# Topologie élémentaire et calcul différentiel (version préliminaire)

Estanislao Herscovich [1]

[1]Institut Fourier, Université Grenoble Alpes, France.

**Résumé**

Ces notes ont été conçues pour être utilisées pour un cours de topologie élémentaire et calcul différentiel basique au niveau L2-L3. Tous les résultats sont standards et se trouvent dans des références basiques sur les sujets (voir par exemple [1]). Par des raisons évidentes, le premier chapitre, sur la topologie élémentaire, est plus abstrait, tandis que le trois derniers chapitres, sur le calcul différentiel de fonctions de plusieurs variables, les équations différentielles ordinaires et les courbes paramétrées, sont plus concrets.

# Table de matières

# Prérequis et notation

Suivant la tradition française, on notera $\mathbb{N} = \{0, 1, 2, \dots\}$ l'ensemble des entiers non négatifs. On rappelle que $\mathbb{Z}$ est l'ensemble de nombres entiers. Pour $a, b \in \mathbb{Z}$ tels que $a < b$ on notera $[\![a, b]\!]$ l'ensemble $\{x \in \mathbb{Z} : a \leq x \leq b\}$. On rappelle que $\mathbb{R}$ (resp., $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$) dénotera l'ensemble de nombres réels (resp., non négatifs, strictement positifs). Étant donnés deux nombres réels $a < b$, $]\,a, b\,[$ (resp., $[a, b]$) est l'**intervalle ouvert** (resp., **fermé**) formé des $x \in \mathbb{R}$ tels que $a < x < b$ (resp., $a \leq x \leq b$). De façon plus générale, un **intervalle** est une partie $I \subseteq \mathbb{R}$ qui satisfait que, étant donné $a, b \in I$ avec $a < b$, alors $x \in I$ pour tout $a < x < b$. Étant donné un intervalle $I \subseteq \mathbb{R}$, on notera $I^+ = \sup(I) \in \mathbb{R} \sqcup \{+\infty\}$ et $I^- = \inf(I) \in \mathbb{R} \sqcup \{+\infty\}$. On dit que l'intervalle $I$ est **fini** si $I^+, I^- \in \mathbb{R}$. On dit sinon que l'intervalle $I$ est **infini**.

On définit la **fonction signe** $\operatorname{sgn} : \mathbb{R} \to \{-1, 0, 1\}$ via

$$\operatorname{sgn}(x) = \begin{cases} 1, & \text{si } x \in \mathbb{R}_{>0}, \\ 0, & \text{si } x = 0, \\ -1, & \text{si } -x \in \mathbb{R}_{>0}. \end{cases}$$

On supposera que la définition d'**espace vectoriel** sur le corps $\mathbb{R}$ de nombres réels, aussi appelé **espace vectoriel réel**, est bien connue. On notera normalement un espace vectoriel réel avec les lettres $\mathbb{E}, \mathbb{F}, \dots$, et un élément d'un espace vectoriel réel par les lettres $v, w, \dots$. On va omettre normalement l'adjectif réel, puisque tous les espaces vectoriels vectoriels que l'on va considérer sont réels, sauf si l'on dit autrement. Le mot **vecteur** n'est qu'un synonyme de "élément d'un espace vectoriel". Le **vecteur nul** de $\mathbb{E}$, *i.e.* l'élément neutre pour l'addition de $\mathbb{E}$, sera noté $\mathbf{0}_{\mathbb{E}}$, ou sinon $\mathbf{0}$, si l'espace vectoriel $\mathbb{E}$ est clair. Pour simplifier, le lecteur/la lectrice pourra penser que $\mathbb{E}$ est l'espace vectoriel $\mathbb{R}^n$, avec $n \in \mathbb{Z}_{>0}$, où $\mathbb{Z}_{>0}$ est l'ensemble de nombres entiers (strictement) positifs. Dans ce cas, un élément générique $x$ de $\mathbb{R}^n$ sera écrit $x = (x_1, \dots, x_n)$, où $x_1, \dots, x_n \in \mathbb{R}$, et on rappelle que la somme et le produit (par un nombre réel) sont donnés par

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad \text{et} \quad \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n),$$

pour tous $\lambda, x_1, y_1, \dots, x_n, y_n \in \mathbb{R}$. On écrira souvent $\lambda(x_1, \dots, x_n)$ au lieu de $\lambda \cdot (x_1, \dots, x_n)$. L'élément neutre $\mathbf{0}_{\mathbb{R}^n}$ pour l'addition de $\mathbb{R}^n$ est le vecteur $(0, \dots, 0)$.

On rappelle que le **produit vectoriel** $\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ est l'application définie par

$$(a, b, c) \wedge (a', b', c') = (bc' - b'c, ca' - ac', ab' - a'b)$$

pour tous $a, a', b, b', c, c' \in \mathbb{R}$. Il est facile à vérifier que

$$v \wedge w = -w \wedge v \text{ et } \langle u \wedge v, w \rangle = \langle v \wedge w, u \rangle = \langle w \wedge u, v \rangle \tag{0.0.1}$$

pour tous $u, v, w \in \mathbb{R}^3$, où $\langle\,,\rangle$ désigne le produit scalaire usuel de $\mathbb{R}^3$, *i.e.* $\langle (a, b, c), (a', b', c') \rangle = aa' + bb' + cc'$ pour tous $a, a', b, b', c, c' \in \mathbb{R}$. En outre, on voit bien de la définition que

$$u \wedge (v \wedge w) = \langle u, w \rangle v - \langle u, v \rangle w \tag{0.0.2}$$

et

$$\|v \wedge w\|^2 = \|v\|^2 \|w\|^2 - \langle u, w \rangle^2, \tag{0.0.3}$$

pour tous $u, v, w \in \mathbb{R}^3$.

En outre, étant donné $n, m \in \mathbb{Z}_{>0}$, on dénote $\mathrm{M}_{n\times m}(\mathbb{R}) = \mathbb{R}^{nm}$ l'espace vectoriel de matrices avec $n$ lignes et $m$ colonnes à coefficients réels, où un élément $A \in \mathrm{M}_{n\times m}(\mathbb{R})$ est donné par $A = (A_{i,j})_{i\in[\![1,n]\!],j\in[\![1,m]\!]}$, avec $A_{i,j} \in \mathbb{R}$ pour tous $i \in [\![1,n]\!]$ et $j \in [\![1,m]\!]$. Si $n = m$ on écrira $\mathrm{M}_n(\mathbb{R})$ au lieu de $\mathrm{M}_{n\times n}(\mathbb{R})$. On rappelle que l'on identifie un vecteur $x \in \mathbb{R}^n$ avec un vecteur ligne $\mathrm{M}_{1\times n}(\mathbb{R})$. Étant donné $A \in \mathrm{M}_{n\times m}(\mathbb{R})$, on dénote $A^t$ la matrice **transposée**, *i.e.* $A^t = (A^t_{i,j})_{i\in[\![1,n]\!],j\in[\![1,m]\!]} \in \mathrm{M}_{m\times n}(\mathbb{R})$ satisfait que $A^t_{i,j} = A_{j,i}$ pour tous $i \in [\![1,n]\!]$ et $j \in [\![1,m]\!]$. On rappelle qu'une matrice $A \in \mathrm{M}_n(\mathbb{R})$ est **diagonale** si $A_{i,j} = 0$ pour tous $i, j \in [\![1,n]\!]$ différents.

Si $X$ et $Y$ sont deux ensembles, on notera $M(X,Y)$ ou $Y^X$ l'ensemble formé de toutes les applications $f : X \to Y$. L'application identité de $X$ sera notée $\mathrm{id}_X$. Étant donné un ensemble $X$, on rappelle qu'une **propriété (sur les éléments)** de $X$ est seulement une partie $Y \subseteq X$. On notera $\mathscr{P}(X)$ l'ensemble formé de toutes les parties de $X$.

On va utiliser les inégalités suivantes, qui sont laissées comme exercice au lecteur/à la lectrice.

**Fait 0.0.1.** *Si $(x_i)_{i\in[\![1,n]\!]}, (y_i)_{i\in[\![1,n]\!]} \in \mathbb{R}^n_{\geq}$ sont deux $n$-uplets et $\mu \geq 0$, alors*

$$\max_{i\in[\![1,n]\!]}(\mu x_i) = \mu \max_{i\in[\![1,n]\!]}(x_i) \text{ et } \max_{i\in[\![1,n]\!]}(x_i + y_i) \leq \max_{i\in[\![1,n]\!]}(x_i) + \max_{i\in[\![1,n]\!]}(y_i). \tag{0.0.4}$$

*Plus généralement, soit $S$ un ensemble, $f, g : S \to \mathbb{R}_{\geq 0}$ deux applications et $\mu \geq 0$, alors*

$$\sup\{\mu f(x) : x \in S\} = \mu \sup\{f(x) : x \in S\} \tag{0.0.5}$$

*et*

$$\sup\{f(x) + g(x) : x \in S\} \leq \sup\{f(x) : x \in S\} + \sup\{g(x) : x \in S\}. \tag{0.0.6}$$

# Chapitre 1

# Espaces vectoriels normés

Dans ce chapitre on étudiera les espaces vectoriels munis des normes. Une norme est essentiellement une forme de mesurer des distances entre les éléments de l'espace vectoriel, ce qui induit une notion d'ensemble ouvert et fermé. On introduira une notion d'équivalence entre normes et on se consacrera après à étudier plusieurs propriétés qui sont préservés par des normes équivalentes. Le résultat le plus important que l'on démontrera c'est que toutes les normes sur un espace vectoriel de dimension finie sont équivalentes, ce qui implique que la topologie d'un espace de dimension finie est canonique.

## 1.1 La définition de norme et quelques propriétés

On commence avec la définition centrale de la notre cours.

**Définition 1.1.1.** *Une **norme** sur un espace vectoriel $\mathbb{E}$ est une application $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$ qui satisfait les propriétés suivantes:*

*(N1) pour tout $v \in \mathbb{E}$, $N(v) = 0$ implique que $v = \mathbf{0}_{\mathbb{E}}$,*

*(N2) pour tous $v \in \mathbb{E}$ et $\lambda \in \mathbb{R}$, on a $N(\lambda \cdot v) = |\lambda| N(v)$,*

*(N3) pour tous $v, w \in \mathbb{E}$, $N(v + w) \leq N(v) + N(w)$.*

*On dira aussi que $(\mathbb{E}, N)$ est un **espace (vectoriel) normé**. La dernière propriété (N3) est appelé l'**inégalité triangulaire** pour la norme $N$.*

**Notation 1.1.2.** *Une autre terminologie typique pour dénoter une norme est du type $\| \ \| : \mathbb{E} \to \mathbb{R}_{\geq 0}$. Dans ce cas, l'évaluation de la norme $\| \ \|$ en un vecteur $v \in \mathbb{E}$ sera notée $\|v\|$.*

**Exemple 1.1.3.** *On considère $\mathbb{E} = \mathbb{R}$ et $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$ donnée par la valeur absolue, i.e. $N(x) = |x|$ pour tout $x \in \mathbb{R}$. Alors $N$ est une norme sur $\mathbb{R}$.*

**Proposition 1.1.4.** *Si $N$ est une norme sur un espace vectoriel $\mathbb{E}$, alors $N(\mathbf{0}_{\mathbb{E}}) = 0$ et $N(-v) = N(v)$, pour tout $v \in \mathbb{E}$.*

***Preuve.*** Comme $\mathbf{0}_{\mathbb{E}} = 0.\mathbf{0}_{\mathbb{E}}$, alors, d'après (N2), $N(\mathbf{0}_{\mathbb{E}}) = |0| N(\mathbf{0}_{\mathbb{E}})$, ce qui implique que $N(\mathbf{0}_{\mathbb{E}}) = 0$. Pour la deuxième partie, il faut remarquer que, d'après (N2), $N(-v) = N((-1) \cdot v) = |-1| N(v) = N(v)$, pour tout $v \in \mathbb{E}$. □

**Définition 1.1.5.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé. On définit la **distance** $d(v, w)$ entre deux éléments $v$ et $w$ dans $\mathbb{E}$ à partir de $d(v, w) = N(v - w)$. C'est clair que $d(v, w) = d(w, v) \geq 0$, pour tous $v, w \in \mathbb{E}$. En outre, d'après (N1), $d(v, w) = 0$ si et seulement si $v = w$.*

*Pour $v \in \mathbb{E}$ et $r \in \mathbb{R}_{>0}$, on définit la **boule ouverte** et la **boule fermée***

$$B_N(v, r) = \{w \in \mathbb{E} : N(w - v) < r\} \text{ et } \bar{B}_N(v, r) = \{w \in \mathbb{E} : N(w - v) \leq r\}$$

*centrées autour de $v$ et de rayon $r$, respectivement. La boule $B_N(\mathbf{0}_{\mathbb{E}}, 1)$ (resp., $\bar{B}_N(\mathbf{0}_{\mathbb{E}}, 1)$) est appelée la **boule unité ouverte** (resp., **fermée**) de $(\mathbb{E}, N)$.*

On rappelle qu'une partie $S \subseteq \mathbb{E}$ d'un espace vectoriel $\mathbb{E}$ est **convexe** si, pour tous $v, w \in S$ et pour tout $t \in [0,1]$, $t \cdot v + (1-t) \cdot w \in S$. Le résultat suivant dit qu'une boule ouverte ou fermée d'un espace vectoriel normé $(\mathbb{E}, N)$ doit être toujours convexe. En particulier, une boule dans un espace vectoriel normé ne peut avoir n'importe quelle forme.

**Proposition 1.1.6.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé. Pour tout $v \in \mathbb{E}$ et $r \in \mathbb{R}_{>0}$, les boules $B_N(v,r)$ et $\bar{B}_N(v,r)$ sont des ensembles convexes de $\mathbb{E}$.*

***Preuve.*** Soient $w$ et $w'$ deux éléments de $B_N(v,r)$ (resp., $\bar{B}_N(v,r)$). Alors, $N(w-v) < r$ et $N(w'-v) < r$ (resp., $N(w-v) \leq r$ et $N(w'-v) \leq r$). Pour tout $t \in [0,1]$, on voit bien que

$$\begin{aligned} N\big(t \cdot w + (1-t) \cdot w' - v\big) &= N\Big(t \cdot w + (1-t) \cdot w' - \big(t \cdot v + (1-t) \cdot v\big)\Big) \\ &= N\big(t \cdot (w-v) + (1-t) \cdot (w'-v)\big) \leq N\big(t \cdot (w-v)\big) + N\big((1-t) \cdot (w'-v)\big) \\ &= tN(w-v) + (1-t)N(w'-v). \end{aligned}$$

Comme $tN(w-v)+(1-t)N(w'-v) < tr+(1-t)r = r$ (resp., $tN(w-v)+(1-t)N(w'-v) \leq tr+(1-t)r = r$), on conclut que $t \cdot w + (1-t) \cdot w' \in B_N(v,r)$ (resp., $t \cdot w + (1-t) \cdot w' \in \bar{B}_N(v,r)$). En conséquence, $B_N(v,r)$ et $\bar{B}_N(v,r)$ sont des ensembles convexes de $\mathbb{E}$. □

## 1.2 Construction de normes à partir d'autres normes

On rappelle qu'une **application linéaire** $T : \mathbb{F} \to \mathbb{E}$ d'un espace vectoriel $\mathbb{F}$ dans un espace vectoriel $\mathbb{E}$ est une application qui satisfait que $T(v + \lambda w) = T(v) + \lambda T(w)$, pour tout $v, w \in \mathbb{F}$ et $\lambda \in \mathbb{R}$.

**Proposition 1.2.1.** *Soient $\mathbb{E}$ et $\mathbb{F}$ deux espaces vectoriels et $T : \mathbb{F} \to \mathbb{E}$ une application linéaire injective. On suppose que $\mathbb{E}$ est muni d'une norme $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$. On définit l'application $N_T : \mathbb{F} \to \mathbb{R}_{\geq 0}$ donnée par $N_T = N \circ T$. Alors, $N_T$ est une norme sur $\mathbb{F}$.*

***Preuve.*** C'est clair que $N_T$ satisfait (N2) et (N3), vu que

$$N_T(\lambda \cdot v) = N\big(T(\lambda \cdot v)\big) = N\big(\lambda \cdot T(v)\big) = |\lambda| N\big(T(v)\big) = |\lambda| N_T(v),$$

et

$$N_T(v+w) = N\big(T(v+w)\big) = N\big(T(v) + T(w)\big) \leq N\big(T(v)\big) + N\big(T(v)\big) = N_T(v) + N_T(w),$$

pour tous $v, w \in \mathbb{F}$ et $\lambda \in \mathbb{R}$, où l'on a utilisé que $N$ satisfait (N2) et (N3).

On va finalement montrer que $N_T$ satisfait (N1). Soit $v \in \mathbb{F}$ tel que $N_T(v) = N(T(v)) = 0$. Comme $N$ satisfait (N1), on voit que $T(v) = \mathbf{0}_{\mathbb{E}}$. Or, $T$ étant injectif, $v = \mathbf{0}_{\mathbb{F}}$, comme on voulait démontrer. □

On rappelle qu'un **sous-espace vectoriel** d'un espace vectoriel $\mathbb{E}$ est une partie $\mathbb{F} \subseteq \mathbb{E}$ qui satisfait que $\mathbf{0}_{\mathbb{E}} \in \mathbb{F}$ et que $v + \lambda w \in \mathbb{F}$ pour tous les éléments $v, w \in \mathbb{F}$ et pour tout $\lambda \in \mathbb{R}$. Dans ce cas, $\mathbb{F}$ est un espace-vectoriel pour la somme et le produit par scalaires induits de ceux de $\mathbb{E}$. Le vecteur nul de $\mathbb{F}$ est $\mathbf{0}_{\mathbb{E}} \in \mathbb{F}$. On commence avec le résultat suivant, dont la preuve est élémentaire.

**Proposition 1.2.2.** *Soit $\mathbb{E}$ un espace vectoriel réel muni d'une norme $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$ et soit $\mathbb{F} \subseteq \mathbb{E}$ un sous-espace vectoriel. Alors, la restriction $N|_{\mathbb{F}} : \mathbb{F} \to \mathbb{R}_{\geq 0}$ de $N$, donnée par $N|_{\mathbb{F}}(v) = N(v)$ pour tout $v \in \mathbb{F}$, est une norme sur $\mathbb{F}$.*

***Preuve.*** Le résultat est une conséquence directe de la Proposition 1.2.1, en employant $T : \mathbb{F} \to \mathbb{E}$ donné par l'inclusion de $\mathbb{F}$ dans $\mathbb{E}$. En effet, le fait que $\mathbb{F} \subseteq \mathbb{E}$ soit un sous-espace vectoriel implique que $T$ est une application linéaire injective. □

Soit $\mathbb{E}$ un espace vectoriel réel. On rappelle qu'un ensemble $S$ est **cône** (au sens de l'analyse convexe) si pour tout $v \in S$ et $t \in \mathbb{R}_{>0}$, $t \cdot v \in S$. Un **cône convexe** dans $\mathbb{E}$ est un cône qui satisfait la propriété de convexité rappelée dans la section précédente. On remarque $S \subseteq \mathbb{E}$ est un cône convexe si et seulement si, pour tous $v, v' \in S$ et $t, t' \in \mathbb{R}_{\geq 0}$ tels que $t + t' > 0$, on a $t \cdot v + t' \cdot v' \in S$. On laisse la vérification de ce fait au lecteur/à la lectrice.

On rappelle que $M(\mathbb{E}, \mathbb{R}_{\geq 0}) = \mathbb{R}_{\geq 0}^{\mathbb{E}}$ est l'ensemble formé de toutes les applications de $\mathbb{E}$ dans $\mathbb{R}_{\geq 0}$. Soit $\operatorname{Norm}(\mathbb{E})$ la partie de $M(\mathbb{E}, \mathbb{R}_{\geq 0})$ formée de toutes les normes sur $\mathbb{E}$. Le résultat suivant nous dit que $\operatorname{Norm}(\mathbb{E})$ est un cône convexe de l'espace vectoriel $M(\mathbb{E}, \mathbb{R}_{\geq 0})$.

**Proposition 1.2.3.** *Soient $N$ et $N'$ deux normes sur un espace vectoriel $\mathbb{E}$ et $\mu, \mu' \in \mathbb{R}_{\geq 0}$ tels que $\mu + \mu' > 0$. L'application $N'' : \mathbb{E} \to \mathbb{R}_{\geq 0}$ donnée par $N''(v) = \mu N(v) + \mu' N'(v)$, pour tout $v \in \mathbb{E}$, est une norme aussi.*

***Preuve.*** Si $N''(v) = \mu N(v) + \mu' N'(v) = 0$ pour $v \in \mathbb{E}$, alors $N(v) = 0$ ou $N'(v) = 0$, ce qui implique que $v = \mathbf{0}_{\mathbb{E}}$, d'après (N1) pour $N$ et $N'$. Cela nous dit que $N''$ vérifié (N1). Par ailleurs, (N2) pour $N$ et $N'$ nous dit que

$$N''(\lambda \cdot v) = \mu N(\lambda \cdot v) + \mu' N'(\lambda \cdot v) = \mu|\lambda|N(v) + \mu'|\lambda|N'(v) = |\lambda|N''(v),$$

pour tous $v \in \mathbb{E}$ et $\lambda \in \mathbb{R}$, ce qui dit que $N''$ vérifié (N2). Finalement, d'après (N3) pour $N$ et $N'$ on voit que

$$N''(v+w) = \mu N(v+w) + \mu' N'(v+w) \leq \mu\big(N(v) + N(w)\big) + \mu'\big(N'(v) + N'(w)\big) = N''(v) + N''(w),$$

pour tous $v, w \in \mathbb{E}$. On a démontré que $N''$ vérifié (N3) et, en conséquence, $N''$ est une norme sur $\mathbb{E}$. $\square$

Soient $\mathbb{E}_1, \dots, \mathbb{E}_n$ des espaces vectoriels. On rappelle que le **produit cartésien** $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$ (aussi noté $\prod_{i=1}^n \mathbb{E}_i$) est l'ensemble formé des $n$-uplets $(x_1, \dots, x_n)$, où $x_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$. Il s'agit d'un espace vectoriel dont la somme et le produit sont donnés par

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad \text{et} \quad \lambda \cdot (x_1, \dots, x_n) = (\lambda \cdot x_1, \dots, \lambda \cdot x_n),$$

pour tous $\lambda \in \mathbb{R}$ et $x_i, y_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$. Le vecteur nul $\mathbf{0}_{\mathbb{E}}$ de $\mathbb{E}$ est le vecteur $(\mathbf{0}_{\mathbb{E}_1}, \dots, \mathbf{0}_{\mathbb{E}_n})$. Si $\mathbb{E}_1, \dots, \mathbb{E}_n$ sont des espaces vectoriels munis des normes $N_1, \dots, N_n$, respectivement. Alors on définit les application $N_{\max}$ et $N_{\text{sum}}$ de $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$ dans $\mathbb{R}_{\geq 0}$ via

$$N_{\max}(x) = \max_{i \in [\![1,n]\!]} \big(N_i(x_i)\big) \tag{1.2.1}$$

et

$$N_{\text{sum}}(x) = \sum_{i=1}^{n} N_i(x_i), \tag{1.2.2}$$

pour tout $x = (x_1, \dots, x_n)$, où $x_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$. On a le résultat suivant.

**Proposition 1.2.4.** *Soient $\mathbb{E}_1, \dots, \mathbb{E}_n$ des espaces vectoriels munis des normes $N_1, \dots, N_n$, respectivement. Alors les applications $N_{\max}$ et $N_{\text{sum}}$ de $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$ dans $\mathbb{R}_{\geq 0}$ définies par* (1.2.1) *et* (1.2.2)*, respectivement, sont de normes sur $\mathbb{E}$.*

***Preuve.*** On montrera d'abord que $N_{\max}$ est une norme sur $\mathbb{E}$. On voit bien que, si $N_{\max}(x) = 0$ pour un élément $x = (x_1, \dots, x_n)$, où $x_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$, alors $N_i(x_i) = 0$, pour tout $i \in [\![1, n]\!]$. Comme $N_i$ est une norme sur $\mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$, (N1) nous dit que $x_i = \mathbf{0}_{\mathbb{E}_i}$ pour tout $i \in [\![1, n]\!]$, ce qui implique que $x = (\mathbf{0}_{\mathbb{E}_1}, \dots, \mathbf{0}_{\mathbb{E}_n}) = \mathbf{0}_{\mathbb{E}}$ est le vecteur nul de $\mathbb{E}$, *i.e.* $N_{\max}$ vérifie (N1). Par ailleurs,

$$\begin{aligned} N_{\max}\big(\lambda \cdot (x_1, \dots, x_n)\big) &= N_{\max}(\lambda \cdot x_1, \dots, \lambda \cdot x_n) = \max_{i \in [\![1,n]\!]} \big(N_i(\lambda \cdot x_i)\big) = \max_{i \in [\![1,n]\!]} \big(|\lambda| N_i(x_i)\big) \\ &= |\lambda| \max_{i \in [\![1,n]\!]} \big(N_i(x_i)\big) = |\lambda| N_{\max}(x_1, \dots, x_n), \end{aligned}$$

pour tous $\lambda \in \mathbb{R}$ et $x_i \in \mathbb{E}_i$ avec $i \in [\![1, n]\!]$, où l'on a utilisé (N2) pour $N_i$ dans la troisième égalité, et la première identité de (0.0.4) dans la quatrième égalité. Cela implique que $N_{\max}$ vérifie (N2). Finalement, on voit bien que

$$\begin{aligned} N_{\max}\big((x_1, \dots, x_n) + (y_1, \dots, y_n)\big) &= N_{\max}(x_1 + y_1, \dots, x_n + y_n) = \max_{i \in [\![1,n]\!]} \big(N_i(x_i + y_i)\big) \\ &\leq \max_{i \in [\![1,n]\!]} \big(N_i(x_i) + N_i(y_i)\big) \leq \max_{i \in [\![1,n]\!]} \big(N_i(x_i)\big) + \max_{i \in [\![1,n]\!]} \big(N_i(y_i)\big) \\ &= N_{\max}(x_1, \dots, x_n) + N_{\max}(y_1, \dots, y_n), \end{aligned}$$

pour tous $x_i, y_i \in \mathbb{E}_i$ avec $i \in [\![1, n]\!]$, où l'on a appliqué (N3) pour $N_i$ au troisième membre, et la deuxième identité de (0.0.4) au quatrième membre. Cela nous dit que $N_{\max}$ vérifie (N3).

On va montrer finalement que $N_{\text{sum}}$ est une norme sur $\mathbb{E}$. On remarque d'abord que, si $N_{\text{sum}}(x) = \sum_{i=1}^{n} N_i(x_i) = 0$ pour un élément $x = (x_1, \dots, x_n)$, où $x_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$, alors $N_i(x_i) = 0$, pour tout $i \in [\![1, n]\!]$, puisque une comme de termes non négatifs est zéro si et seulement si chaque opérande est zéro. Comme $N_i$ est une norme sur $\mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$, (N1) nous dit que $x_i = \mathbf{0}_{\mathbb{E}_i}$ pour tout $i \in [\![1, n]\!]$, ce qui implique que $x = (\mathbf{0}_{\mathbb{E}_1}, \dots, \mathbf{0}_{\mathbb{E}_n}) = \mathbf{0}_{\mathbb{E}}$ est le vecteur nul de $\mathbb{E}$, *i.e.* $N_{\max}$ vérifie (N1). Par ailleurs,

$$N_{\text{sum}}\big(\lambda \cdot (x_1, \dots, x_n)\big) = N_{\text{sum}}(\lambda \cdot x_1, \dots, \lambda \cdot x_n) = \sum_{i=1}^{n} N_i(\lambda \cdot x_i) = \sum_{i=1}^{n} |\lambda| N_i(x_i) =$$
$$= |\lambda| \sum_{i=1}^{n} N_i(x_i) = |\lambda| N_{\text{sum}}(x_1, \dots, x_n),$$

pour tous $\lambda \in \mathbb{R}$ et $x_i \in \mathbb{E}_i$ avec $i \in [\![1, n]\!]$, où l'on a utilisé (N2) pour $N_i$ dans le troisième égalité. Cela implique que $N_{\text{sum}}$ vérifie (N2). Finalement, on voit bien que

$$\begin{aligned} N_{\text{sum}}\big((x_1, \dots, x_n) + (y_1, \dots, y_n)\big) &= N_{\text{sum}}(x_1 + y_1, \dots, x_n + y_n) = \sum_{i=1}^{n} N_i(x_i + y_i) \\ &\leq \sum_{i=1}^{n} \big(N_i(x_i) + N_i(y_i)\big) = \Big(\sum_{i=1}^{n} N_i(x_i)\Big) + \Big(\sum_{i=1}^{n} N_i(y_i)\Big) \\ &= N_{\text{sum}}(x_1, \dots, x_n) + N_{\text{sum}}(y_1, \dots, y_n), \end{aligned}$$

pour tous $x_i, y_i \in \mathbb{E}_i$ avec $i \in [\![1, n]\!]$, où l'on a appliqué (N3) pour $N_i$ au troisième membre. Cela nous dit que $N_{\text{sum}}$ vérifie (N3). □

## 1.3 Construction d'une norme à partir d'un produit scalaire

On rappelle qu'un **produit scalaire** sur un espace vectoriel $\mathbb{E}$ est une application $\langle\, , \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ qui satisfait

(PS1) $\langle v, v \rangle \in \mathbb{R}_{\geq 0}$ pour tout $v \in \mathbb{E}$, et $\langle v, v \rangle = 0$ implique $v = \mathbf{0}_{\mathbb{E}}$,

(PS2) $\langle v, w \rangle = \langle w, v \rangle$, pour tous $v, w \in \mathbb{E}$,

(PS3) $\langle v + \lambda \cdot v', w \rangle = \langle v, w \rangle + \lambda \langle v', w \rangle$, pour tous $v, v', w \in \mathbb{E}$ et $\lambda \in \mathbb{R}$.

Étant donné un produit scalaire $\langle\, , \rangle$ sur un espace vectoriel $\mathbb{E}$, on définit l'application $\|\ \| : \mathbb{E} \to \mathbb{R}_{\geq 0}$ donnée par $\|v\| = \sqrt{\langle v, v \rangle}$, pour tout $v \in \mathbb{E}$.

On a le résultat suivant.

**Proposition 1.3.1.** *Soit $\mathbb{E}$ un espace vectoriel muni d'un produit scalaire $\langle\, , \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$. On a l'**inégalité de Cauchy-Schwarz**, i.e.*

$$|\langle v, w \rangle| \leq \|v\| \|w\| \tag{CS}$$

*pour tous $v, w \in \mathbb{E}$. En plus, on a l'égalité si et seulement si $v$ et $w$ sont linéairement dépendants.*

***Preuve.*** Si $w = \mathbf{0}_{\mathbb{E}}$, alors (CS) est clairement vérifiée. On suppose désormais que $w \neq \mathbf{0}_{\mathbb{E}}$. Dans ce cas, (PS1) nous dit que $\langle w, w \rangle \neq 0$. Soient $\alpha = \langle w, w \rangle = \|w\|^2 \neq 0$ et $\beta = \langle v, w \rangle$. On voit bien que

$$\begin{aligned} 0 \leq \|\alpha \cdot v - \beta \cdot w\|^2 &= \big\langle \alpha \cdot v - \beta \cdot w, \alpha \cdot v - \beta \cdot w \big\rangle = \alpha^2 \langle v, v \rangle - 2\alpha\beta \langle v, w \rangle + \beta^2 \langle w, w \rangle \\ &= \|v\|^2 \|w\|^4 - \|w\|^2 \langle v, w \rangle^2, \end{aligned} \tag{1.3.1}$$

*i.e.*

$$\|w\|^2 \langle v, w \rangle^2 \leq \|v\|^2 \|w\|^2.$$

Comme $\|w\|^2 \neq 0$, l'inégalité précédente équivaut à $\langle v, w \rangle^2 \leq \|v\|^2 \|w\|^2$. La fonction racine carrée étant croissante, on voit bien que l'inégalité précédente est équivalente à

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

comme on voulait démontrer. En plus, elle devient une égalité si et seulement si l'inégalité dans (1.3.1) est une égalité, *i.e.* $0 = \big\|\alpha \cdot v - \beta \cdot w\big\|$. D'après (PS1), cela implique que $\alpha \cdot v - \beta \cdot w$ est le vecteur nul, *i.e.* $v$ et $w$ sont linéairement dépendants. □

**Proposition 1.3.2.** *Soit* $\mathbb{E}$ *un espace vectoriel muni d'un produit scalaire* $\langle\, ,\rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$. *Alors, l'application* $\|\ \| : \mathbb{E} \to \mathbb{R}_{\geq 0}$ *donnée par* $\|v\| = \sqrt{\langle v, v\rangle}$ *est une norme.*

***Preuve.*** C'est clair que (PS1) nous dit exactement que $\|\ \|$ satisfait (N1). En outre,

$$\|\lambda \cdot v\| = \sqrt{\langle \lambda \cdot v, \lambda \cdot v\rangle} = \sqrt{\lambda^2 \langle v, v\rangle} = |\lambda| \sqrt{\langle v, v\rangle} = |\lambda| \|v\|,$$

pour tout $v \in \mathbb{E}$ et $\lambda \in \mathbb{R}$, où l'on a utilisé (PS3) dans la deuxième égalité. Cela nous dit que $\|\ \|$ satisfait (N2).

Finalement, on va démontrer que $\|\ \|$ satisfait (N3). En effet,

$$\begin{aligned}\|v + w\|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle &\leq \|v\|^2 + \|w\|^2 + 2|\langle v, w\rangle| \\ &\leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = \big(\|v\| + \|w\|\big)^2,\end{aligned}$$

pour tous $v, w \in \mathbb{E}$, où l'on a utilisé (PS2) et (PS3) dans la deuxième égalité, et l'inégalité (CS). La fonction racine carrée étant croissante, on voit bien que l'inégalité précédente est équivalente à

$$\|v + w\| \leq \|v\| + \|w\|,$$

comme on voulait démontrer. En conséquence, $\|\ \|$ est une norme. □

L'application $\|\ \| : \mathbb{E} \to \mathbb{R}_{\geq 0}$ est appelée la **norme associée** au produit scalaire $\langle\, ,\rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$.

## 1.4 Premiers exemples

### 1.4.1 Classification des normes sur l'espace vectoriel de dimension 1

On peut classifier toutes les normes sur l'espace vectoriel de dimension 1.

**Proposition 1.4.1.** *Soit* $N$ *une norme sur l'espace vectoriel* $\mathbb{R}$. *Alors il existe* $c \in \mathbb{R}_{>0}$ *unique tel que* $N(x) = c|x|$, *pour tout* $x \in \mathbb{R}$. *Réciproquement, pour tout* $c > 0$, *l'application* $x \mapsto c|x|$ *est une norme sur* $\mathbb{R}$. *Cela nous donne une bijection entre* $\mathbb{R}_{>0}$ *et* $\mathrm{Norm}(\mathbb{R})$.

***Preuve.*** Pour la première partie de l'énoncé, on pose $c = N(1)$. D'après (N1), $c > 0$. En outre,

$$N(x) = N(x.1) = |x|N(1) = |x|c,$$

où l'on a utilisé (N2) dans la deuxième égalité. L'unicité de $c$ est claire, vu que $N(x) = c|x|$, pour tout $x \in \mathbb{R}$, nous dit *a fortiori* que $N(1) = c$.

La dernière partie suit de la proposition précédente et de l'Exemple 1.1.3. □

### 1.4.2 La norme infini

Soit $n \in \mathbb{Z}_{>0}$. On définit la **norme infini** $\|\ \|_\infty : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ sur $\mathbb{R}^n$ de la façon suivante. Étant donné $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, on pose

$$\|x\|_\infty = \max_{i \in [\![1,n]\!]} \big(|x_i|\big). \tag{1.4.1}$$

**Proposition 1.4.2.** *Soit* $n \in \mathbb{Z}_{>0}$ *et soit* $\|\ \|_\infty : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ *l'application définie par* (1.4.1). *Alors,* $\|\ \|_\infty$ *est une norme.*

***Preuve.*** Le résultat est une conséquence directe de la Proposition 1.2.4, vu que la norme infini de $\mathbb{R}^n$ est la norme $N_{\max}$ sur $\mathbb{R} \times \cdots \times \mathbb{R}$ ($n$ facteurs), où chaque $\mathbb{R}$ est muni de la norme donnée par la valeur absolue. □

### 1.4.3 La norme $L^1$

Soit $n \in \mathbb{Z}_{>0}$. On définit la **norme $L^1$**, aussi appelée **norme 1**, $\| \; \|_1 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ sur $\mathbb{R}^n$ de la façon suivante. Étant donné $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on pose

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|. \tag{1.4.2}$$

**Proposition 1.4.3.** *Soit $n \in \mathbb{Z}_{>0}$ et soit $\| \; \|_1 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ l'application définie par* (1.4.2)*. Alors, $\| \; \|_1$ est une norme.*

***Preuve.*** Le résultat est une conséquence directe de la Proposition 1.2.4, vu que la norme infini de $\mathbb{R}^n$ est la norme $N_{\text{sum}}$ sur $\mathbb{R} \times \cdots \times \mathbb{R}$ ($n$ facteurs), où chaque $\mathbb{R}$ est muni de la norme donnée par la valeur absolue. □

### 1.4.4 La norme $L^2$

Soit $n \in \mathbb{Z}_{>0}$. On définit la **norme $L^2$**, aussi appelée **norme 2** ou **norme euclidienne**, $\| \; \|_2 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ sur $\mathbb{R}^n$ de la façon suivante. Étant donné $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on pose

$$\|x\|_2 = \left( \sum_{i=1}^{n} x_i^2 \right)^{1/2} = \left( x_1^2 + \cdots + x_n^2 \right)^{1/2}. \tag{1.4.3}$$

**Proposition 1.4.4.** *Soit $n \in \mathbb{Z}_{>0}$ et soit $\| \; \|_2 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ l'application définie par* (1.4.3)*. Alors, $\| \; \|_2$ est une norme.*

***Preuve.*** Il s'agit d'une conséquence immédiate de la Proposition 1.3.2, vu que la norme 2 de $\mathbb{R}^n$ est la norme associée au produit scalaire $\langle \, , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ donné par

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n,$$

pour tous les vecteurs $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ de $\mathbb{R}^n$. □

## 1.5 Équivalence de normes

**Définition 1.5.1.** *Soient $N$ et $N'$ deux normes sur un espace vectoriel $\mathbb{E}$. On dit que $N$ est **équivalente** à $N'$ s'il existe $C, C' \in \mathbb{R}_{>0}$ tels que*

$$CN(v) \leq N'(v) \leq C'N(v), \tag{Equiv}$$

*pour tout $v \in \mathbb{E}$.*

On la caractérisation suivante de la équivalence en termes de boules.

**Proposition 1.5.2.** *Soient $N$ et $N'$ deux normes sur un espace vectoriel $\mathbb{E}$. Alors $N$ est équivalente à $N'$ si et seulement s'il existe $C, C' \in \mathbb{R}_{>0}$ tels que*

$$\bar{B}_{N'}(\mathbf{0}_{\mathbb{E}}, C) \subseteq \bar{B}_N(\mathbf{0}_{\mathbb{E}}, 1) \subseteq \bar{B}_{N'}(\mathbf{0}_{\mathbb{E}}, C'). \tag{Equiv'}$$

*En fait, de façon équivalente, on peut prendre n'importe quel vecteur $v \in \mathbb{E}$ pour le centre des boules dans* (Equiv')*, au lieu de $\mathbf{0}_{\mathbb{E}}$.*

***Preuve.*** On laisse au lecteur/à la lectrice la vérification directe de l'équivalence entre (Equiv) et (Equiv'). □

**Proposition 1.5.3.** *Soit $\mathrm{Norm}(\mathbb{E})$ la partie de $M(\mathbb{E}, \mathbb{R}_{\geq 0})$ formée de toutes les normes sur $\mathbb{E}$. Pour $N$ et $N'$ dans $\mathrm{Norm}(\mathbb{E})$, on écrit $N \sim N'$ si $N$ est équivalente à $N'$. Alors $\sim$ définit une relation d'équivalence sur $\mathrm{Norm}(\mathbb{E})$.*

***Preuve.*** C'est clair que (Equiv) est vérifiée pour $N = N'$ et $C = C' = 1$. Cela implique que $N \sim N$, pour tout $N \in \mathrm{Norm}(\mathbb{E})$. En outre, supposons que $N \sim N'$, *i.e.* il existe $C, C' \in \mathbb{R}_{>0}$ qui vérifient (Equiv). Alors,

$$(C')^{-1}N'(v) \leq N(v) \leq C^{-1}N'(v),$$

pour tout $v \in \mathbb{E}$, *i.e.* $N' \sim N$. Finalement, soient $N, N', N'' \in \mathrm{Norm}(\mathbb{E})$ tels que $N \sim N'$ et $N' \sim N''$. On suppose en particulier qu'il existe $C, C', C'', C''' \in \mathbb{R}_{>0}$ tels que

$$CN(v) \leq N'(v) \leq C'N(v) \text{ et } C''N'(v) \leq N''(v) \leq C'''N'(v),$$

pour tout $v \in \mathbb{E}$. Cela nous dit que

$$CC''N(v) \leq N''(v) \leq C'C'''N(v),$$

pour tout $v \in \mathbb{E}$, *i.e.* $N \sim N''$. En conséquence, $\sim$ définit une relation d'équivalence sur $\mathrm{Norm}(\mathbb{E})$. □

D'après le résultat précédent, au lieu de dire que $N$ est équivalente à $N'$, on dira plutôt que $N$ et $N'$ sont **équivalentes**.

**Proposition 1.5.4.** *Soient $\mathbb{E}_1, \dots, \mathbb{E}_n$ des espaces vectoriels munis des normes $N_1, \dots, N_n$, respectivement. Alors les normes $N_{\max}$ et $N_{\text{sum}}$ sur $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$, définies par* (1.2.1) *et* (1.2.2)*, respectivement, sont équivalentes. En conséquence, les normes $\| \ \|_\infty$ et $\| \ \|_1$ sur $\mathbb{R}^n$ sont équivalentes.*

***Preuve.*** On voit bien que

$$N_{\max}(x) = \max_{i \in [\![1,n]\!]} \big(N_i(x_i)\big) \leq \sum_{i=1}^{n} N_i(x_i) = N_{\text{sum}}(x),$$

pour tout $x = (x_1, \dots, x_n)$, où $x_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$. Par ailleurs,

$$N_{\text{sum}}(x) = \sum_{i=1}^{n} N_i(x_i) \leq \sum_{i=1}^{n} \max_{j \in [\![1,n]\!]} \big(N_j(x_j)\big) = n \max_{j \in [\![1,n]\!]} \big(N_j(x_j)\big) = nN_{\max}(x),$$

pour tout $x = (x_1, \dots, x_n)$, où $x_i \in \mathbb{E}_i$ pour tout $i \in [\![1, n]\!]$. Cela implique que $N_{\max}$ et $N_{\text{sum}}$ sont équivalentes.

La dernière partie de l'énoncé suit de la première et du fait que la norme $\| \ \|_\infty$ (resp., $\| \ \|_1$) sur $\mathbb{R}^n$ est précisément la norme $N_{\max}$ (resp., $N_{\text{sum}}$) sur $\mathbb{R} \times \cdots \times \mathbb{R}$ ($n$ facteurs), où chaque $\mathbb{R}$ est muni de la norme donnée par la valeur absolue. □

On va démontrer dans la suite que toutes les normes sur un espace vectoriel de dimension finie sont équivalentes. Il s'agit de l'un des résultats les plus importants de ce cours. Pour cela il faut introduire quelques notions topologiques.

## 1.6 Suites

Soit $X$ un ensemble. On rappelle qu'une **suite à valeurs dans $X$** est une application $x : \mathbb{N} \to X$. On écrira la suite $x$ sous la forme $(x_n)_{n \in \mathbb{N}}$, où $x_n = x(n)$. L'ensemble de suites à valeurs dans $X$ est $X^{\mathbb{N}}$.

**Définition 1.6.1.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $(v_n)_{n \in \mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ une suite à valeurs dans $\mathbb{E}$. On dit que $(v_n)_{n \in \mathbb{N}}$ est*

*(S1)* ***convergente*** *(pour $N$) s'il existe $v \in \mathbb{E}$ tel que, pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N(v - v_n) < \varepsilon$;*

*(S2)* ***de Cauchy*** *(pour $N$) si pour tout $\varepsilon > 0$, il existe $n_1 \in \mathbb{N}$ tel que, pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N(v_n - v_m) < \varepsilon$;*

*(S3)* ***bornée*** *(pour $N$) s'il existe $C \in \mathbb{R}_{\geq 0}$ tel que $N(v_n) < C$ pour tout $n \in \mathbb{N}$.*

*Si besoin, dans les deux premières définitions on écrira $n_0(\varepsilon)$ et $n_1(\varepsilon)$ pour remarquer la dépendance de $\varepsilon$. En outre, on ne mentionnera pas la norme si la même est claire du contexte.*

**Proposition 1.6.2.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ une suite convergente. On suppose que $v$ et $v'$ satisfont la condition dans la définition de convergence de $(v_n)_{n\in\mathbb{N}}$, i.e. pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ (resp., $n'_0 \in \mathbb{N}$) tel que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ (resp., $n \geq n'_0$) implique que $N(v - v_n) < \varepsilon$ (resp., $N(v' - v_n) < \varepsilon$). Alors, $v = v'$. On appelle cet élément la **limite** de la suite $(v_n)_{n\in\mathbb{N}}$ et on écrit*

$$\lim_{n\to+\infty} v_n = v.$$

***Preuve.*** On voit bien que

$$N(v - v') = N(v - v_n + v_n - v') \leq N(v - v_n) + N(v_n - v'),$$

pour tout $n \in \mathbb{N}$, d'après (N3) pour $N$. On affirme que $N(v - v') = 0$. Si ce n'est pas le cas, alors $N(v - v') > 0$ et on prend $\varepsilon = N(v - v')/4$. Soient $n_0 \in \mathbb{N}$ et $n'_0 \in \mathbb{N}$ les éléments correspondants dans le définition de convergence de $(v_n)_{n\in\mathbb{N}}$. Alors, la condition $n \geq \max(n_0, n'_0)$ implique que $N(v - v_n) < \varepsilon$ et $N(v' - v_n) < \varepsilon$. En conséquence,

$$N(v - v') = N(v - v_n + v_n - v') \leq N(v - v_n) + N(v_n - v') < \varepsilon + \varepsilon = \frac{N(v - v')}{2},$$

ce qui est absurde si $N(v - v') > 0$. En conséquence, $N(v - v') = 0$. D'après (N1) pour $N$ on conclut que $v = v'$, comme on voulait démontrer. □

On a les résultats standards sur l'arithmétique basique de suites convergentes, de Cauchy et bornées.

**Proposition 1.6.3.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soient $(v_n)_{n\in\mathbb{N}}, (w_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ deux suites. Soit $\lambda \in \mathbb{R}$. On a les résultats suivants:*

*(i) si $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ sont convergentes avec limites $v \in \mathbb{E}$ et $w \in \mathbb{E}$, respectivement, alors la suite $(v_n + \lambda \cdot w_n)_{n\in\mathbb{N}}$ est convergente avec limite $v + \lambda \cdot w$;*

*(ii) si $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ sont suites de Cauchy, alors $(v_n + \lambda \cdot w_n)_{n\in\mathbb{N}}$ est aussi une suite de Cauchy;*

*(iii) si $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ sont suites bornées, alors $(v_n + \lambda \cdot w_n)_{n\in\mathbb{N}}$ est aussi une suite bornée.*

***Preuve.*** Si $\lambda = 0$, il n'y a rien à démontrer (dans les trois cas), donc on va supposer désormais $\lambda \neq 0$.

On démontrera d'abord l'item (i). Si $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ convergent vers $v$ et $w$, respectivement, alors pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ et $n'_0 \in \mathbb{N}$ tels que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N(v - v_n) < \varepsilon$ et la condition $n \geq n'_0$ implique que $N(w - w_n) < \varepsilon$. On fixe $\varepsilon > 0$ et on définit $\bar{n}_0(\varepsilon) = \max(n_0(\varepsilon/2), n_0(\varepsilon/(2|\lambda|))) \in \mathbb{N}$. Pour tout $n \geq \bar{n}_0(\varepsilon)$, on a

$$N\big((v+\lambda\cdot w)-(v_n+\lambda\cdot w_n)\big) = N\big((v-v_n)+\lambda\cdot(w-w_n)\big) \leq N(v-v_n)+|\lambda|N(w-w_n) < \frac{\varepsilon}{2}+|\lambda|\frac{\varepsilon}{2|\lambda|} = \varepsilon,$$

*i.e.* la suite $(v_n + \lambda w_n)_{n\in\mathbb{N}}$ converge vers $v + \lambda \cdot w$.

On va procéder à démontrer l'item (ii). Si $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ sont deux suites de Cauchy, alors pour tout $\varepsilon > 0$, il existe $n_1 \in \mathbb{N}$ et $n'_1 \in \mathbb{N}$ tels que, pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N(v_n - v_m) < \varepsilon$ et la condition $n, m \geq n'_1$ implique que $N(w_n - w_m) < \varepsilon$. On fixe $\varepsilon > 0$ et on définit $\bar{n}_1(\varepsilon) = \max(n_1(\varepsilon/2), n_1(\varepsilon/(2|\lambda|))) \in \mathbb{N}$. Pour tous $n, m \geq \bar{n}_1(\varepsilon)$, on a

$$\begin{aligned} N\big((v_n + \lambda\cdot w_n) - (v_m + \lambda\cdot\cdot w_m)\big) &= N\big((v_n - v_m) + \lambda\cdot(w_n - w_m)\big) \leq N(v_n - v_m) + |\lambda|N(w_n - w_m) \\ &< \frac{\varepsilon}{2} + |\lambda|\frac{\varepsilon}{2|\lambda|} = \varepsilon. \end{aligned}$$

*i.e.* la suite $(v_n + \lambda w_n)_{n\in\mathbb{N}}$ est de Cauchy.

Finalement, on va démontrer l'item (iii). Si $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ sont deux suites bornées, alors il existe $C > 0$ et $C' > 0$ tels que $N(v_n) < C$ et $N(w_n) < C'$, pour tout $n \in \mathbb{N}$. Alors, $N(v_n + \lambda\cdot w_n) \leq N(v_n) + |\lambda|N(w_n) < C + |\lambda C'|$, pour tout $n \in \mathbb{N}$, *i.e.* la suite $(v_n + \lambda\cdot w_n)_{n\in\mathbb{N}}$ est bornée. □

**Proposition 1.6.4.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ une suite.*

*(i) Si $(v_n)_{n\in\mathbb{N}}$ est convergente, elle est de Cauchy.*

*(ii) Si $(v_n)_{n\in\mathbb{N}}$ est de Cauchy, elle est bornée.*

***Preuve.*** On démontrera d'abord l'item (i). On suppose alors que $(v_n)_{n\in\mathbb{N}}$ est convergente et que sa limite est $v \in \mathbb{E}$. Alors, pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N(v - v_n) < \varepsilon$. Soit $\varepsilon > 0$. On pose $n_1(\varepsilon) = n_0(\varepsilon/2)$. Pour tous $n, m \in \mathbb{N}$ tels que $n, m \geq n_1(\varepsilon)$, on trouve alors

$$N(v_n - v_m) = N(v_n - v + v - v_m) \leq N(v_n - v) + N(v - v_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

*i.e.* $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy.

On démontrera finalement l'item (ii). On suppose que $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy. Alors, pour tout $\varepsilon > 0$, il existe $n_1 \in \mathbb{N}$ tel que, pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N(v_n - v_m) < \varepsilon$. On prend $\varepsilon = 1$. Alors pour tous $n, m \in \mathbb{N}$ tels que $n, m \geq n_1(1)$, on a que $N(v_n - v_m) < 1$. En conséquence,

$$N(v_n) = N(v_n - v_{n_1(1)} + v_{n_1(1)}) \leq N(v_n - v_{n_1(1)}) + N(v_{n_1(1)}) < 1 + N(v_{n_1(1)}),$$

pour tout $n \in \mathbb{N}$ tel que $n \geq n_1(1)$. Soit

$$C' = \max_{i \in [\![1, n_1(1)]\!]} \big(N(x_i)\big) \geq 0$$

et $C = \max(C', 1 + N(v_{n_1(1)}))$. C'est clair que $C > 0$ et que $N(v_n) < C$, pour tout $n \in \mathbb{N}$, ce qui dit que la suite $(v_n)_{n\in\mathbb{N}}$ est bornée. $\square$

Soit $\mathbb{E}$ un espace vectoriel. On rappelle qu'une **propriété** sur les suites à valeurs dans $\mathbb{E}$ est par définition une partie de $\mathbb{E}^{\mathbb{N}}$. Soit

$$\mathcal{P} : \mathrm{Norm}(\mathbb{E}) \to \mathscr{P}(\mathbb{E}^{\mathbb{N}})$$

une application. De façon intuitive, pour une norme $N \in \mathrm{Norm}(\mathbb{E})$ sur $\mathbb{E}$, $\mathcal{P}(N)$ dénote une propriété sur les suites $(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ à valeurs dans $\mathbb{E}$ qui est définie en termes de $N$. Par exemple, on peut considérer l'application

$$\mathcal{P}_{\mathrm{conv}} : \mathrm{Norm}(\mathbb{E}) \to \mathscr{P}(\mathbb{E}^{\mathbb{N}})$$

donnée par

$$\mathcal{P}_{\mathrm{conv}}(N) = \big\{(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}} : (v_n)_{n\in\mathbb{N}} \text{ est convergent}\big\}$$

pour tout $N \in \mathrm{Norm}(\mathbb{E})$. De façon analogue, on peut aussi considérer les applications

$$\mathcal{P}_{\mathrm{dC}}, \mathcal{P}_{\mathrm{bor}} : \mathrm{Norm}(\mathbb{E}) \to \mathscr{P}(\mathbb{E}^{\mathbb{N}})$$

données par

$$\mathcal{P}_{\mathrm{dC}}(N) = \big\{(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}} : (v_n)_{n\in\mathbb{N}} \text{ est de Cauchy}\big\}$$

et

$$\mathcal{P}_{\mathrm{bor}}(N) = \big\{(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}} : (v_n)_{n\in\mathbb{N}} \text{ est borné}\big\}$$

pour tout $N \in \mathrm{Norm}(\mathbb{E})$. On dira qu'une application $\mathcal{P}$ comme ci-dessus est **fortement topologique** si $\mathcal{P}(N) = \mathcal{P}(N')$ pour toute paire de normes $N, N' \in \mathrm{Norm}(\mathbb{E})$ équivalentes. Le résultat suivant nous dit que les propriétés $\mathcal{P}_{\mathrm{conv}}$ d'être convergent, $\mathcal{P}_{\mathrm{dC}}$ d'être de Cauchy et $\mathcal{P}_{\mathrm{bor}}$ d'être borné sont fortement topologiques.

**Proposition 1.6.5.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $N'$ une norme sur $\mathbb{E}$ équivalente à $N$. Soit $(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ une suite. On a les résultats suivants:*

*(i) $(v_n)_{n\in\mathbb{N}}$ converge vers $v \in \mathbb{E}$ pour $N$ si et seulement si $(v_n)_{n\in\mathbb{N}}$ converge vers $v \in \mathbb{E}$ pour $N'$;*

*(ii) $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour $N$ si et seulement si $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour $N'$;*

*(iii)* $(v_n)_{n\in\mathbb{N}}$ *est une suite bornée pour* $N$ *si et seulement si* $(v_n)_{n\in\mathbb{N}}$ *est une suite bornée pour* $N'$.

***Preuve.*** On suppose qu'il existe $C, C' \in \mathbb{R}_{>0}$ tels que

$$CN(v) \leq N'(v) \leq C'N(v),$$

pour tout $v \in \mathbb{E}$. C'est clair que pour tous les items, il suffit de démontrer que si la propriété est vérifiée pour $N$ alors elle est aussi vérifiée pour $N'$.

On démontrera d'abord l'item (i). Si $(v_n)_{n\in\mathbb{N}}$ converge vers $v$, alors pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N(v - v_n) < \varepsilon$. On fixe $\varepsilon > 0$ et on définit $n_0'(\varepsilon) = n_0(\varepsilon/C') \in \mathbb{N}$. Pour tout $n \geq n_0'(\varepsilon)$, on a

$$N'(v - v_n) \leq C'N(v - v_n) < C'\frac{\varepsilon}{C'} = \varepsilon,$$

*i.e.* la suite $(v_n)_{n\in\mathbb{N}}$ converge vers $v$ pour $N'$.

On va procéder à démontrer l'item (ii). Si $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy, alors pour tout $\varepsilon > 0$, il existe $n_1 \in \mathbb{N}$ tel que, pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N(v_n - v_m) < \varepsilon$. On fixe $\varepsilon > 0$ et on définit $n_1'(\varepsilon) = n_1(\varepsilon/C') \in \mathbb{N}$. Pour tous $n, m \geq n_1'(\varepsilon)$, on a

$$N'(v_n - v_m) \leq C'N(v_n - v_m) < C'\frac{\varepsilon}{C'} = \varepsilon,$$

*i.e.* la suite $(v_n)_{n\in\mathbb{N}}$ est de Cauchy pour $N'$.

Finalement, on va démontrer l'item (iii). Si $(v_n)_{n\in\mathbb{N}}$ est une suites bornée, alors il existe $\bar{C} > 0$ tel que $N(v_n) < \bar{C}$, pour tout $n \in \mathbb{N}$. Alors, $N'(v_n) \leq C'N(v_n) < C'\bar{C}$, pour tout $n \in \mathbb{N}$, *i.e.* la suite $(v_n)_{n\in\mathbb{N}}$ est bornée. □

Le résultat suivant est immédiat. On laisse la preuve au lecteur/à la lectrice.

**Proposition 1.6.6.** *Soient* $\mathbb{E}$ *et* $\mathbb{F}$ *deux espaces vectoriels et* $T : \mathbb{F} \to \mathbb{E}$ *une application linéaire injective. On suppose que* $\mathbb{E}$ *est muni d'une norme* $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$. *On considère la norme* $N_T : \mathbb{F} \to \mathbb{R}_{\geq 0}$ *donnée par* $N_T = N \circ T$ *(voir la Proposition 1.2.1). Soit* $(v_n)_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ *une suite. Alors,* $(v_n)_{n\in\mathbb{N}}$ *est convergente vers* $v \in \mathbb{F}$ *(resp., de Cauchy, bornée) pour* $N_T$ *si et seulement si* $(T(v_n))_{n\in\mathbb{N}}$ *est convergente vers* $T(v) \in \mathbb{E}$ *(resp., de Cauchy, bornée) pour* $N$.

**Proposition 1.6.7.** *Soient* $\mathbb{E}_1, \ldots, \mathbb{E}_n$ *des espaces vectoriels munis des normes* $N_1, \ldots, N_n$*, respectivement. On considère le produit cartésien* $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$. *Soit* $(x_n) \in \mathbb{E}^{\mathbb{N}}$ *une suite. Pour tout* $n \in \mathbb{N}$*, on écrit* $x_n = x_{n,i} \in \mathbb{E}_i$ *avec* $i \in [\![1, n]\!]$. *Alors* $(x_{n,i}) \in \mathbb{E}_i^{\mathbb{N}}$ *est une suite pour tout* $i \in [\![1, n]\!]$. *On a les résultats suivants:*

*(i)* $(x_n) \in \mathbb{E}^{\mathbb{N}}$ *converge vers* $x_{\lim} = (x_{\lim,1}, \ldots, x_{\lim,n}) \in \mathbb{E}$ *pour* $N_{\max}$ *(ou* $N_{\text{sum}}$*) si et seulement si* $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ *converge vers* $x_{\lim,i} \in \mathbb{E}_i$ *pour* $N_i$ *pour tout* $i \in [\![1, n]\!]$*;*

*(ii)* $(x_n) \in \mathbb{E}^{\mathbb{N}}$ *est une suite de Cauchy pour* $N_{\max}$ *(ou* $N_{\text{sum}}$*) si et seulement si* $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ *est une suite de Cauchy pour* $N_i$ *pour tout* $i \in [\![1, n]\!]$*;*

*(iii)* $(x_n) \in \mathbb{E}^{\mathbb{N}}$ *est une suite bornée pour* $N_{\max}$ *(ou* $N_{\text{sum}}$*) si et seulement si* $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ *est une suite bornée pour* $N_i$ *pour tout* $i \in [\![1, n]\!]$.

***Preuve.*** On note d'abord que, vu que $N_{\max}$ et $N_{\text{sum}}$ sont des normes équivalentes sur $\mathbb{E}$ (voir la Proposition 1.5.4), la Proposition 1.6.5 nous dit qu'une suite est convergente (resp., de Cauchy, bornée) pour $N_{\max}$ si et seulement si elle est convergente (resp., de Cauchy, bornée) pour $N_{\text{sum}}$. Cela nous dit qu'il suffit de démontrer les items précédents avec $N_{\max}$.

On montrera d'abord l'item (i). Supposons que $(x_n) \in \mathbb{E}^{\mathbb{N}}$ converge vers $x_{\lim} = (x_{\lim,1}, \ldots, x_{\lim,n}) \in \mathbb{E}$ pour $N_{\max}$. Alors, pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N_{\max}(x_{\lim} - x_n) < \varepsilon$. Comme $N_i(x_{\lim,i} - x_{n,i}) \leq N_{\max}(x_{\lim} - x_n)$ pour tout $i \in [\![1, n]\!]$, alors pour tout $\varepsilon > 0$ et tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N_i(x_{\lim,i} - x_{n,i}) < \varepsilon$, pour tout $i \in [\![1, n]\!]$, *i.e.* $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ converge vers $x_{\lim,i} \in \mathbb{E}_i$ pour $N_i$ pour tout $i \in [\![1, n]\!]$. On suppose maintenant que $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ converge vers $x_{\lim,i} \in \mathbb{E}_i$ pour $N_i$ pour tout $i \in [\![1, n]\!]$. Cela veut dire que pour tous

$\varepsilon > 0$ et $i \in [\![1, n]\!]$, il existe $n_{0,i} \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, la condition $n \geq n_{0,i}$ implique que $N_i(x_{\lim,i} - x_{n,i}) < \varepsilon$. On fixe $\varepsilon > 0$ et on prend $n_{0,i} \in \mathbb{N}$ comme ci-dessus, pour tout $i \in [\![1, n]\!]$. On pose

$$n_0 = \max_{i \in [\![1,n]\!]} (n_{0,i}).$$

On voit bien que pour tout $n \in \mathbb{N}$, la condition $n \geq n_0$ implique que $N_i(x_{\lim,i} - x_{n,i}) < \varepsilon$ pour tout $i \in [\![1, n]\!]$, ce qui nous dit que $N_{\max}(x_{\lim} - x_n) < \varepsilon$ par la définition de maximum. En conséquence, $(x_n) \in \mathbb{E}^{\mathbb{N}}$ converge vers $x_{\lim} = (x_{\lim,1}, \dots, x_{\lim,n}) \in \mathbb{E}$ pour $N_{\max}$.

On montrera maintenant l'item (ii). Supposons que $(x_n) \in \mathbb{E}^{\mathbb{N}}$ est une suite de Cauchy pour $N_{\max}$. Alors, pour tout $\varepsilon > 0$, il existe $n_1 \in \mathbb{N}$ tel que, pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N_{\max}(x_n - x_m) < \varepsilon$. Comme $N_i(x_{n,i} - x_{m,i}) \leq N_{\max}(x_n - x_m)$ pour tout $i \in [\![1, n]\!]$, alors pour tout $\varepsilon > 0$ et tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N_i(x_{n,i} - x_{m,i}) < \varepsilon$, pour tout $i \in [\![1, n]\!]$, *i.e.* $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ est une suite de Cauchy pour $N_i$ pour tout $i \in [\![1, n]\!]$. On suppose maintenant que $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ est une suite de Cauchy pour $N_i$ pour tout $i \in [\![1, n]\!]$. Cela veut dire que pour tous $\varepsilon > 0$ et $i \in [\![1, n]\!]$, il existe $n_{1,i} \in \mathbb{N}$ tel que, pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_{0,i}$ implique que $N_i(x_{n,i} - x_{m,i}) < \varepsilon$. On fixe $\varepsilon > 0$ et on prend $n_{1,i} \in \mathbb{N}$ comme ci-dessus, pour tout $i \in [\![1, n]\!]$. On pose

$$n_1 = \max_{i \in [\![1,n]\!]} (n_{1,i}).$$

On voit bien que pour tous $n, m \in \mathbb{N}$, la condition $n, m \geq n_1$ implique que $N_i(x_{n,i} - x_{m,i}) < \varepsilon$ pour tout $i \in [\![1, n]\!]$, ce qui nous dit que $N_{\max}(x_n - x_m) < \varepsilon$ par la définition de maximum. En conséquence, $(x_n) \in \mathbb{E}^{\mathbb{N}}$ est une suite de Cauchy pour $N_{\max}$.

