



Calcul différentiel

Olivier Cots

5 septembre 2023

Table des matières

Chapitre 1. Applications différentiables	1
1.1 Préliminaires	1
1.2 Différentielle de Fréchet	5
1.3 Dérivée directionnelle	8
1.4 Applications de classe \mathcal{C}^1	9
1.5 Théorème de dérivation des applications composées	10
1.6 Fonctions à valeurs dans un espace produit	12
1.7 Fonctions définies sur un espace produit	14
1.8 Fonctions définies et à valeurs dans des espaces produit	17
1.9 Différentielle d'ordre k	19
1.10 Théorèmes des accroissements finis	23
1.11 Dérivabilité versus dérivée partielle	27
Chapitre 2. Grands théorèmes	33
2.1 Théorème du point fixe	33
2.2 Application inversion	34
2.3 Théorème d'inversion locale	37
2.4 Théorème des fonctions implicites	41
Chapitre 3. Extrémum libre et contraintes affines	43
3.1 Extrémum libre	43
3.2 Extrémum lié : contraintes affines	48
3.3 "Two-norm discrepancy"	49
Chapitre 4. Corrections des exercices	55
Bibliographie	63
Index	65

Applications différentiables

1.1	Préliminaires	1
1.2	Différentielle de Fréchet	5
1.3	Dérivée directionnelle	8
1.4	Applications de classe \mathcal{C}^1	9
1.5	Théorème de dérivation des applications composées	10
1.6	Fonctions à valeurs dans un espace produit	12
1.7	Fonctions définies sur un espace produit	14
1.8	Fonctions définies et à valeurs dans des espaces produit	17
1.9	Différentielle d'ordre k	19
1.10	Théorèmes des accroissements finis	23
1.11	Dérivabilité versus dérivée partielle	27

1.1 Préliminaires

Considérons deux espaces vectoriels normés¹ $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sur le corps des réels \mathbb{R} (pour simplifier) et Ω un ouvert de E contenant 0_E . Soit $g: \Omega \subset E \rightarrow F$ une application. On introduit, 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'ensemble $o(v)$ est l'ensemble des fonctions de Ω dans F qui sont négligeables devant l'application identité. 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 $\lim_{\|v\|_E \rightarrow 0} \|\varepsilon(v)\|_F = 0$.

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

- 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 [9, 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 $\mathcal{GL}(E)$.

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.

- 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 normés $(E, \|\cdot\|_E)$ 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 [9, 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.

4. Voir [9, Section 1.6, page 29] pour la définition de base et le résultat d'existence.

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

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

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 [9, 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** [9, 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 2.3.1 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

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 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 le terme d'opérateur pour 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 . ■

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

1.3 Dérivée directionnelle

Définition 1.3.1 – 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.

Remarque 1.3.1. $\mathcal{V}_{\mathbb{R}}(0)$ 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 1.3.2

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 1.3.1. 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 Applications de classe \mathcal{C}^1

Définition 1.4.1 – 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 1.4.1. 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 1.4.2 – 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 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 dans** U si f est \mathcal{C}^1 en tout point de U , autrement dit, si f est différentiable en tout point de U et si $f': U \rightarrow \mathcal{L}(E, F)$ est continue.

Remarque 1.4.2. Dire que f est \mathcal{C}^1 dans U , c'est avoir $f' \in \mathcal{C}^0(U, \mathcal{L}(E, F))$. 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 $\|\cdot\|_{\mathcal{L}(E, F)}$.

1.5 Théorème de dérivation des applications composées

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

Théorème 1.5.1 – 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

$$\begin{aligned} h': U &\longrightarrow \mathcal{L}(E, G) \\ x &\longmapsto h'(x) = g'(f(x)) \circ f'(x). \end{aligned}$$

Notons

$$\begin{aligned} \psi: U &\longrightarrow \mathcal{L}(F, G) \times \mathcal{L}(E, F) \\ x &\longmapsto \psi(x) := (g'(f(x)), f'(x)) \end{aligned}$$

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

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

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 1.5.1 (solution page 55): Soit

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbf{M}_{m,n}(\mathbb{R}), \quad a_i = (a_{i,j})_{1 \leq j \leq n} \in \mathbb{R}^n, \quad i = 1, \dots, m.$$

On pose

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto f(x) := Ax = \sum_{i=1}^m (a_i | x) e_i, \end{aligned}$$

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

1. Montrer que f est différentiable sur \mathbb{R}^n et que $f'(x) \cdot v = Av$, pour tout $v \in \mathbb{R}^n$.
2. Donner f' et en déduire que f est de classe \mathcal{C}^1 sur \mathbb{R}^n .

Exercice 1.5.2 (solution page 55): Soit f une application constante de E dans F . Montrer que f est de classe \mathcal{C}^1 sur E et que $f' = 0_{\mathcal{L}(E, F)}$.

Exercice 1.5.3 (solution page 56): Soit L une application linéaire continue de E dans F . Montrer que L est de classe \mathcal{C}^1 sur E et que $L'(x) = L$.

Exercice 1.5.4 (solution page 56): Soit A une application affine de E dans F telle que $A(x) = L(x) + p$, avec $L \in \mathcal{L}(E, F)$ et $p \in F$. Montrer que A est de classe \mathcal{C}^1 sur E et que $A'(x) = L$.

Exercice 1.5.5 (solution page 56): Soit $B: E \times F \rightarrow G$ (evn), bilinéaire continue. On rappelle que B est continue si et seulement si il existe K tel que pour tout $(x, y) \in E \times F$, $\|B(x, y)\|_G \leq K \|x\|_E \|y\|_F$.

Montrer que B est de classe \mathcal{C}^1 sur $E \times F$ et que $B'(x, y) \cdot (v, w) = B(v, y) + B(x, w)$.

1.6 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

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

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 1.6.1. 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 :

$$\begin{aligned} q_j: \quad F &\longrightarrow F_j \\ y &\longmapsto q_j(y) := y_j \end{aligned}$$

et

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

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) 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

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

Utiliser cet isomorphisme signifie que l'on peut écrire

$$f = (f_j)_{1 \leq j \leq m}.$$

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

Proposition 1.6.1

On a

$$f \text{ dérivable en } x \iff \forall j \in \llbracket 1, m \rrbracket : f_j \text{ dérivable en } x.$$

Cela reste vrai pour la dérivabilité sur U et la classe \mathcal{C}^1 sur U .

► Supposons f dérivable en x . Alors, d'après le Théorème 1.5.1, $f_j = q_j \circ f$ 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

$$f'(x) \cdot v = (f'_j(x) \cdot v)_{1 \leq j \leq m} \in F. \quad (1.1)$$

Remarque 1.6.2. 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 1.6.1 (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, et

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

□

Exemple 1.6.2 (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

$$f'(x) \cdot v = \sum_{j=1}^m (f'_j(x) \cdot v) e_j = \sum_{j=1}^m (\nabla f_j(x) | v) e_j.$$

□

1.7 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 :

$$\begin{aligned} p_i: E &\longrightarrow E_i \\ x &\longmapsto p_i(x) := x_i \end{aligned}$$

et

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

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

$$f \circ \theta_{\bar{x}, i} =: f(\bar{x}_1, \dots, \bar{x}_{i-1}, \cdot, \bar{x}_{i+1}, \dots, \bar{x}_n).$$

Remarque 1.7.1. 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 $f \circ \theta_{x[i],i}$. 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 1.7.1 – 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 1.7.2. On utilisera aussi la notation $\partial_{x_i} f(x)$ pour écrire la dérivée partielle de f par rapport à la variable x_i au point x .

Nous avons alors le résultat suivant et on peut noter que la réciproque du point i) de la Proposition 1.7.2 est fausse !

Proposition 1.7.2

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) $\forall i \in \llbracket 1, n \rrbracket$,

$$\frac{\partial f}{\partial x_i}(x) = f'(x) \circ \lambda_i.$$

$$\text{iii) } f'(x) = \sum_{i=1}^n \partial_{x_i} f(x) \circ p_i.$$

► 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 1.7.3. 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 1.7.1 (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 $\partial_{x_i} f(x) \in \mathcal{L}(\mathbb{R}, F) \simeq \bar{F}$ 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 1.7.4. 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 1.7.2 (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,

$$f'(x) \cdot v = (\nabla f(x) | v)$$

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

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

□

1.8 Fonctions définies et à valeurs dans des espaces produit

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 1.6.2, 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 1.7.2, 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

$$\frac{\partial f_j}{\partial x_i}(x) = q_j \circ f'(x) \circ \lambda_i \in \mathcal{L}(E_i, F_j). \quad (1.2)$$

D'après l'isomorphisme vu en Section 1.6, 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**) :

$$f'(x) = \left(\frac{\partial f_j}{\partial x_i}(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \quad \text{avec} \quad \frac{\partial f_j}{\partial x_i}(x) \in \mathcal{L}(E_i, F_j).$$

Proposition 1.8.1 – 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 $i \in \llbracket 1, n \rrbracket$ et $k \in \llbracket 1, p \rrbracket$ 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 :

$$\begin{aligned} r_k: G &\longrightarrow G_k \\ z &\longmapsto r_k(z) := z_k \end{aligned}$$

et

$$\begin{aligned} \nu_k: G_k &\longrightarrow G \\ a &\longmapsto \nu_k(a) := (0_{G_1}, \dots, 0_{G_{k-1}}, a, 0_{G_{k+1}}, \dots, 0_{G_p}). \end{aligned}$$

On a alors

$$\begin{aligned} \frac{\partial(g \circ f)_k}{\partial x_i}(x) &= r_k \circ (g \circ f)'(x) \circ \lambda_i, && \text{cf. Eq. (1.2)} \\ &= r_k \circ g'(f(x)) \circ f'(x) \circ \lambda_i, && \text{cf. Théorème 1.5.1} \\ &= r_k \circ g'(f(x)) \circ \text{Id}_F \circ f'(x) \circ \lambda_i \\ &= r_k \circ g'(f(x)) \circ \left(\sum_{j=1}^m \mu_j \circ q_j \right) \circ f'(x) \circ \lambda_i, && \text{cf. Section 1.6} \\ &= \sum_{j=1}^m (r_k \circ g'(f(x)) \circ \mu_j) \circ (q_j \circ f'(x) \circ \lambda_i), \\ &&& \text{car } r_k \text{ et } g'(f(x)) \text{ sont linéaires} \\ &= \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \circ \frac{\partial f_j}{\partial x_i}(x). && \text{cf. Eq. (1.2)} \end{aligned}$$

■

Exemple 1.8.1 (Cas fondamental : $E = \mathbb{R}^n$, $F = \mathbb{R}^m$). Dans ce cas,

$$\frac{\partial f_j}{\partial x_i}(x) \in \mathcal{L}(\mathbb{R}, \mathbb{R}) \simeq \mathbb{R}$$

et alors la matrice

$$f'(x) = \left(\frac{\partial f_j}{\partial x_i}(x) \right)_{i,j} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} \in \mathbf{M}_{m,n}(\mathbb{R})$$

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 la base canonique de \mathbb{R}^m $(e_j)_{1 \leq j \leq m}$ et $(dx_i)_{1 \leq i \leq n}$ celle de $(\mathbb{R}^n)^*$. Soit un vecteur $v = (v_i)_{1 \leq i \leq n} \in E = \mathbb{R}^n$. Récapitulons ce que l'on sait déjà :

- $f'(x) \cdot v = \sum_{j=1}^m (f'_j(x) \cdot v) e_j \in \mathbb{R}^m$, cf. Exemple 1.6.2, Section 1.6 ;
- $f'_j(x) \cdot v = \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(x) dx_i(v) \in \mathbb{R}$, $dx_i(v) = v_i$, cf. Exemple 1.7.2, Section 1.7.

