CHAPITRE 38

Calcul différentiel

Soit E, F, G des $\mathbb{R}$-espaces vectoriels normés de **dimension finie** munis de produits scalaires et normes respectives $\langle\ ,\ \rangle_E$, $\langle\ ,\ \rangle_F$, $\langle\ ,\ \rangle_G$, $\|.\|_E$, $\|.\|_F$, $\|.\|_G$. Lorsqu'il n'y a pas ambiguïté sur l'espace où l'on travaille, on notera la norme et le produit scalaire tout simplement $\|.\|$ et $\langle\ ,\ \rangle$. Soit Ω un ouvert de E.

I Applications linéaires et hyperplans

On commence par voir quelques éléments qui nous seront utiles par la suite.

Proposition I.1.

Soit $a \in E$ et soit f l'application

$$f : \begin{cases} E & \longrightarrow \mathbb{R} \\ x & \longmapsto \langle a,\ x\rangle\,. \end{cases}$$

L'ensemble $H(f) = \{(x, f(x)),\ x \in E\}$ est un hyperplan de l'espace vectoriel $E \times \mathbb{R}$.

Démonstration. On a pour tout $(u, v) \in E \times \mathbb{R}$,

$$(u, v) \in H(f) \iff v = f(u) \iff v - \langle a,\ u\rangle = 0.$$

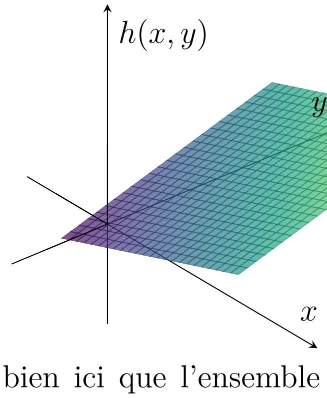
L'application $\psi : (u, v) \longmapsto v - \langle a,\ u\rangle$ est une forme linéaire de $\mathbb{E} \times \mathbb{R}$. L'équivalence précédente nous donne que $H = \operatorname{Ker}\psi$, *i.e.* H est le noyau d'une forme linéaire, ce qui signifie que H est un hyperplan de $E \times \mathbb{R}$. □

Exemple. Observons les deux exemples suivants.

Considérons le cas où $E = \mathbb{R}^2$, et l'application

$$h : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto \langle (x, y),\ (1, 1)\rangle = x + y. \end{cases}$$

La surface représentative de h, $H(h)$ est la suivante.

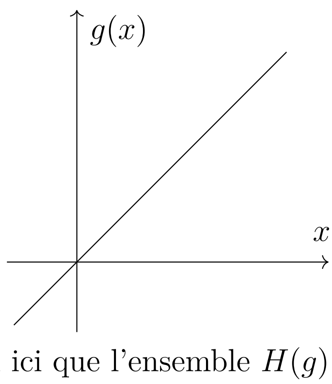


On voit bien ici que l'ensemble $H(h)$ est un plan (hyperplan en dimension 3) de $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$.

Considérons le cas où $E = \mathbb{R}$, et l'application

$$g : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto 2x. \end{cases}$$

La courbe représentative de g, $H(g)$ est la suivante.



On voit bien ici que l'ensemble $H(g)$ est une droite (hyperplan en dimension 2) de $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Remarque. Bien entendu, dans la proposition I.1, si f était de la forme $x \longmapsto b + \langle a,\ x\rangle$ avec $b \in \mathbb{R}$, $H(f)$ serait un hyperplan affine de $E \times \mathbb{R}$.

(cc) BY-NC-SA

II Différentiabilité

Dans cette partie, on suppose que f est une application de Ω dans E.

Définition II.1.

Soit $g : \Omega \longrightarrow F$ et $h : \Omega \longmapsto F$ deux applications de Ω dans F et $a \in \Omega$. On dit que h est dominée par g au voisinage de a lorsque

$$\exists \eta > 0,\ \exists C \geq 0,\ \forall x \in \Omega,\ x \in B(a, \eta) \Longrightarrow \|h(x)\|_F \leq C \|g(x)\|_F .$$

Dans ce cas, on note $h(x) \underset{x \to a}{=} O(g(x))$. De même, on dit que h est négligeable devant g au voisinage de a lorsque

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \Omega,\ x \in B(a, \eta) \Longrightarrow \|h(x)\|_F \leq \varepsilon \|g(x)\|_F .$$

Dans ce cas, on note $h(x) \underset{x \to a}{=} o(g(x))$. Lorsqu'il n'y a pas d'ambiguïté sur le point a, on note tout simplement dans les deux cas respectivement $h(x) = O(g(x))$ et $h(x) = o(g(x))$.

Remarques.

→ Bien entendu, on a

$$\begin{aligned} h(x) \underset{x \to a}{=} o(g(x)) &\Longleftrightarrow \frac{\|h(x)\|_F}{\|g(x)\|_F} \underset{x \to a}{\longrightarrow} 0 \text{ lorsque } g \text{ est non nulle sur un voisinage de } a \\ &\Longleftrightarrow \exists \varepsilon \in F^E,\ \exists \eta > 0, \forall x \in B_E(a, \eta),\ h(x) = \varepsilon(x) \|g(x)\|_F ,\ \text{ et } \varepsilon(x) \underset{x \to a}{\longrightarrow} 0, \end{aligned}$$

où, $B_E(a, \eta)$ est la boule de rayon η centrée en a pour la distance $(x, y) \longmapsto \|x - y\|_E$. Remarquons que lorsque $g(x) = x - a$ pour tout x, on n'a pas besoin que x soit dans un voisinage de a. En effet, la condition imposant que x soit dans un voisinage de a nous permet d'éviter le cas où $h(x) \neq 0$ et $g(x) = 0$, ce qui donne $h(x) = 0$. Dans le cas où $g(x) = x - a$, g est toujours non nulle (excepté en a où par définition $h(a) = 0$ et donc on définit $\varepsilon(a) = 0$), et donc on peut définir $\varepsilon(x)$ pour tout $x \neq a$ comme $\dfrac{h(x)}{\|g(x)\|_F}$.

→ En dimension finie, on peut remplacer les normes dans la définition par n'importe quelles normes sur E et sur F, car toutes les normes sont équivalentes en dimension finie (voir chapitre 11.4).

→ Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction réelle et $a \in \mathbb{R}$. On sait que g est dérivable en a si et seulement s'il existe $l \in \mathbb{R}$ tel que

$$\frac{g(x) - g(a)}{x - a} \underset{x \to a}{\longrightarrow} l \Longleftrightarrow g(x) \underset{x \to a}{=} g(a) + l \times (x - a) + o(x - a).$$

On veut étendre cette définition de dérivée aux fonctions dont l'espace de départ est un espace vectoriel de dimension supérieure à 1. La définition suivant cette remarque nous donne un moyen de le faire.

(cc) BY-NC-SA

Définition II.2.

Soit $a \in \Omega$. On dit que f est différentiable en a si et seulement s'il existe une application linéaire $\varphi \in \mathcal{L}(E, F)$ telle que

$$f(x) \underset{x \to a}{=} f(a) + \varphi(x-a) + o(x-a),$$

ou d'une manière équivalente

$$f(a+h) \underset{h \to 0}{=} f(a) + \varphi(h) + o(h).$$

Lorsque φ existe, il s'agit de la seule application linéaire qui vérifie la propriété ci-dessus. on dit que φ est la différentielle de f en a, et on la note df_a.

Démonstration. Montrons l'unicité de φ. Supposons qu'il existe φ et ψ vérifiant la même propriété. Dans ce cas, on a

$$f(a) + \varphi(h) + o(h) = f(a+h) = f(a) + \psi(h) + o(h),$$

et donc $(\varphi - \psi)(h) = o(h)$. Posons $v = \varphi - \psi$. v est clairement une application linéaire de E dans F, et on a $v\left(\frac{h}{\|h\|_F}\right) \xrightarrow[h \to 0]{} 0$. Soit $x \in E \setminus \{0\}$. Posons $x_n = \frac{x}{n+1}$. On a, étant donné que $x_n \xrightarrow[n \to +\infty]{} 0$,

$$v\left(\frac{x}{\|x\|}\right) = v\left(\frac{x_n}{\|x_n\|}\right) \xrightarrow[n \to +\infty]{} 0$$

et donc $v(x) = 0$. On en déduit donc que v et nulle sur E et que finalement $\varphi = \psi$.

$\square$

Remarque.

→ En dimension 1, i.e. lorsque $\Omega \subset \mathbb{R}$, la définition II.2 est cohérente avec celle de la dérivée. En effet, si $f : \Omega \longrightarrow \mathbb{R}$ est différentiable en a, on écrit

$$f(x) \underset{x \to a}{=} f(a) + \varphi(x-a) + o(x-a). \tag{1}$$

φ est une forme linéaire de $\mathbb{R}$, donc il existe $u \in \mathbb{R}$ tel que pour tout $x \in \mathbb{R}$, $\varphi(x) = ux$ (le lecteur non convaincu de cette affirmation peut la montrer lui-même). On en déduit donc que l'égalité 1 s'écrit

$$f(x) \underset{x \to a}{=} f(a) + u \times (x-a) + o(x-a) \quad i.e. \quad \frac{f(x) - f(a)}{x-a} \xrightarrow[x \to a]{} u.$$

On en déduit que la différentiabilité est équivalente à la dérivabilité sur $\mathbb{R}$, et donc qu'il s'agit bien d'une généralisation de la dérivabilité aux $\mathbb{R}$-espaces vectoriels normés quelconques.

→ En réexaminant la définition II.2, on voit qu'écrire $f(x) \underset{x \to a}{=} f(a) + \varphi(x-a) + o(x-a)$ signifie qu'au voisinage de a, on peut dire que $f(x) \simeq f(a) + \varphi(x-a)$, c'est-à-dire qu'on peut approximer f par la somme d'une constante et une application linéaire appliquée à $x - a$ (*i.e.* une application linéaire affine dont la courbe représentative est une hyperplan affine dans le cas $F = \mathbb{R}$, comme vu dans la première section de ce chapitre). Autrement dit, à une constante près, l'application linéaire φ est l'application linéaire qui approxime f le mieux au voisinage de a. En particulier, écrire $g(x) \underset{x \to a}{=} g(a) + l \times (x-a) + o(x)$ signifie qu'au voisinage de a, on peut dire que $g(x) \simeq g(a) + l \times (x-a)$, c'est-à-dire qu'on peut approximer g au voisinage de a par une fonction affine, *i.e.* une droite. Une

(cc) BY-NC-SA

autre manière de le voir est de dire que lorsque x se déplace légèrement à partir de a, la variation de $f(x)$ (égale à $f(x) - f(a)$) est approximativement linéaire en la variation de x (égale à $x - a$) à l'ordre 1 près (c'est-à-dire en négligeant les termes négligeables devant $x - a$ au voisinage de a).

Définition II.3.

Lorsque $F = \mathbb{R}$, φ est une forme linéaire, et donc il existe $u \in E$ tel que

$$f(a+h) \underset{h\to 0}{=} f(a) + \langle u,\ h\rangle_E + o(h).$$

Dans ce cas, le vecteur u est noté ∇f_a et est appelé gradient de f en a.

Remarques.

→ Attention, le gradient dépend du produit scalaire considéré. Lorsqu'on mentionnera le gradient d'une fonction sans préciser pour quel produit scalaire, il s'agira du produit scalaire duquel E est muni, $\langle .,\ .\rangle$.

→ En fait, le gradient de f en a ∇f_a, lorsqu'il est non nul, est la direction dans laquelle f croît le plus rapidement au voisinage de a. En effet, pour une petite variation $\delta \in E$, de norme fixée $\|\delta\| = r > 0$, la variation de f, $f(a+\delta) - f(a)$ est approximativement égale à $\langle \nabla f_a,\ \delta\rangle$. Cette variation est maximale lorsque δ est positivement proportionnel à ∇f_a, *i.e.* $\delta = r\dfrac{\nabla f_a}{\|\nabla f_a\|}$.

Exemples.

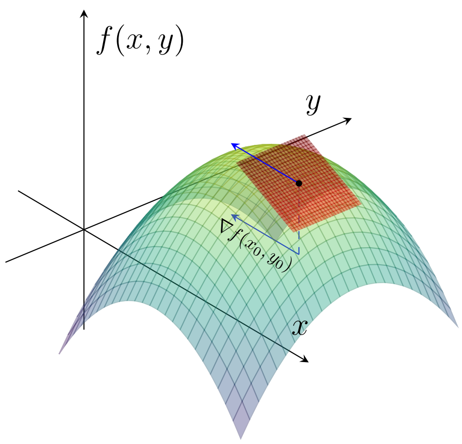
→ Observons l'exemple suivant. On considère l'application

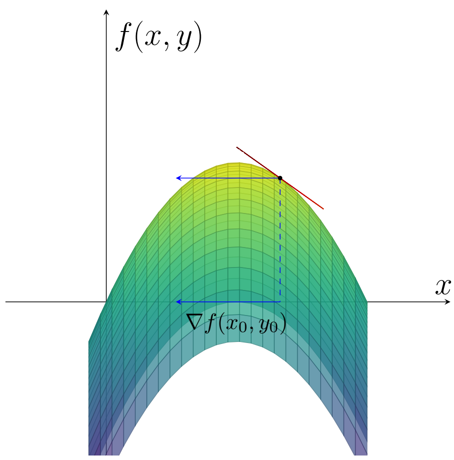
$$h : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto 1-(x-1)^2-(y-1)^2. \end{cases}$$

On a pour tout $(\delta_1,\delta_2) \in \mathbb{R}^2$ et $(x_0,y_0) \in \mathbb{R}^2$, $dh_{(x_0,y_0)}(\delta_1,\delta_2) = 2(1-x_0)\delta_1 + 2(1-y_0)\delta_2$ i.e. $\nabla h_{(x_0,y_0)} = (2(1-x_0), 2(1-y_0))$. Nous montrerons un peu plus tard un moyen pratique de trouver cette différentielle. On en déduit par exemple lorsque $(x_0,y_0) = (3,1)$, et donc $df_{(x_0,y_0)}(\delta_1,\delta_2) = -4\delta_1$. On peut donc écrire

$$\begin{aligned} f(x,y) &\underset{(x,y)\to(x_0,y_0)}{=} f(x_0,y_0) + df_{(x_0,y_0)}(x-x_0, y-y_0) + o((x-x_0,y-y_0)) \\ &\underset{(x,y)\to(3,1)}{=} -3 - 4(x-3) + o((x-3,y-1)). \end{aligned}$$

On en déduit qu'on peut écrire au voisinage de $(3,1)$, $f(x,y) \simeq -3-4(x-3) = 9 + \langle (x,y),\ (-4,0)\rangle$. Observons ces éléments sur les deux figures ci-dessus : la courbe représentative de f (en vert), la courbe de la fonction à droite de l'égalité précédente (en rouge) en (x_0,y_0) et le gradient de f en (x_0,y_0) (en bleu).





BY-NC-SA

→ Si $u \in \mathcal{L}(E, F)$, alors pour tout $a \in E$, $du_a = u$. En effet, on a bien $u(a+h) = u(a) + u(h) = u(a) + u(h) + o(h)$. Ce résultat est très naturel, car la meilleure application linéaire qui approxime u au voisinage de tout point est u elle-même.

→ Soit $u \in E$. Considérons l'application de E dans $\mathbb{R}$, $h : x \longmapsto \langle u,\, x\rangle_E$. Il s'agit d'une application linéaire, et en utilisant le point précédent, on peut facilement voir que pour tout $a \in E$, $\nabla h_a = u$.

→ Si φ est une application linéaire et $C \in F$ un vecteur constant, alors pour tout $a \in E$, $d(\varphi + C)_a = d\varphi_a$.

