Notes de cours

MATH-105(a)
Analyse avancée II

(Section MA)

Boris Buffoni, Fabio Nobile

2020-2021

Dernière mise à jour : 20 mars 2021

Table des matières

Notations

$\mathbb{R}_+ = \{x \in \mathbb{R} : \ x \geq 0\}$

$\mathbb{R}_+^* = \{x \in \mathbb{R} : \ x > 0\}$

$\overline{\mathbb{R}} = \mathbb{R} \cup \{- + \infty, +\infty\}$

$\mathbb{N} = \{0, 1, 2, \ldots\}$

$\mathbb{N}^* = \{1, 2, \ldots\}$

Chapitre 0

Intégrales Généralisées

Ce chapitre reprend le dernier sujet du cours d'Analyse Avancée I, notamment la construction de l'intégrale définie $\int_a^b f(x)dx$ d'une fonction $f : [a, b] \to \mathbb{R}$ continue où continue par morceaux sur un intervalle borné et fermé. On rappelle qu'une fonction $f : [a, b] \to \mathbb{R}$ avec $a < b$ dans $\mathbb{R}$ est dite *continue par morceaux* si $\lim_{t\to x^+} f(t)$ existe (sous-entendu, dans $\mathbb{R}$) pour tout $x \in [a, b[$, $\lim_{t\to x^-} f(t)$ existe pour tout $x \in]a, b]$ et f est continue en x pour tout $x \in [a, b]$ avec au plus un nombre fini d'exceptions. Cette définition se généralise à un intervalle I quelconque avec une infinité de points : une fonction $f : I \to \mathbb{R}$ est dite continue par morceaux (c.p.m.) si f est continue par morceaux sur tout intervalle $[a, b] \subset I$ avec $a < b$.

On se pose ici la question de comment généraliser la définition de l'intégrale $\int_a^b f(x)dx$ dans le cas où l'intervalle d'intégration est borné mais pas fermé, et la fonction f n'est pas définie en a ou en b, comme dans les exemples suivantes :

$$\int_0^1 \sin\left(\frac{1}{x}\right) dx, \qquad \int_0^1 \ln(x)dx, \qquad \int_0^1 \frac{1}{x}dx,$$

ou encore, comment généraliser la définition de l'intégrale sur un intervalle non borné, par exemple :

$$\int_0^\infty \sin(x)dx, \qquad \int_0^\infty \sin(x^2)dx, \qquad \int_0^\infty e^{-x}dx, \qquad \int_1^\infty \frac{1}{x^2}dx, \qquad \int_{-\infty}^\infty \frac{1}{1+x^2}dx,$$

où on entendra toujours ∞ par $+\infty$, dans ce chapitre.

0.1 Intégrale généralisée sur un intervalle borné

On commence par définir l'intégrale généralisée sur un intervalle borné et "demi-ouvert" (ouvert à droite et fermé à gauche ou bien ouvert à gauche et fermé à droite).

Définition 0.1. *Pour $a < b$ dans $\mathbb{R}$, soit $f : [a, b[\to \mathbb{R}$ c.p.m. et $F(x) = \int_a^x f(t)dt$, $x \in [a, b[$. Si $\lim_{x\to b^-} F(x)$ existe, on dit que $\int_a^b f(x)dx$ existe (ou converge) et on pose*

$$\int_a^b f(x)dx = \lim_{x\to b^-} F(x).$$

Si $\lim_{x\to b^-} F(x)$ *n'existe pas, on dit que* $\int_a^b f(x)dx$ *diverge.*

De façon similaire, soit $f :]a,b] \to \mathbb{R}$ *c.p.m., et* $F(x) = \int_x^b f(t)dt$*,* $x \in]a,b]$*. Si* $\lim_{x\to a^+} F(x)$ *existe, on dit que* $\int_a^b f(x)dx$ *existe (ou converge) et on pose*

$$\int_a^b f(x)dx = \lim_{x\to a^+} F(x).$$

Autrement on dit que l'intégrale $\int_a^b f(x)dx$ *diverge.*

Exemple 0.2. *On étudie l'existence de l'intégrale généralisée*

$$\int_0^1 \frac{1}{x^\alpha}dx$$

Pour $\alpha \in \mathbb{R}$*.*

Soit $F(x) = \int_x^1 \frac{1}{t^\alpha}dt$*,* $x \in]0,1]$ *qui est bien définie car la fonction* $f(t) = t^{-\alpha}$ *est continues sur* $[x,1]$ *pour tout* $x \in]0,1]$*. Pour* $\alpha \neq 1$ *on a*

$$F(x) = \left.\frac{t^{-\alpha+1}}{-\alpha+1}\right|_x^1 = \frac{1}{1-\alpha} - \frac{x^{1-\alpha}}{1-\alpha}$$

tandis que pour $\alpha = 1$ *on a*

$$F(x) = \ln t\big|_x^1 = -\ln x.$$

Donc

$$\lim_{x\to 0^+} F(x) = \begin{cases} \frac{1}{1-\alpha} & \alpha < 1 \\ +\infty & \alpha \geq 1 \end{cases}$$

et l'intégrale généralisée $\int_0^1 \frac{1}{x^\alpha}dx$ *existe pour tout* $\alpha < 1$*.*

On vérifie facilement que si la fonction f admet une extension par continuité sur $[a,b]$ alors l'intégrale généralisé de f existe et coïncide avec l'intégrale sur l'intervalle fermé $[a,b]$ de l'extension par continuité de la fonction. Plus précisément, on a le résultat suivant, dont la preuve est laissée comme exercice.

Lemme 0.3. *Pour* $a < b$ *dans* $\mathbb{R}$*, soit* $f : [a,b[\to \mathbb{R}$ *c.p.m. telle que* $\lim_{x\to b^-} f(x)$ *existe et définissons la fonction c.p.m. sur* $[a,b]$

$$f_b(x) = \begin{cases} f(x), & x \in [a,b[, \\ \lim_{t\to b^-} f(t), & x = b. \end{cases}$$

Alors, l'intégrale généralisée $\int_a^b f(x)dx$ *existe et* $\int_a^b f(x)dx = \int_a^b f_b(x)dx$*.*

On a le même résultat pour une fonction $f :]a,b] \to \mathbb{R}$ *c.p.m. telle que* $\lim_{x\to a^+} f(x)$ *existe.*

On considère maintenant le cas d'un intervalle ouvert (à gauche et à droite).

Définition 0.4. *Soit* $f :]a,b[\to \mathbb{R}$ *c.p.m.,* $a < b$*. Pour* $c \in]a,b[$*, si* $\int_a^c f(x)dx$ *et* $\int_c^b f(x)dx$ *existent, alors on pose* $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$*, auquel cas* $\int_a^b f(x)dx$ *est dite exister ou converger, sinon* $\int_a^b f(x)dx$ *est dite diverger.*

Il est facile de montrer que l'existence (ou non) de $\int_a^b f(x)dx$ ne dépend pas du choix de $c \in]a,b[$. En fait, soit $\tilde{c} \in]a,b[$, $\tilde{c} \neq c$. Alors

$$\int_a^{\tilde{c}} f(x)dx = \int_a^c f(x)dx + \int_c^{\tilde{c}} f(x)dx$$

existe ssi $\int_a^c f(x)dx$ existe, car f est c.p.m. sur $[c,\tilde{c}] \cup [\tilde{c},c]$. De même,

$$\int_{\tilde{c}}^b f(x)dx = \int_{\tilde{c}}^c f(x)dx + \int_c^b f(x)dx$$

existe ssi $\int_c^b f(x)dx$ existe, et

$$\int_a^{\tilde{c}} f(x)dx + \int_{\tilde{c}}^b f(x)dx = \int_a^c f(x)dx + \int_c^{\tilde{c}} f(x)dx + \int_{\tilde{c}}^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx.$$

Exemple 0.5. *Dire si l'intégrale suivante*

$$\int_{-\pi/2}^{\pi/2} \tan(x)dx$$

existe ou non.

On est tenté de calculer

$$\begin{aligned}\lim_{x\to\pi/2^-} \int_{-x}^x \tan(t)dt &= \lim_{x\to\pi/2^-} \int_{-x}^x (-\ln(\cos t))'dt = \lim_{x\to\pi/2^-} \Big(-\ln(\cos x) + \ln(\cos(-x)) \Big) \\ &= \lim_{x\to\pi/2^-} \ln\left(\frac{\cos(-x)}{\cos x}\right) = \lim_{x\to\pi/2^-} \ln(1) = 0\end{aligned}$$

mais

$$\begin{aligned}\lim_{x\to\pi/2^-} \int_0^x \tan(t)dt &= \lim_{x\to\pi/2^-} \ln\left(\frac{\cos(0)}{\cos x}\right) = +\infty \\ \textit{et } \lim_{x\to\pi/2^-} \int_{-x}^0 \tan(t)dt &= -\infty.\end{aligned}$$

Donc $\int_{-\pi/2}^{\pi/2} \tan(x)dx$ *diverge. En effet,* $\displaystyle\lim_{\substack{x_1\to-\pi/2^+ \\ x_2\to\pi/2^-}} \int_{x_1}^{x_2} \tan(x)dx$ *dépend de comment* x_1 *et* x_2 *tendent respectivement vers* $-\pi/2$ *et* $\pi/2$. *Prendre par exemple* $\lim_{\epsilon\to0^+} \int_{-\pi/2+\epsilon}^{\pi/2-\epsilon^2} \tan(x)dx = +\infty$ *et* $\lim_{\epsilon\to0^+} \int_{-\pi/2+\epsilon^2}^{\pi/2-\epsilon} \tan(x)dx = -\infty$.

L'intégrale généralisée a les propriétés suivantes :

Lemme 0.6 (Propriétés de l'intégrale généralisée)**.** *Soit* $a < b$ *dans* $\mathbb{R}$, I *de la forme* $[a,b[$, $]a,b]$ *ou* $]a,b[$, *et des fonctions c.p.m.* $f,g : I \to \mathbb{R}$. *Alors*

— $\int_a^b(\alpha f(x) + \beta g(x))dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$, $\forall \alpha,\beta \in \mathbb{R}$, *l'intégrale généralisée à gauche étant convergente si les deux intégrales généralisées à droite convergent.*

— $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$, $\forall c \in]a,b[$, *l'intégrale généralisée à gauche étant convergente ssi chacune des deux intégrales à droite converge au cas où elle est généralisée.*

— *si* $f(x) \leq g(x)$, $\forall x \in I$ *et* $\int_a^b f(x)dx$ *et* $\int_a^b g(x)dx$ *existent, alors* $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.

Pour établir si une intégrale généralisée existe, le critère de comparaison suivant est souvent très utile.

Lemme 0.7 (Critère de comparaison)**.** *Soit* $f, g : [a,b[\to \mathbb{R}$ *deux fonctions c.p.m. et supposons qu'il existe* $c \in [a,b[$ *tel que*

$$0 \leq f(x) \leq g(x) \quad \forall x \in [c,b[.$$

— *Si* $\int_a^b g(x)dx$ *existe, alors* $\int_a^b f(x)dx$ *existe aussi ;*

— *si* $\int_a^b f(x)dx$ *diverge, alors* $\int_a^b g(x)dx$ *diverge aussi.*

Démonstration. Si $\int_a^b g(x)dx$ existe, alors $\int_c^b g(x)dx$ existe aussi. De plus, pour tout $x \in [c,b[$

$$F(x) = \int_c^x f(t)dt \leq \int_c^x g(t)dt \leq \int_c^b g(t)dt < +\infty$$

Puisque F est une fonction croissante et bornée supérieurement sur $[c,b[$ on a que $\lim_{x\to b^-} F(x)$ existe. Ainsi $\int_c^b f(t)dt$ converge et

$$\int_a^b f(t)dt = \lim_{x\to b^-} \int_a^x f(t)dt = \int_a^c f(t)dt + \lim_{x\to b^-} \int_c^x f(t)dt$$

existe aussi. La seconde affirmation est la contraposée de la première. □

Corollaire 0.8. *Soit* $f : [a,b[\to \mathbb{R}$ *c.p.m., non négative et bornée. Alors* $\int_a^b f(x)dx$ *existe.*

Démonstration. Il suit immédiatement du fait que $0 \leq f(x) \leq M < +\infty, \quad \forall x \in [a,b[$ et $\int_a^b M dx$ converge. □

Des versions analogues du Lemme 0.7 et Corollaire 0.8 existent pour $f, g :]a,b] \to \mathbb{R}$ ou bien $f, g :]a,b[\to \mathbb{R}$.

Exemple 0.9. *Grâce au Corollaire 0.8, on montre facilement que* $\displaystyle\int_0^1 \frac{(\sin\frac{1}{x})^2}{\sqrt{x}}dx$ *converge.*

En effet, $0 \leq f(x) = \dfrac{(\sin\frac{1}{x})^2}{\sqrt{x}} \leq \dfrac{1}{\sqrt{x}}$ *et* $\int_0^1 \frac{1}{\sqrt{x}}dx$ *converge.*

0.2 Intégrale généralisée absolument convergente

Définition 0.10. *Soit* I *un intervalle de la forme* $[a,b[$, $]a,b]$ *ou* $]a,b[$ *et* $f : I \to \mathbb{R}$ *c.p.m. On dit que l'intégrale généralisée est absolument convergente si l'intégrale* $\int_a^b |f(x)|dx$ *existe.*

Théorème 0.11. *Si l'intégrale $\int_a^b f(x)dx$ est absolument convergente, alors elle existe.*

Démonstration. Soit $f_+(x) = \max\{f(x), 0\}$ et $f_-(x) = -\min\{f(x), 0\}$. On a $f(x) = f_+(x) - f_-(x)$, $|f(x)| = f_+(x) + f_-(x)$,

$$0 \leq f_+(x) \leq |f(x)|, \quad 0 \leq f_-(x) \leq |f(x)|.$$

Donc, par le critère de comparaison, $\int_a^b f_+(x)dx$ et $\int_a^b f_-(x)dx$ existent, et $\int_a^b f(x)dx = \int_a^b (f_+(x) - f_-(x))dx$ existe aussi. □

Corollaire 0.12. *Soit $f : I \to \mathbb{R}$ c.p.m. et* bornée*, où I est un intervalle* borné *comme ci-dessus. Alors $\int_a^b f(x)dx$ converge absolument et par conséquent, existe.*

Démonstration. On a $0 \leq |f(x)| \leq M < +\infty, \quad \forall x \in I$. Comme $\int_a^b M dx$ converge, $\int_a^b |f(x)|dx$ converge aussi et $\int_a^b f(x)dx$ converge absolument. □

Exercice 0.13. *Montrer que les intégrales généralisées*

$$\int_0^1 \sin\left(\frac{1}{x}\right) dx, \qquad \int_0^{2\pi} \frac{\sin x}{x^{3/2}} dx$$

convergent absolument.

Théorème 0.14. *Une condition suffisante pour que $\int_a^b f(x)dx$ converge absolument avec $f : [a, b[\to \mathbb{R}$ c.p.m. est qu'il existe $\alpha \in]-\infty, 1[$ tel que*

$$\lim_{x \to b^-} (b - x)^\alpha f(x) = l \in \mathbb{R}.$$

Si $\exists \alpha \geq 1$ tel que $\lim_{x \to b^-} (b - x)^\alpha f(x) = l \neq 0$, alors $\int_a^b f(x)dx$ diverge.

Démonstration. Soit $\alpha \in]-\infty, 1[$ tel que $\lim_{x \to b^-} (b - x)^\alpha f(x) = l \in \mathbb{R}$. Alors, pour tout $\epsilon > 0$ il existe $\delta_\epsilon \in]0, b - a[$ tel que

$$\forall x \in [b - \delta_\epsilon, b[, \quad (l - \epsilon)(b - x)^{-\alpha} < f(x) < (l + \epsilon)(b - x)^{-\alpha}.$$

Si $\alpha < 1$, alors $\int_a^b (b - x)^{-\alpha} dx$ converge et, puisque $0 \leq |f(x)| < (|l| + \epsilon)(b - x)^{-\alpha}$ pour $x \in [b - \delta_\epsilon, b[$, on en déduit grâce au Lemme 0.7 que $\int_a^b |f(x)|dx$ converge, donc $\int_a^b f(x)dx$ converge absolument.

Si, par contre, $\alpha \geq 1$ et $l \neq 0$, en prenant $\epsilon \in]0, |l|[$ on a que $\int_{b-\delta_\epsilon}^b (b - x)^{-\alpha} dx$ diverge, f est non nul et de signe constant sur $[b - \delta_\epsilon, b[$ et $|f(x)| > (|l| - \epsilon)(b - x)^{-\alpha}, \quad \forall x \in [b - \delta_\epsilon, b[$. Grâce au Lemme 0.7 on déduit que $\int_{b-\delta_\epsilon}^b |f(x)|dx$ diverge, et $\int_{b-\delta_\epsilon}^b f(x)dx$ et $\int_a^b f(x)dx$ divergent aussi. □

Remarque 0.15. *On a présenté le Théorème 0.14 pour un intervalle I de la forme $[a,b[$. Il y a deux conditions suffisantes analogues en a pour I de la forme $]a,b]$. Si I est de la forme $]a,b[$, on choisit $c\in]a,b[$ et on étudie séparément l'intégrale généralisée sur $[c,b[$ et celle sur $]a,c]$.*

Exercice 0.16. *Étudier, en utilisant le critère de puissance du Théorème 0.14 si l'intégrale généralisée $\int_0^{\pi/2}\frac{1}{\sqrt{\tan x}}dx$ converge ou non.*

0.3 Intégrale généralisée sur un intervalle non borné

Soit I de la forme $[a,\infty[$, $]-\infty,a]$, $]a,\infty[$, $]-\infty,a[$ ou $]-\infty,\infty[$, et soit $f:I\to\mathbb{R}$ c.p.m. On définit l'intégrale généralisée de f sur I de façon similaire à ce qu'on a fait pour I borné.

Définition 0.17. *Si $I=[a,\infty[$ et $\lim_{x\to\infty}\int_a^x f(t)dt$ existe, on dit que l'intégrale généralisée $\int_a^\infty f(x)dx$ existe et on pose*

$$\int_a^\infty f(t)dt=\lim_{x\to\infty}F(x).$$

On dit que $\int_a^\infty f(x)dx$ est absolument convergente si $\int_a^\infty |f(x)|dx$ existe.

Si $I=]-\infty,a]$ et $\lim_{x\to-\infty}\int_x^a f(t)dt$ existe, on dit que l'intégrale généralisée $\int_{-\infty}^a f(x)dx$ existe et on pose

$$\int_{-\infty}^a f(t)dt=\lim_{x\to-\infty}\int_x^a f(t)dt.$$

On dit que $\int_{-\infty}^a f(x)dx$ est absolument convergente si $\int_\infty^a |f(x)|dx$ existe.

Soit I de la forme $]a,\infty[$, $]-\infty,a[$ ou $\mathbb{R}$, et soit $c\in I$. On dit que l'intégrale généralisée $\int_I f(t)dt$ existe si les intégrales généralisées de f sur $I_1=I\cap]-\infty,c]$ et sur $I_2=I\cap[c,\infty[$ existent les deux, auquel cas on pose

$$\int_I f(t)dt=\int_{I_1}f(t)dt+\int_{I_2}f(t)dt$$

(ceci ne dépend pas du choix de c dans I). On dit que $\int_I f(x)dx$ est absolument convergente si $\int_I |f(x)|dx$ converge.

Si une intégrale généralisée existe, on dit aussi qu'elle converge, sinon on dit qu'elle diverge.

Exemple 0.18. *On étudie l'existence de l'intégrale généralisée*

$$\int_1^\infty \frac{1}{x^\alpha}dx$$

pour $\alpha \in \mathbb{R}$.

On a

$$F(x) = \int_1^x \frac{1}{t^\alpha} dt = \left.\frac{t^{-\alpha+1}}{-\alpha+1}\right|_1^x = \frac{x^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha}, \qquad \alpha \neq 1$$

$$F(x) = \ln x \qquad \alpha = 1$$

Donc

$$\lim_{x\to\infty} F(x) = \begin{cases} \frac{1}{\alpha-1} < \infty, & \alpha > 1 \\ +\infty, & \alpha \in]-\infty, 1] \end{cases}$$

et l'intégrale généralisée $\int_1^\infty \frac{1}{x^\alpha} dx$ existe pour tout $\alpha > 1$.

Il est intéressant de comparer ce dernier exemple avec l'exemple 0.2. On voit que l'intégrale généralisée $\int_0^1 \frac{1}{x^\alpha} dx$ existe pour tout $\alpha < 1$ alors que l'intégrale généralisée $\int_1^\infty \frac{1}{x^\alpha} dx$ existe pour tout $\alpha > 1$.

L'intégrale généralisée sur un intervalle non borné a les mêmes propriétés que celui sur un intervalle borné. En particulier, étant donné un intervalle I non borné et des fonctions c.p.m. $f, g : I \to \mathbb{R}$.

- linéarité : $\int_I (\alpha f(x) + \beta g(x))dx = \alpha \int_I f(x)dx + \beta \int_I g(x)dx, \quad \forall \alpha, \beta \in \mathbb{R}$, l'intégrale généralisée à gauche étant convergente si les deux intégrales généralisées à droite convergent.
- $\int_I f(x)dx = \int_{I\cap]-\infty,c]} f(x)dx + \int_{I\cap[c,\infty[} f(x)dx, \quad \forall c \in I$, l'intégrale généralisée à gauche étant convergente ssi chacune des deux intégrales à droite converge au cas où elle est généralisée.
- relation d'ordre : si $f(x) \leq g(x)$ $\forall x \in I$ et $\int_I f(x)dx$ et $\int_I g(x)dx$ existent, alors $\int_I f(x)dx \leq \int_I g(x)dx$.

Les critères de comparaison énoncés au Lemme 0.7 et au Théorème 0.14 se généralisent aussi au cas d'un intervalle non borné.

Lemme 0.19 (Critères de comparaison – intervalles non bornés)**.**

- *Soit $f, g : [a, +\infty[\to \mathbb{R}$ c.p.m. et supposons qu'il existe $c \in [a, \infty[$ tel que*

$$0 \leq f(x) \leq g(x), \quad \forall x \in [c, +\infty[.$$

  Si $\int_a^\infty g(x)dx$ existe alors $\int_a^\infty f(x)dx$ existe. Si $\int_a^\infty f(x)dx$ diverge alors $\int_a^\infty g(x)dx$ diverge.
- *Soit $f : [a, +\infty[\to \mathbb{R}$ c.p.m.*

$$\text{Si } \exists \beta > 1 : \lim_{x\to\infty} x^\beta f(x) = l \in \mathbb{R} \quad \text{alors } \int_a^\infty f(x)dx \text{ converge absolument.}$$

$$\text{Si } \exists \beta \in]-\infty, 1] : \lim_{x\to\infty} x^\beta f(x) = l \neq 0 \quad \text{alors } \int_a^\infty f(x)dx \text{ diverge.}$$

Remarque 0.20. *Attention, comparez les conditions sur β avec les conditions du Théorème 0.14 sur α ! Il y a des résultats analogues pour I de la forme $]-\infty, a]$. Si I est de la forme $]a, \infty[$, $]-\infty, a[$ ou $]-\infty, \infty[$, on choisit $c \in I$ et on étudie séparément l'intégrale généralisée sur $I \cap [c, +\infty[$ et celle sur $I\cap]-\infty, c]$.*

Exercice 0.21. *Étudier à l'aide du Lemme 0.19 l'existence des intégrales suivantes*

$$\int_1^\infty x^\alpha e^{-x} dx,\ \alpha \in \mathbb{R}, \qquad \int_1^\infty \frac{1}{\sqrt{x}} \ln\left(1+\frac{1}{x}\right) dx.$$

On remarque qu'une intégrale généralisée peut être convergente mais pas absolument convergente, comme l'exemple suivant le montre.

