Calcul Intégral et Différentiel

J. Melleray

Université Lyon I
Semestre de printemps 2013-2014

Avertissement. Ces notes comportent certainement leur lot d'erreurs plus ou moins graves. Le cours ne suivra pas exactement les notes, qui sont suceptibles d'évoluer au cours du semestre. Remarques, commentaires, questions, corrections, etc. sont les bienvenus.

Les mots qui apparaissent dans un cadre rouge sur le fichier pdf sont cliquables (par exemple, dans la table des matières, pour passer à la section correspondante ; ou dans l'index, pour retrouver les endroits où apparaît un terme particulier).

Table des matières

Chapitre 1

Intégrale des fonctions continues par morceaux

1.1 Rappels

Définition 1.1. Soit I un intervalle de $\mathbb{R}$. Une fonction $f\colon I \to \mathbb{C}$ est *continue en* $x \in I$ si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in I\ |x-y| \leq \delta \Rightarrow |f(y)-f(x)| \leq \varepsilon\ .$$

On dit que f est *continue sur* I si f est continue en x pour tout $x \in I$.

Remarque 1.2. Cette définition s'applique bien sûr également pour les fonctions à valeurs réelles, qui sont un cas particulier de fonctions à valeurs dans $\mathbb{C}$. Dans ces notes, à chaque fois qu'il sera écrit « Soit $f\colon I \to \mathbb{C}$ », il faut penser qu'on parle d'une fonction à valeurs réelles ou complexes.

Attention, la notion de continuité est *locale* : elle dépend du point x où l'on se place. En particulier, dans la définition ci-dessus, δ dépend à la fois de x et de ε. Quand on peut choisir δ ne dépendant que de ε, on parle d'*uniforme continuité*.

Définition 1.3. Soit I un intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{C}$. On dit que f est *uniformément continue* sur I si

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in I\ |x-y| \leq \delta \Rightarrow |f(y)-f(x)| \leq \varepsilon\ .$$

En général, une fonction continue sur un intervalle I n'est pas uniformément continue, comme le montre l'exercice suivant.

Exercice 1.4. Soit $f\colon]0,+\infty[\to \mathbb{R}$ définie par $f(x) = \frac{1}{x}$. Montrer que f est continue mais pas uniformément continue. Montrer qu'il en va de même de la fonction $x \mapsto x^2$, définie sur $\mathbb{R}$ tout entier [i].

Néanmoins, il existe un cas très important où la continuité est équivalente à l'uniforme continuité.

Théorème 1.5. *Soit $I = [a,b]$ un* segment*, i.e. un intervalle fermé borné de $\mathbb{R}$, et $f\colon I \to \mathbb{C}$ une fonction continue sur I. Alors f est uniformément continue sur I.*

Démonstration. Supposons que f ne soit pas uniformément continue sur $[a,b]$: alors il existe $\varepsilon > 0$ tel que, pour tout entier n, on peut trouver x_n et y_n dans I avec $|x_n - y_n| \leq \frac{1}{n}$ mais $|f(x_n) - f(y_n)| \geq \varepsilon$. Grâce au théorème de Bolzano-Weierstrass, on peut trouver une application strictement croissante $\varphi\colon \mathbb{N} \to \mathbb{N}$ telle que les suites $(x_{\varphi(n)})$ et $(y_{\varphi(n)})$ soient toutes les deux convergentes. Comme $\varphi(n) \geq n$, on a $|x_{\varphi(n)} - y_{\varphi(n)}| \leq \frac{1}{\varphi(n)} \to 0$ quand n tend vers $+\infty$, donc $(x_{\varphi(n)})$ et $(y_{\varphi(n)})$ convergent vers le même point x. Puisque $|f(x_{\varphi(n)}) - f(y_{\varphi(n)})| \geq \varepsilon$, il est impossible que les deux suites $f(x_{\varphi(n)})$ et $f(y_{\varphi(n)})$ convergent toutes deux vers $f(x)$. Par conséquent f n'est pas continue en x, donc f n'est pas continue sur I.

On vient de montrer que si f n'est pas uniformément continue sur I alors f n'est pas continue sur I, ce qui est la même chose que montrer que si f est continue sur I alors f est uniformément continue sur I. $\square$

i. Question subsidiaire : quels polynômes sont uniformément continus sur $\mathbb{R}$?

On a vu que l'idée de la continuité uniforme était que, pour un ε fixé, le δ qui témoigne de la continuité devient indépendant du point où l'on se place. La même idée se retrouve quand on considère des suites de fonctions.

Définition 1.6. Soit I un intervalle de $\mathbb{R}$, (f_n) une suite de fonctions définies sur I à valeurs dans $\mathbb{C}$ et $f\colon I \to \mathbb{C}$. On dit que (f_n) *converge simplement* vers f sur I si pour tout $x \in I$ la suite $(f_n(x))$ converge vers $f(x)$. En utilisant des quantificateurs :

$$\forall \varepsilon > 0 \; \forall x \in I \; \exists N \; \forall n \geq N \; |f_n(x) - f(x)| \leq \varepsilon \; .$$

Comme dans la définition de la continuité, N dépend a priori de ε *et* de x ; quand il est possible de choisir un N qui ne dépend que de ε, on dit qu'il y a *convergence uniforme.*

Définition 1.7. Soit I un intervalle de $\mathbb{R}$, (f_n) une suite de fonctions définies sur I à valeurs dans $\mathbb{C}$ et $f\colon I \to \mathbb{C}$. On dit que (f_n) *converge uniformément* vers f sur I si

$$\forall \varepsilon > 0 \; \exists N \; \forall x \in I \; \forall n \geq N \; |f_n(x) - f(x)| \leq \varepsilon \; .$$

De manière équivalente, (f_n) converge uniformément vers f sur I si la suite $\sup_I |f_n - f|$ converge vers 0 quand n tend vers $+\infty$.

La convergence uniforme est bien plus forte que la convergence simple ; si les f_n sont des fonctions sympathiques (par exemple, continues) convergeant uniformément vers f, on peut espérer que f ait également des propriétés sympathiques. Le théorème suivant en est un exemple important.

Théorème 1.8. *Soit I un intervalle de $\mathbb{R}$, et (f_n) une suite de fonctions continues sur I à valeurs dans $\mathbb{C}$ convergeant uniformément vers $f\colon I \to \mathbb{C}$. Alors f est continue.*

Démonstration. Fixons $x \in I$ et $\varepsilon > 0$. Il existe N tel que

$$\forall y \in I \; \forall n \geq N \; |f_n(y) - f(y)| \leq \varepsilon \; .$$

Fixons un tel N ; comme f_N est continue en x, il existe δ tel que pour tout $y \in I$ satisfaisant $|x - y| \leq \delta$ on ait $|f_N(y) - f_N(x)| \leq \varepsilon$.

Par conséquent, pour tout $y \in I$ satisfaisant $|x - y| \leq \delta$ on a

$$\begin{aligned} |f(x) - f(y)| &= |f(x) - f_N(x) + f_N(x) - f_N(y) + f_N(y) - f(y)| \\ &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| \\ &\leq 3\varepsilon \; . \end{aligned}$$

Comme $\varepsilon > 0$ et x étaient quelconques, cela suffit à prouver que f est continue sur I. $\square$

Remarque 1.9. Dans la preuve, on a eu besoin de l'*inégalité triangulaire*, qui affirme que, étant donnés deux complexes a et b, on a $|a + b| \leq |a| + |b|$. Cette inégalité est fondamentale en analyse ; elle tire son nom du fait qu'elle exprime analytiquement le fait que, dans un triangle, la longueur d'un côté est toujours plus courte que la somme des longueurs des deux autres côtés - conséquence du fait que le plus court chemin entre deux points est une ligne droite.

Exercice 1.10. Soit I un segment, de $\mathbb{R}$, (f_n) une suite de fonctions à valeurs complexes convergeant uniformément vers f sur I, et (g_n) une suite de fonctions à valeurs complexes convergeant uniformément vers g sur I. Montrer que $(f_n g_n)$ converge uniformément vers fg sur I.

1.2 Intégrale des fonctions en escalier

Définition 1.11. Soient $a < b$ deux réels. Une *subdivision* de $[a, b]$ est une suite finie $(a_0, a_1, \ldots, a_n)$ telle que $a_0 = a$, $a_n = b$ et $a_i < a_{i+1}$ pour tout $i \in \{0, \ldots, n-1\}$. On définit le *pas* d'une subdivision $(a_0, a_1, \ldots, a_n)$ comme étant égal à la quantité $\max\{a_{i+1} - a_i \colon i \in \{0, \ldots, n-1\}\}$.

Intuitivement, considérer une subdivision $(a_0, \ldots, a_n)$ revient à considérer un découpage de $[a, b]$ en n intervalles $[a_0, a_1], \ldots, [a_{n-1}, b]$; dire que le pas de la subdivision est petit signifie que tous les intervalles créés lors du découpage sont petits.

Définition 1.12. Soient $a < b$ deux réels, et $(a_0, \dots, a_n)$, $(b_0, \dots, b_m)$ deux subdivisions de $[a, b]$. On dit que $(b_0, \dots, b_m)$ *raffine* $(a_0, \dots, a_n)$ si chaque intervalle $[b_j, b_{j+1}]$ est contenu dans un invervalle de la forme $[a_k, a_{k+1}]$.

Cela signifie que la subdivision $(b_0, \dots, b_m)$ a été obtenue en découpant les intervalles de la subdivision $(a_0, \dots, a_n)$.

Exercice 1.13. Soients $a < b$ deux réels, et $(a_0, \dots, a_n)$ et $(b_0, \dots, b_m)$ deux subdivisions de $[a, b]$. Alors il existe une subdivision $(c_0, \dots, c_p)$ qui raffine *à la fois* $(a_0, \dots, a_n)$ et $(b_0, \dots, b_m)$.

(Indication : $c_0, \dots, c_p$ peuvent par exemple être obtenus en écrivant dans l'ordre croissant l'ensemble $\{a_0, \dots, a_n; b_0, \dots, b_m\}$)

Définition 1.14. Soient $a < b$ deux réels ; $f : [a, b] \to \mathbb{C}$ est une *fonction en escalier* s'il existe une subdivision $(a_0, \dots, a_n)$ de $[a, b]$ telle que f soit constante sur chaque intervalle $]a_i, a_{i+1}[$. On dit que $(a_0, \dots, a_n)$ *témoigne* du fait que f est en escalier, ou encore est une *subdivision adaptée* à f.

Proposition 1.15.
1. *Une fonction en escalier ne prend qu'un nombre fini de valeurs.*
2. *Une combinaison linéaire de fonctions en escalier sur $[a, b]$ est une fonction en escalier sur $[a, b]$.*
3. *Un produit de fonctions en escalier sur $[a, b]$ est une fonction en escalier sur $[a, b]$.*

Démonstration. La première propriété découle immédiatement de la définition. Les preuves des deuxièmes et troisième propriétés sont très similaires, on va simplement montrer la troisième. Soient donc f, g deux fonctions en escalier, $(a_0, \dots, a_n)$ une subdivision qui témoigne du fait que f est en escalier et $(b_0, \dots, b_m)$ une subdivision qui témoigne du fait que g est en escalier. Par l'exercice précédent, on peut trouver une subdivision $(c_0, \dots, c_p)$ qui raffine ces deux subdivisions. Etant donné i entre 0 et $p - 1$, il existe j, k tel que $[c_i, c_{i+1}]$ soit contenu dans $[a_j, a_{j+1}]$ et dans $[b_k, b_{k+1}]$. En particulier, les deux fonctions f et g sont constantes sur $]c_i, c_{i+1}[$, donc fg y est constante aussi. Ainsi, la subdivision $(c_0, \dots, c_p)$ témoigne du fait que fg est une fonction en escalier. □

Définition 1.16. Soient $a < b$ deux réels, $f : [a, b] \to \mathbb{C}$ une fonction en escalier, et $\sigma = (a_0, \dots, a_n)$ une subdivision adaptée à f. On pose

$$I(f, \sigma) = \sum_{k=0}^{n-1} (a_{k+1} - a_k) f\left(\frac{a_k + a_{k+1}}{2}\right) .$$

Remarque 1.17. Dans la définition de $I(a, \sigma)$, on aurait pu remplacer $\frac{a_k + a_{k+1}}{2}$ par n'importe quel point de $]a_k, a_{k+1}[$ sans changer la valeur de $I(a, \sigma)$.

Lemme 1.18. *Soient $a < b$ deux réels, $f : (a, b] \to \mathbb{C}$ une fonction en escalier, et σ, τ deux subdivisions adaptées à f. Alors $I(f, \sigma) = I(f, \tau)$.*

Démonstration. Commençons par le cas où $\tau = (b_0, \dots, b_m)$ raffine $\sigma = (a_0, \dots, a_n)$. Alors il existe $j_0, \dots, j_n$ tels que pour tout $k \in \{0, \dots, n\}$ on ait $b_{j_k} = a_k$ (en particulier $j_0 = 0$, $j_n = m$). Alors on a

$$\begin{aligned}
I(f, \tau) &= \sum_{j=0}^{m-1} (b_{j+1} - b_j) f\left(\frac{b_j + b_{j+1}}{2}\right) \\
&= \sum_{k=0}^{n-1} \sum_{j=i_k}^{i_{k+1}-1} (b_{j+1} - b_j) f\left(\frac{b_j + b_{j+1}}{2}\right) \\
&= \sum_{k=0}^{n-1} \sum_{j=i_k}^{i_{k+1}-1} (b_{j+1} - b_j) f\left(\frac{a_k + a_{k+1}}{2}\right) \\
&= \sum_{k=0}^{n-1} f\left(\frac{a_k + a_{k+1}}{2}\right) \left(\sum_{j=i_k}^{i_{k+1}-1} b_{j+1} - b_j\right) \\
&= \sum_{k=0}^{n-1} f\left(\frac{a_k + a_{k+1}}{2}\right) (a_{k+1} - a_k) \\
&= I(f, \sigma) .
\end{aligned}$$

On a donc démontré le résultat désiré dans le cas où τ raffine σ. Si maintenant τ et σ sont deux subdivisions quelconques adaptées à f, il existe une subdivision γ qui raffine à la fois τ et σ. Cette subdivison est encore adaptée à f, donc par le cas précédent on a $I(f,\sigma) = I(f,\gamma)$ et $I(f,\gamma) = I(f,\tau)$. On en déduit bien que $I(f,\sigma) = I(f,\tau)$.