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 1.8.1, 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

$$J_{g \circ f}(x) = J_g(f(x)) J_f(x) \quad \text{ou} \quad J_{g \circ f} = (J_g \circ f) \cdot J_f.$$

autrement dit, la matrice jacobienne de la composée de deux fonctions est le produit des matrices jacobiniennes des deux fonctions. \square

Exemple 1.8.2 (Cas particulier : $E = \mathbb{R}^n$, $F = \mathbb{R}^m$, $m = 1$). Voir l'Exemple 1.7.2 de la Section 1.7. Dans ce cas,

$$J_f(x) \in \mathbf{M}_{1,n}(\mathbb{R}) \simeq (\mathbb{R}^n)^*,$$

autrement dit $f'(x)$ est une forme linéaire sur \mathbb{R}^n et $J_f(x)$ est un vecteur ligne. Ainsi,

$$J_f(x)^T = \nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix} \in \mathbf{M}_{n,1}(\mathbb{R}) \simeq \mathbb{R}^n,$$

où A^T est la matrice transposée de la matrice A , et donc $\nabla f(x)$ est un vecteur colonne. \square

1.9 Différentielle d'ordre k

Définition 1.9.1 – 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 1.9.1. 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 1.9.2. 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 1.9.1 (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 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 1.9.2 (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 :

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

avec

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{bmatrix} \in \mathbf{M}_n(\mathbb{R}).$$

\square

Différentielle d'ordre k . On introduit la notation

$$\mathcal{E}_0 := F, \quad \mathcal{E}_1 := \mathcal{L}(E, F) = \mathcal{L}(E, \mathcal{E}_0),$$

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

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

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 $((T \cdot x_1) \cdot x_2) \dots \cdot x_k$. 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 1.9.2 – 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 **infinitement 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 1.9.3. Une application de classe \mathcal{C}^k est aussi de classe \mathcal{C}^j , pour tout $j \leq k$.

1.10 Théorèmes des accroissements finis

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 Rolle⁶ 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 1.10.1. 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, $f'(x) = 2\pi(-\sin(2\pi x), \cos(2\pi x))$, 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 1.10.1 – 1^{re} forme des accroissements finis

Soient $a < b$ deux réels, $(F, \|\cdot\|_F)$ 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 $I_\varepsilon = [a, b]$. Nous aurons alors l'inégalité pour $x = b$, et ce, quelque soit $\varepsilon > 0$. 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

6. cf. <https://tinyurl.com/Rolle-wikipedia>.

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 $x - c_\varepsilon > 0$:

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

Ensuite, puisque nous avons $g'(c_\varepsilon) \geq 0$, alors, le théorème de Rolle nous donne $g(x) - g(c_\varepsilon) \geq 0$ et 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

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

En prenant $x = y$ il vient que $y \in I_\varepsilon$, ce qui est 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. ■

Définition 1.10.2

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 $k \in [0, 1[$.

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

Corollaire 1.10.3

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 1.10.4

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 1.10.1, on obtient :

Corollaire 1.10.5

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. 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 1.10.6 – 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 1.10.1, 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 1.10.7

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 :

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

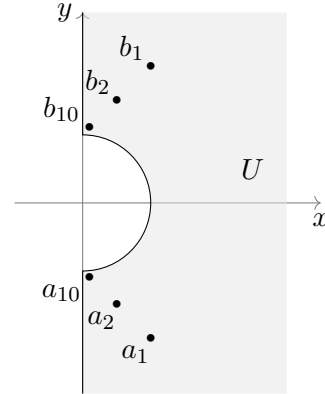
► Puisque U convexe, alors pour tout $(x, y) \in U^2$, le segment $[x, y] \subset U$ et on peut appliquer le Théorème 1.10.6. ■

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 1.10.2. 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 $k < \pi/2$. \square

Sur un ouvert connexe, nous avons le résultat suivant.

Corollaire 1.10.8

*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 1.10.7 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 1.10.1. 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.

1.11 Dérivabilité versus dérivée partielle

Proposition 1.11.1

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 1.7.2. 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.

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

* 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) suivante⁷ :

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

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 1.10.6 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 1.11.1. Attention, une application qui admet des dérivées partielles n'est pas nécessairement dérivable, ni même continue, cf. Exemple 1.3.1.

Exercice 1.11.1 (solution page 57): Soit

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto f(x, y, z) = \begin{bmatrix} y \cos x - z \sin x \\ x y z \end{bmatrix}. \end{aligned}$$

Montrer que f est de classe \mathcal{C}^1 sur \mathbb{R}^3 et donner l'expression de f' .

Exercice 1.11.2 (solution page 57): Montrer que toute application bilinéaire continue, puis multilinéaire continue, est de classe \mathcal{C}^1 , et donner leurs différentielles.

Exercice 1.11.3 (solution page 58): Soient $A \in \mathcal{L}(\mathbb{R}^n)$, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$. On pose

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \frac{1}{2} (Ax | x) + (b | x) + c. \end{aligned}$$

Montrer que f est dérivable 2 fois sur \mathbb{R}^n et calculer ∇f et $\nabla^2 f$. Même calculs en supposant $A \in \text{Sym}(n, \mathbb{R})$.

Exercice 1.11.4 (solution page 58): Soient $A \in \text{Sym}(n, \mathbb{R})$, $b \in \mathbb{R}^n$ et l'applica-

tion $x: \mathbb{R} \rightarrow \mathbb{R}^n$ deux fois dérivable. On considère l'application f définie par :

$$\begin{aligned} f: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto f(t) = \frac{1}{2} (A x(t) | x(t)) - (b | x(t)). \end{aligned}$$

1. Montrer que f est deux fois dérivable sur \mathbb{R} .
2. Exprimer $f'(t)$ et $f''(t)$.

Exercise 1.11.5 (solution page 59): Soit

$$\begin{aligned} \varphi: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \varphi(x) = \frac{1}{2} \|x\|^2 = \frac{1}{2} \sum_{i=1}^n x_i^2. \end{aligned}$$

Montrer que φ est deux fois dérivable sur \mathbb{R}^n et exprimer son développement limité à l'ordre 2 en tout point.

Exercise 1.11.6 (solution page 59): Soit

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \cos\left(\frac{1}{2} \|x\|^2\right). \end{aligned}$$

Montrer que f est deux fois dérivable sur \mathbb{R}^n , et calculer $\nabla f(x)$ ainsi que la matrice hessienne $\nabla^2 f(x)$ en tout point $x \in \mathbb{R}^n$.

Exercise 1.11.7 (solution page 60): Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \|x\|$.

1. Montrer que f est dérivable sur l'ouvert $\mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ et que $\nabla f(x) = \frac{x}{\|x\|}$.
2. Montrer que f n'est pas dérivable en $0_{\mathbb{R}^n}$.
3. Montrer que f est deux fois dérivable sur $\mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ et donner $\nabla^2 f$.

Exercise 1.11.8 (solution page 61): Soit $F: \mathbb{R}^n \rightarrow \mathbb{R}^p$ dont chaque application composante F_i , $i = 1, \dots, p$, est deux fois dérivable sur \mathbb{R}^n . On définit

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \frac{1}{2} \|F(x)\|^2. \end{aligned}$$

1. Montrer que f est dérivable deux fois sur \mathbb{R}^n et calculer ∇f et $\nabla^2 f$.
2. Application au cas particulier où $f(x) = \frac{1}{2} \|Ax - b\|^2$.

Proposition 1.11.2

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 1.6.1, 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 1.11), donc en particulier en $f(\bar{x})$. Par composition, cf. Théorème 1.5.1, 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 (1.1). Par l'Exercice 1.11, 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 1.11.3. Si $E = E_1 = E_2 = F = \mathbb{R}$, et si B est le produit usuel \times entre réels, alors la formule de la Proposition 1.11.2 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 1.11.4. 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 la Proposition 1.11.2 ($U = E_1 = E_2 = E$ 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 $Q'(x) \cdot v = 2B(x, v)$.

La proposition précédente 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 1.11. Pour x et v dans E^n , on a :

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

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

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

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

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

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 :

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

Grands théorèmes

2.1 Théorème du point fixe	33
2.2 Application inversion	34
2.3 Théorème d'inversion locale	37
2.4 Théorème des fonctions implicites	41

2.1 Théorème du point fixe

Théorème 2.1.1 – du point fixe de Picard

Soit $f : X \rightarrow X$ une application contractante sur un espace métrique complet¹ (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) 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.

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

1. Soit (X, d) un espace métrique. Il est complet si toute suite de Cauchy converge. Une suite (x_n) de X vérifiant $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n, m \geq N, d(x_n, x_m)_E \leq \varepsilon$, est dite de Cauchy.

