangente.

Définition 1 *Soit E et F des espaces de Banach. Soit U un ouvert de E, et $f : U \longrightarrow F$ une application (qu'on ne suppose pas nécessairement continue). Soit x un point de U. On dit que f est* dérivable en x, *s'il existe une application linéaire continue f'_x de E vers F, telle que, pour tout h dans un voisinage de 0, on ait :*

$$f(x+h) = f(x) + f'_x(h) + o(h),$$

où $o(h)$ est négligeable devant h, quand h tend vers 0, ce qui signifie que :

$$\forall_{\varepsilon>0}\ \exists_{\eta>0}\ \|h\| < \eta \Rightarrow \|o(h)\| < \varepsilon\|h\|.$$

On notera que la fonction $h \mapsto o(h)$ dépend de f et de x. On remarquera également que "être néligeable devant h quand h tend vers 0" est plus fort que "tendre vers 0 quand h tend vers 0". En effet, en particularisant l'hypothèse que $o(h)$ est négligeable devant h quand h tend vers 0 à $\varepsilon = 1$, on obtient $\|o(h)\| \leq \|h\|$ pour h assez petit.

Notation de Landau : Quand on manipule plusieurs fonctions telles que la fonction o ci-dessus, on devrait normalement leur donner des noms différents : o_1, o_2, ... Toutefois, il est d'usage de les noter toutes o. C'est ce qu'on appelle la "notation de Landau". Il faut évidemment garder cette convention présente à l'esprit. Elle a pour conséquence entre autres, que :

$$o(h) + o(h) = o(h) \qquad\qquad o(h) - o(h) = o(h) \qquad\qquad ko(h) = o(h)$$

(où k est un réel). Ceci signifie tout simplement qu'une somme, une différence ou un multiple réel d'expressions négligeables devant h quand h tend vers 0 est encore une expression négligeable devant h quand h tend vers 0. Toutefois, pour plus de clarté, il nous arrivera de donner des noms différents aux expressions négligeables devant h quand h tend vers 0.

La fonction $h \mapsto f(x+h)$ est donc "presque" une fonction affine continue (à savoir la fonction $h \mapsto f(x) + f'_x(h)$), en ce sens que la différence entre ces deux fonctions est négligeable devant h quand h tend vers 0. On peut aussi dire que la fonction $h \mapsto f(x+h) - f(x)$ est "presque" linéaire continue.

Lemme 1 *Si f est dérivable en x, alors elle est continue en x.*

En effet, on a $f(x+h) - f(x) = f'_x(h) + o(h)$, et comme f'_x est continue, le membre de droite de cette égalité tend vers 0 quand h tend vers 0.

Lemme 2 *L'application linéaire continue f'_x de la définition précédente est unique (si elle existe).*

En effet, supposons que g soit une autre application linéaire continue, telle que $f(x+h) = f(x) + g(h) + o(h)$. On aurait alors $f'_x(h) - g(h) = o(h)$. Soit $\varepsilon > 0$, et h un élément de E. Pour h assez petit, on a :

$$\|f'_x(h) - g(h)\| < \varepsilon\|h\|.$$

Comme ε est arbitraire, on voit que $\|f'_x - g\| = 0$, ce qui signifie $g = f'_x$. $\square$

L'application linéaire continue f'_x est appelée "application linéaire tangente à f en x", ou "dérivée de f en x", ou "différentielle de f en x".

L'application $x \mapsto f'_x$, qui envoie U dans l'espace des applications linéaires continues de E vers F s'appelle "dérivée de f", et sera notée f'. On ne doit pas confondre la dérivée $f' : U \longrightarrow \mathcal{L}(E, F)$, avec la dérivée en x, $f'_x : E \longrightarrow F$. On évitera la notation $f'(x)$, et on conservera la notation f'_x pour désigner l'image de x par f'. La notation $f'(x)$ quant–à elle est réservée pour un autre usage, comme on va le voir ci–après.

Certains ouvrages utilisent l'une des notations $(df)_x$ ou $(Df)_x$ au lieu de f'_x.

1.2 Lien avec la dimension 1.

Pour une fonction f définie sur un intervalle U de $\mathbb{R}$, et à valeurs dans un espace de Banach E, pas nécessairement de dimension 1, on a déjà une notion de dérivée en un point x_0. Il s'agit de la limite du rapport suivant (appelé taux d'accroissement dans le cas où $E = \mathbb{R}$) :

$$\frac{f(x) - f(x_0)}{x - x_0}$$

quand x tend vers x_0.

Cette dérivée est notée $f'(x_0)$, et est un élément de E, et non pas une *application linéaire* de $\mathbb{R}$ vers E. C'est la raison pour laquelle, nous réservons la notation $f'(x)$ pour désigner la dérivée de f en x (limite du rapport ci–dessus), et adoptons la notation f'_x pour désigner l'application linéaire tangente à f en x (malheureusement aussi appelée "dérivée").

Bien sûr, dans cette situation, l'application linéaire tangente a encore un sens. Il y a donc certainement un lien entre $f'(x)$ (élément de E) et l'application linéaire tangente f'_x (élément de $\mathcal{L}(\mathbb{R}, E)$). Ce lien s'exprime par l'égalité :

$$f'(x) = f'_x(1).$$

Remarquons d'abord que cette égalité a un sens, puisque f'_x étant une application (linéaire) de $\mathbb{R}$ vers E, $f'_x(1)$ est un élément de E, tout comme $f'(x)$.

Maintenant $f'_x : \mathbb{R} \longrightarrow E$ est caractérisé par :

$$f(x+h) = f(x) + f'_x(h) + o(h),$$

où $o(h)$ est négligeable devant h quand h tend vers 0, et $f'(x)$ est caractérisé par :

$$f(x+h) = f(x) + hf'(x) + o(h)$$

(développement limité à l'ordre 1 de f en x). Comme $f'_x(h) = hf'_x(1)$, par linéarité de f'_x, on voit, par unicité de la dérivée $f'(x)$ que $f'(x) = f'_x(1)$.

1.3 Exemples de calcul de dérivées.

La méthode standard est de tenter de "développer" $f(x+h)$ sous la forme $f(x)+l(h)+o(h)$, avec l linéaire continue, et $o(h)$ négligeable devant h quand h tend vers 0. Si on y parvient, l sera nécessairement f'_x, par unicité de l'application linéaire tangente en x.

Noter que toute application contante a une dérivée nulle en tout point.

1.3.1 Applications linéaires continues.

Si $u : E \longrightarrow F$ est une application linéaire continue, elle est dérivable en tout point de E. Un effet, on peut écrire

$$u(x+h) = u(x) + u(h) + o(h)$$

(avec ici $o(h) = 0$). La dérivée de u en x est donc u (et ceci quel que soit le point x de E).

1.3.2 L'application $x \mapsto x^n$.

Soit $\mathcal{A}$ une algèbre de Banach. L'application $x \mapsto x^n$ de $\mathcal{A}$ vers $\mathcal{A}$ est dérivable en tout x de E. En effet, on a, en developpant le produit $(x+h)^n$,

$$(x+h)(x+h)\dots(x+h) = x^n + x^{n-1}h + x^{n-2}hx + \dots + xhx^{n-2} + hx^{n-1} + o(h),$$

où $o(h)$ contient tous les termes du developpement qui sont de degré au moins 2 en h. On notera que tout monôme de degré au moins 2 en h est négligeable devant h. En effet, la norme d'un tel monôme est majorée par le produit d'une constante réelle par une puissance de $\|h\|$ d'exposant au moins 2.

Comme l'application $h \mapsto x^{n-1}h + x^{n-2}hx + \dots + xhx^{n-2} + hx^{n-1}$ est linéaire et continue, on voit qu'elle doit être la dérivée de $x \mapsto x^n$ en x.

Bien sûr dans le cas d'une algèbre *commutative*, on trouve la formule plus familière $h \mapsto nx^{n-1}h$.

1.3.3 L'application $x \mapsto x^{-1}$.

Soit $\mathcal{A}$ une algèbre de Banach. On sait que $\mathcal{A}^*$ est un ouvert de $\mathcal{A}$. Par ailleurs, l'application $x \mapsto x^{-1}$ est définie sur $\mathcal{A}^*$, et envoie $\mathcal{A}^*$ dans $\mathcal{A}$. Nous allons la dériver en un point x.

Commençons par la dériver en 1. On a

$$(1+h)(1-h) = 1 - h^2,$$

ce qui donne

$$(1+h)^{-1} = 1 - h + (1+h)^{-1}h^2,$$

ce qui a un sens pour h assez petit. Or, comme $(1+h)^{-1}$ est borné pour h voisin de 0 (continuité de $x \mapsto x^{-1}$ en 1), et comme h^2 est négligeable devant h quand h tend vers 0, on voit que $(1+h)^{-1}h^2$ est négligeable devant h. On a donc

$$(1+h)^{-1} = 1 - h + o(h).$$

Comme $h \mapsto -h$ est linéaire continue, la dérivée de $x \mapsto x^{-1}$ en 1 est donc $h \mapsto -h$.([1])

[1]On pouvait aussi procéder comme suit : On a pour $\|h\| < 1$,

$$(1+h)^{-1} = 1 - h + h^2 - h^3 + \dots = 1 - h + o(h),$$

puisque $\|h^2 - h^3 + \dots\| \leq \|h\|^2 \, \|1 - h + h^2 - \dots\| \leq \|h\|^2 \, \dfrac{1}{1-\|h\|}$.

Soit maintenant x un élément quelconque de $\mathcal{A}^*$. On a

$$\begin{aligned}(x+h)^{-1} &= (x(1+x^{-1}h))^{-1} \\ &= (1+x^{-1}h)^{-1}x^{-1} \\ &= (1-x^{-1}h+o(h))x^{-1} \\ &= x^{-1}-x^{-1}hx^{-1}+o(h).\end{aligned}$$

La dérivée de $x \mapsto x^{-1}$ en x est donc l'application (linéaire continue) $h \mapsto -x^{-1}hx^{-1}$.

Dans le cas d'une algèbre *commutative*, cette formule peut se réécrire $h \mapsto \dfrac{-h}{x^2}$. On retrouve ainsi le résultat bien connu dans le cas de l'algèbre $\mathbb{R}$.

1.3.4 Applications bilinéaires continues.

Soit $f : E \times F \longrightarrow G$ une application bilinéaire continue (où E, F et G sont des espaces de Banach). f est dérivable en tout point (x, y) de $E \times F$. En effet, on peut écrire

$$f(x+h, y+k) = f(x,y) + f(x,k) + f(h,y) + f(h,k).$$

Comme f est continue, on a $\|f(h,k)\| \leq \|f\|\,\|h\|\,\|k\|$. La norme de (h,k) est par définition $\sup(\|h\|, \|k\|)$. On a donc $\|f(h,k)\| \leq \|f\|\,\|(h,k)\|^2 = o((h,k))$.

Il en résulte que la dérivée de f en (x,y) est l'application linéaire continue $(h,k) \mapsto f(x,k) + f(h,y)$.

En particulier, si $E = F = G = \mathbb{R}$, et si f est le produit (multiplication des nombres réels), on a pour dérivée du produit en (x,y) :

$$(h,k) \mapsto xk + hy.$$

Des exemples importants sont bien sûr les produits scalaires, le produit vectoriel, et d'une manière générale, tout ce qui porte le nom de "produit", qui est généralement bilinéaire continu.

1.4 Dérivation des applications à valeurs dans un produit.

Supposons maintenant que f soit une application d'un ouvert U d'un espace de Banach E, vers un produit $F = F_1 \times \dots \times F_n$ d'espaces de Banach (lequel est un espace de banach avec la norme $\|(y_1, \dots, y_n)\| = \sup(\|y_1\|, \dots, \|y_n\|)$).

Il existe alors des applications uniques $f_i : U \longrightarrow F_i$, telles que pour tout x de U, on ait : $f(x) = (f_1(x), \dots, f_n(x))$. Les applications f_i sont appelées les *composantes* de f (relativement à la décomposition $F_1 \times \dots \times F_n$ de F).

Lemme 3 *Sous les conditions ci–dessus, f est dérivable en $x \in U$, si et seulement si chaque composante f_i est dérivable en x. De plus, on a pour tout h de E :*

$$f'_x(h) = ((f_1)'_x(h), \dots, (f_n)'_x(h)).$$

Autrement–dit, la dérivée de f en x a pour composantes les dérivées en x des composantes de f.

Supposons f dérivable en x, c'est–à–dire

$$f(x+h) = f(x) + f'_x(h) + o(h),$$

où $o(h)$ est négligeable devant h quand h tend vers 0. Notons $(f'_x)_i$ la $i^{ième}$ composante de l'application linéaire f'_x. $(f'_x)_i$ est une application linéaire continue de E vers F_i. En projetant la relation ci–dessus sur le $i^{ième}$ facteur de F, on obtient :

$$f_i(x+h) = f_i(x) + (f'_x)_i(h) + o_i(h),$$

où $o_i(h)$ est négligeable devant h quand h tend vers 0, ce qui montre que f_i est dérivable en x, et que $(f_i)'_x = (f'_x)_i$.

Réciproquement, si chaque f_i est dérivable en x, on a :

$$f_i(x+h) = f_i(x) + (f_i)'_x(h) + o_i(h),$$

où $o_i(h)$ est négligeable devant h quand h tend vers 0, ce qui donne immédiatemment :

$$f(x+h) = f(x) + ((f_i)'_x(h), \dots, (f_i)'_x(h)) + o(h),$$

où $o(h)$ est négligeable devant h quand h tend vers 0, et prouve donc que f est dérivable en x, avec pour dérivée l'application linéaire continue :

$$h \mapsto ((f_i)'_x(h), \dots, (f_i)'_x(h)). \ \square$$

1.5 Le théorème de dérivation des fonctions composées.

1.5.1 Énoncé et démonstration.

Théorème 1 *Soient E, F et G trois espaces de Banach. Soit U un ouvert de E, et V un ouvert de F. Soit $f : U \longrightarrow F$ et $g : V \longrightarrow G$ des applications, telles que $f(U) \subset V$. Soit enfin x un point de U.*

Si f est dérivable en x, et si g est dérivable en $f(x)$, alors $g \circ f$ est dérivable en x, et on a :

$$(g \circ f)'_x = g'_{f(x)} \circ f'_x.$$

Ce théorème dit simplement que la dérivation *commute* à la composition.

En effet, comme f est dérivable en x, on a

$$f(x+h) = f(x) + f'_x(h) + o_1(h),$$

où $o_1(h)$ est négligeable devant h quand h tend vers 0. De même, on a puisque g est dérivable en $f(x)$,

$$g(f(x)+k) = g(f(x)) + g'_{f(x)}(k) + o_2(k),$$

où $o_2(k)$ est négligeable devant k quand k tend vers 0.

Dans l'égalité ci–dessus, qui est valable pour tout k assez petit, on peut remplacer k par $f(x+h)-f(x)$ (lui aussi petit quand h est petit, et qui est égal à $f'_x(h) + o_1(h)$), ce qui donne

$$g(f(x+h)) = g(f(x)) + g'_{f(x)}(f'_x(h)) + g'_{f(x)}(o_1(h)) + o_2(f'_x(h) + o_1(h)).$$

Il suffit donc pour démontrer le théorème de montrer que

$$g'_{f(x)}(o_1(h)) + o_2(f'_x(h) + o_1(h))$$

est négligeable devant h, quand h tend vers 0. Or $o_1(h)$ étant négligeable devant h quand h tend vers 0, il en est de même de $g'_{f(x)}(o_1(h))$, puisque $g'_{f(x)}$ est une application linéaire continue, et $h \mapsto f'_x(h)+o_1(h)$ étant dominée par h au voisinage de 0 (car f'_x est linéaire continue ; en fait, on a clairement $\|f'_x(h) + o_1(h)\| \leq (\|f'_x\| + 1)\|h\|$ pour h assez petit), $o_2(f'_x(h) + o_1(h))$, qui est négligeable devant $f'_x(h) + o_1(h)$ quand h tend vers 0, est donc négligeable devant h quand h tend vers 0. $\square$

1.5.2 Cas de la dimension 1.

Dans le cas où les espaces E et F sont de dimension 1, on peut écrire :

$$(g \circ f)'_x(1) = g'_{f(x)} \circ f'_x(1),$$

ce qui donne :

$$(g \circ f)'(x) = g'_{f(x)}(f'(x)) = f'(x)g'_{f(x)}(1) = f'(x)g'(f(x)),$$

c'est–à–dire, la formule bien connue de dérivation des fonctions composées d'une seule variable réelle.

Exercices

$\boxed{1}$ Soit $a, b \in \mathbb{R}$ avec $a < b$ et soit $t \in]a, b[$. On considère les apllications :

$$\begin{array}{ccc} C([a,b],\mathbb{R}) & \xrightarrow{\varphi_t} & \mathbb{R} \\ f & \longmapsto & f(t) \end{array} \quad \text{et} \quad \begin{array}{ccc} C([a,b],\mathbb{R}) & \xrightarrow{\psi} & \mathbb{R} \\ f & \longmapsto & \displaystyle\int_a^b f(x)dx \end{array}$$

Montrer que φ_t et ψ sont dérivables.

$\boxed{2}$ On considère $\mathcal{M}_2(\mathbb{R})$ l'ensemble des matrices 2×2 à coéfficients réels. Soit l'application θ sur $\mathcal{M}_2(\mathbb{R})$ définie par :

$$\begin{array}{ccc} \mathcal{M}_2(\mathbb{R}) & \xrightarrow{\theta} & \mathcal{M}_2(\mathbb{R}) \\ A & \longmapsto & A^2 \end{array}$$

Montrer que θ est dérivable sur $\mathcal{M}_2(\mathbb{R})$ et donner une éxpression de sa dérivée.

$\boxed{3}$ Soit une application $\varphi : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^2$. Soient $a, b \in \mathbb{R}$ tels que $a \leq b$. On pose alors :

$$\begin{array}{ccc} C([a,b],\mathbb{R}) & \xrightarrow{\varphi_*} & C([a,b],\mathbb{R}) \\ f & \longmapsto & \varphi \circ f \end{array}$$

Montrer que φ_* est dérivable et donner sa dérivée.

$\boxed{4}$ On considère l'application :

$$\begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \xrightarrow{f} & \mathbb{R}^3 \\ (x, y) & \longmapsto & x \wedge y \end{array}$$

où $\wedge$ est le produit vectoriel de $\mathbb{R}^3$.

Montrer que f est dérivable et donner sa sérivée.

1.6 Nouveaux exemples de dérivées.

1.6.1 Produit de deux fonctions.

Lemme 4 *(Dérivation d'un produit) Soit E un espace de Banach, U un ouvert de E, et $\mathcal{A}$ une algèbre de Banach. Soit $f, g : U \longrightarrow \mathcal{A}$ deux applications dérivables en $x_0 \in U$. Alors l'application produit fg est* *dérivable en x_0, et a pour dérivée :*

$$h \mapsto f'_{x_0}(h)g(x_0) + f(x_0)g'_{x_0}(h).$$

L'application fg, c'est–à–dire $x \mapsto f(x)g(x)$ est la composée suivante :

$$\begin{array}{ccccc} E & \xrightarrow{(f,g)} & \mathcal{A}\times\mathcal{A} & \xrightarrow{.} & \mathcal{A} \\ x & \longmapsto & (f(x),g(x)) & \longmapsto & f(x)g(x) \end{array}$$

La première application envoie x_0 sur le couple $(f(x_0), g(x_0))$, et a pour dérivée en x_0 l'application linéaire :

$$h \mapsto (f'_{x_0}(h), g'_{x_0}(h)),$$

et la seconde, qui est une application bilinéaire continue, a pour dérivée en $(f(x_0), g(x_0))$, l'application linéaire :

$$(k,l) \mapsto hg(x_0) + f(x_0)l.$$

On obtient le résultat annoncé en appliquant le théorème de dérivation des fonctions composées.

