



Département Sciences du Numérique

Calcul différentiel

Olivier Cots

2 octobre 2021

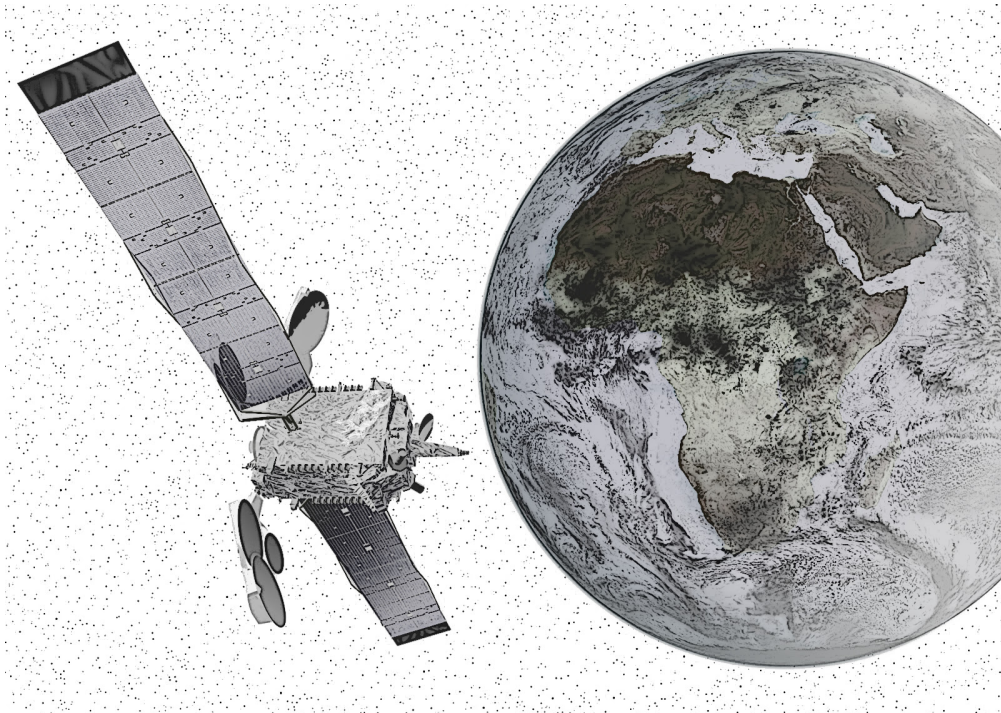


Table des matières

Chapitre 1. Applications différentiables	1
1.1 Préliminaires	1
1.2 Différentielle de Fréchet	5
1.2.1 Définition et propriétés	5
1.2.2 Trois cas fondamentaux	7
1.3 Applications de classe \mathcal{C}^1	9
1.3.1 Définitions	9
1.3.2 Propriétés et exemples fondamentaux	9
1.4 Dérivée directionnelle et dérivation	13
1.4.1 Dérivée directionnelle	13
1.4.2 Dérivation	14
1.5 Fonctions définies sur et à valeur dans un espace produit	15
1.5.1 Cas des fonctions à valeurs dans un espace produit	15
1.5.2 Cas des fonctions définies sur un espace produit	16
1.5.3 Cas général : synthèse	19
1.6 Théorèmes des accroissements finis et applications	20
1.7 Dérivabilité vs dérivée partielle et cas fondamentaux	25
1.8 Dérivabilité dans le cas fonctionnel	29
Chapitre 2. Différentielle d'ordre supérieur	33
2.1 Différentielle du second ordre	33
2.2 Différentielle d'ordre k	35
Chapitre 3. Inversion locale et équations implicites	37
3.1 Théorème du point fixe	37
3.2 Théorème d'inversion locale	39
3.3 Théorème des fonctions implicites	42
Chapitre 4. Sous-variétés différentielles de \mathbb{R}^n	45
4.1 Formes normales des submersions et immersions	45
4.2 Sous-variétés différentielles de \mathbb{R}^n et exemples	48
4.3 Sous-espaces tangents d'une sous-variété de \mathbb{R}^n	52
4.4 Discussion : sous-variétés de \mathbb{R}^n comme préimage d'une submersion	55
4.5 Optimisation en dimension finie avec contraintes d'égalité : CN1-CN2	56
Chapitre 5. Optimisation	58
Bibliographie	61

Applications différentiables

1.1	Préliminaires	1
1.2	Différentielle de Fréchet	5
1.2.1	Définition et propriétés	5
1.2.2	Trois cas fondamentaux	7
1.3	Applications de classe \mathcal{C}^1	9
1.3.1	Définitions	9
1.3.2	Propriétés et exemples fondamentaux	9
1.4	Dérivée directionnelle et dérivation	13
1.4.1	Dérivée directionnelle	13
1.4.2	Dérivation	14
1.5	Fonctions définies sur et à valeur dans un espace produit	15
1.5.1	Cas des fonctions à valeurs dans un espace produit	15
1.5.2	Cas des fonctions définies sur un espace produit	16
1.5.3	Cas général : synthèse	19
1.6	Théorèmes des accroissements finis et applications	20
1.7	Dérivabilité vs dérivée partielle et cas fondamentaux	25
1.8	Dérivabilité dans le cas fonctionnel	29

1.1 Préliminaires

Considérons deux espaces vectoriels normés¹ $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sur le corps des réels \mathbb{R} (pour simplifier, un peu) et Ω un ouvert de E contenant 0_E . Soit $g: \Omega \subset E \rightarrow F$ une application. On rappelle, pour $p \in \mathbb{N}$, la notation de Landau :

$$o(v^p) := \{g: \Omega \subset E \rightarrow F \mid \forall \varepsilon > 0, \exists \eta_\varepsilon > 0, \forall v \in \Omega: \|v\|_E \leq \eta_\varepsilon \Rightarrow \|g(v)\|_F \leq \varepsilon \|v\|_E^p\}.$$

L'usage historique et commode veut que l'on note $g \in o(v^p)$ par $g(v) = o(v^p)$, qui se lit “g est un petit o de v puissance p”.

Remarque 1.1.1. Nous utiliserons les notations F^Ω et $\mathcal{F}(\Omega, F)$ pour désigner l'ensemble des applications de Ω dans F .

Proposition 1.1.1

- i) L'ensemble $o(v^p)$ est un espace vectoriel.
- ii) $g(v) = o(v^p) \implies g(0_E) = 0_F$ et g est continue en 0_E .
- iii) $g(v) = o(v^p) \iff$ Il existe $\varepsilon \in F^\Omega$ continue en 0_E telle que $g(v) = \|v\|_E^p \varepsilon(v)$ et

1. Les espaces vectoriels normés E et F ne sont pas nécessairement de dimension finie.

$$\lim_{\|v\|_E \rightarrow 0} \|\varepsilon(v)\|_F = 0.$$

- i) Soient f, g dans $o(v^p)$ et $\lambda \in \mathbb{R}$. Soit $\varepsilon > 0$. On pose $\varepsilon_\lambda := \varepsilon/(1 + |\lambda|) > 0$. Puisque f et g sont dans $o(v^p)$, il existe $\eta_{f,\varepsilon_\lambda} > 0$ et $\eta_{g,\varepsilon_\lambda} > 0$, tels que pour tout $v \in \Omega$:

$$\|v\|_E \leq \eta_{f,\varepsilon_\lambda} \Rightarrow \|f(v)\|_F \leq \varepsilon_\lambda \|v\|_E^p \quad \text{et} \quad \|v\|_E \leq \eta_{g,\varepsilon_\lambda} \Rightarrow \|g(v)\|_F \leq \varepsilon_\lambda \|v\|_E^p.$$

On introduit $\eta_\varepsilon := \min(\eta_{f,\varepsilon_\lambda}, \eta_{g,\varepsilon_\lambda})$. Alors, pour tout $v \in \Omega$, on a :

$$\begin{aligned} \|v\|_E \leq \eta_\varepsilon &\Rightarrow \|(f + \lambda g)(v)\|_F \leq \|f(v)\|_F + |\lambda| \|g(v)\|_F \\ &\leq \varepsilon_\lambda \|v\|_E^p + |\lambda| \varepsilon_\lambda \|v\|_E^p \\ &= (1 + |\lambda|) \varepsilon_\lambda \|v\|_E^p = \varepsilon \|v\|_E^p, \end{aligned}$$

ce qui montre que $f + \lambda g \in o(v^p)$, i.e. $o(v^p)$ est un espace vectoriel.

- ii) Soit $g(v) = o(v^p)$. Soit $\varepsilon > 0$. Puisque $0 = \|0_E\|_E \leq \eta_\varepsilon$, alors $\|g(0_E)\|_F \leq \varepsilon \|0_E\|_E^p$, donc si $p > 0$, $\|g(0_E)\|_F \leq 0$ et donc $g(0_E) = 0_F$, et si $p = 0$ alors $\|g(0_E)\|_F \leq \varepsilon$ et comme ceci est vrai quelque soit $\varepsilon > 0$, on a aussi $g(0_E) = 0_F$. Pour le $\varepsilon > 0$ choisi, on pose $\varepsilon_0 := \varepsilon/\eta_\varepsilon^p > 0$. Alors, pour tout $v \in \Omega$ tel que $\|v\|_E \leq \min(\eta_{\varepsilon_0}, \eta_\varepsilon)$, on a :

$$\|g(0_E + v) - g(0_E)\|_F = \|g(v)\|_F \leq \varepsilon_0 \|v\|_E^p \leq \varepsilon_0 \eta_\varepsilon^p = \varepsilon,$$

ce qui montre que g est continue en 0_E .

- iii) Pour $g(v) = o(v^p)$, on pose sur Ω , $\varepsilon(v) := g(v)/\|v\|_E^p$ si $v \neq 0_E$ et $\varepsilon(0_E) := 0_F$. Alors, $g(v) = \|v\|_E^p \varepsilon(v)$ et puisque $g(v) = o(v^p)$, pour tout $\varepsilon_0 > 0$ et tout $v \in \Omega$ tel que $0 < \|v\|_E \leq \eta_{\varepsilon_0}$, on a $\|g(v)\|_F \leq \varepsilon_0 \|v\|_E^p$, i.e. $\|\varepsilon(v)\|_F \leq \varepsilon_0$. De même, $0 = \|\varepsilon(0_E)\|_F \leq \varepsilon_0$. Ainsi pour tout $v \in \Omega$ tel que $\|v\|_E \leq \eta_{\varepsilon_0}$, on a $\|\varepsilon(v)\|_F \leq \varepsilon_0$, i.e. $\lim_{\|v\|_E \rightarrow 0} \|\varepsilon(v)\|_F = 0$ et de plus ε est continue en 0_E . La réciproque est évidente. ■

Nous noterons $\mathcal{L}(E, F)$ l'ensemble des applications linéaires et continues de E dans F et nous munirons $\mathcal{L}(E, F)$ de la norme d'opérateur

$$\|T\|_{\mathcal{L}(E, F)} := \sup_{\|x\|_E \leq 1} \|T(x)\|_F.$$

Nous rappelons que si E est de dimension finie, alors toute application linéaire de E dans F est continue, car une application linéaire est continue si et seulement si

$$\sup_{\|x\|_E \leq 1} \|T(x)\|_F < \infty.$$

On rappelle la définition

Définition 1.1.2 – Espace de Banach

Un *espace de Banach* est un espace vectoriel normé $(E, \|\cdot\|_E)$ qui est *complet* pour la distance $d_E(x, y) := \|x - y\|_E$, i.e. toute suite de Cauchy² x_n de E converge vers un point x de E .

de telle sorte que $\mathcal{L}(E, F)$ munie de la norme d'opérateur est un espace de Banach si F l'est [4, Théorème 3.10.1]. Rappelons de même que tout sous-espace vectoriel de dimension finie d'un espace vectoriel normé est un espace de Banach. Introduisons maintenant un peu de terminologie.

Définition 1.1.3

Soient E, F deux espaces vectoriels.

- Un *morphisme* (algébrique) entre E et F est une application linéaire de E dans F . On note $L(E, F)$ ou $\text{Hom}(E, F)$ ³ l'ensemble des applications linéaires de E dans F .
- Un *endomorphisme* (algébrique) de E est une application linéaire de E dans E . On note $L(E) := L(E, E)$ l'ensemble des endomorphismes de E .
- Un *isomorphisme* (algébrique) entre E et F est une application de E dans F , linéaire et bijective. C'est donc un morphisme bijectif. On note $\text{Isom}(E, F)$ l'ensemble des isomorphismes de E sur F .
- Un *automorphisme* (algébrique) de E est une application dans $\text{GL}(E) := \text{Isom}(E, E)$. C'est un endomorphisme bijectif. L'ensemble des automorphismes est aussi appelé le *groupe linéaire*.
- Si $F = \mathbb{R}$, on parle de *forme linéaire*. On note $E^* := L(E, \mathbb{R})$, l'ensemble des formes linéaires de E sur \mathbb{R} . Cet espace est appelé *l'espace dual algébrique* de E .

Donnons maintenant la version “topologique” de ces définitions.

Définition 1.1.4

Soient E, F deux espaces vectoriels **topologiques** (par exemple, deux espaces vectoriels normés).

- Un *morphisme* (topologique) entre E et F est une application linéaire **continue** de E dans F . On note $\mathcal{L}(E, F)$ ou $L_c(E, F)$ l'ensemble des applications linéaires continues de E dans F .
- Un *endomorphisme* (topologique) de E est une application linéaire **continue** de E dans E . On note $\mathcal{L}(E) := \mathcal{L}(E, E)$ ou $L_c(E)$ l'ensemble des endomorphismes de E .
- Un *isomorphisme* (topologique) entre E et F est une application de E dans F , linéaire, bijective, **continue**, dont l'application réciproque (ou inverse) est **continue**. C'est donc un *homéomorphisme* (*i.e.* une application bijective continue d'inverse continue) linéaire. On note $\text{Isom}_c(E, F)$ l'ensemble des isomorphismes de E sur F .
- Un *automorphisme* (topologique) de E est une application dans $\text{GL}_c(E) := \text{Isom}_c(E, E)$. L'ensemble des automorphismes peut être aussi noté $\mathcal{GL}(E)$.
- Si $F = \mathbb{R}$, on parle de *forme linéaire continue*. On note $E' := \mathcal{L}(E, \mathbb{R})$, l'ensemble des formes linéaires **continue** de E sur \mathbb{R} . Cet espace est appelé *l'espace dual topologique* de E .

Faisons quelques commentaires sur ces notions. Considérons les deux espaces vectoriels nor-

2. Une suite x_n de E vérifiant $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n, m \geq N, \|x_n - x_m\|_E \leq \varepsilon$, est dite de Cauchy.

3. La notation $\text{Hom}(E, F)$ vient du mot *homomorphisme* qui veut dire morphisme, et à ne pas confondre avec homéomorphisme.

més $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ du début. Si E est de dimension finie, alors $\mathcal{L}(E, F) = L(E, F)$, puisque alors toute application linéaire est continue. En dimension infinie, on a

$$\mathcal{L}(E, F) \subset L(E, F)$$

et cette inclusion est toujours stricte, cf. l'exemple suivant extrait de [4, Exemple 3.3.1].

Exemple 1.1.1. Sur un espace normé de dimension infinie, il existe toujours des formes linéaires non continues. En effet, soit $(e_i)_{i \in I}$ une base⁴ de E ; on peut supposer $\|e_i\|_E = 1$. Soit $(a_i)_{i \in I}$ une famille non bornée de \mathbb{R} (il en existe, I étant infini) et soit T la forme linéaire vérifiant $T(e_i) = a_i$. Il ne peut alors exister de constante $C \geq 0$ telle que $|T(e_i)| = |a_i| \leq C$ pour tout $i \in I$ et T n'est donc pas continue. \square

D'après les définitions précédentes, il est clair que T est un isomorphisme topologique de E dans F si et seulement si T est une application linéaire surjective et s'il existe une constante $C > 0$ telle que

$$\forall x \in E, \quad C^{-1}\|x\|_E \leq \|T(x)\|_F \leq C\|x\|_E.$$

Ainsi, si T est un isomorphisme algébrique *isométrique*, i.e. $\|T(x)\|_F = \|x\|_E$, alors T est un isomorphisme topologique (isométrique).

Remarque 1.1.2. Le qualificatif algébrique ou topologique est souvent sous-entendu et dépend du contexte. On dit que deux espaces E et F sont *isomorphes* s'il existe un isomorphisme de E dans F . Deux espaces vectoriels isomorphes ont donc essentiellement la même structure algébrique, tandis que deux espaces vectoriels topologiques isomorphes ont la même structure algébrique mais aussi la même structure topologique. L'intérêt des isomorphismes réside dans le fait que de nombreuses propriétés importantes sont invariantes par isomorphisme.

Rappelons d'autre part la différence entre application inversible et application bijective dans le cadre "topologique". On dit qu'une application linéaire continue $A \in \mathcal{L}(E, F)$ est inversible, resp. bijective, si il existe $B \in \mathcal{L}(F, E)$, resp. $B \in F^E$, telle que

$$A \circ B = \text{Id}_F \quad \text{et} \quad B \circ A = \text{Id}_E.$$

L'application B si elle existe est unique. On note en général $A^{-1} := B$ et il est facile de montrer que A^{-1} est automatiquement linéaire. Ainsi, pour montrer qu'une application linéaire continue bijective A est inversible, il reste simplement à montrer que A^{-1} (au sens de la théorie des ensembles) est elle-même continue.

Remarque 1.1.3. L'ensemble $\text{Isom}(E, F)$, resp. $\text{Isom}_c(E, F)$, est exactement l'ensemble des applications linéaires, resp. linéaires continues, inversibles.

Remarque 1.1.4. Il est nécessaire d'avoir les deux identités précédentes. On peut par exemple prendre $E = F = \mathbb{R}^{\mathbb{N}}$ et définir

$$\begin{aligned} A: \quad E &\longrightarrow E \\ (u_n)_{n \in \mathbb{N}} &\longmapsto A(u_n)_{n \in \mathbb{N}} := (v_n)_{n \in \mathbb{N}}, \end{aligned}$$

où $v_n := u_{n-1}$ pour $n \geq 1$ et $v_0 := 0$. On définit aussi

$$\begin{aligned} B: \quad E &\longrightarrow E \\ (v_n)_{n \in \mathbb{N}} &\longmapsto B(v_n)_{n \in \mathbb{N}} := (v_{n+1})_{n \in \mathbb{N}}, \end{aligned}$$

4. Voir [4, Section 1.6, page 29] pour la définition de base et le résultat d'existence.

de telle sorte que $B \circ A = \text{Id}_E$ mais A n'est pas inversible (elle n'est pas surjective donc pas bijective).

En vertu du résultat important (mais difficile) suivant, le théorème de Banach [4, Corollaire 3.11.3], on sait qu'une application linéaire continue bijective A est automatiquement inversible, si E et F sont deux espaces de Banach. Dans ce cas, il n'y a donc en pratique pas besoin de montrer la continuité de A^{-1} , elle est automatique.

Théorème 1.1.5 – de Banach

Soient E et F deux espaces de Banach, alors, toute bijection linéaire et continue de E sur F est un isomorphisme (topologique).

Ce résultat est un corollaire du théorème fondamental de l'*application ouverte* [4, Théorème 3.11.1] : si $f \in \mathcal{L}(E, F)$ (E, F deux Banach) est surjective, alors f est ouverte, c'est-à-dire que l'image par f de tout ouvert de E est un ouvert de F . Avec nos définitions, on peut réécrire le résultat du théorème 1.1.5 sous la forme :

$$T \in \text{Isom}(E, F) \text{ et continue} \iff T \in \text{Isom}_c(E, F),$$

ou de manière équivalente

$$T \in \mathcal{L}(E, F) \text{ et bijective} \iff T \in \text{Isom}_c(E, F),$$

si E et F sont deux espaces de Banach.

1.2 Différentielle de Fréchet

1.2.1 Définition et propriétés

Définition 1.2.1 – Différentielle de Fréchet

Soient une application $f: U \subset E \rightarrow F$, U ouvert et un point $x \in U$. On dit que f est *différentiable* (ou *dérivable*) au point x s'il existe une application $T_{f,x} \in \mathcal{L}(E, F)$ telle que pour tout $v \in E$ vérifiant $x + v \in U$, on ait

$$f(x + v) - f(x) - T_{f,x}(v) = o(v).$$

L'application linéaire continue $T_{f,x}$, si elle existe est unique (voir ci-après), et on l'appelle la *différentielle (de Fréchet)* de f au point x .

Remarque 1.2.1. On utilisera les notations

$$df(x) \cdot v := T_{f,x}(v) \quad \text{ou} \quad f'(x) \cdot v := T_{f,x}(v)$$

suivant le contexte et cette expression se lit donc “la différentielle de f au point x appliquée au vecteur v ”.

Proposition 1.2.2

Si f est différentiable au point x alors $f'(x) := T_{f,x}$ est unique et f reste différentiable en x avec $f'(x)$ inchangée si l'on remplace les normes de E et/ou F par des normes équivalentes.

► Soit $x \in U$ (ouvert) $\subset E$. Supposons que f est différentiable en x . Supposons qu'il existe deux différentielles de f au point x notées T_1 et T_2 . Puisque U est ouvert, il existe une boule ouverte $B(0, r_x)$, $r_x > 0$, telle que pour tout $v \in B(0, r_x)$, $x + v \in U$. Par hypothèse, on a pour tout $v \in B(0, r_x)$:

$$f(x + v) = f(x) + T_1(v) + o(v) = f(x) + T_2(v) + o(v).$$

Donc $(T_1 - T_2)(v) =: T(v) = o(v)$. Soit maintenant $\varepsilon > 0$. Par définition du $o(v)$, il existe $r_\varepsilon > 0$ tel que pour tout $v \in B(0, r_\varepsilon)$:

$$\|T(v)\|_F \leq \|v\|_E \varepsilon.$$

Donc pour tout $v \in B(0, r^*)$, $r^* := \min(r_x, r_\varepsilon)$, on a

$$\|T(v)\|_F \leq \|v\|_E \varepsilon \leq r^* \varepsilon.$$

Ainsi, par la linéarité de T , on a que

$$\|T\|_{\mathcal{L}(E,F)} = \sup_{\|v\|_E \leq 1} \|T(v)\|_F = \frac{1}{r^*} \sup_{\|v\|_E \leq r^*} \|T(v)\|_F \leq \frac{1}{r^*} r^* \varepsilon = \varepsilon,$$

et puisque ceci est vrai pour tout $\varepsilon > 0$, finalement $\|T\|_{\mathcal{L}(E,F)} = 0$ donc $T = 0$, ce qui prouve l'unicité de la différentielle de f au point x . La fin de la preuve est laissée en exercice. ■

Proposition 1.2.3

Si f est différentiable en un point $x \in U \subset E$ alors f est continue en x .

► Puisque f est différentiable en $x \in U \subset E$, on a $f(x + v) - f(x) = f'(x) \cdot v + o(v)$, pour $v \in E$ vérifiant $x + v \in U$. Puisque U est ouvert, on a donc sur un voisinage ouvert de x :

$$\begin{aligned} \|f(x + v) - f(x)\|_F &= \|f'(x) \cdot v + o(v)\|_F \\ &\leq \|f'(x) \cdot v\|_F + \|o(v)\|_F \\ &\leq (\|f'(x)\|_{\mathcal{L}(E,F)} + \|\varepsilon(v)\|_F) \|v\|_E, \end{aligned}$$

ce qui montre que f est continue en x . ■

Il nous sera utile de noter que l'opération de dérivation est une opération linéaire et que l'ensemble des applications différentiables en un point forme un espace vectoriel. Plus précisément, on a le résultat suivant :

Proposition 1.2.4

L'ensemble des applications de U dans F différentiables au point $x \in U \subset E$, noté \mathcal{D}_x , est un sous-espace vectoriel de l'ensemble des applications continues en x et l'opérateur D_x défini par

$$\begin{aligned} D_x: \mathcal{D}_x &\longrightarrow \mathcal{L}(E, F) \\ f &\longmapsto D_x(f) := f'(x) \end{aligned}$$

est linéaire.

Remarque 1.2.2. On emploie fréquemment le terme d'opérateur pour mentionner une application linéaire.

► Le fait que \mathcal{D}_x soit un sous-espace vectoriel vient directement du fait que $o(v)$ en est un, cf. proposition 1.1.1. De plus, \mathcal{D}_x est inclus dans l'espace des applications continues en x d'après la proposition 1.2.3. Enfin, la linéarité de D_x vient du fait que $\mathcal{L}(E, F)$ est un sous-espace vectoriel de F^E . ■

1.2.2 Trois cas fondamentaux

Exemple 1.2.1 (Fonction réelle de la variable réelle). Supposons que $E = F = \mathbb{R}$. On rappelle qu'une fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ est dérivable en un point $x \in \mathbb{R}$ s'il existe un nombre réel a tel que :

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = a \quad \text{ou écrit autrement} \quad \lim_{y \rightarrow x} \frac{f(x) - f(y)}{x - y} = a.$$

Ce nombre a est habituellement noté $f'(x)$ et la notion de dérivabilité est équivalente à l'existence du développement limité

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

Dans ce cas différentiable équivaut à dérivable et la différentielle est l'application linéaire $df(x): h \mapsto hf'(x)$ de \mathbb{R} dans \mathbb{R} . La dérivée est simplement le coefficient directeur obtenu à partir de la différentielle par $f'(x) = df(x) \cdot 1$. □

Exemple 1.2.2 (Fonction de la variable réelle). Supposons que $E = \mathbb{R}$. Dans ce cas, $df(x) \in \mathcal{L}(\mathbb{R}, F)$ et l'application $\varphi: T \mapsto T(1)$ est un isomorphisme algébrique (*i.e.* une application linéaire bijective) isométrique (c'est donc un isomorphisme topologique aussi) de $\mathcal{L}(\mathbb{R}, F)$ dans F . On peut alors identifier l'application linéaire continue $df(x)$ à un vecteur y de F . Ce vecteur y étant noté $f'(x)$. Voyons tout d'abord comment introduire $f'(x)$ puis montrons que φ est bien un isomorphisme isométrique.

Comme dans l'exemple précédent, la division par h est possible. La différentiabilité de f en x équivaut à l'existence de la limite

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} =: f'(x) \in F$$

appelé *vecteur dérivé*, et la différentielle $df(x)$ est l'application linéaire $df(x): h \mapsto hf'(x)$. Ainsi, on a de nouveau $f'(x) = df(x) \cdot 1$.

Montrons que φ est bien un isomorphisme isométrique. Soit donc

$$\begin{aligned} \varphi: \mathcal{L}(\mathbb{R}, F) &\longrightarrow F \\ T &\longmapsto \varphi(T) := T(1). \end{aligned}$$

L'ensemble des applications de \mathbb{R} dans F , noté $F^{\mathbb{R}}$, est un espace vectoriel. Il est clair que l'ensemble $\mathcal{L}(\mathbb{R}, F) \subset F^{\mathbb{R}}$ est un sous-espace vectoriel de $F^{\mathbb{R}}$ et on montre alors facilement que φ est linéaire. Soit $T \in \text{Ker } \varphi$. Pour tout $\lambda \in \mathbb{R}$, $T(\lambda) = \lambda T(1) = \lambda \varphi(T) = 0$ donc $T = 0$. Donc φ est injective. Soit maintenant $y \in F$. Posons $T_y: \lambda \mapsto T_y(\lambda) = \lambda y$ définie sur \mathbb{R} . Il est clair que $T_y \in \mathcal{L}(\mathbb{R}, F)$ (elle est continue car pour tout $\lambda \in \mathbb{R}$, $\|T_y(\lambda)\|_F = \|\lambda y\|_F \leq C|\lambda|$, $C := \|y\|_F$) donc φ est surjective ($\varphi(T_y) = T_y(1) = y$). En conclusion, φ est bijective. Montrons enfin que φ est isométrique. D'un côté, on a

$$\|T\|_{\mathcal{L}(\mathbb{R}, F)} = \sup_{|\lambda| \leq 1} \|T(\lambda)\|_F$$

et de l'autre, $\|\varphi(T)\|_F = \|T(1)\|_F$. Donc finalement, puisque T est définie sur \mathbb{R} , on a

$$\|T\|_{\mathcal{L}(\mathbb{R}, F)} = \sup_{|\lambda| \leq 1} \|T(\lambda)\|_F = \sup_{|\lambda| \leq 1} |\lambda| \|T(1)\|_F = \|T(1)\|_F = \|\varphi(T)\|_F$$

et φ est bien un isomorphisme isométrique.

