Intégrales impropres, introduction au calcul différentiel

Cours de L2, Université de Bourgogne

Thomas Dreyfus

20 mars 2024

Résumé

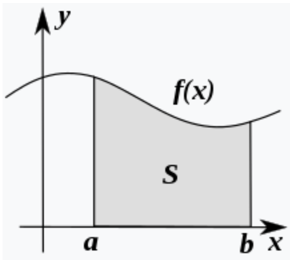
Ce cours comporte essentiellement trois parties. La première sur l'intégration comprend des rappels sur l'intégrale de Riemann ainsi que l'étude des intégrales généralisés. Ensuite nous traiterons la continuité et la dérivabilité des suite et séries de fonctions. Nous terminerons ce cours avec l'étude des fonctions à plusieurs variables et les bases de topologies nécessaires pour l'étude de ces dernières.

Table des matières

1 Intégrales de Riemann

1.1 Définitions, premières propriétés

Soient $a < b \in \mathbb{R}$, et f une fonction de $[a, b]$ dans $\mathbb{R}$. Si f est positive, l'intégrale $\int_a^b f(x)dx$ représente l'air entre la droite $y = 0$ et la courbe $y = f(x)$ pour x entre a et b. Voir par exemple la figure ci dessous.



Si f est négative, l'intégrale sera négative et représentera l'opposé de l'air entre la droite $y = 0$ et la courbe $y = f(x)$ pour x entre a et b. Si f est positive sur un intervalle I et négative sur un intervalle J, alors l'intégrale représente l'air entre $y = f(x)$ et $y = 0$ pour $x \in I$ moins l'air entre les courbes $y = f(x)$ et $y = 0$ pour $x \in J$. L'idée de ce qui suit est de se donner un cadre formel à ceci où l'intégrale est bien définie.

Commençons par définir l'intégrale pour les fonctions en escalier.

Définition 1.1. Soient $a < b$. Une fonction $f : [a, b] \to \mathbb{R}$ est dite en escalier si il existe des réels $a_0, \dots, a_N$ avec $a = a_0 < a_1 < \cdots < a_N = b$ tels que f soit constante sur tous les intervalles ouverts de la forme de $]a_i, a_{i+1}[$. Autrement dit, il existe des réels $c_0, \dots, c_{N-1}$ tels que pour tout $0 \leq i < N$, pour tout $a_i < x < a_{i+1}$, $f(x) = c_i$.

Notons que f peut prendre n'importe quelle valeurs en les a_i. En particulier une fonction en escalier n'est pas nécessairement continues.

Exemple 1.2. Les fonctions constantes sont des fonctions en escaliers.

Définition 1.3. Soit f une fonction en escalier comme définit ci dessus. On définit $\int_a^b f(x)dx := \sum_{i=0}^{N-1} c_i(a_{i+1} - a_i)$, et $\int_b^a f(x)dx = -\int_a^b f(x)dx$.

Ainsi lorsque la fonction en escalier est positive, l'intégrale représente bien l'air sous la courbe, puisque le domaine entre $y = 0$ et $y = f(x)$ est une union de rectangles de côté c_i et $(a_{i+1} - a_i)$. Notons qu'on autorise les c_i à être éventuellement négatifs et que l'intégrale vaut bien l'air entre $y = f(x)$ et $y = 0$ pour x tel que f

est positive, moins l'air entre les courbes $y = f(x)$ et $y = 0$ pour les x tels que f est négative.

Le lemme suivant servira beaucoup dans ce qui suit. Il dit essentiellement que si on a deux fonction en escaliers sur un intervalle, on peut trouver une subdivision commune.

Lemme 1.4. *Soient f, g en escalier sur $[a, b]$. Alors il existe des réels $a_0, \ldots, a_N$ avec $a = a_0 < a_1 < \cdots < a_N = b$ tels que f et g soient constantes sur tous les intervalles ouverts de la forme de $]a_i, a_{i+1}[$.*

Démonstration. Soit $I = \{\alpha_0, \ldots, \alpha_n\}$ tel que f soit constante sur tous les intervalles ouverts de la forme de $]\alpha_i, \alpha_{i+1}[$. Soit $J = \{\beta_0, \ldots, \beta_m\}$ tel que g soit constante sur tous les intervalles ouverts de la forme de $]\beta_i, \beta_{i+1}[$. On considère $I \cup J = \{a_0, \ldots, a_N\}$ avec $a = a_0 < a_1 < \cdots < a_N = b$. Par construction, f et g sont constantes sur les intervalles de la forme $]a_i, a_{i+1}[$. □

On va maintenant voir que l'intégrale est invariante par changement d'un nombre fini de valeurs.

Proposition 1.5. *Soient f, g en escalier en $[a, b]$. Supposons qu'il existe un ensemble fini $I \subset [a, b]$ tel que $x \in [a, b] \backslash I$ entraîne $f(x) = g(x)$. Alors $\int_a^b f(x)dx = \int_a^b g(x)dx$.*

Démonstration. D'après le lemme 1.4, il existe des réels $a_0, \ldots, a_N$ avec $a = a_0 < a_1 < \cdots < a_N = b$ tels que f et g soient constantes sur tous les intervalles ouverts de la forme de $]a_i, a_{i+1}[$ avec valeurs respectives c_i et d_i. Puisque $x \in [a, b] \backslash I$ entraîne $f(x) = g(x)$, on a pour tout i, $c_i = d_i$. Par définition, on a

$$\int_a^b f(x)dx = \sum_{i=0}^{N-1} c_i(a_{i+1} - a_i) = \sum_{i=0}^{N-1} d_i(a_{i+1} - a_i) = \int_a^b g(x)dx.$$

□

Proposition 1.6. *Soient f, g en escalier, et $c, d \in \mathbb{R}$. Alors $cf + dg$ est une fonction en escalier et*

$$\int_a^b cf(x) + dg(x)dx = c\int_a^b f(x)dx + d\int_a^b g(x)dx.$$

Démonstration. D'après le lemme 1.4, il existe des réels $a_0, \ldots, a_N$ avec $a = a_0 < a_1 < \cdots < a_N = b$ tels que f et g soient constantes sur tous les intervalles ouverts de la forme de $]a_i, a_{i+1}[$ avec valeurs respectives c_i et d_i. Donc $cf + dg$ est constantes sur de tels intervalles avec valeur $cc_i + dd_i$. Donc $cf + dg$ est une fonction en escalier et nous avons

$$\int_a^b cf(x) + dg(x)dx = \sum_{i=0}^{N-1} (cc_i + dd_i)(a_{i+1} - a_i).$$

Cette somme vaut

$$c \sum_{i=0}^{N-1} c_i(a_{i+1} - a_i) + d \sum_{i=0}^{N-1} d_i(a_{i+1} - a_i) = c \int_a^b f(x)dx + d \int_a^b g(x)dx.$$

□

Remarque 1.7. On a prouvé que les fonction en escaliers forment un $\mathbb{R}$ espace vectoriel et que l'application $f \mapsto \int_a^b f(x)dx$ est une application linéaire.

Proposition 1.8. *Soient f, g en escalier. On suppose que $f(x) \leq g(x)$ sauf pour un ensemble fini. Alors $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.*

Démonstration. Quitte à changer un nombre fini de valeurs et à utiliser la proposition 1.11, on peut sans pertes de généralités supposer que $f(x) \leq g(x)$ pour tout x. D'après le lemme 1.4, il existe des réels $a_0, \ldots, a_N$ avec $a = a_0 < a_1 < \cdots < a_N = b$ tels que f et g soient constante sur tous les intervalles ouverts de la forme de $]a_i, a_{i+1}[$ avec valeurs respectives c_i et d_i et $c_i \leq d_i$ pour tout i. Alors

$$\int_a^b f(x)dx = \sum_{i=0}^{N-1} c_i(a_{i+1} - a_i) \leq \sum_{i=0}^{N-1} d_i(a_{i+1} - a_i) = \int_a^b g(x)dx.$$

□

On souhaite intégrer un peu plus de fonctions que les seules fonctions en escaliers. Donnons la définitions des fonctions que l'on souhaite intégrer dans ce cours, l'espace des fonctions intégrables au sens de Riemann. Mais avant introduisons les intégrales supérieures et inférieures

Définition 1.9. Soit $f : [a, b] \to \mathbb{R}$ une fonction bornée. On définit

— $I^-(f) = \sup \left\{ \int_a^b g(x)dx, \quad g \quad \text{en escalier} \quad g \leq f \right\}$.

— $I^+(f) = \inf \left\{ \int_a^b g(x)dx, \quad g \quad \text{en escalier} \quad g \geq f \right\}$.

Rappelons que tout ensemble non vide majoré admet un sup et tout ensemble non vide minoré admet un inf. L'hypothèse que l'ensemble est majoré/minoré est importante, par exemple $\mathbb{R}$ ne possède ni inf ni sup. Remarquons que puisque f est borné il existe m, M tels que pour tout $x \in [a, b]$, $m \leq f(x) \leq M$. Ainsi il existe une fonction en escalier telle que $g \geq f$, par exemple la fonction constante égale à M. De plus une telle fonction en escalier sera plus grande que m dont l'intégrale vaut $m(b-a)$. Donc des telles fonctions en escaliers ont une intégrale dans l'intervalle $[m(b-a), +\infty[$. C'est un ensemble minoré (par $m(b-a)$), donc $I^+(f)$ est bien définie. De même, $I^-(f)$ est aussi bien définie. On a

$$m(b-a) = I^-(f) \leq I^+(f) \leq M(b-a).$$

Définition 1.10. Soit $f : [a, b] \to \mathbb{R}$ une fonction bornée. On dira que f est intégrable au sens de Riemann, si

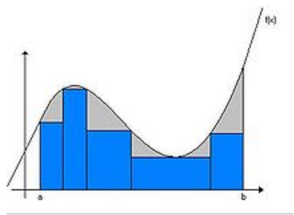
$$I^-(f) = I^+(f).$$

Dans ce cas on pose

$$\int_a^b f(x)dx = I^-(f) = I^+(f).$$

On prendra la convention que $\displaystyle\int_a^b f(x)dx = -\int_b^a f(x)dx$.

Par soucis d'alléger les énoncés, on dira qu'une fonction est intégrable si elle est intégrable au sens de Riemann. Ainsi on approxime l'air sous la courbe par l'air de fonctions en escaliers. Sur la figure ci dessous, l'air en bleu correspond à l'intégrale d'une fonction en escalier plus petite que f.



Voyons maintenant que comme pour les fonctions en escaliers, l'intégrale ne change pas si on modifie un nombre fini de valeurs.

Proposition 1.11. *Soit f, gintégrable au sens de Riemann sur $[a, b]$. Soit g une fonction de $[a, b]$ dans $\mathbb{R}$. Supposons qu'il existe un ensemble fini $I \subset [a, b]$ tel que $x \in [a, b] \backslash I$ entraîne $f(x) = g(x)$. Alors g est intégrable au sens de Riemann et $\displaystyle\int_a^b f(x)dx = \int_a^b g(x)dx$.*

Démonstration. Soit f^- en escalier tel que $f^- \leq f$. Soient $J = \{\alpha_0, \ldots, \alpha_n\}$ tel que f^- soit constantes sur les intervalles de la forme $]\alpha_i, \alpha_{i+1}[$. Soit $I \cup J = \{a = a_0, \ldots, a_N = b\}$. Alors f^- est constantes sur les intervalles de la forme $]a_i, a_{i+1}[$. Soit $\widetilde{f}^-$ fonction en escalier prenant les mêmes valeurs que f^- sur les intervalles de la forme $]a_i, a_{i+1}[$, et valant $\min(f(a_i), g(a_i))$ en les a_i. D'après la proposition 1.11, $\int_a^b f^-(x)dx = \int_a^b \widetilde{f}^-(x)dx$. La fonction en escalier $\widetilde{f}^-$ est plus petite que f et g. On a donc prouvé que pour toute fonction en escalier $f^- \leq f$ il existe une fonction en escalier $\widetilde{f}^-$ plus petite que f et g avec $\int_a^b f^-(x)dx = \int_a^b \widetilde{f}^-(x)dx$. En prenant le sup on trouve que $I^-(f) \leq I^-(g)$.

De même, pour toute fonction en escalier $g^- \leq g$ il existe une fonction en escalier $\widetilde{g}^-$ plus petite que f et g avec $\int_a^b g^-(x)dx = \int_a^b \widetilde{g}^-(x)dx$. En prenant le sup on trouve donc que $I^-(g) \leq I^-(f)$ et donc $I^-(f) = I^-(g)$. De même, $I^+(f) = I^+(g)$. Puisque f est intégrable, on a $I^-(f) = I^+(f)$ et donc $I^-(g) = I^+(g) = \int_a^b f(x)dx$. □

Remarque 1.12. Si $f \geq 0$ est intégrable au sens de Riemann, alors $\int_a^b f(x)dx \geq 0$. Ceci reste vrai si $f(x) \geq 0$ sauf en un nombre fini de points.

Remarque 1.13. Si $f(x)$ est borné avec $m \leq f(x) \leq M$, on a prouvé que

$$\left| \int_a^b f(x)dx \right| \leq \max\{|m|, |M|\}(b-a).$$

Lemme 1.14. *Soit $f \geq 0$ sauf en un nombre fini de point que l'on suppose intégrable au sens de Riemann. Alors $\int_a^b f(x)dx = 0$ implique que f est nul sauf en un nombre finit de point.*

Démonstration. Prouvons le résultat par la contraposée. On suppose que f n'est pas nulle en une infinité de points. Un nombre finit de valeurs n'affectant pas l'intégrale, on peut supposer sans pertes de généralités, que $f \geq 0$ sur $[a, b]$. Une fonction en escalier telle que $g \geq f$ prend une infinité de valeurs non nulles, il existe donc $\varepsilon > 0$ et un intervalle $I \subset [a, b]$, où $g(x) > \varepsilon$ pour tout $x \in I$. Puisque $f \geq 0$, g est positive. On a $\int_a^b g(x)dx \geq \varepsilon(d-c) > 0$. Ainsi $I^+(f) \geq \varepsilon(d-c)$ et donc $\int_a^b f(x)dx > 0$. □

Présentons maintenant quelques propriétés.

Proposition 1.15. *Soient f, g intégrable au sens de Riemann, et $c, d \in \mathbb{R}$. Alors $cf + dg$ est intégrable au sens de Riemann et*

$$\int_a^b cf(x) + dg(x)dx = c\int_a^b f(x)dx + d\int_a^b g(x)dx.$$

Démonstration. Soient f^- et g^- des fonctions en escaliers tels que $f^- \leq f$ et $g^- \leq f$. On a que $cf^- + dg^- \leq cf + dg$ est une fonction en escalier. Ainsi $cI^-(f) + dI^-(g) \leq I^-(cf + dg)$. De même $I^+(cf + dg) \leq cI^+(f) + dI^+(g)$. Puisque $I^-(f) = I^+(f)$, $I^-(g) = I^+(g)$ on obtient $I^+(cf + dg) = I^-(cf + dg) = cI^\pm(f) + dI^\pm(g)$ ce qui prouve que $cf + dg$ est intégrable au sens de Riemann et

$$\int_a^b cf(x) + dg(x)dx = c\int_a^b f(x)dx + d\int_a^b g(x)dx.$$

□

Proposition 1.16. *Soient f, g intégrable au sens de Riemann, avec $f \leq g$ sauf en un nombre finit de valeurs. Alors*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

Démonstration. Soit $I := \int_a^b g(x)dx - \int_a^b f(x)dx = \int_a^b g(x) - f(x)dx$. Il suffit de montrer que $I \geq 0$. La fonction $g(x) - f(x)$ est positive sauf en un nombre finit de valeurs, elle est donc minorée par la fonction nulle sauf en un nombre finie de valeurs. On a ainsi $\int_a^b g(x) - f(x)dx \geq 0$, ce qui conclue la preuve. □

Lemme 1.17. *Si f est intégrable au sens de Riemann, alors il en est de même pour sa valeur absolue $|f|$ et $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|\, dx$.*

Démonstration. Le fait que $|f|$ est intégrable au sens de Riemann est admis. On a $f \leq |f|$. Si $\int_a^b f(x)dx \geq 0$ on trouve par la proposition 1.16, $\left|\int_a^b f(x)dx\right| = \int_a^b f(x)dx \leq \int_a^b |f(x)|\, dx$. Si $\int_a^b f(x)dx \leq 0$ on a $\int_a^b -f(x)dx \geq 0$, $-f \leq |f|$ et donc par la proposition 1.16 $\int_a^b -f(x)dx \leq \int_a^b |f(x)|\, dx$. Dans les deux cas le résultat est prouvé. □

Proposition 1.18 (Relation de Chasles)**.** *Soient $a < b < c$ et f intégrable au sens de Riemann sur $[a, b]$ et $[b, c]$. Alors f est intégrable au sens de Riemann sur $[a, c]$ et*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Démonstration. Soit g_1 est escalier tel que $g_1 \leq f$ sur $[a,b]$ sauf en un nombre finit de point. Soit g_2 est escalier tel que $g_2 \leq f$ sur $[b,c]$ sauf en un nombre finit de point. Alors on construit une fonction en escalier g qui vaut g_1 sur $[a,b[$ et g_2 sur $[b,c]$. On a bien $g \leq f$ sauf en un nombre finit de point et $\int_a^b g(x)dx = \int_a^b g_1(x)dx + \int_b^c g_2(x)dx$. Ainsi pour tout $g_1 \leq f$ et $g_2 \leq f$ en escalier, on peut construire $g \leq f$ en escalier avec $\int_a^b g(x)dx = \int_a^b g_1(x)dx + \int_b^c g_2(x)dx$. En prenant le sup sur $\int_a^b g_1(x)dx$ et $\int_b^c g_2(x)dx$ avec $g_1 \leq f$ et $g_2 \leq f$ en escalier, on trouve que le sup des $\int_a^c g(x)dx$ avec $g \leq f$ en escalier est au moins égale à $\int_a^b f(x)dx + \int_b^c f(x)dx$. Autrement dit, on a $I^-(f) \geq \int_a^b f(x)dx + \int_b^c f(x)dx$. On procède de même avec les fonctions en escaliers plus grandes et on trouve $I^+(f) \leq \int_a^b f(x)dx + \int_b^c f(x)dx$. Ceci prouve que $I^-(f) = I^+(f) = \int_a^b f(x)dx + \int_b^c f(x)dx$. □

On a vu que les fonctions en escaliers sont intégrable. Donnons maintenant deux famille de fonctions qui sont intégrables, au sens de Riemann.

Théorème 1.19. *Les fonctions continues sur $[a,b]$ sont intégrables au sens de Riemann.*

Démonstration. Soit f continue. Quelque soit $x \in [a,b]$, $\varepsilon > 0$, il existe $\eta > 0$ tel que pour $x_0 \in [a,b]$ avec $|x - x_0| < \eta$ on a $|f(x) - f(x_0)| < \varepsilon$. Fixons $\varepsilon > 0$. Soit $\eta(x) > 0$ tel que tel que pour $x_0 \in [a,b]$ avec $|x - x_0| < \eta(x)$ on a $|f(x) - f(x_0)| < \varepsilon$. Comme f est continue sur un compacte, elle est uniformément continue d'après le théorème de Heine, ce qui signifie qu'il existe $\eta > 0$ tel que pour $x_0 \in [a,b]$ avec $|x - x_0| < \eta$ on a $|f(x) - f(x_0)| < \varepsilon$. Prenons une subdivision de $[a,b]$ avec $a_{n+1} - a_n \leq \eta$. Prenons la fonction en escaliers qui sur $]a_n, a_{n+1}[$ prend la valeur $\min_{x \in [a_n, a_{n+1}]} f(x)$ (ce min est bien atteint car on regarde une fonction continue sur un compacte). On a donc construit une fonction en escalier f^- avec $f - \varepsilon \leq f^- \leq f$. De même on construit une fonction en escalier f^+ telle que $f \leq f^+ \leq f + \varepsilon$. Ainsi $f^+ - f^- < 2\varepsilon$. On trouve

$$\int_a^b f^-(x) \leq \int_a^b f^+(x)dx < \int_a^b f^-(x)dx + 2\varepsilon(b-a).$$

Puisque ε est arbitraire, on peut poser $\varepsilon = 1/n$. On a ainsi une suite de fonctions en escaliers tels que $f_n^- < f < f_n^+$ et

$$\int_a^b f_n^-(x)dx \leq \int_a^b f_n^+(x)dx < \int_a^b f_n^-(x)dx + 2\frac{(b-a)}{n}.$$

Puisque $\int_a^b f_n^-(x)dx \leq I^-(f)$, $\int_a^b f_n^+(x)dx \geq I^+(f)$, en faisant tendre n vers l'infinie on trouve bien que $I^-(f) = I^+(f)$. Donc f est intégrable. $\square$

On dira que f est continue par morceau si il existe une subdivision $a = a_0 < a_1 < \cdots < a_N = b$ tel que f soit continue sur les $]a_i, a_{i+1}[$. Grâce à la relation de Chasles on trouve que si on peut prolonger chaque restrictions de f sur $]a_i, a_{i+1}[$ en une fonction continue sur $[a_i, a_{i+1}]$, alors la fonction est intégrable.

Remarque 1.20. Toutes les fonctions ne sont pas intégrable au sens de Riemann. Par exemple la fonction qui vaut 1 si $x \in [0, 1] \cap \mathbb{Q}$ et 0 ailleurs n'est pas intégrable au sens de Riemann. Elle l'est au sens de Lebesgue en revanche, mais cette notion sera pour un autre cours.

1.2 Inégalité de Cauchy Schwartz

On a vu que l'intégrale se comportait bien vis à vis de la somme. Les choses sont plus difficiles vis à vis du produit.

Lemme 1.21. *Soient f, g des fonctions intégrables au sens de Riemann. Alors fg est intégrable.*

Démonstration. Fixons $\varepsilon > 0$. Soient f^-, g^- en escalier de même signe que f et g tels que $|f^-(x)| \leq |f(x)|$, $|g^-(x)| \leq |g(x)|$ si $f(x)$ et $g(x)$ sont de même signes, $|f^-(x)| \geq |f(x)|$, $|g^-(x)| \geq |g(x)|$ si ils sont de signes contraires. Supposons de plus que et $-\varepsilon < \int_a^b f(x) - f^-(x)dx < \varepsilon$ et $-\varepsilon < \int_a^b g(x) - g^-(x)dx < \varepsilon$. On a $f^-g^- \leq fg$. De plus $fg - f^-g^- = f(g-g^-) + g^-(f-f^-)$. Les fonctions f et g^- étant bornées (disons par $M > 0$) on trouve que $fg - f^-g^- < M(g-g^-) + M(f-f^-)$. La fonction $M(g-g^-) + M(f-f^-)$ est intégrable d'intégrale majorée par $2M\varepsilon$. En faisant la même chose avec f^+, g^+ on trouve que $\int_a^b f^+(x)g^+(x) - f^-(x)g^-(x)dx \leq 4M\varepsilon$ avec $f^-g^- \leq fg \leq f^+g^+$. Puisque $\varepsilon > 0$ est arbitrairement petit, on obtient $I^+(fg) = I^-(fg)$ prouvant que fg est intégrable. $\square$

Passons maintenant à l'inégalité de Cauchy Schwartz.

Théorème 1.22 (Inégalité de Cauchy Schwartz)**.** *Si f, g sont intégrables au sens de Riemann sur $[a, b]$ alors*

$$\left| \int_a^b f(x)g(x)dx \right| \leq \left(\int_a^b f^2(x)dx \right)^{1/2} \left(\int_a^b g^2(x)dx \right)^{1/2}.$$

Il y a égalité si et seulement si il existe $c \in \mathbb{R}$ tel que $f = gc$ sauf en un nombre finit de points.