Le résultat du lemme précédent est assez malpratique, car il oblige à faire apparaître h dans l'expression de la dérivée de fg. On aimerait pouvoir écrire :

$$(fg)'_{x_0} = f'_{x_0}g(x_0) + f(x_0)g'_{x_0}$$

ou même la forme plus condensée :

$$(fg)' = f'g + fg'$$

C'est possible, mais à condition de prendre conscience du fait que les produits qui interviennent dans ces nouvelles formules ne sont plus celui de l'algèbre $\mathcal{A}$, mais des produit induits par lui. En effet, f'_{x_0} n'appartient pas à $\mathcal{A}$, mais à $\mathcal{L}(E,\mathcal{A})$, alors que $f'_{x_0}(h)$ appartient bien entendu à $\mathcal{A}$. Le produit $f'_{x_0}g(x_0)$ (qui doit appartenir à $\mathcal{L}(E,\mathcal{A})$) est donc celui d'un élément de $\mathcal{L}(E,\mathcal{A})$ par un élément de $\mathcal{A}$.

Un tel produit existe bien. Il est défini par :

$$(l,x) \mapsto (h \mapsto l(h)x)$$

où le produit de $l(h)$ par x est celui de $\mathcal{A}$. Ce nouveau produit est clairement bilinéaire, et est le seul (de $\mathcal{L}(E,\mathcal{A})\times\mathcal{A}$ vers $\mathcal{L}(E,\mathcal{A})$) qu'on puisse "naturellement" associer au produit dont on dispose déjà sur $\mathcal{A}$. C'est pourquoi il est dit "induit" par celui de $\mathcal{A}$. Notez qu'on a de même un produit $\mathcal{A}\times\mathcal{L}(E,\mathcal{A}) \longrightarrow \mathcal{L}(E,\mathcal{A})$, défini par :

$$(x,l) \mapsto (h \mapsto xl(h))$$

qui est celui qu'on doit utiliser dans l'expression $f(x_0)g'_{x_0}$.

Ces conventions étant admises (c'est à dire qu'on a surdéfini la simple juxtaposition qui sert à noter les produits), la formule $(fg)'_{x_0} = f'_{x_0}g(x_0) + f(x_0)g'_{x_0}$ est correcte.

Pour ce qui est de la seconde formule, il faut aller un peu plus loin. De même qu'on définit le produit de deux fonctions f et g (par exemple de $\mathbb{R}$ vers $\mathbb{R}$ ou de U vers $\mathcal{A}$ comme c'est le cas ici) en posant $(fg)(x) = f(x)g(x)$, on peut définir le produit de deux fonctions :

$$\varphi : U \longrightarrow \mathcal{L}(E,\mathcal{A}) \qquad \text{et} \qquad \psi : U \longrightarrow \mathcal{A}$$

en posant $(\varphi\psi)(x) = \varphi(x)\psi(x)$, où bien entendu, le produit de $\varphi(x) \in \mathcal{L}(E,\mathcal{A})$ par $\psi(x) \in \mathcal{A}$ est celui dont on a parlé ci–dessus. Dans ces conditions, le seconde formule $(fg)' = f'g + fg'$ est elle aussi correcte. On a donc :

Lemme 5 *Soit E un espace de Banach, U un ouvert de E, et $\mathcal{A}$ une algèbre de Banach. Soit $f, g : U \longrightarrow \mathcal{A}$ deux applications dérivables sur U. Alors l'application produit fg est dérivable sur U, et on a :*

$$(fg)' = f'g + fg'$$

où le produit de f' par g est défini par :

$$f'g = x_0 \mapsto (h \mapsto f'_{x_0}(h)g(x_0)) \qquad \textit{c'est à dire} \qquad (f'g)(x_0)(h) = f'_{x_0}(h)g(x_0)$$

et de manière similaire pour le produit de f par g'.

1.6.2 Fonctions polynômiales.

Soit E un espace de Banach. Soit E^* l'espace des formes linéaires continues sur E. E^* est un sous–espace de l'algèbre des fonctions continues de E vers $\mathbb{R}$. Ce sous–espace engendre une sous–algèbre, qu'on appelle *algèbre des fonctions polynômiales sur E*.

Dans le cas où $E = \mathbb{R}^n$, E^* a pour base $(e_1^*, \dots, e_n^*)$, où $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$. L'algèbre des fonctions polynômiales sur $\mathbb{R}^n$ est donc engendrée par ces formes linéaires. Elle contient dont les produits de la forme $e_{i_1}^* \dots e_{i_k}^*$, de même que toutes les combinaisons linéaires de ces produits. Il est facile de vérifier que l'ensemble de ces combinaisons linéaires est une algèbre. C'est donc l'algèbre des fonctions polynômiales.

Il faut remarquer que si on note $(x_1, \dots, x_n)$ une variable représentant un élément quelconque de $\mathbb{R}^n$, la forme linéaire e_i^* n'est autre que :

$$(x_1, \dots, x_n) \mapsto x_i.$$

Avec ces notations, une combinaison linéaire de produits des formes linéaires e_i^* n'est alors pas autre chose qu'une application de la forme :

$$(x_1, \dots, x_n) \mapsto P(x_1, \dots, x_n),$$

où $P(x_1, \dots, x_n)$ est un *polynôme en les variables $x_1, \dots, x_n$*.

Ceci justifie la définition donnée plus haut de l'algèbre des fonctions polynômiales sur E.

Une fonction polynômiale étant une somme de produits de formes linéaires, il suffit pour la dériver de savoir dériver une somme (la dérivée est la somme des dérivées), un produit de deux facteurs (c'est un cas particulier d'application bilinéaire), et une forme linéaire continue (sa dérivée est elle–même).

Par exemple, la dérivée au point (x_0, y_0, z_0) de l'application suivante de $\mathbb{R}^3$ vers R :

$$(x, y, z) \mapsto x^2 + 3xy - 5xyz$$

est :

$$(h, k, l) \mapsto 2hx_0 + 3hy_0 + 3x_0k - 5hy_0z_0 - 5x_0kz_0 - 5x_0y_0l.$$

En effet, $(x, y, z) \mapsto x^2$ est le produit de la forme linéaire $(x, y, z) \mapsto x$ par elle–même. Sa dérivée est donc :

$$(h, k, l) \mapsto x_0h + hx_0 = 2hx_0$$

$(x, y, z) \mapsto 3xy$ est un produit de deux forme linéaires, et a pour dérivée :

$$(h, k, l) \mapsto 3hy_0 + 3x_0k$$

Enfin, la fonction polynômiale $(x, y, z) \mapsto -5xyz$ est un produit de trois forme linéaires, et a pour dérivée :

$$(h, k, l) \mapsto -5hy_0z_0 - 5x_0kz_0 - 5x_0y_0l.$$

Exercices

$\boxed{5}$ Montrer que la norme associée au produit scalaire d'un espace euclidien est dérivable en tout point autre que 0, et calculer cette dérivée.

1.7 Matrice jacobienne.

Nous suppposons maintenant que les espaces de Banach E et F sont de dimensions finies n et p, et sont munis de bases. Soit donc $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, et $\mathcal{B}' = (\varepsilon_1, \dots, \varepsilon_p)$ une base de F.

Soit U un ouvert de E, x un point de U, et $f : U \longrightarrow F$ une application dérivable en x. On notera f_i les composantes de f relativement à la base $\mathcal{B}'$. La dérivée f'_x de f en x est alors une application linéaire de E vers F. f'_x a une matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$.

Définition 2 *La matrice de f'_x dans les bases $\mathcal{B}$ et $\mathcal{B}'$ s'appelle la* matrice jacobienne *de f en x (relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$). Le coefficient de la $i^{\text{ième}}$colonne et $j^{\text{ième}}$ligne de cette matrice est noté :*

$$\left(\frac{\partial f_j}{\partial x_i}\right)_x,$$

et est appelé dérivée partielle en x de la composante f_j relativement à la variable x_i.

Note : La notation $\dfrac{\partial f_j}{\partial x_i}$ contient une référence à une variable x_i qui n'a pas été définie. Ceci est dû au fait qu'en général, la fonction f_j est décrite par une formule de la forme :

$$(x_1, \dots, x_n) \longmapsto f_j(x_1, \dots, x_n).$$

Si la description de cette fonction faisait intervenir d'autres variables, par exemple, comme $u_1, \dots, u_n$, il faudrait noter cette dérivée partielle $\dfrac{\partial f_j}{\partial u_i}$.

Il s'agit donc d'un abus de langage, que seule une tradition séculaire nous oblige à conserver. Dans tous les cas, le contexte permet de savoir de quoi on parle.

La matrice jacobienne de f en x, notée parfois $\mathrm{Jac}\,(f)_x$ peut donc s'écrire comme suit :

$$\mathrm{Jac}\,(f)_x = \begin{pmatrix} \left(\frac{\partial f_1}{\partial x_1}\right)_x & \dots & \left(\frac{\partial f_1}{\partial x_n}\right)_x \\ \vdots & & \vdots \\ \left(\frac{\partial f_p}{\partial x_1}\right)_x & \dots & \left(\frac{\partial f_p}{\partial x_n}\right)_x \end{pmatrix}.$$

1.8 Calcul des dérivées partielles.

Supposons toujours que $x = (x_1, \dots, x_n)$ est un point donné de U, et que f est dérivable en x.

Si on fixe des indices i et j ($1 \leq i \leq n$ et $1 \leq j \leq p$), on peut considérer l'application :

$$t \overset{\varphi}{\longmapsto} f_j(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n),$$

(appelée *fonction partielle à hauteur de x selon la variable x_i*) qui est la composée des trois applications suivantes :

$$t \overset{\iota}{\longmapsto} (x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n), \qquad f, \qquad (y_1, \dots, y_p) \overset{\pi}{\longmapsto} y_j.$$

Il s'agit d'une application d'un ouvert de $\mathbb{R}$ vers $\mathbb{R}$, bien définie au voisinage x_i. Les applications ι et π ci–dessus sont respectivement affine et linéaire. Le théorème de dérivation des fonctions composées montre donc que la dérivée de φ en x_i est la composition d'applications linéaires :

$$\pi \circ f'_x \circ \iota'.$$

Or la matrice de π est la matrice ligne $(0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0)$ (avec le 1 à la $j^{ième}$place), et la matrice de ι' est la matrice colonne :

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(avec le 1 à la $i^{ième}$place). On voit donc tout de suite que la matrice de φ'_{x_i} est la matrice 1×1 suivante :

$$\left(\left(\frac{\partial f_j}{\partial x_i} \right)_x \right).$$

Il en résulte que le calcul de la dérivée partielle $\left(\frac{\partial f_j}{\partial x_i} \right)_x$ se fait en dérivant la fonction d'une seule variable et à valeurs réelles φ.

La notion de dérivée partielle, nous donne un moyen efficace de calculer la dérivée de certaines fonctions, comme par exemple la fonction polynômiale f donnée en exemple précédemment, définie par :

$$f(x, y, z) = x^2 + 3xy - 5xyz.$$

On a très facilement :

$$\frac{\partial f}{\partial x} = 2x + 3y - 5yz, \qquad \frac{\partial f}{\partial y} = 3x - 5xz, \qquad \frac{\partial f}{\partial z} = -5xy,$$

ce qui donne tout de suite la matrice jacobienne. On vérifiera facilement qu'on retrouve ainsi le résultat précédent.

1.9 Un exemple pathologique.

On vient de voir que si f est dérivable en x, elle a des dérivées partielles en x. La réciproque est fausse, comme le montre l'exemple suivant.

Considérons l'application f de $\mathbb{R} \times \mathbb{R}$ vers $\mathbb{R}$ définie par :

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

pour $(x, y) \neq (0, 0)$, et $f(0, 0) = 0$. Intéressons nous à ce qui se passe en $(0, 0)$. La fonction partielle à hauteur de 0, selon la variable x est :

$$x \mapsto 0,$$

puisqu'il suffit de faire $y = 0$ dans la formule donnée. La fonction nulle de $\mathbb{R}$ vers $\mathbb{R}$ est évidemmment dérivable. La dérivée partielle existe donc bien. De la même façon, on peut voir que la dérivée partielle selon la variable y existe et est elle aussi nulle.

Toutefois, la fonction f n'est pas dérivable en 0. En effet, supposons f dérivable en $(0,0)$, et soit $u = f'_{(0,0)} : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ l'application linéaire tangente. Les dérivées partielles étant nulles, la matrice jacobienne de f en $(0,0)$ est nulle, et donc $u = 0$.

En conséquence, la dérivée en 0 de la composée de :

$$x \mapsto (x,x) \qquad \text{et} \qquad f$$

doit elle aussi être nulle. Mais cette dernière fonction est l'identité $x \mapsto x$ de $\mathbb{R}$ et a 1 pour dérivée en 0.

Toutefois, on verra plus loin que si les dérivées partielles sont continues, la fonction est dérivable.

1.10 Dérivée dans la direction d'un vecteur.

Définition 3 *Soit U un ouvert d'un espace de Banach E, et $f : U \longrightarrow F$ une application à valeurs dans l'espace de Banach F. Soit x un point de U, et v un vecteur de E. La dérivée de f en x dans la direction de v est par définition, si elle existe, la limite suivante :*

$$\left(\frac{\partial f}{\partial v}\right)_x = \lim_{h \longrightarrow 0} \frac{f(x+hv) - f(x)}{h}.$$

Autrement–dit, c'est la dérivée $g'(0)$ de la fonction g définie par $g(h) = f(x+hv)$.

Lemme 6 *Dans les mêmes conditions que ci–dessus, et en supposant f dérivable en x, on a :*

$$\left(\frac{\partial f}{\partial v}\right)_x = f'_x(v).$$

En effet, l'application $g : h \mapsto f(x+hv)$ est la composée de l'application affine $h \mapsto x+hv$, et de f. L'application affine en question, qui va de $\mathbb{R}$ vers E, envoie 0 sur x, et a pour dérivée en 0 l'application linéaire $h \mapsto hv$. La dérivée g'_0 de g en 0 est donc, d'après le théorème de dérivation des fonctions composées :

$$h \mapsto f'_x(hv).$$

On a donc $\left(\frac{\partial f}{\partial v}\right)_x = g'(0) = g'_0(1) = f'_x(v).$ □

1.11 Encore des exemples de dérivées.

1.11.1 Le déterminant.

Lemme 7 *Soit $\mathcal{M}$ l'espace des matrices carrées $n \times n$ réelles. L'application "déterminant"* $\det : \mathcal{M} \longrightarrow \mathbb{R}$ *est différentiable, et sa dérivée en A est donnée par la formule :*

$$\det{}'_A(H) = \operatorname{tr}({}^t\tilde{A}H).$$

(où $\tilde{A}$ est la comatrice de A).

Remarquons d'abord que le déterminant étant une application polynômiale, sa dérivée en A existe. On a donc une application linéaire $\det'_A$ de $\mathcal{M}$ vers $\mathbb{R}$, telle que pour tout H, on ait :

$$\det(A+H) = \det(A) + \det'_A(H) + o(H)$$

où $o(H)$ est négligeable devant H quand H tend vers 0.

Une matrice est dite "unicolonne" si toutes ses colonnes sont nulles sauf une. Soit U une matrice unicolonne dont l'unique colonne non nulle est la $i^{ième}$. En développant le déterminant de $A+U$ par rapport à la $i^{ième}$ colonne, on obtient immédiatement :

$$\det(A+U) = \det(A) + \operatorname{tr}({}^t\tilde{A}U).$$

On voit donc que l'appplication linéaire φ définie par $\varphi(H) = \det'_A(H) - \operatorname{tr}({}^t\tilde{A}H)$ vérifie l'égalité : $\varphi(U) = o(U)$ pour tout matrice unicolonne U. Pour tout réel λ non nul, on a donc par linéarité de φ :

$$\varphi(U) = \frac{o(\lambda U)}{\lambda}.$$

Or le membre de droite de cette égalité tend vers 0 quand λ tend vers 0. Il en résulte que φ est nulle sur toute matrice unicolonne. Comme toute matrice est somme de matrices unicolonnes, φ est nulle. □

Dans le cas où A est inversible, cette formule se "simplifie" en :

$$\det'_A(H) = \det(A)\operatorname{tr}(A^{-1}H).$$

car alors

$$\operatorname{tr}({}^t\tilde{A}H) = \operatorname{tr}(A^{-1}A{}^t\tilde{A}H) = \det(A)\operatorname{tr}(A^{-1}H).$$

<u>Exercices</u>

$\boxed{6}$ Soit G un sous–groupe ouvert du groupe des éléments inversibles $\mathcal{A}^*$ d'une algèbre de Banach $\mathcal{A}$. Soit de même G' un sous–groupe ouvert du groupe des éléments inversibles $\mathcal{A}'^*$ d'une algèbre de Banach $\mathcal{A}'$. Soit $g : G \longrightarrow G'$ un morphisme de groupes supposé continu.

Montrer que si g est dérivable en 1, alors il est dérivable en tout point x de G, et on a :

$$g'_x(h) = g(x)g'_1(x^{-1}h) = g'_1(hx^{-1})g(x).$$

$\boxed{7}$ Le but de cet éxercices est de montrer que $\forall x \in \mathbb{R},\ \det(e^{xA}) = e^{\operatorname{tr}(A)}$.

On considère $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices $n \times n$ à coéfficients réels et soit $\mathrm{GL}_n(\mathbb{R})$ l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ inversibles. Soit $A \in \mathcal{M}_n(\mathbb{R})$, on considère les applications :

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\varphi} & \mathrm{GL}_n(\mathbb{R}) \\ x & \longmapsto & e^{xA} \end{array} \quad \text{et} \quad \begin{array}{ccc} \mathbb{R} & \xrightarrow{\psi} & \mathbb{R}_+^* \\ x & \longmapsto & \det(\varphi(x)) = \det(e^{xA}) \end{array}$$

1. Montrer que φ est dérivable en 0 et que $\varphi'(0) = A$.
2. Montrer que ψ est dérivable en 0 et que $\psi'(0) = \operatorname{tr}(A)$.
3. En déduire que $\det(e^{xA}) = e^{\operatorname{tr}(A)x}$ pour tout $x \in \mathbb{R}$.

$\boxed{8}$ Soit l'application :

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R} \\ (x,y) & \longmapsto & \arctan x + \arctan y - \arctan\left(\dfrac{x+y}{1-xy}\right) \end{array}$$

On note $D = \mathbb{R}^2 \backslash \{(x, y) \in \mathbb{R}^2,\ xy = 1\}$ le domaine de définition de f.

Montrer que $f(x, y)$ est égal soit à 0, soit à π, soit à $-\pi$.

2 Le théorème de la moyenne (ou des accroissements finis).

Le théorème de la moyenne a une interprétation *mécanique* simple. Soit A un point mobile dont la trajectoire est une courbe de l'espace $\mathbb{R}^3$. Soit B un deuxième point mobile dont la trajectoire est une courbe de $\mathbb{R}$ (c'est–à–dire une ligne droite). On suppose qu'à tout moment, la vitesse de B est positive et plus grande que le module de la vitesse de A. Alors, à tout moment, B est plus éloigné de son point de départ que ne l'est A. Bien sûr, ceci tient non seulement au fait que B a une vitesse plus élevée que A, mais aussi au fait que B se déplace en ligne droite, alors que A peut par exemple tourner en rond.

Pourtant, ce théorème "évident", qui s'appelle *théorème de la moyenne*, ou *théorème des accroissements finis*, n'est pas si simple à démontrer. La raison en est que les hypothèses portent sur la vitesse (une notion "infinitésimale"), alors que la conclusion porte sur les distances parcourues (une notion "globale"). Pour passer de l'hypothèse à la conclusion, il est donc nécessaire de passer de l'infinitésimal au global (ou de l'infinitésimal au *fini*, d'où le nom de théorème des *accroissements finis*).

2.1 Énoncé et démonstration

Théorème 2 *(théorème de la moyenne) Soit E un espace de Banach. Soit $[a, b]$ $(a < b)$ un intervalle compact de $\mathbb{R}$. Soient :*

$$f : [a, b] \longrightarrow E \qquad \text{et} \qquad g : [a, b] \longrightarrow \mathbb{R}$$

deux applications continues sur $[a, b]$ et dérivables sur $]a, b[$, telles que $\forall x \in]a, b[\ \ \|f'(x)\| \leq g'(x)$. Alors, on a :

$$\|f(b) - f(a)\| \leq g(b) - g(a).$$

On remarquera l'absence de signes de valeurs absolues autour de $g'(x)$ et de $g(b) - g(a)$.