Remarque 1.2.3. Établir un isomorphisme entre deux espaces vectoriels permet d'identifier ces deux espaces. Si l'on se restreint aux propriétés invariantes par isomorphisme, ces deux espaces sont indiscernables, donc deux représentations du même objet. Aucun des deux n'est plus légitime que l'autre et on choisira donc de travailler avec l'espace qui nous arrange. \square

Exemple 1.2.3 (Fonction définie sur un Hilbert et à valeurs dans \mathbb{R}). Supposons que E soit un espace de Hilbert muni du produit scalaire $(\cdot | \cdot)_E$ (i.e. un espace préhilbertien complet pour la distance $d_E(x, y) := \sqrt{(x - y | x - y)_E}$) et supposons que $F = \mathbb{R}$. Dans ce cas, la différentielle $f'(x) \in \mathcal{L}(E, \mathbb{R}) = E'$, i.e. appartient au dual topologique de E (c-a-d à l'ensemble des formes linéaires continues sur E) et on sait alors d'après le théorème de représentation de Riesz qu'il existe un unique $u_x \in E$ tel que pour tout $v \in E$, on ait

$$f'(x) \cdot v = (u_x | v)_E =: (\nabla f(x) | v)_E.$$

Ce vecteur u_x est généralement noté $\text{grad } f(x)$ ou $\nabla f(x)$ et est appelé le *gradient* de f en x . \square

Rappelons le théorème de représentation des formes linéaires dans les Hilbert (isomorphisme entre E et E').

Théorème 1.2.5 – de Riesz

Soit $(E, (\cdot | \cdot)_E)$ un espace de Hilbert réel. Notons $\|\cdot\|_E := \sqrt{(\cdot | \cdot)_E}$ la norme associée. Alors, il existe un isomorphisme isométrique entre E et son dual topologique $E' = \mathcal{L}(E, \mathbb{R})$. Plus précisément on a :

- i) $\forall y \in E$, la forme linéaire $T_y: E \rightarrow \mathbb{R}$ définie par $T_y(x) := (x | y)_E$ est telle que $\|T_y\|_{E'} = \|y\|_E < \infty$ et donc $T_y \in E'$.
- ii) $\forall T \in E'$, $\exists! y_T \in E$, tel que $\forall x \in E$ on a $T(x) = (y_T | x)_E$ et $\|y_T\|_E = \|T\|_{E'} < \infty$.

L'application $\varphi: E \rightarrow E'$ définie par $\varphi(y) := T_y = (\cdot | y)_E$ est l'isomorphisme en question.

1.3 Applications de classe \mathcal{C}^1

1.3.1 Définitions

Définition 1.3.1 – Application dérivée

Si f est différentiable en tout point de $U \subset E$ alors

$$f': U \rightarrow \mathcal{L}(E, F)$$

s'appelle *l'application dérivée* de f sur U .

Remarque 1.3.1. Ne pas confondre l'application dérivée avec sa valeur en un point. La valeur en un point, étant elle-même une application (linéaire continue).

Définition 1.3.2 – Application de classe \mathcal{C}^1

- On dit que f est *différentiable* dans U si f l'est en tout point de U .
- On dit que f est de classe \mathcal{C}^1 (ou aussi *continuellement différentiable*) en x si elle est différentiable sur un voisinage ouvert V de x dans U et si $f': V \rightarrow \mathcal{L}(E, F)$ est continue en x .
- On dit que f est de classe \mathcal{C}^1 (ou aussi *continuellement différentiable*) dans U si f est continuellement différentiable en tout point de U , ou de manière équivalente, si f est différentiable en tout point de U et si $f': U \rightarrow \mathcal{L}(E, F)$ est continue.

Remarque 1.3.2. Dire que f est \mathcal{C}^1 dans U , c'est avoir que $f' \in \mathcal{C}^0(U, \mathcal{L}(E, F))$. On a donc besoin de topologies sur les espaces de départ et d'arrivée. Pas de problème, on munit $U \subset E$ de la topologie induite par la norme $\|\cdot\|_E$ (ce que l'on faisait déjà pour f) et $\mathcal{L}(E, F)$ est bien munie d'une topologie définie par la norme $\|\cdot\|_{\mathcal{L}(E, F)}$.

1.3.2 Propriétés et exemples fondamentaux

La propriété suivante est fondamentale, elle permet de calculer de nombreuses dérivées.

Théorème 1.3.3 – de dérivation des applications composées

Soient E, F et G trois espaces vectoriels normés sur \mathbb{R} , U un ouvert de E , $V \supset f(U)$ un ouvert de F , $x \in U$ et deux applications $f: U \rightarrow F$ et $g: V \rightarrow G$. On a alors :

- i) Si f est dérivable en x et si g est dérivable en $f(x)$ alors $g \circ f$ est dérivable en x et

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x).$$

- ii) Si f est dérivable sur U et si g est dérivable sur V (resp. f et g de classe \mathcal{C}^1) alors $g \circ f$ est dérivable sur U (resp. de classe \mathcal{C}^1).

► Pour $v \in E$ suffisamment petit, on peut écrire $f(x+v) = f(x) + w$, avec $w := f'(x) \cdot v + o(v) = f'(x) \cdot v + \|v\| \varepsilon(v)$ (on omet les indices des normes par commodité d'écriture). Ainsi,

$$\begin{aligned} (g \circ f)(x+v) &= g(f(x+v)) = g(f(x) + w) = (g \circ f)(x) + g'(f(x)) \cdot w + o(w) \\ &= (g \circ f)(x) + g'(f(x)) \cdot (f'(x) \cdot v) + g'(f(x)) \cdot (\|v\| \varepsilon(v)) + o(w) \\ &= (g \circ f)(x) + (g'(f(x)) \circ f'(x)) \cdot v + o(v) + o(w). \end{aligned}$$

Pour conclure, il ne reste plus qu'à montrer que le $o(w)$ est aussi un $o(v)$. On peut écrire le $o(w)$ sous la forme $\varphi(w) := \|w\| \varepsilon_1(w)$ tel que $\|\varepsilon_1(w)\| \rightarrow 0$ quand $\|w\| \rightarrow 0$. Puisque

$$\|w\| = \|f'(x) \cdot v + \|v\| \varepsilon(v)\| \leq (\|f'(x)\| + \|\varepsilon(v)\|) \|v\|,$$

on a $\|w\| \rightarrow 0$ quand $\|v\| \rightarrow 0$. Donc

$$\frac{\|\varphi(w)\|}{\|v\|} \leq (\|f'(x)\| + \|\varepsilon(v)\|) \|\varepsilon_1(w)\| \rightarrow 0 \quad \text{quand} \quad \|v\| \rightarrow 0,$$

et le point $i)$ est démontré.

* Montrons le point $ii)$. Pour la dérivabilité, il n'y a rien à faire. Montrons que si f et g sont \mathcal{C}^1 , alors la composée $g \circ f$ l'est aussi. Introduisons la notation $h := g \circ f$. L'application dérivée est définie par

$$\begin{array}{ll} h': U \longrightarrow \mathcal{L}(E, G) & \text{Notons } \psi: U \longrightarrow \mathcal{L}(F, G) \times \mathcal{L}(E, F) \\ x \longmapsto h'(x) = g'(f(x)) \circ f'(x). & x \longmapsto \psi(x) := (g'(f(x)), f'(x)) \end{array}$$

qui à x associe un couple d'applications linéaires continues. Les applications composantes de $\psi = (\psi_1, \psi_2)$ sont telles que $\psi_1 = g' \circ f \in \mathcal{C}^0(U, \mathcal{L}(F, G))$ car $f \in \mathcal{C}^0(U, F)$, car f de classe \mathcal{C}^1 et $g' \in \mathcal{C}^0(V, \mathcal{L}(F, G))$, car g de classe \mathcal{C}^1 . Chaque application composante est continue donc $\psi \in \mathcal{C}^0(U, \mathcal{L}(F, G) \times \mathcal{L}(E, F))$. Introduisons de plus

$$\begin{array}{ll} \varphi: \mathcal{L}(F, G) \times \mathcal{L}(E, F) \longrightarrow \mathcal{L}(E, G) \\ (u, v) \longmapsto \varphi(u, v) := u \circ v \end{array}$$

de telle sorte que $h' = \varphi \circ \psi$. L'application φ est le produit de composition d'applications linéaires continues donc φ est une application bilinéaire continue (cf. $\|u \circ v\| \leq \|u\| \|v\|$). En conclusion, $h' \in \mathcal{C}^0(U, \mathcal{L}(E, G))$ est le point $ii)$ est démontré. ■



Exercice 1.3.1 (Cas fondamentaux). Montrer que toute application constante, linéaire continue et affine continue est de classe \mathcal{C}^1 . Donner leurs applications dérivées. □

▷ **Correction.** Prenons les cas, les uns après les autres.

- i) Soit une application constante $f: U \rightarrow F$ définie par $f(x) := y_0$, avec $U \subset E$ ouvert et $y_0 \in F$. Alors, pour tout $v \in E$ tel que $x+v \in U$, on a $f(x+v) - f(x) = y_0 - y_0 = 0_F$. L'application f est donc différentiable en x et son application dérivée est donnée par

$$\begin{array}{ll} f': U \longrightarrow \mathcal{L}(E, F) \\ x \longmapsto f'(x) = 0_{\mathcal{L}(E, F)}. \end{array}$$

L'application dérivée est constante donc continue et f est donc de classe \mathcal{C}^1 sur U .

- ii) Soit $T \in \mathcal{L}(E, F)$. On a $T(x + v) - (T(x) + T(v)) = 0 = o(v)$, $T \in \mathcal{L}(E, F)$, donc l'application dérivée est donnée par :

$$\begin{aligned} T' : E &\longrightarrow \mathcal{L}(E, F) \\ x &\longmapsto T'(x) = T. \end{aligned}$$

L'application dérivée est constante donc continue et T est donc de classe \mathcal{C}^1 sur E .

- iii) Soit une application affine de E dans F définie par $A(x) := T(x) + y_0$, avec $T \in \mathcal{L}(E, F)$ et $y_0 \in F$. On a alors

$$A(x + v) - (A(x) + T(v)) = T(x + v) + y_0 - (T(x) + y_0 + T(v)) = 0 = o(v),$$

et $T \in \mathcal{L}(E, F)$, donc l'application dérivée est donnée par :

$$\begin{aligned} A' : E &\longrightarrow \mathcal{L}(E, F) \\ x &\longmapsto A'(x) = T. \end{aligned}$$

L'application dérivée vérifie $A' = T'$, elle est constante donc continue et A est donc de classe \mathcal{C}^1 sur E . □

Exemple 1.3.2 (Application inversion). On suppose que E et F sont deux espaces de Banach et que $\text{Isom}_c(E, F)$ est non vide. D'après [4, Théorème 3.19.8 page 400], l'ensemble $\text{Isom}_c(E, F)$ est ouvert dans $\mathcal{L}(E, F)$. Introduisons l'*application inversion* :

$$\begin{aligned} \text{inv} : \text{Isom}_c(E, F) &\longrightarrow \text{Isom}_c(F, E) \\ T &\longmapsto \text{inv}(T) := T^{-1}. \end{aligned}$$

Toujours d'après le même résultat, l'application inv est un homéomorphisme de $\text{Isom}_c(E, F)$ dans $\text{Isom}_c(F, E)$. Montrons que inv est de classe \mathcal{C}^1 .

* Nous devons dans un premier temps nous ramener à $\text{Isom}_c(E, E) = \text{GL}_c(E)$. Pour cela, considérons $B \in \text{Isom}_c(F, E)$ et introduisons l'isomorphisme (à vérifier) :

$$\begin{aligned} \varphi : \text{Isom}_c(E, F) &\longrightarrow \text{GL}_c(E) \\ T &\longmapsto \varphi(T) := B \circ T. \end{aligned}$$

Ajoutons la notation $\text{inv}_{E,F}$ pour représenter l'application inversion sur $\text{Isom}_c(E, F)$. Avec cette notation, $\text{inv} = \text{inv}_{E,F}$ et l'application inversion sur $\text{GL}_c(E)$ s'écrit $\text{inv}_E := \text{inv}_{E,E}$. Il apparaît alors que :

$$\text{inv}(T) = \varphi(T)^{-1} \circ B = \text{inv}_E(\varphi(T)) \circ B. \quad (1.1)$$

Ainsi, par le théorème 1.3.3, il est clair que inv est de classe \mathcal{C}^1 si et seulement si inv_E l'est. En effet, φ est linéaire continue donc \mathcal{C}^1 tout comme l'application de composition à gauche $A \mapsto A \circ B$, avec $A \in \text{GL}_c(E)$. La vérification de ceci est laissé au lecteur.

* Montrons donc que inv_E est de classe \mathcal{C}^1 . Considérons T, H dans $\text{GL}_c(E)$ et remarquons tout d'abord que

$$T + H = T \circ (\text{Id}_E + \text{inv}_E(T) \circ H) =: T(I + T^{-1}H),$$

où dans le dernier terme de ces égalités, nous utilisons les notations T^{-1} et I qui sont plus concises que $\text{inv}_E(T)$ et Id_E et où nous ne notons pas le symbole de composition pour avoir une

notation semblable à l'écriture matricielle. D'après [4, Proposition 3.19.6], si $\|T^{-1}H\|_{\mathcal{L}(E)} < 1$, alors $I - T^{-1}H$ est inversible et

$$K := (I - T^{-1}H)^{-1} = I - T^{-1}H + (T^{-1}H)^2 + \cdots + (-1)^n (T^{-1}H)^n + \cdots.$$

Ainsi, avec ces notations : $T + H = TK^{-1}$. En combinant cela avec le fait que

$$K = I - T^{-1}H + (T^{-1}H)^2 K,$$

nous obtenons, pour H suffisamment petit,

$$\begin{aligned} (T + H)^{-1} &= KT^{-1} = (I - T^{-1}H + (T^{-1}H)^2 K)T^{-1} \\ &= T - T^{-1}HT^{-1} + (T^{-1}H)^2 KT^{-1} = T - T^{-1}HT^{-1} + o(H). \end{aligned}$$

Autrement dit,

$$\text{inv}_E(T + H) = \text{inv}_E(T) - \text{inv}_E(T) \circ H \circ \text{inv}_E(T) + o(H).$$

Puisque $H \mapsto -\text{inv}_E(T) \circ H \circ \text{inv}_E(T)$ est linéaire continue (de norme inférieure à $2\|T^{-1}\|_{\mathcal{L}(E)}$), on a inv_E dérivable en T , donc sur tout $\text{GL}_c(E)$ (car T quelconque). Sa dérivée est donnée par

$$\text{inv}'_E(T) \cdot H = -\text{inv}_E(T) \circ H \circ \text{inv}_E(T).$$

* Montrons maintenant que $\text{inv}'_E : \text{GL}_c(E) \rightarrow \mathcal{L}(\text{GL}_c(E))$ est continue. Définissons pour cela $B : \text{GL}_c(E) \times \text{GL}_c(E) \rightarrow \mathcal{L}(\text{GL}_c(E))$ par $B(M, N) \cdot H := -M \circ H \circ N$. Ainsi définie, l'application B est une application bilinéaire sur $\text{GL}_c(E) \times \text{GL}_c(E)$. On munit cet espace de la norme

$$\|B\| := \sup \left\{ \|B(M, N)\|_{\mathcal{L}(\text{GL}_c(E))} \mid \|M\|_{\mathcal{L}(E)} \leq 1 \text{ et } \|N\|_{\mathcal{L}(E)} \leq 1 \right\}.$$

À l'aide de B , on peut écrire $\text{inv}'_E(T) = B(\text{inv}_E(T), \text{inv}_E(T))$ et il est alors clair que inv'_E est continue si B l'est. Puisque

$$\|B(M, N)\|_{\mathcal{L}(\text{GL}_c(E))} = \sup_{\|H\|_{\mathcal{L}(E)} \leq 1} \|B(M, N) \cdot H\|_{\mathcal{L}(E)} = \sup_{\|H\|_{\mathcal{L}(E)} \leq 1} \|M \circ H \circ N\|_{\mathcal{L}(E)},$$

et puisque

$$\|M \circ H \circ N\|_{\mathcal{L}(E)} \leq \|M\|_{\mathcal{L}(E)} \|H\|_{\mathcal{L}(E)} \|N\|_{\mathcal{L}(E)},$$

on obtient $\|B(M, N)\|_{\mathcal{L}(\text{GL}_c(E))} \leq \|M\|_{\mathcal{L}(E)} \|N\|_{\mathcal{L}(E)}$ et donc finalement $\|B\| \leq 1$. Nous avons donc montré que B est continue donc inv'_E aussi, donc inv_E est de classe \mathcal{C}^1 . Nous pouvons alors conclure que $\text{inv} = \text{inv}_{E,F}$ est elle-même de classe \mathcal{C}^1 .

* Maintenant que nous savons que inv est dérivable (même \mathcal{C}^1), nous pouvons retrouver facilement sa dérivée. En effet, d'après le corollaire 1.7.2, sachant que l'on a pour tout $T \in \text{Isom}_c(E, F)$, la relation $\Psi(T) := \text{inv}(T) \circ T = \text{Id}_E$, on obtient

$$\Psi'(T) \cdot H = \text{inv}'(T) \cdot H \circ T + \text{inv}(T) \circ H = 0_{\mathcal{L}(E)},$$

que l'on peut transformer en : $\text{inv}'(T) \cdot H = -\text{inv}(T) \circ H \circ \text{inv}(T)$. Il est à noter que l'on retrouve la même expression à partir de la relation (1.1).

□

Remarque 1.3.3. Si $E = F = \mathbb{R}^n$, alors inv est définie sur $\text{GL}_n(\mathbb{R})$, l'ensemble des matrices inversibles. On a la relation plus familière, pour $A \in \text{GL}_n(\mathbb{R})$,

$$\text{inv}'(A) \cdot H = -A^{-1}HA^{-1}.$$

1.4 Dérivée directionnelle et dérivation

1.4.1 Dérivée directionnelle

Définition 1.4.1 – Dérivée directionnelle

Soient une application $f: U \subset E \rightarrow F$, U ouvert, un point $x \in U$ et un vecteur $v \in E$. On dit que f admet une *dérivée directionnelle* au point x dans la direction v si l'application

$$\begin{aligned} \varphi_{x,v}: V \subset \mathcal{V}_{\mathbb{R}}(0) &\longrightarrow F \\ t &\longmapsto \varphi_{x,v}(t) := f(x + tv) \end{aligned}$$

est dérivable en $t = 0$. Dans ce cas, on notera la dérivée directionnelle

$$D_v f(x) := \varphi'_{x,v}(0) = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} \in F,$$

cf. l'exemple fondamental 1.2.2, section 1.2.2.

Remarque 1.4.1. $\mathcal{V}_{\mathbb{R}}(0)$ est l'ensemble des voisinages de 0 dans \mathbb{R} . Le voisinage V de la définition ci-dessus existe toujours car $x \in U$ et U ouvert. La dérivée directionnelle $D_v f(x)$ s'appelle aussi la *dérivée partielle en x suivant le vecteur v* .

Proposition 1.4.2

Si f est dérivable en x alors pour tout $v \in E$, la dérivée directionnelle $D_v f(x)$ existe et

$$D_v f(x) = f'(x) \cdot v.$$

► On note $g: V \subset \mathcal{V}_{\mathbb{R}}(0) \rightarrow E$, $t \mapsto g(t) := x + tv$. L'application g est dérivable en $t = 0$ car c'est une application affine continue. Puisque f est dérivable en $x = g(0)$, alors $\varphi := f \circ g$ est dérivable en $t = 0$ et on a $\varphi'(0) = D_v f(x) = f'(g(0)) \circ g'(0) = f'(x) \cdot v \in F$. ■

Exemple 1.4.1. La réciproque de la proposition est fausse. Considérons l'application

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ x = (x_1, x_2) &\longmapsto f(x) = \begin{cases} 1 & \text{si } x_2 = x_1^2 \text{ et } x_1 \neq 0, \\ 0 & \text{sinon.} \end{cases} \end{aligned}$$

Soit $x = (0, 0)$ et $v \in \mathbb{R}^2 \setminus \{0_{\mathbb{R}^2}\}$. Alors, $x + tv = tv = (tv_1, tv_2)$ et $tv_2 \neq t^2 v_1^2$ pour $|t|$ suffisamment petit. On a donc

$$\frac{f(x + tv) - f(x)}{t} = \frac{f(tv) - f(0_{\mathbb{R}^2})}{t} = \frac{f(tv)}{t} = \frac{0}{t} = 0 \xrightarrow{t \rightarrow 0} 0 = D_v f(0),$$

et donc pour tout $v \in \mathbb{R}^2$, $D_v f(0)$ existe, or f n'est pas dérivable en 0 car elle n'y est même pas continue, cf. $\lim_{x_1 \rightarrow 0} f(x_1, x_1^2) = 1 \neq f(0_{\mathbb{R}^2}) = 0$. □

1.4.2 Dérivation

Considérons l'ensemble des applications de classe \mathcal{C}^1 de U dans F , noté $\mathcal{C}^1(U, F)$. Supposons que F soit muni d'une opération interne (notée multiplicativement) faisant de F une algèbre (par exemple $F = \mathbb{R}$ avec la multiplication usuelle). Notons 1_F l'élément neutre (que l'on suppose exister) pour la multiplication. Notons alors que $\mathcal{C}^1(U, F)$ est une algèbre. Sur cet ensemble d'applications, on définit les opérateurs suivants :

Définition 1.4.3 – Dérivation

Soit $x \in U \subset E$. Une *dérivation en x* est une application linéaire $D_x: \mathcal{C}^1(U, F) \mapsto F$ qui vérifie la règle de Leibniz. Plus précisément, on a pour tout couple $(\alpha, \beta) \in \mathbb{R}^2$, et toutes fonctions f, g dans $\mathcal{C}^1(U, F)$:

- i) $D_x(\alpha f + \beta g) = \alpha D_x(f) + \beta D_x(g)$,
- ii) $D_x(fg) = f(x) D_x(g) + D_x(f) g(x)$.

Toute dérivation vérifie $D_x(1_F) = 0$ car $D_x(1_F) = D_x(1_F \times 1_F) = 2 D_x(1_F)$, donc pour toute fonction constante de la forme $f(x) = \alpha 1_F$, $\alpha \in \mathbb{R}$, on a $D_x(f) = \alpha D_x(1_F) = 0$. On peut noter aussi que l'ensemble des dérivations en x muni des opérations habituelles

$$\begin{aligned} (\widetilde{D}_x + D_x)(f) &= \widetilde{D}_x(f) + D_x(f), \\ (\alpha D_x)(f) &= \alpha D_x(f), \end{aligned}$$

forme un sous-espace vectoriel des applications linéaires. Pour simplifier un peu le résultat suivant, supposons que $F = \mathbb{R}$.

Proposition 1.4.4

L'application $D_{x,v}: \mathcal{C}^1(U, \mathbb{R}) \mapsto \mathbb{R}$, définie par $D_{x,v}(f) := D_v f(x) = f'(x) \cdot v$ dans \mathbb{R} est un opérateur de dérivation.

► L'opérateur $D_{x,v}$ est linéaire par composition d'applications linéaires. En effet, on a $D_{x,v} = \lambda_v \circ D_x$, avec $D_x(f) := f'(x)$ linéaire d'après la proposition 1.2.4, et $\lambda_v: \mathcal{L}(E, \mathbb{R}) \rightarrow \mathbb{R}$ définie par $\lambda_v(T) := T(v)$, qui est bien linéaire.

Montrons que $D_{x,v}$ vérifie la règle de Leibniz. Calculons tout d'abord la différentielle du produit de deux applications f et g de classe \mathcal{C}^1 sur U . Posons, $\psi: U \rightarrow \mathbb{R}^2$ qui à $x \in U$, associe le couple $\psi(x) := (f(x), g(x))$. Posons $B: \mathbb{R}^2 \rightarrow \mathbb{R}$, l'application bilinéaire (continue) définie par $B(y_1, y_2) := y_1 y_2$. Alors, l'application $x \mapsto (fg)(x) = f(x)g(x) = (B \circ \psi)(x)$ a pour différentielle en $x \in U$, l'application $(fg)'(x)$ telle que pour tout $v \in E$:

$$\begin{aligned} (fg)'(x) \cdot v &= B'(\psi(x)) \cdot (\psi'(x) \cdot v) = B(f(x), g'(x) \cdot v) + B(f'(x) \cdot v, g(x)) \\ &= f(x) (g'(x) \cdot v) + (f'(x) \cdot v) g(x). \end{aligned}$$

Finalement, $D_{x,v}$ vérifie la règle de Leibniz car :

$$\begin{aligned} D_{x,v}(fg) &= (fg)'(x) \cdot v = f(x) (g'(x) \cdot v) + (f'(x) \cdot v) g(x) \\ &= f(x) D_{x,v}(g) + D_{x,v}(f) g(x). \end{aligned}$$

■

1.5 Fonctions définies sur et à valeur dans un espace produit

1.5.1 Cas des fonctions à valeurs dans un espace produit

Soit $(F_j, \|\cdot\|_{F_j})_{1 \leq j \leq m}$ une famille finie d'espaces vectoriels normés. Supposons que l'ensemble F s'écrit sous la forme du produit cartésien des espaces F_j , i.e. $F = \prod_{j=1}^m F_j$. Notons \mathcal{N} une norme quelconque sur \mathbb{R}^m parmi l'une de ces trois normes équivalentes⁵ :

$$\forall x \in \mathbb{R}^m, \quad \|x\|_1 := \sum_{j=1}^m |x_j|, \quad \|x\|_2 := \sqrt{\sum_{j=1}^m |x_j|^2}, \quad \|x\|_\infty := \max_{1 \leq j \leq m} |x_j|,$$

et introduisons l'application

$$\begin{aligned} \psi: F &\longrightarrow \mathbb{R}^m \\ y = (y_j)_{1 \leq j \leq m} &\longmapsto \psi(y) := (\|y_j\|_{F_j})_{1 \leq j \leq m} \end{aligned}$$

qui à un vecteur y de F associe le vecteur de \mathbb{R}^m dont les composantes sont données par les normes des composantes de y . On peut munir F de la norme $\|\cdot\|_F := \mathcal{N} \circ \psi$ ce qui fait alors de F un espace vectoriel normé.

Remarque 1.5.1. Soient $\mathcal{N}_1, \mathcal{N}_2$ deux normes sur \mathbb{R}^m parmi les trois. Les normes $\mathcal{N}_1 \circ \psi$ et $\mathcal{N}_2 \circ \psi$ sont alors équivalentes. Elles définissent donc la même topologie qui n'est rien d'autre que la topologie produit des topologies des espaces facteurs.