Exemple 0.22. *On montre dans cet exemple que l'intégrale généralisée $\int_\pi^\infty \frac{\sin x}{x} dx$ est convergente mais pas absolument convergente.*

Soit $F(x) = \int_\pi^x \frac{\sin t}{t} dt$. On a

$$\begin{aligned} F(x) &= -\int_\pi^x \frac{\cos t}{t^2} dt + \left.\frac{-\cos t}{t}\right|_\pi^x \\ &= -\int_\pi^x \frac{\cos t}{t^2} dt + \frac{\cos \pi}{\pi} - \frac{\cos x}{x} \end{aligned}$$

Donc $\lim_{x\to\infty} F(x) = \lim_{x\to\infty} \underbrace{-\int_\pi^x \frac{\cos t}{t^2} dt}_{\substack{\text{absolument convergent} \\ \text{donc la limite existe}}} - \frac{1}{\pi} - \underbrace{\lim_{x\to\infty} \frac{\cos x}{x}}_{=0}$ *existe.*

En revanche, $\int_\pi^\infty \frac{\sin x}{x} dx$ ne converge pas absolument. En effet, pour tout $k \in \mathbb{N}$, $k > 1$

$$\begin{aligned} F(k\pi) &= \int_\pi^{k\pi} \frac{|\sin x|}{x} dx \\ &= \sum_{j=1}^{k-1} \int_{j\pi}^{(j+1)\pi} \frac{|\sin x|}{x} dx \\ &\geq \sum_{j=1}^{k-1} \frac{1}{(j+1)\pi} \underbrace{\int_{j\pi}^{(j+1)\pi} |\sin x| dx}_{=2} \\ &= \sum_{j=1}^{k-1} \frac{2}{(j+1)\pi} \xrightarrow{k\to\infty} +\infty. \end{aligned}$$

Remarque 0.23. *Si $\int_a^\infty f(x)dx$ converge et $\lim_{x\to\infty} f(x)$ existe, alors nécessairement $\lim_{x\to\infty} f(x) = 0$ (cf le critère du Lemme 0.19 avec $\beta = 0$). En revanche le fait que $\int_a^\infty f(x)dx$ converge absolument n'implique pas que f est bornée.*

Par exemple, considérons la fonction $\phi(x) = \max\{1 - |x|, 0\}$ et la fonction $f : [0, \infty[\to \mathbb{R}_+$ définie par

$$f(x) = \sum_{n=1}^\infty n\phi\left(n^3(x-n)\right)$$

qui est continue mais pas bornée sur $[0, \infty[$.

Puisque

$$\int_0^\infty n\phi\left(n^3(x-n)\right)dx = n^{-2}\int_{-n^4}^\infty \phi(y)dy = n^{-2}\int_{-1}^1 \phi(y)dy = n^{-2},$$

on a que pour tout $x \in [0, \infty[$

$$F(x) = \int_0^x |f(t)|dx \leq \int_0^{\lceil x\rceil} \leq \sum_{n=1}^{\lceil x\rceil}\int_0^\infty n\phi\left(n^3(x-n)\right)dx \leq \sum_{n=1}^\infty \frac{1}{n^2} < \infty$$

Donc, F *est non décroissante est bornée supérieurement et la limite* $\lim_{x\to\infty} F(x)$ *existe. Il en suit que l'intégrale généralisée* $\int_0^\infty f(x)dx$ *converge absolument, même si* $\limsup_{x\to\infty} f(x) = +\infty$.

Remarque 0.24. *Soit un intervalle* I *qui n'est pas simultanément borné fermé, et une fonction c.p.m.* $f : I \to [0, \infty[$. *Comme* $f \geq 0$, *la limite (ou chacune des deux limites) intervenant dans la définition d'intégrale généralisée tend vers un nombre réel* ≥ 0 *ou vers* $+\infty$. *On peut donc dans ce cas écrire* $\int_I f(x)dx < \infty$ *si l'intégrale généralisée converge et* $\int_I f(x)dx = \infty = +\infty$ *si elle diverge.*

0.4 Intégrales et séries numériques

L'intégrale généralisée peut être utilisé aussi pour étudier la convergence d'une série numérique $\sum_{n=1}^\infty a_n$. L'idée est de construir la fonction c.p.m. $f : [1, \infty[\to \mathbb{R}$ telle que $f(x) = a_n$ pour $x \in [n, n+1[$ ou bien la fonction c.p.m. $\tilde{f} : [1, \infty[\to \mathbb{R}$, telle que $\tilde{f}(x) = a_n$, pour $x \in [n-1, n[$. Alors, pour tout $N \geq 0$,

$$\sum_{n=1}^N a_n = \sum_{n=1}^N \int_n^{n+1} a_n dx = \int_1^{N+1} f(x)dx = \int_0^N \tilde{f}(x)dx$$

et on peut utiliser les critères de comparaison pour les intégrales généralisées afin d'établir la convergence de la série.

Exemple 0.25. *On veut étudier la convergence de la série numérique* $S = \sum_{n=1}^\infty \frac{1}{n^2}$ *pour* $N \to \infty$.

Soit $f : [1, \infty[\to \mathbb{R}$ *c.p.m. définie par* $f(x) = 1/n^2$ *si* $x \in [n, n+1[$, $n \in \mathbb{N}^*$ *de telle sorte que* $S_N = \sum_{n=1}^N \frac{1}{n^2} = \int_1^{N+1} f(x)dx$. *On vérifie facilement que* $0 \leq f(x) \leq 1/(x-1)^2$ *sur* $[2, \infty[$ *et* $\int_2^\infty \frac{1}{(x-1)^2}dx = \int_1^\infty \frac{1}{x^2}dx$ *converge, d'où on déduit que* $\int_1^\infty f(x)dx$ *converge et*

$$\sum_{n=1}^\infty \frac{1}{n^2} = \lim_{N\to\infty}\sum_{n=1}^N \frac{1}{n^2} = \lim_{N\to\infty}\int_1^{N+1} f(x)dx < +\infty.$$

On peut encore utiliser les propriétés des intégrales généralisées pour donner des bornes par dessous et par dessus à la série S. *En effet, on a* $\frac{1}{x^2} \leq f(x) \leq \frac{1}{(x-1)^2}$ *pour tout* $x \in [1, \infty[$

du coup

$$S = \sum_{n=1}^{\infty} \frac{1}{n^2} = \int_1^{\infty} f(x)dx \leq \int_1^2 f(x)dx + \int_2^{\infty} \frac{1}{(x-1)^2}dx = 2,$$
$$S = \sum_{n=1}^{\infty} \frac{1}{n^2} = \int_1^{\infty} f(x)dx \geq \int_1^{\infty} \frac{1}{x^2}dx = 1,$$

On conclut donc $1 \leq \sum_{n=1}^{\infty} \frac{1}{n^2} \leq 2$.

Exercice 0.26. *Étudier si la série* $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ *converge ou non.*

Chapitre 1

L'espace $\mathbb{R}^n$ et sa topologie

1.1 Espaces vectoriels normés

On rappelle ici les notions générales d'espace vectoriel, norme, distance et produit scalaire.

Définition 1.1 (Espace vectoriel réel). *Un ensemble V est un* espace vectoriel réel *si les opérations de somme et multiplication par un scalaire (réel) sont définies sur V avec les propriétés suivantes :*

1. *somme :* $V \times V \to V, \quad (x, y) \in V \times V \mapsto z = x + y \in V,$
   - $\forall x, y \in V, \quad x + y = y + x,$
   - $\forall x, y, z \in V, \quad (x + y) + z = x + (y + z),$
   - $\exists$ *élément nul* $0 : \quad x + 0 = x,$
   - $\forall x \in V$, $\exists$ *élément opposé* $-x : \quad x + (-x) = 0,$
2. *multiplication par un scalaire :* $\mathbb{R} \times V \to V, \quad (\lambda, x) \in \mathbb{R} \times V \mapsto z = \lambda x \in V,$
   - $\forall \lambda, \mu \in \mathbb{R}, \forall x \in \mathbb{R}, \quad \lambda(\mu x) = (\lambda\mu)x,$
   - $\forall x \in V, \quad 1 \cdot x = x,$
   - $\forall \lambda, \mu \in \mathbb{R}, \forall x \in V, \quad (\lambda + \mu)x = \lambda x + \mu x,$
   - $\forall \lambda \in \mathbb{R}, \forall x, y \in V, \quad \lambda(x + y) = \lambda x + \lambda y.$

Définition 1.2 (Norme). *Soit V un espace vectoriel réel. Une norme sur V est une application $N : V \to \mathbb{R}_+$ qui satisfait les propriétés suivantes :*

1. $\forall x \in V, \quad N(x) \geq 0, \quad$ *et* $\quad N(x) = 0 \Leftrightarrow x = 0,$
2. $\forall x \in V, \lambda \in \mathbb{R}, \quad N(\lambda x) = |\lambda| N(x),$
3. $\forall x, y \in V, \quad N(x + y) \leq N(x) + N(y) \quad$ *(inégalité triangulaire).*

On note souvent une norme par $\|\cdot\| \quad (N(x) = \|x\|)$.

Un espace vectoriel muni d'une norme est appelé **espace vectoriel normé** et souvent noté $(V, \|\cdot\|)$. On dit que deux normes N_1 et N_2 sur un espace vectoriel V sont équivalentes s'ils existent deux constants $\underline{c}, \overline{c} > 0$ telles que $\underline{c} N_1(x) \leq N_2(x) \leq \overline{c} N_1(x)$ pour tout $x \in V$.

Définition 1.3 (distance). *Soit X un ensemble. Une* distance *ou* métrique *sur X est une application $d : X \times X \to \mathbb{R}_+$ qui satisfait les propriétés suivantes :*

1. $\forall x, y \in X, \quad d(x, y) \geq 0, \quad$ *et* $\quad d(x, y) = 0 \Leftrightarrow x = y,$
2. $\forall x, y \in X, \quad d(x, y) = d(y, x),$
3. $\forall x, y, z \in X, \quad d(x, y) \leq d(x, z) + d(z, y) \quad$ *(inégalité triangulaire).*

Un ensemble X muni d'une distance (X, d) est appelé **espace métrique**. Si $(V, \|\cdot\|)$ est un espace vectoriel normé, alors l'application

$$d : V \times V \to \mathbb{R}, \quad d(x, y) = \|x - y\|, \quad \forall x, y \in V,$$

est une distance (vérifiez-le) appelée *la distance induite par la norme* $\|\cdot\|$. Donc $(V, d(x, y) = \|x - y\|)$ est un espace métrique.

Exercice 1.4. *Soit V un espace vectoriel, $d(\cdot, \cdot) : V \times V \to \mathbb{R}_+$ une distance sur V et $h : [0, \infty) \to \mathbb{R}_+$ une fonction différentiable telle que $h(0) = 0$, $h'(x) > 0$ pour $x > 0$ et $h'(x)$ décroissante sur $[0, \infty)$. Montrer que $\tilde{d} = h \circ d$ est aussi une distance sur V. Vérifier que les hypothèses sont satisfaites par la fonction $h(x) = x/(1 + x)$, mais que la distance $\tilde{d} = h \circ d$ n'est pas induite par une norme même si ceci est vrais pour d.*

Définition 1.5 (Produit scalaire). *Un produit scalaire sur un espace vectoriel réel V est une application $b : V \times V \to \mathbb{R}$ qui satisfait les propriétés suivantes :*

1. *(symétrie)* $\quad \forall x, y \in V, \quad b(x, y) = b(y, x),$
2. *(bi-linéarité)* $\quad \forall x, y \in V, \quad \forall \alpha, \beta \in \mathbb{R}, \quad b(\alpha x + \beta y, z) = \alpha b(x, z) + \beta b(y, z),$
3. *(positivité)* $\quad \forall x \in V, \quad b(x, x) \geq 0, \quad$ *et* $\quad b(x, x) = 0 \Leftrightarrow x = 0.$

Un produit scalaire satisfait l'importante inégalité suivante :

Lemme 1.6 (Inégalité de Cauchy–Schwarz). *Soit V un espace vectoriel réel et $b : V \times V \to \mathbb{R}$ un produit scalaire sur V. Alors*

$$\forall x, y \in V, \quad |b(x, y)| \leq b(x, x)^{\frac{1}{2}} b(y, y)^{\frac{1}{2}}.$$

Démonstration. $\forall \alpha \in \mathbb{R}$, $\forall x, y \in V$

$$0 \leq b(\alpha x + y, \alpha x + y) = \alpha^2 b(x, x) + 2\alpha b(x, y) + b(y, y) = p_2(\alpha)$$

où $p_2(\alpha)$ est un polynôme de degré 2 en α. Par la positivité de b, on obtient la condition suivante pour le discriminant : $\Delta \leq 0$ ce qui implique $b^2(x, y) - b(x, x) b(y, y) \leq 0$, d'où la thèse. □

Grâce à cette propriété, un espace vectoriel réel muni d'un produit scalaire est toujours un espace normé.

Théorème 1.7. *Soit $b : V \times V \to \mathbb{R}$ un produit scalaire sur un espace vectoriel réel V. Alors $\|x\|_b = b(x, x)^{\frac{1}{2}} : V \to \mathbb{R}_+$ est une norme et $(V, \|\cdot\|_b)$ un espace vectoriel normé.*

Démonstration. Il faut vérifier que $\|\cdot\|_b$ satisfait toutes les propriétés d'une norme selon la Définition 1.2. Les propriétés 1. et 2. suivent directement de la positivité du produit scalaire (propriété 3. de 1.5) et de sa bi-linéarité (propriété 2. de 1.5).

Quant à l'inégalité triangulaire (propriété 3.) elle est une conséquence de l'inégalité de Cauchy–Schwarz :

$$\begin{aligned}\|x+y\|_b^2 &= b(x+y,x+y) = b(x,x) + 2b(x,y) + b(y,y)\\ &\leq b(x,x) + b(y,y) + 2b(x,x)^{\frac{1}{2}}b(y,y)^{\frac{1}{2}}\\ &= (\|x\|_b + \|y\|_b)^2\end{aligned}$$

□

1.2 L'espace $\mathbb{R}^n$

On note $\mathbb{R}^n = \overbrace{\mathbb{R}\times\mathbb{R}\times\cdots\times\mathbb{R}}^{n \text{ fois}}$ l'ensemble des n-uples $\mathbf{x} = (x_1,\ldots,x_n)$, avec $x_i \in \mathbb{R}$, pour $i = 1,\ldots,n$, muni des opérations de

— somme : pour tout $\mathbf{x} = (x_1,\ldots,x_n)$, $\mathbf{y} = (y_1,\ldots,y_n)$, $\quad \mathbf{x}+\mathbf{y} = (x_1+y_1,\ldots,x_n+y_n)$,

— multiplication par un scalaire : pour tout $\mathbf{x} = (x_1,\ldots,x_n), \lambda \in \mathbb{R}$, $\quad \lambda\mathbf{x} = (\lambda x_1,\ldots,\lambda x_n)$.

Ainsi, $\mathbb{R}^n$ a une structure d'espace vectoriel réel. Sur $\mathbb{R}^n$, on peut introduire plusieurs normes. Voici les plus communes :

— **Norme euclidienne :**

$$\|\mathbf{x}\|_E = \sqrt{\sum_{i=1}^n x_i^2},$$

— **Norme p :**

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}, \quad p \geq 1,$$

— **Norme ∞ :**

$$\|\mathbf{x}\|_\infty = \max_{1\leq j\leq n} |x_j|.$$

La norme euclidienne correspond à la norme p avec $p = 2$, i.e. $\|\mathbf{x}\|_E = \|\mathbf{x}\|_2$. On vérifie (exercice) que toutes les applications $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}_+$, pour tout $p \geq 1$ et $p = \infty$, sont des normes. Par contre, $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ avec $0 < p < 1$ n'est pas une norme lorsque $n \geq 2$ (elle ne vérifie pas l'inégalité triangulaire).

Toutes les normes $\|\cdot\|_p$, $p \geq 1$, sont équivalentes, c'est-à-dire, $\forall p, q \geq 1$, il existe $0 < c_1(p,q) < c_2(p,q)$:

$$c_1(p,q)\|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q \leq c_2(p,q)\|\mathbf{x}\|_p, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Plus généralement, sur $\mathbb{R}^n$, **toutes les normes sont équivalentes**, c'est-à-dire, si $\|\cdot\|$ et $|||\cdot|||$ sont deux normes sur $\mathbb{R}^n$, alors $\exists\, 0 < c_1 < c_2$ tels que.

$$c_1\|\mathbf{x}\| \leq |||\mathbf{x}||| \leq c_2\|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

La preuve de ce résultat sera proposée plus loin dans le cours.

Seule la norme euclidienne $\|\mathbf{x}\|_2 = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$ parmi toutes les normes $\|\cdot\|_p,\ p \geq 1$ est une norme induite par un produit scalaire, nommément le **produit scalaire euclidien** :

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i y_i = \mathbf{y}^\top \cdot \mathbf{x} = \mathbf{y}^\top \mathbf{x}.$$

Dans ces deux dernières expressions, $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs colonnes, $\mathbf{y}^\top$ est la transposée de $\mathbf{y}$ et le produit est la multiplication matricielle. En effet, $\|\mathbf{x}\|_2 = (\mathbf{x}, \mathbf{x})^{\frac{1}{2}}$ et l'inégalité de Cauchy Schwarz sur $\mathbb{R}^n$ devient :

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2.$$

1.3 Suites dans $\mathbb{R}^n$

Définition 1.8 (**Suite convergente**). *Soit* $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$ *une suite d'éléments de* $\mathbb{R}^n$, $\mathbf{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)}) \in \mathbb{R}^n$. *On dit que* $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ *converge s'il existe* $\mathbf{x} \in \mathbb{R}^n$ *tel que* $\lim_{k\to\infty} \|\mathbf{x} - \mathbf{x}^{(k)}\| = 0$*, c.-à-d. :*

$$\forall \epsilon > 0, \quad \exists N > 0 : \quad \forall k \geq N, \quad \|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \epsilon.$$

Dans ce cas, on note $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$.

Puisque toutes les normes sont équivalentes sur $\mathbb{R}^n$, la convergence de la suite $\mathbf{x}^{(k)}$ **ne dépend pas** de la norme choisie. De même que la valeur limite $\mathbf{x}$ (si elle existe) ne dépend pas de la norme. En particulier, si on prend la norme $\|\cdot\|_\infty$ dans la définition de $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$, on en tire la propriété suivante.

Lemme 1.9. *Une suite* $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$ *converge vers* $\mathbf{x} \in \mathbb{R}^n$ *si et seulement si* $\{x_j^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}$ *converge vers* $x_j \in \mathbb{R}$ *pour toute composante* $j = 1, \ldots, n$.

Démonstration. Soit $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$, et considérons la norme $\|\cdot\|_\infty$ dans la définition de convergence d'une suite de $\mathbb{R}^n$. Alors,

$$\forall \epsilon > 0, \quad \exists N > 0 : \quad \forall k \geq N, \quad \|\mathbf{x} - \mathbf{x}^{(k)}\|_\infty = \max_{j=1,\ldots,n} |x_j - x_j^{(k)}| \leq \epsilon,$$

ce qui implique que $\lim_{k\to\infty} x_j^{(k)} = x_j,\ \forall j = 1, \ldots, n$.

Réciproquement, supposons qu'il existe $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tel que $\lim_{k\to\infty} x_j^{(k)} = x_j$, $\forall j = 1, \ldots, n$. Alors,

$$\forall \epsilon > 0, \quad \exists N_j > 0 : \quad \forall k \geq N_j, \quad |x_j - x_j^{(k)}| \leq \epsilon.$$

En prenant $\bar{N} = \max\{N_1, \ldots, N_n\}$ on a $\max_{j=1,\ldots,n} |x_j - x_j^{(k)}| \leq \epsilon$ pour tout $k \geq \bar{N}$, ce qui implique $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$. $\square$

Définition 1.10 (**Suite de Cauchy**)**.** *Une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ est dite de* ***Cauchy*** *si*

$$\forall \epsilon > 0,\ \exists N > 0 : \quad \forall k, j \geq N, \quad \|\mathbf{x}^{(k)} - \mathbf{x}^{(j)}\| \leq \epsilon.$$

En suivant la même démonstration du lemme 1.9, on peut montrer qu'une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ est de Cauchy si et seulement si chaque suite $\{x_j^{(k)}\}_{k\in\mathbb{N}} \subset \mathbb{R},\ j = 1, \ldots, n$ est de Cauchy.

Théorème 1.11. *Une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ est convergente si et seulement si elle est de Cauchy.*

Démonstration. Ceci est vrai pour $n = 1$ (voir cours d'Analyse I). En utilisant la norme $\|\cdot\|_\infty$ et le lemme 1.9, on a : $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ converge $\iff$ $\{x_i^{(k)}\}_{k\in\mathbb{N}}$ converge pour tout i $\iff$ $\{x_i^{(k)}\}_{k\in\mathbb{N}}$ est de Cauchy pour tout i $\iff$ $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ est de Cauchy. □

Théorème 1.12 (Bolzano–Weierstrass sur $\mathbb{R}^n$)**.** *Soit $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ une suite* ***bornée****, c.-à-d., $\exists M \in]0, +\infty[$ tel que $\|\mathbf{x}^{(k)}\| \leq M,\ \forall k \geq 0$. Alors il existe une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j\in\mathbb{N}} \subset \{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ qui est convergente.*

Démonstration. On utilise le théorème de Bolzano–Weierstrass sur $\mathbb{R}$: puisque $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ est bornée, en particulier $\{x_1^{(k)}\}_{k\in\mathbb{N}}$ est bornée où $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$. Donc on peut extraire une sous-suite $x_1^{(k_j)}$ qui converge vers $x_1 \in \mathbb{R}$. Prenons maintenant la suite $y_2^{(j)} = x_2^{(k_j)}$. Puisqu'elle est bornée, on peut extraire une sous-suite $y_2^{(j_\ell)}$ qui converge vers $x_2 \in \mathbb{R}$. Ainsi $\lim_{\ell\to\infty} x_2^{(k_{j_\ell})} = x_2$ et $\lim_{\ell\to\infty} x_1^{(k_{j_\ell})} = x_1$. En itérant ce raisonnement n fois, on peut extraire une sous-suite de $\mathbf{x}^{(k)}$ dont chaque composante converge vers $(x_1, x_2, \ldots, x_n)$. □

Ce qui est important dans la preuve de ce théorème est que l'on fait un nombre **fini** d'itérations (n est fini). Si n était ∞ la preuve ne porterait pas à conclusion.

1.4 Topologie de $\mathbb{R}^n$

1.4.1 Concepts de base

On s'intéresse ici à l'étude et classification des sous-ensembles de $\mathbb{R}^n$. On commence par définir les *boules.*

Définition 1.13 (Boule de $\mathbb{R}^n$)**.** *Pour tout $\mathbf{x} \in \mathbb{R}^n$ et $\delta > 0$, on appelle*

- $B(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < \delta\}$ *: la boule ouverte centrée en $\mathbf{x}$ et de rayon δ,*
- $S(\mathbf{x}, \delta) = \partial B(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| = \delta\}$ *: la sphère centrée en $\mathbf{x}$ et de rayon δ,*
- $\overline{B}(\mathbf{x}, \delta) = B(\mathbf{x}, \delta) \cup S(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| \leq \delta\}$ *: la boule fermée centrée en $\mathbf{x}$ et de rayon δ.*

La Figure 1.1 montre la forme des boules de $\mathbb{R}^2$ selon la norme qu'on choisit. On travaille par la suite avec la norme euclidienne $\|\mathbf{x}\| = \|\mathbf{x}\|_2$ mais toutes les définitions ci après s'appliquent à n'importe quelle norme puisque toutes les normes sont équivalentes sur $\mathbb{R}^n$.

On considère maintenant un sous-ensemble quelconque E de $\mathbb{R}^n$.