Pour prouver ce résultat introduisons l'espace des fonctions f de $[a, b]$ dans $\mathbb{R}$ tel que f^2 soit intégrable. On nomme cet espace $L^2([a, b])$. On définit un produit scalaire sur $L^2([a, b])$ par

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Si $f = g$ on trouve la norme $L^2([a, b])$

$$||f|| = \langle f, f\rangle^{1/2} = \left(\int_a^b f^2(x)dx\right)^{1/2}.$$

Démonstration. On veut montrer que

$$\langle f, g\rangle^2 \leq ||f||^2.||g||^2 = \langle f, f\rangle\langle g, g\rangle.$$

Pour $c \in \mathbb{R}$ on considère

$$\langle cf + g, cf + g\rangle^2 = \int_a^b (cf(x) + g(x))^2 dx.$$

Cette dernière intégrale est positive car $(cf(x) + g(x))^2$ est une fonction positive. Par linéarité de l'intégrale il vient

$$0 \leq c^2 \int_a^b f^2(x)dx + 2c \int_a^b f(x)g(x)dx + \int_a^b g^2(x)dx.$$

On a un polynôme de degrés 2 en c qui est toujours positif ou nul. Son discriminant est donc négatif ou nul. Ainsi

$$4\left(\int_a^b f(x)g(x)dx\right)^2 - 4\left(\int_a^b f^2(x)dx\right)\left(\int_a^b g^2(x)dx\right) \leq 0.$$

On trouve l'inégalité recherchée.

Il y a égalité si et seulement si le discriminant est nul. Il existe alors un unique c_0 tel que le polynôme est nul. Pour cette valeur de c_0 on trouve

$$\langle c_0 f + g, c_0 f + g\rangle^2 = \int_a^b (c_0 f(x) + g(x))^2 dx = 0.$$

La fonction $(c_0 f(x) + g(x))^2$ étant positive elle est nulle sauf en un nombre finit de points d'après le lemme 1.14. On trouve alors que $f = -gc_0$ sauf en un nombre finit de point. On trouve le résultat avec $-c_0$. □

En corollaire on trouve l'inégalité triangulaire de la norme $L^2([a, b])$ (et donc que cette dernière définit bien une norme)

Corollaire 1.23. *Soient $f, g \in L^2([a,b])$. Alors*

$$||f+g|| \leq ||f|| + ||g||.$$

Démonstration. On a $||f+g||^2 = \int_a^b (f(x)+g(x))^2 dx$. Par linéarité cette intégrale vaut $\int_a^b f(x)^2 dx + \int_a^b 2f(x)g(x)dx + \int_a^b g(x)^2 dx$. On reconnaît $||f||^2 + 2\langle f, g\rangle + ||g||^2$. Par l"inégalité de Cauchy-Schwartz, $\langle f, g\rangle \leq ||f||.||g||$ et donc

$$||f+g||^2 \leq (||f|| + ||g||)^2.$$

□

1.3 Calcule d'intégrales en pratique

Soit $f : [a,b] \to \mathbb{R}$ continue (et donc intégrable). On définit $F : [a,b] \to \mathbb{R}$ par

$$F(x) = \int_a^x f(t)dt.$$

Théorème 1.24. *La fonction F est $\mathcal{C}^1$ sur $[a,b]$ est une primitive de f, c'est à dire que*

$$F'(x) = f(x).$$

Réciproquement, si G est $\mathcal{C}^1$ sur $[a,b]$, avec $G'(x) = f(x)$, alors $\int_a^x f(x)dx = G(x) - G(a)$.

Donc pour calculer une intégrale d'une fonction, il suffit de calculer une fonction dont la dérivée est f et de l'évaluer. Bien entendu, calculer une primitive n'a rien d'évident en général.

Exemple 1.25. si $f(x) = x^2$ sur $[0,b]$, alors on peut prendre $F(x) = \frac{x^3}{3}$ et on a $\int_a^x t^2 dt = \frac{x^3}{3}$.

Démonstration. Soit $x \in]a,b[$. Calculons le taux d'accroissement $\frac{F(x+h)-F(x)}{h}$ et pour simplifier supposons $h > 0$. On a

$$\frac{F(x+h)-F(x)}{h} = \int_x^{x+h} \frac{f(t)}{h} dt.$$

La fonction f étant continue, pour tout ε il existe $h_0 > 0$ tel que pour $0 < h < h_0$ on a

$$f(x) - \varepsilon \leq \min_{t \in [x,x+h]} f(t) \leq f(x) \leq \max_{t \in [x,x+h]} f(t) \leq f(x) + \varepsilon.$$

Donc pour tout $t \in [x, x+h]$ on a $f(x) - \varepsilon \leq f(t) \leq f(x) + \varepsilon$. On a donc

$$\int_x^{x+h} \frac{f(x) - \varepsilon}{h} dt - \leq \int_x^{x+h} \frac{f(t)}{h} dt \leq \int_x^{x+h} \frac{f(x) + \varepsilon}{h} dt.$$

Ainsi

$$f(x) - \varepsilon \leq \int_x^{x+h} \frac{f(t)}{h} dt \leq f(x) + \varepsilon.$$

En faisant tendre ε vers 0 on trouve que F est dérivable à droite en x et $F'(x) = f(x)$. On procède de manière similaire avec la dérivée à gauche en prenant $h < 0$.

Si $G'(x) = f(x)$ alors $(G - F)' = 0$ et donc il existe $c \in \mathbb{R}$ tel que $F(x) + c = G(x)$. Par définition, $F(a) = \int_a^a f(t)dt = 0$ et donc $c = G(a)$. On trouve donc $\int_a^x f(x)dx = F(x) = G(x) - G(a)$. □

Terminons par des méthodes de calcules d'intégrales. La première est pratique lorsqu'il s'agit de calculer une intégrale d'un produit.

Proposition 1.26 (Intégration par parties)**.** *Si u, v sont dérivables sur $[a, b]$, tels que u', v' soient intégrables au sens de Riemann, alors uv', $u'v$ sont intégrables au sens de Riemann et*

$$\int_a^b u(t)v'(t)dt = [uv]_a^b - \int_a^b u'(t)v(t)dt,$$

avec la notation $[uv]_a^b = u(b)v(a) - u(b)v(a)$.

Démonstration. La dérivée du produit donne

$$(uv)' = u'v + uv'.$$

Ainsi,

$$[uv]_a^b = \int_a^b (uv)'(t)dt = \int_a^b u(t)v'(t)dt + \int_a^b u'(t)v(t)dt.$$

□

Exemple 1.27. Soit $f(x) = \log(x)$ que l'on veut intégrer sur $[1, b]$. On pose $u' = 1$, $v = \log(x)$. On a $u = x$, $v' = x^{-1}$ et on trouve

$$\int_1^b \log(t)dt = \int_1^b u'(t)v(t)dt = [t \log(t)]_1^b - \int_1^b u(t)v'(t)dt.$$

Finalement, il vient, $\int_1^b u(t)v'(t)dt = \int_1^b dt = [t]_1^b$. Ainsi,

$$\int_1^b \log(t)dt = [t \log(t) - t]_1^b.$$

On trouve donc que $t \log(t) - t$ est une primitive de $\log(t)$

Exemple 1.28. Soit $f(x) = xe^x$ que l'on veut intégrer sur $[a, b]$. On pose $u' = \exp(x)$, $v = x$. On a $u = \exp(x)$, $v' = 1$ et on trouve

$$\int_a^b t\exp(t)dt = \int_a^b u'(t)v(t)dt = [t\exp(t)]_a^b - \int_a^b \exp(t)dt = [t\exp(t) - \exp(t)]_a^b.$$

Exemple 1.29. Soit $f(x) = \exp(x)\cos(x)$ que l'on veut intégrer sur $[a, b]$. On pose

$$I := \int_a^b \exp(t)\cos(t)dt.$$

On pose $u' = \cos(x)$, $v = \exp(x)$. On a $u = \sin(x)$, $v' = \exp(x)$ et on trouve

$$I = \int_a^b u'(t)v(t)dt = [\exp(t)\sin(t)]_a^b - \int_a^b \exp(t)\sin(t)dt.$$

On refait une intégration par parties avec $\int_a^b \exp(t)\sin(t)dt$. En posant, $u' = \sin(x)$, $v = \exp(x)$ on trouve $u = -\cos(x)$, $v' = \exp(x)$,

$$\int_a^b \exp(t)\sin(t)dt = [-\cos(t)\exp(t)]_a^b + \int_a^b \exp(t)\cos(t)dt.$$

On retrouve l'intégrale initiale. En combinant les deux égalités, on trouve

$$I = [\exp(t)\sin(t)]_a^b + [\cos(t)\exp(t)] - I.$$

Ainsi, $2I = [\exp(t)\sin(t) + \cos(t)\exp(t)]_a^b$ et donc

$$I = \left[\frac{\exp(t)(\sin(t) + \cos(t))}{2}\right]_a^b.$$

On peut vérifier que $\left(\frac{\exp(x)(\sin(x)+\cos(x))}{2}\right)' = \exp(x)\cos(x)$.

Terminons ce chapitre par la méthode de changement de variable.

Proposition 1.30 (Changement de variables)**.** *Soient ϕ de classe $\mathcal{C}^1$ sur $[a, b]$ et f continue sur $\phi([a, b])$. Alors $f \circ \phi$ est intégrable et*

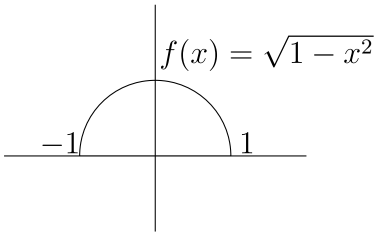
$$\int_{\phi(a)}^{\phi(b)} f(x)dx = \int_a^b f(\phi(t))\phi'(t)dt.$$

Démonstration. Soit F une primitive de f. On a $(F \circ \phi)' = \phi' \times F' \circ \phi = \phi' \times f \circ \phi$. Ainsi,

$$\int_{\phi(a)}^{\phi(b)} f(x)dx = [F]_{\phi(a)}^{\phi(b)} = [F \circ \phi]_a^b = \int_a^b f(\phi(t))\phi'(t)dt.$$

$\square$

Exemple 1.31. Posons $f(x) = \sqrt{1-x^2}$ et intégrons entre -1 et 1.



On pose $x = \phi(t)$ avec $\phi(t) = \sin(t)$. On a $dx = \cos(t)dt$, $\sin(-\frac{\pi}{2}) = -1$, $\sin(\frac{\pi}{2}) = 1$ et donc

$$\int_{-1}^{1} \sqrt{1-x^2}dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin(t)^2}\cos(t)dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\cos(t)|\cos(t)dt.$$

Sur l'intervalle considéré $\cos(t) \geq 0$ et donc l'intégrale vaut $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(t)dt$. On utilise la formule $\cos(t) = \frac{\exp(it)+\exp(-it)}{2}$, pour trouver $\cos^2(t) = \frac{\exp(2it)+2+\exp(-2it)}{4} = \frac{\cos(2t)+1}{2}$. Une primitive est $\frac{\sin(2t)+2t}{4}$. On a donc

$$\int_{-1}^{1} \sqrt{1-x^2}dx = \left[\frac{\sin(2t)+2t}{4}\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{\sin(\pi)+\pi}{4} - \frac{\sin(-\pi)-\pi}{4} = \frac{\pi}{2}.$$

Remarquons que la courbe $\sqrt{1-x^2}$ entre $x = -1$ et 1 représente le demi cercle supérieure de centre $(0,0)$ et de rayon 1. Son air sous la courbe représente donc la moitié de l'air du cercle de centre 0 et de rayon 1. On trouve donc que l'air d'un cercle de rayon 1 est $\pi(1)^2 = \pi$.

2 Intégrales impropres

Dans cette section nous voulons donner un sens à des intégrales plus générales. Par exemple, $\int_0^1 t^{-1/2}dt$ n'est à priori pas bien définit car $t^{-1/2}$ n'est pas continue en 0. Or, $t^{-1/2}$ possède $2t^{1/2}$ comme primitive et $[2t^{1/2}]_0^1 = 1$. On aurait donc envie de poser $\int_0^1 t^{-1/2}dt = 1$. De même, on pourrait vouloir prendre des bornes infinies, par exemple, donner un sens à $\int_1^\infty t^{-2}dt$. Une primitive de t^{-2} est $-t^{-1}$ et donc pour tout $M > 1$, $\int_1^M t^{-2}dx = [-t^{-1}]_1^M = -M^{-1} + 1$. Donc $\lim_{M\to\infty} \int_1^M t^{-2}dt$ existe et vaut 1. On a donc envie de poser $\int_1^\infty t^{-2}dx = 1$. Donnons maintenant un sens formel à tout cela.

2.1 Définitions

Définition 2.1 (Localement intégrable)**.** Soit I un intervalle d'extrémité $a < b$ (avec éventuellement $a = -\infty$, $b = \infty$). On dit que $f : I \to \mathbb{R}$ est localement intégrable si $\forall J \subset I$ fermé borné, la restriction de f à J est intégrable.

Remarque 2.2. Si I est fermé borné alors f est intégrable si et seulement si elle est localement intégrable.

Remarque 2.3. Si f est continue alors f est localement intégrable.

Définition 2.4 (Intégrales impropres)**.** Soit $a \in \mathbb{R}$ et soit $b \in [a, +\infty]$ (on autorise potentiellement $b = +\infty$). Soit $f : [a, b[\to \mathbb{R}$ localement intégrable. On note $\int_a^b f(t)dt$ si $\lim_{\substack{x \to b \\ x \in [a,b[}} \int_a^x f(t)dt$ existe et est finie. Dans ce cas on dira que $\int_a^b f(t)dt$ est l'intégrale impropre de f.

De même, on définit la notation d'intégrales impropres pour les intervalles de la forme $]a, b]$, $b \in \mathbb{R}$ et $a \in [-\infty, b]$.

Soit I un intervalle ouvert de $\mathbb{R}$ d'extrémité a et b (avec éventuellement $a = -\infty$ ou $b = +\infty$). Soit $f :]a, b[\to \mathbb{R}$ localement intégrable. On dira que $\int_a^b f(t)dt$ est l'intégrale impropre de f si et seulement si il existe $a < c < b$ tels que $\int_a^c f(t)dt$ et $\int_c^b f(t)dt$ sont bien définies. Dans ce cas on posera

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Exemple 2.5. L'intégrale $\int_{-\infty}^{+\infty} tdt$ n'est pas définie. En effet, pour tout $c \in \mathbb{R}$, on a $\int_c^x tdt = [\frac{t^2}{2}]_c^x = \frac{x^2}{2} - \frac{c^2}{2}$ qui tend vers l'infini lorsque x tend vers l'infini.

Exemple 2.6. D'après ce qui précède, $\int_0^1 t^{-1/2}dt = 1$ et $\int_1^{+\infty} t^{-2}dt = 1$. Plus généralement, si $\alpha \in \mathbb{R}$, considérons x^α sur $]0, 1]$ et $[1, +\infty[$. Si $\alpha \neq -1$, une primitive est donnée par $(\alpha + 1)^{-1}x^{\alpha+1}$. Cette fonction à une limite en 0 si et seulement si $\alpha > -1$. Cette fonction à une limite en $+\infty$ si et seulement si $\alpha < -1$. Lorsque $\alpha = -1$ une primitive est donnée par $\log(x)$ qui n'admet ni de limites en 0 ni en l'infinie.

Dans l'exemple précédent on a donc montré :

Proposition 2.7 (Intégrales de Riemann)**.** *L'intégrale impropre* $\int_1^{+\infty} t^\alpha dt$ *est définie si et seulement si* $\alpha < -1$. *L'intégrale impropre* $\int_0^1 t^\alpha dt$ *est définie si et seulement si* $\alpha > -1$.

2.2 Premières propriétés

Nous allons voir que les intégrales impropres ont des propriétés très similaires aux intégrales classiques.

Lemme 2.8 (Relation de Chasles)**.** *Soit* $[a, b[$ *et considérons* $a < c < b$*. Supposons que* $f : [a, b[$ *soit localement intégrable. Alors les intégrales impropres* $\int_a^b f(t)dt$ *et* $\int_c^b f(t)dt$ *ont mêmes natures. Si elles convergent, alors*

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Démonstration. Fixons $c < x < b$. La fonction f est intégrable sur $[a, x]$ et par la relation de Chasles pour les fonctions intégrables,

$$\int_a^x f(t)dt = \int_a^c f(t)dt + \int_c^x f(t)dt.$$

Puisque $\int_a^c f(t)dt$ est finie, on a bien que $\lim_{\substack{x \to b \\ x \in [a,b[}} \int_a^x f(t)dt$ existe et est finie si et seulement si $\lim_{\substack{x \to b \\ x \in [c,b[}} \int_c^x f(t)dt$ existe et est finie. En cas d'existence on a bien le relation souhaitée. □

Remarque 2.9. Un énoncé complètement symétrique peut être obtenu pour les fonctions localement intégrables sur $]a, b]$.

Remarque 2.10. Considérons f localement intégrable sur $]a, b[$ et supposons que $\int_a^b f(t)dt$ soit définie. Il existe $c \in]a, b[$ tel que $\int_a^c f(t)dt$ et $\int_c^b f(t)dt$ soient définies. Soit $c' \in]a, b[$. Il vient par la relation de Chasles que $\int_a^{c'} f(t)dt$ et $\int_{c'}^b f(t)dt$ sont définies. Ainsi, $\int_a^b f(t)dt$ est définie si et seulement si pour tout $c \in]a, b[$, $\int_a^c f(t)dt$ et $\int_c^b f(t)dt$ sont définies. Ainsi, pour vérifier qu'une intégrale impropre sur $]a, b[$ est définie, il suffira de prendre $a < c_1 < c_2 < b$, regarder $\int_{c_1}^{c_2} f(t)dt$ et faire tendre c_1 vers a puis c_2 vers b. On peut montrer que les deux limites existent si et seulement si l'intégrale impropre est définie.

Remarque 2.11. En pratique nous nous servirons de ce résultat pour découper l'intervalle suivant les endroits qui posent problèmes.

Puisque toute fonction localement intégrable sur $[a, b[$ (resp. $]a, b]$) l'est aussi sur $]a, b[$, on va se contenter des énoncés pour les intervalles de la forme $]a, b[$.

Exemple 2.12. Considérons $\int_{-\infty}^{+\infty} \frac{dt}{1+t^2}$. Les points à problèmes sont $\pm\infty$. Regardons $+\infty$. Soit $M > 1$ et considérons $\int_1^M \frac{dt}{1+t^2}$. En faisant le changement de variable $t = 1/x$ il vient $dt = -x^{-2}dx$ et $\int_1^{M^{-1}} \frac{-x^{-2}dx}{1+x^{-2}} = \int_{M^{-1}}^1 \frac{dx}{x^2+1}$. Lorsque M tend vers l'infini, M^{-1} tend vers 0. Puisque $\frac{1}{1+x^2}$ est continue sur $[0,1]$, on peut passer à la limite et trouver que $\int_0^1 \frac{dx}{x^2+1}$ converge. Donc $\int_1^{+\infty} \frac{dt}{1+t^2}$ converge. De même $\int_{-\infty}^{-1} \frac{dt}{1+t^2}$ converge et puisque la fonction est intégrable sur $[-1,1]$, il vient que l"intégrale impropre $\int_{-\infty}^{+\infty} \frac{dt}{1+t^2}$ converge. Remarquons qu'on aurait aussi pu utiliser que le fait que $\arctan(x)$ est une primitive de $\frac{1}{1+t^2}$ et que le limite de $\arctan(x)$ en $\pm\infty$ est $\frac{\pm\pi}{2}$. On trouve alors que $\int_{-\infty}^{+\infty} \frac{dt}{1+t^2} = [\arctan(t)]_{-\infty}^{+\infty} = \pi$.

Lemme 2.13 (Linéarité)**.** *Soient f, g localement intégrables sur $]a,b[$, et $c, d \in \mathbb{R}$. Alors $cf + dg$ est une fonction localement intégrable. Si les intégrales impropres $\int_a^b f(t)dt$ et $\int_a^b g(t)dt$ sont bien définies, alors $\int_a^b cf(t) + dg(t)dt$ est bien définie et*

$$\int_a^b cf(x) + dg(x)dx = c\int_a^b f(x)dx + d\int_a^b g(x)dx.$$

Démonstration. Fixons $a < v_1 < v_2 < b$. Par linéarité, on a

$$\int_{v_1}^{v_2} cf(x) + dg(x)dx = c\int_{v_1}^{v_2} f(x)dx + d\int_{v_1}^{v_2} g(x)dx.$$

Il suffit de faire tendre v_1 vers a et v_2 vers b pour obtenir le résultat. $\square$

Remarque 2.14. La réciproque est fausse. Par exemple, les intégrales impropres de $f(x) = \frac{1}{x}$ et $g(x) = \frac{1}{x} - 1$ ne sont pas définies sur $]0,1[$, pourtant celle de $f - g = 1$ l'est.

Lemme 2.15 (Monotonie)**.** *Soient f, g localement intégrable sur $]a,b[$, avec $f \leq g$ sauf en un nombre fini de point. Si les intégrales impropres $\int_a^b f(x)dx$ et $\int_a^b g(x)dx$ sont définies, alors*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

En particulier, si $f = 0$ on trouve que l'intégrale impropre $\int_a^b g(x)dx$, si elle est définie, est positive.

Démonstration. Fixons $a < c_1 < c_2 < b$. Par l'analogue pour les fonction Riemann intégrable, on a

$$\int_{c_1}^{c_2} f(x)dx \leq \int_{c_1}^{c_2} g(x)dx.$$

Il suffit de faire tendre c_1 vers a et c_2 vers b pour trouver que si $\int_a^b f(x)dx$ et $\int_a^b g(x)dx$ sont définies, alors $\int_{c_1}^{c_2} g(x) - f(x)dx$ admet une limite positive lorsque c_1 tend vers a et c_2 tend vers b. Puisque pour tout c_1, c_2, $0 \leq \int_{c_1}^{c_2} g(x) - f(x)dx$, la limite est positive ou nulle. Ainsi

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

□

Proposition 2.16 (Intégration par parties)**.** *Soient u, v de classe $\mathcal{C}^1$ sur $]a, b[$ et supposons que les limites de u et v en b existent et soient finies. Alors les intégrales impropres $\int_a^b u'(t)v(t)dt$ et $\int_a^b u(t)v'(t)dt$ ont mêmes natures. Si elles convergent, on a*

$$\int_a^b u'(t)v(t)dt = \lim_{\substack{x_2 \to b \\ x_1 \to a \\ a < x_1 < x_2 < b}} u(x_2)v(x_2) - u(x_1)v(x_1) - \int_a^b u(t)v'(t)dt.$$

Démonstration. Fixons $a < x_1 < x_2 < b$. La formule d'intégration par parties nous dit que

$$\int_{x_1}^{x_2} u'(t)v(t)dt = u(x_2)v(x_2) - u(x_1)v(x_1) - \int_{x_1}^{x_2} u(t)v'(t)dt.$$

Si uv admet une limite en a et b, et si $\int_{x_1}^{x_2} u(t)v'(t)dt$ admet une limite lorsque x_1, x_2 tendent respectivement vers a et b, alors il en est de même pour $u(x_2)v(x_2) - u(x_1)v(x_1) - \int_{x_1}^{x_2} u(t)v'(t)dt = \int_{x_1}^{x_2} u'(t)v(t)dt$. Le rôle de u et v étant parfaitement symétriques, on obtient le résultat. □

Exemple 2.17. Soit $f(x) = xe^{-x}$ que l'on veut intégrer sur $[0, +\infty[$. Fixons $0 < b < +\infty$. On pose $u' = \exp(-x)$, $v = x$. On a $u = -\exp(-x)$, $v' = 1$ et on trouve

$$\begin{aligned}
\int_0^b t\exp(-t)dt &= \int_0^b u'(t)v(t)dt \\
&= [-t\exp(-t)]_0^b + \int_0^b \exp(-t)dt \\
&= [-t\exp(-t) - \exp(-t)]_0^b.
\end{aligned}$$

Par le théorème des croissances comparés, $-t\exp(-t) - \exp(-t)$ tend vers 0 lors t tend vers $+\infty$. Il vient donc que l'intégrale impropre $\displaystyle\int_0^{+\infty} t\exp(-t)dt$ est bien définie et vaut 1.

Théorème 2.18 (Changement de variables)**.** *Soient ϕ de classe $\mathcal{C}^1$ sur $]a,b[$,* ***bijective*** *et f continue sur $\phi(]a,b[)$. Alors les intégrales $\displaystyle\int_{\phi(a)}^{\phi(b)} f(x)dx$ et $\displaystyle\int_a^b f(\phi(t))\phi'(t)dt$ ont mêmes natures et sont égales.*

Remarque 2.19. Si on ne suppose pas que ϕ est bijective, alors les intégrales $\displaystyle\int_{\phi(a)}^{\phi(b)} f(x)dx$ et $\displaystyle\int_a^b f(\phi(t))\phi'(t)dt$ n'ont pas forcément la même nature. En revanche nous allons voir que si $\displaystyle\int_a^b f(\phi(t))\phi'(t)dt$ est définie, il en va de même pour $\int_{\phi(a)}^{\phi(b)} f(x)dx$.