Introduisons maintenant les projections et injections canoniques suivantes :

$$\begin{aligned} q_j: F &\longrightarrow F_j & \text{et} & \quad \mu_j: F_j &\longrightarrow F \\ y &\longmapsto q_j(y) := y_j & & \quad a &\longmapsto \mu_j(a) := (0_{F_1}, \dots, 0_{F_{j-1}}, a, 0_{F_{j+1}}, \dots, 0_{F_m}) \end{aligned}$$

telles que la j^{e} composante de $\mu_j(a)$ vaut a et les autres sont nulles. Il est clair que les projections et injections canoniques sont des applications linéaires continues. On a de plus $q_j \circ \mu_j = \text{Id}_{F_j}$, $q_i \circ \mu_j = 0_{\mathcal{L}(F_j, F_i)}$ si $i \neq j$ et $\sum_{j=1}^m \mu_j \circ q_j = \text{Id}_F$. Soit $f: U \subset E \rightarrow F$. On pose $f_j := q_j \circ f$, la j^{e} application composante. On peut alors écrire

$$f = \text{Id}_F \circ f = \left(\sum_{j=1}^m \mu_j \circ q_j \right) \circ f = \sum_{j=1}^m \mu_j \circ f_j$$

pour identifier $\mathcal{F}(U, F) = \mathcal{F}(U, \prod_{j=1}^m F_j)$ (l'ensemble des applications de U dans $F = \prod_{j=1}^m F_j$) avec $\prod_{j=1}^m \mathcal{F}(U, F_j)$ (l'ensemble des m-uplets d'applications de U dans F_j) à l'aide de

$$\begin{aligned} \varphi: \mathcal{F}(U, \prod_{j=1}^m F_j) &\longrightarrow \prod_{j=1}^m \mathcal{F}(U, F_j) \\ f &\longmapsto \varphi(f) := (f_j)_{1 \leq j \leq m}, \quad f_j = q_j \circ f. \end{aligned}$$

Utiliser cet isomorphisme signifie que l'on peut écrire $f = (f_j)_{1 \leq j \leq m}$.

Proposition 1.5.1

f dérivable en x (sur $U, \mathcal{C}^1(U, F)$) $\iff \forall j \in \llbracket 1, m \rrbracket, f_j$ dérivable en x (sur $U, \mathcal{C}^1(U, F_j)$).

5. Toutes les normes sont équivalentes sur un espace de dimension finie.

► Supposons f dérivable en x . Alors, d'après le théorème 1.3.3, $f_j = q_j \circ f$ est dérivable en x puisque q_j l'est en $f(x)$, en tant qu'application linéaire continue. De même si pour tout j , f_j est dérivable en x , alors $\mu_j \circ f_j$ l'est aussi puisque μ_j est une application linéaire continue, et donc d'après la proposition 1.2.4, $f = \sum_{j=1}^m \mu_j \circ f_j$ est dérivable en x . Remplacer dérivable en x par dérivable sur U puis par \mathcal{C}^1 sur U pour compléter la preuve. ■

Si f est dérivable en x alors $f'(x) \in \mathcal{L}(E, F) = \mathcal{L}(E, \prod_{j=1}^m F_j) \simeq \prod_{j=1}^m \mathcal{L}(U, F_j)$, cf. l'isomorphisme

$$\begin{aligned} \varphi: \mathcal{L}(U, \prod_{j=1}^m F_j) &\longrightarrow \prod_{j=1}^m \mathcal{L}(U, F_j) \\ T &\longmapsto \varphi(T) := (T_j)_{1 \leq j \leq m}, \quad T_j := q_j \circ T, \end{aligned}$$

et on peut donc écrire $f'(x) = (f'_j(x))_{1 \leq j \leq m}$. Ainsi, pour tout $v \in E$ on a

$$f'(x) \cdot v = (f'_j(x) \cdot v)_{1 \leq j \leq m} \in F. \quad (1.2)$$

Remarque 1.5.2. On retrouve ce dernier résultat à partir de la notation $f = \sum_{j=1}^m \mu_j \circ f_j$. En effet, puisque l'opérateur de dérivation en un point est linéaire, cf. proposition 1.2.4, on a, si f est dérivable en x :

$$f'(x) \cdot v = \sum_{j=1}^m (\mu_j \circ f_j)'(x) \cdot v = \sum_{j=1}^m \mu_j(f'_j(x) \cdot v) \in F,$$

en remarquant que $\mu'_j(y) \cdot w = \mu_j(w)$, puisque μ_j est une application linéaire continue. La dérivée de l'application composante est quant à elle donnée par

$$f'_j(x) = (q_j \circ f)'(x) = q'_j(f(x)) \circ f'(x) = q_j \circ f'(x),$$

car q_j est une application linéaire continue.

Exemple 1.5.1 (Cas $E = \mathbb{R}$). Dans ce cas, $f'_j(x) \in \mathcal{L}(\mathbb{R}, F_j) \simeq F_j$, cf. l'exemple fondamental 1.2.2, section 1.2.2, et $f'(x) = (f'_j(x))_{1 \leq j \leq m} \in F$. □

Exemple 1.5.2 (Cas $F_j = \mathbb{R}$ pour tout $j \in \llbracket 1, m \rrbracket$). Dans ce cas, $f'_j(x) \in \mathcal{L}(E, \mathbb{R}) = E'$ et si E est un espace de Hilbert, alors $f'_j(x) \cdot v = (\nabla f_j(x) | v)$. Soit $(e_j)_{1 \leq j \leq m}$ la base canonique de $F = \mathbb{R}^m$. On peut écrire $f = \sum_{j=1}^m f_j e_j$ et en utilisant les règles de dérivation tout en remarquant que l'application $x \mapsto e_j$ est constante, on obtient $f'(x) \cdot v = \sum_{j=1}^m (f'_j(x) \cdot v) e_j \in F = \mathbb{R}^m$. Si E est un espace de Hilbert, alors $f'(x) \cdot v = \sum_{j=1}^m (f'_j(x) \cdot v) e_j = \sum_{j=1}^m (\nabla f_j(x) | v) e_j$. □

1.5.2 Cas des fonctions définies sur un espace produit

Supposons que $E = \prod_{i=1}^n E_i$, où les E_i sont des espaces vectoriels normés. Introduisons les projections et injections canoniques suivantes :

$$\begin{aligned} p_i: E &\longrightarrow E_i & \text{et} & & \lambda_i: E_i &\longrightarrow F \\ x &\longmapsto p_i(x) := x_i & & & a &\longmapsto \lambda_i(a) := (0_{E_1}, \dots, 0_{E_{i-1}}, a, 0_{E_{i+1}}, \dots, 0_{E_n}) \end{aligned}$$

telles que la i^{e} composante de $\lambda_i(a)$ vaut a et les autres sont nulles. On a alors $p_i \circ \lambda_i = \text{Id}_{E_i}$, $p_i \circ \lambda_j = 0_{\mathcal{L}(E_j, E_i)}$ si $i \neq j$ et $\sum_{i=1}^n \lambda_i \circ p_i = \text{Id}_E$. Soit

$$\begin{aligned} f: U \subset E = \prod_{i=1}^n E_i &\longrightarrow F \\ x = (x_i)_{1 \leq i \leq n} &\longmapsto f(x) = f(x_1, \dots, x_n). \end{aligned}$$

Fixons un point $\bar{x} \in U$ et introduisons l'application

$$\begin{aligned} \theta_{\bar{x},i}: E_i &\longrightarrow E \\ a &\longmapsto \theta_{\bar{x},i}(a) := \bar{x} + \lambda_i(a - \bar{x}_i) = (\bar{x}_1, \dots, \bar{x}_{i-1}, a, \bar{x}_{i+1}, \dots, \bar{x}_n). \end{aligned}$$

L'application $\theta_{\bar{x},i}$ est une application affine continue que l'on peut l'écrire sous la forme $\theta_{\bar{x},i} = \tau_{\bar{x}} \circ \lambda_i \circ \tau_{-\bar{x}_i} = \tau_{\bar{x}_{[i]}} \circ \lambda_i$, où $\tau_x(a) := x + a$ et $x_{[i]} := (x_1, \dots, x_{i-1}, 0_{E_i}, x_{i+1}, \dots, x_n)$. L'application

$$\begin{aligned} f \circ \theta_{\bar{x},i}: \theta_{\bar{x},i}^{-1}(U) \subset E_i &\longrightarrow F \\ a &\longmapsto f \circ \theta_{\bar{x},i}(a) = f(\theta_{\bar{x},i}(a)) = f(\bar{x}_1, \dots, \bar{x}_{i-1}, a, \bar{x}_{i+1}, \dots, \bar{x}_n) \end{aligned}$$

est appelée la *i^e application partielle de f relative au point \bar{x}* , et est notée abusivement $f \circ \theta_{\bar{x},i} =: f(\bar{x}_1, \dots, \bar{x}_{i-1}, \cdot, \bar{x}_{i+1}, \dots, \bar{x}_n)$.

Remarque 1.5.3. Pour fixer les idées, nous avons ici noté \bar{x} le point de U et a la nouvelle variable de l'application partielle. Cependant, on retrouvera vite la notation x pour le point de U et on utilisera la notation un peu ambiguë mais pratique x_i pour la variable de l'application $f \circ \theta_{x,i}$. On remarque de plus que $\theta_{x,i}$ est égale à $\tau_{x_{[i]}} \circ \lambda_i$, donc cette application ne dépend pas vraiment de x mais de $x_{[i]}$. On pourrait donc la noter $\theta_{x_{[i]},i}$ et donc écrire l'application partielle

$$f \circ \theta_{x_{[i]},i}(x_i) = f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$$

et on retiendra que les composantes sont toutes fixées, exceptée x_i .

Définition 1.5.2 – Dérivée partielle

On dit que f admet une dérivée partielle au point $x \in U$ par rapport à la variable x_i si l'application $f \circ \theta_{x_{[i]},i}$ est dérivable en x_i . Si la dérivée partielle existe, on notera

$$\frac{\partial f}{\partial x_i}(x) := (f \circ \theta_{x_{[i]},i})'(x_i) \in \mathcal{L}(E_i, F).$$

Remarque 1.5.4. On utilisera aussi la notation

$$\partial_{x_i} f(x) := \frac{\partial f}{\partial x_i}(x).$$

Nous avons alors le résultat suivant et on peut noter que la réciproque du point i) de la proposition (1.5.3) est fausse !

Proposition 1.5.3

Soient $f: U \subset E = \prod_{i=1}^n E_i \rightarrow F$ et $x \in U$.

- i) f dérivable en $x \Rightarrow \forall i \in \llbracket 1, n \rrbracket, \partial_{x_i} f(x)$ existe.
- ii) $\forall i \in \llbracket 1, n \rrbracket, \partial_{x_i} f(x) = f'(x) \circ \lambda_i$.
- iii) $f'(x) = \sum_{i=1}^n \partial_{x_i} f(x) \circ p_i$.

► L'application $\theta_{x_{[i]},i}$ étant affine et continue, elle est dérivable en x_i . Par hypothèse, f dérivable en $\theta_{x_{[i]},i}(x_i) = x$ donc par composition $f \circ \theta_{x_{[i]},i}$ dérivable en x_i et $\partial_{x_i} f(x)$ existe. Le point i) est démontré. Supposons f dérivable en x . On a alors, par dérivation des fonctions

composées

$$\frac{\partial f}{\partial x_i}(x) = (f \circ \theta_{x_{[i]}, i})'(x_i) = f'(\theta_{x_{[i]}, i}(x_i)) \circ \theta'_{x_{[i]}, i}(x_i) = f'(x) \circ \theta'_{x_{[i]}, i}(x_i),$$

or puisque $\tau_{x_{[i]}} = x_{[i]} + \text{Id}_E$ et λ_i est linéaire continue :

$$\theta'_{x_{[i]}, i}(x_i) = (\tau_{x_{[i]}} \circ \lambda_i)'(x_i) = \tau'_{x_{[i]}}(\lambda_i(x_i)) \circ \lambda'_i(x_i) = \lambda'_i(x_i) = \lambda_i.$$

Le point *ii*) est démontré. Puisque $\sum_{i=1}^n \lambda_i \circ p_i = \text{Id}_E$, on peut écrire $f = f \circ (\sum_{i=1}^n \lambda_i \circ p_i)$. En utilisant les règles de dérivation, avec le fait que $x = (\sum_{i=1}^n \lambda_i \circ p_i)(x)$, on a

$$f'(x) = f'(x) \circ (\sum_{i=1}^n \lambda_i \circ p_i)'(x) = f'(x) \circ (\sum_{i=1}^n \lambda'_i(p_i(x)) \circ p'_i(x)) = f'(x) \circ (\sum_{i=1}^n \lambda_i \circ p_i)$$

puisque λ_i et p_i sont des applications linéaires continues. La linéarité de $f'(x)$ combinée au point *ii*) permet de conclure que $f'(x) = \sum_{i=1}^n \partial_{x_i} f(x) \circ p_i$ et le point *iii*) est démontré. ■

Remarque 1.5.5. D'après le point *iii*) de la proposition, on a $\forall v := (v_i)_{1 \leq i \leq n} \in E$:

$$f'(x) \cdot v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \cdot v_i.$$

Exemple 1.5.3 (Cas $E_i = \mathbb{R}$ pour tout $i \in \llbracket 1, n \rrbracket$). Notons $(e_i)_{1 \leq i \leq n}$ la base canonique de \mathbb{R}^n . Supposons f dérivable en x . On a alors $\partial_{x_i} f(x) \in \mathcal{L}(\mathbb{R}, F) \simeq F$ et donc

$$\frac{\partial f}{\partial x_i}(x) \simeq \frac{\partial f}{\partial x_i}(x) \cdot 1 = (f'(x) \circ \lambda_i) \cdot 1 = f'(x) \cdot \lambda_i(1) = f'(x) \cdot e_i = D_{e_i} f(x).$$

La dérivée partielle $\partial_{x_i} f(x)$ n'est rien d'autre ici que la dérivée directionnelle de f en x dans la direction e_i . □

Remarque 1.5.6. De manière générale, c'est la dérivée partielle appliquée à un vecteur qui est une dérivée directionnelle. En effet, pour $v_i \in E_i$, on a :

$$\frac{\partial f}{\partial x_i}(x) \cdot v_i = f'(x) \cdot \lambda_i(v_i) = D_{\lambda_i(v_i)} f(x).$$

Exemple 1.5.4 (Cas $F = \mathbb{R}$ et $E_i = \mathbb{R}$ pour tout $i \in \llbracket 1, n \rrbracket$). Dans ce cas, $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et d'après le troisième exemple fondamental 1.2.3, section 1.2.2, $f'(x) \cdot v = (\nabla f(x) | v)$ avec $\nabla f(x)$ un vecteur de $E = \mathbb{R}^n$. Les composantes de $\nabla f(x)$ sont alors les dérivées partielles de f ! En effet, pour tout $i \in \llbracket 1, n \rrbracket$, $\partial_{x_i} f(x) = f'(x) \cdot e_i = (\nabla f(x) | e_i) \in \mathbb{R}$ ce qui nous donne bien la i^{e} composante du gradient. Nous avons dans ce cas $f'(x) \in (\mathbb{R}^n)^*$, c'est donc une forme linéaire sur \mathbb{R}^n . Introduisons $(dx_i)_{1 \leq i \leq n}$ la base duale de $(e_i)_{1 \leq i \leq n}$ telle que $dx_i(e_j) = \delta_{ij}$ où δ_{ij} est le symbole de Kronecker (qui vaut 1 ou 0 suivant que i et j sont égaux ou non). Ainsi, $f'(x)$ s'écrit dans cette base sous la forme $f'(x) = \sum_{k=1}^n \alpha_k dx_k$ et on a $\partial_{x_i} f(x) = f'(x) \cdot e_i = \sum_{k=1}^n \alpha_k dx_k(e_i) = \alpha_i \in \mathbb{R}$, donc finalement on peut écrire

$$f'(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) dx_i.$$

□

1.5.3 Cas général : synthèse

Soit $f: U \subset E = \prod_{i=1}^n E_i \rightarrow F = \prod_{j=1}^m F_j$. Supposons f dérivable en x . D'après la remarque 1.5.2, on a

$$f'(x) = \sum_{j=1}^m \mu_j \circ f'_j(x), \quad f'_j(x) = q_j \circ f'(x),$$

et d'après les points *ii*) et *iii*) de la proposition 1.5.3, on a

$$f'_j(x) = \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(x) \circ p_i, \quad \frac{\partial f_j}{\partial x_i}(x) = f'_j(x) \circ \lambda_i,$$

donc en combinant les deux, on obtient finalement

$$f'(x) = \sum_{j=1}^m \mu_j \circ \left(\sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(x) \circ p_i \right) = \sum_{j=1}^m \sum_{i=1}^n \mu_j \circ \frac{\partial f_j}{\partial x_i}(x) \circ p_i \in \mathcal{L}(E, F)$$

et

$$\frac{\partial f_j}{\partial x_i}(x) = q_j \circ f'(x) \circ \lambda_i \in \mathcal{L}(E_i, F_j). \quad (1.3)$$

D'après l'isomorphisme vu en section 1.5.1, on a $f'(x) \in \mathcal{L}(E, \prod_{j=1}^m F_j) \simeq \prod_{j=1}^m \mathcal{L}(E, F_j)$. En combinant avec l'isomorphisme suivant :

$$\begin{aligned} \varphi: \mathcal{L}(\prod_{i=1}^n E_i, F) &\longrightarrow \prod_{i=1}^n \mathcal{L}(E_i, F) \\ T &\longmapsto \varphi(T) := (T_i)_{1 \leq i \leq n}, \quad T_i := T \circ \lambda_i, \end{aligned}$$

on a donc $f'(x) \in \mathcal{L}(E, F) \simeq \prod_{i=1}^n \prod_{j=1}^m \mathcal{L}(E_i, F_j)$ et on peut finalement faire l'identification suivante (on appelle cela la *notation matricielle*) :

$$f'(x) = \left(\frac{\partial f_j}{\partial x_i}(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$$

avec

$$\frac{\partial f_j}{\partial x_i}(x) \in \mathcal{L}(E_i, F_j).$$

Proposition 1.5.4 – Dérivée partielle d'une application composée

Soient $E = \prod_{i=1}^n E_i$, $F = \prod_{j=1}^m F_j$ et $G = \prod_{k=1}^p G_k$ trois espaces vectoriels normés produits. Soient $f: U \rightarrow F$, U ouvert de E et $g: V \rightarrow G$, V ouvert F , tels que $f(U) \subset V$. Soit $x \in U$. Si f est dérivable en x et si g est dérivable en $f(x)$ alors $g \circ f$ est dérivable en x et toutes les applications composantes de $g \circ f$ admettent des dérivées partielles en x par rapport aux variables x_i . Les dérivées partielles sont alors données pour tout $i \in [1, n]$ et $k \in [1, p]$ par :

$$\frac{\partial (g \circ f)_k}{\partial x_i}(x) = \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \circ \frac{\partial f_j}{\partial x_i}(x) \in \mathcal{L}(E_i, G_k).$$

► Introduisons les projections et injections canoniques suivantes :

$$\begin{array}{ccc} r_k: G & \longrightarrow & G_k \\ z & \longmapsto & r_k(z) := z_k \end{array} \quad \text{et} \quad \begin{array}{ccc} \nu_k: G_k & \longrightarrow & G \\ a & \longmapsto & \nu_k(a) := (0_{G_1}, \dots, 0_{G_{k-1}}, a, 0_{G_{k+1}}, \dots, 0_{G_p}). \end{array}$$

On a alors

$$\begin{aligned} \frac{\partial(g \circ f)_k}{\partial x_i}(x) &= r_k \circ (g \circ f)'(x) \circ \lambda_i, && \text{cf. eq. (1.3)} \\ &= r_k \circ g'(f(x)) \circ f'(x) \circ \lambda_i, && \text{cf. théorème 1.3.3} \\ &= r_k \circ g'(f(x)) \circ \text{Id}_F \circ f'(x) \circ \lambda_i \\ &= r_k \circ g'(f(x)) \circ \left(\sum_{j=1}^m \mu_j \circ q_j \right) \circ f'(x) \circ \lambda_i, && \text{cf. section 1.5.1} \\ &= \sum_{j=1}^m (r_k \circ g'(f(x)) \circ \mu_j) \circ (q_j \circ f'(x) \circ \lambda_i), \quad \text{car } r_k \text{ et } g'(f(x)) \text{ sont linéaires} \\ &= \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \circ \frac{\partial f_j}{\partial x_i}(x). && \text{cf. eq. (1.3)} \end{aligned}$$

■

Exemple 1.5.5 (Cas fondamental : $E = \mathbb{R}^n$, $F = \mathbb{R}^m$). Dans ce cas, $\partial_{x_i} f_j(x) \in \mathcal{L}(\mathbb{R}, \mathbb{R}) \simeq \mathbb{R}$ et alors la matrice $f'(x) \in \mathbf{M}_{m,n}(\mathbb{R})$ est appelée la *matrice jacobienne de f en x* . Elle sera parfois notée $J_f(x)$ pour mettre en exergue le fait que c'est une matrice. Si $m = n$, alors le déterminant $\det(J_f(x))$ est appelé le *jacobien de f en x* . Introduisons $(e_j)_{1 \leq j \leq m}$ la base canonique de \mathbb{R}^m et $(dx_i)_{1 \leq i \leq n}$ celle de $(\mathbb{R}^n)^*$. Soit $v = (v_i)_{1 \leq i \leq n} \in E = \mathbb{R}^n$. Récapitulons ce que l'on sait déjà :

1. $f'(x) \cdot v = \sum_{j=1}^m (f'_j(x) \cdot v) e_j \in \mathbb{R}^m$, cf. exemple 1.5.2, section 1.5.1 ;
2. $f'_j(x) \cdot v = \sum_{i=1}^n \partial_{x_i} f_j(x) dx_i(v) \in \mathbb{R}$, $dx_i(v) = v_i$, cf. exemple 1.5.4, section 1.5.2.

En combinant ces deux points, on obtient finalement

$$f'(x) \cdot v = \sum_{j=1}^m \left(\sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(x) dx_i(v) \right) e_j = J_f(x) v,$$

où le dernier terme est bien le produit matrice-vecteur entre $J_f(x)$ et v . Enfin, dans les hypothèses de la proposition 1.5.4, on peut voir que la matrice jacobienne de $g \circ f$ au point x est le produit de la matrice jacobienne de g au point $f(x)$ avec celle f au point x , *i.e.* on a

$$J_{g \circ f}(x) = J_g(f(x)) J_f(x),$$

autrement dit, la jacobienne de la composée de deux fonctions est le produit des jacobienes des deux fonctions. □

Exemple 1.5.6 (Cas particulier : $E = \mathbb{R}^n$, $F = \mathbb{R}^m$, $m = 1$). Voir l'exemple 1.5.4 de la section 1.5.2. Dans ce cas, $J_f(x)^T = \nabla f(x) \in \mathbf{M}_{n,1}(\mathbb{R}) \simeq \mathbb{R}^n$. □

1.6 Théorèmes des accroissements finis et applications

Le théorème des accroissements finis donne une majoration de $f(b) - f(a)$ sous des hypothèses appropriées portant sur la dérivée de f . Il s'agit essentiellement d'un résultat portant sur les

fonctions de la variable réelle. Rappelons que le théorème de Rolle pour des fonctions à valeurs réelles assure l'existence d'un point $c \in]a, b[$, $a < b$ réels, tel que $f(b) - f(a) = f'(c)(b - a)$ sous les hypothèses suivantes : f continue sur $[a, b]$ et dérivable sur $]a, b[$. Ce résultat ne subsiste pas pour les fonctions à valeurs vectorielles, comme le montre l'exemple de la fonction à valeurs dans \mathbb{R}^2 , $f(x) = (\cos(2\pi x), \sin(2\pi x))$ sur $[0, 1]$.

Exemple 1.6.1. La fonction $f: \mathbb{R} \rightarrow \mathbb{R}^2$ définie par $f(x) = (\cos(2\pi x), \sin(2\pi x))$, a, en chaque point une dérivée de norme 2π , et il n'existe aucun point $c \in]0, 1[$ tel que $f(1) - f(0) = f'(c)$. En effet, la dérivée de f en x s'écrit $f'(x) = 2\pi(-\sin(2\pi x), \cos(2\pi x))$, donc pour tout $x \in \mathbb{R}$, $\|f'(x)\| = 2\pi$. Et puisque $f(1) = f(0)$, on ne peut avoir $\|f(1) - f(0)\| = 0 = \|f'(c)\| = 2\pi$. \square

Rappelons qu'il existe aussi une version intégrale du théorème des accroissements finis pour les fonctions à valeurs dans \mathbb{R} . Pour une fonction f à valeurs réelles, définie et continue sur $[a, b]$, il existe $c \in]a, b[$ tel que

$$f(c) = \frac{1}{b-a} \int_a^b f(x) dx.$$

Pour les fonctions de la variable réelle et à valeurs dans un espace vectoriel normés quelconque, nous avons le résultat suivant.

Théorème 1.6.1 – 1^{re} forme des accroissements finis

Soient $]a, b[\subset \mathbb{R}$, $(F, \|\cdot\|_F)$ un espace vectoriel normé, $f \in \mathcal{C}^0([a, b], F)$ et $g \in \mathcal{C}^0([a, b], \mathbb{R})$. Supposons f et g dérivables dans $]a, b[$. Si $\forall x \in]a, b[, \|f'(x)\|_F \leq g'(x)$ alors :

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

► Soit $\varepsilon > 0$. Notons

$$I_\varepsilon := \{y \in [a, b] \mid \forall x \in [a, y], \|f(x) - f(a)\|_F \leq g(x) - g(a) + \varepsilon(x - a) + \varepsilon\}.$$

* Nous allons montrer que $I_\varepsilon = [a, b]$. On aura alors l'inégalité vraie pour $x = b$, et ce, quelque soit $\varepsilon > 0$. Ainsi, la conclusion viendra en fixant $x = b$ et en faisant tendre ε vers 0.