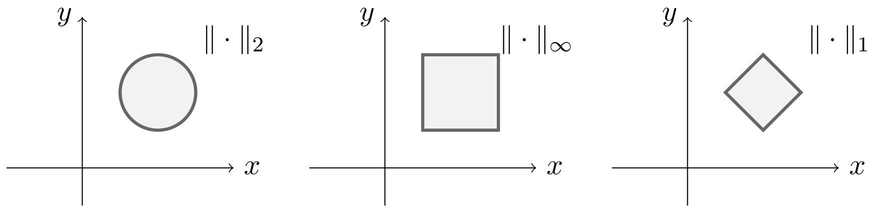


FIGURE 1.1 – Forme des boules de $\mathbb{R}^2$ pour des normes différentes

Définition 1.14 (sous-ensembles ouverts, fermés, bornés)**.** *Soit $E \subset \mathbb{R}^n$. On dit que*

- *E **est ouvert** s'il est vide ou si $\forall \mathbf{x} \in E, \quad \exists \delta > 0$ tel que $B(\mathbf{x}, \delta) \subset E$.*
- *E **est fermé** si son complémentaire $E^c = \mathbb{R}^n \backslash E = \{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \notin E\}$ est ouvert.*
- *E **est borné** s'il existe $M > 0$ tel que $\|\mathbf{x}\| \leq M, \ \forall \mathbf{x} \in E$.*

On vérifie facilement que si un sous-ensemble E est ouvert par rapport a une norme, il est aussi ouvert par rapport à n'importe quelle autre norme puisque toutes les normes sont équivalentes sur $\mathbb{R}^n$. La même conclusion est vraie pour les ensembles fermé ou bornés. L'ensemble des ouverts de $\mathbb{R}^n$ est appelé la **topologie** de $\mathbb{R}^n$ (induite par une norme).

Remarque 1.15. *Pour $\mathbf{x} \in \mathbb{R}^n$ et $\delta > 0$,*

- *$B(\mathbf{x}, \delta)$ est* ouvert *car $\forall \mathbf{z} \in B(\mathbf{x}, \delta), \ B(\mathbf{z}, \delta - \|\mathbf{z} - \mathbf{x}\|) \subset B(\mathbf{x}, \delta)$;*
- *de même, $\overline{B}(\mathbf{x}, \delta)$ est* fermé *car $\forall \mathbf{z} \in \mathbb{R}^n \backslash \overline{B}(\mathbf{x}, \delta), \ B(\mathbf{z}, \|\mathbf{z} - \mathbf{x}\| - \delta) \subset \mathbb{R}^n \backslash \overline{B}(\mathbf{x}, \delta)$.*

Étant donné un sous-ensemble $E \subset \mathbb{R}^n$, on peut classifier les points $\mathbf{x} \in \mathbb{R}^n$ par rapport à E de la façon suivante :

Définition 1.16. *Soit $E \subset \mathbb{R}^n$ et $\mathbf{x} \in \mathbb{R}^n$. On dit que*

- *$\mathbf{x}$ est un **point intérieur** de E si*

$$\exists \delta > 0 : \quad B(\mathbf{x}, \delta) \subset E.$$

  *L'ensemble des points intérieurs de E est noté $\mathring{E}$ ou $\operatorname{int}(E)$ et appelé l'**intérieur** de E.*

- *$\mathbf{x}$ est un **point frontière** si*

$$\forall \delta > 0, \quad B(\mathbf{x}, \delta) \cap E \neq \emptyset \quad \text{et} \quad B(\mathbf{x}, \delta) \cap E^c \neq \emptyset.$$

  *L'ensemble des points frontières de E est noté ∂E et appelé la **frontière** ou le **bord** de E.*

- *$\mathbf{x}$ est un **point adhérent** à E si*

$$\forall \delta > 0, \quad B(\mathbf{x}, \delta) \cap E \neq \emptyset.$$

  *Un point adhérent est soit un point intérieur, soit un point frontière. L'ensemble des points adhérents à E est noté $\overline{E}$, et appelé l'**adhérence** ou la **fermeture** de E, et coïncide avec $\overline{E} = E \cup \partial E = \mathring{E} \cup \partial E$.*

— $\mathbf{x}$ *est un **point isolé** de* E *si*

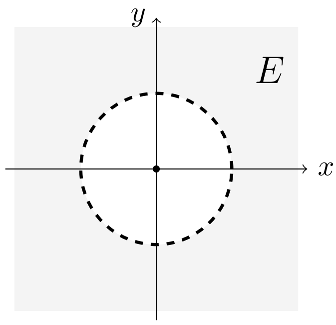
$$\exists \delta > 0 : \quad B(\mathbf{x}, \delta) \cap E = \{\mathbf{x}\}$$

— $\mathbf{x}$ *est un **point d'accumulation** de* E *si* $\forall \delta > 0, \quad B(\mathbf{x}, \delta)$ *contient au moins un point de* E *autre que* $\mathbf{x}$*, c.-à-d.* $B(\mathbf{x}, \delta) \cap (E \backslash \{\mathbf{x}\}) \neq \emptyset$*. (Rappel :* $E \backslash \{\mathbf{x}\} = E$ *si* $\mathbf{x} \notin E$*.)*

Il s'ensuit que si $\mathbf{x}$ *est un point d'accumulation de* E*, alors* $\forall \delta > 0$*,* $B(\mathbf{x}, \delta)$ *contient une infinité de points de* E*. Les points d'accumulation de* E *sont tous les points de* $\overline{E} = E \cup \partial E$ *qui ne sont pas isolés.*

On remarque que pour tout $\mathbf{y} \in \mathbb{R}^n$ et $\delta > 0$, on a $B(\mathbf{y}, \delta) \cap E \neq \emptyset \Leftrightarrow B(\mathbf{y}, \delta) \cap \overline{E} \neq \emptyset$. En effet, l'implication $\Rightarrow$ est claire car $E \subset \overline{E}$. Pour montrer l'implication $\Leftarrow$, soit $\mathbf{z} \in B(\mathbf{y}, \delta) \cap \overline{E}$. Puisque $\mathbf{z}$ est point adhérent à E, il existe $\mathbf{w} \in B(\mathbf{z}, \delta - \|\mathbf{z} - \mathbf{y}\|) \cap E \subset B(\mathbf{y}, \delta) \cap E$. Donc $B(\mathbf{y}, \delta) \cap E \neq \emptyset$.

Exercice 1.17. *Soit* $E = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 > 1\} \cup \{(0, 0)\}$ *le sous-ensemble montré en figure. Déterminer son intérieur* $\mathring{E}$*, sa frontière* ∂E*, son complémentaire* E^c*, son adhérence* $\overline{E}$*, ainsi que tous ses points isolés et d'accumulation.*



Quelques remarques sur les ensemble ouverts :

— $\mathring{E}$ est ouvert.

Démonstration. En effet, si $\mathbf{x} \in \mathring{E}$, il existe $\delta > 0$ tel que $B(\mathbf{x}, \delta) \subset E$. Vérifions que $B(\mathbf{x}, \delta) \subset \mathring{E}$, ce qui prouvera que $\mathring{E}$ est ouvert. Pour tout $\mathbf{z} \in B(\mathbf{x}, \delta)$, on a $B(\mathbf{z}, \delta - \|\mathbf{z} - \mathbf{x}\|) \subset B(\mathbf{x}, \delta) \subset E$, et donc $\mathbf{z} \in \mathring{E}$. □

— E est ouvert si et seulement si $E = \mathring{E}$.

— Toute réunion quelconque (même *infinie*, dénombrable ou non) de sous-ensembles ouverts de $\mathbb{R}^n$ est un sous-ensemble ouvert.

Démonstration. Soit $E = \bigcup_\alpha E_\alpha$ avec E_α ouvert. Pour tout $\mathbf{x} \in E$, il existe $\alpha : \mathbf{x} \in E_\alpha$. Mais, E_α étant ouvert, $\exists \delta > 0 :\ B(\mathbf{x}, \delta) \subset E_\alpha \subset E$. Donc E est ouvert. □

— Toute intersection *finie* de sous-ensembles ouverts de $\mathbb{R}^n$ est un sous-ensemble ouvert.

Démonstration. Soit $E = \bigcap_{i=1}^m E_i$, avec E_i ouvert. Si $\mathbf{x} \in E$, alors $\mathbf{x} \in E_i \quad \forall i = 1, \dots, m$ et, puisque chaque E_i est ouvert, $\exists \delta_i :\ B(\mathbf{x}, \delta_i) \subset E_i$. Soit $\delta = \min\{\delta_1, ..., \delta_m\}$ alors $B(\mathbf{x}, \delta) \subset B(\mathbf{x}, \delta_i) \subset E_i,\ \forall i$ et donc $B(\mathbf{x}, \delta) \subset \bigcap_{i=1}^m E_i = E$. □

— $\emptyset$ et $\mathbb{R}^n$ sont ouverts.

Quelques remarques sur les ensembles fermés :

— $\mathbb{R}^n \backslash \overline{E} = \mathrm{int}(\mathbb{R}^n \backslash E)$ et $\overline{\mathbb{R}^n \backslash E} = \mathbb{R}^n \backslash \mathring{E}$.

— L'adhérence $\overline{E}$ d'un ensemble $E \subset \mathbb{R}^n$ est toujours fermé.

Démonstration. La preuve est par "passage au complémentaire" : son complémentaire $\mathbb{R}^n \backslash \overline{E} = \mathrm{int}(\mathbb{R}^n \backslash E)$ est en effet ouvert. □

— E est fermé si et seulement si $E = \overline{E}$. (Preuve : par passage aux complémentaires.)

— Toute intersection quelconque (même *infinie*, dénombrable ou non) de sous-ensembles fermés est fermée. (On passe aux complémentaires pour montrer cette propriété et la suivante.)

— Toute union *finie* de sous-ensembles fermés est fermée.

— $\emptyset$ et $\mathbb{R}^n$ sont fermés.

Une caractérisation importante des ensembles fermés est la suivante.

Lemme 1.18. *Un ensemble $E \subset \mathbb{R}^n$ non vide est fermé si et seulement si toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ convergente, converge vers un élément de E.*

Démonstration. Soit E fermé et $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ une suite convergente vers $\mathbf{x} \in \mathbb{R}^n$. Alors $\mathbf{x}$ adhère à E et, puisque E est fermé, $\mathbf{x} \in E$.

Réciproquement, supposons que E n'est pas fermé, autrement dit, que $\mathbb{R}^n \backslash E$ n'est pas ouvert. Il existe donc $\overline{\mathbf{x}} \in \mathbb{R}^n \backslash E$ tel que $\forall \delta > 0 \ \ B(\overline{\mathbf{x}}, \delta) \not\subset (\mathbb{R}^n \backslash E)$. En choisissant $\delta = 1/k$ avec $k \in \mathbb{N}^*$, on obtient $\mathbf{x}^{(k)} \in B(\overline{\mathbf{x}}, 1/k) \cap E$. La suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}^*}$ dans E converge alors vers $\overline{\mathbf{x}} \notin E$. □

On montre de même que $\mathbf{x} \in \overline{E}$ ssi il existe une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ qui converge vers $\mathbf{x}$.

1.4.2 Ensembles compacts

On peut donner plusieurs définitions équivalentes d'un ensemble compact en $\mathbb{R}^n$. On présente ici la définition la plus "facile", mais non pas celle qui caractérise le mieux la notion de compacité.

Définition 1.19 (Compacité)**.** *Un ensemble $E \subset \mathbb{R}^n$ est **compact** s'il est à la fois borné et fermé. L'ensemble vide sera considéré comme compact.*

Les deux autres caractérisations (équivalentes en $\mathbb{R}^n$) sont montrées dans les théorèmes suivants.

Théorème 1.20 (Caractérisation de la compacité par sous-suites convergentes)**.** *Un sous-ensemble non vide $E \subset \mathbb{R}^n$ est compact (fermé et borné) si et seulement si de toute suite d'éléments de E on peut extraire une sous-suite qui converge vers un élément de E.*

Démonstration.

1. Soit E compact (fermé et borné). Par le théorème de Bolzano–Weierstrass, de toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ bornée (car E est borné), on peut extraire une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j\in\mathbb{N}} \subset E$ convergente telle que $\lim_{j\to\infty} \mathbf{x}^{(k_j)} = \mathbf{x} \in \mathbb{R}^n$. Puisque E est fermé, $\mathbf{x} \in E$.
2. Supposons que E n'est pas compact, autrement dit, qu'il n'est pas fermé ou qu'il n'est pas borné (ou ni l'un ni l'autre). Si E n'est pas fermé, il existe $\overline{\mathbf{x}} \in \mathbb{R}^n \backslash E$ et une suite $\{\mathbf{y}^{(k)}\}_{k\in\mathbb{N}} \subset E$ telle que $\lim_{k\to\infty} \mathbf{y}^{(k)} = \overline{\mathbf{x}}$. Une telle suite n'a aucune sous-suite qui converge vers un élément de E (car $\overline{\mathbf{x}} \notin E$). Si E n'est pas borné, il existe une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ telle que $\|\mathbf{x}^{(k)}\| > k$ pour tout $k \in \mathbb{N}$. Toute sous-suite $\{\mathbf{x}^{(k_j)}\}_{j\in\mathbb{N}}$ satisfait $\|\mathbf{x}^{(k_j)}\| > k_j \geq j$ et $\{\mathbf{x}^{(k_j)}\}_{j\in\mathbb{N}}$ ne converge pas.

□

Théorème 1.21 (de Heine-Borel-Lebesgue – Caractérisation de la compacité par recouvrements finis)**.** *Un sous-ensemble non vide $E \subset \mathbb{R}^n$ est compact (fermé et borné) si et seulement si de toute famille de sous-ensembles ouverts de $\mathbb{R}^n$ constituant un recouvrement de E, c.-à-d. $E \subset \bigcup_\alpha U_\alpha$, avec U_α ouvert, on peut extraire une famille **finie** qui est encore un recouvrement de E.*

On dit qu'un ensemble E satisfait la *propriété de Heine–Borel* si de toute famille de sous-ensembles ouverts de $\mathbb{R}^n$ constituant un recouvrement de E on peut extraire une famille finie qui est encore un recouvrement de E. Le théorème précédent affirme donc que un ensemble E est compact si et seulement si il satisfait la propriété de Heine–Borel. On montre deux exemples d'application de la propriété de Heine–Borel pour montrer qu'un ensemble n'est pas compact.

Exemple 1.22. $\mathbb{R}^2$ *n'est pas compact. En fait, on peut écrire $\mathbb{R}^2 = \bigcup_{k\in\mathbb{N}^*} B(0,k)$ mais de ce recouvrement, on ne peut pas extraire de sous-recouvrement* fini*.*

Exemple 1.23. $E = \overline{B}(0,1)\backslash\{0\}$ *n'est pas compact. En fait, on peut écrire $E \subset \bigcup_{k\in\mathbb{N}^*} \left(\overline{B}(0,\frac{1}{k})\right)^c$ qui est un recouvrement de E mais duquel on ne peut pas extraire un sous-recouvrement fini.*

La démonstration du théorème 1.21 va au dela du programme de ce cours, mais on la propose quand même ici par souci d'exhaustivité.
<u>Quelques remarques sur les ensembles compacts :</u>

— La caractérisation du théorème 1.20, qui peut être prise comme définition alternative de compacité, dit qu'un ensemble E est compact si et seulement si toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ de E admet une sous-suite qui converge vers un élément de E, c'est-à-dire qu'il existe (au moins) un point $\mathbf{x} \in E$ (point d'accumulation de la suite) tel que toute boule $B(\mathbf{x}, \delta)$, $\delta > 0$ contient une infinité de terms de la suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$. Donc E est suffisamment contraignant (compact) pour que toute suite de E s'accumule quelque part dans E.

— La caractérisation du théorème 1.21 est la définition la plus générale de compacité, mais aussi la plus abstraite. Elle exprime le fait qu'on puisse décrire un ensemble compact par un nombre *fini* de termes et est à la base de toute étape d'approximation. Soit $E \subset \mathbb{R}^n$ un sous-ensemble quelconque. Clairement, pour tout $\varepsilon > 0$, $\bigcup_{\mathbf{x}\in E} B(\mathbf{x}, \varepsilon)$ est un recouvrement de E. Si E est compact, on peut extraire un sous-recouvrement fini, c.-à-d. il existe $s = s(\varepsilon) \in \mathbb{N}^*$ et $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}\} \in E$ tels que $E \subset \bigcup_{i=1}^{s} B(\mathbf{x}^{(i)}, \varepsilon)$. Donc, E est bien approché par l'ensemble fini $\hat{E} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}\}$ au sens que pour tout $\mathbf{x} \in E$, $\mathrm{dist}(\mathbf{x}, \hat{E}) = \inf_{i=1,\dots,s} \|\mathbf{x} - \mathbf{x}^{(i)}\| < \varepsilon$. Le nombre $s = s(\varepsilon)$ est appelé *nombre de recouvrement* de E et est un indicateur de la difficulté d'approcher E par un ensemble fini.

1.4.3 Ensembles connexes et connexes par arcs

Intuitivement, un ensemble $E \subset \mathbb{R}^n$ est connexe s'il est fait "d'un seul morceau". Plus rigoureusement, on dit qu'un ensemble E ouvert est connexe si on ne peut pas le séparer en deux parties *ouvertes* non vides et *disjointes*. La définition générale pour un ensemble quelconque est la suivante :

Définition 1.24 (Connexité)**.** *Soit $E \subset \mathbb{R}^n$. On dit que E est* ***connexe*** *s'il n'existe pas deux ouverts $A, B \subset \mathbb{R}^n$ disjoints ($A \cap B = \emptyset$) tels que $A \cap E \neq \emptyset$, $B \cap E \neq \emptyset$ et $E \subset A \cup B$.*

En particulier $\emptyset$ est connexe. Les ensembles connexes de $\mathbb{R}$ sont les intervalles, par exemple $\emptyset$, $\mathbb{R}$, $[0, 1]$, $]0, 1[$, $[0, 1[$, $]-\infty, 0[$,$[0, \infty[$, etc. Une notion un peu plus forte de connexité est celle de *connexité par arcs.*

Définition 1.25. *Soit $E \subset \mathbb{R}^n$ un ensemble non vide. On appelle chemin de E une application $\boldsymbol{\gamma} : [0, 1] \to E$, $t \mapsto \boldsymbol{\gamma}(t) = (\gamma_1(t), ..., \gamma_n(t)) \in E$, dont les fonctions $\gamma_i : [0, 1] \to \mathbb{R}$ sont continues.*

Définition 1.26 (Connexité par arcs)**.** *Un ensemble non vide $E \subset \mathbb{R}^n$ est connexe par arcs si pour tout $\mathbf{x}, \mathbf{y} \in E$, il existe un chemin $\boldsymbol{\gamma} : [0, 1] \to E$ tel que $\boldsymbol{\gamma}(0) = \mathbf{x}, \boldsymbol{\gamma}(1) = \mathbf{y}$ (et $\boldsymbol{\gamma}(t) \in E, \forall t \in [0, 1]$). Nous considérerons $\emptyset$ comme connexe par arcs.*

On peut montrer que tout ensemble $E \subset \mathbb{R}^n$ connexe par arcs est aussi connexe. Le réciproque n'est toutefois pas vraie. On verra dans le chapitre suivant que les propriétés de compacité, connexité et connexité par arcs sont des propriétés topologiques, préservées par les applications continues. Autrement dit, si $E \subset \mathbb{R}^n$ est un ensemble compact (resp. connexe ou connexe par arcs) et $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ une fonction continue, alors $\mathbf{f}(E) \subset \mathbb{R}^m$ est compact (resp. connexe ou connexe par arcs).

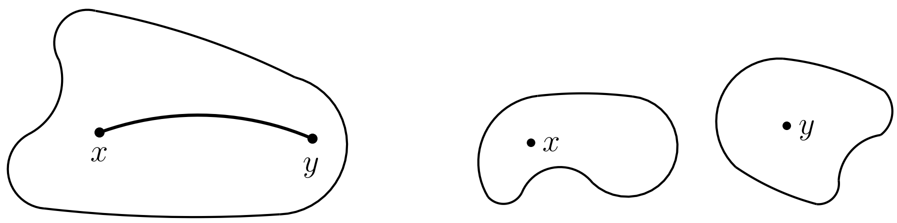


FIGURE 1.2 – Gauche : ensemble connexe par arcs. Droite : ensemble non connexe

Chapitre 2

Fonctions de plusieurs variables réelles ; limites et continuité

Soit $E \subset \mathbb{R}^n$ un ensemble non vide. On appelle *fonction sur E à valeurs réelles* une application $f : E \to \mathbb{R}$. C'est à dire, $\forall \mathbf{x} = (x_1, \ldots, x_n) \in E$, $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$ est l'image de $\mathbf{x}$ par f. La fonction f est donc une fonction de n variables réelles. On note :

— E ou $D(f)$ le domaine de f ;

— $\mathrm{Im}(f) = \{f(\mathbf{x}) \in \mathbb{R} : \mathbf{x} \in E\}$ l'image de f (notée aussi $f(E)$) ;

— $\mathcal{G}(f) = \{(\mathbf{x}, f(\mathbf{x})) \in \mathbb{R}^{n+1} : \mathbf{x} \in E\}$ le graphe de f.

Une fonction de 2 variables réelles à valeurs réelles, $(x, y) \mapsto f(x, y) \in \mathbb{R}$, peut être visualisée par son graphe (surface de $\mathbb{R}^3$), ou par ces lignes de niveau $N_f(c) = \{(x, y) \in \mathbb{R}^2 : \ f(x, y) = c\}$. La figure 2.1 montre le graphe et les lignes de niveau de la fonction $f(x, y) = e^{-x^2-y^2}$, $(x, y) \in [-1, 1]^2$.

2.1 Notions de limite

Définition 2.1 (limite)**.** *Soit $f : E \to \mathbb{R}$ et $\mathbf{x}_0 \in \mathbb{R}^n$ un point d'accumulation de E. On dit que $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ existe et est égale à $l \in \mathbb{R}$ si*

$$\forall \epsilon > 0 \ \exists \delta > 0 : \ \forall \mathbf{x} \in E \ \ \Big(\ 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta \ \Rightarrow \ |f(\mathbf{x}) - l| \leq \epsilon \ \Big).$$

On écrit alors $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = l$

La propriété $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = l$ ne dépend pas du choix de la norme $\|\cdot\|$ sur $\mathbb{R}^n$ car les normes sur $\mathbb{R}^n$ sont deux à deux équivalentes. On note que dans la définition de limite ci dessus, on exclut le point $\mathbf{x}_0$ de l'ensemble $\{\mathbf{x} \in E, \ 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta\}$. Cette limite est parfois appelée *limite épointée* et notée aussi $\lim\limits_{\substack{\mathbf{x}\to\mathbf{x}_0 \\ \neq}} f(\mathbf{x})$ ou $\lim\limits_{\substack{\mathbf{x}\to\mathbf{x}_0 \\ \mathbf{x}\neq\mathbf{x}_0}} f(\mathbf{x})$. Dans ces notes on entendra toujours par $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ la limite épointée. Le fait que $\mathbf{x}_0 \in E$ ou $\mathbf{x}_0 \notin E$ n'intervient pas dans cette définition.

Le théorème suivant donne une caractérisation équivalente de limite par les suites.

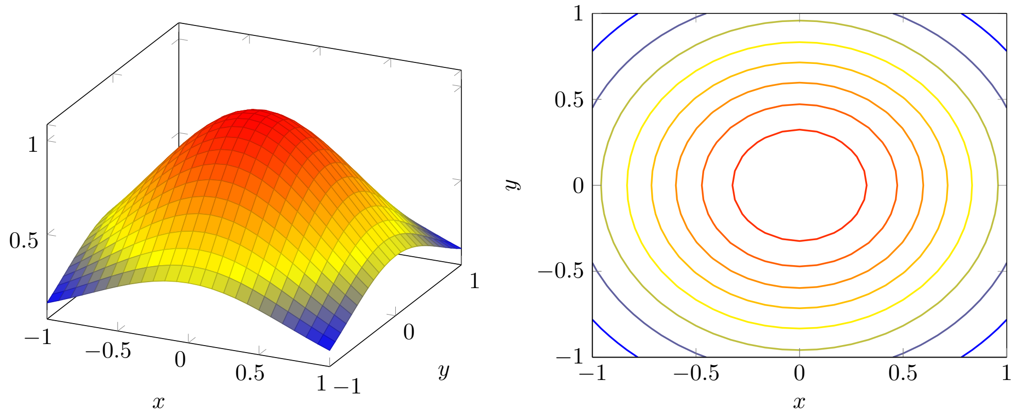


Figure 2.1 – Graphe (gauche) et lignes de niveau (droite) de la fonction $f(x,y) = e^{-x^2-y^2}$, $(x,y) \in [-1,1]^2$.