Démonstration. Fixons $a < c_1 < c_2 < b$. Par la formule de changement de variables, nous avons

$$\int_{\phi(c_1)}^{\phi(c_2)} f(x)dx = \int_{c_1}^{c_2} f(\phi(t))\phi'(t)dt.$$

Soit F une primitive de f. On a $\displaystyle\int_{c_1}^{c_2} f(\phi(t))\phi'(t)dt = [F\circ\phi]_{c_1}^{c_2} = F(\phi(c_2)) - F(\phi(c_1)) = \int_{\phi(c_1)}^{\phi(c_2)} f(x)dx$. Supposons que $\displaystyle\int_a^b f(\phi(t))\phi'(t)dt$ soit convergente. Alors $F(\phi(c_1))$ admet une limite lorsque c_1 tend vers a avec $a < c_1 < b$. Il en va de même pour $F(\phi(c_2))$ lorsque c_2 tend vers b avec $a < c_2 < b$ Par continuité de ϕ ces limites sont $F(\phi(a))$ et $F(\phi(b))$. On a donc que $\int_{\phi(c_1)}^{\phi(c_2)} f(x)dx$ admet une limite lorsque c_1 et c_2 tendent respectivement vers a et b avec $a < c_1 < c_2 < b$ qui est $\int_{\phi(a)}^{\phi(b)} f(x)dx$.

Supposons que $\int_{\phi(a)}^{\phi(b)} f(x)dx$ soit convergente et faisons tendre c_1 vers a par valeurs supérieures, et c_2 vers b par valeur inférieures. Puisque ϕ est bijective et $\mathcal{C}^1$, son inverse est continue, et $\phi(c_1)$ tend vers $\phi(a)$ implique c_1 tend vers a. De même c_2 tend vers b. On trouve donc que $\displaystyle\int_a^b f(\phi(t))\phi'(t)dt$ est convergente et vaut $\displaystyle\int_{\phi(a)}^{\phi(b)} f(x)dx$. □

2.3 Cas des fonctions positives

Si $f :]a,b[\to \mathbb{R}$ est positive et localement intégrable, et si $c \in]a,b[$, alors la fonction $x \mapsto \displaystyle\int_c^x f(t)dt$ est croissante et positive. Soit elle tend vers $+\infty$ lorsque

x tend vers b soit elle admet une limite. On va se donner quelques outils pour déterminer dans quelle situation nous nous trouvons. Il n'y a pas de démarches générales, l'expérience aide simplement à deviner quelle méthode utiliser en pratique. Le premier outil, est la comparaison d'intégrales.

Théorème 2.20 (Comparaison d'intégrales)**.** *Soient $f, g :]a, b[\to \mathbb{R}$ positives avec $f \leq g$ sauf en un nombre finit de point. Si l'intégrale impropre $\int_a^b g(x)dx$ est définie, alors il en est de même pour $\int_a^b f(x)dx$ et $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.*

Démonstration. Par l'absurde, supposons que $\int_a^b f(x)dx$ ne soit pas définie. On fixe $a < c_1 < c_2 < b$. On a $0 \leq \int_{c_1}^{c_2} f(x)dx \leq \int_{c_1}^{c_2} g(x)dx$. Soit $c_2 \mapsto \int_{c_1}^{c_2} f(x)dx$ admet une limite finie lorsque c_2 tend vers b soit elle tend vers l'infini. Puisque $\int_{c_1}^b f(x)dx < \int_a^b g(x)dx < \infty$ la limite est finie. De même, soit $c_1 \mapsto \int_{c_1}^b f(x)dx$ admet une limite finie lorsque c_1 tend vers a soit elle tend vers l'infini. Puisque $\int_{c_1}^b f(x)dx < \int_a^b g(x)dx < \infty$ la limite est aussi finie. On abouti donc à une contradiction. $\square$

La contraposée du résultat est aussi intéressante. Si $\int_a^b f(x)dx$ ne converge pas, elle diverge et il en est de même pour $\int_a^b g(x)dx$.

En pratique, on procède comme suit. On identifie tous les points qui peuvent poser problèmes. Ce sont les points à l'infinie si $a = -\infty$ ou $b = +\infty$ mais aussi tous les points où f n'est pas continue. Au voisinage de ces points problématiques on cherche à majorer f par une fonction localement intégrable. Les intégrales de Riemann, dont nous connaissons la nature, peuvent servir pour les majorations.

Exemple 2.21. Regardons $f(t) = \frac{\exp(-t^2)}{t^{1/2}}$ sur $]0, +\infty[$. Les points à problèmes sont 0 et $+\infty$. Regardons en $+\infty$. On a pour $t \geq 1$, $f(t) \leq \exp(-t) = g(t)$. Comme $\int_1^{+\infty} g(x)dx < \lim_{c\to+\infty}[-\exp(-t)]_1^c = \lim_{c\to+\infty} -\exp(-c) + 1 < +\infty$ il vient que $\int_1^{+\infty} f(x)dx$ converge. Reste à regarder l'intégrale sur $]0, 1[$.On a $f(t) < \frac{1}{t^{1/2}} = g(t)$. On a $\int_0^1 g(x)dx < \lim_{c\to 0}[2t^{1/2}]_c^1 = \lim_{c\to 0} 1 - 2c^{1/2} < +\infty$. Donc $\int_0^1 f(x)dx$ converge.

Théorème 2.22 (Équivalence)**.** *Soient $f, g : [a, b[\to \mathbb{R}$ strictement positives et localement intégrable. On suppose que*

$$\lim_{\substack{x\to b \\ a<x<b}} \frac{f(x)}{g(x)} = 1.$$

Alors $\int_a^b f(x)dx$ et $\int_a^b g(x)dx$ ont même natures.

Soient $f, g :]a, b] \to \mathbb{R}$ strictement positives et localement intégrable. On suppose que

$$\lim_{\substack{x\to a \\ a<x<b}} \frac{f(x)}{g(x)} = 1.$$

Alors $\int_a^b f(x)dx$ et $\int_a^b g(x)dx$ ont même natures.

Démonstration. Démontrons uniquement le première énoncé. Soit $\varepsilon \in]0,1[$. Il existe $a < b_0 < b$ tel que pour tout $b_0 < x < b$, on a $\left|\frac{f(x)}{g(x)} - 1\right| < \varepsilon$. Ainsi $|f(x) - g(x)| < \varepsilon g(x)$ et donc par inégalité triangulaire et positivité des fonction considérées

$$(1-\varepsilon)g(x) < f(x) < (1+\varepsilon)g(x).$$

Si $\int_{b_0}^{b} g(x)dx$ est convergente, il en est de même pour $(1+\varepsilon)\int_{b_0}^{b} g(x)dx$ et donc par majoration, $\int_{b_0}^{b} f(x)dx$ est convergente. Si $\int_{b_0}^{b} f(x)dx$ est convergente, $(1-\varepsilon)\int_{b_0}^{b} g(x)dx$ est convergente, et donc $\int_{b_0}^{b} g(x)dx$ l'est. On conclue grâce à la relation de Chasles. □

Ce théorème permet au voisinage des points problématiques, de remplacer une fonction f par une fonction équivalente plus simple dont on connaît la nature de l'intégrale impropre.

Exemple 2.23. Soit $f(t) = \frac{1+3t}{1+t}\exp(-t)$ sur $]0,+\infty[$. Il n'y a pas de problème en 0. En l'infini $f(t)$ est équivalente à $3\exp(-t)$ dont l'intégrale converge. Ainsi $\int_0^{+\infty} f(x)dx$ est convergente.

Théorème 2.24 (Comparaison série intégrale)**.** *Soit $f : [a,b[\to \mathbb{R}$ positive et continue. Soit $(x_k)_{k\geq 0}$ une suite croissante d'éléments de $[a,b[$ qui convergent vers b. Alors l'intégrale $\int_a^b f(x)dx$ a la même nature que la série $\sum_{k=0}^{\infty}\int_{x_k}^{x_{k+1}} f(x)dx$.*

Démonstration. Soit $X_k = \int_{x_k}^{x_{k+1}} f(x)dx$. Par la relation de Chasles, $\int_a^b f(x)dx$ et $\int_{x_0}^{b} f(x)dx$ ont même natures. On peut donc supposer sans pertes de généralités que $x_0 = a$. On a par la relation de Chasles que pour tout n, $\int_a^{x_{n+1}} f(x)dx = \sum_{k=0}^{n} X_k$. De plus, puisque f est positive, $\int_a^{x_{n+1}} f(x)dx \leq \int_a^b f(x)dx$ et la suite $n \mapsto \int_a^{x_{n+1}} f(x)dx$ est croissante. Si $\int_a^b f(x)dx$ est convergente, $n \mapsto \int_a^{x_{n+1}} f(x)dx$ est majoré et donc converge. Puisque x_n tend vers b, il vient que $\int_a^{x_{n+1}} f(x)dx$ possède une limite lorsque n tend vers l'infini qui est $\int_a^b f(x)dx$. En particulier, on a que $\sum_{k=0}^{\infty}\int_{x_k}^{x_{k+1}} f(x)dx$ converge. Réciproquement, supposons que $\sum_{k=0}^{\infty} X_k$ converge. La fonction f étant positive, les suites, $\int_a^{x_{n+1}} f(x)dx$ et $\sum_{k=0}^{n} X_k$ sont croissantes. Puisqu'elles sont égales pour tout n, $\int_a^{x_{n+1}} f(x)dx$ est majorée (par $\sum_{k=0}^{\infty} X_k$) et donc converge. □

Ce théorème possède un corollaire bien pratique à l'usage.

Corollaire 2.25. *Soit $f : [1, +\infty[\to \mathbb{R}$ positive et continue et décroissante. Alors l'intégrale $\int_1^{+\infty} f(x)dx$ converge si et seulement si $\sum_{k=1}^{n} f(k)$ converge.*

Démonstration. D'après le théorème précédent, $\int_1^{+\infty} f(x)dx$ converge si et seulement si $\sum_{k=1}^{+\infty} \int_k^{k+1} f(x)dx$ converge. Puisque f est décroissante, on a $f(k+1) = \min_{x\in[k,k+1]} f(x) \leq \max_{x\in[k,k+1]} f(x) = f(k)$. Ainsi, $f(k+1) \leq \int_k^{k+1} f(x)dx \leq f(k)$ et on trouve en sommant pour tout n

$$\sum_{k=2}^{n} f(k) \geq \sum_{k=1}^{n} \int_k^{k+1} f(x)dx \geq \sum_{k=1}^{n} f(k).$$

Donc $\sum_{k=1}^{\infty} \int_k^{k+1} f(x)dx$ a la même nature que $\sum_{k=1}^{n} f(k)$. □

Il est souvent plus facile de prouver qu'une intégrale converge plutôt qu'une série converge. Ce corollaire est donc un outil assez puissant pour prouver la convergence de certaines séries.

Exemple 2.26 (Séries de Riemann). En utilisant le critère sur les intégrales de Riemann, il vient que la série $\sum_{k=1}^{n} k^{\alpha}$ converge si et seulement si $\alpha < -1$.

2.4 Intégrales absolument convergentes

Nous allons maintenant regarder les intégrales impropres des fonctions non nécessairement positives.

Définition 2.27. Soit $f :]a, b[\to \mathbb{R}$ localement intégrable. On dit que $\int_a^b f(x)dx$ est absolument convergente si et seulement si $\int_a^b |f(x)|dx$ est convergente.

Remarquons que si f est positive, alors $\int_a^b f(x)dx$ est absolument convergente si et seulement si elle est convergente.

Proposition 2.28. *Si $\int_a^b f(x)dx$ est absolument convergente alors $\int_a^b f(x)dx$ est convergente.*

Démonstration. Fixons $a < c < b$. Considérons une suite c_n telle que $c < c_n < b$ et qui converge vers b. On a pour tout m, n

$$\int_{c_m}^{c_n} f(x)dx \leq \left| \int_{c_m}^{c_n} f(x)dx \right| \leq \int_{c_m}^{c_n} |f(x)|dx.$$

Soit $\varepsilon > 0$. Puisque $\int_a^b f(x)dx$ est absolument convergente, il existe $N \in \mathbb{N}$, tel que $m, n \geq N$ avec $m < n$ entraîne $\int_{c_m}^{c_n} |f(x)|dx < \varepsilon$. Ceci montre que $-\varepsilon <$

$\int_{c_m}^{c_n} f(x)dx < \varepsilon$. On en déduit que a suite $\int_{c}^{c_n} f(x)dx$ est une suite de Cauchy car $\int_{c}^{c_n} f(x)dx - \int_{c}^{c_m} f(x)dx = \int_{c_m}^{c_n} f(x)dx$. Puisque $\mathbb{R}$ est complet, elle converge donc. On considère de même une suite d_n telle que $a < d_n < c$ et qui converge vers a pour compléter la preuve. □

Exemple 2.29. Considérons $f(t) = \frac{\sin(t)}{t^2}$ sur $[1, +\infty[$. On a $|f(t)| = \frac{|\sin(t)|}{t^2} \leq \frac{1}{t^2}$. On a vue que $\int_1^{+\infty} t^{-2}dt$ convergeait, donc par majoration, l'intégrale $\int_1^{+\infty} \frac{|\sin(t)|}{t^2}dt$ est convergente. On a l'absolue convergence (et donc la convergence) de $\int_1^{+\infty} \frac{|\sin(t)|}{t^2}dt$.

2.5 Intégrales semi convergentes

Définition 2.30. Soit $f :]a, b[\to \mathbb{R}$ localement intégrable. On dit que $\int_a^b f(x)dx$ est semi convergente si elle est convergente mais pas absolument convergente.

Cette notion est un peu l'analogue de la série $\sum_n \frac{(-1)^n}{n}$. Cette dernière est convergente mais pas absolument convergente.

Exemple 2.31. Considérons $f(x) = \frac{\cos(2\pi x)}{x}$ sur $]1, +\infty[$. Montrons qu'elle n'est pas absolument convergente. On a $|f(x)| = \frac{|\cos(2\pi x)|}{|x|}$. Pour k entier, essayons de minorer $\int_k^{k+1} \frac{|\cos(2\pi t)|}{|t|}dt$. On a

$$\int_k^{k+1} \frac{|\cos(2\pi t)|}{|t|}dt \geq \frac{1}{(k+1)} \int_k^{k+1} |\cos(2\pi t)|dt \geq \frac{1}{(k+1)} \int_k^{k+1/4} \cos(2\pi t)dt.$$

Or $\frac{1}{(k+1)} \int_k^{k+1/4} \cos(2\pi t)dt = \frac{1}{(k+1)2\pi}[\sin(2\pi t)]_k^{k+1/4} = \frac{1}{(k+1)2\pi}$. Puisque la série $\sum_{k=1}^{+\infty} \frac{1}{(k+1)2\pi}$ diverge, il en va de même pour $\sum_{k=1}^{+\infty} \int_k^{k+1} \frac{|\cos(2\pi t)|}{|t|}dt$. Ceci montre que l'intégrale n'est pas absolument convergente.

Montrons qu'elle est convergente (et donc semi convergente). On va faire une intégration par partie. Fixons $m > 1$. On a en posant $u = \frac{1}{x}$ et $v' = \cos(2\pi x)$, $u' = \frac{-1}{x^2}$, $v = \frac{\sin(2\pi x)}{2\pi}$ et donc

$$\int_1^m \frac{\cos(2\pi t)}{t}dt = \left[\frac{\sin(2\pi x)}{2\pi t}\right]_1^m + \int_1^m \frac{\sin(2\pi t)}{2\pi t^2}dt.$$

Pour montrer que l'intégrale $\int_1^{+\infty} \frac{\sin(2\pi t)}{2\pi t^2}dt$ est convergente on va montrer qu'elle est absolument convergente. Il suffit donc de montrer que $\int_1^m \left|\frac{\sin(2\pi t)}{2\pi t^2}\right| dt$ possède une limite lorsque m tend vers $+\infty$. On a $\displaystyle\int_1^m \left|\frac{\sin(2\pi t)}{2\pi t^2}\right| dt < \int_1^m \frac{1}{2\pi t^2}dt$. Puisque $\displaystyle\int_1^{+\infty} \frac{1}{2\pi t^2}dt < +\infty$, on trouve que $\displaystyle\int_1^{+\infty} \frac{\sin(2\pi t)}{2\pi t^2}dt$ est absolument convergente et donc convergente. De plus, puisque $\left[\frac{\sin(2\pi t)}{2\pi t}\right]_1^m$ tend vers $-\frac{1}{2\pi}$ lorsque m tend vers l'infini, on trouve bien que $\displaystyle\int_1^{+\infty} \frac{\cos(2\pi t)}{t}dt$ converge.

Finissons ce chapitre par un théorème de convergence pour les intégrales semis convergentes.

Théorème 2.32 (Abel)**.** *Soit u une fonction $\mathcal{C}^1$ sur $[a,b[$, positive, décroissante, et tendant vers 0 lorsque x tend vers b. Soit v une fonction $\mathcal{C}^1$ sur $[a,b[$ et qui soit borné. Alors $\displaystyle\int_a^b u(t)v'(t)dt$ est convergente.*

Démonstration. Fixons $a < m < b$ et faisons une intégration par parties.

$$\int_a^m u(t)v'(t)dt = [u(t)v(t)]_a^m - \int_a^m u'(t)v(t)dt.$$

Comme $u(m)$ tend vers 0 lorsque m tend vers b et que v est borné, on trouve que $uv(m)$ tend vers 0 lorsque m tend vers b. Donc $[u(t)v(t)]_a^m$ converge vers $u(a)v(a)$ lorsque m tend vers b. Considérons maintenant $-\displaystyle\int_a^m u'(t)v(t)dt$. On a que $u'(t) < 0$ et qu'il existe $M > 0$ tel que $|v(t)| < M$ pour tout $a < t < b$. On a donc

$$\int_a^m |u'(t)v(t)|\, dt < M\int_a^m -u'(t)dt = [-u(t)M]_a^m = -u(m)M + u(1)M.$$

Puisque u admet une limite lorsque m tend vers b, il vient que $\displaystyle\int_a^m |u'(t)v(t)|\, dt$ admet une limite lorsque m tend vers b. Donc $\displaystyle\int_a^m u'(t)v(t)dt$ est absolument convergente et donc convergente. On a donc montré que $\displaystyle\int_a^m u(t)v'(t)dt$ admet une limite lorsque m tend vers b, ce qui conclue la preuve. □

Remarque 2.33. Pour montrer la convergence de l'intégrale $f(x) = \frac{\cos(2\pi x)}{x}$ o aurait pu appliquer le théorème d'Abel avec $u = \frac{1}{x}$ et $v = \frac{\sin(2\pi x)}{2\pi x}$.

3 Suites et séries de fonctions

3.1 Suites de fonctions

Fixons I un intervalle de $\mathbb{R}$. Pour tout $n \in \mathbb{N}$, on considère $f_n : I \to \mathbb{R}$ une fonction et $(f_n)_{n\in\mathbb{N}}$ la suite de fonction correspondante. Pour tout $x \in I$, $(f_n(x))_{n\in\mathbb{N}}$ est une suite réelle.

Définition 3.1. On dit que la suite de fonction $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur I si pour tout $x \in I$, $f_n(x)$ converge vers $f(x)$. On appellera $f(x)$ la limite simple.

Exemple 3.2. Pour $I = \mathbb{R}$ on peut considérer la suite $(x^n)_{n\in\mathbb{N}}$. Remarquons que si $x \in]-1,1[$, la suite numérique x^n converge vers 0, si $x = 1$ elle converge vers 1 et sinon elle diverge. Donc $(x^n)_{n\in\mathbb{N}}$ n'admet pas de limite simple pour $I = \mathbb{R}$. Si $I =]-1,1]$, alors la limite simple est la fonction f valant 0 sauf en 1 ou elle vaut 1. Ainsi, même si les fonctions sont continue et convergent simplement, la limite n'est pas forcément continue. Nous allons énoncer des théorèmes garantissant les propriétés de la limite (lorsqu'elle existe).

Définition 3.3. On dit que la suite de fonction $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur I si $\lim_{n\to+\infty} \sup_{x\in I} |f_n(x) - f(x)| = 0$. Dans ce cas on dit que f est la limite uniforme de $(f_n)_{n\in\mathbb{N}}$.

Comparons les deux notions de convergence avec des quantificateurs. En effet, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f si et seulement si

$$\forall x \in I, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \text{ tel que } \forall n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon.$$

D'un autre côté, $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f si et seulement si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \text{ tel que } \forall x \in I, \forall n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon.$$

On a inversé l'ordre des quantificateurs, et on trouve que si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f, alors elle converge simplement vers f. En revanche la réciproque n'est pas vrai comme le montre l'exemple suivant.

Exemple 3.4. Reprenons l'exemple précédent avec $I =]-1,1]$. Si il y a convergence uniforme il y a convergence simple, et donc la seule limite potentielle est f, la fonction valant 0 sauf en 1 ou elle vaut 1. Pour $x = 1$ on a $f_n(x) - f(x) = 0$. Pour $x < 1$, on a $|f_n(x) - f(x)| = |x|^n$. Ainsi, $\sup_{x\in I} |f_n(x) - f(x)| = \sup_{-1<x<1} |x|^n = 1 \neq 0$ ce qui prouve que $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément vers f sur $]-1,1]$. En revanche si on fixe $0 < m < 1$ alors $\sup_{-m<x<m} |f_n(x) - f(x)| = |m|^n$ tend vers 0 lorsque n tend vers l'infinie. On obtient donc que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $]-m,m[$.

Un des intérêt de la notion de convergence uniforme est le résultat suivant.

Théorème 3.5. *Soit I un intervalle ouvert. Soit une suite de fonction $(f_n)_{n\in\mathbb{N}}$ qui converge uniformément vers f sur I. On suppose que pour tout n, f_n est continue sur I. Alors f est continue sur I.*

Démonstration. Soit $x_0 \in I$. Montrons que f est continue en x_0. Pour cela on doit pour x proche de x_0 borner $|f(x) - f(x_0)|$. Pour tout n, on a par inégalité triangulaire,

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|.$$

Fixons $\varepsilon > 0$. Par convergence uniforme, il existe N tel que $n > N$ entraîne pour tout $x \in I$, $|f(x) - f_n(x)| < \frac{\varepsilon}{3}$. On a en particulier $|f(x_0) - f_n(x_0)| < \frac{\varepsilon}{3}$. Fixons $n > N$. On utilise la continuité de f_n pour déduire l'existence de $\delta > 0$ tel que $x \in I$ avec $|x - x_0| < \delta$ entraîne, $|f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3}$. On trouve donc $|x - x_0| < \delta$ implique $|f(x) - f(x_0)| < \varepsilon$. La constante $\varepsilon > 0$ étant arbitraire on a prouvé la continuité de f en x_0. Puisque $x_0 \in I$ est arbitraire, on a prouvé la continuité de f sur I. □

On trouve donc par contraposé que si la limite simple n'est pas continue alors la convergence n'est pas uniforme.

Exemple 3.6. Soit $I = \mathbb{R}$ et considérons $f_n(x) = \frac{n}{n+x^2}$. Si $x \in \mathbb{R}$, $f_n(x)$ converge vers 1. Si n est fixé, $f_n(x)$ tend vers 0 lorsque x tend vers l'infini. On a donc

$$0 = \lim_{n\to+\infty} \lim_{x\to+\infty} f_n(x) \neq \lim_{x\to+\infty} \lim_{n\to+\infty} f_n(x) = 1.$$

On ne peut donc pas inverser les limites à priori sans hypothèses supplémentaires.

Théorème 3.7. *Soit x_0 appartenant à $\bar{I}$, l'adhérence de I. Soit une suite de fonction $(f_n)_{n\in\mathbb{N}}$ définit sur I. On suppose que*
- *$(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur I.*
- *pour tout n, $f_n(x)$ converge vers ℓ_n lorsque x tend vers x_0.*

Alors
- *ℓ_n converge vers ℓ lorsque n tend vers l'infinie.*
- *$f(x)$ converge vers ℓ lorsque x tend vers x_0.*

En d'autres termes on peut inverser les limites

$$\lim_{n\to+\infty} \lim_{x\to x_0} f_n(x) = \lim_{x\to x_0} \lim_{n\to+\infty} f_n(x).$$

Démonstration. La preuve est assez similaire à la preuve de la continuité. Pour simplifier supposons que x_0 soit la borne sup de I, le cas de la borne inf se traite de manière identique. Fixons $\varepsilon > 0$. Montrons que ℓ_n converge. Soit $x \in I$.