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) est une suite de Cauchy. Puisque X est complet, la suite (x_n) converge vers un élément de X noté x . Posons

$$u_n := f(x_n) - x_n.$$

La suite (u_n) converge vers $f(x) - x$ (car f continue). Mais puisque pour $n \geq 1$,

$$u_n = f(x_n) - x_n = f(x_n) - f(x_{n-1}),$$

la suite (u_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 2.1.2 – 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 2.1.1, 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. ■

2.2 Application inversion

Rappelons quelques définitions.

Définition 2.2.1 – Espace de Banach

Un **espace de Banach** est un espace vectoriel normé $(E, \|\cdot\|_E)$ qui est complet pour la distance $d_E(x, y) := \|x - y\|_E$.

Définition 2.2.2

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

On considère deux espaces de Banach E et F et on suppose que $\text{Isom}_c(E, F)$ est non vide. D'après [9, Théorème 3.19.8], 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 nous ramenons dans un premier temps 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 :

$$\text{inv}(T) = \varphi(T)^{-1} \circ B = \text{inv}_E(\varphi(T)) \circ B. \quad (2.1)$$

Ainsi, par le théorème de dérivation des applications composées, 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 $A \in \text{GL}_c(E)$.

* 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 [9, Proposition 3.19.6], si $\|T^{-1}H\|_{\mathcal{L}(E)} < 1$, alors $I - T^{-1}H$ est inversible et

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

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

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

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, 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 (2.1).

Remarque 2.2.1. 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 $A \in \text{GL}_n(\mathbb{R})$,

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

* Enfin, montrons que inv est \mathcal{C}^∞ . Nous pouvons écrire,

$$\text{inv}'_E = B \circ \varphi \circ \text{inv}_E$$

avec $\varphi(A, B) := (A, B)$ linéaire continue donc \mathcal{C}^∞ et B bilinéaire continue donc \mathcal{C}^∞ . Ainsi, par induction, inv_E est \mathcal{C}^∞ tout comme inv .

2.3 Théorème d'inversion locale

Nous aurons besoin du résultat suivant, voir [9, Corollaire 3.11.3].

Théorème 2.3.1 – de Banach

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

Définition 2.3.2 – 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, sont des homéomorphismes.

Remarque 2.3.1. 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 2.3.3

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 2.3.1 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

$$T_y(h) = x_h - x + T_y(\psi(x_h - x)) = g(y + h) - g(y) + T_y(\psi(x_h - x)). \quad (2.2)$$

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 $x_h - x = O(h)$,² 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 (2.2), 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}^∞ , cf. 2.2, 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. ■

Remarque 2.3.2. Rappelons que l'ensemble $\text{Isom}_c(E, F)$ est ouvert dans $\mathcal{L}(E, F)$ d'après [9, 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 2.3.3, l'application inv est un \mathcal{C}^∞ -difféomorphisme.

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

3. D'après le Théorème 2.3.1, dans ce cas $f'(x) \in \text{Isom}_c(E, F)$.

Théorème 2.3.4 – 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.³ 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 1.10.6 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 :

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

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 2.1.2 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 2.3.3, f_0 est un \mathcal{C}^k -difféomorphisme de U_0 dans V_0 et finalement le théorème est démontré. ■

2.4 Théorème des fonctions implicites

Théorème 2.4.1 – 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

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

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 2.3.4 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 2.4.1. 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 2.4.1. 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 $x'(b) = A^{-1}$. □

Extrémum libre et contraintes affines

3.1 Extrémum libre

Définition 3.1.1

- i) Soient X un ensemble quelconque, $J: X \rightarrow \mathbb{R}$ et $\bar{x} \in X$. On dit que \bar{x} est un **minimum global** de J si $J(\bar{x}) \leq J(x)$ pour tout $x \in X$.
- ii) Si X est un espace topologique, on dit que \bar{x} est un **minimum local** de J si $\exists V \in \mathcal{V}_X(\bar{x})$ (i.e. un voisinage ouvert de \bar{x} dans X) t.q. $\forall x \in V : J(\bar{x}) \leq J(x)$.
- iii) Un minimum local est dit **strict** si : $\forall x \in V \setminus \{\bar{x}\} : J(\bar{x}) < J(x)$. On a de même les notions de maximum global, local et local strict.
- iv) On dit que \bar{x} est un **extremum** de J (respectivement global, local, local strict) si \bar{x} est un minimum ou maximum de J (respectivement global, local, local strict).
- v) Soient E un espace vectoriel normé (evn) et $J: U \subset E \rightarrow \mathbb{R}$, U un ouvert de E , J dérivable en $\tilde{x} \in U$. On dit que \tilde{x} est un **point critique** de J si $J'(\tilde{x}) = 0_{E'}$, et alors $J(\tilde{x})$ est appelée une **valeur critique**.

Remarque 3.1.1.

- On rappelle que E' est l'ensemble des formes linéaires continues sur E et que la notion de dérivabilité est ici sous-entendue au sens de Fréchet. On notera $\|\cdot\|$ la norme de E .
- Un point critique n'est pas nécessairement un extrémum.
- En un point critique, on dit aussi que J est **stationnaire**.

Proposition 3.1.2 – Condition du premier ordre (CN1)

Soient $J: U \subset E \rightarrow \mathbb{R}$ dérivable sur U ouvert de E evn et $\bar{x} \in U$. Si \bar{x} est un extrémum local de J alors l'équation de Euler est vérifiée :

$$J'(\bar{x}) = 0_{E'}.$$

► Si \bar{x} est par exemple un minimum (c'est similaire pour un max.) local alors :

$$\exists \eta > 0, \text{ t.q. } B(\bar{x}, \eta) \subset U \text{ et } \forall x \in B(\bar{x}, \eta) : J(\bar{x}) \leq J(x),$$

puisque U est un ouvert de E . Soit maintenant $h \in E$. Alors, il existe $\rho_{\eta,h} > 0$ t.q. :

$$\begin{aligned} f: V_{\eta,h} :=]-\rho_{\eta,h}, \rho_{\eta,h}[&\longrightarrow \mathbb{R} \\ s &\longmapsto f(s) := J(\bar{x} + sh) \end{aligned}$$

soit bien définie et t.q. $\bar{x} + sh \in B(\bar{x}, \eta)$, $\forall s \in V_{\eta, h}$. Si $h \neq 0_E$, prendre $\rho_{\eta, h} := \eta/\|h\|$. Ainsi définie, f est dérivable en 0 et $f'(0) = J'(\bar{x}) \cdot h$. Mais puisque \bar{x} est un minimum local de J , alors 0 est un minimum de f sur $V_{\eta, h}$. Ainsi, puisque $V_{\eta, h}$ est un ouvert de \mathbb{R} , il vient $f'(0) = J'(\bar{x}) \cdot h = 0$, car

$$0 \geq \lim_{s \rightarrow 0^-} \frac{f(s) - f(0)}{s} = f'(0) = \lim_{s \rightarrow 0^+} \frac{f(s) - f(0)}{s} \geq 0.$$

■

Proposition 3.1.3 – Cas convexe (CNS)

Soient $J: U \subset E \rightarrow \mathbb{R}$ convexe et dérivable sur U un ouvert convexe de E , et $\bar{x} \in U$. Si \bar{x} est un point critique de J alors \bar{x} est un minimum global de J (sous-entendu sur U). Nous avons donc

$$\bar{x} \in U \text{ minimum global} \Leftrightarrow \bar{x} \in U \text{ point critique.}$$

► Puisque J est dérivable et convexe sur U alors $\forall (x, y) \in U^2 : J(x) - J(y) \geq J'(y) \cdot (x - y)$. En particulier, au point $y := \bar{x}$, on a $\forall x \in U : J(x) - J(\bar{x}) \geq J'(\bar{x}) \cdot (x - \bar{x}) = 0$, puisque \bar{x} est un point critique de J . Il vient donc que \bar{x} est un minimum global de J sur U . ■

Proposition 3.1.4 – Condition nécessaire d'ordre supérieur (CN_k)

Soit $J: U \subset E \rightarrow \mathbb{R}$ une application $k \geq 2$ fois dérivable sur U un ouvert de E (evn) et soit $\bar{x} \in U$ un point critique de J t.q. $\forall i \in \llbracket 1, k-1 \rrbracket : J^{(i)}(\bar{x}) = 0$ et $J^{(k)}(\bar{x}) \neq 0$. On a :

$$\bar{x} \text{ minimum local de } J \Rightarrow k \text{ pair et } \forall h \in E : J^{(k)}(\bar{x}) \cdot h^k \geq 0.$$

► Supposons donc que \bar{x} soit un minimum local de J et considérons $h \in E$. On définit $f(s)$ comme dans la preuve de la Proposition 3.1.2 et on garde les mêmes notations. Alors, $\forall s \in V_{\eta, h}$:

$$f(s) - f(0) = J(\bar{x} + sh) - J(\bar{x}) \geq 0.$$

Puisque J est k fois dérivable sur U alors f l'est sur $V_{\eta, h}$ donc en particulier en 0. D'après la formule de Taylor-Young :

$$f(s) - f(0) = \sum_{i=1}^k \frac{s^i}{i!} f^{(i)}(0) + s^k \varepsilon(s) \quad \text{avec} \quad \lim_{s \rightarrow 0} \varepsilon(s) = 0.$$

Par hypothèse :

$$f(s) - f(0) = \frac{s^k}{k!} f^{(k)}(0) + s^k \varepsilon(s) = s^k \left(\frac{1}{k!} J^{(k)}(\bar{x}) \cdot h^k + \varepsilon(s) \right) \geq 0,$$

En divisant par s^k et en passant à la limite quand $s \rightarrow 0^+$ dans le terme de droite on en déduit que $J^{(k)}(\bar{x}) \cdot h^k \geq 0$. Montrons pour finir que k est pair. Puisque $J^{(k)}(\bar{x}) \neq 0$,

alors $\exists h_0 \in E$ t.q. $J^{(k)}(\bar{x}) \cdot h_0^k \neq 0$ ¹ et donc $J^{(k)}(\bar{x}) \cdot h_0^k > 0$. Posons

$$\Omega := \left\{ s \in \mathbb{R}^* \mid |\varepsilon(s)| < \frac{1}{k!} J^{(k)}(\bar{x}) \cdot h_0^k \right\} \cap V_{\eta, h_0},$$

de telle sorte que $\forall s \in \Omega$:

$$\text{sign}(f(s) - f(0)) = \text{sign}(s^k) \times \text{sign} \left(J^{(k)}(\bar{x}) \cdot h_0^k \right) = \text{sign}(s^k).$$

Puisque $f(s) - f(0) \geq 0$, $\Omega \cap \mathbb{R}_- \neq \emptyset$ et $\Omega \cap \mathbb{R}_+ \neq \emptyset$, on en déduit que k est pair. ■

Proposition 3.1.5 – Conditions suffisantes d'ordre supérieur (CSk)

Soit $J: U \subset E \rightarrow \mathbb{R}$ une application $k \geq 2$ fois dérivable sur U un ouvert de E (evn) et soit $\bar{x} \in U$ un point critique de J t.q. $\forall i \in \llbracket 1, k-1 \rrbracket : J^{(i)}(\bar{x}) = 0$ et $J^{(k)}(\bar{x}) \neq 0$. On a :

i) Si

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E : J^{(k)}(\bar{x}) \cdot h^k \geq \alpha \|h\|^k,$$

alors k est pair et \bar{x} est un minimum local strict.

ii) Si

$$\exists \eta > 0 \quad \text{t.q.} \quad \forall x \in B(\bar{x}, \eta) \subset U \quad \text{et} \quad \forall h \in E : J^{(k)}(x) \cdot h^k \geq 0,$$

alors k est pair et \bar{x} est un minimum local.