On montrera finalement l'item (iii). Supposons que $(x_n) \in \mathbb{E}^{\mathbb{N}}$ est une suite bornée pour $N_{\max}$. Alors, il existe $C > 0$ tel que $N_{\max}(x_n) < C$, pour tout $n \in \mathbb{N}$. Comme $N_i(x_{n,i}) \leq N_{\max}(x_n) < C$ pour tout $i \in [\![1, n]\!]$, alors $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ est une suite bornée pour $N_i$ pour tout $i \in [\![1, n]\!]$. Réciproquement, supposons que $(x_{n,i})_{n\in\mathbb{N}} \in \mathbb{E}_i^{\mathbb{N}}$ est une suite bornée pour $N_i$ pour tout $i \in [\![1, n]\!]$. Alors, pour tout $i \in [\![1, n]\!]$ il existe $C_i > 0$ tel que $N_i(x_{n,i}) < C_i$, pour tout $n \in \mathbb{N}$. Soit $C = \max(C_1, \dots, C_n)$. Comme $N_i(x_{n,i}) < C_i \leq C$ pour tout $i \in [\![1, n]\!]$, alors $N_{\max}(x_n) < C$, par la définition de maximum. Cela nous dit que $(x_n) \in \mathbb{E}^{\mathbb{N}}$ est une suite bornée pour $N_{\max}$. □

## 1.7 Espaces complets

**Définition 1.7.1.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé. On dit que $(\mathbb{E}, N)$ est **complet** si toute suite de Cauchy $(v_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ admet une limite $v \in \mathbb{E}$. Par abus de notation, on dira plus simplement que $\mathbb{E}$ est **complet**, où la norme est sous-entendue.*

**Exemple 1.7.2.** *On admettra que l'espace vectoriel normé $(\mathbb{R}, |\ |)$ est complet. Il s'agit d'une conséquence de la construction des nombres réels.*

Soit $\mathbb{E}$ un espace vectoriel et soit $\mathcal{P}$ une propriété sur les normes sur $\mathbb{E}$, *i.e.* une partie de $\operatorname{Norm}(\mathbb{E})$. De la même façon que dans le cas de suites, on dira que $\mathcal{P}$ est **fortement topologique** si, pour toute norme $N'$ sur $\mathbb{E}$ équivalente à $N$, $N \in \mathcal{P}$ si et seulement si $N' \in \mathcal{P}$. Par exemple, soit

$$\mathcal{P} = \{N \in \operatorname{Norm}(\mathbb{E}) : (\mathbb{E}, N) \text{ est complet}\}.$$

Le résultat suivant nous dit que la propriété précédente est fortement topologique.

**Proposition 1.7.3.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $N'$ une norme sur $\mathbb{E}$ équivalente à $N$. Alors $(\mathbb{E}, N)$ est complet si et seulement si $(\mathbb{E}, N')$ est complet.*

***Preuve.*** Il s'agit d'une conséquence immédiate de la Proposition 1.6.5. □

**Proposition 1.7.4.** *Soient $\mathbb{E}$ et $\mathbb{F}$ deux espaces vectoriels et $T : \mathbb{F} \to \mathbb{E}$ une application linéaire bijective. On suppose que $\mathbb{E}$ est muni d'une norme $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$ et on considère la norme $N_T : \mathbb{F} \to \mathbb{R}_{\geq 0}$ donnée par $N_T = N \circ T$ (voir la Proposition 1.2.1). Alors, $(\mathbb{E}, N)$ est complet si et seulement si $(\mathbb{F}, N_T)$ est complet.*

***Preuve.*** Il s'agit d'une conséquence directe de la Proposition 1.2.1. □

**Proposition 1.7.5.** *Soient $\mathbb{E}_1, \dots, \mathbb{E}_n$ des espaces vectoriels munis des normes $N_1, \dots, N_n$, respectivement. On considère le produit cartésien $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$. Si $(\mathbb{E}_i, N_i)$ est complet pour tout $i \in [\![1, n]\!]$, alors $(\mathbb{E}, N_{\max})$ et $(\mathbb{E}, N_{\mathrm{sum}})$ sont complets.*

***Preuve.*** Il s'agit d'une conséquence immédiate de la Proposition 1.6.7. □

**Exemple 1.7.6.** *Soit $n \in \mathbb{Z}_{>0}$. Comme la norme $\| \ \|_\infty$ (resp., $\| \ \|_1$) sur $\mathbb{R}^n$ est précisément la norme $N_{\max}$ (resp., $N_{sum}$) sur $\mathbb{R} \times \cdots \times \mathbb{R}$ ($n$ facteurs), où chaque $\mathbb{R}$ est muni de la norme donnée par la valeur absolue et $(\mathbb{R}, | \ |)$ est complet, alors $(\mathbb{R}^n, \| \ \|_\infty)$ et $(\mathbb{R}^n, \| \ \|_1)$ sont des espaces vectoriels normés complets.*

**Non-exemple 1.7.7.** *On va voir après que tout espace vectoriel normé de dimension finie est complet (voir Corollaire 1.14.1). Ce résultat n'est pas vrai si l'on considère des espaces vectoriels normés de dimension infinie. Par exemple, soit $\mathbb{E}$ l'espace vectoriel formé de toutes les fonctions polynomiales $f : [0, 1] \to \mathbb{R}$, muni de la norme*

$$\|f\| = \sup_{x \in [0,1]} |f(x)|.$$

*On peut démontrer que $(\mathbb{E}, \| \ \|)$ est un espace vectoriel normé. Par contre, il n'est pas complet. En effet, on considère par exemple la suite $(f_n)_{n \in \mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ donnée par*

$$f_n(x) = \sum_{i=0}^{n} \frac{x^i}{i!},$$

*pour tout $x \in [0, 1]$ et $n \in \mathbb{N}$. Ce n'est pas difficile à montrer que $(f_n)_{n \in \mathbb{N}}$ est une suite de Cauchy (pour $\| \ \|$) mais elle n'a pas de limite dans $\mathbb{E}$. De façon un peu intuitive, la limite "devrait" être la fonction exponentielle $x \mapsto e^x$, pour $x \in [0, 1]$, qui n'est pas polynômiale.*

*On autre exemple d'espace non complet peut s'obtenir si l'on travaille sur des espaces vectoriels définis sur un corps autre que $\mathbb{R}$. En particulier, on remarque d'abord que toutes les définitions que l'on a considéré de norme, suites et complétude ont du sens si l'on travaille avec des espaces vectoriels sur le corps $\mathbb{Q}$ de nombres rationnels. Dans ce cas, le $\mathbb{Q}$-espace vectoriel de dimension 1, muni de la norme donnée par la valeur absolue, n'est pas complet. Sans rentrer dans les détails, qui nécessitent d'une étude plus approfondie de $\mathbb{Q}$ et de $\mathbb{R}$, on peut montrer que la suite $(q_n)_{n \in \mathbb{N}} \in \mathbb{Q}^{\mathbb{N}}$ donnée par*

$$q_n = \left(1 + \frac{1}{n}\right)^n,$$

*pour tout $n \in \mathbb{N}$, est une suite de Cauchy qui n'a pas de limite dans $\mathbb{Q}$. De façon intuitive, la limite de $(q_n)_{n \in \mathbb{N}} \in \mathbb{Q}^{\mathbb{N}}$ serait $e$ (dans $\mathbb{R}$ muni de la valeur absolue), qui n'est pas rationnel.*

## 1.8 Première preuve de l'équivalence de normes sur un espace vectoriel de dimension finie (optionnel)

On va donner une première démonstration de l'équivalence de normes sur un espace vectoriel de dimension finie. Par contre, on ne va pas utiliser ce résultat dans les sections 1.9-1.12 suivantes. On verra plus tard une autre preuve, plus simple, qui utilise la notion de compacité.

**Théorème 1.8.1.** *Soient $\mathbb{E}$ un espace vectoriel de dimension $n \in \mathbb{Z}_{>0}$ et soient $N$ et $N'$ deux normes sur $\mathbb{E}$. Alors, $N$ et $N'$ sont équivalentes.*

***Preuve.*** Soit $\mathscr{B} = \{v_1, \dots, v_n\} \subseteq \mathbb{E}$ une base (ordonnée) de $\mathbb{E}$. On considère l'application $T : \mathbb{E} \to \mathbb{R}^n$ qui associe à chaque $v = \sum_{i=1}^{n} c_i v_i$ (où $c_i \in \mathbb{R}$ pour tout $i \in [\![1, n]\!]$) le $n$-uplet $T(v) = (c_1, \dots, c_n)$. On définit $N_\infty : \mathbb{E} \to \mathbb{R}_{\geq 0}$ via $N_\infty = \| \ \|_\infty \circ T$. D'après la Proposition 1.2.1, $N_\infty$ est une norme sur $\mathbb{E}$. C'est clair que

$$N_\infty(v) = \max_{i \in [\![1,n]\!]} \big(|c_i|\big), \tag{1.8.1}$$

pour tout $v = \sum_{i=1}^{n} c_i v_i \in \mathbb{E}$. Comme l'équivalence de normes est une relation d'équivalence (voir la Proposition 1.5.3), on peut supposer sans perte de généralité que $N' = N_\infty$.

D'abord, on voit bien que

$$N(v) = N\left(\sum_{i=1}^{n} c_i v_i\right) \leq \sum_{i=1}^{n} N(c_i v_i) = \sum_{i=1}^{n} |c_i| N(v_i),$$

pour tout $v = \sum_{i=1}^{n} c_i v_i \in \mathbb{E}$. Soit $C' = \sum_{i=1}^{n} N(v_i)$. Comme $\{v_1, \dots, v_n\} \subseteq \mathbb{E}$ est une base, $N(v_i) > 0$ pour tout $i \in [\![1, n]\!]$, ce qui dit que $C' > 0$. Alors, (1.8.1) implique que

$$N(v) \leq C' N_\infty(v),$$

pour tout $v \in \mathbb{E}$.

On va démontrer qu'il existe $C > 0$ tel que

$$N(v) \geq C N_\infty(v), \tag{1.8.2}$$

pour tout $v \in \mathbb{E}$. On va procéder par induction sur la dimension $n$ de $\mathbb{E}$. Le cas $n = 1$ suit directement de la Proposition 1.4.1. On suppose désormais que $n > 1$ et que le théorème est vrai pour les espaces vectoriels de dimension (strictement) inférieure à $n$. On va procéder par l'absurde. Supposons alors que (1.8.2) n'est pas vérifiée dans $\mathbb{E}$, *i.e.* pour tout $C > 0$, il existe $w \in \mathbb{E}$ (forcement non nul) tel que $N(w) < C N_\infty(w)$. De façon équivalente, pour tout $m \in \mathbb{Z}_{>0}$, il existe $w_m \in \mathbb{E}$ (forcement non nul) tel que

$$N(w_m) < \frac{1}{m} N_\infty(w_m).$$

Comme $w_m \neq \mathbf{0}_{\mathbb{E}}$ pour tout $m \in \mathbb{Z}_{>0}$, $N_\infty(w_m) \neq 0$. On définit $\bar{w}_m = w_m / N_\infty(w_m) \in \mathbb{E}$, pour tout $m \in \mathbb{Z}_{>0}$. Alors, l'hypothèse de l'absurde est équivalente à dire qu'il existe une suite $(\bar{w}_m)_{m \in \mathbb{Z}_{>0}} \in \mathbb{E}^{\mathbb{Z}_{>0}}$ telle que $N_\infty(\bar{w}_m) = 1$ et

$$N(\bar{w}_m) < \frac{1}{m}, \tag{1.8.3}$$

pour tout $m \in \mathbb{Z}_{>0}$. Comme $\mathscr{B}$ est une base, on peut écrire $\bar{w}_m = \sum_{i=1}^{n} d_{i,m} v_i$, pour tout $m \in \mathbb{Z}_{>0}$. On définit $A_j = \{m \in \mathbb{Z}_{>0} : |d_{m,j}| = 1\}$, pour tout $j \in [\![1, n]\!]$. C'est clair qu'il existe $j_0 \in [\![1, n]\!]$ tel que l'ensemble $A_j$ soit infini. En effet, comme $N_\infty(\bar{w}_m) = 1$, pour tout $m \in \mathbb{Z}_{>0}$, $\cup_{j=1}^{n} A_j = \mathbb{Z}_{>0}$. Si $A_j$ est fini pour tout $j \in [\![1, n]\!]$, alors $\cup_{j=1}^{n} A_j = \mathbb{Z}_{>0}$ est fini, ce qui est absurde.

Soit $j_0 \in [\![1, n]\!]$ tel que l'ensemble $A_j$ soit infini. On pose $A_{j_0}^{+} = \{m \in \mathbb{Z}_{>0} : d_{m,j_0} = 1\}$ et $A_{j_0}^{-} = \{m \in \mathbb{Z}_{>0} : d_{m,j_0} = -1\}$. Comme $A_{j_0}^{+} \cup A_{j_0}^{-} = A_{j_0}$ et $A_{j_0}$ est infini, $A_{j_0}^{+}$ est infini ou $A_{j_0}^{-}$ est infini. On suppose sans perte de généralité que $A_{j_0}^{+}$ est infini (sinon, il faut changer $\bar{w}_m$ par $-\bar{w}_m$, pour tout $m \in \mathbb{Z}_{>0}$). Il existe alors une application croissante et bijective $\phi : \mathbb{Z}_{>0} \to A_{j_0}^{+}$. Soit $\tilde{w}_m = \bar{w}_{\phi(m)}$, pour tout $m \in \mathbb{Z}_{>0}$, et soit $\mathbb{F}$ l'espace vectoriel engendré par $\mathscr{B} \setminus \{v_{j_0}\}$. Alors, $\mathbb{F}$ a dimension $n-1$ et, par récurrence, toutes les normes sur $\mathbb{F}$ sont équivalentes. On remarque que $\mathbb{F}$ est complet pour la norme $N_\infty|_{\mathbb{F}}$, d'après la Proposition 1.7.4. Comme les normes $N|_{\mathbb{F}}$ et $N_\infty|_{\mathbb{F}}$ sont équivalentes, $\mathbb{F}$ est complet pour la norme $N|_{\mathbb{F}}$.

On pose $\tilde{u}_m = \bar{w}_{\phi(m)} - v_{j_0}$, pour tout $m \in \mathbb{Z}_{>0}$. C'est clair que $(\tilde{u}_m)_{m \in \mathbb{Z}_{>0}} \in \mathbb{F}^{\mathbb{Z}_{>0}}$ est une suite à valeurs dans $\mathbb{F}$. Soit $\varepsilon > 0$. On prend $n_1 \in \mathbb{N}$ tel que $n_1 > 2/\varepsilon$. Alors,

$$N(\tilde{u}_n - \tilde{u}_m) = N(\tilde{w}_n - \tilde{w}_m) < N(\tilde{w}_n) + N(\tilde{w}_m) < \frac{1}{\phi(n)} + \frac{1}{\phi(m)} \leq \frac{1}{n} + \frac{1}{m} \leq \frac{2}{n_1} < \varepsilon,$$

pour tous $n, m \geq n_1$, où l'on a appliqué (1.8.3) au troisième membre, et le fait que $\phi(n) \geq n$ au quatrième membre. Cela nous dit que $(\tilde{u}_m)_{m \in \mathbb{Z}_{>0}} \in \mathbb{F}^{\mathbb{Z}_{>0}}$ est une suite de Cauchy de $\mathbb{F}$ pour la norme Comme $\mathbb{F}$ est complet pour la norme $N|_{\mathbb{F}}$, alors $(\tilde{u}_m)_{m \in \mathbb{Z}_{>0}}$ a une limite $\tilde{u} \in \mathbb{F}$ pour $N|_{\mathbb{F}}$. En conséquence,

$$\begin{aligned} 0 \leq N(\tilde{u} + v_{j_0}) = N(\tilde{u} - \tilde{u}_m + \tilde{u}_m + v_{j_0}) &\leq N(\tilde{u} - \tilde{u}_m) + N(\tilde{u}_m + v_{j_0}) = N(\tilde{u} - \tilde{u}_m) + N(\tilde{w}_m) \\ &< N(\tilde{u} - \tilde{u}_m) + \frac{1}{\phi(m)} \leq N(\tilde{u} - \tilde{u}_m) + \frac{1}{m}, \end{aligned} \tag{1.8.4}$$

pour tout $m \geq n_1$. Si l'on prend la limite quand $m$ tend vers plus $+\infty$, le dernier membre de (1.8.4) tend vers zéro. Cela implique que la limite de $N(\tilde{u} + v_{j_0})$ quand $m$ tend vers plus $+\infty$ est aussi zéro. Comme $N(\tilde{u} + v_{j_0})$ ne dépend pas de $m$, on conclut que $N(\tilde{u} + v_{j_0}) = 0$, *i.e.* $\tilde{u} + v_{j_0} = \mathbf{0}_{\mathbb{E}}$. Cela donne une contradiction, puisque $\tilde{u} \in \mathbb{F}$ et $v_{j_0}$ forment un ensemble libre. En conséquence, (1.8.2) est vérifiée et $N$ et $N_\infty$ sont équivalentes. □

**Remarque 1.8.2.** *On remarque que le théorème précédent n'est pas valable dans le cas des espaces vectoriels de dimension infinie.*

## 1.9 Premières définitions topologiques: ouverts et fermés

**Définition 1.9.1.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $U \subseteq \mathbb{E}$. On dit que $U$ est* ***ouvert*** *(pour $N$) si, pour tout $v \in U$, il existe $r \in \mathbb{R}_{>0}$ tel que $B_N(v,r) \subseteq U$. Un ensemble $F \subseteq \mathbb{E}$ est* ***fermé*** *(pour $N$) si $\mathbb{E} \setminus F$ est ouvert.*

*Plus généralement, soit $S \subseteq \mathbb{E}$ une partie de l'espace vectoriel normé $(\mathbb{E}, N)$. Un ensemble $U \subseteq S$ est dit un* ***ouvert relatif de $S$*** *(pour $N$) si, pour tout $v \in U$, il existe $r \in \mathbb{R}_{>0}$ tel que $B_N(v,r) \cap S \subseteq U$. Un ensemble $F \subseteq S$ est un* ***fermé relatif de $S$*** *(pour $N$) si $S \setminus F$ est un ouvert relatif de $S$. On dira souvent* ***ouvert de $S$*** *(resp.,* ***fermé de $S$****) au lieu de ouvert relatif de $S$ (resp., fermé relatif de $S$). C'est clair qu'un ensemble $T \subseteq \mathbb{E}$ est un ouvert relatif de $\mathbb{E}$ (resp., fermé relatif de $\mathbb{E}$) si et seulement si $T$ est ouvert (resp., fermé).*

**Exemple 1.9.2.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie quelconque. Alors les ensembles $S$ et $\emptyset$ sont des ouverts relatifs de $S$. Cela implique que $S$ et $\emptyset$ sont aussi des fermés relatifs de $S$.*

Le résultat suivant nous donne l'un des exemples les plus importants d'ensemble ouvert.

**Proposition 1.9.3.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé, $v \in \mathbb{E}$ et $r \in \mathbb{R}_{>0}$. Alors la boule ouverte $B_N(v,r)$ es un ensemble ouvert.*

***Preuve.*** Soit $w \in B_N(v,r)$. Il faut montrer qu'il existe $s \in \mathbb{R}_{>0}$ tel que $B_N(w,s) \subseteq B_N(v,r)$. Comme $w \in B_N(v,r)$, $N(w-v) < r$, *i.e.* $r - N(w-v) > 0$. On pose $s = r - N(w-v)$. On va montrer que $B_N(w,s) \subseteq B_N(v,r)$. Soit $u \in B_N(w,s)$, *i.e.* $N(u-w) < s = r - N(w-v)$. Cela équivaut à $N(u-w)+N(w-v) < r$. Or, $N(u-v) = N(u-w+w-v) \leq N(u-w)+N(w-v) < r$, *i.e.* $u \in B_N(v,r)$, comme on voulait démontrer. □

Avant de préciser quelques propriétés des ensembles ouverts, on rappelle que, étant donné une famille $\{S_i\}_{i \in I}$ de parties d'un ensemble $S$, la **réunion** (ou l'**union**) de la famille est

$$\bigcup_{i \in I} C_i = \{v \in S : \text{ il existe } i \in I \text{ tel que } v \in S_i\},$$

et l'**intersection** de la famille est donnée par

$$\bigcap_{i \in I} C_i = \{v \in S : v \in S_i \text{ pour tout } i \in I\}.$$

**Lemme 1.9.4.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie.*

*(i) Soit $\{U_i\}_{i \in I}$ une famille de parties ouvertes de $S$, indexé par un ensemble arbitraire $I$. Alors la réunion $\cup_{i \in I} U_i$ est un ouvert de $S$. En plus, si $I$ est fini, l'intersection $\cap_{i \in I} U_i$ est un ouvert de $S$.*

*(ii) Soit $\{F_i\}_{i \in I}$ une famille de parties fermées de $S$, indexé par un ensemble arbitraire $I$. Alors l'intersection $\cap_{i \in I} F_i$ est un fermé de $S$. En plus, si $I$ est fini, la réunion $\cup_{i \in I} F_i$ est un fermé de $S$.*

***Preuve.*** On va démontrer seulement les énoncés pour les ensembles ouverts de $S$, puisque les propriétés pour les ensembles fermés de $S$ est un corollaire. En effet, le cas des ensembles fermés suit des égalités standards des ensembles

$$S \setminus \Big(\bigcap_{i \in I} F_i\Big) = \bigcup_{i \in I} (S \setminus F_i) \text{ et } S \setminus \Big(\bigcup_{i \in I} F_i\Big) = \bigcap_{i \in I} (S \setminus F_i).$$

On va montrer d'abord que $\cup_{i \in I} U_i$ est un ouvert de $S$. Soit $v \in \cup_{i \in I} U_i$. Alors, il existe $i \in I$ tel que $v \in U_i$. Comme $U_i$ est ouvert, il existe $r > 0$ tel que $B_N(v,r) \cap S \subseteq U_i \subseteq \cup_{i \in I} U_i$. Cela nous dit que $\cup_{i \in I} U_i$ est un ouvert de $S$.

On suppose par ailleurs que $I$ est fini et soit $v \in \cap_{i \in I} U_i$. Alors, pour tout $i \in I$, $v \in U_i$. Comme $U_i$ est ouvert, il existe $r_i > 0$ tel que $B_N(v, r_i) \cap S \subseteq U_i$. Soit $r = \min(r_i : i \in I)$. Comme $I$ est fini, $r > 0$. En plus, $B_N(v,r) \subseteq B_N(v, r_i) \cap S \subseteq U_i$, pour tout $i \in I$. Cela nous dit que $B_N(v,r) \subseteq \cap_{i \in I} U_i$ et que $\cap_{i \in I} U_i$ est un ouvert de $S$. □

On a la caractérisation suivante d'ensembles ouverts et fermés relatifs.

**Proposition 1.9.5.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. Un ensemble $T \subseteq S$ est un ouvert relatif de $S$ (resp., un fermé relatif de $S$) si et seulement s'il existe un ouvert (resp., un fermé) $\hat{T} \subseteq \mathbb{E}$ tel que $T = \hat{T} \cap S$.*

***Preuve.*** On suppose que $T$ est un ouvert relatif de $S$. Par définition, pour tout $v \in T$, il existe $r_v \in \mathbb{R}_{>0}$ tel que $B_N(v, r_v) \cap S \subseteq T$. Soit $\hat{T} = \cup_{v \in T} B_N(v, r_v)$. Comme les boules ouvertes sont des ensembles ouverts, d'après la Proposition 1.9.3, et la réunion arbitraire d'ouverts est un ensemble ouvert, d'après le Lemme 1.9.4, $\hat{T}$ est un ouvert. Comme $T \subseteq \hat{T}$, c'est clair que $T \subseteq \hat{T} \cap S$. En outre,

$$\hat{T} \cap S = \left( \bigcup_{v \in T} B_N(v, r_v) \right) \cap S = \bigcup_{v \in T} \left( B_N(v, r_v) \cap S \right) \subseteq T,$$

puisque $B_N(v, r_v) \cap S \subseteq T$ pour tout $v \in T$. Cela nous dit que $T = \hat{T} \cap S$. Réciproquement, o suppose qu'il existe un ouvert $\hat{T} \subseteq \mathbb{E}$ tel que $T = \hat{T} \cap S$. Comme $\hat{T}$ est un ouvert, pour tout $v \in T$, il existe $r \in \mathbb{R}_{>0}$ tel que $B_N(v, r) \subseteq \hat{T}$, ce qui implique $B_N(v, r) \cap S \subseteq \hat{T} \cap S = T$, *i.e.* $T$ est un ouvert relatif de $S$.

L'énoncé pour les fermés est un corollaire de l'énoncé pour les ouverts. En effet, $T$ est un fermé relatif de $S$ si et seulement si $S \setminus T$ est un ouvert relatif de $S$. Cela équivaut à l'existence d'un ensemble ouvert $\hat{U} \subseteq \mathbb{E}$ tel que $\hat{U} \cap S = S \setminus T$, qui est équivalent à l'existence d'un ensemble fermé $\hat{T} \subseteq \mathbb{E}$ tel que $(\mathbb{E} \setminus \hat{T}) \cap S = S \setminus T$. En prenant le complémentaire, la dernière égalité équivaut à $\hat{T} \cap S = T$, ce qui montre le résultat. □

D'après le résultat précédent, on voit que l'étude des propriétés d'ensembles ouvertes et fermés relatifs suit généralement de l'étude des propriétés d'ensembles ouvertes et fermés, donc on va se concentrer souvent à considérer la dernière, sans que cela représente une perte de généralité.

**Définition 1.9.6.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. On définit l'**intérieur** $S^\circ$ de $S$ (pour $N$) par*

$$S^\circ = \bigcup_{\substack{U \text{ ouvert}, \\ U \subseteq S}} U.$$

*Par la Proposition 1.9.4, $S^\circ$ est un ensemble ouvert. En outre, on définit l'**adhérence** $\bar{S}$ de $S$ (pour $N$) par*

$$\bar{S} = \bigcap_{\substack{F \text{ fermé}, \\ S \subseteq F}} F.$$

*La Proposition 1.9.4 nous dit que $\bar{S}$ est un ensemble fermé. On définit la **frontière** $\partial S$ de $S$ (pour $N$) par $\partial S = \bar{S} \setminus S^\circ$.*

**Remarque 1.9.7.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. Noter que $S^\circ \subseteq S \subseteq \bar{S}$. En outre, c'est clair que $S$ est ouvert (resp., fermé) si et seulement si $S = S^\circ$ (resp., $S = \bar{S}$). En particulier, $S$ est fermé si et seulement si $\partial S \subseteq S$.*

On a la caractérisation suivante d'intérieur d'une partie.

**Proposition 1.9.8.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. Alors,*

$$S^\circ = \{x \in \mathbb{E} : \textit{ il existe } r > 0 \textit{ tel que } B_N(x, r) \subseteq S\}.$$

***Preuve.*** On définit

$$S^i = \{x \in \mathbb{E} : \text{ il existe } r > 0 \text{ tel que } B_N(x, r) \subseteq S\}.$$

Il faut montrer que $S^\circ = S^i$.

D'abord, on voit bien que, si $x \in S^i$, il existe $r > 0$ tel que $x \in B_N(x, r) \subseteq S$, ce qui implique que $x \in S$. En conséquence, $S^i \subseteq S$. En plus, on affirme que $S^i$ est un ouvert. En effet, si $x \in S^i$, il existe $r > 0$ tel que $B_N(x, r) \subseteq S$. Comme $B_N(x, v)$ est ouvert (voir la Proposition 1.9.3), pour tout $y \in B_N(x, v)$ il existe $s > 0$ tel que $B_N(y, s) \subseteq B_N(x, r) \subseteq S$, *i.e.* $y \in S^i$. Comme $S^i \subseteq S$ et $S^i$ est ouvert, $S^i \subseteq S^\circ$.

Finalement, si $x \in S^\circ$, il existe $r > 0$ tel que $x \in B_N(x, r) \subseteq S^\circ$, puisque $S^\circ$ est ouvert. L'inclusion $S^\circ \subseteq S$ nous dit alors que $B_N(x, r) \subseteq S$, ce qui implique que $x \in S^i$. En conséquence, $S^\circ \subseteq S^i$, comme on voulait démontrer. □

**Définition 1.9.9.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. On dit que $S$ est* ***borné*** *(pour $N$) s'il existe $C \in \mathbb{R}_{>0}$ tel que $N(v) < C$, pour tout $v \in S$.*

Soit $\mathbb{E}$ un espace vectoriel. On rappelle qu'une **propriété** sur les parties de $\mathbb{E}$ est par définition une partie de $\mathscr{P}(\mathbb{E})$. Soit

$$\mathcal{P} : \mathrm{Norm}(\mathbb{E}) \to \mathscr{P}\big(\mathscr{P}(\mathbb{E})\big)$$

une application. De façon intuitive, pour une norme $N \in \mathrm{Norm}(\mathbb{E})$ sur $\mathbb{E}$, $\mathcal{P}(N)$ dénote une propriété sur les parties de $\mathbb{E}$ qui est définie en termes de $N$. Par exemple, on peut considérer les applications

$$\mathcal{P}_{\mathrm{o}}, \mathcal{P}_{\mathrm{f}}, \mathcal{P}_{\mathrm{b}} : \mathrm{Norm}(\mathbb{E}) \to \mathscr{P}\big(\mathscr{P}(\mathbb{E})\big)$$

données par

$$\mathcal{P}_{\mathrm{o}}(N) = \big\{U \in \mathscr{P}(\mathbb{E}) : U \text{ est un ensemble ouvert de } (\mathbb{E}, N)\big\},$$

$$\mathcal{P}_{\mathrm{f}}(N) = \big\{F \in \mathscr{P}(\mathbb{E}) : F \text{ est un ensemble fermé de } (\mathbb{E}, N)\big\}$$

et

$$\mathcal{P}_{\mathrm{b}}(N) = \big\{B \in \mathscr{P}(\mathbb{E}) : B \text{ est un ensemble borné de } (\mathbb{E}, N)\big\},$$

pour tout $N \in \mathrm{Norm}(\mathbb{E})$. On dira qu'une application $\mathcal{P}$ comme ci-dessus est **fortement topologique** si $\mathcal{P}(N) = \mathcal{P}(N')$ pour toute paire de normes $N, N' \in \mathrm{Norm}(\mathbb{E})$ équivalentes. Le résultat suivant nous dit que les notions $\mathcal{P}_{\mathrm{o}}$ d'être ouvert, $\mathcal{P}_{\mathrm{f}}$ d'être fermé et $\mathcal{P}_{\mathrm{b}}$ d'être borné sont fortement topologiques.

**Proposition 1.9.10.** *Soit $\mathbb{E}$ un espace vectoriel et soit $S \subseteq \mathbb{E}$ une partie. Soient $N$ et $N'$ deux normes équivalentes sur $\mathbb{E}$. Alors, $S$ est ouvert (resp., fermé, borné) pour $N$ si et seulement si $S$ est ouvert (resp., fermé, borné) pour $N'$.*

***Preuve.*** On suppose qu'il existe $C, C' \in \mathbb{R}_{>0}$ tels que

$$CN(v) \leq N'(v) \leq C'N(v),$$

pour tout $v \in \mathbb{E}$. C'est clair que pour chaque propriété, il suffit de montrer que si elle est vérifiée pour $N$ alors elle est aussi vérifiée pour $N'$.

On suppose que $S$ est ouvert pour $N$. Alors, pour tout $w \in S$, il existe $r \in \mathbb{R}_{>0}$ tel que $B_N(w, r) \subseteq S$. On affirme que $B_{N'}(w, Cr) \subseteq S$. En effet, soit $u \in B_{N'}(w, Cr)$, *i.e.* $N'(u - w) < Cr$, alors $CN(u - w) \leq N'(u - w) < Cr$, ce qui nous dit que $N(u - w) < r$, *i.e.* $u \in B_N(w, r) \subseteq S$, comme on voulait démontrer. Cela implique que $S$ est ouvert pour $N'$. Si $S$ est fermé pour $N$, *i.e.* $\mathbb{E} \setminus S$ est ouvert pour $N$, alors l'argument précédent nous dit que $\mathbb{E} \setminus S$ est ouvert pour $N'$, *i.e.* $S$ est fermé pour $N'$.

Finalement, on suppose que $S$ est borné pour $N$. Alors, il existe $\bar{C} \in \mathbb{R}_{>0}$ tel que $N(v) \leq \bar{C}$, pour tout $v \in S$. Alors $N'(v) \leq C'N(v) \leq C'\bar{C}$, pour tout $v \in S$, ce qui dit que $S$ est borné pour $N'$. □

## 1.10 Lien entre la convergence de suites et la topologie

Le résultat suivant met en évidence le lien très fort qui existe entre la notion de convergence de suites et les notions topologiques introduites dans la Section 1.9.

**Proposition 1.10.1.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. Alors, $S$ est fermé si et seulement la limite de toute suite convergente $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ à valeurs dans $S$ apparient à $S$.*

***Preuve.*** On suppose d'abord que $S$ est fermé et soit $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ une suite convergente, de limite $v \in \mathbb{E}$. Il faut montrer que sa limite $v$ apparient à $S$. Supposons que $v \in \mathbb{E} \setminus S$. Comme $S$ est fermé, alors $\mathbb{E} \setminus S$ est ouvert, ce qui nous dit qu'il existe $r \in \mathbb{R}_{>0}$ tel que $B_N(v, r) \subseteq \mathbb{E} \setminus S$. Comme $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $v$, il existe $n_0 \in \mathbb{N}$ tel que $N(v_n - v) < r$ pour tout entier $n \geq n_0$. En conséquence, $v_n \in B_N(v, r) \subseteq \mathbb{E} \setminus S$ pour tout entier $n \geq n_0$. Comme $v_n \in S$ pour tout $n \in \mathbb{N}$ on a trouvé un absurde.

Réciproquement, on suppose que la limite de toute suite convergente $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ à valeurs dans $S$ apparient à $S$. On va montrer que $S$ est fermé. Supposons que ce n'est pas le cas. Cela veut dire qu'il existe $v \in \mathbb{E} \setminus S$ tel que, pour tout $r \in \mathbb{R}_{>0}$, il existe $w_r \in S$ tel que $N(w_r - v) < r$. En particulier, pour tout $n \in \mathbb{N}$ il existe $v_n = w_{1/n}$ tel que $v_n \in S$ tel que $N(w_n - v) < 1/n$. Cela nous dit que la suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ à valeurs dans $S$ converge vers un élément $v \in \mathbb{E} \setminus S$, ce qui est absurde. □

**Exemple 1.10.2.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et $v \in \mathbb{E}$. Alors $S = \{v\}$ est un ensemble fermé. En effet, toute suite à valeurs dans $S$ est constante, avec limite $v \in S$. Comme l'union fini de fermés est fermée, tout ensemble fini est fermé.*

Une conséquence assez directe de la caractérisation précédente d'ensemble fermé est l'exemple suivant.

**Exemple 1.10.3.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé, $v \in \mathbb{E}$ et $r \in \mathbb{R}_{>0}$. Alors la boule fermée $\bar{B}_N(v, r)$ est un ensemble fermé. En effet, on affirme que si $(w_n)_{n\in\mathbb{N}} \in \bar{B}_N(v, r)^{\mathbb{N}}$ une suite convergente de limite $w \in \mathbb{E}$, alors $w \in \bar{B}_N(v, r)$. Comme $w_n \in \bar{B}_N(v, r)$, $N(w_n - v) \leq r$, ce qui nous dit que*

$$N(w - v) \leq N(w - w_n) + N(w_n - v) \leq N(w - w_n) + r,$$

*pour tout $n \in \mathbb{N}$. Si l'on prend la limite quand $n$ tend vers $+\infty$, on trouve que $N(w - v) \leq r$, i.e. $w \in \bar{B}_N(v, r)$, comme on voulait montrer.*

Une autre conséquence du lien entre la notion de convergence et la notion de topologie est la suivante.

**Corollaire 1.10.4.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé complet et soit $\mathbb{F} \subseteq \mathbb{E}$ un sous-espace vectoriel. Alors $(\mathbb{F}, N|_{\mathbb{F}})$ est complet si et seulement si $\mathbb{F} \subseteq \mathbb{E}$ est fermé.*

***Preuve.*** On suppose que $(\mathbb{F}, N|_{\mathbb{F}})$ est complet, et soit $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ une suite convergente dans $(\mathbb{E}, N)$ avec limite $v \in \mathbb{E}$. Par la Proposition 1.10.1, il suffit de démontrer que $v \in \mathbb{F}$. Comme $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ est une suite convergente dans $(\mathbb{E}, N)$, elle de Cauchy dans $(\mathbb{E}, N)$, ce qui équivaut à dire qu'elle de Cauchy dans $(\mathbb{F}, N|_{\mathbb{F}})$. Comme $(\mathbb{F}, N|_{\mathbb{F}})$ est complet, il existe $w \in \mathbb{F}$ tel que $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ converge vers $w$ dans $(\mathbb{F}, N|_{\mathbb{F}})$, ce qui implique que $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ converge vers $w$ dans $(\mathbb{E}, N)$. Par l'unicité de la limite, $v = w$, ce qui nous dit que $v \in \mathbb{F}$.

Réciproquement, on suppose que $\mathbb{F} \subseteq \mathbb{E}$ est fermé. Soit $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ une suite de Cauchy dans $(\mathbb{F}, N|_{\mathbb{F}})$, ce qui équivaut à dire que $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ une suite de Cauchy dans $(\mathbb{E}, N)$. Il suffit de démontrer que $v \in \mathbb{F}$. Comme $(\mathbb{E}, N)$ est complet, il existe $v \in \mathbb{E}$ tel que $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ converge vers $v$ dans $(\mathbb{E}, N)$. Comme $\mathbb{F} \subseteq \mathbb{E}$ est fermé et $(v_n)_{n\in\mathbb{N}} \in (\mathbb{F}^{\mathbb{N}}$ converge vers $v$, la Proposition 1.10.1 nous dit que $v \in \mathbb{F}$. □

Une autre version de la relation entre la notion de convergence et la notion de topologie est la suivante.

**Proposition 1.10.5.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. Alors, $v \in \bar{S}$ si et seulement s'il existe une suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ qui converge vers $v$.*

***Preuve.*** Soit $\hat{S} \subseteq \mathbb{E}$ l'ensemble formé de toutes les limites (dans $\mathbb{E}$) des suites convergentes $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ à valeurs dans $S$. Il faut démontrer que $\hat{S} = \bar{S}$.

On remarque d'abord que $S \subseteq \hat{S}$. En effet, pour tout $v \in S$, la suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ à valeurs dans $S$ donnée par $v_n = v$, pour tout $n \in \mathbb{N}$, est convergente et sa limite est $v$, *i.e.* $v \in \hat{S}$. On va maintenant montrer que $\hat{S}$ est un ensemble fermé. D'après la Proposition 1.10.1, il suffit de montrer que la limite $\hat{v}$ de toute suite convergente $(\hat{v}_n)_{n\in\mathbb{N}} \in \hat{S}^{\mathbb{N}}$ appartient à $\hat{S}$. Par définition, $\hat{v}_n$ est la limite d'une suite convergente $(v_{n,m})_{m\in\mathbb{N}} \in S^{\mathbb{N}}$. Cela nous dit que, pour tout $n \in \mathbb{N}$, il existe $i_n \in \mathbb{N}$ tel que $N(\hat{v}_n - v_{n,m}) < 1/n$ pour tout entier $m \geq i_n$. Soit $(u_n)_{n\in\mathbb{N}}$ la suite donnée par $u_n = v_{n,i_n}$, pour tout $n \in \mathbb{N}$. C'est clair que $(u_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$. On affirme que $(u_n)_{n\in\mathbb{N}}$ converge vers $\hat{v}$. En effet, on remarque d'abord que, pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que

$$N(\hat{v} - \hat{v}_n) < \frac{\varepsilon}{2}$$

pour tout entier $n \geq n_0$. Sans perte de généralité, on suppose en plus que $n_0 > 2/\varepsilon$. Alors,

$$\begin{aligned} N(\hat{v} - u_n) &= N(\hat{v} - v_{n,i_n}) = N(\hat{v} - \hat{v}_n + \hat{v}_n - v_{n,i_n}) \leq N(\hat{v} - \hat{v}_n) + N(\hat{v}_n - v_{n,i_n}) \\ &< \frac{\varepsilon}{2} + \frac{1}{n} \leq \frac{\varepsilon}{2} + \frac{1}{n_0} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

pour tout entier $n \geq n_0$, comme on voulait démontrer. Comme $(u_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $\hat{v}$, $\hat{v}$ appartient à $\hat{S}$, par définition de $\hat{S}$, et $\hat{S}$ est donc fermé. Or, d'après l'inclusion $S \subseteq \hat{S}$ et le fait que $\hat{S}$ soit fermé, on conclut que $\bar{S} \subseteq \hat{S}$, par définition d'adhérence.

Finalement, on va démontrer que $\hat{S} \subseteq \bar{S}$. Pour cela, il faut montrer que si $F \subseteq \mathbb{E}$ est un ensemble fermé et $S \subseteq F$, alors $\hat{S} \subseteq F$. Or, comme $F$ est fermé, d'après la Proposition 1.10.1, la limite de toute suite convergente à valeurs dans $F$ appartient à $F$, ce qui nous dit *a fortiori* que la limite de toute suite convergente à valeurs dans $S \subseteq F$ appartient à $F$. Comme $\hat{S}$ est l'ensemble de limites de suites convergentes à valeurs dans $S$, on conclut que $\hat{S} \subseteq F$, comme on voulait démontrer. □

Une version plus topologique du résultat précédent est le suivant. On laisse la preuve au lecteur/à la lectrice.

**Proposition 1.10.6.** *Soit* $(\mathbb{E}, N)$ *un espace vectoriel normé et soit* $S \subseteq \mathbb{E}$ *une partie. Alors,*

$$\bar{S} = \{x \in \mathbb{E} : \text{ pour tout } r > 0, B_N(x,r) \cap S \neq \emptyset\}.$$

## 1.11 Compacité (séquentielle)

Soit $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ une suite à valeurs dans un ensemble $S$. On rappelle qu'une **sous-suite** de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$, aussi appelée une **suite extraite**, est une suite $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ donnée par $w_n = v_{\phi(n)}$ pour tout $n \in \mathbb{N}$, où $\phi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective. On remarque que l'application $\phi$ est aussi une donnée dans la définition de sous-suite. On mentionne le fait suivante, que l'on utilisera souvent dans la suite: toute sous-suite d'une suite convergente dans un espace vectoriel normé est aussi convergente et a la même limite. On laisse la preuve au lecteur/à la lectrice.

**Définition 1.11.1.** *Soit* $(\mathbb{E}, N)$ *un espace vectoriel normé et soit* $S \subseteq \mathbb{E}$ *une partie. On dit que* $S$ *est **séquentiellement compact** (pour* $N$*) si toute suite* $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ *à valeurs dans* $S$ *admet une sous-suite convergente dont la limite appartient à* $S$*. Comme la compacité séquentielle est la seule propriété de compacité que l'on va considérer dans ce cours, à exception de la Section 1.16, on va omettre désormais l'adjectif pour simplicité, sauf dans la Section 1.16, où l'on va comparer la notion de compacité séquentielle avec une notion de compacité obtenue de la propriété de Borel-Lebesgue.*

**Exemple 1.11.2.** *Le Théorème de Bolzano-Weierstrass (pour* $\mathbb{R}$*) affirme précisément que tout intervalle fermé et borné de* $\mathbb{R}$ *est compact.*

On peut donner une version plus topologique équivalente de la définition de compacité séquentielle, basé sur la notion suivante.

**Définition 1.11.3.** *Soit* $(\mathbb{E}, N)$ *un espaces vectoriel normés et soit* $S \subseteq \mathbb{E}$ *une partie infinie. On dit que* $v \in \mathbb{E}$ *est un **point d'accumulation** de* $S$ *si, pour tout* $r \in \mathbb{R}_{>0}$*, il existe* $w \in S \setminus \{v\}$ *tel que* $N(v - w) \leq r$*. De façon équivalente,* $v \in \mathbb{E}$ *est un point d'accumulation de* $S$ *si et seulement si* $v \in \overline{S \setminus \{v\}}$*. C'est aussi clair que* $v \in \mathbb{E}$ *est un point d'accumulation de* $S$ *si et seulement si, pour tout* $r \in \mathbb{R}_{>0}$*, l'ensemble* $\{w \in S \setminus \{v\} : N(v-w) \leq r\}$ *est infini.*

**Proposition 1.11.4.** *Soit* $(\mathbb{E}, N)$ *un espaces vectoriel normés et soit* $S \subseteq \mathbb{E}$ *une partie. Alors,* $S$ *est compact si et seulement si toute partie infinie* $T \subseteq S$ *a un point d'accumulation dans* $S$*.*

***Preuve.*** On suppose que $S$ est compact. On va montrer que toute partie infinie $T \subseteq S$ a un point d'accumulation dans $S$. Comme $T$ est infinie, il existe une application injective $f : \mathbb{N} \to T$. On considère la suite $(v_n)_{n\in\mathbb{N}} \in T^{\mathbb{N}} \subseteq S^{\mathbb{N}}$ donnée par $v_n = f(n)$. D'après la compacité de $S$, $(v_n)_{n\in\mathbb{N}}$ a une sous-suite convergente $(w_n)_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ avec limite $v \in S$. Par définition, $v \in S$ est un point d'accumulation de $T$, ce qui montre que $T$ a un point d'accumulation dans $S$. Réciproquement, on suppose que toute partie infinie $T \subseteq S$ a un point d'accumulation dans $S$. Soit $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ une suite. On va montrer qu'elle possède une sous-suite convergente, dont la limite appartient à $S$. Soit $T = \{v_n : n \in \mathbb{N}\} \subseteq S$ l'ensemble sous-jacent de la suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$. Si $T$ est fini, il existe $v \in S$ tel que $\{n \in \mathbb{N} : v_n = v\}$ est infini. Soit $\phi : \mathbb{N} \to \{n \in \mathbb{N} : v_n = v\}$ la seule fonction bijective et croissante. Alors $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ est une sous-suite constante de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ et *a fortiori* convergente, avec limite $v \in S$. Si $T$ est infini, par hypothèse $T$ possède un point d'accumulation $v \in S$. On construit par récurrence une sous-suite de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ qui converge vers $v$. On prend $w_0 = v_0$ et on suppose que l'on a construit entiers non négatifs $\phi(0) < \cdots < \phi(n)$ tels $N(v_{\phi(i)} - v) \leq 1/2^i$, pour tout $i \in [\![1, n]\!]$. Par définition de point d'accumulation, l'ensemble $\{w \in T \setminus \{v\} : N(w - v) \leq 2^{-(n+1)}\}$ est infini. Il existe alors $m > \phi(n)$ tel que $N(v_m - v) \leq 2^{-(n+1)}$. On pose $\phi(n + 1) = m$. Par récurrence, on a un application $\phi : \mathbb{N} \to \mathbb{N}$ croissante et injective telle que $N(v_{\phi(n)} - v) \leq 1/2^n$, pour tout $n \in \mathbb{N}$, ce qui nous dit que la sous-suite $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $v$, comme on voulait démontrer. $\square$

**Définition 1.11.5.** *Soit $(\mathbb{E}, N)$ un espaces vectoriel normés et soit $S \subseteq \mathbb{E}$ une partie. On dit que $v \in \mathbb{E}$ est un **point isolé** de $S$ s'il existe $r \in \mathbb{R}_{>0}$ tel que $B_N(v, r) \cap S = \{v\}$. La partie $S$ est dite **discrète** si pour tout $s \in S$, $s$ est un point isolé de $S$.*

Le résultat suivant nous dit que la notion de compacité d'une partie d'un espace vectoriel normé est fortement topologique.

**Proposition 1.11.6.** *Soit $\mathbb{E}$ un espace vectoriel et soit $S \subseteq \mathbb{E}$ une partie. Soient $N$ et $N'$ deux normes équivalentes sur $\mathbb{E}$. Alors, $S$ est compact pour $N$ si et seulement si $S$ est compact pour $N'$.*

***Preuve.*** Il s'agit d'une conséquence directe de la Proposition 1.6.5. $\square$

**Proposition 1.11.7.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie compacte. Alors, $S$ est fermé et borné.*

***Preuve.*** On va d'abord démontrer que $S$ est fermé. D'après la Proposition 1.10.1, il suffit de montrer que la limite de toute suite convergente $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ appartient à $S$. Soit $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ une suite convergente de limite $v \in \mathbb{E}$. Comme $S$ est compact, il existe une sous-suite $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$, dont la limite $w$ appartient à $S$. Or, comme $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ est convergente, toute sous-suite est aussi convergente et elle possède la même limite. Cela implique que $v = w \in S$, comme on voulait démontrer.

Finalement, on va montrer que $S$ est borné. Si ce n'est pas le cas, pour tout $C > 0$, il existe une suite $v \in S$ tel que $N(v) \geq C$. On construit par récurrence la suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ suivante. On choisit n'importe quel élément $v_0 \in S$ et on suppose que l'on a construit $v_0, \ldots, v_n \in S$ tels que $N(v_i) \geq N(v_{i-1}) + 1$, pour tout $i \in [\![1, n]\!]$, où $n \in \mathbb{N}$. On pose $C = N(v_n) + 1$. D'après l'hypothèse de l'absurde, il existe $v_{n+1} \in S$ tel que $N(v_n) \geq C = N(v_n) + 1$, ce qui nous donne la récurrence. Cette suite satisfait clairement que $N(v_{n+m}) - N(v_n) \geq m$, pour tous $n, m \in \mathbb{N}$. Para ailleurs, on voit bien que

$$1 \leq m \leq N(v_{n+m}) - N(v_n) = \big|N(v_{n+m}) - N(v_n)\big| \leq N(v_{n+m} - v_n),$$

pour tous $n \in \mathbb{N}$ et $m \in \mathbb{Z}_{>0}$, où la dernière inégalité suit de (N3). Si $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ est une sous-suite de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ donnée par $w_n = v_{\phi(n)}$ pour tout $n \in \mathbb{N}$, où $\phi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective, alors

$$1 \leq \phi(m + n) - \phi(n) \leq N(w_{n+m}) - N(w_n) = \big|N(w_{n+m}) - N(w_n)\big| \leq N(w_{n+m} - w_n),$$

pour tous $n \in \mathbb{N}$ et $m \in \mathbb{Z}_{>0}$. En conséquence, $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ n'est pas de Cauchy, et *a fortiori* elle n'est pas convergente (voir la Proposition 1.6.4). Cela est un absurde, parce que $S$ est compact. $\square$

**Proposition 1.11.8.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie compacte. Soit $T \subseteq S$ un ensemble fermé. Alors, $T$ est compact.*

***Preuve.*** Soit $(v_n)_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ une suite à valeurs dans $T$. Comme elle appartient aussi à $S$, il existe une sous-suite convergente $(w_n)_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ dont la limite $w$ appartient à $S$. Comme $T$ est fermé, d'après la Proposition 1.10.1, $w \in T$, comme on voulait démontrer. $\square$

**Proposition 1.11.9.** *Soient $\mathbb{E}_1, \ldots, \mathbb{E}_n$ des espaces vectoriels munis des normes $N_1, \ldots, N_n$, respectivement. On muni $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$ de la norme $N_{\max}$ (voir la Proposition 1.2.4). Pour tout $i \in [\![1, n]\!]$, soit $S_i \subseteq \mathbb{E}_i$ une partie compacte (pour la norme $N_i$). Alors $S = S_1 \times \cdots \times S_n$ est une partie compacte de $\mathbb{E}$ (pour la norme $N_{\max}$).*

***Preuve.*** Soit $(x_n) \in S^{\mathbb{N}}$ une suite. Pour tout $n \in \mathbb{N}$, on écrit $x_n = x_{n,i} \in \mathbb{E}_i$ avec $i \in [\![1,n]\!]$. Alors $(x_{n,i})_{n\in\mathbb{N}} \in S_i^{\mathbb{N}}$ est une suite pour tout $i \in [\![1,n]\!]$. On va construire une sous-suite convergente de $(x_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$, dont la limite appartient à $S$. On procède par récurrence sur $i \in [\![1,n]\!]$. On suppose que l'on a une application injective et croissante $\phi_{i-1} : \mathbb{N} \to \mathbb{N}$ telle que la sous-suite $(x_{\phi_{i-1}(n),j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ de $(x_{n,j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ converge vers un élément $\bar{x}_j$ de $S_j$, pour tout $j \in [\![1,i-1]\!]$, où $i \in [\![1,n]\!]$. On va montrer que cette affirmation est aussi vérifié pour $i$. Pour cela, on considère la suite $(x_{\phi_{i-1}(n),i})_{n\in\mathbb{N}} \in S_i^{\mathbb{N}}$. Comme $S_i$ est compact, il existe une application injective et croissante $\psi : \mathbb{N} \to \mathbb{N}$ telle que $(x_{\phi_{i-1}(\psi(n)),i})_{n\in\mathbb{N}} \in S_i^{\mathbb{N}}$ converge vers un élément $\bar{x}_i$ de $S_i$. On pose $\phi_i : \mathbb{N} \to \mathbb{N}$ donnée par $\phi_i = \phi_{i-1} \circ \psi$. Comme la composition de fonctions injectives et croissante est aussi injective et croissante, ces propriétés sont vérifiées par $\phi_i$. Si $j \in [\![1,i-1]\!]$, $(x_{\phi_i(n),j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ est une sous-suite de la suite convergente $(x_{\phi_{i-1}(n),j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$, ce qui implique que $(x_{\phi_i(n),j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ converge vers l'élément $\bar{x}_j$ de $S_j$. On conclut que la sous-suite $(x_{\phi_i(n),j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ de $(x_{n,j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ converge vers un élément $\bar{x}_j$ de $S_j$, pour tout $j \in [\![1,i]\!]$. Finalement, on considère $\phi = \phi_n$ et la sous-suite $(x_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(x_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$. Par construction, $(x_{\phi(n),j})_{n\in\mathbb{N}} \in S_j^{\mathbb{N}}$ converge vers un élément $\bar{x}_j$ de $S_j$, pour tout $j \in [\![1,n]\!]$. La Proposition 1.6.7 nous dit que $(x_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n) \in S_1 \times \cdots \times S_n = S$, ce qui montre que $S$ est compact. □

On a la réciproque de la Proposition 1.11.7 dans le cas de la norme infini.

**Proposition 1.11.10.** *Soit $\mathbb{E}$ un espace vectoriel de dimension finie $n$ et soit $S \subseteq \mathbb{E}$ une partie. Soit $\mathscr{B} = \{v_1, \dots, v_n\} \subseteq \mathbb{E}$ une base (ordonnée) de $\mathbb{E}$. On considère l'application $T : \mathbb{E} \to \mathbb{R}^n$ qui associe à chaque $v = \sum_{i=1}^n c_i v_i$ (où $c_i \in \mathbb{R}$ pour tout $i \in [\![1,n]\!]$) le $n$-uplet $T(v) = (c_1, \dots, c_n)$ et on définit la norme $N_\infty : \mathbb{E} \to \mathbb{R}_{\geq 0}$ via $N_\infty = \| \ \|_\infty \circ T$ (voir la Proposition 1.2.1). Si l'ensemble $S$ est fermé et borné pour la norme $N_\infty$, alors il est compact.*

***Preuve.*** On remarque la norme $N_\infty$ coïncide avec la norme $N_{\max}$ du produit des espaces vectoriels normés $\mathbb{E}_1, \dots, \mathbb{E}_n$, où $\mathbb{E}_i$ est le sous-espace vectoriel de $\mathbb{E}$ engendré par $v_i$ muni de la norme $N_i(\lambda \cdot v_i) = |\lambda|$.

Soit $S$ une ensemble fermé et borné de $\mathbb{E}$ pour la norme $N_\infty$. Comme $S$ est borné, il existe $C > 0$ tel que $S \subseteq B_N(\mathbf{0}_{\mathbb{E}}, C)$. D'après les commentaires dans le paragraphe précédent,

$$\bar{B}_N(\mathbf{0}_{\mathbb{E}}, C) = \prod_{i=1}^n \bar{B}_{N_i}(\mathbf{0}_{\mathbb{E}_i}, C).$$

D'après le Théorème de Bolzano-Weierstrass, tout intervalle fermé et borné de $\mathbb{R}$ est un compact, pour la norme donnée par la valeur absolue (voir l'Exemple 1.11.2). La Proposition 1.11.6 nous dit aussi que $\bar{B}_{N_i}(\mathbf{0}_{\mathbb{E}_i}, C)$ est un compact de $\mathbb{E}_i$ pour la norme $N_i$, pour tout $i \in [\![1,n]\!]$. D'après la Proposition 1.11.9, $\bar{B}_N(\mathbf{0}_{\mathbb{E}}, C)$ est un compact de $\mathbb{E}$ pour la norme $N_\infty$. Comme $S$ est un ensemble fermé et il est inclus dans le compact $\bar{B}_N(\mathbf{0}_{\mathbb{E}}, C)$, la Proposition 1.11.8 implique que $S$ est compact, comme on voulait démontrer. □

**Remarque 1.11.11.** *Les Propositions 1.11.6, 1.11.7 et 1.11.10, et le Théorème 1.8.1 impliquent qu'un ensemble d'un espace vectoriel normé de dimension finie est compact si et seulement s'il est fermé et borné. Comme on veut donner une preuve différente du Théorème 1.8.1, basée sur la notion de compacité, on ne va pas énoncer ce résultat pour l'instant.*

*On remarque aussi que l'équivalence entre les notions de compacité d'un côté et celle d'ensemble fermé et borné d'un autre côté n'est pas valable pour des espaces vectoriels normés de dimension infinie.*

## 1.12 Applications continues

Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $S \subseteq \mathbb{E}$. Soit $f : S \to \mathbb{F}$ une application et $v \in S$. On rappelle que $f(x)$ **converge** vers l'élément $w \in \mathbb{F}$ quand $x \in S$ tend vers $v$ (pour $N$ et $N'$) si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $x \in S$ et $N(x - v) < \delta$ impliquent que $N'(w - f(v)) < \varepsilon$.

**Proposition 1.12.1.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $S \subseteq \mathbb{E}$. Soit $f : S \to \mathbb{F}$ une application et $v \in S$. Si $f(x)$ converge vers $w \in \mathbb{F}$ et $w' \in \mathbb{F}$ quand $x \in S$ tend vers $v$ (pour $N$ et $N'$), alors $w = w'$. On appelle cet élément la **limite** de $f(x)$ quand $x \in S$ tend vers $v$ (pour $N$ et $N'$) et on écrit*

$$\lim_{\substack{x \to v \\ x \in S}} f(x) = w.$$

***Preuve.*** On voit bien que

$$N'(w - w') = N'\big(w - f(x) + f(x) - w'\big) \leq N'\big(w - f(x)\big) + N'\big(f(x) - w'\big),$$

pour tout $x \in S$, d'après (N3) pour $N'$. On affirme que $N'(w - w') = 0$. Si ce n'est pas le cas, alors $N'(w - w') > 0$ et on prend $\varepsilon = N'(w - w')/4$. Soient $\delta \in \mathbb{R}_{>0}$ et $\delta' \in \mathbb{R}_{>0}$ les éléments correspondants dans le définition de convergence de $f(x)$ quand $x \in S$ tend vers $v$. Alors, les conditions $x \in S$ et $N(x - v) < \min(\delta, \delta')$ impliquent que $N'(f(x) - w) < \varepsilon$ et $N'(w' - f(x)) < \varepsilon$. En conséquence,

$$N'(w - w') = N'\big(w - f(x) + f(x) - w'\big) \leq N'\big(w - f(x)\big) + N'\big(f(x) - w'\big) < \varepsilon + \varepsilon = \frac{N'(w - w')}{2},$$

ce qui est absurde si $N'(w - w') > 0$. En conséquence, $N'(w - w') = 0$. D'après (N1) pour $N'$ on conclut que $w = w'$, comme on voulait démontrer. □

Si le domaine de définition $S$ de $f : S \to \mathbb{F}$ est un ouvert de $\mathbb{E}$, on l'omettra dans la terminologie et la notation de limite. En particulier, dans ce cas on dira que $f(x)$ converge vers $w \in \mathbb{F}$ quand $x$ tend vers $v$ (pour $N$ et $N'$) et on écrira

$$\lim_{x \to v} f(x) = w.$$

**Définition 1.12.2.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $S \subseteq \mathbb{E}$. Soit $f : S \to \mathbb{F}$ une application. Étant donné $v \in S$, on dit que $f$ est **continue** en $v$ (pour $N$ et $N'$) si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $w \in S$ et $N(w - v) < \delta$ impliquent que $N'(f(w) - f(v)) < \varepsilon$. De façon équivalente, $f$ est continue en $v$ si et seulement si la limite de $f(x)$ quand $x \in S$ tend vers $v$ existe et elle coïncide avec $f(v)$. On dit que $f$ est **continue** (sur $S$, pour $N$ et $N'$) si elle est continue en $v$, pour tout $v \in S$.*

*En outre, l'application $f$ est dite **uniformément continue** (pour $N$ et $N'$) si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que les conditions $v, w \in S$ et $N(w - v) < \delta$ impliquent que $N'(f(w) - f(v)) < \varepsilon$.*

**Remarque 1.12.3.** *C'est clair qu'une application uniformément continue $f : S \to \mathbb{F}$ est aussi continue sur $S$.*

Le résultat suivant nous dit que les notions de continuité et de continuité uniforme d'applications sont fortement topologiques. On laisse au lecteur/à la lectrice la formulation précise de la notion de propriété fortement topologique pour des applications.

**Proposition 1.12.4.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $S \subseteq \mathbb{E}$. Soit $f : S \to \mathbb{F}$ une application et $v \in S$. On considère $\bar{N} : \mathbb{E} \to \mathbb{R}_{\geq 0}$ et $\bar{N}' : \mathbb{F} \to \mathbb{R}_{\geq 0}$ deux normes équivalentes à $N$ et à $N'$, respectivement. Alors, $f$ est continue en $v$ (resp., uniformément continue) pour $N$ et $N'$ si et seulement si $f$ est continue en $v$ (resp., uniformément continue) pour $\bar{N}$ et $\bar{N}'$.*

***Preuve.*** On suppose qu'il existe $C, C', \bar{C}, \bar{C}' \in \mathbb{R}_{>0}$ tels que

$$CN(v) \leq \bar{N}(v) \leq \bar{C}N(v) \text{ et } C'N'(w) \leq \bar{N}'(w) \leq \bar{C}'N'(w),$$

pour tout $v \in \mathbb{E}$ et $w \in \mathbb{F}$. C'est clair que pour chaque propriété, il suffit de montrer que, si elle est vérifiée pour $N$ et $N'$, alors elle est aussi vérifiée pour $\bar{N}$ et $\bar{N}'$. On fera uniquement la preuve pour le cas de la continuité en $v$, la preuve de la continuité uniforme étant donné *mutatis mutandi*.

On suppose que $f$ est continue en $v \in S$ pour $N$ et $N'$. Alors, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $v' \in S$ et $N(v' - v) < \delta$ impliquent que $N'(f(v') - f(v)) < \varepsilon/\bar{C}'$. On fixe $\varepsilon > 0$ et on prend $\delta > 0$ comme ci-dessus. On pose $\bar{\delta} = C\delta$. Alors, pour tout $v' \in S$, la condition $\bar{N}(v' - v) < \bar{\delta}$ implique que $N(v' - v) < \delta$. Par continuité de $f$ en $v$ et le choix des paramètres précédents, on voit que $N'(f(v') - f(v)) < \varepsilon/\bar{C}'$, ce qui implique que $\bar{N}'(f(w) - f(v)) < \varepsilon$, comme on voulait démontrer. □

**Exemple 1.12.5.** *Soit* $(\mathbb{E}, N)$ *un espace vectoriel normé. On considère les applications* $\mathpzc{s} : \mathbb{E} \times \mathbb{E} \to \mathbb{E}$ *et* $\mathpzc{p} : \mathbb{R} \times \mathbb{E} \to \mathbb{E}$ *données par*

$$\mathpzc{s}(x, y) = x + y \text{ et } \mathpzc{p}(\lambda, x) = \lambda \cdot x,$$

*pour tous* $x, y \in \mathbb{E}$ *et* $\lambda \in \mathbb{R}$. *On muni* $\mathbb{E} \times \mathbb{E}$ *et* $\mathbb{R} \times \mathbb{E}$ *avec les normes* $N_{\max}$ *induites par* $N$*, et par la valeur absolue* $|\ |$ *et* $N$*, respectivement (voir la Proposition 1.2.4). C'est facile à vérifier que* $\mathpzc{s}$ *est uniformément continue. En effet, la continuité uniforme de* $\mathpzc{s}$ *est une conséquence de l'équivalence des normes* $N_{\max}$ *et* $N_{\text{sum}}$ *(voir la Proposition 1.5.4). Plus généralement, pour tout* $n \in \mathbb{Z}_{>0}$ *on pourrait aussi considérer l'application* $\mathpzc{s}_n : \mathbb{E} \times \cdots \times \mathbb{E} \to \mathbb{E}$ *(avec* $n$ *facteurs dans le domaine de définition) donnée par* $\mathpzc{s}(x_1, \ldots, x_n) = x_1 + \cdots + x_n$*, pour tous* $x_1, \ldots, x_n \in \mathbb{E}$. *Un argument par récurrence assez direct nous dit que* $\mathpzc{s}_n$ *est uniformément continue.*

*Par ailleurs, l'application* $\mathpzc{p}$ *est continue. Pour montrer la continuité de* $\mathpzc{p}$ *en* $(\lambda_0, x_0) \in \mathbb{R} \times \mathbb{E}$ *on remarque que*

$$\begin{aligned} N\big(\mathpzc{p}(\lambda, x) - \mathpzc{p}(\lambda_0, x_0)\big) &= N(\lambda \cdot x - \lambda_0 \cdot x_0) = N(\lambda \cdot x - \lambda_0 \cdot x + \lambda_0 \cdot x - \lambda_0 \cdot x_0) \\ &\leq N(\lambda \cdot x - \lambda_0 \cdot x) + N(\lambda_0 \cdot x - \lambda_0 \cdot x_0) = |\lambda - \lambda_0| N(x) + |\lambda_0| N(x - x_0) \\ &\leq N_{\max}\big((\lambda, x) - (\lambda_0, x_0)\big)\Big(N(x) + |\lambda_0|\Big), \end{aligned} \tag{1.12.1}$$

*où l'on a utilisé que* $|\lambda - \lambda_0| \leq N_{\max}((\lambda, x) - (\lambda_0, x_0))$ *et* $N(x - x_0) \leq N_{\max}((\lambda, x) - (\lambda_0, x_0))$. *Étant donné* $\epsilon > 0$ *(et* $(\lambda_0, x_0) \in \mathbb{R} \times \mathbb{E}$*), on pose*

$$\delta = \min\left(1, \frac{\epsilon}{2N_{\max}(\lambda_0, x_0) + 1}\right).$$

*Alors, si* $(\lambda, x) \in \mathbb{R} \times \mathbb{E}$ *satisfait que* $N_{\max}((\lambda, x) - (\lambda_0, x_0)) \leq \delta$*, on conclut en particulier que*

$$N(x) - N(x_0) \leq N(x - x_0) \leq N_{\max}\big((\lambda, x) - (\lambda_0, x_0)\big) \leq \delta \leq 1,$$

*ce qui nous dit que* $N(x) \leq N(x_0) + 1$*, tandis que* (1.12.1) *implique finalement que*

$$\begin{aligned} N\big(\mathpzc{p}(\lambda, x) - \mathpzc{p}(\lambda_0, x_0)\big) &\leq N_{\max}\big((\lambda, x) - (\lambda_0, x_0)\big)\Big(N(x) + |\lambda_0|\Big) \\ &\leq N_{\max}\big((\lambda, x) - (\lambda_0, x_0)\big)\Big(N(x_0) + |\lambda_0| + 1\Big) \\ &\leq N_{\max}\big((\lambda, x) - (\lambda_0, x_0)\big)\Big(2N_{\max}(\lambda_0, x_0) + 1\Big) \leq \epsilon, \end{aligned}$$

*ce qui montre que* $\mathpzc{p}$ *est continue en* $(\lambda_0, x_0)$*, pour tout* $(\lambda_0, x_0) \in \mathbb{R} \times \mathbb{E}$. *On remarque finalement que, étant donné* $\lambda \in \mathbb{R}$*, l'application* $\mathpzc{p}_\lambda : \mathbb{E} \to \mathbb{E}$ *donnée par* $\mathpzc{p}_\lambda(x) = \mathpzc{p}(\lambda, x)$ *pour* $x \in \mathbb{E}$ *est uniformément continue, vu que*

$$N\big(\mathpzc{p}_\lambda(x) - \mathpzc{p}_\lambda(x')\big) = |\lambda| N(x - x'),$$

*pour tous* $x, x' \in \mathbb{E}$.