Pour tout n, m et tout $x \in I$, considérons

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)|.$$

Puisque f_n converge uniformément, vers f il existe N tel que $n \geq N$ entraîne $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$. Ainsi pour tout $x \in I$, on a lorsque $n, m > N$, $|f_n(x) -$

$f_m(x)| < \varepsilon$. On peut donc faire tendre x vers x_0 et trouver que pour tout $n, m > N$, $|\ell_n - \ell_m| < \varepsilon$. Donc (ℓ_n) est une suite de Cauchy et puisque $\mathbb{R}$ est complet, $(\ell_n)_{n\in\mathbb{N}}$ converge. Soit $\ell \in \mathbb{R}$ la limite. Il reste à montrer que $\lim_{x\to x_0} f(x) = \ell$. Pour tout n, on a par inégalité triangulaire,

$$|f(x) - \ell| \leq |f(x) - f_n(x)| + |f_n(x) - \ell_n| + |\ell_n - \ell|.$$

Par convergence uniforme, il existe N tel que $n > N_1$ et $x \in I$ entraîne $|f(x) - f_n(x)| < \frac{\varepsilon}{3}$. Comme $f_n(x)$ converge vers ℓ_n lorsque n tend vers l'infini, pour tout n, il existe $\delta > 0$ tel que $|x - x_0| < \delta$ entraîne $|f_n(x) - \ell_n| < \frac{\varepsilon}{3}$. Puisque ℓ_n converge vers ℓ lorsque n tend vers l'infini, il existe N_2 tel que $n \geq N_2$ entraîne $|\ell_n - \ell| < \frac{\varepsilon}{3}$. Soit $n > \max(N_1, N_2)$. On trouve qu'il existe $\delta > 0$ tel que $|x - x_0| < \delta$ entraîne $|f(x) - \ell| < \varepsilon$. Ceci démontre bien la convergence de $f(x)$ vers ℓ lorsque x tend vers x_0. □

On se retrouve souvent à vouloir intégrer une suite de fonction et cela ne passe pas toujours à la limite. Là encore nous allons avoir besoin de la convergence uniforme.

Théorème 3.8. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonction convergent uniformément sur $[a, b]$ vers f. On suppose de plus que chaque f_n et f sont intégrables au sens de Riemann. Alors*

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx.$$

Démonstration. On a par linéarité de l'intégrale

$$\left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| = \left| \int_a^b f_n(x) - f(x)dx \right|.$$

Par monotonie, on a $f_n(x) - f(x) \leq |f_n(x) - f(x)|$ et donc

$$\left| \int_a^b f_n(x) - f(x)dx \right| \leq \left| \int_a^b |f_n(x) - f(x)|dx \right| = \int_a^b |f_n(x) - f(x)|dx.$$

Fixons $\varepsilon > 0$. Puisque $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[a, b]$ il existe N tel que pour tout $x \in [a, b]$, pour tout $n > N$, $|f_n(x) - f(x)| < \varepsilon$. On a donc par monotonie

$$\left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| \leq \int_a^b \varepsilon dx = \varepsilon(b - a).$$

Puisque ε est arbitrairement petit, que

$$\lim_{n\to\infty} \left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| = 0.$$

Ceci conclue la preuve. □

Exemple 3.9. Considérons la suite de fonction définie par $f_n(x) = \sum_{k=0}^{n} x^k$ et prenons un intervalle fermé de la forme $I = [0, 1-\varepsilon]$ avec $0 < \varepsilon < 1$. Puisque la série $f(x) = \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$ converge pour $x \in I$, on a convergence simple de $(f_n)_{n\in\mathbb{N}}$ vers $\frac{1}{1-x}$. La convergence est même uniforme car pour $x \in I$,

$$|f_n(x) - f(x)| \leq \sum_{k=n+1}^{\infty} |x|^k \leq \sum_{k=n+1}^{\infty} (1-\varepsilon)^k = \frac{(1-\varepsilon)^{n+1}}{1-(1-\varepsilon)} = \frac{(1-\varepsilon)^{n+1}}{\varepsilon}.$$

Lorsque n tend vers l'infini cette quantité tend uniformément vers 0, et donc f_n converge uniformément vers f sur I. On peut donc appliquer le théorème d'intégration et trouver que pour tout $0 < x < 1$,

$$\lim_{n\to\infty} \int_0^x f_n(t)dt = \int_0^x f(t)dt = \int_0^x \frac{1}{1-t}dt = [-\log(1-t)]_0^x = -\log(1-x).$$

On retrouve donc le développement en série de $\log(1-x)$:

$$\sum_{k=0}^{+\infty} \frac{x^{k+1}}{k+1} = -\log(1-x).$$

Finissons cette section par la dérivation des suites de fonctions.

Théorème 3.10. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonction sur $]a,b[$. On suppose que*

— *chaque f_n est $\mathcal{C}^1$ sur $]a,b[$.*

— *f_n converge simplement vers f sur $]a,b[$.*

— *la suite $(f_n')_{n\in\mathbb{N}}$ converge uniformément vers g sur $]a,b[$.*

Alors

— *f est de classe $\mathcal{C}^1$ sur $]a,b[$.*

— *$f' = g$.*

En particulier,

$$\lim_{n\to+\infty} f_n' = \left(\lim_{n\to+\infty} f_n\right)'.$$

Démonstration. Considérons un intervalle fermé $K \subset]a,b[$. Soit $\alpha \in K$ et soit $x \in K$. Puisque f_n' converge uniformément vers g, on trouve que par le théorème 3.8,

$$\lim_{n\to+\infty} \int_\alpha^x f_n'(t)dt = \int_\alpha^x g(t)dt.$$

Donc $\lim_{n\to+\infty} f_n(x) - f_n(\alpha) = \int_\alpha^x g(t)dt$. Puisque f_n converge simplement vers f, on obtient

$$f(x) - f(\alpha) = \int_\alpha^x g(t)dt.$$

On trouve en dérivant que $f'(x) = g(x)$. Puisque $(f'_n)_{n\in\mathbb{N}}$ converge uniformément vers g sur K on trouve par le théorème 3.5 que g est continue sur K. Puisque K est un compacte arbitraire, il vient que g est continue sur $]a, b[$. Puisque $f' = g$, il vient que f est $\mathcal{C}^1$ sur $]a, b[$. $\square$

3.2 Séries de fonctions

Définition 3.11. Soit $I \subset \mathbb{R}$ un intervalle, et soit $f_n : I \to \mathbb{R}$ une suite de fonction. Pour $N \in \mathbb{N}$, définissons la somme partiel $S_N(x) = \sum_{n=0}^{N} f_n(x)$. Si $S_N(x)$ converge simplement vers $S(x)$ lorsque x tend vers l'infini, on appellera $S(x)$ la série de fonction $S(x) = \sum_{n=0}^{\infty} f_n(x)$.

On dira que la série de fonction converge uniformément si S_N converge uniformément vers S.

Exemple 3.12. Si $f_n = x^n$ et $I =]-1, 1[$, on a $\sum_{n=0}^{N} x^n = \dfrac{x^{N+1} - 1}{1 - x}$ qu converge vers $\frac{1}{x-1}$ lorsque N tend vers l'infini. On a ici une expression explicite de la série de fonction, ce qui est plutôt rare.

Par soucis de simplification, on notera souvent $\sum_n f_n(x)$ lorsqu'on aura pas encore prouvé que la série de fonction converge. La notation $\sum_{n=0}^{\infty} f_n(x)$ sera réservé aux séries dont on a prouvé la convergence.

Commençons par donner une condition nécessaire de convergence.

Proposition 3.13. *Soit $I \subset \mathbb{R}$ un intervalle, et soit $f_n : I \to \mathbb{R}$ une suite de fonction. Si $\sum f_n(x)$ converge simplement, alors pour tout $x \in I$, $f_n(x)$ tend vers 0 lorsque n tend vers l'infini. Si la convergence est uniforme, alors $f_n(x)$ converge uniformément vers 0 sur I.*

Remarque 3.14. La réciproque est fausse. Par exemple si $f_n(x) = \frac{x}{n}$, alors $f_n(x)$ tend vers 0 lorsque n tend vers l'infini, pourtant pour tout $x \neq 0$, la série de Riemann $\sum_n \dfrac{x}{n}$ diverge.

Démonstration. Fixons $x \in I$. Si la série numérique $\sum f_n(x)$ converge alors son terme général tend vers 0, donc $f_n(x)$ tend vers 0.

Supposons maintenant la convergence uniforme. Soit S_N la somme partiel, S la limite. Fixons maintenant $\varepsilon > 0$ et soit N tel que $n \geq N$ implique pour tout

$x \in I$, $|S_n(x) - S(x)| < \varepsilon$. On a par inégalité triangulaire, pour tout $x \in I$, pour tout $n \geq N$,

$$|f_{n+1}(x)| = |S_n(x) - S_{n+1}(x)| < |S_n(x) - S(x)| + |S(x) - S_{n+1}(x)| < 2\varepsilon.$$

Puisque ε est arbitrairement petit, on a bien que f_n converge uniformément vers 0 lorsque n tend vers l'infini. $\square$

Pour les séries alternés, la situation est plus simple.

Proposition 3.15. *Soit $I \subset \mathbb{R}$ un intervalle, et soit $f_n : I \to \mathbb{R}$ une suite de fonction. On suppose que f_n est de la forme $(-1)^n g_n$ avec*
- *pour tout $x \in I$, $g_n(x)$ est positive et décroissante.*
- *$g_n(x)$ tend uniformément sur I vers 0 lorsque n tend vers l'infini.*

Alors $\sum_n f_n(x)$ converge uniformément sur I.

Démonstration. Par le critère de convergence des suites numériques alterné, on a la convergence simple. Montrons que la convergence est uniforme. Montrons que pour tout $x \in I$, on a $|\sum_{n\geq 2N} f_n(x)| < |f_{2N}(x)| = g_{2N}(x)$. On a

$$\sum_{n\geq 2N} f_n(x) = g_{2N} - (g_{2N+1} - g_{2N+2}) - \dots.$$

Puisque g_n est décroissante et positive, les termes $g_{2N+1} - g_{2N+2}$ sont négatifs. On obtient donc $\sum_{n\geq 2N} f_n(x) \leq g_{2N}$. Puisque $\sum_{n\geq 2N} f_n(x) = (g_{2N} - g_{2N+1}) + \dots$ est une somme de terme positifs, on trouve que $0 \leq \sum_{n\geq 2N} f_n(x) \leq g_{2N}$. On passe donc à la valeur absolue pour trouver $|\sum_{n\geq 2N} f_n(x)| < |f_{2N}(x)| = g_{2N}(x)$.

Puisque $(g_n)_{n\in\mathbb{N}}$ converge uniformément vers 0, on a que $\sum_{n\geq 2N} f_n(x)$ et $\sum_{n\geq 2N+1} f_n(x)$ convergent uniformément vers 0 sur I. Donc pour tout N, le reste $\sum_{n\geq N} f_n(x)$ convergent uniformément vers 0 sur I. On en déduit que $\sum_n f_n(x)$ converge uniformément sur I. $\square$

Avant de donner quelques propriétés des séries de fonctions convergent normalement, étudiant un exemple historiquement très important.

Exemple 3.16 (La fonction ζ de Riemann). Considérons la sérié de fonction dépendant de x : $\sum_n \frac{1}{n^x}$ Par le critère des séries de Riemann, on a que lorsque $x > 1$ la série converge. Montrons que la convergence est normale sur les intervalles de la forme $[a, +\infty[$, $a > 1$. Fixons un tel a. On a pour tout $x \in [a, +\infty[$

$$\left|\frac{1}{n^x}\right| \leq \frac{1}{n^a}.$$

Pour N fixé on a $\left|\sum_{n=1}^{\infty} \frac{1}{n^x} - \sum_{n=N}^{\infty} \frac{1}{n^x}\right| \leq \sum_{n=N+1}^{\infty} \frac{1}{n^a}$. Puisque la série de terme général $\frac{1}{n^a}$ converge, on a la convergence uniforme.

Montrons qu'il n'y a pas convergence uniforme sur $]1,+\infty[$. Soit $S_N = \sum_{n=0}^{N} \frac{1}{n^x}$ la somme partielle. Par l'absurde, supposons qu'il y ait convergence uniforme. Pour tout $\varepsilon > 0$ il existe $N \in \mathbb{N}$ tel que $N' \geq N$ et $x > 1$ entraîne que $\sum_{n=N'}^{\infty} \frac{1}{n^x} < \varepsilon$. Si on remplace ε par $\varepsilon/2$ par inégalité triangulaire, il vient pour pour tout $N'' \geq N' \geq N$ et $x > 1$, on a

$$\sum_{n=N'}^{N''} \frac{1}{n^x} \leq \sum_{n=N'}^{\infty} \frac{1}{n^x} - \sum_{n=N''}^{\infty} \frac{1}{n^x} < \varepsilon.$$

Prenons $N' = N$ et $N'' = 2N$. On a $\sum_{n=N+1}^{2N} \frac{1}{n^x} \geq \frac{N}{(2N)^x}$. On trouve

$$\sup_{x\in]1,\infty[} \sum_{n=N+1}^{2N} \frac{1}{n^x} \geq \sup_{x\in]1,\infty[} \frac{N}{(2N)^x} = \frac{1}{2}.$$

Si on prend $\varepsilon < \frac{1}{2}$ on trouve une contradiction.

Énonçons maintenant quelques propriétés des séries de fonctions convergent uniformément.

Théorème 3.17 (Limites de séries de fonctions)**.** *Soit $I \subset \mathbb{R}$ un intervalle, et soit $f_n : I \to \mathbb{R}$ une suite de fonction. Soit x_0 appartenant à $\bar{I}$, l'adhérence de I. Si $\sum_n f_n(x)$ converge uniformément sur I et que chaque f_n admet une limite lorsque x tend vers x_0, alors la série $\sum_{n=0}^{\infty} f_n(x)$ admet une limite lorsque x tend vers x_0 qui est $\lim_{N\to\infty} \lim_{x\to x_0} \sum_{n=0}^{N} f_n(x)$. En d'autres termes, on peut inverser les limites*

$$\lim_{x\to x_0} \lim_{N\to\infty} \sum_{n=0}^{N} f_n(x) = \lim_{N\to\infty} \lim_{x\to x_0} \sum_{n=0}^{N} f_n(x).$$

Si on suppose de plus que x_0 est dans l'intérieure de I on trouve que si $\sum_n f_n(x)$ converge uniformément sur I et que chaque f_n est continue en x_0, alors $\sum_{n=0}^{\infty} f_n(x)$ est continue en x_0.

Démonstration. C'est le théorème 3.7 appliqué à la suite de fonction $S_N(x) = \sum_{n=0}^{N} f_n(x)$. $\square$

Exemple 3.18. Puisque chaque $\frac{1}{n^x}$ sont continues sur $]1, \infty[$ on trouve que $\zeta(x) = \sum_{n=0}^{\infty} \frac{1}{n^x}$ est continue. Remarquons que puisque ζ n'admet pas de limite en $x = 1$, on trouve par le théorème que la convergence ne peut pas être uniforme sur $]1, \infty[$.

On peut appliquer les autres théorèmes des suites de fonctions dans ce contexte. On trouve l'intégration et la dérivabilité. Les preuves sont laissées en exercices

Théorème 3.19 (Intégration)**.** *Soit I un segment de $\mathbb{R}$ et soit $(f_n)_{n\in\mathbb{N}}$ ne suite de fonctions intégrables sur I. Si $\sum_n f_n(x)$ converge uniformément sur I et $\sum_{n=0}^{\infty} f_n(x)$ est intégrable, alors*

$$\int_I \sum_{n=0}^{\infty} f_n(x)dx = \sum_{n=0}^{\infty} \int_I f_n(x)dx.$$

Théorème 3.20 (Dérivation)**.** *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonction sur $]a, b[$. On suppose que*

- *chaque f_n est $\mathcal{C}^1$ sur $]a, b[$.*
- *$\sum_n f_n$ converge simplement vers S sur $]a, b[$.*
- *$\sum_n f_n'$ converge uniformément vers g sur $]a, b[$.*

Alors

- *S est de classe $\mathcal{C}^1$ sur $]a, b[$.*
- *$S' = g$.*

En particulier,

$$\sum_{n=0}^{\infty} f_n'(x) = \left(\sum_{n=0}^{\infty} f_n(x)\right)'.$$

La convergence uniforme n'est pas toujours aisé à montrer. Introduisons une convergence plus forte mais aussi plus facile à vérifier en pratique.

Définition 3.21. Soit $I \subset \mathbb{R}$ un intervalle, et soit $f_n : I \to \mathbb{R}$ une suite de fonction. On dit que $\sum_n f_n(x)$ converge normalement sur I si il existe une suite numérique (α_n) tel que pour tout n, pour tout $x \in I$, $|f_n(x)| < \alpha_n$ et $\sum_{n=0}^{\infty} \alpha_n$ converge.

Théorème 3.22. *Soit $I \subset \mathbb{R}$ un intervalle, et soit $f_n : I \to \mathbb{R}$ une suite de fonction. Si $\sum_n f_n(x)$ converge normalement sur I alors elle converge uniformément. En particulier elle converge.*

Démonstration. Soit $(\alpha_n)_{n\in\mathbb{N}}$ la suite associé à la convergence normale. Pour tout $x \in I$, la série $\sum_n f_n(x)$ converge absolument car

$$\sum_n f_n(x) \leq \sum_{n=0}^{\infty} \alpha_n < \infty.$$

En particulier, pour tout $x \in I$, la série converge, ce qui prouve la convergence simple. Montrons que la convergence est uniforme. Fixons $\varepsilon > 0$. Pour tout $N \in \mathbb{N}$, on a

$$\left|\sum_{n=0}^{\infty} f_n(x) - \sum_{n=0}^{N} f_n(x)\right| \leq \sum_{n=N}^{\infty} \alpha_n.$$

Comme la série à termes positif $\sum_{n=0}^{\infty} \alpha_n$ converge on à l'existence de $N \in \mathbb{N}$ tel que $N' \geq N$ entraîne $\sum_{n=N'}^{\infty} \alpha_n < \varepsilon$. Pour un tel N on a bien $N' \geq N$ entraîne

$$\left|\sum_{n=0}^{\infty} f_n(x) - \sum_{n=0}^{N'} f_n(x)\right| \leq \varepsilon.$$

Ceci prouve la convergence uniforme. $\square$

Exemple 3.23. On a $(\frac{1}{n^x})' = \frac{-\ln(x)}{n^x}$ et $(\frac{1}{n^x})'' = \frac{\ln^2(x)}{n^x}$. On peut montrer que $\displaystyle\sum_n \frac{-\ln(x)}{n^x}$ converge normalement sur $[a, \infty[$, $a > 1$. On trouve donc que $\zeta'(x) = \displaystyle\sum_{n=0}^{\infty} \frac{-\ln(x)}{n^x} < 0$. De même $\zeta''(x) = \displaystyle\sum_{n=0}^{\infty} \frac{\ln^2(x)}{n^x} > 0$ ce qui prouve que $\zeta(x)$ est décroissante et convexe.

On retrouve un résultat bien connu.

Corollaire 3.24. *Soit $\sum_n f_n x^n$ une série entière de rayon de convergence $R > 0$. Alors pour tout $0 < r < R$, $\sum_n f_n x^n$ converge normalement sur $[-r, r]$.*

Démonstration. On a pour tout $x \in [-r, r]$, $|f_n x^n| \leq |f_n||x|^n \leq |f_n| r^n$. Par définition du rayon de convergence, on a que $\sum_n |f_n| r^n$ converge. $\square$

En particulier, puisque $\sum_n f_n x^n$ et $\sum_n n f_n x^{n-1}$ on mêmes rayons de convergence, on trouve que $\sum_n n f_n x^{n-1}$ converge normalement et donc uniformément. On peut appliquer le résultat ci dessus pour retrouver que dans $[-r, r]$, on peut intégrer termes à termes. Il suffit de remarquer que $\sum_n n f_n x^{n-1}$ a la même rayon de convergence pour retrouver qu'on peut dériver termes à termes.

4 Espaces vectoriels normés

Dans tout ce qui suit, $\mathbb{K}$ désigne les corps $\mathbb{R}$ ou $\mathbb{C}$, et E est un $\mathbb{K}$ espace vectoriel de dimension finie.

4.1 Premières définitions

Définition 4.1. Une norme sur E est une application notée $\|.\|$ de E dans $\mathbb{R}_{\geq 0}$ telle que

1. pour tout $\lambda \in \mathbb{K}$, pour tout $x \in E$, $\|\lambda x\| = |\lambda| \times \|x\|$ (homogénéité).
2. pour tout $x \in E$, $\|x\| = 0$ si et seulement si $x = 0$ (séparation).
3. pour tout $x, y \in E$, $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).

Un espace vectoriel normé est un $\mathbb{K}$-espace vectoriel E munit d'une norme. On le note $(E, \|.\|)$ où E si il n'y a pas d'ambiguïtés.

On notera qu'on a par homogénéité $\|0\| = 0$. C'est $\|x\| = 0$ implique $x = 0$ qui est en général difficile à montrer.

Exemple 4.2. La valeur absolue $|.|$ définie une norme sur $\mathbb{C}$.

Exemple 4.3. Soit $]a, b[$ un intervalle ouvert de $\mathbb{R}$ et soit $L^2(]a, b[)$ l'espace des fonction f de I dans $\mathbb{R}$ tel que f^2 soit intégrable sur I. Pour tout $f \in L^2(]a, b[)$, définissons $||f|| = \left(\int_a^b f^2(t)dt\right)^{1/2}$. Par le lemme 1.14, $f \in L^2(]a, b[)$ tel que $||f|| = 0$ si et seulement si $f = 0$ sauf en un nombre finit de points. En particulier, $\|.\|$ ne définit par une norme sur $L^2(]a, b[)$. Pour définir une norme on va devoir restreindre Soit $\mathcal{L}^2(]a, b[) = L^2(]a, b[) \cap \mathcal{C}^0(]a, b[)$. Montrons que $\|.\|$. elle est une norme sur $\mathcal{L}^2(]a, b[)$. Soient $f \in \mathcal{L}^2(]a, b[)$, $\lambda \in \mathbb{R}$. On a par la proposition 1.15

$$||\lambda f|| = \left(\int_a^b (\lambda f)^2(t)dt\right)^{1/2} = |\lambda| \left(\int_a^b f^2(t)dt\right)^{1/2} = |\lambda| \times ||f||.$$

Par le corollaire 1.23, pour tout $f, g \in \mathcal{L}^2(]a, b[)$, $\|f + g\| \leq \|f\| + \|g\|$. Enfin par le lemme 1.14, $f \in \mathcal{L}^2(]a, b[)$ vérifie $\|f\| = 0$ si et seulement si $f = 0$ sauf en un nombre finit de points. Puisque f est continue, il vient que c'est équivalent à $f = 0$.

On dire que $x \in E$ est un vecteur unitaire ou normé si $\|x\| = 1$. Si $x \neq 0$, on lui associe un vecteur unitaire $\frac{\|x\|}{x}$. Ce dernier sera appelé vecteur unitaire associé.

Proposition 4.4. *[Inégalités triangulaires] Soit $(E, \|.\|)$ un espace vectoriel normé.*

1. *pour tout $x, y \in E$, $|\|x\| - \|y\|| \leq \|x - y\|$.*
2. *Pour tout $\alpha_1, \dots, \alpha_k \in \mathbb{K}$, pour tout $x_1, \dots, x_k \in E$, on a*

$$\left\|\sum_{i=1}^{k} \alpha_i x_i\right\| \leq \sum_{i=1}^{k} |\alpha_i| \|x_i\|.$$

Démonstration. 1. On a $\|x\| = \|x-y+y\| \leq \|x-y\| + \|y\|$. Ceci montre que

$$\|x\| - \|y\| \leq \|x-y\|.$$

Ceci prouve le résultat lorsque $\|x\| - \|y\| \geq 0$. Le rôle de x et y étant symétriques, on trouve $\|y\| - \|x\| \leq \|x-y\|$. Ceci prouve donc le résultat lorsque $\|x\| - \|y\| \leq 0$.