Soit $\varepsilon > 0$. Il suffit de montrer que :

$$\|f(b) - f(a)\| \leq g(b) - g(a) + \varepsilon(b - a) + \varepsilon,$$

car ε peut être pris aussi petit qu'on veut.

Soit A l'ensemble des x de $[a, b]$ tels que :

$$\|f(x) - f(a)\| > g(x) - g(a) + \varepsilon(x - a) + \varepsilon.$$

Si on démontre que A est vide, alors b ne pourra pas appartenir à A, et on aura démontré l'inégalité voulue. On peut donc supposer A non vide.

À cause du terme ε à la fin du second membre de l'inégalité définissant A, et à cause de la continuité des fonctions f et g et de la fonction $x \mapsto \varepsilon(x - a)$, on voit qu'il existe un voisinage de a ne contenant aucun point de A.

Comme A est non vide et minoré par a, A a une borne inférieure. Notons c cette borne inférieure. On a $a < c$ d'après ce qui vient d'être dit plus haut.

Par ailleurs, c ne peut pas être dans A. En effet, s'il y était, un voisinage de c dans $[a, b]$ serait inclus dans A, car A est défini par une inégalité *stricte* dont les deux membres sont des fonctions continues de x.

c est donc distinct de b, car sinon, A serait réduit au point b, et c serait dans A.

On a donc $a < c < b$, et on peut appliquer l'hypothèse au point c, c'est–à–dire qu'on a $\|f'(c)\| \leq g'(c)$.

Comme c n'est pas dans A, on a :

$$\|f(c) - f(a)\| \leq g(c) - g(a) + \varepsilon(c-a) + \varepsilon.$$

Écrivons que f et g sont dérivables en c :

$$f(c+h) = f(c) + hf'(c) + o_1(h), \qquad g(c+h) = g(c) + hg'(c) + o_2(h).$$

où $o_1(h)$ et $o_2(h)$ sont négligeables devant h, quand h tend vers 0.

Il suffit maintenant de prouver que pour $h > 0$ assez petit, $c + h$ n'est pas dans A, car ceci sera en contradiction avec le fait que c est la borne inférieure de A. Or, on a (pour $h > 0$) :

$$\begin{array}{rcl} \|f(c+h) - f(a)\| & \leq & \|f(c+h) - f(c)\| + \|f(c) - f(a)\| \\ & \leq & \|hf'(c)\| + \|o_1(h)\| + g(c) - g(a) + \varepsilon(c-a) + \varepsilon \\ & \leq & hg'(c) + g(c) - g(a) + \varepsilon(c-a) + \varepsilon + \|o_1(h)\| \\ & \leq & g(c+h) - g(a) + \varepsilon(c-a) + \varepsilon + \|o_1(h)\| - o_2(h). \end{array}$$

Or $\|o_1(h)\| - o_2(h)$ étant négligeable devant h, quand h tend vers 0, on peut choisir h assez petit pour que cette quantité soit plus petite que εh. $\square$

2.2 Applications du théorème de la moyenne.

2.2.1 Fonctions lipschitziennes.

Lemme 8 *Soit $f : [a,b] \longrightarrow F$ une application continue d'un intervalle compact de $\mathbb{R}$ vers un espace de Banach F. On suppose que f est dérivable sur $]a,b[$. Alors, s'il existe un réel k tel que $\forall x \in]a,b[\ \|f'(x)\| \leq k$, l'application f est k–lipschitzienne sur $[a,b]$.*

Le théorème de la moyenne appliqué sur un intervalle quelconque $[u,v]$ contenu dans $[a,b]$, en prenant pour g la fonction $x \mapsto kx$, montre que $\|f(u) - f(v)\| \leq k(v-u)$. $\square$

Théorème 3 *Soit U une partie ouverte et convexe d'un espace de Banach E. Soit $f : U \longrightarrow F$ une fonction continue et dérivable, à valeurs dans un espace de Banach F. Alors, s'il existe un réel k tel que $\forall x \in U\ \|f'_x\| \leq k$, f est k–lipschitzienne sur U.*

Soient u et v deux point de U. Comme U est convexe, la fonction affine φ définie par $t \mapsto (1-t)u + tv$ envoie l'intervalle $[0,1]$ dans U. La dérivée de $\varphi'(t)$, en un point quelconque de $[0,1]$ est le vecteur $v - u$. La fonction $f \circ \varphi$ est continue sur $[0,1]$ et dérivable sur $]0,1[$. Sa dérivée en un point t est $f'_{\varphi(t)}(v-u)$. Ce vecteur a une norme inférieure ou égale à $\|f'_{\varphi(t)}\|\,\|v-u\|$, donc à $k\|v-u\|$. Le lemme précédent appliqué à $f \circ \varphi$ et aux point 0 et 1 de $[0,1]$, donne donc $\|f(v) - f(u)\| \leq k\|v-u\|$. $\square$

Note : il arrive souvent que la condition $\|f'_x\| \leq k$ soit obtenue en utilisant la continuité de f' sur un compact.

Corollaire 1 *Soit U un ouvert convexe d'un espace de Banach E. Soit $f : U \longrightarrow F$ une fonction dérivable à valeurs dans un espace de Banach F. Si f'_x est nul pour tout x de U, la fonction f est constante.*

On applique le théorème précédent avec $k = 0$. □

Note : L'hypothèse que U est convexe est indispensable. En effet, une fonction dérivable sur un ouvert U quelconque, et dont la dérivée est nulle n'est pas nécessairement constante. Elle est seulement *localement* constante.

2.2.2 Convergence d'une suite de fonctions dérivables.

On sait que si une suite de fonctions dérivables converge uniformément, il n'y a aucune raison qu'il en soit de même de la suite de leurs dérivées. Par exemple, la suite de fonctions $(f_n)_n$ définies par :

$$x \mapsto \frac{1}{n}\sin(nx),$$

converge uniformément vers 0 (sur un intervalle quelconque de $\mathbb{R}$). Par contre, la suite des dérivées :

$$x \mapsto \cos(nx),$$

ne converge pas (même simplement).

Par contre, comme on va le voir, la convergence de la suite des dérivées, implique (sous certaines conditions) la convergence de la suite de fonctions.

Théorème 4 *Soit U un ouvert convexe d'un espace de Banach E, et x_0 un point de U. Soit $(f_n)_n$ une suite de fonctions dérivables sur U, à valeurs dans un espace de Banach F. On suppose que :*
- *La suite d'éléments de F : $(f_n(x_0))_n$ converge.*
- *La suite de fonctions : $x \mapsto (f_n)'_x$ converge uniformément sur U vers une fonction $g : U \longrightarrow \mathcal{L}(E, F)$.*

Alors, la suite $(f_n)_n$ converge vers une fonction f. La convergence est uniforme sur toute partie bornée de U. De plus, f est dérivable sur U, et on a $f' = g$.

On commence par prouver l'existence de la fonction f. Pour cela, il suffit de montrer que pour chaque x de U, la suite d'éléments de F : $(f_n(x))_n$ est une suite de Cauchy.

La suite de fonctions $(f'_n)_n$ étant uniformément convergente sur U, elle est uniformément de Cauchy sur U. En particulier, on peut poser :

$$k_{p,q} = \sup_{x \in U} \|(f_p)'_x - (f_q)'_x\|,$$

et de plus $k_{p,q}$ tend vers 0 quand p et q tendent vers l'infini.

Comme U est convexe, et $f'_p - f'_q$ borné sur U par la constante $k_{p,q}$, $f_p - f_q$ est $k_{p,q}$–lipschitzienne :

$$\|f_p(x) - f_q(x) - (f_p(x_0) - f_q(x_0))\| \leq k_{p,q}\|x - x_0\|.$$

Si on suppose maintenant que x reste dans une partie bornée B de U, $\|x - x_0\|$ est majoré par une constante d_B indépendante de x.

Soit $\varepsilon > 0$. Pour p et q assez grands, on a $k_{p,q} < \varepsilon$, et $\|f_p(x_0) - f_q(x_0)\| < \varepsilon$, puisque la suite $(f_n(x_0))_n$ est convergente. On a donc :

$$\|f_p(x) - f_q(x)\| \leq \varepsilon(d_B + 1),$$

ce qui montre que $(f_n(x))_n$ est une suite de Cauchy, et que la convergence est uniforme sur toute partie bornée de U. Notons $f(x)$ la limite de $f_n(x)$ quand n tend vers l'infini. La fonction f est continue car elle est localement limite uniforme de fonctions continues (chaque point de U a un voisinage borné).

Il nous reste donc à montrer que f est dérivable sur U, et que sa dérivée est g. Soit donc x un point de U. Nous devons montrer que

$$f(x+h) = f(x) + g(x)(h) + o(h).$$

C'est–à–dire que $\|f(x+h) - f(x) - g(x)(h)\|$ est négligeable devant h. On a :

$$\begin{array}{rcl} \|f(x+h) - f(x) - g(x)(h)\| & \leq & \|f(x+h) - f(x) - (f_n(x+h) - f_n(x))\| \\ & + & \|f_n(x+h) - f_n(x) - (f_n)'_x(h)\| \\ & + & \|(f_n)'_x(h) - g(x)(h)\|. \end{array}$$

Soit $\varepsilon > 0$. Prenons n et p assez grands pour que $\|f_p(x+h) - f_p(x) - (f_n(x+h) - f_n(x))\| \leq k_{p,n}\|h\|$, et pour que $k_{p,n} < \varepsilon$. Quand p tend vers l'infini, $f_p(x)$ tend vers $f(x)$, et $f_p(x+h)$ vers $f(x+h)$. On a donc :

$$\|f(x+h) - f(x) - (f_n(x+h) - f_n(x))\| \leq \varepsilon\|h\|,$$

dès que n est supérieur ou égal à un certain entier n_0.

Comme chaque f_n est dérivable en x, on a :

$$f_n(x+h) - f_n(x) - (f_n)'_x(h) = o_n(h).$$

De plus, comme $(f_n)'_x$ et $g(x)$ sont des applications linéaires continues, on a $\|(f_n)'_x(h) - g(x)(h)\| \leq \|(f_n)'_x - g(x)\|\,\|h\|$. Comme $(f'_n)_n$ converge uniformément vers g, on a $\|(f_n)'_x - g(x)\| \leq \varepsilon$ dès que n est assez grand (disons supérieur ou égal à n_1, lui même supérieur ou égal à n_0).

On a alors, pour tout h :

$$\|f(x+h) - f(x) - g(x)(h)\| \leq 2\varepsilon\|h\| + \|o_{n_1}(h)\|.$$

Soit maintenant un $\eta > 0$ tel que $\|o_{n_1}(h)\| \leq \varepsilon\|h\|$, dès que $\|h\| \leq \eta$. On a alors montré que pour tout $\varepsilon > 0$, il existe $\eta > 0$, tel que $\|f(x+h) - f(x) - g(x)(h)\| \leq 3\varepsilon\|h\|$ dès que $\|h\| \leq \eta$. $\square$

2.2.3 Dérivabilité et dérivées partielles.

Rappelons qu'on dit qu'une fonction est de classe $\mathcal{C}^1$ sur l'ouvert U, si elle est dérivable sur U et si sa dérivée f' est continue sur U.

Si une fonction est dérivable en un point, elle a des dérivées partielles en ce point. La réciproque est fausse comme on l'a vu sur un exemple. Toutefois, on a quand même la propriété suivante :

Théorème 5 *Soit $f : U \longrightarrow \mathbb{R}^p$ une fonction continue définie sur un ouvert U de l'espace de banach $\mathbb{R}^n$. Pour que f soit de classe $\mathcal{C}^1$ sur U, il faut et il suffit que ses dérivées partielles existent et soient continues sur U.*

Si f est de classe $\mathcal{C}^1$, ses dérivées partielles sont continues, puisqu'une composition de fonctions de classe $\mathcal{C}^1$ est de classe $\mathcal{C}^1$.

Réciproquement, supposons les dérivées partielles continues. On peut très bien supposer que f est à valeurs dans $\mathbb{R}$ (ce qui revient à se limiter à une composante de f). Les dérivées partielles sont donc les fonctions $\dfrac{\partial f}{\partial x_i}$ $(1 \leq i \leq n)$, toutes définies sur U.

Soit $x = (x_1, \dots, x_n)$ un point de U, et $h = (h_1, \dots, h_n)$ un vecteur de $\mathbb{R}^n$. On veut montrer que

$$f(x+h) - f(x) - \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_x h_i = o(h).$$

Posons $y_i(t) = (x_1, \dots, x_{i-1}, x_i + th_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n)$, et

$$g_i(t) = f(y_i(t)) - \left(\frac{\partial f}{\partial x_i}\right)_x th_i.$$

Calculons maintenant $g_i(1) - g_i(0)$:

$$g_i(1) - g_i(0) = f(y_i(1)) - f(y_i(0)) - \left(\frac{\partial f}{\partial x_i}\right)_x h_i.$$

En remarquant que $y_i(0) = y_{i+1}(1)$, on voit que :

$$\sum_{i=1}^{n} (g_i(1) - g_i(0)) = f(y_1(1)) - f(y_n(0)) - \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_x h_i,$$

c'est–à–dire exactement l'expression dont on doit montrer qu'elle est négligeable devant h. Il suffit donc de montrer que chacune des expressions $g_i(1) - g_i(0)$ est négligeable devant h.

La fonction g_i est dérivable et sa dérivée est :

$$g_i'(t) = \left(\frac{\partial f}{\partial x_i}\right)_{y_i(t)} h_i - \left(\frac{\partial f}{\partial x_i}\right)_x h_i.$$

Comme la dérivée partielle $\dfrac{\partial f}{\partial x_i}$ est continue sur U, et comme $\|y_i(t) - x\| \leq \|h\|$ (en supposant que la norme sur $\mathbb{R}^n$ soit la norme euclidienne), on voit que pour tout $\varepsilon > 0$, il existe $\eta > 0$, tel que :

$$\left\| \left(\frac{\partial f}{\partial x_i}\right)_{y_i(t)} - \left(\frac{\partial f}{\partial x_i}\right)_x \right\| \leq \varepsilon,$$

dès que $\|h\| \leq \eta$. Pour $\|h\| \leq \eta$, on a donc $\|g_i'(t)\| \leq \varepsilon \|h\|$. Le théorème de la moyenne permet alors de conclure. □

2.2.4 Dérivée d'une curryfiée et règle de Leibnitz.

Soit U un ouvert d'un espace de Banach E, F un espace de Banach, et K un espace compact. Soit enfin $\varphi : U \times K \longrightarrow F$ une application continue. On appelle "curryfiée de φ" l'application $\psi : U \longrightarrow \mathcal{C}(K, F)$ de U vers l'espace des applications continues de K vers F, définie par l'une quelconque des formules suivantes, qui sont équivalentes :

$$\psi(x)(t) = \varphi(x,t) \qquad \psi(x) = t \mapsto \varphi(x,t) \qquad \psi = x \mapsto (t \mapsto \varphi(x,t))$$

On notera que K étant compact, L'espace vectoriel réel $\mathcal{C}(K, F)$ est un espace de Banach quand on le munit de la norme de la convergence uniforme. De plus, φ étant continue, l'application $t \mapsto \varphi(x,t)$ est continue, et ψ est donc bien à valeurs dans $\mathcal{C}(K, F)$.

Lemme 9 *ψ est continue.*

Il s'agit de prouver que, pour tout $x_0 \in U$:

$$\forall_{\varepsilon>0}\ \exists_{\eta>0}\ \forall_{x\in U}\ \|x-x_0\|<\eta \Rightarrow \|\psi(x)-\psi(x_0)\|<\varepsilon.$$

Soit $\varepsilon > 0$. Comme φ est continue, on a pour chaque point t de K, un $\eta_t > 0$, et un voisinage ouvert V_t de t dans K, tels que pour tous $x \in U$ et $t' \in K$:

$$\|x-x_0\|<\eta_t \wedge t'\in V_t \Rightarrow \|\varphi(x,t')-\varphi(x_0,t)\|<\varepsilon/2.$$

La famille d'ouverts $(V_t)_{t\in K}$ recouvre K. Il en est donc de même d'une sous–famille finie $V_{t_1},\dots,V_{t_p}$ de cette famille. Posons $\eta = \inf(\eta_{t_1},\dots,\eta_{t_p})$. On a $\eta > 0$. Il nous reste à prouver que :

$$\forall_{x\in U}\ \|x-x_0\|<\eta \Rightarrow \|\psi(x)-\psi(x_0)\|<\varepsilon,$$

c'est à dire que :

$$\forall_{x\in U}\ \|x-x_0\|<\eta \Rightarrow \forall_{t\in K}\ \|\psi(x)(t)-\psi(x_0)(t)\|<\varepsilon.$$

Soit donc $x \in U$, tel que $\|x-x_0\| < \eta$ et soit $t \in K$. Il reste à prouver que $\|\varphi(x,t)-\varphi(x_0,t)\| < \varepsilon$. Il existe un t_i $(1 \le i \le p)$, tel que $t \in V_{t_i}$. On a donc $\|\varphi(x,t)-\varphi(x_0,t_i)\| < \varepsilon/2$. Pour la même raison on a aussi $\|\varphi(x_0,t)-\varphi(x_0,t_i)\| < \varepsilon/2$, puisque $\|x_0-x_0\| = 0 < \eta$. Donc :

$$\|\varphi(x,t)-\varphi(x_0,t)\| \le \|\varphi(x,t)-\varphi(x_0,t_i)\| + \|\varphi(x_0,t_i)-\varphi(x_0,t)\| < \varepsilon/2+\varepsilon/2 = \varepsilon.\ \square$$

Lemme 10 *(Dérivée d'une curryfiée) Si U est un ouvert d'un espace de Banach E, K un espace compact, F un espace de Banach, $\varphi : U\times K \longrightarrow F$ une application continue, dont la dérivée partielle $d_1\varphi$ par rapport à la première variable (celle qui appartient à U) existe et est continue, alors la curryfiée $\psi : U \longrightarrow \mathcal{C}(K,F)$ de φ est dérivable, et sa dérivée en $x \in U$, $\psi'_x : E \longrightarrow \mathcal{C}(K,F)$ est donnée par :*

$$\psi'_x(h) = t \mapsto ((d_1\varphi)_{(x,t)}(h))$$

Par définition, $(d_1\varphi)_{(x,t)}$ est la dérivée en x de la fonction composée $\varphi\circ\alpha_t$, où α_t est définie par $\alpha_t(x) = (x,t)$, c'est à dire $(d_1\varphi)_{(x,t)} = (\varphi\circ\alpha_t)'_x$.

La seule chose qu'on ait à prouver est que :

$$\psi(x+h)-\psi(x)-(t\mapsto(\varphi\circ\alpha_t)'_x(h))$$

est négligeable devant h, quand h tend vers 0. Cette expression est égale à :

$$t\mapsto \varphi(x+h,t)-\varphi(x,t)-(\varphi\circ\alpha_t)'_x(h).$$

Soit $\varepsilon > 0$. Il s'agit de trouver $\eta > 0$, tel que :

$$\|h\|<\eta \Rightarrow \|t\mapsto\varphi(x+h,t)-\varphi(x,t)-(\varphi\circ\alpha_t)'_x(h)\|<\varepsilon\|h\|.$$

La norme sur les fonctions continues de K vers F étant celle de la convergence uniforme, cette dernière implication est équivalente à :

$$\|h\|<\eta \Rightarrow \sup_{t\in K}\|\varphi(x+h,t)-\varphi(x,t)-(\varphi\circ\alpha_t)'_x(h)\|<\varepsilon\|h\|.$$

La question essentielle est donc de majorer l'expression $\|\varphi(x+h,t)-\varphi(x,t)-(\varphi\circ\alpha_t)'_x(h)\|$ d'une manière indépendante de t. Or cette expression n'est autre que :

$$\|(\varphi\circ\alpha_t)(x+h)-(\varphi\circ\alpha_t)(x)-(\varphi\circ\alpha_t)'_x(h)\|$$

Comme $\varphi\circ\alpha_t$ est par hypothèse dérivable en x, on a :

$$(\varphi\circ\alpha_t)(x+h)-(\varphi\circ\alpha_t)(x)-(\varphi\circ\alpha_t)'_x(h) = o_{x,t}(h),$$

avec $\lim\limits_{h \longrightarrow 0} \dfrac{\|o_{x,t}(h)\|}{\|h\|} = 0$, le point délicat étant que $o_{x,t}(h)$ dépend de t. Si ce n'était pas le cas, la démonstration serait finie ici.