□

Ce lemme nous permet finalement de définir l'intégrale d'une fonction en escalier.

Définition 1.19. Soient $a < b$ deux réels, $f\colon [a,b] \to \mathbb{C}$ une fonction en escalier, et σ une subdivision adaptée à f. On pose

$$\int_a^b f(x)dx = I(f,\sigma) \ .$$

Le lemme précédent nous dit que cette définition est légitime : quelle que soit la subdivision σ adaptée à f que l'on choisisse, $I(f,\sigma)$ a toujours la même valeur.

Quelle est l'intuition derrière cette définition ? Pour une fonction f à valeurs positives, l'intégrale est censée représenter « l'aire sous la courbe de f ». Dans le cas où f est en escalier, le domaine sous la courbe de f est une union finie de rectangles, et la formule que l'on a donnée pour l'intégrale de f correspond à la somme des aires de ces rectangles.

Evidemment, on ne veut pas intégrer que des fonctions en escalier ; l'idée de la construction de l'intégrale présentée dans ces notes est que l'intégrale ici définie se comporte suffisamment bien pour que l'on puisse l'étendre aux fonctions qui sont *limite uniforme* d'une suite de fonctions en escalier. Ce sont ces fonctions qui joueront un rôle clé - en particulier, on montrera que toute fonction continue a cette propriété. Avant cela, faisons une liste de propriétés remarquables de l'intégrale des fonctions en escalier.

Proposition 1.20.
1. *Etant donnés deux réels $a < b$, on a $\int_a^b 1 = b - a$.*
2. *Etant donnés trois réels $a < b < c$ et f une fonction en escalier sur $[a,c]$, on a* **(relation de Chasles)**

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

3. *Etant donnés deux réels $a < b$ et une fonction f en escalier sur $[a,b]$ et à valeurs positives, $\int_a^b f \geq 0$* **(positivité de l'intégrale)***)=.*
4. *Etant donnés deux réels $a < b$, deux fonctions f, g en escalier sur $[a,b]$ et deux constantes α, β, on a* **(linéarité de l'intégrale)** *:*

$$\int_a^b (\alpha f(x) + \beta g(x))\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx \ .$$

5. *Etant donnés deux réels $a < b$, et f une fonction en escalier sur $[a,b]$, on a* **(inégalité triangulaire)**

$$\left| \int_a^b f(x)\,dx \right| \leq \int_a^b |f(x)|\,dx \ .$$

6. *Etant donnés deux réels $a < b$, et f une fonction en escalier sur $[a,b]$, on a*

$$\int_a^b f(x)\,dx = \int_a^b \mathit{Ré}(f)(x)\,dx + i\int_a^b \mathit{Im}(f)(x)\,dx \ .$$

Toutes ces propriétés découlent immédiatement de la définition de l'intégrale des fonctions en escalier ; leur vérification est laissée en exercice (important pour s'assurer que les définitions ont été bien comprises...). On utilisera dans les suites ces propriétés en les combinant entre elles, par exemple sous la forme suivante : si f et g sont deux fonctions en escalier sur $[a,b]$ à valeurs réelles, et $f(x) \leq g(x)$ pour tout $x \in [a,b]$, alors $\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$ (cette inégalité découle des propriétés de linéarité et de positivité).

1.3 Intégrale des fonctions réglées

Définition 1.21. Soient $a < b$ deux réels, et $f\colon [a,b] \to \mathbb{C}$. On dit que f est *réglée* s'il existe une suite de fonctions (f_n) en escalier sur $[a,b]$ et convergeant uniformément vers f.

De manière équivalente, f est réglée sur $[a,b]$ si pour tout $\varepsilon > 0$ il existe une fonction en escalier g telle que pour tout x de $[a,b]$ on ait $|f(x) - g(x)| \leq \varepsilon$. Notons qu'une fonction réglée f est nécessairement bornée : en effet, il existe une fonction en escalier g telle que $|f(x) - g(x)| \leq 1$ pour tout x de $[a,b]$ (caractérisation qu'on vient d'énoncer appliquée avec $\varepsilon = 1$). Comme g ne prend qu'un nombre fini de valeurs, il existe M tel que $|g(x)| \leq M$ pour tout x de $[a,b]$, ce dont on déduit que

$$\forall x \in [a,b] \ |f(x)| \leq |f(x) - g(x)| + |g(x)| \leq 1 + M \ .$$

Un dernier effort avant de pouvoir définir l'intégrale des fonctions réglées.

Lemme 1.22. *Soient $a < b$ deux réels.*

1. *Soit (f_n) une suite de fonctions en escalier sur $[a,b]$ convergeant uniformément vers une fonction f. Alors la suite $(\int_a^b f_n(x)\,dx)$ est une suite de Cauchy (donc convergente).*
2. *Si (f_n) et (g_n) sont deux suites de fonctions en escalier sur $[a,b]$ convergeant uniformément vers la même fonction f, alors*
$$\lim_{n\to+\infty} \int_a^b f_n(x)\,dx = \lim_{n\to+\infty} \int_a^b g_n(x) \ , dx \ .$$

Démonstration. Commençons par prouver (1). Fixons $\varepsilon > 0$. Puisque (f_n) converge uniformément vers f, il existe N tel que
$$\forall m,n \geq N \ \forall x \in [a,b] \ |f_n(x) - f_m(x)| \leq \varepsilon \ .$$
On en déduit, par linéarité et positivité de l'intégrale des fonctions en escalier, que :

$$\forall m,n \geq N \int_a^b |f_n(x) - f_m(x)|\,dx \leq \int_a^b \varepsilon\,dx = \varepsilon(b-a) \ .$$

En appliquant l'inégalité triangulaire, on a finalement

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

Ainsi, pour tout $\varepsilon > 0$ il existe N tel que

$$\forall m,n \geq N \ \left|\int_a^b f_n(x)\,dx - \int_a^b f_m(x)\,dx\right| \leq \varepsilon(b-a) \ .$$

Ceci achève la démonstration de (1) ; pour prouver (2), nous devons simplement montrer que $\int_a^b f_n(x)\,dx - \int_a^b g_n(x)\,dx$ converge vers 0. Fixons à nouveau $\varepsilon > 0$; puisque (f_n) et (g_n) convergent uniformément vers la même fonction, $(f_n - g_n)$ converge uniformément vers 0, par conséquent il existe N tel que $|f_n(x) - g_n(x)| \leq \varepsilon$ pour tout $x \in [a,b]$ et tout $n \geq N$. On a alors, pour tout $n \geq N$:

$$\begin{aligned}
\left|\int_a^b f_n(x)\,dx - \int_a^b g_n(x)\,dx\right| &= \left|\int_a^b (f_n(x) - g_n(x))\,dx\right| \\
&\leq \int_a^b |f_n(x) - g_n(x)|\,dx \\
&\leq \varepsilon(b-a)
\end{aligned}$$

Ceci prouve que $\int_a^b f_n(x)\,dx - \int_a^b g_n(x)\,dx$ converge vers 0.

□

Définition 1.23. Soient $a < b$ deux réels, f une fonction réglée sur $[a, b]$ et (f_n) une suite de fonctions en escalier sur $[a, b]$ qui converge uniformément vers f. Alors on pose

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

La limite apparaissant dans la définition ne dépend pas du choix de la suite (f_n) de fonctions en escalier convergeant uniformément vers f, donc notre définition a bien un sens. Comme pour les fonctions en escalier, on pourrait énumérer les propriétés fondamentales de l'intégrale des fonctions réglées, ce qu'on fera dans la section suivante. Pour l'instant, introduisons la classe de fonctions que nous voulons vraiment pouvoir intégrer.

Exercice 1.24. Soient $a < b$ deux réels.Montrer qu'une combinaison linéaire et un produit de fonctions réglées sur $[a, b]$ sont encore des fonctions réglées. Montrer qu'une limite uniforme de fonctions réglées sur $[a, b]$ est encore une fonction réglée sur $[a, b]$.

Définition 1.25. Soient $a < b$ deux réels. On dit qu'une fonction $f \colon [a, b] \to \mathbb{C}$ est *continue par morceaux* sur $[a, b]$ s'il existe une subdivision $(a_0, \dots, a_n)$ de $[a, b]$ telle que la restriction de f à chaque intervalle $]a_i, a_{i+1}[$ soit continue et admette une limite à droite en a_i et une limite à gauche en a_{i+1} (autrement dit, il existe une fonction continue sur $[a_i, a_{i+1}]$ qui coïncide avec f sur $]a_i, a_{i+1}[$).

Comme premiers exemples, notons qu'une fonction continue sur $[a, b]$ est bien sûr continue par morceaux, et qu'une fonction en escalier est également continue par morceaux.

Définition 1.26. Soient $a < b$ deux réels. Une *subdivision pointée* $\sigma = (a_0, \dots, a_n; \xi_0, \dots, \xi_{n-1})$ de $[a, b]$ est la donnée :

- d'une subdivision $(a_0, \dots, a_n)$ de $[a, b]$;
- de points $\xi_0, \dots, \xi_{n-1}$ de $[a, b]$ tels que pour tout $i \in \{1, \dots, n-1\}$ on ait $\xi_i \in [a_i, a_{i+1}]$.

Théorème 1.27. *Soient $a < b$ deux réels, $f \colon [a, b] \to \mathbb{C}$ une fonction continue et (σ_p) une suite de subdivisions pointées $(a_0^p, \dots, a_{n_p}^p; \xi_0^p, \dots, \xi_{n_p-1}^p)$ dont le pas tend vers 0 quand p tend vers $+\infty$. Soit f_p la fonction en escalier sur $[a, b]$ égale à $f(\xi_i^p)$ sur $[a_i^p, a_{i+1}^p[$. Alors la suite (f_p) converge uniformément vers f.*

Démonstration. Fixons $\varepsilon > 0$. Comme f est continue sur $[a, b]$, elle est uniformément continue sur $[a, b]$. Ainsi, il existe δ tel que

$$\forall x, y \in I \ |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon \ .$$

Pour p suffisamment grand, le pas de σ_p est inférieur à δ. Alors, pour tout $i \in \{0, \dots, n_p - 1\}$ et tout x dans $[a_i, a_{i+1}[$ on a $|f(x) - f_p(x)| = |f(x) - f(\xi_i^p)| \leq \varepsilon$ puisque $|x - \xi_i^p| \leq \delta$. On en déduit que, pour p suffisamment grand :

$$\forall x \in I \ |f(x) - f_p(x)| \leq \varepsilon \ .$$

Ceci prouve que (f_p) converge uniformément vers f. □

On en déduit immédiatement le résultat suivant.

Théorème 1.28. *Soient $a < b$ deux réels. Toute fonction continue par morceaux sur $[a, b]$ est réglée.*

En effet, le résultat précédent entraîne que toute fonction continue sur $[a, b]$ est réglée ; le cas des fonctions continues par morceaux est laissé en exercice. Dans la suite de ce cours, nous nous concentrerons sur l'intégrale des fonctions continues par morceaux. L'intégrale d'une fonction continue par morceaux peut se calculer en découpant son intervalle de définition en intervalles témoignant de la continuité par morceaux, comme le précise l'énoncé ci-dessous dont la démonstration est laissée en exercice.

Proposition 1.29. *Soit f une fonction continue par morceaux sur $[a, b]$, et $a_0 < \ldots < a_n$ des éléments de $[a, b]$ tels que $a_0 = a, a_n = b$ et f coïncide sur chaque intervalle $]a_i, a_{i+1}[$ avec une fonction f_i qui est continue sur $[a_i, a_{i+1}]$. Alors on a*

$$\int_a^b f(x)\, dx = \sum_{i=0}^{n-1} \int_{a_i}^{a_{i+1}} f_i(x)\, dx \ .$$

Définition 1.30. Soient $a < b$ deux réels, $f\colon [a,b] \to \mathbb{C}$ continue par morceaux et $\sigma = (a_0, \ldots, a_n; \xi_0, \ldots, \xi_{n-1})$ une subdivision pointée de $[a,b]$. On appelle *somme de Riemann* associée à f et σ le nombre

$$S(f,\sigma) = \sum_{k=0}^{n-1} (a_{k+1} - a_k) f(\xi_k) \ .$$

Théorème 1.31. *Soient $a < b$ deux réels, f une fonction continue par morceaux sur $[a,b]$ et (σ_p) une suite de subdivisions pointées de $[a,b]$ dont le pas tend vers 0 quand p tend vers $+\infty$. Alors $S(f,\sigma_p)$ tend vers $\int_a^b f(x)\,dx$ quand p tend vers $+\infty$.*

Démonstration. Dans le cas où f est continue, ce résultat est une conséquence immédiate du théorème 1.27, dont on reprend ici les notations : en effet, la suite (f_p) converge uniformément vers f, et $\int_a^b f_p(x)\,dx = S(f,\sigma_p)$. Par définition de l'intégrale d'une fonction réglée, $S(f,\sigma_p)$ tend donc vers $\int_a^b f(x)\,dx$.

Le cas des fonctions continues par morceaux est laissé en exercice (diviser $[a,b]$ en intervalles témoignant du fait que f est continue par morceaux, puis appliquer le résultat obtenu pour les fonctions continues et la proposition 1.29). $\square$

Un cas particulier très important en pratique est celui ou l'on divise l'intervalle $[a,b]$ en n intervalles de même longueur $\frac{b-a}{n}$. On obtient alors, par exemple, les formules suivantes :

– Si f est une fonction continue sur $[a,b]$, alors

$$\int_a^b f(x)\,dx = \lim_{n\to+\infty} \left(\frac{b-a}{n}\right) \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) \ .$$

– Si f est une fonction continue sur $[a,b]$, alors

$$\int_a^b f(x)\,dx = \lim_{n\to+\infty} \left(\frac{b-a}{n}\right) \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right) \ .$$

En effet, la première somme correspond à une somme de Riemann pour une subdivision pointée où l'on découpe $[a,b]$ en n intervalles de même longueur, et où le point choisi dans chacun des intervalles est son extrémité de gauche ; la deuxième formule correspond au même découpage de l'intervalle, en choisissant dans chaque intervalle son extrémité de droite.