* Notons $c_\varepsilon := \sup I_\varepsilon$. Montrons que $c_\varepsilon = b$. Tout d'abord, l'application

$$x \mapsto \varphi(x) := \|f(x) - f(a)\|_F - (g(x) - g(a) + \varepsilon(x - a)), \quad x \in [a, b],$$

est continue en a , puisque f et g le sont, et $\varphi(a) = 0$, donc pour tout $x > a$ suffisamment proche de a , $\varphi(x) \leq \varepsilon$, donc $c_\varepsilon > a$. Supposons maintenant que $a < c_\varepsilon < b$ et raisonnons par l'absurde. Puisque $c_\varepsilon \in]a, b[$, alors, par hypothèse, f et g sont dérivables en c_ε . Ainsi, il existe $y \in]c_\varepsilon, b[$, suffisamment proche de c_ε , tel que pour tout $x \in]c_\varepsilon, y]$, on a

$$\left\| \frac{f(x) - f(c_\varepsilon)}{x - c_\varepsilon} - f'(c_\varepsilon) \right\|_F \leq \frac{\varepsilon}{2} \quad \text{et} \quad \left| \frac{g(x) - g(c_\varepsilon)}{x - c_\varepsilon} - g'(c_\varepsilon) \right| \leq \frac{\varepsilon}{2},$$

d'où, puisque $x - c_\varepsilon > 0$:

$$\begin{aligned} \|f(x) - f(c_\varepsilon)\|_F &= \|f(x) - f(c_\varepsilon) - f'(c_\varepsilon)(x - c_\varepsilon) + f'(c_\varepsilon)(x - c_\varepsilon)\|_F \\ &\leq \|f(x) - f(c_\varepsilon) - f'(c_\varepsilon)(x - c_\varepsilon)\| + \|f'(c_\varepsilon)(x - c_\varepsilon)\|_F \\ &\leq \left(\frac{\varepsilon}{2} + \|f'(c_\varepsilon)\|_F \right) (x - c_\varepsilon) \end{aligned}$$

et puisque aussi $g'(c_\varepsilon) \geq 0$, et $g(x) - g(c_\varepsilon) \geq 0$ par le théorème de Rolle, il vient :

$$\begin{aligned} g'(c_\varepsilon)(x - c_\varepsilon) &= |g(x) - g(c_\varepsilon) - g'(c_\varepsilon)(x - c_\varepsilon) - (g(x) - g(c_\varepsilon))| \\ &\leq \frac{\varepsilon}{2}(x - c_\varepsilon) + |g(x) - g(c_\varepsilon)| \\ &= \frac{\varepsilon}{2}(x - c_\varepsilon) + g(x) - g(c_\varepsilon). \end{aligned}$$

Par hypothèse, puisque $a < c_\varepsilon < b$, on a $\|f'(c_\varepsilon)\|_F \leq g'(c_\varepsilon)$. Ainsi, pour tout $x \in]c_\varepsilon, y]$:

$$\|f(x) - f(c_\varepsilon)\|_F \leq \varepsilon(x - c_\varepsilon) + g(x) - g(c_\varepsilon).$$

En outre, la continuité de f et g nous assure que $c_\varepsilon \in I_\varepsilon$. Et finalement, on obtient

$$\begin{aligned} \|f(x) - f(a)\|_F &= \|f(x) - f(c_\varepsilon) + f(c_\varepsilon) - f(a)\|_F \leq \|f(x) - f(c_\varepsilon)\|_F + \|f(c_\varepsilon) - f(a)\|_F \\ &\leq \varepsilon(x - c_\varepsilon) + g(x) - g(c_\varepsilon) + g(c_\varepsilon) - g(a) + \varepsilon(c_\varepsilon - a) + \varepsilon \\ &= g(x) - g(a) + \varepsilon(x - a) + \varepsilon, \end{aligned}$$

autrement dit, $y \in I_\varepsilon$, ce qui impossible puisque $y > c_\varepsilon = \sup I_\varepsilon$. On a donc $c_\varepsilon = b$. En fixant $x = b$ et en faisant tendre ε vers 0, on obtient le résultat attendu. ■

En prenant $g(x) = kx$, $k \in \mathbb{R}_+$, on obtient le corollaire suivant.

Corollaire 1.6.2

Soit $f \in \mathcal{C}^0([a, b], F)$, f dérivable dans $]a, b[$ et telle que $\|f'(x)\|_F \leq k$ pour tout $x \in]a, b[$, alors f est k -lipschitzienne :

$$\|f(y) - f(x)\|_F \leq k|y - x|, \quad \forall x, y \in [a, b].$$

Lorsque $k = 0$, on obtient en particulier :

Corollaire 1.6.3

Soit $f \in \mathcal{C}^0([a, b], F)$, admettant dans $]a, b[$ une dérivée nulle, alors f est constante.

Enfin, en prenant $f \equiv 0$ dans le théorème 1.6.1, on obtient :

Corollaire 1.6.4

Soit $g \in \mathcal{C}^0([a, b], \mathbb{R})$, admettant dans $]a, b[$ une dérivée ≥ 0 , alors g est croissante.

Revenons au cas des fonctions définies sur un ouvert d'un espace vectoriel normé E . Comme pour les fonctions dérivables d'une variable réelle, le résultat fondamental du calcul différentiel est celui qui, à partir de la connaissance des différentielles d'une fonction en chaque point de son domaine, permet d'estimer l'accroissement de cette fonction entre deux points fixés à l'avance. On se ramène pour cela au cas des fonctions d'une variable, en considérant la restriction de la fonction au segment qui joint les deux points. Rappelons donc que pour une paire $(a, b) \in E^2$, on note

$$[a, b] := \nu([0, 1]) \quad \text{et} \quad]a, b[:= \nu(]0, 1[),$$

où

$$\begin{aligned} \nu: [0, 1] \subset \mathbb{R} &\longrightarrow E \\ t &\longmapsto \nu(t) := (1-t)a + tb = a + t(b-a). \end{aligned}$$

Nous avons alors le résultat suivant.

Théorème 1.6.5 – 2^e forme des accroissements finis

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Soient U un ouvert de E , $(a, b) \in E^2$ tel que $[a, b] \subset U$ et $f: U \rightarrow F$. Si f est continue dans $[a, b]$ et différentiable dans $]a, b[$, alors :

$$\|f(b) - f(a)\|_F \leq \left(\sup_{x \in]a, b[} \|f'(x)\|_{\mathcal{L}(E, F)} \right) \|b - a\|_E \in \overline{\mathbb{R}}_+.$$

► Posons $M := \sup_{x \in]a, b[} \|f'(x)\|_{\mathcal{L}(E, F)}$. Si $M = +\infty$, alors le résultat trivialement vrai. Supposons donc $M < +\infty$. Posons sur l'intervalle $[0, 1]$, l'application $\varphi := f \circ \nu$, $\nu(t) = a + t(b - a)$, de telle sorte que $\varphi(0) = f(a)$ et $\varphi(1) = f(b)$. Introduisons de même sur $[0, 1]$, l'application $g(t) := tM\|b - a\|_E$. Ainsi définies, $\varphi \in \mathcal{C}^0([0, 1], F)$ par composition d'applications continues, $g \in \mathcal{C}^0([0, 1], \mathbb{R})$ et pour tout $t \in]0, 1[$:

$$\begin{aligned} \|\varphi'(t)\|_F &= \|f'(\nu(t)) \cdot \nu'(t)\|_F = \|f'(\nu(t)) \cdot (b - a)\|_F \\ &\leq \|f'(\nu(t))\|_{\mathcal{L}(E, F)} \|b - a\|_E \leq M\|b - a\|_E = g'(t). \end{aligned}$$

D'après le théorème 1.6.1, on obtient finalement

$$\|\varphi(1) - \varphi(0)\|_F = \|f(b) - f(a)\|_F \leq g(1) - g(0) = M\|b - a\|_E,$$

ce qui donne le résultat attendu. ■

Corollaire 1.6.6

Soit $f: U \rightarrow F$ une fonction différentiable définie sur un ouvert **convexe** telle que $\|f'(x)\|_{\mathcal{L}(E, F)} \leq k$ pour tout $x \in U$. Alors, f est k -lipschitzienne :

$$\|f(y) - f(x)\|_F \leq k \|y - x\|_E, \quad \forall x, y \in U.$$

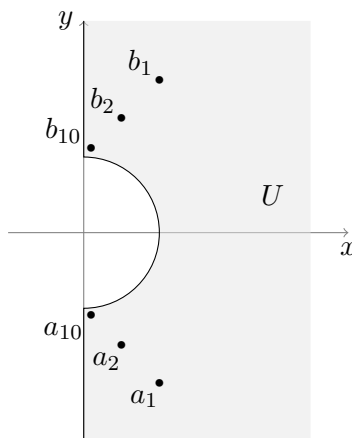
► Puisque U convexe, alors pour tout $(x, y) \in U^2$, le segment $[x, y] \subset U$ et on peut appliquer le théorème 1.6.5. ■

L'hypothèse de convexité de U dans ce dernier résultat ne peut être remplacé par la connexité de U . On a en effet l'exemple suivant.

Exemple 1.6.2. Soient U l'ouvert de \mathbb{R}^2 défini par

$$U := \{(x, y) \mid x > 0 \text{ et } x^2 + y^2 > 1\}$$

et $\psi: U \rightarrow \mathbb{R}$ telle que $\psi(x, y) = \arctan(y/x)$. Alors U est un ouvert connexe, la différentielle de ψ vérifie $\|\psi'(x, y)\| < 1$ pour tout (x, y) dans U et ψ n'est pas k -lipschitzienne pour $k < \pi/2$.



□

► **Correction.** Il est clair que U est connexe mais pas convexe. Sur U , l'application ψ est dérivable et on a :

$$\psi'(x, y) = \frac{1}{1 + \left(\frac{y}{x}\right)^2} \left(-\frac{y}{x^2}, \frac{1}{x} \right),$$

donc

$$\|\psi'(x, y)\| = \frac{x^2}{x^2 + y^2} \left(\frac{y^2}{x^4} + \frac{1}{x^2} \right)^{\frac{1}{2}} = \frac{x^2}{x^2 + y^2} \left(\frac{y^2 + x^2}{x^4} \right)^{\frac{1}{2}} = (x^2 + y^2)^{-\frac{1}{2}} < 1.$$

Montrons que ψ n'est pas k -lipschitzienne pour $k < \pi/2$. On définit dans U , les points :

$$a_n := \left(\frac{1}{n}, -\frac{1+n}{n} \right) \quad \text{et} \quad b_n := \left(\frac{1}{n}, \frac{1+n}{n} \right),$$

pour $n \in \mathbb{N}^*$. Alors, $\|b_n - a_n\| \rightarrow 2$ et

$$|\psi(b_n) - \psi(a_n)| = 2 \arctan(n+1) \rightarrow \pi,$$

donc ψ ne peut pas être k -lipschitzienne sur U pour $k < \pi/2$. □

Sur un ouvert connexe, nous avons le résultat suivant.

Corollaire 1.6.7

Soit $f: U \rightarrow F$ une fonction différentiable définie sur un ouvert **connexe**, admettant une dérivée nulle. Alors, f est constante.

► Soit $x_0 \in U$. Posons

$$A := \{x \in U \mid f(x) = f(x_0)\}.$$

L'application f étant différentiable sur U , elle y est continue. L'ensemble A est donc fermé comme préimage d'un fermé par une application continue. Pour tout $x \in A$, le corollaire 1.6.6 montre que toute boule ouverte $B(x, r)$, $r > 0$, contenue dans U est aussi contenue dans A , puisque f est 0-lipschitzienne sur la boule $B(x, r)$ qui est convexe. Ainsi, l'ensemble A est ouvert dans U . Finalement, A est ouvert et fermé dans U , U connexe, donc $A = U$, et on obtient le résultat attendu. ■

Remarque 1.6.1. Sur un ouvert formé de plusieurs composantes connexes, si la dérivée est nulle, alors l'application est constante sur chaque composante connexe, mais rien oblige à ce que les constantes soient égales.

1.7 Dérivabilité vs dérivée partielle et cas fondamentaux

Proposition 1.7.1

Soit $f: U \subset E = \prod_{i=1}^n E_i \rightarrow F$, U ouvert de E , avec E_i et F , des espaces vectoriels normés. On a :

$$f \in \mathcal{C}^1(U, F) \iff \forall i \in \llbracket 1, n \rrbracket, \frac{\partial f}{\partial x_i} \in \mathcal{C}^0(U, \mathcal{L}(E_i, F)).$$

► Démontrons l'implication de gauche à droite en premier, puis la réciproque, plus difficile.

1. Si $f \in \mathcal{C}^1(U, F)$, alors il est clair que $\forall i \in \llbracket 1, n \rrbracket, \partial_{x_i} f \in \mathcal{C}^0(U, \mathcal{L}(E_i, F))$, puisque $\partial_{x_i} f(x) = f'(x) \circ \lambda_i$, avec $\lambda_i: E_i \rightarrow F$, $\lambda_i(a) = (0_{E_1}, \dots, 0_{E_{i-1}}, a, 0_{E_{i+1}}, \dots, 0_{E_n})$, cf. proposition 1.5.3. En effet, $\partial_{x_i} f$ est la composée de l'application continue $x \mapsto f'(x)$ (car $f \in \mathcal{C}^1$) avec l'application sur $\mathcal{L}(E, F)$ définie par $T \mapsto \varphi(T) := T \circ \lambda_i$, qui est linéaire (évident) et continue, cf. $\|\varphi(T)\|_{\mathcal{L}(E_i, F)} \leq \|\lambda_i\|_{\mathcal{L}(E_i, E)} \|T\|_{\mathcal{L}(E, F)}$.

2. Supposons que $\forall i \in \llbracket 1, n \rrbracket, \partial_{x_i} f \in \mathcal{C}^0(U, \mathcal{L}(E_i, F))$, et montrons la réciproque.

* Introduisons :

$$\begin{aligned} \psi: U &\longrightarrow \prod_{i=1}^n \mathcal{L}(E_i, F) \\ x &\longmapsto \psi(x) := \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \end{aligned}$$

et

$$\begin{aligned} \varphi: G := \prod_{i=1}^n \mathcal{L}(E_i, F) &\longrightarrow \mathcal{L}(E, F) \\ T := (T_1, \dots, T_n) &\longmapsto \varphi(T) := \sum_{i=1}^n T_i \circ p_i, \end{aligned}$$

avec $p_i: E \rightarrow E_i$, la projection canonique $p_i(x) = x_i$, pour $x := (x_1, \dots, x_n) \in E$. On note $g := \varphi \circ \psi$ de telle sorte que

$$\begin{aligned} g: U &\longrightarrow \mathcal{L}(E, F) \\ x &\longmapsto g(x) = (\varphi \circ \psi)(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \circ p_i. \end{aligned}$$

L'application ψ est continue par hypothèse. L'application φ est linéaire (évident). Montrons qu'elle est continue. Pour $T := (T_1, \dots, T_n) \in G$, on définit la norme

$$\|T\|_G := \max_{1 \leq i \leq n} \|T_i\|_{\mathcal{L}(E_i, F)}.$$

Ainsi, φ est continue sur G puisque

$$\begin{aligned} \|\varphi(T)\|_{\mathcal{L}(E, F)} &\leq \sum_{i=1}^n \|T_i \circ p_i\|_{\mathcal{L}(E, F)} \\ &\leq \sum_{i=1}^n \|T_i\|_{\mathcal{L}(E_i, F)} \|p_i\|_{\mathcal{L}(E, E_i)} \leq \left(\sum_{i=1}^n \|p_i\|_{\mathcal{L}(E, E_i)} \right) \|T\|_G. \end{aligned}$$

Finalement, l'application $g = \varphi \circ \psi$ est continue sur U .

* Soit $x \in U$. Posons $h_x(v) := f(x+v) - f(x) - g(x) \cdot v$, définie sur un ouvert tel que $x+v \in U$. Si $h_x(v) = o(v)$, alors f est différentiable en x , de différentielle en x , $f'(x) = g(x)$, donc d'application dérivée $f' = g$ continue, et donc alors f est de classe \mathcal{C}^1 sur U .

* Montrons donc que $h_x(v) = o(v)$. Tout d'abord, on peut décomposer $h_x(v)$ sous la forme (vérifiez-le) suivante ⁶ :

$$h_x(v) = \sum_{i=1}^n (h_i(v_i) - h_i(0_{E_i})),$$

avec $v := (v_1, \dots, v_n)$ et $h_i(a) := f(x_1, \dots, x_{i-1}, x_i + a, x_{i+1} + v_i, \dots, x_n + v_n) - \partial_{x_i} f(x) \cdot a$. Ensuite, étant donné $\varepsilon > 0$, il existe $\delta > 0$ tel que $\Omega := \prod_{i=1}^n B(x_i, \delta)$ est contenu dans U et de plus, tel que pour tout $i \in \llbracket 1, n \rrbracket$ et pour tout $v \in E$ vérifiant $x+v \in \Omega$, on a :

$$\left\| \frac{\partial f}{\partial x_i}(x+v) - \frac{\partial f}{\partial x_i}(x) \right\|_{\mathcal{L}(E_i, F)} \leq \varepsilon,$$

car $\partial_{x_i} f$ est continue sur U par hypothèse, et où l'on a noté $B(x_i, \delta)$ la boule ouverte de rayon δ , centrée en x_i . Soit maintenant $v := (v_1, \dots, v_n) \in E$ tel que $x+v \in \Omega \subset U$. Alors, pour tout i , l'application h_i est définie et dérivable (car $\partial_{x_i} f$ existe sur U) sur $B(0, \delta)$, sa différentielle est donnée par :

$$h'_i(a) = \frac{\partial f}{\partial x_i}(x_1, \dots, x_{i-1}, x_i + a, x_{i+1} + v_{i+1}, \dots, x_n + v_n) - \frac{\partial f}{\partial x_i}(x),$$

et puisque $(x_1, \dots, x_{i-1}, x_i + a, x_{i+1} + v_{i+1}, \dots, x_n + v_n) \in \Omega$, alors, sur $B(0, \delta)$, on a $\|h'_i(a)\|_{\mathcal{L}(E_i, F)} \leq \varepsilon$. On peut donc appliquer le théorème des accroissements finis 1.6.5 à h_i , ce qui donne

$$\|h_i(v_i) - h_i(0_{E_i})\|_F \leq \varepsilon \|v_i\|_{E_i}, \quad \forall v_i \in B(0, \delta).$$

Finalement, pour tout $v \in E$ tel que $x+v \in \Omega$, i.e. tel que $\|v\|_E := \max_{1 \leq i \leq n} \|v_i\|_{E_i} \leq \delta$, on a :

$$\|h_x(v)\|_F \leq \varepsilon \|v\|_E,$$

c'est-à-dire, $h_x(v) = o(v)$. ■

Remarque 1.7.1. Attention, une application qui admet des dérivées partielles n'est pas nécessairement dérivable, ni même continue, cf. exemple 1.4.1.



Exercice 1.7.1 (Cas fondamentaux). Montrer que toute application bilinéaire continue, puis multilinéaire continue, est de classe \mathcal{C}^1 , et donner leurs différentielles. □

▷ **Correction.**

- i) Soient E_1, E_2 et F trois evn. Montrons que $\mathcal{L}_2(E_1 \times E_2, F) \subset \mathcal{C}^1(E_1 \times E_2, F)$, où $\mathcal{L}_2(E_1 \times E_2, F)$ est l'ensemble des applications bilinéaires de $E_1 \times E_2$ à valeurs dans F , muni de la norme

$$\|B\|_{\mathcal{L}_2(E_1 \times E_2, F)} := \sup \{ \|B(x, y)\|_F \mid \|x\|_{E_1} \leq 1 \text{ et } \|y\|_{E_2} \leq 1 \}.$$

Soit $\bar{x} := (\bar{x}_1, \bar{x}_2) \in E_1 \times E_2$. L'application partielle $x_1 \mapsto B(x_1, \bar{x}_2)$, notée $B(\cdot, \bar{x}_2)$, est linéaire (évident) et continue, car (on omet les indices des normes par commodité

6. Il est à noter que h_i dépend de x et v mais nous n'écrivons pas cette dépendance pour alléger les notations.

d'écriture) :

$$\|B(x_1, \bar{x}_2)\| \leq \|B\| \|x_1\| \|\bar{x}_2\| = C \|x_1\|, \quad C := \|B\| \|\bar{x}_2\|.$$

Ainsi $B(\cdot, \bar{x}_2)$ est dérivable et $\partial_{x_1} B(\bar{x}) = B(\cdot, \bar{x}_2)$. On a donc :

$$\begin{aligned} \frac{\partial B}{\partial x_1} : E_1 \times E_2 &\longrightarrow \mathcal{L}(E_1, F) \\ x := (x_1, x_2) &\longmapsto \frac{\partial B}{\partial x_1}(x) = B(\cdot, x_2). \end{aligned}$$

On peut écrire $\partial_{x_1} B := L \circ p_2$ avec $p_2(x) := x_2$ la projection canonique dans E_2 et

$$\begin{aligned} L : E_2 &\longrightarrow \mathcal{L}(E_1, F) \\ x_2 &\longmapsto L(x_2) := B(\cdot, x_2). \end{aligned}$$

L'application p_2 est linéaire et continue (évident). L'application L est linéaire (évident) et continue car :

$$\|L(x_2)\| = \|B(\cdot, x_2)\| = \sup_{\|x_1\| \leq 1} \|B(x_1, x_2)\| \leq \|B\| \|x_2\|.$$

Par composition, $\partial_{x_1} B$ est continue sur $E_1 \times E_2$. Par un raisonnement symétrique, on peut montrer que $\partial_{x_2} B$ est continue sur $E_1 \times E_2$, donc par la proposition 1.7.1, B est \mathcal{C}^1 sur $E_1 \times E_2$. La différentielle est alors donnée par :

$$\begin{aligned} B'(x) \cdot v &= \left(\frac{\partial B}{\partial x_1}(x) \circ p_1 + \frac{\partial B}{\partial x_2}(x) \circ p_2 \right) \cdot v = \frac{\partial B}{\partial x_1}(x) \cdot p_1(v) + \frac{\partial B}{\partial x_2}(x) \cdot p_2(v) \\ &= \frac{\partial B}{\partial x_1}(x) \cdot v_1 + \frac{\partial B}{\partial x_2}(x) \cdot v_2 = B(v_1, x_2) + B(x_1, v_2) \end{aligned}$$

avec $x := (x_1, x_2)$ et $v := (v_1, v_2)$ dans $\mathcal{L}_2(E_1 \times E_2, F)$. On peut noter que l'on peut utiliser la notation matricielle et écrire :

$$B'(x) \cdot v = \begin{pmatrix} \frac{\partial B}{\partial x_1}(x) & \frac{\partial B}{\partial x_2}(x) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = B(v_1, x_2) + B(x_1, v_2).$$

ii) La démonstration précédente se généralise bien au cas multilinéaire. On a donc

$$\mathcal{L}_n(\prod_{i=1}^n E_i, F) \subset \mathcal{C}^1(\prod_{i=1}^n E_i, F),$$

avec $\mathcal{L}_n(\prod_{i=1}^n E_i, F)$ l'ensemble des applications multilinéaire $\prod_{i=1}^n E_i$ muni de la norme

$$\|T\|_{\mathcal{L}_n(\prod_{i=1}^n E_i, F)} := \sup \{ \|T(x_1, \dots, x_n)\|_F \mid \|x_1\|_{E_1} \leq 1, \dots, \|x_n\|_{E_n} \leq 1 \}.$$

La différentielle de $T \in \mathcal{L}_n(\prod_{i=1}^n E_i, F)$ est donnée par

$$T'(x) \cdot v = \begin{pmatrix} \frac{\partial T}{\partial x_1}(x) & \cdots & \frac{\partial T}{\partial x_n}(x) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^n T(x_1, \dots, x_{i-1}, v_i, x_{i+1}, \dots, x_n),$$

avec $x := (x_1, \dots, x_n)$ et $v := (v_1, \dots, v_n)$ dans $\mathcal{L}_n(\prod_{i=1}^n E_i, F)$.

□

Corollaire 1.7.2

Soient E_1, E_2, E et F quatre espaces vectoriels normés, U un ouvert de E et $\bar{x} \in U$. Soient deux applications $f_1: U \rightarrow E_1$ et $f_2: U \rightarrow E_2$ dérivable en \bar{x} , et une application bilinéaire continue $B \in \mathcal{L}_2(E_1 \times E_2, F)$. Alors l'application

$$\begin{aligned} \Psi: U &\longrightarrow F \\ x &\longmapsto \Psi(x) := B(f_1(x), f_2(x)) \end{aligned}$$

est dérivable en \bar{x} et pour tout $v \in E$, $\Psi'(\bar{x}) \cdot v = B(f'_1(\bar{x}) \cdot v, f_2(\bar{x})) + B(f_1(\bar{x}), f'_2(\bar{x}) \cdot v)$.

► On peut écrire $\Psi := B \circ f$, avec

$$\begin{aligned} f: U &\longrightarrow E_1 \times E_2 \\ x &\longmapsto f(x) := (f_1(x), f_2(x)). \end{aligned}$$

D'après la proposition 1.5.1, f dérivable en \bar{x} car f_1 et f_2 le sont. Puisque $B \in \mathcal{L}_2(E_1 \times E_2, F)$, alors B est dérivable sur $E_1 \times E_2$ (cf. exercice 1.7.1), donc en particulier en $f(\bar{x})$. Par composition, cf. théorème 1.3.3, Ψ est dérivable en \bar{x} et pour tout $v \in E$, $\Psi'(\bar{x}) \cdot v = B'(f(\bar{x})) \cdot (f'(\bar{x}) \cdot v)$, avec $f'(\bar{x}) \cdot v = (f'_1(\bar{x}) \cdot v, f'_2(\bar{x}) \cdot v)$ d'après l'équation (1.2). Par l'exercice 1.7.1, on obtient

$$\Psi'(\bar{x}) \cdot v = B'(f(\bar{x})) \cdot (f'(\bar{x}) \cdot v) = B(f'_1(\bar{x}) \cdot v, f_2(\bar{x})) + B(f_1(\bar{x}), f'_2(\bar{x}) \cdot v).$$

■

Remarque 1.7.2. Si $E = E_1 = E_2 = F = \mathbb{R}$, et si B est le produit usuel \times entre réels, alors la formule du corollaire 1.7.2 n'est rien d'autre que la formule usuelle de dérivation du produit de deux fonctions réelles f_1 et f_2 de la variable réelle :

$$(f_1 \times f_2)'(x) = f'_1(x) \times f_2(x) + f_1(x) \times f'_2(x).$$

Remarque 1.7.3. Soit $B \in \mathcal{L}_2(E^2, F)$, E et F deux espaces vectoriels normés. On définit l'application quadratique associée à B :

$$\begin{aligned} Q: E &\longrightarrow F \\ x &\longmapsto Q(x) := B(x, x). \end{aligned}$$

D'après le corollaire 1.7.2 ($U = E_1 = E_2 = E$ et $f_1 = f_2 = \text{Id}_E$), l'application Q est de classe \mathcal{C}^1 sur E et pour tout x, v dans E , on a :

$$Q'(x) \cdot v = B(x, v) + B(v, x),$$

et si B est symétrique, alors $Q'(x) \cdot v = 2B(x, v)$.

Le corollaire précédent se généralise bien aux applications multilinéaires. Considérons par exemple un espace vectoriel normé E de dimension n muni d'une base $B := (e_1, \dots, e_n)$ et notons \det , l'application déterminant en base B . Cette application est une forme n -linéaire alternée continue (E^n est de dimension finie) et sa dérivée est donnée par l'exercice 1.7.1. Pour

x et v dans E^n , on a :

$$\det'(x) \cdot v = \sum_{i=1}^n \det(x_1, \dots, x_{i-1}, v_i, x_{i+1}, \dots, x_n).$$

En utilisant la notation matricielle $A := (x_1 \dots x_n)$ et $H := (v_1 \dots v_n)$, on peut écrire :

$$\det'(A) \cdot H = \text{tr}((\text{com } A)^T H),$$

autrement dit, $\det(A + H) = \det(A) + \text{tr}((\text{com } A)^T H) + o(H)$. On rappelle que $\text{com } A$ est la comatrice (ou matrice de cofacteurs) de A . Elle est de même dimension que A et elle est donnée par

$$(\text{com } A)_{i,j} := \det(\tilde{A}_{i,j}) = (-1)^{i+j} \det(A_{i,j}),$$

où $\tilde{A}_{i,j}$ est la matrice carrée de taille n déduite de A en remplaçant la j^{e} colonne de A par une colonne constituée uniquement de zéros, sauf un 1 sur la i^{e} ligne, et où $A_{i,j}$ est la sous-matrice carrée de taille $n-1$ déduite de A en supprimant la i^{e} ligne et la j^{e} colonne. Soient maintenant un autre espace vectoriel normé F et une application $f: F \rightarrow E^n$, dérivable en un point $t \in F$. Alors, l'application $\Psi := \det \circ f$ est dérivable en t et pour tout $h \in F$, on a :

$$\Psi'(t) \cdot h = \sum_{i=1}^n \frac{\partial \det}{\partial x_i}(f(t)) \cdot (f'(t) \cdot h) = \sum_{i=1}^n \det(f_1(t), \dots, f_{i-1}(t), f'_i(t) \cdot h, f_{i+1}(t), \dots, f_n(t)).$$

1.8 Dérivabilité dans le cas fonctionnel

On s'intéresse ici à la dérivabilité de la composition à droite d'une application \mathcal{C}^1 avec une application continue définie sur un intervalle compact de \mathbb{R} . Concrètement, on considère deux espaces vectoriels normés $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$, et une application $f \in \mathcal{C}^1(E, F)$. On définit l'application

$$\begin{aligned} \Phi_f: \mathcal{C}^0([0, 1], E) &\longrightarrow \mathcal{C}^0([0, 1], F) \\ g &\longmapsto \Phi_f(g) := f \circ g, \end{aligned}$$

où $\mathcal{C}^0([0, 1], E)$ et $\mathcal{C}^0([0, 1], F)$ sont munis de la norme de la convergence uniforme $\|\cdot\|_\infty$,⁷ et on s'intéresse à la dérivabilité de Φ_f . Si f est linéaire, *i.e.* $f \in \mathcal{L}(E, F)$, alors la réponse est plutôt simple : l'application Φ_f est linéaire (évident) et continue puisque

$$\forall t \in [0, 1], \|\Phi_f(g)(t)\|_F = \|(f(g(t)))\|_F \leq \|f\|_{\mathcal{L}(E, F)} \|g(t)\|_E \leq \|f\|_{\mathcal{L}(E, F)} \|g\|_\infty$$

donc

$$\|\Phi_f(g)\|_\infty \leq \|f\|_{\mathcal{L}(E, F)} \|g\|_\infty.$$

D'après les exemples fondamentaux, cf. exercice 1.3.1, Φ_f est donc de classe \mathcal{C}^1 et sa différentielle est donnée pour tout g et v dans $\mathcal{C}^0([0, 1], E)$, par :

$$\Phi'_f(g) \cdot v = \Phi_f(v) = f \circ v.$$

Dans le cas non linéaire, nous avons le résultat suivant.