On a :

$$\begin{aligned} o_{x,t}(h) &= (\varphi \circ \alpha_t)(x+h) - (\varphi \circ \alpha_t)(x) - (\varphi \circ \alpha_t)'_x(h) \\ &= [(\varphi \circ \alpha_t)(x+h) - (\varphi \circ \alpha_t)'_x(h)] - [(\varphi \circ \alpha_t)(x) - (\varphi \circ \alpha_t)'_x(0)] \\ &= \beta_t(h) - \beta_t(0), \end{aligned}$$

où on a posé $\beta_t(h) = (\varphi \circ \alpha_t)(x+h) - (\varphi \circ \alpha_t)'_x(h)$.

Pour chaque $t \in K$, la fonction β_t est dérivable sur un voisinage de 0 dans E, puisque pour h assez petit, $x + h$ appartient à U (U est ouvert). Sa dérivée en h est :

$$(\beta_t)'_h = (\varphi \circ \alpha_t)'_{x+h} - (\varphi \circ \alpha_t)'_x.$$

Par hypothèse, la fonction $(x,t) \mapsto (\varphi \circ \alpha_t)'_x$ est continue sur $U \times K$. Il résulte du lemme précédent que sa curryfiée $x \mapsto (t \mapsto (\varphi \circ \alpha_t)'_x)$ est continue sur U, et donc que la fonction :

$$h \mapsto (t \mapsto (\beta_t)'_h)$$

est continue sur un voisinage de 0 dans E. Or, cette dernière fonction vaut 0 (le 0 de l'espace de Banach $\mathcal{C}(K, \mathcal{L}(E,F))$) en 0 (le 0 de E). Il existe donc $\eta > 0$, tel que :

$$\|h\| < \eta \Rightarrow \forall_{t \in K} \ \|(\beta_t)'_h\| < \varepsilon.$$

Le théorème de la moyenne nous montre donc que :

$$\|h\| < \eta \Rightarrow \forall_{t \in K} \ \|\beta_t(h) - \beta_t(0)\| < \varepsilon \|h\|,$$

c'est à dire :

$$\|h\| < \eta \Rightarrow \|t \mapsto o_{x,t}(h)\| < \varepsilon \|h\|. \ \square$$

Corollaire 2 *(Règle de Leibnitz) Soit U un ouvert d'un espace de Banach E, $[a,b]$ un intervalle compact de $\mathbb{R}$, F un espace de Banach, et $\varphi : U \times [a,b] \longrightarrow F$ une fonction continue, ayant une dérivée partielle $d_1\varphi$ par rapport à la première variable, continue sur $U \times [a,b]$. Alors, la fonction θ définie par :*

$$\theta(x) = \int_a^b \varphi(x,t) \ dt$$

est dérivable sur U et sa dérivée en un point x de U est donnée par :

$$\theta'_x = \int_a^b (d_1\varphi)_{(x,t)} \ dt.$$

On a $\theta = I \circ \psi$, où ψ est la curryfiée de φ, et où $I(f) = \displaystyle\int_a^b f(t) \ dt$, pour toute $f \in \mathcal{C}([a,b], F)$. On sait par le lemme précédent que ψ est dérivable. Par ailleurs, I est dérivable, puisque c'est une application linéaire continue (on a d'ailleurs $\|I\| = |b-a|$). Il en résulte que θ est dérivable, et que

$$\theta'_x(h) = I(\psi'_x(h)) = \int_a^b (d_1\varphi)_{(x,t)}(h) \ dt = \left(\int_a^b (d_1\varphi)_{(x,t)} \ dt \right)(h).$$

On remarquera que la première des deux intégrales ci–dessus est celle d'une fonction à valeurs dans F, alors que la deuxième est celle d'une fonction à valeurs dans $\mathcal{L}(E,F)$. La dernière égalité est légitimée par le fait que l'évaluation en h : $f \mapsto f(h)$ de $\mathcal{L}(E,F)$ vers F est linéaire continue. $\square$

Exercices

$\boxed{\mathbf{9}}$ *Soit $f : U \longrightarrow F$ une application de classe $\mathcal{C}^1$ d'un ouvert U d'un espace de Banach E non réduit à 0, vers un espace de Banach F. Soit x_0 un point de U.*

a) Soit k un réel strictement positif. Montrer qu'il existe un voisinage V de x_0 (contenu dans U), sur lequel la fonction $\psi = (x \mapsto f(x) - f(x_0) - f'_{x_0}(x - x_0))$ est k–lipschitzienne.

On suppose qu'il existe une application linéaire continue $g : F \longrightarrow E$, telle que $g \circ f'_{x_0}$ soit l'application identique de E.

b) Montrer que la norme de g n'est pas nulle, et que $\|f'_{x_0}(h)\| \geq \dfrac{\|h\|}{\|g\|}$, pour tout h de E.

c) Montrer qu'il existe un voisinage V de x_0 dans U, sur lequel f est injective.

$\boxed{\mathbf{10}}$ Montrer que le système :

$$\begin{cases} x &= \dfrac{1}{4}\sin(x+y) \\ y &= 1 + \dfrac{2}{3}\arctan(x-y) \end{cases}$$

admet une unique solution sur $\mathbb{R}^2$.

$\boxed{\mathbf{11}}$ On considère $\mathcal{C}([0,1], \mathbb{R})$ l'espace vectoriel réel des fonctions continue de $[0,1]$ dans $\mathbb{R}$ et on le munit de la norme 1 noté $\|\ \|_1$ ($\|f\|_1 = \displaystyle\int_0^1 f(t)dt$).

Soit la fonction :

$$\begin{array}{ccc}]0,+\infty[& \xrightarrow{\varphi} & \mathcal{C}\left([0,1],\mathbb{R}\right) \\ \alpha & \longmapsto & \left(t \mapsto t^{\alpha}\right) \end{array}$$

Montrar que φ est dérivable et calculer φ', puis montrer que φ' est 2-lipschtzienne.

3 Dérivées d'ordre supérieur.

3.1 Définition.

On a vu que si f est une application dérivable de l'ouvert U de l'espace de Banach E, vers l'espace de Banach F, alors pour tout point x de E on a une *dérivée en x* ou *application linéaire tangente en x* : $f'_x \in \mathcal{L}(E,F)$. On a appelé *dérivée de f* la fonction $f' = (x \mapsto f'_x)$ de U vers $\mathcal{L}(E,F)$, qui à tout x de U associe l'application linéaire tangente à f en x.

f' est donc une application de l'ouvert U vers l'espace vectoriel $\mathcal{L}(E,F)$. Comme E et F sont des espaces de Banach, il en est de même de $\mathcal{L}(E,F)$, et cela a donc un sens de se demander si f' est dérivable en un point x de U.

Si tel est le cas, l'application linéaire tangente à f' en x sera notée f''_x. C'est une application linéaire de E vers $\mathcal{L}(E,F)$, autrement–dit un élément de $\mathcal{L}(E,\mathcal{L}(E,F))$.

Par ailleurs, l'espace $\mathcal{L}(E,\mathcal{L}(E,F))$ est canoniquement isomorphe à l'espace des applications bilinéaires de $E\times E$ vers F. La correspondance canonique (curryfication/décurryfication) est donnée par les formules :

$$\begin{array}{rcl} \mathcal{L}(E,\mathcal{L}(E,F)) & \longrightarrow & \mathcal{L}(E,E;F) \\ f & \longmapsto & (x,y)\mapsto f(x)(y) \end{array} \qquad \begin{array}{rcl} \mathcal{L}(E,E;F) & \longrightarrow & \mathcal{L}(E,\mathcal{L}(E,F)) \\ g & \longmapsto & x\mapsto (y\mapsto g(x,y)) \end{array}$$

On peut donc considérer f''_x comme une application bilinéaire de $E\times E$ vers F. En principe on devrait adopter une autre notation que f''_x pour cette application bilinéaire, mais aucune confusion n'est à craindre, car le contexte indiquera toujours clairement s'il s'agit d'un élément de $\mathcal{L}(E,\mathcal{L}(E,F))$ ou d'un élément de $\mathcal{L}(E,E;F)$. En particulier, si h et k sont deux vecteurs de E, les deux expressions :

$$f''_x(h)(k) \qquad \text{et} \qquad f''_x(h,k),$$

représentent le même élément de F. La seule différence est que dans la première expression f''_x est considéré comme un élément de $\mathcal{L}(E,\mathcal{L}(E,F))$, alors que dans la seconde expression, f''_x est considéré comme un élément de $\mathcal{L}(E,E;F)$.

Comme on l'a fait pour f', on peut maintenant considérer l'application :

$$\begin{array}{rcl} U & \longrightarrow & \mathcal{L}(E,E;F) \\ x & \longmapsto & f''_x \end{array}$$

qu'on notera f'', et qu'on appelera *dérivée seconde de f*.

Comme $\mathcal{L}(E,E;F)$ est encore un espace de Banach, on peut envisager de dériver f'', si f'' est dérivable. Ceci nous donnera une application $f''' : U \longrightarrow \mathcal{L}(E,\mathcal{L}(E,E;F))$. Comme $\mathcal{L}(E,\mathcal{L}(E,E;F))$ s'identifie canoniquement à $\mathcal{L}(E,E,E;F)$, on peut considérer f'''_x comme une application *trilinéaire* de $E\times E\times E$ vers F.

D'une manière générale, on pourra considérer, si elle existe, la dérivée $n^{ième}$de f, qui sera une application de U vers l'espace des applications n–multilinéaires de E^n vers F. La dérivée $n^{ième}$de f sera notée $f^{(n)}$. En particulier, on peut utiliser les notations $f^{(0)}$, $f^{(1)}$, $f^{(2)}$, etc... pour représenter f, f', f'', etc...

Définition 4 *Une application f d'un ouvert U d'un espace de Banach E vers un espace de Banach F est dite* de classe $\mathcal{C}^n$ *(sur U), si elle admet une dérivée $n^{ième}$continue (sur U). On dit aussi que f est n fois continument dérivable (ou différentiable) sur U. Elle est dite de classe $\mathcal{C}^\infty$, si elle est de classe $\mathcal{C}^n$ pour tout entier naturel n.*

Si une fonction f est de classe $\mathcal{C}^n$, elle admet donc des dérivées successives f', f'',...,$f^{(n)}$, toutes continues sur U. Ces dérivées sont à valeurs dans des espaces de Banach de plus en plus gros, sauf si la dimension de E est 1.

Définition 5 *Une application f d'un ouvert U d'un espace de banach E vers un espace de Banach F est dite n fois dérivable en x, si elle est de classe $\mathcal{C}^{n-1}$ dans un voisinage de x, et si $f^{(n-1)}$ est dérivable en x.*

3.2 Exemples de dérivées d'ordre supérieur.

3.2.1 Applications linéaires.

Une application linéaire continue $f : E \longrightarrow F$ a elle–même pour dérivée en tout point x de E. Ceci signifie que la dérivée f' de f est l'application *constante* $x\mapsto f'_x = f$, de E vers $\mathcal{L}(E,F)$.

La dérivée seconde f'' est donc nulle en tout point x, de même que les dérivées suivantes.

3.2.2 Applications bilinéaires.

On a vu que la dérivée au point (x, y) de l'application bilinéaire continue $f : E \times F \longrightarrow G$ est :

$$(h, k) \mapsto f(h, y) + f(x, k).$$

Or il est clair que $f' : E \times F \longrightarrow \mathcal{L}(E \times F, G)$, qui n'est autre que :

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

est linéaire continue, et a donc elle–même pour dérivée en tout point (x, y). Ceci donne :

$$f''_{(x,y)}((h, k), (u, v)) = f(h, v) + f(u, k).$$

On notera que $f''_{(x,y)}$ est une application bilinéaire de $(E \times F) \times (E \times F)$ vers G.

L'application f'' est constante, puisque $f''_{(x,y)}$ ne dépend pas de (x, y), ce qui montre que f''' est nulle, de même que toutes les dérivées suivantes.

La formule ci–dessus met clairement en évidence le fait que f'' est une application bilinéaire *symétrique*. En effet, on a :

$$f''((h, k), (u, v)) = f(h, v) + f(u, k) = f(u, k) + f(h, v) = f''((u, v), (h, k)).$$

La forme quadratique (en supposant que $G = \mathbb{R}$) associée à f'' est donc :

$$(h, k) \mapsto f''((h, k), (h, k)) = 2f(h, k).$$

Remarque : Toute forme bilinéaire f (même non symétrique) sur $E \times F$ est une *forme quadratique sur* $E \times F$, dont la forme polaire est la forme bilinéaire *symétrique* sur $(E \times F) \times (E \times F)$ donnée par :

$$((h, k), (u, v)) \mapsto \frac{1}{2}(f(h, v) + f(u, k)).$$

3.2.3 Composition avec une fonction affine.

Lemme 11 *Si f est affine, et g n fois dérivable en $f(x)$, on a :*

$$(g \circ f)^{(n)}_x = g^{(n)}_{f(x)} \circ (f' \times \cdots \times f').$$

Si g est affine, et f n fois dérivable en x, on a :

$$(g \circ f)^{(n)}_x = g' \circ f^{(n)}_x.$$

On a écrit f' au lieu de f'_x, car f étant affine, sa dérivée première est constante, de même pour g'. La vérification se fait par récurrence sur n. □

3.2.4 Fonctions composées.

Théorème 6 *Soient E, F et G des espaces de Banach. Soit U un ouvert de E, et V un ouvert de F. Soit $x \in U$. Soient $f : U \longrightarrow V$, et $g : V \longrightarrow G$ des applications continues. Alors, si f est n fois dérivable en x, et si g est n fois dérivable en $f(x)$, $g \circ f$ est n fois dérivable en x.*

Ceci se montre par récurrence sur n. Le cas $n = 1$ résulte immédiatement du théorème de dérivation des fonctions composées.

Supposons donc $n \geq 2$, et le théorème valable jusqu'à l'ordre $n - 1$. On a pour tout y de U :

$$(g \circ f)'_y = g'_{f(y)} \circ f'_y.$$

La fonction $y \mapsto (g \circ f)'_y$ est donc la composées des deux fonctions suivantes :

$$y \mapsto (g'_{f(y)}, f'_y) \qquad \text{et} \qquad (u, v) \mapsto u \circ v.$$

g étant n fois dérivable en $f(x)$, g' est $n - 1$ fois dérivable en $f(x)$. Comme f est n fois dérivable en x, elle est a fortiori $n - 1$ fois dérivable en x. Le composé $y \mapsto g'_{f(y)}$ est donc $n - 1$ fois dérivable en x, par hypothèse de récurrence.

Comme par ailleurs, f' est $n-1$ fois dérivable en x, l'application $y \mapsto (g'_{f(y)}, f'_y)$ est $n-1$ fois dérivable en x.

Par ailleurs, $(u, v) \mapsto u \circ v$ est de classe $\mathcal{C}^\infty$ (c'est une application bilinéaire continue). Elle est a fortiori $n - 1$ fois dérivable en tout point.

En appliquant une deuxième fois l'hypothèse de récurrence, on voit que $y \mapsto g'_{f(y)} \circ f'_y$ est $n - 1$ fois dérivable en x. Mais ceci signifie exactement que $g \circ f$ est n fois dérivable en x. $\square$

Théorème 7 *La composition de deux fonctions de classe $\mathcal{C}^n$ est de classe $\mathcal{C}^n$.*

Il suffit d'adapter légèrement la démonstration du thèorème précédent. $\square$

3.2.5 L'application $x \mapsto x^{-1}$.

Soit $\mathcal{A}$ une algèbre de Banach, et Ψ l'application $x \mapsto x^{-1}$ de $\mathcal{A}^*$ vers $\mathcal{A}$. On a déjà dérivé Ψ une fois, et obtenu :

$$\Psi'_x(h) = -x^{-1}hx^{-1}.$$

Lemme 12 *La fonction Ψ est indéfiniment dérivable.*

Il sagit de montrer que Ψ est de classe $\mathcal{C}^n$ pour tout n. On démontre ceci par récurrence sur n. Noter que le résultat est vrai pour $n = 1$ (et d'ailleurs aussi pour $n = 0$). Supposons donc $n \geq 2$, et le lemme démontré pour toute valeur strictement inférieure à n.

Considérons l'application Γ de $\mathcal{A} \times \mathcal{A}$ vers $\mathcal{L}(\mathcal{A}, \mathcal{A})$, définie par

$$(x, y) \overset{\Gamma}{\longmapsto} (h \mapsto -xhy).$$

On a donc $\Gamma(x, y)(h) = -xhy$.

Noter que l'application Γ est bilinéaire et continue (la continuité étant assurée par le fait que la norme de l'application linéaire $h \mapsto -xhy$ est majorée par $\|x\| \, \|y\|$). Γ est donc de classe $\mathcal{C}^\infty$.

Par ailleurs, on voit que $\Psi'_x(h) = -x^{-1}hx^{-1} = \Gamma(\Psi(x), \Psi(x))(h)$. C'est–à–dire qu'on a :

$$\Psi'_x = \Psi'(x) = \Gamma(\Psi(x), \Psi(x)).$$

De là il résulte que si Ψ est de classe $\mathcal{C}^{n-1}$, il en est de même de Ψ', ce qui montre que Ψ est de classe $\mathcal{C}^n$. Ψ est donc de classe $\mathcal{C}^\infty$. $\square$

Corollaire 3 *Soient E et F deux espaces de banach. L'application :*

$$f \mapsto f^{-1},$$

de $\mathrm{Iso}(E, F)$ *vers* $\mathrm{Iso}(F, E)$ *est indéfiniment dérivable.*

Si $\mathrm{Iso}(E, F)$ est vide, il n'y a rien à montrer. Sinon, il existe un isomorphisme $f_0 : E \longrightarrow F$ entre les espaces de Banach E et F. Ceci nous donne un isomorphisme *linéaire* (donc indéfiniment dérivable) Λ entre $\mathrm{Iso}(E, F)$ et l'algèbre de Banach $\mathrm{Iso}(E, E)$:

$$f \stackrel{\Lambda}{\longmapsto} f_0^{-1} \circ f.$$

De même on a un isomorphisme linéaire Δ entre $\mathrm{Iso}(E, E)$ et $\mathrm{Iso}(F, E)$ donné par :

$$g \stackrel{\Delta}{\longmapsto} g \circ (f_0)^{-1}.$$

Notons Ψ l'application $x \mapsto x^{-1}$ de l'algèbre de Banach $\mathrm{Iso}(E, E)$, et Φ l'application $x \mapsto x^{-1}$ de l'énoncé. On a alors clairement $\Phi = \Delta \circ \Psi \circ \Lambda$, ce qui montre que Φ est de classe $\mathcal{C}^\infty$. $\square$

3.3 Symétrie des dérivées d'ordre supérieur.

Théorème 8 *(Théorème de Schwarz) Soit U un ouvert d'un espace de Banach E. Soit $f : U \longrightarrow F$ une application deux fois dérivable en x. Alors l'application bilinéaire :*

$$(h, k) \mapsto f''_x(h, k),$$

de $E \times E$ vers F, est symétrique.

Il suffit de prouver que l'expression $f''_x(h, k) - f''_x(k, h)$ est $o((\|h\| + \|k\|)^2)$. En effet, le quotient

$$\frac{\|f''_x(th, tk) - f''_x(tk, th)\|}{(\|th\| + \|tk\|)^2}$$

tendra alors vers 0 quand le réel t tend vers 0, mais comme il est indépendant de t (car numérateur et dénominateur sont homogènes de degré 2), il ne peut être que nul, ce qui montrera le théorème.