Cette constatation nous permettrait de calculer les intégrales de certaines fonctions continues très simples : par exemple,

$$\int_0^1 x\,dx = \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{k}{n} = \lim_{n\to+\infty} \frac{1}{n^2} \frac{n(n+1)}{2} = \frac{1}{2} \ .$$

Bien sûr, on est habitué à calculer ce type d'intégrales en utilisant des primitives, ce qu'on fera dans la section suivante, et il est donc assez rare qu'on utilise une somme de Riemann pour calculer une intégrale ; par contre, calculer une intégrale peut permettre de calculer des limites de suites, si les suites en question sont des sommes de Riemann (ou sont proches d'être des sommes de Riemann ; cf. feuille d'exercices).

Exercice 1.32. En utilisant une somme de Riemann, calculer $\int_0^\pi e^{ix}\,dx$, $\int_0^\pi \cos(x)\,dx$ et $\int_0^\pi \sin(x)\,dx$.

1.4 Propriétés fondamentales de l'intégrale

Notation. Soient $a > b$ deux réels, et $f\colon [b,a] \to \mathbb{C}$ une fonction continue par morceaux. On pose $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$. On pose aussi $\int_a^a f(x)\,dx = 0$.

Proposition 1.33. – *Soient a, b, c trois réels et f une fonction continue par morceaux sur un intervalle qui contient a,b et c. Alors on a* **(Relation de Chasles)**

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

– *Etant donnés deux réels $a < b$ et une fonction f continue par morceaux sur $[a, b]$ et à valeurs positives, $\int_a^b f \geq 0$* **(positivité de l'intégrale)***.*
– *Etant donnés deux réels $a < b$, deux fonctions f, g continue par morceauxs sur $[a, b]$ et deux constantes α, β, on a* **(linéarité de l'intégrale)** *:*

$$\int_a^b (\alpha f(x) + \beta g(x))\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx \ .$$

– *Etant donnés deux réels $a < b$, et f une fonction continue par morceaux sur $[a, b]$, on a* **(inégalité triangulaire)**

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx \ .$$

– *Etant donnés deux réels $a < b$, et f une fonction continue par morceaux sur $[a, b]$, on a*

$$\int_a^b f(x)\, dx = \int_a^b Ré(f)(x)\, dx + i \int_a^b Im(f)(x)\, dx \ .$$

Toutes ces propriétés se déduisent facilement à partir de la définition de l'intégrale d'une fonction continue par morceaux et des propriétés analogues pour l'intégrale d'une fonction en escalier ; leur vérification est laissée en exercice.

Théorème 1.34. *Soient $a < b$ deux réels et f une fonction continue par morceaux sur $[a, b]$* **à valeurs réelles***. Alors on a*

$$(b - a) \inf_{[a,b]} f \leq \int_a^b f(x)\, dx \leq (b - a) \sup_{[a,b]} f(x)\, dx \ .$$

Si f est de plus supposée continue, alors il existe $c \in [a, b]$ tel que $f(x) = \frac{1}{b-a} \int_a^b f(x)\, dx$.

Notons que l'inf et le sup dans le théorème ci-dessus sont bien définis puisqu'une fonction continue par morceaux sur $[a, b]$ est nécessairement bornée.

Démonstration. Par définition d'un inf et d'un sup, on a, pour tout $x \in [a, b]$, $\inf_{[a,b]} f \leq f(x) \leq \sup_{[a,b]} f$. Par positivité de l'intégrale, on en déduit

$$\int_a^b \inf_{[a,b]} f\, dx \leq \int_a^b f(x)\, dx \leq \int_a^b \sup_{[a,b]} f\, dx \ .$$

Ceci prouve l'inégalité désirée. Si maintenant f est continue sur $[a, b]$, alors l'inf et le sup sont un min et un max puisqu'une fonction continue sur un intervalle fermé borné est bornée et atteint ses bornes sur cet intervalle. Appelons m le minimum de f sur $[a, b]$ et M le maximum de f sur $[a, b]$. On a alors $f([a, b]) = [m, M]$ par le théorème des valeurs intermédiaires, et l'inégalité ci-dessus donne

$$m \leq \frac{1}{b - a} \int_a^b f(x)\, dx \leq M \ .$$

Ceci achève la démonstration. □

Théorème 1.35. *Soient $a < b$ deux réels, et f une fonction continue par morceaux sur $[a, b]$ et à valeurs positives. Alors $\displaystyle\int_a^b f(x)\, dx = 0$ si, et seulement si, f est nulle partout sauf peut-être en un nombre fini de points de $[a, b]$.*

Démonstration. Si f est nulle partout sauf peut-être en un nombre fini de points de $[a, b]$, alors il découle de la proposition 1.29 que f est d'intégrale nulle.

Réciproquement, supposons f d'intégrale nulle. Commençons par le cas où f est continue ; si f prend une valeur strictement positive en $x_0 \in [a, b]$, alors il existe un intervalle $[c, d] \subseteq [a, b]$ avec $c < d$ sur lequel f est à valeurs strictement positives, donc le minimum de f sur $[c, d]$ est strictement positif, et le théorème précédent nous permet de conclure que $\displaystyle\int_c^d f(x)\,dx > 0$; la relation de Chasles et la positivité de l'intégrale nous donnent alors :

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^d f(x)\,dx + \int_d^b f(x)\,dx > 0 \ .$$

On vient de montrer que, si f est continue, à valeurs positives et prend une valeur non nulle, alors son intégrale est strictement positive ; autrement dit, si $\displaystyle\int_a^b f(x)\,dx = 0$ et f est continue à valeurs positives sur $[a,b]$ alors f est la fonction nulle.

Reste le cas où f n'est que supposée continue par morceaux : alors il existe $a_1 < \ldots < a_n$ avec $a_1 = a$, $a_n = b$ tels que f coïncide sur chaque intervalle $]a_i, a_{i+1}[$ avec une fonction f_i qui se prolonge continûment à $[a_i, a_{i+1}]$, et on a

$$\int_a^b f(x)\,dx = \sum_{i=1}^{n-1} \int_{a_i}^{a_{i+1}} f_i(x)\,dx \ .$$

Par positivité de l'intégrale, si $\displaystyle\int_a^b f(x)\,dx = 0$ alors on doit avoir $\displaystyle\int_{a_i}^{a_{i+1}} f_i(x)\,dx = 0$ pour tout $i \in \{1, \ldots, n-1\}$; comme les f_i sont continues on en déduit que chaque f_i est nulle sur $[a_i, a_{i+1}]$. Par conséquent, f est nulle partout sauf peut-être en $a_1, \ldots, a_n$. □

Le théorème suivant, qui peut s'avérer très utile pour comprendre le comportement des intégrales de produits de fonctions, en particulier lorsqu'on étudie des intégrales généralisées.

Théorème 1.36 (Première formule de la moyenne)**.** *Soient $a < b$ deux réels, $f \colon [a,b] \to \mathbb{R}$ une fonction continue et $g \colon [a,b] \to \mathbb{R}$ une fonction continue par morceaux et à valeurs positives. Alors il existe $c \in [a,b]$ tel que*

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

Démonstration. Comme g est continue par morceaux et à valeurs positives, l'intégrale de g ne peut valoir 0 que si g est nulle partout sauf peut-être en un nombre fini de points, auquel cas il en va de même de fg, dont l'intégrale est donc nulle elle aussi. Par conséquent, tout $c \in [a,b]$ est tel que $\displaystyle\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx$ dans ce cas (très) particulier.

On peut donc supposer que $\displaystyle\int_a^b g(x)\,dx \neq 0$, auquel cas on doit montrer qu'il existe c tel que

$$\frac{1}{\int_a^b g(x)\,dx}\int_a^b f(x)g(x)\,dx = f(x) \ .$$

Comme f est continue sur $[a,b]$, l'image de $[a,b]$ par f est un intervalle $[m,M]$ (où m est le minimum de f sur $[a,b]$ et M son maximum), et l'existence d'un c comme ci-dessus est équivalente à l'inégalité

$$m \le \frac{1}{\int_a^b g(x)\,dx}\int_a^b f(x)g(x)\,dx \le M$$

Autrement dit, il nous suffit de démontrer que

$$m\int_a^b g(x)\,dx \le \int_a^b f(x)g(x)\,dx \le M\int_a^b g(x)\,dx \ .$$

Comme g est à valeurs positives, on a pour tout $x \in [a,b]$ $mg(x) \le f(x)g(x) \le Mg(x)$, et on obtient l'inégalité désirée par positivité et linéarité de l'intégrale.

□

Exercice 1.37. Montrer que, si g est continue et à valeurs strictement positives, on peut obtenir $c \in]a,b[$ dans la conclusion du théorème ci-dessus (on pourra introduire la fonction $G \colon t \mapsto \int_a^t g(x)\,dx$ et utiliser le changement de variables $u = a + \dfrac{G(t)(b-a)}{G(b)}$).

Théorème 1.38. *Soient $a < b$ deux réels, et f une fonction continue par morceaux sur $[a, b]$. Alors la fonction F définie sur $[a, b]$ par $F(t) = \int_a^t f(x)\, dx$ est continue sur $[a, b]$.*

Démonstration. Une fonction continue par morceaux est nécessairement bornée, autrement dit il existe M tel que $|f(x)| \leq M$ pour tout $x \in [a, b]$. On a alors, pour tout $s, t \in [a, b]$:

$$|F(t) - F(s)| = \left| \int_s^t f(x)\, dx \right| \leq \left| \int_s^t |f(x)|\, dx \right| \leq M|t - s| \ .$$

En particulier, $F(s)$ tend vers $F(t)$ quand s tend vers t, donc F est continue sur $[a, b]$ [ii] □

Il serait tentant de penser que la fonction F définie ci-dessus est toujours dérivable, et que $F' = f$. C'est faux en général : par exemple, pour des fonctions à valeurs réelles, le théorème de Darboux nous dit qu'une fonction dérivée doit vérifier la conclusion du théorème des valeurs intermédiaires (i.e. l'image d'un intervalle par une fonction dérivée est un intervalle) donc une fonction en escalier non constante ne peut jamais être une dérivée. Il existe néanmoins un cas essentiel où ce résultat est vrai.

Théorème 1.39 (Théorème fondamental de l'analyse)**.** *Soient $a < b$ deux réels, f une fonction continue sur $[a, b]$, et F la fonction définie sur $[a, b]$ par $F(t) = \int_a^t f(x)\, dx$. Alors F est dérivable sur $[a, b]$, $F' = f$, et F est l'unique primitive de f sur $[a, b]$ qui s'annule en a.*

Démonstration. Commençons par montrer que $F' = f$; pour cela, fixons $t \in [a, b]$ (bien sûr en a on ne considèrera qu'une dérivée à droite, et de même en b on ne considèrera qu'une dérivée à gauche) Pour tout $s \in [a, b]$, on a :

$$F(t) - F(s) - (t - s)f(t) = \int_s^t f(x)\, dx - \int_s^t f(t)\, dx = \int_s^t (f(x) - f(t))\, dx \ .$$

Fixons $\varepsilon > 0$. Comme f est continue en t, il existe $\delta > 0$ tel que, pour tout $s \in [a, b]$, on ait

$$|t - s| \leq \delta \Rightarrow |f(t) - f(s)| \leq \varepsilon \ .$$

Notons que si $|t-s| \leq \delta$ alors $|t-x| \leq \delta$ pour tout x appartenant au segment d'extrémités t et s. Par conséquent, $|f(t) - f(x)| \leq \varepsilon$ pour tout x appartenant à ce segment, et l'inégalité triangulaire nous donne, pour tout s tel que $|t - s| \leq \delta$:

$$\left| \int_s^t (f(x) - f(t))\, dx \right| \leq \varepsilon |t - s| \ .$$

On obtient donc finalement, pour tout $s \in [a, b]$ tel que $|t - s| \leq \delta$, que

$$\left| \frac{F(t) - F(s)}{t - s} - f(t) \right| \leq \varepsilon \ .$$

Ceci prouve que $\lim_{s \to t} \frac{F(s) - F(t)}{s - t} = f(t)$, autrement dit que F est dérivable en t et $F'(t) = f(t)$.

Si maintenant G est une autre primitive de f sur $[a, b]$ qui s'annule en a, alors $(G - F)' = 0$, donc l'inégalité des accroissements finis appliquée à $G - F$ (qui est continue et dérivable sur $[a, b]$) entraîne que $G - F$ est constante sur $[a, b]$. Comme $G(a) = F(a) = 0$ par hypothèse, on obtient bien que $G(x) = F(x)$ pour tout x de $[a, b]$. □

Remarque 1.40. La raison pour laquelle on a utilisé l'inégalité des accroissements finis (qu'on reverra plus tard dans ce cours) est que, pour des fonctions à valeurs complexes, il n'y pas d'égalité des accroissements finis, comme le montre l'exemple de la fonction $t \mapsto e^{it}$.

Exercice 1.41. En utilisant le théorème fondamental de l'analyse, donner une nouvelle démonstration du théorème 1.35.

On peut maintenant se rappeler de notre technique habituelle pour calculer des intégrales : utiliser les primitives (technique qui ne marche, hélas, pas toujours...).

Corollaire 1.42. Soient $a < b$ deux réels, f une fonction continue sur $[a, b]$ à valeurs dans $\mathbb{C}$ et F une primitive de f sur $[a, b]$. Alors on a $\int_a^b f(x)\, dx = F(b) - F(a)$.

ii. On vient en fait de montrer que F est *lipschitzienne* sur $[a, b]$.

On note souvent $[F]_a^b$ la quantité $F(b) - F(a)$.