**Exemple 1.12.6.** *Soient* $\mathbb{E}_1, \ldots, \mathbb{E}_n$ *des espaces vectoriels munis des normes* $N_1, \ldots, N_n$*, respectivement. Soit* $\mathbb{E} = \mathbb{E}_1 \times \cdots \times \mathbb{E}_n$ *l'espace vectoriel muni de la norme* $N_{\max}$ *définie par* (1.2.1). *On considère les applications* $\iota_i^{\mathbb{E}} : \mathbb{E}_i \to \mathbb{E}$ *et* $\pi_i^{\mathbb{E}} : \mathbb{E} \to \mathbb{E}_i$ *données par l'inclusion canonique et la projection canonique. Alors, elles sont uniformément continues.*

**Lemme 1.12.7.** *Soient* $(\mathbb{E}, N)$*,* $(\mathbb{F}, N')$ *et* $(\mathbb{G}, N'')$ *trois espaces vectoriels normés et soient* $S \subseteq \mathbb{E}$ *et* $T \subseteq \mathbb{F}$ *deux parties. Soit* $v_0 \in S$. *Soit* $f : S \to \mathbb{F}$ *une application continue en* $v_0$ *(resp., uniformément continues) et* $g : T \to \mathbb{G}$ *une application continue en* $f(v_0)$ *(resp., uniformément continue), où l'on suppose en plus que* $f(S) \subseteq T$. *Alors, la composition* $g \circ f : S \to \mathbb{G}$ *donnée par* $(g \circ f)(v) = g(f(v))$*, pour tout* $v \in S$*, est continue en* $v_0$ *(resp., uniformément continue).*

***Preuve.*** On va prouver les cas de la continuité simple et on laissera au lecteur/à la lectrice la preuve de l'énoncé pour la continuité uniforme: il s'agit seulement du même argument, sans fixer le point $v_0$.

Comme $f$ est continue en $v_0$, pour tout $\varepsilon > 0$, il existe $\delta = \delta(\varepsilon) > 0$ tel que $w \in S$ et $N(w - v_0) < \delta$ impliquent que $N'(f(w) - f(v_0)) < \varepsilon$. En outre, comme $g$ est continue en $f(v_0)$, pour tout $\epsilon > 0$, il existe

$\eta > 0$ tel que $u \in T$ et $N'(u - f(v_0)) < \eta$ impliquent que $N''(g(u) - g(f(v_0))) < \epsilon$. On fixe $\epsilon > 0$ et on prend $\hat{\delta} = \delta(\eta)$, où $\eta$ est déterminé par la continuité de $g$ en $f(v_0)$. Alors, $w \in S$ et $N(w - v_0) < \hat{\delta}$ impliquent que $N'(f(w) - f(v_0)) < \eta$, ce qui nous donne que $N''(g(f(w)) - g(f(v_0))) < \epsilon$, *i.e.* $g \circ f$ est continue en $v_0$. $\square$

Voici un critère assez simple pour construire des application à partir d'autres application continues.

**Proposition 1.12.8.** *Soient $\mathbb{E}, \mathbb{F}_1, \dots, \mathbb{F}_n$ des espaces vectoriels munis des normes $N, N'_1, \dots, N'_n$, respectivement. On pose $\mathbb{F} = \mathbb{F}_1 \times \dots \times \mathbb{F}_n$ muni de la norme $N'_{\max}$, définie par* (1.2.1) *à partir des normes des facteurs respectifs. On suppose que pour tout $i \in [\![1, n]\!]$, on a une application $f_i : S \to \mathbb{F}_i$, où $S \subseteq \mathbb{E}$. On considère l'application $f : S \to \mathbb{F}$ donnée par $f(x) = (f_1(x), \dots, f_n(x))$, où $x \in S$. Alors, $f$ est continue en $v_0 \in S$ (resp., uniformément continue) si et seulement si $f_i$ est continue en $v_0 \in S$ (resp., uniformément continue) pour tout $i \in [\![1, n]\!]$.*

***Preuve.*** On suppose que $f$ est continue en $v_0$ (resp., uniformément continue). On note que $f_i = \pi_i^{\mathbb{F}} \circ f$, avec les applications uniformément continues $\pi_i^{\mathbb{F}}$ définies dans l'Exemple 1.12.6. D'après le Lemme 1.12.7, on conclut que $f_i$ est continue en $v_0$ (resp., uniformément continue) pour tout $i \in [\![1, n]\!]$. Réciproquement, on suppose que $f_i$ est continue en $v_0$ (resp., uniformément continue) pour tout $i \in [\![1, n]\!]$. On montrera que $f$ est continue en $v_0$, le cas de la continuité uniforme étant obtenu *mutatis mutandi*. Comme $f_i$ est continue en $v_0$, pour tout $\varepsilon > 0$, il existe $\delta_i > 0$ tel que $w \in S$ et $N(w - v_0) < \delta_i$ impliquent que $N'_i(f(w) - f(v_0)) < \varepsilon$. On fixe $\varepsilon > 0$ et on prend $\delta = \min(\delta_1, \dots, \delta_n) > 0$. Alors, $w \in S$ et $N(w - v_0) < \delta$ impliquent que

$$N'_{\max}\big(f(w) - f(v_0)\big) \leq \varepsilon,$$

comme on voulait démontrer. $\square$

**Corollaire 1.12.9.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés, soit $S \subseteq \mathbb{E}$ une partie et $v_0 \in S$. Soient $f, g : S \to \mathbb{F}$ et $\lambda : S \to \mathbb{R}$ applications continues en $v_0$ (resp., uniformément continues avec $\lambda$ constante) Alors, l'application $f + \lambda \cdot g : S \to \mathbb{F}$ donnée par $(f + \lambda \cdot g)(v) = f(v) + \lambda(v)g(v)$, pour tout $v \in S$, est continue en $v_0$ (resp., uniformément continue). Par ailleurs, si $\lambda(v) \neq 0$ pour tout $v \in S$, alors l'application $\mu : S \to \mathbb{R}$ donnée par $\mu(v) = 1/\lambda(v)$ pour $v \in S$ est continue en $v_0$.*

***Preuve.*** On va démontrer d'abord la continuité de $f + \lambda \cdot g$ en $v_0$. C'est clair que $f + \lambda \cdot g$ est la composition de l'application $(f, \lambda, g) : S \to \mathbb{F} \times \mathbb{R} \times \mathbb{F}$ donnée par $x \mapsto (f(x), \lambda(x), g(x))$ pour tout $x \in S$, $\mathrm{id}_{\mathbb{F}} \times \mathfrak{p} : \mathbb{F} \times \mathbb{R} \times \mathbb{F} \to \mathbb{F} \times \mathbb{F}$ et $\mathfrak{s} : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$, où $\mathfrak{p} : \mathbb{R} \times \mathbb{F} \to \mathbb{F}$ est l'application $(c, v) \mapsto cv$ pour tous $c \in \mathbb{R}$ et $v \in \mathbb{F}$. La Proposition 1.12.8 nous dit que $(f, \lambda, g)$ est continue $v_0$. De même, le Lemme 1.12.7, la Proposition 1.12.8 et l'Exemple 1.12.5 nous dissent que $\mathrm{id}_{\mathbb{F}} \times \mathfrak{p}$ est continue. En outre, l'application $\mathfrak{s}$ est uniformément continue, d'après l'Exemple 1.12.5. La continuité de $f + \lambda \cdot g$ suit alors du Lemme 1.12.7.

Pour démontrer la continuité uniforme on remarque que $f + \lambda \cdot g$ est la composition de l'application $(f, g) : S \to \mathbb{F} \times \mathbb{F}$ donnée par $x \mapsto (f(x), g(x))$ pour tout $x \in S$, $\mathrm{id}_{\mathbb{F}} \times \mathfrak{p}_\lambda : \mathbb{F} \times \mathbb{F} \to \mathbb{F} \times \mathbb{F}$ et $\mathfrak{s} : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$. La Proposition 1.12.8 nous dit que $(f, g)$ est uniformément continue. De même, le Lemme 1.12.7, la Proposition 1.12.8 et l'Exemple 1.12.5 nous dissent que $\mathrm{id} \times \mathfrak{p}_\lambda$ est uniformément continue. En effet, on peut écrire $\mathrm{id}_{\mathbb{F}} \times \mathfrak{p}_\lambda = (q_1, q_2)$ avec $q_1 = \pi_1^{\mathbb{F} \times \mathbb{F}}$ et $q_2 = \mathfrak{p}_\lambda \circ \pi_2^{\mathbb{F} \times \mathbb{F}}$, où $\pi_1^{\mathbb{F} \times \mathbb{F}} : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ et $\pi_2^{\mathbb{F} \times \mathbb{F}} : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ sont les projections canoniques. Les résultats mentionnés nous disent que $\mathrm{id} \times \mathfrak{p}_\lambda$ est uniformément continue. En outre, l'application $\mathfrak{s}$ est uniformément continue, d'après l'Exemple 1.12.5. L'affirmation suit alors du Lemme 1.12.7.

Pour démontrer la dernière affirmation il suffit de noter que $\mu$ est la composition de l'application $\lambda : S \to \mathbb{R} \setminus \{0\}$ avec l'application continue $\mathrm{inv} : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$ donnée par $\mathrm{inv}(x) = x^{-1}$ pour tout $x \in \mathbb{R} \setminus \{0\}$. Le Lemme 1.12.7 nous dit alors que $\mu$ est continue en $v_0$. $\square$

**Exemple 1.12.10** (Une famille de fonctions continues sur le plan $\mathbb{R}^2$). *Soient $a, b, c, d \in \mathbb{R}_{>0}$ et soit $f : \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}$ l'application donnée par*

$$f(x, y) = \frac{|x|^a |y|^b}{|x|^c + |y|^d}$$

*pour tout $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Le Corollaire 1.12.9 nous dit que $f$ est continue. On affirme qu'il existe une application continue $\bar{f} : \mathbb{R}^2 \to \mathbb{R}$ telle que $\bar{f}|_{\mathbb{R}^2\setminus\{(0,0)\}} = f$ si et seulement si*

$$\frac{a}{c} + \frac{b}{d} > 1.$$

*On note d'abord que, par définition de $f$, on a*

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{\substack{(x,\,y)\,\to\,(0,\,0)\\(x,\,y)\,\in\,\mathbb{R}^2_{\geq 0}}} f(x,y).$$

*En outre, vu que l'application $\Psi : \mathbb{R}^2_{\geq 0} \to \mathbb{R}^2_{\geq 0}$ donnée par $\Psi(X,Y) = (X^d, Y^c)$ pour tout $(X,Y) \in \mathbb{R}^2_{\geq 0}$ est continue avec réciproque $\Psi^{-1} : \mathbb{R}^2_{\geq 0} \to \mathbb{R}^2_{\geq 0}$ donnée par $\Psi^{-1}(x,y) = (x^{1/d}, y^{1/c})$ pour tout $(x,y) \in \mathbb{R}^2_{\geq 0}$ aussi continue, on conclut que*

$$\begin{aligned}\lim_{(x,y)\to(0,0)} f(x,y) &= \lim_{\substack{(x,\,y)\,\to\,(0,\,0)\\(x,\,y)\,\in\,\mathbb{R}^2_{\geq 0}}} f(x,y) = \lim_{\substack{(X,\,Y)\,\to\,(0,\,0)\\(X,\,Y)\,\in\,\mathbb{R}^2_{\geq 0}}} f\big(\Psi(X,Y)\big)\\ &= \lim_{\substack{(X,\,Y)\,\to\,(0,\,0)\\(X,\,Y)\,\in\,\mathbb{R}^2_{\geq 0}}} f(X^d, X^c) = \lim_{(X,Y)\to(0,0)} \frac{|X|^{ad}|Y|^{bc}}{|X|^{cd} + |Y|^{cd}}.\end{aligned}$$

*Or, l'inégalité $a/c + b/d > 1$ équivaut à $ad + bc > cd$. Soient $\alpha, \beta \in \mathbb{R}_{>0}$ tels que $\alpha < ad$, $\beta < bc$ et $\alpha + \beta = cd$ (par exemple, il suffit de choisir $\alpha$ tel que $cd - bc < \alpha < ad$ et $\beta = cd - \alpha$). On a alors*

$$0 \leq \frac{|X|^{ad}|Y|^{bc}}{|X|^{cd} + |Y|^{cd}} \leq \frac{|X|^{\alpha}|X|^{ad-\alpha}|Y|^{\beta}|Y|^{bc-\beta}}{(|X|^{\alpha+\beta} + |Y|^{\alpha+\beta})^{\alpha/(\alpha+\beta)}(|X|^{\alpha+\beta} + |Y|^{\alpha+\beta})^{\beta/(\alpha+\beta)}} \leq |X|^{ad-\alpha}|Y|^{bc-\beta},$$

*où l'on a utilisé que $|X|^{\alpha} \leq (|X|^{\alpha+\beta} + |Y|^{\alpha+\beta})^{\alpha/(\alpha+\beta)}$ et $|Y|^{\beta} \leq (|X|^{\alpha+\beta} + |Y|^{\alpha+\beta})^{\beta/(\alpha+\beta)}$. Comme le dernier membre dans la chaîne d'inégalités précédentes tend vers zéro lorsque $(X,Y)$ tend vers zéro, on conclut que $f(x,y)$ tend vers zéro lorsque $(x,y)$ tend vers zéro, par le théorème des gendarmes.*

*On suppose désormais que $f(x,y)$ converge lorsque $(x,y)$ tend vers zéro. On remarque d'abord que dans ce cas la limite de $f(x,y)$ lorsque $(x,y)$ tend vers zéro doit être nulle, car*

$$\lim_{x\to 0} f(x,0) = \lim_{x\to 0} 0 = 0.$$

*En outre, l'existence de la limite de $f(x,y)$ lorsque $(x,y)$ tend vers zéro implique à fortiori que la limite*

$$\lim_{t\to 0+} f(t^d, t^c) = \lim_{t\to 0+} \frac{|t^d|^a|t^c|^b}{|t^d|^c + |t^c|^d} = \lim_{t\to 0+} \frac{|t|^{ad+bc}}{2|t|^{cd}} = \lim_{t\to 0+} \frac{|t|^{ad+bc-cd}}{2}$$

*existe, ce qui implique que $ad + bc \geq cd$. La limite précédente est dans ce cas $1/2$ si $ad + bc = cd$ et zéro si $ad + bc > cd$, ce qui équivaut à $a/c + b/d > 1$. Comme la limite est nulle si elle existe, on conclut que la convergence de $f(x,y)$ lorsque $(x,y)$ tend vers zéro nous dit que $a/c + b/d > 1$, comme on voulait démontrer.*

On a la caractérisation topologique suivante de la continuité d'une application.

**Proposition 1.12.11.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soient $S \subseteq \mathbb{E}$. Soit $f : S \to \mathbb{F}$ une application. Alors, $f$ est continue sur $S$ si et seulement si $f^{-1}(U)$ est ouvert (resp., fermé) de $S$ pour tout ensemble ouvert (resp., fermé) $U \subseteq \mathbb{F}$.*

***Preuve.*** On va démontrer d'abord l'énoncé pour les ouverts. On suppose que $f$ est continue sur $S$. Soit $U \subseteq \mathbb{F}$ un ensemble ouvert et soit $v \in f^{-1}(U)$. Comme $U$ est ouvert et $f(v) \in U$, il existe $\epsilon > 0$ tel que $B_{N'}(f(v), r) \subseteq U$. Par la continuité de $f$ en $v$, il existe $\delta > 0$ tel que $w \in S$ et $N(w - v) < \delta$ impliquent que $N'(f(w) - f(v)) < \varepsilon$, *i.e.* $B_N(v,\delta) \cap S \subseteq f^{-1}(B_{N'}(f(v),\delta)) \subseteq f^{-1}(U)$. Cela nous dit que $f^{-1}(U)$ est un ouvert de $S$. Réciproquement, on suppose que $f^{-1}(U)$ est ouvert de $S$ pour tout ensemble ouvert $U \subseteq \mathbb{F}$. Soit $v \in S$ et $U = B_{N'}(f(v), \varepsilon)$, où $\varepsilon > 0$ est arbitraire. En particulier, $U$ est ouvert d'après la Proposition 1.9.3. Par hypothèse, $f^{-1}(U)$ est ouvert de $S$. Comme $v \in f^{-1}(U)$ et $f^{-1}(U)$ est ouvert de $S$, il existe $\delta > 0$ tel que $B_N(v,\delta) \cap S \subseteq f^{-1}(U)$, *i.e.* $f$ et continue en $v$.

L'énoncé pour les ensembles fermés suit du cas des ensembles ouverts et de l'égalité $f^{-1}(\mathbb{F} \setminus F) = S \setminus f^{-1}(F)$, pour toute partie $F$ de $\mathbb{F}$. □

On a aussi une caractérisation séquentielle de la continuité d'une application.

**Proposition 1.12.12.** *Soient* $(\mathbb{E}, N)$ *et* $(\mathbb{F}, N')$ *deux espaces vectoriels normés et soient* $S \subseteq \mathbb{E}$. *Soit* $f : S \to \mathbb{F}$ *une application. Alors,* $f$ *est continue en* $v$ *si et seulement si la image directe* $(f(v_n))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ *de toute suite* $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ *qui converge vers* $v$ *est convergente de limite* $f(v)$.

***Preuve.*** On suppose que $f$ est continue en $v$ et soit $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ une suite qui converge vers $v$. On veut montrer que $(f(v_n))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ converge vers $f(v)$. On fixe $\varepsilon > 0$. D'après la continuité de $f$, il existe $\delta > 0$ tel que $w \in S$ et $N(w - v) < \delta$ impliquent que $N'(f(w) - f(v)) < \varepsilon$. Par ailleurs, comme $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $v$, pour $\delta > 0$ ci-dessus, il existe $n_0 \in \mathbb{N}$ tel que $N(v_n - v) < \delta$ pour tout entier $n \geq n_0$. En conséquence, $N'(f(v_n) - f(v)) < \varepsilon$ pour tout entier $n \geq n_0$, *i.e.* $(f(v_n))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ converge vers $f(v)$.

On suppose maintenant que la image directe $(f(v_n))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ de toute suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ qui converge vers $v$ est convergente de limite $f(v)$. On va montrer que $f$ est continue en $v$. On procède par l'absurde. Si ce n'est pas vrai, il existe $\varepsilon > 0$ tel que pour tout $\delta > 0$, il existe $w \in S$ avec $N(w - v) < \delta$ mais $N'(f(w)-f(v)) \geq \varepsilon$. De façon équivalente, il existe une suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ tel que $N(v_n-v) < 1/n$ mais $N'(f(v_n) - f(v)) \geq \varepsilon$. Cela veut dire que $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ qui converge vers $v$ mais $(f(v_n))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ ne converge pas vers $f(v)$, ce qui est absurde. $\square$

Le résultat suivante est clé dans la suite, pour démontrer que toutes les normes sur un espace vectoriel de dimension finie sont équivalentes.

**Proposition 1.12.13.** *Soient* $(\mathbb{E}, N)$ *et* $(\mathbb{F}, N')$ *deux espaces vectoriels normés,* $S \subseteq \mathbb{E}$ *une partie compacte et* $f : S \to \mathbb{F}$ *une application continue. Alors* $f(S) \subseteq \mathbb{F}$ *est compact aussi.*

***Preuve.*** Soit $(w_n)_{n\in\mathbb{N}} \in f(S)^{\mathbb{N}}$ une suite. On va monter qu'elle possède une sous-suite convergente dont la limite appartient à $f(S)$. Comme $w_n \in f(S)$, soit $v_n \in S$ tel que $f(v_n) = w_n$, pour tout $n \in \mathbb{N}$. Comme $S$ est compact, la suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ possède une sous-suite $(u_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ convergente dont la limite $\bar{v}$ appartient à $S$. Comme $f$ est continue, la Proposition 1.12.12 nous dit que la suite $(f(u_n))_{n\in\mathbb{N}} \in f(S)^{\mathbb{N}}$ converge vers $f(\bar{v}) \in f(S)$, comme on voulait démontrer. $\square$

## 1.13 Espaces des fonctions bornées et continues (optionnel)

Soient $S$ un ensemble et $(\mathbb{F}, N')$ un espace vectoriels normé. On définit l'ensemble

$$\mathrm{B}(S, \mathbb{F}) = \{f : S \to \mathbb{F} : \text{ il existe } C > 0 \text{ tel que } N'(f(v)) \leq C \text{ pour tout } v \in S\}$$

C'est clair que $\mathrm{B}(S, \mathbb{F})$ est un espace vectoriel avec les opérations $(f + \lambda g)(v) = f(v) + \lambda g(v)$ pour tout $v \in S$, où $f, g \in \mathrm{B}(S, \mathbb{F})$ et $\lambda \in \mathbb{R}$, vu que, si $N'(f(v)) \leq C$ et $N'(g(v)) \leq D$ pour tout $v \in S$, avec $C, D \in \mathbb{R}_{>0}$, alors $N'(f(v) + \lambda g(v)) \leq C + |\lambda| D$, pour tout $v \in S$. L'élément neutre de $\mathrm{B}(S, \mathbb{F})$ est la fonction nulle, qui associe $\mathbf{0}_{\mathbb{F}}$ à tout élément $v \in S$.

On suppose en plus que $(\mathbb{E}, N)$ est un espace vectoriel normé et $S \subseteq \mathbb{E}$ est une partie On définit les ensemble

$$\mathrm{BC}^0(S, \mathbb{F}) = \{f \in \mathrm{B}(S, \mathbb{F}) : f \text{ est continue } \}.$$

Comme une combinaison linéaire de fonctions continue est continue, d'après le Corollaire 1.12.9, on conclut que $\mathrm{BC}^0(S, \mathbb{F})$ est un sous-espace vectoriel de $\mathrm{B}(S, \mathbb{F})$. En plus, la Proposition 1.12.13 nous dit que, si $S$ est compact, alors $\mathrm{BC}^0(S, \mathbb{F})$ coïncide avec l'ensemble formé des applications continues de $S$ dans $\mathbb{F}$.

**Proposition 1.13.1.** *L'application* $N_\infty : \mathrm{B}(S, \mathbb{F}) \to \mathbb{R}_{\geq 0}$ *donnée par* $N_\infty(f) = \sup\{N'(f(v)) : v \in S\}$ *est bien définie et donne une norme sur l'espace vectoriel* $\mathrm{B}(S, \mathbb{F})$. *L'espace vectoriel normé ainsi défini est complet si* $\mathbb{F}$ *est complet.*

***Preuve.*** On voit bien que $N_\infty(f) = 0$ implique $f(v) = \mathbf{0}_{\mathbb{F}}$ pour tout $v \in S$, ce qui nous dit que $f$ est la fonction nulle. En outre, on voit bien que

$$N_\infty(\lambda f) = \sup\Big\{N'\big(\lambda f(v)\big) : v \in S\Big\} = \sup\Big\{|\lambda| N'\big(f(v)\big) : v \in S\Big\} = |\lambda| \sup\Big\{N'\big(f(v)\big) : v \in S\Big\} N_\infty(f),$$

et

$$\begin{aligned}N_\infty(f+g) &= \sup\Big\{N'\big(f(v)+g(v)\big) : v\in S\Big\} \le \sup\Big\{N'\big(f(v)\big)+N'\big(g(v)\big) : v\in S\Big\}\\ &\le \sup\Big\{N'\big(f(v)\big) : v\in S\Big\} + \sup\Big\{N'\big(f(v)\big) : v\in S\Big\} = N_\infty(f)+N_\infty(g),\end{aligned}$$

pour tous $f, g \in$ et $\lambda \in \mathbb{R}$, ce qui nous dit que $N_\infty$ est une norme.

Pour la deuxième, soit $(f_n)_{n\in\mathbb{N}} \in \mathrm{B}(S,\mathbb{F})^{\mathbb{N}}$ une suite de Cauchy. En conséquence, étant donné $\epsilon > 0$, il existe $n_0(\epsilon) \in \mathbb{N}$ tel que $N_\infty(f_n - f_m) \le \epsilon$ pour tous $n, m \ge n_0(\epsilon)$, ce qui implique que

$$N'\big(f_n(v)-f_m(v)\big) \le \sup\Big\{N'\big(f_n(v)-f_m(v)\big) : v\in S\Big\} = N_\infty(f_n-f_m) \le \epsilon$$

pour tous $v \in S, n, m \ge n_0(\epsilon)$, *i.e.* la suite $(f_n(v))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ est une suite de Cauchy de $\mathbb{F}$ pour tout $v \in S$. Comme $\mathbb{F}$ est complet, la limite de $(f_n(v))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ existe dans $\mathbb{F}$ pour tout $v \in S$. On définie ainsi $f(v) \in \mathbb{F}$ comme la limite de la suite convergente $(f_n(v))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$, pour tout $v \in S$. On va démontrer d'abord que $f \in \mathrm{B}(S,\mathbb{F})$. On note pour cela que, vu que $(f_n)_{n\in\mathbb{N}} \in \mathrm{B}(S,\mathbb{F})^{\mathbb{N}}$ une suite de Cauchy, elle est bornée, *i.e.* il existe $C > 0$ tel que $N_\infty(f_n) \le C$ pour tout $n \in \mathbb{N}$. Or, comme $(f_n(v))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ converge ver $f(v)$ lorsque $n$ tend vers $+\infty$, on choisit $n_v(\epsilon) \in \mathbb{N}$ tel que $N(f(v) - f_n(v)) \le \epsilon$ pour tout $n \ge n_v(\epsilon)$ et $v \in S$. Alors,

$$\begin{aligned}N'\big(f(v)\big) &= N'\big(f(v)-f_n(v)+f_n(v)\big) \le N'\big(f(v)-f_n(v)\big)+N'\big(f_n(v)\big)\\ &\le N'\big(f(v)-f_n(v)\big) + \Big\{N'\big(f_n(v)\big) : v\in S\Big\} \le C + C = 2C\end{aligned}$$

pour $n \ge n_v(C)$ et $v \in S$, ce qui implique que $N'(f(v)) \le 2C$ pour tout $v \in S$, *i.e.* $f \in \mathrm{B}(S,\mathbb{F})^{\mathbb{N}}$.

On va finalement démontrer que $(f_n)_{n\in\mathbb{N}} \in \mathrm{B}(S,\mathbb{F})^{\mathbb{N}}$ converge vers $f$ lorsque $n$ tend vers $+\infty$. Comme $(f_n(v))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ converge ver $f(v)$ lorsque $n$ tend vers $+\infty$, on choisit $n_v(\epsilon) \in \mathbb{N}$ tel que $N(f(v) - f_n(v)) \le \epsilon$ pour tout $n \ge n_v(\epsilon)$. Or, pour $n \ge n_0(\epsilon/2)$, on a

$$\begin{aligned}N'\big(f(v)-f_n(v)\big) &= N'\big(f(v)-f_m(v)+f_m(v)-f_n(v)\big) \le N'\big(f(v)-f_m(v)\big)+N'\big(f_m(v)-f_n(v)\big)\\ &\le N'\big(f(v)-f_m(v)\big) + \Big\{N'\big(f_m(v)-f_m(v)\big) : v\in S\Big\} \le \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon,\end{aligned}$$

pour $m \ge n_v(\epsilon/2)$ et $v \in S$. Cela nous dit que $N'(f(v) - f_n(v)) \le \epsilon$ pour tout $v \in S$ et $n \ge n_0(\epsilon/2)$, *i.e.* $(f_n)_{n\in\mathbb{N}} \in \mathrm{B}(S,\mathbb{F})^{\mathbb{N}}$ converge vers $f$ lorsque $n$ tend vers $+\infty$. □

**Proposition 1.13.2.** *La restriction de l'application $N_\infty : \mathrm{B}(S,\mathbb{F}) \to \mathbb{R}_{\ge 0}$ définie dans la Proposition 1.13.1 donne une norme sur le sous-espace $\mathrm{BC}^0(S,\mathbb{F})$ de $\mathrm{B}(S,\mathbb{F})$. Ce sous-espace est fermé dans $\mathrm{B}(S,\mathbb{F})$, et, en conséquence, complet.*

***Preuve.*** La première partie est une conséquence des Propositions 1.2.2 et 1.13.1. Pour montrer que $\mathrm{BC}^0(S,\mathbb{F})$ est fermé dans $\mathrm{B}(S,\mathbb{F})$, soit $(f_n)_{n\in\mathbb{N}} \in \mathrm{BC}^0(S,\mathbb{F})^{\mathbb{N}}$ une suite convergente et soit $f \in \mathrm{B}(S,\mathbb{F})$ sa limite. Il suffit de démontrer que $f$ est continue. Soit $v_0 \in S$. On va montrer que $f$ est continue en $v_0$. Soit $\epsilon > 0$. Comme $(f_n)_{n\in\mathbb{N}} \in \mathrm{BC}^0(S,\mathbb{F})^{\mathbb{N}}$ converge vers $f \in \mathrm{B}(S,\mathbb{F})$, il existe $n_0 \in \mathbb{N}$ tel que $N_\infty(f - f_n) \le \epsilon/3$, ce qui implique que $N'(f(v) - f_n(v)) \le \epsilon/3$ pour tout $v \in S$. En plus, comme $f_{n_0}$ est continue, il existe $\delta > 0$ tel que $N'(f_{n_0}(v) - f_{n_0}(v_0)) \le \epsilon/3$ pour tout $v \in S$ tel que $N(v - v_0) \le \delta$. En conséquence,

$$\begin{aligned}N'\big(f(v)-f(v_0)\big) &= N'\big(f(v)-f_{n_0}(v)+f_{n_0}(v)-f_{n_0}(v_0)+f_{n_0}(v_0)-f(v_0)\big)\\ &\le N'\big(f(v)-f_{n_0}(v)\big)+N'\big(f_{n_0}(v)-f_{n_0}(v_0)\big)+N'\big(f_{n_0}(v_0)-f(v_0)\big)\\ &\le \frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3} = \epsilon\end{aligned}$$

si $N(v-v_0) \le \delta$, ce qui implique que $f$ est continue en $v_0$. La complétude de $\mathrm{BC}^0(S,\mathbb{F})$ suit directement du Corollaire 1.10.4. □

## 1.14 Une autre preuve de l'équivalence de normes sur un espace vectoriel de dimension finie

Dans cette section on va donner une autre preuve du Théorème 1.8.1, indépendante de celle que l'on a donnée dans la Section 1.8. Plus précisément, on ne va pas modifier les deux premiers paragraphes dans la preuve du Théorème 1.8.1. Cela démontre en particulier que

$$N(v) \leq C' N_\infty(v), \tag{1.14.1}$$

pour tout $v \in \mathbb{E}$, où l'on utilise la même notation que celle considérée dans les deux premiers paragraphes de la preuve du Théorème 1.8.1.

Il reste à démontrer qu'il existe $C > 0$ tel que

$$N(v) \geq C N_\infty(v), \tag{1.14.2}$$

pour tout $v \in \mathbb{E}$. C'est cette partie qui sera différente à ce que l'on a fait dans la Section 1.8. C'est clair que (1.14.2) est vérifiée pour $v = \mathbf{0}_\mathbb{E}$. On suppose désormais $v \neq \mathbf{0}_\mathbb{E}$.

Or, (1.14.1) nous dit que l'application $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$ est continue pour $N_\infty$ et $|\ \ |$. En effet, l'inégalité triangulaire pour $N$ et (1.14.1) impliquent que

$$\big|N(v) - N(w)\big| \leq N(v-w) \leq C' N_\infty(v-w),$$

pour tous $v, w \in \mathbb{E}$, ce qui donne la continuité de l'application $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$ pour $N_\infty$ et $|\ \ |$.

Soit $S = \{w \in \mathbb{E} : N_\infty(w) = 1\}$. C'est clair que $S$ est fermé et borné pour $N_\infty$, et donc compact pour $N_\infty$, d'après la Proposition 1.11.10. On remarque que $\mathbf{0}_\mathbb{E} \notin S$, puisque $N_\infty$ est une norme. La Proposition 1.12.13 nous dit que l'image $N(S) \subseteq \mathbb{R}_{\geq 0}$ est un compact de $\mathbb{R}$, donc *a fortiori* fermé. Comme $N$ est une norme et $\mathbf{0}_\mathbb{E} \notin S$, $0 \notin f(S)$, ce qui dit qu'il existe $C > 0$ tel que $N(w) \geq C$ pour tout $w \in S$. Pour $v \in \mathbb{E}$ non nul, on considère $w = v/N_\infty(v) \in S$. D'après nos arguments précédents, $N(v/N_\infty(v)) \geq C$, *i.e.* $N(v) \geq C N_\infty(v)$, comme on voulait démontrer.

**Corollaire 1.14.1.** *Tout espace vectoriel normé $(\mathbb{E}, N)$ de dimension finie est complet. En plus, une partie $S$ de $\mathbb{E}$ est compacte si et seulement si elle est fermée et bornée.*

***Preuve.*** La première partie est une conséquence immédiate des Propositions 1.7.3, 1.7.4 et 1.7.5, et du Théorème 1.8.1. La deuxième partie suit directement des Propositions 1.11.6, 1.11.7 et 1.11.10, et du Théorème 1.8.1. □

## 1.15 D'autres applications de la compacité

**Définition 1.15.1.** *Soient $(\mathbb{E}, N)$ un espace vectoriel normé, $S \subseteq \mathbb{E}$ une partie et $f : S \to \mathbb{R}$ une application On dit que $v_0 \in S$ est un **maximum** (resp., **minimum**) de $f$ si $f(v_0) \geq f(v)$ (resp., $f(v_0) \leq f(v)$) pour tout $v \in S$. Un **extremum** de $f$ est un maximum ou minimum.*

**Proposition 1.15.2.** *Soient $(\mathbb{E}, N)$ un espace vectoriel normé, $S \subseteq \mathbb{E}$ une partie compacte et $f : S \to \mathbb{R}$ une application continue. Alors, $f$ possède un maximum et un minimum.*

***Preuve.*** Comme $f$ est continue et $S$ est compact, la Proposition 1.12.13 nous dit que $f(S) \subseteq \mathbb{R}$ est un compact de $\mathbb{R}$, *i.e.*, il est fermé et borné (voir Propositions 1.11.7 et 1.11.10). Soient $c = \inf f(S)$ et $d = \sup f(S)$. Comme $f(S)$ est compact, l'infimum et le supremum précédents sont achevés, *i.e.* il existe $v, w \in S$ tel que $c = f(v)$ et $d = f(w)$. En effet, il existe des suites $(v_n)_{n\in\mathbb{N}} \in S^\mathbb{N}$ et $(w_n)_{n\in\mathbb{N}} \in S^\mathbb{N}$ telles que $(f(v_n))_{n\in\mathbb{N}}$ et $(f(w_n))_{n\in\mathbb{N}}$ convergent vers $c$ et $d$, respectivement. Comme $S$ est compact, il existe des sous-suites convergentes $(v_{\varphi(n)})_{n\in\mathbb{N}} \in S^\mathbb{N}$ et $(w_{\varphi'(n)})_{n\in\mathbb{N}} \in S^\mathbb{N}$ de limites $v \in S$ et $w \in S$, respectivement. Alors,

$$f(v) = f\Big(\lim_{n\to\infty} v_{\varphi(n)}\Big) = \lim_{n\to\infty} f(v_{\varphi(n)}) = c$$

et

$$f(w) = f\Big(\lim_{n\to\infty} w_{\varphi'(n)}\Big) = \lim_{n\to\infty} f(w_{\varphi'(n)}) = d.$$

En conséquence, $v$ est un minimum de $f$ et $w$ est un maximum de $f$. □

**Proposition 1.15.3.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés, $S \subseteq \mathbb{E}$ une partie compacte et $f : S \to \mathbb{F}$ une application continue. Alors, $f$ est uniformément continue.*

***Preuve.*** On procède par l'absurde. Si $f$ n'est pas uniformément continue, il existe $\varepsilon_0 > 0$ tel que, pour tout $\delta > 0$, il existe $v, w \in S$ avec $N(v-w) < \delta$ mais $N'(f(v)-f(w)) \geq \varepsilon_0$. En particulier, pour tout $n \in \mathbb{N}$, il existe $v_n, w_n \in S$ avec $N(v_n - w_n) < 1/2^n$ mais $N'(f(v_n)-f(w_n)) \geq \varepsilon_0$. Comme $S$ est compact, il existe une sous-suite convergente $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de la suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ avec limite $\bar{v} \in S$, où $\phi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective. Le même raisonnement nous dit qu'il existe une sous-suite convergente $(w_{\phi(\psi(n))})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de la suite $(w_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ avec limite $\bar{w} \in S$, où $\psi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective. Comme toute sous-suite d'une suite convergente est convergente avec la même limite, $(v_{\phi(\psi(n))})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $\bar{v} \in S$. Soit l'application croissante et injective $\rho = \phi \circ \psi : \mathbb{N} \to \mathbb{N}$ et les sous-suites convergentes $(v_{\rho(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ et $(w_{\rho(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ avec limites $\bar{v}$ et $\bar{w}$, respectivement, Comme $f$ est continue, $(f(v_{\rho(n)}))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ et $(f(w_{\rho(n)}))_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}}$ convergent vers $f(\bar{v})$ et $f(\bar{w})$, respectivement. Or,

$$\begin{aligned} 0 \leq N(\bar{v} - \bar{w}) &\leq N(\bar{v} - v_{\rho(n)}) + N(v_{\rho(n)} - w_{\rho(n)}) + N(w_{\rho(n)} - \bar{w}) \\ &\leq N(\bar{v} - v_{\rho(n)}) + \frac{1}{2^{\rho(n)}} + N(w_{\rho(n)} - \bar{w}) \leq N(\bar{v} - v_{\rho(n)}) + \frac{1}{2^n} + N(w_{\rho(n)} - \bar{w}), \end{aligned}$$

pour tout $n \in \mathbb{N}$, où l'on a utilisé que $\rho(n) \geq n$. Si l'on prend la limite quand $n$ tend vers $+\infty$ du dernier membre, on conclut que $N(\bar{v} - \bar{w}) = 0$, *i.e.* $\bar{v} = \bar{w}$. Par contre, l'inégalité triangulaire nous dit que

$$\begin{aligned} N'\big(f(\bar{v}) - f(\bar{w})\big) &\geq N'\big(f(v_{\rho(n)}) - f(w_{\rho(n)})\big) - N'\big(f(v_{\rho(n)}) - f(\bar{v})\big) - N'\big(f(w_{\rho(n)}) - f(\bar{w})\big) \\ &\geq \varepsilon_0 - N'\big(f(v_{\rho(n)}) - f(\bar{v})\big) - N'\big(f(w_{\rho(n)}) - f(\bar{w})\big), \end{aligned}$$

pour tout $n \in \mathbb{N}$. Si l'on prend la limite quand $n$ tend vers $+\infty$ du dernier membre, on conclut que $N'\big(f(\bar{v}) - f(\bar{w})\big) \geq \varepsilon_0$, ce qui est absurde, car $\bar{v} = \bar{w}$. □

**Définition 1.15.4.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés, $S \subseteq \mathbb{E}$ et $T \subseteq \mathbb{F}$ Soit $f : S \to T$ une application continue sur $S$. On dit que $f$ est un* ***homéomorphisme*** *(pour $N$ et $N'$) s'il existe une application continue $g : T \to S$ (pour $N'$ et $N$) telle que $g \circ f = \mathrm{id}_S$ et $f \circ g = \mathrm{id}_T$.*

**Remarque 1.15.5.** *De façon équivalente, une application continue $f : S \to T$ est un homéomorphisme si $f$ est bijective et l'application réciproque $f^{-1} : T \to S$ est continue (pour $N'$ et $N$).*

**Proposition 1.15.6.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés, $S \subseteq \mathbb{E}$ et $T \subseteq \mathbb{F}$. Soit $f : S \to T$ une application continue et bijective. Si $S$ est compact, alors $f$ est un homéomorphisme.*

***Preuve.*** Il suffit de démontrer que l'application réciproque $f^{-1} : T \to S$ est continue. D'après, la Proposition 1.12.12, cela équivaut à montrer que, pour toute suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ à valeurs dans $S$, si $(f(v_n))_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ converge vers $w \in T$, alors $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $f^{-1}(w) \in S$. On pose $v = f^{-1}(w)$, *i.e.* $f(v) = w$. On va procéder par l'absurde. Si $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ ne converge pas vers $v$, il existe $\varepsilon_0 > 0$ et une sous-suite $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ telle que $N(v_{\phi(n)} - v) \geq \varepsilon_0$, pour tout $n \in \mathbb{N}$, où $\phi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective. Or, $S$ étant compact, $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ a une sous-suite convergente $(v_{\phi(\psi(n))})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ avec limite $v'$ dans $S$, où $\psi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective. On considère l'application croissante et injective $\rho = \phi \circ \psi : \mathbb{N} \to \mathbb{N}$. Comme $N(v_{\phi(n)} - v) \geq \varepsilon_0$, pour tout $n \in \mathbb{N}$, alors

$$N(v - v') \geq N(v - v_{\rho(n)}) - N(v_{\rho(n)} - v') \geq \varepsilon_0 - N(v_{\rho(n)} - v'),$$

pour tout $n \in \mathbb{N}$. Si l'on prend la limite du dernier membre quand $n$ tend vers $+\infty$, on trouve que $N(v - v') \geq \varepsilon_0 > 0$. En conséquence, $v' \neq v$, ce qui nous dit que $f(v') \neq f(v)$, puisque $f$ est injective. Par ailleurs, comme $f$ est continue, la sous-suite $(f(v_{\rho(n)}))_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(f(v_n))_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ converge vers $f(v')$. Comme $(f(v_n))_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ est convergente avec limite $f(v)$ et toute sous-suite d'une suite convergente est aussi convergente avec la même limite, $f(v') = f(v)$, ce qui est absurde. □

## 1.16 Compacité par recouvrements (optionnel)

Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. Un **recouvrement ouvert** de $S$ est une famille $\{U_i\}_{i\in I}$ d'ensembles ouverts de $\mathbb{E}$ telle que $S \subseteq \cup_{i\in I} U_i$. Dans ce cas, on dit aussi que $\{U_i\}_{i\in I}$ **recouvre** $S$. Si $I$ est fini, on dit que le recouvrement est **fini**.

**Définition 1.16.1.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé et soit $S \subseteq \mathbb{E}$ une partie. On dit que $S$ est* **compact** *(pour $N$), ou qu'il satisfait la* **propriété de Borel-Lebesgue***, si, pour tout recouvrement ouvert $\{U_i\}_{i\in I}$ de $S$, il existe $J \subseteq I$ fini tel que $\{U_i\}_{i\in J}$ recouvre $S$ aussi.*

Même si cette propriété s'avère compliquée, elle est en fait très souple et permet de démontrer plusieurs résultats très puissants en topologie, analyse et algèbre. Le but de cette section sera de démontrer que la notion introduite ci-dessus est équivalente avec la notion de compacité séquentielle introduite dans la Section 1.11: c'est aussi la raison pour laquelle on peut se permettre d'utiliser le mot "compacité" pour les deux notions. Par contre, dans cette section, pour des raisons de clarté, on va distinguer les deux notions et on ne va pas omettre l'adjectif "séquentielle" (resp., l'adverbe "séquentiellement") devant le nom "compacité" (resp., l'adjectif "compact(e)") si l'on veux faire allusion à la notion introduite dans la Définition 1.11.1. On remarque aussi que, dans tous les résultats dans les autres sections de ces notes, où l'on a utilisé le mot "compact", cela ne fait allusion qu'à la notion de compacité séquentielle.

On commence avec un résultat préliminaire.

**Proposition 1.16.2.** *Soient $(\mathbb{E}, N)$ un espaces vectoriel normé et $S \subseteq \mathbb{E}$ une partie séquentiellement compacte. Pour tout $r \in \mathbb{R}_{>0}$, il existe un ensemble fini $S' \subseteq S$ tel que $\{B_N(v, r))\}_{v\in S'}$ recouvre $S$.*

***Preuve.*** On procède par l'absurde. L'hypothèse de l'absurde nous dit que l'on peut construire par récurrence une suite $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ telle que, pour tout $n \in \mathbb{N}$, $v_n \in S \setminus \cup_{i=1}^{n-1} B_N(v_i, r)$. Comme $S$ est séquentiellement compact $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ admet une sous-suite convergente $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ avec limite $w \in S$. C'est clair que $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ satisfait aussi que, pour tout $n \in \mathbb{N}$, $w_n \in S \setminus \cup_{i=1}^{n-1} B_N(w_i, r)$. La convergence de $(w_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ nous dit qu'il existe $n_0 \in \mathbb{N}$ tel que $N(w_n - w) < r/2$, pour tout entier $n \geq n_0$. Cela implique que

$$N(w_{n+1} - w_n) \leq N(w_{n+1} - w) + N(w - w_n) < \frac{r}{2} + \frac{r}{2} = r,$$

pour tous entier $n \geq n_0$, ce qui est absurde. □

On va démontrer l'équivalence entre les deux notions de compacité que l'on a introduites.

**Théorème 1.16.3.** *Soient $(\mathbb{E}, N)$ un espaces vectoriel normé et $S \subseteq \mathbb{E}$ une partie. Alors $S$ est séquentiellement compact si et seulement si $S$ est compact.*

***Preuve.*** On suppose d'abord que $S$ est séquentiellement compact. On va montrer que $S$ satisfait la propriété de Borel-Lebesgue. Soit $\{U_i\}_{i\in I}$ un recouvrement ouvert de $S$. On affirme d'abord qu'il existe $n \in \mathbb{N}$ tel que, pour tout $v \in S$, il existe $i \in I$ satisfaisant $B_N(v, 1/2^n) \subseteq U_i$. Si ce n'est pas vrai, pour tout $n \in \mathbb{N}$, il existe $v_n \in S$ tel que $B_N(v_n, 1/2^n) \cap (\mathbb{E} \setminus U_i) \neq \emptyset$, pour tout $i \in I$. Par compacité séquentielle, il existe une sou-suite convergente $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ de $(v_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ avec limite $w \in S$, où $\phi : \mathbb{N} \to \mathbb{N}$ est une application croissante et injective. Comme $\{U_i\}_{i\in I}$ est un recouvrement ouvert de $S$ et $w \in S$, il existe $i_0 \in I$ tel que $w \in U_{i_0}$. Soit $m \in \mathbb{N}$ tel que $B_N(w, 2^{-m}) \subseteq U_{i_0}$, dont l'existence est assurée puisque $U_{i_0}$ est ouvert. Comme $(v_{\phi(n)})_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ converge vers $w$, il existe $n_0 \in \mathbb{N}$ tel que $v_{\phi(n)} \in B_N(w, 2^{-m}) \subseteq U_{i_0}$, pour tout entier $n \geq n_0$. Sans perte de généralité, on peut supposer que $n_0 \geq m$. Cela implique que $\phi(n) \geq \phi(n_0) \geq \phi(m) \geq m$ et en conséquence

$$B_N(v_{\phi(n)}, 1/2^{\phi(n)}) \subseteq B_N(v_{\phi(n)}, 1/2^m) \subseteq U_{i_0},$$

pour tout entier $n \geq n_0$. Cela contredit le fait que $B_N(v_{\phi(n)}, 1/2^{\phi(n)}) \cap (\mathbb{E} \setminus U_i) \neq \emptyset$ pour tout $i \in I$. Par conséquent, il existe $n \in \mathbb{N}$ tel que, pour tout $v \in S$, il existe $i \in I$ satisfaisant $B_N(v, 1/2^n) \subseteq U_i$. On fixe $n$ comme ci-dessus. D'après la Proposition 1.16.2, il existe un ensemble fini $S' \subseteq S$ tel que $\{B_N(v, 1/2^n)\}_{v\in S'}$ recouvre $S$. Pour tout $v \in S'$, on choisit $i_v \in I$ tel que $B_N(v, 1/2^n) \subseteq U_{i_v}$. Alors $\{U_{i_v}\}_{v\in S'}$ est un recouvrement fini de $S$, ce qui dit que $S$ est compact.

On va démontrer la réciproque. On suppose que $S$ satisfait la propriété de Borel-Lebesgue. On va montrer que $S$ est séquentiellement compact. Pour le faire, on va utiliser la description de compacité séquentielle donnée par la Proposition 1.11.4, *i.e.* on va montrer que toute partie infinie $T \subseteq S$ admet un point d'accumulation dans $S$. On va procéder par l'absurde, *i.e.* on suppose que pour tout point $v \in S$, il existe $r_v \in \mathbb{R}_{>0}$ tel que $B_N(v, r_v) \cap T$ est fini. On considère le recouvrement ouvert $\{B_N(v, r_v)\}_{v \in S}$ de $S$. Comme $S$ est compact, *i.e.* il satisfait la propriété de Borel-Lebesgue, il existe une partie finie $S' \subseteq S$ telle que $\{B_N(v, r_v)\}_{v \in S'}$ couvre $S$. En conséquence,

$$T = T \cap S \subseteq T \cap \bigg( \bigcup_{v \in S'} B_N(v, r_v) \bigg) = \bigcup_{v \in S'} \big( T \cap B_N(v, r_v) \big).$$

Comme le dernier ensemble est une réunion finie d'ensembles finis, il est fini. Cela qui implique que $T$ est fini, ce qui est absurde. En conséquence, $S$ est séquentiellement compact. □

## 1.17 Normes sur l'espace des applications linéaires

Soient $\mathbb{E}$ et $\mathbb{F}$ deux espaces vectoriels. On définit $\mathrm{Lin}(\mathbb{E}, \mathbb{F})$ l'espace vectoriel formé des applications linéaires $T : \mathbb{E} \to \mathbb{F}$. On rappelle que, étant donnés $T, T' \in \mathrm{Lin}(\mathbb{E}, \mathbb{F})$ et $\lambda \in \mathbb{R}$, $T + \lambda T' : \mathbb{E} \to \mathbb{F}$ est l'application linéaire définie par $(T + \lambda T')(v) = T(v) + \lambda \cdot T'(v)$, pour tout $v \in \mathbb{E}$, et l'élément neutre pour l'addition de $\mathrm{Lin}(\mathbb{E}, \mathbb{F})$ est la fonction nulle de $\mathbb{E}$ dans $\mathbb{F}$. Si $\mathbb{E} = \mathbb{F}$, on écrira souvent $\mathrm{Lin}(\mathbb{E})$ au lieu de $\mathrm{Lin}(\mathbb{E}, \mathbb{E})$.

Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés. On définit maintenant $\mathrm{L}(\mathbb{E}, \mathbb{F}) \subseteq \mathrm{Lin}(\mathbb{E}, \mathbb{F})$ la partie formée des applications linéaires continues $T : \mathbb{E} \to \mathbb{F}$. D'après le Lemme 1.12.7, $\mathrm{L}(\mathbb{E}, \mathbb{F})$ est un sous-espace vectoriel de $\mathrm{Lin}(\mathbb{E}, \mathbb{F})$. Si $\mathbb{E} = \mathbb{F}$, on écrira souvent $\mathrm{L}(\mathbb{E})$ au lieu de $\mathrm{L}(\mathbb{E}, \mathbb{E})$.

**Proposition 1.17.1.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $T : \mathbb{E} \to \mathbb{F}$ une application linéaire. Les conditions suivantes sont équivalentes:*

*(i) il existe $C > 0$ tel que*

$$N'\big(T(v)\big) \leq C N(v), \tag{1.17.1}$$

*pour tout $v \in \mathbb{E}$;*

*(ii) $T$ est uniformément continue;*

*(iii) $T$ est continue;*

*(iv) $T$ est continue en $\mathbf{0}_{\mathbb{E}}$;*

*(v) $T$ est borné sur un ouvert de $\mathbf{0}_{\mathbb{E}}$.*

***Preuve.*** Les implications (i) → (ii) → (iii) → (iv) → (v) sont évidentes. Il suffit de démontrer que la condition (v) implique (i). On suppose qu'il existe $r, C' > 0$ tels que $N'(T(v)) \leq C'$ pour tout $v \in B_N(\mathbf{0}_{\mathbb{E}}, r)$. C'est clair que (1.17.1) est vérifiée pour $v = \mathbf{0}_{\mathbb{E}}$, puis $T(\mathbf{0}_{\mathbb{E}}) = \mathbf{0}_{\mathbb{F}}$. Soit $v \in \mathbb{E} \setminus \{\mathbf{0}_{\mathbb{E}}\}$ et on considère $w = r \cdot v / (2N(v))$. Alors $w \in B_N(\mathbf{0}_{\mathbb{E}}, r)$, ce qui implique que $N'(T(w)) \leq C'$, *i.e.*

$$N'\big(T(v)\big) \leq \frac{2C'}{r} N(v).$$

Cela nous que (1.17.1) est vérifiée avec $C = 2C'/r$. □

On a encore une autre conséquence du Théorème 1.8.1.

**Proposition 1.17.2.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $T : \mathbb{E} \to \mathbb{F}$ une application linéaire. Si la dimension de $\mathbb{E}$ est finie, alors $T$ est continue.*

***Preuve.*** Soit $\{v_1, \dots, v_n\}$ une base $\mathbb{E}$ et soit $U : \mathbb{E} \to \mathbb{R}^n$ la bijection linéaire donnée par $U(v_i) = e_i$, pour tout $i \in [\![1, n]\!]$, où $\{e_1, \dots, e_n\}$ est la base canonique de $\mathbb{R}^n$. On considère la norme $N_U$ donnée par $U$ et la norme $\| \ \|_\infty$ (voir la Proposition 1.2.1). D'après le Théorème 1.8.1, il existe $K > 0$ tel que

$N_U(v) \leq KN(v)$, pour tout $v \in \mathbb{E}$. On pose $C' = \max\{N'(T(v_i)) : i \in [\![1,n]\!]\}$. Soit $v \in \mathbb{E}$. On écrit $v = \sum_{i=1}^n c_i v_i$, avec $c_i \in \mathbb{R}$ pour tout $i \in [\![1,n]\!]$. C'est clair que

$$N'\big(T(v)\big) \leq \sum_{i=1}^n |c_i| N'\big(T(v_i)\big) \leq C' \sum_{i=1}^n |c_i| \leq nC' \max_{i \in [\![1,n]\!]}\big(|c_i|\big) = nC' N_U(v) \leq nC'KN(v),$$

ce qui implique que $T$ est continue. □

On définit l'application $\mathbf{N}_{N,N'} : \mathrm{L}(\mathbb{E},\mathbb{F}) \to \mathbb{R}_{\geq 0}$ par

$$\mathbf{N}_{N,N'}(T) = \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big(T(v)\big)}{N(v)}. \tag{1.17.2}$$

La Proposition 1.17.1 nous dit que ce supremum est fini et que l'application $\mathbf{N}_{N,N'}$ est bien définie.

**Proposition 1.17.3.** *Soient $(\mathbb{E}, N)$ et $(\mathbb{F}, N')$ deux espaces vectoriels normés et soit $\mathbf{N}_{N,N'} : \mathrm{L}(\mathbb{E},\mathbb{F}) \to \mathbb{R}_{\geq 0}$ l'application définie dans* (1.17.2). *Alors $\mathbf{N}_{N,N'}$ est une norme sur $\mathrm{L}(\mathbb{E},\mathbb{F})$, appelée la **norme associée (ou subordonnée) aux normes $N$ et $N'$**.*

***Preuve.*** Soit $T : \mathbb{E} \to \mathbb{F}$ une application linéaire continue. C'est direct à voir que $\mathbf{N}_{N,N'}(T) = 0$ implique $N'(T(v)) = 0$, *i.e.* $T(v) = \mathbf{0}_\mathbb{F}$, pour tout $v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}$. En outre, comme $T$ est linéaire, $T(\mathbf{0}_\mathbb{E}) = \mathbf{0}_\mathbb{F}$. Cela nous dit que $T$ est le vecteur nul de $\mathrm{L}(\mathbb{E},\mathbb{F})$, *i.e.* $\mathbf{N}_{N,N'}$ satisfait (N1). Par ailleurs,

$$\begin{aligned}\mathbf{N}_{N,N'}(\lambda T) &= \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big((\lambda T)(v)\big)}{N(v)} = \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big(\lambda \cdot T(v)\big)}{N(v)} = \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} |\lambda| \frac{N'\big(T(v)\big)}{N(v)} \\ &= |\lambda| \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big(T(v)\big)}{N(v)} = |\lambda| \mathbf{N}_{N,N'}(T),\end{aligned}$$

pour tout $\lambda \in \mathbb{R}$, où l'on a utilisé (0.0.5) dans la quatrième égalité. En conséquence, $\mathbf{N}_{N,N'}$ satisfait (N2).

Finalement, soient $T, T' : \mathbb{E} \to \mathbb{F}$ deux applications linéaires continues.

$$\begin{aligned}\mathbf{N}_{N,N'}(T+T') &= \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big((T+T')(v)\big)}{N(v)} = \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big(T(v)+T'(v)\big)}{N(v)} \\ &\leq \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \left(\frac{N'\big(T(v)\big)}{N(v)} + \frac{N'\big(T'(v)\big)}{N(v)}\right) \leq \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big(T(v)\big)}{N(v)} + \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N'\big(T'(v)\big)}{N(v)} \\ &= \mathbf{N}_{N,N'}(T) + \mathbf{N}_{N,N'}(T'),\end{aligned}$$

où l'on a appliqué (0.0.6) au quatrième membre. Par conséquent, $\mathbf{N}_{N,N'}$ satisfait (N3). □

Par définition de $\mathbf{N}_{N,N'}$, on voit bien que

$$N'\big(T(v)\big) \leq \mathbf{N}_{N,N'}(T) N(v), \tag{1.17.3}$$

pour tout $v \in \mathbb{E}$.

**Proposition 1.17.4.** *Soit $(\mathbb{E}, N)$ un espace vectoriel normé. On muni $\mathrm{L}(\mathbb{E},\mathbb{E})$ de la norme $\mathbf{N}_{N,N}$. Soient $T, T' \in \mathrm{L}(\mathbb{E},\mathbb{E})$. Alors, la norme $\mathbf{N}_{N,N}$ est **sous-multiplicative**, i.e.*

$$\mathbf{N}_{N,N}(T \circ T') \leq \mathbf{N}_{N,N}(T)\mathbf{N}_{N,N}(T'). \tag{1.17.4}$$

*En plus, $\mathbf{N}_{N,N}(\mathrm{id}_\mathbb{E}) = 1$.*

***Preuve.*** Soit $v \in \mathbb{E}$ non nul. Alors, (1.17.3) nous dit que

$$\frac{N\big((T \circ T')(v)\big)}{N(v)} \leq \frac{\mathbf{N}_{N,N}(T) N\big(T'(v)\big)}{N(v)} \leq \frac{\mathbf{N}_{N,N}(T)\mathbf{N}_{N,N}(T')N(v)}{N(v)} = \mathbf{N}_{N,N}(T)\mathbf{N}_{N,N}(T').$$

La définition de supremum implique alors

$$\mathbf{N}_{N,N}(T \circ T') = \sup_{v \in \mathbb{E} \setminus \{\mathbf{0}_\mathbb{E}\}} \frac{N\big((T \circ T')(v)\big)}{N(v)} \leq \mathbf{N}_{N,N}(T)\mathbf{N}_{N,N}(T'),$$

comme on voulait démontrer. L'égalité $\mathbf{N}_{N,N}(\mathrm{id}_\mathbb{E}) = 1$ est immédiate de la définition. □

Si l'on note $\| \ \|_{\mathbb{E}}$ et $\| \ \|_{\mathbb{F}}$ les normes sur $\mathbb{E}$ et $\mathbb{F}$, respectivement, au lieu de $N$ et $N'$, on écrira $\| \ \|_{\mathbb{E},\mathbb{F}}$ (ou simplement $\| \ \|$) au lieu de $\mathbf{N}_{N,N'}$.

**Remarque 1.17.5.** *Soient $\mathbb{E} = \mathbb{R}^n$ et $\mathbb{F} = \mathbb{R}^m$, avec $n, m \in \mathbb{Z}_{>0}$. Il existe un isomorphisme $\phi_{\mathbb{E},\mathbb{F}} : \mathrm{L}(\mathbb{E},\mathbb{F}) \to \mathrm{M}_{n\times m}(\mathbb{R})$ d'espaces vectoriels, où $\phi_{\mathbb{E},\mathbb{F}}(T) = (a_{i,j})_{i\in[\![1,n]\!],j\in[\![1,n]\!]}$ est donnée par*

$$T(e_j^{\mathbb{R}^n}) = \sum_{i=1}^{m} a_{i,j} e_i^{\mathbb{R}^m},$$

*où $\{e_j^{\mathbb{R}^\ell} : j \in [\![1,\ell]\!]\}$ est la base canonique de $\mathbb{R}^\ell$ pour tout $\ell \in \mathbb{Z}_{>0}$. En plus, c'est clair que $\phi_{\mathbb{D},\mathbb{F}}(T \circ T') = \phi_{\mathbb{E},\mathbb{F}}(T)\phi_{\mathbb{D},\mathbb{E}}(T')$, pour tous $T' \in \mathrm{L}(\mathbb{D},\mathbb{E})$ et $T \in \mathrm{L}(\mathbb{E},\mathbb{F})$, où $\mathbb{D} = \mathbb{R}^p$ avec $p \in \mathbb{Z}_{>0}$ et $\phi_{\mathbb{E},\mathbb{F}}(T)\phi_{\mathbb{D},\mathbb{E}}(T')$ dénote le produit matriciel.*

*D'après l'identification évidente $\mathrm{M}_{n\times m}(\mathbb{R}) \simeq \mathbb{R}^{nm}$, on peut munir $\mathrm{L}(\mathbb{E},\mathbb{F})$ de la norme $N_\infty = \| \ \|_\infty \circ \phi_{\mathbb{E},\mathbb{F}}$, où $\| \ \|_\infty$ est la norme infini de $\mathbb{R}^{nm}$.*

*On pose désormais $\mathbb{E} = \mathbb{F}$, et donc $n = m$. Soit $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ la matrice telle que toutes les coefficients valent $1$ et $T = \phi_{\mathbb{E},\mathbb{E}}^{-1}(A)$. Comme $N_\infty(T) = 1$ et $N_\infty(T \circ T) = n$, on voit bien que $N_\infty$ ne satisfait pas la condition* (1.17.4) *de la Proposition 1.17.4, ce qui implique que $N_\infty$ n'est pas la norme associée à une norme sur $\mathbb{R}^n$ si $n > 1$.*

*Si l'on utilise l'identification précédente $\mathrm{M}_{n\times n}(\mathbb{R}) \simeq \mathbb{R}^{n^2}$, on peut aussi munir $\mathrm{L}(\mathbb{E},\mathbb{E})$ de la norme $N_2 = \| \ \|_2 \circ \phi_{\mathbb{E},\mathbb{E}}$, où $\| \ \|_2$ est la norme 2 de $\mathbb{R}^{n^2}$. Dans ce cas, la norme $N_2$, appelée **norme de Frobenius**, est sous-multiplicative, comme le lecteur/la lectrice pourra démontrer en employant l'inégalité de Cauchy-Schwarz. En effet, on considère la produit scalaire $\langle\, ,\rangle : \mathrm{L}(\mathbb{E},\mathbb{E}) \times \mathrm{L}(\mathbb{E},\mathbb{E}) \to \mathbb{R}$ donné par $\langle T, T'\rangle = \sum_{i=1}^n \langle T(v_i), T'(v_i)\rangle_2$, pour $T, T' \in \mathrm{L}(\mathbb{E},\mathbb{E})$, où $\langle\, ,\rangle_2$ est le produit scalaire euclidien de $\mathbb{R}^n$ et $\{v_i\}_{i\in[\![1,n]\!]}$ est une base de $\mathbb{R}^n$. Alors, c'est facile à démontrer que la définition précédente est indépendante de la base $\{v_i\}_{i\in[\![1,n]\!]}$ de $\mathbb{R}^n$ et $N_2$ est la normé associée à $\langle\, ,\rangle$. En plus, si $\phi_{\mathbb{E},\mathbb{E}}(T) = (a_{i,j})_{i,j\in[\![1,n]\!]}$ et $\phi_{\mathbb{E},\mathbb{E}}(T') = (b_{i,j})_{i,j\in[\![1,n]\!]}$, on a*

$$\langle T, T'\rangle = \sum_{i,j=1}^{n} a_{i,j} b_{i,j},$$

*et en conséquence*

$$N_2(T\circ T')^2 = \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n} a_{i,k}b_{k,j}\right)^2 \le \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n} a_{i,k}^2 \sum_{k'=1}^{n} b_{k',j}^2\right) = \sum_{i,k=1}^{n} a_{i,k}^2 \sum_{j,k'=1}^{n} b_{j,k'}^2 = N_2(T)^2 N_2(T')^2,$$

*comme on voulait démontrer, où l'on a utilisé l'inégalité de Cauchy-Schwarz du produit scalaire euclidien de $\mathbb{R}^n$. Par contre, $N_2$ n'est pas associée à une norme sur $\mathbb{R}^n$ si $n > 1$, puisque $N_2(\mathrm{id}_{\mathbb{R}^2}) = \sqrt{n} > 1$ (voir la dernière condition dans la Proposition 1.17.4).*

# Chapitre 2

# Calcul différentiel

Le but du deuxième chapitre c'est d'étudier le comportement local des fonctions définies sur des ouverts des espaces vectoriels de dimension finie. En particulier, on étudiera le problème d'optimisation de fonctions. Dans l'espace vectoriel $\mathbb{R}^n$ on considérera toujours la norme euclidienne, que l'on notera $\| \ \|$, et, étant donné un espace vectoriel normé $\mathbb{E}$ avec norme $\| \ \|$, l'espace vectoriel $\mathrm{L}(\mathbb{E})$ avec la norme associée à la norme de $\mathbb{E}$ sera notée aussi $\| \ \|$.