► i) Pour tout $h \in E$, $J^{(k)}(\bar{x}) \cdot (-h)^k = (-1)^k J^{(k)}(\bar{x}) \cdot h^k \geq \alpha \|h\|^k \geq 0$ donc k doit être pair. La formule de Taylor appliquée à J nous donne pour $h \in E$ t.q. $\bar{x} + h \in U$:

$$J(\bar{x} + h) - J(\bar{x}) = \frac{1}{k!} J^{(k)}(\bar{x}) \cdot h^k + \|h\|^k \varepsilon(h) \quad \text{avec} \quad \lim_{\|h\| \rightarrow 0} |\varepsilon(h)| = 0.$$

Par hypothèse :

$$J(\bar{x} + h) - J(\bar{x}) \geq \left(\frac{\alpha}{k!} + \varepsilon(h) \right) \|h\|^k$$

et donc il existe $\eta > 0$ tel que pour tout $h \neq 0_E$ t.q. $\|h\| < \eta$ on ait $\bar{x} + h \in U$ et

$$J(\bar{x} + h) - J(\bar{x}) > 0.$$

Autrement dit, \bar{x} est un minimum local strict.

ii) Il est évident que k soit pair car $J^{(k)}(\bar{x}) \neq 0$ (cf. Proposition précédente). Par hypothèse, J est donc convexe sur la boule ouverte $B(\bar{x}, \eta)$ et puisque \bar{x} est un point critique, alors \bar{x} est un minimum global de J sur $B(\bar{x}, \eta)$, autrement dit un minimum local de J , cf. Proposition 3.1.3. ■

1. Voir [4, Chapitre IV, § 3, Problème 3, page 92].

Définition 3.1.6

Soit $B \in \mathcal{L}_2(E^2, \mathbb{R})$ une forme bilinéaire symétrique définie sur $(E, \|\cdot\|)$ un espace vectoriel normé.

- B est dite **semi-définie positive**, notée $B \succeq 0$, si $\forall h \in E : B(h, h) \geq 0$.
- B est dite **définie positive**, notée $B \succ 0$, si $\forall h \in E \setminus \{0_E\} : B(h, h) > 0$.
- B est dite **elliptique** si $\exists \alpha > 0$ t.q. $\forall h \in E : B(h, h) \geq \alpha \|h\|^2$.

D'après la Proposition 3.1.4, en un point critique \bar{x} , une condition nécessaire d'ordre deux est donc que $J''(\bar{x})$ soit semi-définie positive, et d'après la Proposition 3.1.5, une condition suffisante d'ordre deux est que $J''(\bar{x})$ soit elliptique :

$$\begin{aligned} \text{(CN2)} \quad & \forall h \in E : J''(\bar{x}) \cdot (h, h) \geq 0, \\ \text{(CS2)} \quad & \exists \alpha > 0 \text{ t.q. } \forall h \in E : J''(\bar{x}) \cdot (h, h) \geq \alpha \|h\|^2. \end{aligned} \quad (3.1)$$

Une autre condition suffisante d'ordre deux est que $J''(x)$ soit semi-définie positive pour tout x contenu dans une boule ouverte centrée en \bar{x} . Attention, la définie positivité ne suffit pas en général en dimension infinie mais suffit en dimension finie comme le montre le résultat suivant.

Proposition 3.1.7 – Cas particulier de la dimension finie

Soit $J : U \subset E \rightarrow \mathbb{R}$ une application $k \geq 2$ fois dérivable sur U un ouvert de E (evn) et soit $x \in U$. Si $\dim E \in \mathbb{N}$ alors :

$$\exists \alpha > 0 \text{ t.q. } \forall h \in E : J^{(k)}(x) \cdot h^k \geq \alpha \|h\|^k \Leftrightarrow \forall h \in E \setminus \{0_E\} : J^{(k)}(x) \cdot h^k > 0.$$

► La conditions nécessaire est évidente et vraie même si $\dim E = +\infty$. Supposons donc que $\forall h \in E \setminus \{0_E\} : J^{(k)}(x) \cdot h^k > 0$. Introduisons

$$\begin{aligned} \varphi : E &\longrightarrow \mathbb{R} & \psi : E &\longrightarrow E^k \\ h &\longmapsto \varphi(h) := f^{(k)}(x) \cdot h^k & \text{et} & & h &\longmapsto \psi(h) = (h, \dots, h) = h^k \end{aligned}$$

de telle sorte que $\varphi = f^{(k)}(x) \circ \psi$, avec ψ une application linéaire continue. Puisque $f^{(k)}(x) \in \mathcal{L}_k(E^k, \mathbb{R})$ est aussi continue, alors φ est continue. Or, $\dim E < +\infty$, donc sa sphère unité $S(0_E, 1)$ est compacte, donc φ est bornée sur $S(0_E, 1)$ et atteint sa borne inférieure, c'est-à-dire :

$$\exists h_0 \in S(0_E, 1) \quad \text{t.q.} \quad \forall h \in S(0_E, 1) : \varphi(h_0) \leq \varphi(h).$$

Notons $\alpha := f^{(k)}(x) \cdot h_0^k > 0$ (car $h_0 \neq 0_E$). Alors, $\forall h \in S(0_E, 1)$ on a $f^{(k)}(x) \cdot h^k \geq \alpha$ et donc par homogénéité, on en déduit :

$$\forall h \in E \setminus \{0_E\} : \frac{h}{\|h\|} \in S(0_E, 1) \Rightarrow f^{(k)}(x) \cdot \left(\frac{h}{\|h\|} \right)^k \geq \alpha \Rightarrow f^{(k)}(x) \cdot h^k \geq \alpha \|h\|^k.$$

■

Remarque 3.1.2. Dans les preuves des propositions 3.1.2 (CN1) et 3.1.4 (CNk), nous nous sommes ramené dans \mathbb{R} , et nous avons utilisé directement les conditions nécessaires pour une fonctionnelle dans \mathbb{R} pour obtenir les conditions nécessaires en dimension infinie. Ce procédé ne fonctionne pas pour obtenir les conditions suffisantes de la Proposition 3.1.5 (CSk). En effet, même si \bar{x} est un minimum local le long de n'importe quelle direction h , c'est-à-dire un minimum local de $f(s) = J(\bar{x} + sh)$, alors il est toujours possible que \bar{x} ne soit pas un minimum local de J , et ce, même en dimension finie (supérieure ou égale à 2), cf. exercice 3.1. De plus, comme nous l'avons vu, il y a une différence notable entre les dimensions finie et infinie concernant les conditions suffisantes. La définie positivité de la dérivée seconde n'est pas une condition suffisante d'optimalité locale en dimension infinie, cf. Exemple 3.1.1. C'est pour ces raisons que nous avons reporté la formalisation des concepts d'optimisation. Dans les sections précédentes, pour obtenir les équations de Euler-Lagrange et la condition de Legendre, seules les conditions nécessaires étaient utiles. De plus, comme nous allons vite le voir, ces conditions suffisantes ne nous seront pas utiles, c'est pourquoi nous allons en donner d'autres dans la prochaine section.

Voici un exercice montrant que l'optimalité "directionnelle" n'implique pas l'optimalité.

Exercice 3.1.1: (Extrait de [5]) On considère la fonction

$$\begin{aligned} J: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ x &\longmapsto J(x) := 2x_1^4 - 4x_1^2x_2 + x_2^2. \end{aligned}$$

1. Montrer que l'unique point critique de J est $\tilde{x} := (0, 0)$.
2. La condition nécessaire de solution du deuxième ordre est-elle vérifiée ?
3. Les conditions suffisantes de solution du deuxième ordre sont-elles vérifiées ?
4. On fixe $h \in \mathbb{R}^2$ et on considère la fonction $f: \mathbb{R} \rightarrow \mathbb{R}$, $s \mapsto f(s) := J(\tilde{x} + sh)$. Montrer que $\bar{s} = 0$ est un minimum local de f .
5. Calculer $J(x_1, 2x_1^2)$. Conclusion (faire le lien avec la remarque 3.1.2) ?

L'exemple suivant montre que la définie positivité de la dérivée seconde n'est pas une condition suffisante d'optimalité locale.