Démonstration. Si F est une primitive de f sur $[a, b]$, alors $F - F(a)$ est encore une primitive de f sur $[a, b]$, par conséquent le théorème précédent nous donne que, pour tout $t \in [a, b]$, on a $F(t) - F(a) = \int_a^t f(x)\,dx$. On obtient le résultat désiré en appliquant cette formule pour $t = b$. □

Corollaire 1.43 (Formule d'intégration par parties)**.** Soient $a < b$ deux réels et f, g deux fonctions de classe C^1 sur $[a, b]$ à valeurs dans $\mathbb{C}$. Alors on a

$$\int_a^b f(x)g'(x)\,dx = [fg]_a^b - \int_a^b f'(x)g(x)\,dx \ .$$

Démonstration. On utilise la formule $(fg)' = f'g + g'f$. Comme $f'g + g'f$ est continue, et a pour primitive fg, on peut lui appliquer le résultat précédent et obtenir (par linéarité de l'intégrale)

$$[fg]_a^b = \int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx \ .$$

C'est la formule qu'on souhaitait démontrer. □

Corollaire 1.44 (Formule de changement de variables)**.** Soient $a < b, c < d$ quatre réels, f une fonction continue sur $[a, b]$ à valeurs dans $\mathbb{C}$ et φ une fonction de classe C^1 sur $[c, d]$ et telle que $\varphi([c, d]) \subseteq [a, b]$. Alors on a

$$\int_c^d f(\varphi(t))\varphi'(t)\,dt = \int_{\varphi(c)}^{\varphi(d)} f(x)\,dx \ .$$

Démonstration. Soit F une primitive de f sur $[a, b]$. Alors, par la formule de dérivation des fonctions composées, $F \circ \varphi$ est une primitive de $(f \circ \varphi)\varphi'$ sur $[c, d]$, et le théorème fondamental de l'analyse nous permet d'écrire :

$$\int_c^d f(\varphi(t))\varphi'(t)\,dt = [F \circ \varphi]_c^d = F(\varphi(d)) - F(\varphi(c)) = \int_{\varphi(c)}^{\varphi(d)} f(x)\,dx \ .$$

□

Il est **très important**, pour la formule de changement de variables écrite ci-dessus, que f soit continue et que φ soit de classe C^1. Il n'est pas nécessaire, par contre, que φ soit une bijection de $[c, d]$ sur son image. Souvent, on applique la formule ci-dessus en « partant de la droite vers la gauche », i.e. on veut poser $x = \varphi(t)$. Il est alors un peu délicat de trouver les bonnes bornes pour l'intégrale de gauche, sauf dans le cas où φ est une bijection. Ce cas particulier est particulièrement utile en pratique.

Corollaire 1.45 (Cas particulier de la formule de changement de variables)**.** Soient $a < b, c < d$ quatre réels, f une fonction continue sur $[a, b]$ à valeurs complexes, et φ une bijection de classe C^1 de $[c, d]$ sur $[a, b]$. Alors on a

$$\int_a^b f(t)\,dt = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(x))\varphi'(x)\,dx \ .$$

Un autre intérêt de la formule ci-dessus est qu'elle se généralise au cas où f est continue par morceaux et aux intégrales de fonctions de plusieurs variables.

Concluons ce chapitre avec la formule de Taylor avec reste intégral, qui nous servira plus tard.

Théorème 1.46 (Formule de Taylor avec reste intégral)**.** *Soit I un intervalle de $\mathbb{R}$, $x, y \in I$ et $f : I \to \mathbb{C}$ une fonction de classe C^{n+1} sur I (pour un entier n). Alors on a*

$$f(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y - x)^k + \int_x^y \frac{(y - t)^n}{n!} f^{(n+1)}(t)\,dt \ .$$

Démonstration. On procède par récurrence sur n. Pour $n=0$, on souhaite montrer que, si f est une fonction de classe C^1 sur $[a,b]$, alors pour tout $x, y \in [a,b]$ on a $f(y) = f(x) + \int_x^y f'(t)\,dt$. C'est une conséquence immédiate du théorème fondamental de l'analyse.

Supposons maintenant la formule établie au rang n, et considérons une fonction f de classe C^{n+2} sur I. Fixons également $x, y \in I$. Puisque f est de classe C^{n+1} sur I, on peut lui appliquer la formule au rang n et obtenir

$$f(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k + \int_x^y \frac{(y-t)^n}{n!} f^{(n+1)}(t)\,dt \ .$$

Puisque $f^{(n+1)}$ est de classe C^1, on peut utiliser une intégration par parties dans l'intégrale exprimant le reste, ce qui donne :

$$\begin{aligned}
\int_x^y \frac{(y-t)^n}{n!} f^{(n+1)}(t)\,dt &= \left[-\frac{(y-t)^{n+1}}{(n+1)!} f^{(n+1)}(t) \right]_x^y + \int_x^y \frac{(y-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,dt \\
&= \frac{(y-x)^{n+1}}{(n+1)!} f^{(n+1)}(x) + \int_x^y \frac{(y-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,dt \ .
\end{aligned}$$

En réinjectant cela dans la formule au rang n, on obtient

$$\begin{aligned}
f(y) &= \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k + \frac{(y-x)^{n+1}}{(n+1)!} f^{(n+1)}(x) + \int_x^y \frac{(y-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,dt \\
&= \sum_{k=0}^{n+1} \frac{f^{(k)}(x)}{k!}(y-x)^k + \int_x^y \frac{(y-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,dt \ .
\end{aligned}$$

□

Remarque 1.47. Il existe d'autres formules de Taylor pour une fonction $f\colon [a,b] \to \mathbb{C}$, qu'on va revoir au début du chapitre suivant.

Chapitre 2

Rappels : équivalents, développements limités...

Dans ce chapitre, I désigne un intervalle ouvert de $\mathbb{R}$, et x_0 un élément de I. Notre objectif principal ici est de rappeler, sans démonstrations, la définition d'un développement limité et les opérations qu'on peut faire sur les développements limités, ainsi que celles qu'on ne doit surtout pas faire. On va surtout s'intéresser aux fonctions définies sur un intervalle ouvert et à valeurs dans $\mathbb{R}$; avant cela, quelques mots sur les équivalents et la comparaison de fonctions ne sont sans doute pas de trop.

2.1 Equivalents et comparaison

Définition 2.1.
Pour f, g deux fonctions de I dans $\mathbb{C}$, on écrit $f(x) = o_{x_0}(g)$ pour dire que $f(x)$ s'écrit sous la forme $g(x)\varepsilon(x)$, où ε est une fonction qui tend vers 0 quand x tend vers x_0 ; on dit que f est *négligeable devant* g au voisinage de x_0.

Remarque 2.2.
- Si g ne s'annule pas au voisinage de x_0 (sauf peut-être en x_0) écrire que $f = o_{x_0}(g)$ est la même chose qu'écrire que $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$. En particulier, écrire $f = o_{x_0}(1)$ signifie simplement que $f(x)$ tend vers 0 quand x tend vers x_0.
- Souvent, x_0 est sous-entendu, et on écrit simplement $f = o(g)$. Dans ce cas, il est important de savoir en quel point on est en train d'essayer de comparer f et g...

Définition 2.3. Pour f, g deux fonctions de I dans $\mathbb{C}$, on écrit $f(x) \sim_{x_0} (g)$ pour dire que $f(x)$ s'écrit sous la forme $g(x)\varepsilon(x)$, où ε est une fonction qui tend vers 1 quand x tend vers x_0 ; on dit que f est *équivalente à* g au voisinage de x_0.

Remarque 2.4.
- Comme son nom l'indique, la relation $\sim_{x_0}$ est une relation d'équivalence sur les fonctions : en particulier, c'est la même chose d'écrire que $f \sim_{x_0}$ ou qu $g \sim_{x_0} f$.
- Si g ne s'annule pas au voisinage de x_0 (sauf peut-être en x_0) écrire que $f = o_{x_0}(g)$ est la même chose qu'écrire que $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$.
- Attention, écrire $f \sim_{x_0} 0$ est une condition très contraignante : cela signifie que f est nulle sur un voisinage de x_0, ce qui est beaucoup plus fort que le fait de dire que $f(x)$ tend vers 0 quand x tend vers x_0. Si jamais vous écrivez dans un exercice $f \sim 0$, il y a une très forte probabilité que vous soyiez en train de commettre une erreur.
- Comme pour o, souvent x_0 est sous-entendu et on écrit simplement $f \sim g$.
- Un équivalent n'est pas une égalité ou une identité magique. En particulier, en général on ne peut pas ajouter des équivalents. Par exemple, définissons des fonctions f_1, g, h sur $\mathbb{R}$ en posant $f(x) = x$, $g(x) = -x$, $h(x) = x + x^2$ pour tout $x \in \mathbb{R}$. Alors, en 0, on a $f(x) \sim h(x)$, mais $f(x) + g(x) = 0 \nsim x^2 h(x) + g(x)$...

Notons qu'on pourrait aussi définir les notions de $o, \sim$ en $\pm\infty$, et aussi les définir pour des suites. On utilisera ces notions sans rappels dans la suite ; si elles posent problème il serait important que vous les révisiez.

2.2 Développements limités

Définition 2.5.
On dit que $f : I \to \mathbb{R}$ admet un développement limité d'ordre n en x_0 s'il existe un polynôme P de degré n, à coefficients réels, et une fonction $\varepsilon : I \to \mathbb{R}$ tels que :

$$\forall x \in I,\ f(x) = P(x - x_0) + (x - x_0)^n \varepsilon(x) \ \textit{et} \ \lim_{x \to x_0} \varepsilon(x) = 0$$

Autrement dit, on a $f(x) = P(x - x_0) + o((x - x_0)^n)$.

On dit alors que P est la *partie régulière* d'ordre n du développement limité, et $f - P$ le *reste* d'ordre n.

Notons que la définition entraîne immédiatement que, si f admet un développement limité à l'ordre n en x_0, de partie régulière P, et $m \leq n$ est un entier, alors f admet aussi un développement limité à l'ordre m en x_0, dont la partie régulière est formée par les termes de P de dégré inférieur ou égal à m.

On se contente souvent ci-dessous de traiter le cas des développements limités au voisinage de 0, puisqu'une simple translation permet de s'y ramener. Rappelons que, si $0 \in I$ et si f admet un développement limité d'ordre $n \geq 1$ en 0, alors f est dérivable en 0 (en fait, admettre un développement limité à l'ordre 1 en x_0 est équivalent à être dérivable en x_0, comme on le vérifie facilement à partir des définitions). Par contre, f' n'est **pas continue** en 0 a priori (considérer par exemple $x \mapsto x^2 \sin(1/x)$).

Proposition 2.6.
Si $f : I \to \mathbb{R}$ admet un développement limité d'ordre n en 0, alors la partie régulière du développement limité est unique (donc le reste est unique également).
Si f est paire ($f(x) = f(-x)$), alors le polynôme P est pair. Si f est impaire ($f(x) = -f(-x)$) alors P est impair.

Maintenant que nous nous souvenons un peu mieux de ce qu'est un développement limité, il est temps d'énoncer les théorèmes qui permettent en pratique de calculer les D.Ls.

Proposition 2.7.
- Formule de Taylor-Lagrange :
Si f est $n + 1$ fois dérivable sur I, alors f admet un développement limité d'ordre n en 0, de partie régulière

$$P(X) = f(0) + f'(0)X + \ldots + \frac{f^{(n)}(0)}{n!} X^n$$

et de reste $f(x) - P(x) = \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1}$ pour un certain c compris entre 0 et x. *(bien sûr c dépend de x)*
- Formule de Taylor-Young : Si $f^{(n)}(0)$ existe, alors f admet un développement limité d'ordre n en 0 de partie régulière

$$P(X) = f(0) + f'(0)X + \ldots + \frac{f^{(n)}(0)X^n}{n!}$$

Remarquez que la deuxième formule a des hypothèses plus faibles, et donne un résultat plus faible aussi puisqu'elle ne permet pas d'évaluer le reste. Les deux formules montrent que, pour calculer le développement limité d'une fonction à l'ordre n en 0, on peut se contenter de calculer ses n dérivées successives en 0 ; en pratique, c'est une très mauvaise méthode dès que n dépasse 2 ou 3, car très lourde en calculs.

On peut sans (trop) risquer de se tromper ajouter, multiplier des développements limités :

Proposition 2.8.
Si f, g admettent des développements limités à l'ordre n en 0, de parties régulières respectives P et Q, alors :

– *$\lambda f + \mu g$ admet un développement limité d'ordre n en 0, de partie régulière $\lambda P + \mu Q$*
– *$f.g$ admet un développement limité d'ordre n en 0, de partie régulière les termes de degré $\leq n$ de $P.Q$.*

Pour calculer le développement limité de $\frac{f}{g}$ dans le cas où $g(0) \neq 0$, on peut également diviser des développements limités selon la méthode des puissances croissantes (voir exemple du DL de tan à la fin de la fiche).

Rappelons un résultat très important : on peut intégrer un développement limité, mais **on ne peut pas dériver un développement limité** en général : il se peut que f admette un développement limité d'ordre n, et que f' n'ait pas de développement limité d'ordre $n-1$.

Proposition 2.9.
Si f est dérivable et f' admet un développement limité d'ordre n en 0 de partie régulière $a_0 + a_1X + \ldots + a_nX^n$, alors f admet un D.L d'ordre $n+1$ en 0, de partie régulière $P(X) = f(0) + a_0X + \frac{a_1}{2}X^2 + \ldots \frac{a_n}{n+1}X^{n+1}$.

Remarquons que, dans l'énoncé ci-dessus, il est primordial de ne pas oublier le terme « $f(x_0)$ », qui est la constante d'intégration !

Il ne reste plus qu'un théorème important à énoncer sur les développement limités : on peut *composer* des développement limités.

Proposition 2.10.
Si f admet un développement limité en x_0 d'ordre n et de partie régulière P, et g admet un développement limité d'ordre n en $f(x_0)$ de partie régulière Q, alors $g \circ f$ admet un développement limité d'ordre n en x_0, de partie régulière constituée par les termes de degré $\leq n$ de $Q \circ P$.

Il y a un piège : il faut veiller à bien utiliser le développement limité de g en $f(x_0)$, et pas en x_0... Il faut aussi faire attention à composer des développements limités de même ordre : si on connaît par exemple le développement limité de f en x_0 à l'ordre 3 en x_0, et le développement limité de g à l'ordre 42 en $f(x_0)$, on n'a assez d'information que pour calculer le développement limité de $g \circ f$ à l'ordre 3 en x_0.

Ce théorème peut être utilisé pour calculer le développement limité d'un quotient $\frac{f}{g}$, où $g(0) \neq 0$: on commence par calculer le développement limité de $\frac{1}{g}$ par composition, puis on le multiplie avec celui de f.
Il est très important de se rappeler qu'un développement limité est une égalité mathématique, pas une identification magique : il faut toujours indiquer le reste et savoir à quel ordre on calcule le développement.