## 2.1 Dérivée, dérivées directionnelles et différentielle

On rappelle d'abord la notion de dérivabilité des fonctions d'une seule variable.

**Définition 2.1.1.** *Soient $m \in \mathbb{Z}_{>0}$, $J \subseteq \mathbb{R}$ un intervalle ouvert, $f : J \to \mathbb{R}^m$ une application et $t_0 \in J$. On rappelle que la **dérivée** de $f$ en $t_0$ est donné par la limite*

$$f'(t_0) = \lim_{t \to 0} \frac{f(t_0 + t) - f(t_0)}{t} \in \mathbb{R}^m,$$

*si elle existe. Dans ce cas on dit que $f$ est **dérivable** en $t_0$.*

On introduit les notions fondamentales de dérivée directionnelle et de différentiabilité.

**Définition 2.1.2.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$ un point de $U$. Étant donné une vecteur non nul $w \in \mathbb{R}^n$, la **dérivée directionnelle** de $f$ en $v_0$ dans la direction de $w$, aussi appelée **dérivée au sens de Gateaux**, est donnée par la limite*

$$\frac{\partial f}{\partial w}(v_0) = \lim_{t \to 0} \frac{f(v_0 + t \cdot w) - f(v_0)}{t} \in \mathbb{R}^m,$$

*si elle existe. Pour $i \in [\![1, n]\!]$, la **$i$-ème dérivée partielle** de $f$ en $v_0$, aussi appelée la **dérivée partielle de $f$ en $v_0$ selon $x_i$** et notée*

$$\frac{\partial f}{\partial x_i}(v_0) \text{ ou } D_i f(v_0),$$

*est donnée par la dérivée directionnelle de $f$ en $v_0$ dans la direction du vecteur $e_i$, où $\{e_1, \dots, e_n\}$ est la base canonique de $\mathbb{R}^n$.*

**Remarque 2.1.3.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. Soit $\alpha : \mathbb{R} \to \mathbb{R}^n$ l'application $\alpha(t) = v_0 + t \cdot w$, où $w \in \mathbb{R}^n$ est une vecteur non nul. Comme $\alpha$ est évidemment continue, $\alpha(0) = v_0 \in U$ et $U$ est ouvert, d'après la Proposition 1.12.11, $\alpha^{-1}(U)$ est un ouvert de $\mathbb{R}$ incluant $0$. On choisit un intervalle ouvert $J \subseteq \alpha^{-1}(U)$ incluant $0$. Alors, on voit bien que la dérivée directionnelle de $f$ en $v_0$ dans la direction du vecteur non nul $w \in \mathbb{R}^n$ existe si et seulement la fonction $f_w : J \to \mathbb{R}^m$ donnée par $f_w = f \circ \alpha|_J$ est différentiable en $0$, et la dérivée directionnelle est précisément $f'_w(0)$.*

Le résultat suivant est standard.

**Proposition 2.1.4.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $v_0 \in U$, $w \in \mathbb{R}^n$ non nul, $f, g : U \to \mathbb{R}^m$ et $\lambda : U \to \mathbb{R}$ des applications dont les dérivées directionnelles en $v_0$ dans le direction de $w$ existent. Alors, la dérivée directionnelle de l'application $f + \lambda \cdot g : U \to \mathbb{R}^m$ en $v_0$ dans le direction de $w$ existe et*

$$\frac{\partial(f+\lambda\cdot g)}{\partial w}(v_0) = \frac{\partial f}{\partial w}(v_0) + \frac{\partial \lambda}{\partial w}(v_0)g(v_0) + \lambda(v_0)\frac{\partial g}{\partial w}(v_0).$$

*En outre, si $\lambda(v) \neq 0$ pour tout $v \in U$, la dérivée directionnelle de l'application $\mu : U \to \mathbb{R}$ donnée par $\mu(v) = 1/\lambda(v)$ pour $v \in U$ en $v_0$ dans la direction de $w$ existe et*

$$\frac{\partial \mu}{\partial w}(v_0) = -\frac{1}{\lambda(v_0)^2}\frac{\partial \lambda}{\partial w}(v_0).$$

***Preuve.*** Soit $\alpha : \mathbb{R} \to \mathbb{R}^n$ l'application $\alpha(t) = v_0 + t \cdot w$, où $w \in \mathbb{R}^n$ est une vecteur non nul. On choisit un intervalle ouvert $J \subseteq \alpha^{-1}(U)$ incluant 0, et soient $f_w, g_w : J \to \mathbb{R}^m$ et $\lambda_w : J \to \mathbb{R}$ les fonctions données par $f_w = f \circ \alpha|_J$, $g_w = g \circ \alpha|_J$ et $\lambda_w = \lambda \circ \alpha|_J$. D'après la Remarque 2.1.3, on sait que les dérivées directionnelles de $f, g : U \to \mathbb{R}^m$ et $\lambda : U \to \mathbb{R}$ en $v_0$ dans le direction de $w$ existent, et elles coïncident avec $f'_w(0)$, $g'_w(0)$ et $\lambda'_w(0)$, respectivement. En outre, soit $h = f + \lambda g : U \to \mathbb{R}^m$ et $h_w = h \circ \alpha|_J$. Alors, la dérivée directionnelle de $h$ en $v_0$ dans la direction de $w$ est donnée par $h'_w(0)$, si elle existe. Or, $h_w = f_w + \lambda_w g_w$ nous dit que $h'_w = f'_w + \lambda'_w g_w + \lambda_w g'_w$, ce qui implique que

$$\frac{\partial(f+\lambda g)}{\partial w}(v_0) = h'_w(0) = f'_w(0) + \lambda'_w(0)g_w(0) + \lambda_w(0)g'_w(0) = \frac{\partial f}{\partial w}(v_0) + \frac{\partial \lambda}{\partial w}(v_0)g(v_0) + \lambda(v_0)\frac{\partial g}{\partial w}(v_0).$$

En conséquence, la dérivée directionnelle de l'application $f + \lambda g : U \to \mathbb{R}^m$ en $v_0$ dans le direction de $w$ existe et

$$\frac{\partial(f+\lambda g)}{\partial w}(v_0) = \frac{\partial f}{\partial w}(v_0) + \frac{\partial \lambda}{\partial w}(v_0)g(v_0) + \lambda(v_0)\frac{\partial g}{\partial w}(v_0).$$

Pour la dernière partie, il suffit de noter que, si $\mu_w = \mu \circ \alpha|_J$, on a alors $\mu_w = 1/\lambda_w$, ce qui implique que

$$\frac{\partial \mu}{\partial w}(v_0) = \mu'_w(0) = -\frac{1}{\lambda_w(0)^2}\lambda'_w(0) = -\frac{1}{\lambda(v_0)^2}\frac{\partial \lambda}{\partial w}(v_0).$$

□

On présente d'abord les deux définitions suivantes de différentiabilité.

**Définition 2.1.5.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$ un point de $U$. On dit que $f$ est **différentiable au sens de Gateaux** en $v_0$ si la dérivée directionnelle de $f$ en $v_0$ dans la direction de $w$ existe pour tout vecteur non nul $w \in \mathbb{R}^n$ et s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ telle que*

$$\frac{\partial \mu}{\partial w}(v_0) = L(w)$$

*pour tout vecteur non nul $w \in \mathbb{R}^n$.*

**Définition 2.1.6.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. On dit que $f$ est **différentiable au sens de Fréchet** en $v_0$, ou simplement que $f$ est **différentiable** en $v_0$, s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ telle que*

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \frac{f(v_0+h) - f(v_0) - L(h)}{\|h\|} = \mathbf{0}_{\mathbb{R}^m}, \tag{2.1.1}$$

*où $\| \ \|$ dénote la norme euclidienne de $\mathbb{R}^n$.*

**Proposition 2.1.7.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}^m$ une application différentiable en $v_0 \in U$. Soient $L, L' : \mathbb{R}^n \to \mathbb{R}^m$ deux applications linéaires telles que*

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \frac{f(v_0+h) - f(v_0) - L(h)}{\|h\|} = \lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \frac{f(v_0+h) - f(v_0) - L'(h)}{\|h\|} = \mathbf{0}_{\mathbb{R}^m}, \tag{2.1.2}$$

*où $\| \ \|$ dénote la norme euclidienne de $\mathbb{R}^n$. Alors $L = L'$.*

***Preuve.*** D'après (2.1.2) on voit bien que

$$\begin{aligned}\lim_{h\to \mathbf{0}_{\mathbb{R}^n}} \frac{L'(h)-L(h)}{\|h\|} &= \lim_{h\to \mathbf{0}_{\mathbb{R}^n}} \left( \frac{f(v_0+h)-f(v_0)-L(h)}{\|h\|} - \frac{f(v_0+h)-f(v_0)-L'(h)}{\|h\|} \right) \\ &= \lim_{h\to \mathbf{0}_{\mathbb{R}^n}} \frac{f(v_0+h)-f(v_0)-L(h)}{\|h\|} - \lim_{h\to \mathbf{0}_{\mathbb{R}^n}} \frac{f(v_0+h)-f(v_0)-L'(h)}{\|h\|} \\ &= \mathbf{0}_{\mathbb{R}^m} - \mathbf{0}_{\mathbb{R}^m} = \mathbf{0}_{\mathbb{R}^m}.\end{aligned}$$

Or, $(L'(h)-L(h))/\|h\| = (L'-L)(h/\|h\|)$, vu que $L'-L$ est une application linéaire. Cela nous dit que

$$\lim_{h\to \mathbf{0}_{\mathbb{R}^n}} (L'-L)\left(\frac{h}{\|h\|}\right) = \mathbf{0}_{\mathbb{R}^m}.$$

En particulier, soit $v \in \mathbb{R}^n$ un vecteur quelconque de norme 1 et on pose $h = tv$, avec $t \in \mathbb{R}_{>0}$. La limite précédente implique *a fortiori* que

$$\mathbf{0}_{\mathbb{R}^m} = \lim_{t\to 0^+} (L'-L)\left(\frac{tv}{\|tv\|}\right) = \lim_{t\to 0^+} (L'-L)\left(\frac{tv}{t\|v\|}\right) = \lim_{t\to 0^+} (L'-L)(v) = (L'-L)(v),$$

car le dernier membre ne dépend pas de $t$. On a bien montré que $L$ et $L'$ coïncident en tout vecteur de norme 1. Soit $w$ un vecteur non nul et on pose $v = w/\|w\|$, *i.e.* $w = \|w\|v$. Alors,

$$L(w) = L\big(\|w\|v\big) = \|w\|L(v) = \|w\|L'(v) = L'\big(\|w\|v\big) = L'(w),$$

Finalement, comme $L(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathbb{R}^m} = L(\mathbf{0}_{\mathbb{R}^n})$, car toute application linéaire envoie le vecteur nul dans le vecteur nul, on conclut que $L = L'$. □

**Définition 2.1.8.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}^m$ une application différentiable en $v_0 \in U$. D'après la proposition précédente, l'application linéaire $L$ dans* (2.1.1) *est unique. Elle est appelée la* ***différentielle (de Fréchet)*** *de $f$ en $v_0$ et sera notée $Df(v_0)$.*

*On dira que $f$ est* ***différentiable (sur $U$)*** *si $f$ est différentiable en tout point $v \in U$. Cela induit une fonction $Df : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ donnée par $v \mapsto Df(v)$, pour tout $v \in U$, que l'on appelle* ***différentielle*** *de $f$.*

**Remarque 2.1.9.** *C'est facile à vérifier que si* (2.1.1) *est vérifiée pour la norme euclidienne, elle est aussi vérifiée pour une norme équivalente. Comme toutes les normes de $\mathbb{R}^n$ son équivalentes, d'après le Théorème 1.8.1, la définition de différentiabilité d'une fonction $f : U \to \mathbb{R}^m$ est indépendante de la norme choisie.*

**Remarque 2.1.10.** *Soit $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$, avec $U$ ouvert. Il existe alors $r > 0$ tel que $B_{\|\ \|}(v_0, r) \subseteq U$. C'est facile à voir que $f$ est différentiable en $v_0$ si et seulement s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ et une fonction $\psi : B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \to \mathbb{R}^m$ telles que*

$$f(v_0+h) = f(v_0) + L(h) + \psi(h)\|h\|,$$

*pour tout $h \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$ et que*

$$\lim_{h\to \mathbf{0}_{\mathbb{R}^n}} \psi(h) = \mathbf{0}_{\mathbb{R}^m}.$$

**Exemple 2.1.11.** *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une application constante. Alors, $f$ est différentiable sur $\mathbb{R}^n$ et $Df(v) = \mathbf{0}_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)}$, pour tout $v \in \mathbb{R}^n$.*

Le résultat suivant montre que les notions de différentiabilité et de dérivabilité des fonctions d'une seule variable sont équivalentes.

**Proposition 2.1.12.** *Soient $m \in \mathbb{Z}_{>0}$, $J \subseteq \mathbb{R}$ un intervalle ouvert, $f : J \to \mathbb{R}^m$, et $t_0 \in J$. Alors, $f$ est différentiable en $t_0$ si et seulement si $f$ est dérivable en $t_0$. Dans ce cas, on a $Df(t_0)(\lambda) = \lambda \cdot f'(t_0)$, pour tout $\lambda \in \mathbb{R}$.*

***Preuve.*** On suppose d'abord que $f$ est différentiable en $t_0$, *i.e.*

$$\lim_{t\to 0} \frac{f(t_0+t)-f(t_0)-Df(t_0)(t)}{|t|} = \mathbf{0}_{\mathbb{R}^m}.$$

Comme $Df(t_0) : \mathbb{R} \to \mathbb{R}^m$ est une application linéaire, $Df(t_0)(t) = t \cdot Df(t_0)(1)$ pour tout $t \in \mathbb{R}$, ce qui nous dit que

$$\lim_{t \to 0} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{|t|} = \mathbf{0}_{\mathbb{R}^m}.$$

Or, l'existence de la limite précédente nous donne les identités

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^m} &= \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{|t|} = \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{t} \\ &= \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0)}{t} - Df(t_0)(1) \end{aligned}$$

et

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^m} &= \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{|t|} = \lim_{t \to 0-} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{-t} \\ &= \lim_{t \to 0+} -\frac{f(t_0 + t) - f(t_0)}{t} + Df(t_0)(1), \end{aligned}$$

ce qui implique que

$$Df(t_0)(1) = \lim_{t \to 0} \frac{f(t_0 + t) - f(t_0)}{t}.$$

En conséquence, $f$ est dérivable en $t_0$.

De façon réciproque, on suppose que $f$ est dérivable en $t_0$, *i.e.*

$$f'(t_0) = \lim_{t \to 0} \frac{f(t_0 + t) - f(t_0)}{t}$$

existe. Soit $Df(t_0) : \mathbb{R} \to \mathbb{R}^m$ l'application linéaire donnée par $Df(t_0)(\lambda) = \lambda \cdot f'(t_0)$, pour tout $\lambda \in \mathbb{R}$. On affirme que

$$\lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{|t|} = \lim_{t \to 0} \frac{f(t_0 + t) - f(t_0) - Df(t_0)(t)}{|t|} = \mathbf{0}_{\mathbb{R}^m}.$$

En effet, la limite précédente est équivalente aux identités

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^m} &= \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{|t|} = \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{t} \\ &= \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0)}{t} - Df(t_0)(1) = \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0)}{t} - f'(t_0) \end{aligned}$$

et

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^m} &= \lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{|t|} = \lim_{t \to 0-} \frac{f(t_0 + t) - f(t_0) - t \cdot Df(t_0)(1)}{-t} \\ &= \lim_{t \to 0+} -\frac{f(t_0 + t) - f(t_0)}{t} + Df(t_0)(1) = -\lim_{t \to 0+} \frac{f(t_0 + t) - f(t_0)}{t} + f'(t_0), \end{aligned}$$

qui suivent de dérivabilité de $f$ en $t_0$. En conséquence, $f$ est différentiable en $t_0$. Finalement, l'identité $Df(t_0)(\lambda) = \lambda \cdot f'(t_0)$, pour tout $\lambda \in \mathbb{R}$, suit directement des identités précédentes. □

**Proposition 2.1.13.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. Si $f$ est différentiable en $v_0$, alors $f$ est continue en $v_0$.*

***Preuve.*** D'après la Remarque 2.1.10, il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ et une fonction $\psi : B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \to \mathbb{R}^m$ (où $r > 0$ est donné par la condition $B_{\| \ \|}(v_0, r) \subseteq U$) telles que

$$f(v_0 + h) - f(v_0) = L(h) + \psi(h)\|h\|, \tag{2.1.3}$$

pour tout $h \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$ et que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \psi(h) = \mathbf{0}_{\mathbb{R}^m}. \tag{2.1.4}$$

Si l'on prend la limite de (2.1.3) quand $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$, on utilise que toute application linéaire entre espaces vectoriels de dimension finie est continue (voir la Proposition 1.17.2) et (2.1.4), on conclut que $f(v_0 + h) - f(v_0)$ tend vers $\mathbf{0}_{\mathbb{R}^m}$, quand $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$, *i.e.* $f$ est continue en $v_0$. □

**Exemple 2.1.14.** *Soit $L : \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire. C'est facile à vérifier que la dérivée directionnelle de $L$ en tout point $v \in \mathbb{R}^n$ dans la direction de tout vecteur non nul $w \in \mathbb{R}^n$ existe et elle est égale à $L(w)$. En plus, $L$ est différentiable en tout point $v \in \mathbb{R}^n$ et sa différentielle $DL(v)$ est précisément $L$. En conséquence, les fonction considérées dans l'Exemple 1.12.6 sont différentiables.*

**Exemple 2.1.15.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application polynomiale. Pour $i \in [\![1, n]\!]$ on écrira $\pi_i : \mathbb{R}^n \to \mathbb{R}$ la projection canonique sur la i-ème coordonnée, i.e. $\pi(x_1, \dots, x_i, \dots, x_n) = x_i$, pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$. Noter que $\pi_i$ possède toutes les dérivées directionnelles en tout point et elle aussi différentiable en tout point, d'après l'Exemple 2.1.14. En outre, comme $f$ est une fonction polynomiale, elle obtenue à partir de faire sommes et produits de fonctions du type $\pi_i : \mathbb{R}^n \to \mathbb{R}$ avec $i \in [\![1, n]\!]$. D'après la Proposition 2.1.4, $f$ possède toutes les dérivées directionnelles en tout point.*

**Exemple 2.1.16.** *On va donner un exemple d'une fonction dont toutes les dérivées directionnelles existent en tout point mais elle n'est pas continue, donc a fortiori elle n'est pas différentiable. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction donnée par*

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4}, & \text{si } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}, \\ 0, & \text{si } (x, y) = (0, 0). \end{cases}$$

*Comme l'application $f|_{\mathbb{R}^2 \setminus \{(0,0)\}}$ est un quotient de deux fonctions polynomiales, la Proposition 2.1.4 et l'Exemple 2.1.15 nous disent que $f|_{\mathbb{R}^2 \setminus \{(0,0)\}}$ possède toutes les dérivées directionnelles en tout point de $\mathbb{R}^2 \setminus \{(0, 0)\}$, i.e. $f$ possède toutes les dérivées directionnelles en tout point de $\mathbb{R}^2 \setminus \{(0, 0)\}$. On va montrer que $f$ possède toutes les dérivées directionnelles en $(0, 0)$. Soit $w = (a, b) \in \mathbb{R}^2$ non nul. Or,*

$$\frac{\partial f}{\partial w}(0, 0) = \lim_{t \to 0} \frac{f(ta, tb) - f(0, 0)}{t} = \lim_{t \to 0} \frac{ta(tb)^2}{t\big((ta)^2 + (tb)^4\big)} = \lim_{t \to 0} \frac{ab^2}{a^2 + t^2b^4}.$$

*C'est clair que si $a = 0$, alors la dérivée directionnelle précédente vaut zéro. En outre, si $a \neq 0$, on voit bien que*

$$\frac{\partial f}{\partial w}(0, 0) = \frac{b^2}{a}.$$

*En conséquence, toutes les dérivées directionnelles de $f$ existent en tout point de $\mathbb{R}^2$.*

*Par contre, on remarque que $f$ n'est pas continue en $(0, 0)$. Soit $\beta : \mathbb{R} \to \mathbb{R}^2$ l'application continue donnée par $\beta(t) = (t^2, t)$ pour $t \in \mathbb{R}$. Si $f$ est continue en $(0, 0)$, alors $f \circ \beta$ est continue en $t = 0$, vu que $\beta(0) = (0, 0)$. Or,*

$$(f \circ \beta)(t) = \frac{t^4}{t^4 + t^4} = \frac{1}{2}$$

*pour tout $t \in \mathbb{R} \setminus \{0\}$ et $(f \circ \beta)(0) = f(0, 0) = 0$, ce qui implique que $f \circ \beta$ n'est pas continue en $t = 0$. En conséquence, $f$ n'est pas continue en $(0, 0)$.*

On va démontrer le résultat le plus important de fonctions différentiables.

**Théorème 2.1.17.** *Soient $n, m, p \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ et $V \subseteq \mathbb{R}^m$ des parties ouvertes, $f : U \to V$, $g : V \to \mathbb{R}^p$ des applications. Soit $v_0 \in U$ et $w_0 = f(v_0) \in V$. Si $f$ est différentiable en $v_0$ et $g$ est différentiable en $w_0$, alors $g \circ f : U \to \mathbb{R}^p$ est différentiable en $v_0$ et*

$$D(g \circ f)(v_0) = Dg(w_0) \circ Df(v_0). \tag{2.1.5}$$

***Preuve.*** On choisit $s > 0$ tel que $B_{\| \ \|}(w_0, s) \subseteq V$. Comme $f$ est continue en $v_0$ (voir la Proposition 2.1.13), il existe $r > 0$, tel que $f(B_{\| \ \|}(v_0, r)) \subseteq B_{\| \ \|}(w_0, s)$. Or, comme $f$ est différentiable en $v_0$ et $g$ est différentiable en $w_0 = f(v_0)$, il existe des applications linéaires $L : \mathbb{R}^n \to \mathbb{R}^m$ et $L' : \mathbb{R}^m \to \mathbb{R}^p$, et des fonctions $\psi : B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \to \mathbb{R}^m$ et $\psi' : B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^m}, s) \to \mathbb{R}^p$ telles que

$$f(v_0 + h) = f(v_0) + L(h) + \psi(h)\|h\| \text{ et } g(w_0 + k) = g(w_0) + L'(k) + \psi'(k)\|k\| \tag{2.1.6}$$

pour tout $h \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$ et $k \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^m}, s)$ et que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \psi(h) = \mathbf{0}_{\mathbb{R}^m} \text{ et } \lim_{k \to \mathbf{0}_{\mathbb{R}^m}} \psi'(k) = \mathbf{0}_{\mathbb{R}^p}. \tag{2.1.7}$$

On remarque que $L = Df(v_0)$ et $L' = Dg(w_0)$.

On considère maintenant la deuxième égalité dans (2.1.6) avec $k = f(v_0+h) - f(v_0)$, *i.e.*

$$\begin{aligned}(g\circ f)(v_0+h) &= (g\circ f)(v_0) + L'\big(f(v_0+h)-f(v_0)\big) + \psi'(f(v_0+h)-f(v_0))\|f(v_0+h)-f(v_0)\| \\ &= (g\circ f)(v_0) + (L'\circ L)(h) + L'\big(\psi(h)\big)\|h\| + \psi'\big(f(v_0+h)-f(v_0)\big)\|f(v_0+h)-f(v_0)\|,\end{aligned}$$

pour tout $h \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$. On pose $\bar{\psi} : B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \to \mathbb{R}^p$ via

$$\bar{\psi}(h) = L'\big(\psi(h)\big) + \psi'\big(f(v_0+h)-f(v_0)\big)\frac{\|f(v_0+h)-f(v_0)\|}{\|h\|},$$

pour tout $h \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$. Or, comme $L'$ est une application linéaire entre espaces vectoriels de dimension finie, elle est continue (voir la Proposition 1.17.2). Cela implique que le premier opérande dans la définition de $\bar{\psi}(h)$ tend vers $\mathbf{0}_{\mathbb{R}^p}$, quand $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$, d'après la première égalité dans (2.1.7). En outre, à partir de la continuité de $f$ en $v_0$ (voir la Proposition 2.1.13), $f(v_0+h) - f(v_0)$ tend vers $\mathbf{0}_{\mathbb{R}^m}$, quand $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$. D'après la deuxième égalité dans (2.1.7), on conclut que $\psi'\big(f(v_0+h)-f(v_0)\big)$ tend vers $\mathbf{0}_{\mathbb{R}^p}$, quand $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$. Pour démontrer que le deuxième opérande dans la définition de $\bar{\psi}(h)$ tend vers $\mathbf{0}_{\mathbb{R}^p}$, quand $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$, il suffit de montrer qu'il existe $r > 0$ et $K > 0$ tels que

$$\frac{\|f(v_0+h)-f(v_0)\|}{\|h\|} \leq K$$

pour tout $h \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$. Or,

$$\begin{aligned}\frac{\|f(v_0+h)-f(v_0)\|}{\|h\|} &= \frac{\big\|Df(v_0)(h) + \psi(h)\|h\|\big\|}{\|h\|} \leq \frac{\|Df(v_0)(h)\| + \|\psi(h)\|\|h\|}{\|h\|} \\ &\leq \frac{\|Df(v_0)(h)\|}{\|h\|} + \|\psi(h)\| \leq C + \|\psi(h)\|,\end{aligned}$$

où $C > 0$ est une constante telle que $\|Df(v_0)(h)\| \leq C\|h\|$ pour tout $h \in \mathbb{R}^n$ (voir Propositions 1.17.1 et 1.17.2). D'après la première égalité dans (2.1.7), on trouve qu'il existe $r > 0$ et $K > 0$ tels que

$$\frac{\|f(v_0+h)-f(v_0)\|}{\|h\|} \leq K,$$

comme on voulait démontrer. En conséquence,

$$\lim_{h\to\mathbf{0}_{\mathbb{R}^n}} \bar{\psi}(h) = \mathbf{0}_{\mathbb{R}^p},$$

ce qui nous dit que $g\circ f$ est différentiable en $v_0$ et sa différentielle en $v_0$ est $L'\circ L = Dg(w_0)\circ Df(v_0)$. $\square$

**Proposition 2.1.18.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. Si $f$ est différentiable en $v_0$, alors pour tout $w \in \mathbb{R}^n$ non nul, la dérivée directionnelle de $f$ en $v_0$ dans la direction de $w$ existe et*

$$\frac{\partial f}{\partial w}(v_0) = Df(v_0)(w).$$

*En particulier, si $f$ est différentiable en $v_0$, toutes les dérivés partielles de $f$ en $v_0$ existent et $f$ est différentiable au sens de Gateaux en $v_0$.*

***Preuve.*** Il s'agit d'une conséquence directe du Théorème 2.1.17, appliqué à la fonction $f_w : J \to \mathbb{R}^m$ définie dans la Remarque 2.1.3. $\square$

**Définition 2.1.19.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. On suppose que les dérivées partielles de $f$ en $v_0$ existent. On écrit en plus $f(x) = (f_1(x), \dots, f_m(x))$, pour tout $x \in U$. La* ***matrice jacobienne*** *$J_f(v_0)$ de $f$ en $v_0$ est la matrice de taille $m \times n$ donnée par*

$$J_f(v_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(v_0) & \dots & \frac{\partial f_1}{\partial x_n}(v_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(v_0) & \dots & \frac{\partial f_m}{\partial x_n}(v_0) \end{pmatrix} \in \mathrm{M}_{n\times m}(\mathbb{R}).$$

*Si $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ est un vecteur, la* ***matrice colonne transposée*** *$h^t$ de $h$ est la matrice de taille $n \times 1$ donnée par $h_{i,1} = h_i$, pour tout $i \in [\![1, n]\!]$, i.e.*

$$h^t = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

*Si $m = 1$, i.e. $f : U \to \mathbb{R}$, la matrice $J_f(v_0)$ a taille $1 \times n$ et elle est précisément le vecteur*

$$\nabla f(v_0) = \Big(\frac{\partial f}{\partial x_1}(v_0), \ldots, \frac{\partial f}{\partial x_n}(v_0)\Big),$$

*appelé le* ***gradient*** *de $f$ en $v_0$.*

*D'après la Proposition 2.1.18, si $f$ est différentiable en $v_0 \in U$, alors*

$$\Big(Df(v_0)(h)\Big)^t = J_f(v_0) \cdot h^t,$$

*ou $\cdot$ dénote le produit matriciel.*

**Remarque 2.1.20.** *En termes des matrices jacobiennes, l'identité* (2.1.5) *dans le Théorème 2.1.17 s'écrit sous la forme*

$$J_{g\circ f}(v_0) = J_g\big(f(v_0)\big) \cdot J_f(v_0),$$

*où $\cdot$ dénote le produit matriciel.*

Soient $n \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}$ une application et $v_0 \in U$ dont toutes les dérivées directionnelles en $v_0$ existent. Une **direction de croissance maximale de $f$ en $v_0$** est un vecteur $w \in \mathbb{R}^n$ tel que $||w|| = 1$ et qui satisfait que

$$\frac{\partial f}{\partial w}(v_0) = \max\Big\{\frac{\partial f}{\partial u}(v_0) : u \in \mathbb{R}^n \text{ tel que } ||u|| = 1\Big\}.$$

**Corollaire 2.1.21.** *Soient $n \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}$ une application et $v_0 \in U$ Si $f$ est différentiable au sens de Gateaux en $v_0$ (e.g. $f$ est différentiable en $v_0$) et $\nabla f(v_0) \neq \mathbf{0}_{\mathbb{R}^n}$, alors $f$ possède une unique direction de croissance maximale en $v_0$ et elle est donnée par $\nabla f(v_0)/||\nabla f(v_0)||$.*

***Preuve.*** D'après la Proposition 2.1.18 et la Remarque 2.1.20, on a

$$\frac{\partial f}{\partial u}(v_0) = \langle \nabla f(v_0), u\rangle$$

pour tout $u \in \mathbb{R}^n$. L'inégalité de Cauchy-Schwarz (CS) nous dit alors que

$$\frac{\partial f}{\partial u}(v_0) = \langle \nabla f(v_0), u\rangle \leq ||\nabla f(v_0)||.||u|| = ||\nabla f(v_0)||$$

pour tout $u \in \mathbb{R}^n$ tel que $||u|| = 1$. En plus, l'égalité est vérifié si et seulement si $\nabla f(v_0) = \lambda u$ avec $\lambda \in \mathbb{R}_{>0}$, *i.e.* $u = \nabla f(v_0)/||\nabla f(v_0)||$. □

**Exemple 2.1.22.** *On va donner un exemple d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ différentiable au sens de Gateaux en $\mathbf{0}_{\mathbb{R}^n}$, i.e. toutes les dérivées directionnelles existent en $\mathbf{0}_{\mathbb{R}^n}$ et elles satisfont l'identité de la Proposition 2.1.18:*

$$\Big(\frac{\partial f}{\partial w}(\mathbf{0}_{\mathbb{R}^n})\Big)^t = J_f(\mathbf{0}_{\mathbb{R}^n}) \cdot w^t \tag{2.1.8}$$

*pour tout $w \in \mathbb{R}^n$ non nul, mais la fonction $f$ n'est pas continue en $\mathbf{0}_{\mathbb{R}^n}$, donc a fortiori pas différentiable en $\mathbf{0}_{\mathbb{R}^n}$.*

*Soit $C = \{(x, y) \in \mathbb{R}^2 : y = x^2\}$ et soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction donnée par*

$$f(x, y) = \begin{cases} 1, & \text{si } (x, y) \in C \setminus \{(0, 0)\}, \\ 0, & \text{si } (x, y) = (0, 0) \text{ ou } (x, y) \in \mathbb{R}^2 \setminus C. \end{cases}$$

*Soient $v_0 = (x_0, y_0) \in \mathbb{R}^2$ et $w = (a, b) \in \mathbb{R}^2$ non nul. On voit bien que $C \cap \{v_0 + tw : t \in \mathbb{R}\}$ est de cardinalité inférieure ou égal à 2, car $v_0 + tw \in C$ si et seulement si $(y_0 + tb)^2 - (y_0 + tb) = 0$. En conséquence,*

$$\frac{\partial f}{\partial w}(v_0) = \lim_{t \to 0} \frac{f(x_0 + ta, y_0 + tb) - f(x_0, y_0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0,$$

*pour tout $v_0 \in (\mathbb{R}^2 \setminus C) \cup \{(0,0)\}$, ce qui nous dit a fortiori que toutes les dérivées directionnelles de $f$ existent en $(0,0)$ et elles valent zéro. En outre, (2.1.8) est vérifiée immédiatement, car la matrice jacobienne de $f$ en $(0,0)$ est nulle, vu que les dérivées partielles de $f$ en $(0,0)$ s'annulent.*

*Par contre, on remarque que $f$ n'est pas continue en $(0,0)$. Soit $\beta : \mathbb{R} \to \mathbb{R}^2$ l'application continue donnée par $\beta(t) = (t, t^2)$ pour $t \in \mathbb{R}$. Si $f$ est continue en $(0,0)$, alors $f \circ \beta$ est continue en $t = 0$, vu que $\beta(0) = (0,0)$. Or,*

$$(f \circ \beta)(t) = 1$$

*pour tout $t \in \mathbb{R} \setminus \{0\}$ et $(f \circ \beta)(0) = f(0,0) = 0$, ce qui implique que $f \circ \beta$ n'est pas continue en $t = 0$. En conséquence, $f$ n'est pas continue en $(0,0)$, donc a fortiori pas différentiable en $(0,0)$.*

**Exemple 2.1.23.** *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une application différentiable en $(0,1)$ telle que*

$$\frac{\partial f}{\partial x}(0,1) = \sqrt{2} \text{ et } \frac{\partial f}{\partial v}(0,1) = 1,$$

*pour $v = (\sqrt{2}, \sqrt{2})/2$. On veut déterminer la direction $w \in \mathbb{R}^2$, avec $||w|| = 1$, telle que $\partial f / \partial w(0,1)$ soit maximal. Comme $f$ est différentiable en $(0,1)$, la Proposition 2.1.18 nous dit que*

$$\frac{\partial f}{\partial v}(0,1) = \langle \nabla f(0,1), v \rangle = \frac{\partial f}{\partial x}(0,1)\frac{\sqrt{2}}{2} + \frac{\partial f}{\partial y}(0,1)\frac{\sqrt{2}}{2},$$

*ce qui implique que*

$$1 = 1 + \frac{\partial f}{\partial y}(0,1)\frac{\sqrt{2}}{2}.$$

*En conséquence, $\nabla f(0,1) = (1,0)$. En outre, le Corollaire 2.1.21 nous dit que la direction de croissance maximale de $f$ en $(0,1)$ est $\nabla f(0,1) = (1,0)$, vu que ce vecteur a norme 1.*

## 2.2 Continuité des dérivées partielles et différentiabilité

**Définition 2.2.1.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. On dit que $f : U \to \mathbb{R}^m$ est de* ***classe*** $\mathrm{C}^1$ *en $v_0$ s'il existe un ouvert $V \subseteq U$ avec $v_0 \in V$ tel que toutes les dérivées partielles de $f$ existent en tout point $x \in V$ et les fonctions $D_i f : V \to \mathbb{R}^m$ sont continues sur $V$ pour tout $i \in [\![1, n]\!]$. En plus, $f : U \to \mathbb{R}^m$ est de* ***classe*** $\mathrm{C}^1$ *(sur $U$) si $f$ est de classe $\mathrm{C}^1$ en tout point $v \in U$.*

**Exemple 2.2.2.** *On continue avec l'Exemple 2.1.14. Soit $L : \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire. Comme la différentielle $DL(v)$ de $L$ en $v \in \mathbb{R}^n$ est précisément $L$, l'application $DL : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$ est constante, ce qui implique que les dérivées partielles $D_i L : \mathbb{R}^n \to \mathbb{R}^m$ sont aussi constantes pour tout $i \in [\![1, n]\!]$, donc continues. En conséquence, $L$ est de classe $\mathrm{C}^1$.*

*De la même, façon toute* ***application affine*** *$A : \mathbb{R}^n \to \mathbb{R}^m$ est de classe $\mathrm{C}^1$, où l'on rappelle qu'une application est dite affine s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ et un vecteur $w \in \mathbb{R}^m$ tels que $A(v) = L(v) + w$, pour tout $v \in \mathbb{R}^n$.*

**Exemple 2.2.3.** *C'est facile à vérifier que l'application $\mathfrak{s} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ donnée par l'addition vectorielle (voir l'Exemple 1.12.5) est aussi une application de classe $\mathrm{C}^1$.*

Le résultat suivant relie la notion de fonction de classe $\mathrm{C}^1$ et la différentiabilité.

**Proposition 2.2.4.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. Si $f$ est de classe $\mathrm{C}^1$ en $v_0$, alors $f$ est différentiable en $v_0$. En plus, si $V \subseteq U$ est une partie ouverte incluant $v_0$ telle que les dérivées partielles de $f$ en $x$ existent pour tout $x \in V$ et qu'elles sont continues, l'application $Df : V \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ est continue.*

***Preuve.*** Soit $V \subseteq U$ une partie ouverte qui inclut $v_0 = (v_{0,1}, \dots, v_{0,n})$ telle que les dérivées partielles de $f$ en $x$ existent pour tout $x \in V$ et qu'elles soient continues. Soit $r > 0$ tel que $B_{\| \ \|}(v_0, r) \subseteq V$ et soit $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ tel que $\|h\| < r$. Le Théorème des accroissements finis pour des fonctions d'une seule variable (2.3.2) nous dit que

$$\begin{aligned} f(v_{0,1} + h_1, \dots, v_{0,n}{+}h_n) - f(v_{0,1}, \dots, v_{0,n}) \\ &= \sum_{i=1}^{n} \Big( f(v_{0,1}, \dots, v_{0,i-1}, v_{0,i} + h_i, v_{0,i+1} + h_{i+1}, \dots, v_{0,n} + h_n) \\ &\qquad - f(v_{0,1}, \dots, v_{0,i-1}, v_{0,i}, v_{0,i+1} + h_{i+1}, \dots, v_{0,n} + h_n) \Big) \\ &= \sum_{i=1}^{n} D_i f(v_{0,1}, \dots, v_{0,i-1}, c_i, v_{0,i+1} + h_{i+1}, \dots, v_{0,n} + h_n) h_i, \end{aligned}$$

où $c_i \in \mathbb{R}$ appartient au segment déterminé par $v_{0,i}$ et $v_{0,i} + h_i$, pour tout $i \in [\![1, n]\!]$. La continuité des dérivées partielles implique qu'il existe une fonction $\varphi_i : B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \to \mathbb{R}^m$ telle que

$$D_i f(v_{0,1}, \dots, v_{0,i-1}, c_i, v_{0,i+1} + h_{i+1}, \dots, v_{0,n} + h_n) = D_i f(v_{0,1}, \dots, v_{0,n}) + \varphi_i(h)$$

et

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \varphi_i(h) = \mathbf{0}_{\mathbb{R}^m}, \tag{2.2.1}$$

pour tout $i \in [\![1, n]\!]$. En conséquence,

$$f(v_{0,1} + h_1, \dots, v_{0,n} + h_n) - f(v_{0,1}, \dots, v_{0,n}) = \sum_{i=1}^{n} D_i f(v_{0,1}, \dots, v_{0,n}) h_i + \sum_{i=1}^{n} \varphi_i(h) h_i.$$

On pose

$$\psi(h) = \sum_{i=1}^{n} \varphi_i(h) \frac{h_i}{\|h\|}$$

pour tout $h \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$ et $L(h) = \sum_{i=1}^{n} D_i f(v_{0,1}, \dots, v_{0,n}) h_i$ pour tout $h \in \mathbb{R}^n$. Comme

$$\|\psi(h)\| \le \sum_{i=1}^{n} \|\varphi_i(h)\| \frac{|h_i|}{\|h\|} \le \sum_{i=1}^{n} \|\varphi_i(h)\|,$$

où l'on a utilisé que $|h_i| \le \sqrt{h_1^2 + \dots + h_n^2} = \|h\|$, (2.2.1) nous dit que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \psi(h) = \mathbf{0}_{\mathbb{R}^m}.$$

En conséquence, $f$ est différentiable en $v_0$, comme on voulait démontrer.

Pour montrer le dernier résultat, on remarque que les applications $D_1 f, \dots, D_n f : V \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ sont continues, ce qui implique que l'application $J_f : V \to \mathrm{M}_{n \times m}(\mathbb{R})$ qui associe à $x \in V$ la matrice jacobienne $J_f(x)$ est continue (voir Proposition 1.12.8). En employant l'isomorphisme de la Remarque 1.17.5 et le fait que toutes les normes sur un espace vectoriel de dimension fini sont équivalentes, on conclut que $Df : V \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ est continue. □

**Remarque 2.2.5.** *La proposition précédente nous dit que, étant donné $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$, $f$ est de classe $\mathrm{C}^1$ en $v_0$ si et seulement s'il existe $V \subseteq U$ est une partie ouverte incluant $v_0$ telle que $f$ est différentiable en tout point $v \in V$ et l'application $Df : V \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ est continue.*

**Corollaire 2.2.6.** *Soient $n, m, p \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ et $V \subseteq \mathbb{R}^m$ des parties ouvertes, $f : U \to V$, $g : V \to \mathbb{R}^p$ des applications. Soit $v_0 \in U$ et $w_0 = f(v_0) \in V$. Si $f$ est de classe $\mathrm{C}^1$ en $v_0$ et $g$ est de classe $\mathrm{C}^1$ en $w_0$, alors $g \circ f : U \to \mathbb{R}^p$ est de classe $\mathrm{C}^1$ en $v_0$.*

***Preuve.*** Il s'agit d'une conséquence immédiate du Théorème 2.1.17, de la Proposition 2.2.4 et de la Remarque 2.2.5. □

**Exemple 2.2.7.** *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ donnée par $f(x_1, x_2) = x_1^2 + x_2^3$, et soit $w = (1, 2)$. On va montrer que la dérivée directionnelle de $f$ en $v_0 = (-1, 3)$ dans la direction de $w$ existe et on va calculer sa valeur. On note d'abord que $f$ est une fonction polynomiales, ce qui nous dit qu'elle est de classe $\mathrm{C}^1$. Les dérivées partielles sont*

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 2x_1, \qquad \textit{et} \qquad \frac{\partial f}{\partial x_2}(x_1, x_2) = 3x_2^2.$$

*Le gradient de $f$ en $(-1, 3)$ est alors $\nabla f(-1, 3) = (-2, 27)$. D'après Proposition 2.1.18 on a*

$$\frac{\partial f}{\partial w}(v_0) = \langle \nabla f(v_0), w \rangle = \langle (-2, 27), (1, 2) \rangle = 52.$$

**Proposition 2.2.8.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. On écrit $f(x) = (f_1(x), \dots, f_m(x))$, pour tout $x \in U$. Alors, $f$ est différentiable (resp., de classe $\mathrm{C}^1$) en $v_0$ si et seulement si $f_i$ est différentiable (resp., de classe $\mathrm{C}^1$) en $v_0$ pour tout $i \in [\![1, m]\!]$.*

***Preuve.*** On va faire seulement la preuve pour l'énoncé sur la différentiabilité, la preuve de l'énoncé sur la propriété $\mathrm{C}^1$ étant similaire.

Comme $f_i = \pi_i \circ f$, où $\pi_i : \mathbb{R}^n \to \mathbb{R}$ est l'application $(y_1, \dots, y_m) \mapsto y_i$, et $\pi_i$ est une application linéaire, le Théorème 2.1.17 nous dit que $f_i$ est différentiable en $v_0$ pour tout $i \in [\![1, m]\!]$ si $f$ est différentiable en $v_0$. Réciproquement, on suppose que $f_i$ est différentiable en $v_0$ pour tout $i \in [\![1, m]\!]$. D'après la Remarque 2.1.10, il existe une application linéaire $L_i : \mathbb{R}^n \to \mathbb{R}$ et une fonction $\psi_i : B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r_i) \to \mathbb{R}$ (où $r_i > 0$ est donné par la condition $B_{\|\ \|}(v_0, r_i) \subseteq U$) telles que

$$f_i(v_0 + h) - f_i(v_0) = L_i(h) + \psi_i(h)\|h\|,$$

pour tout $h \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r_i))$ et que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \psi_i(h) = 0.$$

On pose $r = \min(r_1, \dots, r_n) > 0$, $L : \mathbb{R}^n \to \mathbb{R}^m$ via $L(h) = (L_1(h), \dots, L_m(h))$, pour tout $h \in \mathbb{R}^n$ et $\psi : B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r) \to \mathbb{R}^m$ via $\psi(h) = (\psi_1(h), \dots, \psi_m(h))$, pour tout $h \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, r)$. Alors, $f(v_0 + h) - f(v_0) = L(h) + \psi(h)\|h\|$ et

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \psi(h) = \mathbf{0}_{\mathbb{R}^m},$$

ce qui implique que $f$ est différentiable en $v_0$. □

Le résultat suivant est standard.

**Proposition 2.2.9.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f, g : U \to \mathbb{R}^m$ et $\lambda : U \to \mathbb{R}$ des applications différentiables (resp., de classe $\mathrm{C}^1$) en $v_0$. Alors, l'application $f + \lambda \cdot g : U \to \mathbb{R}^m$ est différentiable (resp., de classe $\mathrm{C}^1$) en $v_0$, et*

$$D(f + \lambda \cdot g)(v_0)(w) = Df(v_0)(w) + D\lambda(v_0)(w)g(v_0) + \lambda(v_0)Dg(v_0)(w),$$

*pour tout $w \in \mathbb{R}^n$. En outre, si $\lambda(v) \neq 0$ pour tout $v \in U$, l'application $\mu : U \to \mathbb{R}$ donnée par $\mu(v) = 1/\lambda(v)$ pour $v \in U$ est différentiable (resp., de classe $\mathrm{C}^1$) en $v_0$, et*

$$D\mu(v_0) = -\frac{1}{\lambda(v_0)^2} D\lambda(v_0),$$

*pour tout $v \in U$.*

***Preuve.*** Il s'agit d'une conséquence du Théorème 2.1.17, car $f + \lambda g$ s'écrit comme la composition de l'application $(f, \lambda, g) : U \to \mathbb{R}^{2m+1}$ donnée par $x \mapsto (f(x), \lambda(x), g(x))$, $\mathrm{id}_{\mathbb{R}^m} \times \mathfrak{p} : \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ et $\mathfrak{s} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ (voir l'Exemple 1.12.5). On remarque que $\mathrm{id}_{\mathbb{R}^m} \times \mathfrak{p}$ et $\mathfrak{s}$ sont de classe $\mathrm{C}^1$ (voir les Exemples 2.2.2 et 2.2.3, ainsi que l'argument dans la preuve du Corollaire 1.12.9), et que $(f, \lambda, g)$ est différentiable (resp., de classe $\mathrm{C}^1$) en $v_0$ d'après la Proposition 2.2.8.

Pour démontrer la dernière partie, on remarque que $\mu$ est la composition de l'application $\lambda : U \to \mathbb{R}$ et de l'application $\mathrm{inv} : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$ donnée par $\mathrm{inv}(x) = x^{-1}$ pour tout $x \in \mathbb{R} \setminus \{0\}$. Comme $\mathrm{inv}$ est de classe $\mathrm{C}^1$ et $\mathrm{inv}'(x) = -x^{-2}$, $\mu$ est différentiable (resp., de classe $\mathrm{C}^1$) en $v_0$ et $D\mu(v_0) = -\lambda(v_0)^{-2} D\lambda(v_0)$, d'après le Théorème 2.1.17. □

**Exemple 2.2.10.** *On va donner un exemple d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ dont toutes les dérivées directionnelles existent en tout point et elles satisfont l'identité de la Proposition 2.1.18, i.e.*

$$\Big(\frac{\partial f}{\partial w}(v_0)\Big)^t = J_f(v_0) \cdot w^t \tag{2.2.2}$$

*pour tous $v_0 \in \mathbb{R}^n$ et $w \in \mathbb{R}^n$ non nul, mais $f$ n'est pas différentiable en tout point.*

*Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction donnée par*

$$f(x,y) = \begin{cases} \frac{x^5y^5}{x^8+y^{12}}, & \text{si } (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\} \text{ avec } x \neq 0, \\ 0, & \text{si } (x,y) = (0,0). \end{cases}$$

*Comme l'application $f|_{\mathbb{R}^2\setminus\{(0,0)\}}$ est un quotient de deux polynômes, qui sont a fortiori de classe $\mathrm{C}^1$, la Proposition 2.2.9 nous dit que $f|_{\mathbb{R}^2\setminus\{(0,0)\}}$ est de classe $\mathrm{C}^1$, donc a fortiori différentiable. D'après la Proposition 2.1.18, toutes les dérivées directionnelles de $f$ existent en tout point de $\mathbb{R}^2 \setminus \{(0,0)\}$, vu que* (2.2.2) *est vérifiée pour tout $v_0$ non nul. Or, on affirme que toutes les dérivées partielles de $f$ en $(0,0)$ existent et elles valent zéro. En effet, étant donné $w = (a,b)$ non nul, on a*

$$\frac{\partial f}{\partial w}(0,0) = \lim_{t\to 0}\frac{t^2a^5b^5}{a^8+t^4b^{12}} = 0.$$

*En particulier,* (2.2.2) *est vérifiée pour $v_0 = (0,0)$.*

*Finalement, on va montrer que $f$ n'est pas différentiable en $(0,0)$. En effet, $f$ est différentiable en $(0,0)$ si et seulement si*

$$\lim_{(h,k)\to(0,0)}\frac{f(h,k)}{\sqrt{h^2+k^2}} = 0,$$

*vu que $f(0,0) = 0$ et $Df(0,0) : \mathbb{R}^2 \to \mathbb{R}$ est l'application linéaire nulle, car les dérivées partielles de $f$ en $(0,0)$ s'annulent. Or, si la limite précédente est vérifiée, on a a fortiori*

$$0 = \lim_{t\to 0}\frac{f(t^3,t^2)}{\sqrt{t^4+t^6}} = \lim_{t\to 0}\frac{1}{2t\sqrt{1+t^2}} = +\infty,$$

*ce qui est absurde. En conséquence, $f$ n'est pas différentiable en $(0,0)$. Noter que $f$ est continue en $(0,0)$ par l'Exemple 1.12.10.*

## 2.3 Théorème des accroissements finis

### 2.3.1 Théorème des accroissements finis pour fonctions d'une seule variable

On rappelle les résultats suivants pour des fonctions scalaires d'une seule variable.

**Proposition 2.3.1.** *Soient $a, b \in \mathbb{R}$ avec $a < b$ et $g : [a,b] \to \mathbb{R}$ une fonction continue (sur $[a,b]$) et dérivable sur $]\,a,b\,[$. Soit $x_0 \in\, ]\,a,b\,[$ tel que*

$$f(x_0) \geq f(x) \ \Big(\textit{resp.}, f(x_0) \leq f(x)\Big) \tag{2.3.1}$$

*pour tout $x \in\, ]\,a,b\,[$. Alors $f'(x_0) = 0$*

***Preuve.*** L'inégalité (2.3.1) nous dit que

$$f'(x_0) = \lim_{h\to 0+}\frac{f(x_0+h)-f(x_0)}{h} \leq 0 \text{ et } f'(x_0) = \lim_{h\to 0-}\frac{f(x_0+h)-f(x_0)}{h} \geq 0,$$

resp.,

$$f'(x_0) = \lim_{h\to 0+}\frac{f(x_0+h)-f(x_0)}{h} \geq 0 \text{ et } f'(x_0) = \lim_{h\to 0-}\frac{f(x_0+h)-f(x_0)}{h} \leq 0,$$

par définition de dérivée. En conséquence, $f'(x_0) = 0$. □

**Proposition 2.3.2** (Théorème des accroissements finis de Rolle). *Étant donné $a, b \in \mathbb{R}$ avec $a < b$ et une fonction $h : [a, b] \to \mathbb{R}$ continue (sur $[a, b]$) et dérivable sur $]\, a, b\, [$, telle que $h(a) = h(b)$. Alors, il existe $c \in \,]\, a, b\, [$ tel que $h'(c) = 0$.*

***Preuve.*** Si $h$ est constante, alors $h'(x) = 0$ pour tout $x \in \,]\, a, b\, [$, et on peut donc choisir $c \in \,]\, a, b\, [$ un point quelconque. On suppose que $h$ n'est pas constante. Alors, on peut supposer sans perte de généralité qu'il existe $x \in \,]\, a, b\, [$ tel que $h(x) > h(a)$, sinon on remplace $h$ par $-h$. D'après la Proposition 1.15.2, $h$ possède un maximum $c \in \,]\, a, b\, [$. La proposition précédente nous dit alors que $h'(c) = 0$. □

**Proposition 2.3.3** (Théorème des accroissements finis de Lagrange). *Étant donné $a, b \in \mathbb{R}$ avec $a < b$ et une fonction $g : [a, b] \to \mathbb{R}$ continue (sur $[a, b]$) et dérivable sur $]\, a, b\, [$, il existe $c \in \,]\, a, b\, [$ tel que*

$$g(b) - g(a) = (b - a)g'(c). \tag{2.3.2}$$

***Preuve.*** Soit $h : [a, b] \to \mathbb{R}$ l'application donnée par

$$h(x) = g(x) - \frac{g(b) - g(a)}{b - a}(x - a)$$

pour $x \in [a, b]$. On voit bien que $h(a) = g(a) = h(b)$. La Proposition 2.3.2 nous dit qu'il existe $c \in \,]\, a, b\, [$ tel que $h'(c) = 0$. Or,

$$0 = h'(c) = f'(c) - \frac{g(b) - g(a)}{b - a},$$

ce qui montre le résultat. □

### 2.3.2 Théorème des accroissements finis pour fonctions de plusieurs variables

On peut améliorer légèrement la Proposition 2.3.3.

**Proposition 2.3.4** (Théorème des accroissements finis I). *Soient $n \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}$ une application et $v_0 \in U$. Soit $v \in U$ tel que $I_{v_0,v} = \{tv + (1 - t)v_0 : t \in [0, 1]\}$ est inclus dans $U$. Alors, si $f$ est différentiable, il existe $t_0 \in \,]\, 0, 1\, [$ tel que*

$$f(v) - f(v_0) = Df\big(t_0 v + (1 - t_0)v_0\big)(v - v_0).$$

***Preuve.*** Soit $\alpha : [0, 1] \to U$ l'application donnée par $\alpha(t) = tv + (1 - t)v_0$. C'est clair que $\alpha$ est différentiable, sur $]\, 0, 1\, [$, car $\alpha'(t) = v - v_0$ pour tout $t \in \,]\, 0, 1\, [$. Comme $f$ est différentiable, $g = f \circ \alpha$ est différentiable sur $]\, 0, 1\, [$. De la même façon, on note que $g$ est continue sur $[0, 1]$. En conséquence, (2.3.2) nous dit qu'il existe $t_0 \in \,]\, 0, 1\, [$ tel que

$$f(v) - f(v_0) = g(1) - g(0) = \Big(Df_i\big(\alpha(t_0)\big) \circ D\alpha(t_0)\Big)(1) = Df\big(\alpha(t_0)\big)(v - v_0),$$

où l'a utilisé le Théorème 2.1.17. □

En général, pour des fonctions $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ avec $m > 1$, on ne peut pas obtenir une identité de la forme (2.3.2), comme l'exemple suivant le montre.

**Exemple 2.3.5.** *Soit $\alpha : \mathbb{R} \to \mathbb{R}^2$ l'application $\alpha(t) = (\cos(t), \sin(t))$ pour $t \in \mathbb{R}$. C'est clair que $\alpha$ est différentiable. Alors, $\mathbf{0}_{\mathbb{R}^2} = \alpha(2\pi) - \alpha(0)$ mais $(2\pi - 0)\alpha'(t) \neq \mathbf{0}_{\mathbb{R}^2}$ pour tout $t \in [0, 2\pi]$, vu que $\|\alpha'(t)\|^2 = (-\sin(t))^2 + (\cos(t))^2 = 1$ pour tout $t \in \mathbb{R}$.*

On peut par contre obtenir le résultat suivant.

**Proposition 2.3.6.** *[Théorème des accroissements finis II] Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. Soit $v \in U$ tel que $I_{v_0,v} = \{tv + (1 - t)v_0 : t \in [0, 1]\}$ est inclus dans $U$. Alors, si $f$ est de classe $\mathrm{C}^1$,*

$$\big\|f(v) - f(v_0)\big\| \leq \|v - v_0\| \sup\big\{\|Df(w)\| : w \in I_{i_0,v}\big\},$$

*où $\|Df(w)\|$ désigne la norme de $Df(w) \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ associée aux normes euclidiennes de $\mathbb{R}^n$ et $\mathbb{R}^m$ (voir la Proposition 1.17.3).*

*Preuve.* Soit $\alpha : [0,1] \to U$ l'application donnée par $\alpha(t) = tv + (1-t)v_0$. C'est clair que $\alpha$ est de classe $C^1$, sur $]0,1[$, car $\alpha'(t) = v - v_0$ pour tout $t \in ]0,1[$. On écrit $f(v) = (f_1(v), \dots, f_m(v))$ pour tout $v \in U$. Comme $f$ est de classe $C^1$, $g = f \circ \alpha$ est de classe $C^1$, sur $]0,1[$. De la même façon, on note que $g$ est continue sur $[0,1]$. En conséquence, (2.3.2) nous dit que

$$f(v) - f(v_0) = g(1) - g(0) = \int_0^1 g'(s)ds = \int_0^1 Dg(s)(1)ds = \int_0^1 Df\big(\alpha(s)\big)(v - v_0)ds,$$

où l'a utilisé le Théorème 2.1.17. En conséquence,

$$\begin{aligned}\|f(v) - f(v_0)\| &= \left\| \int_0^1 Df(\alpha(s))(v - v_0)ds \right\| \leq \int_0^1 \left\| Df\big(\alpha(s)\big)(v - v_0)\right\| ds \\ &\leq \|v - v_0\| \int_0^1 \left\|Df\big(\alpha(s)\big)\right\| ds \leq \|v - v_0\| \sup\left\{\|Df(w)\| : w \in I_{i_0,v}\right\},\end{aligned}$$

comme on voulait démontrer. □

**Proposition 2.3.7.** *Soient $n, m \in \mathbb{Z}_{>0}$ et $U \subseteq \mathbb{R}^n$ une partie ouverte. On suppose que, étant donné $v, w \in U$ il existe une application $\alpha : [a,b] \to U$ continue et dérivable par morceaux avec $a < b$ réels (i.e. il existe $\{a = t_0 < \dots < t_n = b\}$ avec $n \in \mathbb{Z}_{>0}$ telle que $\alpha|_{]t_i,t_{i+1}[}$ est dérivable pour tout $i \in [\![0, n-1]\!]$) telle que $\alpha(a) = v$ et $\alpha(b) = w$. Soit $f : U \to \mathbb{R}^m$ une application différentiable telle que $Df(v) = \mathbf{0}_{\mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)}$ pour tout $v \in U$. Alors, $f$ est constante, i.e. $f(v) = f(w)$ pour tous $v, w \in U$.*

*Preuve.* Il suffit de montrer que, étant donné $v, w \in U$, $f(v) = f(w)$. Soit $\alpha : [a,b] \to U$ continue et dérivable par morceaux telle que $\alpha(a) = v$ et $\alpha(b) = w$. Soit $g = f \circ \alpha : [a,b] \to \mathbb{R}^m$. Alors, $g$ est dérivable par morceaux. En outre, $g'(t) = Dg(t)(1) = (Df(\alpha(t)) \circ D\alpha(t))(1) = Df(\alpha(t))(\alpha'(t)) = \mathbf{0}_{\mathbb{R}^m}$ pour tout $t \in ]t_i, t_{i+1}[$ et $i \in [\![0, n-1]\!]$, ce qui implique, $g(t_i) = g(t_{i+1})$ pour tout $i \in [\![0, n-1]\!]$, en employant (2.3.2). En particulier, $f(v) = g(a) = g(b) = f(w)$. □

On rappelle qu'une partie $S \subseteq \mathbb{R}^n$ est dite **étoilée** s'il existe $v_0 \in S$ tel que $I_{v_0,v} = \{tv + (1-t)v_0 : t \in [0,1]\}$ est inclus dans $S$ pour tout $v \in S$. On dit dans ce cas que $v_0$ est un **centre** de $S$. Noter qu'une partie étoilée satisfait les hypothèses de la proposition précédente.

## 2.4 Quelques applications géométriques

Étant donné une partie non vide $S \subseteq \mathbb{R}^n$ et $v_0 \in S$, on dira qu'un vecteur $w \in \mathbb{R}^n$ est **tangent** à $S$ en $v_0$ s'il existe une application $\alpha : J \to S$ de classe $C^1$ telle que $\alpha(0) = v_0$ et $\alpha'(0) = w$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert tel que $0 \in J$. On notera $T_{v_0}S \subseteq \mathbb{R}^n$ l'ensemble de $\mathbb{R}^n$ formé de tous les vecteurs tangents à $S$ en $v_0 \in S$. On l'appelle **espace tangent** à $S$ en $v_0$. C'est clair que $\mathbf{0}_{\mathbb{R}^n} \in T_{v_0}S$, car l'application $\alpha : \mathbb{R} \to S$ donnée par $\alpha(t) = v_0$ est de classe $C^1$ et $\alpha'(t) = \mathbf{0}_{\mathbb{R}^n}$ pour tout $t \in \mathbb{R}$.

### 2.4.1 Ensemble de niveau

Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une fonction différentiable. On rappelle que l'**ensemble de niveau** $c \in \mathbb{R}$ de $f$ est donné par

$$L_{f,c} = \{x \in U : f(x) = c\}.$$

Si $n = 2$, on appelle $L_{f,c}$ la **ligne de niveau** $c \in \mathbb{R}$ de $f$.

Le lemme suivant décrit l'espace tangent des ensembles de niveau de fonctions différentiables $f : U \to \mathbb{R}$ en tout point où le gradient est non nul.

**Lemme 2.4.1.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une fonction différentiable. Soit $v_0 \in L_{f,c}$ tel que $\nabla f(v_0) \neq \mathbf{0}_{\mathbb{R}^n}$. Alors, $T_{v_0}L_{f,c} = \{v \in \mathbb{R}^n : \langle v, \nabla f(v_0)\rangle = 0\}$.*

***Preuve.*** On va démontrer la première partie. Soit $\alpha : J \to L_{f,c}$ une application de classe $\mathrm{C}^1$ telle que $\alpha(0) = v_0$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert tel que $0 \in J$. Alors, comme la fonction $f \circ \alpha$ est constante (avec valeur $c$), sa dérivée est nulle. Le Théorème 2.1.17 nous dit que

$$0 = (f \circ \alpha)'(t) = \langle \nabla f\big(\alpha(t)\big), \alpha'(t) \rangle,$$

pour tout $t \in J$. En particulier, pour $t = 0$ on a

$$0 = (f \circ \alpha)'(0) = \langle \nabla f(v_0), \alpha'(0) \rangle,$$

*i.e.* $T_{v_0} L_{f,c} \subseteq \{v \in \mathbb{R}^n : \langle v, \nabla f(v_0)\rangle = 0\}$. Pour démontrer l'inclusion $T_{v_0} L_{f,c} \supseteq \{v \in \mathbb{R}^n : \langle v, \nabla f(v_0)\rangle = 0\}$, il suffit de montrer que, étant donné $w \in \mathbb{R}^n$ tel que $0 = \langle \nabla f(v_0), w\rangle$, il existe une application $\alpha : J \to L_{f,c}$ de classe $\mathrm{C}^1$ (avec $J \subseteq \mathbb{R}$ un intervalle ouvert qui inclut 0) tel que $\alpha(0) = v_0$ et $w = \alpha'(0)$. Cela est un cas particulier du Lemme 2.7.15. □

Le lemme précédent peut s'interpréter de la façon suivante.