7. Il y a ici un léger abus de notation (sans conséquences) car dans $\mathcal{C}^0([0, 1], E)$, la norme sup est donnée par $\|g\|_\infty = \sup_{t \in [0, 1]} (\|g(t)\|_E)$, tandis que dans $\mathcal{C}^0([0, 1], F)$, on a $\|h\|_\infty = \sup_{t \in [0, 1]} (\|h(t)\|_F)$.

Proposition 1.8.1

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Soit $f \in \mathcal{C}^1(E, F)$. On définit l'application

$$\begin{aligned} \Phi_f: \mathcal{C}^0([0, 1], E) &\longrightarrow \mathcal{C}^0([0, 1], F) \\ g &\longmapsto \Phi_f(g) := f \circ g, \end{aligned}$$

où $\mathcal{C}^0([0, 1], E)$ et $\mathcal{C}^0([0, 1], F)$ sont munis de la norme de la convergence uniforme $\|\cdot\|_\infty$. Alors, l'application Φ_f est de classe \mathcal{C}^1 et pour tout g, v dans $\mathcal{C}^0([0, 1], E)$, la différentielle est donnée par :

$$\begin{aligned} \Phi'_f(g) \cdot v: [0, 1] &\longrightarrow F \\ t &\longmapsto (\Phi'_f(g) \cdot v)(t) = f'(g(t)) \cdot v(t). \end{aligned}$$

Remarque 1.8.1. On a donc $\Phi_f(g)(t) = f(g(t))$ et sa différentielle est donnée par $(\Phi'_f(g) \cdot v)(t) = f'(g(t)) \cdot v(t) = (f' \circ g)(t) \cdot v(t)$ que l'on peut donc noter $\Phi'_f(g) \cdot v = (f' \circ g) \cdot v$.

Nous aurons besoin du lemme suivant.

Lemme 1.8.1. Soient $F: (X, d_X) \rightarrow (Y, d_Y)$ une application continue entre deux espaces métriques et $K \subset X$ un sous-ensemble compact. On pose pour $\delta \in \mathbb{R}_+$:

$$M(\delta) := \sup \{d_Y(F(x_1), F(x_2)) \mid x_1 \in K, x_2 \in X, d_X(x_1, x_2) \leq \delta\}.$$

Alors,

$$M(\delta) = o(1) \text{ quand } \delta \rightarrow 0.$$

► Soit $\varepsilon > 0$. Si la conclusion est fausse, alors il existe deux suites $(x_{1,n})$ dans K et $(x_{2,n})$ dans X telles que $d_X(x_{1,n}, x_{2,n}) \rightarrow 0$ et $d_Y(F(x_{1,n}), F(x_{2,n})) > \varepsilon$ pour tout n . Quitte à extraire une sous-suite, on peut supposer que $x_{1,n} \rightarrow x \in K$. Alors, $x_{2,n} \rightarrow x$ et puisque F est continue sur X , $F(x_{1,n})$ et $F(x_{2,n})$ converge vers $F(x)$ d'où la contradiction. ■

Nous pouvons maintenant démontrer le théorème.

► Soit $g \in \mathcal{C}^0([0, 1], E)$. On introduit dans $\mathcal{C}^0([0, 1], E)$, l'application

$$h(v) := \Phi_f(g + v) - \Phi_f(g) - (f' \circ g) \cdot v = f \circ (g + v) - f \circ g - (f' \circ g) \cdot v.$$

* L'idée est donc de montrer dans un premier temps que $h(v) = o(v)$. Soit $v \in \mathcal{C}^0([0, 1], E)$. On introduit l'application

$$\begin{aligned} \varphi: [0, 1]^2 &\longrightarrow F \\ (t, s) &\longmapsto \varphi(t, s) := h(sv)(t) = f(g(t) + sv(t)) - f(g(t)) - s f'(g(t)) \cdot v(t) \end{aligned}$$

de telle sorte que $\|h(v)\|_\infty = \sup_{t \in [0, 1]} (\|\varphi(t, 1)\|_F)$. Pour $t \in [0, 1]$ fixé, l'application partielle $s \mapsto \varphi(t, s)$ est continue sur $[0, 1]$, dérivable sur $]0, 1[$, de dérivée

$$\frac{\partial \varphi}{\partial s}(t, s) = (f'(g(t) + sv(t)) - f'(g(t))) \cdot v(t),$$

donc d'après le théorème 1.6.5 des accroissements finis :

$$\begin{aligned} \|\varphi(t, 1) - \varphi(t, 0)\|_F &= \|\varphi(t, 1)\|_F \leq \sup_{s \in]0, 1[} (\|(f'(g(t) + sv(t)) - f'(g(t))) \cdot v(t)\|_F) \\ &\leq \sup_{s \in]0, 1[} (\|(f'(g(t) + sv(t)) - f'(g(t)))\|_{\mathcal{L}(E, F)}) \|v\|_\infty. \end{aligned}$$

Ainsi,

$$\|h(v)\|_\infty \leq \sup_{(t,s) \in [0,1]^2} (\|(f'(g(t) + sv(t)) - f'(g(t)))\|_{\mathcal{L}(E,F)}) \|v\|_\infty \leq M(\|v\|_\infty) \|v\|_\infty,$$

où l'on a introduit pour $\delta \in \mathbb{R}_+$:

$$M(\delta) := \sup \left\{ \|f'(y) - f'(x)\|_{\mathcal{L}(E,F)} \mid x \in g([0,1]), y \in E, \|y - x\|_E \leq \delta \right\}. \quad (1.4)$$

On applique le lemme précédent avec $K := g([0,1])$ compact de E (car g continue) et $F := f'$ continue (car f de classe \mathcal{C}^1). Ainsi, $M(\|v\|_\infty) = o(1)$ quand $\|v\|_\infty \rightarrow 0$, donc $M(\|v\|_\infty) \|v\|_\infty = o(\|v\|_\infty)$ et puisque $\|h(v)\|_\infty \leq M(\|v\|_\infty) \|v\|_\infty$, finalement on obtient bien que $h(v) = o(v)$.

* Montrons que $v \mapsto (f' \circ g) \cdot v$ est une application linéaire continue, car alors, on aura que Φ_f est dérivable en g de différentielle donnée par $\Phi'_f(g) \cdot v = (f' \circ g) \cdot v$. Il est clair que $v \mapsto (f' \circ g) \cdot v$ est linéaire. Elle est continue car, pour tout $v \in \mathcal{C}^0([0,1], E)$, on a

$$\forall t \in [0,1], \|((f' \circ g) \cdot v)(t)\|_F \leq \|f'(g(t))\|_{\mathcal{L}(E,F)} \|v(t)\|_E \leq \|f' \circ g\|_\infty \|v\|_\infty$$

donc $\|(f' \circ g) \cdot v\|_\infty \leq \|f' \circ g\|_\infty \|v\|_\infty$. On a utilisé le fait que $f' \circ g$ est continue sur le compact $[0,1]$ donc $\|f' \circ g\|_\infty < \infty$. Finalement, on a bien $\Phi'_f(g) \cdot v = (f' \circ g) \cdot v$.

* Montrons enfin que Φ_f est de classe \mathcal{C}^1 , i.e. que Φ'_f est continue sur $\mathcal{C}_E^0 := \mathcal{C}^0([0,1], E)$. Soit $g \in \mathcal{C}_E^0$. Montrons que $\|\Phi'_f(g+v) - \Phi'_f(g)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} = o(1)$ quand $\|v\|_\infty \rightarrow 0$, où l'on a noté $\mathcal{C}_F^0 := \mathcal{C}^0([0,1], F)$. Puisque

$$\|\Phi'_f(g+v) - \Phi'_f(g)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} = \sup_{\|u\|_\infty=1} \|(\Phi'_f(g+v) - \Phi'_f(g)) \cdot u\|_\infty,$$

et puisque

$$\begin{aligned} \|(\Phi'_f(g+v) - \Phi'_f(g)) \cdot u\|_\infty &= \sup_{t \in [0,1]} \|(f'(g(t) + v(t)) - f'(g(t))) \cdot u(t)\|_F \\ &\leq \sup_{t \in [0,1]} \|(f'(g(t) + v(t)) - f'(g(t)))\|_{\mathcal{L}(E,F)}, \end{aligned}$$

on obtient donc

$$\|\Phi'_f(g+v) - \Phi'_f(g)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \sup_{t \in [0,1]} \|(f'(g(t) + v(t)) - f'(g(t)))\|_{\mathcal{L}(E,F)} \leq M(\|v\|_\infty),$$

avec M définie comme en (1.4) et puisque $M(\|v\|_\infty) = o(1)$ quand $\|v\|_\infty \rightarrow 0$ d'après le lemme précédent, on obtient le résultat voulu. ■

Remarque 1.8.2. Montrons d'une autre manière que Φ_f est de classe \mathcal{C}^1 , i.e. que Φ'_f est continue sur $\mathcal{C}_E^0 := \mathcal{C}^0([0,1], E)$. Soient $a \in \mathcal{C}_E^0$ et $\varepsilon > 0$. Il faut montrer qu'il existe $\eta_{\varepsilon,a} > 0$ tel que pour tout $g \in \mathcal{C}_E^0$:

$$\|g - a\|_\infty \leq \eta_{\varepsilon,a} \implies \|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \varepsilon,$$

où $\mathcal{C}_F^0 := \mathcal{C}^0([0, 1], F)$. Or,

$$\|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} = \sup_{\|v\|_\infty=1} \|(\Phi'_f(g) - \Phi'_f(a)) \cdot v\|_\infty,$$

et

$$\begin{aligned} \|(\Phi'_f(g) - \Phi'_f(a)) \cdot v\|_\infty &= \sup_{t \in [0, 1]} \|(f'(g(t)) - f'(a(t))) \cdot v(t)\|_F \\ &\leq \sup_{t \in [0, 1]} \|f'(g(t)) - f'(a(t))\|_{\mathcal{L}(E, F)}, \end{aligned}$$

donc

$$\|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \sup_{t \in [0, 1]} \|f'(g(t)) - f'(a(t))\|_{\mathcal{L}(E, F)}.$$

Mais par hypothèse, $f \in \mathcal{C}^1(E, F)$ donc $f' \in \mathcal{C}^0(E, \mathcal{L}(E, F))$, donc f' est uniformément continue sur le compact $K_a := a([0, 1]) \subset E$ et ainsi pour le $\varepsilon > 0$ choisi : $\exists \eta_{\varepsilon, K_a} > 0$, tel que $\forall (x, y) \in K_a \times E : \|y - x\|_E \leq \eta_{\varepsilon, K_a} \Rightarrow \|f'(y) - f'(x)\|_F \leq \varepsilon$. Alors en prenant $\eta_{\varepsilon, a} := \eta_{\varepsilon, K_a}$, on voit que si $\|g - a\|_\infty \leq \eta_{\varepsilon, a}$, alors pour tout $t \in [0, 1]$, $\|g(t) - a(t)\|_E \leq \eta_{\varepsilon, K_a}$ donc $\|f'(g(t)) - f'(a(t))\|_F \leq \varepsilon$. D'où : $\|g - a\|_\infty \leq \eta_{\varepsilon, a} \Rightarrow \|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \varepsilon$.

Donnons un corollaire à ce résultat, utile pour démontrer la différentiabilité de solutions d'équations différentielles ordinaires par rapport par exemple à la condition initiale. Dans ce corollaire, nous reprenons des notations plus proches des équations différentielles et nous appliquons la proposition précédente avec $E = F = \mathbb{R}^n$ muni de la norme euclidienne usuelle. Il est à noter que l'on peut bien évidemment définir f sur un ouvert Ω de \mathbb{R}^n et considérer des applications $x(\cdot)$ à valeurs dans Ω .

Corollaire 1.8.2

On note $X := (\mathcal{C}^0([0, 1], \mathbb{R}^n), \|\cdot\|_\infty)$. Soit $f \in \mathcal{C}^1(\mathbb{R}^n)$. On définit $F: X \rightarrow X$ par

$$\begin{aligned} F(x): [0, 1] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto F(x)(t) := f(x(t)), \end{aligned}$$

c'est-à-dire, $F(x) = f \circ x$. Alors, l'application $F \in \mathcal{C}^1(X)$ et pour tout x, v dans X :

$$\begin{aligned} F'(x) \cdot v: [0, 1] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto (F'(x) \cdot v)(t) = f'(x(t)) \cdot v(t). \end{aligned}$$

Différentielle d'ordre supérieur

2.1	Différentielle du second ordre	33
2.2	Différentielle d'ordre k	35

Ce chapitre n'est pour le moment qu'une ébauche.

2.1 Différentielle du second ordre

Définition 2.1.1 – Dérivée seconde

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, U un ouvert de E , une application $f: U \subset E \rightarrow F$ et un point $x \in U$. On dit que f est *2 fois dérivable* (ou 2 fois différentiable) au point x si f est dérivable sur un ouvert $\Omega \subset U$ contenant x et si $f'|_{\Omega}$, l'application dérivée de f restreinte à Ω est dérivable en x . Si f est 2 fois dérivable en x alors $f''(x) := (f')'(x) \in \mathcal{L}(E, \mathcal{L}(E, F))$ est appelée la *dérivée seconde de f au point x* .

Comme il n'est pas très commode de considérer des applications à valeurs dans un espace d'applications, on utilise habituellement, à la place de $\mathcal{L}(E, \mathcal{L}(E, F))$, un espace qui lui est isomorphe, même mieux isomorphe et isométrique. Notons $\mathcal{L}_2(E \times E, F)$ l'ensemble des applications bilinéaires continues de $E \times E$ dans F et munissons cet espace de la norme

$$\|B\|_{\mathcal{L}_2(E \times E, F)} := \sup \{ \|B(u, v)\|_F \mid \|u\|_E \leq 1 \text{ et } \|v\|_E \leq 1 \}.$$

Introduisons maintenant

$$\begin{aligned} \varphi: \mathcal{L}(E, \mathcal{L}(E, F)) &\longrightarrow \mathcal{L}_2(E \times E, F) \\ T &\longmapsto \varphi(T), \end{aligned}$$

telle que

$$\begin{aligned} \varphi(T): E \times E &\longrightarrow F \\ (u, v) &\longmapsto \varphi(T) \cdot (u, v) := (T \cdot u) \cdot v = (T(u))(v) \end{aligned}$$

et montrons que φ est un isomorphisme isométrique, auquel cas $\varphi(T) \simeq T$, on pourra identifier $\mathcal{L}(E, \mathcal{L}(E, F))$ avec $\mathcal{L}_2(E \times E, F)$ et on pourra donc noter $T(u, v) := \varphi(T) \cdot (u, v)$. En particulier, on aura $f''(x) \in \mathcal{L}_2(E^2, F)$ et pour tout $(u, v) \in E^2$, on notera $f''(x) \cdot (u, v)$ la dérivée seconde de f en x appliquée aux vecteurs u et v , au lieu de $f''(x) \cdot u \cdot v$. Montrons donc que φ est un isomorphisme isométrique, pour cela il suffit de montrer que φ est linéaire, surjective et isométrique. Tout d'abord, on vérifie que φ est bien définie, ce qui est le cas, car $\varphi(T)$ est bien une application bilinéaire continue de E^2 dans F . Ensuite, il est clair que φ est linéaire car un ensemble d'applications linéaires continues est un sous-espace vectoriel. Soit maintenant

$B \in \mathcal{L}_2(E^2, F)$. Alors l'application $T_B: E \rightarrow \mathcal{L}(E, F)$, $T_B(u) := B(u, \cdot)$, est bien linéaire, définie sur E et à valeurs dans $\mathcal{L}(E, F)$ donc φ est surjective. Enfin,

$$\begin{aligned} \|\varphi(T)\|_{\mathcal{L}_2(E^2, F)} &= \sup \{ \|(T \cdot u) \cdot v\|_F \mid \|u\|_E \leq 1 \text{ et } \|v\|_E \leq 1 \} \\ &= \sup_{\|u\|_E \leq 1} \left(\sup_{\|v\|_E \leq 1} \|(T \cdot u) \cdot v\|_F \right) = \sup_{\|u\|_E \leq 1} \|T \cdot u\|_{\mathcal{L}(E, F)} = \|T\|_{\mathcal{L}(E, \mathcal{L}(E, F))}, \end{aligned}$$

donc φ est bien isométrique. En conclusion, φ est un isomorphisme (topologique) isométrique.

Remarque 2.1.1. Nous n'avons pas besoin de montrer que φ est injective car le fait que φ soit isométrique implique son injectivité. En effet, si $\varphi(T_1) = \varphi(T_2)$, alors $\varphi(T_1 - T_2) = 0$ car φ linéaire, donc par isométrie, $\|T_1 - T_2\| = 0$, donc $T_1 = T_2$.

Remarque 2.1.2. Si f est 2 fois dérivable en x , alors pour $v \in E$ proche de x :

$$f'(x+v) = f'(x) + (f')'(x) \cdot v + o(v),$$

autrement dit on a pour $u \in E$,

$$f'(x+v) \cdot u = f'(x) \cdot u + f''(x) \cdot (v, u) + o(v).$$

Il en résulte que l'application $g: x \mapsto g(x) := f'(x) \cdot u$ vérifie $g'(x) \cdot v = f''(x) \cdot (v, u)$. D'un point de vue pratique, la dérivée seconde en un point x appliquée à une paire (v, u) peut être obtenue ainsi : on calcule d'abord l'application dérivée appliquée en u , i.e. $x \mapsto f'(x) \cdot u$, puis on calcule la dérivée de cette application en x appliquée en v . Pour insister une fois de plus, au lieu de dériver 2 fois puis appliquer 2 fois, on dérive, puis applique, et on fait cela 2 fois.

Exemple 2.1.1 (Fonction définie sur un Hilbert et à valeurs dans \mathbb{R}). Soit $(H, (\cdot | \cdot)_H)$ un espace de Hilbert et $f: U \subset H \rightarrow \mathbb{R}$, U ouvert de H , une application 2 fois dérivable en $x \in U$. On a alors $f''(x) \in \mathcal{L}(H, \mathcal{L}(H, \mathbb{R})) \simeq \mathcal{L}(H, H) = \mathcal{L}(H)$, en utilisant l'isomorphisme entre H et $H' = \mathcal{L}(H, \mathbb{R})$, cf. exemple 1.2.3. Ainsi, $f''(x)$ s'identifie à un endomorphisme (topologique) de H , i.e. une application linéaire continue de H dans H . Construisons cet endomorphisme. Puisque $f''(x) \cdot (u, v) \simeq (f''(x) \cdot u) \cdot v$ et $f''(x) \cdot u \in \mathcal{L}(H, \mathbb{R}) = H'$, d'après le théorème 1.2.5 de Riesz, pour tout $u \in H$, il existe un unique vecteur $z_u \in H$ tel que $f''(x) \cdot (u, v) = (z_u | v)_H$. Posons

$$\begin{aligned} A_{f,x}: H &\longrightarrow H \\ u &\longmapsto A_{f,x}(u) := z_u \end{aligned}$$

et montrons que $A_{f,x} \in \mathcal{L}(H)$. Tout d'abord, $A_{f,x}$ est bien linéaire car

$$\begin{aligned} (z_{u_1+u_2} | v)_H &= f''(x) \cdot (u_1 + u_2, v) = f''(x) \cdot (u_1, v) + f''(x) \cdot (u_2, v) \\ &= (z_{u_1} | v)_H + (z_{u_2} | v)_H = (z_{u_1} + z_{u_2} | v)_H. \end{aligned}$$

De plus $A_{f,x}$ est continue car d'après le théorème de Riesz, l'isomorphisme étant isométrique on a $\|z_u\|_H = \|f''(x) \cdot u\|_{H'}$ et donc :

$$\|A_{f,x} \cdot u\|_H = \|z_u\|_H = \|f''(x) \cdot u\|_{H'} \leq \|f''(x)\|_{\mathcal{L}(H, H')} \|u\|_H.$$

L'endomorphisme dans $\mathcal{L}(H)$ que l'on vient de construire s'appelle *l'opérateur hessien de f en x* et est noté

$$\nabla^2 f(x) := A_{f,x}.$$

Nous avons même plus : $\|A_{f,x}\|_{\mathcal{L}(H)} = \|f''(x)\|_{\mathcal{L}(H,H')}$. Finalement, nous avons partiellement démontré que l'application

$$\begin{aligned} \varphi: \mathcal{L}(H, H') &\longrightarrow \mathcal{L}(H) \\ T &\longmapsto \varphi(T) := A_T, \end{aligned}$$

telle que $(T \cdot u) \cdot v = (A_T \cdot u | v)_H$, avec A_T définie comme précédemment à partir de T et non plus $f''(x)$, est un isomorphisme (topologique) isométrique. \square

Exemple 2.1.2 (Fonction définie sur \mathbb{R}^n et à valeurs dans \mathbb{R}). On munit \mathbb{R}^n du produit scalaire usuel $(\cdot | \cdot)$ de telle sorte que \mathbb{R}^n est un espace de Hilbert de dimension finie, *i.e.* un espace euclidien. Dans ce cas $f''(x) \in \mathcal{L}(\mathbb{R}^n) \simeq \mathbf{M}_n(\mathbb{R})$ et alors $f''(x) = \nabla^2 f(x)$ s'appelle la *matrice hessienne de f au point x* . Dans ce cas on a :

$$f''(x) \cdot (u, v) = v^T \nabla^2 f(x) u.$$

\square

2.2 Différentielle d'ordre k

On introduit la notation $\mathcal{E}_0 := F$, $\mathcal{E}_1 := \mathcal{L}(E, F) = \mathcal{L}(E, \mathcal{E}_0)$, puis par récurrence, pour tout $k \in \mathbb{N}$, $\mathcal{E}_{k+1} := \mathcal{L}(E, \mathcal{E}_k)$. Nous avons alors un isomorphisme isométrique entre \mathcal{E}_k et l'ensemble $\mathcal{L}_k(E^k, F)$ des applications k -linéaires de E^k dans F , défini par l'application :

$$\begin{aligned} \varphi: \mathcal{E}_k &\longrightarrow \mathcal{L}_k(E^k, F) \\ T &\longmapsto \varphi(T), \end{aligned}$$

telle que

$$\begin{aligned} \varphi(T): E^k &\longrightarrow F \\ x := (x_1, \dots, x_k) &\longmapsto \varphi(T) \cdot x := (((T \cdot x_1) \cdot x_2) \cdots) \cdot x_k, \end{aligned}$$

où $\mathcal{E}_k = \mathcal{L}(E, \mathcal{E}_{k-1})$ est muni de la norme $\|\cdot\|_{\mathcal{L}(E, \mathcal{E}_{k-1})}$, et où l'ensemble $\mathcal{L}_k(E^k, F)$ est muni de la norme

$$\|T\|_{\mathcal{L}_k(E^k, F)} := \sup \{ \|T(x_1, \dots, x_k)\|_F \mid \|x_1\|_E \leq 1, \dots, \|x_k\|_E \leq 1 \}.$$

Puisque ces espaces sont isomorphes, on a $\varphi(T) \simeq T$ et on peut donc noter $T(x)$ ou $T \cdot x$ à la place de $((T \cdot x_1) \cdot x_2) \cdots \cdot x_k$.

Faisons l'hypothèse de récurrence suivante : on suppose définies les notions d'applications $k-1$ fois dérivable en x , et de dérivée $(k-1)^e$ en x , notée $f^{(k-1)}(x) \in \mathcal{E}_{k-1} \simeq \mathcal{L}_{k-1}(E^{k-1}, F)$. Nous pouvons introduire les définitions suivantes.

Définition 2.2.1 – Dérivée d'ordre k et applications de classe \mathcal{C}^k , $k \in \overline{\mathbb{N}}$

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, U un ouvert de E , une application $f: U \subset E \rightarrow F$ et un point $x \in U$.

- On dit que f est k fois dérivable (ou k fois différentiable) au point x si f est $k-1$ fois dérivable sur un ouvert $\Omega \subset U$ contenant x et si $f|_{\Omega}^{(k-1)}$, l'application dérivée

$(k-1)^e$ de f restreinte à Ω est dérivable en x .

Si f est k fois dérivable en x alors $f^{(k)}(x) := (f^{(k-1)})'(x) \in \mathcal{L}(E, \mathcal{E}_{k-1}) = \mathcal{E}_k \simeq \mathcal{L}_k(E^k, F)$ est appelée la *dérivée k^e de f au point x* .

- On dit que f est de classe \mathcal{C}^k , $k \in \mathbb{N}$, si f est k fois dérivable dans U et si son application dérivée k^e , $f^{(k)}$, est continue de U dans \mathcal{E}_k .

On note $\mathcal{C}^k(U, F) := \{f \in F^U \mid f^{(k)} \in \mathcal{C}^0(U, \mathcal{E}_k)\}$, ssev de $\mathcal{C}^{k-1}(U, F)$.

- On dit que f est de classe \mathcal{C}^∞ (ou *infiniment dérivable*), si f est de classe \mathcal{C}^k , pour tout k dans \mathbb{N} .

On note $\mathcal{C}^\infty(U, F) := \cap_{k \in \mathbb{N}} \mathcal{C}^k(U, F)$

Remarque 2.2.1. Bien entendu, une application de classe \mathcal{C}^k est aussi de classe \mathcal{C}^j , pour tout $j \leq k$.



Exercice 2.2.1 (Cas fondamentaux). Montrer que toute application constante, linéaire continue, affine continue, bilinéaire continue, multilinéaire continue, quadratique associée à une application bilinéaire continue et tout carré de norme hilbertienne, est de classe \mathcal{C}^∞ . Donner pour tout $k \in \mathbb{N}$, leurs applications dérivées k^e . \square

Inversion locale et équations implicites

3.1	Théorème du point fixe	37
3.2	Théorème d'inversion locale	39
3.3	Théorème des fonctions implicites	42

3.1 Théorème du point fixe

Définition 3.1.1

Soient (X, d_X) et (Y, d_Y) deux espaces métriques. Une application $f: X \rightarrow Y$ est dite *k-lipschitzienne*, $k \geq 0$ dans \mathbb{R} , si pour tout x_1, x_2 dans X ,

$$d_Y(f(x_1), f(x_2)) \leq k d_X(x_1, x_2).$$

$f: X \rightarrow X$ est dite *contractante* si elle est *k-lipschitzienne* pour un $k \in [0, 1[$.