Pour terminer ces notes, mentionnons le cas des fonctions d'une variable réelle **à valeurs complexes**. Tous les énoncés donnés plus haut restent corrects (en considérant des polynômes à coefficients complexes, bien sûr), sauf un : **la formule de Taylor-Lagrange est fausse pour les fonctions à valeurs complexes**.

En effet, la preuve de cette formule se base sur l'égalité des accroissements finis (qui est en fait la formule de Taylor-Lagrange à l'ordre 1) et celle-ci est fausse pour les fonctions à valeurs complexes. Un exemple pour se convaincre : considérons la fonction $f: \mathbb{R} \to \mathbb{C}$ définie par $f(t) = e^{it}$. Alors f est de classe C^∞, et $f'(t) = ie^{it}$ ne s'annule jamais (le module de $f'(t)$ vaut toujours 1). On a $f(2\pi) - f(0) = 0$, donc il est impossible qu'il existe c tel que $f(2\pi) - f(0) = 2\pi f'(c)$. La raison de cette différence entre $\mathbb{R}$ et $\mathbb{C}$ est que $\mathbb{R}$ est muni d'un ordre naturel, compatible avec les opérations algébriques (tout particulièrement, un nombre est positif ssi c'est un carré) ; de plus cet ordre a de nombreuses propriétés (borne sup, borne inf, archimédianité...). Sur $\mathbb{C}$, il n'existe pas de telle relation d'ordre.

Tout cela ne pose pas de problème particulier : sur $\mathbb{C}$, on peut toujours utiliser la formule de Taylor-Young, ainsi que la formule de Taylor avec reste intégral, qu'on reverra plus tard (et qui est à connaître pour les écrits du CAPES !). La théorie serait par contre totalement différente si on considérait des dérivées de fonctions de variable complexe à valeurs complexes - c'est la théorie des fonctions holomorphes.

En analyse, les développements limités sont particulièrement importants pour calculer des limites, mais aussi pour obtenir des équivalents de fonctions, méthode fréquemment utilisée pour étudier la convergence d'une série ou d'une intégrale. Il faut absolument les maîtriser et connaître les développements limités des fonctions classiques, ou savoir les retrouver rapidement.

Quelques développements importants :

- $\dfrac{1}{1+x} = 1 - x + x^2 - \ldots + (-1)^n x^n + o(x^n)$.
- $\ln(1+x) = x - \dfrac{x^2}{2} + \ldots + (-1)^{n+1}\dfrac{x^n}{n} + o(x^n)$.
- $e^x = 1 + x + \dfrac{x^2}{2} + \dfrac{x^3}{6} + \ldots + \dfrac{x^n}{n!} + o(x^n)$.
- $\cos(x) = 1 - \dfrac{x^2}{2} + \ldots + (-1)^n \dfrac{x^{2n}}{(2n)!} + o\left(x^{2n+1}\right)$.
- $\sin(x) = x - \dfrac{x^3}{6} + \ldots + (-1)^n \dfrac{x^{2n+1}}{(2n+1)!} + o\left(x^{2n+2}\right)$.
- $\arctan(x) = x - \dfrac{x^3}{3} + \ldots + (-1)^n \dfrac{x^{2n+1}}{2n+1} + o\left(x^{2n+2}\right)$.
- $(1+x)^\alpha = 1 + \alpha.x + \ldots + \alpha.(\alpha-1)\ldots(\alpha-n+1)\dfrac{x^n}{n!} + o(x^n)$ (valable pour tout $\alpha \in \mathbb{R}$).

Un exemple : trois méthodes pour calculer le développement limité de tan **en** 0 **à l'ordre** 5.

Première méthode : la division selon les puissances croissantes.

On commence par écrire

$$\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)}{1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)} .$$

Puis on utilise la méthode des puissances croissantes :

$$\begin{array}{cccccccc|l}
 & x & - & \frac{x^3}{6} & + & \frac{x^5}{120} & + & o(x^5) & 1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5) \\ \cline{9-9}
- & x & + & \frac{x^3}{2} & - & \frac{x^5}{24} & + & o(x^5) & x + \frac{x^3}{3} + \frac{2}{15}x^5 \\ \cline{1-8}
 & & & \frac{x^3}{3} & - & \frac{x^5}{30} & + & o(x^5) & \\
 & & - & \frac{x^3}{3} & + & \frac{x^5}{6} & + & o(x^5) & \\ \cline{3-8}
 & & & & & \frac{2}{15}x^5 & + & o(x^5) &
\end{array}$$

On obtient donc finalement que $\tan(x) = x + \frac{x^3}{3} + \frac{2}{15}x^5 + o(x^5)$.

Deuxième méthode : par composition, en utilisant le développement limité de $\frac{1}{1-u}$ (ci-dessous, les termes en gris clair sont ceux qu'on aurait pu se passer d'écrire, puisqu'ils font apparaître des termes de degré > 5).

$$\begin{aligned}
\tan(x) &= \frac{x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)}{1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)} \\
&= \frac{x - \frac{x^3}{6} + \frac{x^5}{120}}{1 - (\frac{x^2}{2} - \frac{x^4}{24}} + o(x^5)) \\
&= (x - \frac{x^3}{6} + \frac{x^5}{120})(1 + (\frac{x^2}{2} - \frac{x^4}{24}) + (\frac{x^2}{2} - \frac{x^4}{24})^2 + (\frac{x^2}{2} - \frac{x^4}{24})^3 + (\frac{x^2}{2} - \frac{x^4}{24})^4 + (\frac{x^2}{2} - \frac{x^4}{24})^5) + o(x^5) \\
&= (x - \frac{x^3}{6} + \frac{x^5}{120})(1 + \frac{x^2}{2} + \frac{5x^4}{24}) + o(x^5) \\
&= x + \frac{x^3}{3} + \frac{2x^5}{15} + o(x^5) .
\end{aligned}$$

(Note : dans la dernière ligne, on n'a pas fait apparaître les termes de degré > 5, puisque ce sont tous des $o(x^5)$)

Troisième méthode : en utilisant une équation différentielle. On a $\tan'(x) = 1 + \tan^2(x)$. En écrivant le développement de $\tan'$ en 0 à l'ordre 4 sous la forme $a + bx^2 + cx^4 + o(x^4)$ (il n'y a pas de termes d'ordre impair : tan est impaire, donc $\tan'$ est paire), le fait que $\tan(0) = 0$ et le théorème d'intégration des développements limités nous donne que le développement limité de tan en 0 à l'ordre 5 est égal à $ax + b\frac{x^3}{3} + c\frac{x^5}{5} + o(x^5)$. La formule $\tan'(x) = 1 + \tan^2(x)$ nous donne alors (par composition) :

$$a + bx^2 + cx^4 + o(x^4) = 1 + (ax + b\frac{x^3}{3} + c\frac{x^5}{5})^2 + o(x^4) \ .$$

Autrement dit, on a

$$a + bx^2 + cx^4 + o(x^4) = 1 + a^2x^2 + \frac{2ab}{3}x^4 + \frac{b^2}{9}x^6 + o(x^4)$$

Le théorème d'unicité des développements limités nous permet d'identifier les deux développements terme à terme : ceci donne $a = 1$, $b = a^2 = 1$, et $c = \frac{2ab}{3} = \frac{2}{3}$. En reportant cela dans la formule donnant le développement de tan à l'ordre 5 en 0 en fonction de a, b, c, on obtient de nouveau $\tan(x) = x - \frac{x^3}{3} + \frac{2}{15}x^5 + o(x^5)$.

Chapitre 3

Intégrales impropres

3.1 Définition et premières propriétés

Définition 3.1. Soit I un intervalle de $\mathbb{R}$ et $f\colon I \to \mathbb{C}$. On dit que f est continue par morceaux sur I si, pour tout segment J contenu dans I, la restriction de f à J est continue par morceaux.

Remarque 3.2.
- La restriction de f à J désigne simplement la fonction définie sur J par $x \mapsto f(x)$. On utilise la notation $f_{|J}$ pour la restriction de la fonction f à l'ensemble J.
- Si I est un segment, alors on retrouve la notion de fonction continue par morceaux vue précédemment.
- Si par exemple $I =]0,1]$, et $f\colon x \mapsto \frac{1}{x}$, alors f est continue sur I donc continue par morceaux, mais f ne se prolonge pas en une fonction continue par morceaux sur $[0,1]$, puisque f n'est pas bornée.

Notation. Dans la suite, quand on écrit un intervalle semi-ouvert $[a,b[$, la notation signifie que $a \in \mathbb{R}$ et $b \in [a,+\infty[\cup\{+\infty\}$; de même pour un intervalle $]a,b]$ on autorise $a = -\infty$, et pour un intervalle $]a,b[$ on autorise $a = -\infty$ et $b = +\infty$.

On commence par définir les intégrales impropres sur les intervalles semi-ouverts, i.e $I = [a,b[$ ou $]a,b]$. Comme ces deux cas sont symétriques, dans les démonstrations je traiterai toujours le cas $I = [a,b[$.

Définition 3.3. Soit $I = [a,b[$ un intervalle semi-ouvert de $\mathbb{R}$, et $f\colon I \to \mathbb{C}$ une fonction continue par morceaux. On dit que $\int_a^b f(t)\,dt$ est *convergente* si $\lim_{x\to b^-} \int_a^x f(t)\,dt$ existe.

Dans ce cas, on pose $\int_a^b f(t)\,dt = \lim_{x\to b^-} \int_a^x f(t)\,dt$ et on appelle cette limite *intégrale impropre* (ou *généralisée*) de f sur I.

Bien sûr, on définit symétriquement la convergence d'une intégrale sur $]a,b]$, en considérant $\lim_{x\to a^+} \int_x^b f(t)\,dt$.

Ici, remarquons qu'il y a un possible conflit de notations : si f est une fonction continue par morceaux sur $[a,b]$, alors le symbole $\int_a^b f(t)\,dt$ peut désigner l'intégrale de f comme fonction continue par morceaux sur $[a,b]$, ou alors l'intégrale de f comme fonction continue par morceaux sur $[a,b[$ dont l'intégrale converge. Il nous faut donc vérifier que nos deux notations coïncident quand elles ont toutes les deux un sens, ce pourquoi on énonce et démontre le lemme suivant.

Lemme 3.4. *Soit f une fonction continue par morceaux sur $[a,b]$. Alors $\int_a^b f(t)\,dt = \lim_{x\to b^-} \int_a^x f(t)\,dt$.*

Démonstration. Comme f est continue par morceaux sur $[a,b]$ il existe M tel que $|f(t)| \leq M$ pour tout $t \in [a,b]$, et on a donc

$$\begin{aligned}
\left| \int_a^b f(t)\,dt - \int_a^x f(t)\,dt \right| &= \left| \int_x^b f(t)\,dt \right| \\
&\leq \int_x^b |f(t)|\,dt \\
&\leq \int_x^b M\,dt \\
&= M(b-x)\ .
\end{aligned}$$

Par conséquent on a $\int_a^b f(t)\,dt - \int_a^x f(t)\,dt \to 0$ quand x tend vers b, ce qu'on voulait démontrer. $\square$

Exemple. 1. On a $\int_0^x e^{-t}\,dt = [-e^{-t}]_0^x = 1 - e^{-x}$. On en déduit que $\int_0^{+\infty} e^{-t}\,dt = 1$ (cette notation signifiant que l'intégrale converge et qu'elle vaut 1).

2. $\int_x^1 \frac{dt}{t} = -\ln(x)$ pour x dans $]0,1]$, qui tend vers $+\infty$ quand x tend vers 0. Par conséquent $\int_0^1 \frac{dt}{t}$ diverge.

3. Il ne faut pas croire que $\lim_{t\to 0} f(t) = +\infty$ entraîne que $\int_0^1 f(t)\,dt$ diverge ! Par exemple, en utilisant les mêmes méthodes que dans les deux exemples ci-dessus, vérifiez que $\int_0^1 \frac{dt}{\sqrt{t}} = 2$.

Remarque 3.5. Pour $I = [a,b[$, et $f\colon I \to \mathbb{C}$ une fonction continue par morceaux, dire que $\int_a^b f(t)\,dt$ converge est équivalent à dire que, pour toute suite (x_n) d'éléments de $[a,b[$ qui converge vers b, la suite $\int_a^{x_n} f(t)\,dt$ est convergente.

Que faire maintenant si on étudie une intégrale sur un intervalle ouvert ? Eh bien, on coupe l'intervalle en deux, et on se ramène à l'étude de deux intervalles semi-ouverts.

Définition 3.6. Si $I =]a,b[$ et $f\colon I \to \mathbb{C}$ est continue par morceaux sur I, on dit que $\int_a^b f(t)\,dt$ *converge* s'il existe $x_0 \in I$ tel que $\int_a^{x_0} f(t)\,dt$ **et** $\int_{x_0}^b f(t)\,dt$ convergent.

On pose alors $\int_a^b f(t)\,dt = \int_a^{x_0} f(t)\,dt + \int_{x_0}^b f(t)\,dt$.

Lemme 3.7. *La définition ci-dessus ne dépend pas du choix de x_0, c'est-à-dire que s'il existe $x_0 \in]a,b[$ tel que $\int_a^{x_0} f(t)\,dt$ et $\int_{x_0}^b f(t)\,dt$ convergent toutes les deux, alors :*

- *Pour tout $x \in]a,b[$ les intégrales $\int_a^x f(t)\,dt$ et $\int_x^b f(t)\,dt$ convergent toutes les deux, et*
- $\int_a^{x_0} f(t)\,dt + \int_{x_0}^b f(t)\,dt = \int_a^x f(t)\,dt + \int_x^b f(t)\,dt.$

Démonstration. En exercice, avec la relation de Chasles. □

Maintenant, la notation $\int_0^1 \sin(x)\,dx$, par exemple, peut avoir 4 sens différents : intégrale sur $[0,1]$, sur $[0,1[$, sur $]0,1]$ ou sur $]0,1[$. Comme dans le lemme 3.4, c'est un bon exercice de vérifier que, dès que deux de ces notations sont définies simultanément, elles coïncident.

En pratique, pour déterminer si une intégrale converge sur $]a,b[$, on choisit arbitrairement un point $x_0 \in]a,b[$, et on étudie séparément les deux intégrales sur $]a,x_0]$ et sur $[x_0,b[$.