Théorème 2.2. *Soit $f : E \to \mathbb{R}$ et $\mathbf{x}_0 \in \mathbb{R}^n$ un point d'accumulation de E. Alors f admet pour limite l lorsque $\mathbf{x}$ tend vers $\mathbf{x}_0$ si et seulement si, pour toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E\backslash\{\mathbf{x}_0\}$ telle que $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}_0$, on a $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = l$. De plus la limite l est unique (si elle existe).*

Démonstration. Identique au cas d'une fonction $f : \mathbb{R} \to \mathbb{R}$ en remplaçant la boule 1D $]x_0 - \delta, x_0 + \delta[$ par la boule de $\mathbb{R}^n$, $B(\mathbf{x}_0, \delta)$. Voici la démonstration complète.

1. Supposons que f admet pour limite l lorsque $\mathbf{x}$ tend vers $\mathbf{x}_0$ et donc, pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que, pour tout $\mathbf{x} \in E$ vérifiant $0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta$, on a $|f(\mathbf{x}) - l| \leq \epsilon$. Soit, maintenant, $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E \setminus \{\mathbf{x}_0\}$ une suite telle que $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}_0$. Alors il existe $N \in \mathbb{N}$ tel que $\|\mathbf{x}^{(k)} - \mathbf{x}_0\| \leq \delta$ pour tout $k \geq N$ et donc $|f(\mathbf{x}^{(k)}) - l| \leq \epsilon$, ce qui montre que $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = l$.

2. Supposons que $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = l$ pour toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}$ qui converge vers $\mathbf{x}_0$. En raisonnant par l'absurde, supposons que f n'admet pas pour limite l lorsque $\mathbf{x}$ tend vers $\mathbf{x}_0$. Alors il existe $\epsilon > 0$ tel que pour tout $\delta > 0$ on a l'existence de $\mathbf{x}_0 \neq \mathbf{x} \in E$, $\|\mathbf{x} - \mathbf{x}_0\| \leq \delta$ (puisque $\mathbf{x}_0$ est un point d'accumulation de E) tel que $|f(\mathbf{x}) - l| > \epsilon$. Prenons $\delta = \frac{1}{k}$, $k \in \mathbb{N}^*$. Alors il existe une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}^*}$ telle que $\|\mathbf{x}^{(k)} - \mathbf{x}_0\| \leq \frac{1}{k}$ (et donc $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}_0$) et $|f(\mathbf{x}^{(k)}) - l| > \epsilon$, ce qui est contradictoire.

$\square$

Bien que la définition de limite soit la même pour des fonctions d'une seule ou de plusieurs variables réelles, le calcul des limites pour des fonctions de plusieurs variables réelles est bien plus compliqué. Prenons la caractérisation de limite par les suites :

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = l \quad \iff \quad \begin{array}{c} \forall\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E\backslash\{\mathbf{x}_0\},\ \lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}_0 \\ \text{on a } \lim\limits_{k\to\infty} f(\mathbf{x}^{(k)}) = l \end{array}$$

Pour affirmer que $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = l$ existe, il faut s'assurer que $f(\mathbf{x}^{(k)}) \xrightarrow{k\to\infty} l$ pour *n'importe quelle* suite s'approchant de $\mathbf{x}_0$.

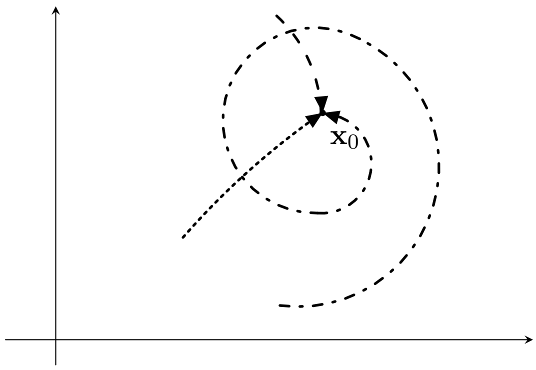


Figure 2.2 – Exemples de chemins possibles qu'on peut suivre pour atteindre $\mathbf{x}_0$.

Exemple 2.3. *Considérons la fonction* $f : E = \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$, $f(x,y) = \frac{x}{\sqrt{x^2+y^2}}$ *et* $\mathbf{x}_0 = (0,0)$, *qui est un point d'accumulation de* E. *Est-ce que* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ *existe ?*

Prenons la suite

$$\mathbf{x}^{(k)} = (\frac{1}{k}, 0) \quad k \in \mathbb{N}^* \quad \mathbf{x}^{(k)} \xrightarrow{k\to\infty} (0,0).$$

On a que

$$f(\mathbf{x}^{(k)}) = \frac{\frac{1}{k}}{\sqrt{\frac{1}{k^2}}} = 1 \quad \forall k \quad \Rightarrow \lim_{k\to\infty} f(\mathbf{x}^{(k)}) = 1.$$

Prenons maintenant la suite

$$\mathbf{y}^{(k)} = (0, \frac{1}{k}) \quad k \in \mathbb{N}^* \quad \mathbf{y}^{(k)} \xrightarrow{k\to\infty} (0,0).$$

On a que

$$f(\mathbf{y}^{(k)}) = 0 \quad \forall k \quad \Rightarrow \lim_{k\to\infty} f(\mathbf{y}^{(k)}) = 0.$$

On a donc trouvé deux suites différentes $\{(\frac{1}{k}, 0)\}_{k\in\mathbb{N}^*}$ *et* $\{(0, \frac{1}{k})\}_{k\in\mathbb{N}^*}$ *qui donnent des limites différentes. On conclut donc que* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ *n'existe pas.*

Exemple 2.4. *Considérons la fonction* $f : E = \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$, $f(x,y) = \frac{xy^2}{x^2+y^4}$, *et* $\mathbf{x}_0 = (0,0)$, *qui est un point d'accumulation de* E. *Est-ce que* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ *existe ?*

Prenons la suite $\mathbf{x}^{(k)} = (\frac{\alpha}{k}, \frac{\beta}{k}), k \geq 1$, *avec* $\alpha, \beta \in \mathbb{R}$ *non nuls en même temps. On a que*

$$f(\mathbf{x}^{(k)}) = \frac{\frac{\alpha\beta^2}{k^3}}{\frac{\alpha^2}{k^2} + \frac{\beta^4}{k^4}} = \frac{\alpha\beta^2 k}{\alpha^2 k^2 + \beta^2} \xrightarrow{k\to\infty} 0, \quad \forall(\alpha,\beta) \neq (0,0).$$

Donc si on se rapproche de $\mathbf{x}_0$ *par un chemin "droit" la limite est 0. Toutefois, si on prend la suite* $\mathbf{x}^{(k)} = (\frac{1}{k^2}, \frac{1}{k}) \xrightarrow{k\to\infty} (0,0)$, *on a que*

$$f(\mathbf{x}^{(k)}) = \frac{\frac{1}{k^4}}{\frac{1}{k^4} + \frac{1}{k^4}} = \frac{1}{2} \quad \forall k \geq 1 \quad \Longrightarrow \quad \lim_{k\to\infty} f(\mathbf{x}^{(k)}) = \frac{1}{2}.$$

Donc $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ *n'existe pas !*

2.1.1 Propriétés de l'opération de limite

L'opération de limite $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ pour des fonctions $f : E \subset \mathbb{R}^n \to \mathbb{R}$ de plusieurs variables réelles a les mêmes propriétés que pour des fonctions $f : \mathbb{R} \to \mathbb{R}$ d'une seule variable réelle.

Théorème 2.5. *Soit $E \subset \mathbb{R}^n$, $\mathbf{x}_0$ un point d'accumulation de E et $f, g : E \to \mathbb{R}$ tels que $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = l_1$ et $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = l_2$. Alors*
- $\forall \alpha, \beta \in \mathbb{R}$, $\lim_{\mathbf{x}\to\mathbf{x}_0} (\alpha f(\mathbf{x}) + \beta g(\mathbf{x})) = \alpha l_1 + \beta l_2$
- $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) g(\mathbf{x}) = l_1 l_2$
- *Si* $l_2 \neq 0$, $\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{l_1}{l_2}$

Comme pour les fonctions $f : \mathbb{R} \to \mathbb{R}$ d'une seule variable réelle, on a un critère de comparaison qui peut être très utile pour établir l'existence d'une limite.

Théorème 2.6 (des deux gendarmes)**.** *Soient $f, g, h : E \subset \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x}_0$ un point d'accumulation de E et $\lim_{\mathbf{x}\to\mathbf{x}_0} h(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = l$. S'il existe $\alpha > 0$ tel que*

$$h(\mathbf{x}) \leq f(\mathbf{x}) \leq g(\mathbf{x}) \quad \forall \mathbf{x} \in E,\ 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \alpha$$

alors $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = l$.

Remarque 2.7. *Dans le théorème des deux gendarmes, on utilise souvent des fonctions h et g qui dépendent uniquement de la distance $\|\mathbf{x} - \mathbf{x}_0\|$. Soit par exemple $g : E \to \mathbb{R}$ avec $E = \mathbb{R}^n$ ou $E = \mathbb{R}^n \backslash \{\mathbf{x}_0\}$, et supposons que*

$$\forall \mathbf{x} \in E \backslash \{\mathbf{x}_0\} \quad g(\mathbf{x}) = \tilde{g}(\|\mathbf{x} - \mathbf{x}_0\|),$$

où $\tilde{g} :]0, \infty[\to \mathbb{R}$. Alors $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = l$ si et seulement si $\lim_{r\to 0^+} \tilde{g}(r) = l$. En effet

$$\forall \epsilon > 0\ \exists \delta > 0 :\ \forall \mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \backslash \{\mathbf{x}_0\} \quad |g(\mathbf{x}) - l| = |\tilde{g}(\|\mathbf{x} - \mathbf{x}_0\|) - l| \leq \epsilon$$

si et seulement si

$$\forall \epsilon > 0\ \exists \delta > 0 :\ \forall r \in]0, \delta] \quad |\tilde{g}(r) - l| \leq \epsilon.$$

Exemple 2.8. *Soit $f : E = \mathbb{R}^2 \setminus \{(1,0)\} \to \mathbb{R}$, $f(x,y) = y \ln((x-1)^2 + y^2)$. Calculer si elle existe $\lim_{(x,y)\to(1,0)} f(x,y)$.*

Prenons la suite $\mathbf{x}^{(k)} = (1, \frac{1}{k})$ Alors $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = \lim_{k\to\infty} \frac{1}{k} \ln \frac{1}{k^2} = 0$. Donc si la limite existe elle doit être égale à $l = 0$. On a de plus

$$\begin{aligned} 0 \leq |f(x,y)| &= |y|\, |\ln((x-1)^2 + y^2)| \\ &\leq \sqrt{(x-1)^2 + y^2}\, |\ln((x-1)^2 + y^2)|. \end{aligned}$$

Notons $\rho = \sqrt{(x-1)^2 + y^2} = \|(x,y) - (1,0)\|$. Puisque

$$\lim_{(x,y)\to(1,0)} \sqrt{(x-1)^2 + y^2} |\ln((x-1)^2 + y^2)| = \lim_{\rho\to 0^+} \rho |\ln \rho^2| = \lim_{\rho\to 0^+} \rho(-\ln \rho^2) = 0$$

(par la remarque appliquée à la norme euclidienne), on conclut par le théorème des deux gendarmes que $\lim_{(x,y)\to(1,0)} |f(x,y)| = 0$ et $\lim_{(x,y)\to(1,0)} f(x,y) = 0$.

Théorème 2.9 (Critère de Cauchy). *Soit* $f : E \subset \mathbb{R}^n \to \mathbb{R}$ *et* $\mathbf{x}_0$ *un point d'accumulation de* E. *Alors* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ *existe (dans* $\mathbb{R}$*) si et seulement si*

$$\forall\epsilon > 0\ \exists\delta > 0 :\ \forall\mathbf{x},\mathbf{y} \in \overline{B}(\mathbf{x}_0,\delta) \cap (E\backslash\{\mathbf{x}_0\})\ \ |f(\mathbf{x}) - f(\mathbf{y})| \leq \epsilon.$$

Démonstration. Le sens $\Rightarrow$ est clair. Montrons le sens $\Leftarrow$. Il existe $\delta_1 > 0$ tel que

$$\forall\mathbf{x},\mathbf{y} \in \overline{B}(\mathbf{x}_0,\delta_1) \cap (E\backslash\{\mathbf{x}_0\})\ \ |f(\mathbf{x}) - f(\mathbf{y})| \leq 1.$$

Pour $\delta \in]0,\delta_1]$, soit les nombres réels

$$\alpha(\delta) = \inf\{f(\mathbf{x}) : \mathbf{x} \in \overline{B}(\mathbf{x}_0,\delta)\cap(E\backslash\{\mathbf{x}_0\})\},\ \ \beta(\delta) = \sup\{f(\mathbf{x}) : \mathbf{x} \in \overline{B}(\mathbf{x}_0,\delta)\cap(E\backslash\{\mathbf{x}_0\})\}$$

(deux fonctions monotones en δ). Il en résulte que

$$\forall\epsilon > 0\ \exists\delta \in]0,\delta_1]\ \ 0 \leq \beta(\delta) - \alpha(\delta) \leq \epsilon$$

et donc, comme $\beta - \alpha$ est croissante en δ, $\lim_{r\to 0^+}(\beta(r) - \alpha(r)) = 0$. Posons $\ell = \lim_{r\to 0^+}\beta(r) = \lim_{r\to 0^+}\alpha(r)$. Comme

$$\forall\mathbf{x} \in \overline{B}(\mathbf{x}_0,\delta_1) \cap (E\backslash\{\mathbf{x}_0\})\ \ \alpha(\|\mathbf{x} - \mathbf{x}_0\|) \leq f(\mathbf{x}) \leq \beta(\|\mathbf{x} - \mathbf{x}_0\|),$$

le théorème des deux gendarmes assure que $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell$. □

2.1.2 Limite de fonctions à valeurs dans $\mathbb{R}^m$

La définition de limite s'étend sans difficultés aux fonctions à valeurs dans $\mathbb{R}^m$. Soit $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$, c'est-à-dire,

$$\forall\mathbf{x} \in E \quad \mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1,\ldots,x_n) = (f_1(x_1,\ldots,x_n),\ldots,f_m(x_1,\ldots,x_n))$$

où $f_i : E \to \mathbb{R}$. Donc une fonction $\mathbf{f} : E \to \mathbb{R}^m$ est une collection de m fonctions $f_i : E \to \mathbb{R}$. Lorsque $n = m$, la norme dans l'espace de départ n'est pas nécessairement la même que celle dans l'espace d'arrivée. Dans ce qui suit, nous choisirons la norme euclidienne, sauf mention du contraire.

Définition 2.10. *Soit* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *et* $\mathbf{x}_0$ *un point d'accumulation de* E. *On dit que* $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l} \in \mathbb{R}^m$ *si*

$$\forall\epsilon > 0,\ \exists\delta > 0 : \quad \forall\mathbf{x} \in E,\ 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta \quad \|\mathbf{f}(\mathbf{x}) - \mathbf{l}\| \leq \epsilon.$$

Comme pour les fonctions à valeurs dans $\mathbb{R}$, $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l}$ existe si et seulement si pour toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E \setminus \{\mathbf{x}_0\}$ telle que $\mathbf{x}^{(k)} \xrightarrow{k\to\infty} \mathbf{x}_0$ on a $\lim_{k\to\infty} \mathbf{f}(\mathbf{x}^{(k)}) = \mathbf{l}$ (limite dans $\mathbb{R}^m$). Il est aussi facile de montrer que $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l}$ si et seulement si toutes les limites $\lim_{\mathbf{x}\to\mathbf{x}_0} f_i(\mathbf{x}) = l_i$, $i = 1,\ldots,m$ existent.

2.2 Fonctions continues

Soit $E \subset \mathbb{R}^n$ un ensemble non vide et $f : E \to \mathbb{R}$.

Définition 2.11 (fonction continue en un point)**.** *Soit* $\mathbf{x}_0 \in E$.

- *Si* $\mathbf{x}_0$ *est un point isolé, on admettra (par définition) que* f *est continue en* $\mathbf{x}_0$.
- *Si* $\mathbf{x}_0$ *n'est pas isolé (il est donc un point d'accumulation de* E*) on dit que* f *est continue en* $\mathbf{x}_0$ *si* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ *existe et* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$.

De la définition 2.1 de limite d'une fonction ainsi que du théorème 2.2, caractérisant les limites par les suites, il en suit que, pour tout $\mathbf{x}_0 \in E$, les trois affirmations suivantes sont équivalentes :

i. $f : E \to \mathbb{R}$ est continue en $\mathbf{x}_0$;

ii. $\forall \epsilon > 0 \ \exists \delta = \delta(\epsilon, \mathbf{x}_0) > 0 : \ \forall \mathbf{y} \in E \ \Big(\|\mathbf{y} - \mathbf{x}_0\| \le \delta \Rightarrow |f(\mathbf{y}) - f(\mathbf{x}_0)| \le \epsilon \Big)$;

iii. $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = f(\mathbf{x}_0)$ pour toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ qui converge vers $\mathbf{x}_0$.

Définition 2.12 (fonction continue sur un ensemble)**.** *On dit que* $f : E \to \mathbb{R}$ *est continue sur* E *si elle est continue en tout point* $\mathbf{x} \in E$. *Dans ce cas, on note* $f \in C^0(E)$ *(ou* $f \in C^0(E, \mathbb{R})$*).*

Il en résulte que f est continue en tout $\mathbf{x} \in E$ si et seulement si

$$\forall \epsilon > 0 \ \forall \mathbf{x} \in E \ \exists \delta = \delta(\epsilon, \mathbf{x}) > 0 : \ \forall \mathbf{y} \in E \ \Big(\|\mathbf{x} - \mathbf{y}\| \le \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| \le \epsilon \Big).$$

Définition 2.13 (fonction uniformément continue)**.** *On dit que* $f : E \to \mathbb{R}$ *est* uniformément *continue sur* E *si*

$$\forall \epsilon > 0 \ \exists \delta = \delta(\epsilon) > 0 : \ \forall \mathbf{x}, \mathbf{y} \in E \ \Big(\|\mathbf{x} - \mathbf{y}\| \le \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| \le \epsilon \Big).$$

On note que ici δ *peut être choisi de manière qui ne dépend pas de* $\mathbf{x}$, *contrairement à la caractérisation précédente de continuité sur* E.

On remarque qu'une fonction $f : E \to \mathbb{R}$ uniformément continue sur E est aussi continue sur E. Le contraire n'est pas nécessairement vrai.

Exemple 2.14. *La fonction* $\mathbf{x} \mapsto \|\mathbf{x}\| \in \mathbb{R}_+$ *est uniformément continue sur* $\mathbb{R}^n$ *(et donc continue) car, pour tous* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\Big| \|\mathbf{x}\| - \|\mathbf{y}\| \Big| \le \|\mathbf{x} - \mathbf{y}\|$$

(inégalité triangulaire inverse, *qui découle de l'inégalité triangulaire).*

Exemple 2.15. *Toute fonction constante* $\mathbf{x} \mapsto \mathbf{C} \in \mathbb{R}^m$ *et, pour* $i \in \{1, \dots, n\}$ *fixé, la* i*-ème projection* $\mathbf{x} \mapsto x_i \in \mathbb{R}$ *sont des fonctions uniformément continues sur* $\mathbb{R}^n$. *En effet* $|x_i - y_i| \le \|\mathbf{x} - \mathbf{y}\|$ *(norme euclidienne) pour tous* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Les définitions de continuité et continuité uniforme s'étendent sans difficultés à des fonctions $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$.

Définition 2.16. *Soit* $\mathbf{x}_0 \in E$.

— *Si* $\mathbf{x}_0$ *est un point isolé, on admettra (par définition) que* $\mathbf{f}$ *est continue en* $\mathbf{x}_0$.

— *Si* $\mathbf{x}_0$ *n'est pas isolé, on dit que* $\mathbf{f}$ *est continue en* $\mathbf{x}_0$ *si* $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x})$ *existe et* $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)$.

Il en suit que, pour tout $\mathbf{x}_0 \in E$, les quatre affirmations suivantes sont équivalentes :

i. $\mathbf{f} = (f_1, \ldots, f_m) : E \to \mathbb{R}^m$ est continue en $\mathbf{x}_0$;

ii. $\forall \epsilon > 0\ \exists \delta = \delta(\epsilon, \mathbf{x}_0) > 0 :\ \forall \mathbf{y} \in E\ \Big(\|\mathbf{y} - \mathbf{x}_0\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}_0)\| \leq \epsilon\Big)$;

iii. $\lim_{k\to\infty} \mathbf{f}(\mathbf{x}^{(k)}) = \mathbf{f}(\mathbf{x}_0)$ pour toute suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ qui converge vers $\mathbf{x}_0$;

iv. pour chaque $i \in \{1, \ldots, m\}$ la fonction $f_i : E \to \mathbb{R}$ est continue en $\mathbf{x}_0$.

Il en résulte aussi que $\mathbf{f}$ est continue en tout $\mathbf{x} \in E$ si et seulement si

$$\forall \epsilon > 0\ \forall \mathbf{x} \in E\ \exists \delta = \delta(\epsilon, \mathbf{x}) > 0 :\ \forall \mathbf{y} \in E\ \Big(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon\Big).$$

Dans ce cas, on note $f \in C^0(E, \mathbb{R}^m)$ (ou simplement $f \in C^0(E)$).

Définition 2.17. *On dira que* $\mathbf{f} : E \to \mathbb{R}$ *est* uniformément *continue sur* E *si*

$$\forall \epsilon > 0\ \exists \delta = \delta(\epsilon) > 0 :\ \forall \mathbf{x}, \mathbf{y} \in E\ \Big(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon\Big). \tag{2.1}$$

Remarque 2.18. *Soit* $\emptyset \neq A \subset E$ *et la restriction* $\mathbf{f} : A \to \mathbb{R}^m$ *de* $\mathbf{f}$ *à* A *(notée aussi* $\mathbf{f}|_A$*). Si* $\mathbf{f} : E \to \mathbb{R}^m$ *est continue sur* E*, alors* $\mathbf{f} : A \to \mathbb{R}^m$ *est continue sur* A.

2.2.1 Propriétés des fonctions continues

Théorème 2.19. *Soient* f *et* g *deux fonctions de* $E \subset \mathbb{R}^n$ *dans* $\mathbb{R}$ *continues en* $\mathbf{x}_0 \in E$.

— $\forall \alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ *est continue en* $\mathbf{x}_0$ *;*

— $f \cdot g$, $|f|$, $|g|$ *sont continues en* $\mathbf{x}_0$ *;*

— *si* $g(\mathbf{x}_0) \neq 0$*, alors* $\frac{f}{g}$ *est continue en* $\mathbf{x}_0$.

Théorème 2.20 (Composition de fonctions continues)**.** *Soit* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *continue en* $\mathbf{x}_0 \in E$ *et* $\mathbf{g} : A \subset \mathbb{R}^p \to \mathbb{R}^n$ *continue en* $\mathbf{y}_0 \in A$ *et telle que* $\mathbf{x}_0 = \mathbf{g}(\mathbf{y}_0)$*. Alors* $\mathbf{h} = \mathbf{f} \circ \mathbf{g} : B = \{\mathbf{y} \in A : \mathbf{g}(\mathbf{y}) \in E\} \to \mathbb{R}^m$ *est continue en* $\mathbf{y}_0$.