Considérons la fonction φ définie par :

$$\varphi(t) = f(x + th + k) - f(x + th),$$

où $t \in [0, 1]$. En appliquant le théorème de la moyenne à la fonction $t \mapsto \varphi(t) - t\varphi'(0)$, sur l'intervalle $[0, 1]$, on obtient :

$$\|\varphi(1) - \varphi(0) - \varphi'(0)\| \leq \sup_{t \in [0,1]} \|\varphi'(t) - \varphi'(0)\|.$$

Calculons $\varphi'(t)$. On a, en appliquant le théorème de dérivation des fonctions composées :

$$\varphi'(t) = f'_{x+th+k}(h) - f'_{x+th}(h).$$

Par ailleurs, par définition de la dérivée seconde de f en x, on a :

$$\begin{aligned} f'_{x+th+k} &= f'_x + f''_x(th + k) + o(th + k) \\ f'_{x+th} &= f'_x + f''_x(th) + o(th). \end{aligned}$$

On peut majorer les normes de $th+k$ et de th par $\|h\|+\|k\|$ (t appartient à $[0,1]$), et donc remplacer $o(th+k)$ et $o(th)$ par $o(\|h\|+\|k\|)$. Noter que chaque membre des égalités ci–dessus est une application linéaire continue. En particulier, il en est ainsi de $o(\|h\|+\|k\|)$. L'expression $o(\|h\|+\|k\|)(h)$ peut s'écrire $o((\|h\|+\|k\|)^2)$. En effet, le quotient :

$$\frac{o(\|h\|+\|k\|)(h)}{(\|h\|+\|k\|)^2}$$

tend vers 0 quand $\|h\|+\|k\|$ tend vers 0, car la norme de $o(\|h\|+\|k\|)(h)$ est majorée par $\|o(\|h\|+\|k\|)\|\,\|h\|$.

On a donc :

$$\begin{aligned} f'_{x+th+k}(h) &= f'_x(h)+f''_x(th+k)(h)+o((\|h\|+\|k\|)^2) \\ f'_{x+th}(h) &= f'_x(h)+f''_x(th)(h)+o((\|h\|+\|k\|)^2). \end{aligned}$$

d'où :

$$\varphi'(t)=f''_x(k)(h)+o((\|h\|+\|k\|)^2),$$

cette égalité étant valable pour tout $t\in[0,1]$.

On en déduit que $\varphi'(t)-\varphi'(0)=o((\|h\|+\|k\|)^2)$, donc $\varphi(1)-\varphi(0)-\varphi'(0)=o((\|h\|+\|k\|)^2)$, et finalement $\varphi(1)-\varphi(0)-f''_x(k)(h)=o((\|h\|+\|k\|)^2)$.

Par ailleurs, l'expression $\varphi(1)-\varphi(0)$ est symétrique en h et k, car elle est égale à :

$$f(x+h+k)-f(x+k)-f(x+h)+f(x).$$

Si on échange les rôles de h et k, on obtient donc $\varphi(1)-\varphi(0)-f''_x(h)(k)=o((\|h\|+\|k\|)^2)$, d'où :

$$f''_x(h)(k)-f''_x(k)(h)=o((\|h\|+\|k\|)^2).\ \square$$

Corollaire 4 *Si $f:U\longrightarrow F$ est n fois dérivable en x, l'application n–multilinéaire $f_x^{(n)}$ est symétrique.*

Rappelons qu'une application $f:E\times\cdots\times E\longrightarrow F$ n–multilinéaire est *symétrique*, si elle vérifie l'égalité :

$$f(x_1,\dots,x_n)=f(x_{\sigma(1)},\dots,x_{\sigma(n)}),$$

pour tout permutation σ de l'ensemble $\{1,\dots,n\}$, et tous $x_1,\dots,x_n$.

Le corollaire est déjà démontré pour $n\leq 2$. Supposons donc $n\geq 3$, et procédons par récurrence sur n.

$f^{(n)}$ n'est autre que la dérivée seconde de $g=f^{(n-2)}$. g est deux fois dérivable en x. L'application :

$$(h_1,h_2)\mapsto g''_x(h_1,h_2)$$

est donc symétrique. Ceci dit, $g''_x(h_1,h_2)$ est une application $(n-2)$–multilinéaires de $E\times\cdots\times E$ vers F. On peut donc l'appliquer à des vecteurs $h_3,\dots,h_n$, on obtient l'élément suivant de F :

$$g''_x(h_1,h_2)(h_3,\dots,h_n)=f_x^{(n)}(h_1,\dots,h_n).$$

Ceci montre que $f_x^{(n)}(h_1,\dots,h_n)$ est invariant par permutation de h_1 et h_2.

Mais on peut aussi considérer $f^{(n)}$ comme la dérivée première de $j=f^{(n-1)}$. On obtient de même :

$$j'_x(h_1)(h_2,\dots,h_n)=f^{(n)}(h_1,\dots,h_n),$$

ce qui montre que $f_x^{(n)}(h_1,\dots,h_n)$ est invariant par permutation des $n-1$ dernières variables, car par hypothèse de récurrence, $j'_x(h_1)$ est une application $(n-1)$–multilinéaire symétrique.

Le corollaire résulte maintenant du fait que le groupe des permutations de l'ensemble $\{1,\dots,n\}$ est engendré par la transposition de 1 et 2, et par les permutation de $2,\dots,n$.([2]) $\square$

[2] Soit σ est une permutation quelconque de $\{1,\dots,n\}$. Si $\sigma(1)=1$, la question est réglée. Sinon, $1=\sigma(i)$ $(i\neq 1)$. Commençons par une permutation des $n-1$ derniers éléments amenant i à la deuxième place, puis transposons les deux

3.4 Formule de Taylor–Young.

Théorème 9 *(formule de Taylor–Young en dimension 1) Soit I un intervalle de $\mathbb{R}$, x un point de I, et f une fonction de I vers un espace de Banach F, n fois dérivable en x. Alors, on a la formule de Taylor–Young à l'ordre n suivante :*

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \cdots + \frac{h^n}{n!}f^{(n)}(x) + o(h^n).$$

Démontrons ceci par récurrence sur n. Pour $n = 1$, la formule se résume à la définition de la dérivée. Supposons donc $n \geq 2$, et la formule démontrée à l'ordre $n-1$.

Posons $\varphi(h) = f(x+h) - hf'(x) - \frac{h^2}{2!}f''(x) - \cdots - \frac{h^n}{n!}f^{(n)}(x)$. On a $\varphi(0) = f(x)$, donc on doit simplement montrer que $\varphi(h) - \varphi(0)$ est un $o(h^n)$.

Dérivons φ. On obtient :

$$\varphi'(h) = f'(x+h) - f'(x) - hf''(x) - \cdots - \frac{h^{n-1}}{(n-1)!}f^{(n)}(x).$$

L'hypothèse de récurrence montre donc que $\varphi'(h)$ est un $o(h^{n-1})$.

Soit $\varepsilon > 0$. Il existe $\eta > 0$, tel que $|h| \leq \eta$ entraine $|\varphi'(h)| \leq \varepsilon|h^{n-1}|$. Pour tout t tel que $|t| \leq |h| \leq \eta$, on a donc $|\varphi'(t)| \leq \varepsilon|t^{n-1}| \leq \varepsilon|h^{n-1}|$. φ est donc $\varepsilon|h^{n-1}|$–lipschitzienne sur l'intervalle $[0, h]$ (où l'intervalle $[h, 0]$ si h est négatif). On a donc $|\varphi(h) - \varphi(0)| \leq \varepsilon|h^{n-1}|\,|h|$. □

Remarque : Le reste de Young $o(h^n)$ peut être noté $h^n\varepsilon(h)$, où $\varepsilon(h)$ tend vers 0 quand h tend vers 0.

Théorème 10 *(formule de Taylor–Young) Soit U un ouvert d'un espace de Banach E, et F un espace de Banach. Soit $x \in E$. Si $f : U \longrightarrow F$ est n fois dérivable en x, on a la formule de Taylor–Young à l'ordre n suivante :*

$$f(x+h) = f(x) + f'_x(h) + \frac{1}{2!}f''_x(h,h) + \cdots + \frac{1}{n!}f^{(n)}(h,\dots,h) + o(\|h\|^n).$$

Nous déduisons ce théorème du précédent. Soit h un vecteur non nul de E. Posons $u = \frac{h}{\|h\|}$, et $g(t) = f(x+tu)$, où t est un réel. La fonction g est définie dans un voisinage de 0 dans $\mathbb{R}$, et à valeurs dans F, et on a $g(0) = f(x)$. Elle est n fois dérivable en 0, et ses dérivées en 0 sont :

$$g'(0) = f'_x(u), \qquad g''(0) = f''_x(u,u), \qquad \dots \qquad g^{(n)}(0) = f^{(n)}_n(u,\dots,u).$$

Le théorème précédent appliqué à g en 0 donne :

$$g(t) = g(0) + tg'(0) + \frac{t^2}{2!}g''(0) + \cdots + \frac{t^n}{n!}g^{(n)}(0) + o(t^n).$$

On a donc :

$$f(x+tu) = f(x) + f'_x(tu) + \frac{1}{2!}f''_x(tu,tu) + \cdots + \frac{1}{n!}f^{(n)}(tu,\dots,tu) + o(t^n).$$

La formule désirée est obtenue en remplaçant t par $\|h\|$ dans cette expression. □

premières places, ce qui a pour effet d'amener i à la première place, et enfin terminons par une permutation des $n-1$ derniers éléments, pour amener les autres éléments à leur place.

Lemme 13 *Si $f : U \longrightarrow F$ est continue et n fois dérivable en x, les conditions suivantes sont équivalentes :*
- *$f(x+h) = o(\|h\|^n)$,*
- *$f_x^{(p)} = 0$ pour $0 \leq p \leq n$.*

La formule de Taylor–Young montre que si $f_x^{(p)} = 0$ pour tout p tel que $0 \leq p \leq n$, alors $f(x+h)$ est $o(\|h\|^n)$.

Démontrons la réciproque par récurrence sur n. Pour $n = 0$, on a $f(x+h) = o(1)$, donc $f(x) = 0$, car f est continue.

Supposons maintenant $n \geq 1$. Par hypothèse de récurrence, on a $f_x^{(p)} = 0$, pour tout p tel que $0 \leq p \leq n-1$. Il nous reste seulement à montrer que $f_x^{(n)}$ est nul. La formule de Taylor–Young donne :

$$f(x+h) = \frac{1}{n!} f_x^{(n)}(h, \dots, h) + o(\|h\|^n).$$

$f_x^{(n)}(h, \dots, h)$ est donc négligeable devant $\|h\|^n$. Donc le quotient :

$$\frac{f_x^{(n)}(th, \dots, th)}{\|th\|^n}$$

tend vers 0 quand le réel t tend vers 0. Comme ce quotient est indépendant de t (le numérateur et le dénominateur sont homogènes de degré n), il ne peut être que nul.

On a donc $f_x^{(n)}(h, \dots, h) = 0$ pour tout h. Le théorème résulte alors immédiatement du lemme suivant :

Lemme 14 *Si $f : E \times \dots \times E \longrightarrow F$ est une application n–multilinéaire continue* symétrique*, et si pour tout x de E $f(x, \dots, x) = 0$, alors f est nulle.*

On va montrer ce lemme par récurrence sur n. Il est évident pour $n = 1$, la conclusion étant alors contenue dans les hypothèses.

Soient x et y deux vecteurs quelconques de E. Posons $\varphi(t) = f(x+(t-1)y, \dots, x+(t-1)y)$. φ est une application de $\mathbb{R}$ vers F. Elle est nulle par hypothèse. Sa dérivée φ' est donc elle aussi nulle. On peut par ailleurs la calculer directement sur la définition, ce qui donne, en tenant compte de la symétrie de f :

$$\varphi'(t) = nf(y, x+(t-1)y, \dots, x+(t-1)y).$$

On a en particulier $\varphi'(1) = 0$, c'est–à–dire $f(y, x, \dots, x) = 0$, ceci étant valable pour tout x et tout y.

Pour y fixé, l'application $(n-1)$–multilinéaire ψ définie par :

$$(x_2, \dots, x_n) \overset{\psi}{\longmapsto} f(y, x_2, \dots, x_n)$$

est symétrique, et vérifie $\psi(x, \dots, x) = 0$, pour tout x, comme on l'a montré plus haut. Par hypothèse de récurrence, ψ est nulle, et comme y est arbitraire, f est nulle. $\square$

3.5 Extrémas de fonctions deux fois dérivables.

Théorème 11 *Soit $f : U \longrightarrow \mathbb{R}$ une fonction définie sur un ouvert U de $\mathbb{R}^n$. Soit $x \in U$. Si f est deux fois dérivable en x, si $f'_x = 0$, et si la forme quadratique $h \mapsto f''_x(h,h)$ est définie positive (resp. négative), f admet un minimum (resp. maximum) local en x.*

La formule de Taylor donne :

$$f(x+h) - f(x) = \frac{1}{2} f''_x(h,h) + o(\|h\|^2).$$

Il s'agit de montrer que $h \mapsto f(x+h) - f(x)$ garde un signe constant dans un voisinage de 0. Comme la sphère unité S de $\mathbb{R}^n$ est compacte, la fonction $h \mapsto f''_x(h,h)$, qui est strictement positive sur S, y est minorée par un réel stritement positif α. On a donc $f''_x(h,h) \geq \alpha$ sur S, ou encore $f''_x(h,h) \geq \alpha \|h\|^2$ pour tout h de $\mathbb{R}^n$. Prenons $\eta > 0$ assez petit pour que $\|h\| < \eta$ entraine $o(\|h\|^2) \leq \frac{\alpha}{4}\|h\|^2$. On voit alors que $f(x+h) - f(x)$ reste pour $\|h\| \leq \eta$, plus grand que $\frac{\alpha}{4}\|h\|^2$, c'est–à–dire positif. □

Lemme 15 *Soit $f : U \longrightarrow \mathbb{R}$ une fonction définie sur un ouvert U de $\mathbb{R}^n$. Soit $x \in U$. Si f est deux fois dérivable en x, si $f'_x = 0$, et si la forme quadratique $h \mapsto f''_x(h,h)$ n'est ni positive ni négative, f n'admet pas d'extremum en x.*

Comme f''_x n'est pas positive, sa restriction a une droite vectorielle D_- est définie négative. De même, comme elle n'est pas négative, sa restriction à une droite vectorielle D_+ est définie positive. Le théorème précédent, appliqué aux restrictions de f à ces deux droites, montre alors qu'il existe des points aussi proche qu'on veut de x, en lesquels f prend des valeurs strictement plus grandes ou strictement plus petites que $f(x)$. □

Exercices

12 Soit E un espace de Banach, et $q : E \longrightarrow \mathbb{R}$ une forme quadratique continue de forme polaire $f : E \times E \longrightarrow F$. Montrer que q est indéfiniment dérivable, et que :

$$q'_x(h) = 2f(x,h), \qquad q''_x(h,k) = 2f(h,k), \qquad q^{(n)} = 0 \text{ (pour } n \geq 3)$$

13 Soient E, F et G des espaces de Banach. Soit U un ouvert de E, et V un ouvert de F. Soit $x \in U$. Soit $f : U \longrightarrow V \subset F$ une application deux fois dérivable en x. Soit $g : V \longrightarrow G$ une application deux fois dérivable en $f(x)$. Montrer que la dérivée seconde de $g \circ f$ en x, est donnée par :

$$(g \circ f)''_x = g''_{f(x)} \circ (f'_x \times f'_x) + g'_{f(x)} \circ f''_x,$$

où $(f'_x \times f'_x)(h,k)$ est par définition $(f'_x(h), f'_x(k))$.

14 Soit E un espace euclidien de dimension finie. On notera $x.y$ le produit scalaire de x et y, et $\|x\|$ la norme de x (qui vaut $\sqrt{x.x}$). Soit $f : E \longrightarrow E$ une application deux fois dérivable, telle que pour tout x, f'_x respecte la norme, c'est–à–dire telle que $\|f'_x(h)\| = \|h\|$, pour tout x et tout h de E.

a) Montrer que pour tout x de E, f'_x respecte le produit scalaire, c'est–à–dire que $f'_x(h).f'_x(k) = h.k$, pour tous x, h et k de E. Montrer de plus que pour tout x, f'_x est bijectif.

Pour h, k et l des éléments de E, on pose $A_x(h,k,l) = f'_x(h).f''_x(l,k)$.

b) Montrer que $A_x(h,k,l) + A_x(k,h,l) = 0$.

c) Montrer que pour tous h, k, l et x, on a $A_x(h,k,l) = 0$ et en déduire que $f''_x = 0$ pour tout x de E.

d) Montrer que f' est une application constante, et en déduire que f est une application affine.

$\boxed{\mathbf{15}}$ Soient $\mathcal{A}$ et $\mathcal{B}$ deux algèbres de Banach. On note $\mathcal{A}^*$ (respectivement : $\mathcal{B}^*$) le groupe des éléments inversibles de $\mathcal{A}$ (respectivement : $\mathcal{B}$). Soit $g : \mathcal{A}^* \longrightarrow \mathcal{B}^*$ un morphisme de groupes deux fois dérivable. Pour tous x et y dans $\mathcal{A}$ ou dans $\mathcal{B}$, on pose $[x, y] = xy - yx$. Montrer que $g'_1([x, y]) = [g'_1(x), g'_1(y)]$. (Utiliser la formule de Taylor–Young à l'ordre 2 pour développer les expressions $g(1 + tx)$, $g(1 + ty)$ et $g(1 + tx + ty + t^2xy)$, où t est un réel.)

$\boxed{\mathbf{16}}$ Soit E un espace euclidien de dimension finie. Soit $\{a_0, \dots, a_p\}$ une famille finie non vide ($p \geq 0$) de points de E. Soit $f : E \longrightarrow \mathbb{R}$ la fonction définie par $f(x) = \displaystyle\sum_{i=0}^{p} \|x - a_i\|^2$.

a) Calculer les deux premières dérivées f'_x et f''_x de f en x.

b) Déterminer les extrema relatifs de f.

$\boxed{\mathbf{17}}$ On note $\mathcal{M}$ l'espace des matrices carrées $n \times n$ à coefficients réels. Soient $A_1, \dots, A_p$ un ensemble fini (éventuellement vide pour $p = 0$), de vecteurs de $\mathcal{M}$. On considère la fonction $\Psi : \mathcal{M} \longrightarrow \mathbb{R}$ définie par $\Psi(X) = \mathrm{tr}\,({}^tXX) + \displaystyle\sum_{i=1}^{p} \mathrm{tr}\,({}^t(X - A_i)(X - A_i))$.

a) Calculer $\Psi'_X(H)$ et $\Psi''_X(H, K)$.

b) Déterminer les extremas relatifs de Ψ.

4 Difféomorphismes et inversion locale.

4.1 Difféomorphismes et difféomorphismes locaux.

Définition 6 *Soit U un ouvert d'un espace de Banach E, et V un ouvert d'un espace de Banach F. Une application $f : U \longrightarrow V$ est appelée difféomorphisme de classe $\mathcal{C}^n$, si f est bijective, et si f et f^{-1} sont de classe $\mathcal{C}^n$.*

On notera que le composé de deux difféomorphismes est un difféomorphisme, que si un composé est un difféomorphisme, et si l'une des deux fonctions de la composition est un difféomorphisme, l'autre en est un aussi. Si f est un difféomorphisme, f'_x est un isomorphisme d'espaces de Banach, comme le montre le théorème de dérivation des fonctions composées, appliqué aux composition $f \circ f^{-1}$ et $f^{-1} \circ f$.

Théorème 12 *Soit U un ouvert d'un espace de Banach E, et V un ouvert d'un espace de Banach F. Soit $f : U \longrightarrow V$ un homéomorphisme. Pour que f soit un difféomorphisme de classe $\mathcal{C}^n$ ($n \geq 1$), il suffit que f soit de classe $\mathcal{C}^n$, et que pour tout x de U, $f'_x \in \mathrm{Iso}\,(E, F)$.*

Posons $g = f^{-1}$. La seule chose à montrer est que g est de classe $\mathcal{C}^n$. Soit $y \in V$. Il existe un unique $x \in U$, tel que $y = f(x)$. On a :

$$f(x + h) = f(x) + f'_x(h) + o(h).$$

Posons $k = f(x + h) - f(x)$. On a alors $y + k = f(x + h)$, donc $g(y + k) = x + h = g(y) + h$. Par ailleurs, $k = f'_x(h) + o(h)$, donc $h = (f'_x)^{-1}(k) + (f'_x)^{-1}(o(h))$. On obtient donc :

$$g(y + k) = g(y) + (f'_x)^{-1}(k) + (f'_x)^{-1}(o(h)).$$

Pour montrer que g est dérivable en y, avec pour application linéaire tangente $(f'_x)^{-1}$, il suffit donc de prouver que $(f'_x)^{-1}(o(h))$ est $o(k)$. Comme $(f'_x)^{-1}$ est linéaire continue, il suffit de prouver que $o(h)$ est $o(k)$, autrement–dit que tout ce qui est négligeable devant h, l'est aussi devant k.