Remarque 3.8. Une autre façon de présenter les choses serait de dire que, pour une fonction f continue par morceaux sur $]a,b[$, l'intégrale de f sur $]a,b[$ existe si, et seulement si, $\lim_{x\to a^+, y\to b^-} \int_x^y f(t)\,dt$ existe (et l'intégrale est alors égale à cette limite). Cela peut-être utile pour calculer facilement une intégrale généralisée quand on connaît une primitive de f, par exemple : mais cela impose de considérer **séparément** les variables x et y, i.e de considérer une limite d'une fonction de deux variables.

Par exemple, il est tout à fait possible que $\lim_{x\to+\infty} \int_{-x}^x f(t)\,dt$ existe sans que $\int_{-\infty}^{+\infty} f(t)\,dt$ converge : regardez ce qui se passe pour la fonction $f\colon t \mapsto t$ (ou pour n'importe quelle fonction impaire...).

Notation. Si $I =]a,b[$ et f est continue par morceaux sur I, d'intégrale convergente, on pose $\int_b^a f(t)\,dt = -\int_a^b f(t)\,dt$. Si une intégrale ne converge pas on convient que l'autre ne converge pas non plus. On retrouve alors la relation de Chasles pour les intégrales généralisées *convergentes.*

Il est parfois utile, pour étudier la convergence d'une intégrale impropre, d'utiliser un changement de variables. On va énoncer un théorème de changement de variables adapté à ce contexte.

Théorème 3.9 (Formule de changement de variables pour les intégrales impropres)**.** *Soient $a < b$, $c < d$ quatre réels, f une fonction continue sur I et φ une* **bijection de classe** C^1 *de $]c,d[$ sur $]a,b[$. Alors on a :*

- *Si φ est croissante, $\int_a^b f(t)\,dt = \int_c^d f(\varphi(t))\varphi'(t)\,dt$.*
- *Si φ est décroissante, $\int_a^b f(t)\,dt = \int_d^c f(\varphi(t))\varphi'(t)\,dt$*

Remarquons que, comme tout egalité entre intégrales généralisées énoncée dans ce cours, la notation $\int_a^b f(t)\,dt = \int_c^d f(\varphi(t))\varphi'(t)\,dt$ signifie qu'une des intégrales converge si, et seulement si, l'autre converge, et qu'alors elles sont égales. **Il ne faut pas manipuler des intégrales généralisées dans des calculs avant d'avoir prouvé qu'elles convergent !**

Démonstration. On va se contenter de traiter le cas où φ est croissante, l'autre cas étant similaire. Alors, on a $\lim_{x\to a^+} \varphi^{-1}(x) = c$ et $\lim_{x\to b^-} \varphi^{-1}(x) = d$ (si φ était décroissante on aurait $\lim_{x\to a^+} \varphi^{-1}(x) = d$ et $\lim_{x\to b^-} \varphi^{-1}(x) = c$) .

De plus, par le théorème de changement de variables pour des fonctions continues sur un segment, on a, pour tout $x < y \in]a, b[$, que

$$\int_x^y f(t)\,dt = \int_{\varphi^{-1}(x)}^{\varphi^{-1}(y)} f(\varphi(t))\varphi'(t)\,dt \ .$$

Si on suppose que que $\int_c^d f(\varphi(t))\varphi'(t)\,dt$ existe, on en déduit, en faisant tendre x vers a et y vers b, que $\int_a^b f(t)\,dt$ existe et vaut $\int_c^d f(\varphi(t))\varphi'(t)\,dt$.

Pour voir la réciproque, on utilise le même raisonnement, en considérant cette fois-ci $x < y \in [c, d]$, en utilisant la formule de changement de variables sous la forme

$$\int_{\varphi(x)}^{\varphi(y)} f(t)\,dt = \int_x^y f(\varphi(t))\varphi'(t)\,dt$$

et en notant que, quand x tend vers c et y tend vers d, $\varphi(x)$ tend vers a et $\varphi(y)$ tend vers b.

□

Comme toujours quand on étudie des questions de convergence, il est très utile de disposer d'un critère permettant de vérifier que la limite existe sans avoir besoin de calculer la limite explicitement.

Théorème 3.10 (Critère de Cauchy)**.** *Soit $I = [a, b[$ et $f : I \to \mathbb{C}$ une fonction continue par morceaux. Alors $\int_a^b f(t)\,dt$ converge si, et seulement si :*

$$\forall \varepsilon > 0\ \exists c \in [a, b[\ \forall x, y \in [c, b[\ \left| \int_x^y f(t)\,dt \right| \leq \varepsilon \ .$$

Démonstration. Supposons que le critère soit vérifié, fixons une suite (x_n) d'élements de $[a, b[$ qui converge vers b et $\varepsilon > 0$. On trouve $c \in [a, b[$ témoignant du fait que le critère est vérifié ; il existe N tel que $x_n \in [c, b[$ pour tout $n \geq N$. Alors, pour tout $n, m \geq N$, on a

$$\left| \int_a^{x_n} f(t)\,dt - \int_a^{x_m} f(t)\,dt \right| = \left| \int_{x_m}^{x_n} f(t)\,dt \right| \leq \varepsilon \ .$$

Ceci prouve que la suite $\left(\int_a^{x_n} f(t)\,dt\right)$ est de Cauchy, donc convergente. Par conséquent, si le critère est vérifié, alors $\int_a^b f(t)\,dt$ converge.

Supposons maintenant que $\int_a^b f(t)\,dt$ converge, et fixons $\varepsilon > 0$. Alors il existe $c \in [a, b[$ tel que $\left| \int_a^b f(t)\,dt - \int_a^x f(t)\,dt \right| \leq \varepsilon$ pour tout $x \in [c, b[$.

Par conséquent, pour tout $x, y \in [c, b[$, on a

$$\begin{aligned}
\left| \int_x^y f(t)\,dt \right| &= \left| \int_a^x f(t)\,dt - \int_a^b f(t)\,dt + \int_a^b f(t)\,dt - \int_a^y f(t)\,dt \right| \\
&\leq \left| \int_a^x f(t)\,dt - \int_a^b f(t)\,dt \right| + \left| \int_a^b f(t)\,dt - \int_a^y f(t)\,dt \right| \\
&\leq 2\varepsilon \ .
\end{aligned}$$

Ceci étant vrai pour tout $\varepsilon > 0$, on voit que le critère de Cauchy est vérifié. □

Exercice 3.11. Enoncer le critère de Cauchy pour la convergence de l'intégrale d'une fonction continue par morceaux sur $I =]a, b]$.

Définition 3.12. Soit I un intervalle de $\mathbb{R}$ d'extrémités a et b, et $f : I \to \mathbb{C}$ une fonction continue par morceaux. On dit que $\int_a^b f(t)\,dt$ *converge absolument* si $\int_a^b |f(t)|\,dt$ converge.

Théorème 3.13. *Soit I un intervalle de $\mathbb{R}$ d'extrémités a et b, et $f : I \to \mathbb{C}$ une fonction continue par morceaux. Si $\int_a^b f(t)\,dt$ converge absolument alors $\int_a^b f(t)\,dt$ converge, et $\left| \int_a^b f(t)\,dt \right| \leq \int_a^b |f(t)|\,dt$.*

Démonstration. On ne va traiter que le cas $I = [a, b[$: le cas $I =]a, b]$ se traite de la même manière, et le cas $I =]a, b[$ se déduit de la conjonction des deux cas précédents.

Fixons donc $I = [a, b[$ et $\varepsilon > 0$. En apliquant le critère de Cauchy pour $|f|$, on sait qu'il existe $c \in [a, b[$ tel que

$$\forall x < y \in [c, b[\int_x^y |f(t)|\, dt \leq \varepsilon \ .$$

Alors, pour tout $x < y \in [c, b[$ on a, grâce à l'inégalité triangulaire :

$$\left| \int_x^y f(t)\, dt \right| \leq \int_x^y |f(t)|\, dt \leq \varepsilon \ .$$

Par conséquent, $\int_a^b f(t)\, dt$ vérifie le critère de Cauchy et est donc convergente.

Il est maintenant immédiat de vérifier l'inégalité de l'énoncé : pour tout $x, y \in]a, b[$ on a $\left|\int_x^y f(t)\, dt\right| \leq \int_x^y |f(t)|\, dt$; en faisant tendre x vers a et y vers b, on obtient l'inégalité désirée (mais il a d'abord fallu montrer que l'intégrale de f était convergente pour pouvoir faire ce passage à la limite !) ☐

A cause de ce théorème, il est particulièrement important de comprendre la convergence des intégrales impropres de fonctions positives.

3.2 Intégrales impropres de fonctions positives

La théorie des intégrales impropres de fonctions positives se base en grande partie sur l'observation suivante : si $I = [a, b[$ et $f \colon I \to \mathbb{R}^+$ est continue par morceaux, alors la fonction $F \colon x \mapsto \int_a^x f(t)\, dt$ est croissante. Donc deux cas sont possibles :

1. F est bornée sur $[a, b[$. Dans ce cas $\int_a^b f(t)\, dt$ converge.
2. F n'est pas bornée sur $[a, b[$. Dans ce cas $\int_a^b f(t)\, dt$ diverge, et on note $\int_a^b f(t)\, dt = +\infty$ (attention, cette notation n'est définie que pour des fonctions à valeurs positives !).

La même observation est bien sûr valide pour un intervalle de la forme $]a, b]$ et, en découpant, pour un intervalle de la forme $]a, b[$. On en déduit le résultat suivant.

Théorème 3.14. *Soit I un intervalle de $\mathbb{R}$ d'extrémités a et b, f une fonction continue par morceaux sur I à valeurs positives. Alors $\int_a^b f(t)\, dt$ existe si, et seulement si, il existe M tel que $\int_x^y f(t)\, dt \leq M$ pour tous $x < y \in]a, b[$.*

Théorème 3.15 (Premier théorème de comparaison)**.** *Soit I un intervalle de $\mathbb{R}$ d'extrémités a et b, et f, g deux fonctions continues par morceaux sur I, à valeurs positives, telles que $f(x) \leq g(x)$ pour tout $x \in I$. Supposons que $\int_a^b g(t)\, dt$ converge ; alors $\int_a^b f(t)\, dt$ converge.*

Notons que, par contraposée, on obtient que, sous les mêmes hypothèses que ci-dessus, si $\int_a^b f(t)\, dt$ diverge alors $\int_a^b g(t)\, dt$ diverge.

Démonstration. Puisque g est à valeurs positives et $\int_a^b g(t)\, dt$ converge, il existe M tel que $\int_x^y g(t)\, dt \leq M$ pour tous $x < y$ de $]a, b[$. Par positivité et linéarité de l'intégrale, le fait que $f(t) \leq g(t)$ pour tout $t \in I$ entraîne que $\int_x^y f(t)\, dt \leq \int_x^y g(t)\, dt$ pour tous $x < y \in]a, b[$. Par conséquent on a aussi $\int_x^y f(t)\, dt \leq M$ pour tous $x < y$ de $]a, b[$, ce qui prouve, comme f est à valeurs positives, que $\int_a^b f(t)\, dt$ converge. ☐

Remarquons qu'on pourrait énoncer un théorème analogue pour les fonctions à valeurs négatives : ce qui compte, c'est que les fonctions considérées gardent un signe constant. En réalité, il suffit qu'elles gardent un signe constant près des bornes de l'intervalle pour qu'on puisse appliquer ce théorème (en découpant judicieusement l'intervalle d'intégration). Les mêmes observations s'appliquent pour le théorème suivant.

Théorème 3.16 (Second théorème de comparaison)**.** *Soit $I = [a, b[$, et f, g deux fonctions continues par morceaux, à valeurs réelles, gardant un signe constant sur I et telles que $f(x) \sim_{b^-} g(x)$. Alors $\int_a^b f(t)\, dt$ converge si, et seulement si, $\int_a^b g(t)\, dt$ converge.*

Démonstration. Quitte à multiplier f et g par -1, on peut supposer que f et g sont toutes deux positives sur I. Par définition de $\sim$, il existe $c < b$ tel que

$$\forall x \in [c, b[\ \frac{1}{2}f(x) \leq g(x) \leq 2f(x) \ .$$

Fixons un tel c. A l'aide du premier théorème de comparaison, on voit que $\int_c^b f(t)\,dt$ converge si, et seulement si, $\int_c^b g(t)\,dt$ converge. Ceci donne le résultat désiré. $\square$

Remarque 3.17. Dans le second théorème de comparaison, on ne considère qu'un intervalle semi-ouvert ; la raison est que l'information sur l'équivalent n'est donnée qu'en b, par conséquent on ne sait rien sur le comportement de f et g au voisinage de a.

Bien sûr, un théorème similaire est vrai pour $I =]a, b]$, pour des fonctions de signe constant sur I et équivalentes en a (en réalité, par découpage, il suffit que les fonctions soient de signe constant au voisinage de a et équivalentes en a).

3.3 Intégrales impropres et séries

Soit $f\colon [0, +\infty[$ une fonction continue par morceaux. Alors $\int_0^{+\infty} f(t)\,dt$ converge si, et seulement si, la suite $\left(\int_0^{x_n} f(t)\,dt\right)$ converge pour toute suite (x_n) qui tend vers $+\infty$. Puisque la relation de Chasles nous donne

$$\int_0^{x_n} f(t)\,dt = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(t)\,dt \ ,$$

on voit que $\left(\int_0^{x_n} f(t)\,dt\right)$ converge si, et seulement si, la série de terme général $\int_{x_k}^{x_{k+1}} f(t)\,dt$ est convergente pour toute suite (x_n) qui tend vers $+\infty$.

Ce lien entre séries et intégrales est plus intéressant pour les fonctions à valeurs positives : en effet, pour une fonction à valeurs positives, la convergence de $\int_0^{+\infty} f(t)\,dt$ est équivalente au fait qu'il existe *une* suite (x_n) tendant vers $+\infty$ et telle que $\int_0^{x_n} f(t)\,dt$ converge dans $\mathbb{R}$.

Sans hypothèse supplémentaires sur f, on ne peut pas faire mieux. Mais il existe un cas particulier très important.