Démonstration. Pour toute suite $\{\mathbf{y}^{(k)}\}_{k\in\mathbb{N}} \subset B$ telle que $\lim_{k\to\infty} \mathbf{y}^{(k)} = \mathbf{y}_0$, on a par la continuité de $\mathbf{g}$ que $\mathbf{g}(\mathbf{y}^{(k)}) \xrightarrow{k\to\infty} \mathbf{g}(\mathbf{y}_0) = \mathbf{x}_0$ et par la continuité de $\mathbf{f}$ en $\mathbf{x}_0$, $\mathbf{f}(\mathbf{g}(\mathbf{y}^{(k)})) \xrightarrow{k\to\infty} \mathbf{f}(\mathbf{x}_0)$. Donc, $\mathbf{h}(\mathbf{y}^{(k)}) \xrightarrow{k\to\infty} \mathbf{f}(\mathbf{x}_0) = \mathbf{f}(\mathbf{g}(\mathbf{y}_0)) = \mathbf{h}(\mathbf{y}_0)$ ce qui montre la continuité de $\mathbf{h}$ en $\mathbf{x}_0$. $\square$

2.3 Prolongement de fonctions par continuité

Définition 2.21. *Soit* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *et* $\mathbf{x}_0 \in \overline{E} \backslash E$. *Une fonction* $\tilde{\mathbf{f}} : E \cup \{\mathbf{x}_0\} \to \mathbb{R}^m$ *est appelée un prolongement par continuité de* $\mathbf{f}$ *en* $\mathbf{x}_0$ *si* $\tilde{\mathbf{f}} = \mathbf{f}$ *sur* E *et* $\tilde{\mathbf{f}}$ *est continue en* $\mathbf{x}_0$.

Il est facile de vérifier qu'un prolongement $\tilde{\mathbf{f}}$ existe si et seulement si $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x})$ existe (dans $\mathbb{R}^m$), auquel cas $\tilde{\mathbf{f}}$ est uniquement déterminée par $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$, $\forall \mathbf{x} \in E$ et $\tilde{\mathbf{f}}(\mathbf{x}_0) = \lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x})$.

Théorème 2.22. *Soit* $E \subset \mathbb{R}^n$ *un ensemble non vide et* $\mathbf{f} : E \to \mathbb{R}^m$ *une fonction continue sur* E. *Supposons que, pour tout* $\mathbf{x} \in \bar{E} \setminus E$, *la limite* $\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y})$ *existe. Alors la fonction* $\tilde{\mathbf{f}} : \bar{E} \to \mathbb{R}^m$ *définie par* $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ *si* $\mathbf{x} \in E$ *et* $\tilde{\mathbf{f}}(\mathbf{x}) = \lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y})$ *si* $\mathbf{x} \in \bar{E} \setminus E$ *est continue et appelée le prolongement de* $\mathbf{f}$ *par continuité sur* $\overline{E}$.

Démonstration. Soit $\mathbf{x} \in \bar{E}$ et prouvons la continuité de $\tilde{\mathbf{f}}$ en $\mathbf{x}$ à l'aide de suites. Soit donc une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset \bar{E}$ qui converge vers $\mathbf{x}$. Remplaçons-la par une suite $\{\mathbf{y}^{(k)}\}_{k\in\mathbb{N}} \subset E$ qui converge aussi vers $\mathbf{x}$ et telle que $\lim_{k\to+\infty} \|\mathbf{f}(\mathbf{y}^{(k)}) - \tilde{\mathbf{f}}(\mathbf{x}^{(k)})\| = 0$: si $\mathbf{x}^{(k)} \in E$, on pose $\mathbf{y}^{(k)} = \mathbf{x}^{(k)}$ et, si $\mathbf{x}^{(k)} \in \bar{E}\backslash E$, on choisit $\mathbf{y}^{(k)} \in E$ tel que

$$\|\mathbf{y}^{(k)} - \mathbf{x}^{(k)}\| \leq 2^{-k} \quad \text{et} \quad \|\mathbf{f}(\mathbf{y}^{(k)}) - \tilde{\mathbf{f}}(\mathbf{x}^{(k)})\| \leq 2^{-k}.$$

Par définition de $\tilde{\mathbf{f}}(\mathbf{x})$, on a $\lim_{k\to+\infty} \mathbf{f}(\mathbf{y}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{x})$; en effet, si $\mathbf{x} \in E$, alors $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ et ceci découle de la continuité de $\mathbf{f}$ en $\mathbf{x}$ et, si $\mathbf{x} \in \bar{E}\backslash E$, ceci découle de la définition de $\tilde{\mathbf{f}}(\mathbf{x})$. Donc $\lim_{k\to+\infty} \tilde{\mathbf{f}}(\mathbf{x}^{(k)}) = \lim_{k\to+\infty} \left(\tilde{\mathbf{f}}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})\right) + \lim_{k\to+\infty} \mathbf{f}(\mathbf{y}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{x})$. □

Le prochain théorème est un résultat important. Il montre que sous l'hypothèse que la fonction soit *uniformément continue* sur E, on n'a pas besoin de vérifier l'existence des limites au bord pour pouvoir prolonger la fonction par continuité. L'hypothèse d'uniforme continuité est bien plus facile à vérifier que l'existence de la limite en chaque point du bord. On verra, par exemple, dans le chapitre suivant, que si la fonction est dérivable sur un ensemble E convexe avec dérivées partielles bornées, alors elle est uniformément continue sur E.

Théorème 2.23 (Prolongement de fonctions uniformément continues)**.** *Soit* $E \subset \mathbb{R}^n$ *un ensemble non vide et* $\mathbf{f} : E \to \mathbb{R}^m$ *une fonction uniformément continue sur* E. *Alors* $\mathbf{f}$ *peut être prolongée par continuité sur* $\overline{E}$ *et son prolongement continu* $\tilde{\mathbf{f}} : \overline{E} \to \mathbb{R}^m$ *est uniformément continu.*

Démonstration. Vérification que $\mathbf{f}$ *peut être prolongée par continuité sur* $\overline{E}$. Il faut vérifier que $\lim_{\mathbf{y}\to\mathbf{x}} f(\mathbf{x})$ existe en tout $\mathbf{x} \in \overline{E} \setminus E$ (la limite en $\mathbf{x} \in E$ existe car $\mathbf{f}$ est continue sur E). Par hypothèse, la fonction $\mathbf{f}$ est uniformément continue sur E. Pour tout $\epsilon > 0$, soit $\delta > 0$ la valeur correspondante dans la définition (2.1) de continuité uniforme.

Pour chaque $\mathbf{a} \in \overline{E} \setminus E$, choisissons une suite $\{\mathbf{a}^{(k)}\}_{k\in\mathbb{N}} \subset E$ qui converge vers $\mathbf{a}$. Pour $\mathbf{a} \in \overline{E}\backslash E$ fixé, la suite choisie $\{\mathbf{a}^{(k)}\}_{k\in\mathbb{N}}$ est une suite de Cauchy. Il existe donc $N = N(\delta)$ tel que $\forall j, k \geq N \;\; \|\mathbf{a}^{(j)} - \mathbf{a}^{(k)}\| \leq \delta$. Il s'ensuit que, pour tous $j, k \geq N$, $\|\mathbf{f}(\mathbf{a}^{(j)}) - \mathbf{f}(\mathbf{a}^{(k)})\| \leq \epsilon$. D'où $\{\mathbf{f}(\mathbf{a}^{(k)})\}_{k\in\mathbb{N}}$ est une suite de Cauchy dans $\mathbb{R}^m$, qui admet pour limite un certain

$\mathbf{l} = \lim_{k\to\infty} \mathbf{f}(\mathbf{a}^{(k)}) \in \mathbb{R}^m$. Cette limite ne dépend pas de la suite choisie. En effet, soit $\{\mathbf{b}^{(k)}\}_{k\in\mathbb{N}} \subset E$ une autre suite telle que $\lim_{k\to\infty} \mathbf{b}^{(k)} = \mathbf{a}$ et $\mathbf{m} = \lim_{k\to\infty} \mathbf{f}(\mathbf{b}^{(k)})$. Alors il existe $\tilde{N} > N : \ \forall k \geq \tilde{N}, \ \|\mathbf{m} - f(\mathbf{b}^{(k)})\| \leq \epsilon, \ \|\mathbf{l} - f(\mathbf{a}^{(k)})\| \leq \epsilon, \ \|\mathbf{a}^{(k)} - \mathbf{b}^{(k)}\| \leq \delta$. Par conséquent,

$$\|\mathbf{l} - \mathbf{m}\| \leq \|\mathbf{l} - \mathbf{f}(\mathbf{a}^{(k)})\| + \|\mathbf{f}(\mathbf{a}^{(k)}) - \mathbf{f}(\mathbf{b}^{(k)})\| + \|\mathbf{m} - \mathbf{f}(\mathbf{b}^{(k)})\| \leq 3\epsilon$$

ce qui implique $\mathbf{l} = \mathbf{m}$ par l'arbitrarité de ϵ et donc la limite $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x})$ existe en tout $\mathbf{a} \in E$ et $\mathbf{f}$ peut être prolongée par continuité. On dénote $\tilde{\mathbf{f}}$ le prolongement par continuité de $\mathbf{f}$.

Vérification que $\tilde{\mathbf{f}}$ est uniformément continue sur $\overline{E}$. Soit $\epsilon > 0$ et $\mathbf{x}, \mathbf{y} \in \overline{E}$ tels que $\|\mathbf{x} - \mathbf{y}\| \leq \frac{\delta}{3}$, où $\delta = \delta(\epsilon)$ est comme ci-dessus. Introduisons deux suites $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}}, \{\mathbf{y}^{(k)}\}_{k\in\mathbb{N}} \subset E$ qui convergent à $\mathbf{x}$ et $\mathbf{y}$, respectivement. Alors, il existe $M > 0$ tel que, pour tout $k \geq M$, $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{\delta}{3}$ et $\|\mathbf{y} - \mathbf{y}^{(k)}\| \leq \frac{\delta}{3}$, d'où $\|\mathbf{x}^{(k)} - \mathbf{y}^{(k)}\| \leq \|\mathbf{x}^{(k)} - \mathbf{x}\| + \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{y}^{(k)}\| \leq \delta$.

Puisque $\tilde{\mathbf{f}}$ est continue sur $\overline{E}$ on a $\lim_{k\to\infty} \mathbf{f}(\mathbf{x}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{x})$, $\lim_{k\to\infty} \mathbf{f}(\mathbf{y}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{y})$ et il existe $\tilde{M} > M$ tel que, pour tout $k \geq \tilde{M}$, $\|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^{(k)})\| \leq \epsilon$ et $\|\tilde{\mathbf{f}}(\mathbf{y}) - \mathbf{f}(\mathbf{y}^{(k)})\| \leq \epsilon$. On a alors

$$\|\tilde{\mathbf{f}}(\mathbf{x}) - \tilde{\mathbf{f}}(\mathbf{y})\| \leq \|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^{(k)})\| + \|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})\| + \|\mathbf{f}(\mathbf{y}^{(k)}) - \tilde{\mathbf{f}}(\mathbf{y})\| \leq 3\epsilon.$$

Ainsi, si $\mathbf{x}, \mathbf{y} \in \overline{E}$ satisfont $\|\mathbf{x} - \mathbf{y}\| \leq \frac{\delta}{3}$, on a $\|\tilde{\mathbf{f}}(\mathbf{x}) - \tilde{\mathbf{f}}(\mathbf{y})\| \leq 3\epsilon$, ce qui prouve que $\tilde{\mathbf{f}}$ est uniformément continue sur $\overline{E}$. $\square$

2.4 Fonctions continues sur un compact

On commence par introduire la notion de fonction bornée et de borne supérieure et inférieure.

Définition 2.24 (fonction bornée)**.** *On dit que $f : E \subset \mathbb{R}^n \to \mathbb{R}$ est bornée s'il existe $C > 0$ tel que $|f(\mathbf{x})| \leq C$, $\forall \mathbf{x} \in E$.*

Définition 2.25 (bornes supérieure et inférieure)**.** *Soit $E \subset \mathbb{R}^n$ non vide et $f : E \to \mathbb{R}$.*

- *Soit $M = \sup_{\mathbf{x}\in E} f(\mathbf{x})$. Si $M < +\infty$, alors on a $f(\mathbf{x}) \leq M$, $\forall \mathbf{x} \in E$ et il existe une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ telle que $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = M$. On dit que M est la borne supérieure ou le supremum de f sur E.*
- *S'il existe $\mathbf{x}_M \in E$ tel que $f(\mathbf{x}_M) = M$, alors on dit que M est le* maximum *de f sur E, $M = \max_{\mathbf{x}\in E} f(\mathbf{x})$ et f atteint son maximum au point $\mathbf{x}_M$. On dit aussi que $\mathbf{x}_M$ (pas nécessairement unique) est un point de maximum de f.*
- *Si $M = +\infty$, on dit que f n'est pas bornée supérieurement.*
- *On a des définitions du même type pour la borne inférieure ou l'infimum $m = \inf_{\mathbf{x}\in E} f(\mathbf{x})$, le minimum $m = \min_{\mathbf{x}\in E} f(\mathbf{x})$ et un point de minimum $\mathbf{x}_m \in E$ tel que $f(\mathbf{x}_m) = m$.*

Théorème 2.26. *Soit $E \subset \mathbb{R}^n$ un ensemble non vide et* compact*, et $f : E \to \mathbb{R}$ une fonction continue. Alors f est bornée et atteint ses bornes, c'est-à-dire, il existe $\mathbf{x}_M, \mathbf{x}_m \in E$ tels que $f(\mathbf{x}_M) = \sup_{\mathbf{x}\in E} f(\mathbf{x}) = \max_{\mathbf{x}\in E} f(\mathbf{x})$ et $f(\mathbf{x}_m) = \inf_{\mathbf{x}\in E} f(\mathbf{x}) = \min_{\mathbf{x}\in E} f(\mathbf{x})$.*

Démonstration. Similaire au cas des fonctions $f : [a, b] \to \mathbb{R}$ sur un intervalle fermé et borné. Montrons d'abord que f est bornée. Ab absurdo, si f n'était pas bornée, alors $\forall k \in \mathbb{N}$, $\exists \mathbf{x}^{(k)} \in E : \ |f(\mathbf{x}^{(k)})| > k$. Puisque E est compact, on peut extraire une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j\in\mathbb{N}}$ qui converge vers un certain $\mathbf{x} \in E$. Mais, f étant continue en $\mathbf{x}$, pour tout $\epsilon > 0$ il existe $K_\epsilon \in \mathbb{N}$ tel que $|f(\mathbf{x}^{(k_j)})| \le |f(\mathbf{x})| + \epsilon$ pour tout $j > K_\epsilon$, ce qui contredit $|f(\mathbf{x}^{(k_j)})| > k_j \ge j$, $\forall j$.

Soit maintenant $M = \sup_{\mathbf{x}\in E} f(\mathbf{x})$. Il existe une suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}} \subset E$ telle que $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = M$. A nouveau, on peut extraire une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j\in\mathbb{N}}$ qui converge vers un certain $\mathbf{x}_M \in E$ et, par continuité de f, on a $f(\mathbf{x}_M) = M$, ce qui prouve que $M = \sup_{\mathbf{x}\in E} f(\mathbf{x}) = f(\mathbf{x}_M) = \max_{\mathbf{x}\in E} f(\mathbf{x})$. Idem pour le minimum. □

Théorème 2.27. *Soit $E \subset \mathbb{R}^n$ un ensemble non vide,* compact *et* connexe par arcs*, et $f : E \to \mathbb{R}$ une fonction continue sur E. Alors f atteint toutes les valeurs entre son minimum m et maximum M sur E, et* $\mathrm{Im}(f) = [m, M]$.

Démonstration. Puisque E est compact et $f : E \to \mathbb{R}$ continue, il existe $\mathbf{x}_m$ et $\mathbf{x}_M$ t.q. $f(\mathbf{x}_m) = \min_{\mathbf{x}\in E} f(\mathbf{x})$ et $f(\mathbf{x}_M) = \max_{\mathbf{x}\in E} f(\mathbf{x})$. Puisque E est connexe par arcs, il existe un chemin $\boldsymbol{\gamma}(t) = (\gamma_1(t), \ldots, \gamma_n(t))$ avec $\gamma_i : [0, 1] \to \mathbb{R}$ continues et $\boldsymbol{\gamma}(t) \in E$, $\forall t \in [0, 1]$, $\boldsymbol{\gamma}(0) = \mathbf{x}_m$, $\boldsymbol{\gamma}(1) = \mathbf{x}_M$. Soit $g : [0, 1] \to \mathbb{R}$, $g(t) = f(\boldsymbol{\gamma}(t)) = f(\gamma_1(t), \ldots, \gamma_n(t))$, qui est continue sur $[0, 1]$ puisqu'elle est la composition de fonctions continues. D'après le théorème de la valeur intermédiaire d'Analyse I, $\mathrm{Im}(g)$ est un intervalle, et il contient $g(0) = m$ et $g(1) = M$, et donc tout $[m, M]$. D'où

$$[m, M] \subset \mathrm{Im}(g) \subset \mathrm{Im}(f) \subset [m, M]$$

et $\mathrm{Im}(f) = [m, M]$. □

Les deux théorèmes précédents montrent deux propriétés importantes des fonctions continues, qui se généralisent comme suit aux fonctions à valeurs dans $\mathbb{R}^m$.

Soit $\mathbf{f} : E \to \mathbb{R}^m$ continue et $\emptyset \neq A \subset E \subset \mathbb{R}^n$.

— Si A est compact, alors $\mathbf{f}(A) \subset \mathbb{R}^m$ est aussi compact.

— Si A est connexe (resp. connexe par arcs), alors $\mathbf{f}(A)$ est connexe (resp. connexe par arcs).

On conclut par une propriété importante des fonctions continues sur un compact.

Théorème 2.28 (Cantor-Heine)**.** *Soit $E \subset \mathbb{R}^n$ un ensemble non vide et* compact*, et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction continue. Alors $\mathbf{f}$ est uniformément continue sur E.*

Démonstration. Similaire au cas des fonctions $f : [a, b] \to \mathbb{R}$ sur un intervalle fermé et borné. Ab absurdo supposons que $\mathbf{f}$ ne soit pas uniformément continue, autrement dit, il n'est pas vrai que

$$\forall \epsilon > 0 \ \exists \delta > 0 : \ \forall \mathbf{x}, \mathbf{y} \in E \ \ \Big(\|\mathbf{x} - \mathbf{y}\| \le \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| \le \epsilon\Big).$$

Ainsi

$$\exists \epsilon > 0 : \ \forall \delta > 0 \ \exists \mathbf{x}, \mathbf{y} \in E \ \ \Big(\|\mathbf{x} - \mathbf{y}\| \le \delta \ \text{ et } \ |f(\mathbf{x}) - f(\mathbf{y})| > \epsilon\Big).$$

Pour un tel $\epsilon > 0$, considérons $\delta = 1/k$ avec $k \in \mathbb{N}^*$: il existe $\mathbf{x}^{(k)}, \mathbf{y}^{(k)} \in E$ tels que

$$\|\mathbf{x}^{(k)} - \mathbf{y}^{(k)}\| \leq \frac{1}{k} \quad \text{et} \quad \|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})\| > \epsilon.$$

La suite $\{\mathbf{x}^{(k)}\}_{k\in\mathbb{N}^*}$ étant bornée (puisque E est borné), il existe une sous-suite $\{\mathbf{x}^{(k_i)}\}_{i\in\mathbb{N}}$ qui converge vers un certain $\mathbf{x} \in E$ puisque E est fermé. On a alors

$$\|\mathbf{y}^{(k_i)} - \mathbf{x}\| \leq \|\mathbf{y}^{(k_i)} - \mathbf{x}^{(k_i)}\| + \|\mathbf{x}^{(k_i)} - \mathbf{x}\| \leq \frac{1}{k_i} + \|\mathbf{x}^{(k_i)} - \mathbf{x}\| \to 0, \quad \text{si} \quad i \to \infty.$$

Ainsi $\lim_{i\to\infty} \mathbf{y}^{(k_i)} = \mathbf{x}$. Puisque $\mathbf{f}$ est continue sur E on a

$$\lim_{i\to\infty} \mathbf{f}(\mathbf{x}^{(k_i)}) = \mathbf{f}(\mathbf{x}) = \lim_{i\to\infty} \mathbf{f}(\mathbf{y}^{(k_i)}),$$

ce qui contredit $\|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})\| > \epsilon$ pour tout k. $\square$

Chapitre 3

Dérivabilité

On commence ce chapitre en rappelant la notion de dérivée d'une fonction réelle d'une seule variable réelle, $f : E \subset \mathbb{R} \to \mathbb{R}$, en un point $x_0 \in \mathring{E}$:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad \text{si la limite existe.}$$

Elle représente le taux d'accroissement de la fonction en x_0. De la définition, il suit immédiatement qu'il existe $\delta > 0$ tel que $]x_0 - \delta, x_0 + \delta[\subset E$ et

$$\forall h \in]-\delta, \delta[\quad f(x_0 + h) = f(x_0) + f'(x_0)h + o(|h|)$$

c'est-à-dire,

$$f(x_0 + h) - f(x_0) = \underbrace{f'(x_0)h}_{\text{application linéaire en } h} + o(|h|)$$

et l'incrément $f(x_0 + h) - f(x_0)$ est bien approché par une application linéaire en h. Ici $o(|h|)$ dénote une fonction g définie sur un ouvert contenant 0 (un "voisinage" de 0) et telle que $\lim_{h \to 0} \frac{g(h)}{h} = 0$ (voir plus loin pour une définition plus générale). En fait il suffit de poser $g(h) = f(x_0 + h) - f(x_0) - f'(x_0)h$ pour $|h| < \delta$.

On a donc une double interprétation de la notion de dérivée : comme limite du taux d'accroissement en x_0, ainsi que comme application linéaire approchant localement la fonction dans un voisinage de x_0 (on appelle voisinage de x_0 un ouvert contenant x_0). En particulier, f dérivable en x_0 implique que f est aussi continue en x_0. Dans ce chapitre, on va généraliser ces deux notions pour des fonctions $f : \mathbb{R}^n \to \mathbb{R}$ ou $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ de plusieurs variables réelles.