Posons $A = \|(f'_x)^{-1}\|$. On a $h = (f'_x)^{-1}(k) + (f'_x)^{-1}(o(h))$. Soit $\eta > 0$, tel que $\|h\| < \eta$ entraine $\|o(h)\| < \frac{\|h\|}{2A}$. On a alors :

$$\|h\| \leq A\|k\| + \frac{\|h\|}{2},$$

c'est–à–dire $\|h\| \leq 2A\|k\|$, ce qui montre que $o(h)$ est $o(k)$.

g est donc dérivable sur V, et a $(f'_x)^{-1}$ pour dérivée en $f(x)$.

Si f est de classe $\mathcal{C}^n$, f' est de classe $\mathcal{C}^{n-1}$, donc le composé de f' avec l'application $x \mapsto x^{-1}$, de $\mathrm{Iso}\,(E,F)$ vers $\mathrm{Iso}\,(F,E)$, qui est de classe $\mathcal{C}^\infty$, est de classe $\mathcal{C}^{n-1}$. Mais ce composé est $(f^{-1})'$. $\square$

Définition 7 *Soient E et F deux espaces de Banach, U un ouvert de E, $f : U \longrightarrow F$ une application continue, et x_0 un point de U. On dit que f est un* difféomorphisme local de classe $\mathcal{C}^k$ en x_0*, s'il existe un ouvert V inclus dans U et contenant x_0, et un ouvert W de F, tels que f soit un difféomorphisme de classe $\mathcal{C}^k$ de V vers W.*

4.2 Applications contractantes.

Définition 8 *Soit E un espace métrique, dont la distance sera notée $(x,y) \mapsto d(x,y)$. Une application $f : E \longrightarrow E$ est dite k–*contractante*, si $0 \leq k < 1$, et si*

$$\forall x \in E\ \forall y \in E\ d(f(x), f(y)) \leq kd(x,y).$$

On dit que f est contractante*, s'il existe k tel que f soit k–contractante.*

On notera qu'une application k–contractante est k–lipschitzienne, donc en particulier continue.

Théorème 13 *Si E est complet et non vide, toute application contractante de E vers E a un unique point fixe.*

En effet, comme E n'est pas vide, il existe un point x dans E. Considérons la suite $(f^n(x))_{n\in\mathbf{N}}$ dans E. C'est une suite de Cauchy. En effet, f diminuant les distances au moins d'un facteur k, on a :

$$\begin{aligned} d(f^{p+q}(x), f^p(x)) &\leq d(f^{p+q}(x), f^{p+q-1}(x)) + \cdots + d(f^{p+1}(x), f^p(x)) \\ &\leq (k^{q-1} + \cdots + 1)d(f^{p+1}(x), f^p(x)) \\ &\leq (k^{q-1} + \cdots + 1)k^p d(f(x), x) \\ &\leq \frac{k^p}{1-k} d(f(x), x). \end{aligned}$$

Or, cette dernière expression expression est plus petite qu'un ε donné à l'avance, dès que p est assez grand.

La suite ci–dessus a donc une limite y (E est complet). Comme la suite $(f(f^n(x)))_{n\in\mathbf{N}}$ a même limite que $(f^n(x))_{n\in\mathbf{N}}$ (c'est la même suite décalée d'un cran), et comme f est continue, on voit que y est un point fixe de f. Si z était un autre point fixe de f, on aurait :

$$d(y,z) = d(f(y), f(z)) \leq kd(y,z),$$

ce qui montre que $d(y, z) = 0$, donc que $y = z$. □

Si l'espace métrique E est borné (c'est–à–dire si la distance de deux points quelconques de E est majorée par un réel A (la borne inférieure desquels est appelée *diamètre de* E) indépendant de ces points), alors on peut introduire une notion de distance entre deux fonctions k–contractantes f et g de E vers E, par la formule :

$$d(f, g) = \sup_{x \in E} d(f(x), g(x)).$$

Ce sup est d'ailleurs majoré par la constante A.

On voit alors que l'ensemble $C_k(E)$ des applications k–contractantes de E vers E est un espace métrique.([3])

Théorème 14 *Soit E un espace métrique borné non vide . L'application $\Psi : C_k(E) \longrightarrow E$, qui à toute fonction k–contractante de E vers E associe son unique point fixe est $\dfrac{1}{1-k}$–lipschitzienne.*

Soient f et g deux applications k–contractantes de E vers E. Soit x le point fixe de f. On a :

$$\begin{array}{rcl} d(g^n(x), x) & \leq & d(g^n(x), g^{n-1}(x)) + \cdots + d(g(x), x) \\ & \leq & (k^{n-1} + \cdots + 1) d(g(x), x) \\ & \leq & (k^{n-1} + \cdots + 1) d(g(x), f(x)) \\ & \leq & \dfrac{1}{1-k} d(f, g). \end{array}$$

Comme la suite $\left(g^n(x)\right)_n$ converge vers le point fixe de g, on a le résultat. □

4.3 Énoncé et démonstration du théorème d'inversion locale.

Théorème 15 *(théorème d'inversion locale) Soit E et F deux espaces de Banach. Soit U un ouvert de E, et x_0 un point de U. Soit $f : U \longrightarrow F$ une application de classe $\mathcal{C}^n$ $(n \geq 1)$, telle que f'_{x_0} soit un isomorphisme d'espaces de Banach de E vers F. Alors f est un difféomorphisme local de classe $\mathcal{C}^n$ en x_0.*

L'hypothèse que f est de classe $\mathcal{C}^1$ dans un voisinage de x_0, c'est–à–dire que f' est continue dans un voisinage de x_0, est essentielle. En effet, considérons l'application f de $\mathbb{R}$ vers $\mathbb{R}$ définie par :

$$f(x) = x + 2x^2 \cos\left(\frac{1}{x}\right),$$

pour $x \neq 0$, et $f(0) = 0$. Elle admet $\mathrm{Id}_{\mathbf{R}}$ comme dérivée en 0. Pour le voir, il suffit de chercher la limite de :

$$\frac{x + 2x^2 \cos\left(\frac{1}{x}\right)}{x},$$

qui est clairement 1. Toutefois, f' n'est pas continue en 0. En effet, pour $x \neq 0$, on a $f'(x) = 1 + 4x \cos\left(\dfrac{1}{x}\right) + 2 \sin\left(\dfrac{1}{x}\right)$, et cette expression n'a pas de limite quand x tend vers 0. La fonction f n'est par ailleurs injective dans aucun voisinage de 0. En effet, l'expression $1 + 4x \cos\left(\dfrac{1}{x}\right)$ reste voisine de 1 quand x est proche de 0, alors que $2 \sin\left(\dfrac{1}{x}\right)$ oscille une infinité de fois entre 2 et -2. En conséquence, f

[3]On pourra montrer à titre d'exercice que cet espace est complet si E est complet.

change de sens de variation une infinité de fois au voisinage de 0, ce qui pour une fonction continue montre qu'elle ne peut être injective dans aucun voisinage de 0. Elle ne saurait donc être une difféomorphisme local en 0.

4.3.1 Réduction du problème

Nous commençons par nous ramener au cas où $E = F$, $x_0 = f(x_0) = 0$ et $f'_{x_0} = \mathrm{Id}_E$. Pour cela, posons pour tout x de E :

$$g(x) = {f'_{x_0}}^{-1} \left(f(x_0 + x) - f(x_0) \right).$$

L'application g est donc la composée des applications suivantes :

$$\begin{array}{cccc} E \longrightarrow E & E \longrightarrow F & F \longrightarrow F & F \longrightarrow E \\ x \mapsto x_0 + x & f & y \mapsto y - f(x_0) & {f'_{x_0}}^{-1} \end{array}$$

g envoie donc E dans E, et envoie 0 sur 0 :

$$0 \mapsto x_0 \mapsto f(x_0) \mapsto 0 \mapsto 0$$

La première et la troisième applications sont des translations, qui sont donc dérivables, et admettent l'application identique (de E ou de F) comme dérivée en un point quelconque. La deuxième, qui est f, est dérivable en x_0, et la dernière qui est linéaire continue est dérivable en 0. g est donc dérivable en 0 et admet pour dérivée le composé ${f'_{x_0}}^{-1} \circ \mathrm{Id}_F \circ f'_{x_0} \circ \mathrm{Id}_E$, c'est–à–dire Id_E.

Toutes ces applications, sauf f, étant des difféomorphismes, on voit que g est un difféomorphisme local en 0 si et seulement si f est un difféomorphisme local en x_0.

Noter que g est de classe $\mathcal{C}^n$ dans un voisinage de 0, précisément dans le translaté de U par la translation $x \mapsto x - x_0$, qu'on notera U'.

Il nous reste donc à prouver que $g : E \longrightarrow E$, vérifiant $g(0) = 0$ $g'_0 = \mathrm{Id}_E$, et de classe $\mathcal{C}^n$ dans un voisinage de 0, est un difféomorphisme local en 0.

4.3.2 Une application contractante.

Lemme 16 *Il existe un voisinage de* 0 *dans* E*, dans lequel l'application*

$$x \mapsto \psi(x) = x - g(x)$$

est $\dfrac{1}{2}$*–contractante.*

Il s'agit de montrer qu'il existe un $\eta > 0$, tel que pour x et y dans la boule de centre 0 et de rayon η, on a :

$$\|\psi(x) - \psi(y)\| \leq \frac{1}{2}\|x - y\|.$$

Ceci résulte du théorème de la moyenne. En effet, on a pour tout x de U' :

$$\psi'_x = \mathrm{Id}_E - g'_x.$$

Comme $x \mapsto g'_x$ est continue, et $g'_0 = \mathrm{Id}_E$, l'expression $\mathrm{Id}_E - g'_x$ tend vers 0 quand x tend vers 0. Il existe donc $\eta > 0$, tel que $\|\psi'_x\| \leq \dfrac{1}{2}$, dès que $\|x\| \leq \eta$. Le théorème de la moyenne donne alors le résultat. □

4.3.3 Injectivité locale.

Notons V la boule ouverte de centre 0 et de rayon η (sur laquelle ψ est $\frac{1}{2}$–contractante). Alors la restrition de g à cette boule est injective. En effet, pour x et y dans V, on a :

$$\|\psi(x) - \psi(y)\| \leq \frac{1}{2}\|x - y\|.$$

Si on avait $g(x) = g(y)$, on aurait alors :

$$\|x - y\| \leq \frac{1}{2}\|x - y\|,$$

donc $x = y$. □

4.3.4 Surjectivité locale.

On désire montrer que si y est assez proche de 0, l'équation en x suivante :

$$y = g(x),$$

a une solution. Cette équation peut encore s'écrire :

$$y - g(x) + x = x,$$

ce qui fait qu'une solution apparaît comme un *point fixe* de la fonction $x \mapsto \varphi(x) = y - g(x) + x$. Pour montrer que cette fonction a un point fixe, il suffit de montrer qu'elle envoie une boule fermée F de centre 0 et de rayon non nul (qui est une partie complète de E) dans elle–même, et que sa restriction à F est contractante. On va montrer qu'il en est ainsi pour y assez proche de 0.

On va construire F dans la boule V de centre 0 et de rayon η. De cette façon, on est assuré que φ est $\frac{1}{2}$–contractante. En effet $\varphi(x) = y - \psi(x)$, donc pour a et b dans V, on a :

$$\|\varphi(a) - \varphi(b)\| = \|\psi(a) - \psi(b)\| \leq \frac{1}{2}\|a - b\|.$$

Prenons pour F n'importe quelle boule fermée de centre 0 et de rayon η' strictement plus petit que η. De cette façon F est contenue dans V, et la condition ci dessus est vérifiée.

Si maintenant y est de norme plus petite que $\frac{\eta'}{2}$, F est stable par φ. En effet, pour $\|x\| \leq \eta'$:

$$\|\varphi(x)\| = \|y - \psi(x)\| \leq \|y\| + \|\psi(x)\| \leq \frac{\eta'}{2} + \frac{\eta'}{2} = \eta',$$

puisque ψ est $\frac{1}{2}$–contractante, et $\psi(0) = 0$.

φ a donc un point fixe unique dans F. Tout y de norme inférieure ou égale à $\frac{\eta'}{2}$ est donc l'image par g d'un unique x de norme inférieure ou égale à η'. Ceci montre la surjectivité locale de g.

4.3.5 L'inverse local et sa continuité.

D'après ce qu'on vient de voir, tout élément de la boule fermée F' de centre 0 et de rayon $\frac{\eta'}{2}$ est l'image par g d'un unique élément de la boule fermée F de centre 0 et de rayon η'.

g n'est pas nécessairement une bijection entre F et F', car des éléments x de F peuvent très bien avoir une image par g qui ne soit pas dans F'. Toutefois, l'assertion ci–dessus définit une application $h : F' \longrightarrow F$. Pour tout y de F', $h(y)$ est l'unique élément de F qui vérifie $g(h(y)) = y$.

Par ailleurs, h est continue, car elle est 2–lipschitzienne. En effet, si on pose $\varphi_y(x) = y - g(x) + x$ et $\varphi_{y'}(x) = y' - g(x) + x$, on a $\|\varphi_y - \varphi_{y'}\| = \|y - y'\|$. Or l'application qui à φ_y associe son unique point fixe $h(y)$ est $\dfrac{1}{1-\frac{1}{2}}$–lipschitzienne, c'est–à–dire 2–lipschitzienne.

Comme h est 2–lipschitzienne, et envoie 0 sur 0, elle envoie en fait la boule *ouverte* V' de centre 0 et de rayon $\dfrac{\eta'}{2}$ dans la boule *ouverte* V'' de centre 0 et de rayon η'.

Par ailleurs, on a

$$h(V') = g^{-1}(V') \cap V''.$$

En effet, si $x \in h(V')$, alors $x = h(y)$ avec $y \in V'$. Donc x est dans V'' et $g(x) = g(h(y)) = y$ est dans V'. Réciproquement, si $x \in V''$ et $g(x) \in V'$, on a $g(h(g(x))) = g(x)$. Mais comme x et $h(g(x))$ sont dans V'', et que la restriction de g à V'' est injective, on voit que $x = h(g(x))$, donc que $x \in h(V')$.

Ceci montre que $h(V')$ est un *ouvert* de E.

On a donc établi que h est un *homéomorphisme* de l'ouvert V' (contenant 0) sur l'ouvert $h(V')$ (contenant 0). Cet homéomorphisme est l'inverse de $g : h(V') \longrightarrow V'$.

4.3.6 Fin de la démonstration.

Pour terminer la démonstration du théorème d'inversion locale, il nous reste à voir que $g : h(V') \longrightarrow V'$ et son inverse $h : V' \longrightarrow h(V')$ sont de classe $\mathcal{C}^n$. Pour g il s'agit d'une hypothèse, et pour h, cela est une conséquence d'un théorème vu plus haut. On a donc démontré le théorème d'inversion locale.

Exercices

$\boxed{18}$ On considère §n la sphère de Riemann de $\mathbb{R}^{n+1}$ ($\mathcal{S}^n = \{x \in \mathbb{R}^{n+1} |\ \|x\| = 1\}$). Soit X un champ de vecteur sur $\mathcal{S}^n$ de classe $\mathcal{C}^\infty$ et tel que $\forall x \in \mathcal{S}^n,\ \langle x, X(x)\rangle = 0$ (le champ est orthogonal à la sphère) et $\forall x \in \mathcal{S}^n,\ X(x) \neq 0$.

Le but de cet éxercice va être de montrer que n est paire.

Pour $t \geq 0$, on introduit les fonctions :

$$\begin{array}{ccc} \mathcal{S}^n & \xrightarrow{Y} & \mathcal{S}^n \\ x & \longmapsto & \dfrac{X(x)}{\|X(x)\|} \end{array} \quad \text{et} \quad \begin{array}{ccc} \mathcal{S}^n & \xrightarrow{\varphi_t} & \mathbb{R}^{n+1} \\ x & \longmapsto & x + tY(x) \end{array}$$

1. Montrer que pour t suffisament petit, φ_t est injective.
2. Montrer qu'alors $\forall x \in \mathcal{S}^n,\ \varphi_t(x) \in S(0, \sqrt{1+t^2})$.
3. Montrer que φ_t est un difféomorphisme local.
4. Conclure.

Solutions des exercices.

$\boxed{1}$ On voit facilement que φ_t est linéaire par définition de l'évaluation des fonctions. Donc si φ_t est continue, alors elle sera sa propre dérivée. Il faut donc montrer que φ_t est linéaire continue.

Or si on munis $C([a,b],\mathbb{R})$ de la norme de convergence uniforme que l'on notera $\| \ \|_\infty^{[a,b]}$, alors on a :

$$\begin{array}{cc} \forall f \in C([a,b],\mathbb{R}), & |\varphi_t(f)| = |f(t)| \leq \|f\|_\infty^{[a,b]} \\ \Longrightarrow & \|\varphi_t\| \leq 1 \end{array}$$

D'où on déduit que φ_t est continue et que donc φ_t est dérivable et on a $\forall f \in C([a,b],\mathbb{R})$, $(\varphi_t)'_f = \varphi_t$.

De même, par linéarité de l'intégrale, ψ est une forme linéaire. Donc il suffit de montrer que ψ est continue pour montrer qu'elle est dérivable et on auras aussi que $\forall f \in C([a,b],\mathbb{R})$, $\psi'_f = \psi$.

$$\begin{array}{cc} \forall f \in C([a,b],\mathbb{R}), & |\psi(f)| = \left| \displaystyle\int_a^b f(t)dt \right| \leq (b-a)\|f\|_\infty^{[a,b]} \\ \Longrightarrow & \|\psi\| \leq b-a \end{array}$$

En fait on a même que $\|\psi\| = b-a$ en considérant la fonction constante égale à 1 par exemple.

Donc ψ est une forme linéaire continue et d'où ce qu'il fallait montrer.

$\boxed{2}$ Soit $A \in \mathcal{M}_2(\mathbb{R})$, on alors :

$$\theta(A+H) = (A+H)^2 = A^2 + AH + HA + H^2$$

Et donc pour montrer que θ est dérivable, il faut montrer que $H^2 = o(H)$ et que $H \mapsto AH + HA$ est une application linéaire continue.

On a : $\dfrac{\|H^2\|}{\|H\|} \leq \|H\| \xrightarrow[\|H\|\to 0]{} 0$ c'est à dire $H^2 = o(H)$. On pose maintenant $\gamma_A(H) = AH + HA$, on a que γ_A est linéaire car le produit est linéaire et la somme aussi. Il reste donc à montrer que γ_A est continue. On a :

$$\|\gamma_A(H)\| \leq 2\|A\|\|H\| \ \Rightarrow \ \|\gamma_A\| \leq 2\|A\|$$

et donc γ_A est continue et d'où θ est dérivable et $\theta'_A = H \mapsto AH + HA$.