2. Le deuxième point s'obtient par une récurrence en utilisant l'inégalité triangulaire.

$\square$

Définition 4.5. Soient N_1, N_2 deux normes sur E. On dit que N_1 et N_2 sont équivalentes, si et seulement si il existe $c_1, c_2 \in \mathbb{R}_{>0}$ tels que pour tout $x \in E$

$$c_1 N_1(x) \leq N_2(x) \leq c_2 N_1(x).$$

Définition 4.6. Soit E un espace vectoriel. On appelle distance une application $d : E \times E \to \mathbb{R}_{\geq 0}$ vérifiant

1. pour tout $(x,y) \in E \times E$, $d(x,y) = 0$ si et seulement si $x = y$ (séparation).
2. pour tout $x, y \in E$, $d(x,y) = d(y,x)$ (symétrie).
3. pour tout $x, y, z \in E$, $d(x,z) \leq d(x,y) + d(y,z)$ (inégalité triangulaire).

Un espace métrique est un espace vectoriel E munit d'une distance d. On le notera (E,d) ou E lorsqu'il n'y a pas d'ambiguïtés.

La proposition suivante dit que tout espace vectoriel normée induit un espace métrique.

Proposition 4.7. *Soit $(E, \|.\|)$ un espace vectoriel normé.L'application $d(x,y) = \|x-y\|$ est une distance sur E. On l'appellera distance induite.*

Démonstration. Il suffit de vérifier les trois axiomes.

1. Soient $x, y \in E$. On a $d(x,y) = 0$ si et seulement si $\|x-y\| = 0$. Par séparation de la norme ceci est équivalent à $x = y$.
2. Soient $x, y \in E$. On a pa homogénéité,

$$d(x,y) = \|x-y\| = |-1|\|y-x\| = d(y,x).$$

3. Soient $x, y, z \in E$. On a par inégalité triangulaire

$$d(x,z) = \|x-y+y-z\| \leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z).$$

$\square$

De la même manière que pour le norme, on obtient

Lemme 4.8. *Soit (E, d) un espace métrique. Pour tout $x, y, z \in E$, $|d(x,y) - d(x,z)| \leq d(y,z)$.*

Démonstration. On a $d(x,y) \leq d(x,z) + d(z,y)$ et donc $d(x,y) - d(x,z) \leq d(z,y)$. Si $d(x,y) - d(x,z) \geq 0$ le résultat est prouvé, sinon on interverti y et z. $\square$

Définition 4.9. Soit (E, d) un espace métrique. Soit $a \in E$ et soit $r > 0$. On note $B(a,r) = \{x \in E | d(x,a) < r\}$ la boule ouverte de centre a et de rayon r. On note $\overline{B(a,r)} = \{x \in E | d(x,a) \leq r\}$ la boule fermée. On note $S(a,r) = \overline{B(a,r)} \setminus B(a,r) = \{x \in E | d(x,a) = r\}$ la sphere de centre a et de rayon r.

On dira qu'une partie $A \subset E$ est bornée si elle est incluse dans une boule.

Lemme 4.10. *Si E est un espace vectoriel normée, A est borné si et seulement si $\sup_{x \in A} \|x\| < \infty$.*

Démonstration. Si A est borné, alors il existe $a \in E$, $r > 0$ tel que $x \in A$ implique $d(x,a) < r$. En particulier, $\|x\| \leq \|x - a\| + \|a\| = d(x,a) + \|a\|$. Donc $\sup_{x \in A} \|x\| \leq r + \|a\| < +\infty$. Réciproquement, si $\sup_{x \in A} \|x\| = k < \infty$, alors pour tout $a \in A$, on a par inégalité triangulaire, $d(x,a) \leq \|x\| + \|a\|$. Fixons $a \in A$. On trouve donc que pour tout $x \in A$, $d(x,a) \leq k + \|a\|$. Ce qui prouve que $d(x,a) \leq k + \|a\|$ et donc que A est borné.

$\square$

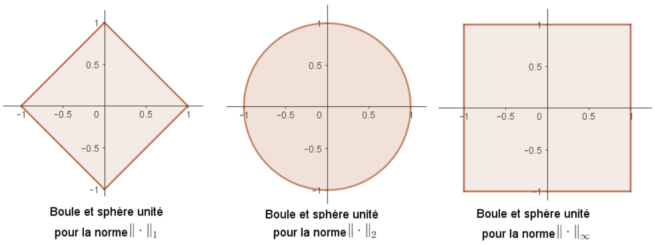
4.2 Premiers exemples

Commençons par les espaces vectoriels de dimension finie. Nous allons voire que l'on peut définir plusieurs normes.

Exemple 4.11 (Dimension finie). Soit $E = \mathbb{K}^n$, $\mathbb{K}$ désignant $\mathbb{R}$ ou $\mathbb{C}$, et pour tout $x \in E$, soient $(x_1, \ldots, x_n) = x$ les coordonnées de x dans une base fixée. On verra que les applications suivantes définissent des normes.

- On définit la norme infinie par $\|x\|_\infty = \max(|x_1|, \ldots, |x_n|)$.
- On définit la norme 1 par $\|x\|_1 = |x_1| + \cdots + |x_n|$.
- On définit la norme 2 par $\|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$.
- Plus généralement pour $1 \leq p < \infty$, on définit la norme p par $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$.

La figure ci dessous décrit la boule unité $B(0,1)$ pour ces trois normes.



Remarque 4.12. Si $E = \mathbb{R}$ ou $\mathbb{C}$, alors pour tout $1 < p < \infty$, $\|.\|_\infty = \|.\|_p$ qui correspond à la valeur absolue $|.|$.

Théorème 4.13. *Pour $p \in [1, +\infty]$, $\|.\|_p$ est une norme. Pour tout $p, q \in [1, +\infty]$, $\|.\|_p$ et $\|.\|_q$ sont équivalente.*

Démonstration. Soit $x \in E$ et soit $\lambda \in \mathbb{K}$. On a $\|\lambda x\|_\infty = \max(|\lambda x_1|, \ldots, |\lambda x_n|) = |\lambda| \max(|x_1|, \ldots, |x_n|)$. De même, $\|\lambda x\|_p = \left(\sum_{i=1}^n |\lambda x_i|^p\right)^{1/p} = |\lambda| \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. Donc $\|\lambda x\|_p = |\lambda| \|\lambda x\|_p$.

Le fait que $\|x\|_p = 0$ implique si $p = \infty$, $\max(|x_1|, \ldots, |x_n|) = 0$ et donc $|x_1| = \cdots = |x_n| = 0$. Si $p < \infty$, on a une somme de terme positif nulle $\sum_{i=1}^n |x_i|^p = 0$, et donc $|x_1| = \cdots = |x_n| = 0$. On trouve bien que $\|x\|_p = 0$ si et seulement si $x = 0$.

Il reste l'inégalité triangulaire que nous allons seulement prouver pour $p = 1, 2, \infty$. La preuve pour p quelconque (théorème de Minkowski) n'est pas spécialement difficile à suivre mais simplement technique. Soient $x, y \in E$. On a pour $p = 1$, par inégalité trnaigulaire,

$$\|x + y\|_1 = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = \|x\|_1 + \|y\|_1.$$

Pour $p = 2$, c'est une conséquence directe de l'inégalité de Cauchy Schwartz, en effet

$$\sum_{i=1}^n (x_i + y_i)^2 = (\|x\|_2)^2 + (\|y\|_2)^2 + 2\sum_{i=1}^n x_i y_i \leq (\|x\|_2)^2 + (\|y\|_2)^2 + 2\|x\|_2\|y\|_2.$$

Cette expression vaut $(\|x\|_2 + \|y\|_2)^2$ et en passant à la racine carrée, on trouve bien que $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$. Pour $p = \infty$, c'est une conséquence directe de l'inégalité triangulaire $|x_i + y_i| \leq |x_i| + |y_i|$.

Montrons l'équivalence des normes avec $p = \infty$. Quitte à réordonner la base, supposons que $\max(|x_1|, \ldots, |x_n|) = |x_1|$. Pour $q < \infty$ on a

$$(\|x\|_\infty)^q = (\max(|x_1|, \ldots, |x_n|))^q = |x_1|^q \leq \sum_{i=1}^n |x_i|^q = (\|x\|_q)^q.$$

De même,

$$(\|x\|_q)^q = \sum_{i=1}^n |x_i|^q \leq \sum_{i=1}^n |x_1|^q = n|x_1|^q = n(\|x\|_\infty)^q.$$

En prenant la racine q eme on trouve bien pour tout $x \in E$,

$$\|x\|_\infty \leq \|x\|_q \leq n^{1/q}\|x\|_\infty.$$

□

En fait l'équivalence des norme ici n'est pas un hasard. Nous verrons qu'en dimension finie toutes les normes sont équivalentes.

Présentons maintenant un exemple de normes sur un espace plus compliqué de dimension infinie (et même non dénombrable).

Exemple 4.14 (Espace de fonctions). Soit $I = [a, b]$, et soit $\mathcal{C}(I)$ l'espace des fonctions continue sur I à valeurs dans $\mathbb{R}$. Soit f dans cet espace.

- On définit la norme infinie par $\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$. Le maximum est atteint car on regarde une fonction continue sur un compacte. Le fait que cela définit bien une norme est laissé en exercice.
- On définit la norme 1 par $\|f\|_1 = \displaystyle\int_a^b |f(t)|dt$. L'homogénéité est une conséquence de la proposition 1.15. La séparation est conséquence du lemme 1.14, et l'inégalité triangulaire provient de $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ couplé avec la proposition 1.16.
- On définit la norme 2 par $\|f\|_2 = \sqrt{\displaystyle\int_a^b |f(t)|^2 dt}$, cf l'exemple 4.3.
- Plus généralement pour $1 \leq p < \infty$, on définit la norme p par

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt\right)^{1/p}.$$

  Le point délicat, à savoir l'inégalité triangulaire, sera admis.

Remarque 4.15. Les normes $\|.\|_\infty$ et $\|.\|_1$ ne sont pas équivalentes. En effet si $I = [0, 1]$, définissons la suite de fonctions f_n, affine par morceaux et qui relient $(0, 0)$, $(1 - 1/n, 0)$ et $(1, 0)$. Autrement dit, f_n qui vaut 0 sur $[0, 1 - 1/n]$ et qui vaut $nx - (n - 1)$ sur $[1 - 1/n, 1]$.Pour tout n, on a $\|f_n\|_\infty = 1$ et $\|f_n\|_1 = \frac{1}{2n}$ (c'est l'air du triangle rectangle de sommet $(1 - 1/n, 0)$, $(1, 0)$ et $(1, 1)$. On ne peut donc pas avoir de constantes $C > 0$ tel que pour tout n on ait $C\|f_n\|_\infty \leq \|f_n\|_1$. Ceci implique que les deux normes ne sont pas équivalentes.

Une suite peut être vu comme la restriction d'une fonction sur les entiers. On trouve un autre espace d'exemples.

Exemple 4.16 (Espace de suites). Soit $\mathbb{R}^{\mathbb{N}}$ l'espace des réelles à valeurs dans $\mathbb{R}$. Soit $u = (u_n)_{n\in\mathbb{N}}$ dans cet espace.

- On définit $\ell_\infty \subset \mathbb{R}^{\mathbb{N}}$ comme étant le sous espace vectoriel tel que $\|u\|_\infty = \sup_{n\in\mathbb{N}} |u_n| < \infty$. C'est un espace vectoriel normé.
- On définit $\ell_2 \subset \mathbb{R}^{\mathbb{N}}$ comme étant le sous espace vectoriel tel que $\|u\|_2 = \sqrt{\sum_{n=0}^{\infty} |u_n|^2} < \infty$. C'est un espace vectoriel normé.
- Plus généralement pour $1 \le p < \infty$, $\ell_p \subset \mathbb{R}^{\mathbb{N}}$ comme étant le sous espace vectoriel tel que $\|u\|_p = \left(\sum_{n=0}^{\infty} |u_n|^p\right)^{1/p} < \infty$. C'est un espace vectoriel normé.

Revenons enfin à un exemple important en dimension finie.

Exemple 4.17 (Espace de matrices). Soient m, n des entiers et considérons $\mathcal{M}_{n,m}(\mathbb{K})$ les matrices de taille $n \times m$. Soit $A = (a_{i,j})$ une matrice.

- On définit la norme infinie par $\|A\|_\infty = \max |a_{i,j}|$.
- On définit la norme 1 par $\|A\|_1 = \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{i,j}|$.
- On définit la norme 2 par $\|A\|_2 = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{i,j}|^2}$.
- Plus généralement pour $1 \le p < \infty$, on définit la norme p par

$$\|A\|_p = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{i,j}|^p\right)^{1/p}.$$

On prenant comme base de $\mathcal{M}_{n,m}(\mathbb{C})$ les matrices ayant des 0 partout sauf en (i, j) où elles ont un coefficient 1, on retrouve les normes de $\mathbb{K}^n$ ce qui prouve à posteriori qu'on a bien définie des normes. Notons $E = \mathbb{K}^n$. Lorsque $p = 1$, on peut montrer que l'on a $\|A\|_1 = \sup_{x\in E} \frac{\|Ax\|_1}{\|x\|_1}$. On dira que la norme matricielle est une norme subordonnée.

4.3 Suites et séries dans un espace vectoriel normé

Commençons par la notion de convergence des suites à valeurs dans un espace vectoriel normé $(E, \|.\|)$ et supposons que $\mathbb{K} = \mathbb{R}$.

Définition 4.18. Soit $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ une suite d'éléments de E. On dira que u_n converge vers $u \in E$ si la suite réelle $\|u_n - u\|$ tend vers 0 lorsque n tend vers l'infini.

Lemme 4.19 (Unicité de la limite)**.** *Si $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ converge vers u et v, alors $u = v$.*

Démonstration. Par l'absurde, supposons que $u \neq v$. On a par séparation et par inégalité triangulaire

$$0 < \|u - v\| \leq \|u - u_n\| + \|u_n - v\|.$$

Puisque $\|u - u_n\| + \|u_n - v\|$ tend vers 0 on a une contradiction. □

Remarque 4.20. Supposons que N_1 et N_2 soient deux normes équivalentes de E et supposons que $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ converge pour la norme N_2. Soient $c_1, c_2 \in \mathbb{R}_{>0}$ tels que pour tout $x \in E$

$$c_1 N_1(x) \leq N_2(x) \leq c_2 N_2(x).$$

On a en particulier

$$c_1 N_1(u_n - u) \leq N_2(u_n - u) \leq c_2 N_1(u_n - u),$$

ce qui prouve que $N_1(u_n - u)$ tend vers 0. La situation étant symétrique on trouve que $(u_n)_{n\in\mathbb{N}}$ converge pour N_2 si et seulement si $(u_n)_{n\in\mathbb{N}}$ converge pour N_1. Les limites sont alors les mêmes.

Exemple 4.21. Dans $\mathcal{C}(I)$, la convergence en norme pour $\|.\|_\infty$ correspond à la convergence uniforme. Montrons qu'ici la convergence dépend du choix de la norme. Reprenons la suite de fonction f_n que l'on a considéré dans la remarque 4.15. On a $\|f_n\|_1 = \frac{1}{2n}$ donc f_n converge vers 0 pour la norme $\|.\|_1$. En revanche, $\max_{x\in[0,1]} |f_n(x)| = 1$ donc f_n ne converge pas vers 0 pour la norme $\|.\|_\infty$.

Définition 4.22. Soit $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ une suite d'éléments de E et soit $S_N = \sum_{n=0}^{N} u_n$ la somme partielle. On dira que la série $\sum_n u_n$ converge si la somme partielle S_N converge. Dans ce cas on notera $\sum_{n=0}^{\infty} u_n$ la limite.

Comme pour les séries, si $\sum_{n=0}^{N} u_n$ converge alors $\|u_n\|$ tend vers 0, mais la réciproque est fausse.

Enfin finissons par une notion qui sera utile par la suite. On appelle sous suite extraite de $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ une suite de la forme $(u_{\phi(n)})_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, avec $\phi : \mathbb{N} \to \mathbb{N}$ strictement croissante. Une sous suite convergente d'une suite convergente reste convergente. On appellera valeur d'adhérence de (u_n) tout élément de E qui peut être la limite d'une suite extraite.

Exemple 4.23. Si $u_n = (-1)^n$ les valeurs d'adhérences sont -1 et 1.

4.4 Introduction à la topologie

Commençons par la définition des parties ouvertes.

Définition 4.24. Soit $(E, \|.\|)$ un espace vectoriel normé.

Une partie $V \subset E$ est une voisinage du point $a \in E$ si il existe $r > 0$ tel que $B(a, r) \subset V$.

On dit qu'une partie U est ouverte si c'est un voisinage de chacun de ses points. En d'autres termes, de chacun des points $a \in U$, on peut trouver $r > 0$ tel que $B(a, r) \subset U$.

Exemple 4.25. Les boules ouvertes sont ouvertes. En effet Soit $b \in B(a, r)$ et soit $r' = \|b - a\| < r$. Alors par inégalité triangulaire, pour tout $c \in E$, on a $\|a - c\| \leq \|a - b\| + \|b - c\|$. Si $\|b - c\| < r - r'$, on trouve que $\|a - c\| < r$. Donc $B(b, r - r') \subset B(a, r)$. Si on considère $\mathbb{R}$, munit de la norme $\|.\|_\infty = |.|$ on trouve par exemple, que les intervalles ouverts $]a, b[$ avec $a < b$ sont des ouverts. De même, $]b, +\infty[$ et $] - \infty, a[$ sont des ouverts. Par contre $]a, b]$ n'est pas ouvert, puisque pour tout r, $]a, b]$ n'est pas inclue dans la boule ouverte $]b - r, b + r[$.

Dans ce qui suit on omettra la norme de E par soucis d'allégement des notations.

Proposition 4.26. *Soit E un espace vectoriel normé.*

1. *$\varnothing$ et E sont des ouverts.*
2. *Une union d'ouverts est un ouvert.*
3. *Une intersection finie d'ouvert est un ouvert.*

Remarque 4.27. Pour la troisième point le fait que l'intersection soit finie est crucial. Par exemple dans $\mathbb{R}$, les ouverts $] - 1/n, 1/n[$ ont une intersection égale à $\{0\}$ qui n'est pas un ouvert de $\mathbb{R}$.

Démonstration.

1. Le premier point est une conséquence directe de la définition. La preuve est laissée est exercice.
2. Soit $(U_i)_{i \in I}$ des ouverts et montrons que $\bigcup_{i \in I} U_i$ est une partie ouverte. Soit $x \in \bigcup_{i \in I} U_i$. Alors il existe $i \in I$ tel que $x \in U_i \subset \bigcup_{i \in I} U_i$. Puisque U_i est un ouvert, il existe r tel que $B(x, r) \subset U_i$. On a donc bien un voisinage de x dans $\bigcup_{i \in I} U_i$.
3. Soit $V_1, \ldots, V_n$ des ouverts. Si $\bigcap_{i=1}^{n} V_i = \varnothing$ c'est un ouvert par le point 1. Sinon soit $x \in \bigcap_{i=1}^{n} V_i$. Puisque x est dans chaque V_i qui est un ouvert, il

existe $r_i > i$ tel que pour tout $i = 1, \dots, n$, on a $B(x, r_i) \subset V_i$. Si on prend $r = \min(r_1, \dots, r_n) > 0$ on a bien $B(x, r) \subset \bigcap_{i=1}^{n} V_i$.

$\square$

Définition 4.28. Une partie de E est dite fermé si on complémentaire est ouvert.

Exemple 4.29. Soit $a < b$. Puisque les intervalles $]b, +\infty[$ et $]-\infty, a[$ sont ouverts, on trouve que $[a, b]$ est fermé. De même, $[b, +\infty[$ et $]-\infty, a]$ sont des fermés. Par contre $[a, b[$ n'est ni fermé ni ouvert.

Exemple 4.30. Les boules fermés sont des fermés. En effet, considérons $\overline{B(a, r)}$ et soit $x \in E \setminus \overline{B(a, r)}$. On a $\|x - a\| = r' > r$. Montrons que $B(x, \|x - a\| - r) \subset E \setminus \overline{B(a, r)}$. Soit $x' \in B(x, \|x - a\| - r)$. Il suffit de montrer que $\|x' - a\| > r$. On a $\|x - a\| \leq \|x - x'\| + \|x' - a\|$. Par conséquent $\|x' - a\| \geq \|x - a\| - \|x - x'\| > r$. On trouve donc que $B(x, \|x - a\| - r) \subset E \setminus \overline{B(a, r)}$.

Proposition 4.31. *Soit E un espace vectoriel normé.*

1. *$\varnothing$ et E sont des fermés.*
2. *Une intersection de fermés est un fermé.*
3. *Une union finie de fermé est un fermé.*

Démonstration. Le point 1 est une conséquence directe de son homologue de la proposition 4.26. Soient F_i des fermés. On a

$$E \setminus \left(\bigcap_i F_i \right) = \bigcup_i (E \setminus F_i) \,,$$

$$E \setminus \left(\bigcup_i F_i \right) = \bigcap_i (E \setminus F_i) \,.$$

Comme les F_i sont des fermés, les $E \setminus F_i$ sont des ouverts. D'après le point 2 de la proposition 4.26, $\bigcup_i (E \setminus F_i)$ est un fermé. De même si les F_i sont en nombre fini, $\bigcap_i (E \setminus F_i)$ est aussi un fermé. $\square$

Exemple 4.32. L'union de fermés $[-1 + 1/n, 1 - 1/n]$, $n \in \mathbb{N}$, est $]-1, 1[$ n'est pas un fermé. L'hypothèse qu'on en prend un nombre fini est donc nécessaire.

Enfin, finissons par une caractérisation bien pratique des fermés.

Proposition 4.33. *Une partie F est fermé si et seulement toute suite de F qui converge dans E a une limite dans F.*

Exemple 4.34. L'intervalle $]0,2[$ n'est pas fermé car $1/n$, $n \in \mathbb{N}$ est à valeurs dans $]0,2[$, converge dans $\mathbb{R}$, mais sa limite 0 n'est pas dans $]0,2[$.

Démonstration. Supposons que F soit fermé et soit $(u_n)_{n\in\mathbb{N}}$ à valeurs dans F qui converge dans E vers u. Par l'absurde, supposons que la limite u n'est pas dans F. Alors, elle est dans l'ouvert $E \setminus F$ et donc il existe $r > 0$ tel que $B(u,r) \subset E \setminus F$. Comme $(u_n)_{n\in\mathbb{N}}$ converge vers u, il existe $N \in \mathbb{N}$ tel que $n \geq N$ entraîne $\|u - u_n\| < r$. Donc $n \geq N$ entraîne $u_n \in B(u,r) \subset E \setminus F$. Ceci contredit le fait que la suite est à valeurs dans F et montre que $u \in F$.

Réciproquement, supposons que toutes suites convergentes à valeurs dans F a une limite dans F. Montrons que $E \setminus F$ est un ouvert. Par l'absurde, supposons que ce ne soit pas le cas. Il existe $a \in E \setminus F$ tel que pour tout $r > 0$, la boule $B(a,r)$ n'est pas inclue dans $E \setminus F$. Considérons les boules $B(a,1/n)$ et soit u_n qui soit dans $B(a,1/n)$ mais pas dans $E \setminus F$. La suite $(u_n)_{n\in\mathbb{N}}$ est à valeurs dans F et converge vers $a \in E \setminus F$. On trouve donc une contradiction et ceci achève la preuve. $\square$

Terminons cette section par la notion d'adhérence et d'intérieure.

Définition 4.35. Soit E un espace vectoriel normé et soit $A \subset E$ un sous ensemble. On dit que $x \in E$ est dans l'intérieure de A si il existe $r > 0$ tel que $B(a,r) \subset A$. L'intérieure de A, noté $\mathring{A}$ est l'ensemble des points à l'intérieure de A.

On dit que $x \in E$ est dans l'adhérence de A, si pour tout $r > 0$, on a $B(a,r) \cap A \neq \varnothing$. L'adhérence de A, noté $\overline{A}$ est l'ensemble des points dans l'adhérence.

La frontière de A, noté $\mathrm{Fr}(A)$ est $\overline{A} \setminus \mathring{A}$.