3.1 Dérivées partielles, dérivées directionnelles, différentielle

Définition 3.1. *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables réelles, $\mathbf{x}_0 \in \mathring{E}$ un point intérieur de E et $\mathbf{v}$ un vecteur arbitraire de $\mathbb{R}^n$. On dit que f est dérivable dans la direction $\mathbf{v}$ au point $\mathbf{x}_0$ si la limite $\lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$ existe. Dans ce cas on pose*

$$D_{\mathbf{v}} f(\mathbf{x}_0) = \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$

*Si $\|\mathbf{v}\| = 1$ pour la norme euclidienne, on appelle $D_{\mathbf{v}}f(\mathbf{x}_0)$ la **dérivée directionnelle** de f dans la direction $\mathbf{v}$ au point $\mathbf{x}_0$.*

Si on définit la fonction $\tilde{f}(t) = f(\mathbf{x}_0 + t\mathbf{v})$ dans un voisinage de 0 (ce qui est toujours possible car $\mathbf{x}_0$ est un point intérieur de E), alors $D_{\mathbf{v}}f(\mathbf{x}_0) = \tilde{f}'(0)$. On peut donc interpréter la dérivée directionnelle $D_{\mathbf{v}}f(\mathbf{x}_0)$ comme la limite du taux d'accroissement en suivant la direction $\mathbf{v}$. En particulier, si on prend $\mathbf{v} = \mathbf{e}_i$, le i-ème vecteur de la base

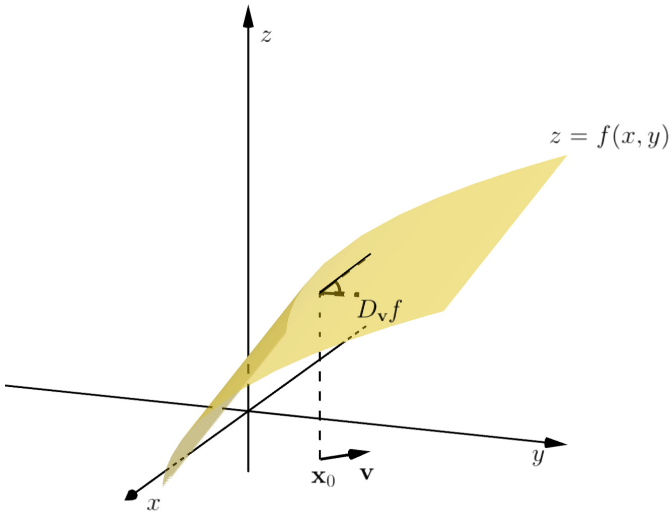


Figure 3.1 – Interprétation géométrique de dérivée directionnelle pour une fonction $f(x, y)$ de deux variables réelles

canonique de $\mathbb{R}^n$, alors la dérivée directionnelle correspondante est appelée *i-ème dérivée partielle* et est notée

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = D_{\mathbf{e}_i}f(\mathbf{x}_0) = \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_i) - f(\mathbf{x}_0)}{t}.$$

Autrement dit, pour $\mathbf{x} = (x_1, \ldots, x_n) \in \mathring{E}$ point intérieur,

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{t \to 0} \frac{f(x_1, \ldots, x_i + t, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{t},$$

ce qui revient à calculer la dérivée de la fonction $x_i \mapsto f(\cdot, x_i, \cdot)$ pensée comme une fonction uniquement de la variable x_i, en traitant les autres variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ comme des paramètres. Pour cela, on peut donc utiliser les règles de dérivation des fonctions d'une seule variable réelle.

Définition 3.2 (vecteur gradient)**.** *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et $\mathbf{x}_0$ un point intérieur de E. Si toutes les dérivées partielles de f existent en $\mathbf{x}_0$, on appelle **matrice jacobienne** $Df(\mathbf{x}_0) \in \mathbb{R}^{n\times 1}$ (vecteur ligne) le vecteur*

$$Df(\mathbf{x}_0) = \left[\tfrac{\partial f}{\partial x_1}(\mathbf{x}_0), \ldots, \tfrac{\partial f}{\partial x_n}(\mathbf{x}_0)\right] = \left(\tfrac{\partial f}{\partial x_1}(\mathbf{x}_0) \quad \ldots \quad \tfrac{\partial f}{\partial x_n}(\mathbf{x}_0)\right)$$

et ***vecteur gradient****, noté* $\nabla f(\mathbf{x}_0)$*, son transposé*

$$\nabla f(\mathbf{x}_0) = Df(\mathbf{x}_0)^\top = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}.$$

Exemple 3.3. *Soit* $f(x, y) = x^2 \cos y$*. Alors*

$$\frac{\partial f}{\partial x}(x, y) = 2x \cos y, \quad \frac{\partial f}{\partial y}(x, y) = -x^2 \sin y, \quad \nabla f(x, y) = \begin{bmatrix} 2x \cos y \\ -x^2 \sin y \end{bmatrix}.$$

Une fonction peut avoir des dérivées partielles ou directionnelles en un point $\mathbf{x}_0$ sans pour autant être continue en ce point, comme les exemples suivants le montrent.

Exemple 3.4. *Soit* $f : \mathbb{R}^2 \to \mathbb{R}$ *donnée par*

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$$

Alors,

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t\to 0} \frac{f(t, 0) - f(0, 0)}{t} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{t\to 0} \frac{f(0, t) - f(0, 0)}{t} = 0$$

mais f *n'est pas continue en* $(0, 0)$*. En effet, on a* $\lim_{t\to 0} f(t, t) = \lim_{t\to 0} \frac{t^2}{2t^2} = \frac{1}{2} \neq 0$*. Pour cette fonction, les autres dérivées directionnelles n'existent pas :*

$$\lim_{t\to 0} \frac{f(tv_1, tv_2) - f(0, 0)}{t} = \lim_{t\to 0} \frac{t^2 v_1 v_2}{t^3(v_1^2 + v_2^2)} = \lim_{t\to 0} \frac{v_1 v_2}{t(v_1^2 + v_2^2)}$$

n'existe pas si $v_1 v_2 \neq 0$ *(ici* $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$*,* $\|\mathbf{v}\| = 1$ *pour la norme euclidienne)*

Exemple 3.5. *Soit* $f : \mathbb{R}^2 \to \mathbb{R}$ *donnée par*

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

Toutes les dérivées directionnelles existent. En effet,

$$D_{\mathbf{v}} f(0, 0) = \lim_{t\to 0} \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4} = \begin{cases} \frac{v_2^2}{v_1}, & si\ v_1 \neq 0, \\ 0, & si\ v_1 = 0. \end{cases}$$

Toutefois, f *n'est pas continue en* $(0, 0)$ *car* $\lim_{t\to 0} f(t^2, t) = \frac{1}{2} \neq f(0, 0)$*.*

Comme l'exemple précédent le montre, la notion de dérivée partielle est trop faible et n'implique pas la continuité de la fonction. Si on veut introduire une notion de dérivabilité qui permet d'approcher une fonction au voisinage de $\mathbf{x}_0$ par une fonction affine, on a besoin d'une définition de dérivée plus forte. On rappelle qu'une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ est telle que $L(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ et $\forall \alpha, \beta \in \mathbb{R}$. Elle peut être représentée par la matrice $A \in \mathbb{R}^{m\times n}$ (m lignes et n colonnes) telle que $L(\mathbf{e}_i)$ est la $i^{\text{ème}}$ colonne de A, où $\{\mathbf{e}_i\}_{i=1}^n$ est la base canonique de $\mathbb{R}^n$, de telle sorte que $L(\mathbf{x}) = A \cdot \mathbf{x} = A\mathbf{x}$, $\forall \mathbf{x} \in \mathbb{R}^n$ (produit matriciel entre une matrice $m \times n$ et un vecteur colonne dans $\mathbb{R}^n$). Le produit matriciel $A\mathbf{x} \in \mathbb{R}^m$ donne dans cette formule un vecteur colonne.

Définition 3.6 (Dérivabilité et différentielle)**.** *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et $\mathbf{x}_0$ un point intérieur de E. On dit que f est différentiable (ou dérivable) en $\mathbf{x}_0$ s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}$ et une fonction $g : E \to \mathbb{R}$ tels que*

$$\forall \mathbf{x} \in E \quad f(\mathbf{x}) = f(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + g(\mathbf{x}), \quad \textit{et} \quad \lim_{\mathbf{x}\to\mathbf{x}_0} \frac{g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

L'application linéaire L est alors appelée la ***différentielle*** *de f en $\mathbf{x}_0$.*

On remarque que la fonction g satisfait nécessairement $g(\mathbf{x}_0) = 0$ et est continue en $\mathbf{x}_0$.

Notation. *Pour $p \geq 0$, on utilise souvent la notation $g(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|^p)$ pour indiquer une fonction g définie au moins sur $B(\mathbf{x}_0, \delta)\backslash\{\mathbf{x}_0\}$ pour un certain $\delta > 0$ et telle que* $\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{g(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|^p} = 0$.

Avec la notation $o(\cdot)$, dans la définition 3.6 on peut ainsi écrire $f(\mathbf{x}) = f(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$ sur E (ici $p = 1$ et g est définie sur E).

Le théorème suivant met en relation l'application linéaire L de la définition précédente avec la matrice jacobienne (ou bien le gradient) de f en $\mathbf{x}_0$ et montre aussi que la différentielle de f en $\mathbf{x}_0$, si elle existe, est unique.

Théorème 3.7. *Soit $f : E \to \mathbb{R}$ différentiable en $\mathbf{x}_0 \in \mathring{E}$. Alors, toutes les dérivées partielles de f existent en $\mathbf{x}_0$ et la différentielle de f en $\mathbf{x}_0$ est unique, donnée par*

$$L(\mathbf{x} - \mathbf{x}_0) = Df(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

(produit matriciel entre un vecteur ligne et un vecteur colonne), ce qui s'écrit aussi $L(\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0)^\top \cdot (\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0)^\top(\mathbf{x} - \mathbf{x}_0)$ (produit matriciel). On peut aussi utiliser le produit scalaire usuel entre deux vecteurs colonnes : $L(\mathbf{x}-\mathbf{x}_0) = \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)$ (produit scalaire). De plus, f est continue en $\mathbf{x}_0$.

Démonstration. On note $a_i = L(\mathbf{e}_i)$ de sorte que $L(\mathbf{x} - \mathbf{x}_0) = \sum_{i=1}^n a_i(x_i - x_{0_i})$. Par définition de la différentiabilité en $\mathbf{x}_0$, on a $f(\mathbf{x}_0 + t\mathbf{e}_i) = f(\mathbf{x}_0) + tL(\mathbf{e}_i) + g(\mathbf{x}_0 + t\mathbf{e}_i)$ et, si on note $\mathbf{x}_t = \mathbf{x}_0 + t\mathbf{e}_i$,

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) &= \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_i) - f(\mathbf{x}_0)}{t} \\ &= L(\mathbf{e}_i) + \lim_{t\to 0} \frac{g(\mathbf{x}_0 + t\mathbf{e}_i)}{t} \\ &= L(\mathbf{e}_i) + \underbrace{\lim_{t\to 0} \frac{\operatorname{sign}(t)g(\mathbf{x}_t)}{\|\mathbf{x}_t - \mathbf{x}_0\|}}_{=0} = a_i, \end{aligned}$$

donc toutes les dérivées partielles existent et $L(\mathbf{x}-\mathbf{x}_0) = Df(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)$. De plus, on a

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \underbrace{\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)(x_i - x_{0_i})}_{=0} + \underbrace{\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x})}_{=0} = f(\mathbf{x}_0),$$

ce qui montre que f est continue en $\mathbf{x}_0$.

$\square$

Une conséquence immédiate du théorème précédent est que, si $f : E \subset \mathbb{R}^n \to \mathbb{R}$ est différentiable en $\mathbf{x}_0 \in \mathring{E}$, alors toutes les dérivées directionnelles existent et $D_{\mathbf{v}}f(\mathbf{x}_0) = Df(\mathbf{x}_0)\cdot\mathbf{v} = \nabla f(\mathbf{x}_0)^\top\mathbf{v}$ (produit matriciel). En effet,

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \lim_{t\to 0}\frac{f(\mathbf{x}_0+t\mathbf{v}) - f(\mathbf{x}_0)}{t} = L(\mathbf{v}) + \underbrace{\lim_{t\to 0}\frac{g(\mathbf{x}_0+t\mathbf{v})}{t}}_{=0} = L(\mathbf{v}) = \nabla f(\mathbf{x}_0)^\top\mathbf{v}.$$

En particulier, $D_{\mathbf{v}}f(\mathbf{x}_0)$ est dans ce cas une application linéaire en $\mathbf{v}$. On a de plus, pour la norme euclidienne,

$$\|\nabla f(\mathbf{x}_0)\| = \max_{\substack{\mathbf{v}\in\mathbb{R}^n\\ \|\mathbf{v}\|=1}} \nabla f(\mathbf{x}_0)^\top\mathbf{v} = \max_{\substack{\mathbf{v}\in\mathbb{R}^n\\ \|\mathbf{v}\|=1}} D_{\mathbf{v}}f(\mathbf{x}_0),$$

le maximum étant atteint en $\mathbf{v} = \frac{\nabla f(\mathbf{x}_0)}{\|\nabla f(\mathbf{x}_0)\|}$ si $\nabla f(\mathbf{x}_0)$ n'est pas nul. S'il n'est pas nul, le vecteur gradient donne donc la direction de croissance maximale pour la fonction f. On a aussi que toute direction $\mathbf{w} \perp \nabla f(\mathbf{x}_0)$ est une direction de croissance nulle.

Exemple 3.8. *Considérons la fonction* $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = x^2+y^2$. *Son gradient est donné par* $\nabla f(x,y) = (2x,2y)^\top$. *En* $(x,y) \in \mathbb{R}^2\backslash\{(0,0)\}$, *la direction de croissance maximale est* $\mathbf{v} = \frac{\nabla f(x,y)}{\|\nabla f(x,y)\|} = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right)^\top$ *(norme euclidienne normalisée à* 1 *ici), la direction de croissance minimale est* $\mathbf{v} = -\frac{\nabla f(x,y)}{\|\nabla f(x,y)\|} = -\left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right)^\top$, *tandis que la direction de croissance nulle est* $\mathbf{w} = \pm\left(-\frac{y}{\sqrt{x^2+y^2}}, \frac{x}{\sqrt{x^2+y^2}}\right)^\top$.

Enfin, si f est différentiable en $\mathbf{x}_0$, on peut construir l'hyperplan tangent au graph de f en $(\mathbf{x}_0, f(\mathbf{x}_0))$, ainsi que un vecteur normal au graphe de f en $(\mathbf{x}_0, f(\mathbf{x}_0))$.

Définition 3.9. *Soit* $f : E \subset \mathbb{R}^n \to \mathbb{R}$ *différentiable en* $\mathbf{x}_0 \in \mathring{E}$ *et notons* $y_0 = f(\mathbf{x}_0)$ *et* $\mathcal{G}_f = \{(\mathbf{x},y) \in \mathbb{R}^{n+1} :\ \mathbf{x}\in E,\ y = f(\mathbf{x})\}$ *le graphe de* f. *On définie* **hyperplan tangent** *à* $\mathcal{G}_f$ *en* $(\mathbf{x}_0, y_0)$ *le sous-ensemble de* $\mathbb{R}^{n+1}$

$$\Pi_{(\mathbf{x}_0,y_0)}(\mathcal{G}_f) = \{(\mathbf{x},y)\in\mathbb{R}^{n+1} :\ y = y_0 + \nabla f(\mathbf{x}_0)^\top(\mathbf{x}-\mathbf{x}_0)\}$$

Le vecteur

$$\mathbf{n} = \frac{1}{\sqrt{1+\|\nabla f(\mathbf{x}_0)\|^2}}\left(-\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \cdots, -\frac{\partial f}{\partial x_n}(\mathbf{x}_0), 1\right)^\top \subset \mathbb{R}^{n+1}$$

est appelé le **vecteur normal** *au graphe de* f *en* $(\mathbf{x}_0, y_0)$.

En effet, on vérifie facilement que le vecteur $\mathbf{n}$ est normal à l'hyperplan tangent $\Pi_{(\mathbf{x}_0,y_0)}(\mathcal{G}_f)$ en $(\mathbf{x}_0, y_0)$, c'est à dire il satisfait

$$\mathbf{n} \cdot ((\mathbf{x}, y) - (\mathbf{x}_0, y_0)) = 0, \qquad \forall (\mathbf{x}, y) \in \Pi_{(\mathbf{x}_0,y_0)}(\mathcal{G}_f).$$

Les définitions de différentiabilité et dérivées partielles/directionnelles s'étendent sans difficultés aux fonctions à valeurs dans $\mathbb{R}^m$.

Définition 3.10. *Soit* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$,

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))^\top = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} = \begin{pmatrix} f_1(\mathbf{x}) \\ \cdots \\ f_m(\mathbf{x}) \end{pmatrix},$$

noté aussi sous forme ligne (nous préciserons si nécessaire))

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})) = \begin{bmatrix} f_1(\mathbf{x}) & \ldots & f_m(\mathbf{x}) \end{bmatrix} = \begin{pmatrix} f_1(\mathbf{x}) & \ldots & f_m(\mathbf{x}) \end{pmatrix},$$

et soit $\mathbf{x}_0 \in \mathring{E}$. *On appelle* dérivée partielle de la i-ème composante de $\mathbf{f}$ par rapport à la j-ème variable *en* $\mathbf{x}_0$*, notée* $\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$*, la quantité*

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) = \lim_{t \to 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{e}_j) - f_i(\mathbf{x}_0)}{t},$$

si la limite existe, et matrice jacobienne *de* $\mathbf{f}$ *en* $\mathbf{x}_0$ *la matrice* $D\mathbf{f}(\mathbf{x}_0) \in \mathbb{R}^{m \times n}$ *de composantes* $(D\mathbf{f}(\mathbf{x}_0))_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$*, autrement dit,*

$$D\mathbf{f}(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \cdots & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \cdots & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix},$$

si toutes les dérivées partielles existent.

De façon similaire, on introduit la notion de différentiabilité.

Définition 3.11. *On dit que* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *est différentiable (ou dérivable) en* $\mathbf{x}_0 \in \mathring{E}$ *s'il existe une application linéaire* $L : \mathbb{R}^n \to \mathbb{R}^m$ *et une fonction* $\mathbf{g} : E \to \mathbb{R}^m$ *tels que*

$$\forall \mathbf{x} \in E \quad \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + \mathbf{g}(\mathbf{x})$$

et $\mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|)$*, c.-à-d.* $\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|\mathbf{g}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$.

Si plus généralement $\|\mathbf{g}(\mathbf{x})\| = o(\|\mathbf{x}-\mathbf{x}_0\|^p)$ dans le sens déjà introduit, on notera simplement $\mathbf{g}(\mathbf{x}) = o(\|\mathbf{x}-\mathbf{x}_0\|^p)$. Si $\mathbf{f}$ est différentiable en $\mathbf{x}_0$, alors toutes les dérivées partielles et directionnelles existent, et $\mathbf{f}$ est continue en $\mathbf{x}_0$. De plus, l'application linéaire L est unique et donnée par

$$L(\mathbf{x}-\mathbf{x}_0) = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0_j})\mathbf{e}_i$$

où $\mathbf{e}_i$ est le $i^{\text{ème}}$ vecteur de la base canonique de $\mathbb{R}^m$. Ceci s'écrit aussi

$$L(\mathbf{x}-\mathbf{x}_0) = D\mathbf{f}(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) = D\mathbf{f}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0),$$

où intervient ici le produit matriciel entre une matrice $\mathbb{R}^{m\times n}$ et un vecteur colonne dans $\mathbb{R}^n$, ce qui donne un vecteur colonne dans $\mathbb{R}^m$. La dérivée de $\mathbf{f}$ dans la direction $\mathbf{v}$ au point $\mathbf{x}_0$, $\mathbf{v} \mapsto D_{\mathbf{v}}\mathbf{f}(\mathbf{x}_0) = L(\mathbf{v}) = D\mathbf{f}(\mathbf{x}_0)\mathbf{v}$, est linéaire en $\mathbf{v} \in \mathbb{R}^n$ (si $\mathbf{f}$ est différentiable en $\mathbf{x}_0$).

La condition de différentiabilité d'une fonction $f : E \subset \mathbb{R}^n \to \mathbb{R}$ (ou $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$) n'est pas immédiate à vérifier. Heureusement, on a une condition suffisante, facile à vérifier, qui nous permet de conclure si f est différentiable en $\mathbf{x}_0 \in \mathring{E}$. On donne ici la version du résultat pour une fonction à valeurs dans $\mathbb{R}$ mais le résultat se généralise sans difficulté au cas d'une fonction à valeurs dans $\mathbb{R}^m$.

Théorème 3.12. *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et $\mathbf{x}_0 \in \mathring{E}$. S'il existe $\delta > 0$ tel que $B(\mathbf{x}_0,\delta) \subset E$ et les dérivées partielles $\frac{\partial f}{\partial x_i}(\mathbf{x})$ existent pour tout $\mathbf{x} \in B(\mathbf{x}_0,\delta)$ et sont continues en $\mathbf{x}_0$, alors f est différentiable en $\mathbf{x}_0$.*

Démonstration. On utilise la norme euclidienne et, pour alléger la notation, on renomme $\mathbf{x}_0 = \mathbf{a} = (a_1,\dots,a_n)$. Pour un $\mathbf{x} = (x_1,\dots,x_n) \in B(\mathbf{a},\delta)$ donné, on introduit la notation $\mathbf{x}^k = (x_1,\dots,x_k,a_{k+1},\dots,a_n)$ de sorte que $\mathbf{x}^n = \mathbf{x}$ et $\mathbf{x}^0 = \mathbf{a}$. Alors, la différence $f(\mathbf{x}) - f(\mathbf{a})$ peut être écrite comme somme télescopique

$$f(\mathbf{x}) - f(\mathbf{a}) = \sum_{k=1}^{n}(f(\mathbf{x}^k) - f(\mathbf{x}^{k-1})).$$

Par le théorème des accroissements finis on a que $\forall k = 1,\dots,n$, il existe un $\theta_k \in]0,1[$ tel que

$$\begin{aligned} f(\mathbf{x}^k) - f(\mathbf{x}^{k-1}) &= f(x_1,\dots,x_{k-1},x_k,a_{k+1},\dots,a_n) - f(x_1,\dots,x_{k-1},a_k,a_{k+1},\dots,a_n) \\ &= \frac{\partial f}{\partial x_k}(x_1,\dots,x_{k-1},a_k+\theta_k(x_k-a_k),a_{k+1},\dots,a_n)(x_k-a_k) \\ &= \frac{\partial f}{\partial x_k}(\mathbf{x}^{k-1}+\theta_k(\mathbf{x}^k-\mathbf{x}^{k-1}))(x_k-a_k). \end{aligned}$$

Puisque les dérivées partielles sont continues en $\mathbf{a}$, on a que

$$g_k(\mathbf{x}) := \frac{\partial f}{\partial x_k}(\mathbf{x}^{k-1}+\theta_k(\mathbf{x}^k-\mathbf{x}^{k-1})) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \xrightarrow{\mathbf{x}\to\mathbf{a}} 0.$$

En effet, la continuité de $\frac{\partial f}{\partial x_k}$ en $\mathbf{a}$ implique que

$$\forall \epsilon > 0 \ \exists \widetilde{\delta} \in]0, \delta/2] : \ \forall \mathbf{y} \in B(\mathbf{a}, 2\widetilde{\delta}) \quad \left| \frac{\partial f}{\partial x_k}(\mathbf{y}) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \right| < \epsilon,$$

et de plus, pour $\mathbf{y}(\mathbf{x}) = \mathbf{x}^{k-1} + \theta_k(\mathbf{x}^k - \mathbf{x}^{k-1})$ avec $1 \leq k \leq n$, on a

$$\|\mathbf{y}(\mathbf{x}) - \mathbf{a}\| = \|(1-\theta_k)(\mathbf{x}^{k-1} - \mathbf{a}) + \theta_k(\mathbf{x}^k - \mathbf{a})\| \leq \|\mathbf{x}^{k-1} - \mathbf{a}\| + \|\mathbf{x}^k - \mathbf{a}\| \leq 2\|\mathbf{x} - \mathbf{a}\|$$

et donc

$$\forall \epsilon > 0 \ \exists \widetilde{\delta} \in]0, \delta/2] : \ \forall \mathbf{x} \in B(\mathbf{a}, \widetilde{\delta}) \quad |g_k(\mathbf{x})| = \left| \frac{\partial f}{\partial x_k}(\mathbf{y}(\mathbf{x})) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \right| < \epsilon,$$

ce qui montre que $\lim_{\mathbf{x} \to \mathbf{a}} g_k(\mathbf{x}) = 0$ pour tout $k = 1, \ldots, n$. Ainsi,

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{a}) &= \sum_{k=1}^{n} (f(\mathbf{x}^k) - f(\mathbf{x}^{k-1})) \\ &= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{x}^{k-1} + \theta_k(\mathbf{x}^k - \mathbf{x}^{k-1}))(x_k - a_k) \\ &= \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(\mathbf{a})(x_k - a_k) + g_k(\mathbf{x})(x_k - a_k) \right) \\ &= Df(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + g(\mathbf{x}) \end{aligned}$$

(produit matriciel entre un vecteur ligne et un vecteur colonne), avec $g(\mathbf{x}) = \sum_{k=1}^{n} g_k(\mathbf{x})(x_k - a_k)$. Par l'inégalité de Cauchy-Schwarz, on a $|g(\mathbf{x})| \leq \left(\sum_{k=1}^{n} g_k(\mathbf{x})^2\right)^{\frac{1}{2}} \|\mathbf{x} - \mathbf{a}\|$ et donc

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{|g(\mathbf{x})|}{\|\mathbf{x} - \mathbf{a}\|} \leq \lim_{\mathbf{x} \to \mathbf{a}} \sqrt{\sum_{k=1}^{n} g_k(\mathbf{x})^2} = 0$$

ce qui montre que $g(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|)$. □

3.1.1 L'espace C^1

Le théorème 3.12 montre que si toutes les dérivées partielles existent dans un voisinage d'un point intérieur $\mathbf{x}_0$ et sont continues en $\mathbf{x}_0$, alors la fonction est différentiable (et continue) en $\mathbf{x}_0$ et toutes les dérivées directionnelles existent aussi et sont continues.