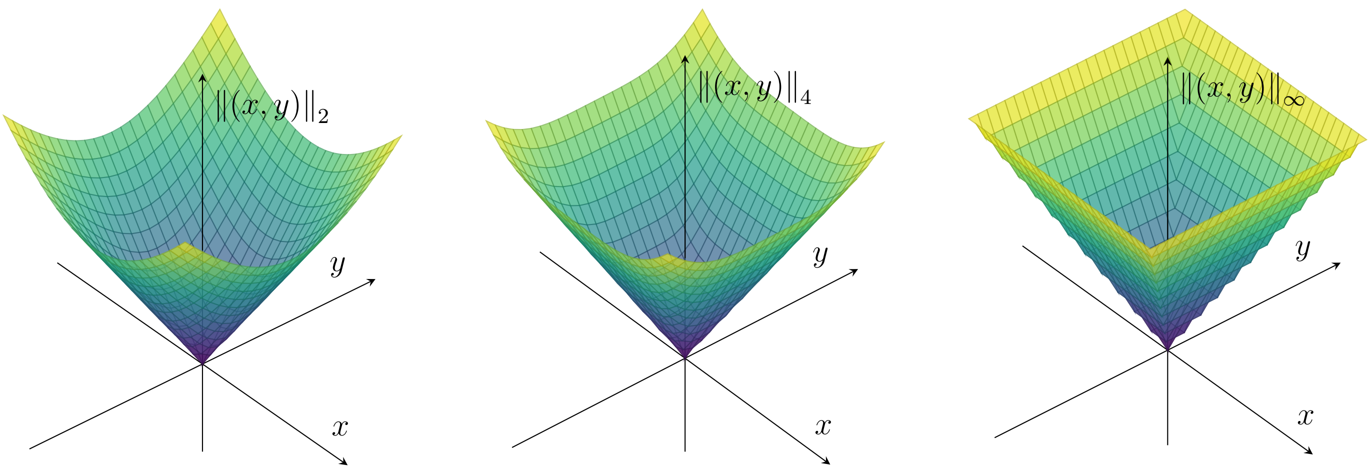
Remarque. Toute norme $N : E \longrightarrow \mathbb{R}$ sur E n'est pas différentiable en 0. Montrons ce résultat par l'absurde. Supposons que N est différentiable en 0. On pose $\varphi = dN_0$. Soit $\varepsilon : E \longrightarrow \mathbb{R}$ telle que $\varepsilon(x) \xrightarrow[x \to 0]{} 0$ et

$$N(x) = N(0) + \varphi(x) + o(x) = \varphi(x) + \|x\|_E\, \varepsilon(x).$$

On en déduit donc que pour tout $t \in \mathbb{R}^*$,

$$N(tx) = \varphi(tx) + \varepsilon(tx)\, \|tx\|_E \;\; i.e. \;\; \frac{|t|}{t} N(x) = \varphi(x) + \frac{|t|}{t}\, \|x\|_E\, \varepsilon(tx). \tag{2}$$

L'égalité 2 devient, lorsqu'on fait tendre t vers 0^+ et 0^- respectivement $N(x) = \varphi(x)$ et $N(x) = -\varphi(x)$. Ceci signifie que pour tout x, $N(x) = 0$, ce qui est absurde. On retrouve en particulier le fait que la valeur absolue n'est pas dérivale en 0, car il s'agit d'une norme sur $\mathbb{R}$.



Les trois figures ci-dessus nous donnent une idée lorsque $E = \mathbb{R}^2$. En effet, pour ces 3 normes, on remarque que leurs courbes forment un point anguleux en 0, ce qui reflète le fait que ces fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ ne sont pas différentiables en 0. En général, si l'on trace la courbe d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ et qu'on observe un point anguleux en $a \in \mathbb{R}^2$, alors nécessairement cette fonction n'est pas différentiable en a. L'inverse n'est pas nécessairement vrai, car en général, lorsqu'une fonction n'est pas différentiable en $a \in \mathbb{R}^2$, on peut observer quelque chose de plus compliqué qu'un point anguleux, comme des oscillations.

De plus, les courbes des normes ci-dessus sont symétriques par rapport à l'axe vertical (l'axe des z). En effet, ceci est cohérent, car pour tout $u \in \mathbb{R}^2$ et toute norme $\|.\|$ sur $\mathbb{R}^2$, $\|-u\| = \|u\|$.

CPGE paradise
(cc) BY-NC-SA

Exercice II.4.

Montrer que chaque application ci-dessous est différentiable et calculer sa différentielle (et son gradient lorsque l'espace d'arrivée est une partie de $\mathbb{R}$) en tout point.

1. $\pi_i : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x_1, \ldots, x_n) & \longmapsto x_i, \end{cases}$ pour tout $i \in [\![1; n]\!]$ où $n \in \mathbb{N}^*$.
2. $\psi_1 : \begin{cases} E & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto N(x)^2, \end{cases}$ où N est une norme euclidienne dans E.
3. $\psi_2 : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow \mathcal{M}_n(\mathbb{R}) \\ A & \longmapsto A^2, \end{cases}$ où $n \in \mathbb{N}^*$.

On énonce maintenant quelques propriétés importantes de la différentielle et du gradient avant de passer à la suite.

Proposition II.5.

Si $f : E \longrightarrow F$ et $g : E \longrightarrow F$ sont différentiables en $a \in E$, alors pour tout $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ est également différentiable en a, et

$$d(\lambda f + \mu g)_a = \lambda df_a + \mu dg_a.$$

De même, lorsque $G = \mathbb{R}$,

$$\nabla(\lambda f + \mu g)_a = \lambda \nabla f_a + \mu \nabla g_a.$$

Démonstration. Cette preuve ne présente pas de difficulté et est laissée comme exercice au lecteur. □

Proposition II.6.

Soit $p \in \mathbb{N}^*$, et $f_1, \ldots, f_p$ des applications de E dans F. Supposons que $f : E \longmapsto F^p$ est définie sur l'espace produit de la manière suivante

$$\forall x \in E,\ f(x) = (f_1(x), \ldots, f_p(x)).$$

Pour tout $a \in E$, f est différentiable si et seulement si $f_1, \ldots, f_p$ le sont. De plus, on a pour tout $h \in E$,

$$df_a(h) = (df_{1,a}(h), \ldots, df_{p,a}(h)).$$

Démonstration.

→ (⇒) Supposons que f est différentiable en $a \in E$ et posons $df_a = \varphi = (\varphi_1, \ldots, \varphi_p)$. On a

$$f(a+h) - f(a) - \varphi(h) = o(h),$$

ce qui donne, en substituant f et φ par leurs expressions,

$$(f_1(a+h) - f_1(a) - \varphi_1(h), \ldots, f_p(a+h) - f_p(a) - \varphi_p(h)) = o(h) = \|h\|_E\, \varepsilon(h).$$

(cc) BY-NC-SA

où ε est une fonction de E dans F^p qui tend vers 0 en 0. On note dans ce cas pour tout $h \in E$, $\varepsilon(h) = (\varepsilon_1(h), \dots, \varepsilon_p(h))$. On en déduit donc que pour tout $i \in [\![1;p]\!]$,

$$f_i(a+h) - f_i(a) - \varphi_i(a) = \|h\|_E \, \varepsilon_i(h) = o(h).$$

Ceci nous permet d'affirmer que pour tout $i \in [\![1;p]\!]$, f_i est différentiable et sa différentielle est égale à φ_i.

→ ($\Leftarrow$) Supposons que pour tout $i \in [\![1;p]\!]$, f_i est différentiable en $a \in E$ et posons $\varphi_i = df_{i,a}$. Posons $\varphi : E \longrightarrow F^p$ l'application définie par

$$\forall h \in E,\ \varphi(h) = (\varphi_1(h), \dots, \varphi_p(h)).$$

On a

$$\begin{aligned} f(a+h) - f(h) - \varphi(h) &= (f_1(a+h) - f_1(a) - \varphi_1(h), \dots, f_p(a+h) - f_p(a) - \varphi_p(h)) \\ &= (\|h\|_E \, \varepsilon_1(h), \dots, \|h\|_E \, \varepsilon_p(h)), \end{aligned}$$

où $\varepsilon_1, \dots, \varepsilon_p$ sont des applications de E dans F qui tendent vers 0 en 0. On en déduit finalement que

$$f(a+h) - f(h) - \varphi(h) = \|h\|_E \underbrace{(\varepsilon_1(h), \dots, \varepsilon_p(h))}_{\xrightarrow[h\to 0]{} 0} = o(h).$$

f est donc bien différentiable et sa différentielle est égale à φ.

□

Proposition II.7.

Si $f : E \longrightarrow F$ est différentiable en $a \in E$, alors f est continue en a.

Démonstration. On a

$$f(x) = f(a) + \underbrace{df_a(x-a)}_{\xrightarrow[x\to a]{} 0} + \underbrace{o(x-a)}_{\xrightarrow[x\to a]{} 0}.$$

Bien entendu, $df_a(x-a) \xrightarrow[x\to a]{} 0$ car df_a est une application linéaire en dimension finie (voir le chapitre 11.6 sur les applications linéaires continues). □

Proposition II.8.

Soit $\Omega_1 \subset E$, $\Omega_2 \subset F$, et $\Omega_3 \subset G$ des ouverts. Si $f : \Omega_1 \longrightarrow \Omega_2$ est différentiable en $a \in \Omega_1$, et $g : \Omega_2 \longrightarrow \Omega_3$ est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et

$$d(g \circ f)_a = dg_{f(a)} \circ df_a.$$

Démonstration. On note $|||.|||$ la norme d'opérateur associée à $\|.\|_E$ (pour l'espace de départ) et à $\|.\|_F$ (pour l'espace d'arrivée). Soit $a \in \Omega_1$. On a

$$g(f(x)) - g(f(a)) \underset{(*)}{=} g(f(a)) + dg_{f(a)}(f(x) - f(a)) + o(f(x) - f(a)).$$

BY-NC-SA

L'égalité $(*)$ est vraie, car g est différentiable en $f(a)$ et $f(x) \xrightarrow[x\to a]{} f(a)$. On a de plus

$$f(x) - f(a) = df_a(x-a) + o(x-a),$$

et pour tout $h \in E$, $\|df_a(h)\|_F \leq |||df_a|||\, \|h\|_E$, et alors $df_a(h) \underset{h\to 0}{=} O(h)$, et de même $dg_{f(a)}(h) \underset{h\to 0}{=} O(h)$. On en déduit donc que

$$\begin{aligned} g(f(x)) - g(f(a)) &= g(f(a)) + dg_{f(a)}(df_a(x-a)) + dg_{f(a)}(o(x-a)) + o(df_a(x-a)) + o(o(x-a)) \\ &= g(f(a)) + dg_{f(a)} \circ df_a(x-a) + \underbrace{O(o(x-a)) + o(O(x-a)) + o(x-a)}_{o(x-a)} \\ &= g(f(a)) + dg_{f(a)} \circ df_a(x-a) + o(x-a). \end{aligned}$$

On a donc bien montré le résultat voulu.

□

Exemple. Dans le cas où $E = F = G = \mathbb{R}$, regardons ce que devient la relation démontrée ci-dessus. On a pour tout $(a,b) \in \Omega_1 \times \Omega_2$ et $h \in \mathbb{R}$, $df_a(h) = f'(a)h$ et $dg_b(h) = g'(b)h$. On en déduit donc que pour tout $h \in \mathbb{R}$,

$$d(g\circ f)_a(h) = dg_{f(a)} \circ df_a(h) = g'(f(a))f'(a)h, \ \textit{i.e.}\ (g\circ f)'(a) = g'(f(a))f'(a).$$

On retrouve bien donc la formule de la dérivée d'une composée de deux fonctions.

La propriété suivante est un cas particulier de la précédente, qui est assez utile en pratique, en particulier en physique.

Proposition (Règle de la chaine) II.9.

Soit I un intervalle ouvert de $\mathbb{R}$, $\gamma : I \longrightarrow E$ dérivable en $t_0 \in I$, et $f : \Omega \longrightarrow F$ différentiable en $\gamma(t_0)$ qu'on suppose dans Ω. $f\circ\gamma$ est dérivable en t_0, et

$$(f\circ\gamma)'(t_0) = df_{\gamma(t_0)}(\gamma'(t_0)).$$

Lorsque $F = \mathbb{R}$, cette égalité devient

$$(f\circ\gamma)'(t_0) = \left\langle \nabla f_{\gamma(t_0)},\, \gamma'(t_0) \right\rangle.$$

Remarque. La deuxième égalité ci-dessus est très similaire à la formule de la dérivée d'une composée de deux fonctions, sauf qu'on remplace le produit sur $\mathbb{R}$ par le produit scalaire sur E.

Démonstration. En utilisant la formule de la proposition II.8, on obtient pour tout $h \in \mathbb{R}$,

$$d(f\circ\gamma)_{t_0}(h) = df_{\gamma(t_0)}(d\gamma_{t_0}(h)) = df_{\gamma(t_0)}(\gamma'(t_0))h.$$

Ceci nous donne bien la formule voulue. La deuxième égalité est une conséquence directe de la première.

□

Exemples.

→ Lorsque $a, b \in \mathbb{R}$, $a \leq b$, et γ est définie sur $I =]0,1[$ par

$$\forall t \in]0,1[,\ \gamma(t) = tb + (1-t)a,$$

alors on a pour tout $t_0 \in]0,1[$

$$(f\circ\gamma)'(t_0) = df_{t_0b+(1-t_0)a}(b-a).$$

(cc) BY-NC-SA

→ On peut également retrouver une formule assez connue en physique de deuxième année. Soit Γ une courbe de $\mathbb{R}^3$ et soit $\gamma : [0,1] \longmapsto \Gamma$ un arc continu bijectif de classe $\mathcal{C}^1$ (il s'agit d'un arc continu parcourant Γ). Soit $\overrightarrow{E} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ un champ vectoriel. On définit, comme en physique, mais cette fois-ci formellement, l'intégrale de E sur l'arc Γ de la manière suivante.

$$\int_\Gamma \overrightarrow{E}(M)\overrightarrow{dM} = \int_0^1 \left\langle \overrightarrow{E}(\gamma(t)),\ \gamma'(t) \right\rangle dt.$$

Lorsqu'il existe un potentiel $V : \mathbb{R}^3 \longrightarrow \mathbb{R}$ tel que $\overrightarrow{E} = (x \longmapsto \nabla V_x)$, alors on a

$$\int_\Gamma \overrightarrow{E}(M)\overrightarrow{dM} = \int_0^1 \left\langle \nabla V_{\gamma(t)},\ \gamma'(t) \right\rangle dt \underset{\substack{\uparrow \\ \text{règle de la chaine}}}{=} \int_0^1 \frac{d}{dt} V(\gamma(t))dt = V(\gamma(1)) - V(\gamma(0)).$$

On retrouve alors l'égalité ci-dessus bien connue en physique.

Proposition II.10.

Lorsque $F = \mathbb{R}$, et f, g sont des applications de Ω dans $\mathbb{R}$, alors si f et g sont différentiables en $a \in \Omega$, $f \times g$ l'est aussi.

Démonstration. L'application

$$\Phi : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto xy \end{cases}$$

est différentiable en tout point de $\mathbb{R}^2$. De plus, si f, g sont différentiables en $a \in \Omega$, alors d'après la proposition II.6, $h : x \longmapsto (f(x), g(x))$ est aussi différentiable en a. Enfin, Φ est différentiable en $(f(a), g(a))$, donc d'après la proposition II.8, on peut affirmer que $\Phi \circ h = f \times g$ est différentiable en a. □

Exercice II.11.

Soit $f : \Omega \longrightarrow \mathbb{R}$ une application différentiable sur Ω.

1. Calculer la différentielle de f^2.
2. Supposons que $f(\Omega) \subset]0, +\infty[$. Calculer la différentielle de $\sqrt{f}$.

Exercice II.12.

Supposons que E est un espace euclidien, et que $\|.\|_E$ est la norme euclidienne associée à $\langle , \rangle$, et soit $b \in E$. Calculer la différentielle en tout point lorsqu'elle existe de l'application

$$d : \begin{cases} E & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto d(a,x) = \|x-b\|_E \,. \end{cases}$$

III Dérivées partielles

1. Généralités

On commence par introduire une notion assez utile.

(cc) BY-NC-SA

Définition III.1.

Soit $u \in E$ et $f : \Omega \longrightarrow F$ une application de E dans F. On appelle dérivée de f en $x \in \Omega$ selon la direction u, lorsqu'elle existe, la quantité définie par

$$D_u(f)(x) = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t}.$$

Une fois cette notion définie, on peut caractériser d'une manière assez élégante la différentielle d'une fonction.

Proposition III.2.

Soit $n = \dim E \in \mathbb{N}^*$, $(e_1, \ldots, e_n)$ une base de E, et $f : \Omega \longrightarrow F$ différentiable en $a \in \Omega$. Pour tout u, la dérivée de f en a selon u existe et pour tout $h = h_1 e_1 + \cdots + h_n e_n \in E$,

$$df_a(h) = \sum_{k=1}^{n} D_{e_k}(f)(a) h_k.$$

Lorsque $F = \mathbb{R}$ et $(e_1, \ldots, e_n)$ est une base orthonormée, l'égalité ci-dessus devient,

$$\nabla f_a = D_{e_1}(f)(a) e_1 + \cdots + D_{e_n}(f)(a) e_n. \tag{3}$$

Démonstration. Pour tout $u \in E$ et $a \in \Omega$, on a

$$f(a+tu) - f(a) = df_a(tu) + o(tu) = t \cdot df_a(u) + o(t).$$

Ceci implique que $\dfrac{f(a+tu) - f(a)}{t} \underset{t \to 0}{\longrightarrow} df_a(u)$, *i.e.* la dérivée directionnelle en a selon u de f existe et $D_u(f)(a) = df_a(u)$. Montrons maintenant la formule de la proposition. On a pour tout $h = h_1 e_1 + \cdots + h_n e_n \in E$,

$$df_a(h) = df_a\left(\sum_{k=1}^{n} h_k e_k\right) = \sum_{k=1}^{n} df_a(e_k) h_k = \sum_{k=1}^{n} D_{e_k}(f)(a) h_k.$$

Ceci est bien la formule voulue. La seconde inégalité est une conséquence directe de la première (le lecteur ayant un doute peut essayer de la montrer lui-même en utilisant les définitions). □

Remarques.