Exemple 3.1.1 (Extrait de [2]). On considère le problème

$$\inf_x \left\{ J(x) := \sum_{n=1}^{\infty} \left(\frac{x_n^2}{n} - x_n^3 \right) \mid x \in l_2 \right\},$$

où $l_2 := \{(x_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ est l'espace de Hilbert des suites de carrés sommables. Alors, le point $\tilde{x} := 0$ est un point critique tel que pour tout $h := (h_n) \in l_2$ non nul :

$$J''(\tilde{x}) \cdot (h, h) = \sum_{n=1}^{\infty} \frac{2h_n^2}{n} > 0.$$

Cependant, considérons les suites $h_k := (h_{k,n})$ telles que toutes les coordonnées sont nulles sauf la k^{e} qui vérifie $h_{k,k} = 2k^{-1}$. Alors, $J(\tilde{x} + h_k) = J(h_k) = -4k^{-3} < 0 = J(\tilde{x})$ et donc \tilde{x} n'est pas localement optimal. Voici donc un exemple de dérivée seconde définie positive mais non elliptique. \square

3.2 Extremum lié : contraintes affines

Considérons un espace vectoriel normé $(E, \|\cdot\|_E)$ et une fonctionnelle $J: E \rightarrow \mathbb{R}$. Considérons de plus un autre espace vectoriel normé $(F, \|\cdot\|_F)$, une application linéaire continue $T \in \mathcal{L}(E, F)$ et un point $y \in F$. On s'intéresse au problème aux contraintes affines suivant :

$$\inf_x \{J(x) \mid x \in \mathcal{K}\}, \quad (P_A)$$

où $\mathcal{K} := \{x \in E \mid \Phi(x) := T(x) - y = 0_F\} = \Phi^{-1}(\{0_F\})$ avec $\Phi: E \rightarrow F$ une application affine continue de E dans F . Supposons que $y \in \text{Im } T$, c'est-à-dire que :

$$\exists u \in E \text{ t.q. } T(u) = y, \text{ i.e. t.q. } \Phi(u) = 0_F.$$

Introduisons alors

$$E_0 := \text{Ker } T \subset E$$

de telle sorte que $\forall v \in E_0 : \Phi(u+v) = T(u+v) - y = T(u) + T(v) - y = \Phi(u) + T(v) = 0_F$, c'est-à-dire :

$$\Phi(u+v) = 0_F \Leftrightarrow v \in E_0.$$

On munit E_0 de la norme induite par la norme de E , i.e. $\|v\|_{E_0} = \|v\|_E$ pour tout $v \in E_0$. Considérons alors la fonctionnelle

$$\begin{aligned} F: E_0 &\longrightarrow \mathbb{R} \\ v &\longmapsto F(v) := J(u+v), \end{aligned}$$

de telle sorte que le problème (P_A) est équivalent au problème sans contraintes suivant :

$$\inf_{v \in E_0} F(v). \quad (P)$$

Ici, "équivalent" signifie : $\bar{x} := u + \bar{v}$ solution de $(P_A) \Leftrightarrow \bar{v}$ solution de (P) .

Remarque 3.2.1. Si on restreint la recherche de la solution de (P_A) à un ouvert $U \subset E$, c'est-à-dire si on s'intéresse au problème

$$\inf_x \{J(x) \mid x \in \mathcal{K}\}, \quad \mathcal{K} := \Phi^{-1}(\{0_F\}) \cap U,$$

alors ce problème est équivalent au problème sans contraintes :

$$\inf_v \{F(v) \mid v \in U_0 \subset E_0\},$$

où $U_0 := E_0 \cap U$ est un ouvert de E_0 . On est alors précisément dans le cadre de l'item v) de la définition 3.1.1.

Écrit de manière schématique, les conditions nécessaires et suffisantes d'ordres 1 et 2 pour le problème avec contraintes affines (P_A) sont donc :

$$\begin{aligned} (\text{CN1}) \quad & F'(\bar{v}) = J'(\bar{x}) = 0_{E_0'}, \text{ i.e. } \forall h \in E_0 : J'(\bar{x}) \cdot h = 0 ; \\ (\text{CN2}) \quad & \forall h \in E_0 : F''(\bar{v}) \cdot (h, h) = J''(\bar{x}) \cdot (h, h) \geq 0 ; \\ (\text{CS2}) \quad & \exists \alpha > 0 \text{ t.q. } \forall h \in E_0 : F''(\bar{v}) \cdot (h, h) = J''(\bar{x}) \cdot (h, h) \geq \alpha \|h\|_{E_0}^2 ; \end{aligned} \quad (3.2)$$

où bien sûr $\bar{x} := u + \bar{v}$ doit être un point admissible, c'est-à-dire $\bar{x} \in \mathcal{K}$.

3.3 “Two-norm discrepancy”

Comme nous allons le voir dans l'exemple suivant, la condition suffisante (CS2) présentée ci-avant n'est pas toujours utile en pratique pour des problèmes en dimension infinie.

Exemple 3.3.1. On s'intéresse au problème suivant :

$$\inf_x \left\{ J(x) := \int_0^1 \sin(\dot{x}(t)) dt \mid x \in \mathcal{C}^1([0, 1], \mathbb{R}), x(0) = 0, x(1) = -\frac{\pi}{2} \right\},$$

dont le lagrangien associé est donné par $L(t, x, u) := \sin(u)$. Il est facile de voir que la fonction définie par $\bar{x}(t) := -\frac{\pi}{2}t$ est la solution globale du problème et vérifie

$$J(\bar{x}) = -1.$$

La CN1 est bien vérifiée car :

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J'(\bar{x}) \cdot h = \int_0^1 \cos(\dot{\bar{x}}(t)) \dot{h}(t) dt = 0.$$

La CN2 est aussi vérifiée car :

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J''(\bar{x}) \cdot h^2 = - \int_0^1 \sin(\dot{\bar{x}}(t)) \dot{h}^2(t) dt = \int_0^1 \dot{h}^2(t) dt = \|\dot{h}\|_{L^2}^2 \geq 0,$$

où pour toute fonction mesurable (sous-entendu pour la mesure de Lebesgue) $f: [0, 1] \rightarrow \mathbb{R}$, on a défini :

$$\|f\|_{L^2} := \left(\int_0^1 |f(t)|^2 dt \right)^{\frac{1}{2}} \in \mathbb{R}_+,$$

de telle sorte que $\|\cdot\|_{L^2}$ est la norme usuelle associée à l'espace de Hilbert $L^2([0, 1], \mathbb{R})$ des (classes de) fonctions mesurables de $[0, 1]$ à valeurs dans \mathbb{R} de carrés intégrables, c'est-à-dire :

$$L^2([0, 1], \mathbb{R}) := \{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ mesurable et } \|f\|_{L^2} < +\infty\}.$$

En revanche, la CS2 d'ellipticité de la dérivée seconde $J''(\bar{x})$ n'est pas vérifiée car on a dans notre exemple :

$$\|\dot{h}\|_{L^2} \leq \|\dot{h}\|_{\infty} \leq \|h\|_{\mathcal{C}^1} = \|h\|_{\infty} + \|\dot{h}\|_{\infty},$$

mais pas l'inverse, c'est-à-dire, il n'existe pas de $\alpha > 0$ tel que $\|\dot{h}\|_{L^2} \geq \alpha \|h\|_{\mathcal{C}^1}$. \square

Remarque 3.3.1. Dans l'exemple précédent, nous ne pouvons pas montrer que \bar{x} est un minimum local strict en utilisant la condition suffisante d'ordre deux d'ellipticité de $J''(\bar{x})$. Cependant, il est aisé de montrer que \bar{x} est un minimum local à l'aide de la condition suffisante de semi-définie positivité de $J''(x)$ pour tout $x \in B(\bar{x}, \eta)$, pour un certain $\eta > 0$. Il suffit de prendre $\eta := \pi/2$ par exemple (la preuve est laissée en exercice).

Poursuivons avec l'exemple précédent. Nous avons donc

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J''(\bar{x}) \cdot h^2 = \|\dot{h}\|_{L^2}^2.$$

Ce qui est gênant ici, c'est qu'il y a dans le terme de gauche h tandis que dans celui de droite, on retrouve \dot{h} . On aimerait avoir quelque chose qui ressemble à :

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) \subset H_0 : J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_{H_0}^2,$$

où $\|\cdot\|_{H_0}$ est une norme à définir sur un espace H_0 lui-même à définir. Introduisons pour cela l'espace de Sobolev $H^1([0, 1], \mathbb{R}) := W^{1,2}([0, 1], \mathbb{R})$ des fonctions f dans $L^2([0, 1], \mathbb{R})$ qui s'écrivent² sous la forme $f(t) = f(0) + \int_0^t g(s) ds$ avec $g \in L^2([0, 1], \mathbb{R})$, c'est-à-dire :

$$H^1([0, 1], \mathbb{R}) := \left\{ f \in L^2([0, 1], \mathbb{R}) \mid \exists g \in L^2([0, 1], \mathbb{R}) \text{ t.q. } f(t) = f(0) + \int_0^t g(s) ds \right\}.$$

Il est important de rappeler que l'application g est unique (modulo le fait que l'on identifie deux fonctions intégrables qui coïncident p.p. = presque partout = sauf sur un ensemble négligeable). Ainsi défini, nous avons

$$\mathcal{C}^1([0, 1], \mathbb{R}) \subset H^1([0, 1], \mathbb{R}) \subset \mathcal{C}^0([0, 1], \mathbb{R}),$$

autrement dit, toute fonction de H^1 (sous-entendu $H^1([0, 1], \mathbb{R})$) est continue mais possède moins de régularité que les fonctions de classe \mathcal{C}^1 . En effet, les fonctions de H^1 ne possèdent pas de dérivée continue puisqu'elles ne sont même pas dérivables en tout point $t \in [0, 1]$. Les fonctions de H^1 sont dérivables presque partout sur $[0, 1]$ et on a :

$$\forall f \in H^1 : f'(t) = g(t) \text{ p.p. sur } [0, 1].$$

On fera donc clairement l'abus de notation $f' = g$ et puisque l'on dérive par rapport à la variable t qui correspond au "temps" pour nous, on écrira plutôt

$$\dot{f} = g.$$

Remarque 3.3.2. Nous pouvons voir g comme une dérivée en un sens plus faible que d'ordinaire. Cependant, si g est de classe \mathcal{C}^0 , alors la notion de dérivée faible coïncide avec celle de dérivée classique et f est alors de classe \mathcal{C}^1 . Ainsi, nous pouvons voir les fonctions de $H^1 = W^{1,2}$ comme des primitives de fonctions L^2 , de la même manière que nous pouvons voir les fonctions de \mathcal{C}^1 comme des primitives de fonctions \mathcal{C}^0 , et nous avons pour rappel :

$$\mathcal{C}^1([0, 1], \mathbb{R}) \subset \mathcal{C}^0([0, 1], \mathbb{R}) \quad \text{et} \quad H^1([0, 1], \mathbb{R}) \subset L^2([0, 1], \mathbb{R}).$$

Il nous faut maintenant une norme sur H^1 . Il existe deux possibilités équivalentes suivant si l'on considère la norme usuelle de $W^{1,2}$:

$$\|f\|_{W^{1,2}} := \|f\|_{L^2} + \|\dot{f}\|_{L^2},$$

ou si l'on considère la norme

$$\|f\|_{H^1} := (\|f\|_{L^2}^2 + \|\dot{f}\|_{L^2}^2)^{\frac{1}{2}}$$

induite par le produit scalaire usuel sur H^1 :

$$(f_1 | f_2)_{H^1} := (f_1 | f_2)_{L^2} + (\dot{f}_1 | \dot{f}_2)_{L^2},$$

où $(\cdot | \cdot)_{L^2}$ est le produit scalaire usuel sur L^2 . Les deux normes ainsi définies sont équivalentes et on peut noter que l'espace H^1 muni du produit scalaire $(\cdot | \cdot)_{H^1}$ forme un espace de Hilbert.