**Interpretation 2.4.2.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une fonction différentiable. Soit $v_0 \in L_{f,c}$ dans l'ensemble de niveau de $f$, où $c \in \mathbb{R}$, tel que $\nabla f(v_0) \neq \mathbf{0}_{\mathbb{R}^n}$. Alors, le vecteur gradient $\nabla f(v_0)$ est orthogonal à l'ensemble de niveau $L_{f,c}$ en $v_0$.*

### 2.4.2 Le plan tangent à une surface paramétrée

**Définition 2.4.3.** *Soit $U \subseteq \mathbb{R}^2$ une partie ouverte. Une* ***surface paramétrique*** *de classe $\mathrm{C}^1$ dans $\mathbb{R}^n$ est une application $\sigma : U \to \mathbb{R}^n$ de classe $\mathrm{C}^1$. On écrira $\sigma(x_1, x_2) = (\sigma_1(x_1,x_2), \dots, \sigma_n(x_1,x_2))$, pour tout $x = (x_1, x_2) \in U$. On dit que la surface paramétrique de classe $\mathrm{C}^1$ est* ***régulière*** *si l'application linéaire $D\sigma(x) : \mathbb{R}^2 \to \mathbb{R}^3$ est injective pour tout $x \in U$, i.e. l'ensemble*

$$\left\{ \frac{\partial \sigma}{\partial x_1}(x), \frac{\partial \sigma}{\partial x_2}(x) \right\} \subseteq \mathbb{R}^n$$

*est libre pour tout $x \in U$, où*

$$\frac{\partial \sigma}{\partial x_i}(x) = \left( \frac{\partial \sigma_1}{\partial x_i}(x), \dots, \frac{\partial \sigma_n}{\partial x_i}(x) \right),$$

*pour $i = 1, 2$.*

**Remarque 2.4.4.** *Soit $U \subseteq \mathbb{R}^2$ une partie ouverte et soit $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^1$. Alors, l'application $\sigma : U \to \mathbb{R}^3$ définie via $\sigma(x_1,x_2) = (x_1, x_2, f(x_1,x_2))$ est une surface paramétrique de classe $\mathrm{C}^1$ régulière dans $\mathbb{R}^3$. La plupart des exemples que l'on va considérer seront de cette forme.*

**Exemple 2.4.5.** *Soit $U = \mathbb{R}^2$ et soit $f : \mathbb{R}^2 \to \mathbb{R}$ l'application donnée par $f(x_1,x_2) = x_1^2 - x_2^2$. La remarque précédente nous donne une surface paramétrique de classe $\mathrm{C}^1$ régulière $\sigma : \mathbb{R}^2 \to \mathbb{R}^3$ via $\sigma(x_1,x_2) = (x_1, x_2, x_1^2 - x_2^2)$.*

Le lemme suivant décrit l'espace tangent à une surface paramétrique régulière. La preuve de la deuxième partie se base sur le Théorème 2.8.6, que l'on démontrera plus tard.

**Lemme 2.4.6.** *Soit $U \subseteq \mathbb{R}^2$ une partie ouverte et $\sigma : U \to \mathbb{R}^n$ surface paramétrique de classe $\mathrm{C}^1$ régulière dans $\mathbb{R}^n$. On pose $S = \mathrm{Img}(\sigma)$. Alors, l'espace tangent $T_{v_0}S$ à $S$ en $v_0 = \sigma(a_0)$, où $a_0 \in U$, coïncide avec $\mathrm{Img}(D\sigma(a_0))$, i.e. le sous-espace vectoriel de $\mathbb{R}^n$ donné par engendré par*

$$\left\{ \frac{\partial \sigma}{\partial x_1}(a_0), \frac{\partial \sigma}{\partial x_2}(a_0) \right\} \subseteq \mathbb{R}^n. \tag{2.4.1}$$

***Preuve.*** Soit $r > 0$ tel que $B_{||\ ||}(a_0, 2r) \subseteq U$. Étant donné $b = (b_1, b_2) \in \mathbb{R}^2$ tel que $||b|| < r$, soit $\beta : J \to U$ l'application de classe $\mathrm{C}^1$ donnée par $\beta(t) = a_0 + tb$ pour $t \in J$, où $J = ]-2, 2[ \subseteq \mathbb{R}$ est un intervalle ouvert tel que $0 \in J$. Noter que $\beta'(t) = b$ pour tout $t \in J$. Alors, $\alpha = \sigma \circ \beta : J \to S$ est une application de classe $\mathrm{C}^1$ telle que $\alpha(0) = v_0$. Le Théorème 2.1.17 nous dit que

$$\alpha'(0) = (\sigma \circ \beta)'(0) = \big(D\sigma(a_0) \circ D\beta(0)\big)(1) = D\sigma(a_0)(\beta'(0)) = D\sigma(a_0)(b) = b_1 \frac{\partial \sigma}{\partial x_1}(a_0) + b_2 \frac{\partial \sigma}{\partial x_2}(a_0),$$

ce qui implique que (2.4.1) est inclus dans $T_{v_0}S$, et, en conséquence, $\mathrm{Img}(D\sigma(a_0))$ est inclus dans $T_{v_0}S$.

On va montrer que $T_{v_0}S$ est inclus dans (2.4.1). Soit $W \subseteq \mathbb{R}^n$ un sous-espace vectoriel de dimension $n-2$ tel que $W \cap \mathrm{Img}(D\sigma(a_0)) = \{\mathbf{0}_{\mathbb{R}^n}\}$ et soit $\hat{\sigma} : U \times W \to \mathbb{R}^n$ l'application donnée par $\hat{\sigma}(x,w) = \sigma(x) + w$, pour tout $x \in U$ et $w \in W$. Comme $D\hat{\sigma}(a_0, \mathbf{0}_W)$ est inversible, vu que

$$J_{\hat{\sigma}}(a_0, \mathbf{0}_W) = \begin{pmatrix} D\sigma(a_0) & \mathbf{0}_{\mathbb{R}^{2\times(n-2)}} \\ \mathbf{0}_{\mathbb{R}^{(n-2)\times 2}} & \mathrm{I}_{n-2} \end{pmatrix},$$

où $I_{n-2}$ est la matrice identité d'ordre $n-2$, le Théorème 2.8.6 nous dit qu'il existe des ouverts $U' \subseteq U$ et $U'' \subseteq W$ tels que $a_0 \in U$, $\mathbf{0}_W \in U''$, $\hat{\sigma}(U' \times U'') \subseteq \mathbb{R}^n$ est ouvert, l'application $\hat{\sigma}|_{U' \times U''} : U' \times U'' \to \hat{\sigma}(U' \times U'')$ est bijective, et son application inverse $\hat{\pi} : \hat{\sigma}(U' \times U'') \to U' \times U''$ est de classe $\mathrm{C}^1$. Soit $\alpha : J \to S$ une application de classe $\mathrm{C}^1$ telle que $\alpha(0) = v_0$, où $J$ est un intervalle ouvert incluant $0$. Si l'on remplace $J$ par $\alpha^{-1}(S \cap \hat{\sigma}(U' \times U''))$ et $\alpha$ par $\alpha|_{\alpha^{-1}(S \cap \hat{\sigma}(U' \times U''))}$, on peut supposer sans perte de généralité que l'image de $\alpha : J \to S$ est incluse dans $S \cap \hat{\sigma}(U' \times U'')$. On pose $\beta : \pi_U \circ \hat{\pi} \circ \alpha$, où $\pi_U : U \times W \to U$ est la projection canonique qui envoie $(u,w) \in U \times W$ dans $u \in U$. Alors, $\beta$ est de classe $\mathrm{C}^1$, $\beta(0) = a_0$ et $\alpha = \sigma \circ \beta$, vu que

$$\begin{aligned} \alpha(t) &= (\hat{\sigma} \circ \hat{\pi})\big(\alpha(t)\big) = (\hat{\sigma} \circ \pi_U \circ \hat{\pi})\big(\alpha(t)\big) + (\hat{\sigma} \circ \pi_W \circ \hat{\pi})\big(\alpha(t)\big) \\ &= (\sigma \circ \pi_U \circ \hat{\pi})\big(\alpha(t)\big) + (\pi_W \circ \hat{\pi})\big(\alpha(t)\big) = (\sigma \circ \pi_U \circ \hat{\pi})\big(\alpha(t)\big) \\ &= (\sigma \circ \beta)(t), \end{aligned}$$

pour tout $t \in J$, où $\pi_W : U \times W \to W$ est la projection canonique qui envoie $(u,w) \in U \times W$ dans $w \in W$, et l'on a utilisé que $(\pi_W \circ \hat{\pi})(\alpha(t)) = \mathbf{0}_{\mathbb{R}^n}$ vu que $\alpha(t) \in S = \mathrm{Img}(\sigma)$. Comme $\alpha = \sigma \circ \beta$, $\alpha'(0) = D\sigma(a_0)(\beta'(0))$, ce qui nous dit que $T_{v_0}S$ est inclus dans $\mathrm{Img}(D\sigma(a_0))$. $\square$

On suppose les mêmes hypothèse que dans le lemme précédent. Le **plan tangent affine** à $\sigma$ en $\sigma(a)$, où $a \in U$, est le sous-espace affine de $\mathbb{R}^n$ donné par $\{\sigma(a) + v : v \in T_{\sigma(a)}S\}$. En fait, le Lemme 2.4.6 nous dit que le plan tangent affine à $\sigma$ en $\sigma(a)$ est l'image donnée par la surface paramétrique de classe $\mathrm{C}^1$ régulière $\varsigma : \mathbb{R}^2 \to \mathbb{R}^n$ donnée par

$$\varsigma(s,t) = \sigma(a) + s \cdot \frac{\partial \sigma}{\partial x_1}(a) + t \cdot \frac{\partial \sigma}{\partial x_2}(a)$$

pour tous $s, t \in \mathbb{R}$. Si $n = 3$, ce plan est donné par

$$\Big\{ x \in \mathbb{R}^3 : \Big\langle \big(x - \sigma(a)\big), N \Big\rangle = 0 \Big\},$$

où $N$ est le produit vectoriel

$$N = \frac{\partial \sigma}{\partial x_1}(a) \wedge \frac{\partial \sigma}{\partial x_2}(a) \in \mathbb{R}^3.$$

## 2.5 Différentiabilité d'ordre supérieure

**Définition 2.5.1.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. On dira qu'une application $f : U \to \mathbb{R}^m$ est de **classe** $\mathrm{C}^0$ si elle est continue. On procède par récurrence pour définir la notion suivante. On dit que $f : U \to \mathbb{R}^m$ est de **classe** $\mathrm{C}^{p+1}$ (sur $U$), pour $p \in \mathbb{N}$, si les applications $D_1 f, \dots, D_n f : U \to \mathbb{R}^m$ existent et elles sont de classe $\mathrm{C}^p$ (sur $U$). Finalement, on dit que $f : U \to \mathbb{R}^m$ est de **classe** $\mathrm{C}^\infty$ (sur $U$) si $f$ est de classe $\mathrm{C}^p$ (sur $U$) pour tout $p \in \mathbb{N}$.*

**Remarque 2.5.2.** *On remarque que le Corollaire 2.2.6, et les Propositions 2.2.8 et 2.2.9 sont valables si l'remplace "de classe $\mathrm{C}^1$" par "de classe $\mathrm{C}^p$ avec $p \in \mathbb{Z}_{>0} \cup \{\infty\}$", avec les mêmes arguments.*

Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^p$ sur $U$. On rappelle que les dérivées partielles successives $D_{i_1} \dots D_{i_\ell} f(x)$, où $i_1, \dots, i_\ell \in [\![1, n]\!]$, $\ell \in [\![1, p]\!]$ et $x \in U$, sont définies par récurrence via

$$D_{i_1} \dots D_{i_\ell} f(x) = D_{i_1}\Big(D_{i_2} \dots D_{i_\ell} f(x)\Big),$$

pour tout $x \in U$. Normalement, on note $D_{i_1} \dots D_{i_\ell} f(x)$ aussi

$$\frac{\partial^\ell f}{\partial x_{i_1} \dots \partial x_{i_\ell}}(x).$$

**Exemple 2.5.3.** *On continue avec l'Exemple 2.2.2. Soit $L : \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire. On a vu dans l'Exemple 2.2.2 que les dérivées partielles $D_i L : \mathbb{R}^n \to \mathbb{R}^m$ sont constantes pour tout $i \in [\![1, n]\!]$, donc linéaires. Un argument direct par récurrence nous dit que $L$ est de classe $\mathrm{C}^\infty$. De la même, façon toute application affine $A : \mathbb{R}^n \to \mathbb{R}^m$ est de classe $\mathrm{C}^\infty$.*

Le résultat suivant est immédiat, et généralise la Remarque 2.2.5.

**Proposition 2.5.4.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}^m$ une application. Alors, étant donné $p \in \mathbb{Z}_{>0}$, $f$ est de classe $\mathrm{C}^p$ si et seulement si $f$ est différentiable et l'application $Df : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ est de classe $\mathrm{C}^{p-1}$.*

***Preuve.*** Le cas $p = 1$ est suite directement de la Remarque 2.2.5. On suppose désormais $p \geq 2$. On suppose que $f$ est différentiable et l'application $Df : U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ est de classe $\mathrm{C}^{p-1}$. Étant donné $i \in [\![1, n]\!]$, soit $\mathrm{ev}_{e_i} : \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}^m$ l'application donnée par $\mathrm{ev}_{e_i}(L) = L(e_i)$ pour tout pour tout $L \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, où $\{e_1, \dots, e_n\}$ désigne la base canonique de $\mathbb{R}^n$. Alors, l'Exemple 2.5.3 nous dit que $\mathrm{ev}_{e_i}$ est de classe $\mathrm{C}^{p-1}$. D'après la Remarque 2.5.2, la composition $\mathrm{ev}_{e_i} \circ Df$ est de classe $\mathrm{C}^{p-1}$ pour tout $i \in [\![1, n]\!]$. Comme $D_i f = \mathrm{ev}_{e_i} \circ Df$ pour tout $i \in [\![1, n]\!]$, on conclut que ces applications sont de classe $\mathrm{C}^{p-1}$, ce qui nous dit que $f$ est de classe $\mathrm{C}^p$.

Réciproquement, si $f$ est de classe $\mathrm{C}^p$, la Proposition 2.2.4 nous dit que $f$ est différentiable. La même proposition nous donne aussi l'identité

$$Df(v) = \sum_{i=1}^{n} D_i f(v) \pi_i$$

pour tout $v \in U$, où $\pi_i : \mathbb{R}^n \to \mathbb{R}$ est la projection sur la $i$-ème coordonnée pour tout $i \in [\![1, n]\!]$. Noter que $\pi_i$ est de classe $\mathrm{C}^{p-1}$ pour tout $i \in [\![1, n]\!]$, d'après l'Exemple 2.5.3. La Remarque 2.5.2 nous dit finalement que $Df$ est de classe $\mathrm{C}^{p-1}$. $\square$

Soient $p \in \mathbb{Z}_{>0}$ et $\mathbb{E}_1, \dots, \mathbb{E}_p$ et $\mathbb{F}$, des espaces vectoriels réels. On rappelle que $\mathrm{Mult}(\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F})$ désigne l'ensemble formé des **applications multilinéaires** $b : \mathbb{E}_1 \times \dots \times \mathbb{E}_p \to \mathbb{F}$, *i.e.*

$$b(v_1, \dots, v_{i-1}, v_i + \lambda v_i', v_{i+1}, \dots, v_p) = b(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_p) + \lambda b(v_1, \dots, v_{i-1}, v_i', v_{i+1}, \dots, v_p)$$

pour tous $v_i, v_i' \in \mathbb{E}_i$ avec $i \in [\![1, p]\!]$ et $\lambda \in \mathbb{R}$. Noter que $\mathrm{Mult}(\mathbb{E}; \mathbb{F}) = \mathrm{Lin}(\mathbb{E}, \mathbb{F})$. C'est facile à vérifier que $\mathrm{Mult}(\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F})$ est un espace vectoriel avec

$$(b + \lambda b')(v_1, \dots, v_p) = b(v_1, \dots, v_p) + \lambda b'(v_1, \dots, v_p)$$

pour tous $b, b' \in \mathrm{Mult}(\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F})$, $v_i \in \mathbb{E}_i$ avec $i \in [\![1, p]\!]$ et $\lambda \in \mathbb{R}$. Si $p = 2$, les éléments de $\mathrm{Mult}(\mathbb{E}_1, \mathbb{E}_2; \mathbb{F})$ sont appelés **applications bilinéaires**.

On rappelle le résultat suivant d'algèbre linéaire. La preuve est immédiate.

**Proposition 2.5.5.** *Soit $p \geq 2$. L'application*

$$\psi_{\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F}} : \mathrm{Mult}(\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F}) \to \mathrm{L}\left(\mathbb{E}_1, \mathrm{Mult}(\mathbb{E}_2, \dots, \mathbb{E}_p; \mathbb{F})\right)$$

*qui associe $\psi_{\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F}}(b)$ à $b \in \mathrm{Mult}(\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F})$ avec $\psi_{\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F}}(b)(v_1)(v_2, \dots, v_p) = b(v_1, \dots, v_p)$ pour tous $v_i \in \mathbb{E}_i$ avec $i \in [\![1, p]\!]$, est bien définie, linéaire et bijective. L'application inverse $\varphi_{\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F}}$ est donnée par $\varphi_{\mathbb{E}_1, \dots, \mathbb{E}_p; \mathbb{F}}(L)(v_1, \dots, v_p) = L(v_1)(v_2, \dots, v_p)$ pour tous $L \in \mathrm{L}\left(\mathbb{E}_1, \mathrm{Mult}(\mathbb{E}_2, \dots, \mathbb{E}_p; \mathbb{F})\right)$ et $v_i \in \mathbb{E}_i$ avec $i \in [\![1, p]\!]$.*

**Définition 2.5.6.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^p$, avec $p \in \mathbb{Z}_{>0} \cup \{\infty\}$. On définit*

$$D^q f : U \to \mathrm{Mult}(\underbrace{\mathbb{R}^n, \dots, \mathbb{R}^n}_{p \text{ fois}}; \mathbb{R}^m)$$

*par récurrence via $D^1 f = Df$ et $D^q f = \varphi_{\mathbb{R}^n, \dots, \mathbb{R}^n; \mathbb{R}^m} \circ D(D^{q-1} f)$ pour tout $q \in [\![2, p]\!]$.*

**Remarque 2.5.7.** *L'application $D^q f$ dans la définition précédente est bien définie et de classe $\mathrm{C}^{p-q}$ pour tout $q \in [\![2,p]\!]$. En effet, pour $q=2$, comme $Df$ est de classe $\mathrm{C}^{p-1}$, elle est de classe $\mathrm{C}^1$, ce qui implique que $D(Df)$ existe et elle de classe $\mathrm{C}^{p-2}$, donc aussi $D^2 f = \varphi_{\mathbb{R}^n,\mathbb{R}^n;\mathbb{R}^m} \circ D(Df)$, par la Proposition 2.5.4, l'Exemple 2.5.3 et la Remarque 2.5.2. On suppose que $D^{q-1}f$ est défini et de classe $\mathrm{C}^{p-q+1}$ avec $q \in [\![3,p]\!]$. Alors, $D(D^{q-1}f)$ est défini et de classe $\mathrm{C}^{p-q}$, ce qui implique que $D^q f = \varphi_{\mathbb{R}^n,\dots,\mathbb{R}^n;\mathbb{R}^m} \circ D(D^{q-1}f)$ est défini et de classe $\mathrm{C}^{p-q}$, par la Proposition 2.5.4, l'Exemple 2.5.3 et la Remarque 2.5.2.*

**Lemme 2.5.8.** *Soient $n,m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^p$. Alors, étant donné $q \in [\![1,p-1]\!]$ et $h_1,\dots,h_q \in \mathbb{R}^n$, l'application $g : U \to \mathbb{R}^m$ donnée par*

$$g(v) = D^q f(v)(h_1,\dots,h_q)$$

*est de classe $\mathrm{C}^{p-q}$ et elle satisfait que $Dg(v)(h) = D^{q+1}f(v)(h,h_1,\dots,h_q)$.*

***Preuve.*** Comme $g$ est la composition de l'application $D^q f$, qui est de classe $\mathrm{C}^{p-q}$ (voir Remarque 2.5.7), et l'application linéaire

$$\mathrm{ev}_{h_1,\dots,h_q} : \mathrm{Mult}(\underbrace{\mathbb{R}^n,\dots,\mathbb{R}^n}_{q \text{ fois}};\mathbb{R}^m) \to \mathbb{R}^m,$$

qui est de classe $\mathrm{C}^\infty$ (voir Exemple 2.5.3), on conclut que $g$ est classe $\mathrm{C}^{p-q}$, par la Remarque 2.5.2. En plus, le théorème 2.1.17 nous dit que

$$\begin{aligned} Dg(v)(h) &= D(\mathrm{ev}_{h_1,\dots,h_q} \circ D^q f)(v)(h) = (\mathrm{ev}_{h_1,\dots,h_q} \circ D(D^q f)(v))(h) = D(D^q f)(v)(h)(h_1,\dots,h_q) \\ &= D^{q+1}f(v)(h,h_1,\dots,h_q), \end{aligned}$$

comme on voulait démontrer. □

Le résultat suivant nous montre le lien entre l'application $D^q f$ dans la définition précédente et les dérivées partielles d'ordre supérieur.

**Proposition 2.5.9.** *Soient $n,m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^p$. Alors, étant donné $q \in [\![1,p]\!]$ et $i_1,\dots,i_q \in [\![1,n]\!]$, on a*

$$D^q f(v)(e_{i_1},\dots,e_{i_q}) = D_{i_1}\dots D_{i_q} f(v),$$

*pour tout $v \in U$, où $\{e_1,\dots,e_n\}$ désigne la base canonique de $\mathbb{R}^n$. De façon équivalente,*

$$D^q f(v)(h_1,\dots,h_q) = \sum_{i_1=1}^{n} \cdots \sum_{i_q=1}^{n} h_{1,i_1}\dots h_{q,i_q} D_{i_1}\dots D_{i_q} f(v),$$

*pour tous $v \in U$ et $h_i = (h_{i,1},\dots,h_{i,n}) \in \mathbb{R}^n$ avec $i \in [\![1,q]\!]$.*

***Preuve.*** On va montrer le résultat par récurrence sur $q \in [\![1,p]\!]$. Si $q=1$, le résultat suit de la Proposition 2.1.18. On suppose que c'est vrai pour $q \in [\![1,p-1]\!]$. On va montrer que c'est vrai pour $q+1$ et $i_0,\dots,i_q \in [\![1,n]\!]$. Or,

$$D^q f(v)(e_{i_1},\dots,e_{i_q}) = D_{i_1}\dots D_{i_q} f(v),$$

pour tous $v \in U$ et $i_1,\dots,i_q \in [\![1,n]\!]$. Soit $g : U \to \mathbb{R}^m$ l'application donnée par

$$g(v) = D^q f(v)(e_{i_1},\dots,e_{i_q}),$$

pour $v \in U$. Le Lemme 2.5.8, la Proposition 2.1.18 et l'hypothèse de la récurrence nous disent que

$$D^{q+1}f(v)(e_{i_0},e_{i_1},\dots,e_{i_q}) = Dg(v)(e_{i_0}) = D_{i_0}g(v) = D_{i_0}D_{i_1}\dots D_{i_q} f(v),$$

comme on voulait démontrer. □

Le résultat suivant est normalement attribué à K. Schwarz ou à A.-C. Clairaut, et montre l'utilité de la propriété $\mathrm{C}^2$.

**Proposition 2.5.10.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}^m$ une application et $v_0 \in U$. Si $f$ de classe $\mathrm{C}^2$, alors,*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(v) = \frac{\partial^2 f}{\partial x_j \partial x_i}(v),$$

*pour tout $v = (v_1, \dots, v_n) \in U$ et $i, j \in [\![1, n]\!]$. En conséquence, $D^2 f : U \to \mathrm{Mult}(\mathbb{R}^n, \mathbb{R}^n; \mathbb{R}^m)$ satisfait que $Df(v)(h,k) = Df(v)(k,h)$ pour tous $v \in U$ et $h, k \in \mathbb{R}^n$.*

***Preuve.*** On remarque d'abord qu'il suffit de démontrer le résultat pour $m = 1$. En effet, on écrit comme d'habitude $f(v) = (f_1(v), \dots, f_m(v))$ pour tout $v \in U$, où $f_k : U \to \mathbb{R}$ est la $i$-ème composante de $f$. Alors $f$ est de classe $\mathrm{C}^2$ si et seulement si $f_k$ est de classe $\mathrm{C}^2$ pour tout $k \in [\![1, m]\!]$. Si l'on admet le résultat pour $m = 1$, alors

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j}(v) = \frac{\partial^2 f_k}{\partial x_j \partial x_i}(v),$$

pour tout $v = (v_1, \dots, v_n) \in U$, $i, j \in [\![1, n]\!]$ et $k \in [\![1, m]\!]$, ce qui implique que

$$\Big(\frac{\partial^2 f_1}{\partial x_i \partial x_j}(v), \dots, \frac{\partial^2 f_m}{\partial x_i \partial x_j}(v)\Big) = \frac{\partial^2 f}{\partial x_i \partial x_j}(v) = \frac{\partial^2 f}{\partial x_j \partial x_i}(v) = \Big(\frac{\partial^2 f_1}{\partial x_j \partial x_i}(v), \dots, \frac{\partial^2 f_m}{\partial x_j \partial x_i}(v)\Big),$$

pour tout $v = (v_1, \dots, v_n) \in U$ et $i, j \in [\![1, n]\!]$. En conséquence, on peut supposer désormais $m = 1$.

Si $i = j$ il n'y a rien à démontrer donc on va supposer $i < j$. Dans ce cas, il suffit de démontrer l'énoncé pour $n = 2$, $i = 1$ et $j = 2$. En effet, soit $\mathrm{inc} : \mathbb{R}^2 \to \mathbb{R}^n$ l'inclusion donnée par

$$\mathrm{inc}(x, y) = (v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_{j-1}, y, v_{j+1}, \dots, v_n),$$

pour tout $(x, y) \in \mathbb{R}^2$. Comme $\mathrm{inc}$ est continue, $U' = \mathrm{inc}^{-1}(U) \subseteq \mathbb{R}^2$ est ouvert et $(v_i, v_j) \in U'$. Si l'on pose $\bar{f} = f \circ \mathrm{inc}|_{U'} : U' \to \mathbb{R}$, c'est clair que

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(v) = \frac{\partial^2 \bar{f}}{\partial x \partial y}(v_i, v_j) \text{ et } \frac{\partial^2 f}{\partial x_j \partial x_i}(v) = \frac{\partial^2 \bar{f}}{\partial y \partial x}(v_i, v_j),$$

ce qui réduit la preuve de l'énoncé au cas $n = 2$, $i = 1$ et $j = 2$, comme on avait affirmé.

Comme $U$ est ouvert et $(v_1, v_2) \in U$, il existe $r > 0$ tel que $R = [v_1 - r, v_1 + r] \times [v_2 - r, v_2 + r] \subseteq U$. Pour $x_2 \in [v_2 - r, v_2 + r] \setminus \{v_2\}$, soit $g_{x_2} : [v_1 - r, v_1 + r] \setminus \{v_1\} \to \mathbb{R}$ donné par $g_{x_2}(x_1) = f(x_1, x_2) - f(x_1, v_2)$. C'est clair que $g_{x_2}$ est de classe $\mathrm{C}^2$, pour tout $x_2 \in [v_2 - r, v_2 + r] \setminus \{v_2\}$. D'après le Théorème des accroissements finis pour des fonctions d'une seule variable (2.3.2),

$$g_{x_2}(x_1) - g_{x_2}(v_1) = h_1 g'_{x_2}(c_1) = h_1 \Big(\frac{\partial f}{\partial x_1}(c_1, x_2) - \frac{\partial f}{\partial x_1}(c_1, v_2)\Big),$$

où $h_1 = x_1 - v_1 \neq 0$ et $c_1$ est dans le segment déterminé par $v_1$ et $x_1$. Si l'on utilise la définition de $g_{x_2}$, l'équation précédente s'écrit

$$\begin{aligned} &f(x_1, x_2) - f(x_1, v_2) - f(v_1, x_2) + f(v_1, v_2) = g_{x_2}(x_1) - g_{x_2}(v_1) \\ &= h_1 \Big(\frac{\partial f}{\partial x_1}(c_1, x_2) - \frac{\partial f}{\partial x_1}(c_1, v_2)\Big) = h_1 h_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(c_1, c_2), \end{aligned} \tag{2.5.1}$$

où $h_2 = x_2 - v_2 \neq 0$, l'on a appliqué le Théorème des accroissements finis (2.3.2) à la fonction $x_2 \mapsto (\partial f / \partial x_1)(c_1, x_2)$ et $c_2$ est dans le segment déterminé par $v_2$ et $x_2$.

De façon similaire, si l'on considère la fonction $h_{x_1} : [v_2 - r, v_2 + r] \setminus \{v_2\} \to \mathbb{R}$ donnée par $h_{x_1}(x_2) = f(x_1, x_2) - f(v_1, x_2)$, où $x_1 \in [v_1 - r, v_1 + r] \setminus \{v_1\}$, $h_{x_1}$ est de classe $\mathrm{C}^2$ pour tout $x_1 \in [v_1 - r, v_1 + r] \setminus \{v_1\}$ et

$$\begin{aligned} &f(x_1, x_2) - f(v_1, x_2) - f(x_1, v_2) + f(v_1, v_2) = h_{x_1}(x_2) - h_{x_1}(v_2) \\ &= h_2 \Big(\frac{\partial f}{\partial x_2}(x_1, d_2) - \frac{\partial f}{\partial x_2}(v_1, d_2)\Big) = h_1 h_2 \frac{\partial^2 f}{\partial x_1 \partial x_2}(d_1, d_2), \end{aligned} \tag{2.5.2}$$

où $d_1$ est dans le segment déterminé par $v_1$ et $x_1$ et $d_2$ est dans le segment déterminé par $v_2$ et $x_2$. Comme les premiers membres de (2.5.1) et (2.5.2) coïncident, les derniers membres aussi. Si l'on divise l'égalité entre les derniers membres de (2.5.1) te (2.5.2) par $h_1 h_2$ et on prend la limite quand $h_1$ et $h_2$ tendent vers $0$, la continuité des dérivées partielles secondes nous dit que

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(v) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(v),$$

comme on voulait démonter.

La dernière partie est une conséquence immédiate de la première partie et la Proposition 2.5.9. □

**Remarque 2.5.11.** *De façon analogue à la Proposition 2.5.10, si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ est de classe $C^p$, alors $D^p f(v)(h_1, \dots, h_p) = D^p f(v)(h_{\sigma(1)}, \dots, h_{\sigma(p)})$ pour tout $v \in U$, $h_1, \dots, h_p \in \mathbb{R}^n$ et toute permutation $\sigma \in \mathbb{S}_p$ (voir* [1], *XVII, Thm. 6.2).*

## 2.6 Développement de Taylor

### 2.6.1 Rappel du développement de Taylor d'une fonction d'une seule variable

Soit $f : ]\,a, b\,[ \to \mathbb{R}^m$ une application de classe $C^{p+1}$, avec $p \in \mathbb{N}$ et $x_0 \in ]\,a, b\,[$ . On rappelle alors l'identité

$$f(x) = \sum_{i=0}^{q} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i + \int_{x_0}^{x} \frac{(x-t)^q}{q!} f^{(q+1)}(t)dt, \tag{2.6.1}$$

pour tout $x \in ]\,a, b\,[$ et $q \in [\![0, p]\!]$, où $f^{(i)}$ dénote la $i$-ème dérivée de $f$ pour $i \in \mathbb{Z}_{>0}$ et $f^{(0)} = f$. La preuve de ce résultat est par récurrence sur $q \in [\![0, p]\!]$, où le cas $q = 0$ correspond au théorème fondamental de l'analyse

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt.$$

En effet, si l'admet (2.6.1) pour $q \in [\![0, p-1]\!]$, on définit les fonctions $u(t) = -(x-t)^{q+1}/(q+1)!$ et $v(t) = f^{(q+1)}(t)$, ce qui implique que

$$\begin{aligned}\int_{x_0}^{x} \frac{(x-t)^q}{q!} f^{(q+1)}(t)dt &= \Big[ -\frac{(x-t)^{q+1}}{(q+1)!} f^{(q+1)}(t) \Big]_{t=x_0}^{t=x} + \int_{x_0}^{x} \frac{(x-t)^{q+1}}{(q+1)!} f^{(q+2)}(t)dt \\ &= \frac{f^{(q+1)}(x_0)}{(q+1)!}(x-x_0)^{q+1} + \int_{x_0}^{x} \frac{(x-x_0-t)^{q+1}}{(q+1)!} f^{(q+2)}(t)dt,\end{aligned}$$

ce qui donne (2.6.1) pour $q + 1$.

### 2.6.2 Développement de Taylor de fonctions de plusieurs variables

Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}^m$ une application de classe $C^{p+1}$, avec $p \in \mathbb{Z}_{>0}$. Soit $v_0 \in U$. Il existe $r > 0$ tel que $\bar{B}_{\|\ \|}(v_0, 2r) \subseteq U$. On prend $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ non nul tel que $\|h\| < r$. Soit $\alpha : J \to \mathbb{R}^n$ l'application $\alpha(t) = v_0 + t \cdot h$ pour tout $t \in J$, où $J = ]-2, 2\,[ \subseteq \mathbb{R}$. Comme $\alpha$ est de classe $C^\infty$ (voir l'Exemple 2.5.3), $f_h = f \circ \alpha : ]-2, 2\,[ \to \mathbb{R}^m$ est classe $C^{p+1}$. En outre, $\alpha'(t) = h$ pour tout $t \in J$. L'expression (2.6.1) nous dit alors que

$$f_h(1) = \sum_{i=0}^{q} \frac{f_h^{(i)}(0)}{i!} + \frac{1}{q!} \int_0^1 (1-s)^q f_h^{(q+1)}(s)ds \tag{2.6.2}$$

pour tout $q \in [\![1, p]\!]$. On voit bien que $f_h(0) = f(v_0)$ et $f_h(1) = f(v_0 + h)$.

En outre, on affirme que

$$f_h^{(i)}(t) = D^i f(v_0 + t \cdot h)(\underbrace{h, \dots, h}_{i \text{ fois}}) \tag{2.6.3}$$

pour tout $t \in J$ et $i \in [\![1, q+1]\!]$. En effet, si $i = 1$, le Théorème 2.1.17 donne

$$f_h'(t) = Df_h(t)(1) = D(f \circ \alpha)(t)(1) = Df(\alpha(t))(\alpha'(t)) = Df(v_0 + t \cdot h)(h).$$

On suppose que (2.6.3) est vérifié pour $i \in [\![1, q]\!]$. On va montrer qu'elle est aussi vérifié pour $i+1$. Soit

$$\mathrm{ev}_h^i : \mathrm{Mult}(\underbrace{\mathbb{R}^n, \dots, \mathbb{R}^n}_{i \text{ fois}}; \mathbb{R}^m) \to \mathbb{R}^m$$

l'application linéaire qui associe à l'application multilinéaire $L$ le vecteur $L(h, \dots, h)$. Alors, l'hypothèse de la récurrence nous dit que $f_h^{(i)} = \mathrm{ev}_h^i \circ D^i f \circ \alpha$, ce qui implique que, par le Théorème 2.1.17,

$$\begin{aligned} f_h^{(i+1)}(t) &= Df_h^{(i)}(t)(1) = D(\mathrm{ev}_h^i \circ D^i f \circ \alpha)(t)(1) = \Big(\mathrm{ev}_h^i \circ D(D^i f)\big(\alpha(t)\big) \circ D\alpha(t)\Big)(1) \\ &= D(D^i f)\big(\alpha(t)\big)(h)(\underbrace{h, \dots, h}_{i \text{ fois}}) = D^{i+1} f\big(\alpha(t)\big)(\underbrace{h, \dots, h}_{i+1 \text{ fois}}), \end{aligned}$$

comme on voulait démontrer.

En conséquence, (2.6.2) avec (2.6.3) nous disent que

$$f(v_0 + h) = f_h(1) = \sum_{i=0}^{q} \frac{1}{i!} D^i f(v_0)(\underbrace{h, \dots, h}_{i \text{ fois}}) + \frac{1}{q!} \int_0^1 (1-s)^q D^{q+1} f(v_0 + s \cdot h)(\underbrace{h, \dots, h}_{q+1 \text{ fois}}) ds \qquad (2.6.4)$$

pour tout $q \in [\![1, p]\!]$.

### 2.6.3 Développement de Taylor de d'ordre $2$ et estimation du résidu

Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^3$. Soit $v_0 \in U$. Il existe $r > 0$ tel que $\bar{B}_{\| \ \|}(v_0, 2r) \subseteq U$. On prend $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ non nul tel que $\|h\| < r$. L'expression (2.6.4) pour $p = 2$ nous dit que

$$f(v_0 + h) = f(v_0) + Df(v_0)(h) + \frac{1}{2} D^2 f(v_0)(h, h) + \frac{1}{2} \int_0^1 (1-s)^2 D^3 f(v_0 + s \cdot h)(h, h, h) ds.$$

Par ailleurs, la Proposition 2.5.9 nous dit que

$$Df(v_0)(h) = \big\langle \nabla f(v_0), h \big\rangle \quad \text{et} \quad D^2 f(v_0)(h, h) = h \cdot H_f(v_0) \cdot h^t,$$

où

$$H_f(v_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(v_0) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(v_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(v_0) & \dots & \frac{\partial^2 f}{\partial x_n^2}(v_0) \end{pmatrix}$$

est la **matrice hessienne** de $f$ en $v_0$. On remarque que le produit $\cdot$ est le produit matriciel standard, où l'on considère le vecteur $h = (h_1, \dots h_n)$ de façon canonique comme une matrice ligne de taille $1 \times n$. D'après la Proposition 2.5.10, comme $f$ est *a fortiori* de classe $\mathrm{C}^2$, la matrice hessienne $H_f(v_0)$ est **symétrique**, *i.e.* $H_f(v_0)^t = H_f(v_0)$.

En plus, la Proposition 2.5.9 nous dit que

$$D^3 f(v_0 + s \cdot h)(h, h, h) = \sum_{i,j,k=1}^{n} h_i h_j h_k \underbrace{\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(v_0 + s \cdot h)}_{\star_{i,j,k}(s)},$$

où $s \in J$. Comme $f$ est de classe $\mathrm{C}^3$, il existe $C \in \mathbb{R}_{>0}$ tel que $|\star_{i,j,k}(s)| \leq C$ pour tout $s \in J$ et tous $i, j, k = 1, \dots, n$. En conséquence,

$$\left| \frac{1}{2} \int_0^1 (1-s)^2 D^3 f(v_0 + s \cdot h)(h, h, h) ds \right| \leq \frac{1}{2} \sum_{i,j,k=1}^{n} \int_0^1 |h_i| \, |h_j| \, |h_k| \, C ds \leq \frac{C}{2} \sum_{i,j,k=1}^{n} |h_i| \, |h_j| \, |h_k|,$$

d'où on obtient que

$$\left|\frac{1}{2\|h\|^2}\int_0^1 (1-s)^2 D^3 f(v_0+s\cdot h)(h,h,h)ds\right| \leq \frac{C}{2}\sum_{i,j,k=1}^{n}\underbrace{\frac{|h_i|}{\|h\|}}_{\leq 1}\underbrace{\frac{|h_j|}{\|h\|}}_{\leq 1}|h_k| \leq \frac{Cn^3}{2}\|h\|,$$

où l'on a utilisé l'inégalité évidente $|h_i| \leq \|h\|$ pour tout $i = 1,\dots,n$ et tout vecteur $h = (h_1,\dots,h_n) \in \mathbb{R}^n$.

La discussion précédente démontre le résultat suivant.

**Proposition 2.6.1.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^3$. Soient $v_0 \in \bar{B}_{\|\ \|}(v_0, 2r) \subseteq U$ et $h = (h_1,\dots,h_n) \in \mathbb{R}^n$ non nul tel que $\|h\| < r$. Alors,*

$$f(v_0+h) = f(v_0) + \langle h, \nabla f(v_0)\rangle + \frac{1}{2}h\cdot H_f(v_0)\cdot h^t + \phi(h),$$

*où $\phi : B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r) \to \mathbb{R}$ est une application qui satisfait la condition $\mathrm{o}(t^2)$, i.e.*

$$\lim_{h\to\mathbf{0}_{\mathbb{R}^n}} \frac{\phi(h)}{\|h\|^2} = 0.$$

## 2.7 Optimisation de fonctions

### 2.7.1 Définitions basiques et quelques propriétés

Dans cette dernière section on va étudier le problème d'optimisation de fonctions scalaires qui satisfont quelques propriétés de régularité. Pour cela on commence avec quelques définitions préliminaires.

**Définition 2.7.1.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une application différentiable. On dit que $v_0 \in U$ est un **point critique** de $f$ si $\nabla f(v_0) = \mathbf{0}_{\mathbb{R}^n}$.*

**Définition 2.7.2.** *Soit $S \subseteq \mathbb{R}^n$ une partie et $f : S \to \mathbb{R}$ une application. On dit que $v_0 \in U$ est*

*(i) un **maximum local** (resp., **minimum local**) de $f$ s'il existe une partie ouverte $V \subseteq S$ de $S$ telle que $v_0 \in V$ et $f(v_0) \geq f(v)$ (resp., $f(v_0) \leq f(v)$) pour tout $v \in V$;*

*(ii) un **maximum local strict** (resp., **minimum local strict**) de $f$ s'il existe une partie ouverte $V \subseteq S$ de $S$ telle que $v_0 \in V$ et $f(v_0) > f(v)$ (resp., $f(v_0) < f(v)$) pour tout $v \in V \setminus \{v_0\}$.*

*Un **extremum local** (resp., **extremum local strict**) de $f$ est un maximum local ou minimum local (resp., maximum local strict ou minimum local strict).*

**Remarque 2.7.3.** *Le point $v_0 \in U$ est un maximum local (resp., maximum local strict) de $f : U \to \mathbb{R}$ si et seulement si $v_0$ est un minimum local (resp., minimum local strict) de $-f$.*

Le résultat suivant est le point de départ pour l'étude de l'optimisation de fonctions différentiable.

**Proposition 2.7.4.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte, $f : U \to \mathbb{R}$ une application différentiable et $v_0 \in U$. Si $v_0$ est un extremum local de $f$, alors c'est un point critique.*

***Preuve.*** Si $n = 1$, le résultat suit directement de la Proposition 2.3.1. On suppose désormais $n > 1$. Soit $\{e_1,\dots,e_n\}$ la base canonique de $\mathbb{R}^n$. Il existe $r > 0$ tel que $B_{\|\ \|}(v_0, 2r) \subseteq U$. Pour $i \in [\![1,n]\!]$, on considère l'application $\alpha_i : J \to \mathbb{R}^n$ donnée par $\alpha_i(t) = v_0 + t\cdot e_i$ pour tout $t \in J$, où $J \subseteq ]-r, r[$ est un intervalle ouvert incluant 0. On remarque que $\alpha_i$ est de classe $\mathrm{C}^\infty$ (voir l'Exemple 2.2.2), ce qui implique que l'application $g_i : J \to \mathbb{R}$ donnée par $g_i = f\circ\alpha_i$ est différentiable. Comme $v_0$ est un extremum local de $f$, $t = 0$ est un extremum local de $g_i$ pour tout $i \in [\![1,n]\!]$, ce qui implique que $g_i'(0) = 0$, car $g_i$ est une fonction dérivable d'une seule variable. En outre, (2.6.3) nous dit que

$$0 = g_i'(0) = \frac{\partial f}{\partial x_i}(v_0),$$

pour tout $i \in [\![1,n]\!]$, comme on voulait démontrer. □

**Définition 2.7.5.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^2$. Un point critique $v_0 \in U$ de $f$ est **non dégénéré** si $\det(H_f(v_0)) \neq 0$. Noter que cela équivaut à dire que toutes les valeurs propres de $H_f(v_0)$ sont non nulles.*

### 2.7.2 Optimisation de fonctions sans contraintes

Le but de cette sous-section est de donner un critère suffisant pour qu'un point critique soit un extremum ou non. On suppose désormais que $f : U \to \mathbb{R}$ une application est une application de classe $\mathrm{C}^3$ sur une partie ouverte $U \subseteq \mathbb{R}^n$ et que $v_0 \in U$ est un point critique de $f$. La Proposition 2.6.1 nous dit que

$$f(v_0 + h) = f(v_0) + \big\langle h, \underbrace{\nabla f(v_0)}_{=\mathbf{0}_{\mathbb{R}^n}} \big\rangle + \frac{1}{2} h \cdot H_f(v_0) \cdot h^t + \phi(h), \tag{2.7.1}$$

pour tout $h \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)$, où $r > 0$ satisfait que $B_{\| \ \|}(v_0, 2r) \subseteq U$, et $\phi : B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r) \to \mathbb{R}$ satisfait que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \frac{\phi(h)}{\|h\|^2} = 0.$$

L'annulation du gradient est une conséquence du fait que $v_0$ est un point critique de $f$. L'expression (2.7.1) suggère que le comportement local de $f$ autour de $v_0$ est déterminé par

$$\frac{1}{2} h \cdot H_f(v_0) \cdot h^t.$$

On rappelle d'abord que $H_f(v_0)$ est une matrice symétrique, d'après la Proposition 2.5.10. On va utiliser le résultat suivant, que l'on ne va pas démontrer. Il suffit de dire que la preuve se base sur une récurrence sur la dimension $n$.

**Proposition 2.7.6.** *Soit $A \in \mathrm{M}_n(\mathbb{R})$ une matrice symétrique avec $n \in \mathbb{Z}_{>0}$. Alors il existe une matrice **orthogonale** $P \in \mathrm{M}_n(\mathbb{R})$, i.e. $P^t \cdot P = P \cdot P^t = \mathrm{Id}_n$ est la matrice identité de $\mathrm{M}_n(\mathbb{R})$, telle que*

$$A = P \cdot D \cdot P^t,$$

*où $D$ est une matrice diagonale. Dans ce cas, la matrice $D$ est formée des valeurs propres de $A$, répétées selon la multiplicité.*

On remarque que la proposition précédente nous dit que

$$\begin{aligned} \det(A) &= \det(P \cdot D \cdot P^t) = \det(P)\det(D)\det(P^t) = \det(P)\det(D)\det(P^{-1}) \\ &= \det(P)\det(D)\det(P)^{-1} = \det(D) = \prod_{i=1}^{n} D_{i,i}, \end{aligned} \tag{2.7.2}$$

pour tout matrice $A \in \mathrm{M}_n(\mathbb{R})$ symétrique.

Le résultat suivant donne une condition suffisante pour décider si un point critique est un extremum local et déterminer sa nature. La preuve est une application de la Proposition 2.6.1.

**Théorème 2.7.7.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^3$. Soit $v_0 \in U$ un point critique de $f$. On a les résultats suivants:*

*(i) si toutes les valeurs propres de $H_f(v_0)$ sont (strictement) positives, alors $v_0$ est un minimum local strict de $f$;*

*(ii) si toutes les valeurs propres de $H_f(v_0)$ sont (strictement) négatives, alors $v_0$ est un maximum local strict de $f$;*

*(iii) si $H_f(v_0)$ admet une valeur propre négative et une valeur propre positive, alors $v_0$ n'est pas un extremum.*

*Dans le dernier cas on dit que $v_0$ est un **point-selle** ou un **point col**.*

***Preuve.*** On peut réécrire (2.7.1) sous la forme suivante

$$f(v_0 + h) - f(v_0) = \frac{1}{2} h \cdot H_f(v_0) \cdot h^t + \phi(h), \tag{2.7.3}$$

pour $h = (h_1, \dots, h_n) \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)$, où $r > 0$ satisfait que $B_{\| \ \|}(v_0, 2r) \subseteq U$, et $\phi : B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r) \to \mathbb{R}$ satisfait que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \frac{\phi(h)}{\sum_{i=1}^{n} h_i^2} = 0. \tag{2.7.4}$$

L'identité (2.7.3) nous dit que

(Ext.1) $v_0$ est un minimum local strict si et seulement s'il existe $0 < s < r$ tel que $\frac{1}{2} h \cdot H_f(v_0) \cdot h^t + \phi(h) > 0$ pour tout $h \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$;

(Ext.2) $v_0$ est un maximum local strict si et seulement s'il existe $0 < s < r$ tel que $\frac{1}{2} h \cdot H_f(v_0) \cdot h^t + \phi(h) < 0$ pour tout $h \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$.

On rappelle d'abord que $H_f(v_0)$ est une matrice symétrique, d'après la Proposition 2.5.10. D'après la Proposition 2.7.6, il existe existe une matrice orthogonale $P \in \mathrm{M}_n(\mathbb{R})$ telle que

$$H_f(v_0) = P \cdot D \cdot P^t,$$

où $D$ est une matrice diagonale formée des valeurs propres de $A$, répétées selon leur multiplicité. Par conséquent,

$$\frac{1}{2} h \cdot H_f(v_0) \cdot h^t = \frac{1}{2} h \cdot P \cdot D \cdot P^t \cdot h^t = \frac{1}{2} k \cdot D \cdot k^t = \frac{\sum_{i=1}^{n} D_{i,i} k_i^2}{2},$$

pour tout $h \in \mathbb{R}^n$, où $k = hP \in \mathbb{R}^n$ et l'on a écrit $k = (k_1, \dots, k_n)$.

Comme $P$ est une matrice orthogonale, l'application $f_P : \mathbb{R}^n \to \mathbb{R}^n$ donnée par $f_P(h) = h \cdot P$ est linéaire et bijective, et elle satisfait que

$$\langle f_P(h'), f_P(h) \rangle = \langle h' \cdot P, h \cdot P \rangle = h' \cdot P \cdot \big(h \cdot P\big)^t = h' \cdot \underbrace{P \cdot P^t}_{\mathrm{Id}_n} \cdot h^t = h' \cdot h^t = \langle h', h \rangle,$$

pour tous $h', h \in \mathbb{R}^n$, où $\langle \, , \rangle$ est le produit scalaire euclidien de $\mathbb{R}^n$, on a utilisé que $\langle k', k \rangle = k' \cdot k^t$ pour $k', k \in \mathbb{R}^n$ et $(A \cdot B)^t = B^t \cdot A^t$ pour des matrices $A, B$ telles que le produit soit défini. La réciproque $f_P^{-1}$ de $f_P$ est donnée par $f_P^{-1}(k) = k \cdot P^t$, pour tout $k \in \mathbb{R}^n$. En particulier, $\|f_P(h)\| = \|h\|$, où $\| \ \|$ dénote la norme euclidienne, ce qui nous dit que $f_P(B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)) = B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)$ et aussi que $h$ tend vers $\mathbf{0}_{\mathbb{R}^n}$ si et seulement si $k = h \cdot P \in \mathbb{R}^n$ tend vers $\mathbf{0}_{\mathbb{R}^n}$. En conséquence, (2.7.3) équivaut à

$$f(v_0 + k \cdot P^t) - f(v_0) = \frac{\sum_{i=1}^{n} D_{i,i} k_i^2}{2} + \tilde{\phi}(k),$$

pour tout $k = (k_1, \dots, k_n) \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)$, où $\tilde{\phi} = \phi \circ f_P$, tandis que (2.7.4) équivaut à

$$\lim_{k \to \mathbf{0}_{\mathbb{R}^n}} \frac{\tilde{\phi}(k)}{\sum_{i=1}^{n} k_i^2} = 0. \tag{2.7.5}$$

On remarque que $f_P(B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)) = B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, 2r)$ nous dit que

(Ext*.1) $v_0$ est un minimum local strict si et seulement s'il existe $0 < s < r$ tel que $\frac{1}{2} k \cdot D \cdot k^t + \tilde{\phi}(k) > 0$ pour tout $k \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$;

(Ext*.2) $v_0$ est un maximum local strict si et seulement s'il existe $0 < s < r$ tel que $\frac{1}{2} k \cdot D \cdot k^t + \tilde{\phi}(k) < 0$ pour tout $k \in B_{\| \ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$.

Supposons que $D_{i,i} > 0$ pour tout $i \in [\![1, n]\!]$, ou $D_{i,i} < 0$ pour tout $i \in [\![1, n]\!]$. La limite (2.7.5) nous dit que

$$\lim_{k \to \mathbf{0}_{\mathbb{R}^n}} \frac{\tilde{\phi}(k)}{\sum_{i=1}^{n} D_{i,i} k_i^2 / 2} = \lim_{k \to \mathbf{0}_{\mathbb{R}^n}} \frac{\tilde{\phi}(k)}{\sum_{i=1}^{n} k_i^2} \frac{\sum_{i=1}^{n} k_i^2}{\sum_{i=1}^{n} D_{i,i} k_i^2 / 2} = 0, \tag{2.7.6}$$

vu que

$$\left|\frac{\sum_{i=1}^{n} k_i^2}{\sum_{i=1}^{n} D_{i,i}k_i^2/2}\right| \leq 2\sum_{i=1}^{n}\frac{1}{|D_{i,i}|},$$

où l'on a utilisé que $x_i^2 \leq \sum_{j=1}^n x_j^2$, pour tout $i \in [\![1,n]\!]$. La limite (2.7.6) nous dit que, étant donné $\epsilon = 1/2$, il existe $0 < s < r$ tel que

$$|\tilde{\phi}(k)| < \frac{1}{2}\frac{\sum_{i=1}^{n}|D_{i,i}|k_i^2}{2} = \frac{1}{4}\left|k \cdot D \cdot k^t\right|,$$

pour tout $k \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$. On remarque le fait élémentaire que, étant donné $x \in \mathbb{R} \setminus \{0\}$ et $y \in \mathbb{R}$ tels que $|y| \leq |x|/2$, alors $x + y$ et $y$ ont le même signe, *i.e.* $x + y > 0$ si et seulement si $x > 0$, et $x + y < 0$ si et seulement si $x < 0$. Si l'on applique ce fait élémentaire au cas $x = k \cdot D \cdot k^t/2$ et $y = \tilde{\phi}(k)$, on conclut que

(Ext**.1) $\frac{1}{2}k \cdot D \cdot k^t + \tilde{\phi}(k) > 0$ si et seulement si $\frac{1}{2}k \cdot D \cdot k^t > 0$, pour tout $k \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$;

(Ext**.2) $\frac{1}{2}k \cdot D \cdot k^t + \tilde{\phi}(k) < 0$ si et seulement si $\frac{1}{2}k \cdot D \cdot k^t < 0$, pour tout $k \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$.

Si l'on combine (Ext*.1) avec (Ext**.1), on trouve précisément (i), tandis que (Ext*.2) avec (Ext**.2) est exactement (ii).

On suppose finalement qu'il existe $i, j \in [\![1,n]\!]$ tels que $D_{i,i} < 0$ et $D_{j,j} > 0$. Soit $k' = \lambda e_i$ et $k'' = \lambda e_j$, avec $\lambda \in \mathbb{R}$ et soit $\{e_1, \dots, e_n\}$ la base canonique de $\mathbb{R}^n$. Dans ce cas, la limite (2.7.5) nous dit que

$$\lim_{\lambda \to 0} \frac{\tilde{\phi}(\lambda e_\ell)}{D_{\ell,\ell}\lambda^2/2} = \lim_{\lambda \to 0} \frac{\tilde{\phi}(\lambda e_\ell)}{\lambda^2}\frac{2}{D_{\ell,\ell}} = 0, \tag{2.7.7}$$

pour tout $\ell \in [\![1,n]\!]$. La limite (2.7.7) nous dit que, étant donné $\epsilon = 1/2$, il existe $0 < s < r$ tel que

$$|\tilde{\phi}(\lambda e_\ell)| < \frac{1}{2}\frac{|D_{\ell,\ell}|\lambda^2}{2},$$

pour tous $\lambda \in ]-s, s[\setminus\{0\}$ et $\ell \in [\![1,n]\!]$. En particulier,

(PS.1) $\frac{1}{2}k \cdot D \cdot k^t + \tilde{\phi}(k) < 0$ pour tout $k = \lambda e_i \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$;

(PS.2) $\frac{1}{2}k \cdot D \cdot k^t + \tilde{\phi}(k) > 0$ pour tout $k = \lambda e_j \in B_{\|\ \|}(\mathbf{0}_{\mathbb{R}^n}, s) \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$.

Cela nous dit exactement que $v_0$ n'est pas un extremum local, *i.e.* le cas dans l'item (iii). □

Si $n = 2$, on peut reformuler le théorème précédent sous la forme suivante.

**Corollaire 2.7.8.** *On admet les mêmes hypothèses que dans le théorème précédent, avec $n = 2$. On écrit*

$$H_f(v_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(v_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(v_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(v_0) & \frac{\partial^2 f}{\partial x_2^2}(v_0) \end{pmatrix} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

*La condition de non dégénération sur $v_0$ équivaut à l'inégalité $ad - b^2 \neq 0$. On a les résultats suivants:*

*(i) si $ad - b^2 > 0$ et $a > 0$, alors $v_0$ est un minimum local strict de $f$;*

*(ii) si $ad - b^2 > 0$ et $a < 0$, alors $v_0$ est un maximum local strict de $f$;*

*(iii) si $ad - b^2 < 0$, alors $v_0$ est un point-selle.*

***Preuve.*** L'identité (2.7.2) nous dit que le déterminant de $H_f(v_0)$ est positif si et seulement si les deux valeurs propres de $H_f(v_0)$ ont le même signe, et que le déterminant de $H_f(v_0)$ est négatif si et seulement si les deux valeurs propres de $H_f(v_0)$ ont signe contraire. On obtient en particulier l'énoncé dans l'item (iii) à partir de l'item (iii) dans le Théorème 2.7.7.

On suppose désormais que le déterminant $ad - b^2$ de $H_f(v_0)$ est positif. En particulier, $ad > 0$, car sinon $ad - b^2 \leq 0$, vu que $-b^2 \leq 0$. Cela nous dit que $a$ et $d$ ont le même signe. Or, on sait que le polynôme caractéristique de la matrice

$$H_f(v_0) = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

est $X^2 - (a+d)X + (ad - b^2)$ et que ses racines $\lambda_-$ et $\lambda_+$, *i.e.* les valeurs propres de $H_f(v_0)$, satisfont que $\lambda_+ + \lambda_- = a + d$. Comme $a$ et $d$ ont le même signe, et $\lambda_-$ et $\lambda_+$ ont le même signe aussi, on conclut que le signe $\lambda_-$ (ou de $\lambda_+$) coïncide avec le signe de $a$ (ou de $d$). On obtient en particulier l'énoncé dans l'item (i) (resp., (ii)) à partir de l'item (i) (resp., (ii)) dans le Théorème 2.7.7. □

**Remarque 2.7.9.** *Le Théorème 2.7.7 (ou, de façon équivalente, le Corollaire 2.7.8 pour le cas $n = 2$) donne des conditions suffisantes mais pas nécessaires pour qu'un point critique d'une fonction $f : U \to \mathbb{R}$ soit un extremum ou non. Par exemple, la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ donnée par $f(x_1, x_2) = x_1^4 + x_2^4$ pour $(x_1, x_2) \in \mathbb{R}^2$ est de classe $\mathrm{C}^3$ et possède le seul point critique $(0,0)$ tel que la matrice hessienne en $(0,0)$ est nulle, mais comme $f(x_1, x_2) = x_1^4 + x_2^4 > 0 = f(0,0)$ pour tout $(x_1, x_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ on voit que $(0,0)$ est un minimum local ou global strict de $f$. Si l'on remplace $f$ par $-f$ on obtient un exemple d'un point critique avec matrice hessienne aussi nulle qui est un maximum local ou global strict. Finalement, la fonction $g : \mathbb{R}^2 \to \mathbb{R}$ donnée par $g(x_1, x_2) = x_1^3 + x_2^3$ pour $(x_1, x_2) \in \mathbb{R}^2$ est de classe $\mathrm{C}^3$ et possède le seul point critique $(0,0)$ tel que la matrice hessienne en $(0,0)$ est aussi nulle, mais comme $f(x_1, 0) = x_1^3 > 0 = f(0,0)$ pour tout $x_1 \in \mathbb{R}_{>0}$ et $f(x_1, 0) = x_1^3 < 0 = f(0,0)$ pour tout $x_1 \in \mathbb{R}_{<0}$, on voit que $(0,0)$ n'est pas un extremum local de $f$.*

**Exemple 2.7.10.** *On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ donnée par*

$$f(x_1, x_2) = \ln(1 + x_1^2 + x_2^2)$$

*pour tout $(x_1, x_2) \in \mathbb{R}^2$. C'est clair que $f$ est de classe $\mathrm{C}^3$, puisqu'elle s'écrit comme composition, somme et produit de fonctions de classe $\mathrm{C}^3$. Le seul point critique de $f$ est l'origine, puisque*

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \frac{2x_1}{1 + x_1^2 + x_2^2} \quad \textit{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = \frac{2x_1}{1 + x_1^2 + x_2^2}.$$

*Les dérivées partielles secondes sont*

$$\begin{aligned} \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) &= 2\frac{1 - x_1^2 + x_2^2}{(1 + x_1^2 + x_2^2)^2}, \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) &= 2\frac{1 + (x_1 - x_2)^2}{(1 + x_1^2 + x_2^2)^2}, \\ \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) &= 2\frac{1 - x_2^2 + x_1^2}{(1 + x_1^2 + x_2^2)^2}. \end{aligned}$$

*La matrice hessienne en $(0,0)$ est donc*

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

*Le Corollaire 2.7.8 nous dit alors que $(0,0)$ est un minimum local strict de $f$. En fait, vu que $1 + x_1^2 + x_2^2 \geq 1$ pour tout $(x_1, x_2) \in \mathbb{R}^2$ et la fonction $\ln$ est strictement croissante sur $\mathbb{R}_{>0}$, on voit que $f(x_1, x_2) \geq \ln(1) = f(0,0) = 0$, ce qui nous dit que $(0,0)$ est aussi un minimum global de $f$.*

**Exemple 2.7.11.** *On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ donnée par*

$$f(x_1, x_2) = e^{x_1 x_2 - 1} - \frac{x_1^2}{2} - \frac{x_2^2}{2}$$

*pour tout* $(x_1, x_2) \in \mathbb{R}^2$*. C'est clair que* $f$ *est de classe* $\mathrm{C}^3$*, puisqu'elle s'écrit comme composition, somme et produit de fonctions de classe* $\mathrm{C}^3$*. Par ailleurs, l'ensemble de points critiques est formé par les points* $(x_1, x_2) \in \mathbb{R}^2$ *tels que*

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = x_2 e^{x_1 x_2 - 1} - x_1 = 0 \quad \textit{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1 e^{x_1 x_2 - 1} - x_2 = 0.$$

*On note que* $(0, 0)$ *est donc un point critique. En outre, on voit bien que, si* $(x_1, x_2) \in \mathbb{R}^2$ *est un point critique,* $x_1 \neq 0$ *si et seulement si* $x_2 \neq 0$*. Si* $(x_1, x_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ *est un point critique, alors* $x_1 = x_2 e^{x_1 x_2 - 1} = x_1 e^{2(x_1 x_2 - 1)}$*, ce qui implique que* $e^{2(x_1 x_2 - 1)} = 1$*, vu que* $x_1 \neq 0$*, ce qui nous dit que* $2(x_1 x_2 - 1) = 0$*, i.e.* $x_1 x_2 = 1$*. Si l'on utilise cette identité dans* $x_1 = x_2 e^{x_1 x_2 - 1}$ *on conclut que* $x_1 = x_2$*, ce qui nous dit que* $(x_1, x_2) \in \{(-1, -1), (1, 1)\}$*. Par ailleurs, on voit bien que* $(-1, -1)$ *et* $(1, 1)$ *sont des points critiques de* $f$*, ce qui implique que l'ensemble de points critiques de* $f$ *est* $\{(-1, -1), (0, 0), (1, 1)\}$*.*

*Les dérivées partielles secondes sont*

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) &= x_2^2 e^{x_1 x_2 - 1} - 1, \\
\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) &= e^{x_1 x_2 - 1}(1 + x_1 x_2), \\
\frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) &= x_1^2 e^{x_1 x_2 - 1} - 1.
\end{aligned}$$

*La matrice hessienne en* $(0, 0)$ *est donc*

$$H_f(0, 0) = \begin{pmatrix} -1 & e^{-1} \\ e^{-1} & -1 \end{pmatrix}.$$

*Le Corollaire 2.7.8 nous dit alors que* $(0, 0)$ *est un maximum local strict de* $f$*, vu que* $\det(H_f(0, 0)) = 1 - e^{-2} > 0$ *et* $-1 < 0$*. Par ailleurs,*

$$H_f(-1, -1) = H_f(1, 1) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

*Le Corollaire 2.7.8 nous dit alors que* $(-1, -1)$ *et* $(1, 1)$ *sont des points-selle de* $f$*, vu que* $\det(H_f(-1, -1)) = \det(H_f(1, 1)) = -4 < 0$*. On note finalement que*

$$\lim_{x \to +\infty} f(x, 0) = -\infty \textit{ et } \lim_{x \to +\infty} f(x, x) = +\infty,$$

*ce qui nous dit que* $f$ *ne possède ni de maximum global ni de minimum global.*

**Exemple 2.7.12.** *On considère la fonction* $f : \mathbb{R}^2 \to \mathbb{R}$ *donnée par*

$$f(x_1, x_2) = x_1^4 + x_2^2$$

*pour tout* $(x_1, x_2) \in \mathbb{R}^2$*. C'est clair que* $f$ *est de classe* $\mathrm{C}^3$*, puisqu'elle s'écrit comme somme et produit de fonctions de classe* $\mathrm{C}^3$*. Le seul point critique de* $f$ *est l'origine, puisque*

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 4x_1^3 \quad \textit{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = 2x_2.$$

*Or, les dérivées partielles secondes sont*

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) &= 12x_1^2, \\
\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) &= 0, \\
\frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) &= 2.
\end{aligned}$$

*La matrice hessienne en* $(0, 0)$ *est donc*

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

*Dans ce cas, on ne peut pas utiliser le Corollaire 2.7.8. Par contre, on voit bien que* $f(x_1, x_2) > 0 = f(0, 0)$ *pour tout* $(x_1, x_2) \in \mathbb{R}^2 \setminus \{0\}$*. En conséquence,* $(0, 0)$ *un minimum global et local strict de* $f$*.*

**Exemple 2.7.13.** *On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ donnée par*

$$f(x_1, x_2) = x_2^2 - x_1^4$$

*pour tout $(x_1, x_2) \in \mathbb{R}^2$. C'est clair que $f$ est de classe $\mathrm{C}^3$, puisqu'elle s'écrit comme somme et produit de fonctions de classe $\mathrm{C}^3$. Le seul point critique de $f$ est l'origine, puisque*

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = -4x_1^3 \quad \text{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = 2x_2.$$

*Or, les dérivées partielles secondes sont*

$$\begin{aligned} \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) &= -12x_1^2, \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) &= 0, \\ \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) &= 2. \end{aligned}$$

*La matrice hessienne en $(0, 0)$ est donc*

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

*Dans ce cas, on ne peut pas utiliser le Corollaire 2.7.8. Par contre, on note que $f(0, x) = x^2 > 0 = f(0, 0)$ pour tout $x \in \mathbb{R} \setminus \{0\}$, et $f(x, 0) = -x^4 < 0 = f(0, 0)$ pour tout $x \in \mathbb{R} \setminus \{0\}$. En conséquence, $(0, 0)$ n'est pas un extremum local de $f$.*

On a vu dans les Exemples 2.7.12 et 2.7.13 des fonctions $f$ qui possèdent un (unique) point critique $v_0$ tel que toutes les valeurs propres de la matrice hessienne $H_f(v_0)$ sont non négatives, et le point critique est un minimum local strict ou il n'est pas un extremum local. Même si l'on ne peut pas utiliser le Théorème 2.7.7 pour distinguer entre ces deux cas, au moins on peut affirmer dans les exemples précédents que $v_0$ ne peut pas être un maximum local, comme le résultat suivant l'indique.