→ Lorsque $E = \mathbb{R}^n$ muni du **produit scalaire canonique** $((x_1, \ldots, x_n), (y_1, \ldots, y_n)) \longmapsto \sum_{k=1}^{n} x_i y_i$, et $(e_1, \ldots, e_n)$ est la base canonique de $\mathbb{R}^n$, alors en notant pour tout $i \in [\![1; n]\!]$, $x = (x_1, \ldots, x_n) \in E$,

$$\frac{\partial f}{\partial x_i}(x) = D_{e_i}(f)(x) = \lim_{t \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{t},$$

alors, on peut réécrire l'égalité 3 pour tout $a \in \Omega$ où a est différentiable de la manière suivante

$$\nabla f_a = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}. \tag{4}$$

(cc) BY-NC-SA

Pour tout $i \in [\![1; n]\!]$, on appelle $\dfrac{\partial f}{\partial x_i}(a)$ la dérivée partielle de f par rapport à la $i-$ème variable en a. En fait, l'égalité 4 nous permet de calculer la différentielle (ou le gradient) de fonctions à valeurs dans $\mathbb{R}$ assez facilement en pratique. En pratique, pour $i \in [\![1; n]\!]$, pour calculer $\dfrac{\partial f}{\partial x_i}$, il suffit de considérer toutes les variables autres que x_i comme constantes et de dériver uniquement par rapport à x_i.

→ Attention, la proposition III.2 ferait croire que la réciproque est vraie. La réciproque est en fait fausse : admettre une dérivée directionnelle, même selon toutes les directions d'une base, n'implique pas être différentiable. En effet, on peut le voir via le contrexemple suivant. Considérons l'application

$$g : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto \begin{cases} \frac{x^2 y}{x^4+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon.} \end{cases} \end{cases}$$

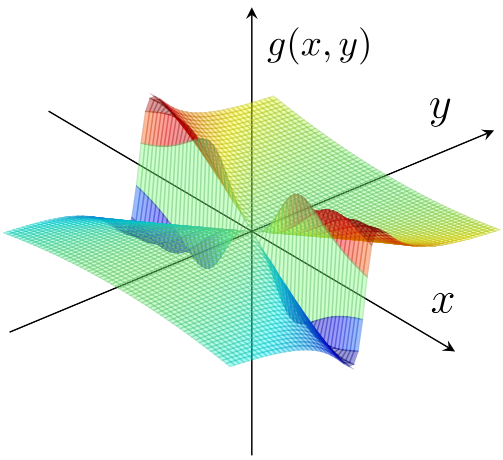
On remarque qu'en $(0, 0)$, g admet des dérivées directionnelles selon $(1, 0)$ et $(0, 1)$. En effet, on a

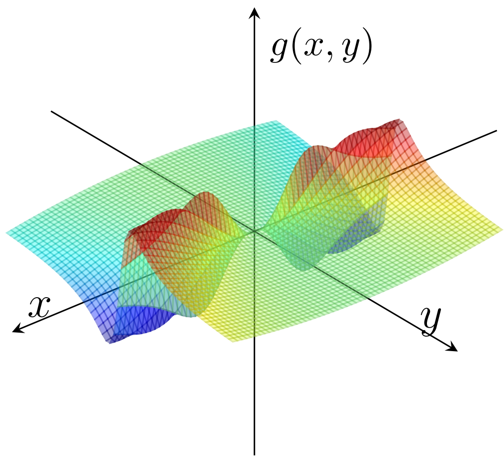
$$\frac{g(t, 0) - g(0, 0)}{t} = 0 \underset{t \to 0}{\longrightarrow} 0 \text{ et } \frac{g(0, t) - g(0, 0)}{t} = 0 \underset{t \to 0}{\longrightarrow} 0.$$

Cependant, g n'est pas différentiable en 0. En effet, en supposant que g est différentiable en 0 et en utilisant ce qui précède et la proposition 3, on peut affirmer que $dg_{(0,0)} = 0$. Ceci implique que la dérivée en $(0, 0)$ selon n'importe quelle direction doit être nulle, mais pour une direction $u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{R}^2$ telle que $\beta \neq 0$, on a

$$\frac{g(tu) - g(0, 0)}{t} = \frac{t^3\alpha^2\beta/(t^4\alpha^4 + t^2\beta^2)}{t} = \frac{\alpha^2\beta}{t^2\alpha^4 + \beta^2} \underset{t \to 0}{\longrightarrow} \underbrace{\frac{\alpha^2}{\beta}}_{\neq 0 \text{ si } \alpha \neq 0},$$

ce qui est absurde, donc g n'est pas différentiable en $(0, 0)$.





On observe sur le dessin ci-dessus qu'en 0, il y a des oscillations qui font que g n'est pas "lisse" et donc ne peut pas être différentiable en 0. Il s'agit donc d'un autre cas de figure que le point anguleux en 0 qu'on a observé auparavant pour les normes. Remarquons que dans des directions particulières au voisinage de 0, ces oscillations divergent vers l'infini, mais on ne peut pas l'observer sur le dessin, car on aurait besoin d'une précision beaucoup plus grande.

→ En réexaminant la définition II.2 et la proposition III.2, on a jusqu'à présent deux manières de montrer qu'une application $f : \Omega \longmapsto F$ est différentiable et deux moyens pratiques de calculer la différentielle d'une application en un point $a \in \Omega$.

(cc) BY-NC-SA

- Si f est "simple", on peut potentiellement montrer grâce aux propositions II.6, II.8, II.5, II.10 que f est différentiable. Par exemple, la fonction $M \longmapsto \det M$ est somme et produit d'applications identiques à celles de la question 1 de l'exercice II.11, et est donc différentiable. On pourra calculer la différentielle de cette fonction dans un exercice ultérieur (exercice III.9)

- Si on sait déjà que f est différentiable en a (ou on l'a montré via le point précédent), pour calculer la différentielle de f, il suffit de calculer les dérivées selon les vecteurs d'une base convenable de E (en général, on calculera plutôt les dérivées partielles dans le cas où $E = \mathbb{R}^n$ avec $n \in \mathbb{N}^*$ muni de son produit scalaire canonique ou est un espace euclidien). Ensuite, il suffit d'écrire pour tout $h = h_1 e_1 + \cdots + h_n e_n \in E$,

$$df_a(h) = \sum_{k=1}^{n} D_{e_k}(f)(a) h_k,$$

et dans le cas plus fréquent où $E = \mathbb{R}^n$, pour tout $h = (h_1, \dots, h_n) \in \mathbb{R}^n$,

$$df_a(h) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) h_k.$$

Si de plus $F = \mathbb{R}$, alors $df_a(h) = \langle \nabla f_a,\, h \rangle$ où ∇f_a est donné par l'égalité 4.

- Si on ne sait pas si f est différentiable à priori, on calcule $f(a+h) - f(a)$, et on essaye d'écrire cette différence conformément à la définition II.2. Une fois cela fait, on aura montré que f est différentiable en a, et on aura potentiellement aussi calculé la différentielle de f en a en même temps.

→ Le premier point de cette remarque nous montre également qu'un moyen possible pour démontrer que f n'est pas différentiable est de démontrer que f n'admet pas une dérivée selon une certaine direction $u \in E$, ou alors montrer par l'absurde que si c'était le cas, les valeurs de la différentielle et des dérivées directionnelles ne sont pas cohérentes entre elles.

Exercice III.3.

Soit $n \in \mathbb{N}$. On munit $\mathcal{M}_n(\mathbb{R})$ du produit scalaire $\langle ., . \rangle : (A, B) \longmapsto \mathrm{Tr}(AB^T)$ et de la norme $\|.\| : A \longmapsto \mathrm{Tr}(AA^T)$. Montrer que les applications suivantes sont différentiables (sauf potentiellement en un point) et calculer leurs différentielles.

1. $h_1 : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R}^+ \\ (x_1, \dots, x_n) & \longmapsto \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2}, \end{cases}$ où $a_1, \dots, a_n \in \mathbb{R}$.
2. $h_2 : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow \mathbb{R} \\ M & \longmapsto \det M. \end{cases}$
3. $h_3 : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow \mathcal{M}_n(\mathbb{R}) \\ M & \longmapsto MAM \end{cases}$ où $A \in \mathcal{M}_n(\mathbb{R})$.

BY-NC-SA

Exercice III.4.

Soit $k \in \mathbb{N}^*$ et f une fonction de E^k dans F. On dit que f est $k-$linéaire lorsque pour tout $x_1, \dots, x_k \in E$ et tout $i \in [\![1;k]\!]$, l'application $x \longmapsto f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_k)$ est linéaire.

1. Montrer que f est continue sur E.
2. Montrer qu'il existe $C \geq 0$ tel que pour tout $x_1, \dots, x_k \in E$,
$$\|f(x_1, \dots, x_n)\|_F \leq C \|x_1\|_E \times \dots \times \|x_k\|_E .$$
3. Calculer la différentielle de l'application
$$g : \begin{cases} E & \longrightarrow F \\ x & \longmapsto f(x, \dots, x). \end{cases}$$
On dit que f est symétrique lorsque pour tout $x_1, \dots, x_k \in E$ et toute permutation $\sigma \in \mathcal{S}_k$, $f(x_1, \dots, x_k) = f(x_{\sigma(1)}, \dots, x_{\sigma(k)})$. Que devient cette différentielle de g lorsque f est symétrique ?

2. Fonctions de classe $\mathcal{C}^1$

Dans cette sous-partie, on suppose que $E = \mathbb{R}^n$ avec $n \in \mathbb{N}^*$ muni de son produit scalaire canonique et Ω est un ouvert de E.

Définition III.5.

Soit f une application de $\mathbb{R}^n$ dans F. On dit que f est de classe $\mathcal{C}^1$ lorsque f admet des dérivées partielles $\dfrac{\partial f}{\partial x_i}$ pour tout $i \in [\![1;n]\!]$ (par rapport à toutes ses variables), et que pour tout $i \in [\![1;n]\!]$, $x \longmapsto \dfrac{\partial f}{\partial x_i}(x)$ est continue sur Ω.

Exemple. Les fonctions polynomiales, i.e. les fonctions de l'ensemble

$$\mathbb{R}[x_1, \dots, x_n] = \left\{ (x_1, \dots, x_n) \longmapsto \sum_{\alpha=(\alpha_1,\dots,\alpha_n)\in A} a_\alpha x_1^{\alpha_1} \dots x_n^{\alpha_n},\ A \subset \mathbb{N}^n \text{ fini}, (a_\alpha)_{\alpha\in A} \in \mathbb{R}^A \right\}.$$

sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$.

Proposition III.6.

Si $f : \Omega \longrightarrow F$ est de classe $\mathcal{C}^1$, alors elle est différentiable en tout point de Ω.

Démonstration. Soit $a \in \Omega$. Si f est différentiable en a, alors d'après la proposition III.2, sa différentielle en a doit être égale à la fonction φ définie par

$$\forall h = (h_1, \dots, h_n) \in \mathbb{R}^n,\ \varphi(h) = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k}(a).$$

On va donc essayer de montrer que $f(a+h) - f(a) - \varphi(h) \underset{h\to 0}{=} o(h)$. On sait que les dérivées partielles sont bien définies et continues, on a donc des informations utiles sur les variations du type $f(a + te_i) - f(a)$

(cc) BY-NC-SA

(approximativement égales, à l'ordre 1, à $t\dfrac{\partial f}{\partial x_i}(a)$) où e_i est le $i-$ème vecteur de la base canonique de $\mathbb{R}^n$ et t un nombre réel assez petit en valeur absolue. On va donc essayer de décomposer la variation $f(a+h)-f(a)$ en somme de variations égales à $f(b+te_i)-f(b)$ où b est proche de a.
Soit $\varepsilon>0$. Par continuité des dérivées partielles et le fait que Ω est ouvert, on peut affirmer qu'il existe $\eta>0$ tel que pour tout $y=(y_1,\dots,y_n)\in\mathbb{R}^n$,

$$\max_{i\in[\![1;n]\!]}|a_i-y_i|<\eta \implies y\in\Omega \text{ et } \max_{i\in[\![1;n]\!]}\left(\left\|\frac{\partial f}{\partial x_i}(a)-\frac{\partial f}{\partial x_i}(y)\right\|_F\right)<\varepsilon. \tag{5}$$

Posons pour tout $h\in\mathbb{R}^n$ tel que $\|h\|_\infty<\eta$ (cette condition sur h nous donne en particulier que $a+h\in\Omega$),

$$\Delta(h)=f(a+h)-f(a)-\varphi(h).$$

On cherche à montrer que $\Delta(h)=o(h)$. On écrit

$$\begin{aligned}\Delta(h)&=\overbrace{f(a_1+h_1,a_2+h_2,\dots,a_n+h_n)-f(a_1,a_2+h_2,\dots,a_n+h_n)-h_1\frac{\partial f}{\partial x_1}(a)}^{\Delta_1(h)}\\
&\quad+\underbrace{f(a_1,a_2+h_2,\dots,a_n+h_n)-f(a_1,a_2,a_3+h_3,\dots,a_n+h_n)-h_2\frac{\partial f}{\partial x_2}(a)}_{\Delta_2(h)}+\dots\\
&\quad+\underbrace{f(a_1,a_2,\dots,a_{n-1},a_n+h_n)-f(a_1,a_2,\dots,a_n)-h_n\frac{\partial f}{\partial x_n}(a)}_{\Delta_n(h)}\\
&=\sum_{k=1}^n\Delta_k(h).\end{aligned}$$

Notre but maintenant est de trouver une bonne majoration pour $\Delta_k(h)$ pour tout $k\in[\![1;n]\!]$. On écrit pour tout $k\in[\![1;n]\!]$,

$$\Delta_k(h)=\underbrace{f(a_1,\dots,a_{k-1},a_k+h_k,\dots,a_n+h_n)-f(a_1,\dots,a_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)}_{\delta_k(h)}-h_k\frac{\partial f}{\partial x_k}(a).$$

On voit que $\delta_k(h)$ ressemble fortement à une variation du type $f(b+te_i)-f(b)$ où b est proche de a, on veut donc trouver une relation entre cette quantité et la dérivée partielle. Un outil assez pratique pour le faire (plutôt de le faire directement, ce qui peut être assez pénible) est l'inégalité des accroissements finis. En effet, en posant pour tout $t\in[0,1]$

$$\Phi(t)=f(a_1,\dots,a_{k-1},a_k+th_k,\dots,a_n+h_n)-f(a_1,\dots,a_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)-th_k\frac{\partial f}{\partial x_k}(a),$$

qui est une fonction continue sur $[0,1]$ et dérivable sur $]0,1[$, on peut dire que

$$\begin{aligned}\forall c\in]0,1[,\ \|\Phi'(c)\|_F&\underset{\substack{\uparrow\\ \text{règle de la chaine}}}{=}\left\|h_k\left(\frac{\partial f}{\partial x_k}(a_1,\dots,a_k+ch_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)-\frac{\partial f}{\partial x_k}(a)\right)\right\|_F\\
&=|h_k|\left\|\frac{\partial f}{\partial x_k}\underbrace{(a_1,\dots,a_k+ch_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)}_{\alpha_{k,c}}-\frac{\partial f}{\partial x_k}(a)\right\|_F\\
&\underset{(*)}{\leq}|h_k|\,\varepsilon.\end{aligned}$$

(cc) BY-NC-SA

L'inégalité $(*)$ est vraie car $\|a-\alpha_{k,c}\|_\infty < \eta$ (on peut facilement le voir en utilisant le fait que $\|h\|_\infty < \eta$, $c \in]0,1[$ et en utilisant l'implication 5).
Maintenant, on peut appliquer l'inégalité des accroissements finis (voir le chapitre sur les fonctions à valeurs vectorielles) sur l'intervalle $[0,1]$, ce qui nous permet de dire que

$$\|\Delta_k(h)\|_F = \|\Phi(1)-\Phi(0)\|_F \leq |h_k|\,\varepsilon.$$

On en déduit donc que

$$\|\Delta(h)\|_F \leq \sum_{k=1}^{n} \|\Delta_k(h)\|_F \leq \varepsilon \sum_{k=1}^{n} |h_k| \leq n\varepsilon \|h\|_\infty .$$

Ceci nous permet de dire finalement que $\Delta(h) = o(h)$, ce qui est bien le résultat voulu : f est différentiable en a et sa différentielle est égale à φ.

$\square$

Notation. On note $\mathcal{C}^1(\Omega, F)$ l'ensemble des applications de Ω dans F de classe $\mathcal{C}^1$.

Proposition III.7.

Soit f une application de Ω dans F. Les propositions suivantes sont équivalentes.

1. f est de classe $\mathcal{C}^1$ sur Ω.
2. f est différentiable sur Ω et l'application $\psi : \begin{cases} \Omega & \longrightarrow \mathcal{L}(E,F) \\ a & \longmapsto df_a \end{cases}$ est continue.

Démonstration. On note $|||.|||$ la norme d'opérateur associée à $\|.\|_\infty$ (pour l'espace de départ) et $\|.\|_F$ (pour l'espace d'arrivée).

→ (1) ⇒ (2) D'après la proposition III.6, f est différentiable, car elle est de classe $\mathcal{C}^1$. De plus, pour tout $a, b \in \Omega$ et $h \in E$, on a

$$\|(df_a - df_b)(h)\|_F = \left\| \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(a) - \frac{\partial f}{\partial x_k}(b) \right) h_k \right\|_F \leq \left(\sum_{k=1}^{n} \left\| \frac{\partial f}{\partial x_k}(a) - \frac{\partial f}{\partial x_k}(b) \right\|_F \right) \|h\|_\infty .$$

On en déduit donc par définition de la norme d'opérateur que

$$|||df_a - df_b||| \leq \sum_{k=1}^{n} \left\| \frac{\partial f}{\partial x_k}(a) - \frac{\partial f}{\partial x_k}(b) \right\|_F \xrightarrow[b \to a]{} 0.$$

On a donc bien montré la continuité de ψ.