Théorème 3.1.2 – du point fixe de Picard

Soit $f: X \rightarrow X$ une application contractante sur un espace métrique complet (X, d) non vide. Il existe alors un unique point fixe $x \in X$ de f , c'est-à-dire tel que $f(x) = x$.

► On suppose f contractante de constante de Lipschitz $k \in [0, 1[$.

* (Unicité). Soit $(x_1, x_2) \in X^2$ t.q. $f(x_1) = x_1$ et $f(x_2) = x_2$. Alors

$$d(f(x_1), f(x_2)) = d(x_1, x_2) \leq k d(x_1, x_2) < d(x_1, x_2),$$

ce qui n'est possible que pour $x_1 = x_2$. On a donc l'unicité.

* (Existence). On fixe $x_0 \in X$ (X est non vide). On définit pour $n \in \mathbb{N}$, $x_{n+1} := f(x_n)$. Montrons que la suite $(x_n)_n$ est de Cauchy. Nous avons pour $n \in \mathbb{N}$:

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq k d(x_n, x_{n-1}) \leq k^n d(x_1, x_0) =: C k^n.$$

Ainsi, pour $(n, p) \in \mathbb{N}^2$:

$$\begin{aligned} d(x_{n+p}, x_n) &\leq \sum_{i=0}^{p-1} d(x_{n+i+1}, x_{n+i}) && \text{(d'après l'inégalité triangulaire)} \\ &\leq C \sum_{i=0}^{p-1} k^{n+i} = C k^n \frac{1 - k^p}{1 - k} \leq \frac{C}{1 - k} k^n. && \text{(car } k \in [0, 1]) \end{aligned}$$

Soit $\varepsilon > 0$. On choisit $N_\varepsilon \in \mathbb{N}$ t.q.

$$\frac{C}{1-k} k^{N_\varepsilon} \leq \varepsilon,$$

alors pour tous $n \geq N_\varepsilon$ et $p \in \mathbb{N}$, on a

$$d(x_{n+p}, x_n) \leq \frac{C}{1-k} k^n \leq \frac{C}{1-k} k^{N_\varepsilon} \leq \varepsilon,$$

c'est-à-dire $(x_n)_n$ est une suite de Cauchy. Puisque X est complet, la suite $(x_n)_n$ converge vers un élément de X noté x . Posons $u_n := f(x_n) - x_n$. La suite $(u_n)_n$ converge vers $f(x) - x$ (car f continue). Mais puisque $u_n = f(x_n) - x_n = f(x_n) - f(x_{n-1})$ (pour $n \geq 1$), la suite $(u_n)_n$ converge vers $f(x) - f(x) = 0$. Par unicité de la limite, on obtient $f(x) = x$, c'est-à-dire l'existence d'un point fixe. ■

Théorème 3.1.3 – du point fixe de Picard à paramètre

Soient (X, d) un espace métrique complet, Λ un espace topologique, $\varphi: X \times \Lambda \rightarrow X$ une application continue et $0 < k < 1$. Supposons que pour tout $\lambda \in \Lambda$, l'application $\varphi(\cdot, \lambda)$ soit k -contractante. Alors, pour tout $\lambda \in \Lambda$, il existe un unique point fixe $x(\lambda) \in X$ de $\varphi(\cdot, \lambda)$ et l'application $x(\cdot): \Lambda \rightarrow X$ est continue.

► On définit $(\mathcal{C}_b^0(\Lambda, X), d_\infty)$, l'ensemble des applications continues et bornées de Λ dans X muni de la distance de la convergence uniforme :

$$d_\infty(f, g) = \sup_{\lambda \in \Lambda} d(f(\lambda), g(\lambda)).$$

Puisque (X, d) est complet, $(\mathcal{C}_b^0(\Lambda, X), d_\infty)$ l'est aussi. On définit

$$\begin{aligned} \Phi: (\mathcal{C}_b^0(\Lambda, X), d_\infty) &\longrightarrow (\mathcal{C}_b^0(\Lambda, X), d_\infty) \\ f &\longmapsto \Phi(f) := \varphi(f(\cdot), \cdot). \end{aligned}$$

Ainsi,

$$d_\infty(\Phi(f), \Phi(g)) = \sup_{\lambda \in \Lambda} d(\Phi(f)(\lambda), \Phi(g)(\lambda)) = \sup_{\lambda \in \Lambda} d(\varphi(f(\lambda), \lambda), \varphi(g(\lambda), \lambda)),$$

or

$$d(\varphi(f(\lambda), \lambda), \varphi(g(\lambda), \lambda)) \leq k d(f(\lambda), g(\lambda)) \leq k d_\infty(f, g),$$

donc

$$d_\infty(\Phi(f), \Phi(g)) \leq k d_\infty(f, g),$$

c'est-à-dire, Φ est k -contractante. D'après le théorème du point fixe de Picard 3.1.2, Φ admet un unique point fixe noté x . Pour tout $\lambda \in \Lambda$, $\Phi(x)(\lambda) = \varphi(x(\lambda), \lambda) = x(\lambda)$. Puisque pour tout $\lambda \in \Lambda$, $\varphi(\cdot, \lambda)$ est k -contractante, elle admet un unique point fixe $x_\lambda \in X$, qui vérifie donc $x_\lambda = \varphi(x_\lambda, \lambda) = x(\lambda)$. Le théorème est démontré. ■

3.2 Théorème d'inversion locale

Définition 3.2.1 – Difféomorphisme, \mathcal{C}^k -difféomorphisme

Soient E, F deux espaces de Banach et U, V deux ouverts respectivement de E, F . Soit une application $f: U \rightarrow V$.

- i) On dit que f est un difféomorphisme de U dans V si f est une bijection de U dans V , dérivable dans U et dont l'inverse $f^{-1}: V \rightarrow U$ est dérivable dans V .
- ii) On dit que f est un \mathcal{C}^k -difféomorphisme, $k \in \overline{\mathbb{N}}^*$, de U dans V si f est une bijection de U dans V , de classe \mathcal{C}^k et dont l'inverse $f^{-1}: V \rightarrow U$ est de classe \mathcal{C}^k .

Un difféomorphisme, et *a fortiori* un \mathcal{C}^k -difféomorphisme, est un homéomorphisme.

Remarque 3.2.1. En fait, on peut voir un homéomorphisme, un difféomorphisme ou un \mathcal{C}^k -difféomorphisme comme une application inversible, respectivement, dans l'ensemble des applications continues, dérivables ou de classe \mathcal{C}^k .

Proposition 3.2.2

Un homéomorphisme $f: U \rightarrow V$ est un \mathcal{C}^k -difféomorphisme, $k \in \overline{\mathbb{N}}^*$, de U dans V si et seulement si f est de classe \mathcal{C}^k et, pour tout $x \in U$, $f'(x) \in \mathcal{L}(E, F)$ est bijective.

► Soit un homéomorphisme $f: U \rightarrow V$.

1. Supposons que f soit un \mathcal{C}^k -difféomorphisme de U dans V . En particulier, f est \mathcal{C}^k . En particulier aussi, f est dérivable sur U et son inverse $f^{-1}: V \rightarrow U$ est elle dérivable sur V . Montrons que $f'(x)$ est bijective. Par hypothèse,

$$f^{-1} \circ f = \text{Id}_U \quad \text{et} \quad f \circ f^{-1} = \text{Id}_V,$$

donc pour tout $x \in U$, en définissant $y := f(x) \in V$, on obtient par dérivation :

$$(f^{-1} \circ f)'(x) = (f^{-1})'(y) \circ f'(x) = \text{Id}_E \quad \text{et} \quad (f \circ f^{-1})'(y) = f'(x) \circ (f^{-1})'(y) = \text{Id}_F,$$

autrement dit $f'(x)$ est bijective avec $(f'(x))^{-1} = (f^{-1})'(y)$.

2. Supposons maintenant que f soit de classe \mathcal{C}^k et que pour tout $x \in U$, $f'(x)$ soit bijective. Nous voulons montrer que f est un \mathcal{C}^k -difféomorphisme, autrement dit que f est \mathcal{C}^k (ceci est vrai par hypothèse) et que f est une bijection (vraie par hypothèse) dont l'inverse $f^{-1}: V \rightarrow U$ est de classe \mathcal{C}^k .

2.a. Montrons dans un premier temps que $g := f^{-1}$ est dérivable sur V . D'après ce qui précède, si g est dérivable en $y \in V$, alors nécessairement $g'(y)$ serait donnée par $(f'(x))^{-1}$ où l'on a noté $x := f^{-1}(y) \in U$. Remarquons tout de suite que, d'après le théorème 1.1.5 de Banach, $f'(x) \in \text{Isom}_c(E, F)$, autrement dit $(f'(x))^{-1} \in \mathcal{L}(F, E)$. Il ne nous reste plus qu'à montrer que pour $y \in V$ donné :

$$\varphi(h) := g(y + h) - g(y) - T_y(h) = o(h),$$

où l'on a introduit la notation $T_y := (f'(x))^{-1} = (f'(f^{-1}(y)))^{-1}$ et où φ est définie pour h suffisamment petit.

* Posons $x_h := g(y + h) = f^{-1}(f(x) + h)$. Alors, $h = f(x_h) - f(x)$. Ainsi,

$$h = f(x_h) - f(x) = f'(x) \cdot (x_h - x) + o(x_h - x).$$

Introduisons la notation $\psi(x_h - x) := f(x_h) - f(x) - f'(x) \cdot (x_h - x)$. Puisque T_y est l'inverse de $f'(x)$, on obtient

$$T_y(h) = x_h - x + T_y(\psi(x_h - x)) = g(y + h) - g(y) + T_y(\psi(x_h - x)). \quad (3.1)$$

Supposons que l'on ait montré que $\psi(x_h - x)$, comme fonction de h , est un $o(h)$, alors on peut écrire $\psi(x_h - x) =: \|h\|_F \varepsilon(h)$ avec $\|\varepsilon(h)\|_F \rightarrow 0$ quand $\|h\|_F \rightarrow 0$. On obtiendrait finalement que

$$\|\varphi(h)\|_E = \|T_y(\psi(x_h - x))\|_E \leq \|T_y\|_{\mathcal{L}(F,E)} \|\varepsilon(h)\|_F \|h\|_F = o(h),$$

autrement dit que $\varphi(h) = o(h)$, c'est-à-dire ce que l'on voulait démontrer dans un premier temps, au point 2.a.

* Montrons donc que $\psi(x_h - x) = o(h)$. On sait déjà que $\psi(x_h - x) = o(x_h - x)$. On peut donc écrire $\psi(x_h - x) =: \|x_h - x\|_E \varepsilon(x_h - x)$, avec $\|\varepsilon(x_h - x)\|_F \rightarrow 0$ quand $\|x_h - x\|_E \rightarrow 0$. Puisque $g = f^{-1}$ est continue, on a $x_h - x = g(y + h) - g(y) = o(1)$ quand $\|h\|_F \rightarrow 0$. Supposons que l'on ait montré que $x_h - x = O(h)$,¹ alors il existe $\eta > 0$ et $C > 0$, tels que pour tout $\|h\|_F \leq \eta$:

$$\|\psi(x_h - x)\|_F = \|x_h - x\|_E \|\varepsilon(x_h - x)\|_F \leq C \|\varepsilon(x_h - x)\|_F \|h\|_F = o(h),$$

car $\|\varepsilon(x_h - x)\|_F \rightarrow 0$ quand $\|h\|_F \rightarrow 0$ (car $\|x_h - x\|_E \rightarrow 0$ quand $\|h\|_F \rightarrow 0$). On aurait donc $\psi(x_h - x) = o(h)$.

* Pour finir de prouver que $\varphi(h) = o(h)$, il nous reste à montrer que $x_h - x = O(h)$. Puisque $0 < \|T_y\|_{\mathcal{L}(F,E)}^{-1} < +\infty$ ($T_y \in \mathcal{L}(F,E)$ et $T_y \neq 0$) et puisque $\psi(x_h - x) = o(x_h - x)$, on a pour h suffisamment petit :

$$\|\psi(x_h - x)\|_F \leq \frac{1}{2} \|T_y\|_{\mathcal{L}(F,E)}^{-1} \|x_h - x\|_E.$$

Ainsi, d'après (3.1), on a :

$$\begin{aligned} \|x_h - x\|_E &\leq \|T_y\|_{\mathcal{L}(F,E)} \|h\|_F + \|T_y\|_{\mathcal{L}(F,E)} \|\psi(x_h - x)\|_F \\ &\leq \|T_y\|_{\mathcal{L}(F,E)} \|h\|_F + \frac{1}{2} \|x_h - x\|_E, \end{aligned}$$

ce qui nous donne $\|x_h - x\|_E \leq 2 \|T_y\|_{\mathcal{L}(F,E)} \|h\|_F$, autrement dit $x_h - x = O(h)$. En conclusion, on a bien $\psi(x_h - x) = o(h)$ ce qui entraîne $\varphi(h) = o(h)$. Le point 2.a est démontré : $g = f^{-1}$ est dérivable sur V .

2.b. L'application réciproque $g = f^{-1}$ admet donc une application dérivée (cf. 2.a) donnée par $g' : y \mapsto g'(y) = (f'(g(y)))^{-1}$. Introduisons

$$\begin{aligned} \text{inv} : \text{Isom}_c(E, F) &\longrightarrow \text{Isom}_c(F, E) \\ T &\longmapsto \text{inv}(T) := T^{-1}, \end{aligned}$$

de telle sorte que $g' = \text{inv} \circ f' \circ g$. Sachant que inv est \mathcal{C}^∞ (admis), que f' est \mathcal{C}^{k-1} et que g est continue, on obtient que g' est continue donc g de classe \mathcal{C}^1 , donc finalement g' de classe \mathcal{C}^1 , donc g de classe \mathcal{C}^2 et ainsi de suite jusqu'à obtenir g de classe \mathcal{C}^k . Nous avons finalement démontré que f est un \mathcal{C}^k -difféomorphisme. Le théorème est démontré.

1. Soit $g : E \rightarrow F$. On dit que $g(v) = O(v)$ si $\exists \eta > 0$ et $C > 0$ t.q. $\|v\|_E \leq \eta \Rightarrow \|g(v)\|_F \leq C\|v\|_E$. ■

Remarque 3.2.2. Rappelons que $\text{Isom}_c(E, F)$ est ouvert dans $\mathcal{L}(E, F)$ d'après [4, Théorème 3.19.8 page 400]. L'application inv est même un homéomorphisme de $\text{Isom}_c(E, F)$ dans $\text{Isom}_c(F, E)$. Et puisque pour tous T et H dans $\text{Isom}_c(E, F)$, on a $\text{inv}'(T) \cdot H = -\text{inv}(T) \circ H \circ \text{inv}(T)$, on peut montrer que $\text{inv}'(T) \in \mathcal{L}(\text{Isom}_c(E, F), \text{Isom}_c(F, E))$ est bijective. Donc, d'après la proposition 3.2.2, l'application inv est un \mathcal{C}^∞ -difféomorphisme.

Théorème 3.2.3 – d'inversion locale

Soient E, F deux espaces de Banach et $f: E \rightarrow F$ une application de classe \mathcal{C}^k , $k \in \overline{\mathbb{N}}^*$, sur un voisinage de $x \in E$. On suppose que $f'(x) \in \mathcal{L}(E, F)$ est bijective.² Il existe alors un ouvert $U \subset E$ contenant x et un ouvert $V \subset F$ contenant $f(x)$, tels que f est un \mathcal{C}^k -difféomorphisme de U dans $V = f(U)$.

► L'idée est de montrer que f est un homéomorphisme de classe \mathcal{C}^k de U sur V t.q. $\forall x \in U$, $f'(x) \in \mathcal{L}(E, F)$ est bijective. On appliquera ensuite la proposition précédente.

* Notons $y := f(x) \in F$. Introduisons

$$f_0(u) := f'(x)^{-1} \cdot (f(x + u) - y),$$

de telle sorte que f est un \mathcal{C}^k -difféomorphisme d'un voisinage de x sur un voisinage de y si et seulement si f_0 est un \mathcal{C}^k -difféomorphisme d'un voisinage de 0_E sur un voisinage de 0_E . On se ramène donc à l'étude de f_0 . Remarquons que f_0 est de classe \mathcal{C}^k , que $f_0(0_E) = 0_E$ et que $f'_0(0_E) = \text{Id}_E$. Puisque f'_0 est continue sur E , alors $\forall \varepsilon > 0$, $\exists \eta_\varepsilon > 0$ t.q. $\forall u \in E$ vérifiant $\|u\|_E \leq \eta_\varepsilon$ on a $f'_0(u) \in B(\text{Id}_E, \varepsilon)$. Enfin, puisque $\text{GL}_c(E) = \text{Isom}_c(E, E)$ est ouvert dans $\mathcal{L}(E)$, il existe $\varepsilon > 0$ t.q. $B(\text{Id}_E, \varepsilon) \subset \text{GL}_c(E)$.

* Il reste donc à montrer que f_0 est un homéomorphisme d'un voisinage de 0_E sur un autre voisinage de 0_E . Puisque f_0 est continue, il reste à montrer que f_0 est bijective d'un voisinage U_0 de 0_E sur un autre voisinage V_0 de 0_E et que sur le voisinage V_0 , l'application inverse est continue. Introduisons pour cela, pour u, v dans E

$$\varphi_v(u) := v + u - f_0(u),$$

de telle sorte que $\varphi_v(u) = u$ si et seulement si $f_0(u) = v$, c'est-à-dire si et seulement si φ_v admet un point fixe. Vérifions tout d'abord que φ_v est contractante dans un voisinage de 0_E . D'après le théorème 1.6.5 des accroissements finis (et son corollaire), pour tous $r > 0$ et u_1, u_2 dans $B(0, r)$, on a :

$$\|\varphi_v(u_1) - \varphi_v(u_2)\|_E \leq \left(\sup_{u \in B(0, r)} \|\text{Id}_E - f'_0(u)\|_{\mathcal{L}(E)} \right) \|u_1 - u_2\|_E.$$

Ainsi, pour $\rho := \min(\varepsilon, 1/2) \in [0, 1[$ avec $\varepsilon > 0$ t.q. $B(\text{Id}_E, \varepsilon) \subset \text{GL}_c(E)$, et pour $\eta := \eta_\rho$, alors pour tous u_1, u_2 dans $B(0, \eta)$, on a :

$$\|\varphi_v(u_1) - \varphi_v(u_2)\|_E \leq \rho \|u_1 - u_2\|_E,$$

2. D'après le théorème 1.1.5, dans ce cas $f'(x) \in \text{Isom}_c(E, F)$.

c'est-à-dire, φ_v est ρ -contractante. On a de plus en particulier (pour $u_2 = 0_E$)

$$\|\varphi_v(u_1) - v\|_E \leq \rho \|u_1\|_E,$$

et donc pour tous $u \in B_f(0, \eta/2)$ et $v \in B_f(0, (1 - \rho)\eta/2)$, on a

$$\|\varphi_v(u)\|_E \leq \|v\|_E + \rho \|u\|_E \leq (1 - \rho) \frac{\eta}{2} + \rho \frac{\eta}{2} = \frac{\eta}{2},$$

autrement dit $\varphi_v(B_f(0, \eta/2)) \subset B_f(0, \eta/2)$ si $v \in B_f(0, (1 - \rho)\eta/2)$. Posons

$$V_0 := B(0, (1 - \rho)\eta/2)$$

et introduisons

$$\begin{aligned} \varphi: B_f(0, \eta/2) \times V_0 &\longrightarrow B_f(0, \eta/2) \\ (u, v) &\longmapsto \varphi(u, v) := \varphi_v(u), \end{aligned}$$

de telle sorte que φ vérifie les hypothèses du théorème 3.1.3 du point fixe de Picard à paramètre ; φ est continue et uniformément ρ -contractante. Ainsi, $\forall v \in V_0$, $\exists! u_v \in B_f(0, \eta/2)$ t.q. $\varphi(u_v, v) = \varphi_v(u_v) = u_v$, et $v \mapsto u_v$ est continue sur V_0 . Finalement, f_0 est bijective de $U_0 := f_0^{-1}(V_0)$ dans V_0 et l'application réciproque $f_0^{-1}(v) = u_v$ est continue.

* Nous avons donc démontré que f_0 est un homéomorphisme de U_0 dans V_0 de classe \mathcal{C}^k t.q. $\forall u \in U_0$, $f'_0(u) \in \mathcal{L}(E)$ est bijective, donc d'après la proposition 3.2.2, f_0 est un \mathcal{C}^k -difféomorphisme de U_0 dans V_0 et finalement le théorème est démontré. ■

3.3 Théorème des fonctions implicites

Théorème 3.3.1 – des fonctions implicites

Soient E, F et G trois espaces de Banach, Ω un ouvert de $E \times F$, $f: \Omega \subset E \times F \rightarrow G$, une application de classe \mathcal{C}^k , $k \in \mathbb{N}^*$, et $(\bar{x}, \bar{y}) \in \Omega$, tels que $\partial_y f(\bar{x}, \bar{y}) \in \mathcal{L}(F, G)$ est bijective. Il existe alors $U \times V \subset \Omega$ un voisinage ouvert du point (\bar{x}, \bar{y}) tel que :

- i) $\forall x \in U$, $\exists! y_x \in V$, tels que $f(x, y_x) = f(\bar{x}, \bar{y})$;
- ii) L'application (unique) $\varphi: U \rightarrow V$, $x \mapsto \varphi(x) := y_x$ est de classe \mathcal{C}^k et $\forall x \in U$:

$$\varphi'(x) = -(\partial_y f(x, \varphi(x)))^{-1} \circ \partial_x f(x, \varphi(x)) ;$$

► On pose

$$\begin{aligned} \Phi: \Omega &\longrightarrow E \times G \\ (x, y) &\longmapsto \Phi(x, y) := (x, f(x, y)). \end{aligned}$$

* L'application Φ est \mathcal{C}^k sur Ω et sa différentielle en (\bar{x}, \bar{y}) est donnée par

$$\Phi'(\bar{x}, \bar{y}) \cdot (u, v) = \begin{pmatrix} \text{Id}_E & 0_{\mathcal{L}(F, E)} \\ \frac{\partial f}{\partial x}(\bar{x}, \bar{y}) & \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \end{pmatrix} \cdot (u, v) = \left(u, \frac{\partial f}{\partial x}(\bar{x}, \bar{y}) \cdot u + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \cdot v \right) \in E \times G,$$

pour tout $(u, v) \in E \times F$. On rappelle que $\Phi'(\bar{x}, \bar{y}) \in \mathcal{L}(E \times F, E \times G)$. De plus, pour tout $(a, b) \in E \times G$, il existe un unique couple $(u, v) \in E \times F$ tel que $\Phi'(\bar{x}, \bar{y}) \cdot (u, v) = (a, b)$. En

effet, $u = a$ et $v = -(\partial_y f(\bar{x}, \bar{y}))^{-1} \cdot (\partial_x f(\bar{x}, \bar{y}) \cdot a)$, car $\partial_y f(\bar{x}, \bar{y})$ inversible par hypothèse, donc $\Phi'(\bar{x}, \bar{y}) \in \mathcal{L}(E \times F, E \times G)$ est bijective et puisque $E \times F$ et $E \times G$ sont deux espaces de Banach, nous pouvons appliquer le théorème 3.2.3 d'inversion locale à Φ .

* D'après le théorème d'inversion locale, il existe $\Omega' := U' \times V \subset \Omega$, U' et V deux ouverts contenant respectivement \bar{x} et \bar{y} , tel que Φ est un \mathcal{C}^k -difféomorphisme de Ω' dans $\Phi(\Omega')$. On note $\Phi^{-1} = (g_1, g_2)$ l'inverse de Φ défini sur $\Phi(\Omega') \subset E \times G$ et à valeurs dans $\Omega' \subset E \times F$. Les applications g_1 et g_2 sont de classe \mathcal{C}^k . Alors, pour tout $(x, z) \in \Phi(\Omega')$,

$$(x, z) = (\Phi \circ \Phi^{-1})(x, z) = \Phi(\Phi^{-1}(x, z)) = (g_1(x, z), f(g_1(x, z), g_2(x, z))).$$

Ceci nous donne $g_1(x, z) = x$ et $f(x, g_2(x, z)) = z$ pour tout $(x, z) \in \Phi(\Omega')$. En particulier, pour $\bar{z} := f(\bar{x}, \bar{y})$, en introduisant $\varphi := g_2(\cdot, \bar{z})$ et $U := \{x \in U' \mid (x, \bar{z}) \in \Phi(\Omega')\}$, nous avons :

$$\forall x \in U, \quad f(x, \varphi(x)) = \bar{z}.$$

Défini ainsi, U est bien un ouvert de E (car préimage de l'ouvert $\Phi(\Omega')$ par l'application continue $x \mapsto (x, \bar{z})$ définie sur U') contenant \bar{x} et $\varphi(U) \subset V$ (par définition).

* L'application φ est \mathcal{C}^k sur U donc par composition $x \mapsto f(x, \varphi(x))$ aussi, cette application est constante sur U donc de dérivée nulle et sa dérivée est donnée par :

$$\forall x \in U, \quad \frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x)) \circ \varphi'(x) = 0_{\mathcal{L}(E, G)},$$

ce qui conclut la démonstration. ■

Remarque 3.3.1. On peut réécrire les conclusions i) et ii) du théorème des fonctions implicites sous la forme suivante. L'application $\varphi: U \rightarrow V$ est unique, de classe \mathcal{C}^k , on a :

$$x \in U, y \in V \text{ et } f(x, y) = f(\bar{x}, \bar{y}) \iff x \in U \text{ et } y = \varphi(x)$$

et la différentielle de φ est donnée pour tout $x \in U$ par :

$$\varphi'(x) = -(\partial_y f(x, \varphi(x)))^{-1} \circ \partial_x f(x, \varphi(x)).$$

Exemple 3.3.1. Soit $A \in \mathbf{M}_n(\mathbb{R})$ inversible. On pose $f: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f(x, b) = Ax - b$. Ici $\partial_x f(x, b) = A \in \mathbf{GL}_n(\mathbb{R})$ est indépendante de x et b , et il existe une fonction implicite (ici globale), notée $x(b)$, telle que, par exemple, pour tout $b \in \mathbb{R}^n$, $f(x(b), b) = f(0_{\mathbb{R}^n}, 0_{\mathbb{R}^n}) = 0_{\mathbb{R}^n}$. Sa différentielle est donnée par $x'(b) = -(\partial_x f(x(b), b))^{-1} \circ \partial_b f(x(b), b) = A^{-1}$, ce que l'on savait déjà car la solution de $f(\cdot, b) = 0$ est $x(b) = A^{-1}b$, donc $x'(b) = A^{-1}$. □

Sous-variétés différentielles de \mathbb{R}^n

4.1	Formes normales des submersions et immersions	45
4.2	Sous-variétés différentielles de \mathbb{R}^n et exemples	48
4.3	Sous-espaces tangents d'une sous-variété de \mathbb{R}^n	52
4.4	Discussion : sous-variétés de \mathbb{R}^n comme préimage d'une submersion	55
4.5	Optimisation en dimension finie avec contraintes d'égalité : CN1-CN2	56

Cette section s'inspire fortement de [2, 3]. Des figures sont reprises de ces notes de cours.