**Théorème 2.7.14.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^3$. Soit $v_0 \in U$ un maximum (resp., minimum) local de $f$. Alors, toutes les valeurs propres de $H_f(v_0)$ sont inférieures (resp., supérieures) ou égales à zéro.*

***Preuve.*** On va démontrer le cas où $v_0 \in U$ est un maximum local de $f$, car le cas où $v_0 \in U$ est un minimum local de $f$ suit de remplacer $f$ par $-f$. On va procéder par l'absurde. On suppose que $H_f(v_0)$ possède une valeur propre $\lambda > 0$, et soit $w \in \mathbb{R}^n \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$ un vecteur propre de $H_f(v_0)$ de valeur propre $\lambda$, *i.e.* $H_f(v_0) \cdot w^t = \lambda w^t$. On suppose sans perte de généralité que $||w|| = 1$. Soit $r > 0$ tel que $B_{||\ ||}(v_0, 2r) \subseteq U$. On considère l'application $\alpha : ]-2, 2[ \to U$ donnée par $\alpha(t) = v_0 + tw$ pour $t \in ]-2, 2[$. C'est clair que $\alpha$ est de classe $\mathrm{C}^3$, car $\alpha'(t) = w$ pour tout $t \in ]-2, 2[$, ce qui implique que $f_w = f \circ \alpha :: ]-2, 2[ \to \mathbb{R}$ est de classe $\mathrm{C}^3$. Or, comme $v_0$ est un maximum local de $f$, $0$ est un maximum local de $f_w$. En outre, le (2.6.3) nous dit que

$$0 = f_w'(0) = \langle \nabla f(v_0), w \rangle \text{ et } f_w''(0) = w \cdot H_f(v_0) \cdot w^t = w \cdot \lambda w^t = \lambda > 0.$$

Le Théorème 2.7.7 appliqué à la fonction $f_w$ nous dit alors que $0$ est un minimum local strict de $f_w$, ce qui est absurde car on a supposé que $v_0$ est un maximum local de $f$ et cela implique que $0$ est un maximum local de $f_w$. □

### 2.7.3 Optimisation de fonctions sous contraintes

On va démontrer le résultat suivant comme corollaire du Théorème 2.8.10.

**Lemme 2.7.15.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte, et $g : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^1$. Soit*

$$S = \big\{ v \in U : g(v) = \mathbf{0}_{\mathbb{R}^m} \text{ et } Dg(v) : \mathbb{R}^n \to \mathbb{R}^m \text{ est surjectif} \big\}.$$

*Alors, $T_v S = \mathrm{Ker}(Dg(v))$.*

Le résultat suivant nous dit un critère nécessaire pour trouver des extrema des fonctions avec contraintes.

**Proposition 2.7.16.** *Soient $n, m \in \mathbb{Z}_{>0}$ tels que $n > m$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $g : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^1$ et $f : U \to \mathbb{R}$ une application différentiable. On considère*

$$S = \big\{ v \in U : g(v) = \mathbf{0}_{\mathbb{R}^m} \text{ et } Dg(v) : \mathbb{R}^n \to \mathbb{R}^m \text{ est surjectif} \big\}.$$

*Si $v_0 \in S$ est un extremum local de $f|_S : S \to \mathbb{R}$, alors il existe $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ tel que*

$$\nabla f(v_0) = \lambda \cdot J_g(v_0). \tag{2.7.8}$$

*Si l'on écrit comme d'habitude $g(v) = (g_1(v), \dots, g_m(v))$ pour tout $v \in U$, où $g_i : U \to \mathbb{R}$ est la i-ème fonction coordonnée de $g$, l'identité* (2.7.8) *équivaut à $\nabla f(v_0) = \sum_{i=1}^m \lambda_i \nabla g_i(v_0)$*

***Preuve.*** On écrira $g(v) = (g_1(v), \dots, g_m(v)) \in \mathbb{R}^m$ pour tout $v \in U$. Soit $W_v \subseteq \mathbb{R}^n$ le sous-espace vectoriel engendré par $\mathscr{B}_v = \{\nabla g_1(v), \dots, \nabla g_m(v)\}$, pour tout $v \in S$. Comme $Dg(v)$ est surjectif pour $v \in S$, l'ensemble $\mathscr{B}_v$ est libre, ce qui nous dit que $W_v$ est de dimension $m$. Il suffit de montrer que, si $v_0 \in S$ est un extremum local de $f|_S$, $\nabla f(v_0) \in W_{v_0}$. Or, $\nabla f(v_0) \in W_{v_0}$ si et seulement si $\nabla f(v_0)$ est orthogonal à tout vecteur dans $W_{v_0}^{\perp} = \{u \in \mathbb{R}^n : \langle w, u \rangle = 0 \text{ pour tout } w \in W_{v_0}\}$, vu que $(W_{v_0}^{\perp})^{\perp} = W_{v_0}$. Par définition, $W_{v_0}^{\perp} = \operatorname{Ker}(Dg(v_0))$. D'après le Lemme 2.7.15, étant donné $u' \in W_{v_0}^{\perp} = \operatorname{Ker}(Dg(v_0))$, il existe une application $\alpha : J \to S$ de classe $\mathrm{C}^1$ tel que $\alpha(0) = v_0$ et $\alpha'(0) = u'$, où $J$ est un intervalle ouvert incluant $0$. Cela implique que $f \circ \alpha : J \to \mathbb{R}$ possède un extremum en $t = 0$. En conséquence,

$$0 = (f \circ \alpha)'(0) = \big\langle \nabla f(v_0), \alpha'(0) \big\rangle = \big\langle \nabla f(v_0), u' \big\rangle,$$

ce qui nous dit que $\nabla f(v_0) \in (W_{v_0}^{\perp})^{\perp}$, comme on voulait démontrer. □

On considère les mêmes hypothèse que dans la Proposition 2.7.16. Un point $v_0 \in S$ qui satisfait que $\nabla f(v_0) = \lambda \cdot J_g(v_0)$ sera appelé **point critique lié** de $f$ par rapport à la contrainte déterminée par $g$ et le vecteur $\lambda$ est appelé le vecteur de **multiplicateurs de Lagrange**.

**Définition 2.7.17.** *Soient $n, m \in \mathbb{Z}_{>0}$ tels que $n > m$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $g : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^2$ et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^2$. On écrira comme d'habitude $g(v) = (g_1(v), \dots, g_m(v))$ pour tout $v \in U$, où $g_i : U \to \mathbb{R}$ est la i-ème fonction coordonnée de $g$. On définit la* ***matrice hessienne à bord***

$$H_{f,g}(\lambda, v) = \left( \begin{array}{c|c} \mathbf{0}_{\mathrm{M}_{m\times m}(\mathbb{R})} & -J_g(v) \\ \hline -J_g(v)^t & H_f(v) - \sum_{i=1}^m \lambda_i H_{g_i}(v) \end{array} \right) \in \mathrm{M}_{(m+n)\times(m+n)}(\mathbb{R})$$

*pour tout $\lambda \in \mathbb{R}^m$ et $v \in U$, où $\mathbf{0}_{\mathrm{M}_{m\times m}(\mathbb{R})}$ est la matrice nulle de $\mathrm{M}_{m\times m}(\mathbb{R})$.*

Le critère suivant nous donne des conditions suffisantes pour déterminer la nature d'un point critique lié.

**Théorème 2.7.18.** *Soient $n, m \in \mathbb{Z}_{>0}$ tels que $n > m$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $g : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^3$ et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^3$. On considère*

$$S = \big\{ v \in U : g(v) = \mathbf{0}_{\mathbb{R}^m} \text{ et } Dg(v) : \mathbb{R}^n \to \mathbb{R}^m \text{ est surjectif} \big\}.$$

*Soit $v_0 \in S$ un point critique lié de $f$ par rapport à la contrainte déterminée par $g$ avec vecteur de multiplicateurs de Lagrange $\lambda_0$, i.e. $\nabla f(v_0) = \lambda_0 \cdot J_g(v_0)$. Alors,*

*(L.1) si $w \cdot H_{f,g}(\lambda_0, v_0) \cdot w^t > 0$ pour tout $w \in T_{v_0}S \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, alors $v_0$ est un minimum local strict de $f|_S : S \to \mathbb{R}$;*

*(L.2) si $w \cdot H_{f,g}(\lambda_0, v_0) \cdot w^t < 0$ pour tout $w \in T_{v_0}S \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$, alors $v_0$ est un maximum local strict de $f|_S : S \to \mathbb{R}$;*

*(L.3) s'il existe $w, u \in T_{v_0}S \setminus \{\mathbf{0}_{\mathbb{R}^n}\}$ tels que $w \cdot H_{f,g}(\lambda_0, v_0) \cdot w^t > 0$ et $u \cdot H_{f,g}(\lambda_0, v_0) \cdot u^t < 0$, alors $v_0$ n'est pas un extremum local de $f|_S : S \to \mathbb{R}$.*

*Preuve.* D'après le théorème d'inversion locale 2.8.6, il existe un $V$ est un ouvert de $\mathbb{R}^{n-m}$ incluant $\mathbf{0}_{\mathbb{R}^{n-m}}$, une application injective $\phi : V \to S$ de classe $\mathrm{C}^3$ et un ouvert $W$ de $\mathbb{R}^n$ incluant $v_0$, tels que $\psi(V) = W \cap S$, $g \circ \psi = \mathrm{id}_V$ et $\psi \circ g|_{W\cap S} = \mathrm{id}_{W\cap S}$. Noter en plus que $D\psi(\mathbf{0}_{\mathbb{R}^{n-m}}) : \mathbb{R}^{n-m} \to \mathbb{R}^n$ est une application linéaire injective dont l'image est précisément $T_{v_0}S$. Soit $F = f \circ \psi : V \to \mathbb{R}$. On voit bien que $v_0$ est un maximum (resp., minimum) local strict de $f|_S$ si et seulement si $\mathbf{0}_{\mathbb{R}^{n-m}}$ est un maximum (resp., minimum) local strict de $f \circ \psi$. Alors, le résultat suit du Théorème 2.7.7. □

Un point critique lié $v_0 \in S$ de $f$ par rapport à la contrainte déterminée par $g$ qui satisfait la condition (L.3) est appelé **point-selle lié** ou **point col lié** de $f$.

Le critère suivant réexprime les conditions (L.1)-(L.3) dans le théorème précédent dans un langage numériques plus facile de calculer. La preuve suit des considérations élémentaires mais un peu longues d'algèbre linéaire.

**Corollaire 2.7.19.** *Soient $n, m \in \mathbb{Z}_{>0}$ tels que $n > m$, $U \subseteq \mathbb{R}^n$ une partie ouverte, $g : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^3$ et $f : U \to \mathbb{R}$ une application de classe $\mathrm{C}^3$. On considère*

$$S = \{v \in U : g(v) = \mathbf{0}_{\mathbb{R}^m} \text{ et } Dg(v) : \mathbb{R}^n \to \mathbb{R}^m \text{ est surjectif}\}.$$

*Soit $v_0 \in S$ un point critique lié de $f$ par rapport à la contrainte déterminée par $g$ avec vecteur de multiplicateurs de Lagrange $\lambda_0$, i.e. $\nabla f(v_0) = \lambda_0 \cdot J_g(v_0)$. Pour $j \in [\![1, n+m]\!]$, on notera $H_j$ la matrice carrée formée des éléments de $H_{f,g}(\lambda_0, x_0)$ dans les premières $j$ lignes et $j$ colonnes. Alors,*

*(LD.1) si $(-1)^m \det(H_j) > 0$ pour tout $j \in [\![2m+1, n+m]\!]$, alors $v_0$ est un minimum local strict de $f|_S : S \to \mathbb{R}$;*

*(LD.2) si $(-1)^{j-m} \det(H_j) > 0$ pour tout $j \in [\![2m+1, n+m]\!]$, alors $v_0$ est un maximum local strict de $f|_S : S \to \mathbb{R}$;*

*(LD.3) si $\det(H_j) \neq 0$ pour tout $j \in [\![2m+1, n+m]\!]$ mais on n'est pas dans les cas (LD.1) ou (LD.2), alors $v_0$ n'est pas un extremum local de $f|_S : S \to \mathbb{R}$.*

Dans la plupart des cas, on travaillera avec une seule contrainte, donc ce sera utile de réexprimer le corollaire précédent dans ce cas particulier.

**Corollaire 2.7.20.** *Soient $n \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ une partie ouverte et $f, g : U \to \mathbb{R}$ des applications de classe $\mathrm{C}^3$. On considère*

$$S = \{v \in U : g(v) = 0 \text{ et } \nabla g(v) \neq \mathbf{0}_{\mathbb{R}^n}\}.$$

*Soit $v_0 \in S$ un point critique lié de $f$ par rapport à la contrainte déterminée par $g$ avec multiplicateur de Lagrange $\lambda_0 \in \mathbb{R}$, i.e. $\nabla f(v_0) = \lambda_0 \cdot \nabla g(v_0)$. Pour $j \in [\![1, n+1]\!]$, on notera $H_j$ la matrice carrée formée des éléments de $H_{f,g}(\lambda_0, x_0)$ dans les premières $j$ lignes et $j$ colonnes. Alors,*

*(LD1.1) si $\det(H_j) < 0$ pour tout $j \in [\![3, n+1]\!]$, alors $v_0$ est un minimum local strict de $f|_S : S \to \mathbb{R}$;*

*(LD1.2) si $(-1)^j \det(H_j) < 0$ pour tout $j \in [\![3, n+1]\!]$, alors $v_0$ est un maximum local strict de $f|_S : S \to \mathbb{R}$;*

*(LD1.3) si $\det(H_j) \neq 0$ pour tout $j \in [\![3, n+1]\!]$ mais on n'est pas dans les cas (LD1.1) ou (LD1.2), alors $v_0$ n'est pas un extremum local de $f|_S : S \to \mathbb{R}$.*

**Exemple 2.7.21.** *On veut construire le parallélépipède de volume maximal dont la surface extérieur soit de valeur fixe $S_0 > 0$. Soient $f, g : \mathbb{R}^3_{>0} \to \mathbb{R}$ les applications donnés par $f(x_1, x_2, x_3) = x_1x_2x_3$ et $g(x_1, x_2, x_3) = 2(x_1x_2 + x_2x_3 + x_3x_1) - S_0$ pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3_{>0}$. C'est clair que $f$ et $g$ sont de classe $\mathrm{C}^1$. En outre,*

$$\nabla g(x_1, x_2, x_3) = 2(x_2 + x_3, x_3 + x_1, x_1 + x_2) \neq (0, 0, 0)$$

*pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3_{>0}$. Soit $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3_{>0} : g(x_1, x_2, x_3) = 0\}$. On veut trouver un maximum de $f|_S : S \to \mathbb{R}$. Soit $(x_{1,0}, x_{2,0}, x_{3,0}) \in S$ un maximum de $f|_S$. Par la Proposition 2.7.16 il existe $\lambda \in \mathbb{R}$ tel que $\nabla f(x_{1,0}, x_{2,0}, x_{3,0}) = \lambda \nabla g(x_{1,0}, x_{2,0}, x_{3,0})$, i.e.*

$$\begin{cases} x_{2,0}x_{3,0} = 2\lambda(x_{2,0} + x_{3,0}), \\ x_{3,0}x_{1,0} = 2\lambda(x_{3,0} + x_{1,0}), \\ x_{1,0}x_{2,0} = 2\lambda(x_{1,0} + x_{2,0}), \end{cases} \tag{2.7.9}$$

*Comme $(x_{1,0}, x_{2,0}, x_{3,0}) \in S$, on a aussi*

$$2(x_{1,0}x_{2,0} + x_{2,0}x_{3,0} + x_3x_{1,0}) = S_0. \tag{2.7.10}$$

*Si l'on fait la somme des trois identités dans* (2.7.9) *et on utilise l'identité précédente, on trouve que*

$$8\lambda(x_{1,0} + x_{2,0} + x_{3,0}) = S_0.$$

*Comme $S_0 > 0$, l'identité précédente nous dit que $\lambda > 0$. Si l'on fait le produit de la $i$-ème identité dans* (2.7.9) *par $x_{i,0}$ pour $i \in \{1, 2, 3\}$, on conclut que*

$$2\lambda(x_{1,0}x_{2,0} + x_{3,0}x_{1,0}) = 2\lambda(x_{2,0}x_{3,0} + x_{1,0}x_{2,0}) = 2\lambda(x_{3,0}x_{1,0} + x_{2,0}x_{3,0}),$$

*ce qui implique que*

$$x_{1,0}x_{2,0} = x_{2,0}x_{3,0} = x_{3,0}x_{1,0}.$$

*Si l'on utilise l'identité précédente avec* (2.7.9)*, on trouve*

$$x_{2,0} + x_{3,0} = x_{3,0} + x_{1,0} = x_{1,0} + x_{2,0}$$

*ce qui implique $x_{1,0} = x_{2,0} = x_{3,0}$. L'identité* (2.7.10) *nous dit alors que $x_{1,0} = x_{2,0} = x_{3,0} = \sqrt{S_0/6}$.*

*On va montrer que le point $(\sqrt{S_0/6}, \sqrt{S_0/6}, \sqrt{S_0/6})$ est un maximum local strict de $f|_S$. Le multiplicateur de Lagrange $\lambda = \sqrt{S_0/6}/4$ suit directement de* (2.7.9)*. En outre, on voit bien que la matrice hessienne de l'application associée $F(\Lambda, x_1, x_2, x_3) = f(x_1, x_2, x_3) - \Lambda g(x_1, x_2, x_3)$ est donnée par*

$$H_F(\Lambda, x_1, x_2, x_3) = \begin{pmatrix} 0 & -2(y+z) & -2(z+x) & -2(x+y) \\ -2(y+z) & 0 & z-2\Lambda & y-2\Lambda \\ -2(z+x) & z-2\Lambda & 0 & x-2\Lambda \\ -2(x+y) & y-2\Lambda & x-2\Lambda & 0 \end{pmatrix},$$

*ce qui implique que*

$$H_F(\lambda_0, x_{1,0}, x_{2,0}, x_{3,0}) = \begin{pmatrix} 0 & -8a & -8a & -8a \\ -8a & 0 & a\Lambda & a \\ -8a & a & 0 & a \\ -8a & a & a & 0 \end{pmatrix},$$

*avec $a = \sqrt{S_0/24} > 0$. Cela implique que $\det(H_3) = 128a^3 > 0$ et $\det(H_4) = -192a^4 < 0$. Cela correspond au cas (LD1.2), ce qui implique que $(\sqrt{S_0/6}, \sqrt{S_0/6}, \sqrt{S_0/6})$ est un maximum local strict de $f|_S$.*

## 2.8 Les théorèmes d'inversion locale et de la fonction implicite

### 2.8.1 Quelques résultats préliminaires: un théorème du point fixe et les applications linéaires inversibles

On commence avec ce résultat de S. Banach sur l'existence des points fixes des applications contractantes.

**Théorème 2.8.1.** *Soient $\mathbb{E}$ un espace vectoriel complet muni d'une norme $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$, $S \subseteq \mathbb{E}$ une partie fermé et $f : S \to S$ une application **contractante**, i.e. il existe $K \in ]\,0, 1\,[$ tel que*

$$N\big(f(v) - f(w)\big) \leq KN(v - w)$$

*pour tous $v, w \in S$. Alors, il existe un unique **point fixe** $v_0 \in S$, i.e. $f(v_0) = v_0$. En plus, étant donné $v \in S$, la suite $(f^n(v))_{n\in\mathbb{N}}$ est un suite de Cauchy qui converge vers $v_0$.*

***Preuve.*** Pour démontrer l'unicité du point fixe, soient $v_0, w_0 \in S$ tels que $f(v_0) = v_0$ et $f(w_0) = w_0$. Alors,

$$N(v_0 - w_0) = N\big(f(v_0) - f(w_0)\big) \leq KN(v_0 - w_0),$$

ce qui implique $N(v_0 - w_0) = 0$, vu que $K \in ]\,0, 1\,[$. En conséquence, $v_0 = w_0$.

Soit $v \in S$. On considère la suite $(f^n(v))_{n\in\mathbb{N}} \in S^{\mathbb{N}}$. Alors, pour des entiers $n > m > 0$, on a

$$N\big(f^n(v) - f^m(v)\big) = N\bigg(\sum_{i=m}^{n-1} f^{i+1}(v) - f^i(v)\bigg) \leq \sum_{i=m}^{n-1} N\big(f^{i+1}(v) - f^i(v)\big) \leq \sum_{i=m}^{n-1} K^i N\big(f(v) - v\big)$$
$$\leq \sum_{i=m}^{+\infty} K^i N\big(f(v) - v\big) = \frac{K^m N\big(f(v) - v\big)}{1 - K}.$$

En conséquence, la suite $(f^n(v))_{n\in\mathbb{N}}$ est un suite de Cauchy. Soit

$$v_0 = \lim_{n\to+\infty} f^n(v).$$

Comme $S$ est fermé, $v_0 \in S$. En outre,

$$f(v_0) = f\bigg(\lim_{n\to+\infty} f^n(v)\bigg) = \lim_{n\to+\infty} f^{n+1}(v) = v_0,$$

vu que $f$ est continue, vu qu'elle est contractante. En conséquence, $v_0$ est un point fixe de $f$. □

En outre, on aurait besoin du résultat suivant.

**Proposition 2.8.2.** *Soit $\mathbb{E}$ un espace vectoriel de dimension finie, muni d'une norme $N : \mathbb{E} \to \mathbb{R}_{\geq 0}$. On rappelle que* $\mathrm{L}(\mathbb{E}) = \{f : \mathbb{E} \to \mathbb{E}$ *tel que $f$ est linéaire*$\}$, *qui sera muni de la norme associée* $\mathbf{N}_{N,N} : \mathrm{L}(\mathbb{E}) \to \mathbb{R}_{\geq 0}$. *On pose* $\mathrm{Inv}(\mathbb{E}) = \{f \in \mathrm{L}(\mathbb{E}) : f$ *est inversible* $\}$. *Si $f : \mathbb{E} \to \mathbb{E}$ une application linéaire tel que $\mathbf{N}_{N,N}(f) < 1$, alors* $\mathrm{id}_{\mathbb{E}} - f \in \mathrm{Inv}(\mathbb{E})$. *En plus,* $\mathrm{Inv}(\mathbb{E})$ *est un ouvert de* $\mathrm{L}(\mathbb{E})$ *et l'application* $\mathrm{inv} : \mathrm{Inv}(\mathbb{E}) \to \mathrm{Inv}(\mathbb{E})$ *qui associe $h^{-1}$ à $h \in \mathrm{Inv}(\mathbb{E})$ est de classe* $\mathrm{C}^{\infty}$.

***Preuve.*** On démontre d'abord le premier énonce. Pour $n \in \mathbb{N}$, on pose $g_n = \sum_{i=0}^{n} f^i \in \mathrm{L}(\mathbb{E})$. Alors,

$$\mathbf{N}_{N,N}(g_n - g_m) = \mathbf{N}_{N,N}\bigg(\sum_{i=m+1}^{n} f^i\bigg) \leq \sum_{i=m+1}^{n} \mathbf{N}_{N,N}(f)^i \leq \sum_{i=m+1}^{+\infty} \mathbf{N}_{N,N}(f)^i \leq \frac{\mathbf{N}_{N,N}(f)^{m+1}}{1 - \mathbf{N}_{N,N}(f)}$$

pour tous $n, m \in \mathbb{N}$ avec $n \geq m$, ce qui nous dit que la suite $(g_n)_{n\in\mathbb{N}}$ est un suite de Cauchy de $\mathrm{L}(\mathbb{E})$. Comme $\mathrm{L}(\mathbb{E})$ est complet, vu qu'il s'agit d'un espace vectoriel normé de dimension finie, la limite de la suite $(g_n)_{n\in\mathbb{N}}$ existe. On la notera $g \in \mathrm{L}(\mathbb{E})$. On affirme que $g(\mathrm{id}_{\mathbb{E}} - f) = (\mathrm{id}_{\mathbb{E}} - f)g = \mathrm{id}_{\mathbb{E}}$. On va démontrer $g(\mathrm{id}_{\mathbb{E}} - f) = \mathrm{id}_{\mathbb{E}}$ et on laisse l'autre égalité au lecteur/à la lectrice. Soit $R_{\mathrm{id}_{\mathbb{E}} - f} : \mathrm{L}(\mathbb{E}) \to \mathrm{L}(\mathbb{E})$ l'application donnée par $R_{\mathrm{id}_{\mathbb{E}} - f}(h) = h(\mathrm{id}_{\mathbb{E}} - f)$ pour $h \in \mathrm{L}(\mathbb{E})$. On remarque que $R_{\mathrm{id}_{\mathbb{E}} - f}$ est linéaire, donc continue. Or,

$$gf = R_{\mathrm{id}_{\mathbb{E}} - f}(g) = R_{\mathrm{id}_{\mathbb{E}} - f}\bigg(\lim_{n\to+\infty} \sum_{i=0}^{n} f^i\bigg) = \lim_{n\to+\infty} R_{\mathrm{id}_{\mathbb{E}} - f}\bigg(\sum_{i=0}^{n} f^i\bigg)$$
$$= \lim_{n\to+\infty} \sum_{i=0}^{n} f^i(\mathrm{id}_{\mathbb{E}} - f) = \lim_{n\to+\infty} (\mathrm{id}_{\mathbb{E}} - f^{n+1}) = \mathrm{id}_{\mathbb{E}}.$$

En conséquence, $g$ est l'inverse de $\mathrm{id}_{\mathbb{E}} - f$.

On va montrer que $\mathrm{Inv}(\mathbb{E})$ est ouvert. Étant donné, $h \in \mathrm{Inv}(\mathbb{E})$, on remarque d'abord que $\mathbf{N}_{N,N}(h) \neq 0$ et $\mathbf{N}_{N,N}(h^{-1}) \neq 0$, vu que $h, h^{-1} \neq \mathbf{0}_{\mathrm{L}(\mathbb{E})}$ car l'application nulle de $\mathbb{E}$ n'est pas inversible. Soit $h \in \mathrm{Inv}(\mathbb{E})$ et $r = \mathbf{N}_{N,N}(h^{-1})^{-1} > 0$. On affirme que $B_{\mathbf{N}_{N,N}}(h, r) \subseteq \mathrm{Inv}(\mathbb{E})$. En effet, si $u \in B_{\mathbf{N}_{N,N}}(h, r)$, alors $\mathbf{N}_{N,N}(u - h) < \mathbf{N}_{N,N}(h^{-1})^{-1}$. Cela nous dit que

$$\mathbf{N}_{N,N}(h^{-1}u - \mathrm{id}_{\mathbb{E}}) = \mathbf{N}_{N,N}\big(h^{-1}(u - h)\big) \leq \mathbf{N}_{N,N}(h^{-1})\mathbf{N}_{N,N}(u - h) < \mathbf{N}_{N,N}(h^{-1})\mathbf{N}_{N,N}(h^{-1})^{-1} = 1,$$

et d'après la première partie de la preuve, on a $h^{-1}u \in \mathrm{Inv}(\mathbb{E})$, ce qui implique que $u \in \mathrm{Inv}(\mathbb{E})$, comme on voulait démontrer.

Pour montrer que l'application $\operatorname{inv} : \operatorname{Inv}(\mathbb{E}) \to \operatorname{Inv}(\mathbb{E})$ est de classe $C^\infty$, on montre d'abord qu'elle est différentiable et que $D\operatorname{inv}(u)(h) = -u^{-1}hu^{-1}$ pour $u \in \operatorname{Inv}(\mathbb{E})$ et $h \in L(\mathbb{E})$. Si $||h|| < ||u^{-1}||^{-1}$, alors $||u^{-1}h|| < 1$ et la première partie de cette preuve nous dit que

$$\begin{aligned}(u+h)^{-1} - u^{-1} + u^{-1}hu^{-1} &= \big((\mathrm{id}_{\mathbb{E}} + u^{-1}h)^{-1} - \mathrm{id}_{\mathbb{E}} + + u^{-1}h\big)u^{-1} \\ &= \Bigg(\sum_{i=0}^{+\infty}(-1)^i(u^{-1}h)^i - \mathrm{id}_{\mathbb{E}} + u^{-1}h\Bigg)u^{-1} = \sum_{i=2}^{+\infty}(-1)^i(u^{-1}h)^i u^{-1}.\end{aligned}$$

En outre,

$$\Bigg|\Bigg|\sum_{i=2}^{+\infty}(-1)^i(u^{-1}h)^i u^{-1}\Bigg|\Bigg| \leq \sum_{i=2}^{+\infty}||u^{-1}h||^{i-2}||u^{-1}||^3||h||^2 \leq \sum_{i=0}^{+\infty}\frac{1}{2^i}||u^{-1}||^3||h||^2 = ||u^{-1}||^3||h||^2,$$

ce qui implique que $\operatorname{inv} : \operatorname{Inv}(\mathbb{E}) \to \operatorname{Inv}(\mathbb{E})$ est différentiable et sa différentielle satisfait $D\operatorname{inv}(u)(h) = -u^{-1}hu^{-1}$ pour $u \in \operatorname{Inv}(\mathbb{E})$ et $h \in L(\mathbb{E})$. Comme $\operatorname{inv}$ est différentiable, elle est continue, ce qui implique que $D\operatorname{inv}(u)(h) = -u^{-1}hu^{-1}$ est continue, et, en conséquence, $\operatorname{inv}$ est de classe $C^1$. Un argument par récurrence direct nous dit alors que $\operatorname{inv}$ est de classe $C^p$ pour tout $p \in \mathbb{Z}_{>0}$, vu que la différentielle est formé comme un produit de fonctions qui font intervenir seulement la même application $\operatorname{inv}$. En conséquence, $\operatorname{inv}$ est de classe $C^\infty$. □

**Remarque 2.8.3.** *La Proposition 2.8.2 ainsi que sa preuve sont encore valables pour tout espace vectoriel normé complet $\mathbb{E}$, car dans ce cas l'espace vectoriel normé $L(\mathbb{E})$ muni de la norme associée à la norme de $\mathbb{E}$ est complet (voir* [1], *Thm. XVII.1.1).*

### 2.8.2 Le théorème d'inversion locale

**Définition 2.8.4.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte, $v_0 \in U$ et $f : U \to \mathbb{R}^m$ une application et $p \in \mathbb{Z}_{>0} \cup \{\infty\}$. On dit que $f$ est* ***localement de classe $C^p$ en $v_0$*** *s'il existe un ouvert $U' \subseteq U$ tel que $v_0 \in U'$ et $f|_{U'} : U' \to \mathbb{R}^m$ est de classe $C^p$. On dit que $f : U \to f(U)$ est un* ***$C^p$-difféomorphisme*** *si $f(U)$ est un ouvert, $f : U \to f(U)$ est bijective et sa réciproque $g : f(U) \to U$ est aussi de classe $C^p$. Finalement, on dira que $f$ est un* ***$C^p$-difféomorphisme local*** *en $v_0$ s'il existe un ouvert $U' \subseteq U$ tel que $v_0 \in U'$ et $f|_{U'} : U' \to f(U')$ est un $C^p$-difféomorphisme.*

**Lemme 2.8.5.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte, $v_0 \in U$ et $f : U \to \mathbb{R}^m$ une application et $p \in \mathbb{Z}_{>0} \cup \{\infty\}$. Si $f$ est $C^p$-difféomorphisme local, alors $n = m$ et, si $U' \subseteq U$ est l'ouvert tel que $v_0 \in U'$, $f(U')$ ouvert, $f|_{U'} : U' \to f(U')$ est bijective et de classe $C^p$, avec réciproque $g : f(U') \to U'$ de classe $C^p$, alors l'application linéaire $Df(v_0)$ est inversible et*

$$Dg\big(f(v_0)\big) = Df(v_0)^{-1}.$$

***Preuve.*** On écrira $f$ au lieu de $f|_{U'} : U' \to f(U')$ pour réduire espace. Comme $f \circ g = \mathrm{id}_{f(U')}$ et $g \circ f = \mathrm{id}_{U'}$, le Théorème 2.1.17 nous dit que $Df(g(w)) \circ Dg(w) = \mathrm{id}_{\mathbb{R}^m}$ et $Dg(f(v)) \circ Df(v) = \mathrm{id}_{\mathbb{R}^n}$ pour tous $v \in U'$ et $w \in f(U')$, ce qui donne une bijection linéaire entre $\mathbb{R}^n$ et $\mathbb{R}^m$ et, en particulier, $n = m$. En outre, les identités précédentes pour $v = v_0$ et $w = f(v_0)$ nous disent précisément que l'application linéaire $Df(v_0)$ est inversible et

$$Dg\big(f(v_0)\big) = Df(v_0)^{-1}.$$

□

Le but de cette section sera de démontrer les deux résultats fondamentaux suivants, qui donnent essentiellement les réciproques du Lemme 2.8.5.

**Théorème 2.8.6** (Théorème d'inversion locale I)**.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte, $v_0 \in U$ et $f : U \to \mathbb{R}^m$ une application de classe $C^1$ telle que $Df(v_0) : \mathbb{R}^n \to \mathbb{R}^m$ est une matrice inversible. Alors, $f$ est $C^1$-difféomorphisme local en $v_0$.*

***Preuve.*** Soit $L = Df(v_0)$. Comme $L$ est inversible, on a une bijection linéaire entre $\mathbb{R}^n$ et $\mathbb{R}^m$ et, en particulier, $n = m$. Pour $v \in \mathbb{R}^n$, on pose $T_v : \mathbb{R}^n \to \mathbb{R}^n$ l'application donnée par $T_v(w) = v + w$ pour tout $w \in \mathbb{R}^n$. Les applications $L$ et $T_v$ sont de classe $C^\infty$, d'après l'Exemple 2.5.3. On pose

$F = L^{-1} \circ T_{-f(v_0)} \circ f \circ T_{v_0}|_{T_{-v_0}(U)} : T_{-v_0}(U) \to \mathbb{R}^n$. Comme $f = T_{f(v_0)} \circ L \circ F \circ T_{-v_0}|_U$, $F$ est de localement de classe $C^1$ en $\mathbf{0}_{\mathbb{R}^n}$ si et seulement si $f$ est localement de classe $C^1$ en $v_0$, et dans ce cas $DF(\mathbf{0}_{\mathbb{R}^n}) = \mathrm{id}_{\mathbb{R}^n} = L^{-1} \circ Df(v_0) = \mathrm{id}_{\mathbb{R}^n}$, d'après le Théorème 2.1.17. Comme $F(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathbb{R}^n}$ et $DF(\mathbf{0}_{\mathbb{R}^n}) = \mathrm{id}_{\mathbb{R}^n}$, il suffit de démontrer le théorème dans le cas $v_0 = \mathbf{0}_{\mathbb{R}^n}$, $f(v_0) = \mathbf{0}_{\mathbb{R}^n}$ et $Df(v_0) = \mathrm{id}_{\mathbb{R}^n}$, que l'on supposera désormais.

Soit $g : U \to \mathbb{R}^n$ l 'application donnée par $g(v) = v - f(v)$ pour tout $v \in U$. Noter que $g(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathbb{R}^n}$. Comme $f$ est de classe $C^1$, $g$ l'est aussi, et $Dg(\mathbf{0}_{\mathbb{R}^n}) = \mathrm{id}_{\mathbb{R}^n} - Df(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathrm{L}(\mathbb{R}^n)}$. Comme $Dg(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathrm{L}(\mathbb{R}^n)}$, $||Dg(\mathbf{0}_{\mathbb{R}^n})|| = 0$ et $Dg$ est continue, il existe $r > 0$ tel que $||Dg(v)|| \leq 1/2$ pour tout $v \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$. Comme $Df(v) = \mathrm{id}_{\mathbb{R}^n} - Dg(\mathbf{0}_{\mathbb{R}^n})$, la Proposition 2.8.2 nous dit aussi que $Df(v)$ est inversible pour tout $v \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$. Par la Proposition 2.3.6, $||g(v)|| = ||g(v) - g(\mathbf{0}_{\mathbb{R}^n})|| \leq ||v||/2$ pour tout $v \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$, ce qui implique que $g(\bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)) \subseteq \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r/2)$.

Soit $w \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r/2)$. On affirme qu'il existe un unique $v \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$ tel que $f(v) = w$. En effet, soit l'application $f_w : U \to \mathbb{R}^n$ donnée par $f_w(u) = g(u) + w$ pour tout $v \in U$. Alors, $f_w(\bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)) \subseteq \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$, vu que $||f_w(u)|| = ||g(u) + w|| \leq ||g(u)|| + ||w|| \leq r/2 + r/2 = r$ pour tout $u \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$. Noter que $||f_w(u) - f_w(u')|| = ||g(u) - g(u')|| \leq ||u - u'||/2$ pour tous $u, u' \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$, par la Proposition 2.3.6 appliquée à $g$. En conséquence, $f_w|_{\bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)} : \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r) \to \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$ est contractante et le Théorème 2.8.1 nous dit que cette application possède un seul point fixe $v \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$, *i.e.* $v - f(v) + w = f_w(v) = v$, ce qui équivaut à $f(v) = w$. Par conséquent, il existe un unique $v \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$ tel que $f(v) = w$.

Soit $U' = \{v \in B_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r) : f(v) < r/2\}$. C'est clair que $U'$ est ouvert et $\mathbf{0}_{\mathbb{R}^n} \in U'$. L'argument dans le paragraphe précédent montre que la restriction $f|_{U'} : U' \to f(U') \subseteq B_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r/2)$ est bijective. Soit $g : f(U') \to U'$ l'application réciproque.

On va démontrer que $f(U')$ est ouvert et $g$ est de classe $C^1$. Or,

$$||v'-v|| = ||v'-f(v')-v+f(v)+f(v')-f(v)|| \leq ||g(v')-g(v)||+||f(v')-f(v)|| \leq \frac{1}{2}||v-v'||+||f(v')-f(v)||$$

pour tous $v, v' \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$, ce qui implique que

$$||v' - v|| \leq 2||f(v') - f(v)|| \tag{2.8.1}$$

pour tous $v, v' \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$. Soit $w = f(v) \in f(U')$, avec $v \in U'$. Comme $U'$ est ouvert, il existe $s > 0$ tel que $B_{||\ ||}(v, s) \subseteq U'$, et comme $f(U') \subseteq B_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r/2)$, il existe $t > 0$ avec $t < s/2$ tel que $B_{||\ ||}(w, t) \subseteq B_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r/2)$. En outre, comme tout point $w'$ de $B_{||\ ||}(w, t) \subseteq \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r/2)$ possède un unique antécédent dans $\bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$ pour $f$, il existe $v' \in \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$ tel que $f(v') = w'$. L'inégalité (2.8.1) nous dit alors que $||v' - v|| \leq 2||w - w'|| < 2t < s$ pour tout $w' \in B_{||\ ||}(w, t)$, ce qui implique que $v' \in B_{||\ ||}(v, s) \subseteq U'$ et, en particulier, $w' = f(v') \in f(U')$. En conséquence, $B_{||\ ||}(w, t) \subseteq f(U')$, ce qui nous dit que $f(U')$ est ouvert. En plus, l'inégalité (2.8.1) nous dit aussi que $g$ est continue, en remplaçant $f(v)$ et $f(v')$ par $w$ et $w'$, resp., et $v$ et $v'$ par $g(w)$ et $g(w')$, resp.

On va finalement démontrer que $g$ est différentiable et $Dg(f(v)) = Df(v)^{-1}$ pour tout $v \in U'$. On rappelle d'abord que, comme $U' \subseteq \bar{B}_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, r)$, $Df(v)$ est inversible pour tout $v \in U'$. Comme $f$ est différentiable, on peut écrire

$$f(v + h) - f(v) = Df(v)(h) + ||h||\varphi(h) \tag{2.8.2}$$

pour tout $v \in U'$ et $h \in \mathbb{R}^n$ tel que $||h|| < s$ où $B_{||\ ||}(v, s) \subseteq U'$ et $\varphi : B_{||\ ||}(\mathbf{0}_{\mathbb{R}^n}, s) \to \mathbb{R}^n$ satisfait que

$$\lim_{h \to \mathbf{0}_{\mathbb{R}^n}} \varphi(h) = \mathbf{0}_{\mathbb{R}^n}.$$

On fixe désormais $v \in U'$. Comme $\mathrm{inv} : \mathrm{Inv}(\mathbb{E}) \to \mathrm{Inv}(\mathbb{E})$ est continue, il existe $s' > 0$ tel que $s' < s$ et $C > 0$ tel que $||Df(v + h)^{-1}|| \leq C$ pour tout $h$ tel que $||h|| < s'$. Or, si l'on choisit $w = f(v)$ et $k = f(v + h) - f(v)$ dans $g(w + k) - g(w) - Df(v)^{-1}(k)$ on trouve

$$\begin{aligned} g(w + k) - g(w) - Df(v)^{-1}(k) &= v + h - v - Df(v)^{-1}\big(f(v + h) - f(v)\big) \\ &= h - Df(v)^{-1}\big(f(v + h) - f(v)\big) = ||h||Df(v)^{-1}\big(\varphi(h)\big), \end{aligned}$$

où l'on a utilisé (2.8.2). Cela nous dit que

$$||g(w + k) - g(w) - Df(v)^{-1}(k)|| = \Big\| ||h||Df(v)^{-1}\big(\varphi(h)\big)\Big\| \leq ||h||C\big\|\varphi(h)\big\|,$$

pour tout $||h|| < s'$, ce qui implique que $g$ est différentiable en $w = f(v)$ et $Dg(w) = Df(g(w))^{-1}$. Comme l'application $\mathrm{inv}$ est continue d'après la Proposition 2.8.2 et $g$ aussi, on conclut que $Dg(w) = Df(g(w))^{-1}$ est une fonction continue de $w$, ce qui nous dit que $g$ est de classe $\mathrm{C}^1$. □

**Corollaire 2.8.7** (Théorème d'inversion locale II)**.** *Soit $U \subseteq \mathbb{R}^n$ une partie ouverte, $v_0 \in U$ et $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^p$ pour $p \in \mathbb{Z}_{>0}$ telle que $Df(v_0) : \mathbb{R}^n \to \mathbb{R}^m$ est une matrice inversible. Alors, $f$ est $\mathrm{C}^p$-difféomorphisme local en $v_0$.*

***Preuve.*** On procède par récurrence, avec le cas $p = 1$ donné par le Théorème 2.8.6. On suppose que le résultat est vrai pour toute fonction de classe $\mathrm{C}^p$. Soit $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^{p+1}$ telle que $Df(v_0) : \mathbb{R}^n \to \mathbb{R}^m$ est une matrice inversible pour $v_0 \in U$. Alors, par récurrence, $f$ est $\mathrm{C}^p$-difféomorphisme local en $v_0$, soit $U' \subseteq U$ l'ouvert tel que $v_0 \in U'$, $f(U')$ est ouvert, $f|_{U'} : U' \to f(U')$ est bijective et sa réciproque $g : f(U') \to U'$ est de classe $\mathrm{C}^p$. L'application $Dg(w) = Df(g(w))^{-1}$ est une fonction de classe $\mathrm{C}^p$ pour $w \in f(U')$, car c'est la composition de l'application $g$ de classe $\mathrm{C}^p$ et de l'application $\mathrm{inv}$, qui est de classe $\mathrm{C}^\infty$ d'après la Proposition 2.8.2. Comme $Dg(w)$ est de classe $\mathrm{C}^p$ pour $w \in f(U')$, $g$ est de classe $\mathrm{C}^{p+1}$, comme on voulait démontrer. □

**Exemple 2.8.8.** *L'hypothèse que $f$ soit au moins de classe $\mathrm{C}^1$ dans le Théorème 2.8.6 est nécessaire. En effet, si l'on suppose que $f$ est seulement différentiable, alors le résultat n'est plus vrai en général. Par exemple, si l'on considère l'application $f : \mathbb{R} \to \mathbb{R}$ donnée par*

$$f(x) = \begin{cases} x + 2x^2 \sin\left(\frac{1}{x}\right), & \text{si } x \in \mathbb{R} \setminus \{0\}, \\ 0, & \text{si } x = 0, \end{cases}$$

*alors $f$ est différentiable en tout point et $f'(0) = 1$. Par contre, étant donné $r > 0$ arbitraire, la restriction $f|_{]-r,r[}$ n'est pas injective. En effet, comme $f'(x) = 1 + 4x\sin(1/x) - 2\cos(1/x)$ pour tout $x \in \mathbb{R} \setminus \{0\}$, elle est continue sur $\mathbb{R} \setminus \{0\}$, et l'on a*

$$f'\left(\frac{1}{n\pi}\right) = 1 - 2(-1)^n = \begin{cases} -1, & \text{si } n \in \mathbb{Z}_{>0} \text{ est pair,} \\ 3, & \text{si } n \in \mathbb{Z}_{>0} \text{ est impair,} \end{cases}$$

*pour tout $n \in \mathbb{Z}_{>0}$. Le théorème des valeurs intermédiaires nous dit alors qu'il existe des suites $(x_n)_{n\in\mathbb{N}}, (r_n)_{n\in\mathbb{N}} \in \mathbb{R}_{>0}^{\mathbb{N}}$ tels que $x_n$ est décroissante et tend vers zéro lorsque $n$ tend vers l'infini, $f'(x) < 0$ pour tout $x \in [x_n - r_n, x_n[$ et $f'(x) > 0$ pour tout $x \in ]x_n, x_n + r_n]$. Cela nous dit que $x_n \in \mathbb{R}_{>0}$ est un minimum local strict de $f$ pour tout $n \in \mathbb{N}$ et, en particulier, la restriction $f|_{[x_n - r_n, x_n + r_n]}$ ne peut pas être injective. Noter que la continuité de $f'|_{\mathbb{R}_{>0}}$ et le fait que $f'$ est une fonction paire nous disent que $f'(x_n) = f'(-x_n) = 0$ pour tout $n \in \mathbb{N}$, et par définition on a $0 < x_{n+1} < x_{n+1} + r_{n+1} < x_n - r_n$ pour tout $n \in \mathbb{N}$. Finalement, étant donné $r > 0$, comme $x_n$ tend vers zéro lorsque $n$ tend vers l'infini, il existe $n_0 > 0$ tel que $x_{n+1} < r$ pour tout $n > n_0$, ce qui implique que $[x_n - r_n, x_n + r_n] \subseteq [0, r[$ pour tout $n > n_0$. Comme la restriction $f|_{[x_n - r_n, x_n + r_n]}$ n'est pas injective, la restriction $f|_{]-r,r[}$ n'est pas injective non plus.*

**Remarque 2.8.9.** *Malgré l'Exemple 2.8.8, le Théorème 2.8.6 reste vrai si l'on remplace l'hypothèse d'être de classe $\mathrm{C}^p$ par la simple différentiabilité de $f$ mais l'inversibilité de $Df(v_0)$ par l'inversibilité de $Df(v)$ pour tout $v$ dans un ouvert $V \subseteq U$ incluant $v_0$ (voir* [2]*).*

### 2.8.3 Le théorème de la fonction implicite

On présente d'abord l'une des applications les plus importantes du théorème d'inversion locale.

**Théorème 2.8.10** (Théorème de la fonction implicite)**.** *Soient $n, m \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^{m+n}$ une partie ouverte, $v_0 \in U$ que l'on écrira $v_0 = (v_0', v_0'')$ avec $v_0' \in \mathbb{R}^m$ $v_0'' \in \mathbb{R}^n$, et $f : U \to \mathbb{R}^m$ une application de classe $\mathrm{C}^p$ avec $p \in \mathbb{Z}_{>0}$. On suppose que $f(v_0) = \mathbf{0}_{\mathbb{R}^m}$ et que la sous-matrice carrée $J_f'(v_0) \in \mathrm{M}_{m\times m}(\mathbb{R})$ de $J_f(v_0) \in \mathrm{M}_{m\times(m+n)}(\mathbb{R})$ formée des premières $m$ colonnes est inversible. Alors, il existe des ouverts $U' \subseteq U$ et $W \subseteq \mathbb{R}^n$ tels que $v_0 \in U'$ et $v_0'' \in W$, et une unique application $g : W \to \mathbb{R}^m$ de classe $\mathrm{C}^p$ telle que $g(v_0'') = v_0'$, $(g(w), w) \in U$ pour tout $w \in W$ et*

$$\big\{(g(w), w) : w \in W\big\} = U' \cap \big\{x \in U : f(x) = \mathbf{0}_{\mathbb{R}^m}\big\}. \tag{2.8.3}$$

***Preuve.*** On considère l'application $F : U \to \mathbb{R}^{n+m}$ donnée par $F(v', v'') = (f(v', v''), v'')$ pour tout $v = (v', v'') \in U$ avec $v' \in \mathbb{R}^m$ et $v'' \in \mathbb{R}^n$. Comme $f$ est de classe $\mathrm{C}^p$, alors $F$ est aussi de classe $\mathrm{C}^p$, et $F(v_0) = (\mathbf{0}_{\mathbb{R}^m}, v_0'')$. En outre, si l'on écrit la matrice jacobienne de $f$ en $v_0$

$$J_f(v_0) = \left( \begin{array}{c:c} J_f'(v_0) & J_f''(v_0) \end{array} \right) \in \mathrm{M}_{m\times(m+n)}(\mathbb{R}),$$

où $J_f''(v_0) \in \mathrm{M}_{m\times n}(\mathbb{R})$ est la sous-matrice de $J_f(v_0)$ formée des dernières $n$ colonnes, alors on voit bien que

$$J_F(v_0) = \left( \begin{array}{c:c} J_f'(v_0) & J_f''(v_0) \\ \hdashline \mathbf{0}_{\mathrm{M}_{n\times m}(\mathbb{R})} & \mathrm{I}_n \end{array} \right) \in \mathrm{M}_{(m+n)\times(m+n)}(\mathbb{R}),$$

où $\mathbf{0}_{\mathrm{M}_{n\times m}(\mathbb{R})}$ est la matrice nulle de $\mathrm{M}_{n\times m}(\mathbb{R})$ et $\mathrm{I}_n$ est la matrice identité d'ordre $n$. Comme $J_f'(v_0)$ est inversible, on conclut que $J_F(v_0)$ est aussi inversible.

Le Corollaire 2.8.7 nous dit alors que $F$ est un $\mathrm{C}^p$-difféomorphisme local en $v_0$. Alors, il existe un ouvert $U'' \subseteq U$ tel que $v_0 \in U''$, $F(U'') \subseteq \mathbb{R}^{m+n}$ est un ouvert et $F|_{U''} : U'' \to F(U'')$ est un $\mathrm{C}^p$-difféomorphisme. Comme $F(v_0) = (\mathbf{0}_{\mathbb{R}^m}, v_0'') \in F(U'')$ il existe des ouverts $V \subseteq \mathbb{R}^m$ et $W \subseteq \mathbb{R}^n$ avec $(\mathbf{0}_{\mathbb{R}^m}, v_0'') \in V \times W \subseteq F(U'')$. Soit $U' = F^{-1}(V \times W)$. On voit bien que $U'$ est un ouvert inclus dans $U'' \subseteq U$, car $F$ est continue, et que $v_0 \in U'$. Alors, $F|_{U'} : U' \to V \times W$ est un $\mathrm{C}^p$-difféomorphisme.

Soit $G : V \times W \to U'$ la réciproque de $F|_{U'}$. Comme $F|_{U'} \circ G = \mathrm{id}_{V\times W}$, $f(G(y,z)) = y$ et $\pi'_{\mathbb{R}^n}(G(y,z)) = z$ pour tous $y \in V$ et $z \in W$, où $\pi'_{\mathbb{R}^n} : \mathbb{R}^{m+n} \to \mathbb{R}^m$ est donnée par la projection canonique $\pi'_{\mathbb{R}^n}(x_1, \dots, x_{m+n}) = (x_{m+1}, \dots, x_{m+n})$ pour tous $x_1, \dots, x_{m+n} \in \mathbb{R}$. En conséquence, on peut écrire $G(y,z) = (G_1(y,z), z)$ pour tous $y \in V$ et $z \in W$, avec $G_1 : V \times W \to \mathbb{R}^m$. On définit l'application $g : W \to \mathbb{R}^m$ par $g(z) = G_1(\mathbf{0}_{\mathbb{R}^m}, z) = \pi_{\mathbb{R}^m}(G(\mathbf{0}_{\mathbb{R}^m}, z))$ pour tout $z \in W$. Comme $G$ est de classe $\mathrm{C}^p$, $g$ l'est aussi, car $\pi_{\mathbb{R}^m}$ est une application linéaire et l'application $\iota_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^{m+n}$ qui associe $(\mathbf{0}_{\mathbb{R}^m}, z)$ à $z \in \mathbb{R}^n$ l'est aussi (voir Exemple 2.5.3). En plus, $g(v_0'') = G_1(\mathbf{0}_{\mathbb{R}^m}, v_0'') = v_0'$, vu que $F(v_0', v_0'') = (\mathbf{0}_{\mathbb{R}^m}, v_0'')$.

Soit $S = F|_{U'}^{-1}(\{\mathbf{0}_{\mathbb{R}^m}\} \times W)$. Noter que $v_0 \in S$, car $F(v_0) = (\mathbf{0}_{\mathbb{R}^m}, v_0'')$ et $\mathbf{0}_{\mathbb{R}^n} \in V$, et que

$$S = U' \cap \big\{x \in U : f(x) = \mathbf{0}_{\mathbb{R}^n}\big\}.$$

Comme $F|_{U'}$ est bijective, $F|_S : S \to \{\mathbf{0}_{\mathbb{R}^m}\} \times W$ aussi, avec réciproque $G|_{\{\mathbf{0}_{\mathbb{R}^m}\}\times W} : \{\mathbf{0}_{\mathbb{R}^m}\} \times W \to S$. En conséquence, $S = \{(g(w), w) : w \in W\}$, comme on voulait démontrer. Finalement, l'unicité de $g$ suit de l'unicité de la réciproque $G$ de $F|_{U'} : U' \to V \times W$, vu que (2.8.3) nous dit que $g$ est la composition de $\iota_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^{m+n}$, la réciproque de $F|_S : S \to \{\mathbf{0}_{\mathbb{R}^m}\} \times W$ et la projection $\pi_{\mathbb{R}^m} : \mathbb{R}^{m+n} \to \mathbb{R}^m$. □

**Remarque 2.8.11.** *Si l'on utilise la Remarque 2.8.9, le Théorème 2.8.10 reste vrai si l'on remplace l'hypothèse d'être de classe $\mathrm{C}^p$ par la simple différentiabilité de $f$ mais l'inversibilité de $J_f'(v_0)$ par l'inversibilité de $J_f'(v)$ pour tout $v$ dans un ouvert $V \subseteq U$ incluant $v_0$.*

**Remarque 2.8.12.** *L'identité $f(g(w), w) = \mathbf{0}_{\mathbb{R}^m}$ pour tout $w \in W$ dans le Théorème 2.8.10 nous dit que la différentielle de l'application $w \mapsto f(g(w), w)$ pour $w \in W$ est nulle. En conséquence, le Théorème 2.1.17 implique que*

$$J_f'\big(g(w), w\big) \cdot J_g(w) + J_f''\big(g(w), w\big) = \mathbf{0}_{\mathrm{M}_{m\times n}(\mathbb{R})} \tag{2.8.4}$$

*pour tout $w \in W$, où l'on utilise la terminologie expliquée dans la preuve du Théorème 2.8.10. En particulier,*

$$J_g(v_0'') = -J_f'(v_0)^{-1} \cdot J_f''(v_0). \tag{2.8.5}$$

**Exemple 2.8.13.** *L'hypothèse que $f$ soit au moins de classe $\mathrm{C}^1$ dans le Théorème 2.8.10 est nécessaire. En effet, si l'on suppose que $f$ est seulement différentiable, alors le résultat n'est plus vrai en général. Pour le montrer, soit $f : \mathbb{R} \to \mathbb{R}$ la fonction de l'Exemple 2.8.8. On définit l'application $F : \mathbb{R}^2 \to \mathbb{R}$ donnée par $F(x,y) = f(x) - y$ pour tout $(x,y) \in \mathbb{R}^2$. Alors, c'est facile à vérifier que $F$ est différentiable en tout point de $\mathbb{R}^2$ et $\nabla F(x,y) = (f'(x), -1)$. En particulier, $\nabla F(0,0) = (1,-1)$. S'il existe $g : J \to \mathbb{R}$ une application différentiable telle que $g(0) = 0$ et $F(g(y), y) = 0$ pour tout $y \in J$, où $J$ est un intervalle ouvert incluant $0$, alors* (2.8.4) *nous dit que*

$$0 = \frac{d}{dy}\Big(F\big(g(y), y\big)\Big) = \frac{\partial F}{\partial x}\big(g(y), y\big)g'(y) + \frac{\partial F}{\partial y}\big(g(y), y\big) = f'\big(g(y)\big)g'(y) - 1, \tag{2.8.6}$$

*pour tout $y \in J$. On voit bien que $g(y) = 0$ pour tout $y \in J$ est impossible, car dans ce cas $g'(y) = 0$ pour tout $y \in J$, ce qui implique que l'identité (2.8.6) n'est pas vérifiée. Alors, il existe $y_0 \in J$ tel que $g(y_0) \neq 0$. Comme $g$ est différentiable, elle est continue, et le théorème des valeurs intermédiaires nous dit que l'intervalle $I = \{tg(y_0) : t \in [0,1]\}$ est inclus dans l'image de $g$. On notera $\bar{x}_n = \operatorname{sgn}(g(y_0))x_n$, où $(x_n)_{n\in\mathbb{N}}$ est la suite de l'Exemple 2.8.8. Alors, il existe $n_0 \in \mathbb{N}$ tel que $\bar{x}_n \in I$ pour tout $n > n_0$, i.e. il existe $y_n \in J$ tel que $g(y_n) = \bar{x}_n$ pour tout $n > n_0$. Si l'on remplace $y$ par $y_n$ dans (2.8.6) on trouve un absurde, car $f'(\bar{x}_n) = 0$ pour tout $n \in \mathbb{N}$.*

On va finalement démontrer le Lemme 2.7.15.

***Preuve du Lemme 2.7.15.*** On voit bien que $T_vS \subseteq \operatorname{Ker}(Dg(v))$. En effet, étant donné $w \in T_vS$, il existe $\alpha : J \to S$ est une application de classe $\mathrm{C}^1$ avec $J$ un intervalle ouvert incluant $0$, $\alpha(0) = v$ et $\alpha'(0) = w$. Par définition de $S$, $g \circ \alpha = \mathbf{0}_{\mathbb{R}^m}$, ce qui implique que

$$\mathbf{0}_{\mathbb{R}^m} = (g \circ \alpha)'(0) = D(g \circ \alpha)(0)(1) = Dg(\alpha(0))\big(D\alpha(0)(1)\big) = Dg(v)(w),$$

*i.e.* $w \in \operatorname{Ker}(Dg(v))$.

On va montrer que $\operatorname{Ker}(Dg(v_0)) \subseteq T_{v_0}S$ pour tout $v_0 \in S$. Or, comme $Dg(v_0) : \mathbb{R}^n \to \mathbb{R}^m$ est surjectif, il existe un ensemble $\{i_1, \dots, i_m\} \subseteq [\![1, n]\!]$ de cardinalité $m$ tel que la sous-matrice carrée d'ordre $m$ de $J_g(v_0)$ formée des colonnes $i_1, \dots, i_m$ est inversible. Sans perte de généralité, on va supposer que $\{i_1, \dots, i_m\} = [\![1, m]\!]$. On écrira $v = (v', v'') \in \mathbb{R}^n$ avec $v' \in \mathbb{R}^m$ et $v'' \in \mathbb{R}^{n-m}$. Alors, le Théorème 2.8.10 nous dit qu'il existe des ouverts $U' \subseteq U$ ouvert $W \subseteq \mathbb{R}^{n-m}$ tel que $v_0'' \in W$ et une application $h : W \to \mathbb{R}^m$ de classe $\mathrm{C}^1$ telle que $h(v_0'') = v_0'$ et

$$\big\{(h(z), z) : z \in W\big\} = U' \cap S.$$

Soit $w = (w', w'') \in \operatorname{Ker}(Dg(v_0))$. Si $w = \mathbf{0}_{\mathbb{R}^n}$, alors $w \in T_{v_0}S$. On suppose $w \neq \mathbf{0}_{\mathbb{R}^n}$. Or, comme $w = (w', w'') \in \operatorname{Ker}(Dg(v_0))$,

$$J_g'(v_0) \cdot w'^t + J_g''(v_0) \cdot w''^t = \mathbf{0}_{\mathbb{R}^m},$$

où $J_g'(v_0) \in \mathrm{M}_{m\times m}(\mathbb{R})$ est la sous-matrice de $J_g(v_0)$ formée des premières $m$ colonnes et $J_g''(v_0) \in \mathrm{M}_{m\times(n-m)}(\mathbb{R})$ est la sous-matrice de $J_g(v_0)$ formée des dernières $n - m$ colonnes. Comme $J_g'(v_0)$ est inversible, on a alors $w'^t = -J_g'(v_0)^{-1} \cdot J_g''(v_0) \cdot w''^t$. L'identité (2.8.5) avec $g$ et $h$ au lieu de $f$ et $g$, resp., nous dit que $w'^t = J_h'(v_0'') \cdot w''^t$. Comme $w \neq \mathbf{0}_{\mathbb{R}^n}$, on conclut que $w'' \neq \mathbf{0}_{\mathbb{R}^{n-m}}$. Soit $r > 0$ tel que $B_{||\ ||}(v_0'', r||w''||) \subseteq W$, $J = ]-r, r[$ et soit $\alpha : J \to \mathbb{R}^n$ l'application $\alpha(t) = (h(v_0'' + tw''), v_0'' + tw'')$ pour tout $t \in J$. Alors, $\alpha$ est de classe $\mathrm{C}^1$, $\alpha(t) \in S$ pour tout $t \in J$ et

$$\alpha'(0) = \big(Dh(v_0'')(w''), w''\big) = (w', w'') = w,$$

ce qui implique que $w \in T_{v_0}S$, comme on voulait démontrer. □

**Exemple 2.8.14.** *Soit $f : \mathbb{R}^3 \to \mathbb{R}$ l'application $f(x_1, x_{2,3}) = 4x_1^2 - \cos(x_2) + x_3 + e^{x_3}$ pour tous $x_1, x_2, x_3 \in \mathbb{R}$. On voit bien que $f$ est de classe $\mathrm{C}^\infty$, car elle s'obtient à partir de considérer des sommes et compositions de fonctions de classe $\mathrm{C}^\infty$. Comme*

$$\frac{\partial f}{\partial x_3}(0,0,0) = 1 + e^0 = 2 \neq 0,$$

*le Théorème 2.8.10 nous dit qu'il existe des ouverts $U' \subseteq \mathbb{R}^3$ et $W \subseteq \mathbb{R}^2$ tels que $(0,0) \in W$, et une application $g : W \to \mathbb{R}$ de classe $\mathrm{C}^\infty$ telle que $g(0,0) = 0$, et*

$$\Big\{\big(x_1, x_2, g(x_1, x_2)\big) : w \in W\Big\} = \big\{x \in U' : f(x_1, x_2, x_3) = 0\big\}.$$

*En plus, l'identité (2.8.4) nous dit que*

$$4x_1 + \Big(1 + e^{g(x_1,x_2)}\Big)\frac{\partial g}{\partial x_1}(x_1, x_2) = 0 \text{ et } \sin(x_2) + \Big(1 + e^{g(x_1,x_2)}\Big)\frac{\partial g}{\partial x_2}(x_1, x_2) = 0,$$

*pour tout $(x_1, x_2) \in W$, ce qui nous dit que*

$$\frac{\partial g}{\partial x_1}(x_1, x_2) = -\frac{4x_1}{1 + e^{g(x_1,x_2)}} \text{ et } \frac{\partial g}{\partial x_2}(x_1, x_2) = -\frac{\sin(x_2)}{1 + e^{g(x_1,x_2)}}, \tag{2.8.7}$$

*pour tout* $(x_1, x_2) \in W$*. Cela implique que* $\nabla g(0,0) = (0,0)$*, i.e.* $(0,0)$ *est un point critique de* $g$*. En outre, si l'on dérive les expressions* (2.8.7)*, on trouve*

$$\begin{aligned}
\frac{\partial^2 g}{\partial x_1^2}(x_1,x_2) &= -\frac{4(1+e^{g(x_1,x_2)}) - 4x_1 e^{g(x_1,x_2)} \frac{\partial g}{\partial x_1}(x_1,x_2)}{(1+e^{g(x_1,x_2)})^2} = -4\frac{(1+e^{g(x_1,x_2)})^2 + 4x_1^2 e^{g(x_1,x_2)}}{(1+e^{g(x_1,x_2)})^3}, \\
\frac{\partial^2 g}{\partial x_1 \partial x_2}(x_1,x_2) &= \frac{4x_1}{(1+e^{g(x_1,x_2)})^2}\frac{\partial g}{\partial x_2}(x_1,x_2) = -\frac{4x_1 \sin(x_2)}{(1+e^{g(x_1,x_2)})^3} \\
\frac{\partial^2 g}{\partial x_2^2}(x_1,x_2) &= -\frac{\cos(x_2)(1+e^{g(x_1,x_2)}) - \sin(x_2) e^{g(x_1,x_2)} \frac{\partial g}{\partial x_2}(x_1,x_2)}{(1+e^{g(x_1,x_2)})^2} \\
&= -\frac{\cos(x_2)(1+e^{g(x_1,x_2)})^2 + \sin(x_2)^2 e^{g(x_1,x_2)}}{(1+e^{g(x_1,x_2)})^3},
\end{aligned}$$

*pour tout* $(x_1, x_2) \in W$*. En conséquence,*

$$H_g(0,0) = \begin{pmatrix} -4 & 0 \\ 0 & -1 \end{pmatrix},$$

*ce qui implique que* $(0,0)$ *est un maximum local strict de* $g$*, d'après le Théorème 2.7.7.*

# Chapitre 3

# Équations différentielles ordinaires

Le but du troisième chapitre c'est d'étudier les équations différentielles ordinaires. On continue avec les mêmes conventions que celles du premier paragraphe du chapitre 2

## 3.1 Définitions préliminaires

**Définition 3.1.1.** *Soient $d \in \mathbb{N}$, $n, m \in \mathbb{Z}_{>0}$, et $U \subseteq \mathbb{R}^n$ ouvert. Une* ***équation différentielle ordinaire (EDO) de taille $m$ et d'ordre $d$*** *sur $U$ est une application $F : J \times U \times \mathbb{R}^{n.d} \to \mathbb{R}^m$. Étant donné une équation différentielle ordinaire de taille $m$ et d'ordre $d$ définie par l'application $F$, une* ***condition initiale à temps $t_0 \in J$*** *est un point $(v_0^0, \dots, v_0^{d-1})$, où $v_0^0 \in U$ et $v_0^i \in \mathbb{R}^n$ pour tout $i \in [\![1, d-1]\!]$. On dira qu'une condition initiale $(v_0^0, \dots, v_0^{d-1}) \in U \times \mathbb{R}^{n.(d-1)}$ à temps $t_0 \in J$ est* ***compatible*** *avec $F$ s'il existe $v_0^d \in \mathbb{R}^n$ tel que $F(t_0, v_0^0, \dots, v_0^{d-1}, v_0^d) = \mathbf{0}_{\mathbb{R}^m}$.*

**Définition 3.1.2.** *Étant donné une équation différentielle ordinaire de taille $m$ et d'ordre $d$ sur $U$ définie par une application $F : J \times U \times \mathbb{R}^{n.d} \to \mathbb{R}^m$ et une condition initiale $(v_0^0, \dots, v_0^{d-1}) \in U \times \mathbb{R}^{n.(d-1)}$ à temps $t_0 \in J$ compatible avec $F$, une* ***solution*** *de $F$ avec condition initiale $(v_0^i)_{i \in [\![0,d]\!]}$ à temps $t_0$ est une application $d$-fois différentiable $\alpha : J' \to U$, où $J' \subseteq J$ est un intervalle ouvert avec $t_0 \in J'$, qui satisfait que*

$$F\big(t, \alpha(t), \alpha'(t), \dots, \alpha^{(d)}(t)\big) = \mathbf{0}_{\mathbb{R}^m},$$

*pour tout $t \in J'$ et $\alpha^{(i)}(t_0) = v_0^i$ pour tout $i \in [\![0, d-1]\!]$, où $\alpha^{(i)}(t) \in \mathbb{R}^n$ dénote la $i$-ème dérivée de $\alpha$ en $t$ si $i \in \mathbb{Z}_{>0}$ et $\alpha^{(0)} = \alpha$.*

Même si cette définition d'équation différentielle semble raisonnable, elle possède quelques problèmes, comme l'exemple suivant le montre.