→ (2) ⇒ (1) f est différentiable en tout point de Ω et donc admet des dérivées partielles en tout point de Ω. On a alors pour tout $a, b \in \Omega$, en notant pour $i \in [\![1;n]\!]$, e_i le $i-$ème vecteur de la base canonique de $\mathbb{R}^n$,

$$\left\| \frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right\|_F = \|df_a(e_i) - df_b(e_i)\|_F \leq |||df_a - df_b|||\, \|e_i\|_\infty = |||df_a - df_b||| \xrightarrow[b \to a]{} 0.$$

f est donc bien de classe $\mathcal{C}^1$ sur Ω.

$\square$

On introduit ensuite quelques propriétés générales utiles des fonctions de classe $\mathcal{C}^1$.

(cc) BY-NC-SA

Proposition III.8.

Les propositions suivantes sont vraies.

1. $\mathcal{C}^1(\Omega, F)$ est un $\mathbb{R}-$espace vectoriel.
2. $(\mathcal{C}^1(\Omega, F), +, \times)$ est une $\mathbb{R}-$algèbre.

Démonstration. La preuve de ces deux propositions ne présente pas de difficulté et est laissée comme exercice au lecteur. $\square$

Exemples.

→ Si $f : \Omega \longrightarrow \mathbb{R}$ ne s'annule pas sur Ω et est de classe $\mathcal{C}^1$, alors $\dfrac{1}{f}$ est de classe $\mathcal{C}^1$.

→ Si $f : \Omega \longrightarrow \mathbb{R}$ est strictement positive sur Ω et est de classe $\mathcal{C}^1$, alors $\sqrt{f}$ est de classe $\mathcal{C}^1$.

Exercice III.9.

Soit $\sum a_n z^n$ une série entière complexe de rayon de convergence égal à $R > 0$. Montrer que $(x, y) \longmapsto \sum_{k=0}^{\infty} a_k (x + iy)^k$ est de classe $\mathcal{C}^1$ sur

$$B_{\mathbb{R}^2}(0, R) = \{(x, y) \in \mathbb{R}^2,\ \|(x, y)\|_2 = \sqrt{x^2 + y^2} < R\}.$$

3. Matrices jacobiennes

Dans cette sous-partie, on suppose que $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ où $n, p \in \mathbb{N}^*$ et sont munis de leurs produits scalaires canoniques. On suppose aussi encore une fois que Ω est un ouvert de E.

Définition III.10.

Soit $a \in \Omega$ et $f : \Omega \longrightarrow \mathbb{R}^p$ différentiable en a. On appelle matrice jacobienne de f en a, la matrice de l'application linéaire df_a dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$. On note cette matrice $J_f(a)$.

Proposition III.11.

Si $f : \Omega \longrightarrow \mathbb{R}^p$ est différentiable en $a \in \Omega$, et pour tout $x \in \Omega$, $f(x) = (f_1(x), \dots, f_p(x))$, alors on a

$$J_f(a) = \begin{pmatrix} \nabla f_{1,a}^T \\ \vdots \\ \nabla f_{p,a}^T \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \dots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R}).$$

(cc) BY-NC-SA

Démonstration. Soit $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. On a pour tout $i \in [\![1; n]\!]$,

$$df_a(e_i) = \begin{pmatrix} df_{1,a}(e_i) \\ \vdots \\ df_{p,a}(e_i) \end{pmatrix} = \begin{pmatrix} \langle \nabla f_{1,a},\ e_i \rangle \\ \vdots \\ \langle \nabla f_{p,a},\ e_i \rangle \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_i}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_i}(a) \end{pmatrix}.$$

On a donc bien retrouvé l'image de la base canonique de $\mathbb{R}^n$ dans la base canonique de $\mathbb{R}^p$ par df_a, ce qui nous permet d'écrire la matrice complète dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.

$$J_f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \ldots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \ldots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \nabla f_{1,a}^T \\ \vdots \\ \nabla f_{p,a}^T \end{pmatrix}.$$

$\square$

Remarques.

→ Bien évidemment, on déduit de ce qui précède que si $f : \Omega \longrightarrow \mathbb{R}^p$ est différentiable en $a \in \Omega \subset \mathbb{R}^n$, alors pour toute matrice colonne $h \in \mathbb{R}^n$,

$$f(a+h) - f(a) = J_f(a) \times h + o(h).$$

→ On déduit également de ce qui précède que pour tout $f : \Omega \longrightarrow \mathbb{R}^p$ telle que pour tout $x \in \Omega$, $f(x) = (f_1(x), \ldots, f_p(x))$ où $f_1, \ldots, f_p$ sont des applications de Ω dans $\mathbb{R}$, et $a \in \Omega$ où f est différentiable, on a pour tous $(i,j) \in [\![1; p]\!] \times [\![1; n]\!]$,

$$[J_f(a)]_{i,j} = \frac{\partial f_i}{\partial x_j}(a),$$

et par conséquent, si $g = (g_1, \ldots, g_m)$ est une application de $\mathbb{R}^p$ dans $\mathbb{R}^m$ différentiable en $f(a)$ avec $m \in \mathbb{N}^*$, on a en utilisant la proposition II.8,

$$J_{g \circ f}(a) = J_g(f(a)) \times J_f(a)$$

et donc en posant pour tous $x \in \mathbb{R}^n$,

$$g \circ f(x) = (h_1(x), \ldots, h_m(x)) = (g_1(f(x)), \ldots, g_m(f(x))),$$

on a pour tous $(i,j) \in [\![1; m]\!] \times [\![1; n]\!]$

$$\begin{aligned}
\frac{\partial h_i}{\partial x_j}(a) &= \frac{\partial (g_i \circ f)}{\partial x_j}(a) = [J_g(f(a)) \times J_f(a)]_{i,j} \\
&= \left[\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}(f(a)) & \ldots & \dfrac{\partial g_1}{\partial x_p}(f(a)) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(f(a)) & \ldots & \dfrac{\partial g_m}{\partial x_p}(f(a)) \end{pmatrix} \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(f(a)) & \ldots & \dfrac{\partial f_1}{\partial x_n}(f(a)) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(f(a)) & \ldots & \dfrac{\partial f_p}{\partial x_n}(f(a)) \end{pmatrix} \right]_{i,j} \\
&= \sum_{k=1}^{p} \frac{\partial g_i}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).
\end{aligned}$$

On aurait aussi pu montrer cette égalité en utilisant la règle de la chaine. En effet, l'application

(cc) BY-NC-SA

$\varphi : x_k \longmapsto f(a_1, \ldots, a_{k-1}, x_k, a_k, \ldots, a_n)$ est dérivable en a_k, et

$$\varphi'(a_k) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_j}(a) \end{pmatrix}.$$

On a alors en utilisant la règle de la chaine,

$$\frac{\partial h_i}{\partial x_j}(a) = (g_i \circ \varphi)'(a_k) = \left\langle \nabla g_{i,\varphi(a_k)},\ \varphi'(a_k) \right\rangle = \left\langle \nabla g_{i,f(a)},\ \varphi'(a_k) \right\rangle = \left\langle \begin{pmatrix} \dfrac{\partial g_i}{\partial x_1}(f(a)) \\ \vdots \\ \dfrac{\partial g_i}{\partial x_p}(f(a)) \end{pmatrix}, \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_j}(a) \end{pmatrix} \right\rangle$$

$$= \sum_{k=1}^{p} \frac{\partial g_i}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

→ En physique, on utilise quelques opérateurs différentiels dans quelques situations, par exemple en électromagnétique, en particulier la divergence et le laplacien. En fait, on peut définir ces opérateurs avec ce que l'on vient de voir. En effet, lorsque $p = n$, si f est une application de Ω dans $\mathbb{R}^n$ différentiable en $a \in \Omega$, alors

$$(\mathrm{div} f)(a) = \mathrm{Tr}(J_f(a)) = \sum_{k=1}^{n} \frac{\partial f_k}{\partial x_k}(a),$$

et lorsque f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ admet des dérivées partielles secondes en a (dérivées deux fois, notées $\dfrac{\partial^2 f}{\partial x_i^2} = \left(\dfrac{\partial}{\partial x_i}\dfrac{\partial f}{\partial x_i}\right)(a)$ pour tout $i \in [\![1; n]\!]$), on définit le laplacien Δf de la manière suivante

$$\Delta f(a) = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k^2}(a),$$

et en notant

$$\psi : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R}^n \\ x & \longmapsto \nabla f_x, \end{cases}$$

on obtient la formule connue suivante

$$(\mathrm{div}\ \nabla f)(a) = (\mathrm{div}\ \psi)(a) = \Delta f.$$

IV Accroissements finis

Dans cette partie, on ne suppose plus que $E = \mathbb{R}^n$, mais juste que E est un espace vectoriel normé de dimension finie, et on considère Ω un ouvert de E.

1. Théorème des accroissements finis

On énonce maintenant un équivalent du théorème des accroissements finis pour les fonctions d'un $\mathbb{R}-$espace vectoriel normé dans un autre, assez utile en exercices.

BY-NC-SA

Théorème (Accroissements finis) IV.1.

Soit $f \in \mathcal{C}^1(\Omega, F)$ et $a, b \in \Omega$. Posons $[a,b] = \{(1-t)a + tb,\ t \in [0,1]\}$, et supposons que $[a,b] \subset \Omega$. Les propositions suivantes sont vraies.

1. $f(b) - f(a) = \displaystyle\int_0^1 df_{(1-t)a+tb}(b-a)dt$.
2. Lorsque $F = \mathbb{R}$, $f(b) - f(a) = \displaystyle\int_0^1 \left\langle \nabla f_{(1-t)a+tb},\ b-a \right\rangle dt$.
3. $\|f(b) - f(a)\|_F \leq \displaystyle\sup_{t\in[0,1]} |||df_{(1-t)a+tb}||| \times \|b-a\|_E$.

Démonstration.

1. Considérons l'application

$$\Phi : \begin{cases} [0,1] & \longrightarrow F \\ t & \longmapsto f((1-t)a+tb). \end{cases}$$

Φ est une composition de deux fonctions différentiables et est donc différentiable sur $]0,1[$, et est donc dérivable. De plus, étant donné que $\Phi \in \mathcal{C}^1([0,1], F)$ et donc que Φ' est continue, on a pour tous $a, b \in \Omega$,

$$f(b) - f(a) = \Phi(1) - \Phi(0) = \int_0^1 \Phi'(t)dt \underset{\substack{\uparrow \\ \text{règle de la chaine}}}{=} \int_0^1 df_{(1-t)a+tb}(b-a)dt.$$

2. Cette égalité est une conséquence directe de l'égalité précédente. En effet, lorsque $F = \mathbb{R}$, on a pour tous $a, b \in \Omega$

$$df_{(1-t)a+tb}(b-a) = \left\langle \nabla f_{(1-t)a+bt},\ b-a \right\rangle_E,$$

ce qui donne bien l'égalité voulue.

3. Cette inégalité est une conséquence directe de l'égalité (1). En effet, on a pour tous $a, b \in \Omega$

$$\begin{aligned} \|f(b) - f(a)\|_F = \left\| \int_0^1 df_{(1-t)a+tb}(b-a)dt \right\|_F &\leq \sup_{t\in[0,1]} \left\| df_{(1-t)a+tb}(b-a) \right\|_F \\ &\leq \sup_{t\in[0,1]} |||df_{(1-t)a+tb}||| \times \|b-a\|_E. \end{aligned}$$

$\square$

Conséquences.

→ Si $f \in \mathcal{C}^1$, alors f est localement lipschitzienne. En effet, considérons $a \in \Omega$ et $\eta > 0$ tel que $B_f(a,\eta) \subset \Omega$. $f \in \mathcal{C}^1$, donc d'après la proposition III.7, l'application

$$\varphi : \begin{cases} B_f(a,\eta) & \longrightarrow \mathcal{L}(E,F) \\ x & \longmapsto df_x \end{cases}$$

est continue sur $B_f(a,\eta)$. De plus, E est de dimension finie et $B_f(a,\eta)$ est fermé borné, et est donc compact. On en déduit que φ est bornée sur $B_f(a,\eta)$ par un réel qu'on note K. On a alors pour

(cc) BY-NC-SA

tous $x, y \in B(a, \eta)$, en utilisant la proposition précédente,

$$\|f(x) - f(y)\|_F \leq \sup_{t\in[0,1]} |||df_{(1-t)x+ty}||| \times \|x - y\|_E \leq K \|x - y\|_E .$$

→ Si pour tout $x \in \Omega$, $df_x = 0$ et Ω est connexe, alors f est constante sur Ω. Montrons ce résultat. Considérons $a \in \Omega$ et $\eta > 0$ tel que $B(a, \eta) \subset \Omega$. D'après L'égalité (1) de la proposition précédente, on a pour tout $x \in B(a, \eta)$,

$$f(x) - f(a) = \int_0^1 df_{(1-t)a+tx}(x - a)dt = 0.$$

On en déduit donc que pour tout $x \in B(a, \eta)$, $f(x) = f(a)$. Posons maintenant

$$A = \{x \in \Omega,\ f(x) = f(a)\} = f^{-1}(\{f(a)\}).$$

Cet ensemble est fermé, car il s'agit de l'image réciproque d'un ensemble fermé de Ω par une fonction continue. De plus, le raisonnement précédent nous permet aussi d'affirmer que A est ouvert (car tout point de cet ensemble peut être entouré d'une boule ouverte incluse dans A). finalement $a \in A$, donc A est non vide, et alors par connexité de Ω, $A = \Omega$. Au passage, on aurait également pu montrer ce résultat si on avait supposé que Ω est connexe par arcs (qui est une propriété plus forte que la connexité) sans passer par la connexité. Pour ce faire, il suffit de considérer deux points x, y de Ω et considérer un arc γ de classe $\mathcal{C}^1$ par morceaux (c'est-à-dire qu'on peut subdiviser $[0, 1]$ en un nombre fini d'intervalles disjoints où γ est de classe $\mathcal{C}^1$) les liant (en fait, la définition de connexité par arcs donne uniquement que γ est continu, mais on peut montrer que dans tout $\mathbb{R}-$espace vectoriel, la connexité pour un ouvert implique la connexité par arcs de classe $\mathcal{C}^1$. Nous montrons ce résultat en annexe pour le lecteur curieux). On peut montrer facilement que $f \circ \gamma$ est constante sur son support, car sa dérivée est nulle et que finalement $f(b) = f(a)$.

→ Si Ω est connexe et qu'il existe $u \in \mathcal{L}(E, F)$ telle que pour tout $x \in \Omega$, $df_x = u$, alors il existe une constante $C \in F$ telle que pour tout $x \in \Omega$, $f(x) = u(x) + C$. En effet, pour voir que c'est vrai, il suffit d'appliquer le raisonnement précédent pour $f - u$ qui a une différentielle nulle sur tout Ω (pour tout x, $d(f - u)_x = 0$). On en déduit donc qu'il existe une constante $C \in F$ telle que $f(x) - u(x) = C$ pour tout x, ce qui donne bien l'égalité voulue.

Exercice IV.2.

On suppose que Ω est convexe et on considère $f \in \mathcal{C}^1(\Omega, \mathbb{R})$. Montrer que pour tous $a, b \in \Omega$

$$\exists \theta \in [0, 1],\ f(b) - f(a) = \left\langle \nabla f_{(1-\theta)a+\theta b},\ b - a \right\rangle_E .$$

Exercice IV.3.

Soient $m \in \mathbb{N}^*$ et $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}_+^*$. Montrer que

$$\left(\prod_{k=1}^m a_k\right)^{\frac{1}{m}} + \left(\prod_{k=1}^m b_k\right)^{\frac{1}{m}} \leq \left(\prod_{k=1}^m (a_k + b_k)\right)^{\frac{1}{m}} .$$

2. Inversion locale

Cette partie est hors programme, mais potentiellement utile aux oraux X-ENS. Par souci de simplicité, on suppose dans cette sous-partie que E et F sont égaux à $\mathbb{R}^n$ et munis d'une même norme $\|.\|$.

CC BY-NC-SA

Définition IV.4.

Soit Ω, Ω' deux ouverts de $E = \mathbb{R}^n$. Une fonction $f : \Omega \longrightarrow \Omega'$ est un $\mathcal{C}^1-$difféomorphisme si f est de classe $\mathcal{C}^1$, bijective, et f^{-1} est de classe $\mathcal{C}^1$ sur Ω'.

Remarque. Lorsque f est un $\mathcal{C}^1-$difféomorphisme comme définit dans la définition ci-dessus, on a $f^{-1} \circ f = \mathrm{Id}$ et donc en appliquant la formule de la différentielle pour une composée de deux fonctions (proposition II.8), on obtient pour tout $a \in \Omega$

$$df^{-1}_{f(a)} \circ df_a = \mathrm{Id}, \ \ i.e. \ J_{f^{-1}}(f(a)) \times J_f(a) = I$$

On remarquera en particulier que la différentielle en a (resp. la matrice jacobienne) de f en tout point $a \in \Omega$ est inversible, d'inverse $df^{-1}_{f(a)}$ (resp. $J_{f^{-1}}(f(a))$).