4.1 Formes normales des submersions et immersions

Soient E, F deux espaces vectoriels normés, U un ouvert de E et $f: U \subset E \rightarrow F$ une application. Si f est différentiable en $x \in U$, on note $\text{rank}_x f$ le rang de l'application linéaire $f'(x)$, *i.e.* $\text{rank}_x f := \dim \text{Im } f'(x)$, appelé le *rang de f en x* . On a $\text{rank}_x f \leq \min\{\dim E, \dim F\}$. Si E et F sont de dimension finie, c'est aussi le rang de l'application linéaire $f'(x)$ dans des bases de E et F . Par exemple, si $F = \mathbb{R}^m$, et si f_1, \dots, f_m sont les composantes de f , alors le rang de f en x est le rang du système de formes linéaires $(d_x f_1, \dots, d_x f_m)$, $d_x f_j := f'_j(x) \in \mathcal{L}(E, \mathbb{R}) = E'$, dans le dual topologique E' de E . Si $E = \mathbb{R}^n$, alors le rang de f en x est le rang du système de vecteurs $(\partial_{x_1} f(x), \dots, \partial_{x_n} f(x))$, $\partial_{x_i} f(x) \in \mathcal{L}(\mathbb{R}, F) \simeq F$, dans F . Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$, alors le rang de f en x est le rang de la matrice jacobienne $J_f(x)$.

Définition 4.1.1 – Submersion, immersion, application de rang constant

Soient E, F deux espaces vectoriels normés, U un ouvert de E , $f: U \subset E \rightarrow F$ une application et x un point de U .

- Si f est différentiable en x , on dit que f est une *immersion en x* si la différentielle $f'(x): E \rightarrow F$ est injective. Si E est de dimension finie, ceci équivaut à dire que $\text{rank}_x f = \dim E$. On dit que f est une *immersion* si f est une immersion en tout point de U .
- Si f est différentiable en x , on dit que f est une *submersion en x* si la différentielle $f'(x): E \rightarrow F$ est surjective. Si F est de dimension finie, ceci équivaut à dire que $\text{rank}_x f = \dim F$. On dit que f est une *submersion* si f est une submersion en tout point de U .
- Si f est différentiable sur un voisinage ouvert $V \subset U$ de x , on dit que f est une *application de rang constant en x* si la différentielle $f'(y): E \rightarrow F$ est de rang constant pour tout y dans un voisinage $\Omega \subset V$ de x .

Remarque 4.1.1. Si E et F sont de dimension finie, une immersion ou submersion différentiable sur un voisinage de x est une application de rang constant.

Exemple 4.1.1. Pour $n \leq m$, l'injection canonique

$$\begin{aligned} i: \quad \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x = (x_1, \dots, x_n) &\longmapsto i(x) := (x_1, \dots, x_n, \underbrace{0, \dots, 0}_{m-n \text{ fois}}) \end{aligned}$$

est une immersion. Pour $n \geq m$, la projection canonique

$$\begin{aligned} p: \quad \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x = (x_1, \dots, x_n) &\longmapsto p(x) := (x_1, \dots, x_m) \end{aligned}$$

est une submersion. Pour $r \leq \min(m, n)$, l'application

$$\begin{aligned} \xi: \quad \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x = (x_1, \dots, x_n) &\longmapsto \xi(x) := (i \circ p)(x) = (x_1, \dots, x_r, \underbrace{0, \dots, 0}_{m-r \text{ fois}}) \end{aligned}$$

avec $p: \mathbb{R}^n \rightarrow \mathbb{R}^r$ et $i: \mathbb{R}^r \rightarrow \mathbb{R}^m$ les projections et injections canoniques, est une application de rang constant égal à r . \square

Les résultats suivant nous montrent que ces exemples précédents sont, localement et modulo des changements de coordonnées (en général non linéaires) locaux, les seuls !

Lemme 4.1.1 (Théorème de forme normale locale des immersions). *Soient U un ouvert de \mathbb{R}^n et $f: U \rightarrow \mathbb{R}^m$, $n \leq m$, une application de classe \mathcal{C}^k , qui est une immersion en $\bar{x} \in U$. Alors il existe $V \in \mathcal{V}_{\mathbb{R}^m}(f(\bar{x}))$ et un \mathcal{C}^k -difféomorphisme $\psi: V \rightarrow \psi(V)$ de \mathbb{R}^m tel que pour tout x suffisamment proche de \bar{x} :*

$$(\psi \circ f)(x) = i(x) = i(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0)$$

(“changement de coordonnées au but”).

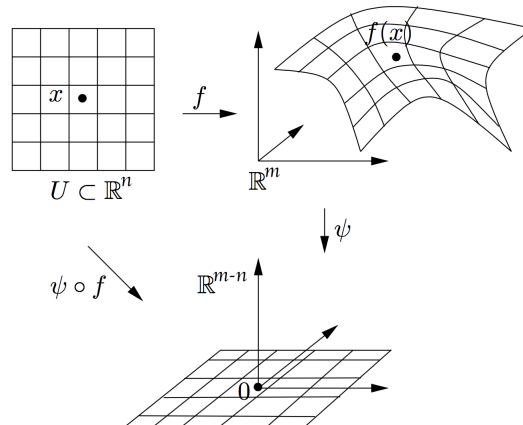


FIGURE 4.1 – Géométriquement, le lemme 4.1.1 dit que l'image d'une immersion peut être, au voisinage de chaque point, redressée en un (ouvert de) sous-espace vectoriel par un difféomorphisme au but.

► **Preuve du lemme 4.1.1.** $A := f'(\bar{x}) \simeq J_f(\bar{x}) \in \mathbf{M}_{m,n}(\mathbb{R})$ injective donc $\text{Im } A$ est un sous-espace vectoriel de dimension n . Montrons qu'il existe une matrice $B \in \mathbf{M}_{m,m-n}(\mathbb{R})$ telle que $\begin{pmatrix} A & B \end{pmatrix} \in \text{GL}_m(\mathbb{R})$. Puisque $\text{Im } A$ est un sev de dimension n , il existe $m - n$ formes linéaires que l'on peut écrire sous la forme $\varphi_i := (v_i | \cdot)_{\mathbb{R}^m}$, $v_i \in \mathbb{R}^m$, telles que $\text{Im } A = \bigcap_{i=1}^{m-n} \text{Ker } \varphi_i$. On introduit

$$B := \begin{pmatrix} v_1 & \cdots & v_{m-n} \end{pmatrix} \in \mathbf{M}_{m,m-n}(\mathbb{R})$$

de telle sorte que $\text{Im } A = \text{Ker } B^T$. Les espaces $\text{Im } A$ et $(\text{Im } A)^\perp$ sont en somme directe, *i.e.* $\mathbb{R}^m = \text{Im } A \oplus (\text{Im } A)^\perp$, et $(\text{Im } A)^\perp = (\text{Ker } B^T)^\perp = \text{Im } B$, cf. remarque 4.1.2. On introduit la matrice $M := \begin{pmatrix} A & B \end{pmatrix} \in \mathbf{M}_m(\mathbb{R})$. Montrons que $M \in \text{GL}_m(\mathbb{R})$, pour cela montrons simplement que M est injective. Soit $x \in \text{Ker } M \subset \mathbb{R}^m$. On écrit $x = (x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^{m-n}$ ce qui nous donne $Mx = Ax_1 + Bx_2 = 0_{\mathbb{R}^m}$. Puisque $Ax_1 \in \text{Im } A$, $Bx_2 \in \text{Im } B$, et puisque $\mathbb{R}^m = \text{Im } A \oplus \text{Im } B$, alors $Ax_1 = Bx_2 = 0_{\mathbb{R}^m}$, *i.e.* $x_1 = 0_{\mathbb{R}^n}$ car A injective par hypothèse, et $x_2 = 0_{\mathbb{R}^{m-n}}$, car les v_i forment une famille libre et donc B est injective. En conclusion $\text{Ker } M = \{0_{\mathbb{R}^m}\}$, donc M injective, donc bijective, *i.e.* $M \in \text{GL}_m(\mathbb{R})$. Posons alors l'application

$$\begin{aligned} g: U \times \mathbb{R}^{m-n} &\longrightarrow \mathbb{R}^m \\ (x, y) &\longmapsto g(x, y) := f(x) + By. \end{aligned}$$

L'application g est \mathcal{C}^k sur $U \times \mathbb{R}^{m-n}$ et

$$g'(\bar{x}, 0_{\mathbb{R}^{m-n}}) = \begin{pmatrix} A & B \end{pmatrix} \in \text{GL}_m(\mathbb{R}),$$

donc d'après le théorème 3.2.3 d'inversion locale, il existe un ouvert $\Omega \subset U \times \mathbb{R}^{m-n}$ contenant $(\bar{x}, 0_{\mathbb{R}^{m-n}})$ tel que g est un \mathcal{C}^k -difféomorphisme de Ω dans $V := g(\Omega) \in \mathcal{V}_{\mathbb{R}^m}(g(\bar{x}, 0_{\mathbb{R}^{m-n}}))$, avec $g(\bar{x}, 0_{\mathbb{R}^{m-n}}) = f(\bar{x})$. On définit alors $\psi := g^{-1}$ sur V qui est un \mathcal{C}^k -difféomorphisme de V dans $\psi(V) = \Omega$, et on a $(\psi \circ f)(x) = g^{-1}(f(x)) = (x, 0_{\mathbb{R}^{m-n}}) = i(x)$, pour tout $x \in f^{-1}(V)$ tel que $(x, 0_{\mathbb{R}^{m-n}}) \in \Omega$. ■

Remarque 4.1.2. Soit $A \in \mathbf{M}_{m,n}(\mathbb{R})$. Alors $\text{Ker } A = (\text{Im } A^T)^\perp$. En effet :

$$\begin{aligned} x \in \text{Ker } A \subset \mathbb{R}^n &\iff Ax = 0_{\mathbb{R}^m} \\ &\iff \forall y \in \mathbb{R}^m, (Ax | y)_{\mathbb{R}^m} = 0 \iff \forall y \in \mathbb{R}^m, (x | A^T y)_{\mathbb{R}^n} = 0 \\ &\iff x \perp \text{Im } A^T. \end{aligned}$$

Remarque 4.1.3. Dans la preuve précédente, nous pouvons construire B plus rapidement. On considère un sous-espace supplémentaire, noté F , de $\text{Im } A$. Il est de dimension $m - n$ par hypothèse. On choisit une base de cet espace, *i.e.* on a $m - n$ vecteurs $v_i \in \mathbb{R}^m$ formant une famille libre de \mathbb{R}^m . On construit alors la matrice B dont les colonnes sont constituées des coordonnées des vecteurs de la base du supplémentaire dans la base canonique de \mathbb{R}^m , *i.e.* $B = \begin{pmatrix} v_1 & \cdots & v_{m-n} \end{pmatrix} \in \mathbf{M}_{m,m-n}(\mathbb{R})$, ainsi $\text{Im } B = \text{Vect}(v_1, \dots, v_{m-n}) = F$. Si on prend un supplémentaire orthogonal, on construit la même matrice B que dans la preuve précédente.

Remarque 4.1.4. Une immersion n'est pas nécessairement injective.

Lemme 4.1.2 (Théorème de forme normale locale des submersions). Soient U un ouvert de \mathbb{R}^n et $f: U \rightarrow \mathbb{R}^m$, $n \geq m$, une application de classe \mathcal{C}^k , qui est une submersion en $\bar{x} \in U$.

Alors il existe $V \in \mathcal{V}_{\mathbb{R}^n}(0)$ et un \mathcal{C}^k -difféomorphisme $\varphi: V \rightarrow \varphi(V)$ de \mathbb{R}^n tel que :

$$\forall y \in V, \quad (f \circ \varphi)(y) - f(\bar{x}) = p(y) = p(y_1, \dots, y_n) = (y_1, \dots, y_m)$$

(“changement de coordonnées à la source”).

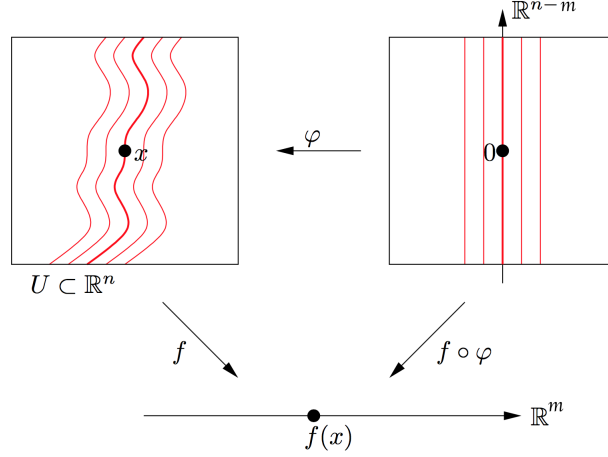


FIGURE 4.2 – Géométriquement, le lemme 4.1.2 implique qu’au voisinage de chaque point de la source, la préimage de l’image de ce point par une submersion peut être redressée en un (ouvert) sous-espace vectoriel par un difféomorphisme à la source.

► **Preuve du lemme 4.1.2.** $A := f'(\bar{x}) \simeq J_f(\bar{x}) \in \mathbf{M}_{m,n}(\mathbb{R})$ surjective donc de rang m . Il existe donc une matrice $B \in \mathbf{M}_{n-m,n}(\mathbb{R})$ telle que

$$M := \begin{pmatrix} A \\ B \end{pmatrix} \in \mathrm{GL}_n(\mathbb{R}).$$

Posons alors l’application

$$\begin{aligned} g: U &\longrightarrow \mathbb{R}^m \times \mathbb{R}^{n-m} \\ x &\longmapsto g(x) := (f(x) - f(\bar{x}), B(x - \bar{x})). \end{aligned}$$

L’application g est \mathcal{C}^k de U dans $\mathbb{R}^m \times \mathbb{R}^{n-m}$ et

$$g'(\bar{x}) = \begin{pmatrix} A \\ B \end{pmatrix} \in \mathrm{GL}_n(\mathbb{R}),$$

donc d’après le théorème 3.2.3 d’inversion locale, il existe un ouvert $\Omega \subset U$ contenant \bar{x} tel que g est un \mathcal{C}^k -difféomorphisme de Ω dans $V := g(\Omega) \in \mathcal{V}_{\mathbb{R}^n}(g(\bar{x}))$, avec $g(\bar{x}) = 0_{\mathbb{R}^n}$. On définit alors $\varphi := g^{-1}$ sur V qui est un \mathcal{C}^k -difféomorphisme de V dans $\varphi(V) = \Omega$. Pour tout $y \in V$, on a $g \circ g^{-1}(y) = y = (f(g^{-1}(y)) - f(\bar{x}), B(g^{-1}(y) - \bar{x}))$, autrement dit, pour tout $y \in V$, on a $(f \circ \varphi)(y) - f(\bar{x}) = f(g^{-1}(y)) - f(\bar{x}) = (y_1, \dots, y_m) = p(y)$. ■

4.2 Sous-variétés différentielles de \mathbb{R}^n et exemples

On identifiera \mathbb{R}^n et $\mathbb{R}^p \times \mathbb{R}^{n-p}$ par $(x_1, \dots, x_n) \mapsto ((x_1, \dots, x_p), (x_{p+1}, \dots, x_n))$ pour $0 \leq p \leq n$. Le premier exemple de sous-variétés de \mathbb{R}^n à garder en tête sont les sous-espaces vectoriels,

par exemple de la forme $\mathbb{R}^p \times \{0\}$, $0 \leq p \leq n$. Nous allons définir une sous-variété générale comme obtenue par difféomorphisme locaux ambiants à partir d'un tel exemple, et donner des caractérisations équivalentes. Rappelons que le *graphe* d'une application $f: X \rightarrow Y$ est la partie de $X \times Y$ formée des couples $(x, f(x))$ pour x dans X . Habituellement, une sous-variété (différentielle) de \mathbb{R}^n est définie de la façon suivante.

Définition 4.2.1 – Sous-variété (différentielle) de \mathbb{R}^n

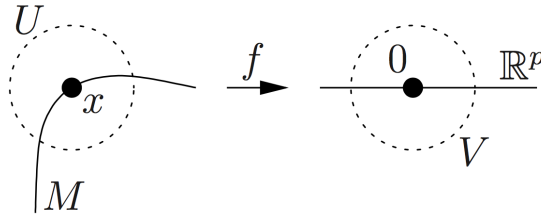
Soit $0 \leq p \leq n$ dans \mathbb{N} et $k \in \overline{\mathbb{N}}^*$. On dit que $M \subset \mathbb{R}^n$ est une sous-variété (différentielle) de \mathbb{R}^n de dimension p et de classe \mathcal{C}^k si pour tout $x \in M$, il existe $U \in \mathcal{V}_{\mathbb{R}^n}(x)$ et un \mathcal{C}^k -difféomorphisme $f: U \rightarrow f(U)$ de \mathbb{R}^n tels que $f(U \cap M) = f(U) \cap (\mathbb{R}^p \times \{0\})$.

Cependant, nous aurions pu choisir une autre définition, d'après le résultat suivant.

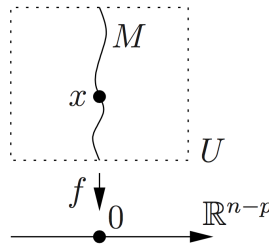
Théorème 4.2.2 – de caractérisation des sous-variétés (différentielles) de \mathbb{R}^n

Soit $0 \leq p \leq n$ dans \mathbb{N} et $k \in \overline{\mathbb{N}}^*$. Les propriétés suivantes pour une partie M de \mathbb{R}^n sont équivalentes :

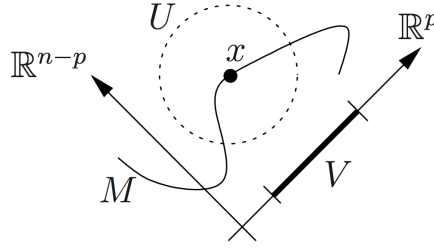
- i) (**Définition locale par redressement**). Pour tout x dans M , il existe un voisinage U de x dans \mathbb{R}^n , un voisinage V de 0 dans \mathbb{R}^n et un \mathcal{C}^k -difféomorphisme $f: U \rightarrow V$, tel que $f(x) = 0$, tels que $f(U \cap M) = V \cap (\mathbb{R}^p \times \{0\})$.



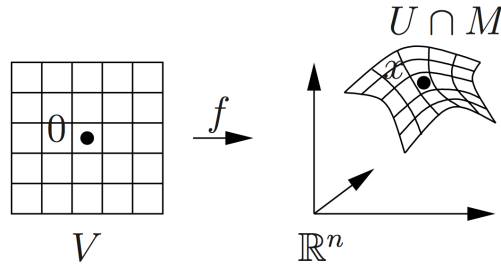
- ii) (**Définition locale par fonction implicite**). Pour tout x dans M , il existe un voisinage U de x dans \mathbb{R}^n , et une application $f: U \rightarrow \mathbb{R}^{n-p}$ de classe \mathcal{C}^k qui est une submersion en x , tels que $U \cap M = f^{-1}(\{0\})$.



- iii) (**Définition locale par graphe**). Pour tout x dans M , il existe un voisinage U de x dans \mathbb{R}^n , un ouvert V de \mathbb{R}^p et une application $f: V \rightarrow \mathbb{R}^{n-p}$ de classe \mathcal{C}^k tels que $U \cap M = \text{graphe}(f)$.



- iv) (**Définition locale par paramétrage**). Pour tout x dans M , il existe un voisinage U de x dans \mathbb{R}^n , un voisinage V de 0 dans \mathbb{R}^p et une application $f: V \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^k tels que $f(0) = x$, f soit une immersion en 0 , et f soit un homéomorphisme de V sur $U \cap M$ (on dit que f est un plongement).



Corollaire 4.2.3

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^k , soit $y \in \text{Im } f$. Si f est une submersion en tout point de $M := f^{-1}(\{y\})$, alors M est une sous-variété de \mathbb{R}^n de dimension $n - m$ et de classe \mathcal{C}^k .

Corollaire 4.2.4

Soit $f: U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^k , U un ouvert de \mathbb{R}^p . Si f est un plongement (i.e. un homéomorphisme de U dans $f(U)$ et une immersion en tout point de U), alors $M := f(U)$ est une sous-variété de \mathbb{R}^n de dimension p et de classe \mathcal{C}^k .

On dit donc qu'une partie M de \mathbb{R}^n est une sous-variété de \mathbb{R}^n de dimension p et de classe \mathcal{C}^k (et tout simplement sous-variété par abus) si elle vérifie l'une des propriétés du théorème 4.2.2. On dit que la sous-variété est *lisse* si elle est de classe \mathcal{C}^∞ . On a de plus les définitions classiques suivantes. On dit que la *codimension* de M dans \mathbb{R}^n est $n - p$. Si $p = 1, 2, n - 1$, on dit que M est une *courbe*, *surface*, *hypersurface* (différentielle) de \mathbb{R}^n respectivement. Soit $x \in M$, un *paramétrage local* de classe \mathcal{C}^k de M en x est une application $f: V \rightarrow M$, où V est un voisinage de 0 dans \mathbb{R}^p , tel que $f(0) = x$, f soit une immersion \mathcal{C}^k sur V et f soit un homéomorphisme de V sur un voisinage ouvert de x dans M .

Remarque 4.2.1. On parle de sous-variété *analytique réelle* lorsqu'elle est \mathcal{C}^ω . La définition 4.2.1 fait encore sens lorsque $k = 0$, on parle alors de sous-variété *topologique*. Cependant, dans ce cours, on ne parlera que de sous-variétés (différentielles) de classe \mathcal{C}^k , pour $k \in \mathbb{N}^*$.

Exemple 4.2.1 (Exemples de sous-variétés (différentielles) de \mathbb{R}^n). Voici une liste de sous-variétés différentielles de \mathbb{R}^n .

- Tout ouvert de $\mathbb{R}^p \times \{0\} \subset \mathbb{R}^n$ est une sous-variété lisse de dimension p ;
- La sphère $\mathbb{S}^n := \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}$ est une sous-variété lisse de dimension n ;

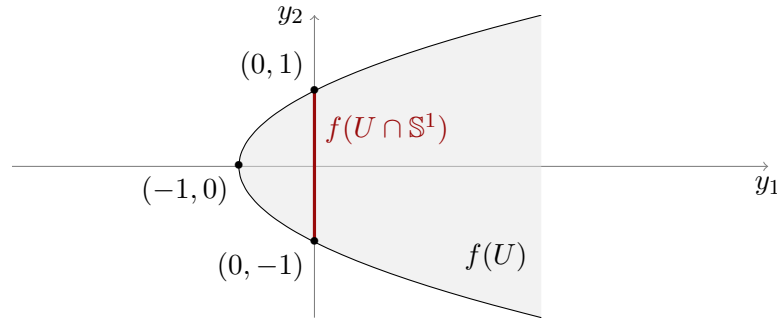
i) (Définition locale par redressement). Soit $\bar{x} \in \mathbb{S}^n$; on peut supposer que $\bar{x}_1 > 0$. Soit $U := \{x \in \mathbb{R}^{n+1} \mid x_1 > 0\}$, un ouvert de \mathbb{R}^{n+1} contenant \bar{x} , et soit

$$\begin{aligned} f: U &\longrightarrow \mathbb{R}^{n+1} \\ x &\longmapsto f(x) := (x_1^2 + \dots + x_{n+1}^2 - 1, x_2 - \bar{x}_2, \dots, x_{n+1} - \bar{x}_{n+1}) \end{aligned}$$

Alors f est un \mathcal{C}^∞ -difféomorphisme sur son image

$$V := f(U) = \left\{ y \in \mathbb{R}^{n+1} \mid y_1 > \sum_{i=2}^{n+1} (y_i + \bar{x}_i)^2 - 1 \right\} \subset]-1, \infty[\times \mathbb{R}^n,$$

V un voisinage de 0 (tel que $f(\bar{x}) = 0$) dans \mathbb{R}^{n+1} , car $\sum_{i=2}^{n+1} \bar{x}_i^2 - 1 < 0$, et $f(U \cap \mathbb{S}^n) = f(U) \cap (\{0\} \times \mathbb{R}^n) = V \cap (\{0\} \times \mathbb{R}^n)$. Illustrons ceci pour $n = 1$ avec $\bar{x} := (1, 0) \in \mathbb{S}^1$:



ii) (Définition locale par fonction implicite). Posons

$$\begin{aligned} f: \mathbb{R}^{n+1} &\longrightarrow \mathbb{R} \\ x &\longmapsto x_1^2 + \dots + x_{n+1}^2 - 1 \end{aligned}$$

Alors f est une submersion \mathcal{C}^∞ en tout point de $\mathbb{S}^n = f^{-1}(\{0\})$, cf. corollaire 4.2.3.

iii) (Définition locale par graphe). Soit $\bar{x} \in \mathbb{S}^n$. Nous pouvons supposer que $\bar{x}_1 > 0$. Posons $U := \{x \in \mathbb{R}^{n+1} \mid x_1 > 0\}$, $V := B(0_{\mathbb{R}^n}, 1)$ la boule ouverte de centre $0_{\mathbb{R}^n}$ et de rayon 1 pour la norme euclidienne, et

$$\begin{aligned} f: V \subset \mathbb{R}^n &\longrightarrow \mathbb{R} \\ y &\longmapsto f(y) := \sqrt{1 - \|y\|_2^2}. \end{aligned}$$

Alors, f est \mathcal{C}^∞ sur V , un ouvert de \mathbb{R}^n , et on a $\text{graphe}(f) = U \cap \mathbb{S}^n$.

iv) (Définition locale par paramétrage). Soit $\bar{x} \in \mathbb{S}^n$. Nous pouvons supposer que $\bar{x}_1 > 0$ et à une rotation près, que $\bar{x}_1 = 1$. Posons $U := \{x \in \mathbb{R}^{n+1} \mid x_1 > 0\} \in \mathcal{V}_{\mathbb{R}^{n+1}}(\bar{x})$, $V :=]-\frac{\pi}{2}, \frac{\pi}{2}[^n \in \mathcal{V}_{\mathbb{R}^n}(0)$, et

$$\begin{aligned} f: V \subset \mathbb{R}^n &\longrightarrow \mathbb{R}^{n+1} \\ \theta &\longmapsto f(\theta) := (\cos(\theta_1), \sin(\theta_1) \cos(\theta_2), \dots, \sin(\theta_1) \sin(\theta_2) \cdots \sin(\theta_n)). \end{aligned}$$

Alors f est de classe \mathcal{C}^∞ , $f(0) = \bar{x}$, f est une immersion en 0 et f est un homéomorphisme de V sur son image $f(V) = U \cap \mathbb{S}^n$.

- L'espace projectif $P^n(\mathbb{R}) = P(\mathbb{R}^{n+1})$ (i.e. l'ensemble des droites vectorielles de \mathbb{R}^{n+1}) est une sous-variété lisse de dimension n ;
- L'ensemble $\mathbf{M}_n(\mathbb{R}) \simeq L(\mathbb{R}^n) \simeq \mathbb{R}^{n^2}$ est une sous-variété lisse de \mathbb{R}^{n^2} de dimension n^2 , à un isomorphisme (topologique) près. De même, $\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathbf{M}_n(\mathbb{R}) \mid \det(A) \neq 0\} \simeq \mathrm{GL}(\mathbb{R}^n)$ est une sous-variété lisse de $\mathbf{M}_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$ de dimension n^2 , comme ouvert de $\mathbf{M}_n(\mathbb{R})$. Notons qu'il est possible de définir la notion de *variété différentielle* sans faire appel à un espace vectoriel ambiant comme \mathbb{R}^n dans la définition 4.2.1. Ceci veut dire qu'une variété différentielle n'est pas nécessairement un espace M contenu dans \mathbb{R}^n , comme supposé dans la définition 4.2.1. Dans ce cadre, M est plus généralement un espace topologique qui vérifie certaines hypothèses pour être appelé variété différentielle. Nous pourrions alors montrer que $\mathbf{M}_n(\mathbb{R})$ est une variété différentielle et nous dirions simplement que $\mathrm{GL}_n(\mathbb{R})$ est une sous-variété de $\mathbf{M}_n(\mathbb{R})$, la notion de sous-variété d'une variété différentielle étant proche de celle de sous-variété de \mathbb{R}^n .