Théorème 3.18 (Comparaison série-intégrale)**.** *Soit $f\colon [0, +\infty[\to \mathbb{R}$ une fonction continue par morceaux, à valeurs positives et décroissante. Alors $\displaystyle\int_0^{+\infty} f(t)\,dt$ converge si, et seulement si, $\displaystyle\sum_{n=0}^{+\infty} f(n)$ converge.*

Démonstration. Soit n un entier. Comme f est décroissante, pour tout $t \in [n, n+1]$ on a $f(n) \geq f(t) \geq f(n+1)$. Par positivité et linéarité de l'intégrale, on en déduit que

$$\int_n^{n+1} f(n)\,dt \geq \int_n^{n+1} f(t)\,dt \geq \int_n^{n+1} f(n+1)\,dt \ ,$$

autrement dit

$$f(n) \geq \int_n^{n+1} f(t)\,dt \geq f(n+1) \ .$$

En sommant ces inégalités pour n compris entre 0 et N, on obtient que, pour tout entier N, on a

$$\sum_{n=0}^{N} f(n+1) \leq \int_0^{N+1} f(t)\,dt \leq \sum_{n=0}^{N} f(n) \ .$$

Si l'intégrale de f converge, on en déduit que $\displaystyle\sum_{n=0}^{N} f(n+1) = \sum_{n=1}^{N+1} f(n)$ est bornée, donc la série est convergente. De même, si la série est convergente alors on voit que $\int_0^{N+1} f(t)\,dt$ est bornée et donc l'intégrale de f converge (n'oublions pas que la fonction est à valeurs positives !). $\square$

Ce théorème est surtout utile pour déterminer si une série converge en se ramenant à la convergence d'une intégrale. Par exemple, en appliquant ce résultat, il est immédiat que $\displaystyle\sum_{k=0}^{\infty} \frac{1}{k+1}$ diverge.

Chapitre 4

Suites d'intégrales ; intégrales à paramètre

4.1 Convergence uniforme et conséquences

Définition 4.1. Soit I un intervalle de $\mathbb{R}$, et (f_n) une suite de fonctions définies sur I et à valeurs complexes. On dit que (f_n) *converge uniformément sur les segments* vers $f\colon I \to \mathbb{C}$ si, pour tout segment $J \subseteq I$, $(f_{n|J})$ converge uniformément vers $f_{|J}$.

Remarque 4.2.
- Si la suite (f_n) converge uniformément sur I, alors (f_n) converge uniformément sur les segments vers f. La réciproque est fausse en général : par exemple, si $I = [0, +\infty[$, la suite de fonctions (f_n) définie par $f_n(x) = \begin{cases} 0 & \text{si } x \leq n \\ x - n & \text{si } x > n \end{cases}$ converge uniformément sur les segments vers la fonction nulle, mais ne converge pas uniformément vers la fonction nulle.
- La convergence uniforme sur les segments entraîne la convergence simple, mais la réciproque est fausse.

Théorème 4.3 (échange limite-intégrale pour la convergence uniforme sur un segment)**.** *Soit $I = [a, b]$ un* **segment** *de $\mathbb{R}$, et (f_n) une suite de fonctions continues par morceaux sur I à valeurs complexes qui converge* **uniformément** *sur I vers une fonction f continue par morceaux. Alors $\int_a^b f_n(x)\,dx$ converge vers $\int_a^b f(x)\,dx$ quand n tend vers $+\infty$. Autrement dit, dans ce cas on peut échanger limite et intégrale :*

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,dx = \int_a^b \left(\lim_{n\to+\infty} f_n(x) \right) dx$$

Démonstration. La fonction f est supposée continue par morceaux, donc on peut considérer son intégrale sur le segment $[a, b]$, et on doit montrer que $\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx$ tend vers 0 quand n tend vers $+\infty$. Pour cela, fixons $\varepsilon > 0$; il existe $N \in \mathbb{N}$ tel que $|f_n(x) - f(x)| \leq \varepsilon$ pour tout $n \geq N$ et tout $x \in I$. On a alors, pour tout $n \geq N$:

$$\begin{aligned} \left| \int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx \right| &= \left| \int_a^b (f(x) - f_n(x)\,dx \right| \\ &\leq \int_a^b |f(x) - f_n(x|\,dx \\ &\leq \int_a^b \varepsilon\,dx \\ &\leq \varepsilon(b-a) \ . \end{aligned}$$

Ceci prouve bien que $\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx$ tend vers 0 quand n tend vers $+\infty$. □

La démonstration ci-dessus a utilisé de manière essentielle que I était de longueur finie ; dans le cas général, on peut énoncer le résultat suivant, de démonstration aussi élémentaire.

Théorème 4.4. *Soit I un intervalle de $\mathbb{R}$ d'extrémités a et b, et (f_n) une suite de fonctions continues par morceaux sur I à valeurs complexes telle que :*

1. (f_n) *converge uniformément sur les segments vers une fonction continue par morceaux* $f\colon I \to \mathbb{C}$.
2. *Il existe une fonction* g *continue par morceaux sur* I*, d'intégrale finie, et telle que pour tout* $n \in \mathbb{N}$ *et tout* $x \in I$ *on ait* $|f_n(x)| \leq g(x)$.

Alors l'intégrale de f *sur* I *est absolument convergente, et* $\int_a^b f_n(x)\,dx$ *converge vers* $\int_a^b f(x)\,dx$ *quand* n *tend vers* $+\infty$.

Avant de donner la preuve de ce résultat, notons qu'il est plus fort que celui du théorème 4.3 (pourquoi ?), mais on va utiliser celui-ci dans notre preuve.

Démonstration. Pour tout x on a $|f_n(x)| \leq g(x)$ pour tout n, donc en passant à la limite on voit que $|f(x)| \leq g(x)$ donc l'intégrale de f sur I est absolument convergente.

Pour prouver que $\int_a^b f_n(x)\,dx$ converge vers $\int_a^b f(x)\,dx$ quand n tend vers $+\infty$, on traite le cas $I = [a, b[$. Alors, fixons $\varepsilon > 0$; le critère de Cauchy appliqué à g nous donne un $c \in [a, b[$ tel que $\int_c^b g(x)\,dx \leq \varepsilon$. On a alors :

$$\begin{aligned}
\left|\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx\right| &= \left|\int_a^b (f(x) - f_n(x))\,dx\right| \\
&= \left|\int_a^c (f(x) - f_n(x))\,dx + \int_c^b (f(x) - f_n(x))\,dx\right| \\
&\leq \left|\int_a^c (f(x) - f_n(x))\,dx\right| + \left|\int_c^b (f(x) - f_n(x))\,dx\right| \\
&\leq \left|\int_a^c (f(x) - f_n(x))\,dx\right| + \int_c^b |f_n(x)|\,dx + \int_c^b |f(x)|\,dx \\
&\leq \left|\int_a^c (f(x) - f_n(x))\,dx\right| + 2\int_c^b g(x)\,dx \\
&\leq \left|\int_a^c (f(x) - f_n(x))\,dx\right| + 2\varepsilon.
\end{aligned}$$

Puisque la suite des restrictions de f_n à $[a, c]$ converge uniformément vers la restriction de f à $[a, c]$, le théorème précédent nous donne que, pour n suffisamment grand, on a $\left|\int_a^c (f(x) - f_n(x))\,dx\right| \leq \varepsilon$. Par conséquent, pour n suffisamment grand on a $\left|\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx\right| \leq 3\varepsilon$. □

Pour raccourcir un peu les énoncés, on appellera une fonction positive g continue par morceaux sur I et d'intégrale finie une fonction positive *intégrable*. Le théorème ci-dessus peut être un peu surprenant à première vue : la nécessité de l'existence d'une fonction g qui majore toutes les $|f_n|$ ne saute pas aux yeux. Mais le théorème deviendrait faux si l'on ne rajoutait pas cette hypothèse : par exemple, considérons la fonction f_n qui vaut $\frac{1}{n}$ sur $[0, n]$ et 0 ailleurs. Alors (f_n) converge uniformément vers la fonction nulle sur $[0, +\infty[$, pourtant $\int_0^\infty f_n(x)\,dx = 1$ ne tend pas vers 0 quand n tend vers $+\infty$...

En fait, si l'on suppose qu'il existe une fonction g comme ci-dessus (ce qu'on appelle souvent une hypothèse de « domination », au sens où la fonction g domine, ou contrôle, le comportement de la fonction f), on peut considérablement améliorer l'énoncé du théorème 4.4 en remplaçant la convergence uniforme par la convergence simple. C'est le *théorème de convergence dominée*, qu'on verra dans la prochaine section.

4.2 Convergence monotone et convergence dominée

On va énoncer deux théorèmes permettant d'échanger limite et intégrale. Ces deux théorèmes sont difficiles à établir dans le cadre de l'intégrale des fonctions continues par morceaux , et à la fois naturels et plus généraux dans le contexte de l'intégrale de Lebesgue. Ils seraient donc une bonne motivation pour mettre à jour notre version de l'intégrale vers l'intégrale de Lebesgue ; nous ne le ferons pas dans ce cours.

Théorème 4.5 (Théorème de convergence monotone)**.** *Soit* I *un intervalle de* $\mathbb{R}$ *d'extrémités* a, b*, et* (f_n) *une suite de fonctions continues par morceaux sur* I*,* **à valeurs positives***, telles que pour tout* $x \in I$ **la suite** $(f_n(x))$

soit croissante *et converge vers* $f(x)$. *On suppose de plus que* f *est continue par morceaux sur* I. *Alors* f *est intégrable sur* I *si, et seulement si, la suite* $\int_a^b f_n(x)\,dx$ *converge, et alors on a* $\int_a^b f(x)\,dx = \lim \int_a^b f_n(x)\,dx$.

Notons que la suite $\int_a^b f_n(x)\,dx$ est croissante, par conséquent soit elle est convergente (et sa limite est égale à $\int_a^b f(x)\,dx$), soit tend vers $+\infty$ (et alors $\int_I f = +\infty$). Donc, si on admet l'écriture « $+\infty = +\infty$ », la conclusion du théorème s'exprime simplement en disant que, dans tous les cas, $\int_a^b f(x)\,dx = \lim \int_a^b f_n(x)\,dx$.

Théorème 4.6 (Théorème de convergence dominée)**.** *Soit* I *un intervalle de* $\mathbb{R}$, *d'extrémités* a, b *et* (f_n) *une suite de fonctions continues par morceaux sur* I *à valeurs complexes telle que :*

1. (f_n) *converge simplement sur* I *vers une fonction* $f\colon I \to \mathbb{C}$.
2. *Il existe une fonction* g *à valeurs positives et intégrable sur* I *telle que* $|f_n(x)| \leq g(x)$ *pour tout* $x \in I$.

Alors $\int_a^b f_n(x)\,dx$ *converge vers* $\int_a^b f(x)\,dx$ *quand* n *tend vers* $+\infty$.

On va essayer, dans la prochaine section, de donner une preuve [i] du théorème de convergence dominée dans le cadre de l'intégrale définie dans ce cours. Cette partie n'est pas au programme de notre cours, mais constitue une lecture intéressante pour celles et ceux qui souhaitent comprendre les outils qu'elles ou ils manipulent.

4.3 Preuve du théorème de convergence dominée

La preuve « naturelle »du théorème de convergence dominée se fait en utilisant des idées de la théorie de la mesure. Pour les fonctions continues, on peut s'en passer avec un peu de travail. Pour s'en sortir, on a d'abord besoin d'établir le théorème suivant.

Théorème 4.7 (Théorème de Dini)**.** *Soit* I *un segment de* $\mathbb{R}$, *et* (f_n) *une suite de fonctions continues sur* I *telle que pour tout* $x \in I$ *la suite* $f_n(x)$ *soit décroissante vers* 0. *Alors la suite* (f_n) *converge uniformément vers* 0 *sur* I.

Remarque 4.8. Ce théorème permet de voir que, dans certains cas, la convergence simple implique la convergence uniforme. Comme on a déjà montré l'analogue du théorème de convergence dominée pour les fonctions convergeant uniformément, ce résultat va nous être très utile.

Démonstration. On raisonne par l'absurde : si (f_n) ne converge pas uniformément vers 0, alors il existe $\varepsilon > 0$ tel que, pour tout n, il existe $x_n \in I$ tel que $f_n(x_n) \geq \varepsilon$. Notons que, puisque $f_n(x)$ est décroissante pour tout $x \in I$, on a $f_k(x_n) \geq f_n(x_n) \geq \varepsilon$ dès que $k \leq n$.

Grâce au théorème de Bolzano-Weierstrass, on peut trouver une application $\varphi\colon \mathbb{N} \to \mathbb{N}$ strictement croissante et telle que $(x_{\varphi(n)})$ converge vers $x \in I$. Fixons un entier k. Pour tout n suffisamment grand, on a $\varphi(n) \geq k$ et, par continuité de la fonction f_k en x, $|f_k(x) - f_k(x_{\varphi(n)})| \leq \frac{\varepsilon}{2}$ pour n grand. Pour n suffisamment grand, on a donc

$$f_k(x) \geq f_k(x_{\varphi(n)}) - \frac{\varepsilon}{2} \geq \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}\,.$$

Ceci montre que la suite $(f_k(x))$ ne tend pas vers 0, contredisant l'hypothèse du théorème. □

On va maintenant montrer un cas particulier du théorème de convergence dominée, à partir duquel il sera relativement facile d'obtenir le théorème général.

Théorème 4.9 (Cas particulier du théorème de convergence dominée)**.** *Soit* $I = [a, b]$ *un segment de* $\mathbb{R}$, *et* (f_n) *une suite de fonctions continues sur* I, *à valeurs positives, telles que :*

1. (f_n) *converge simplement sur* I *vers la fonction nulle.*
2. *Il existe une constante* M *telle que* $|f_n(x)| \leq M$ *pour tout* n *et tout* $x \in I$.