Exemple. Posons $\Omega = \mathbb{R}_+^* \times]-\pi, \pi[$ et $\Omega' = \mathbb{R}^2 \setminus (\mathbb{R}^- \times \{0\})$ et considérons l'application

$$\varphi : \begin{cases} \Omega & \longrightarrow \Omega' \\ (r, \theta) & \longmapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \end{cases}. \tag{6}$$

φ est bijective de classe $\mathcal{C}^1$ sur Ω et $\forall (x,y) \in \Omega'$, $\varphi^{-1}(x,y) = \left(\sqrt{x^2+y^2}, 2\arctan\left(\dfrac{y}{x+1}\right)\right)$. φ^{-1} est également de classe $\mathcal{C}^1$ sur Ω', donc φ est bien in difféomorphisme de Ω dans Ω'.

Le théorème hors programme suivant peut être également vu comme un exercice. Il arrive que quelque chose de similaire tombe aux oraux X-ENS.

Théorème (Théorème d'inversion locale) IV.5.

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$. Soit $a \in \Omega$ tel que $\det J_f(a) \neq 0$. Il existe $U \subset \Omega$ un voisinage ouvert de a et $V \subset \mathbb{R}^n$ un voisinage ouvert de $f(a)$ tels que

$$\tilde{f} : \begin{cases} U & \longrightarrow V \\ x & \longmapsto f(x) \end{cases}$$

est un $\mathcal{C}^1-$difféomorphisme.

Démonstration. On commence par faire les réductions suivantes du problème.

→ Quitte à remplacer f par la fonction $x \longmapsto f(x+a)$, on peut supposer que $a = 0$.

→ Quitte à remplacer f par $x \longmapsto f(x) - f(0)$, on suppose que $f(0) = 0$.

→ Enfin, en posant $u = df_0$, quitte à remplacer f par $x \longmapsto u^{-1} \circ f$, on peut supposer que $df_0 = \mathrm{Id}$.

Le lecteur ayant un doute peut montrer lui-même que ces simplifications ne changent pas le résultat, en particulier, l'image d'un voisinage ouvert d'un point $b \in \mathbb{R}^n$ par une application linéaire inversible $v \in \mathcal{L}(\mathbb{R}^n)$ est un voisinage ouvert de $v(b)$.

Citons les étapes clés de la preuve.

1. Dans un premier temps, on montre qu'il existe un voisinage U de 0 dans lequel f est injective.
2. Ensuite, en posant $V = f(U)$, il est clair que $\tilde{f} : U \longrightarrow V$ est bijective. Il faut alors montrer que V est un voisinage de $f(0) = 0$, et qu'on peut supposer que V est ouvert, quitte à changer U par un sous-ensemble ouvert de U.
3. Enfin, il reste à montrer que $\tilde{f}^{-1} \in \mathcal{C}^1(V, U)$.

(cc) BY-NC-SA

On commence par la première étape. Inspirons-nous du cas où $n = 1$. Dans ce cas, $f'(0) = 1$. Si on se restreint à un voisinage convexe de 0 où $f'(x) > 0$ (par exemple, voisinage assez petit pour qu'on ait $|f'(x) - 1| < \frac{1}{2}$) pour tout x dans ce voisinage, alors f y est strictement croissante et par conséquent injective.

Revenons au cas général $n \in \mathbb{N}^*$. f étant de classe $\mathcal{C}^1$, d'après le théorème III.7, on peut considérer $r > 0$ tel que pour tout $x \in B(0,r)$, $|||df_x - \mathrm{Id}||| < \frac{1}{2}$. Pour montrer que f est nécessairement injective, il faut étudier sa variation entre deux points $x, y \in B(0,r)$ en exploitant la relation entre les variations de f sur le segment $[x,y]$ et sa différentielle. Ceci nous fait penser directement au théorème des accroissements finis (théorème IV.1). Posons $U = B(0,r)$. On a pour tous $x, y \in U$ tels que $f(x) = f(y)$, d'après le théorème IV.1,

$$0 = f(y) - f(x) = \int_0^1 df_{(1-t)x+ty}(y-x)dt.$$

Dans le segment $[x,y] \subset B(0,r)$, df est proche de Id et alors pour tout $t \in [0,1]$, $df_{(1-t)x+ty}$ est également proche de Id. Par conséquent, $df_{(1-t)x+ty}(y-x)$ est proche de $y-x$, et enfin son intégrale (qui est égale à 0) est assez proche de $y-x$, ce qui n'est pas possible sauf si $x = y$. Formellement, ceci s'écrit

$$\begin{aligned}
\underbrace{\left\| \int_0^1 df_{(1-t)x+ty}(y-x)dt - (y-x) \right\|}_{=\|y-x\|} &= \left\| \int_0^1 (df_{(1-t)x+ty} - \mathrm{Id})(y-x)dt \right\| && (7)\\
&\leq \int_0^1 \left\| (df_{(1-t)x+ty} - \mathrm{Id})(y-x) \right\| dt && (8)\\
&\leq \int_0^1 |||df_{(1-t)x+ty} - \mathrm{Id}|||\, \|y-x)\|\, dt && (9)\\
&\leq \int_0^1 \frac{1}{2} \|y-x\|\, dt = \frac{1}{2} \|y-x\|. && (10)
\end{aligned}$$

Ceci implique donc que $\|y-x\| = 0$, c'est-à-dire que $y = x$. On a donc bien montré que F est injective sur U.

Passons maintenant à la deuxième étape. Posons $V = f(U)$. On veut montrer que quitte à remplacer U à un certain sous-ensemble de U, on peut supposer que V est ouvert. On va montrer que $B\left(0, \frac{r}{4}\right) \subset V$, puis remplacer U par $f^{-1}\left(B\left(0, \frac{r}{4}\right)\right)$. Soit $x \in U$ tel que $\|x\| < \frac{r}{2}$ et $y \in \mathbb{R}^n$ tel que $\|y\| < \frac{r}{4}$. Considérons la suite $(x_m)_{m\in\mathbb{N}}$ définie par

$$\begin{cases} x_0 = x \\ \forall m \in \mathbb{N},\ x_{m+1} = x_m + y - f(x_m). \end{cases}$$

On va montrer que (x_n) converge vers un point $x' \in U$ tel que $f(x') = y$. Pour cela, on va montrer que (x_m) est à valeurs dans U et qu'elle converge. Montrons le premier point par récurrence. La propriété

(cc) BY-NC-SA

$\|x_0\| \leq \frac{r}{2}$ est vérifiée pour $m = 0$. Soit $m \in \mathbb{N}$. Supposons que $\|x_m\| \leq \frac{r}{2}$. On a

$$\begin{aligned}
\|x_{m+1}\| = \|x_m + y - f(x_m)\| &\leq \|y\| + \left\| x_m - \int_0^1 df_{tx_m}(x_m)dt \right\| \\
&\leq \frac{r}{4} + \left\| \int_0^1 (df_{tx_m} - \mathrm{Id})(x_m)dt \right\| \\
&\leq \frac{r}{4} + \int_0^1 |||df_{tx_m} - \mathrm{Id}||| \, \|x_m\| \, dt \\
&\leq \frac{r}{4} + \frac{1}{2}\|x_m\| \leq \frac{r}{2}.
\end{aligned}$$

On a donc bien montré par récurrence que pour tout $m \in \mathbb{N}$, $\|x_m\| \leq \frac{r}{2}$. Montrons maintenant qu'elle converge. On a pour tout $m \in \mathbb{N}^*$,

$$\begin{aligned}
\|x_{m+1} - x_m\| &= \|x_m - x_{m-1} - (f(x_m) - f(x_{m-1}))\| \\
&= \left\| \int_0^1 \left(df_{(1-t)x_{m-1}+tx_m} - \mathrm{Id} \right)(x_m - x_{m-1})dt \right\| \\
&\leq \int_0^1 |||df_{(1-t)x_{m-1}+tx_m} - \mathrm{Id}||| \, \|x_m - x_{m-1}\| \, dt \\
&\underset{x_m, x_{m-1} \in U}{\leq} \frac{1}{2}\|x_m - x_{m-1}\|.
\end{aligned}$$

On en déduit donc par une récurrence immédiate que pour tout $m \in \mathbb{N}$,

$$\|x_{m+1} - x_m\| \leq \frac{1}{2^m}\|x_1 - x_0\| \leq \frac{1}{2^m}(\|x_1\| + \|x_0\|) \leq \frac{r}{2^m},$$

et enfin pour tout $\varepsilon > 0$ et tous $k, m \in \mathbb{N}$ tels que $m \geq k$ et $k \geq -\log_2\left(\frac{\varepsilon}{r}\right) + 1$,

$$\|x_m - x_k\| = \left\| \sum_{i=m}^{k-1} x_{i+1} - x_i \right\| \leq \sum_{i=m}^{k-1} \|x_{i+1} - x_i\| \leq \sum_{i=k}^{m} \frac{r}{2^i} \leq \frac{r}{2^{k-1}} \leq \varepsilon.$$

(x_m) est donc une suite de Cauchy dans un espace normé de dimension finie (donc complet), ce qui nous permet d'affirmer qu'elle converge. Notons sa limite $x' \in E$. On a montré précédemment que pour tout $m \in \mathbb{N}$, $\|x_m\| \leq \frac{r}{2}$, et donc $\|x'\| \leq \frac{r}{2}$, ce qui implique $x' \in U$. Enfin, en passant à la limite dans la relation de récurrence de (x_m), par continuité de f, on obtient $y - f(x') = 0$, et alors $y \in f(U) \subset V$. En résumé, on a considéré $y \in B\left(0, \frac{r}{4}\right)$ et on a montré que $y \in V$, c'est-à-dire que $B\left(0, \frac{r}{4}\right) \subset V$. On en déduit donc que V est bien un voisinage de 0. Quitte à restreindre f à $f^{-1}\left(B\left(0, \frac{r}{4}\right)\right)$ (qui est ouvert, car il s'agit de l'image réciproque d'un ouvert par une application continue) et remplacer V par $B\left(0, \frac{r}{4}\right)$, on peut supposer que V est ouvert.

Il reste maintenant à montrer que $\tilde{f}^{-1} \in \mathcal{C}^1(V, U)$. On commence par montrer que $\tilde{f}^{-1}$ est différentiable sur V. Soit $k \in \mathbb{R}^n$ et $b = f(c) \in V$ avec $c \in U$. On suppose que la norme de k est assez petite pour que $b + k \in V$, et on pose

$$h = \tilde{f}^{-1}(b + k) - \tilde{f}^{-1}(b) \text{ ce qui implique } f(c + h) - f(c) = k.$$

Si $\tilde{f}^{-1}$ est différentiable en b, la formule de la différentielle de la composée de deux fonctions nous permet de dire que $d\tilde{f}_b^{-1} = \left(df_{\tilde{f}^{-1}(b)}\right)^{-1} = (df_c)^{-1}$. Cette propriété suggère de procéder comme ce qui suit.

(cc) BY-NC-SA

$$\begin{aligned}\left\|\tilde{f}^{-1}(b+k)-\tilde{f}^{-1}(b)-(df_c)^{-1}(k)\right\| &= \left\|h-(df_c)^{-1}(f(c+h)-f(c))\right\| \\ &= \left\|(df_c)^{-1}(f(c+h)-f(c)-df_c(h))\right\| \\ &\leq |||\,(df_c)^{-1}|||\,\|f(c+h)-f(c)-df_c(h)\| = o(h).\end{aligned}$$

Une suite naturelle à ce qu'on a écrit ci-dessus est de comparer h et k. On a

$$\|h-k\| = \|f(c+h)-f(c)-(c+h-c)\| \underset{(*)}{\leq} \frac{1}{2}\|c+h-c\| = \frac{1}{2}\|h\|,$$

L'inégalité $(*)$ ci-dessus peut être obtenue en utilisant l'inégalité 10. L'inégalité ci-dessus implique que $\dfrac{\|h\|}{2} \leq \|k\| \leq 2\|h\|$ et alors $h = O(k)$. Enfin, on a

$$\left\|\tilde{f}^{-1}(b+k)-\tilde{f}^{-1}(b)-(df_c)^{-1}(k)\right\| = o(h) = o(k),$$

ce qui signifie que $\tilde{f}^{-1}$ est différentiable (et donc continue) sur V de différentielle égale en tout point $b \in V$ à $\left(df_{f^{-\tilde{1}}(b)}\right)^{-1}$. Enfin, il reste à prouver que $\tilde{f}^{-1}$ est de classe $\mathcal{C}^1$, ce qui est clair, car l'application $b \longmapsto \left(df_{\tilde{f}^{-1}(b)}\right)^{-1}$ est continue sur V, étant donné qu'il s'agit d'une composition d'applications continues. □

Remarque. Nous citons sans démonstration une formule hors programme très souvent utilisée en physique pour le changement de variables.

Soit φ un $\mathcal{C}^1-$difféomorphisme d'un ouvert $U \subset \mathbb{R}^n$ vers un ouvert $V \subset \mathbb{R}^n$. Cette formule est souvent utilisée en physique pour des ensembles fermés d'intérieur non vides $K \subset U$ et $L = \varphi(K)$. En fait, on peut formaliser ceci et garder la véracité de la formule ci-haut, pourvu que lesdits fermés vérifient certaines conditions qu'on ne précisera pas (mais qui sont, en général en physique de spé, toujours vérifiées).

Soit $L = \varphi(K) \subset \mathbb{R}^n$, et g une fonction intégrable de L dans $\mathbb{R}$. Lorsque K et L vérifient ces conditions (qu'on ne précisera pas par souci de simplicité), on admet la formule suivante.

$$\int_L g(x)dx = \int_K g(\varphi(y))\,|\det J_\varphi(y)|\,dy.$$

Un exemple d'application de cette formule en physique est le passage en coordonnées polaires. En effet, en reconsidérant l'application φ définie en (6), on a pour toute fonction g d'un ouvert $L \subset \mathbb{R}^* \times]-\pi,\pi[$ dans $\mathbb{R}$, on a

$$\begin{aligned}\int_L g(x,y)dxdy &= \int_{\varphi^{-1}(L)} g(r\cos\theta, r\sin\theta)\,|\det J_\varphi(r,\theta)|\,drd\theta \\ &= \int_{\varphi^{-1}(L)} g(r\cos\theta, r\sin\theta)\left|\det\begin{pmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{pmatrix}\right| drd\theta \\ &= \int_{\varphi^{-1}(L)} g(r\cos\theta, r\sin\theta) r drd\theta.\end{aligned}$$

Exercice IV.6.

Calculer $\int_0^{+\infty} e^{-x^2}$. *Indice : calculer* $\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy.$

(cc) BY-NC-SA

3. Conditions d'optimalité

On généralise les conditions d'optimalité vues pour les fonctions définies sur une partie de $\mathbb{R}$. Il est naturel de déduire de ce qui précède et par analogie avec le cas de dimension 1 que la différentielle en un point est nulle si ce point est un extremum local. On formalise cette propriété via la proposition suivante.

Proposition IV.7.

Soit $f : \Omega \longrightarrow \mathbb{R}$ et $a \in \Omega$. On suppose que f est différentiable en a. Les propositions suivantes sont vraies.

1. Si f admet un extremum local selon la direction $u \in \mathbb{R}^n$ en a, *i.e.* la fonction $\Phi : t \longmapsto f(a+tu)$ admet extremum local en 0, alors $D_u(f)(a) = 0$.
2. Si f admet un extremum local en a, alors $df_a = 0$ (*i.e.* $\nabla f_a = 0$).

Démonstration.

1. Il s'agit d'un résultat classique d'analyse sur $\mathbb{R}$, étant donné que $\Phi'(0) = D_u(f)(a)$ et donc le fait que Φ admette un extremum local en 0 implique directement que $D_u(f)(a) = 0$.

2. L'intuition derrière cette preuve est simple : si ∇f_a est non nul, alors pour h assez petit, on peut dire que $f(a+h) - f(a) \simeq \langle \nabla f_a,\, h\rangle$, et alors si la variation h a une direction opposée à ∇f_a (égal à $-\alpha\nabla f_a / \|\nabla f_a\|$ avec $\alpha > 0$ assez petit), la variation $f(a+h) - f(a)$ est nécessairement strictement négative, ce qui est en contradiction avec le fait que a est un minimum local car cela implique que pour tout $h \in \mathbb{R}^n$ assez petit, la variation $f(a+h) - f(a)$ est positive. Maintenant, formalisons ce raisonnement via la preuve ci-dessous.

   Supposons sans perte de généralité que f admet un minimum local en a. On peut alors considérer $\varepsilon > 0$ tel que pour tout $h \in B(0,\varepsilon)$, $f(a+h) - f(a) \geq 0$. Supposons par l'absurde que $\nabla f_a \neq 0$. On a

   $$\forall h \in B(0,\varepsilon),\ 0 \leq f(a+h) - f(a) = \langle \nabla f_a,\, h\rangle + o(h).$$

   On a alors, pour $\alpha \in]0, \varepsilon[$,

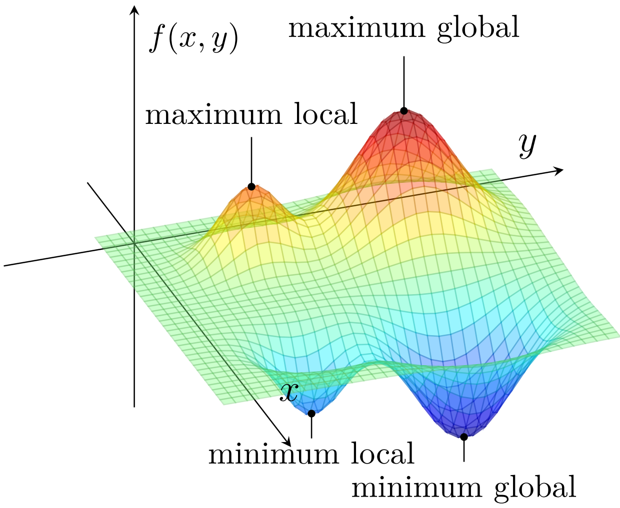
   $$0 \leq \left\langle \nabla f_a,\, -\alpha \frac{\nabla f_a}{\|\nabla f_a\|} \right\rangle + o(\alpha) = -\alpha \|\nabla f_a\| + o(\alpha).$$

   En divisant des deux côtés par α et en faisant tendre α vers 0, on obtient que $-\|\nabla f_a\| \geq 0$, *i.e.* $\nabla f_a = 0$, ce qui est absurde. On a donc bien montré le résultat voulu.