Définition 3.13 (Espace C^1). *Soit $E \subset \mathbb{R}^n$ ouvert non vide. On dit que $f : E \to \mathbb{R}$ est continûment différentiable sur E si toutes les dérivées partielles $\frac{\partial f}{\partial x_i}$ pour $i = 1, \ldots, n$, existent et sont continues en tout point $\mathbf{x} \in E$. Dans ce cas on note $f \in C^1(E)$ (f est de classe C^1). De même, on dit que $\mathbf{f} : E \to \mathbb{R}^m$ est de classe C^1, $\mathbf{f} \in C^1(E, \mathbb{R}^m)$, si chaque composante de $\mathbf{f} = (f_1, \ldots, f_m)$ satisfait $f_k \in C^1(E)$, $k = 1, \ldots, m$.*

Grâce au théorème 3.12, si $f \in C^1(E)$ alors f est différentiable en tout point $\mathbf{x}_0 \in E$ (E étant ouvert) et donc elle est aussi continue, c'est-à-dire $f \in C^1(E) \Rightarrow f \in C^0(E)$; de même, on a $C^1(E, \mathbb{R}^m) \subset C^0(E, \mathbb{R}^m)$. De plus, toutes les dérivées directionnelles $D_{\mathbf{v}} f$ existent et sont continues sur E.

On vérifie facilement que $C^1(E)$ a une structure d'espace vectoriel réel. En effet, pour tous $f, g \in C^1(E)$ et toutes constantes $\lambda, \mu \in \mathbb{R}$ on a $\lambda f + \mu g \in C^1(E)$. On vérifie aussi facilement les règles de dérivation suivantes : Soit $E \subset \mathbb{R}^n$ ouvert, $f, g \in C^1(E)$ et des constantes $\lambda, \mu \in \mathbb{R}$. Alors

$$
\begin{array}{lll}
\text{i)} & \lambda f + \mu g \in C^1(E), & D(\lambda f + \mu g)(\mathbf{x}) = \lambda Df(\mathbf{x}) + \mu Dg(\mathbf{x}), \quad \forall \mathbf{x} \in E, \\
\text{ii)} & fg \in C^1(E), & D(fg)(\mathbf{x}) = Df(\mathbf{x})g(\mathbf{x}) + f(\mathbf{x})Dg(\mathbf{x}), \quad \forall \mathbf{x} \in E.
\end{array}
$$

3.2 Dérivation de fonctions composées

Soit $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ et $\mathbf{g} : F \subset \mathbb{R}^m \to \mathbb{R}^p$ avec E, F ouverts non vides. Si $\mathrm{Im}(\mathbf{f}) \subseteq F$, on peut définir la fonction composée

$$\boldsymbol{\varphi} = \mathbf{g} \circ \mathbf{f} : E \to \mathbb{R}^p, \quad \boldsymbol{\varphi}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})), \quad \forall \mathbf{x} \in E.$$

Par composantes :

$$
\begin{aligned}
&\mathbf{x} = (x_1, \ldots, x_n), \quad \mathbf{f}(\mathbf{x}) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)), \\
&\mathbf{y} = (y_1, \ldots, y_m), \quad \mathbf{g}(\mathbf{y}) = (g_1(y_1, \ldots, y_m), \ldots, g_p(y_1, \ldots, y_m)), \\
&\boldsymbol{\varphi}(\mathbf{x}) = \Big(g_1(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))\,,\ \ldots\ ,\ g_p(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))\Big).
\end{aligned}
$$

Théorème 3.14. *Soient $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m, \mathbf{g} : F \subset \mathbb{R}^m \to \mathbb{R}^p$; E, F des ouverts contenant respectivement $\mathbf{x}_0$ et $\mathbf{y}_0$; $\mathrm{Im}(\mathbf{f}) \subseteq F$ et $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Si $\mathbf{f}$ est différentiable en $\mathbf{x}_0$ et $\mathbf{g}$ est différentiable en $\mathbf{y}_0$, alors $\boldsymbol{\varphi} = \mathbf{g} \circ \mathbf{f} : E \to \mathbb{R}^p$ est différentiable en $\mathbf{x}_0$ et*

$$D\boldsymbol{\varphi}(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \cdot D\mathbf{f}(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0))D\mathbf{f}(\mathbf{x}_0).$$

(produit matriciel des matrices jacobiennes). Par composantes on a :

$$\forall i \in \{1, \ldots, p\}\ \forall j \in \{1, \ldots, n\} \quad \frac{\partial \varphi_i}{\partial x_j}(\mathbf{x}_0) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(\mathbf{f}(\mathbf{x}_0)) \frac{\partial f_k}{\partial x_j}(\mathbf{x}_0).$$

Si de plus $\mathbf{f} \in C^1(E, \mathbb{R}^m)$ et $\mathbf{g} \in C^1(F, \mathbb{R}^p)$ alors $\boldsymbol{\varphi} \in C^1(E, \mathbb{R}^p)$.

Démonstration. Par hypothèse on a

$$
\begin{aligned}
&\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{R}_{\mathbf{f}}(\mathbf{x}), \quad \lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\mathbf{R}_{\mathbf{f}}(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}, \\
&\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{y}_0) + D\mathbf{g}(\mathbf{y}_0) \cdot (\mathbf{y} - \mathbf{y}_0) + \mathbf{R}_{\mathbf{g}}(\mathbf{y}), \quad \lim_{\mathbf{y} \to \mathbf{y}_0} \frac{\mathbf{R}_{\mathbf{g}}(\mathbf{y})}{\|\mathbf{y} - \mathbf{y}_0\|} = \mathbf{0}.
\end{aligned}
$$

On écrit

$$\begin{aligned}
\mathbf{R_f}(\mathbf{x}) &= \|\mathbf{x}-\mathbf{x}_0\|\mathbf{r_f}(\mathbf{x}) \quad \text{avec} \quad \lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{r_f}(\mathbf{x}) = \mathbf{0} \text{ et } \mathbf{r_f}(\mathbf{x}_0) := \mathbf{0},\\
\mathbf{R_g}(\mathbf{y}) &= \|\mathbf{y}-\mathbf{y}_0\|\mathbf{r_g}(\mathbf{y}) \quad \text{avec} \quad \lim_{\mathbf{y}\to\mathbf{y}_0} \mathbf{r_g}(\mathbf{y}) = \mathbf{0} \text{ et } \mathbf{r_g}(\mathbf{y}_0) := \mathbf{0}.
\end{aligned}$$

On veut montrer qu'il existe une application linéaire $\mathbf{L}_\varphi : \mathbb{R}^n \to \mathbb{R}^p$ et une fonction $\mathbf{R}_\varphi : E \to \mathbb{R}^p$ tels que

$$\varphi(\mathbf{x}) = \varphi(\mathbf{x}_0) + \mathbf{L}_\varphi(\mathbf{x}-\mathbf{x}_0) + \mathbf{R}_\varphi(\mathbf{x}) \quad \text{et} \quad \lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\mathbf{R}_\varphi(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|} = \mathbf{0}.$$

On a :

$$\begin{aligned}
\varphi(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})) &= \mathbf{g}\Big(\mathbf{f}(\mathbf{x}_0) + \overbrace{D\mathbf{f}(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) + \|\mathbf{x}-\mathbf{x}_0\|\mathbf{r_f}(\mathbf{x})}^{\mathbf{h}(\mathbf{x})}\Big)\\
&= \mathbf{g}(\mathbf{y}_0) + D\mathbf{g}(\mathbf{y}_0)\cdot\mathbf{h}(\mathbf{x}) + \|\mathbf{h}(\mathbf{x})\|\mathbf{r_g}(\mathbf{f}(\mathbf{x}))\\
&= \mathbf{g}(\mathbf{y}_0) + D\mathbf{g}(\mathbf{y}_0)\cdot D\mathbf{f}(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) + \underbrace{\|\mathbf{x}-\mathbf{x}_0\|D\mathbf{g}(\mathbf{y}_0)\cdot\mathbf{r_f}(\mathbf{x})}_{\mathbf{A}(\mathbf{x})} + \underbrace{\|\mathbf{h}(\mathbf{x})\|\mathbf{r_g}(\mathbf{f}(\mathbf{x}))}_{\mathbf{B}(\mathbf{x})}.
\end{aligned}$$

Il faut montrer que $\mathbf{R}_\varphi(\mathbf{x}) := \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})$ satisfait $\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\mathbf{R}_\varphi(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|} = \mathbf{0}$. Pour $\mathbf{A}(\mathbf{x})$ on a

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\|\mathbf{A}(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{x}_0\|} = \lim_{\mathbf{x}\to\mathbf{x}_0} \|D\mathbf{g}(\mathbf{y}_0)\cdot\mathbf{r_f}(\mathbf{x})\| \leq |||D\mathbf{g}(\mathbf{y}_0)||| \lim_{\mathbf{x}\to\mathbf{x}_0} \|\mathbf{r_f}(\mathbf{x})\| = 0,$$

où on a noté $|||C|||$ la norme matricielle $|||C||| = \sup_{\substack{\mathbf{y}\in\mathbb{R}^m\\ \mathbf{y}\neq\mathbf{0}}} \frac{\|C\mathbf{y}\|}{\|\mathbf{y}\|} < +\infty$ pour $C \in \mathbb{R}^{p\times m}$ (et on utilise finalement le théorème des deux gendarmes).
Pour $\mathbf{B}(\mathbf{x})$ on a pour tout $\mathbf{x}_0 \neq \mathbf{x} \in E$

$$\begin{aligned}
\frac{\|\mathbf{B}(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{x}_0\|} &= \frac{\|\mathbf{h}(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{x}_0\|}\|\mathbf{r_g}(\mathbf{f}(\mathbf{x}))\|\\
&\leq \frac{\|D\mathbf{f}(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)\| + \|\mathbf{x}-\mathbf{x}_0\|\|\mathbf{r_f}(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{x}_0\|}\|\mathbf{r_g}(\mathbf{f}(\mathbf{x}))\|\\
&\leq (|||D\mathbf{f}(\mathbf{x}_0)||| + \|\mathbf{r_f}(\mathbf{x})\|)\|\mathbf{r_g}(\mathbf{f}(\mathbf{x}))\|
\end{aligned}$$

Puisque $\mathbf{f}$ est continue en $\mathbf{x}_0$ et $\mathbf{r_g}$ en $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$, la composée $\mathbf{r_g}\circ\mathbf{f}$ est continue en $\mathbf{x}_0$. D'où $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{r_g}(\mathbf{f}(\mathbf{x})) = \mathbf{r_g}(\mathbf{f}(\mathbf{x}_0)) = \mathbf{r_g}(\mathbf{y}_0) = \mathbf{0}$. D'autre part $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{r_f}(\mathbf{x}) = \mathbf{0}$ (par hypothèse) et on conclut que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\|\mathbf{B}(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{x}_0\|} \leq (|||D\mathbf{f}(\mathbf{x}_0)||| + \lim_{\mathbf{x}\to\mathbf{x}_0} \|\mathbf{r_f}(\mathbf{x})\|) \lim_{\mathbf{x}\to\mathbf{x}_0} \|\mathbf{r_g}(\mathbf{f}(\mathbf{x}))\| = 0.$$

On a donc montré que

$$\varphi(\mathbf{x}) = \varphi(\mathbf{x}_0) + D\mathbf{g}(\mathbf{f}(\mathbf{x}_0))\cdot D\mathbf{f}(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) + \mathbf{R}_\varphi(\mathbf{x}) \quad \text{où} \quad \lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\mathbf{R}_\varphi(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|} = \mathbf{0}.$$

On conclut donc que $\boldsymbol{\varphi}$ est différentiable en $\mathbf{x}_0$ et

$$D\boldsymbol{\varphi}(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \cdot D\mathbf{f}(\mathbf{x}_0).$$

Si maintenant $\mathbf{f} \in C^1(E, \mathbb{R}^m)$ et $\mathbf{g} \in C^1(F, \mathbb{R}^p)$, alors chaque composante des matrices $D\mathbf{f}(\mathbf{x})$ et $D\mathbf{g}(\mathbf{y})$ est une fonction continue de $\mathbf{x}$ (resp. $\mathbf{y}$) et $D\boldsymbol{\varphi}(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x})) \cdot D\mathbf{f}(\mathbf{x})$ est continue en $\mathbf{x}$ pour tout $\mathbf{x} \in E$. Donc $\boldsymbol{\varphi} \in C^1(E, \mathbb{R}^p)$. $\square$

Remarque 3.15. *Dans cette preuve, on a utilisé la norme matricielle* $|||C||| = \sup_{\substack{\mathbf{y}\in\mathbb{R}^m \\ \mathbf{y}\neq\mathbf{0}}} \frac{\|C\mathbf{y}\|}{\|\mathbf{y}\|} < +\infty$ *pour* $C \in \mathbb{R}^{p\times m}$*. C'est effectivement une norme sur l'espace vectoriel des matrices* $\mathbb{R}^{p\times m}$*, appelée plus spécifiquement la norme "spectrale". Il y a d'autres normes matricielles possibles sur* $\mathbb{R}^{p\times m}$*.*

Exemple 3.16. *Soient* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ *et* $\mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^2$ *donnés par*

$$\mathbf{f}(x, y) = \begin{bmatrix} xy \\ x + 2y \\ \sin x \end{bmatrix}, \quad \mathbf{g}(u, v, w) = \begin{bmatrix} u + w \\ v^2 \end{bmatrix},$$

avec matrices jacobiennes

$$D\mathbf{f}(x, y) = \begin{bmatrix} y & x \\ 1 & 2 \\ \cos x & 0 \end{bmatrix} \in \mathbb{R}^{3\times 2}, \quad D\mathbf{g}(u, v, w) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2v & 0 \end{bmatrix} \in \mathbb{R}^{2\times 3}.$$

Alors,

$$\boldsymbol{\varphi}(x, y) = \mathbf{g} \circ \mathbf{f}(x, y) = \begin{bmatrix} xy + \sin x \\ (x + 2y)^2 \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$$

et

$$D\boldsymbol{\varphi}(x, y) = \begin{bmatrix} y + \cos x & x \\ 2(x + 2y) & 4(x + 2y) \end{bmatrix} = D\mathbf{g}(\mathbf{f}(x, y)) \cdot D\mathbf{f}(x, y) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2v & 0 \end{bmatrix} \begin{bmatrix} y & x \\ 1 & 2 \\ \cos x & 0 \end{bmatrix}.$$

Un cas particulier du théorème 3.14 est celui de la composition d'une fonction $f : E \subset \mathbb{R}^n \to \mathbb{R}$, de classe C^1 sur l'ouvert non vide E et un chemin dans E, $\boldsymbol{\gamma} : I \subset \mathbb{R} \to E$, I ouvert non vide et $t \mapsto \boldsymbol{\gamma}(t) = (\gamma_1(t), \dots, \gamma_n(t))$, avec $\gamma_i \in C^1(I)$, $\forall i = 1, \dots, n$. Dans ce cas, la fonction composée $\varphi = f \circ \boldsymbol{\gamma} : I \to \mathbb{R}$, $t \mapsto \varphi(t) = f(\gamma_1(t), \dots, \gamma_n(t))$ est dérivable sur I, $\varphi \in C^1(I)$ et

$$\begin{aligned}
\frac{d}{dt}\varphi(t) &= Df(\boldsymbol{\gamma}(t)) \cdot D\boldsymbol{\gamma}(t) \\
&= \begin{bmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{\gamma}(t)), & \dots, & \frac{\partial f}{\partial x_n}(\boldsymbol{\gamma}(t)) \end{bmatrix} \begin{bmatrix} \dot{\gamma}_1(t) \\ \vdots \\ \dot{\gamma}_n(t) \end{bmatrix} \\
&= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{\gamma}(t))\dot{\gamma}_i(t).
\end{aligned}$$

La quantité $\frac{d}{dt}\varphi(t) = \frac{d}{dt}f(\gamma_1(t), \dots, \gamma_n(t))$ est appelée la dérivée totale de f le long du chemin $\boldsymbol{\gamma}(t)$.

3.3 Théorème des accroissements finis

On rappelle d'abord le théorème des accroissements finis pour une fonction réelle d'une seule variable réelle : soit $f : [a,b] \to \mathbb{R}$ continue et dérivable sur $]a,b[$, alors il existe $c \in]a,b[$ tel que $f(b) - f(a) = f'(c)(b-a)$. On vise à généraliser ce résultat pour des fonctions $f : U \subset \mathbb{R}^n \to \mathbb{R}$ ou $\mathbf{f} : U \subset \mathbb{R}^n \to \mathbb{R}^m$ de plusieurs variables réelles.

On introduit la notation suivante : pour $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{y}$, on note $[\mathbf{x}, \mathbf{y}]$ le segment fermé d'origine $\mathbf{x}$ et d'extrémité $\mathbf{y}$: $[\mathbf{x}, \mathbf{y}] = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}),\ t \in [0,1]\}$ et par $]\mathbf{x}, \mathbf{y}[$ le segment "ouvert" $]\mathbf{x}, \mathbf{y}[= \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}),\ t \in]0,1[\}$ (si $n \geq 2$, $]\mathbf{x}, \mathbf{y}[$ n'est ni ouvert ni fermé).

Théorème 3.17. *Soit $E \subset \mathbb{R}^n$ un ouvert non vide et $f : E \to \mathbb{R}$ une fonction dérivable sur E. Si $\mathbf{x}, \mathbf{y} \in E$ distincts sont tels que $[\mathbf{x}, \mathbf{y}] \subset E$, alors il existe $\mathbf{z} \in]\mathbf{x}, \mathbf{y}[$, tel que*

$$f(\mathbf{y}) - f(\mathbf{x}) = Df(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x})$$

(produit matriciel entre un vecteur ligne et un vecteur colonne).

Démonstration. On pose $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$ avec $t \in [0,1]$. g est continue sur $[0,1]$ et dérivable sur $]0,1[$ (composition de fonctions dérivables $f(\mathbf{v}(t))$, où $\mathbf{v}(t) = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$), et par la règle de dérivation d'une composée, $g'(t) = Df(\mathbf{v}(t)) \cdot \mathbf{v}'(t) = Df(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x})$. Par le théorème des accroissements finis pour des fonctions d'une seule variable, il existe $\theta \in]0,1[$ t.q. $g(1) - g(0) = g'(\theta)$ ce qui équivaut à

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &= Df(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x}), \qquad \theta \in]0,1[\\ &= Df(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x}), \qquad \mathbf{z} = \mathbf{x} + \theta(\mathbf{y} - \mathbf{x}) \in]\mathbf{x}, \mathbf{y}[. \end{aligned}$$

□

Dans la démonstration ci-dessus on a $g'(t) = Df(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x})$. Si on suppose, maintenant, que f est de classe C^1, on a $g \in C^1(]0,1[)$ et $g(1) - g(0) = \int_0^1 g'(t)dt$. On en déduit

$$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 Df(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x})dt. \tag{3.1}$$

Considérons maintenant une fonction $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ avec E ouvert non vide et $\mathbf{f}$ dérivable sur E. L'analogue du théorème des accroissements finis pour une telle fonction vectorielle est *en défaut* ! En fait, on peut appliquer le théorème à chaque composante $f_k,\ k = 1, \dots, m$ de $\mathbf{f}$. Donc il existe $\mathbf{z}_k \in]\mathbf{x}, \mathbf{y}[,\ k = 1, \dots, m$ tels que

$$f_k(\mathbf{y}) - f_k(\mathbf{x}) = Df_k(\mathbf{z}_k) \cdot (\mathbf{y} - \mathbf{x}).$$

Toutefois, on n'a aucune garantie que les $\mathbf{z}_k$ coïncident. On ne peut donc pas trouver, en général, un seul $\mathbf{z}$ pour lequel $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x})$ (produit matriciel entre une matrice $m \times n$ et un vecteur colonne dans $\mathbb{R}^n$, ce qui donne un vecteur colonne dans $\mathbb{R}^m$). Néanmoins, la formule (3.1) se généralise pour toute fonction vectorielle $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ de classe C^1.

Lemme 3.18. *Soit* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *de classe* C^1 *sur* E *ouvert non vide et* $\mathbf{x}, \mathbf{y} \in E$ *tels que* $[\mathbf{x}, \mathbf{y}] \subset E$. *Alors*

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \int_0^1 D\mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x}) dt.$$

Ici l'intégrale d'une fonction continue $[0, 1] \to \mathbb{R}^m$ est comprise comme le vecteur dans $\mathbb{R}^m$ obtenu en prenant l'intégrale de chaque composante (chaque composante étant une fonction continue $[0, 1] \to \mathbb{R}$).

Le Lemme précédent, demande que la fonction $\mathbf{f}$ soit de classe C^1 sur E. On peut obtenir une *version faible* du théorème des accroissements finis qui demande uniquement que les dérivées partielles existent et soient bornées sur le segment $[\mathbf{x}, \mathbf{y}]$.

Lemme 3.19. *Soit* $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *dérivable sur* E *ouvert non vide et* $\mathbf{x}, \mathbf{y} \in E$ *distincts tels que* $[\mathbf{x}, \mathbf{y}] \in E$. *S'il existe* $M > 0$ *tel que* $|||D\mathbf{f}(\mathbf{z})||| \leq M$, $\forall \mathbf{z} \in]\mathbf{x}, \mathbf{y}[$, *alors*

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M \|\mathbf{y} - \mathbf{x}\|.$$

Démonstration. Le résultat est évident si $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \mathbf{0}$. Supposons donc $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) \neq \mathbf{0}$ et considérons $\mathbf{f}(\mathbf{u})$ comme un vecteur colonne, $\mathbf{u} \in E$. Soit un vecteur ligne $\mathbf{w} \in \mathbb{R}^m$ fixé et considérons la fonction $f_{\mathbf{w}} : E \to \mathbb{R}$ définie par $f_{\mathbf{w}}(\mathbf{u}) = \mathbf{w}\mathbf{f}(\mathbf{u})$ pour $\mathbf{u} \in E$, où apparaît le produit matriciel entre $\mathbf{w}$ et $\mathbf{f}(\mathbf{u})$. On sait alors qu'il existe $\mathbf{z} \in]\mathbf{x}, \mathbf{y}[$ (qui peut dépendre de $\mathbf{w}$) tel que $f_{\mathbf{w}}(\mathbf{y}) - f_{\mathbf{w}}(\mathbf{x}) = Df_{\mathbf{w}}(\mathbf{z})(\mathbf{y} - \mathbf{x})$, c'est-à-dire

$$\mathbf{w}(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})) = \mathbf{w}D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x}).$$

Choisissons $\mathbf{w} = (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}))^\top$, ce qui donne

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\|^2 = \mathbf{w}D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x}) \leq \|\mathbf{w}\| \, \|D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x})\|$$

$$\leq \|\mathbf{w}\| \, |||D\mathbf{f}(\mathbf{z})||| \|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \, M \, \|\mathbf{y} - \mathbf{x}\|,$$

et donc $\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M \|\mathbf{y} - \mathbf{x}\|$. □

Chapitre 4

Dérivées d'ordres supérieurs

4.1 Dérivées secondes

Soit $E \subset \mathbb{R}^n$ ouvert, une fonction $f : E \to \mathbb{R}$ et un indice $j \in \{1, \ldots, n\}$ fixé, tels que $\frac{\partial f}{\partial x_j}(\mathbf{x})$ existe en tout $\mathbf{x} \in E$. On considère maintenant la fonction $\frac{\partial f}{\partial x_j} : E \to \mathbb{R}$ et soit $k \in \{1, \ldots, n\}$ fixé. Si cette fonction admet une dérivée partielle par rapport à x_k en tout $\mathbf{x} \in E$, on peut définir la dérivée partielle seconde $\frac{\partial\left(\frac{\partial f}{\partial x_j}\right)}{\partial x_k} : E \to \mathbb{R}$. On utilise la notation $\frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x}) := \frac{\partial\left(\frac{\partial f}{\partial x_j}\right)}{\partial x_k}(\mathbf{x})$. Pour $k = j$ on utilise la notation $\frac{\partial^2 f}{\partial x_j^2}(\mathbf{x}) = \frac{\partial\left(\frac{\partial f}{\partial x_j}\right)}{\partial x_j}(\mathbf{x})$.