2. On rappelle qu'en réalité, $W^{1,2}([0, 1], \mathbb{R})$ est un ensemble de classes de fonctions et non de fonctions. Nous ne donnons pas ici sa définition, mais plutôt une propriété. En effet, d'après [3, Théorème VIII.2], il existe un unique représentant continu pour toute classe dans $W^{1,2}([0, 1], \mathbb{R})$ que nous appelons ici f .

Exemple 3.3.2 (suite). Revenons à l'exemple précédent. Nous avons donc

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J''(\bar{x}) \cdot h^2 = \|\dot{h}\|_{L^2}^2.$$

Rappelons que le problème classique du calcul des variations est un problème d'optimisation avec des contraintes affines et nous avons vu dans la Section 3.2 comment définir un problème équivalent sans contraintes. Appliquons les notations du cadre général de la Section 3.2 à cet exemple. Nous nous intéressons donc au problème :

$$\inf_x \{J(x) \mid x \in \mathcal{K}\},$$

avec $\mathcal{K} = \Phi^{-1}(\{0_{\mathbb{R}^2}\}) \subset E$, $E = \mathcal{C}^1([0, 1], \mathbb{R})$ et où $\Phi(x) = T(x) - y$ définit les contraintes affines. On a dans cet exemple $T(x) = (x(0), x(1))$ et $y = (0, -\pi/2)$. Pour obtenir de nouvelles conditions suffisantes d'optimalité locale, nous venons d'introduire ci-avant l'espace $H := H^1([0, 1], \mathbb{R})$ de telle sorte que $E \subset H$. Dans l'optique de se ramener à un problème sans contraintes, nous avons introduit

$$E_0 = E \cap \text{Ker } T = \mathcal{C}_0^1([0, 1], \mathbb{R}).$$

Toujours dans cette optique, nous introduisons :

$$H_0 := H \cap \text{Ker } T = H_0^1([0, 1], \mathbb{R}) := \{h \in H^1([0, 1], \mathbb{R}) \mid h(a) = h(b) = 0\},$$

de telle sorte que

$$E_0 \subset H_0.$$

L'idée générale est d'introduire un espace vectoriel normé plus grand que l'espace de travail. Ici, en raison des contraintes affines, l'espace de travail est E_0 et nous avons introduit H_0 (nous utilisons cette notation pour mettre un cadre général).

Remarque 3.3.3. Pour être un peu plus précis, l'application linéaire continue³ T appartient à $\mathcal{L}(H, \mathbb{R}^2)$ au lieu de $\mathcal{L}(E, \mathbb{R}^2)$ comme dans la Section 3.2. Nous devons tenir compte ici du rôle du nouvel espace H .

L'espace H_0 est muni de la norme induite par celle de H , *i.e.* $\|h\|_{H_0} = \|h\|_H$ et d'après le lemme 3.3.1 (ci-après), on a :

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E_0 = \mathcal{C}_0^1([0, 1], \mathbb{R}) \subset H_0 : J''(\bar{x}) \cdot h^2 = \|\dot{h}\|_{L^2}^2 \geq \alpha \|h\|_{H_0}^2. \quad (3.3)$$

Ceci est l'objectif voulu. Car, d'après la Proposition 3.3.1 (ci-après), on peut conclure que \bar{x} est un minimum local (faible) strict, c'est-à-dire :

$$\exists \eta > 0 \quad \text{t.q.} \quad \forall x \in B_E(\bar{x}, \eta) \cap \mathcal{K}, x \neq \bar{x} : J(\bar{x}) < J(x),$$

où $B_E(\bar{x}, \eta) := \{x \in E \mid \|x - \bar{x}\|_E < \eta\}$. □

3. L'application T est bien continue car il existe $C > 0$ t.q. $\|T(x)\|_{\mathbb{R}^2} \leq \|x\|_\infty \leq C\|x\|_{H^1}$ puisque H_1 s'injecte de manière continue dans \mathcal{C}^0 , cf. [3]. On peut noter que puisque T est continue, alors H_0^1 est un fermé de l'espace de Hilbert H^1 , donc lui-même un Hilbert, si on le munit du produit scalaire de H^1 .

Remarque 3.3.4. C'est cette idée d'utiliser deux normes, $\|\cdot\|_E$ pour la différentiabilité et $\|\cdot\|_H$ pour la **condition de croissance** $J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_H^2$ (qui remplace la condition d'ellipticité), que l'on nomme "two-norm discrepancy".

Lemme 3.3.1. Soit $a < b$ dans \mathbb{R} . Alors,

$$\exists c > 0 \quad \text{t.q.} \quad \forall h \in H_0^1([a, b], \mathbb{R}) : \|h\|_{H_0^1} \leq c \|\dot{h}\|_{L^2}.$$

► Soit $h \in H_0^1([a, b], \mathbb{R})$. Alors, puisque $h(a) = 0$ et d'après l'inégalité de Cauchy-Schwarz [3] :

$$h^2(t) = \left(\int_a^t \dot{h}(s) ds \right)^2 \leq (t-a) \int_a^t \dot{h}^2(s) ds \leq (b-a) \int_a^b \dot{h}^2(s) ds = (b-a) \|\dot{h}\|_{L^2}^2.$$

Ainsi,

$$\|h\|_{L^2}^2 = \int_a^b h^2(t) dt \leq (b-a)^2 \|\dot{h}\|_{L^2}^2,$$

donc

$$\|h\|_{H_0^1}^2 = \|h\|_{H^1}^2 = \|h\|_{L^2}^2 + \|\dot{h}\|_{L^2}^2 \leq (1 + (b-a)^2) \|\dot{h}\|_{L^2}^2$$

et la conclusion suit avec $c := \sqrt{1 + (b-a)^2} > 0$. ■

Proposition 3.3.1 – Condition suffisante d'ordre 2 généralisée (CS2G)

Soient $(E, \|\cdot\|_E)$ et $(H, \|\cdot\|_H)$ deux espaces vectoriels normés t.q. $E \subsetneq H$, i.e. $E \subset H$ et $\exists c > 0$ t.q. $\forall h \in E : \|h\|_H \leq c \|h\|_E$. Soit $J : U \subset E \rightarrow \mathbb{R}$ une application 2 fois dérivable sur U un ouvert de E et soit $\bar{x} \in U$ un point critique de J . Si

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E : J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_H^2, \quad (3.4)$$

alors \bar{x} est un minimum local strict (sous-entendu pour la norme $\|\cdot\|_E$).

► La formule de Taylor appliquée à J nous donne pour $h \in E$ t.q. $\bar{x} + h \in U$:

$$J(\bar{x} + h) - J(\bar{x}) = \frac{1}{2!} J''(\bar{x}) \cdot h^2 + \|h\|_E^2 \varepsilon(h) \quad \text{avec} \quad \lim_{\|h\|_E \rightarrow 0} |\varepsilon(h)| = 0.$$

Par hypothèse :

$$J(\bar{x} + h) - J(\bar{x}) \geq \frac{\alpha}{2} \|h\|_H^2 + \varepsilon(h) \|h\|_E^2 \geq \left(\frac{\alpha}{2} + \frac{\varepsilon(h)}{c} \right) \|h\|_E^2$$

et donc il existe $\eta > 0$ tel que pour tout $h \neq 0_E$ t.q. $\|h\|_E < \eta$ on ait $\bar{x} + h \in U$ et

$$J(\bar{x} + h) - J(\bar{x}) > 0.$$

Autrement dit, \bar{x} est un minimum local strict (sous-entendu pour la norme $\|\cdot\|_E$). ■

Remarque 3.3.5. Dans la proposition précédente, il suffit de faire l'hypothèse $E \subset H$. L'hypothèse d'injection continue n'est pas nécessaire car elle est automatiquement vérifiée

dès lors que l'on a l'équation (3.4). En effet, $J''(\bar{x})$ est continue car J est deux fois dérivable en \bar{x} et donc, il existe une constante $K > 0$ telle que

$$\forall h \in E : |J''(\bar{x}) \cdot h^2| \leq K \|h\|_E^2$$

et finalement puisque (3.4) est vérifiée, on obtient automatiquement :

$$\forall h \in E : \alpha \|h\|_H^2 \leq J''(\bar{x}) \cdot h^2 \leq K \|h\|_E^2, \quad (3.5)$$

et la constante $c := \sqrt{K/\alpha}$ fait l'affaire pour l'hypothèse de la proposition précédente.

Remarque 3.3.6. Dans le cas particulier où $E = H$ et $\|\cdot\|_E = \|\cdot\|_H$, on retrouve la condition suffisante d'ordre deux (CS2) classique, cf. conditions (3.1). Introduisons dans ce cas la quantité $\|h\| := \sqrt{J''(\bar{x}) \cdot h^2}$. Il est alors intéressant de noter (ce que l'on aurait pu mentionner avant) que $\|\cdot\|$ définit une norme équivalente à la norme $\|\cdot\|_E$, cf. l'équation (3.5), mais que cette nouvelle norme est induite par le produit scalaire

$$(h | k) := J''(\bar{x}) \cdot (h, k).$$

L'ensemble E munit du produit scalaire $(\cdot | \cdot)$ est donc un espace préhilbertien (un espace de Hilbert si E est un espace vectoriel normé complet, *i.e.* si E est un espace de Banach), on dit alors que $(E, \|\cdot\|_E)$ est un espace vectoriel normé **homéomorphe** à un espace préhilbertien. On arrive alors à la conclusion remarquable que si E n'est pas homéomorphe à un espace préhilbertien, alors la condition d'ellipticité ne peut jamais être vérifiée, quel que soit le point critique ! Et donc dans ce cas, il est nécessaire d'utiliser une nouvelle norme pour obtenir la condition de croissance.