   Remarquons qu'on aurait pu prouver ce second point plus rapidement en utilisant le premier point : a étant un extremum local, sa dérivée selon toutes les directions est nulle, ce qui signifie que pour tout vecteur $u \in \mathbb{R}^n$, $df_a(u) = 0$ et donc $\nabla f_a = 0$.

$\square$

(cc) BY-NC-SA



Remarque. Bien entendu, comme sur $\mathbb{R}$, un minimum (resp. maximum) local d'une fonction convexe (resp. concave) est un minimum (resp. maximum) global.

Exercice (Régression linéaire) IV.8.

Soit $A \in \mathcal{M}_{r,n}(\mathbb{R})$, $r, n \in \mathbb{N}^*$ tels que $r \leq n$, et $B \in \mathbb{R}^r$.

1. Montrer que si A est de rang maximal (*i.e.* $\operatorname{Ker} A = \{0\}$), alors $A^T A$ est inversible.
2. On suppose que A est de rang maximal. Trouver le minimum global sur $\mathbb{R}^n$ de la fonction

$$g : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ X & \longmapsto \|AX - B\|_2^2 . \end{cases}$$

V Dérivées d'ordre supérieur à 2

Dans cette partie, on considère $n, p \in \mathbb{N}^*$ et on suppose que $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ sont respectivement munis de leurs normes euclidiennes canoniques.

Définition V.1.

Soit Ω un ouvert de $E = \mathbb{R}^n$, $a \in \Omega$ et f une fonction de Ω dans F. Lorsque cela a un sens, on définit récursivement la quantité suivante.

$$\forall k \in \mathbb{N}^*,\ \forall (i_1, \ldots, i_k) \in [\![1; n]\!]^k,\ \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial^{k-1} f}{\partial x_{i_2} \ldots \partial x_{i_k}} \right).$$

Toutes les dérivées de cette forme sont appelées dérivées d'ordre k. L'ensemble des fonctions admettant des dérivées à l'ordre $m \in \mathbb{N}^*$ continues en tout point est appelé ensemble des fonctions de classe $\mathcal{C}^m$, et est noté, lorsque les ensembles de départ et d'arrivée sont respectivement Ω et F, $\mathcal{C}^m(\Omega, F)$.

En fait, on peut montrer qu'on peut librement permuter les dérivées d'ordre inférieur à $k \in \mathbb{N}^*$ pour les fonctions de classe $\mathcal{C}^k$.

BY-NC-SA

Proposition V.2.

Soit $k \in \mathbb{N}^*$ et $f \in \mathcal{C}^k(\Omega, F)$. Pour tout $(i_1, \ldots, i_k) \in \mathbb{N}^k$ et $\sigma \in \mathcal{S}_k$ permutation de $[\![1;k]\!]$, on a

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \ldots \partial x_{i_{\sigma(k)}}}.$$

Démonstration. Étant donné que pour tout $k \geq 2$, l'ensemble de permutations de $\mathcal{S}_k$ est généré par l'ensemble des transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ avec $i \in [\![1; k-1]\!]$, il suffit de montrer le lemme suivant.

Lemme V.3.

Pour tout $g \in \mathcal{C}^2(\Omega', \mathbb{R})$ avec $\Omega' \subset \mathbb{R}^2$ ouvert, on a

$$\frac{\partial^2 g}{\partial x_1 \partial x_2} = \frac{\partial^2 g}{\partial x_2 \partial x_1}.$$

Démonstration du lemme. Sans perte de généralité, on suppose que $(0,0) \in \Omega'$. Pour montrer le résultat voulu, on peut remplacer tout point quelconque de Ω par $(0,0)$, car la preuve est la même, et alors il suffit de montrer l'égalité

$$\frac{\partial^2 g}{\partial x_1 \partial x_2}(0,0) = \frac{\partial^2 g}{\partial x_2 \partial x_1}(0,0).$$

On va utiliser le théorème de Fubini, qui nous permet de permuter les intégrales, pour montrer qu'on peut permuter les dérivées. On pose pour tout $(x,y) \in \Omega'$,

$$F(x,y) = \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) dt \right) ds \text{ et } G(x,y) = \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) dt \right) ds.$$

On a alors pour tous $(x,y) \in \Omega'$,

$$F(x,y) = \int_0^x \left(\frac{\partial g}{\partial x_1}(s,y) - \frac{\partial g}{\partial x_1}(s,0) \right) ds = g(x,y) - g(0,y) - g(x,0) + g(0,0).$$

$$\begin{aligned} G(x,y) = \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) dt \right) ds &\underset{\text{Fubini}}{=} \int_0^y \left(\int_0^x \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) ds \right) dt \\ &= \int_0^y \left(\frac{\partial g}{\partial x_2}(x,t) - \frac{\partial g}{\partial x_2}(0,t) \right) dt \\ &= g(x,y) - g(x,0) - g(0,y) + g(0,0). \end{aligned}$$

Ces deux égalités entrainent

$$\forall (x,y) \in \Omega', \; F(x,y) = G(x,y) \; i.e. \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) ds \right) dt = 0.$$

Supposons maintenant par l'absurde, sans perte de généralité que

$$\frac{\partial^2 g}{\partial x_1 \partial x_2}(0,0) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(0,0) = \lambda > 0.$$

BY-NC-SA

f est de classe $\mathcal{C}^2$, donc la fonction ci-dessus est continue. Il existe donc $\eta > 0$ tel que pour tout $(s,t) \in \Omega'$, on ait

$$\|(s,t)\|_\infty < \eta \implies \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) > \frac{\lambda}{2}.$$

On en déduit donc que

$$\int_0^\eta \left(\int_0^\eta \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) ds \right) dt \geq \frac{\eta^2 \lambda}{2} > 0,$$

ce qui est absurde. On a donc bien démontré lemme voulu.

Revenons maintenant à la preuve du théorème. Soit $k \in \mathbb{N}^*$ et $\sigma \in \mathcal{S}_k$. On sait que l'ensemble des permutations est généré par l'ensemble des transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ avec $i \in [\![1; k-1]\!]$, il existe donc $m \in \mathbb{N}$ et $\tau_1, \dots, \tau_m \in \mathcal{S}_k$ des transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ telles que $\sigma = \tau_1 \circ \dots \circ \tau_m$. Étant donné que toute dérivée d'ordre au plus $k-2$ de f est de classe $\mathcal{C}^2$ par définition, on peut affirmer d'après le lemme V.3 qu'on peut permuter les dérivées voisines c'est-à-dire pour toute transposition $\tau \in \mathcal{S}_k$ de la forme $\begin{pmatrix} d & d+1 \end{pmatrix}$ avec $d \in [\![1; k-1]\!]$,

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_{d-1}} \partial x_{i_{d+1}} \partial x_{i_d} \partial x_{i_{d+2}} \dots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\tau(1)}} \dots \partial x_{i_{\tau(k)}}}.$$

Enfin, en appliquant successivement les permutations, on peut affirmer que pour tous $i_1, \dots, i_k \in [\![1; n]\!]^k$,

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\tau_1 \circ \dots \circ \tau_m(1)}} \dots \partial x_{i_{\tau_1 \circ \dots \circ \tau_m(k)}}} = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \dots \partial x_{i_{\sigma(k)}}}.$$

$\square$

Conséquence. Cette propriété de permutation des dérivées a plusieurs applications intéressantes. Citons-en une qui est courante en physique de deuxième année.
Soit $g_1, g_2, g_3 \in \mathcal{C}^1(\Omega_3, \mathbb{R})$ et $g : x \longmapsto (g_1(x), g_2(x), g_3(x))$ où Ω_3 est un ouvert de $\mathbb{R}^3$. On définit l'opérateur rotationnel, comme définit en physique, de la manière suivante.

$$\operatorname{Rot} g \underset{\substack{\uparrow \\ \text{abus de notation}}}{=} \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \wedge \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial g_3}{\partial x_2} - \dfrac{\partial g_2}{\partial x_3} \\ \dfrac{\partial g_1}{\partial x_3} - \dfrac{\partial g_3}{\partial x_1} \\ \dfrac{\partial g_2}{\partial x_1} - \dfrac{\partial g_1}{\partial x_2} \end{pmatrix}$$

Le fait de pouvoir permuter les variables nous permet d'avoir l'égalité suivante pour toute fonction $f \in \mathcal{C}^2(\Omega_3, \mathbb{R})$, connue au programme de physique de deuxième année.

$$\operatorname{Rot} \nabla f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_2 \partial x_3} - \dfrac{\partial^2 f}{\partial x_3 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_3 \partial x_1} - \dfrac{\partial^2 f}{\partial x_1 \partial x_3} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} - \dfrac{\partial^2 f}{\partial x_2 \partial x_1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

(cc) BY-NC-SA

Exercice V.4.

Considérons $u : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}$ définie par

$$\forall (t,x) \in \mathbb{R}^+ \times \mathbb{R},\ u(t,x) = \sum_{k=0}^{\infty} \alpha_k e^{-k^2 t} \cos(kx),$$

où (α_k) est le terme général d'une série absolument convergente. Montrer que u est continue sur $[0,+\infty[\times\mathbb{R}$ et que $u \in \mathcal{C}^2(]0,+\infty[\times\mathbb{R},\mathbb{R})$.

Exercice V.5.

Déterminer toutes les fonctions $u \in \mathcal{C}^1(\mathbb{R}^n,\mathbb{R})$ telles qu'il existe $(a_1,\dots,a_n) \in \mathbb{R}^n \setminus \{0\}$ tel que

$$\forall x \in \mathbb{R}^n,\ \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i}(x) = 0.$$

Exercice V.6.

Soit O un ouvert de $\mathbb{R}^n$ tel que pour tout $x \in O$ et $\lambda > 0$, $\lambda x \in O$. On dit qu'une fonction $f : O \longrightarrow \mathbb{R}$ est homogène de degré $r > 0$ si pour tout $x \in \mathbb{R}^n$, et $\lambda > 0$, $f(\lambda x) = \lambda^r f(x)$. Montrer que $f \in \mathcal{C}^1(O,\mathbb{R})$ est homogène de degré $r > 0$ si et seulement si

$$\forall x = (x_1,\dots,x_n) \in O,\ \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(x) = r f(x).$$

(cc) BY-NC-SA

Correction de l'exercice II.4 :

1. π_i est une application linéaire en dimension finie, elle est donc différentiable. De plus, pour tout $a \in \mathbb{R}^n$, $d\pi_{i,a} = \pi_i$ et

$$\nabla \pi_{i,a} = \begin{pmatrix} \dfrac{\partial \pi_i}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial \pi_i}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial \pi_i}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial \pi_i}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \mathbb{1}_{i=1} \\ \vdots \\ \mathbb{1}_{i=n} \end{pmatrix} = e_i,$$

où e_i est le $i-$ème élément de la base canonique de $\mathbb{R}^n$.

2. On a pour tout $x, h \in E$, en notant $\langle ., .\rangle_N$ le produit scalaire associé à la norme euclidienne N,

$$N(x+h)^2 - N(x)^2 = \langle x+h,\ x+h\rangle_N - \langle x,\ x\rangle_N = 2\langle x,\ h\rangle_N + \langle h,\ h\rangle_N = 2\langle x,\ h\rangle_N + N(h)^2.$$

Toutes les normes sont équivalentes en dimension finie, il existe donc $C > 0$ tel que pour tout $h \in E$, $N(h) \leq C\,\|h\|_E$, et alors on a $N(h)^2 = o(h)$, ce qui donne

$$N(x+h)^2 - N(x)^2 = 2\langle x,\ h\rangle_N + o(h),$$

et finalement pour tout $x, h \in E$, $d\psi_{1,x}(h) = 2\langle x,\ h\rangle$, et $\nabla\psi_{1,x} = 2x$.

3. On travaille avec une norme d'opérateur $|||.|||$ quelconque, associée à la même norme pour l'espace de départ et l'espace d'arrivée (en dimension finie le choix de la norme importe peu). On a pour tout $A, H \in \mathcal{M}_n(\mathbb{R})$,

$$(A+H)^2 - A^2 = AH + HA + H^2.$$

De plus, on a $|||H^2||| \leq |||H|||^2 = o(H)$, et alors

$$(A+H)^2 - A^2 = AH + HA + o(H),$$

ce qui donne $d\psi_{2,A}(H) = AH + HA$.

Correction de l'exercice II.11 :

1. f^2 est la composition de $\psi : x \longmapsto x^2$ et de f. On a alors, en appliquant la proposition II.8, pour tout $a \in \Omega$, ψ est différentiable en $f(a)$ et f est différentiable en a, donc $\psi \circ f$ est différentiable en a et

$$df_a^2 = d\psi \circ f_a = d\psi_{f(a)} \circ df_a = 2f(a)df_a.$$

2. $\sqrt{f}$ est la composition de $\omega : x \longmapsto \sqrt{x}$ et de f. On a alors, en appliquant la proposition II.8, pour tout $a \in \Omega$, f est différentiable en a et $f(a) > 0$ donc ω est différentiable en $f(a)$ et

$$d\omega \circ f_a = d\omega_{f(a)} \circ df_a = \frac{1}{2\sqrt{f(a)}}df_a.$$

Correction de l'exercice II.12 :

(cc) BY-NC-SA

f n'est pas différentiable en b car $x \longmapsto \|x\|_E$ n'est pas différentiable en 0. Soit $a \in E \setminus \{b\}$. On a

$$\begin{aligned}
\|a+h-b\|_E &= \sqrt{\langle a+h-b,\ a+h-b\rangle} \\
&= \sqrt{\|a-b\|_E^2 + 2\langle a-b,\ h\rangle + \|h\|_E^2} \\
&= \|a-b\|_E \sqrt{1 + \frac{2\langle a-b,\ h\rangle}{\|a-b\|_E^2} + \frac{\|h\|_E^2}{\|a-b\|_E^2}} \\
&= \|a-b\|_E \left(1 + \frac{\langle a,\ h\rangle}{\|a-b\|_E^2} + \underbrace{\frac{\|h\|_E^2}{2\|a-b\|_E^2} + o\left(\frac{2\langle a,\ h\rangle}{\|a-b\|_E^2} + \frac{\|h\|_E^2}{\|a-b\|_E^2}\right)}_{=o(h)}\right) \\
&= \|a-b\|_E + \left\langle \frac{a-b}{\|a-b\|_E},\ h\right\rangle + o(h).
\end{aligned}$$

On en déduit donc que pour tout $h \in E$, $df_a(h) = \left\langle \frac{a-b}{\|a-b\|_E},\ h\right\rangle$ et $\nabla f_a = \frac{a-b}{\|a-b\|_E}$.

Correction de l'exercice III.3 :

1. h_1 est composition de la fonction racine et de la fonction polynomiale $(x_1, \ldots, x_n) \longmapsto (x_1-a_1)^2 + \cdots + (x_n-a_n)^2$, qui est strictement positive sur $\mathbb{R}^n \setminus \{a\}$. On en déduit que h_1 est différentiable sur $\mathbb{R}^n \setminus \{a\}$, et en utilisant la proposition 3, on peut affirmer que

$$\forall u = (u_1, \ldots, u_n) \in \mathbb{R}^n \setminus \{a\},\ \forall h \in \mathbb{R}^n,\ dh_{1,u}(h) = \langle \nabla h_{1,u},\ h\rangle,$$

où

$$\nabla h_{1,u} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(u) \\ \vdots \\ \frac{\partial h_1}{\partial x_n}(u) \end{pmatrix} = \begin{pmatrix} \frac{u_1-a_1}{\sqrt{(u_1-a_1)^2+\cdots+(u_n-a_n)^2}} \\ \vdots \\ \frac{u_n-a_n}{\sqrt{(u_1-a_1)^2+\cdots+(u_n-a_n)^2}} \end{pmatrix} = \frac{u-a}{\|u-a\|_2}.$$

Remarquons que ce résultat est cohérent avec celui de l'exercice précédent.