On a par définition
$$\mathring{A} \subset A \subset \overline{A}.$$

Exemple 4.36. L'adhérence de $[0,1[$ est $[0,1]$, son intérieure est $]0,1[$, sa frontière est $\{0,1\}$.

Si $E = \mathbb{R}$, l'adhérence de $\mathbb{Q}$ est $\mathbb{R}$, son intérieure est $\varnothing$, et sa frontière est $\mathbb{R}$.

Lemme 4.37. *L'adhérence de A est le plus petit fermé contenant A, l'intérieure de A est le plus grand ouvert contenu dans A.*

Démonstration. Le fait que l'intérieure soit un ouvert découle de la définition. De plus si $\mathring{A} \subset O \subset A$ est un ouvert, alors tout élement de O possède un voisnage ouvert dans A. Ainsi tout élément de O est dans $\mathring{A}$ et par conséquent, $O \subset \mathring{A}$ (et donc $O = \mathring{A}$). Donc $\mathring{A}$ est bien le plus grand ouvert contenu dans A.

Montrons que l'adhérence est un fermé. On va montrer que $E \setminus \overline{A}$ est ouvert. Soit $a \notin \overline{A}$. Il existe $\varepsilon > 0$ tel que $B(a,\varepsilon) \cap A = \varnothing$. Si $x \in B(a,\varepsilon)$ et $b \in A$, on a

$\|b-a\| \leq \|b-x\| + \|x-a\|$ et donc $\|b-x\| \geq \varepsilon - \|x-a\| > 0$. Donc pour tout $x \in B(a,\varepsilon)$, il existe $r > 0$ tel que $B(x,r) \cap A = \varnothing$. Donc a possède un voisinage ouvert de point dans $E \setminus \overline{A}$, ce qui prouve que $E \setminus \overline{A}$ est un ouvert. Donc l'adhérence est un fermé. Soit $A \subset F \subset \overline{A}$ un fermé. En passant au complémentaire, on trouve que $E \setminus \overline{A} \subset E \setminus F \subset E \setminus A$ sont des ouverts. De plus $E \setminus \overline{A}$ est l'ensemble des éléments de $E \setminus A$ possédant une voisinage dans $E \setminus \overline{A}$. C'est l'intérieur de $E \setminus \overline{A}$ et donc $E \setminus \overline{A} = E \setminus F$. Ceci prouve que $\overline{A} = F$. □

4.5 Continuité, limite

Soient $(E, \|.\|_E)$ et $(F, \|.\|_F)$ des espaces vectoriels normés.

Définition 4.38. Soit $A \subset E$. Soit $x_0 \in \overline{A}$. On dira que $f : A \to F$ tend vers $\ell \in F$ lorsque x tend vers x_0, si pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $x \in B(x_0, \eta) \cap A$ implique $f(x) \in B(\ell, \varepsilon)$.

Si $x_0 \in A$ et f tend vers $f(x_0)$ lorsque x tend vers x_0, on dira que f est continue en x_0. On dira que f est continue sur A si f est continue en tout point de A.

Remarque 4.39. Si $a \in \overline{A} \setminus A$, et que f continue sur A, admet une limite b en a alors on peut prolonger f sur $A \cup \{a\}$ avec comme valeur b. On dira alors qu'on a prolongé f par continuité en a.

Si $A = E$ on trouve la notation plus familière que pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que

$$\|x - x_0\|_E < \eta \Rightarrow \|f(x) - f(x_0)\|_F < \varepsilon.$$

Si $E = F = \mathbb{R}$ que l'on munit de la norme $|.|$, on retrouve la notion classique de convergence. On va pouvoir appliquer cette définition à des exemples plus compliqués faisant intervenir plusieurs variables.

Exemple 4.40. Supposons que $E = \mathbb{R}^2$ que l'on munit de la nome $\|.\|_\infty$. Soit $A = \mathbb{R}^2 \setminus \{(0,0)\}$. Soit $F = \mathbb{R}$, munit de la norme $|.|$. Considérons $f : \mathbb{R}^2 \to \mathbb{R}$ tel que $f(x,y) = \frac{x+y}{x^2+y^2}$. La fonction f est bien définie sur A. Montrons qu'elle n'admet pas de limite en $(0,0)$. Sur $\mathbb{R}^2$ on a vu que les normes ∞ et 2 sont équivalentes, on peut donc remplacer par la norme 2. Passons en coordonnées polaires, en remplaçant $(x,y) \in A$, par $(r\cos(\theta), r\sin(\theta))$, avec $r > 0, \theta \in \mathbb{R}$. On a alors $\|(x,y) - (0,0)\|_2 < \eta$ équivalent à $|r| < \eta$. Donc on est ramené à un problème à une variable. On a

$$\frac{r\cos(\theta) + r\sin(\theta)}{(r\cos(\theta))^2 + (r\sin(\theta))^2} = \frac{r(\cos(\theta) + \sin(\theta))}{r^2}.$$

Ainsi quand r tend vers 0, $f(r\cos(\theta), r\sin(\theta))$ tend vers l'infini, ce qui prouve que f ne possède pas de limite en $(0,0)$.

Exemple 4.41. Si on considère $f(x,y) = \frac{xy}{x^2+y^2}$ on a un tout autre phénomène. En effet, $(x,0)$ et (x,x) tendent tout deux vers $(0,0)$ lorsque x tend vers 0. Pourtant $f(x,0) = 0$ et $f(x,x) = \frac{x^2}{2x^2} = \frac{1}{2}$. Ici, la limite dépend de la manière dont (x,y) tend vers $(0,0)$, on en déduit donc que f n'admet pas de limites en $(0,0)$. Si on pense au cas unidimensionnel, on a deux manière de faire tendre x vers x_0 : à droite et à gauche, ici dans $\mathbb{R}^2$ il y a une infinité de manières et pour avoir la convergence il faut que pour toutes ces manières on trouve la même limite.

Donnons maintenant quelques propriétés des fonctions continues. Les deux premières sont immédiates et laissées en exercice.

Proposition 4.42. *La somme et le produit de deux fonctions continues est continue. Les fonctions constantes sont continues.*

Proposition 4.43 (Composition)**.** *Soient E, F, G des espaces vectoriels normées. Soient $A \subset E$ et $B \subset F$. Soit $f : A \to B$ et $g : B \to G$. Si f et g sont continues, alors $g \circ f$ est continue.*

La proposition suivante, très utile en pratique n'est pas du tout immédiate.

Proposition 4.44 (Caractérisation séquentielle des limites)**.** *Soit $(E, \|.\|_E)$ un espace vectoriel normé et soit $A \subset E$ une partie. Soit $(F, \|.\|_F)$ un espace vectoriel normé. Considérons $f : A \to F$. Alors $f(x)$ tend vers b quand x tend vers $a \in \overline{A}$ si et seulement si pour toute suite a_n de A qui converge vers a, la suite $b_n := f(a_n)$ converge vers b.*

Démonstration. Supposons que $f(x)$ tend vers b quand x tend vers a. Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que $x \in B(a,\eta) \cap A$ entraîne $f(x) \in B(b,\varepsilon)$. Puisque a est dans l'adhérence de A on a $A \cap B(a,\eta) \neq \varnothing$. Soit a_n une suite d'éléments de A convergent vers a. Il existe $N > 0$ tel que $n \geq N$ entraîne $a_n \in B(a,\eta)$. Par continuité de f on a donc $f(a_n) \in B(b,\varepsilon)$. Ceci prouve que $f(a_n)$ tend vers $f(b)$.

Montrons la réciproque par contraposée. Supposons que f ne tende pas vers b lorsque x tend vers a. Il existe $\varepsilon > 0$ tel que pour tout $\eta > 0$ on ait l'existence de $x \in B(x_0,\eta) \cap A$ tel que $f(x) \notin B(b,\varepsilon)$. Si on prend $\eta = 1/n$, on a une suite $a_n \in B(x_0,\eta) \cap A$ tel que $f(a_n) \notin B(b,\varepsilon)$. Ainsi, on a une suite a_n qui converge vers a mais $f(a_n)$ ne converge pas vers b. □

Les application continues sont les applications qui préservent les ouverts et les fermés, comme le montre le théorème suivant.

Théorème 4.45. *Soit $(E, \|.\|_E)$ un espace vectoriel normé et soit $A \subset E$ une partie. Soit $(F, \|.\|_F)$ un espace vectoriel normé. Considérons $f : A \to F$. Alors f est continue si et seulement si l'image réciproque de tout ouvert (resp. fermé) est un ouvert (resp. fermé).*

Démonstration. Supposons que f est continue et soit V un ouvert de F. Pour tout $x \in f^{-1}(V)$, $f(x) \in V$ et comme V est ouvert, il existe $\varepsilon_x > 0$ tel que $B(f(x), \varepsilon_x) \subset V$. Puisque f est continue, il existe $\eta_x > 0$ tel que $f(B(x, \eta_x)) \subset B(f(x), \varepsilon_x)$. Donc x possède un voisinage ouvert dans $f^{-1}(V)$. Donc $f^{-1}(V)$ est un ouvert.

Réciproquement, supposons que l'image réciproque de tout ouvert est un ouvert. Soit $V \subset F$ un ouvert. Quitte à changer d'ouvert on peut supposer qu'il existe $v \in V$ qui soit image d'un point $u \in A$ par f (c'est à dire $f(u) = v$). Soit $\varepsilon > 0$ tel que $B(v, \varepsilon) \subset V$. Alors $f^{-1}(B(v, \varepsilon))$ est un ouvert qui contient u. On a donc l'existence de $\eta > 0$ tel que $x \in B(u, \eta) \cap A$ entraîne $f(x) \in B(v, \varepsilon)$. Ceci prouve que f est continue.

La preuve pour les fermés est similaire et laissée en exercice. □

Finissons par la notion de fonction k-lipschitzienne. Soit $k > 0$. On dit que $f : A \to F$ est k-lipschitzienne si pour tout $(x, y) \in A \times A$ on a

$$\|f(x) - f(y)\|_F \leq k\|x - y\|_E.$$

Si f est k-lipschitzienne, il vient de manière assez directe, que f est continue. En effet, si pour tout $\varepsilon > 0$, on prend $\eta = k^{-1}\varepsilon$ et on trouve $\|x-y\|_E < \eta$ implique $\|f(x) - f(y)\|_F < \varepsilon$.

4.6 Compacité

Commençons par la définition.

Définition 4.46. Soit E un espace vectoriel normé. On dit qu'une partie $K \subset E$ est un compacte si toute suite à valeur dans K possède une valeur d'adhérence. Autrement dit si toute suite à valeurs dans K possède une sous suite extraite qui converge.

Par exemple $[0, 1[$ n'est pas compacte car la suite $1 - 1/n$ ne possède pas de sous suite extraite qui converge dans $[0, 1[$. En revanche $[0, 1]$ est bien compacte.

Commençons par quelques propriétés des espaces compactes.

Proposition 4.47. *Soit E un espace vectoriel normé. Si $K \subset E$ est compacte, alors K est fermé et borné.*

Démonstration. Montrons que K est fermé. Soit $(a_n)_{n\in\mathbb{N}}$ une suite d'élément de K qui converge dans E vers a. Il s'agit de montrer que $a \in K$. Puisque K est compacte, la suite extraite a_n converge dans K. Puisque cette limite est a, on a bien $a \in K$. Montrons par contraposé que K non borné implique K non compacte. Si K n'est pas borné, pour tout n, on peut construire une fonction croissante ϕ tel que $\|a_{\phi(n)}\| \geq n$. Ainsi, la suite extraite $(a_{\phi(n)})_{n\in\mathbb{N}}$ diverge. Ainsi K n'est pas un compacte. □

Proposition 4.48. *Soit E un espace vectoriel normé et soit $f : E \to E$ une application continue. Soit K un compacte. Alors $f(K)$ est compacte.*

Démonstration. Soit $(b_n)_{n\in\mathbb{N}}$ une suite d'éléments de $f(K)$. Pour tout n il existe donc $a_n \in K$ tel que $f(a_n) = b_n$. Puisque K est compacte il existe une suite extraite $\phi(n)$ tel que $(a_{\phi(n)})_{n\in\mathbb{N}}$ converge vers $a \in K$. Par la proposition 4.44, $(b_{\phi(n)})_{n\in\mathbb{N}}$ converge vers $f(a) \in f(K)$. □

rapellons que l'on a définit la norme $\|.\|_\infty$ sur $\mathbb{R}^p$ par $\|(x_1, \dots, x_p)\|_\infty = \max(|x_1|, \dots, |x_p|)$.

Théorème 4.49 (Théorème de Bolzano Weierstrass)**.** *Considérons $\mathbb{R}^p$ munit de la norme $\|.\|_\infty$. Alors toute suite borné possède une sous suite extraite convergente.*

Démonstration. Soit $(a_n)_{n\in\mathbb{N}}$ une suite borné et écrivons $a_n = (a_{n,1}, \dots, a_{n,p})$. Par le théorème de Bolzano Weierstrass dans $\mathbb{R}$, il existe une suite extraite $\phi_1(n)$ tel que $a_{\phi_1(n),1}$ converge vers a_1. Considérons maintenant la suite extraite $a_{\phi_1(n)} = (a_{\phi_1(n),1}, \dots, a_{\phi_1(n),p})$. Par le théorème de Bolzano Weierstrass dans $\mathbb{R}$, il existe une suite extraite ϕ_2 tel que $a_{\phi_2\circ\phi_1(n),2}$ converge vers a_2. Bien entendu, $a_{\phi_2\circ\phi_1(n),1}$ converge toujours vers a_1. On procède ainsi pour toutes les coordonnés pour obtenir une suite extraite qui converge. □

Remarque 4.50. On a vu qu'un compacte est fermé et borné. Montrons la réciproque pour $E = \mathbb{R}^p$ munit de la norme $\|.\|_\infty$. Soit K fermé et borné et soit $(a_n)_{n\in\mathbb{N}}$ une suite d'éléments de K. Par le théorème de Bolzano Weierstrass, il existe une suite extraite $\phi(n)$ tel que $(a_{\phi(n)})_{n\in\mathbb{N}}$ converge. Puisque K est fermé la limite est dans K, ce qui montre bien que K est compacte.

On obtient la proposition très importante.

Proposition 4.51 (Equivalence des normes)**.** *Considérons $\mathbb{R}^p$ munit de deux normes N_1 et N_2. Alors N_1 et N_2 sont équivalentes.*

Démonstration. Soit $(e_1, \dots, e_p)$ la base canonique de $\mathbb{R}^p$. Soit $x \in \mathbb{R}^p$ et écrivons $x = x_1e_1 + \cdots + x_pe_p$. Rappelons que l'on a définit la norme $\|.\|_\infty$ par $\|x\|_\infty = \max(|x_1|, \dots, |x_p|)$. Il suffit de montrer que toute norme N est équivalente à $\|.\|_\infty$. On a par inégalité triangulaire

$$N(x) \leq \sum_{i=1}^{p} N(x_ie_i) \leq \sum_{i=1}^{p} |x_i|N(e_i) \leq b\|x\|_\infty,$$

avec $b = \sum_{i=1}^{p} N(e_i)$. Ceci montre que N est borné par $\|.\|_\infty$. De plus on a pour tout $x, y \in \mathbb{R}^p$, par inégalité triangulaire

$$|N(x) - N(y)| \leq N(x - y) \leq b|x - y|.$$

Donc l'application qui a x associe $N(x)$ est Lipschitzienne de $(\mathbb{R}^p, \|.\|_\infty)$ dans $(\mathbb{R}, |.|)$. Elle est en particulier continue.

Montrons que la sphère unité

$$S(0,1) = \{x \in \mathbb{R}^p | \|x\|_\infty = 1\}$$

est fermée. Définissons les boules ouvertes et fermées

$$B(0,1) = \{x \in \mathbb{R}^p | \|x\|_\infty < 1\}$$

$$\overline{B(0,1)} = \{x \in \mathbb{R}^p | \|x\|_\infty \leq 1\}.$$

On a $S(0,1) = \overline{B(0,1)} \setminus B(0,1)$, est égale à un fermé privé d'un ouvert. Soit $(a_n)_{n\in\mathbb{N}}$ une suite de $S(0,1)$ convergent dans $\mathbb{R}^p$. Puisque $\overline{B(0,1)}$ est fermé, la limite est dans $\overline{B(0,1)}$. Si la limite était dans l'ouvert $B(0,1)$, pour n assez grand, a_n appartiendrait à $B(0,1)$ et donc pas à $S(0,1)$. Ceci contredit le fait que la suite $(a_n)_{n\in\mathbb{N}}$ est à valeurs dans $S(0,1)$. Donc $S(0,1)$ est fermé. Puisque $S(0,1)$ est borné, c'est un compacte.

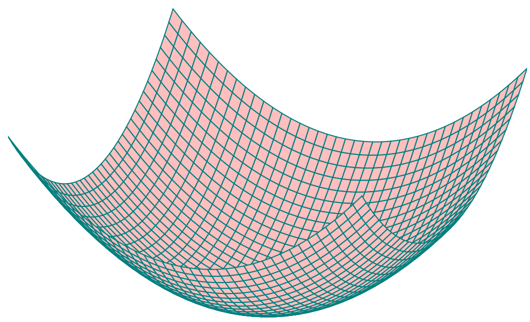
Donc la restriction à $S(0,1)$ de l'application continue N est compacte (et donc fermée) et atteint sa borne inférieure $a \geq 0$. Par l'axiome de séparabilité, $a \neq 0$. Pour tout $x \in \mathbb{R}^p$, on a $\|x\|_\infty^{-1} x \in S(0,1)$ et

$$a \leq N(\|x\|_\infty^{-1} x) = \|x\|_\infty^{-1} N(x).$$

On trouve donc que $a\|x\|_\infty \leq N(x)$, ceci achève la preuve que N est équivalente à $\|.\|_\infty$. □

5 Fonctions de plusieurs variables

Le but de cette section est d'étudier les fonctions de plusieurs variables. Dans ce qui suit, de tels fonctions seront à valeur de $A \subset \mathbb{R}^m$ dans $\mathbb{R}$ (on pourrait dans certain énoncé remplacer $\mathbb{R}$ par $\mathbb{R}^n$ mais par soucis de simplification on restera sur $\mathbb{R}$). Par exemple, si $m = 2$, on peut considérer la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ qui a (x, y) associe $x^2 + y^2$. Pour $m = 1$, le graphe d'une fonction était la courbe $y = f(x)$, pour $m = 2$ c'est la surface $z = f(x, y)$. Par exemple la surface $z = x^2 + y^2$ est représentée par



Le but de ce qui suit va être de donner des outils pour étudier de tels fonctions, comme on étudierait une fonction d'une variable classique. Nous allons voire que toutes les notions sont ici plus délicates à manipuler.

Rappelons qu'en dimension finie toutes les normes sont équivalentes. On notera $\|.\|$ la norme $\|.\|_\infty$ mais dans tout ce qui suit on pourra la remplacer par n'importe quel autre norme. Rappelons la définition 4.38 dans ce contexte. Une fonction $f : A \to \mathbb{R}$ est continue si pour tout $x_0 \in A$, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $\|x - x_0\| < \eta$ et $x \in A$ entraîne $|f(x) - f(x_0)| < \varepsilon$. Regardons maintenant la notion de dérivabilité.

5.1 Fonctions de classe $\mathcal{C}^1$

Soit f une fonction de $A \subset \mathbb{R}^m$ dans $\mathbb{R}$. On notera $\vec{x} = (x_1, \dots, x_m)$ les variables de f. Lorsque $m = 2$ on notera souvent $\vec{x} = (x, y)$ plutôt que (x_1, x_2) les variables.

Définition 5.1 (Dérivées partielles)**.** Supposons que A est un ouvert. On dit que f admet une dérivée partiel par rapport à la variable x_i, en le point $\vec{u} = (u_1, \dots, u_m)$ si la fonction d'une variable

$$x_i \mapsto f(u_1, \dots, u_{i-1}, x_i, u_{i+1}, \dots, u_m)$$

est dérivable au point u_i. Dans ce cas on notera

$$\partial_{x_i} f(\vec{u}) = \lim_{h \to 0} \frac{f(u_1, \dots, u_{i-1}, u_i + h, u_{i+1}, \dots, u_m) - f(u_1, \dots, u_{i-1}, u_i, u_{i+1}, \dots, u_m)}{h}.$$

Il existe donc m dérivées partielles. Comme dans le cas d'une variable on va les considérer comme des fonctions. En pratique, pour calculer une dérivée partielle, on traite toutes les variables sauf une comme des constantes et on dérive par rapport à la dernière variable.

Exemple 5.2. Soit $m = 2$ et $f(x, y) = \sin(x^2 + y)$. On a $\partial_x f(x, y) = 2x\cos(x^2 + y)$ et $\partial_y f(x, y) = \cos(x^2 + y)$ Les dérivées partielles sont définies comme f sur $\mathbb{R}^2$.

Exemple 5.3. Soit $m = 2$ et $f(x, y) = \sqrt{x - y}$. La fonction f est définie sur l'ensemble $\{(x, y) \in \mathbb{R}^2, x \geq y\}$. En revanche, les dérivées partielles $\partial_x f(x, y) = \frac{1}{2\sqrt{x-y}}$ et $\partial_y f(x, y) = \frac{-1}{2\sqrt{x-y}}$ sont définies sur l'ouvert $\{(x, y) \in \mathbb{R}^2, x > y\}$.

Définition 5.4. Supposons que A est un ouvert et que f admet une dérivée partiel par rapport à chaque variable x_i, en le point $\vec{u}$. On appellera gradient de f en $\vec{u}$ le vecteur colonne $\mathrm{grad} f(\vec{u}) := \begin{pmatrix} \partial_{x_1} f(\vec{u}) \\ \vdots \\ \partial_{x_m} f(\vec{u}) \end{pmatrix}$.

Pour une fonction d'une variable, le h du taux d'accroissement $\frac{f(x+h)-f(x)}{h}$ peut tendre vers 0 par deux chemins : par valeurs positives ou négatives. Dans le cas des fonctions de plusieurs variables, il y a une infinité de chemins. La notion de différentielle va encoder cela.

Définition 5.5 (Différentiabilité)**.** Supposons que A est un ouvert et soit $\|.\|$ une norme sur $\mathbb{R}^m$. On dit que f est différentiable en $\vec{u} = (u_1, \ldots, u_m) \in A$ si il existe une application linéaire $\ell : \mathbb{R}^m \to \mathbb{R}$ telle que

$$\lim_{\|\vec{h}\| \to 0} \frac{f(\vec{u} + \vec{h}) - f(\vec{u}) - \ell(\vec{h})}{\|\vec{h}\|} = 0.$$

L'application ℓ est la différentielle de f en $\vec{u}$ et se note $df_{\vec{u}}$.

Remarque 5.6. Rappelons, que puisqu'on est en dimension finie, par la proposition 4.51, les normes sont équivalentes. Ainsi une fonction sera différentiable pour une norme si et seulement si elle le sera pour n'importe laquelle. On peut donc prendre n'importe quelle norme. En générale on prendra la norme $\|.\|_\infty$.

Notons que $\vec{u}$ représente un vecteur de $\mathbb{R}^m$. Si $m = 1$, $df_{\vec{u}}$ n'est autre que la fonction linéaire $h_1 \mapsto f'(u_1)h_1$. Ainsi dans ce cas f est différentiable si et seulement si f est dérivable. L'expression $f(\vec{u}) + \ell(\vec{h})$ est alors celle de la tangente de f en u_1. Si $m = 2$, $f(\vec{u}) + \ell(\vec{h})$ représente l'équation du plan tangent de f en $\vec{u}$. En générale, les fonctions différentiables sont donc les fonctions admettant un développement limité à l'ordre 1. On donnera plus loin une expression plus explicite de ℓ. Comme pour le cas des fonctions à une variable, la différentiabilité entraîne la continuité.