**Exemple 3.1.3.** *Soit $F : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ l'application donnée par $F(t, x, u) = xu + t$ pour tous $t, x, u \in \mathbb{R}$, qui définie une équation différentielle de taille 1 et d'ordre 1 sur $\mathbb{R}$. On voit bien que la condition initiale $0 \in \mathbb{R}$ à temps $t = 0$ est compatible avec $F$. Alors, une solution de l'équation différentielle ordinaire de taille 1 et ordre 1 sur $\mathbb{R}$ définie par $F$ avec condition initiale $0 \in \mathbb{R}$ à temps $t = 0$ est une application $\alpha : J' \to \mathbb{R}$ telle que $0 \in J'$, $\alpha(0) = 0$ et*

$$t + \alpha(t)\alpha'(t) = 0$$

*pour tout $t \in J'$. En conséquence, la fonction $t \mapsto t^2 + \alpha(t)^2$ pour $t \in J'$ est constante, vu que sa dérivée en $t$ coïncide avec $2(t + \alpha(t)\alpha'(t))$. Or, la condition initiale nous dit alors que $t^2 + \alpha(t)^2 = 0^2 + \alpha(0)^2 = 0$ pour tout $t \in J'$, ce qui implique que $t = 0$ pour tout $t \in J'$, ce qui est absurde car $J'$ est un intervalle ouvert non vide. En conséquence, cette équation différentielle ne possède pas de solution avec condition initiale $0 \in \mathbb{R}$ à temps $t = 0$.*

Même si une équation différentielle possède une solution pour une certaine condition initiale et à un temps fixe, on peut avoir plusieurs solutions, comme l'exemple suivant le montre.

**Exemple 3.1.4.** *Soit $F : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ l'application donnée par $F(t, x, u) = u - 3x^{2/3}$, pour tous $t, x, u \in \mathbb{R}$, qui définie une équation différentielle de taille 1 et d'ordre 1 sur $\mathbb{R}$. On voit bien que toute condition initiale $x_0$ à tout temps $t_0$ est compatible. En particulier, la condition initiale $0 \in \mathbb{R}$ à temps $t = 0$ est compatible. Or, une* *solution de l'équation différentielle ordinaire définie par $F$ sur $\mathbb{R}$ avec condition initiale $0 \in \mathbb{R}$ à temps $t = 0$ est une application $\alpha : J' \to \mathbb{R}$ telle que $0 \in J'$, $\alpha(0) = 0$ et*

$$\alpha'(t) = 3\alpha(t)^{2/3}$$

*pour tout $t \in J'$. Or, étant donné, $C \in \mathbb{R}_{>0}$, c'est facile à voir que la fonction $\alpha : \mathbb{R} \to \mathbb{R}$ donnée par*

$$\alpha(t) = \begin{cases} 0, & \text{si } t < C, \\ (t-C)^3, & \text{si } t \geq C, \end{cases}$$

*est de classe $\mathrm{C}^1$ et une solution de l'équation avec condition initiale $0$ à temps $0$, ce qui implique que cette équation différentielle ordinaire ne possède pas d'unique solution pour la condition initiale indiquée.*

## 3.2 Courbes intégrales

Pour éviter les problèmes de non existence ou non unicité de solutions dans les exemples de la section précédente, on va se restreindre à des équations différentielles particulières.

**Définition 3.2.1.** *Soient $n, m \in \mathbb{Z}_{>0}$. Un **champ vectoriel (indépendant du temps)** est une application $f : U \to \mathbb{R}^n$, où $U \subseteq \mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$. Un **champ vectoriel dépendant du temps** est une application $\bar{f} : J \times U \to \mathbb{R}^n$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. On va regarder tout champ vectoriel indépendant du temps $f : U \to \mathbb{R}^n$ comme un champ vectoriel dépendant du temps $\bar{f} : J \times U \to \mathbb{R}^n$, avec $\bar{f}(t, v) = f(v)$ pour tout $t \in J$ et $v \in U$.*

**Définition 3.2.2.** *Étant donné un champ vectoriel dépendant du temps $\bar{f} : J \times U \to \mathbb{R}^n$, $t_0 \in J$ et $v_0 \in U$, une **courbe intégrale de $\bar{f}$ avec condition initiale $v_0$ à temps $t_0$** est une application dérivable $\alpha : J' \to U$, où $J' \subseteq J$ est un intervalle ouvert incluant $t_0$, telle que*

$$\alpha'(t) = \bar{f}\big(t, \alpha(t)\big) \tag{3.2.1}$$

*pour tout $t \in J'$ et $\alpha(t_0) = v_0$.*

Le résultat suivant nous dit que les notions de champs vectoriels dépendant du temps et indépendants du temps ne sont pas vraiment très éloignées, en ce qui respecte leurs courbes intégrales.

**Lemme 3.2.3.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps (resp. de classe $\mathrm{C}^p$), où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. On définit $\hat{f} : J \times U \to \mathbb{R}^{n+1}$ via $\hat{f}(t, v) = (1, \bar{f}(t, v))$ pour tout $t \in J$ et $v \in U$. Alors, $\hat{f}$ est un champ vectoriel indépendant du temps (resp. de classe $\mathrm{C}^p$) défini sur l'ouvert $J \times U$ de $\mathbb{R}^{n+1}$.*

*En plus, étant donné $t_0 \in J$ et $v_0 \in U$, si $\alpha : J' \to U$ est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$, l'application $\hat{\alpha} : J' \to J \times U$ donnée par $\hat{\alpha}(t) = (t, \alpha(t))$ pour tout $t \in J'$ est une courbe intégrale de $\hat{f}$ avec condition initiale $(t_0, v_0)$ à temps $t_0$, et, réciproquement, toute courbe intégrale $\hat{\alpha} : J' \to J \times U$ de $\hat{f}$ avec condition initiale $(t_0, v_0)$ à temps $t_0$ satisfait que $\hat{\alpha}(t) = (t, \alpha(t))$ pour tout $t \in J'$, où $\alpha : J' \to U$ est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$.*

***Preuve.*** La première partie de l'énoncé est immédiate, vu que $\hat{f}$ est de classe $\mathrm{C}^p$ si $\bar{f}$ l'est. On démontrera la deuxième partie. Soit $\alpha : J' \to U$ une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$. On définit $\hat{\alpha} : J' \to J \times U$ via $\hat{\alpha}(t) = (t, \alpha(t))$ pour tout $t \in J'$. Alors,

$$\hat{\alpha}'(t) = \big(1, \alpha'(t)\big) = \Big(1, f\big(t, \alpha(t)\big)\Big) = \hat{f}\big(\hat{\alpha}(t)\big)$$

pour tout $t \in J'$ et $\hat{\alpha}(t_0) = (t_0, \alpha(t_0)) = (t_0, v_0)$, ce qui implique que $\hat{\alpha}$ est une courbe intégrale de $\hat{f}$ avec condition initiale $(t_0, v_0)$ à temps $t_0$. Réciproquement, soit $\hat{\alpha} : J' \to J \times U$ est une courbe intégrale de $\hat{f}$ avec condition initiale $(t_0, v_0)$ à temps $t_0$. On écrit $\hat{\alpha}(t) = (\alpha_1(t), \alpha_2(t))$ pour $t \in J'$, avec $\alpha_1 : J' \to J$ et $\alpha_2 : J' \to U$. Alors,

$$\big(\alpha_1'(t), \alpha_2'(t)\big) = \hat{\alpha}'(t) = \hat{f}\big(\hat{\alpha}(t)\big) = \Big(1, f\big(\alpha_1(t), \alpha_2(t)\big)\Big) \tag{3.2.2}$$

pour tout $t \in J'$ et $(\alpha_1(t_0), \alpha_2(t_0)) = \hat{\alpha}(t_0) = (t_0, v_0)$. Si l'on regarde la première coordonné de l'identité (3.2.2) ainsi que de la condition initiale précédente, on conclut que $\alpha_1'(t) = 1$ et $\alpha_1(t_0) = t_0$, ce qui implique que $\alpha_1(t) = t$ pour tout $t \in J'$. La deuxième coordonné de (3.2.2) ainsi que de la condition initiale nous donnent précisément que $\alpha_2 : J' \to U$ est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$. □

On a la suivante description équivalente de la condition de courbe intégrale.

**Lemme 3.2.4.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Étant donné $t_0 \in J$ et $v_0 \in U$, alors une application $\alpha : J' \to U$, où $J' \subseteq J$ est un intervalle ouvert incluant $t_0$, est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$ si et seulement si*

$$\alpha(t) = v_0 + \int_{t_0}^{t} \bar{f}\big(s, \alpha(s)\big) ds \tag{3.2.3}$$

*pour tout $t \in J'$.*

***Preuve.*** On suppose que $\alpha$ satisfait (3.2.3). Cette identité nous dit alors que $\alpha(t_0) = v_0$. En outre, en dérivant (3.2.3) on trouve précisément (3.2.1), ce qui nous dit que $\alpha$ est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$. Réciproquement, si $\alpha$ est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$, alors (2.6.1) nous dit que

$$\alpha(t) = \alpha(t_0) + \int_{t_0}^{t} \alpha'(s) ds = v_0 + \int_{t_0}^{t} \bar{f}\big(s, \alpha(s)\big) ds$$

pour tout $t \in J'$, comme on voulait démontrer. □

## 3.3 Réduction au cas des champs vectoriels (optionnel)

**Définition 3.3.1.** *Soit $F : J \times U \times \mathbb{R}^{n.d} \to \mathbb{R}^m$ une application définissant une équation différentielle ordinaire de taille $m < n$ et d'ordre $d$ sur $U$ et soit $(v_0^0, \dots, v_0^{d-1}) \in U \times \mathbb{R}^{n.(d-1)}$ un condition initiale à temps $t_0 \in J$ compatible avec $F$. Soit $v_0^d \in \mathbb{R}^n$ tel que $F(t_0, v_0^0, \dots, v_0^d) = \mathbf{0}_{\mathbb{R}^m}$. On suppose en plus que $F$ est de classe $\mathrm{C}^1$. On dira que $F$ est **localement admissible** en $(t_0, v_0^0, \dots, v_0^d)$ si la sous-matrice $J_F^{(d)}(t_0, v_0^0, \dots, v_0^d)$ formée des dernières $n$ colonnes de $J_F(t_0, v_0^0, \dots, v_0^d)$ est surjective.*

**Définition 3.3.2.** *On suppose les hypothèses de la Définition 3.3.1. D'après le Théorème 2.8.10, si $F$ est localement admissible en $(t_0, v_0^0, \dots, v_0^d)$, il existe une unique application $\hat{F} : J' \times U' \times W \to \mathbb{R}^n$ avec $J' \subseteq J$, $U' \subseteq U$, $W \subseteq \mathbb{R}^{n.(d-1)}$ et $W' \subseteq \mathbb{R}^n$ ouverts tels que $(t_0, v_0^0, \dots, v_0^d) \in J' \times U' \times W \times W'$ et*

$$\begin{aligned}\Big\{\big(t, v^0, \dots, v^{d-1}, \hat{F}(t, v^0, \dots, v^{d-1})\big) : (t, v^0, \dots, v^{d-1}) \in J' \times U' \times W\Big\} \\ = \Big\{F(t, v^0, \dots, v^d) = \mathbf{0}_{\mathbb{R}^m} : (t, v^0, \dots, v^{d-1}) \in J' \times U' \times W \times W'\Big\}.\end{aligned}$$

*Soit $\bar{f} : J' \times U' \times W \to \mathbb{R}^{n.d}$ l'application définie par $\bar{f}(t, v^0, \dots, v^{d-1}) = (v^1, \dots, v^{d-1}, \hat{F}(t, v^0, \dots, v^{d-1}))$ pour tout $(t, v^0, \dots, v^{d-1}) \in J' \times U' \times W$. On appelle $\bar{f}$ le **champ vectoriel dépendant du temps** associé à $F$ et à $(t_0, v_0^0, \dots, v_0^d)$.*

Le résultat suivant est maintenant immédiat. Il nous dit essentiellement que, avec certaines hypothèses de régularité, on peut se restreindre à considérer des équations différentielles ordinaires de la forme (3.2.1).

**Lemme 3.3.3.** *Soit $F : J \times U \times \mathbb{R}^{n.d} \to \mathbb{R}^m$ une application de classe $\mathrm{C}^1$ définissant une équation différentielle ordinaire de taille $m < n$ et d'ordre $d$ sur $U$ et soit $(v_0^0, \dots, v_0^{d-1}) \in U \times \mathbb{R}^{n.(d-1)}$ une condition initiale à temps $t_0 \in J$ compatible avec $F$. Soit $v_0^d \in \mathbb{R}^n$ tel que $F(t_0, v_0^0, \dots, v_0^d) = \mathbf{0}_{\mathbb{R}^m}$. On suppose que $F$ est localement admissible en $(t_0, v_0^0, \dots, v_0^d)$, et soit $\bar{f}$ le champ vectoriel dépendant du temps associé à $F$ et à $(t_0, v_0^0, \dots, v_0^d)$.*

*Alors, si $\alpha : J' \to U$ est une solution de l'équation différentielle ordinaire associée à $F$ avec condition initiale $(v_0^0, \dots, v_0^{d-1})$ à temps $t_0 \in J$ définie sur un sous-intervalle de $J'$ et qui satisfait $\alpha^{(d)}(t_0) = v_0^d$, l'application*

*$\tilde{\alpha} : J' \to U \times \mathbb{R}^{n.(d-1)}$ donnée par $\tilde{\alpha}(t) = (\alpha(t), \alpha'(t), \ldots, \alpha^{(d-1)}(t))$ pour $t \in J'$ est une courbe intégrale de $\bar{f}$ de condition initiale $(v_0^0, \ldots, v_0^{d-1})$ à temps $t_0 \in J$. De façon réciproque, si $\tilde{\alpha} : J' \to U \times \mathbb{R}^{n.(d-1)}$ est une courbe intégrale de $\bar{f}$ de condition initiale $(v_0^0, \ldots, v_0^{d-1})$ à temps $t_0 \in J$, et on écrit $\tilde{\alpha}(t) = (\alpha(t), \alpha^1(t), \ldots, \alpha^{d-1}(t))$ pour $t \in J'$, alors $\alpha^{(i)} = \alpha^i$ pour tout $i \in [\![1, d-1]\!]$ et $\alpha$ est une solution de l'équation différentielle ordinaire associée à $F$ avec condition initiale $(v_0^0, \ldots, v_0^{d-1})$ à temps $t_0 \in J$.*

**Exemple 3.3.4.** *Soit $g : J \times U \times \mathbb{R}^{d-1} \to \mathbb{R}$ une application de classe $\mathrm{C}^1$, où $J, U \subseteq \mathbb{R}$ sont des intervalles ouverts non vides, et soit $F : J \times U \times \mathbb{R}^d \to \mathbb{R}$ l'application*

$$F(t, x^0, \ldots, x^d) = x^d - g(t, x^0, \ldots, x^{d-1})$$

*pour tout $t \in J$, $x^0 \in U$ et $x^i \in \mathbb{R}$ pour $i \in [\![1, d]\!]$. On considère le champ vectoriel dépendant du temps $\bar{f} : J \times (U \times \mathbb{R}^{d-1}) \to \mathbb{R}^d$ donné par*

$$\bar{f}(t, x^0, \ldots, x^{d-1}) = \big(x^0, \ldots, x^{d-1}, g(t, x^0, \ldots, x^{d-1})\big)$$

*pour tout $t \in J$, $x^0 \in U$ et $x^i \in \mathbb{R}$ pour $i \in [\![1, d-1]\!]$. On fixe $t_0 \in J$, $x_0 \in U$ et $x_0^i \in \mathbb{R}$ pour $i \in [\![1, d-1]\!]$. D'après le Lemme 3.3.3, étant donné une application $x : J' \to U$, où $J' \subseteq J$ est un intervalle ouvert incluant $t_0$, si $x$ est un solution de l'équation différentielle déterminée par $F$ avec condition initiale $(x_0, x_0^1, \ldots, x_0^{d-1})$ à temps $t_0$ alors l'application $\alpha : J' \to U \times \mathbb{R}^{d-1}$ donnée par $\alpha(t) = (x(t), x'(t), \ldots, x^{(d-1)}(t))$ pour $t \in J'$ est une courbe intégrale de $\bar{f}$ avec condition initiale $(x_0, x_0^1, \ldots, x_0^{d-1})$ à temps $t_0$, et, de façon réciproque, toute courbe intégrale de $\bar{f}$ avec condition initiale $(x_0, x_0^1, \ldots, x_0^{d-1})$ à temps $t_0$ est de la forme précédente.*

## 3.4 Existence et unicité du flot local

On peut regrouper des courbes intégrales dans un objet commun de la façon suivante.

**Définition 3.4.1.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Étant donné $t_0 \in J$ et $v_0 \in U$, un* ***flot local*** *de $\bar{f}$ en $(t_0, v_0)$ est une application $\hat{\alpha} : J' \times U' \to U$ telle que $t_0 \in J' \subseteq J$, $v_0 \in U' \subseteq U$, $J'$ est un intervalle ouvert, $U'$ est un ouvert, et pour tout $v \in U'$ l'application $J' \to \mathbb{R}^n$ donnée par $t \mapsto \hat{\alpha}(t, v)$ est une courbe intégrale avec condition initiale $v$ à temps $t_0$.*

Étant donné un flot local $\hat{\alpha} : J' \times U' \to U$ de $\bar{f}$ en $(t_0, v_0)$ et $v \in U'$, on écrira $\hat{\alpha}_v : J' \to U$ l'application donnée par $\hat{\alpha}_v(t) = \hat{\alpha}(t, v)$ pour tout $t \in J'$.

Soit $J \subseteq \mathbb{R}$ un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ un ouvert non vide de $\mathbb{R}^n$. Soit $K > 0$. On dira qu'un champ vectoriel dépendant du temps $\bar{f} : J \times U \to \mathbb{R}^n$ est **$K$-lipschitzien sur $U$ uniformément pour $J$** si

$$||\bar{f}(t, v) - \bar{f}(t, u)|| \leq K||v - u||$$

pour tous $v, u \in U'$ et $t \in I$. De façon plus générale, on dira qu'un champ vectoriel dépendant du temps $\bar{f} : J \times U \to \mathbb{R}^n$ est **localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$** si, étant donné $v_0 \in U$ il existe un ouvert $U' \subseteq U$ incluant $v_0$ tel que, étant donné un intervalle fermé et borné $I \subseteq J$, il existe $K > 0$ qui satisfait que

$$||\bar{f}(t, v) - \bar{f}(t, u)|| \leq K||v - u||$$

pour tous $v, u \in U'$ et $t \in I$.

**Remarque 3.4.2.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps de classe $\mathrm{C}^1$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Alors, la Proposition 2.3.6 nous dit que $\bar{f}$ est localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$.*

Le but de cette section est de démontrer le résultat suivant.

**Proposition 3.4.3** (Existence et unicité du flot local)**.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Étant donné $t_0 \in J$ et $v_0 \in U$, soient $a \in ]0, 1[$ tel que $\bar{B}_{||\ ||}(v_0, 2a) \subseteq U$, et soit $I \subseteq J$ un intervalle fermé et borné dont sont intérieur inclut*

$t_0$. *Comme $f$ est localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, on suppose que $\bar{B}_{||\ ||}(v_0, 2a)$ satisfait que*

$$||\bar{f}(t,v) - \bar{f}(t,u)|| \leq K||v-u||$$

*pour tous $v, u \in \bar{B}_{||\ ||}(v_0, 2a)$ et $t \in I$, où $K > 0$. Comme $f$ est continu, il existe $C > 0$ tel que $||\bar{f}(t,v)|| \leq C$ pour tous $t \in I$ et $v \in \bar{B}_{||\ ||}(v_0, 2a)$. Soit $b > 0$ tel que $t_0 + b, t_0 - b \in I$ et $b < \min(1/K, a/C)$.*

*Alors, il existe un unique flot local $\hat{\alpha} : ]t_0 - b, t_0 + b[ \times B_{||\ ||}(v_0, a) \to U$ de $\bar{f}$ en $(t_0, v_0)$. Si $\bar{f}$ est de classe $C^p$ avec $p \in \mathbb{Z}_{>0}$, alors les courbes intégrales $\hat{\alpha}_v$ sont de classe $C^p$.*

***Preuve.*** Étant donné $v \in \bar{B}_{||\ ||}(v_0, a)$, soit

$$\mathscr{F}_{t_0,v,b} = \big\{\alpha : [t_0 - b, t_0 + b] \to \bar{B}_{||\ ||}(v_0, 2a) \text{ continue telle que } \alpha(t_0) = v\big\} \subseteq C^0([t_0 - b, t_0 + b], \mathbb{R}^n),$$

où $C^0([t_0-b, t_0+b], \mathbb{R}^n)$ désigne l'espace vectoriel des applications continues de $[t_0-b, t_0+b]$ dans $\mathbb{R}^n$. On considère que $C^0([t_0 - b, t_0 + b], \mathbb{R}^n)$ est muni de la norme infini $||f|| = \sup\{||f(t)|| : t \in [t_0 - b, t_0 + b]\}$. Alors, $C^0([t_0 - b, t_0 + b], \mathbb{R}^n)$ avec la norme précédente est un espace vectoriel normé complet (voir Proposition 1.13.2), et $\mathscr{F}_{t_0,v,b}$ est une partie fermée de $C^0([t_0-b, t_0+b], \mathbb{R}^n)$. Soit $I_{t_0,v,b} : \mathscr{F}_{t_0,v,b} \to \mathscr{F}_{t_0,v,b}$ l'application donnée par

$$I_{t_0,v,b}(\alpha)(t) = v + \int_{t_0}^{t} \bar{f}\big(s, \alpha(s)\big) ds$$

pour tout $t \in [t_0 - b, t_0 + b]$. Cette application est bien définie. En effet, c'est clair que $I_{t_0,v,b}(\alpha)$ est une application continue sur $[t_0 - b, t_0 + b]$, et on voit bien que

$$\big\|I_{t_0,v,b}(\alpha)(t) - v\big\| = \bigg\| \int_{t_0}^{t} \bar{f}\big(s, \alpha(s)\big) ds \bigg\| \leq \int_{\min(t_0,t)}^{\max(t_0,t)} \big\|\bar{f}\big(s, \alpha(s)\big)\big\| ds \leq C2b < 2a$$

pour tout $t \in [t_0 - b, t_0 + b]$, *i.e.* $I_{t_0,v,b}(\alpha)(t) \in \bar{B}_{||\ ||}(v_0, 2a)$ pour tout $t \in [t_0 - b, t_0 + b]$. En conséquence, $I_{t_0,v,b}(\alpha) \in \mathscr{F}_{t_0,v,b}$ pour tout $\alpha \in \mathscr{F}_{t_0,v,b}$. En plus, comme

$$\begin{aligned} ||I_{t_0,v,b}(\alpha - \beta)|| &= \sup_{t \in [t_0-b, t_0+b]} \bigg|\bigg| \int_{t_0}^{t} \Big(\bar{f}\big(s, \alpha(s)\big) - \bar{f}\big(s, \beta(s)\big)\Big) ds \bigg|\bigg| \\ &\leq \sup_{t \in [t_0-b, t_0+b]} K \int_{\min(t_0,t)}^{\max(t_0,t)} ||\ \alpha(s) - \beta(s)|| ds \leq Kb \sup_{t \in [t_0-b, t_0+b]} ||\ \alpha(t) - \beta(t)||, \end{aligned}$$

$I_{t_0,v,b}$ est une application contractante avec constante $Kb < 1$. Le Théorème 2.8.1 nous dit que $I_{t_0,v,b}$ admet un unique point fixe $\alpha_v$. Le Lemme 3.2.4 nous dit que le point fixe $\alpha_v$ de $I_{t_0,v,b}$ est une courbe intégrale de $\bar{f}$ définie sur $]t_0-b, t_0+b[$ avec condition initiale $v$ à temps $t_0$. En conséquence, l'application $\hat{\alpha} : ]t_0-b, t_0+b[ \times B_{||\ ||}(v_0, a) \to U$ donnée par $\hat{\alpha}(t, v) = \alpha_v(t)$ pour tout $t \in ]t_0-b, t_0+b[$ et $v \in B_{||\ ||}(v_0, a)$ est un flot local de $\bar{f}$ en $(t_0, v_0)$ et il est unique. Le fait que $\hat{\alpha}_v$ est de classe $C^p$ si $\bar{f}$ l'est suit directement de l'identité (3.2.1) pour $\hat{\alpha}_v$. □

**Remarque 3.4.4.** *Dans la preuve précédente, on peut aussi considérer l'ensemble*

$$\mathscr{F}_{t_0,v,b',b''} = \big\{\alpha : [t_0 - b', t_0 + b''] \to \bar{B}_{||\ ||}(v_0, 2a) \textit{ continue telle que } \alpha(t_0) = v\big\} \subseteq C^0([t_0 - b, t_0 + b], \mathbb{R}^n),$$

*avec $b', b'' \in [0, b]$. C'est clair que $\mathscr{F}_{t_0,v,b',b''}$ est aussi fermé dans $C^0([t_0 - b, t_0 + b], \mathbb{R}^n)$. En plus, l'application $I_{t_0,v,b',b''} : \mathscr{F}_{t_0,v,b',b''} \to \mathscr{F}_{t_0,v,b',b''}$ donnée par*

$$I_{t_0,v,b',b''}(\alpha)(t) = v + \int_{t_0}^{t} \bar{f}\big(s, \alpha(s)\big) ds$$

*pour tout $t \in [t_0 - b', t_0 + b'']$ est aussi contractante avec constante $Kb < 1$. En conséquence, $I_{t_0,v,b',b''}$ possède aussi un seul point fixe $\alpha_{v,b',b''}$. Comme $I_{t_0,v,b}(\alpha)|_{[t_0-b',t_0+b'']} = I_{t_0,v,b',b''}(\alpha|_{[t_0-b',t_0+b'']})$ pour tout $\alpha \in \mathscr{F}_{t_0,v,b}$, on voit bien que $\alpha_{v,b',b''} = \alpha_v|_{[t_0-b',t_0+b'']}$, où $\alpha_v \in \mathscr{F}_{t_0,v,b}$ est le point fixe de $I_{t_0,v,b}$.*

**Exemple 3.4.5** (Solution générale d'une EDO indépendante du temps sur $\mathbb{R}$). *Soient $U \subseteq \mathbb{R}$ un intervalle ouvert non vide et $f : U \to \mathbb{R}$ une fonction de classe $\mathrm{C}^1$, que l'on regarde comme un champ indépendant du temps. Soient $t_0 \in \mathbb{R}$ et $x_0 \in U$. Dans ce cas on s'intéresse à trouver une courbe intégrale $\alpha : J \to U$ du champ $f$ avec condition initiale $x_0$ à temps $t_0$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert incluant $t_0$, i.e.*

$$\alpha'(t) = f\big(\alpha(t)\big)$$

*et $\alpha(t_0) = x_0$. Comme $f$ est de classe $\mathrm{C}^1$, la Proposition 3.4.3 nous dit qu'il existe une unique courbe intégrale dans un voisinage $J'$ de $t_0$. Si $f(x_0) = 0$, on voit bien que l'application $\alpha : J' \to U$ donnée par $\alpha(t) = x_0$ pour tout $t \in J'$ est la courbe intégrale de $f$ avec condition initiale $x_0$. Si $f(x_0) \neq 0$, et suppose que $\alpha : J' \to U$ est la courbe intégrale de $f$ avec condition initiale $x_0$, on peut supposer par continuité de $f$ et $\alpha$ que $f(\alpha(t)) \neq 0$ pour tout $t \in J'$. Alors,*

$$F\big(\alpha(t)\big)' = \frac{\alpha'(t)}{f\big(\alpha(t)\big)} = 1 = t'$$

*où*

$$F(x) = t_0 + \int_{x_0}^{x} \frac{dy}{f(y)},$$

*pour tout $y$ dans un intervalle ouvert $J''$ incluant $x_0$ tel que $f(y) \neq 0$ pour tout $y \in J''$. En conséquence, $F(\alpha(t)) = t + C$, avec $C \in \mathbb{R}$ constante. Comme $F(\alpha(t_0)) = F(x_0) = t_0$, on conclut que $C = 0$, ce qui implique que*

$$F\big(\alpha(t)\big) = t$$

*pour tout $t \in J'$. Comme $F'(x) = 1/f(x) \neq 0$ pour tout $x \in J''$, le Théorème 2.8.6 nous dit que $F$ est un $\mathrm{C}^1$-difféomorphisme local en $x_0$, ce qui nous dit que $\alpha(t) = F^{-1}(t)$ pour tout $t \in J'$, où $F^{-1}$ désigne l'inverse locale de $F$ en $x_0$.*

## 3.5 Courbes intégrales approximatives et continuité du flot local (optionnel)

**Définition 3.5.1.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Étant donné $t_0 \in J$ et $\epsilon \in \mathbb{R}_{\geq 0}$, une **courbe intégrale $\epsilon$-approximative** de $\bar{f}$ autour de $t_0$ est une application $\alpha : J' \to U$ différentiable, où $J' \subseteq J$ est un intervalle ouvert incluant $t_0$, telle que*

$$\big\|\alpha'(t) - \bar{f}\big(t, \alpha(t)\big)\big\| \leq \epsilon$$

*pour tout $t \in J'$.*

On commence avec ce résultat que l'on va utiliser dans la suite.

**Lemme 3.5.2.** *Soit $h : J \to \mathbb{R}_{\geq 0}$ une application continue telle que $h(t) \leq C$ pour tout $t \in J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $C > 0$. Soit $t_0 \in J$. On suppose qu'il existe $A, B \geq 0$ tels que*

$$h(t) \leq A + B\left|\int_{t_0}^{t} h(s)ds\right|$$

*pour tout $t \in J$. Alors, étant donné $n \in \mathbb{N}$, on a*

$$h(t) \leq \frac{CB^{n+1}|t-t_0|^{n+1}}{(n+1)!} + A\sum_{i=0}^{n} \frac{B^i|t-t_0|^i}{i!} \tag{3.5.1}$$

*pour tout $t \in J$. En conséquence,*

$$h(t) \leq Ae^{B|t-t_0|} \tag{3.5.2}$$

*pour tout $t \in J$.*

***Preuve.*** On voit bien que

$$h(t) \leq A + B\left|\int_{t_0}^{t} h(s)ds\right| = A + B\int_{\min(t,t_0)}^{\max(t,t_0)} h(s)ds \leq A + CB|t-t_0|,$$

pour tout $t \in J$, ce qui nous donne l'inégalité (3.5.1) pour $n = 0$. On suppose que l'inégalité (3.5.1) est vérifiée pour $n \in \mathbb{N}$. Alors,

$$h(t) \leq A + B\left|\int_{t_0}^{t} h(s)ds\right| \leq A + B\int_{\min(t,t_0)}^{\max(t,t_0)} \left(\frac{CB^{n+1}|s-t_0|^{n+1}}{(n+1)!} + A\sum_{i=0}^{n}\frac{B^i|s-t_0|^i}{i!}\right)ds$$
$$= \frac{CB^{n+2}|t-t_0|^{n+2}}{(n+2)!} + A\sum_{i=0}^{n+1}\frac{B^i|t-t_0|^i}{i!},$$

pour tout $t \in J$, ce qui montre (3.5.1) pour $n+1$. L'inégalité (3.5.2) est une conséquence directe de (3.5.1), vu que la limite pour $t$ fixe lorsque $n$ tend vers $+\infty$ du premier opérande dans le membre de droite de (3.5.1) est zéro. □

Le premier résultat intéressant sur les courbes intégrales approximatives est le suivant.

**Proposition 3.5.3.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et $K$-lipschitzien sur $U$ uniformément pour $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert et $U \subseteq \mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$. Soient $\alpha : J' \to U$ et $\beta : J' \to U$ une courbes intégrale $\epsilon$-approximative et $\varepsilon$-approximative, respectivement. Soit $\bar{\epsilon} = \epsilon + \varepsilon$. Alors,*

$$\|\alpha(t) - \beta(t)\| \leq \left(\|\alpha(t_0) - \beta(t_0)\| + \frac{\bar{\epsilon}}{K}\right)e^{K|t-t_0|}$$

*pour tout $t \in J'$.*

***Preuve.*** On voit bien que

$$\|\alpha'(t) - \beta'(t) - \bar{f}(t,\alpha(t)) + \bar{f}(t,\beta(t))\| \leq \|\alpha'(t) - \bar{f}(t,\alpha(t))\| + \|\beta'(t) - \bar{f}(t,\beta(t))\| \leq \epsilon + \varepsilon = \bar{\epsilon},$$

pour tout $t \in J'$. L'inégalité précédente avec l'inégalité triangulaire nous disent que

$$\left|\left\|\big(\alpha(t)-\beta(t)\big) - \big(\alpha(t_0)-\beta(t_0)\big)\right\| - \left\|\int_{t_0}^{t} \bar{f}(s,\alpha(s)) - \bar{f}(s,\beta(s))ds\right\|\right|$$
$$\leq \left|\left\|\int_{t_0}^{t}\big(\alpha'(s)-\beta'(s)\big)ds\right\| - \left\|\int_{t_0}^{t} \bar{f}(s,\alpha(s)) - \bar{f}(s,\beta(s))ds\right\|\right|$$
$$\leq \left\|\int_{t_0}^{t}\big(\alpha'(s)-\beta'(s)\big)ds - \int_{t_0}^{t}\Big(\bar{f}(s,\alpha(s)) - \bar{f}(s,\beta(s))\Big)ds\right\|$$
$$\leq \int_{\min(t,t_0)}^{\max(t,t_0)} \left\|\big(\alpha'(s)-\beta'(s)\big) - \Big(\bar{f}(s,\alpha(s)) - \bar{f}(s,\beta(s))\Big)\right\| ds \leq \bar{\epsilon}|t-t_0|$$

pour tout $t \in J'$, ce qui implique que

$$\|\alpha(t)-\beta(t)\| - \|\alpha(t_0)-\beta(t_0)\| \leq \left\|\int_{t_0}^{t} \bar{f}(s,\alpha(s)) - \bar{f}(s,\beta(s))ds\right\| + \bar{\epsilon}|t-t_0|$$
$$\leq \int_{\min(t,t_0)}^{\max(t,t_0)} \left\|\bar{f}(s,\alpha(s)) - \bar{f}(s,\beta(s))\right\| ds + \bar{\epsilon}|t-t_0|$$
$$\leq K\int_{\min(t,t_0)}^{\max(t,t_0)} \|\alpha(t)-\beta(t)\|ds + \bar{\epsilon}|t-t_0|$$
$$= K\int_{\min(t,t_0)}^{\max(t,t_0)} \left(\|\alpha(t)-\beta(t)\| + \frac{\bar{\epsilon}}{K}\right)ds$$
$$= K\left|\int_{t_0}^{t}\left(\|\alpha(t)-\beta(t)\| + \frac{\bar{\epsilon}}{K}\right)\right|$$

pour tout $t \in J'$.

On définit l'application $h : J' \to \mathbb{R}_{\geq 0}$ via

$$h(t) = \big\|\alpha(t) - \beta(t)\,\big\| + \frac{\bar{\epsilon}}{K}$$

pour tout $t \in J$. Alors $h$ est continue, vu que $\alpha$ et $\beta$ le sont. En outre, si l'on applique (3.5.2) à $h|_I$ avec $I \subseteq J$ intervalle fermé et borné dont l'intérieur inclut $t_0$, $A = h(t_0)$ et $B = K$, on trouve que

$$\big\|\alpha(t) - \beta(t)\,\big\| \leq \big\|\alpha(t) - \beta(t)\,\big\| + \frac{\bar{\epsilon}}{K} \leq \left(\big\|\alpha(t_0) - \beta(t_0)\,\big\| + \frac{\bar{\epsilon}}{K}\right)e^{K|t-t_0|}$$

pour tout $t \in J$, comme on voulait démontrer. $\square$

On va donner une autre preuve de l'unicité du flot local, ainsi qu'une preuve de sa continuité.

**Proposition 3.5.4** (Unicité et continuité du flot local). *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. On fixe $a, b, K, C > 0$ comme dans la Proposition 3.4.3 et on admet que le flot local $\hat{\alpha} : ]t_0 - b, t_0 + b[\, \times B_{\|\ \|}(v_0, a) \to U$ de $\bar{f}$ en $(t_0, v_0)$ existe. Alors, il est unique, et l'application $\hat{\alpha}$ est continue et lipschitzienne sur $B_{\|\ \|}(v_0, a)$ uniformément pour $]t_0 - b, t_0 + b[$.*

***Preuve.*** Pour montrer l'unicité, soient $\hat{\alpha}, \hat{\beta} : ]t_0 - b, t_0 + b[\, \times B_{\|\ \|}(v_0, a) \to U$ deux flots locaux de $\bar{f}$ en $(t_0, v_0)$. Pour tout $v \in B_{\|\ \|}(v_0, a)$, on considère les courbes intégrales $\hat{\alpha}_v, \hat{\beta}_v : ]t_0 - b, t_0 + b[ \to U$ de $\bar{f}$, que l'on regarde comme courbes intégrales $\epsilon$-approximative et $\varepsilon$-approximative, respectivement, avec $\epsilon = \varepsilon = 0$. Comme $\hat{\alpha}_v(t_0) = \hat{\beta}_v(t_0) = v$ pour tout $v \in B_{\|\ \|}(v_0, a)$, la Proposition 3.5.3 nous dit que $\hat{\alpha}_v(t) = \hat{\beta}_v(t)$ pour tout $t \in ]t_0 - b, t_0 + b[$.

Par ailleurs, on a aussi

$$\begin{aligned}
\big\|\hat{\alpha}(t, v) - \hat{\alpha}(s, u)\big\| &\leq \big\|\hat{\alpha}(t, v) - \hat{\alpha}(t, u)\big\| + \big\|\hat{\alpha}(t, u) - \hat{\alpha}(s, u)\big\| \\
&\leq \big\|v - u\big\|\frac{K+1}{K}e^{K|t-t_0|} + \left\|\int_s^t \bar{f}(x, \hat{\alpha}(x, u)) - \bar{f}(x, \hat{\alpha}(x, u))dx\right\| \\
&\leq \big\|v - u\big\|\frac{K+1}{K}e^{K|t-t_0|} + 2C|t - s|
\end{aligned}$$

pour tout $(t, v), (s, u) \in ]t_0 - b, t_0 + b[\, \times B_{\|\ \|}(v_0, a)$, où l'a appliqué la Proposition 3.5.3 dans le premier opérande du deuxième membre en regardant $\hat{\alpha}(t, v)$ comme une courbe intégrale $\|v-u\|$-approximative de $\bar{f}$. En conséquence, $\hat{\alpha}$ est continue et lipschitzienne sur $B_{\|\ \|}(v_0, a)$ uniformément pour l'intervalle $]t_0 - b, t_0 + b[$. $\square$

## 3.6 Unicité globale et prolongement des courbes intégrales

**Proposition 3.6.1** (Unicité globale de courbes intégrales). *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert et $U \subseteq \mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$. Soient $\alpha : J' \to U$ et $\beta : J'' \to U$ deux courbes intégrales de $\bar{f}$ avec condition initiale $v_0 \in U$ à temps $t_0$, où $t_0 \in J', J''$ et $J', J'' \subseteq J$ sont des intervalles ouverts. Alors $\alpha|_{J' \cap J''} = \beta|_{J' \cap J''}$.*

***Preuve.*** On va montrer l'égalité $\alpha|_{J' \cap J'' \cap \mathbb{R}_{\geq t_0}} = \beta|_{J' \cap J'' \cap \mathbb{R}_{\geq t_0}}$ et on va laisser au lecteur/à la lectrice l'identité analogue $\alpha|_{J' \cap J'' \cap \mathbb{R}_{\leq t_0}} = \beta|_{J' \cap J'' \cap \mathbb{R}_{\leq t_0}}$. Soit $E = \{t \in J' \cap J'' : \alpha(s) = \beta(s)$ pour tout $s \in [t_0, t]\}$. En particulier, $t_0 \in E$. La Proposition 3.4.3 nous dit qu'il existe $b > 0$ tel que $t_0 + b \in E$. Si $E$ n'est pas majoré, alors le l'identité $\alpha|_{J' \cap J'' \cap \mathbb{R}_{\geq t_0}} = \beta|_{J' \cap J'' \cap \mathbb{R}_{\geq t_0}}$ est immédiate. On suppose désormais que $E$ est majoré et on pose $e = \sup E \in \mathbb{R}$. On affirme que $e$ est la borne supérieure de l'intervalle $J' \cap J''$. Si ce n'est pas le cas, alors $\alpha(e) = \beta(e)$, vu que $\alpha$ et $\beta$ sont continues, ce qui implique que $e \in E$, et les applications $\alpha$ et $\beta$ sont des courbes intégrales de $\bar{f}$ avec condition initiale $\alpha(e) = \beta(e)$ à temps $e$. Or, la Proposition 3.4.3 nous dit que, localement, $\bar{f}$ admet une unique courbe intégrale avec condition initiale $\alpha(e) = \beta(e)$ à temps $e$, *i.e.* il existe $r > 0$ tel que la courbe intégrale $\gamma : ]e - t, e + t[ \to U$

de $\bar{f}$ avec condition initiale $\alpha(e) = \beta(e)$ à temps $e$ est unique. En conséquence, $\alpha(t) = \beta(t)$ pour tout $t \in ]e-r, e+r[$, ce qui implique que $e < \sup E \in \mathbb{R}$, ce qui est absurde. En conséquence, $e$ est la borne supérieure de l'intervalle $J' \cap J''$, ce qui nous dit que $\alpha|_{J' \cap J'' \cap \mathbb{R}_{\geq t_0}} = \beta|_{J' \cap J'' \cap \mathbb{R}_{\geq t_0}}$. $\square$

**Définition 3.6.2.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert $0$ et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Étant donné $t_0 \in J$ et $v_0 \in U$, on définit $J(t_0, v_0) \subseteq \mathbb{R}$ comme la réunion de tous les intervalles ouverts $J' \subseteq J$ incluant $t_0$ tels qu'il existe une courbe intégrale $\alpha : J' \to U$ de $\hat{f}$ avec condition initiale $v_0$ à temps $t_0$.*

*Noter que, par la Proposition 3.6.1, étant donné $t_0 \in J$ et $v_0 \in U$, il existe une unique courbe intégrale $\alpha : J(t_0, v_0) \to U$ de $\hat{f}$ avec condition initiale $v_0$ à temps $t_0$.*

**Exemple 3.6.3.** *Soit $\bar{f} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ le champ vectoriel indépendant du temps donné par $\bar{f}(t, x) = x^2$ pour tous $t, x \in \mathbb{R}$. Comme $\bar{f}$ est de classe $C^1$, étant donné $x_0 \in \mathbb{R}$ et $t_0 \in \mathbb{R}$, on sait que $\bar{f}$ possède une unique courbe intégrale avec condition initiale $x_0$ à temps $t_0$. Soit $\alpha : J \to \mathbb{R}$ une courbe intégrale de $\bar{f}$ qui satisfait que $\alpha(t_0) = x_0$, où $J$ est l'intervalle ouvert maximal incluant $t_0$. Si $x_0 = 0$, on voit bien que $J(t_0, x_0) = \mathbb{R}$ et $\alpha(t) = 0$ pour tout $t \in \mathbb{R}$ est la courbe intégrale respective. Si $\alpha(t_0) = x_0 \neq 0$, alors $\alpha(t) \neq 0$ dans un intervalle ouvert $J' \subseteq J$ incluant $t_0$, ce qui implique que*

$$t' = 1 = \frac{\alpha'(t)}{\alpha(t)^2} = \left( -\frac{1}{\alpha(t)} \right)'$$

*pour tout $t \in J'$, i.e.*

$$t + \frac{1}{\alpha(t)} = C,$$

*pour tout $t \in J'$, avec $C \in \mathbb{R}$ une constante. En conséquence,*

$$\alpha(t) = \frac{1}{t_0 - t + x_0^{-1}} \tag{3.6.1}$$

*pour tout $t \in J'$. On voit bien que l'expression* (3.6.1) *pour tout $t \in \mathbb{R}_{>t_0 + x_0^{-1}}$ donne une courbe intégrale de $\bar{f}$ avec condition initiale $x_0$ à temps $t_0$ si $x_0 < 0$, tandis que l'expression* (3.6.1) *pour tout $t \in \mathbb{R}_{<t_0 + x_0^{-1}}$ donne une courbe intégrale de $\bar{f}$ avec condition initiale $x_0$ à temps $t_0$ si $x_0 > 0$. Comme la limite de* (3.6.1) *lorsque $t$ tend vers $t_0 + x_0^{-1}$ ne converge pas, on conclut que $J(t_0, x_0) = \mathbb{R}_{>t_0 + x_0^{-1}}$ si $x_0 < 0$ et $J(t_0, x_0) = \mathbb{R}_{<t_0 + x_0^{-1}}$ si $x_0 > 0$. En conséquence, l'intervalle maximal d'une courbe intégrale d'un champ vectoriel ne coïncide pas forcément avec l'intervalle dans la domaine de définition du champ vectoriel (dans ce cas, $\mathbb{R}$).*

Le résultat suivant nous donne un critère nécessaire et suffisant pour pouvoir comparer l'intervalle du domaine de définition du champ vectoriel et celui des courbes intégrales associées.

**Proposition 3.6.4** (Prolongement de courbes intégrales)**.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Soit $\alpha : J' \to U$ une courbe intégrale de $\bar{f}$ avec condition initiale $v_0$ à temps $t_0 \in J'$ définie sur un intervalle ouvert maximal $J' \subseteq J$. On suppose qu'il existe un intervalle ouvert $J'' \subseteq J'$ tel que les bornes supérieures (resp., inférieures) de $J'$ et de $J''$ coïncident, $\overline{\alpha(t)} \subseteq U$ et qu'il existe $C > 0$ tel que $||f(t, \alpha(t))|| \leq C$ pour tout $t \in J''$. Alors, les bornes supérieures (resp., inférieures) de $J$ et de $J'$ coïncident.*

***Preuve.*** On va démontrer l'énoncé pour les bornes supérieures, le cas des bornes inférieures étant complètement analogue. On suppose que $J = ]a, b[$, $J' = ]a', b'[$ et $J'' = ]a'', b''[$, avec $a, a', a'' \in \mathbb{R} \sqcup \{-\infty\}$ et $b, b', b'' \in \mathbb{R} \sqcup \{+\infty\}$. Par hypothèse, $b' = b''$. Si $b' = b$ il n'y a rien à démontrer. On va supposer que $b' < b$ si $b \in \mathbb{R}$, et $b' \in \mathbb{R}$ si $b = +\infty$. Comme $\alpha$ est un courbe intégrale de $\bar{f}$ avec condition initiale $v_0$ à temps $t_0 \in J'$, alors

$$\alpha(t) = v_0 + \int_{t_0}^{t} \bar{f}\big(s, \alpha(s)\big) ds$$

pour tout $t \in J'$, ce qui nous dit que

$$\|\alpha(t) - \alpha(t')\| = \left\| \int_t^{t'} \bar{f}\big(s, \alpha(s)\big) ds \right\| \leq \int_{\min(t,t')}^{\max(t,t')} \|\bar{f}\big(s, \alpha(s)\big)\| ds \leq C|t' - t|$$

pour tous $t, t' \in J''$. En conséquence, la limite

$$\lim_{t \to b'} \alpha(t)$$

existe, vu que toute suite de la forme $(\alpha(t_n))_{n\in\mathbb{N}}$, avec $(t_n)_{n\in\mathbb{N}} \in J''^{\mathbb{N}}$ convergente vers $b'$ lorsque $n$ tend vers $+\infty$, est une suite de Cauchy. En outre, par hypothèse, cette limite appartient à $U$. Soit $u_0 \in U$ la limite de $\alpha(t)$ lorsque $t$ tend vers $b'$. On conclut que l'on peut prolonger $\alpha$ en une application $\alpha : J' \sqcup \{b'\} \to U$ donnée par

$$\alpha(t) = u_0 + \int_{b'}^{t} \bar{f}\big(s, \alpha(s)\big) ds$$

pour tout $t \in J' \sqcup \{b'\}$. Par la Proposition 3.4.3, il existe une unique courbe intégrale $\beta : ]b' - \epsilon, b' + \epsilon[ \to U$ de $\bar{f}$ avec condition initiale $u_0$ à temps $b'$, où l'on suppose que $a' < b' - \epsilon$. En employant la terminologie et les résultat de la Remarque 3.4.4, $\beta|_{[b'-\epsilon, b'+\epsilon]}$ est un point fixe de l'application $I_{b', u_0, \epsilon/2}$ et $\alpha|_{[b'-\epsilon, b']}$ est un point fixe de l'application $I_{b', u_0, \epsilon/2, 0}$, ce qui nous dit que $\beta|_{[b'-\epsilon, b']} = \alpha|_{[b'-\epsilon, b']}$. En conséquence, l'application $\bar{\alpha} : J' \cup [b', b' + \epsilon[ \to U$ donnée par $\bar{\alpha}(t) = \alpha(t)$ si $t \in J'$ et $\bar{\alpha}(t) = \beta(t)$ si $t \in [b', b' + \epsilon[$ est une courbe intégrale de $\bar{f}$ avec condition initiale $v_0$ à temps $t_0$. Comme on avait supposé que $J'$ était maximal, on trouve une contradiction. En conséquence, $b' = b$. □

## 3.7 Champs vectoriel dépendant des paramètres et flot global

On peut considérer des familles de champs vectoriels dépendant du temps.

**Définition 3.7.1.** *Soient $n, m \in \mathbb{Z}_{>0}$. Une **famille de champs vectoriels dépendant du temps** (ou **champ vectoriel dépendant du temps et des paramètres**) est une application $\hat{f} : J \times U \times W \to \mathbb{R}^n$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert non vide, $U \subseteq \mathbb{R}^n$ et $W \subseteq \mathbb{R}^m$ sont des ouverts non vides. On va regarder tout champ vectoriel dépendant du temps $\bar{f} : J \times U \to \mathbb{R}^n$ comme un champ vectoriel dépendant du temps et des paramètres $\hat{f} : J \times U \times W \to \mathbb{R}^n$, avec $\hat{f}(t, v, w) = \bar{f}(t, v)$ pour tout $t \in J$, $v \in U$ et $w \in W$.*

**Remarque 3.7.2.** *Pour toute famille de champs vectoriels dépendant du temps $\hat{f} : J \times U \times W \to \mathbb{R}^n$ et $w \in W$, on considère le champ vectoriel dépendant du temps $\hat{f}_w : J \times U \to \mathbb{R}^n$ défini par $\hat{f}_w(t, v) = \hat{f}(t, v, w)$ pour tout $t \in J$ et $v \in U$. Tous les résultats des Sections 3.1-3.6 sont valables pour toute famille de champs vectoriels dépendant du temps $\hat{f} : J \times U \times W \to \mathbb{R}^n$ si l'on remplace dans les énoncés respectifs $\bar{f}$ par $\hat{f}_w$ pour $w \in W$.*

**Définition 3.7.3.** *Soit $\hat{f} : J \times U \times W \to \mathbb{R}^n$ un champ vectoriel continu dépendant du temps et des paramètres, tel que, pour tout $w \in W$, $\hat{f}_w$ est localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert 0 et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. Étant donné $v \in U$ et $w \in W$, on définit $J(v, w) \subseteq \mathbb{R}$ comme la réunion de tous les intervalles $J' \subseteq J$ incluant 0 tels qu'il existe une courbe intégrale $\alpha_w : J' \to U$ de $\hat{f}_w$ avec condition initiale $v$ à temps $t = 0$, et $\mathscr{D}(\hat{f}) = \{(t, v, w) \in J \times U \times W : t \in J(v, w)\}$. On pose aussi $t^-(v, w) \in \mathbb{R} \sqcup \{-\infty\}$ et $t^+(v, w) \in \mathbb{R} \sqcup \{+\infty\}$ tels que $J(v, w) = ]t^-(v, w), t^+(v, w)[$.*

*Noter que, par la Proposition 3.6.1, étant donné $v \in U$ et $w \in W$, il existe une unique courbe intégrale $\alpha_w : J(v, w) \to U$ de $\hat{f}_w$ avec condition initiale $v$ à temps $t = 0$. Le **flot global** de $\hat{f}$ est l'application $\breve{\alpha} : \mathscr{D}(\hat{f}) \to U$ telle que l'application associée $\breve{\alpha}_{v,w} : J(v, w) \to U$ donnée par $\breve{\alpha}_{v,w}(t) = \breve{\alpha}(t, v, w)$ soit l'unique courbe intégrale de $\hat{f}_w$ avec condition initiale $v$ à temps 0.*

**Remarque 3.7.4.** *Soit $\bar{f} : J \times U \to \mathbb{R}^n$ un champ vectoriel dépendant du temps continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$, où $J \subseteq \mathbb{R}$ est un intervalle ouvert 0 et $U \subseteq \mathbb{R}^n$ est un ouvert non vide de $\mathbb{R}^n$. On remarque que la restriction du flot global n'est pas en général un flot local. En effet, si $\bar{\alpha} : J' \times U' \to U$ est un flot local de $\bar{f}$ en $(t_0, x_0)$, où $J' \subseteq J$ est un intervalle ouvert incluant $t_0$ et $U' \subseteq U$ est un ouvert de $\mathbb{R}^n$ incluant $x_0$, et $\breve{\alpha} : \mathscr{D}(\bar{f}) \to U$ est le flot global de $\bar{f}$, alors $\breve{\alpha}|_{J' \times U'} \neq \bar{\alpha}$, car $\bar{\alpha}(t_0, x) = x$ pour tout $x \in U'$ tandis que $\breve{\alpha}(t_0, x) \neq x$ en général. La relation entre les deux est plutôt de la forme $\breve{\alpha}(t, y) = \bar{\alpha}(t, \breve{\alpha}(t_0, y))$ pour tout $t \in J'$ et $y \in U$ tel que $\breve{\alpha}(t_0, y) \in U'$.*

**Exemple 3.7.5.** *On continue avec l'Exemple 3.6.3. Dans ce cas, on voit bien que* $\mathscr{D}(\bar{f}) = \{(t,x) \in \mathbb{R}^2 : tx < 1\}$ *et le flot global* $\breve{\alpha} : \mathscr{D}(\bar{f}) \to \mathbb{R}$ *est donné par*

$$\breve{\alpha}(t,x) = \frac{x}{1-tx}$$

*pour tout* $(t,x) \in \mathscr{D}(\bar{f})$.

## 3.8 Équations différentielles linéaires

**Définition 3.8.1.** *Soit* $\ell : J \times W \to \mathrm{L}(\mathbb{R}^n)$ *une application, où* $J \subseteq \mathbb{R}$ *est un intervalle ouvert non vide et* $W \subseteq \mathbb{R}^m$ *une partie ouverte non vide. Le* ***champ vectoriel dépendant du temps et dépendant des paramètres associé à*** $\ell$ *est donné par l'application* $\hat{f} : J \times \mathbb{R}^n \times W \to \mathbb{R}^n$ *définie par* $\hat{f}(t,v,w) = \ell(t,w)(v)$*, pour tout* $t \in J$*,* $v \in \mathbb{R}^n$ *et* $w \in W$*.*

**Proposition 3.8.2** (Domaine d'équations différentielles linéaires)**.** *Soit* $\ell : J \times W \to \mathrm{L}(\mathbb{R}^n)$ *une application continue, où* $J \subseteq \mathbb{R}$ *est un intervalle ouvert incluant* $0$ *et* $W \subseteq \mathbb{R}^m$ *une partie ouverte non vide, et soit* $\hat{f} : J \times \mathbb{R}^n \times W \to \mathbb{R}^n$ *le champ vectoriel dépendant du temps et dépendant des paramètres associé à* $\ell$*. Alors,* $\mathscr{D}(\hat{f}) = J \times \mathbb{R}^n \times W$*.*

***Preuve.*** Soit $w \in W$. On va montrer que le domaine de définition maximal de chaque courbe intégrale $\alpha : J' \to \mathbb{R}^n$ de $\hat{f}_w$ est $J$. Pour cela il suffit de vérifier que les hypothèses de la Proposition 3.6.4 sont vérifiées. On va vérifier les hypothèses pour la borne supérieure de $J$, le cas de la borne inférieure étant analogue. On note d'abord que, comme $\ell : J \to \mathrm{L}(\mathbb{R}^n)$ est continue, alors $\hat{f}_w$ est continu et localement lipschitzien sur $U$ uniformément pour tout compact inclus dans $J$. En effet, la continuité de $\hat{f}_w$ suit directement de celle de $\ell$, et si $I \subseteq J$ est un fermé et borné, $||\ell(t)|| \leq K$ pour tout $t \in I$, où $K > 0$, ce qui implique que

$$\big\|\hat{f}_w(t,v) - \hat{f}_w(t,u)\big\| = \big\|\ell(t)(v) - \ell(t)(v)\big\| \leq \big\|\ell(t)\big\|.||v-u|| \leq K||v-u||$$

pour tout $v,u \in \mathbb{R}^n$ et $t \in I$. La condition $\overline{\alpha(t)} \subseteq U$ pour $t \in J''$ est immédiate, vu que $U = \mathbb{R}^n$ dans ce cas. Il reste à vérifier la dernière condition dans la Proposition 3.6.4. Si la borne supérieure $b'$ de $J'$ est $+\infty$ il n'y a rien à démontrer. On suppose alors que $b' \in \mathbb{R}$ et soit $I \subseteq \mathbb{R}$ un intervalle fermé et borné dont l'intérieur inclut $b'$. Par continuité de $\ell$, il existe $K > 0$ tel que $||\ell(t)|| \leq K$ pour tout $t \in I$. Comme $\alpha$ est une courbe intégrale de $\hat{f}_w$ on a lors que

$$\alpha(t) = \alpha(0) + \int_0^t \ell(s)\big(\alpha(s)\big)ds,$$

ce qui implique que

$$\big\|\alpha(t)\big\| = \big\|\alpha(0)\big\| + \int_{\min(0,t)}^{\max(0,t)} \big\|\ell(s)\big(\alpha(s)\big)\big\|ds \leq \big\|\alpha(0)\big\| + K\int_{\min(0,t)}^{\max(0,t)} \big\|\big(\alpha(s)\big)\big\|ds$$

pour tout $t \in I$. Lemme 3.5.2 nous dit alors que $||\alpha(t)|| \leq ||\alpha(0)||e^{K|t-t_0|}$ pour tout $t \in I$, ce qui implique que $\hat{f}_w(t,\alpha(t))$ est borné sur $I$. La Proposition 3.6.4 nous dit alors que le domaine de définition maximal de la courbe intégrale $\alpha : J' \to \mathbb{R}^n$ de $\hat{f}_w$ est $J$. En conséquence, $\mathscr{D}(\hat{f}_w) = J \times \mathbb{R}^n$. On conclut que $\mathscr{D}(\hat{f}) = J \times \mathbb{R}^n \times W$. □

**Proposition 3.8.3** (Solutions d'équations différentielles linéaires)**.** *Soient* $p_0, \dots, p_{n-1} : J \to \mathbb{R}$ *des fonctions continues, où* $J \subseteq \mathbb{R}$ *est un intervalle ouvert non vide. Soit* $S$ *l'ensemble formé des applications* $x : J \to \mathbb{R}$ *qui satisfont*

$$x^{(n)}(t) + \sum_{i=0}^{n-1} p_i(t)x^{(i)}(t) = 0$$

*pour tout* $t \in J$*, où* $x^{(0)}(t) = x(t)$*. Alors* $S$ *est un espace vectoriel réel de dimension* $n$*.*

*Preuve.* C'est clair que $S$ est un espace vectoriel réel, vu que toute combinaison linéaire d'éléments de $S$ est un élément de $S$. Soit $A : J \to \mathrm{M}_{n\times n}(\mathbb{R})$ l'application donnée par

$$A(t) = \begin{pmatrix} -p_0(t) & -p_1(t) & -p_2(t) & \dots & -p_{n-1}(t) \\ 0 & 1 & 0 & \dots & 0 \\ \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

pour tout $t \in J$, et soit $C$ l'ensemble formé des applications $\alpha : J \to \mathrm{M}_{n\times 1}(\mathbb{R})$ qui satisfont que

$$\alpha'(t) = A(t)\alpha(t)$$

pour tout $t \in J$. D'après l'Exemple 3.3.4, l'application linéaire $S \to C$ qui associe à $x \in S$ l'application $\alpha \in C$ donnée par $\alpha(t) = (x(t), x'(t), \dots, x^{(n-1)}(t))$ pour tout $t \in J$ est une bijection. Soit $t_0 \in J$. La Proposition 3.6.1 nous dit que l'application linéaire $C \to \mathrm{M}_{n\times 1}(\mathbb{R})$ qui associe $\alpha(t_0)$ à $\alpha \in C$ est une bijection. On conclut que $S$ a dimension $n$. □

Le **wronskian** d'une famille de fonctions $\{f_1, \dots, f_n\}$, où $f_1, \dots, f_n : J \to \mathbb{R}$ sont des fonctions $(n-1)$-fois différentiables, et $J$ est un intervalle ouvert non vide, est l'application $W(f_1, \dots, f_n) : J \to \mathbb{R}$ donnée par

$$W(f_1, \dots, f_n)(t) = \det \begin{pmatrix} f_1(t) & \dots & f_n(t) \\ \vdots & \ddots & \vdots \\ f_1^{(d-1)}(t) & \dots & f_n^{(d-1)}(t) \end{pmatrix}$$

pour tout $t \in J$.