2. det est différentiable sur $\mathcal{M}_n(\mathbb{R})$ car il s'agit d'une fonction polynomiale. On veut calculer la différentielle de det. Pour toute matrice $M = (m_{i,j})_{i,j \in [\![1;n]\!]}$ et $i_0, j_0 \in [\![1;n]\!]$, on note $\Delta_{i_0,j_0}(M)$ le mineur de la matrice M d'indices (i_0, j_0), *i.e.* le déterminant de la matrice M à laquelle on a enlevé le i_0-ème ligne et la j_0-ème colonne, qui s'écrit

$$\Delta_{i_0,j_0}(M) = \begin{vmatrix} m_{1,1} & \ldots & m_{1,j_0-1} & m_{1,j_0+1} & \ldots & m_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ m_{i_0-1,1} & \ldots & m_{i_0-1,j_0-1} & m_{i_0-1,j_0+1} & \ldots & m_{i_0-1,n} \\ m_{i_0+1,1} & \ldots & m_{i_0+1,j_0-1} & m_{i_0+1,j_0+1} & \ldots & m_{i_0+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ m_{n,1} & \ldots & m_{n,j_0-1} & m_{n,j_0+1} & \ldots & m_{n,n} \end{vmatrix}.$$

Soit $i_0, j_0 \in [\![1;n]\!]$. On développant par rapport à colonne j_0, on peut écrire pour toute matrice

(cc) BY-NC-SA

$A = (a_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{R})$

$$\det A = \sum_{i=1}^{n} \underbrace{(-1)^{i+j_0}\, a_{i,j_0} \quad \underbrace{\Delta_{i,j_0}(A).}_{\text{independant de } a_{i_0,j_0}}}_{H_{i,j_0}(A)}$$

On a alors, en notant $E_{i_0,j_0} = (\delta_{i,i_0}\delta_{j,j_0})_{i,j\in[\![1;n]\!]}$ la matrice avec un 1 dans la coordonnée i_0, j_0 et 0 partout ailleurs, et

$$\frac{\partial \det}{\partial x_{i_0,j_0}}(A) = \lim_{t\to 0} \frac{\det(A + tE_{i_0,j_0}) - \det A}{t}$$

la dérivée de det par rapport à la variable de coordonnées i_0, j_0, on peut écrire

$$\frac{\partial \det}{\partial x_{i_0,j_0}}(A) = \sum_{i=1}^{n} (-1)^{i+j_0} \frac{\partial H_{i,j_0}}{\partial x_{i_0,j_0}}(A) = (-1)^{i_0+j_0} \Delta_{i_0,j_0}(A).$$

On en déduit donc que pour tout $H = (h_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{R})$,

$$d\det_A(H) = \sum_{i,j\in[\![1;n]\!]} h_{i,j} \frac{\partial \det}{\partial x_{i,j}}(A) = \sum_{i,j\in[\![1;n]\!]} (-1)^{i+j} h_{i,j}\Delta_{i,j}(A) = \mathrm{Tr}(H(\mathrm{Com}\ A)^T),$$

et alors, on peut écrire $\nabla \det_A = \mathrm{Com}A$.

3. Soit $M \in \mathcal{M}_n(\mathbb{R})$. On a

$$(M + H)A(M + H) = MAM + HAM + MAH + HAH.$$

$HAH = o(H)$ car on a $\|HAH\| \leq \|A\|\,\|H\|^2$, et $H \longmapsto HAM + MAH$ est linéaire, on en déduit alors que

$$h_3(M + H) - h_3(M) = HAM + MAH + o(H),$$

i.e. h_3 est différentiable et pour tout $H \in \mathcal{M}_n(\mathbb{R})$, $dh_{3,M}(H) = MAH + HAM$.

Correction de l'exercice III.4 :

1. Soit $e_1, \ldots, e_n$ une base de E. On a pour tout $x_1, \ldots, x_k \in E$, en posant pour tout $i \in [\![1;k]\!]$, $x_i = x_{i,1}e_1 + \ldots, x_{i,n}e_n$,

$$\begin{aligned} f(x_1, \ldots, x_n) &= \sum_{r_1=1}^{n} x_{1,r_1} f(e_{r_1}, x_2, \ldots, x_k) \\ &= \sum_{r_1=1}^{n}\sum_{r_2=1}^{n} x_{1,r_1}x_{2,r_2} f(e_{r_1}, e_{r_2}, x_3, \ldots, x_k) \\ &\ \ \vdots \\ &= \sum_{r_1=1}^{n}\sum_{r_2=1}^{n} \ldots \sum_{r_k=1}^{n} x_{1,r_1}x_{2,r_2}\ldots x_{k,r_k} \underbrace{f(e_{r_1}, e_{r_2}, \ldots, e_{r_k})}_{\text{indépendant de } x_1,\ldots,x_n} \end{aligned}$$

On en déduit alors que f est une fonction polynomiale (somme et produit de fonctions de la forme $(x_1, \ldots, x_n) \longmapsto x_{i,r_i}$, $i \in [\![1;n]\!]$) en les coordonnées de ses arguments en dimension finie, ce qui nous permet d'affirmer que f est continue sur E^k.

(cc) BY-NC-SA

2. Soit $x_1, \dots, x_k \in E$ non nuls. On écrit

$$\|f(x_1, \dots, x_n)\|_F = \|x_1\|_E \dots \|x_k\|_E \underbrace{\left\| f\left(\frac{x_1}{\|x_1\|_E}, \dots, \frac{x_k}{\|x_k\|_E}\right) \right\|_F}_{A}.$$

On veut montrer que A est bornée. Pour ce faire, on remarque que A est égale à f appliquée à des éléments de l'ensemble $S(0,1) = \{x \in E,\ \|x\|_E = 1\}$. D'après la question précédente, f est continue sur E^k et alors, elle est bornée sur le compact $S(0,1)^k$ (car le produit de compacts est compact d'après le chapitre sur les compacts). En posant alors C un majorant de $(y_1, \dots, y_k) \longmapsto \|f(y_1, \dots, y_k)\|_F$ sur $S(0,1)^k$, on obtient

$$\|f(x_1, \dots, x_n)\|_F = \|x_1\|_E \dots \|x_k\|_E \left\| f\left(\frac{x_1}{\|x_1\|_E}, \dots, \frac{x_k}{\|x_k\|_E}\right) \right\|_F \leq C \|x_1\|_E \times \dots \times \|x_k\|_E.$$

L'inégalité reste vraie lorsque $x_1, \dots, x_k$ sont potentiellement nuls.

3. Soit $a, h \in E$. On veut écrire $f(a+h, \dots, a+h) - f(a, \dots, a)$ comme somme d'une application linéaire appliquée à h et un terme négligeable devant h. Lorsqu'on développe $f(a+h, \dots, a+h)$ comme un "produit" non commutatif, on obtient trois types de termes : le terme $f(a, \dots, a)$, les termes de la forme $f(a, \dots, a, h, a, \dots, a)$ où h apparaît une fois, et les termes où h apparait plus d'une fois. Les termes où h apparait une fois sont linéaires en h, il est donc naturel d'essayer de montrer que les termes où h apparait plus d'une fois sont négligeables devant h. Soit $r \in [\![2; k]\!]$. On va montrer que $f(a, \dots, a, \underbrace{h, \dots, h}_{r \text{ fois}}) = o(h)$. Une fois cela fait, la preuve pour tous les autres termes où h apparait plus d'une fois se fait exactement de la même manière.

D'après la question précédente, il existe $C \geq 0$ tel que pour tout $h \in E$,

$$\|f(a, \dots, a, h \dots, h)\|_F \leq C \|h\|_E^r \|a\|_E^{k-r} = o(h).$$

On en déduit donc que

$$f(a+h, \dots, a+h) = f(a, \dots, a) + f(h, a, \dots, a) + f(a, h, a, \dots, a) + \dots + f(a, \dots, a, h) + o(h),$$

et alors

$$dg_a(h) = f(h, a, \dots, a) + f(a, h, a, \dots, a) + \dots + f(a, \dots, a, h).$$

Enfin, lorsque f est symétrique, on a $dg_a(h) = kf(a, \dots, a, h)$.

Remarques.

→ En fait, on aurait pu faire la question 3 de l'exercice précédent en utilisant le résultat ci-dessus. En effet, pour une matrice A de taille n, considérons l'application $g : M \longmapsto f(M, M)$ où pour toutes matrices U, V de taille n, $f(U, V) = UAV$. f est $2-$linéaire, donc la différentielle de g en une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est égale à $H \longmapsto f(M, H) + f(H, M) = MAH + HAM$.

→ Dans le cas où f est symétrique, la formule du binôme de Newton s'applique, *i.e.* on peut écrire pour tout $a, b \in E$,

$$f(a+b, \dots, a+b) = \sum_{i=0}^{k} \binom{k}{i} f(\underbrace{a, \dots, a}_{i \text{ fois}}, \underbrace{b, \dots, b}_{k-i \text{ fois}}),$$

et dans le cas général où f n'est pas forcément symétrique,

$$f(a+b, \dots, a+b) = \sum_{U \subset [\![1;k]\!]} f(\mathbb{1}_{1 \in U} a + \mathbb{1}_{1 \notin U} b, \dots, \mathbb{1}_{k \in U} a + \mathbb{1}_{k \notin U} b).$$

(cc) BY-NC-SA

→ On aurait aussi pu faire la question 2 de la manière suivante. Pour tout $x_1, \dots, x_n \in E$, on a

$$\begin{aligned}
\|f(x_1,\dots,x_n)\|_F &= \left\| \sum_{r_1=1}^{n}\sum_{r_2=1}^{n}\cdots\sum_{r_k=1}^{n} x_{1,r_1}x_{2,r_2}\dots x_{k,r_k} f(e_{r_1},e_{r_2},\dots,e_{r_k}) \right\|_F \\
&\leq \sum_{r_1=1}^{n}\sum_{r_2=1}^{n}\cdots\sum_{r_k=1}^{n} |x_{1,r_1}x_{2,r_2}\dots x_{k,r_k}| \, \|f(e_{r_1},e_{r_2},\dots,e_{r_k})\|_F \\
&\leq \sum_{r_1=1}^{n}\sum_{r_2=1}^{n}\cdots\sum_{r_k=1}^{n} \left(\max_{r\in[\![1;n]\!]}|x_{1,r}|\right)\times\cdots\times\left(\max_{r\in[\![1;n]\!]}|x_{k,r}|\right)\times \max_{m_1,\dots,m_k\in[\![1;n]\!]}\|f(e_{m_1},e_{m_2},\dots,e_{m_k})\|_F \\
&= n^k \times \left(\max_{r\in[\![1;n]\!]}|x_{1,r}|\right)\times\cdots\times\left(\max_{r\in[\![1;n]\!]}|x_{k,r}|\right)\times \max_{m_1,\dots,m_k\in[\![1;n]\!]}\|f(e_{m_1},e_{m_2},\dots,e_{m_k})\|_F \\
&\underset{(*)}{\leq} \underbrace{M^k n^k \times \left(\max_{m_1,\dots,m_k\in[\![1;n]\!]}\|f(e_{m_1},e_{m_2},\dots,e_{m_k})\|_F\right)}_{C} \times \|x_1\|_E\times\cdots\times\|x_k\|_E \\
&= C\,\|x_1\|_E\times\cdots\times\|x_k\|_E\,.
\end{aligned}$$

où M est une constante positive telle que pour tout $y = y_1e_1 + \dots + y_ne_n \in E$, $\max_{r\in[\![1;n]\!]}|y_r| \leq M\|y\|_F$ (qui existe par équivalence des normes en dimension finie). L'inégalité $(*)$ est vraie d'après l'inégalité $\max_{r\in[\![1;n]\!]}|y_r| \leq M\|y\|_F$.

→ Nous encourageons le lecteur à montrer ce résultat plus général à titre d'exercice. La différentielle de la fonction $(x_1,\dots,x_k) \longmapsto f(x_1,\dots,x_n)$ en tout point $a = (a_1,\dots,a_k) \in E^k$ est définie par

$$\forall (h_1,\dots,h_k)\in E^k,\ df_a(h_1,\dots,h_k) = f(h_1,a_2,\dots,a_k) + f(a_1,h_2,\dots,a_k) + \dots + f(a_1,\dots,a_{k-1},h_k).$$

Correction de l'exercice III.9 :

L'application $g : (x,y) \longmapsto x+iy$ de $\mathbb{R}^2$ dans $\mathbb{C}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ (où $\mathbb{C}$ est vu comme un $\mathbb{R}-$espace vectoriel de dimension 2). De plus, d'après le cours des séries entières, $f : z \longmapsto \sum_{k=0}^{\infty} a_k z^k$ est de classe $\mathcal{C}^1$ sur $B_{\mathbb{C}}(0,R) = \{z\in\mathbb{C},\ |z|<R\}$ et en posant $B_{\mathbb{R}^2}(0,R) = \{(x,y)\in\mathbb{R}^2,\ \sqrt{x^2+y^2}<R\}$, on a clairement $g(B_{\mathbb{R}^2}(0,R)) = B_{\mathbb{C}}(0,\mathbb{R})$ et donc $f\circ g$ est de classe $\mathcal{C}^1$ sur $B_{\mathbb{R}^2}(0,R)$.

Correction de l'exercice IV.2 :

On pose $\Phi : t \longmapsto f((1-t)a+tb)$, qu'on définit sur l'intervalle $[0,1]$. Φ est dérivable sur $]0,1[$ et continue sur $[0,1]$. On a donc, d'après l'égalité des accroissements finis,

$$\exists\theta\in]0,1[,\ \Phi(1)-\Phi(0) = \Phi'(\theta) \underset{\substack{\uparrow \\ \text{règle de la chaine}}}{=} \left\langle \nabla f_{(1-\theta)a+\theta b},\ b-a\right\rangle,$$

et finalement

$$f(b)-f(a) = \left\langle \nabla f_{(1-\theta)a+\theta b},\ b-a\right\rangle.$$

(cc) BY-NC-SA

Correction de l'exercice IV.3 :

On considère la fonction

$$f : \begin{cases}]0,+\infty[^n & \longrightarrow \mathbb{R} \\ (x_1,\dots,x_n) & \longmapsto \left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}} \end{cases}.$$

En rejetant un coup d'oeil à l'inégalité qu'on veut montrer, on voit qu'elle est équivalente à

$$\forall a_1,\dots,a_n,b_1,\dots,b_n \in \mathbb{R}^+,\ f(a_1+b_1,\dots,a_n+b_n) - f(a_1,\dots,a_n) \geq f(b_1,\dots,b_n).$$

Cette forme nous faire penser au théorème des accroissements finis sur le segment $[(a_1,\dots,a_n),(a_1+b_1,\dots,a_n+b_n)]$. f est clairement différentiable sur $]0,+\infty[^n$, car il s'agit de la composition de $x \longmapsto x^{\frac{1}{n}}$ et $(x_1\dots,x_n) \longmapsto \prod_{k=1}^{n} x_k$ qui sont toutes les deux des fonctions différentiables en respectivement tout point de $]0,+\infty[$ et $]0,+\infty[^n$. On a de plus, pour tout $i \in [\![1;n]\!]$ et $x_1,\dots,x_n \in]0,+\infty[^n$,

$$\frac{\partial f}{\partial x_i}(x_1,\dots,x_n) = \frac{1}{n}\left(\prod_{\substack{k=1\\k\neq i}}^{n} x_k\right)\left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}-1} = \frac{1}{n}\frac{1}{x_i}\left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}}.$$

En appliquant l'égalité des accroissements finis montrée dans l'exercice précédent, on peut dire qu'il existe $\theta \in]0,1[$ tel qu'en posant pour tout $k \in [\![1;n]\!]$, $x_k = (1-\theta)a_k + \theta(a_k+b_k) = a_k + \theta b_k$,

$$\begin{aligned} f(a_1+b_1,\dots,a_n+b_n) - f(a_1,\dots,a_n) &= \left\langle \nabla f_{(x_1,\dots,x_n)},\ (b_1,\dots,b_n)\right\rangle \\ = \frac{1}{n}\sum_{k=1}^{n} b_k \frac{1}{x_k}\left(\prod_{h=1}^{n} x_h\right)^{\frac{1}{n}} \underset{\substack{\uparrow\\ \text{AM-GM}}}{\geq} \left(\prod_{k=1}^{n} b_k \frac{1}{x_k}\left(\prod_{h=1}^{n} x_h\right)^{\frac{1}{n}}\right)^{\frac{1}{n}} &= \left(\prod_{k=1}^{n} b_k\right)^{\frac{1}{n}} = f(b_1,\dots,b_k), \end{aligned}$$

ce qui est bien l'inégalité voulue.

Correction de l'exercice IV.6 :

L'indice fourni dans cet exercice est de calculer $\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy$. Essayons de comprendre pourquoi. Il est clair que cette intégrale est bien définie, car la fonction sous le signe intégral est intégrable. De plus, on a

$$\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy = \left(\int_0^{+\infty} e^{-x^2}dx\right)\left(\int_0^{+\infty} e^{-y^2}dx\right) = \left(\int_0^{+\infty} e^{-x^2}dx\right)^2.$$

Il est donc clair que la valeur de cette intégrale double nous donne directement la valeur de l'intégrale qu'on souhaite calculer. On va donc maintenant trouver la valeur de cette intégrale double. En la regardant de plus près, la quantité x^2+y^2 nous fait penser à un changement de variable polaire. En effet, on a, en utilisant l'application φ définie en 6,

$$\begin{aligned} \int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy &= \int_0^{\frac{\pi}{2}}\int_0^{+\infty} e^{-r^2}\left|\det J_\phi(r,\theta)\right| drd\theta \\ &= \int_0^{\frac{\pi}{2}}\int_0^{+\infty} e^{-r^2} r drd\theta = \frac{\pi}{4}. \end{aligned}$$

CC BY-NC-SA

On en déduit alors finalement que

$$\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Correction de l'exercice IV.8 :

1. On a pour tout $X \in \mathbb{R}^n$,

$$A^T AX = 0 \implies X^T A^T AX = 0 \implies \|AX\|_2 = 0 \implies AX = 0 \implies X = 0.$$

On en déduit que $\operatorname{Ker} A^T A = \{0\}$ et donc $A^T A$ étant une matrice carrée (de taille n), elle est inversible.