Lemme 5.7. *Si f est différentiable en $\vec{u}$ alors elle est continue en $\vec{u}$.*

Démonstration. Soit $\varepsilon > 0$. On a l'existence de $\eta > 0$ tel que $\|\vec{h}\| < \eta$ entraîne $\left|\frac{f(\vec{u}+\vec{h})-f(\vec{u})-\ell(\vec{h})}{\|\vec{h}\|}\right| < \varepsilon$. On a donc $\left|f(\vec{u}+\vec{h}) - f(\vec{u}) - \ell(\vec{h})\right| < \varepsilon\|\vec{h}\|$. Grâce à l'inégalité triangulaire on trouve $\left|f(\vec{u}+\vec{h}) - f(\vec{u})\right| < |\ell(\vec{h})| + \varepsilon\|\vec{h}\|$. Puisque ℓ est linéaire, il existe $\ell_1, \ldots, \ell_m \in \mathbb{R}$ tel que $\ell(\vec{h}) = \ell_1 h_1 + \cdots + \ell_m h_m$ avec $\vec{h} = (h_1, \ldots, h_m)$. En prenant $L = \max(|\ell_1|, \ldots, |\ell_m|)$ on trouve que $|\ell(\vec{h})| + \varepsilon\|\vec{h}\| \leq (mL+\varepsilon)\|\vec{h}\| \leq (mL+\varepsilon)\eta$. En prenant $\eta > 0$ tel que $(mL+\varepsilon)\eta = \varepsilon$ on trouve bien que f est continue en $\vec{u}$. □

Donnons maintenant une expression explicite de la différentielle.

Proposition 5.8. *Si f est différentiable en $\vec{u}$, alors ses dérivées partielles existent en $\vec{u}$ et on a*

$$df_{\vec{u}}(\vec{h}) = h_1 \partial_{x_1} f(\vec{u}) + \cdots + h_m \partial_{x_m} f(\vec{u}).$$

Démonstration. Puisque f est différentiable, la fonction linéaire ℓ se représente de la forme $\ell(\vec{h}) = \ell_1 h_1 + \cdots + \ell_m h_m$. Il s'agit de prouver que les ℓ_i sont les $\partial_{x_i} f(\vec{u})$. Prenons $\vec{h}$ de la forme $(h_1, 0 \ldots, 0)$. On a $\|(h_1, 0 \ldots, 0)\|$ tend vers 0 implique que $|h_1|$ tend vers 0. On trouve donc

$$\lim_{|h_1| \to 0} \frac{f(u_1+h_1, u_2, \ldots, u_m) - f(u_1, u_2, \ldots, u_m) - \ell_1 h_1}{|h_1|} = 0.$$

Comme $\lim_{|h_1|\to 0} \frac{f(u_1+h_1,u_2,\ldots,u_m)-f(u_1,u_2,\ldots,u_m)}{|h_1|} = \partial_{x_1} f(\vec{u})$ on voit bien que $\ell_1 = \partial_{x_1} f(\vec{u})$. On procède de même avec les autres variables pour trouver le résultat. □

Si les dérivées partielles sont non définies, la fonction n'est pas différentiable. Si elles sont définies, on a un candidat pour la différentielle potentielle. Il suffit de vérifier que la fraction correspondante tend vers 0. Ce n'est pas toujours une chose simple. Pour cela on utilisera la notion suivante, encore plus forte que la différentiabilité, mais plus simple à vérifier.

Définition 5.9. Soit A un ouvert de $\mathbb{R}^m$. On dit que $f : A \to \mathbb{R}$ est $\mathcal{C}^1$ si ses dérivées partielles existent et sont continues sur A.

Bien entendu lorsque $m = 1$ on retrouve la notion de $\mathcal{C}^1$.

Proposition 5.10. *Si f est $\mathcal{C}^1$ alors elle est différentiable.*

Ainsi pour montrer qu'une fonction est différentiable, il suffit de montrer que ses dérivées partielles sont continues.

Exemple 5.11. La fonction $f(x, y) = x\exp(y)$ possède des dérivées partielles qui sont $\partial_x f(x, y) = \exp(y)$ et $\partial_y f(x, y) = x\exp(y)$. Puisqu'elles sont continues, f est $\mathcal{C}^1$ et donc différentiable.

Démonstration. Pour simplifier prouvons le résultat pour $m = 2$, la preuve à m quelconque est similaire. Supposons que f est $\mathcal{C}^1$ en $\vec{u} = (u_1, u_2)$. On a alors

$$f(u_1+h_1, u_2+h_2)-f(u_1, u_2) = f(u_1+h_1, u_2+h_2)-f(u_1+h_1, u_2)+f(u_1+h_1, u_2)-f(u_1, u_2).$$

Puisque f est $\mathcal{C}^1$, ses dérivées partielles existent et on trouve

$$f(u_1 + h_1, u_2) - f(u_1, u_2) = h_1\partial_{x_1} f(\vec{u}) + h_1\varepsilon_1(h_1),$$

$$f(u_1 + h_1, u_2 + h_2) - f(u_1 + h_1, u_2) = h_2\partial_{x_2} f(u_1 + h_1, u_2) + h_2\varepsilon_2(h_2),$$

avec $\varepsilon_i(h_i)$ qui tend vers 0 lorsque h_i tend vers 0. Comme f est $\mathcal{C}^1$ ses dérivées partielles sont continues et on a $\partial_{x_2} f(u_1 + h_1, u_2) - \partial_{x_2} f(u_1, u_2) = h_1\varepsilon_3(h_1)$ avec $\varepsilon_3(h_1)$ qui tend vers 0 lorsque h_1 tend vers 0. Ceci montre bien que

$$\lim_{\|\vec{h}\|\to 0} \frac{f(\vec{u} + \vec{h}) - f(\vec{u}) - \ell(\vec{h})}{\|\vec{h}\|_\infty} = 0$$

avec $\ell = h_1\partial_{x_1} f(\vec{u}) + h_2\partial_{x_2} f(\vec{u})$. □

Pour résumer on a donc prouvé les implication suivantes.

$f \in \mathcal{C}^1$	par définition ses dérivées partielles sont continues
$\Downarrow$	
f est différentiable	par définition f admet un développement limité d'ordre 1
$\Downarrow$	
$f$est continue	

5.2 Matrice Jacobienne

Soient $f_1, \ldots, f_n$ des fonctions de $A \subset \mathbb{R}^m$ dans $\mathbb{R}$ avec A ouvert. On considère f de A dans $\mathbb{R}^n$ qui a $\vec{x}$ associe $(f_1(\vec{x}), \ldots, f_n(\vec{x}))$. On dire que les f_i sont les composantes de f. On dira que f est différentiable, ou $\mathcal{C}^1$, si chacune de ses composantes le sont.

On suppose que pour $i \leq m$ et $j \leq n$ les dérivées partielles $\partial_{x_i} f_j$ existent.

Définition 5.12. On appelle matrice Jacobienne de f en $\vec{u} \in A$ la matrice

$$\mathrm{Jac}_f(\vec{u}) := \begin{pmatrix} \partial_{x_1} f_1(\vec{u}) & \ldots & \partial_{x_m} f_1(\vec{u}) \\ \vdots & & \vdots \\ \partial_{x_1} f_n(\vec{u}) & & \partial_{x_m} f_n(\vec{u}) \end{pmatrix}.$$

C'est donc une matrice à n lignes et m colonnes.

Exemple 5.13. Si $m = n = 2$ et $f_1(x,y) = x^2 + y^2$, $f_2(x,y) = \exp(xy)$ on a

$$\mathrm{Jac}_f(x,y) := \begin{pmatrix} 2x & 2y \\ y\exp(xy) & x\exp(xy) \end{pmatrix}.$$

Définition 5.14. De manière analogue à la dimension 1, on désignera par $o(\|\vec{h}\|)$ une fonction définie au voisinage de $(0,\ldots,0) \in \mathbb{R}^m$ et qui tend vers 0 lorsque $\|\vec{h}\|$ tend vers 0.

Remarque 5.15. Par la définition de la différentielle, on trouve que si $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$ est une fonction de classe $\mathcal{C}^1$ avec A un ouvert, alors pour tout $\vec{u} \in A$ et $\vec{h} \in \mathbb{R}^m$ proche de $(0,\ldots,0)$, on a

$$f(\vec{u} + \vec{h}) = f(\vec{u}) + \mathrm{J}_f(\vec{u})\vec{h} + o(\|\vec{h}\|).$$

Ceci généralise le formule de Taylor à l'ordre 1.

L'intérêt principal de la Jacobienne est qu'elle permet de calculer la dérivé d'une composé de fonctions de plusieurs variables. Rappelons que si f, g fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$, alors $f \circ g$ est dérivable et $(f \circ g)' = g' \times f' \circ g$. Voyons maintenant la généralisation en dimension supérieure. Puisque la matrice Jacobienne encode les dérivées partielles, on cherchera simplement à calculer cette dernière.

Proposition 5.16. *Soit f une fonction de $\mathbb{R}^p$ dans $\mathbb{R}^n$ et g une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^p$. Alors $f \circ g$ est une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Si f et g sont $\mathcal{C}^1$, alors $f \circ g$ l'est et*

$$\mathrm{Jac}_{f\circ g}(\vec{u}) = \mathrm{Jac}_f(g(\vec{u})) \times \mathrm{Jac}_g(\vec{u}).$$

Ici $\times$ représente la multiplication matricielle. Si $m = n = 1$, on retrouve la formule de la dérivé de la composition classique.

Démonstration. Fixons $i \leq m$ et $j \leq n$. On cherche à calculer la j eme composante de $f \circ g$. Cette dernière vaut $f_j(g_1(\vec{x}), \ldots, g_p(\vec{x}))$. Pour simplifier, supposons ici que $p = 2$. Pour $\vec{h} = (0, \ldots, h_i, \ldots, 0)$ on trouve que

$$\frac{f_j(g_1(\vec{x}+\vec{h}), g_2(\vec{x}+\vec{h})) - f_j(g_1(\vec{x}), g_2(\vec{x}))}{|h_i|}$$
$$= \frac{f_j(g_1(\vec{x}+\vec{h}), g_2(\vec{x}+\vec{h})) - f_j(g_1(\vec{x}+\vec{h}), g_2(\vec{x}))}{|h_i|} + \frac{f_j(g_1(\vec{x}+\vec{h}), g_2(\vec{x})) - f_j(g_1(\vec{x}), g_2(\vec{x}))}{|h_i|}.$$

On utilise alors la dérivée d'une composé de fonctions d'une variable pour déduire que $\frac{f_j(g_1(\vec{x}+\vec{h}), g_2(\vec{x})) - f_j(g_1(\vec{x}), g_2(\vec{x}))}{|h_i|}$ tend vers $\partial_{x_1} f_j(g_1(\vec{x}), g_2(\vec{x})) \times \partial_{x_i} g_1(\vec{x})$

lorsque h_i tend vers 0. Comme dans la preuve de la proposition 5.10 on utilise la continuité des dérivées partielles pour déduire que lorsque h_i tend vers 0, $\frac{f_j(g_1(\vec{x}+\vec{h}),g_2(\vec{x}+\vec{h}))-f_j(g_1(\vec{x}+\vec{h}),g_2(\vec{x}))}{|h_i|}$ tend vers $\partial_{x_2} f_j(g_1(\vec{x}), g_2(\vec{x})) \times \partial_{x_i} g_2(\vec{x})$ Plus généralement, lorsque p est quelconque on trouve que On trouve que

$$\partial_{x_i}(f_j(g_1(\vec{x}), \ldots, g_p(\vec{x}))) = \sum_{k=1}^{p} \partial_{x_k} f_j(g_1(\vec{x}), g_2(\vec{x})) \times \partial_{x_i} g_k(\vec{x}).$$

Ceci conclue la preuve. □

Remarque 5.17. Le théorème reste vrai si f est une fonction de $B \subset \mathbb{R}^p$ dans $\mathbb{R}^n$ et g une fonction de $A \subset \mathbb{R}^m$ dans B, avec A et B des ouverts.

Exemple 5.18. Considérons $f : \mathbb{R}^2 \setminus \{(0,0)\}$ dans $\mathbb{R}$ et $g :]0, \infty[\times \mathbb{R} \to \mathbb{R}^2 \setminus \{(0,0)\}$ qui a (r, θ) associe $(r\cos(\theta), r\sin(\theta))$. C'est le changement de variables en coordonnées polaires. La Jacobienne de g est $\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$, celle de $\mathrm{Jac}_f(g(r,\theta))$ est $\begin{pmatrix} \partial_x f(r\cos(\theta), r\sin(\theta)) & \partial_y f(r\cos(\theta), r\sin(\theta)) \end{pmatrix}$. Alors

$$\mathrm{Jac}_{f\circ g}(r,\theta) = \begin{pmatrix} \partial_x f(r\cos(\theta), r\sin(\theta)) & \partial_y f(r\cos(\theta), r\sin(\theta)) \end{pmatrix} \times \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

Cette matrice est par définition $\begin{pmatrix} \partial_r(f\circ g) \\ \partial_\theta(f\circ g) \end{pmatrix}$.

On trouve donc que $f \circ g = f(r\cos(\theta), r\sin(\theta))$ admet des dérivées partielles par rapport à r et θ et on trouve

$$\begin{aligned} \partial_r(f\circ g) &= \cos(\theta)\partial_x f(r\cos(\theta), r\sin(\theta)) + \sin(\theta)\partial_y f(r\cos(\theta), r\sin(\theta)) \\ \partial_\theta(f\circ g) &= -r\sin(\theta)\partial_x f(r\cos(\theta), r\sin(\theta)) + r\cos(\theta)\partial_y f(r\cos(\theta), r\sin(\theta)). \end{aligned}$$

Par exemple, si $f(x,y) = xy$, on trouve $f\circ g(r,\theta) = r\cos(\theta) r\sin(\theta) = \frac{r^2}{2}\sin(2\theta)$. C'est une fonction de r et θ que l'on peut dériver $\partial_r(f\circ g)(r,\theta) = r\sin(2\theta)$ et $\partial_\theta(f\circ g)(r,\theta) = r^2\cos(2\theta)$. En utilisant les formules ci dessus, $\partial_x f(x,y) = y$, $\partial_y f(x,y) = x$ et on trouve bien

$$\begin{aligned} \partial_r(f\circ g)(r,\theta) &= \cos(\theta) r\sin(\theta) + \sin(\theta) r\cos(\theta) = r\sin(2\theta) \\ \partial_\theta(f\circ g)(r,\theta) &= -r\sin(\theta) r\sin(\theta) + r\cos(\theta) r\cos(\theta) = r^2\cos(2\theta). \end{aligned}$$

5.3 Théorèmes des accroissements finis

Le théorème des accroissements finis en dimension 1 s'énonce ainsi. Soit $f : [a,b] \to \mathbb{R}$ une fonction continue sur $[a,b]$, dérivable sur $]a,b[$. Il existe $a < c < b$

tel que $f'(c)(b-a) = f(b) - f(a)$. Écrivons maintenant un résultat en dimension supérieure avec une hypothèse légèrement plus contraignante sur f.

Si $\vec{a}, \vec{b} \in \mathbb{R}^m$, on définit $[\vec{a}, \vec{b}]$ comme étant le segment d'extrémité $\vec{a}$ et $\vec{b}$. Autrement dit

$$[\vec{a}, \vec{b}] = \{\vec{a} + \theta(\vec{b} - \vec{a}), \theta \in [0,1]\}.$$

Rappelons que si A est un ouvert de $\mathbb{R}^n$ et que $f : A \to \mathbb{R}$ est $\mathcal{C}^1$, alors on définit son gradient par $\mathrm{grad} f(\vec{u}) := \begin{pmatrix} \partial_{x_1} f(\vec{u}) \\ \vdots \\ \partial_{x_m} f(\vec{u}) \end{pmatrix}$. Enfin, si $\vec{u} = (u_1, \ldots, u_m)$ et $\vec{v} = (v_1, \ldots, v_m)$ sont des vecteurs de $\mathbb{R}^m$, on définit leurs produit scalaire par

$$\langle \vec{u} | \vec{v} \rangle := \sum_{i=1}^{m} u_i v_i.$$

Théorème 5.19 (Théorème des accroissement finis)**.** *Soit A un ouvert de $\mathbb{R}^m$ et soit $f : A \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur A. Soient $\vec{a}, \vec{b} \in A$ et supposons que le segment $[\vec{a}, \vec{b}]$ soit inclue dans A. Alors il existe $\vec{c} \in [\vec{a}, \vec{b}]$ tel que*

$$\langle \mathrm{grad} f(\vec{c}) | \vec{b} - \vec{a} \rangle = f(\vec{b}) - f(\vec{a}).$$

D'après la proposition 5.8, si f est différentiable, alors $df_{\vec{u}}(\vec{h}) = \langle \mathrm{grad} f(\vec{u}) | \vec{h} \rangle$. On trouve donc $df_{\vec{c}}(\vec{b} - \vec{a}) = f(\vec{b}) - f(\vec{a})$.

Démonstration. Regardons la fonction d'une variable de $[0,1] \subset \mathbb{R}$ dans $\mathbb{R}$ définie par $g(\theta) = f(\vec{a} + \theta(\vec{b} - \vec{a}))$. On peut écrire sous forme d'une composé $g = f \circ \gamma$ avec $\gamma(\theta) = \vec{a} + \theta(\vec{b} - \vec{a})$ qui est une fonction $\mathcal{C}^1$ de $[0,1] \subset \mathbb{R}$ dans $\mathbb{R}^n$. D'après le théorème des accroissement finie à une variable, il existe $\theta_0 \in]0,1[$ tel que $g'(\theta_0)(1-0) = g(1) - g(0)$. Autrement dit $g'(\theta_0) = f(\vec{b}) - f(\vec{a})$. Il reste maintenant à calculer la dériver de g ou autrement dit sa Jacobienne. Par la proposition 5.16, on peut calculer la Jacobienne de g par la formule suivante :

$$\mathrm{Jac}_g(\theta) = \mathrm{Jac}_f(\gamma(\theta)) \times \mathrm{Jac}_\gamma(\theta).$$

La fonction γ est à valeur de $[0,1] \subset \mathbb{R}$ dans $\mathbb{R}^n$. En dérivant coordonnées par coordonnées, on trouve que $\mathrm{Jac}_\gamma(\theta) = \begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_m - a_m \end{pmatrix}$, où l'on a noté a_i et b_i les coordonnées de $\vec{a}$ et $\vec{b}$. Par définition, on a $\mathrm{Jac}_f(\gamma(\theta)) = (\partial_{x_1} f(\gamma(\theta)), \ldots, \partial_{x_m} f(\gamma(\theta)))$. On trouve donc que $g'(\theta) = \langle \mathrm{grad} f(\gamma(\theta)) | \vec{b} - \vec{a} \rangle$. En prenant $\theta = \theta_0$ on trouve

$$f(\vec{b}) - f(\vec{a}) = \langle \mathrm{grad} f(\gamma(\theta_0)) | \vec{b} - \vec{a} \rangle.$$

En posant $\vec{c} = \gamma(\theta_0) \in [\vec{a}, \vec{b}]$ on trouve le résultat. $\square$

On trouve quelques corollaire bien utiles.

Corollaire 5.20 (Inégalité des accroissement finis). *Soit A un ouvert de $\mathbb{R}^m$ et soit $f : A \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur A. Soient $\vec{a}, \vec{b} \in A$ et supposons que le segment $[\vec{a}, \vec{b}]$ soit inclue dans A. Si pour tout $\vec{c} \in [\vec{a}, \vec{b}]$, $\|\mathrm{grad} f(\vec{c})\|_2 \leq k$, alors*

$$|f(\vec{b}) - f(\vec{a})| \leq k\|\vec{b} - \vec{a}\|_2.$$

Démonstration. Soit $\vec{c} \in [\vec{a}, \vec{b}]$ tel que $\langle \mathrm{grad} f(\vec{c}) | \vec{b} - \vec{a} \rangle = f(\vec{b}) - f(\vec{a})$. Par l'inégalité de Cauchy Schwartz, on a

$$|f(\vec{b}) - f(\vec{a})| = |\langle \mathrm{grad} f(\vec{c}) | \vec{b} - \vec{a} \rangle| \leq \|\mathrm{grad} f(\vec{c})\|_2 \|\vec{b} - \vec{a}\|_2 \leq k\|\vec{b} - \vec{a}\|_2.$$

On trouve le résultat souhaité. □

Corollaire 5.21 (Inégalité des accroissement finis). *Soit A un ouvert de $\mathbb{R}^m$ et soit $f : A \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur A. Soient $\vec{a}, \vec{b} \in A$ et supposons que le segment $[\vec{a}, \vec{b}]$ soit inclue dans A. Si pour tout $\vec{c} \in [\vec{a}, \vec{b}]$, $\mathrm{grad} f(\vec{c})$ est le vecteur nul, alors f est constante sur $[\vec{a}, \vec{b}]$.*

Corollaire 5.22. *Si $\mathrm{grad} f(\vec{c})$ est le vecteur nul alors $\|\mathrm{grad} f(\vec{c})\|_2 = 0$. Par l'inégalité des accroissement finis on trouve $|f(\vec{b}) - f(\vec{a})| = 0$, c'est à dire $f(\vec{b}) = f(\vec{a})$. Il suffit de remplacer $\vec{b}$ par n'importe quel élément de $[\vec{a}, \vec{b}]$ pour déduire que f est constante. sur $[\vec{a}, \vec{b}]$.*

5.4 Fonctions de classe $\mathcal{C}^2$

Considérons A un ouvert de $\mathbb{R}^m$ et $f : A \to \mathbb{R}$ de classe $\mathcal{C}^1$. Ses dérivées partielles $\partial_{x_i} f$ sont des fonctions continues de A dans $\mathbb{R}$. Si elles sont différentiables, on peut donc les dériver. On arrive naturellement à la définition suivante.

Définition 5.23. On dit que $f : A \to \mathbb{R}$ est de classe $\mathcal{C}^2$ si elle est de classe $\mathcal{C}^1$ et ses dérivées partielles sont de classe $\mathcal{C}^1$.

Exemple 5.24. Définissons $f : \mathbb{R}^2 \to \mathbb{R}$ par $f(x, y) = x^2 \cos(y) + \exp(x - y)$. Les dérivées partielles sont $\partial_x f(x, y) = 2x \cos(y) + \exp(x - y)$ et $\partial_y f(x, y) = -x^2 \sin(y) - \exp(x - y)$. Ces dérivées partielles sont dérivables, et on trouve quatre dérivées partielles d'ordre 2, puisqu'on peut dériver chaque dérivée partielle par rapport à deux variables.On trouve donc $\partial_x(\partial_x f(x, y)) = 2\cos(y) + \exp(x - y)$ et $\partial_y(\partial_x f(x, y)) = -2x \sin(y) - \exp(x - y)$. D'un autre côté on a $\partial_x(\partial_y f(x, y)) = -2x \sin(y) - \exp(x - y)$ et $\partial_y(\partial_x f(x, y)) = -x^2 \cos(y) + \exp(x - y)$. Ces dérivées partielles d'ordre 2 étant continues, on trouve que f est $\mathcal{C}^2$. En réalité, on pourrait prouver qu'elle est $\mathcal{C}^\infty$ si on prenait la peine de définir les espaces $\mathcal{C}^3, \dots$. Remarquons que l'on a $\partial_y(\partial_x f(x,y)) = \partial_x(\partial_y f(x,y))$. On verra que dans un cadre large, cette égalité reste vrai : on peut en général dériver dans l'ordre que l'on veut, et il n'y a pas 4 mais 3 dérivées partielles d'ordre 2 à regarder.

Dans ce qui suit, si $A \subset \mathbb{R}^m$ est un ouvert et si $f : A \to \mathbb{R}$ est de classe $\mathcal{C}^2$ on notera

$$\partial_{x_i}^2 f = \partial_{x_i}(\partial_{x_i} f), \quad \partial_{x_i}\partial_{x_j} f = \partial_{x_i}(\partial_{x_j} f)$$

les dérivées partielles d'ordre deux. Il y en a à priori m^2. Mais en réalité, certains sont égales, c'est une conséquence du théorème de Schwarz dont la preuve sera admise.

Théorème 5.25 (Théorème de Schwarz)**.** *Soit $A \subset \mathbb{R}^m$ un ouvert et soit $f : A \to \mathbb{R}$ est de classe $\mathcal{C}^2$. Alors pour tout $i \neq j$, on a*

$$\partial_{x_i}\partial_{x_j} f = \partial_{x_j}\partial_{x_i} f.$$

La matrice Jacobienne encodait les dérivées partielles d'ordre 1, la matrice Hessienne va encoder les dérivées partielles d'ordre 2.