□

Remarque 4.2.2. Toute sous-variété lisse de \mathbb{R}^n connexe de dimension 1 est difféomorphe soit à \mathbb{R} , soit à \mathbb{S}^1 , voir [1].

4.3 Sous-espaces tangents d'une sous-variété de \mathbb{R}^n

On se donne une sous-variété M de \mathbb{R}^n de dimension p de classe \mathcal{C}^1 , et un point $x \in M$.

Définition 4.3.1 – Vecteurs tangents à une sous-variété de \mathbb{R}^n

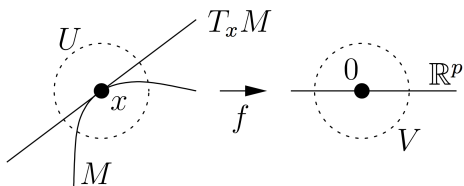
Un vecteur $v \in \mathbb{R}^n$ est *tangent* à M en x s'il existe $\delta > 0$ et $c:]\delta, \delta[\rightarrow \mathbb{R}^n$ une courbe \mathcal{C}^1 à valeurs dans M telle que : $c(0) = x$ et $\dot{c}(0) = v$. On note $T_x M$ l'ensemble des vecteurs tangents à M en x , que l'on appelle *sous-espace (vectoriel) tangent* à M en x , $x + T_x M$ étant le sous-espace *affine* tangent à M en x .

Le résultat suivant donne le calcul de $T_x M$ en fonction des différentes caractérisations locales des sous-variétés de \mathbb{R}^n , voir théorème 4.2.2.

Proposition 4.3.2 – Sous-espaces tangents d'une sous-variété de \mathbb{R}^n

Soit M une sous-variété différentielle de \mathbb{R}^n de dimension p . Alors, le sous-espace tangent $T_x M$ est un sous-espace vectoriel de \mathbb{R}^n de dimension p et l'on a :

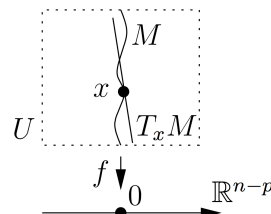
- i) (**Définition locale par redressement**). Si U est un voisinage ouvert de x dans \mathbb{R}^n , si V est un voisinage ouvert de 0 dans \mathbb{R}^p et $f: U \rightarrow V$ un \mathcal{C}^1 -difféomorphisme tels que $f(x) = 0$ et $f(U \cap M) = V \cap (\mathbb{R}^p \times \{0\})$, alors

$$T_x M = (f'(x))^{-1} \cdot (\mathbb{R}^p \times \{0\})$$


- ii) (**Définition locale par fonction implicite**). Si U est un voisinage ouvert de

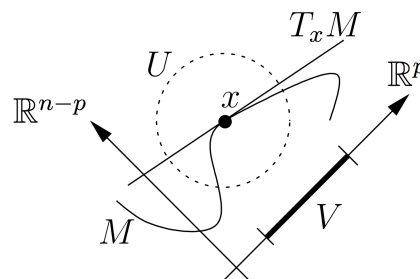
x dans \mathbb{R}^n , et si $f: U \rightarrow \mathbb{R}^{n-p}$ est une submersion en x , de classe \mathcal{C}^1 , telle que $f(x) = 0$ et le tout tel que $U \cap M = f^{-1}(\{0\})$, alors

$$T_x M = \text{Ker } f'(x)$$



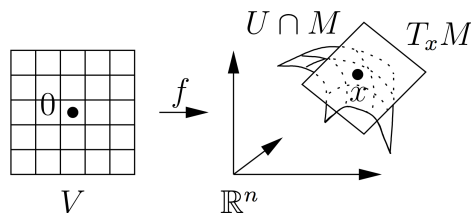
iii) **(Définition locale par graphe).** Si U est un voisinage ouvert de x dans $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}$, si V est un voisinage ouvert de 0 dans \mathbb{R}^p et $f: V \rightarrow \mathbb{R}^{n-p}$ une application de classe \mathcal{C}^1 , tels que $U \cap M = \text{graphe}(f)$ et $x = (0, f(0))$, alors

$$T_x M = \{(v, f'(0) \cdot v) \mid v \in \mathbb{R}^p\}$$



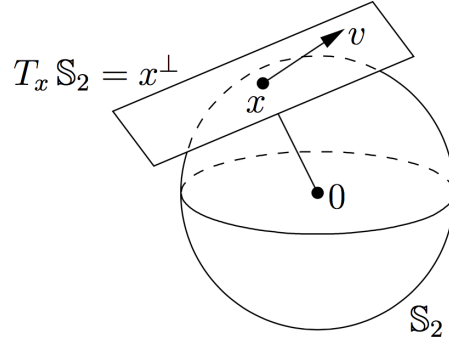
iv) **(Définition locale par paramétrage).** Si U est un voisinage ouvert de x dans \mathbb{R}^n , si V est un voisinage ouvert de 0 dans \mathbb{R}^p et $f: V \rightarrow \mathbb{R}^n$, avec $f(0) = x$, est un paramétrage local \mathcal{C}^1 de $U \cap M$ en x , alors

$$T_x M = \text{Im } f'(0)$$



Exemple 4.3.1. Voici deux exemples d'espaces tangents.

- La sphère \mathbb{S}^n est définie comme la préimage de 1 par la submersion $x \rightarrow \|x\|_2^2$ de $\mathbb{R}^{n+1} \setminus \{0\}$ dans \mathbb{R} . Donc le sous-espace tangent à \mathbb{S}^n en x est le noyau de l'application linéaire $v \rightarrow (x | v)$, i.e. $T_x \mathbb{S}^n$ est l'orthogonal de x (pour le produit scalaire usuel de \mathbb{R}^{n+1}).



- Soit $f: U \rightarrow \mathbb{R}^{n-p}$, U un ouvert de \mathbb{R}^p , de classe \mathcal{C}^1 . Posons $M := \text{graphe}(f) \subset \mathbb{R}^n \simeq \mathbb{R}^p \times \mathbb{R}^{n-p}$. Soit $x_0 \in U$, notons $y_0 := f(x_0)$ et $u_0 = (x_0, y_0)$. Alors, le sous-espace affine tangent en u_0 est donné par $u_0 + T_{u_0}M = \{(x, y) \in \mathbb{R}^p \times \mathbb{R}^{n-p} \mid y - y_0 = f'(x_0) \cdot (x - x_0)\}$.

□

La réunion disjointe des sous-espaces affines tangents aux points d'une sous-variété de \mathbb{R}^n est naturellement munie d'une structure de sous-variété.

Proposition 4.3.3 – Fibré tangent

Soit M une sous-variété de classe \mathcal{C}^{r+1} , $r \geq 1$, de dimension p . Le fibré tangent associé à M défini par

$$TM := \{(x, v) \mid x \in M, v \in T_x M\} \subset \mathbb{R}^n \times \mathbb{R}^n,$$

est une sous-variété de classe \mathcal{C}^r de $\mathbb{R}^n \times \mathbb{R}^n$ de dimension $2p$.

Exemple 4.3.2. Le fibré tangent au cercle $\mathbb{S}^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ est donc la sous-variété $T\mathbb{S}^1 = \{(x, y, v_x, v_y) \in \mathbb{R}^4 \mid x^2 + y^2 = 1, xv_x + yv_y = 0\}$ difféomorphe au cylindre $\mathbb{S}^1 \times \mathbb{R}$.

□

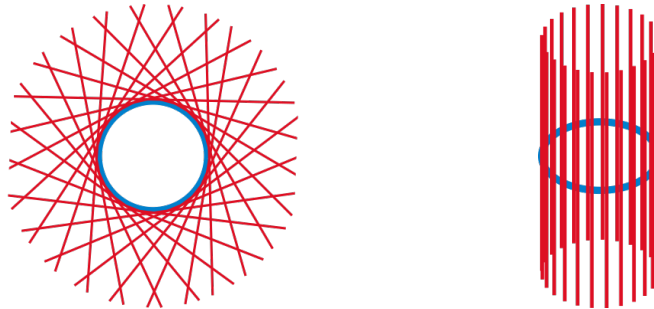


FIGURE 4.3 – Deux manières de représenter le fibré tangent d'un cercle : tous les espaces tangents (à gauche) sont regroupés de manière continue et sans se recouvrir (à droite).

Définition 4.3.4 – Champ de vecteurs

Soit M une sous-variété (différentielle) de \mathbb{R}^n . Un *champ de vecteurs* sur M est une application v définie sur M telle que pour tout $x \in M$, $v(x) \in T_x M$.

4.4 Discussion : sous-variétés de \mathbb{R}^n comme préimage d'une submersion

Soit $A \in \mathbf{M}_{m,n}(\mathbb{R})$, $n \geq m$. On pose

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto f(x) := Ax, \end{aligned}$$

alors $M := f^{-1}(\{b\})$, $b \in \mathbb{R}^m$, est l'ensemble des solutions du système d'équations $Ax = b$. Il y a n inconnues et m équations. Si $f'(x) = A$ est de rang plein, *i.e.* A surjective, alors f est une submersion et donc M est une sous-variété de \mathbb{R}^n de dimension $n - m$. Dans ce cas-ci, M est sous-espace affine de \mathbb{R}^n et si $b = 0_{\mathbb{R}^m}$, alors M est un sous-espace vectoriel. Supposons que $b = 0_{\mathbb{R}^m}$. Alors, si $m = 1$, c'est-à-dire si l'on n'a qu'une seule équation, alors f est une forme linéaire sur \mathbb{R}^n et $M = f^{-1}(\{0_{\mathbb{R}^m}\}) = \text{Ker } A$ est un hyperplan. Sinon, on peut écrire $f = (f_j)_{1 \leq j \leq m}$ et M est alors l'intersection des m hyperplans $H_j := \text{Ker } f_j$.

Parlons maintenant de l'espace tangent à M . Supposons toujours que $b = 0_{\mathbb{R}^m}$. Puisque M est un espace vectoriel, alors son espace tangent est égal à lui-même. En effet, d'après la proposition 4.3.2, puisque M est donnée par une submersion, pour tout $x \in M$, $T_x M = \text{Ker } f'(x) = \text{Ker } A$. Si $m = 1$, alors $v \in T_x M$ si et seulement si $f'(x) \cdot v = (\nabla f(x) | v)_{\mathbb{R}^n} = 0$. Ici A est un vecteur ligne et donc $Av = (A^T | v)_{\mathbb{R}^n}$. Autrement dit, $v \in T_x M$ si et seulement si v est orthogonal au gradient de f en x . On se souvient que le gradient est bien orthogonal à l'hyperplan si $x \mapsto f(x) = Ax$ est linéaire. De manière générale, dans le cas non linéaire, on a que le gradient est orthogonal à l'espace tangent. Tout ceci était pour $m = 1$. Pour m quelconque, $v \in T_x M$ si et seulement si pour tout $j \in \llbracket 1, m \rrbracket$, $(\nabla f_j(x) | v)_{\mathbb{R}^n} = 0$, *i.e.* v est orthogonal aux gradients des f_j en x .

Finissons avec le résultat suivant :


Corollaire 4.4.1

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $n \geq m$. Soit $x \in M := f^{-1}(\{y\})$, avec $y \in \mathbb{R}^m$. Si f est une submersion en x , alors

$$w \perp T_x M \iff \exists \lambda \in \mathbb{R}^m, \text{ t.q. } w = J_f(x)^T \lambda = \sum_{j=1}^m \lambda_j \nabla f_j(x).$$

► f est une submersion en x donc $T_x M = \text{Ker } f'(x) = \text{Ker } J_f(x)$, $J_f(x) \in \mathbf{M}_{m,n}(\mathbb{R})$. Donc $w \perp T_x M$ ssi $w \in (T_x M)^\perp = (\text{Ker } J_f(x))^\perp = \text{Im } J_f(x)^T$, cf. remarque 4.1.2. On rappelle de plus que $J_f(x)^T = (\nabla f_1(x) \cdots \nabla f_m(x))$. ■

4.5 Optimisation en dimension finie avec contraintes d'égalité : CN1-CN2

 **Exercice 4.5.1** (correction page 57). Soit le problème d'optimisation suivant :

$$(P) \quad \min \{f(x) \mid x \in C \subset \mathbb{R}^n\},$$

où $C = h^{-1}(\{0_{\mathbb{R}^m}\})$, f et h sont supposées de classe \mathcal{C}^2 sur un ouvert contenant C , f est à valeurs réelles et h à valeurs dans \mathbb{R}^m , $m \leq n$, et où h est supposée être une **submersion** en tout point de C . On se propose de retrouver le résultat suivant : si $\bar{x} \in C$ est une solution (locale ou globale) de (P) alors $\exists \bar{\lambda} \in \mathbb{R}^m$ t.q. :

- i) $\nabla_x L(\bar{x}, \bar{\lambda}) = 0$,
- ii) $(\nabla_{xx}^2 L(\bar{x}, \bar{\lambda}) v \mid v) \geq 0, \forall v \in T(C, \bar{x})$,

où $L: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $L(x, \lambda) := f(x) + (\lambda \mid h(x))$, est le *lagrangien* associé à (P) et où $T(C, x)$ est l'espace tangent aux contraintes en x .

Remarque 4.5.1. D'après les hypothèses, l'espace des contraintes C est une sous-variété de \mathbb{R}^n de dimension $n - m$ et de classe \mathcal{C}^2 . Nous pourrions affaiblir l'hypothèse f submersion sur C en f submersion en \bar{x} et travailler localement au voisinage de \bar{x} .

Soit $v \in \text{Ker } h'(\bar{x}) = T(C, \bar{x}) = T_{\bar{x}}C$ (voir proposition 4.3.2), avec \bar{x} solution de (P) . Soit $\gamma(\cdot)$ une courbe¹ de classe \mathcal{C}^2 dans C t.q. $\gamma(0) = \bar{x}$ et $\gamma'(0) = v$, $\gamma(\cdot)$ définie sur un intervalle ouvert de la forme $]-\varepsilon, \varepsilon[$, $\varepsilon > 0$. On pose $\varphi_v := f \circ \gamma$ et $\psi_v := h \circ \gamma$. Il est clair que $\bar{t} := 0$ est une solution au problème :

$$(P_v) \quad \min \{\varphi_v(t) \mid t \in]-\varepsilon, \varepsilon[\}.$$

1. Montrons la CN1, *i.e.* le point *i*) du théorème.

1.1. Montrer que $\nabla f(\bar{x}) \perp \text{Ker } h'(\bar{x})$.

1.2. En déduire que : $\exists \bar{\lambda} \in \mathbb{R}^m$ t.q. $\nabla_x L(\bar{x}, \bar{\lambda}) = 0$.

2. Montrons la CN2, *i.e.* le point *ii*) du théorème.

2.1. Montrer que : $\forall t \in]-\varepsilon, \varepsilon[$, $h'(\gamma(t)) \cdot \gamma''(t) = -h''(\gamma(t)) \cdot (\gamma'(t), \gamma'(t))$.

2.2. Calculer $\varphi_v''(t)$ et en déduire la condition *ii*) du théorème. □

1. Il est toujours possible de construire une telle courbe.

▷ **Correction** (de l'exercice 4.5.1, page 56).

1.1. D'après la CN1 d'optimalité appliquée au problème sans contraintes (P_v), $\varphi'_v(\bar{t}) = 0$. Puisque $\varphi'_v(t) = f'(\gamma(t)) \cdot \gamma'(t) = (\nabla f(\gamma(t)) | \gamma'(t))$, on a $\varphi'_v(\bar{t}) = \varphi'_v(0) = (\nabla f(\bar{x}) | v) = 0$, or puisque ceci est vrai pour tout $v \in \text{Ker } h'(\bar{x})$, la conclusion suit.

1.2. $\nabla f(\bar{x}) \in (\text{Ker } h'(\bar{x}))^\perp = \text{Im } h'(\bar{x})^T$, donc $\exists \lambda \in \mathbb{R}^m$ t.q. $\nabla f(\bar{x}) = h'(\bar{x})^T \lambda$ (voir corollaire 4.4.1 et la remarque 4.1.2). La conclusion suit en prenant $\bar{\lambda} = -\lambda$.

Remarque 4.5.2. Si $L(x, \lambda) = f(x) + (\lambda | h(x))$, alors pour tout $v \in \mathbb{R}^n$ on a :

$$(\nabla_x L(x, \lambda) | v) = \frac{\partial L}{\partial x}(x, \lambda) \cdot v = f'(x) \cdot v + (\lambda | h'(x) \cdot v) = (\nabla f(x) | v) + (h'(x)^T \lambda | v),$$

donc $\nabla_x L(x, \lambda) = \nabla f(x) + h'(x)^T \lambda$.

2.1. $\forall t, \psi_v(t) = 0 \Rightarrow \psi'_v(t) = h'(\gamma(t)) \cdot \gamma'(t) = 0 \Rightarrow \psi''_v(t) = h'(\gamma(t)) \cdot \gamma''(t) + h''(\gamma(t)) \cdot (\gamma'(t), \gamma'(t)) = 0$.

2.2. Puisque $\varphi''_v(t) = f'(\gamma(t)) \cdot \gamma''(t) + f''(\gamma(t)) \cdot (\gamma'(t), \gamma'(t))$, on a :

$$\begin{aligned} \varphi''_v(\bar{t}) = \varphi''_v(0) &= f'(\bar{x}) \cdot \gamma''(0) + f''(\bar{x}) \cdot (v, v) \\ &= (\nabla f(\bar{x}) | \gamma''(0)) + f''(\bar{x}) \cdot (v, v) \\ &= - (h'(\bar{x})^T \bar{\lambda} | \gamma''(0)) + f''(\bar{x}) \cdot (v, v), \quad \text{car } \nabla_x L(\bar{x}, \bar{\lambda}) = 0, \\ &= - (\bar{\lambda} | h'(\bar{x}) \cdot \gamma''(0)) + f''(\bar{x}) \cdot (v, v), \quad \text{par passage à l'adjoint,} \\ &= (\bar{\lambda} | h''(\bar{x}) \cdot (v, v)) + f''(\bar{x}) \cdot (v, v), \quad \text{d'après la question précédente,} \\ &= \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) \cdot (v, v) \geq 0, \quad \text{car } \bar{t} = 0 \text{ minimum de } (P_v). \end{aligned}$$

La conclusion suit puisque $v \in \text{Ker } h'(\bar{x}) = T(C, \bar{x})$.

□

Optimisation

On rappelle le résultat [4, Théorème 3.11.1] :

Théorème 5.0.1 – de l’application ouverte

Soient E et F deux espaces de Banach et $f \in \mathcal{L}(E, F)$. Si f est surjective, alors il existe une constante $\alpha > 0$ telle que $f(B_E(0, 1)) \supset B_F(0, \alpha)$, ce qui est équivalent à dire que f est ouverte (par linéarité).

que l’on va généraliser dans le cas non linéaire (restreint au cas F de dimension finie). Nous avons tout d’abord besoin du lemme suivant.

Lemme 5.0.1. *Soit E un Banach tel que E est somme directe algébrique¹ de deux sous-espaces vectoriels G et H , tous deux fermés, alors, E est somme directe topologique de G et H .*

► On pose

$$\begin{aligned} \varphi: G \times H &\longrightarrow E \\ (x, y) &\longmapsto \varphi(x, y) := x + y. \end{aligned}$$

Par hypothèse, φ est un isomorphisme algébrique. Montrons que c’est un isomorphisme topologique. Les sous-espaces G et H sont munis de la norme induite par E , notée $\|\cdot\|$. On munit $G \times H$ de la norme produit $\|(x, y)\|_{G \times H} := \|x\| + \|y\|$. L’application φ est continue car $\|\varphi(x, y)\| = \|x + y\| \leq \|x\| + \|y\| = \|(x, y)\|_{G \times H}$. Le sous-espace G est fermé dans un espace complet donc G complet (et fermé) donc G Banach, de même H Banach. Un produit fini de Banach étant un Banach, $G \times H$ Banach. On a donc que l’application $\varphi \in \text{Isom}(G \times H, E)$ est continue, $G \times H$ et E Banach, donc d’après le théorème 1.1.5 de Banach, φ^{-1} est continue, et $\varphi \in \text{Isom}_c(G \times H, E)$. ■

Théorème 5.0.2 – de l’application ouverte non linéaire

Soit $f: U \subset E \rightarrow F := \mathbb{R}^n$ une application de classe \mathcal{C}^1 dans U , définie sur un ouvert U d’un espace de Banach E . Soit $\bar{x} \in U$ un point régulier de f (i.e. $f'(\bar{x})$ surjective), alors f est localement ouverte² en \bar{x} .

► Puisque $f'(\bar{x}) \in \mathcal{L}(E, F)$ surjective, $E/\text{Ker } f'(\bar{x}) \simeq F$ (isomorphisme algébrique). Donc en notant $G := \text{Ker } f'(\bar{x})$, il existe un supplémentaire algébrique $H \simeq F$, de G dans E , i.e. $E = G \oplus H$ (algébrique). Le sous-espace G est fermé car $f'(\bar{x})$ est continue et H est fermé car c’est un sous-espace vectoriel de dimension finie (comme F). D’après le lemme 5.0.1, $E = G \oplus H$ (topologique) et donc en particulier, la projection $p: E \rightarrow G$ est continue. Introduisons l’application $\psi: U \subset E \rightarrow F \times G$, définie par $\psi := (f, p)$, de classe \mathcal{C}^1 et montrons que $\psi'(\bar{x}) = (f'(\bar{x}), p) \in \mathcal{L}(E, F \times G)$ est bijective. L’application $f'(\bar{x})$ est surjective

1. C’est à dire $E = G \oplus H$, i.e. $E = \text{Vect}(G \cup H)$, et $G \cap H = \{0\}$.

par hypothèse, donc $\psi'(\bar{x})$ est surjective. Maintenant, si $\psi'(\bar{x}) \cdot v = (0, 0)$, alors $v \in \text{Ker } f'(\bar{x}) \cap \text{Ker } p = G \cap H = \{0\}$, donc $\psi'(\bar{x})$ est injective. Finalement, $\psi'(\bar{x})$ est bijective, donc d'après le théorème 3.2.3 d'inversion locale, il existe un ouvert $\Omega \in \mathcal{V}_U(\bar{x})$ tel que $\psi|_\Omega$ est un \mathcal{C}^1 -difféomorphisme de Ω sur $\psi(\Omega)$. L'application $\psi|_\Omega$ est donc ouverte, tout comme $f|_\Omega = q \circ \psi|_\Omega$, $q: F \times G \rightarrow F$ projection canonique, car q ouverte. Pour conclure, pour tout $V \in \mathcal{V}_U(\bar{x})$, $V \cap \Omega \in \mathcal{V}_\Omega(\bar{x})$ donc $f|_\Omega(V \cap \Omega) \in \mathcal{V}_F(f(\bar{x}))$ (car $f|_\Omega$ ouverte), et puisque $f|_\Omega(V \cap \Omega) \subset f(V)$, on a $f(V) \in \mathcal{V}_F(f(\bar{x}))$. ■

Théorème 5.0.3 – CN1 d'optimalité, Optimisation

Soit E un espace de Banach. Soient $f \in \mathcal{C}^1(U, \mathbb{R})$ et $h \in \mathcal{C}^1(U, \mathbb{R}^m)$, avec $m \in \mathbb{N}^*$ et U un ouvert de E . Soit $\bar{x} \in U$ une solution (locale ou globale) au problème

$$\min \{f(x) \mid x \in U, h(x) = 0_{\mathbb{R}^m}\}.$$

Il existe alors $(\lambda, \lambda^0) \in (\mathbb{R}^{m+1})^*$, $(\lambda, \lambda^0) \neq 0_{(\mathbb{R}^{m+1})^*}$, tel que $\lambda^0 f'(\bar{x}) + \lambda \circ h'(\bar{x}) = 0_{E'}$.

► Soit $\bar{x} \in U \subset E$ un minimum de f sous la contrainte $h = 0_{\mathbb{R}^m}$. Il existe alors $V \in \mathcal{V}_U(\bar{x})$ (l'ensemble des voisinages de \bar{x} dans U) tel que $\forall x \in V : h(x) = 0_{\mathbb{R}^m} \Rightarrow f(x) \geq f(\bar{x})$. Posons $\varphi: U \subset E \rightarrow \mathbb{R}^m \times \mathbb{R}$, $\varphi(x) := (h(x), f(x))$ de sorte que $\varphi(\bar{x}) = (h(\bar{x}), f(\bar{x})) = (0_{\mathbb{R}^m}, f(\bar{x})) \in \mathbb{R}^{m+1}$. L'application φ ne peut être localement ouverte en \bar{x} car sinon $\varphi(V) \in \mathcal{V}_{\mathbb{R}^{m+1}}(\varphi(\bar{x}))$ et il existerait alors $\varepsilon > 0$ tel que $(0_{\mathbb{R}^m}, f(\bar{x}) - \varepsilon) \in \varphi(V)$, ce qui contredirait l'optimalité de \bar{x} . Puisque φ n'est pas localement ouverte en \bar{x} , d'après le théorème 5.0.2, par contraposée, $\varphi'(\bar{x}) \in \mathcal{L}(E, \mathbb{R}^{m+1})$ n'est pas surjective ; ainsi $\text{Im } \varphi'(\bar{x}) \subsetneq \mathbb{R}^{m+1}$ et il existe donc $\bar{\lambda} := (\lambda, \lambda^0) \in (\mathbb{R}^{m+1})^*$, $\bar{\lambda} \neq 0_{(\mathbb{R}^{m+1})^*}$ tel que $\text{Im } \varphi'(\bar{x}) \subset \text{Ker } \bar{\lambda}$ (puisque $\text{Im } \varphi'(\bar{x})$ est inclus dans un hyperplan, et tout hyperplan est le noyau d'une forme linéaire), i.e. $\forall v \in E$, $\bar{\lambda}(\varphi'(\bar{x}) \cdot v) = 0$, c-a-d $\bar{\lambda} \circ \varphi'(\bar{x}) = 0_{E'}$. ■

Remarque 5.0.1. Dans le théorème précédent, on ne fait pas d'hypothèses de qualification des contraintes, le paramètre λ^0 peut être nul.

2. i.e. envoie les voisinages de \bar{x} dans U sur les voisinages de $f(\bar{x})$ dans $\mathbb{R}^n : \forall V \in \mathcal{V}_U(\bar{x}) \Rightarrow f(V) \in \mathcal{V}_{\mathbb{R}^n}(f(\bar{x}))$.

Bibliographie

- [1] J. W. Milnor, *Topology from the differentiable viewpoint*, Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Based on notes by David W. Weaver, Revised reprint of the 1965 original \leftrightarrow 52.
- [2] F. Paulin, *Géométrie différentielle élémentaire*, Notes de cours, 2006-2007. \leftrightarrow 45.
- [3] F. Paulin, *Topologie, analyse et calcul différentiel*, Notes de cours, 2008-2009. \leftrightarrow 45.
- [4] C. Wagschal, *topologie et analyse fonctionnelle*, Hermann, 2012. \leftrightarrow 3, 4, 5, 11, 12, 41 et 58.