Remarque 3.3.7. La Proposition 3.3.1 donne une nouvelle condition suffisante d'ordre deux d'optimalité locale pour un minimum libre, *i.e.* dans le cadre d'un problème d'optimisation sans contraintes. Bien entendu, nous pouvons donner une version de cette proposition dans le cas de contraintes affines de manière similaire à ce qui a été présenté dans la Section 3.2 (voir les conditions (3.2)) :

$$(CS2G) \quad \exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E_0 : J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_{H_0}^2,$$

où $E_0 := E \cap \text{Ker } T$, $H_0 := H \cap \text{Ker } T$ et où $\bar{x} \in \mathcal{K}$. Cette version ayant longtemps été détaillée dans cette section, cf. notamment l'équation (3.3).

Corrections des exercices

Solution de l'exercice 1.5.1 page 11:

1. Soit $x \in \mathbb{R}^n$ et $v \in \mathbb{R}^n$. On a

$$f(x+v) - f(x) = A(x+v) - Ax = Av + o(v),$$

donc f est différentiable en x et $f'(x) \cdot v = Av$.

2. On a $f'(x) = A$ et f' est constante donc continue. Donc f est de classe \mathcal{C}^1 sur \mathbb{R}^n .

Solution de l'exercice 1.5.2 page 11:

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

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

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 .

Solution de l'exercice 1.5.3 page 11:

Soit $x \in E$ et $v \in E$. On a $L(x+v) - L(x) = L(v) = Lv + 0_F = Lv + o(v)$, donc L est différentiable en x et $L'(x) \cdot v = Lv$, autrement dit $L'(x) = L$. L'application L' est constante donc continue et L est de classe \mathcal{C}^1 sur E .

Solution de l'exercice 1.5.4 page 11:

Soit $x \in E$ et $v \in E$. On a $A(x+v) - A(x) = L(x+v) + p - (L(x) + p) = L(v) = Lv + 0_F = Lv + o(v)$, donc A est différentiable en x et $A'(x) \cdot v = Lv$, autrement dit $A'(x) = L$. L'application A' est constante donc continue et A est de classe \mathcal{C}^1 sur E .

Solution de l'exercice 1.5.5 page 11:

Soit $(x, y) \in E \times F$ et $(v, w) \in E \times F$. On a

$$\begin{aligned} B(x+v, y+w) - B(x, y) &= B(x, y+w) + B(v, y) + B(v, w) - B(x, y) \\ &= B(x, y) + B(x, w) + B(v, y) + B(v, w) - B(x, y) \\ &= B(x, w) + B(v, y) + B(v, w) \\ &= B'(x, y) \cdot (v, w) + o((v, w)), \end{aligned}$$

cf. $\|B(v, w)\|_G \leq K\|v\|_E\|w\|_F$ et

$$K \frac{\|v\|_E\|w\|_F}{\|(v, w)\|_{E \times F}} \rightarrow 0 \quad \text{quand} \quad \|(v, w)\|_{E \times F} = \max(\|v\|_E, \|w\|_F) \rightarrow 0.$$

Ainsi définie, $B'(x, y)$ est linéaire et continue donc B est différentiable en (x, y) et $B'(x, y) \cdot (v, w) = B(v, y) + B(x, w)$.

L'application B' est la somme de deux applications que l'on va montrer être continues pour conclure que B est de classe \mathcal{C}^1 sur $E \times F$. On note

$$\begin{aligned} B_1: E \times F &\longrightarrow \mathcal{L}(E \times F, G) \\ (x, y) &\longmapsto B_1(x, y) := ((v, w) \mapsto B(v, y)), \end{aligned}$$

et

$$\begin{aligned} B_2: E \times F &\longrightarrow \mathcal{L}(E \times F, G) \\ (x, y) &\longmapsto B_2(x, y) := ((v, w) \mapsto B(x, w)). \end{aligned}$$

L'application B_1 est linéaire et continue car

$$\begin{aligned} \|B_1(x, y)\|_{\mathcal{L}(E \times F, G)} &= \sup_{\|(v, w)\|=1} \|B_1(x, y)(v, w)\|_G = \sup_{\|(v, w)\|=1} \|B(v, y)\|_G \\ &\leq K\|y\|_F \leq K\|(x, y)\|_{E \times F}, \end{aligned}$$

par la continuité de B et $\|y\|_F \leq \|(x, y)\|_{E \times F} = \max(\|x\|_E, \|y\|_F)$. On fait de même pour B_2 . Donc B_1 et B_2 sont continues et $B' = B_1 + B_2$ est continue.

Solution de l'exercice 1.11.1 page 29:

Les applications composantes de f admettent des dérivées partielles à tout ordre, continues en tout point de \mathbb{R}^3 donc f est \mathcal{C}^∞ sur \mathbb{R}^3 , en particulier elle est \mathcal{C}^1 . $\forall (x, y, z) \in \mathbb{R}^3$,

$$f'(x, y, z) = J_f(x, y, z) = \begin{bmatrix} -y \sin x - z \cos x & \cos x & -\sin x \\ y z & x z & y x \end{bmatrix}.$$

Solution de l'exercice 1.11.2 page 29:

- 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 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 1.11.1, 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

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

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

Solution de l'exercice 1.11.3 page 29:

On peut écrire

$$f(x) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c,$$

donc f admet des dérivées partielles continues à tout ordre sur \mathbb{R}^n donc f est \mathcal{C}^∞ sur \mathbb{R}^n , en particulier elle est 2 fois dérivable. $\forall (x, h) \in (\mathbb{R}^n)^2$,

$$\begin{aligned} f(x+h) &= \frac{1}{2} (A(x+h) \mid x+h) + (b \mid x+h) + c \\ &= \frac{1}{2} [(Ax \mid x) + (Ax \mid h) + (Ah \mid x) + (Ah \mid h)] + (b \mid x) + (b \mid h) + c \\ &= f(x) + \left(\frac{1}{2} (A + A^T) x + b \mid h \right) + \frac{1}{2} (Ah \mid h), \quad \text{et } \frac{1}{2} (Ah \mid h) = o(\|h\|). \end{aligned}$$

$$\Rightarrow \nabla f(x) = \frac{1}{2} (A + A^T) x + b \text{ donc } \nabla f(x) = Ax + b \text{ si } A \in \text{Sym}(n, \mathbb{R}).$$

$$\Rightarrow \nabla^2 f(x) = (\nabla f)'(x) = \frac{1}{2} (A + A^T) \text{ donc } \nabla^2 f(x) = A \text{ si } A \in \text{Sym}(n, \mathbb{R}).$$

Solution de l'exercice 1.11.4 page 29:

1. On note $f = Q \circ x$ avec $Q(x) = \frac{1}{2} (Ax \mid x) - (b \mid x)$. f est deux fois dérivable sur \mathbb{R} car $x(\cdot)$ l'est et Q est deux fois dérivable sur \mathbb{R}^n , cf. Exercice 1.11.
2. $\forall t \in \mathbb{R}$ on a

$$f'(t) = Q'(x(t)) \cdot x'(t) = (Ax(t) - b \mid x'(t)), \quad \text{cf. Exercice 1.11.}$$

On note $f'(t) = (B \circ \psi)(t)$ avec

$$\begin{aligned} \psi: \mathbb{R} &\longrightarrow \mathbb{R}^n \times \mathbb{R}^n \\ t &\longmapsto \psi(t) = ((A \circ x)(t) - b, x'(t)) \end{aligned} \quad \text{et} \quad \begin{aligned} B: \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (x, y) &\longmapsto B(x, y) = (x | y). \end{aligned}$$

Remarque 4.0.1. On peut voir la matrice A comme une application linéaire.

On a alors

$$\begin{aligned} f''(t) &= B'(\psi(t)) \cdot \psi'(t) \\ &= B(Ax(t) - b, x''(t)) + B(Ax'(t), x'(t)) \\ &= (Ax(t) - b | x''(t)) + (Ax'(t) | x'(t)). \end{aligned}$$

Solution de l'exercice 1.11.5 page 30:

- $\forall x \in \mathbb{R}^n$, $\frac{\partial \varphi}{\partial x_i}(x) = x_i$, $i = 1, \dots, n$ et $\frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) = \delta_{ij}$, $1 \leq i, j \leq n$, donc φ admet des dérivées partielles à l'ordre 1 et 2 continues sur \mathbb{R}^n donc φ est \mathcal{C}^2 sur \mathbb{R}^n .
- $\forall x \in \mathbb{R}^n$, $\forall h \in \mathbb{R}^n$,

$$\varphi(x+h) = \frac{1}{2} (x+h | x+h) = \frac{1}{2} (x | x) + (x | h) + \frac{1}{2} (h | h) = \varphi(x) + (x | h) + \varphi(h) \quad (4.1)$$

et $\varphi(h) = o(\|h\|)$. On retrouve tout d'abord le développement limité à l'ordre 1 avec $\varphi'(x) \cdot h = (x | h)$ et puisque $\varphi(h) = \frac{1}{2} \varphi''(x) \cdot (h, h)$, cf. point précédent, alors la formule (4.1) est le développement limité (exact) de φ en x à l'ordre 2.

Solution de l'exercice 1.11.6 page 30:

$f = \cos \circ \varphi$ est 2 fois dérivable sur \mathbb{R}^n car φ l'est et \cos l'est sur \mathbb{R} .