$\boxed{3}$ On a $\varphi_*(f+h) = \varphi \circ (f+h)$ mais ceci n'est pas égal à $\varphi \circ f + \varphi \circ h$ car φ n'est peut–être pas linéaire. Il faut utiliser autre chose que les méthodes précédentes. En fait, si $x \in [a,b]$, on a :

$$\begin{aligned} \varphi_*(f+h)(x) &= \varphi(f(x)+h(x)) \\ &= \varphi(f(x)) + \varphi'(f(x))h(x) + \frac{h(x)^2}{2}\varphi''\left(f(x)+\theta h(x)\right), \ \theta \in [0,1] \qquad \text{(Taylor-Lagrange)} \\ &= \varphi_*(f)(x) + \left(h\varphi'_*(f)\right)(x) + \left(x \mapsto \frac{h(x)^2}{2}\varphi''(f(x)+o(f(x),h(x))h(x)\right)(x) \end{aligned}$$

On a donc :

$$\varphi_*(f+h) = \varphi_*(f) + (x \mapsto h(x)\varphi'(f(x))) + \left(x \mapsto \frac{h(x)^2}{2}\varphi''(f(x)+o(f(x),h(x)))h(x)\right)$$

et on voit aisément que $h \mapsto (x \mapsto h(x)\varphi'(f(x)))$ est linéaire. On va donc maintenant montrer la continuité de cette application linéaire. On a successivement :

$$\begin{array}{rcl} |h(x)\varphi'(f(x))| & \leq & \|h\|_\infty |\varphi'(f(x))| \\ \|x \mapsto h(x)\varphi'(f(x))\|_\infty & \leq & \|h\|_\infty \|x \mapsto \varphi'(f(x))\|_\infty \\ \|h \mapsto (x \mapsto h(x)\varphi'(f(x)))\,\| & \leq & \|\varphi \circ f\|_\infty \end{array}$$

et la continuité est démontrée. Il reste à s'occuper du o. On suppose $h \neq 0$, alors :

$$\frac{\left\| \frac{h^2}{2}\varphi''(f+\theta h)\right\|_\infty}{\|h\|_\infty} \leq \frac{\|h\|_\infty}{2} ||\varphi''(f+\theta h)||_\infty \xrightarrow[\|h\|_\infty \to 0]{} 0$$

car φ étant de classe $\mathcal{C}^2$, $\varphi''(f+\theta h)$ est continue sur $[a,b]$ et donc bornée.

φ_* est donc dérivable et $(D\varphi_*)_f = h \mapsto h(\varphi' \circ f)$.

$\boxed{\mathbf{4}}$ On a que $f(x+h, y+k) = f(x,y) + f(h,y) + f(x,k) + f(h,k)$ car le produit vectoriel est bilinéaire. Il suffit de montrer que $f(h,k) = o(h,k)$. Or on sait que $\|h \wedge k \leq |h||k|$ et comme $\|(h,k)\| = \max\{|h|,|k|\}$ on a le résultat avec $f'_{(x,y)}(h,k) = f(h,y) + f(x,k)$.

$\boxed{\mathbf{5}}$ Notons $x \mapsto \|x\|$ la norme (norme euclidienne) associée au produit scalaire de l'espace euclidien E, lui–même noté $(x,y) \mapsto x.y$. La norme est la composée des fonctions suivantes :

$$x \overset{\Delta}{\longmapsto} (x,x) \overset{\cdot}{\longmapsto} x.x \overset{\sqrt{\ }}{\longmapsto} \sqrt{x.x} = \|x\|.$$

Δ est linéaire continue, le produit scalaire bilinéaire continue, et $\sqrt{\ }$ continue sur $]0,+\infty[$. Comme $\sqrt{\ }$ n'est pas dérivable en 0, et comme seul 0 est envoyé sur 0 par $. \circ \Delta$, on voit que la norme euclidienne est dérivable en tout point autre que 0.

On applique le théorème de dérivation des fonctions composées en un point $x \neq 0$ de E. En un tel point, la dérivée de cette composition est la composition des trois dérivées suivantes :

$$h \overset{\Delta}{\longmapsto} (h,h) \qquad (h,k) \longmapsto h.x + x.k \qquad h \longmapsto \frac{h}{2\sqrt{x.x}}$$

ce qui donne (compte tenu de la symétrie du produit scalaire) $h \mapsto \dfrac{x.h}{\|x\|}$.

$\boxed{\mathbf{6}}$ On a :

$$\begin{array}{rcl} g(x+h) & = & g(x(1+x^{-1}h)) \\ & = & g(x)g(1+x^{-1}h) \\ & = & g(x)(1+g'_1(x^{-1}h)+o(h)) \\ & = & g(x)+g(x)g'_1(x^{-1}h)+o(h). \end{array}$$

Compte tenu des hypothèses faites sur les groupes G et G', l'application $h \mapsto g(x)g'_1(x^{-1}h)$ est linéaire et continue. La deuxième égalité s'obtient de même, en commençant le calcul par $g(x+h) = g((1+hx^{-1})x)$.

$\boxed{7}$ **1.** Tout d'abord on voit que $\varphi(0) = e^0 = I_n$. Et si $h \in \mathbb{R}$, alors :

$$\begin{aligned} \varphi(0+h) &= e^{hA} \\ &= I_n + hA + \sum_{p=2}^{\infty} \frac{h^p A^p}{p!} \end{aligned}$$

Il est évident que $h \mapsto hA$ est linéaire. Il faut donc montrer que $\displaystyle\sum_{p=0}^{\infty} \frac{h^p A^p}{p!} = o(h)$.

$$\| \sum_{p=2}^{\infty} \frac{h^n A^n}{n!} \| \leq |h|^2 \sum_{p=2}^{\infty} \frac{|h|^{p-2} \|A^p}{\|} p!$$

Et d'où on a bien que φ est dérivable en 0 et que $\varphi'(0) = A$

2. Le déterminant est dérivable sur $\mathcal{M}_n(\mathbb{R})$ entier donc en particulier en $\varphi(0) = I_n$, et par composition on a : $\psi'(0) = \det'_{\varphi(0)}(\varphi'(0)) = \det'_{I_n}(A) = tr(A)$.

3. Comme l'exponentielle est un morphisme de groupe $(\mathbb{R}, +)$ dans $(\mathrm{GL}_n(\mathbb{R}), \circ)$ et que le déterminant est un morphisme de groupe $(\mathrm{GL}_n(\mathbb{R}), \circ)$ dans $(\mathbb{R}^*, .)$, par composition ψ est un morphisme de groupe de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, .)$. Or on montre facilement que les applications de la forme $x \mapsto \pm \exp(\lambda x)$ pour $\lambda \in \mathbb{R}$ sont déjà de tel morphisme. On va montrer que tous les morphismes de groupe de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, .)$ sont de cette forme et on auras alors que $\exists \lambda \in \mathbb{R}$ tel que $\forall x \in \mathbb{R}$, $\det(e^{xA}) = e^{\lambda x}$ car $\det(e^0) = \det(I_n) = 1 > 0$.

Soit donc f un morphisme de groupe de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, .)$. Supposons que $f(1) > 0$, sinon $-f$ est aussi un morphisme de groupe et change f en $-f$. Alors on a :

$$\forall n \in \mathbb{N},\ f(n) = f(\underbrace{1 + \cdots + 1}_{n}) = \underbrace{f(1) \times \cdots \times f(1)}_{n} = (f(1))^n,$$

$$\begin{aligned} \forall n \in \mathbb{Z}, \quad f(n) &= \begin{cases} (f(1))^n & \text{si } n \geq 0 \\ f(-(-n)) = (f(-n))^{-1} = \left(f(1)^{-n}\right)^{-1} & \text{si } n < 0 \end{cases} \\ \iff f(n) &= (f(1))^n \end{aligned}$$

$$\begin{aligned} \forall (p,q) \in \mathbb{Z} \times \mathbb{N}^*, \qquad & f(1) = f\left(\frac{q}{q}\right) = \left(f\left(\frac{1}{q}\right)\right)^q \\ \implies \qquad & f\left(\frac{1}{q}\right) = (f(1))^{\frac{1}{q}} \\ \implies \qquad & f\left(\frac{p}{q}\right) = \left(f\left(\frac{1}{q}\right)\right)^p = \left(f(1)^{\frac{1}{q}}\right)^p = f(1)^{\frac{p}{q}} \end{aligned}$$

Comme $f(1) > 0$ par hypothèse, $\exists \lambda \in \mathbb{R}$, $f(1) = e^{\lambda}$. On considère la fonction g définie sur $\mathbb{R}$ par $g(x) = e^{\lambda x}$. On a alors que $f = g$ sur $\mathbb{Q}$ et comme $\overline{\mathbb{Q}} = \mathbb{R}$, on a que $f = g$. D'où ce qu'il fallait démontrer.

On a donc l'éxistence d'un $\lambda \in \mathbb{R}$ tel que $\forall x \in \mathbb{R}$, $\det(e^{xA}) = e^{\lambda x}$, d'où $\psi'(x) = \lambda e^{\lambda x}$ et de $\psi'(0) = \mathrm{tr}\,(A)$ on déduit que $\lambda = \mathrm{tr}\,(A)$. On vient donc bien de démontrer que $\det(e^{xA}) = e^{\mathrm{tr}\,(A)x}$.

$\boxed{8}$ f est de classe $\mathcal{C}^\infty$ sur D et on a :

$$\begin{aligned}
\frac{\partial f}{\partial x}(x,y) &= \frac{1}{1+x^2} - \frac{\frac{1-xy+y(x+y)}{(1-xy)^2}}{1+\left(\frac{x+y}{1-xy}\right)^2} \\
&= \frac{1}{1+x^2} - \frac{1-xy+y(x+y)}{(1-xy)^2+(x+y)^2} \\
&= \frac{1}{1+x^2} - \frac{1+y^2}{(1-xy)^2+(x+y)^2}
\end{aligned}$$

$$\begin{aligned}
\text{or} \quad (1-xy)^2+(x+y)^2 &= 1-2xy+x^2y^2+x^2+y^2+2xy \\
&= 1+x^2+y^2+x^2y^2 \\
&= (1+x^2)(1+y^2)
\end{aligned}$$

$$\begin{aligned}
\text{donc} \quad \frac{\partial f}{\partial x} &= \frac{1+y^2-y^2-1}{(1+x^2)(1+y^2)} \\
&= 0
\end{aligned}$$

De même, on a pour la dérivée par rapport à y :

$$\begin{aligned}
\frac{\partial f}{\partial y} &= \frac{1}{1+y^2} - \frac{1+x^2}{(1-xy)^2+(x+y)^2} \\
&= \frac{1}{1+y^2} - \frac{1+x^2}{(1+x^2)(1+y^2)} \\
&= 0
\end{aligned}$$

On a donc que $f'=0$ et donc que f est constante sur chaques composantes connexes de D.

On note :

$$\begin{aligned}
D_1 &= \left\{(x,y)\in\mathbb{R}^2,\ x<0,\ y<\frac{1}{x}\right\} \\
D_2 &= \left\{(x,y)\in\mathbb{R}^2,\ y>\frac{1}{x} \text{ si } x<0,\ y\in\mathbb{R} \text{ si } x=0,\ y<\frac{1}{x} \text{ si } x>0\right\} \\
D_3 &= \left\{(x,y)\in\mathbb{R}^2,\ x>0,\ y>\frac{1}{x}\right\}
\end{aligned}$$

$\lim\limits_{x\to\infty} f(x,x)=\pi$ donc $f\Big|_{D_1}=\pi$.

On a $(0,0)\in D_2$ et $f(0,0)=0$ donc $f\Big|_{D_2}=0$.

$\lim\limits_{x\to-\infty} f(x,x)=-\pi$ donc $f\Big|_{D_3}=-\pi$.

$\boxed{9}$ a) La dérivée en x de ψ est $h\mapsto(f'_x-f'_{x_0})(h)$. Or, par hypothèse, $x\mapsto f'_x$ est continue. Il en résulte que pour x dans un voisinage convenable V de x_0, on a $\|f'_x-f'_{x_0}\|\le k$. Comme il est possible de prendre V convexe, il en résulte immédiatement (cours) que ψ est k–lipschitzienne sur V.

b) Si la norme de g était nulle, g serait la fonction nulle, et E serait réduit à 0. On a $h=g(f'_{x_0}(h))$, donc $\|h\|\le\|g\|\,\|f'_{x_0}(h)\|$, d'où le résultat.

c) Prenons $k < \dfrac{1}{\|g\|}$. On a un voisinage V de x_0 sur lequel ψ est k–lipschitzienne. Soient x et y deux points de V. On a alors, en supposant $x \neq y$, $\|\psi(x) - \psi(y)\| < \dfrac{1}{\|g\|}\|x - y\|$. Si $f(x) = f(y)$, on a $\|\psi(x) - \psi(y)\| = \|f'_{x_0}(x - y)\| \geq \dfrac{1}{\|g\|}\|x - y\|$, ce qui est impossible.

$\boxed{\mathbf{10}}$ Soit la fonction :

$$\begin{array}{ccc} \mathbb{R}^2 & \overset{F}{\longrightarrow} & \mathbb{R}^2 \\ (x,y) & \longmapsto & \left(\dfrac{1}{4}\sin(x+y),\ 1 + \dfrac{2}{3}\arctan(x-y)\right) \end{array}$$

Pour montrer que le système admet une unique solution, il suffit de montrer que F a un unique point fixe. On va donc montrer que F est contractante et en appliquant le théorème du point fixe, on auras ce qu'il fallait démontrer car $\mathbb{R}^2$ est complet et non vide.

On voit aisément que F est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$. En nottant alors $\mathrm{Jac}_{(x,y)}F$ la matrice jacobienne de F en (x,y), on a :

$$\mathrm{Jac}_{(x,y)}F = \begin{pmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\[2ex] \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{1}{4}\cos(x+y) & \dfrac{1}{4}\cos(x+y) \\[2ex] \dfrac{2}{3\,(1+(x-y)^2)} & \dfrac{-2}{3\,(1+(x-y)^2)} \end{pmatrix}$$

On pose alors $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \mathrm{Jac}_{(x,y)}F$ pour simplifier les écritures. Et on a donc que $\forall \begin{pmatrix} h \\ k \end{pmatrix} \in \mathbb{R}^2$,

$$\left\| M \begin{pmatrix} h \\ k \end{pmatrix} \right\| \leq \|M\| \left\| \begin{pmatrix} h \\ k \end{pmatrix} \right\| \text{ avec } \left\| \begin{pmatrix} h \\ k \end{pmatrix} \right\| = |h| + |k|$$

$$\text{et } M \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} ah + bk \\ ch + dk \end{pmatrix}$$

d'où l'on a :

$$\begin{aligned} \left\| M \begin{pmatrix} h \\ k \end{pmatrix} \right\| &= |ah + bk| + |ch + dk| \\ &\leq \big(|a| + |c|\big)|h| + \big(|b| + |d|\big)|k| \\ &= \big(|a| + |c|\big)\big(|h| + |k|\big) \text{ car } a = b \text{ et } c = -d \end{aligned}$$

or :

$$\begin{aligned} |a| + |c| &= \frac{1}{4}\,|\cos(x+y)| + \frac{2}{3\,(1+(x-y)^2)} \\ &\leq \frac{1}{4} + \frac{2}{3} = \frac{3+8}{12} = \frac{11}{12} < 1 \\ |b| + |d| &\leq \frac{1}{4} + \frac{2}{3} = \frac{11}{12} \end{aligned}$$

d'où on déduit que $\|M\| \leq \dfrac{11}{12}$ et par inégalitée de la moyenne, F est $\dfrac{11}{12}$ -contractante.

$\boxed{\mathbf{11}}$ Soient $\alpha \in]0, +\infty[$, $h \in \mathbb{R}$ et $t \in [0,1]$, on a :

$$\begin{aligned}
\varphi(\alpha+h)(t) &= t^{\alpha+h} \\
&= t^{\alpha}t^{h} \\
&= t^{\alpha}e^{h\ln(t)} \\
&= t^{\alpha}(1+h\ln(t)+o_t(h)) \\
&= t^{\alpha}+ht^{\alpha}\ln(t)+t^{\alpha}o_t(h)
\end{aligned}$$

La fonction $h \mapsto (t \mapsto ht^{\alpha}\ln(t))$ est évidemment linéaire et continue puisque son espace de départ est de dimension 1. On a donc :

$$\varphi(\alpha+h) = \varphi(\alpha) + (t \mapsto ht^{\alpha}\ln(t)) + (t \mapsto t^{\alpha}o_t(h)).$$

Pour montrer que φ est dérivable, il suffit donc de montrer que la fonction $t \mapsto t^{\alpha}o_t(h)$ est négligeable devant h quand h tend vers 0. Or, on a $o_t(h) = t^h - 1 - h\ln(t)$. Le signe de cette expression ne dépend pas de t. En effet, en la dérivant par rapport à t, on obtient $h\dfrac{t^h-1}{t}$ qui ne s'annule pas dans l'intervalle ouvert $]0,1[$. Pour majorer l'intégrale :

$$\int_0^1 |t^{\alpha}(t^h-1-h\ln(t))|dt$$

il suffit donc de majorer

$$\left|\int_0^1 t^{\alpha}(t^h-1-h\ln(t))dt\right|$$

Or, on a

$$\int_0^1 t^{\alpha+h}dt = \frac{1}{\alpha+1+h} = \frac{1}{\alpha+1} - \frac{h}{(\alpha+1)^2} + o(h)$$

où $o(h)$ est négligeable devant h quand h tend vers 0. En particulier, on a :

$$\int_0^1 t^{\alpha}dt = \frac{1}{\alpha+1},$$

et enfin, en intégrant par parties :

$$\int_0^1 t^{\alpha}\ln(t)dt = \frac{1}{\alpha+1}\left[t^{\alpha+1}\ln(t)\right]_0^1 - \frac{1}{\alpha+1}\int_0^1 t^{\alpha+1}\frac{1}{t}dt = -\frac{1}{(\alpha+1)^2}$$

On voit donc finalement que :

$$\int_0^1 t^{\alpha}(t^h-1-h\ln(t))dt = o(h)$$

On a donc $\varphi'(\alpha) = t \mapsto t^{\alpha}\ln(t)$.

12 La forme polaire $f : E \times E \longrightarrow \mathbb{R}$ est continue, car donnée par la formule $f(x,y) = \frac{1}{4}(q(x+y) - q(x) - q(y))$. Comme $q(x) = f(x,x) = f \circ \Delta$ (où $\Delta(x) = (x,x)$), le théorème de dérivation des fonctions composées donne :

$$q'_x(h) = f'_{\Delta(x)}(\Delta(h)) = f'_{(x,x)}(h,h).$$

(Noter que Δ est linéaire.) Par ailleurs, f étant bilinéaire continue et symétrique, on a $f'_{(x,x)}(h,h) = f(x,h) + f(h,x) = 2f(x,h)$. Le reste est évident, puisque q' est une fonction linéaire de x.

13 Il suffit d'appliquer le théorème de dérivation des fonctions composées. On a d'abord :

$$(g \circ f)'_x = g'_{f(x)} \circ f'_x.$$

Noter que $(g\circ f)'$ est une application de U vers $\mathcal{L}(E,G)$. C'est cette application qu'on veut dériver en x. Pour ce faire, on va la décomposer en un produit de deux applications, comme suit :

$$\begin{array}{ccccc} U & \longrightarrow & \mathcal{L}(E,F)\times\mathcal{L}(F,G) & \xrightarrow{\circ} & \mathcal{L}(E,G) \\ x & \mapsto & (f'_x, g'_{f(x)}) & \mapsto & (g\circ f)'_x \end{array}$$

La première de ces applications a deux composantes qui sont respectivement $x\mapsto f'_x$ et $x\mapsto g'_{f(x)}$. Pour la dériver en x, il suffit de dériver chaque composante en x. La première composante se dérive facilement, puisqu'elle est simplement f'. Sa dérivée en x est donc f''_x. La deuxième composante est la composition :

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g'} & \mathcal{L}(F,G) \\ x & \mapsto & f(x) & \mapsto & g'_{f(x)} \end{array}$$

Sa dérivée en x s'obtient donc par le théorème de dérivation des fonctions composées, ce qui donne : $g''_{f(x)}\circ f'_x$.