Définition 5.26. Soit $f : A \subset \mathbb{R}^m \to \mathbb{R}$ une matrice de classe $\mathcal{C}^2$. Alors la matrice Hessienne est définie par

$$\mathrm{H}_f(\vec{u}) = \Big(\partial_{x_i}\partial_{x_j} f(\vec{u})\Big)_{i,j \leq m}.$$

Remarquons que d'après le théorème de Schwarz la matrice Hessienne est une matrice symétrique.

Exemple 5.27. Soit $m = 2$ et $f(x,y) = \sin(x)\exp(y)$. Alors la matrice Hessienne vaut

$$\begin{pmatrix} -\sin(x)\exp(y) & \cos(x)\exp(y) \\ \cos(x)\exp(y) & \sin(x)\exp(y) \end{pmatrix}.$$

L'intérêt de la matrice Hessienne est d'écrire un développement limité à l'ordre 2. Ceci sera crucial pour la recherche d'extremum qui sera l'objectif final de ce cours.

Théorème 5.28. *Soit $f : A \subset \mathbb{R}^m \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ avec A un ouvert. Alors pour tout $\vec{u} \in A$ et $\vec{h} \in \mathbb{R}^m$ petit, on a*

$$f(\vec{u} + \vec{h}) = f(\vec{u}) + \mathrm{J}_f(\vec{u})\vec{h} + \frac{1}{2}(\vec{h})^t \mathrm{H}_f(\vec{u})\vec{h} + o(\|\vec{h}\|^2).$$

Ici $\|.\|$ désigne la norme 2 mais on pourrait prendre n'importe quelle norme de $\mathbb{R}^m$ car elles sont toutes équivalentes. Comme pour le développement limité classique on approxime une fonction suffisamment régulière par un polynômes de degrés 2 en m variables $h\vec{h}$.

Exemple 5.29. Si $m = 1$ on a $\mathrm{J}_f(u) = f'(u)$ et $\mathrm{H}_f(u) = f''(u)$. On trouve donc la formule de Taylor à l'ordre 2 :

$$f(u+h) = f(u) + f'(u)h + \frac{1}{2}f''(u)h^2 + o(h^2).$$

Exemple 5.30. Si $m = 2$, on a

$$f(u_1+h_1, u_2+h_2) = f(u_1,u_2) + \partial_{x_1} f(\vec{u})h_1 + \partial_{x_2} f(\vec{u})h_2 + \\ \frac{1}{2}\left(\partial^2_{x_1} f(\vec{u})h_1^2 + 2\partial_{x_1}\partial_{x_2} f(\vec{u})h_1h_2 + \partial^2_{x_2} f(\vec{u})h_2^2\right) + o(\|\vec{h}\|^2). \quad (5.1)$$

Démonstration. Considérons la fonction qui a $\vec{h}$ associe $g(\vec{h}) = f(\vec{u}+\vec{h}) - f(\vec{u}) - \mathrm{J}_f(\vec{u})\vec{h} - \frac{1}{2}(\vec{h})^t\mathrm{H}_f(\vec{u})\vec{h}$. Elle vaut 0 en $\vec{h} = \vec{0}$. Étudions les dérivées partielles d'ordre 1. Le terme $f(\vec{u})$ est constant, sa dérivée est nulle. D'après le théorème de dérivation des fonction composés, on a $\partial_{h_i}(f(\vec{u}+\vec{h})) = \partial_{x_i} f(\vec{u}+\vec{h})$. Donc la dérivée partielle en $\vec{h} = \vec{0}$ est $\partial_{x_i} f(\vec{u})$. Les dérivées partielles de $\mathrm{J}_f(\vec{u})\vec{h}$ sont les coefficients de la Jacobienne. Donc $\partial_{h_i}\mathrm{J}_f(\vec{u})\vec{h} = \partial_{x_i} f(\vec{u})$. Le terme $(\vec{h})^t\mathrm{H}_f(\vec{u})\vec{h}$ ne contient que des termes en h_ih_j et $(h_i)^2$ donc ses dérivées partielles d'ordre 1 sont nuls en $\vec{h} = \vec{0}$. Ainsi les dérivées partielle d'ordre 1 de $g(\vec{h})$ sont nuls en $\vec{h} = \vec{0}$. De même, les dérivées partielle d'ordre 2 (en $\vec{h}$ de $\mathrm{J}_f(\vec{u})\vec{h}$ sont nuls, et on a $\partial_{h_i}\partial_{h_j}(\vec{h})^t\mathrm{H}_f(\vec{u})\vec{h} = \partial_{x_i}\partial_{x_j} f(\vec{u})$. On trouve que les dérivées d'ordre 2 de $g(\vec{h})$ sont nuls en $\vec{h} = \vec{0}$.

Par l'absurde, supposons que $g(\vec{h})\|\vec{h}\|^{-2}$ ne tend pas vers 0. On a l'existence d'une suite $\vec{h_n}$ qui tend vers $\vec{0}$ et de $\varepsilon > 0$ tel que $g(\vec{h_n}) > \varepsilon\|\vec{h_n}\|^2$. Ceci contredit le fait que les dérivées partielles d'ordre 2 sont nuls en $\vec{0}$. □

5.5 Recherche d'extremums

Quand on étudie une fonction d'une variable réelle, on cherche à déterminer le tableau de variation. En d'autres termes on cherche à savoir si des réelles correspondent à des valeurs minimum (maximum) de la fonction. Formalisons cela pour les fonctions de plusieurs variables.

Définition 5.31. Soit $f : A \to \mathbb{R}$ une fonction avec $A \subset \mathbb{R}^m$ un ouvert.

On dit que $\vec{x} \in A$ est un minimum local si il existe un ouvert U, avec $\vec{x} \in U \subset A$ tel que pour tout $\vec{y} \in U$, on a $f(\vec{y}) \geq f(\vec{x})$.

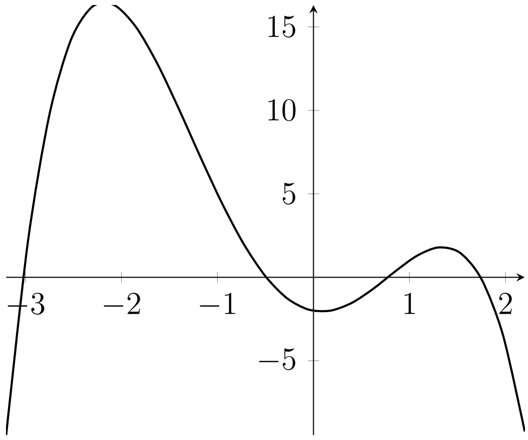
On dit que $\vec{x} \in A$ est un minimum global si tout $\vec{y} \in A$, on a $f(\vec{y}) \geq f(\vec{x})$.

On dit que $\vec{x} \in A$ est un maximum local si il existe un ouvert U, avec $\vec{x} \in U \subset A$ tel que pour tout $\vec{y} \in U$, on a $f(\vec{y}) \leq f(\vec{x})$.

On dit que $\vec{x} \in A$ est un maximum global si tout $\vec{y} \in A$, on a $f(\vec{y}) \leq f(\vec{x})$.

Un extremum local est un élément qui est soit un minimum local, soit un maximum locale. Un extremum global est un élément qui est soit un minimum global, soit un maximum global.

Exemple 5.32. Pour la courbe $y = -x^4 - x^3 + 6x^2 - x - 2$ dont le graphe est racé ci dessous on a un maximum global, un autre local, et un minimum local.



Exemple 5.33. Considérons la fonction $\sin(x)$ définie sur $\mathbb{R}$. Le réel $\pi/2$ est un maximum local et même global, le point $-\pi/2$ est un minimum local, et même global.

Pour les fonctions d'une variable, une condition nécessaire pour avoir un extremum local est que la dérivée s'annule. Ce n'est pas une condition suffisante, comme le montre la fonction $f(x) = x^3$, dont la dérivée s'annule en 0 mais qui n'est pas un extremum. Plus précisément, si $f'(x_0) = 0$ et f est $\mathcal{C}^2$ la formule de Taylor donne

$$f(x) = f(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + o(x^2).$$

Si $f''(x_0) > 0$, alors au voisinage de x_0 on a $f(x) \geq f(x_0)$. On a un minimum local, comme pour la fonction $f(x) = x^2$ en $x_0 = 0$. Si $f''(x_0) < 0$, alors au voisinage de x_0 on a $f(x) \leq f(x_0)$. On a un maximum local, comme pour la fonction $f(x) = -x^2$ en $x_0 = 0$. Si $f''(x_0) = 0$, on ne peut à priori pas conclure. Par exemple $f(x) = x^4$ admet un minimum local en 0, mais $f(x) = x^3$ n'a pas d'extremum en 0, alors que leur dérivés secondes s'annulent en 0. Voire la figure 1.

On va voir que le même phénomène arrive dans le cas des fonctions de plusieurs variables.

Définition 5.34. Soit $f : A \to \mathbb{R}$ une fonction avec $A \subset \mathbb{R}^m$ un ouvert. Supposons que f soit $\mathcal{C}^2$.

On dit que $\vec{x} \in A$ est un point critique si

$$\partial_{x_1} f(\vec{x}) = \cdots = \partial_{x_m} f(\vec{x}) = 0.$$

Exemple 5.35. Considérons $f(x, y) = x^2 + y^2$. Ses dérivées partielles valent $\partial_x f(x, y) = 2x$ et $\partial_y f(x, y) = 2y$. Le seul point critique est $(0, 0)$. D'un autre

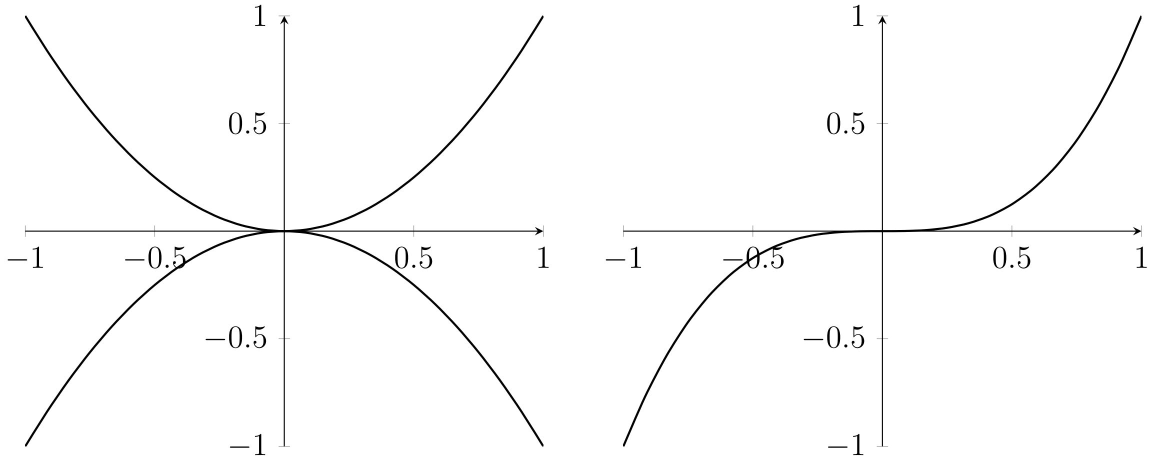


FIGURE 1 – A gauche, tracé des fonction $y = x^2$ (cas de la dérivée seconde strictement positive) et $y = -x^2$ (cas de la dérivée seconde strictement négative). A droite tracé de la fonction $y = x^3$ (cas où la dérivée seconde est nulle, il y a ici un point d'inflexion en 0).

côté, on voit que $(0,0)$ est un minimum local (et même global) car $f(0,0) = 0$ et $0 \leq f(x,y)$. Cela est en cohérence avec le résultat suivant.

Proposition 5.36. *Soit $f : A \to \mathbb{R}$ une fonction avec $A \subset \mathbb{R}^m$ un ouvert. Si f admet un extremum local en $\vec{x} \in A$, alors $\vec{x}$ est un point critique.*

Démonstration. Fixons toutes les variables sauf une et considérons la fonction d'une variable $h_i \mapsto f(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_m)$. Par définition elle admet un extremum local en $h_i = x_i$. Donc $\partial_{x_i} f(\vec{x}) = 0$. □

Définition 5.37. Un point selle est un point critique qui n'est pas un extremum local.

Par exemple, 0 est un point selle de $f(x) = x^3$.

Exemple 5.38. La fonction $f(x,y) = x^2 - y^2$ admet comme dérivées partielles $2x$ et $-2y$ et donc comme unique point critique le point $(0,0)$ où elle vaut 0. Puisque $f(x,0) > 0$ lorsque $x \neq 0$ et $f(0,y) < 0$ lorsque $y \neq 0$, c'est un point selle.

Limitons nous maintenant à $m = 2$. Soit donc $f : A \to \mathbb{R}$ une fonction avec $A \subset \mathbb{R}^2$ un ouvert. Supposons que f soit $\mathcal{C}^2$. On recherche les extremum de f. La première étape consiste à chercher les points critiques. Pour chaque point critique on doit déterminer si c'est un maximum local, minimum local ou point selle. Pour cela, on peut utiliser le critère de Monge.

Théorème 5.39 (Critères de Monge)**.** *Soit $f : A \to \mathbb{R}$ une fonction avec $A \subset \mathbb{R}^m$ un ouvert. Supposons que f soit $\mathcal{C}^2$. Soit (x_0, y_0) un point critique et soient*

$$r = \partial_x\partial_x f(x_0, y_0), \quad s = \partial_x\partial_y f(x_0, y_0), \quad t = \partial_y\partial_y f(x_0, y_0).$$

- *Si $r > 0$ et $rt - s^2 > 0$ alors (x_0, y_0) est un minimum local.*
- *Si $r < 0$ et $rt - s^2 > 0$ alors (x_0, y_0) est un maximum local.*
- *Se $rt - s^2 < 0$, alors (x_0, y_0) est un point selle.*

Démonstration. Puisque (x_0, y_0) est un point critique, les coefficients de la matrices Jacobienne $\mathrm{J}_f(x_0, y_0)$ sont nuls. Le développement limité de f à l'ordre 2 donne donc

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{1}{2}(h, k)\mathrm{H}_f(x_0, y_0)\begin{pmatrix} h \\ k \end{pmatrix} + o(\|(h, k)\|^2).$$

D'après les définitions de r, s, t, nous avons

$$\mathrm{H}_f(x_0, y_0) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}.$$

Si on veut prouver qu'un point est extremum, il suffit donc de prouver que $(h, k)\mathrm{H}_f(x_0, y_0)\begin{pmatrix} h \\ k \end{pmatrix}$ et suffisamment positive (ou négative) au voisinage de $(0, 0)$. C'est une forme quadratique, et on peut l'écrire comme somme de carrées. Sans refaire la théorie des formes quadratiques, on va le faire à la main dans ce cas. On a donc $(h, k)\mathrm{H}_f(x_0, y_0)\begin{pmatrix} h \\ k \end{pmatrix} = h^2r + 2hks + k^2t$.

Supposons que $r > 0$. On a

$$h^2r + 2hks + k^2t = \left(h\sqrt{r} + k\frac{s}{\sqrt{r}}\right)^2 + k^2\left(t - \frac{s^2}{r}\right).$$

Si $t - \frac{s^2}{r} > 0$ (ou encore $rt - s^2 > 0$), on a une somme de deux carrés qui est donc positive. Prenons pour $\|.\|$ la norme 2 (rappelons que d'après la proposition 4.51 en dimension finie les normes sont équivalentes). On a donc $\|(h, k)\|^2 = h^2 + k^2$. Puisque par définition $\|(h, k)\|^{-2}o(\|(h, k)\|^2)$ tend vers 0 lorsque (h, k) tend vers $(0, 0)$, on veut donc étudier $g(h, k) = (h^2r + 2hks + k^2t)\|(h, k)\|^{-2} = \frac{\left(h\sqrt{r}+k\frac{s}{\sqrt{r}}\right)^2 + k^2\left(t - \frac{s^2}{r}\right)}{h^2+k^2}$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Montrons qu'il existe $c > 0$ tel que $\frac{\left(h\sqrt{r}+k\frac{s}{\sqrt{r}}\right)^2 + k^2\left(t - \frac{s^2}{r}\right)}{h^2+k^2} > c$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Avec la formule du développement limité à l'ordre 2

$$\frac{f(x_0 + h, y_0 + k)}{\|(h, k)\|^2} = \frac{f(x_0, y_0)}{\|(h, k)\|^2} + \frac{1}{2\|(h, k)\|^2}(h, k)\mathrm{H}_f(x_0, y_0)\begin{pmatrix} h \\ k \end{pmatrix} + \frac{o(\|(h, k)\|^2)}{\|(h, k)\|^2}$$

on trouvera que (x_0, y_0) est un minimum local.

Notons que $g(h,k) = g(\lambda h, \lambda k)$ pour tout $\lambda \in]0,\infty[$. Donc $\inf_{(h,k)\in\mathbb{R}^2\setminus\{(0,0)\}} g(h,k) = \inf_{S(0,1)} g(h,k)$, où $S(0,1) = \{(h,k)\in\mathbb{R}^2 | \|(h,k)\| = 1\}$ est la sphère unité. Puisque la sphère unité est un compacte (voire la preuve de la proposition 4.51), on trouve que la fonction continue $g(h,k)$ atteint son min sur $S(0,1)$. Comme $g(h,k)$ est à valeur strictement positive, ce min est strictement positif. On a donc montré l'existence de $c > 0$ tel que $\frac{\left(h\sqrt{r}+k\frac{s}{\sqrt{r}}\right)^2+k^2\left(t-\frac{s^2}{r}\right)}{h^2+k^2} > c$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$. Comme expliqué plus haut, par la formule du développement limité à l'ordre 2 on trouve que (x_0, y_0) est un minimum local.

Si $rt - s^2 < 0$ on a un carré moins un autre. En particulier $(h^2r + 2hks + k^2t)\|(h,k)\|^{-2}$ prend des valeurs positives et négatives pour des valeurs arbitrairement proches de $(0,0)$. En raisonnant comme précédemment, on en déduit qu'on a un point selle. Si $rt - s^2 = 0$ on ne peut pas tout de suite conclure, il faudra affiner l'étude.

Si $r < 0$, on peut refaire le raisonnement avec l'opposé de $(h^2r + 2hks + k^2t)$. La condition $rt - s^2 > 0$ donnera l'opposé de deux carré donc (x_0, y_0) est un maximum local. De même, si $rt - s^2 < 0$ on trouve un point selle.

Si $r = 0$ on voit que puisque r et t jouent des rôles symétriques, on peut regarder les cas $t > 0$ et $t < 0$. Dans tous les cas on a $rt - s^2 < 0$ et on en déduit qu'on a un point selle. Si $r = t = 0$, alors $2hks$ prend des valeurs positives et négatives et on a un point selle. □

Remarque 5.40. — Nous n'avons pas couvert tous les sous cas. Par exemple si $rt - s^2 = 0$, on ne peut pas conclure et il faut affiner l'étude. Dans ce cas tout peut encore se produire.
— On pourrait faire la même chose avec des fonctions de m variables. On aurait alors des sommes et différences de au plus m carrées.
— Remarquons que la quantité $rt - s^2$ n'est autre que le determinant de la Hessienne. Une caractérisations des extremum locaux/ point selle peut aussi se faire en fonction des valeurs propres.

Exemple 5.41. Considérons $f(x,y) = x^2 + y^2$. Les dérivées partielles sont $\partial_x f = 2x$ et $\partial_y f = 2y$, et donc le seul point critique vérifie $2x = 2y = 0$, c'est donc $(0,0)$. On a $\partial_x^2 f = 2$, $\partial_y^2 f = 2$, $\partial x \partial_y f = \partial y \partial_x f = 0$ et la fonction est bien $\mathcal{C}^2$. Donc en prenant les notation de Monge, on a $r = t = 2$ etr $s = 0$. Donc $rt - s^2 = 4$ et $r > 0$ on a donc un minimum locale.

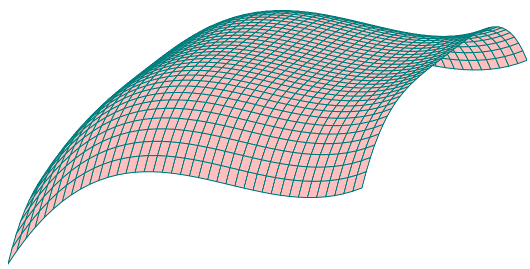
Exemple 5.42. Considérons $f(x,y) = x^2 - y^2$. En reprenant les calculs précédent on trouve $(0,0)$ comme seul point critique, $r = 2$, $s = 0$ et $t = -2$. Donc $rt - s^2 < 0$ et on trouve que $(0,0)$ est un point selle.

Les deux derniers exemples pouvaient se traiter à la main. Regardons maintenant deux exemples plus compliqués.

Exemple 5.43. Considérons $f(x, y) = x^3 + y^3 - 3xy$. Les dérivées partielles sont $\partial f_x(x, y) = 3x^2 - 3y$ et $\partial f_y(x, y) = 3y^2 - 3x$. Recherchons les points critiques. Pour cela il faut résoudre $\begin{cases} 0 = 3x^2 - 3y \\ 0 = 3y^2 - 3x. \end{cases}$ C'est un système non linéaire mais en injectant $y = x^3$ (première équation) dans la seconde, on trouve, $0 = 3x^6 - 3x$. En factorisant il vient $0 = 3x(x^5 - 1)$. Les seuls solutions réelles sont $x = 0$ et $x = 1$ (les autres racines 5ème de l'unité sont non réelles). Puisque $y = x^3$ on trouve que $x = 0$ et $x = 1$ correspondent à $y = 0$ et $y = 1$. Les deux points critiques sont donc $(0, 0)$ et $(1, 1)$. Étudions maintenant les dérivées partielles d'ordre deux. On a $\partial_x^2 f(x, y) = 6x$, $\partial_y^2 f(x, y) = 6y$, $\partial x \partial_y f(x, y) = \partial y \partial_x f(x, y) = -3$ et la fonction est bien $\mathcal{C}^2$. Pour le point critique $(0, 0)$ on a $r = t = 0$ et $s = -3$. Donc $rt - s^2 < 0$ et $(0, 0)$ est un point selle. En $(1, 1)$ on a $r = t = 6$ et $s = -3$. Donc $rt - s^2 = 36 - 9 = 27 > 0$ et $(1, 1)$ est un minimum local.

Finissons enfin par un exemple ou le critère de Monge ne s'applique pas.

Exemple 5.44. Considérons $f(x, y) = 2x^3 - y^4 - 3x^2$.



Les dérivées partielles sont $\partial f_x(x, y) = 6x^2 - 6x$ et $\partial f_y(x, y) = -4y^3$. Recherchons les points critiques. Pour cela il faut résoudre $\begin{cases} 0 = 6x^2 - 6x \\ 0 = -4y^3. \end{cases}$ On trouve donc $(0, 0)$ et $(1, 0)$. Calculons les dérivées partielles d'ordre deux. On a $\partial_x^2 f(x, y) = 12x - 6$, $\partial_y^2 f(x, y) = -12y^2$, $\partial x \partial_y f(x, y) = \partial y \partial_x f(x, y) = 0$ et la fonction est bien $\mathcal{C}^2$. Pour le point critique $(0, 0)$ on a $r = -6$ et $s = t = 0$. On est dans un cas où le critère de Monge ne permet pas de conclure. La fonction $g(x) = 2x^3 - 3x^2$ admet un maximum local en 0 car sa dérivée est nulle en 0 et sa dérivée seconde $12x - 6$ y est négative. De même, $h(y) = -y^4$ admet aussi un maximum local en 0. Donc pour (x, y) proche de $(0, 0)$ on a $f(x, y) = g(x) + h(y) \leq 0$. Donc le point critique $(0, 0)$ est un maximum local.

Regardons le point critique $(1, 0)$. On a $r = 6$, $s = t = 0$. On est encore dans un cas où $rt - s^2 = 0$ et le critère de Monge ne s'applique pas. On a $f(1, 0) = -1$.

On peut encore écrire $f(x, y) = g(x) + h(y)$. Ici puisque $g'(1) = 0$ et $g''(1) = 6 > 0$ on a que 1 est un minimum local de g. Puisque 0 est un maximum local de h, on s'attend à avoir un point selle pour $f(x, y) = g(x) + h(y)$ en $(1, 0)$. En effet pour y proche de 0 (mais différent de 0), on a $f(1, y) < f(1, 0) = -1$ et pour x proche de 1 (mais différent de 1), on a $f(x, 0) > f(1, 0) = -1$.