**Lemme 3.8.4.** *Soient $f_1, \dots, f_n : J \to \mathbb{R}$ des fonctions $(n-1)$-fois différentiables, où $J$ est un intervalle ouvert non vide. S'il existe $t_0 \in J$ tel que $W(f_1, \dots, f_n)(t_0) \neq 0$, alors $\{f_1, \dots, f_n\}$ est libre.*

*Preuve.* On va démontrer la proposition contraposée. On suppose que $\{f_1, \dots, f_n\}$ n'est pas libre. On va montrer que $W(f_1, \dots, f_n)(t) \neq 0$ pour tout $t \in J$. Or, comme $\{f_1, \dots, f_n\}$ n'est pas libre, il existe $(c_1, \dots, c_n) \in \mathbb{R}^n$ non nul tel que

$$\sum_{i=1}^{n} c_i f_i(t) = 0,$$

pour tout $t \in J$. Cela implique que

$$\sum_{i=1}^{n} c_i f_i^{(j)}(t) = 0,$$

pour tout $t \in J$ et tout $j \in [\![1, d-1]\!]$. En conséquence,

$$\begin{pmatrix} f_1(t) & \dots & f_n(t) \\ \vdots & \ddots & \vdots \\ f_1^{(d-1)}(t) & \dots & f_n^{(d-1)}(t) \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

pour tout $t \in J$. En conséquence, le noyau de cette matrice est non trivial, pour tout $t \in J$, ce qui implique que

$$W(f_1, \dots, f_n)(t) = \det \begin{pmatrix} f_1(t) & \dots & f_n(t) \\ \vdots & \ddots & \vdots \\ f_1^{(d-1)}(t) & \dots & f_n^{(d-1)}(t) \end{pmatrix} = 0$$

pour tout $t \in J$, comme on voulait démontrer. □

**Remarque 3.8.5.** *La proposition contraposée du Lemme 3.8.4 n'est pas vraie en général. Par exemple, c'est facile à vérifier que les applications $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ données par $f_1(t) = t^2$ et $f_2(t) = t.|t|$ pour tout $t \in \mathbb{R}$ forment un ensemble libre mais $W(f_1, f_2)(t) = 0$ pour tout $t \in \mathbb{R}$.*

## 3.9 Résolution explicite des EDO linéaires avec terme inhomogène

### 3.9.1 Systèmes d'équations différentielles linéaires

Soit $n \in \mathbb{Z}_{>0}$. Dans cette sous-section on fixe un intervalle ouvert non vide $J \subseteq \mathbb{R}$ et des applications continues $A : J \to \mathrm{M}_n(\mathbb{R})$ et $b : J \to \mathrm{M}_{n\times 1}(\mathbb{R})$. On fixe $t_0 \in J$ et $v_0 \in \mathrm{M}_{n\times 1}(\mathbb{R})$. Pour simplifier, la bijection usuelle $\mathbb{R}^n \to \mathrm{M}_{n\times 1}(\mathbb{R})$ donnée par $v \mapsto v^t$ sera utilisée de façon implicite dans notre écriture, et on écrira alors $v_0 \in \mathbb{R}^n$. Même si ce n'est pas nécessaire, on va supposer que $A$ est diagonalisable de façon continue, *i.e.*, qu'il existe des applications continues $D, P : J \to \mathrm{M}_n(\mathbb{R})$ telles que $D(t)$ est diagonalisable, $\det(P(t)) \neq 0$ et

$$A(t) = P(t).D(t).P(t)^{-1}$$

pour tout $t \in J$. Pour $(d_1, \dots, d_n) \in \mathbb{R}^n$, on écrira $\mathrm{diag}(d_1, \dots, d_n)$ la matrice diagonale dont la $i$-ème entrée sur la diagonale est $d_i$, avec $i \in [\![1, n]\!]$. Soient $d_i : J \to \mathbb{R}$ $i \in [\![1, n]\!]$ les applications continues telles que $D(t) = \mathrm{diag}(d_1(t), \dots, d_n(t))$ pour tout $t \in J$.

Dans cette section on va déterminer de façon explicite l'application $\alpha : J \to \mathbb{R}^n$ qui satisfait l'équation différentielle

$$\alpha'(t) = A(t).\alpha(t) + b(t) \tag{3.9.1}$$

pour tout $t \in J$ et condition initiale $\alpha(t_0) = v_0$, associée au champ vectoriel dépendant du temps $\bar{f}(t, v) = A(t).v + b(t)$ pour tout $t \in J$ et $v \in \mathbb{R}^n$.

#### 3.9.1.1 Systèmes d'équations différentielles linéaires homogènes

On va commencer pour trouver une application $\alpha_h : J \to \mathrm{M}_n(\mathbb{R})$ différentiable qui satisfait que

$$\alpha_h'(t) = A(t).\alpha_h(t), \tag{3.9.2}$$

pour tout $t \in J$ avec la condition initiale $\alpha_h(t_0) = \mathrm{I}_n$, où $\mathrm{I}_n$ est la matrice identité.

Soit $\bar{\alpha}_h(t) = P(t)^{-1}.\alpha_h(t).P(t)$ pour tout $t \in J$. Alors, (3.9.2) équivaut à

$$\bar{\alpha}_h'(t) = D(t).\bar{\alpha}_h(t),$$

et $\bar{\alpha}_h(t_0) = \mathrm{I}_n$, ce qui nous dit que

$$\bar{\alpha}_h(t) = \mathrm{diag}\left(e^{\int_{t_0}^t d_1(s)ds}, \dots, e^{\int_{t_0}^t d_n(s)ds}\right)$$

pour tout $t \in J$. En conséquence,

$$\alpha_h(t) = P(t).\,\mathrm{diag}\left(e^{\int_{t_0}^t d_1(s)ds}, \dots, e^{\int_{t_0}^t d_n(s)ds}\right).P(t)^{-1} \tag{3.9.3}$$

pour tout $t \in J$. On dira que $\alpha_h : J \to \mathrm{M}_n(\mathbb{R})$ est la **solution fondamentale** de (3.9.2). Comme

$$\begin{aligned}\det\big(\alpha_h(t)\big) &= \det\big(P(t)\big).\det\left(\mathrm{diag}\left(e^{\int_{t_0}^t d_1(s)ds}, \dots, e^{\int_{t_0}^t d_n(s)ds}\right)\right).\det\big(P(t)^{-1}\big) \\ &= \det\left(\mathrm{diag}\left(e^{\int_{t_0}^t d_1(s)ds}, \dots, e^{\int_{t_0}^t d_n(s)ds}\right)\right) = \prod_{i=1}^n e^{\int_{t_0}^t d_i(s)ds} > 0,\end{aligned}$$

on conclut que $\alpha_h(t)$ est inversible pour tout $t \in J$.

#### 3.9.1.2 Cas général: méthode de variation des constantes

Soit $\alpha_h : J \to \mathrm{M}_n(\mathbb{R})$ la solution fondamentale de (3.9.2) donnée par (3.9.3) et soit

$$\alpha(t) = \alpha_h(t).w(t) \tag{3.9.4}$$

pour $t \in J$, avec $w : J \to \mathbb{R}^n$ une application différentiable. Alors, (3.9.2) nous dit que

$$\big(\alpha_h(t).w(t)\big)' = \alpha_h'(t).w(t) + \alpha_h(t).w'(t) = A(t).\alpha_h(t).w(t) + \alpha_h(t).w'(t) = A(t).\alpha(t) + \alpha_h(t).w'(t)$$

pour $t \in J$, ce qui implique que $\alpha$ donnée par (3.9.4) est une solution de (3.9.1) si et seulement si

$$\alpha_h(t).w'(t) = b(t)$$

pour $t \in J$, *i.e.*

$$w'(t) = \alpha_h(t)^{-1}.b(t)$$

pour $t \in J$. La condition initiale $\alpha(t_0) = v_0$ équivaut à $w(t_0) = \alpha_h(t_0)^{-1}.v_0$. Cela nous dit que

$$w(t) = \alpha_h(t_0)^{-1}.v_0 + \int_{t_0}^{t} \alpha_h(s)^{-1}.b(s)ds$$

pour $t \in J$, où l'intégrale précédente est calculée pour chaque entrée du vecteur $\alpha_h(s)^{-1}.b(s)$. On conclut que

$$\alpha(t) = \alpha_h(t).w(t) = \alpha_h(t).\alpha_h(t_0)^{-1}.v_0 + \int_{t_0}^{t} \alpha_h(t).\alpha_h(s)^{-1}.b(s)ds \tag{3.9.5}$$

pour $t \in J$.

**Exemple 3.9.1.** *On va résoudre l'équation différentielle associée* (3.9.1) *avec condition initiale* $(1,1)$ *à temps* $t_0 = 0$*, pour* $J = \mathbb{R}$*,*

$$A(t) = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} \text{ et } b(t) = \begin{pmatrix} e^t \\ t \end{pmatrix}$$

*pour tout* $t \in \mathbb{R}$*. Or, on voit bien que*

$$A = \underbrace{\begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}}_{P} . \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{D} . \underbrace{\begin{pmatrix} -1/2 & 1/2 \\ 3/2 & -1/2 \end{pmatrix}}_{P^{-1}} .$$

*En conséquence,* (3.9.3) *nous dit que la solution fondamentale de l'équation différentielle homogène est*

$$\alpha_h(t) = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} . \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} . \begin{pmatrix} -1/2 & 1/2 \\ 3/2 & -1/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3e^t - e^{-t} & e^{-t} - e^t \\ 3(e^t - e^{-t}) & 3e^{-t} - e^t \end{pmatrix}$$

*pour tout* $t \in \mathbb{R}$*. On notera* $\hat{D}(t) = \operatorname{diag}(e^{-t}, e^t)$ *pour tout* $t \in \mathbb{R}$*. Ce sera plus utile pour les calculs d'utiliser l'expression*

$$\alpha_h(t) = P.\hat{D}(t).P^{-1},$$

*pour tout* $t \in \mathbb{R}$*. Noter que*

$$\alpha_h(t)^{-1} = P.\hat{D}(t)^{-1}.P^{-1} = P.\hat{D}(-t).P^{-1},$$

*pour tout* $t \in \mathbb{R}$*.*

*Finalement,* (3.9.5) *nous dit que la solution* $\alpha$ *de* (3.9.1) *avec condition initiale* $(1,1)$ *à temps* $t_0 = 0$ *pour le choix de* $A$ *et* $b$ *précédents est donnée par*

$$\begin{aligned}
\alpha(t) &= \frac{1}{2} \begin{pmatrix} 3e^t - e^{-t} & e^{-t} - e^t \\ 3(e^t - e^{-t}) & 3e^{-t} - e^t \end{pmatrix} . \begin{pmatrix} 1 \\ 1 \end{pmatrix} + P. \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} .P^{-1}. \int_0^t P. \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} .P^{-1}. \begin{pmatrix} e^s \\ s \end{pmatrix} ds \\
&= \begin{pmatrix} e^t \\ e^t \end{pmatrix} + P. \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} . \int_0^t \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} . \begin{pmatrix} (s - e^s)/2 \\ (3e^s - s)/2 \end{pmatrix} ds \\
&= \begin{pmatrix} e^t \\ e^t \end{pmatrix} + P. \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} . \int_0^t \begin{pmatrix} e^s(s - e^s)/2 \\ (3 - se^{-s})/2 \end{pmatrix} ds \\
&= \begin{pmatrix} e^t \\ e^t \end{pmatrix} + P. \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} . \begin{pmatrix} (2te^t - e^{2t} - 2e^t + 3)/4 \\ (3t + te^{-t} + e^{-t} - 1)/2 \end{pmatrix} = \begin{pmatrix} e^t \\ e^t \end{pmatrix} + P. \begin{pmatrix} (2t - e^t - 2 + 3e^{-t})/4 \\ (3te^t + t + 1 - e^t)/2 \end{pmatrix} \\
&= \begin{pmatrix} e^t \\ e^t \end{pmatrix} + \begin{pmatrix} t + 3(2te^t - e^t + e^{-t})/4 \\ 2t - 1 + (6te^t - 5e^t + 9e^{-t})/4 \end{pmatrix} = \begin{pmatrix} t + (6te^t + e^t + 3e^{-t})/4 \\ 2t - 1 + (6te^t - e^t + 9e^{-t})/4 \end{pmatrix}
\end{aligned}$$

*pour tout* $t \in \mathbb{R}$*.*

#### 3.9.1.3 Application aux équations différentielles linéaires d'ordre 2

Soient $p, q, r : J \to \mathbb{R}$ des fonctions continues, où $J \subseteq \mathbb{R}$ est un intervalle ouvert, et soit $t_0 \in J$. On veut trouver l'application $x : J \to \mathbb{R}$ qui satisfait l'équation différentielle

$$x''(t) + p(t)x'(t) + q(t)x(t) = r(t) \tag{3.9.6}$$

pour tout $t \in J$ et les conditions initiales $x(t_0) = x_0$ et $x'(t_0) = v_0$. Soient

$$A(t) = \begin{pmatrix} -p(t) & -q(t) \\ 1 & 0 \end{pmatrix} \text{ et } b(t) = \begin{pmatrix} r(t) \\ 0 \end{pmatrix}$$

pour $t \in J$. C'est clair que $A : J \to \mathrm{M}_{2\times 2}(\mathbb{R})$ et $b : J \to \mathrm{M}_{2\times 1}(\mathbb{R})$ sont des application continues, vu que $p, q, r$ le sont. D'après l'Exemple 3.3.4, étant donné une application $x : J \to \mathbb{R}$, si elle est solution de (3.9.6) avec la condition initiale $x(t_0) = x_0$ et $x'(t_0) = v_0$, alors l'application $\alpha : J \to \mathbb{R}^2$ de la forme $\alpha(t) = (x(t), x'(t))$ pour $t \in J$ est solution de

$$\alpha'(t) = A(t).\alpha(t) + b(t) \tag{3.9.7}$$

pour tout $t \in J$ et condition initiale $\alpha(t_0) = (x_0, v_0)$, et toute solution $\alpha$ de (3.9.7) avec la condition initiale $\alpha(t_0) = (x_0, v_0)$ est obtenue de cette forme. La méthode expliquée dans les paragraphes précédents nous permet alors de résoudre (3.9.6), si la matrice $A(t)$ est diagonalisable de façon continue.

### 3.9.2 Une autre méthode pour les équations différentielles linéaires d'ordre 2

Soient $p, q, r : J \to \mathbb{R}$ des fonctions continues, où $J \subseteq \mathbb{R}$ est un intervalle ouvert, et soit $t_0 \in J$. Dans cette section on va donner une méthode directe pour trouver de façon explicite l'application $x : J \to \mathbb{R}$ qui satisfait l'équation différentielle

$$x''(t) + p(t)x'(t) + q(t)x(t) = r(t) \tag{3.9.8}$$

pour tout $t \in J$ et les conditions initiales $x(t_0) = x_0$ et $x'(t_0) = v_0$.

On va supposer que l'on connaît une famille libre $\{x_1, x_2\}$ de fonctions 2-fois différentiables définies sur $J$ qui satisfont que

$$x''(t) + p(t)x'(t) + q(t)x(t) = 0$$

pour tout $t \in J$.

Avant de continuer, on peut donner la réciproque suivante du Lemme 3.9.2, sous des hypothèses supplémentaires.

**Lemme 3.9.2.** *Soient $x_1, x_2 : J \to \mathbb{R}$ deux solutions de l'équation différentielle*

$$x''(t) + p(t)x'(t) + q(t)x(t) = 0 \tag{3.9.9}$$

*pour tout $t \in J$, où $p, q : J \to \mathbb{R}$ sont des applications continues et $J$ est un intervalle ouvert non vide. Si $\{x_1, x_2\}$ est libre, alors $W(x_1, x_2)(t) \neq 0$ pour tout $t \in J$.*

***Preuve.*** Un calcul long mais direct nous dit que $W(x_1, x_2)$ est solution de l'équation différentielle

$$W'(t) + p(t)W(t) = 0,$$

pour tout $t \in J$. En conséquence, la Proposition 3.6.1 nous dit que, si $W(x_1, x_2)(t_0) \neq 0$ pour $t_0 \in J$ quelconque, alors $W(x_1, x_2)(t) \neq 0$ pour tout $t \in J$. On suppose que $W(x_1, x_2)(t) = 0$ pour tout $t \in J$. Comme $x_i$ est une solution non nulle de (3.9.9) pour $i = 1, 2$, car $\{x_1, x_2\}$ est libre, la Proposition 3.6.1 nous dit qu'il existe un intervalle ouvert $J' \subseteq J$ tel que $x_i(t) \neq 0$ pour tout $t \in J'$ et $i = 1, 2$. Or, $W(x_1, x_2)(t) = 0$ pour tout $t \in J$ nous dit qu'il existe une application 2-fois différentiable $a : J' \to \mathbb{R}$ telle que $x_2(t) = a(t)x_1(t)$ et $x_2'(t) = a(t)x_1'(t)$ pour tout $t \in J'$, vu que les vecteurs $(x_1(t), x_1'(t))$ et $(x_2(t), x_2'(t))$ son linéairement dépendants pour tout $t \in J$. Alors, $a(t)x_1'(t) = x_2'(t) = a'(t)x_1(t) + a(t)x_1'(t)$, ce qui implique que $a'(t)x_1(t) = 0$ pour tout $t \in J'$, *i.e.* $a'(t) = 0$ pour tout $t \in J'$, car $x_1(t) \neq 0$ pour tout $t \in J'$. En conséquence, il existe $a \in \mathbb{R}$ tel que $a(t) = a$ pour tout $t \in J'$, *i.e.* $x_2(t) = ax_1(t)$ pour tout $t \in J'$. Comme $x_2 - ax_1$ est une solution de (3.9.9) et elle s'annule sur un intervalle ouvert non vide, elle est nulle sur $J$, ce qui implique que $\{x_1, x_2\}$ n'est pas libre, ce qui est absurde. En conséquence, $W(x_1, x_2)(t) \neq 0$ pour tout $t \in J$. $\square$

On continue maintenant avec (3.9.8). Pour cela, on va proposer que la solution de (3.9.8) soit de la forme

$$x(t) = a_1(t)x_1(t) + a_2(t)x_2(t) \tag{3.9.10}$$

pour tout $t \in J$, où $a_1, a_2 : J \to \mathbb{R}$ sont deux fonctions 2-fois différentiables. On suppose aussi que ces applications $a_1$ et $a_2$ satisfont l'identité

$$a_1'(t)x_1(t) + a_2'(t)x_2(t) = 0 \tag{3.9.11}$$

pour tout $t \in J$.

Or, (3.9.10) et (3.9.11) nous disent que

$$x'(t) = a_1'(t)x_1(t) + a_2'(t)x_2(t) + a_1(t)x_1'(t) + a_2(t)x_2'(t) = a_1(t)x_1'(t) + a_2(t)x_2'(t),$$

pour tout $t \in J$. Alors,

$$x''(t) = \big(a_1(t)x_1'(t) + a_2(t)x_2'(t)\big)' = a_1(t)x_1''(t) + a_2(t)x_2''(t) + a_1'(t)x_1'(t) + a_2'(t)x_2'(t)$$

pour tout $t \in J$. En conséquence, (3.9.10) satisfait (3.9.8) si et seulement si

$$a_1'(t)x_1'(t) + a_2'(t)x_2'(t) = r(t) \tag{3.9.12}$$

pour tout $t \in J$. Les équations (3.9.11) et (3.9.12) nous disent alors que

$$\begin{pmatrix} x_1(t) & x_2(t) \\ x_1'(t) & x_2'(t) \end{pmatrix} \cdot \begin{pmatrix} a_1'(t) \\ a_2'(t) \end{pmatrix} = \begin{pmatrix} 0 \\ r(t) \end{pmatrix}$$

pour tout $t \in J$. En conséquence,

$$\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} = \frac{1}{W(x_1, x_2)(t)} \begin{pmatrix} x_2'(t) & -x_2(t) \\ -x_1'(t) & x_1(t) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ r(t) \end{pmatrix} = \frac{1}{W(x_1, x_2)(t)} \begin{pmatrix} -r(t)x_2(t) \\ r(t)x_1(t) \end{pmatrix}$$

pour tout $t \in J$, où l'on a utilisé le Lemme 3.9.2. On conclut que

$$a_1(t) = -\int \frac{r(t)x_2(t)}{W(x_1, x_2)(t)} dt \text{ et } a_2(t) = \int \frac{r(t)x_1(t)}{W(x_1, x_2)(t)} dt \tag{3.9.13}$$

pour tout $t \in J$ et, par conséquent,

$$x(t) = -x_1(t) \int \frac{r(t)x_2(t)}{W(x_1, x_2)(t)} dt + x_2(t) \int \frac{r(t)x_1(t)}{W(x_1, x_2)(t)} dt$$

pour tout $t \in J$.

**Exemple 3.9.3.** *On veux trouver une application $x : \mathbb{R}_{>0} \to \mathbb{R}$ telle que*

$$x''(t) - \frac{2}{t}x'(t) + \frac{2}{t^2}x(t) = 3$$

*pour $t \in \mathbb{R}_{>0}$ et satisfait les conditions initiales $x(1) = 1$ et $x'(1) = 1$. C'est facile à voir que les applications $x_1, x_2 : \mathbb{R}_{>0} \to \mathbb{R}$ données par $x_1(t) = t$ et $x_2(t) = t^2$ pour $t \in \mathbb{R}_{>0}$ sont des solutions de l'équation*

$$x''(t) - \frac{2}{t}x'(t) + \frac{2}{t^2}x(t) = 0$$

*pour $t \in \mathbb{R}_{>0}$. Noter que $W(x_1, x_2)(t) = t^2 \neq 0$, et, d'après le Lemme 3.8.4, $\{x_1, x_2\}$ est un ensemble libre. Or,* (3.9.13) *nous dit que, dans ce cas*

$$a_1(t) = -\int \frac{r(t)x_2(t)}{W(x_1, x_2)(t)} dt = -\int \frac{3t^2}{t^2} dt = -3t + c_1$$

*et*

$$a_2(t) = \int \frac{r(t)x_1(t)}{W(x_1, x_2)(t)} dt = \int \frac{3t}{t^2} dt = 3\ln(t) + c_2$$

*for $t \in \mathbb{R}_{>0}$. En conséquence,*

$$x(t) = -3t^2 + 3t^2\ln(t) + c_1 t + c_2 t^2 = 3t^2\ln(t) + C_1 t + C_2 t^2$$

*pour tout $t \in J$, où $C_1 = c_1$ et $C_2 = c_2 - 3$. Les conditions initiales impliquent que $C_1 = 4$ et $C_2 = -3$.*

# Chapitre 4

# Théorie élémentaire de courbes paramétrées

Le but du quatrième chapitre c'est d'étudier les courbes paramétrées. On continue avec les mêmes conventions que celles du premier paragraphe du chapitre 2.

## 4.1 Premières définitions

On rappelle d'abord la définition suivante.

**Définition 4.1.1.** *Soient $n, p \in \mathbb{Z}_{>0}$ et $I \subseteq \mathbb{R}$ un intervalle non vide. On rappelle que $I^- = \inf(I)$ et $I^+ = \sup(I)$. Soit $\alpha : I \to \mathbb{R}^n$ une application. Si $I$ est un intervalle fini, on dit que $\alpha : I \to \mathbb{R}^n$ est de **classe $\mathrm{C}^p$ par morceaux** si $\alpha$ est continue et s'il existe $a_1 < \cdots < a_N \in I$ tel que $\alpha|_{]\,a_i,a_{i+1}\,[}$ est de classe $\mathrm{C}^p$ pour tout $i \in [\![0, N]\!]$, où $a_0 = I^-$ et $a_{N+1} = I^+$. Si $I$ est un intervalle infini, on dit que $\alpha : I \to \mathbb{R}^n$ est de **classe $\mathrm{C}^p$ par morceaux** si $\alpha|_{I'}$ est de **classe $\mathrm{C}^p$ par morceaux** pour tout intervalle fini $I' \subseteq I$.*

En particulier, étant donné une application $\alpha : I \to \mathbb{R}^n$ de classe $\mathrm{C}^p$ par morceaux, il existe une partie discrète $D \subseteq I$ telle que $\alpha|_{I^\circ \setminus D}$ est de classe $\mathrm{C}^p$. On pose

$$D_\alpha = \bigcap \big\{ D \subseteq I^\circ : \alpha|_{I^\circ \setminus D} \text{ est de classe } \mathrm{C}^p \big\}.$$

Il s'agit de la plus petite partie discrète $D_\alpha \subseteq I$ telle que $\alpha|_{I^\circ \setminus D_\alpha}$ est de classe $\mathrm{C}^p$. Noter que $\alpha$ est de classe $\mathrm{C}^p$ si et seulement si $D_\alpha = \emptyset$.

**Pour simplifier, dans toutes les définitions suivantes, le lecteur/la lectrice pourrait se restreindre au cas où toutes les applications sont de classe $\mathrm{C}^p$ sur l'intérieur de leurs domaines de définition.**

**Définition 4.1.2.** *Soient $n, p \in \mathbb{Z}_{>0}$ et $I \subseteq \mathbb{R}$ un intervalle non vide. Une **courbe (paramétrée) de classe $\mathrm{C}^p$ par morceaux** est une application continue $\alpha : I \to \mathbb{R}^n$ de classe $\mathrm{C}^p$ par morceaux. Dans le cas $p = 1$, on dira simplement que $\alpha$ est une **courbe**.*

**Définition 4.1.3.** *Soient $n, p \in \mathbb{Z}_{>0}$ et $I \subseteq \mathbb{R}$ un intervalle non vide. On dit qu'une courbe $\alpha : I \to \mathbb{R}^n$ de classe $\mathrm{C}^p$ par morceaux est **$p$-régulière** si $\{\alpha'(t), \dots, \alpha^{(p)}(t)\}$ est libre pour tout $t \in I^\circ \setminus D_\alpha$. Noter que dans ce cas $p \leq n$. Un courbe 1-régulière sera appelée plus simplement **régulière**, i.e. $\alpha'(t) \neq \mathbf{0}_{\mathbb{R}^n}$ pour tout $t \in I^\circ \setminus D_\alpha$, et une courbe de classe $\mathrm{C}^2$ par morceaux 2-régulière sera appelée plus simplement **birégulière**. Une courbe régulière $\alpha : I \to \mathbb{R}^n$ est **paramétrée par longueur d'arc** si $||\alpha'(s)|| = 1$ pour tout $s \in I^\circ \setminus D_\alpha$.*

*Si $\alpha$ est régulière, on définit l'application $\mathbf{t}_\alpha : I^\circ \setminus D_\alpha \to \mathbb{R}^n$*

$$\mathbf{t}_\alpha(s) = \frac{\alpha'(s)}{||\alpha'(s)||}$$

*pour tout $s \in I^\circ \setminus D_\alpha$. On dit que $\mathbf{t}_\alpha(s)$ est le **vecteur tangent** à $\alpha$ au temps $s$.*

**Définition 4.1.4.** *Soient $n \in \mathbb{Z}_{>0}$ et $I \subseteq \mathbb{R}$ un intervalle. La* ***longueur*** *d'une courbe régulière $\alpha : I \to \mathbb{R}^n$ est définie par*

$$\ell(\alpha) = \int_{I^-}^{I^+} ||\alpha'(s)||ds \in \mathbb{R}_{>0} \sqcup \{+\infty\}.$$

**Définition 4.1.5.** *Soient $n, p \in \mathbb{Z}_{>0}$ et $I, I' \subseteq \mathbb{R}$ des intervalles non vides. Un* ***reparamétrage*** *d'une courbe $\alpha : I \to \mathbb{R}^n$ classe $\mathrm{C}^p$ par morceaux est un homéomorphisme $g : I' \to I$ tel que $g|_{I'^\circ} : I'^\circ \to I^\circ$ est de classe $\mathrm{C}^p$ par morceaux avec $g'(s) \neq 0$ pour tout $s \in I'^\circ \setminus D_g$ et $g(D_g) \subseteq D_\alpha$. On dit que le reparamétrage $g$ est* ***positif*** *(resp.,* ***négatif****) si $g'(s) > 0$ (resp., $g'(s) < 0$) pour tout $s \in I'^\circ \setminus D_g$. La courbe* ***reparamétrée*** *est la composition $\alpha \circ g : I' \to \mathbb{R}^n$.*

**Proposition 4.1.6.** *Soient $n, p \in \mathbb{Z}_{>0}$ et $I, I' \subseteq \mathbb{R}$ des intervalles non vides. Si $\alpha : I \to \mathbb{R}^n$ est une courbe classe $\mathrm{C}^p$ par morceaux et $g : I' \to I$ est un reparamétrage de $\alpha$, alors*

$$\ell(\alpha) = \ell(\alpha \circ g).$$

*En plus, si $\alpha$ est $p$-régulière, $\alpha \circ g$ l'est aussi, $D_\alpha = g(D_{\alpha\circ g})$ et*

$$\mathbf{t}_\alpha\big(g(s)\big) = \pm \mathbf{t}_{\alpha\circ g}(s)$$

*pour tout $s \in I'^\circ \setminus D_{\alpha\circ g}$, où le signe $+$ (resp., $-$) correspond au cas où le reparamétrage $g$ est positif (resp., négatif).*

***Preuve.*** On fera la preuve pour $p = 1$ et on laisse le cas $p > 1$ au lecteur/à la lectrice. On voit bien que

$$\begin{aligned}\ell(\alpha \circ g) &= \int_{I'^-}^{I'^+} \big\|(\alpha \circ g)'(s)\big\|ds = \int_{I'^-}^{I'^+} \big\|\alpha'(g(s)).g'(s)\big\|ds \\ &= \int_{I'^-}^{I'^+} \big\|\alpha'(g(s))\big\|.\big|g'(s)\big|ds = \int_{I^-}^{I^+} ||\alpha'(r)||dr = \ell(\alpha),\end{aligned}$$

où l'on a utilisé Théorème 2.1.17 dans la deuxième égalité et le théorème de changement de variable dans la quatrième égalité.

Par ailleurs, comme $D_g \subseteq g^{-1}(D_\alpha)$ et $(\alpha \circ g)'(s) = \alpha'(g(s)).g'(s)$ pour tout $s \in I'^\circ \setminus D_{\alpha\circ g}$, d'après le Théorème 2.1.17, $D_{\alpha\circ g} \subseteq g^{-1}(D_\alpha)$, *i.e.* $g(D_{\alpha\circ g}) \subseteq D_\alpha$. En employant que $g|_{I'^\circ \setminus D_g} : I'^\circ \setminus D_g \to I^\circ \setminus g(D_g)$ est un $\mathrm{C}^1$-difféomorphisme et l'inclusion évidente $D_g \subseteq D_{\alpha\circ g}$, si l'on remplace $g$ par $g^{-1}$ et $\alpha$ par $\alpha \circ g$ dans l'argument précédent, on conclut que $g^{-1}(D_\alpha) \subseteq D_{\alpha\circ g}$, ce qui nous dit que $g(D_{\alpha\circ g}) = D_\alpha$. En outre, l'égalité $(\alpha \circ g)'(s) = \alpha'(g(s)).g'(s)$ et le fait que $g'(s) \neq 0$ pour tout $s \in I' \setminus D_{\alpha\circ g}$ nous disent que $\alpha$ est régulière si et seulement si $\alpha \circ g$ est régulière. Finalement, comme $(\alpha \circ g)'(s) = \alpha'(g(s)).g'(s)$, on voit bien que

$$\mathbf{t}_{\alpha\circ g}(s) = \frac{\alpha'(g(s)).g'(s)}{||\alpha'(g(s)).g'(s)||} = \frac{g'(s)}{|g'(s)|}\frac{\alpha'(g(s))}{||\alpha'(g(s))||} = \frac{g'(s)}{|g'(s)|}\mathbf{t}_\alpha(g(s))$$

pour tout $s \in I'^\circ \setminus D_{\alpha\circ g}$, ce qui montre le résultat, vu que $g'(s)/|g'(s)| = 1$ (resp., $g'(s)/|g'(s)| = -1$) pour tout $s \in I'^\circ \setminus D_{\alpha\circ g}$ si le reparamétrage est positif (resp., négatif). □

**Proposition 4.1.7.** *Soient $n, p \in \mathbb{Z}_{>0}$ et $I \subseteq \mathbb{R}$ un intervalle non vide. Si $\alpha : I \to \mathbb{R}^n$ est une courbe $p$-régulière, alors il existe un unique reparamétrage positif $g : I' \to I$ de $\alpha$ tel que $0 = I'^-$ et $\alpha \circ g : I' \to \mathbb{R}^n$ est paramétrée par longueur d'arc.*

***Preuve.*** On fera le cas $p = 1$ et on laisse le cas $p > 1$ au lecteur/à la lectrice Pour montrer l'unicité de $g$, soient $g : I' \to I$ et $\mathscr{g} : I'' \to I$ deux reparamétrages de $\alpha$ par longueur d'arc, où $I', I'' \subseteq \mathbb{R}$ sont des intervalles non vides avec $0 = I'^- = I''^-$. Comme $g$ et $\mathscr{g}$ sont des homéomorphismes, $\hbar = g^{-1} \circ \mathscr{g} : I'' \to I'$ est un homéomorphisme, $0 = I'^- \in I'$ si et seulement si $0 = I''^- \in I''$, et $I'^- \in I'$ si et seulement si $I''^- \in I''$. En plus, on voit bien que $\hbar$ est de classe $\mathrm{C}^1$ par morceaux. On écrira $\mathscr{g} = g \circ \hbar$ et on fixe $D = D_{\mathscr{g}} \cup \hbar^{-1}(D_g)$. Noter que $D \subseteq I''$ est une partie discrète. En conséquence, le Théorème 2.1.17 nous dit que

$$\mathscr{g}'(t) = (g \circ \hbar)'(t) = g'(\hbar(t))\hbar'(t)$$

pour tout $t \in I''^{\circ} \setminus D$, ce qui implique que $\hbar'(t) > 0$ pour tout $t \in I''^{\circ} \setminus D$. En outre, le Théorème 2.1.17 et le fait que les reparamétrages sont positifs nous disent que

$$\begin{aligned} 1 &= \|(\alpha \circ g)'(t)\| = \|\alpha'(g(t))\| g'(t) = \|\alpha'(g(t))\|(g \circ \hbar)'(t) = \|\alpha'(g(\hbar(t)))\| g'(\hbar(t))\hbar'(t) \\ &= \|(\alpha \circ g)'(\hbar(t))\|\hbar'(t) = \hbar'(t) \end{aligned}$$

pour tout $t \in I''^{\circ} \setminus D$. En conséquence, $\hbar(t) = t$ pour tout $t \in I''$, ce qui implique que $I' = I''$ et $g = g$.

On montrera maintenant l'existence. Soit $I' \subseteq \mathbb{R}$ l'unique intervalle tel que $I'^{-} = 0$ et $I'^{+} = \ell(\alpha)$, $0 \in I'$ si et seulement si $I^{-} \in I$ et $I'^{+} \in I'$ si et seulement si $I^{+} \in I$ et $\ell(\alpha) \in \mathbb{R}$. On définit $h : I \to I'$ via

$$h(t) = \int_{I^{-}}^{t} ||\alpha'(s)||ds \tag{4.1.1}$$

pour tout $t \in I$. Alors $h$ est continue et strictement croissante, ce qui nous dit que $h$ est un homéomorphisme. En plus, $h$ et de classe $C^1$ par morceaux, $D_h \subseteq D_\alpha$ et $h'(t) = ||\alpha'(t)|| > 0$ pour tout $t \in I^{\circ} \setminus D_h$. Soit $g : I' \to I$ l'application réciproque de $h$. Alors, $D_g = h(D_h)$ et le Théorème 2.1.17 nous dit que

$$g'(s) = \frac{1}{||\alpha'(g(s))||} > 0 \tag{4.1.2}$$

pour tout $s \in I'^{\circ} \setminus D_g$, et aussi

$$(\alpha \circ g)'(s) = \alpha'\big(g(s)\big)\frac{1}{||\alpha'(g(s))||}$$

pour tout $s \in I'^{\circ} \setminus D_g$, *i.e.* $\alpha \circ g$ est paramétrée par longueur d'arc. □

Étant donné une courbe $p$-régulière $\alpha : I \to \mathbb{R}^n$, où $I \subseteq \mathbb{R}$ un intervalle non vide, la courbe $p$-régulière $\alpha \circ g : I' \to \mathbb{R}^n$ construite dans la preuve de la Proposition 4.1.7 sera appelée la **courbe reparamétrée par longueur d'arc associée à $\alpha$** et sera notée ${}_\ell\alpha$.

## 4.2 Courbes planes

On rappelle qu'une partie non vide $\Pi \subseteq \mathbb{R}^n$ est dit **plane affine** s'il existe une partie libre $\{v, w\} \subseteq \mathbb{R}^n$ et un point $u_0 \in \Pi$ tel que

$$\Pi = \{u_0 + tv + sw : t, s \in \mathbb{R}\}.$$

De façon équivalente, une partie non vide $\Pi \subseteq \mathbb{R}^n$ est un plane affine s'il existe une application injective affine $L : \mathbb{R}^2 \to \mathbb{R}^n$ telle que $\Pi = \mathrm{Img}(L)$.

**Définition 4.2.1.** *Une courbe $\alpha : I \to \mathbb{R}^n$ est dite* **plane** *s'il existe une courbe $\bar{\alpha} : I \to \mathbb{R}^2$ et une application affine injective $L : \mathbb{R}^2 \to \mathbb{R}^n$ telles que $\alpha = L \circ \bar{\alpha}$. Noter que $\alpha$ est régulière si et seulement si $\bar{\alpha}$ l'est et $D_\alpha = D_{\bar{\alpha}}$.*

**Lemme 4.2.2.** *Soit $\alpha : I \to \mathbb{R}^n$ une courbe, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Alors $\alpha$ est plane si et seulement s'il existe un plan affine $\Pi \subseteq \mathbb{R}^n$ tel que $\mathrm{Img}(\alpha) \subseteq \Pi$.*

***Preuve.*** C'est clair que si $\alpha$ est plane, avec $\alpha = L \circ \bar{\alpha}$, où $\bar{\alpha} : I \to \mathbb{R}^2$ est une courbe et $L : \mathbb{R}^2 \to \mathbb{R}^n$ est une application affine injective, alors $\mathrm{Img}(\alpha) \subseteq \mathrm{Img}(L)$. Comme $\mathrm{Img}(L)$ est un plan affine, le résultat suit. Réciproquement, soit $\Pi \subseteq \mathbb{R}^n$ un plan affine incluant l'image de $\alpha$. Alors, il existe une application injective affine $L : \mathbb{R}^2 \to \mathbb{R}^n$ telle que $\Pi = \mathrm{Img}(L)$. Soit $\bar{L} : \mathbb{R}^2 \to \Pi$ la bijection déterminée par $L$ et soit $\bar{L}^{-1}$ sa réciproque. On pose $\bar{\alpha} = \bar{L}^{-1} \circ \alpha : I \to \mathbb{R}^2$. Alors, $\bar{\alpha}$ est une courbe et $\alpha = L \circ \bar{\alpha}$. □

**On va supposer désormais dans cette section que l'espace ambiant de toutes nos courbes planes est $\mathbb{R}^2$, *i.e.* toute courbe plane que l'on va considérer dans cette section sera de la forme $\alpha : I \to \mathbb{R}^2$, où $I \subseteq \mathbb{R}$ est un intervalle non vide.**

**Définition 4.2.3.** *Soit $\alpha : I \to \mathbb{R}^2$ une courbe régulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. On définit l'application $\mathbf{n}_\alpha : I^{\circ} \setminus D_\alpha \to \mathbb{R}^2$ tel que $(\mathbf{t}_\alpha(s), \mathbf{n}_\alpha(s))$ est une base orthonormale orientée de $\mathbb{R}^2$ pour tout $s \in I^{\circ} \setminus D_\alpha$. On dit que $\mathbf{n}_\alpha(s)$ est le* **vecteur normal** *à $\alpha$ en la longueur d'arc $s$.*

*Soit ${}_\ell\alpha : I' \to \mathbb{R}^2$ la courbe reparamétrée par longueur d'arc associée à $\alpha$. Alors, on définit la* ***courbure*** *$\kappa_{{}_\ell\alpha}(s) \in \mathbb{R}$ de $\alpha$ en la longueur d'arc $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$ par*

$$\mathbf{t}'_{{}_\ell\alpha}(s) = \kappa_{{}_\ell\alpha}(s)\mathbf{n}_{{}_\ell\alpha}(s)$$

*pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$.*

Le calcule de la courbure d'une courbe plane à partir de la Définition 4.2.3 est en principe compliqué, vu qu'elle reposent sur un reparamétrage par longueur d'arc. On a par contre l'expression suivante, qui ne nécessite pas cette dernière condition.

**Lemme 4.2.4.** *Soit $\alpha : I \to \mathbb{R}^2$ une courbe régulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. On suppose que $\alpha(t) = (\alpha_1(t), \alpha_2(t))$ pour tout $t \in I$. Soit $g : I' \to I$ l'unique reparamétrage positif de $\alpha$ tel que ${}_\ell\alpha = \alpha \circ g$ est paramétrée par longueur d'arc et $I'^{-} = 0$, donné dans la Proposition 4.1.7. On suppose que ${}_\ell\alpha(s) = ({}_\ell\alpha_1(s), {}_\ell\alpha_2(s))$ pour tout $s \in I'$. Alors,*

$$\kappa_{{}_\ell\alpha}(s) = {}_\ell\alpha_1'(s){}_\ell\alpha_2''(s) - {}_\ell\alpha_2'(s){}_\ell\alpha_1''(s),$$

*et les vecteurs $\mathbf{t}'_{{}_\ell\alpha}(s)$ et $\mathbf{n}'_{{}_\ell\alpha}(s)$ sont orthogonaux, pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$.*

*Si l'on pose*

$$\kappa_\alpha(t) = \frac{\alpha_1'(t)\alpha_2''(t) - \alpha_2'(t)\alpha_1''(t)}{[(\alpha_1'(t))^2 + (\alpha_2'(t))^2]^{3/2}},$$

*pour tout $t \in I^{\circ} \setminus D_\alpha$, alors $\kappa_\alpha(g(s)) = \kappa_{{}_\ell\alpha}(s)$ pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$.*

***Preuve.*** On remarque d'abord que, comme la courbe ${}_\ell\alpha$ est paramétrée par longueur d'arc,

$$\mathbf{t}_\alpha\big(g(s)\big) = \mathbf{t}_{{}_\ell\alpha}(s) = ({}_\ell\alpha_1'(s), {}_\ell\alpha_2'(s))$$

pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$, ce qui implique que $\mathbf{n}_\alpha(g(s)) = (-{}_\ell\alpha_2'(s), {}_\ell\alpha_1'(s))$ pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$.

Par définition de $\kappa_{{}_\ell\alpha}$, on a que $\mathbf{t}'_{{}_\ell\alpha}(s) = \kappa_{{}_\ell\alpha}(s)\mathbf{n}_{{}_\ell\alpha}(s)$, ce qui nous dit que ${}_\ell\alpha_1''(s) = -\kappa_{{}_\ell\alpha}(s){}_\ell\alpha_2'(s)$ et ${}_\ell\alpha_2''(s) = \kappa_{{}_\ell\alpha}(s){}_\ell\alpha_1'(s)$ pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$. En particulier,

$$\mathbf{n}'_{{}_\ell\alpha}(s) = \big(-{}_\ell\alpha_2''(s), {}_\ell\alpha_1''(s)\big) = \big(-\kappa_{{}_\ell\alpha}(s){}_\ell\alpha_1'(s), -\kappa_{{}_\ell\alpha}(s){}_\ell\alpha_2'(s)\big) = -\kappa_{{}_\ell\alpha}(s)\mathbf{t}_{{}_\ell\alpha}(s) \tag{4.2.1}$$

pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$. D'après (4.2.1), on voit bien que

$$\big\langle \mathbf{t}'_{{}_\ell\alpha}(s), \mathbf{n}'_{{}_\ell\alpha}(s)\big\rangle = \big\langle \kappa_{{}_\ell\alpha}(s)\mathbf{n}_{{}_\ell\alpha}(s), -\kappa_{{}_\ell\alpha}(s)\mathbf{t}'_{{}_\ell\alpha}(s)\big\rangle = -\kappa_{{}_\ell\alpha}(s)^2\big\langle \mathbf{n}_{{}_\ell\alpha}(s), (s)\big\rangle = 0$$

pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$.

Comme ${}_\ell\alpha = \alpha \circ g$, on voit bien que

$${}_\ell\alpha_i'(s) = \alpha_i'\big(g(s)\big)g'(s) \text{ et } {}_\ell\alpha_i''(s) = \alpha_i''\big(g(s)\big)g'(s)^2 + \alpha_i'\big(g(s)\big)g''(s),$$

pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$. En conséquence,

$$\begin{aligned}
\kappa_{{}_\ell\alpha}(s) &= {}_\ell\alpha_1'(s){}_\ell\alpha_2''(s) - {}_\ell\alpha_2'(s){}_\ell\alpha_1''(s) \\
&= \alpha_1'\big(g(s)\big)g'(s)\Big(\alpha_2''\big(g(s)\big)g'(s)^2 + \alpha_2'\big(g(s)\big)g''(s)\Big) \\
&\quad - \alpha_2'\big(g(s)\big)g'(s)\Big(\alpha_1''\big(g(s)\big)g'(s)^2 + \alpha_1'\big(g(s)\big)g''(s)\Big) \\
&= \Big(\alpha_1'\big(g(s)\big)\alpha_2''\big(g(s)\big) - \alpha_2'\big(g(s)\big)\alpha_1''\big(g(s)\big)\Big)g'(s)^3 \\
&= \frac{\alpha_1'\big(g(s)\big)\alpha_2''\big(g(s)\big) - \alpha_2'\big(g(s)\big)\alpha_1''\big(g(s)\big)}{\big\|\alpha'\big(g(s)\big)\big\|^3} = \kappa_\alpha\big(g(s)\big)
\end{aligned}$$

pour tout $s \in I'^{\circ} \setminus D_{{}_\ell\alpha}$. □

## 4.3 Courbes dans l'espace

**On suppose désormais que l'espace ambiant de toutes nos courbes est $\mathbb{R}^3$, *i.e.* toute courbe que l'on va considérer sera de la forme $\alpha : I \to \mathbb{R}^3$, où $I \subseteq \mathbb{R}$ est un intervalle non vide.**

**Définition 4.3.1.** *Soit $\alpha : I \to \mathbb{R}^3$ une courbe birégulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Soit ${}_\ell\alpha : I' \to \mathbb{R}^3$ la courbe reparamétrée par longueur d'arc associée à $\alpha$. On définit les applications $\mathbf{n}_{\ell\alpha}, \mathbf{b}_{\ell\alpha} : I'^{\circ} \setminus D_{\ell\alpha} \to \mathbb{R}^3$ et $\kappa_{\ell\alpha} : I'^{\circ} \setminus D_{\ell\alpha} \to \mathbb{R}_{>0}$ par*

$$\mathbf{n}_{\ell\alpha}(s) = \frac{{}_\ell\alpha''(s)}{\|{}_\ell\alpha''(s)\|}, \qquad \mathbf{b}_{\ell\alpha}(s) = \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s) \qquad \text{et} \qquad \kappa_{\ell\alpha}(s) = \|{}_\ell\alpha''(s)\|,$$

*pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$, où $\wedge$ désigne le produit vectoriel de $\mathbb{R}^3$. On appelle $\mathbf{n}_{\ell\alpha}(s)$ et $\mathbf{b}_{\ell\alpha}(s)$ le **vecteur normal** et le le **vecteur binormal** à $\alpha$ en la longueur d'arc $s$, respectivement. En plus, on appelle $\kappa_{\ell\alpha}(s)$ la **courbure** de $\alpha$ en la longueur d'arc $s \in I'^{\circ} \setminus D_{\ell\alpha}$.*

Noter que, dans la définition précédente, si $\alpha$ est de classe $\mathrm{C}^3$ par morceaux, alors les applications $\mathbf{n}_{\ell\alpha}$, $\mathbf{b}_{\ell\alpha}$ et $\kappa_{\ell\alpha}$ sont de classe $\mathrm{C}^1$. Noter aussi que les définitions de courbure et du vecteur normal nous disent que

$$\mathbf{t}'_{\ell\alpha}(s) = \kappa_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s), \tag{4.3.1}$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$.

**Lemme 4.3.2.** *Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $\mathrm{C}^3$ par morceaux birégulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Soit ${}_\ell\alpha : I' \to \mathbb{R}^3$ la courbe reparamétrée par longueur d'arc associée à $\alpha$. Alors $\{\mathbf{t}_{\ell\alpha}(s), \mathbf{n}_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s)\}$ est une base orthonormée de $\mathbb{R}^3$ pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$ pour le produit scalaire usuel.*

***Preuve.*** Par définition, $\mathbf{t}_{\ell\alpha}(s)$ et $\mathbf{n}_{\ell\alpha}(s)$ ont norme 1 pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$. En outre, l'identité $||\mathbf{t}_{\ell\alpha}(s)||^2 = 1$ pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$ nous dit que

$$\kappa_{\ell\alpha}(s)\langle \mathbf{n}_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\rangle = \langle \mathbf{t}'_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\rangle = 0$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$. Comme $\alpha$ est birégulière, $\kappa_{\ell\alpha}(s) \neq 0$ pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$, ce qui implique que

$$\langle \mathbf{n}_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\rangle = 0 \tag{4.3.2}$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$.

La définition de $\mathbf{b}_{\ell\alpha}(s)$ ainsi que les propriétés (0.0.1) et (4.3.2) nous disent que $\mathbf{b}_{\ell\alpha}(s)$ est orthogonal à $\mathbf{t}_{\ell\alpha}(s)$ et $\mathbf{n}_{\ell\alpha}(s)$, pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$. En outre, comme

$$\begin{aligned}\big\langle \mathbf{b}_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s)\big\rangle &= \big\langle \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s)\big\rangle = \Big\langle \mathbf{n}_{\ell\alpha}(s) \wedge \big(\mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s)\big), \mathbf{t}_{\ell\alpha}(s)\Big\rangle \\ &= \big\langle \mathbf{t}_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\big\rangle = \big\langle \mathbf{t}_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\big\rangle = 1,\end{aligned}$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$, où l'on a utilisé (0.0.2) dans la troisième égalité et (0.0.1) dans la troisième égalité. En conséquence, $\mathbf{b}_{\ell\alpha}(s)$ a norme 1 pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$. □

**Proposition 4.3.3.** *Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $\mathrm{C}^3$ par morceaux birégulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Soit ${}_\ell\alpha : I' \to \mathbb{R}^3$ la courbe reparamétrée par longueur d'arc associée à $\alpha$. Alors, $\mathbf{b}'_{\ell\alpha}(s)$ est orthogonal à $\mathbf{t}_{\ell\alpha}(s)$ et à $\mathbf{b}_{\ell\alpha}(s)$ pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$.*

***Preuve.*** Comme $\mathbf{b}_{\ell\alpha}(s) = \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s)$ pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$, on voit bien que

$$\begin{aligned}\mathbf{b}'_{\ell\alpha}(s) &= \mathbf{t}'_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s) + \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}'_{\ell\alpha}(s) \\ &= \kappa_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s) + \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}'_{\ell\alpha}(s) = \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}'_{\ell\alpha}(s),\end{aligned}$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$. Noter aussi que l'identité $||\mathbf{b}_{\ell\alpha}(s)||^2 = 1$ pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$ nous dit que

$$\langle \mathbf{b}'_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s)\rangle = 0$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$. En outre, l'identité (0.0.1) nous dit que

$$\langle \mathbf{b}'_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\rangle = \langle \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}'_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s)\rangle = \langle \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{t}_{\ell\alpha}(s), \mathbf{n}'_{\ell\alpha}(s)\rangle = 0,$$

pour tout $s \in I'^{\circ} \setminus D_{\ell\alpha}$, comme on voulait démontrer. □

**Définition 4.3.4.** *Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $C^3$ par morceaux birégulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Soit ${}_\ell\alpha : I' \to \mathbb{R}^3$ la courbe reparamétrée par longueur d'arc associée à $\alpha$. On définit l'application $\kappa_{\ell\alpha} : I'^\circ \setminus D_{\ell\alpha} \to \mathbb{R}$ par*

$$\tau_{\ell\alpha}(s) = -\langle \mathbf{b}'_{\ell\alpha}(s), \mathbf{n}_{\ell\alpha}(s) \rangle$$

*pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$. D'après la Proposition 4.3.3,*

$$\mathbf{b}'_{\ell\alpha}(s) = -\tau_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s), \tag{4.3.3}$$

*pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$. On appelle $\tau_{\ell\alpha}(s)$ la* ***torsion*** *de $\alpha$ en la longueur d'arc $s \in I'^\circ \setminus D_{\ell\alpha}$.*

**Proposition 4.3.5.** *Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $C^3$ par morceaux birégulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Soit ${}_\ell\alpha : I' \to \mathbb{R}^3$ la courbe reparamétrée par longueur d'arc associée à $\alpha$. Alors,*

$$\mathbf{n}'_{\ell\alpha}(s) = -\kappa_{\ell\alpha}(s)\mathbf{t}_{\ell\alpha}(s) + \tau_{\ell\alpha}(s)\mathbf{b}_{\ell\alpha}(s) \tag{4.3.4}$$

*pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$.*

***Preuve.*** Noter que l'identité $||\mathbf{n}_{\ell\alpha}(s)||^2 = 1$ pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$ nous dit que

$$\langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{n}_{\ell\alpha}(s) \rangle = 0$$

pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$. Alors, le Lemme 4.3.2 nous dit que

$$\mathbf{n}'_{\ell\alpha}(s) = \langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s) \rangle \mathbf{t}_{\ell\alpha}(s) + \langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s) \rangle \mathbf{b}_{\ell\alpha}(s)$$

pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$. Or, l'identité $\langle \mathbf{n}_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s) \rangle = 0$ et $\langle \mathbf{n}_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s) \rangle = 0$ pour tous $s \in I'^\circ \setminus D_{\ell\alpha}$ nous disent que

$$\langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s) \rangle + \langle \mathbf{n}_{\ell\alpha}(s), \mathbf{t}'_{\ell\alpha}(s) \rangle = 0$$

et

$$\langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s) \rangle + \langle \mathbf{n}_{\ell\alpha}(s), \mathbf{b}'_{\ell\alpha}(s) \rangle = 0$$

pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$, ce qui implique que

$$\langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{t}_{\ell\alpha}(s) \rangle = -\langle \mathbf{n}_{\ell\alpha}(s), \mathbf{t}'_{\ell\alpha}(s) \rangle = -\langle \mathbf{n}_{\ell\alpha}(s), \kappa_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s) \rangle = -\kappa_{\ell\alpha}(s)$$

et

$$\langle \mathbf{n}'_{\ell\alpha}(s), \mathbf{b}_{\ell\alpha}(s) \rangle = -\langle \mathbf{n}_{\ell\alpha}(s), \mathbf{b}'_{\ell\alpha}(s) \rangle = \langle \mathbf{n}_{\ell\alpha}(s), \tau_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s) \rangle = \tau_{\ell\alpha}(s)$$

pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$. Le résultat est démontré. □

Les identités (4.3.1), (4.3.3) et (4.3.4) sont appelées **identités de Frenet-Serret**.

Le calcul de la courbure et la torsion d'une courbe à partir de la Définition 4.3.1 est en principe compliqué, vu qu'elles reposent sur un reparamétrage par longueur d'arc. On a par contre les expressions suivantes, qui ne nécessitent pas cette dernière condition.

**Proposition 4.3.6.** *Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $C^3$ par morceaux birégulière, où $I \subseteq \mathbb{R}$ est un intervalle non vide. Soit $g : I' \to I$ l'unique reparamétrage positif de $\alpha$ tel que ${}_\ell\alpha = \alpha \circ g$ est paramétrée par longueur d'arc et $I'^- = 0$, donné dans la Proposition 4.1.7. On définit les applications $\kappa_\alpha : I^\circ \setminus D_\alpha \to \mathbb{R}_{>0}$ et $\tau_\alpha : I^\circ \setminus D_\alpha \to \mathbb{R}$ par*

$$\kappa_\alpha(t) = \frac{\|\alpha'(t) \wedge \alpha''(t)\|}{\|\alpha'(t)\|^3} \text{ et } \tau_\alpha(t) = \frac{\langle \alpha'(t) \wedge \alpha''(t), \alpha'''(t) \rangle}{\|\alpha'(t) \wedge \alpha''(t)\|^2}. \tag{4.3.5}$$

*pour tout $t \in I^\circ \setminus D_\alpha$. Alors, $\kappa_\alpha(g(s)) = \kappa_{\ell\alpha}(s)$ et $\tau_\alpha(g(s)) = \tau_{\ell\alpha}(s)$ pour tout $s \in I'^\circ \setminus D_{\ell\alpha}$.*

***Preuve.*** Soit $h$ l'application réciproque de $g$ définie dans l'équation (4.1.1). Alors,

$$h'(t) = \|\alpha'(t)\|, h''(t) = \frac{\langle \alpha'(t), \alpha''(t) \rangle}{\|\alpha'(t)\|} \text{ et } h'''(t) = \frac{\langle \alpha'(t), \alpha'''(t) \rangle + \|\alpha''(t)\|^2}{\|\alpha'(t)\|} - \frac{\langle \alpha'(t), \alpha''(t) \rangle^2}{\|\alpha'(t)\|^3},$$

pour tout $t \in I^\circ \setminus D_\alpha$, où l'on a utilisé le Théorème 2.1.17 et que l'application $\| \ \| : \mathbb{R}^3 \setminus \{\mathbf{0}_{\mathbb{R}^3}\} \to \mathbb{R}_{>0}$ est différentiable et sa différentielle est $d\| \ \|(x)(v) = \langle x, v \rangle / \|x\|$, pour tout $x \in \mathbb{R}^3 \setminus \{\mathbf{0}_{\mathbb{R}^3}\}$ et $v \in \mathbb{R}^3$. En outre, comme $g$ et $h$ sont des applications réciproques, on a vu dans (4.1.2) que si l'on dérive l'identité $h(g(s)) = s$ pour tout $s \in s \in I'^{\circ} \setminus D_{\alpha\circ g}$ on obtient que $h'(g(s))g'(s) = 1$ pour tout $s \in s \in I'^{\circ} \setminus D_{\alpha\circ g}$, ce qui nous dit que

$$g'(s) = \frac{1}{\|\alpha'(g(s))\|},$$

pour tout $s \in I'^{\circ} \setminus D_g$. De la même façon, si l'on dérive $h'(g(s))g'(s) = 1$ encore une fois on trouve $h''(g(s))g'(s)^2 + h'(g(s))g''(s) = 0$ pour tout $s \in s \in I'^{\circ} \setminus D_{\alpha\circ g}$, ce qui implique que

$$g''(s) = -\frac{\Big\langle \alpha'\big(g(s)\big), \alpha''\big(g(s)\big)\Big\rangle}{\big\|\alpha'\big(g(s)\big)\big\|^4}, \tag{4.3.6}$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$. Finalement, si l'on dérive $h''(g(s))g'(s)^2 + h'(g(s))g''(s) = 0$ encore une fois on trouve $h'''(g(s))g'(s)^3 + 3h''(g(s))g'(s)g''(s) + h'(s)g'''(s) = 0$ pour tout $s \in s \in I'^{\circ} \setminus D_{\alpha\circ g}$, ce qui implique que

$$g'''(s) = \frac{4\Big\langle \alpha'\big(g(s)\big), \alpha''\big(g(s)\big)\Big\rangle^2}{\big\|\alpha'\big(g(s)\big)\big\|^5} - \frac{\big\|\alpha''\big(g(s)\big)\big\|^2 + \Big\langle \alpha'\big(g(s)\big), \alpha'''\big(g(s)\big)\Big\rangle}{\big\|\alpha'\big(g(s)\big)\big\|^4},$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$.

Si l'on applique le Théorème 2.1.17 à l'identité ${}_{\ell}\alpha = \alpha \circ g$, on trouve que

$${}_{\ell}\alpha'(s) = \alpha'\big(g(s)\big)g'(s) = \frac{\alpha'\big(g(s)\big)}{\big\|\alpha'\big(g(s)\big)\big\|}, \tag{4.3.7}$$

ainsi que

$${}_{\ell}\alpha''(s) = \alpha''\big(g(s)\big)g'(s)^2 + \alpha'\big(g(s)\big)g''(s)$$

et

$${}_{\ell}\alpha'''(s) = \alpha'''\big(g(s)\big)g'(s)^3 + 3\alpha''\big(g(s)\big)g'(s)g''(s) + \alpha'\big(g(s)\big)g'''(s)$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$. En conséquence, (4.1.2) et (4.3.6) impliquent que

$${}_{\ell}\alpha''(s) = \frac{1}{\big\|\alpha'\big(g(s)\big)\big\|^2}\left(\alpha''\big(g(s)\big) - \alpha'\big(g(s)\big)\frac{\big\langle \alpha'\big(g(s)\big), \alpha''\big(g(s)\big)\big\rangle}{\big\|\alpha'\big(g(s)\big)\big\|^2}\right) \tag{4.3.8}$$

et

$$\begin{aligned}{}_{\ell}\alpha'''(s) = \frac{1}{\big\|\alpha'\big(g(s)\big)\big\|^3}\Bigg(&\alpha'''\big(g(s)\big) - 3\alpha''\big(g(s)\big)\frac{\Big\langle \alpha'\big(g(s)\big), \alpha''\big(g(s)\big)\Big\rangle}{\Big\|\alpha'\big(g(s)\big)\Big\|^2} \\ &+ \alpha'\big(g(s)\big)\left\{\frac{4\Big\langle \alpha'\big(g(s)\big), \alpha''\big(g(s)\big)\Big\rangle^2}{\big\|\alpha'\big(g(s)\big)\big\|^5} - \frac{\big\|\alpha''\big(g(s)\big)\big\|^2 + \Big\langle \alpha'\big(g(s)\big), \alpha'''\big(g(s)\big)\Big\rangle}{\big\|\alpha'\big(g(s)\big)\big\|^4}\right\}\Bigg),\end{aligned} \tag{4.3.9}$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$. Par les propriétés du produit vectoriel, (4.3.7), (4.3.8) et (4.3.9), on trouve que

$$\frac{\big\|{}_{\ell}\alpha'(s) \wedge {}_{\ell}\alpha''(s)\big\|}{\big\|{}_{\ell}\alpha'(s)\big\|^3} = \frac{\big\|\alpha'\big(g(s)\big) \wedge \alpha''\big(g(s)\big)\big\|}{\big\|\alpha'\big(g(s)\big)\big\|^3}$$

et

$$\frac{\big\langle \alpha'(t) \wedge \alpha''(t), \alpha'''(t)\big\rangle}{\big\|\alpha'(t) \wedge \alpha''(t)\big\|^2} = \frac{\big\langle \alpha'\big(g(s)\big) \wedge \alpha''\big(g(s)\big), \alpha'''\big(g(s)\big)\big\rangle}{\big\|\alpha'\big(g(s)\big) \wedge \alpha''\big(g(s)\big)\big\|^2}.$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$. Il suffit donc de démontrer les identités (4.3.5) pour le cas où $\alpha$ est paramétrée par longueur d'arc, *i.e.* $\alpha = {}_{\ell}\alpha$. Or, par définition on a que

$$\kappa_{\ell\alpha}(s)^2 = \left\|{}_{\ell}\alpha''(s)\right\|^2 = \frac{\left\|{}_{\ell}\alpha'(s) \wedge {}_{\ell}\alpha''(s)\right\|}{\left\|{}_{\ell}\alpha'(s)\right\|^3},$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$, où l'on a utilise (0.0.3) dans la dernière égalité. Cela montre que la première égalité dans (4.3.5) est vérifiée.

Il nous reste à montrer que la dernière égalité dans (4.3.5) est vérifiée. Comme

$${}_{\ell}\alpha'''(s) = \kappa'_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s) + \kappa_{\ell\alpha}(s)\mathbf{n}'_{\ell\alpha}(s),$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$, et comme $\kappa_{\ell\alpha}(s) \neq 0$, on trouve que

$$\mathbf{n}'_{\ell\alpha}(s) = \frac{{}_{\ell}\alpha'''(s) - \kappa'_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s)}{\kappa_{\ell\alpha}(s)},$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$. En conséquence, $\mathbf{b}_{\ell\alpha}(s) = \mathbf{t}_{\ell\alpha}(s) \wedge \mathbf{n}_{\ell\alpha}(s) = {}_{\ell}\alpha'(s) \wedge \mathbf{n}_{\ell\alpha}(s)$ implique que

$$\mathbf{b}'_{\ell\alpha}(s) = {}_{\ell}\alpha''(s) \wedge \mathbf{n}_{\ell\alpha}(s) + {}_{\ell}\alpha'(s) \wedge \mathbf{n}'_{\ell\alpha}(s) = \frac{{}_{\ell}\alpha'(s) \wedge {}_{\ell}\alpha'''(s)}{\kappa_{\ell\alpha}(s)} - \kappa'_{\ell\alpha}(s)\frac{{}_{\ell}\alpha'(s) \wedge \mathbf{n}_{\ell\alpha}(s)}{\kappa_{\ell\alpha}(s)},$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$. En outre, $\mathbf{b}'_{\ell\alpha}(s) = -\tau_{\ell\alpha}(s)\mathbf{n}_{\ell\alpha}(s)$, nous dit que

$$\begin{aligned}\tau_{\ell\alpha}(s) = -\langle \mathbf{b}'_{\ell\alpha}(s), \mathbf{n}_{\ell\alpha}(s)\rangle &= -\frac{\left\langle {}_{\ell}\alpha'(s) \wedge {}_{\ell}\alpha'''(s), {}_{\ell}\alpha''(s)\right\rangle}{\kappa_{\ell\alpha}(s)^2} = \frac{\left\langle {}_{\ell}\alpha'(s) \wedge {}_{\ell}\alpha''(s), {}_{\ell}\alpha'''(s)\right\rangle}{\kappa_{\ell\alpha}(s)^2} \\ &= \frac{\left\langle {}_{\ell}\alpha'(t) \wedge {}_{\ell}\alpha''(t), {}_{\ell}\alpha'''(t)\right\rangle}{\left\|{}_{\ell}\alpha'(t) \wedge {}_{\ell}\alpha''(t)\right\|^2},\end{aligned}$$

pour tout $s \in I'^{\circ} \setminus D_{\alpha\circ g}$, où l'on a utilisé (0.0.1) dans la troisième égalité. $\square$

# Références


[1] Serge Lang, *Undergraduate analysis*, 2nd ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1997. ↑i, 51, 64

[2] Sorin Rădulescu and Marius Rădulescu, *Local inversion theorems without assuming continuous differentiability*, J. Math. Anal. Appl. **138** (1989), no. 2, 581–590. ↑66