- $\forall x \in \mathbb{R}^n$, $\forall h \in \mathbb{R}^n$,

$$f'(x) \cdot h = \cos'(\varphi(x)) \cdot (\varphi'(x) \cdot h) = -\sin(\varphi(x)) \cdot (x | h) = (-\sin(\varphi(x)) x | h).$$

$$\Rightarrow \nabla f(x) = -\sin(\varphi(x)) x.$$

- On note $\nabla f = B \circ \psi$ où

$$\begin{aligned} \psi: \mathbb{R}^n &\longrightarrow \mathbb{R} \times \mathbb{R}^n \\ x &\longmapsto \psi(x) = (-\sin(\varphi(x)), x) \end{aligned} \quad \text{et} \quad \begin{aligned} B: \mathbb{R} \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (\alpha, y) &\longmapsto B(\alpha, y) = \alpha y. \end{aligned}$$

Puisque $\nabla^2 f(x) = (\nabla f)'(x)$, on a $\forall x \in \mathbb{R}^n$, $\forall h \in \mathbb{R}^n$,

$$\begin{aligned} \nabla^2 f(x) \cdot h &= (\nabla f)'(x) \cdot h = B'(\psi(x)) \cdot (\psi'(x) \cdot h) \\ &= B'(\psi(x)) \cdot (-\cos(\varphi(x)) \cdot (\varphi'(x) \cdot h), h) \\ &= B(-\sin(\varphi(x)), h) + B(-\cos(\varphi(x)) (x | h), x) \\ &= -(\sin(\varphi(x)) I_n + \cos(\varphi(x)) x x^T) h. \end{aligned}$$

$$\Rightarrow \nabla^2 f(x) = -(\sin(\varphi(x)) I_n + \cos(\varphi(x)) x x^T).$$

(Variante). On peut utiliser la relation $f''(x) \cdot (h, k) = g'(x) \cdot k$, où l'on définit $g(x) = f'(x) \cdot h$, pour x, h et k dans \mathbb{R}^n . Dans notre cas $g(x) = f'(x) \cdot h = -\sin(\varphi(x)) (x | h)$. La hessienne est alors donnée par

$$\begin{aligned} f''(x) \cdot (h, k) &= -(\cos(\varphi(x)) (x | k) (x | h) + \sin(\varphi(x)) (k | h)) \\ &= k^T \left[-(\cos(\varphi(x)) x x^T + \sin(\varphi(x)) I_n) \right] h \\ &= k^T \nabla^2 f(x) h \end{aligned}$$

Solution de l'exercice 1.11.7 page 30:

1. (Variante). On peut aussi noter $f = \sqrt{\cdot} \circ 2\varphi$. Alors puisque 2φ est 2 fois dérivable sur \mathbb{R}^n , que $\sqrt{\cdot}$ est lisse sur $\mathbb{R} \setminus \{0\}$ et $2\varphi(x) = 0 \Leftrightarrow x = 0_{\mathbb{R}^n}$, on peut conclure que f est 2 fois dérivable sur $\mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$.

On a pour $x \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ et $h \in \mathbb{R}^n$:

$$f'(x) \cdot h = (\sqrt{\cdot})'(2\varphi(x)) \cdot (2\varphi'(x) \cdot h) = 2 \frac{1}{2} \frac{1}{\sqrt{2\varphi(x)}} (x | h) = \left(\frac{x}{\|x\|} | h \right).$$

On peut écrire $f(x) = (\sum_{i=1}^n x_i^2)^{1/2}$. Pour tout $1 \leq i \leq n$,

$$\frac{\partial f}{\partial x_i}(x) = \frac{1}{2} 2 x_i \frac{1}{\|x\|} = \frac{x_i}{\|x\|},$$

donc les dérivées partielles de f existent et sont continues sur $\mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ donc f y est dérivable. $\Rightarrow \nabla f(x) = \frac{x}{\|x\|}$.

2. Pour tout $t \in \mathbb{R}$, $h \in \mathbb{R}^n$, on a $f(0_{\mathbb{R}^n} + th) - f(0_{\mathbb{R}^n}) = f(th) = |t| \|h\|$ donc

$$\lim_{t \rightarrow 0^+} \frac{f(0_{\mathbb{R}^n} + th) - f(0_{\mathbb{R}^n})}{t} \neq \lim_{t \rightarrow 0^-} \frac{f(0_{\mathbb{R}^n} + th) - f(0_{\mathbb{R}^n})}{t},$$

donc f n'est pas dérivable en $0_{\mathbb{R}^n}$.

3. $f'(x) = \frac{x^T}{\|x\|}$ est dérivable sur $\mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ en tant que composée entre $x \mapsto (x^T, \|x\|)$ dérivable sur $\mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ et $(a, b) \mapsto a/b$ dérivable si $b \neq 0$, et $b = \|x\| \neq 0 \Leftrightarrow x \neq 0_{\mathbb{R}^n}$.

On rappelle que pour tout $x \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, h et k dans \mathbb{R}^n , $f''(x) \cdot (h, k) = (f'(\cdot) \cdot h)'(x) \cdot k$ et $f'(x) \cdot h = \frac{(x | h)}{\|x\|}$. On a donc

$$f''(x) \cdot (h, k) = \frac{1}{\|x\|^2} \left((k | h) \|x\| - (x | h) \frac{(x | k)}{\|x\|} \right) = k^T \left[\frac{I_n}{\|x\|} - \frac{x x^T}{\|x\|^3} \right] h.$$

$$\Rightarrow \nabla^2 f(x) = \frac{I_n}{\|x\|} - \frac{x x^T}{\|x\|^3}.$$

(Variante). On peut aussi faire : $\nabla^2 f(x) \cdot h = (\nabla f)'(x) \cdot h$ avec $\nabla f(x) = \frac{x}{\|x\|}$.

Solution de l'exercice 1.11.8 page 30:

1. $f = \frac{1}{2} \|\cdot\|^2 \circ F : f$ est deux fois dérivable sur \mathbb{R}^n en tant que composée de fonctions deux fois dérivable.

On note $\varphi(y) = \frac{1}{2} \|y\|^2$ pour $y \in \mathbb{R}^p$. Alors, $f = \varphi \circ F$ et pour tout x, h dans \mathbb{R}^n on a

$$f'(x) \cdot h = \varphi'(F(x)) \cdot F'(x) \cdot h = (F(x) | F'(x) \cdot h) = (J_F(x)^T F(x) | h),$$

en identifiant la différentielle en x avec sa matrice jacobienne et en utilisant l'adjoint.

$$\Rightarrow \nabla f(x) = J_F(x)^T F(x).$$

Rappelons que si l'on note $F(x) = \sum_{i=1}^p F_i(x) e_i$, où (e_1, \dots, e_p) forme la base canonique de \mathbb{R}^p , alors $F''(x) \cdot (h, k) = \sum_{i=1}^p F_i''(x) \cdot (h, k) e_i \in \mathbb{R}^p$. Ainsi, pour tout x, h et k dans \mathbb{R}^n , on a

$$\begin{aligned} f''(x) \cdot (h, k) &= (f'(\cdot) \cdot h)'(x) \cdot k \\ &= (F'(x) \cdot k | F'(x) \cdot h) + (F(x) | F''(x) \cdot (h, k)) \\ &= k^T \left[J_F(x)^T J_F(x) + \sum_{i=1}^p F_i(x) \nabla^2 F_i(x) \right] h. \end{aligned}$$

$$\Rightarrow \nabla^2 f(x) = J_F(x)^T J_F(x) + \sum_{i=1}^p F_i(x) \nabla^2 F_i(x).$$

2.

$$\nabla f(x) = J_F(x)^T F(x) = A^T (Ax - b),$$

$$\nabla^2 f(x) = J_F(x)^T J_F(x) + \sum_{i=1}^p F_i(x) \nabla^2 F_i(x) = J_F(x)^T J_F(x) = A^T A.$$

Bibliographie

- [1] J. Appell & P. P. Zabrejko, *Nonlinear Superposition Operators*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 1990.
- [2] J. F. Bonnans & A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research and Financial Engineering, Springer-Verlag New York 2000, XVIII, 601. [↔ 47](#).
- [3] H. Brezis, *Analyse fonctionnelle, Théorie et applications*, Collection Mathématiques appliquées pour la maîtrise, Masson, 1987. [↔ 50](#), [51](#) et [52](#).
- [4] W. Greub, *Multilinear algebra*, Springer, New York (1967). [↔ 45](#).
- [5] J.-B. Hiriart-Urruty, *L'optimisation*, Que sais-je, Presses Universitaires de France, 1996. [↔ 47](#).
- [6] 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.
- [7] F. Paulin, *Géométrie différentielle élémentaire*, Notes de cours, 2006–2007.
- [8] F. Paulin, *Topologie, analyse et calcul différentiel*, Notes de cours, 2008–2009.
- [9] C. Wagschal, *Topologie et analyse fonctionnelle*, Hermann, 2012. [↔ 3](#), [4](#), [5](#), [35](#), [36](#), [37](#) et [39](#).

Index

Application

contractante, 24, 33, 40
de classe \mathcal{C}^1 , 9
de classe \mathcal{C}^k , 22
dérivée, 9
dérivée d'ordre k , 22
dérivée seconde, 19
linéaire, 3
linéaire continue, 2
lipschitzienne, 24

Cas fondamentaux, 29

Condition

d'ellipticité, 46, 52, 53
de croissance, 52, 53
nécessaire d'ordre deux, 46, 48
nécessaire d'ordre supérieur, 44
nécessaire du premier ordre, 43, 48
nécessaire et suffisante, 44
suffisante d'ordre deux, 46, 48
suffisante d'ordre deux généralisée, 52
suffisante d'ordre supérieur, 45

Contraintes

affines, 48

Dérivée directionnelle, 8

Dérivée partielle, 15

Difféomorphisme, \mathcal{C}^k -difféomorphisme, 37

Différentielle de Fréchet, 5

Équation

de Euler, 43

Espace

de Banach, 2, 35
métrique, 24
vectoriel normé, 1

Extrémum

libre, 43
lié, 48

Forme bilinéaire

définie positive, 46

elliptique, 46

semi-définie positive, 46

Gradient, 8, 16

Homéomorphisme, 3, 35

Isomorphisme, 3, 7, 19, 35

Matrice

hessienne, 21

jacobienne, 18

Minimum

global, 43

local, 43

local strict, 43

Notation de Landau, 1

Opérateur hessien, 20

Point critique, 43

Théorème

d'inversion locale, 40

de Banach, 5, 37

de dérivation des composées, 10

des accroissements finis 1, 23

des accroissements finis 2, 25

des fonctions implicites, 41

du point fixe, 33

Two-norm discrepancy, 52

Valeur critique, 43