Définition 4.1 (espace $C^2(E)$)**.** *Soit $E \subset \mathbb{R}^n$ ouvert non vide. On dit que $f : E \to \mathbb{R}$ est de classe C^2, $f \in C^2(E)$, si toutes les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j} : E \to \mathbb{R}$, $i, j = 1, \ldots, n$ existent et sont continues sur E.*

Définition 4.2 (matrice hessienne)**.** *Soit $E \subset \mathbb{R}^n$ ouvert non vide et $f \in C^2(E)$. On appelle matrice hessienne de f en $\mathbf{x} \in E$ la matrice $H_f(\mathbf{x}) \in \mathbb{R}^{n \times n}$:*

$$H_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{bmatrix}$$

$$= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}.$$

On peut aussi définir les dérivées directionnelles mixtes : soit $f : E \to \mathbb{R}$ différentiable sur E et $D_{\mathbf{v}} f(\mathbf{x}) = \nabla f(\mathbf{x})^\top \mathbf{v}$ la dérivée directionnelle de f dans la direction fixée $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$ (norme euclidienne). Si $D_{\mathbf{v}} f : E \to \mathbb{R}$ admet une dérivée directionnelle dans la direction fixée $\mathbf{w}$ en tout point $\mathbf{x} \in E$, on note

$$D^2_{\mathbf{w}\mathbf{v}} f(\mathbf{x}) = D_{\mathbf{w}}(D_{\mathbf{v}} f)(\mathbf{x})$$

la dérivée directionnelle mixte.

Lemme 4.3. *Soit $f \in C^2(E)$, $E \subset \mathbb{R}^n$ ouvert non vide et $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\|\mathbf{v}\| = \|\mathbf{w}\| = 1$. Alors*

$$D^2_{\mathbf{wv}} f(\mathbf{x}) = \mathbf{w}^\top H_f(\mathbf{x})\mathbf{v},$$

où il y a deux produits matriciels, et $\mathbf{v}$ et $\mathbf{w}$ sont des vecteurs colonnes. Ceci s'écrit aussi

$$D^2_{\mathbf{wv}} f(\mathbf{x}) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) v_j w_i.$$

Démonstration. Puisque $f \in C^2(E)$, toutes les dérivées partielles $\frac{\partial f}{\partial x_j}$, $j = 1, \ldots, n$ sont continues sur E et f est différentiable sur E. Donc $D_{\mathbf{v}} f(\mathbf{x}) = Df(\mathbf{x}) \cdot \mathbf{v} = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}) v_j$ et $D_{\mathbf{v}} f : E \to \mathbb{R}$ est continue. De plus, ses dérivées partielles $\frac{\partial (D_{\mathbf{v}} f)}{\partial x_i}$, $i = 1, \ldots, n$ existent et sont continues sur E car

$$\frac{\partial (D_{\mathbf{v}} f)}{\partial x_i} = \frac{\partial \left(\sum_{j=1}^n \frac{\partial f}{\partial x_j} v_j \right)}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} v_j$$

et $\frac{\partial^2 f}{\partial x_i \partial x_j} \in C^0(E)$ pour tout $i, j = 1, \ldots, n$ puisque $f \in C^2(E)$. Donc $D_{\mathbf{v}} f \in C^1(E)$ et

$$D_{\mathbf{w}}(D_{\mathbf{v}} f)(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial (D_{\mathbf{v}} f)}{\partial x_i}(\mathbf{x}) w_i = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) v_j w_i = \mathbf{w}^\top H_f(\mathbf{x})\mathbf{v}.$$

□

Exemple 4.4. *Soit $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y) = x^2 y$, et considérons les deux directions $\mathbf{v} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ et $\mathbf{w} = (\frac{1}{2}, -\frac{\sqrt{3}}{2})$. On a*

$$\frac{\partial f}{\partial x}(x, y) = 2xy, \qquad \frac{\partial f}{\partial y}(x, y) = x^2,$$
$$\frac{\partial^2 f}{\partial x^2}(x, y) = 2y, \quad \frac{\partial^2 f}{\partial y \partial x}(x, y) = 2x, \qquad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 2x, \qquad \frac{\partial^2 f}{\partial y^2}(x, y) = 0.$$

Donc

$$Df(x, y) = (2xy, x^2), \qquad H_f(x, y) = \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}$$

et

$$D^2_{\mathbf{wv}} f(x, y) = \mathbf{w}^\top H_f(x, y)\mathbf{v} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}(x(1 - \sqrt{3}) + y).$$

On remarque de plus que $D^2_{\mathbf{wv}} f(x, y) = D^2_{\mathbf{vw}} f(x, y)$ puisque H_f est symétrique.

Dans l'exemple précédent on a $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. Sous des conditions assez générales, on aura toujours $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$, $i, j = 1, \ldots, n$. Ceci est garanti par l'important résultat suivant :

Théorème 4.5 (de Schwarz)**.** *Soit deux indices fixés $i, j \in \{1, \dots, n\}$, $E \subset \mathbb{R}^n$ ouvert non vide avec $n \geq 2$, $\mathbf{x} \in E$ fixé et $f : E \to \mathbb{R}$ telle que $\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j}, \frac{\partial^2 f}{\partial x_i \partial x_j}, \frac{\partial^2 f}{\partial x_j \partial x_i}$ existent dans E et $\frac{\partial^2 f}{\partial x_i \partial x_j}, \frac{\partial^2 f}{\partial x_j \partial x_i}$ sont continues en cet $\mathbf{x} \in E$. Alors $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x})$.*

Démonstration. Soient $s, t > 0$ fixés suffisamment petits de sorte que le rectangle rempli fermé de sommets $\mathbf{x}, \mathbf{x} + s\mathbf{e}_i, \mathbf{x} + t\mathbf{e}_j, \mathbf{x} + s\mathbf{e}_i + t\mathbf{e}_j$ est contenu dans E. Ceci est toujours possible car $\mathbf{x}$ est un point intérieur de E. On considère la quantité

$$\Delta(s,t) = f(\mathbf{x} + s\mathbf{e}_i + t\mathbf{e}_j) - f(\mathbf{x} + s\mathbf{e}_i) - f(\mathbf{x} + t\mathbf{e}_j) + f(\mathbf{x})$$

et les deux fonctions auxiliaires

$$g(\xi) = f(\mathbf{x}+\xi\mathbf{e}_i+t\mathbf{e}_j)-f(\mathbf{x}+\xi\mathbf{e}_i),\ \xi \in [0,s], \quad h(\xi) = f(\mathbf{x}+s\mathbf{e}_i+\xi\mathbf{e}_j)-f(\mathbf{x}+\xi\mathbf{e}_j),\ \xi \in [0,t].$$

La fonction g est dérivable dans $]0, s[$ par hypothèse ($\frac{\partial f}{\partial x_i}$ existe dans E) et donc, par le théorème des accroissements finis, il existe $\tilde{s} \in \,]0, s[$ tel que

$$g(s) - g(0) = g'(\tilde{s})s = \left(\frac{\partial f}{\partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + t\mathbf{e}_j) - \frac{\partial f}{\partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i) \right) s.$$

Soit maintenant

$$\varphi(\eta) = \frac{\partial f}{\partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \eta\mathbf{e}_j), \quad \eta \in [0,t].$$

La fonction φ est dérivable dans $]0, t[$ car $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existe dans E, et donc il existe $\tilde{t} \in \,]0, t[$ tel que

$$\varphi(t) - \varphi(0) = \varphi'(\tilde{t})t = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \tilde{t}\mathbf{e}_j)t.$$

On a finalement

$$\Delta(s,t) = g(s) - g(0) = g'(\tilde{s})s = \big(\varphi(t) - \varphi(0)\big)s = \varphi'(\tilde{t})st = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \tilde{t}\mathbf{e}_j)st.$$

On répète maintenant les mêmes arguments pour la fonction h :

$$\exists \hat{t} \in \,]0,t[\ : \quad \Delta(s,t) = h(t) - h(0) = h'(\hat{t})t = \left(\frac{\partial f}{\partial x_j}(\mathbf{x} + s\mathbf{e}_i + \hat{t}\mathbf{e}_j) - \frac{\partial f}{\partial x_j}(\mathbf{x} + \hat{t}\mathbf{e}_j) \right) t$$

Soit $\psi(\eta) = \frac{\partial f}{\partial x_j}(\mathbf{x} + \eta\mathbf{e}_i + \hat{t}\mathbf{e}_j),\ \eta \in [0,s]$, alors

$$\exists \hat{s} \in \,]0,s[\ : \quad \psi(s) - \psi(0) = \frac{\partial f}{\partial x_j}(\mathbf{x} + s\mathbf{e}_i + \hat{t}\mathbf{e}_j) - \frac{\partial f}{\partial x_j}(\mathbf{x} + \hat{t}\mathbf{e}_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} + \hat{s}\mathbf{e}_i + \hat{t}\mathbf{e}_j)s$$

et donc

$$\Delta(s,t) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} + \hat{s}\mathbf{e}_i + \hat{t}\mathbf{e}_j)st = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \tilde{t}\mathbf{e}_j)st.$$

Si on prend maintenant $s = t \to 0^+$, on obtient $\hat{s}, \tilde{s} \to 0^+$ et $\hat{t}, \tilde{t} \to 0^+$. Puisque $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ sont continues en $\mathbf{x}$, on a

$$\lim_{s \to 0^+} \frac{1}{s^2} \Delta(s, s) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}).$$

□

Corollaire 4.6. *Soit $E \subset \mathbb{R}^n$ ouvert non vide avec $n \geq 2$ et $f \in C^2(E)$. Alors*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}), \quad \forall \mathbf{x} \in E, \quad \forall i, j = 1, \ldots, n.$$

Le résultat du théorème de Schwarz n'est pas forcément vrai sans l'hypothèse de continuité des dérivées partielles secondes, comme l'exemple suivant le montre.

Exercice 4.7. *Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par*

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2+y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

On vérifie que $\frac{\partial^2 f}{\partial x \partial y}(0, 0)$, $\frac{\partial^2 f}{\partial y \partial x}(0, 0)$ existent mais elles ne sont pas égales.

Grâce au théorème de Schwarz on a que si $f : E \to \mathbb{R}$ admet toutes les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j} : E \to \mathbb{R}$ $i, j = 1, \ldots, n$ et qu'elles sont continues en $\mathbf{x} \in E$, alors $H_f(\mathbf{x})$ est une matrice symétrique. Donc, en fait, il suffit de calculer seulement $\frac{n(n+1)}{2}$ dérivées partielles secondes au lieu de n^2. Si de plus f est de classe $C^2(E)$, il s'ensuit aussi que les dérivées directionnelles secondes $D_{\mathbf{vw}} f(\mathbf{x})$ existent pour tout $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ de norme 1 et que $D_{\mathbf{vw}} f(\mathbf{x}) = D_{\mathbf{wv}} f(\mathbf{x})$.

4.2 Dérivées partielles d'ordres supérieurs à 2

Soit $f : E \to \mathbb{R}$, avec $E \subset \mathbb{R}^n$ ouvert non vide, et $(i_1, \ldots, i_p) \in \{1, \ldots, n\}^p$. On généralise facilement la notion de dérivée partielle d'ordre p de f par rapport aux variables $x_{i_1}, \ldots, x_{i_p}$:

$$\frac{\partial^p f}{\partial x_{i_p} \partial x_{i_{p-1}} \ldots \partial x_{i_1}}(\mathbf{x}) = \frac{\partial \left(\frac{\partial \left(\cdots \left(\frac{\partial f}{\partial x_{i_1}} \right) \cdots \right)}{\partial x_{i_{p-1}}} \right)}{\partial x_{i_p}}(\mathbf{x}).$$

Définition 4.8 (espace $C^p(E)$)**.** *Soit $E \subset \mathbb{R}^n$ ouvert non vide et $f : E \to \mathbb{R}$. On dit que f est de classe C^p, $f \in C^p(E)$, si toutes les dérivées partielles d'ordre p, $\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}} : E \to \mathbb{R}$, existent pour tout $(i_1, \ldots, i_p) \in \{1, \ldots, n\}^p$ et sont continues sur E.*

Grâce au théorème de Schwarz, si $f \in C^p(E)$, alors

$$\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}} = \frac{\partial^p f}{\partial x_{i_{\sigma(1)}} \cdots \partial x_{i_{\sigma(p)}}}$$

où $\sigma(1), \ldots, \sigma(p)$ est une permutation arbitraire des indices $1, \ldots, p$. Autrement dit, l'ordre de dérivation n'est pas importante si $f \in C^p(E)$.

On remarque que si $f \in C^p(E)$, alors $f \in C^q(E)$, $\forall 0 \leq q \leq p$, c.-à-d., $C^p(E) \subset C^{p-1}(E) \subset \cdots \subset C^1(E) \subset C^0(E)$.

Notation par multi-entiers. Considérons une dérivée partielle d'ordre p, $\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}$. Soit $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ où l'entier $\alpha_j \geq 0$ est le nombre de fois que l'indice $j \in \{1, \ldots, n\}$ apparaît dans la suite $i_1, \ldots, i_p$. On note $|\boldsymbol{\alpha}| = \sum_{j=1}^n \alpha_j$ (dans ce cas $|\boldsymbol{\alpha}| = p$). On utilise souvent la notation

$$\frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}} = \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} = \frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}$$

où $\partial x_j^{\alpha_j}$ signifie qu'on dérive α_j fois par rapport à x_j. Cette notation ne distingue pas l'ordre de dérivation.

Exemple 4.9. $f(\mathbf{x}) = x_1 x_2^2 x_3^3$. *Calculons* $\frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}$ *pour* $\boldsymbol{\alpha} = (1, 2, 0)$ *et* $\boldsymbol{\alpha} = (1, 1, 1)$.

$$\boldsymbol{\alpha} = (1,2,0) \Rightarrow \frac{\partial^3 f}{\partial x_1 \partial x_2^2} = \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_2} = \frac{\partial^3 f}{\partial x_2 \partial x_1 \partial x_2} = \frac{\partial^3 f}{\partial x_2 \partial x_2 \partial x_1} = 2x_3^3$$

$$\begin{aligned}\boldsymbol{\alpha} = (1,1,1) \Rightarrow \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3} &= \frac{\partial^3 f}{\partial x_3 \partial x_1 \partial x_2} = \frac{\partial^3 f}{\partial x_2 \partial x_3 \partial x_1} = \frac{\partial^3 f}{\partial x_1 \partial x_3 \partial x_2} \\ &= \frac{\partial^3 f}{\partial x_2 \partial x_1 \partial x_3} = \frac{\partial^3 f}{\partial x_3 \partial x_2 \partial x_1} = 6x_2 x_3^2\end{aligned}$$

Remarquons que le nombre de dérivées partielles d'ordre $|\boldsymbol{\alpha}|$ qui correspondent à un multi-entier $\boldsymbol{\alpha} \in \mathbb{N}^n$, est

$$\binom{|\boldsymbol{\alpha}|}{\boldsymbol{\alpha}} = \frac{|\boldsymbol{\alpha}|!}{\alpha_1! \cdots \alpha_n!}.$$

4.3 Formule de Taylor en dimension $n > 1$

Rappelons d'abord la formule de Taylor pour une fonction d'une seule variable réelle définie sur un intervalle ouvert $I =]a, b[\subset \mathbb{R}$ et p fois continûment dérivable sur I, $f \in C^p(I)$: $\forall x, y \in I$

$$\begin{aligned}f(y) &= f(x) + f'(x)(y-x) + \frac{f''(x)}{2}(y-x)^2 + \cdots + \frac{f^{(p)}(x)}{p!}(y-x)^p + R_p(y) \\ &= \underbrace{\sum_{k=0}^{p} \frac{f^{(k)}(x)}{k!}(y-x)^k}_{=T_x^p(y)} + R_p(y)\end{aligned}$$

où $T_x^p(y) = \sum_{k=0}^p \frac{f^{(k)}(x)}{k!}(y-x)^k$ est le polynôme de Taylor de degré p et $R_p(y)$ est le reste de la formule de Taylor t.q. $\lim_{y\to x} \frac{R_p(y)}{|y-x|^p} = 0$ (ou avec notation “petit o”, $R_p(y) = o(|y-x|^p)$). Si, de plus, la dérivée d'ordre $p+1$ existe sur I et est continue, alors le reste de la formule peut être caractérisé par

$$\text{Reste de Lagrange :} \qquad R_p(y) = \frac{f^{(p+1)}(x+t(y-x))}{(p+1)!}(y-x)^{p+1} \quad \text{pour un } t \in]0,1[,$$

$$\text{Reste (sous forme d') intégrale :} \qquad R_p(y) = \int_0^1 (1-t)^p \frac{f^{(p+1)}(x+t(y-x))}{p!}(y-x)^{p+1}dt.$$

Considérons maintenant le **cas $\boldsymbol{n \geq 2}$**. Soit $E \subset \mathbb{R}^n$ ouvert non vide et $\mathbf{x}, \mathbf{y} \in E$ tels que le segment $[\mathbf{x}, \mathbf{y}] = \{\mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}),\ t \in [0,1]\} \subset E$. Pour un multi-entier $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, on utilise les notations suivantes :

$$\boldsymbol{\alpha}! = \prod_{i=1}^n \alpha_i!, \qquad \mathbf{x}^{\boldsymbol{\alpha}} = \prod_{i=1}^n x_i^{\alpha_i}.$$

Soit $f \in C^{p+1}(E)$. On note $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$, qui est bien définie et de classe C^{p+1} sur un intervalle ouvert $I =]-\delta, 1+\delta[$ qui contient $[0,1]$, grâce au fait que $\mathbf{x}, \mathbf{y} \in E$ sont des points intérieurs. L'idée pour dériver une formule de Taylor pour f est d'utiliser la formule de Taylor pour $g : I \to \mathbb{R}$:

$$g(t) = g(0) + g'(0)t + \frac{g''(0)}{2}t^2 + \cdots + \frac{g^p(0)}{p!}t^p + R_p^g(t), \quad t \in]0,1]$$

où

$$R_p^g(t) = \int_0^t \frac{(t-\tau)^p}{p!} g^{(p+1)}(\tau)d\tau = \frac{g^{(p+1)}(\theta)}{(p+1)!}t^{p+1} \quad \text{pour un } \theta \in]0,t[.$$

Or, si on note $\mathbf{x}_t = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$:

$$\begin{aligned}
g(t) &= f(\mathbf{x}_t) \\
g'(t) &= \sum_{i_1=1}^{n} \frac{\partial f}{\partial x_{i_1}}(\mathbf{x} + t(\mathbf{y}-\mathbf{x}))(y_{i_1} - x_{i_1}) = \sum_{|\boldsymbol{\alpha}|=1} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x}_t)(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}} \\
g''(t) &= \sum_{i_2=1}^{n} \frac{\partial}{\partial x_{i_2}} \left(\sum_{i_1=1}^{n} \frac{\partial f}{\partial x_{i_1}}(\mathbf{x} + t(\mathbf{y}-\mathbf{x}))(y_{i_1} - x_{i_1}) \right) (y_{i_2} - x_{i_2}) \\
&= \sum_{i_1, i_2 = 1}^{n} \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(\mathbf{x}_t)(y_{i_1} - x_{i_1})(y_{i_2} - x_{i_2}) \\
&= \sum_{|\boldsymbol{\alpha}|=2} \frac{2!}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x}_t)(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}} \\
&\vdots \\
g^{(p)}(t) &= \sum_{i_1, \dots, i_p = 1}^{n} \frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}(\mathbf{x} + t(\mathbf{y}-\mathbf{x}))(y_{i_1} - x_{i_1}) \cdots (y_{i_p} - x_{i_p}) \\
&= \sum_{|\boldsymbol{\alpha}|=p} \frac{p!}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x}_t)(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}}.
\end{aligned}$$

Finalement, on peut écrire la formule de Taylor

$$\begin{aligned}
f(\mathbf{y}) = g(1) &= \sum_{k=0}^{p} \frac{1}{k!} g^{(k)}(0) + R_p^g(1) \\
&= \sum_{k=0}^{p} \frac{1}{k!} \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x})(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}} + R_p(\mathbf{y}) \\
&= \underbrace{\sum_{|\boldsymbol{\alpha}| \le p} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x})(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}}}_{=T_{\mathbf{x}}^p(\mathbf{y})} + R_p(\mathbf{y}),
\end{aligned}$$

où on identifie :

$$T_{\mathbf{x}}^p(\mathbf{y}) = \sum_{|\boldsymbol{\alpha}| \le p} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x})(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}} \qquad \textbf{(polynôme de Taylor de degré } p\textbf{)}$$

et

$$\begin{aligned}
R_p(\mathbf{y}) &= \sum_{|\boldsymbol{\alpha}|=p+1} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{(p+1)} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x} + \theta(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}} \quad \text{pour un } \theta \in]0,1[\quad \textbf{(reste de Lagrange)}, \\
&= \sum_{|\boldsymbol{\alpha}|=p+1} \frac{(p+1)}{\boldsymbol{\alpha}!} \int_0^1 (1-t)^p \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x} + t(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x})^{\boldsymbol{\alpha}} dt \qquad \textbf{(reste intégrale)}.
\end{aligned}$$

4.3.1 Formule de Taylor pour des fonctions vectorielles

Soit $E \subset \mathbb{R}^n$ ouvert non vide, $\mathbf{f} \in C^{p+1}(E, \mathbb{R}^m)$ et $\mathbf{x}, \mathbf{y} \in E$ t.q. $[\mathbf{x}, \mathbf{y}] \subset E$. On peut appliquer la formule de Taylor à chaque composante f_k de $\mathbf{f} = (f_1, \ldots, f_m)$: il existe $\theta_1, \theta_2, \ldots, \theta_m \in]0, 1[$ t.q.

$$f_k(\mathbf{y}) = \sum_{|\boldsymbol{\alpha}| \leq p} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f_k}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x})(\mathbf{y} - \mathbf{x})^{\boldsymbol{\alpha}} + \sum_{|\boldsymbol{\alpha}| = p+1} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{(p+1)} f_k}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x} + \theta_k(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^{\boldsymbol{\alpha}}$$

Toutefois, en général les $\theta_1, \ldots, \theta_m$ ne sont pas égaux entre eux et on ne peut pas trouver un seul $\theta \in]0, 1[$ tel que $\mathbf{R}_p(\mathbf{y}) = \sum_{|\boldsymbol{\alpha}|=p+1} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{(p+1)} \mathbf{f}}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^{\boldsymbol{\alpha}}$. D'un autre côté, la formule de Taylor avec reste intégrale est valable aussi pour des fonctions vectorielles :

$$\mathbf{f}(\mathbf{y}) = \sum_{|\boldsymbol{\alpha}| \leq p} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} \mathbf{f}}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x})(\mathbf{y} - \mathbf{x})^{\boldsymbol{\alpha}} + \sum_{|\boldsymbol{\alpha}| = p+1} \frac{p+1}{\boldsymbol{\alpha}!} \int_0^1 (1-t)^p \frac{\partial^{(p+1)} \mathbf{f}}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^{\boldsymbol{\alpha}} dt.$$