La dérivée en x de $x\mapsto(f'_x, g'_{f(x)})$ est donc :

$$h\mapsto (f''_x(h), g''_{f(x)}(f'_x(h))).$$

Nous devons maintenant dériver la composition $\circ : \mathcal{L}(E,F)\times\mathcal{L}(F,G)\longrightarrow\mathcal{L}(E,G)$. Noter que compte tenu des conventions courantes sur la composition, cette application envoie le couple (u,v) sur $v\circ u$, et non pas $u\circ v$. Comme elle est bilinéaire, sa dérivée en (u,v) est :

$$(h,k)\mapsto v\circ h + k\circ u.$$

Dans la situation présente, on doit la dériver en $(f'_x, g'_{f(x)})$, ce qui donne :

$$(h,k)\mapsto g'_{f(x)}\circ h + k\circ f'_x.$$

En rassemblant maintenant tous ces résultats, on obtient la formule suivante pour $h\mapsto(g\circ f)''_x(h)$:

$$h\mapsto g'_{f(x)}\circ f''_x(h) + g''_{f(x)}(f'_x(h))\circ f'_x,$$

ce qui fait que

$$\begin{aligned} (g\circ f)''_x(h,k) &= (g\circ f)''_x(h)(k) \\ &= g'_{f(x)}(f''_x(h)(k)) + g''_{f(x)}(f'_x(h))(f'_x(k)) \\ &= g'_{f(x)}(f''_x(h,k)) + g''_{f(x)}(f'_x(h), f'_x(k)). \end{aligned}$$

$\boxed{\mathbf{14}}$ a) C'est immédiat, puisque le produit scalaire s'exprime en fonction de la norme par la formule : $x.y = \dfrac{\|x+y\|-\|x\|-\|y\|}{2}$. f'_x a un noyau réduit à 0, puisque qu'il respecte la norme. Comme c'est un endomorphisme d'un espace de dimension finie, f'_x est bijectif.

b) Pour h et k fixés, dérivons les deux fonctions $x\mapsto f'_x(h).f'_x(k)$ et $x\mapsto h.k$ (qui sont égales d'après la question précédente). La deuxième a une dérivée nulle, puisqu'elle est constante. La première est la composition des deux fonctions suivantes : $x\mapsto(f'_x(h), f'_x(k))$ et $(u,v)\mapsto u.v$.

La fonction $x\mapsto f'_x(h)$ est elle même la composée des deux fonctions : $x\mapsto f'_x$ et $\varphi\mapsto\varphi(h)$, la deuxième étant linéaire continue.

$x\mapsto f'_x$ a pour dérivée en x : $l\mapsto f''_x(l)$, et $\varphi\mapsto\varphi(h)$ a pour dérivée elle–même en tout point. Donc $x\mapsto f'_x(h)$ a pour dérivée en x : $l\mapsto f''_x(l)(h)$.

Comme par ailleurs $(u,v) \mapsto u.v$ a pour dérivée en (u,v) : $(\alpha,\beta) \mapsto \alpha.v + u.\beta$, on voit que la dérivée en x cherchée est : $l \mapsto f''_x(l,h).f'_x(k) + f'_x(h).f''_x(l,k)$.

Or cette dérivée étant nulle, on a $A_x(h,k,l) + A_x(k,h,l) = 0$.

c) La question précédente a montré que $(h,k,l) \mapsto A_x(h,k,l)$ est antisymétrique relativement aux deux premières variables. Par ailleurs, le théorème de Schwarz montre qu'elle est symétrique par rapport aux deux dernières. On a donc $A_x(h,k,l) = A_x(h,l,k) = -A_x(l,h,k) = -A_x(l,k,h) = A_x(k,l,h) = A_x(k,h,l) = -A_x(h,k,l)$, ce qui montre que $A_x(h,k,l) = 0$.

On a donc $f'_x(h).f''_x(l,k) = 0$. Comme f'_x est surjectif, et le produit scalaire non dégénéré, on en déduit que $f''_x(k,l) = 0$ pour tous k, l et x, donc $f''_x = 0$, pour tout x.

d) Comme f'' est la fonction nulle, f' est constante (application du théorème de la moyenne). On a donc $f'_x(h) = u(h)$, où u est une application linéaire de E vers E. La fonction $x \mapsto f(x) - u(x)$ a donc elle aussi une dérivée nulle. Elle est donc constante (disons de valeur C) et on a $f(x) = u(x) + C$, c'est–à–dire que f est une fonction affine.

$\boxed{\textbf{15}}$ La formule de Taylor–Young à l'ordre 2, donne pour x dans $\mathcal{A}$ et t réel :

$$g(1+tx) = 1 + g'_1(tx) + b(tx,tx) + o(\|tx\|^2) = 1 + tg'_1(x) + t^2 b(x,x) + o(t^2),$$

où b est une application bilinéaire symétrique de $\mathcal{A} \times \mathcal{A}$ vers $\mathcal{B}$. Dans cette formule, on peut remplacer tx par ty, mais aussi par $tx + ty + t^2xy$. On a alors, puisque $g(1 + tx + ty + t^2xy) = g(1+tx)g(1+ty)$, et en isolant les coefficients de t^2 :

$$g'_1(xy) - g'_1(x)g'_1(y) = 2b(x,y).$$

Le résultat en découle en soustrayant des deux membres de cette égalité les deux membres de celle obtenue en échangeant x et y.

$\boxed{\textbf{16}}$ a) La fonction étant une somme, il suffit de calculer les dérivées de chaque terme de cette somme. Posons donc $g(x) = \|x-a\|^2 = (x-a).(x-a)$, où le point représente le produit scalaire.

L'application $x \mapsto x - a$ étant affine de dérivée $h \mapsto h$ en tout point, et le produit scalaire étant bilinéaire, on a :

$$g'_x(h) = (x-a).h + h.(x-a) = 2(x-a).h.$$

L'application $g' : E \longrightarrow \mathcal{L}(E,\mathbb{R})$ est donc $x \mapsto (h \mapsto 2(x-a).h)$. C'est la composée de l'application affine $x \mapsto x - a$ et de l'application linéaire $x \mapsto (h \mapsto 2x.h))$. On a donc : $g''(k) = (h \mapsto 2k.h)$, ou encore : $g''(k,h) = 2k.h$.

La dérivée première de f est donc

$$h \mapsto \left(2\sum_{i=0}^{p}(x-a_i)\right).h,$$

et sa dérivée seconde est

$$(h,k) \mapsto 2(p+1)h.k.$$

b) En un extremum relatif, la dérivée première doit s'annuler, ce qui ne peut arriver que si $\displaystyle\sum_{i=0}^{p}(x-a_i) = 0$, c'est–à–dire au point $\displaystyle\frac{1}{p+1}\sum_{i=0}^{p} a_i$. Par ailleurs, la dérivée seconde étant $2(p+1)$ fois le produit scalaire, elle est définie positive, ce qui montre que f a un minimum relatif en ce point.

$\boxed{\mathbf{17}}$ a) $\Psi(X)$ étant une somme, il suffit de dériver chaque terme de la somme. Noter que $\mathrm{tr}\,({}^tXX)$ n'est autre que $\mathrm{tr}\,({}^t(X-0)(X-0))$.

Considérons donc la fonction Γ définie par $\Gamma(X) = \mathrm{tr}\,({}^t(X-A)(X-A))$, où A est une matrice quelconque. Cette fonction est une composition de fonctions. Notons que $X \mapsto X - A$ est affine et a pour dérivée l'application identique. La transposition et la trace sont linéaires. Enfin, le produit de matrices est bilinéaire. La dérivée de Γ est donc donnée par : $\Gamma'_X(H) = \mathrm{tr}\,({}^t(X-A)H + {}^tH(X-A)) = 2\mathrm{tr}\,({}^t(X-A)H)$. On a donc

$$\Psi'_X(H) = 2\mathrm{tr}\,({}^tXH) + 2\sum_{i=1}^{p} \mathrm{tr}\,({}^t(X-A_i)H) = 2\mathrm{tr}\,({}^t((p+1)X - \sum_{i=1}^{p} A_i)H).$$

Calculons maintenant $\Gamma''_X(H,K)$. Ici on doit dériver Γ', c'est–à–dire l'application : $X \mapsto (H \mapsto 2\mathrm{tr}\,({}^t(X-A)H))$, qui est la composée de l'application affine $X \mapsto {}^t(X-A)$, dont la dérivée est la transposition, et de l'application linéaire $U \mapsto (H \mapsto 2\mathrm{tr}\,(UH))$. On a donc : $\Gamma''_X(K) = (H \mapsto 2\mathrm{tr}\,({}^tKH))$, c'est–à–dire $\Gamma''_X(H,K) = 2\mathrm{tr}\,({}^tKH) = 2\mathrm{tr}\,({}^tHK)$. Ceci donne :

$$\Psi''_X(H,K) = 2(p+1)\mathrm{tr}\,({}^tHK).$$

b) Remarquons que $p+1$ n'est jamais nul, même si $p=0$. Pour que Ψ ait un extremum relatif en X, il est nécessaire que Ψ'_X soit nul. On doit donc avoir pour tout H : $\mathrm{tr}\,({}^t((p+1)X - \sum_{i=1}^{p} A_i)H) = 0$. Comme la forme bilinéaire symétrique $(H,K) \mapsto \mathrm{tr}\,({}^tHK)$ est définie positive, on doit avoir : $(p+1)X - \sum_{i=1}^{p} A_i = 0$, c'est–à–dire $X = \dfrac{1}{p+1}\sum_{i=1}^{p} A_i$. Il y a donc au plus un seul extremum local, se situant au point qu'on vient de calculer. (Si $p=0$, ce point est 0.)

Comme on l'a vu, la dérivée seconde de Ψ en X est indépendante de X. Elle est le produit par $2(p+1)$ (qui n'est pas nul) du produit scalaire $(H,K) \mapsto \mathrm{tr}\,({}^tHK)$. Elle est donc définie positive, ce qui montre que Ψ a, au point $\dfrac{1}{p+1}\sum_{i=1}^{p} A_i$ un minimum local.

$\boxed{\mathbf{18}}$ **1.** On a $\forall x \in \mathcal{S}^n, Y(x) = \dfrac{X(x)}{\|X(x)\|}$, et comme X est de classe $\mathcal{C}^\infty$ et que la norme aussi, on en déduit que Y est aussi de classe $\mathcal{C}^\infty$. Et par la bilinéarité du produit scalaire, on a que Y est un champ de vecteur normé orthogonal à la sphère.

Du fait que Y soit $\mathcal{C}^\infty$, on déduit que φ_t est également de classe $\mathcal{C}^\infty$ ($\varphi_t = \mathrm{Id}_{\mathcal{S}^n} + tY$). On considère $x, y \in \mathcal{S}^n$ tels que $\varphi_t(x) = \varphi_t(y)$ et on veut montrer que cela implique $x = y$. Or :

$$\begin{aligned} \varphi_t(x) = \varphi_t(y) &\iff x + tY(x) = y + tY(y) \\ &\iff \|x + tY(x) - y - tY(y)\| = 0 \end{aligned}$$

De plus on a $\|x + tY(x) - y - tY(y)\| \geq \big|\|x-y\| - t\|Y(x) - Y(y)\|\big|$. On introduit alors :

$$\begin{array}{ccc} \mathbb{R}^{n+1}\backslash\{0\} & \xrightarrow{\overline{Y}} & \mathbb{R}^{n+1} \\ x & \mapsto & \|x\| Y\left(\dfrac{x}{\|x\|}\right) \end{array}$$

et $\overline{Y}$ est toujours de classe $\mathcal{C}^\infty$ avec $\overline{Y}\Big|_{\mathcal{S}^n} = Y$. On a donc pour $x, y \in \mathcal{S}^n$, $\|Y(x) - Y(y)\| = \|\overline{Y}(x) - \overline{Y}(y)\|$. De plus comme $\overline{Y}$ est de classe $\mathcal{C}^\infty$, on a en particulier que $\overline{Y}'$ est continue sur $\mathbb{R}^{n+1}\backslash\{0\}$, donc bornée sur tous compact de $\mathbb{R}^{n+1}\backslash\{0\}$ et on a donc que $\forall z \in [x,y],\ \exists k \geq 0,\ \|\overline{Y}'_z\| \leq k$. Puis par inégalitée de la moyenne, on obtient que $\|\overline{Y}(x) - \overline{Y}(y)\| \leq k\|x-y\|$, ce qui implique que $\|x + tY(x) - y - tY(y)\| \geq |1-tk|\|x-y\|$. Il suffit alors de prendre $t > 0$ tel que $tk < 1$ pour avoir l'injectivité. En effet, avec un tel t on a $1 - tk > 0$ et donc :

$$\begin{aligned}
\varphi_t(x) = \varphi_t(y) &\iff \|\varphi_t(x) - \varphi_t(y)\| = 0 \\
&\iff 0 \geq (1-tk)\|x-y\| \geq 0 \\
&\iff x = y
\end{aligned}$$

2. Soit $x \in \mathcal{S}^n$. On a :

$$\begin{aligned}
\|\varphi_t(x)\|^2 &= \|x + tY(x)\|^2 \\
&= \|x\|^2 + t^2\|Y(x)\|^2 \text{ car } Y \perp \mathcal{S}^n \\
&= 1 + t^2 \text{ car } x \in \mathcal{S}^n \text{ et } Y \text{ normé par définition} \\
&\Rightarrow \|\varphi_t(x)\| = \sqrt{1+t^2}
\end{aligned}$$

D'où on a bien que $\varphi_t(x) \in \mathcal{S}(0, \sqrt{1+t^2})$.

3. Soit $x \in \mathcal{S}^n$. D'après la première question, $\forall K \subset B(0,1),\ K$ compact, $\exists k > 0,\ \forall y \in K,\ \|\overline{Y}'_y\| \leq k$. Et comme pour $x \in \mathcal{S}^n$, $\varphi_t = \mathrm{Id}_{\mathcal{S}^n} + \overline{Y}$, on a que $\forall h \in \mathbb{R}^{n+1},\ (\varphi_t)'_x(h) = h + t\overline{Y}'_x(h)$. D'où $\forall x \in \mathcal{S}^n$, il éxiste $K \subset \mathcal{S}^n$ tel que $x \in K$ et il éxiste $k > 0$ tel que $\|\overline{Y}'_x\| \leq k$. Puis en choisissant le t comme à la question **1.**, on a que $\|t\overline{Y}'_x\| \leq tk < 1$. On en déduit pour $h \in \mathbb{R}^{n+1}$, $h \neq 0$,

$$\begin{aligned}
(\varphi_t)'_x(h) = 0 &\iff -h = t\overline{Y}'_x(h) \\
&\iff \|h\| \leq t\|h\|\|\overline{Y}'_x\| \\
&\Longrightarrow \|h\| < \|h\| \quad \text{car } h \neq 0
\end{aligned}$$

Ce qui est impossible. D'où $(\varphi_t)'_x(h) = 0 \iff h = 0$. On vient donc de montrer que $(\varphi_t)'_x$ est injective.

Or c'est une application linéaire de $\mathbb{R}^{n+1}$ dans lui même et donc, puisque nous sommes en dimension finie, $(\varphi_t)'_x$ est un isomorphisme. Enfin, on déduit du théorème d'inversion locale que φ_t est un difféomorphisme local.

4. D'après la question **2.** on sait que $\forall x \in \mathcal{S}^n$, $\varphi_t(x) \in \mathcal{S}(0, \sqrt{1+t^2})$. Alors en notant $\mathcal{C}_{\alpha,\beta}$ la couronne de centre 0 comprise entre les rayons α et β avec $\beta > \alpha$ ($\mathcal{C}_{\alpha,\beta} = \{x \in \mathbb{R}^{n+1} \mid \alpha < \|x\| < \beta\}$), montrons que l'image par φ_t de $\mathcal{C}_{\alpha,\beta}$ est $\mathcal{C}_{\alpha\sqrt{1+t^2},\beta\sqrt{1+t^2}}$.

En effet comme φ_t est un difféomorphisme local, on a que $\varphi_t(\mathcal{C}_{\alpha,\beta})$ est un ouvert de $\mathcal{C}_{\alpha\sqrt{1+t^2},\beta\sqrt{1+t^2}}$ car $\mathcal{C}_{\alpha,\beta}$ est un ouvert et si $r \in]\alpha,\beta[$, alors $\varphi_t(\mathcal{C}_{\alpha,\beta} \cap \mathcal{S}_r) = \varphi_t(\mathcal{C}_{\alpha,\beta}) \cap \mathcal{S}_{r\sqrt{1+t^2}}$ est un ouvert de $\mathcal{S}_{r\sqrt{1+t^2}}$, en notant $\mathcal{S}_m$ la sphère de centre 0 et de rayon m. Or comme $r \in]\alpha,\beta[$, $\mathcal{C}_{\alpha,\beta} \cap \mathcal{S}_r = \mathcal{S}_r$ est donc compact et comme φ_t est continue, $\varphi_t(\mathcal{C}_{\alpha,\beta} \cap \mathcal{S}_r)$ est compact donc fermé. On vient donc de montrer que $\varphi_t(\mathbb{C}_{\alpha,\beta}) \cap \mathcal{S}_{r\sqrt{1+t^2}}$ est un non vide à la fois ouvert et fermé dans $\mathcal{S}_{r\sqrt{1+t^2}}$. On déduit, par connexité de $\mathcal{S}_m$, que $\varphi_t(\mathcal{C}_{\alpha,\beta}) \cap \mathcal{S}_{r\sqrt{1+t^2}} = \mathcal{S}_{r\sqrt{1+t^2}}$.

On a donc que $\forall r \in]\alpha,\beta[$, $\varphi_t(\mathcal{C}_{\alpha,\beta} \cap \mathcal{S}_r) = \mathcal{S}_{r\sqrt{1+t^2}}$. Or $\mathcal{C}_{\alpha,\beta} = \displaystyle\bigcup_{\alpha<r<\beta} \mathcal{S}_r$, d'où on a bien que $\varphi_t(\mathcal{C}_{\alpha,\beta}) =$

$\mathcal{C}_{\alpha\sqrt{1+t^2},\beta\sqrt{1+t^2}}$. Et de plus on a :

$$\begin{aligned}\mathcal{C}_{\alpha\sqrt{1+t^2},\beta\sqrt{1+t^2}} &= \{x \in \mathbb{R}^{n+1},\ \mid \alpha\sqrt{1+t^2} < \|x\| < \beta\sqrt{1+t^2}\} \\ &= \{x \in \mathbb{R}^{n+1},\ \mid \frac{\|x\|}{\sqrt{1+t^2}} \in \mathcal{C}_{\alpha,\beta}\} \\ &= \sqrt{1+t^2}\mathcal{C}_{\alpha,\beta}\end{aligned}$$

D'où $\varphi_t(\mathcal{C}_{\alpha,\beta}) = \sqrt{1+t^2}\mathcal{C}_{\alpha,\beta}$.

Alors si on note V le volume de $\mathcal{C}_{\alpha,\beta}$ et V' celui de $\varphi_t(\mathcal{C}_{\alpha,\beta})$, on a alors que $V' = \left(\sqrt{1+t^2}\right)^{n+1} V$. D'autre part,

$$\begin{aligned}V' &= \int_{\mathcal{C}_{\alpha\sqrt{1+y^2},\beta\sqrt{1+t^2}}} du \\ &= \int_{\mathcal{C}_{\alpha,\beta}} \det((\varphi_t)'_x)dx\end{aligned}$$

Or, de $(\varphi_t)'_x = \mathrm{Id} + t\overline{Y}'_x$ on a $\mathrm{Jac}_x(\varphi_t) = \mathrm{I}_{n+1} + t\mathrm{Jac}_x(Y)$ où Jac $f$est la matrice jacobienne de f, d'où on a pour t suffisament petit, $\det((\varphi_t)'_x) > 0$. On voit alors que $\det((\varphi_t)'_x)$ est un polynôme en t de degré au plus $n+1$. On pose $\det((\varphi_t)'_x) = \sum_{i=0}^{n+1} A_i(x)t^i$. Alors, par linéarité de l'intégrale, $V' = \sum_{i=0}^{n+1} t^i \int_{\mathcal{C}_{\alpha,\beta}} A_i(x)dx$, d'où V' est un polynôme en t de degré $n+1$ au plus.

Mais comme on avait déjà que $V' = \left(\sqrt{1+t^2}\right)^{n+1} V$ et que V est une constante, on déduit que $\left(\sqrt{1+t^2}\right)^{n+1}$ est un polynôme en t.

On a donc $\left(\sqrt{1+t^2}\right)^{n+1} = P(t)$, d'où $P(t)^2 = \left(1+t^2\right)^{n+1}$. Puis, par unicité de la décomposition en produit de polynômes irréductibles dans $\mathbb{R}[X]$, et comme il y a un nombre forcément paire de facteur $\left(1+t^2\right)$ dans la décomposition de P^2, on a qu'il y a également un nombre paire de facteur $\left(1+t^2\right)$ dans le polynôme $\left(1+t^2\right)^{n+1}$, autrement dit $n+1$ est paire et donc n est impaire.