2. g est la composition d'une application affine $g_1 : X \longmapsto AX - B$ et de $g_2 : X \longmapsto \|X\|_2^2$ qui est clairement convexe (le lecteur ayant un doute peut essayer de le montrer lui-même). On en déduit donc que g est convexe. Il suffit alors de trouver où son gradient s'annule pour trouver son minimum. On a pour tout $U, H \in \mathbb{R}^n$,

$$\begin{aligned} dg_U(H) &= dg_{2,g_1(U)} \circ dg_{1,U}(H) = \left\langle \nabla g_{2,g_1(U)},\ dg_{2,U}(H) \right\rangle \\ &= 2 \langle AU - B,\ AH \rangle = (AU - B)^T AH = (A^T AU - A^T B)^T H. \end{aligned}$$

On en déduit donc que

$$dg_U = 0 \iff A^T AU - A^T B = 0 \iff U = (A^T A)^{-1} A^T B.$$

Correction de l'exercice V.4 :

Dans cet exercice, on utilise des propriétés du cours des suites et séries de fonctions. Posons pour tout $k \in \mathbb{N}$ et $(x, t) \in \mathbb{R} \times \mathbb{R}^+$, $u_k(t, x) = \alpha_k e^{-k^2 t} \cos(kx)$. On a pour tout $k \in \mathbb{N}$, $\sup\limits_{(t,x) \in \mathbb{R}^+ \times \mathbb{R}} |u_k(t, x)| \leq |\alpha_k|$, donc la série de fonctions (continues) converge normalement, car $\sum |\alpha_k|$ converge. On en déduit donc que u est continue sur $\mathbb{R}^+ \times \mathbb{R}$. On a de même pour tout $(t, x) \in \mathbb{R}_+^* \times \mathbb{R}$,

$$\forall k \in \mathbb{N},\ \frac{\partial u_k}{\partial x}(t, x) = -k\alpha_k e^{-k^2 t} \sin(kt) \text{ et } \frac{\partial u}{\partial t}(t, x) = -k^2 \alpha_k e^{-k^2 t} \cos(kx).$$

On a alors pour tout $k \in \mathbb{N}$ et $\tau > 0$,

$$\begin{aligned} \sup_{(t,x) \in [\tau, +\infty[\times \mathbb{R}} \left| \frac{\partial u}{\partial x}(t, x) \right| &\leq |\alpha_k|\, k^2 e^{-k^2 \tau} = o(\alpha_k), \\ \sup_{(t,x) \in [\tau, +\infty[\times \mathbb{R}} \left| \frac{\partial u}{\partial t}(t, x) \right| &\leq |\alpha_k|\, k e^{-k^2 \tau} = o(\alpha_k), \end{aligned}$$

et donc les deux séries de fonctions $(t, x) \longmapsto \sum \frac{\partial u}{\partial x}(t, x)$ et $(t, x) \longmapsto \sum \frac{\partial u}{\partial t}(t, x)$ convergent normalement sur $[\tau, +\infty[$ pour tout $\tau > 0$. On en déduit donc que u est dérivable en x et en t et ses dérivées sont continues sur $[\tau, +\infty[\times \mathbb{R}$ pour tout $\tau > 0$ et par conséquent sur $]0, +\infty[\times \mathbb{R}$. On a alors bien montré que $u \in \mathcal{C}^1(]0, +\infty[\times \mathbb{R}, \mathbb{R})$. Pour montrer que u est de classe $\mathcal{C}^2$, il suffit de réappliquer le même argument que précédemment.

(cc) BY-NC-SA

Correction de l'exercice V.5 :

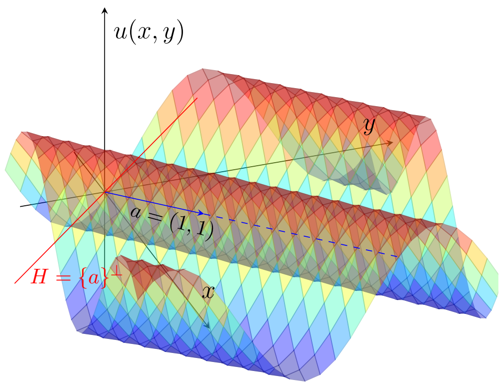
En fait, l'égalité de cet exercice peut être réécrite de la manière suivante, en posant $a = (a_1, \ldots, a_n)$, $\langle \nabla u_x,\, a\rangle = 0$. On déduit donc que localement, au voisinage de tout $x \in \mathbb{R}^n$ et pour tout $t \in \mathbb{R}$, $u(x + ta) - u(x) \simeq \langle \nabla u_x,\, ta\rangle = 0$, c'est-à-dire que u est approximativement constante dans la direction a. On va montrer qu'en fait u est constante dans la direction a, et que donc toutes ses valeurs dans $\mathbb{R}^n$ peuvent être entièrement caractérisées par ses valeurs dans l'hyperplan $H = \{a\}^\perp$. Pour cela, on va considérer un élément $x \in \mathbb{R}^n$ et montrer qu'ajouter un vecteur proportionnel à a ne change pas son image par u. Soit $x \in \mathbb{R}^n$. On pose pour tout $t \in \mathbb{R}$, $v(t) = u(x + ta)$. On a, en utilisant la règle de la chaine

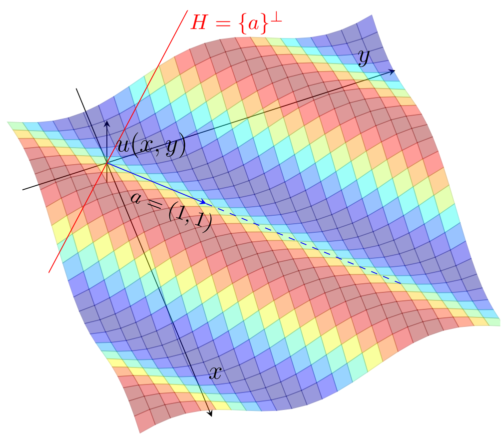
$$\forall t \in \mathbb{R},\ u'(t) = \langle \nabla u_{x+ta},\, a\rangle = 0.$$

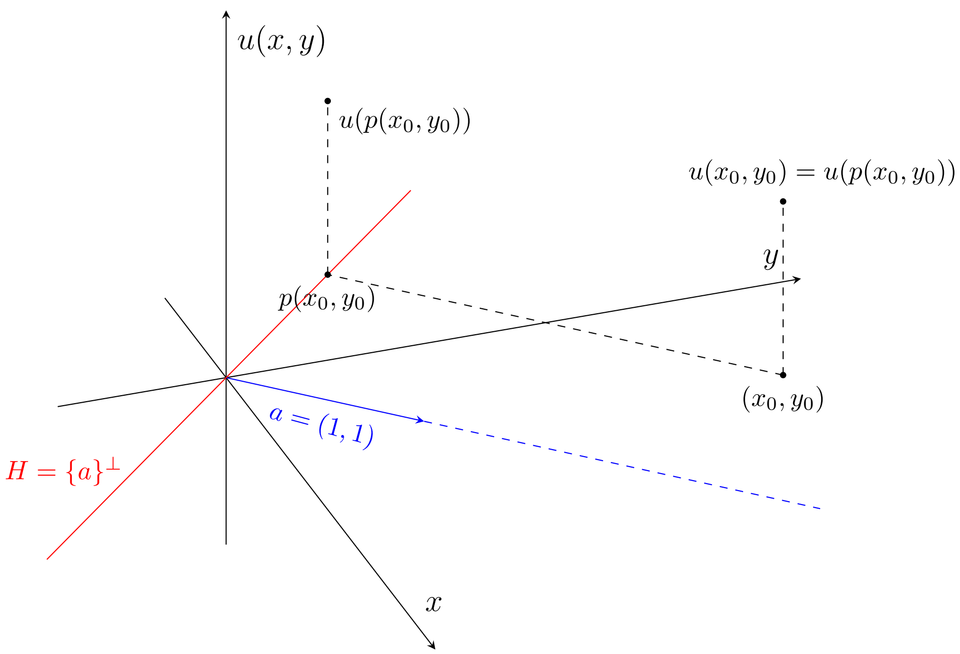
On en déduit donc que v est constante, et que u ne varie pas dans la direction de a et qu'alors si on considère $p : x \longmapsto x - \left\langle x,\, \dfrac{a}{\|a\|}\right\rangle \dfrac{a}{\|a\|}$ la projection sur $H = \{a\}^\perp$ parallèlement à $\mathrm{Vect}\{a\}$, on a pour tout $x \in \mathbb{R}^n$, $u(p(x)) = u(x)$.

Montrons à présent la réciproque. Supposons que u est constante sur la direction de a, *i.e.* pour tout $x \in \mathbb{R}^n$, $v_x : t \longmapsto u(x + ta)$ est constante. On a alors pour tout $x \in \mathbb{R}^n$, $v_x'(0) = 0$, *i.e.* $\langle \nabla u_x,\, a\rangle = 0$, ce qui bien l'égalité voulue. En conclusion, l'ensemble des fonctions u qui vérifient l'égalité sont celles qui sont constantes selon la direction de a.

Pour avoir une meilleure intuition géométrique de ce qui ce passe, observons la figure ci-dessous qui illustre l'exemple en dimension 2 où $u = (x \longmapsto \sin(x - y))$ et $a = (1, 1)$.







CC BY-NC-SA

Correction de l'exercice V.6 :
Procédons par double implication.

→ (⇒) On dérive des deux côtés l'égalité $f(\lambda x) = \lambda^r f(x)$ pour tout $\lambda > 0$, et on obtient en appliquant la règle de la chaine, pour tout $x \in O$,

$$\langle \nabla f_{\lambda x},\ x\rangle = r\lambda^{r-1} f(x) \ \textit{i.e.}\ \sum_{k=1}^{n} x_i \frac{\partial f}{\partial x_i}(\lambda x_i) = r\lambda^{r-1} f(x).$$

Ensuite, en évaluant en $\lambda = 1$, on obtient l'égalité désirée.

→ (⇐) Soit $x = (x_1, \dots, x_k) \in O$. Posons pour tout $t > 0$, $\gamma(t) = f(tx)$. Le domaine de définition de γ est ouvert car égal à $\mathbb{R}_+^*$. γ est dérivable, car f est différentiable et pour tout $t > 0$,

$$\gamma'(t) = \langle \nabla f_{tx},\ x\rangle = \sum_{k=1}^{n} x_k \frac{\partial f}{\partial x_k}(tx) = \frac{r}{t} f(tx) = \frac{r}{t}\gamma(t).$$

On a de plus

$$\begin{aligned} \gamma'(t) = \frac{r}{t}\gamma(t) &\implies e^{-r\log t}\Big(\gamma'(t) - \frac{r}{t}\gamma(t)\Big) = 0 \\ &\implies \Big(\gamma(t)e^{-r\log t}\Big)' = 0 \\ &\implies \exists C \in \mathbb{R},\ \gamma(t) = Ce^{r\log t} = Ct^r. \end{aligned}$$

En évaluant en $t = 1$, on obtient que $C = f(x)$, et finalement, on a $f(tx) = t^r f(x)$ pour tout $t > 0$ tel que $tx \in O$.

Annexe

On présente dans cette annexe un résultat intéressant de connexité pour le lecteur curieux. On rappelle que les annexes ne sont pas nécessaires pour se préparer aux concours et sont uniquement à but culturel.

Définition V.7.

Un arc (ou chemin) $\mathcal{C}^1$ par morceaux liant (ou entre) $x, y \in E$ est une application $\gamma : [0,1] \to E$ vérifiant les propriétés suivantes.

→ γ est continue.

→ $\gamma(0) = x$ et $\gamma(1) = y$ (ou $\gamma(0) = y$ et $\gamma(1) = x$).

→ Il existe $m \geq 2$ et $\sigma_1, \dots, \sigma_m \in [0,1]$ tels que $\sigma_0 = 0 < \sigma_1 < \dots < \sigma_n = 1$, et pour tout $i \in [\![1; m-1]\!]$ la restriction de γ à $]\sigma_i, \sigma_{i+1}[$ est de classe $\mathcal{C}^1$.

Si de plus, pour tout $i \in [\![1; m-1]\!]$, la restriction de γ à $]\sigma_i, \sigma_{i+1}[$ est affine, *i.e.* la restriction de γ à $]\sigma_i, \sigma_{i+1}[$ est une application de la forme $t \longmapsto u + tv$ où $u, v \in E$, alors γ est dit être un arc (ou chemin) affine par morceaux (ou polygonal).

Vocabulaire. On dit qu'un arc γ est contenu dans un sous-ensemble S de E si $\gamma([0,1]) \subset S$.

Définition V.8.

Soit S un sous-ensemble non vide de E. On dit que S est connexe par arcs $\mathcal{C}^1$ par morceaux (resp. connexe par arcs affines par morceaux) si pour tout $x, y \in S$, il existe un arc de classe $\mathcal{C}^1$ par morceaux (resp. affine par morceaux) liant x et y et contenu dans S.

BY-NC-SA

Proposition V.9.

Soit U un ouvert non vide de E. Les propriétés suivantes sont équivalentes.

1. U est connexe.
2. U est connexe par arcs C^1 par morceaux.
3. U est connexe par arcs affines par morceaux.

Démonstration.

→ (3) ⇒ (2) ⇒ (1) Ces deux implications sont évidentes.

→ (1) ⇒ (3) On rappelle qu'ici, la connexité est définie en utilisant la topologie induite sur U *i.e.* en considérant comme espace métrique (U, d_U) où d_U désigne la restriction de d, la distance induite par la norme de $\|.\|_E$ à U^2.

Notons que même si un fermé de U ne l'est pas forcément dans E, les ouverts de U sont bien des ouverts de E, étant donné que U est ouvert dans E. Pour $a \in U$ et $r > 0$, On note

$$B_E(a, r) = \{x \in E,\ d(x, a) < r\}$$

la boule ouverte de centre a et de rayon r dans E, et par $B_U(a, r) = B_E(a, r) \cap U$ la boule centrée en a de rayon r dans U.

Supposons désormais que U connexe. Soit $x \in U$. Remarquons d'abord que pour montrer que U est connexe par arcs polygonaux, il suffit de montrer qu'on peut lier x à tout autre élément $y \in U$ par un arc polygonal contenu dans U. En effet, si y et z peuvent être liés à x par des arcs γ et ψ $(\gamma(0) = y, \gamma(1) = x, \psi(0) = x, \psi(1) = z)$, alors l'arc polygonal

$$\gamma \bullet \psi : \begin{cases} [0,1] & \longrightarrow A \\ t & \longmapsto \begin{cases} \gamma(2t) & \text{si } t \in [0, 1/2] \\ \psi(1-2t) & \text{sinon} \end{cases} \end{cases}$$

lie y et z dans u. Posons F l'ensemble des $z \in U$ tel qu'on puisse lier x à z par un arc polygonal contenu dans U. On a alors,

- F est non vide. En effet, F contient x, étant donné que l'arc polygonal

$$\begin{cases} [0,1] & \longrightarrow U \\ t & \longmapsto x \end{cases}$$

  lie x à x et donc $x \in F$.

- F est un ouvert de U. Soit $y \in F$ et $r > 0$ assez petit pour que la boule $B_E(y, r)$ soit incluse dans U (possible vu que U est un ouvert de E). Soit $z \in B_E(y, r)$, l'arc polygonal

$$\begin{cases} [0,1] & \longrightarrow U \\ t & \longmapsto y + t(z - y) \end{cases}$$

  lie y à z et est inclus dans $B(y, r) \subset U$. En liant cet arc à un autre arc polygonal liant x et y (qui existe par définition de F), on obtient un arc polygonal entre x et z et alors $B(y, r) \subset F$. F est alors bien ouvert dans Ω.

- F est un fermé de U. Soit y un élément de l'adhérence de F dans U et $r > 0$ assez petit tel que $B_E(y, r)$ soit incluse dans U (possible, car U est un ouvert de E). Par définition de l'adhérence,

CC BY-NC-SA

l'intersection $F \cap B_E(y,r) \neq \varnothing$. Soit $z \in F \cap B_E(y,r)$. L'arc polygonal

$$\begin{cases} [0,1] & \longrightarrow U \\ t & \longmapsto z + t(y-z) \end{cases}$$

lie y à z dans U, et alors en liant cet arc à un autre arc polygonal liant x et y (qui existe par définition de F), on obtient un arc polygonal entre x et z ce qui donne que $y \in F$, et finalement l'adhérence de F dans Ω est bien incluse dans F.

On conclut donc par connexité de U, que $F = U$, ce qui est bien le résultat voulu.

□

* *
*

Document compilé par Omar Bennouna et révisé par Issam Tauil le 30/08/2023 pour cpge-paradise.com.
Si vous repérez une erreur, ou avez des remarques, prière de me contacter via l'adresse contact@cpge-paradise.com.

(cc) BY-NC-SA