Alors $\int_a^b f_n(x)\,dx$ *converge vers* 0.

i. basée sur des notes de M. Alain Frisch disponibles sur Internet à l'adresse `http://alain.frisch.fr/math/tcd.ps`

Démonstration. Fixons $\varepsilon > 0$. Etant donnés deux entiers n, k et $x \in I$, on pose

$$g_{n,k}(x) = \max(f_n(x), f_{n+1}(x), \ldots, f_{n+k}(x)) \ .$$

Pour un n fixé, la suite $(g_{n,k}(x))_k$ est croissante pour tout $x \in I$; de plus, $g_{n,k}$ est à valeurs positives et majorée par M. Par conséquent (toujours à n fixé) la suite $\int_a^b g_{n,k}(x)\,dx$ converge quand k tend vers $+\infty$, et on peut donc trouver k_n tel que

$$\forall k \in \mathbb{N} \quad \int_a^b g_{n,k}(x)\,dx \leq \int_a^b g_{n,k_n}(x)\,dx + \frac{\varepsilon}{2^n} \ .$$

On fixe un tel k_n, et on pose $g_n(x) = g_{n,k_n}(x)$. On peut également faire en sorte que la suite (k_n) soit croissante, ce qu'on fait dans la suite.

Faisons une liste des propriétés de la suite (g_n) que nous allons utiliser :

1. Chaque g_n est une fonction continue.
2. $g_n(x) \geq f_n(x)$ pour tout n et tout $x \in I$.
3. $\max(g_n(x), g_{n+1}(x)) = g_{n,k_{n+1}}(x)$ (ici on utilise que (k_n) est croissante), ce dont on déduit que
$$\int_a^b \max(g_n(x), g_{n+1}(x))\,dx \leq \int_a^b g_{n,k_n}(x)\,dx + \frac{\varepsilon}{2^n} = \int_a^b g_n(x)\,dx + \frac{\varepsilon}{2^n} \ .$$
4. Pour tout $x \in I$ la suite $g_n(x)$ tend vers 0.

Maintenant, on voudrait bien que (g_n) soit une suite décroissante, pour pouvoir lui appliquer le théorème de Dini... Hélas, ce n'est pas tout à fait le cas. On introduit une nouvelle suite de fonctions, en posant pour tout n et tout $x \in I$

$$h_n(x) = \min(g_0(x,), \ldots, g_n(x)) \ .$$

Pour tout x la suite $h_n(x)$ décroît vers 0 ; de plus chaque h_n est continue. Grâce au théorème de Dini, on peut donc conclure que la suite (h_n) converge uniformément vers 0 sur I et donc, puisque I est un segment, $\int_a^b h_n(x)\,dx$ converge vers 0. On en déduit qu'il existe $N \in \mathbb{N}$ tel que $\int_a^b h_n(x)\,dx \leq \varepsilon$ pour tout $n \geq N$.

Notons que, pour tout $x \in I$ et tout n, on a

$$h_{n+1}(x) = \min(h_n(x), g_{n+1}(x)) = h_n(x) + g_{n+1}(x) - \max(h_n(x), g_{n+1}(x)) \ .$$

Donc, pour tout $x \in I$ et tout n, on a

$$0 \leq g_{n+1}(x) - h_{n+1}(x) = \max(h_n(x), g_{n+1}(x)) - h_n(x) \leq \max(g_n(x), g_{n+1}(x)) - h_n(x) \ .$$

Finalement, en utilisant l'inégalité obtenue en (3) ci-dessus, on voit que

$$0 \leq \int_a^b (g_{n+1}(x) - h_{n+1}(x))\,dx \leq \int_a^b (g_n(x) - h_n(x)) + \frac{\varepsilon}{2^n} \ .$$

On en déduit, en utilisant le fait que $g_0 = h_0$ et la formule donnant la somme des termes d'une série géométrique, que, pour tout n,

$$\int_a^b (g_n(x) - h_n(x)) \leq \varepsilon(1 - \frac{1}{2^n}) \leq \varepsilon \ .$$

En mettant tout cela ensemble, on obtient finalement que, pour tout $n \geq N$, on a

$$\int_a^b f_n(x)\,dx \leq \int_a^b g_n(x)\,dx \leq \int_a^b (g_n(x) - h_n(x))\,dx + \int_a^b h_n(x)\,dx \leq 2\varepsilon \ .$$

La suite $(\int_a^b f_n(x)\,dx)$ converge donc vers 0, ce qu'on voulait démontrer. □

On a fait le plus gros du travail.

Preuve du théorème de convergence dominée pour une suite de fonctions continues, à valeurs réelles. On se place dans le cas où $I = [a, b[$ et (f_n) est une suite de fonctions continues, à valeurs réelles, majorée sur I par une fonction g intégrable et convergeant simplement vers une fonction f continue par morceaux sur I.

Comme $|f| \leq g$ sur I, f est intégrable. Fixons $\varepsilon > 0$. Il existe $c \in [a, b[$ tel que $\int_c^b g(x)\,dx \leq \varepsilon$.

Sur $[a, c]$, la suite de fonctions $|f_n - f|$ est majorée par $2g$, qui est bornée sur $[a, c]$ car réglée. Donc il existe M tel que $|f_n(x) - f(x)| \leq M$ pour tout x de $[a, c]$. De plus, $|f_n - f|$ converge simplement vers 0 sur $[a, c]$. On peut donc appliquer le théorème qu'on vient de démontrer pour conclure que $\int_a^c |f_n(x) - f(x)|\,dx$ tend vers 0 quand n tend vers $+\infty$. En particulier, il existe $N \in \mathbb{N}$ tel que $\int_a^c |f_n(x) - f(x)|\,dx \leq \varepsilon$ pour tout $n \geq N$.

Alors on obtient, pour tout $n \geq N$:

$$\begin{aligned}
\left|\int_a^b f_n(x)\,dx - \int_a^b f(x)\,dx\right| &\leq \left|\int_a^c f_n(x)\,dx - \int_a^c f(x)\,dx\right| + \left|\int_c^b f_n(x)\,dx - \int_c^b f(x)\,dx\right| \\
&\leq \int_a^b |f_n(x) - f(x)|\,dx + \left|\int_c^b f_n(x)\,dx\right| + \left|\int_c^b f(x)\,dx\right| \\
&\leq \int_a^b |f_n(x) - f(x)|\,dx + +2\left|\int_c^b g(x)\,dx\right| \\
&\leq 3\varepsilon.
\end{aligned}$$

□

Le cas des suites de fonctions continues à valeurs complexes s'en déduit facilement, en décomposant les fonctions en partie réelle et partie imaginaire.

Le théorème général, pour les fonctions continues par morceaux , se déduit à partir de ce qu'on a fait en utilisant le lemme suivant, dont la preuve est laissée en exercice (ainsi que la fin de la preuve du théorème de convergence dominée - si vous avez compris cette section, la fin de la preuve ne devrait pas poser trop de problèmes...).

Lemme 4.10. *Soit $I = [a, b]$ un segment de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction réglée. Alors, pour tout $\varepsilon > 0$ il existe une fonction g continue sur I telle que $g \leq f$ sur I et $\int_a^b g(x)\,dx \geq \int_a^b f(x)\,dx - \varepsilon$.*

4.4 Echanges série-intégrale

Les théorèmes d'échange limite-intégrale vus précédemment peuvent se reformuler comme des théorèmes sur des séries de fonctions : étudier une série de fonctions $\sum_{k=0}^{+\infty} f_k$ revient à essayer de comprendre le comportement de la suite $(\sum_{k=0}^{n} f_k)$ quand n tend vers $+\infty$.

En particulier, notre théorème sur la convergence uniforme sur un segment d'une suite de fonctions continues devient :

Théorème 4.11. *Soit $I = [a, b]$ un* **segment** *de $\mathbb{R}$, et $f_n : I \to \mathbb{C}$ une suite de fonctions continues par morceaux telle que $\sum_{n=0}^{+\infty} f_n$ converge* **uniformément** *vers f continue par morceaux sur I. Alors, $\sum_{n=0}^{+\infty} \int_a^b f_n(x)\,dx$ converge et on a*

$$\int_a^b f(x)\,dx = \sum_{n=0}^{+\infty} \int_a^b f_n(x)\,dx \text{ , ou encore } \int_a^b \left(\sum_{n=0}^{+\infty} f_n(x)\right)\,dx = \sum_{n=0}^{+\infty} \int_a^n f_n(x)\,dx \text{ .}$$

Remarque 4.12. Si les (f_n) sont supposées continues, alors f est automatiquement continue en tant que limite uniforme d'une série de fonctions continues.

Le théorème de convergence monotone s'applique au cas des séries de fonctions positives, donnant le résultat suivant.

Théorème 4.13. *Soit I un intervalle de $\mathbb{R}$ d'extrémités a, b, et (f_n) une suite de fonctions continues par morceaux* **à valeurs positives** *telle que $\sum_{n=0}^{+\infty} f_n$ converge simplement vers une fonction f continue par morceaux.*

Alors $\sum_{n=0}^{+\infty}\left(\int_a^b f_n(x)\,dx\right)$ *converge si, et seulement si, l'intégrale* $\int_a^b f(x)\,dx$ *est convergente, et dans ce cas on a*

$$\int_a^b f(x)\,dx = \sum_{n=0}^{+\infty}\left(\int_a^b f_n(x)\,dx\right) \text{ , ou encore } \int_a^b \left(\sum_{n=0}^{+\infty} f_n(x)\right) dx = \sum_{n=0}^{+\infty}\int_a^b f_n(x)\,dx \ .$$

Exercice 4.14. Montrer les deux théorèmes ci-dessus, en utilisant les théorèmes correspondants pour les suites de fonctions.

Enfin, le théorème de convergence dominée a une conséquence particulièrement importante en pratique ; la démonstration de ce théorème en utilisant la version de l'intégrale que l'on a développée dans ce cours est difficile et on ne la traitera pas.

Théorème 4.15. *Soit* I *un intervalle de* $\mathbb{R}$ *d'extrémités* a, b *et* (f_n) *une suite de fonctions de* I *dans* $\mathbb{C}$*, continues par morceaux, telle que* $\sum_{n=0}^{+\infty} f_n$ *converge simplement sur* I *vers une fonction* f *continue par morceaux. On suppose que la série* $\sum_{n=0}^{+\infty}\int_a^b |f_n(x)|\,dx$ *converge. Alors l'intégrale de* f *sur* I *converge absolument, et on a*

$$\int_a^b f(x)\,dx = \sum_{n=0}^{+\infty}\left(\int_a^b f_n(x)\,dx\right) \text{ , ou encore } \int_a^b \left(\sum_{n=0}^{+\infty} f_n(x)\right) dx = \sum_{n=0}^{+\infty}\int_a^n f_n(x)\,dx \ .$$

Remarque 4.16. Si on fait l'hypothèse supplémentaire que la suite de fonctions $\sum |f_n|$ converge vers une fonction continue par morceaux, alors on peut déduire le théorème ci-dessus des théorèmes de convergence monotone et de convergence dominée qu'on a vus précédemment. Une difficulté technique de la démonstration est qu'il pourrait y avoir des points où la suite $\sum |f_n|$ diverge, ou bien il se pourrait que la limite ne soit pas continue par morceaux ; ces difficultés se lèvent relativement facilement si l'on utilise la théorie de la mesure, mais posent un vrai problème dans le cadre de notre version de l'intégrale...

4.5 Intégrales à paramètre

Dans cette section, on veut considérer des fonctions définies par une intégrale, et étudier la continuité/la dérivabilité de ces fonctions. Considérons un exemple : la fonction F définie sur $\mathbb{R}$ par

$$F(x) = \int_0^{+\infty} \frac{\sin(xt)}{t} dt \ .$$

Il est clair que $F(0) = 0$. Il n'est pas clair que $F(x)$ soit bien définie pour $x \neq 0$; par exemple, si $x > 0$, le changement de variable $u = xt$ donne

$$F(x) = \int_0^{+\infty} \frac{\sin(u)}{u} du = \frac{\pi}{2} \text{ (exemple vu en cours) .}$$

Si $x < 0$, le changement de variables $u = -x$ donne

$$F(x) = \int_0^{+\infty} \frac{\sin(-u)}{u} du = -\frac{\pi}{2} \ .$$

La fonction $f\colon (x,t) \mapsto \frac{\in(xt)}{t}$ est continue sur $\mathbb{R}\times]0,+\infty[$, $F(x) = \int_0^{+\infty} f(x,t)\,dt$ est bien définie pour tout $x \in \mathbb{R}$... mais n'est pas du tout continue puisqu'elle ne prend que 3 valeurs. Etant donné ce qu'on a vu dans les sections précédentes, il n'est pas étonnant qu'on doive introduire une hypothèse de domination pour obtenir la continuité d'une intégrale à paramètres.

Théorème 4.17 (Théorème de continuité des intégrales à paramètre)**.** *Soit* I *un intervalle de* $\mathbb{R}$ *d'extrémités* a, b*,* J *un intervalle de* $\mathbb{R}$ *et* $f\colon J \times I \to \mathbb{C}$ *une fonction telle que :*

1. *Pour tout* $x \in J$*,* $t \mapsto f(x,t)$ *est continue par morceaux sur* I*.*
2. *Pour tout* $t \in I$*,* $x \mapsto f(x,t)$ *est continue sur* J*.*
3. *Il existe une fonction* g *intégrable à valeurs positives et telle que* $|f(x,t)| \leq g(t)$ *pour tout* $(x,t) \in J \times I$*.*

Alors la fonction $F\colon x \mapsto \int_a^b f(x,t)\,dt$ *est bien définie et continue sur* J*.*

Remarque 4.18. Les deux premières hypothèses ci-dessus sont automatiquement satisfaites si f est continue sur $J \times I$. Par ailleurs, on peut affaiblir l'hypothèse de domination : pour obtenir la conclusion, il suffit que, pour tout segment J' contenu dans J, il existe une fonction g positive et intégrable telle que $|f(x,t)| \leq g(t)$ pour tout $x \in J'$ et tout $t \in I$. En effet, le théorème ci-dessus appliqué sur J' montre que F est continue sur J' ; comme J est une union de segments, cela montre que F est continue sur J'.

Démonstration. Pour voir que F est bien définie sur J, il suffit de remarquer que, puisque $t \mapsto f(x,t)$ est continue par morceaux et $|f(x,t)| \leq g(t)$, l'intégrale $\int_a^b f(x,t)\,dt$ est absolument convergente, donc convergente.

Soit maintenant (x_n) une suite d'éléments de J qui converge vers $x \in J$. Alors la suite de fonctions (h_n) définie par $h_n(t) = f(x_n,t)$ converge simplement vers $t \mapsto f(x,t)$ puisque, à t fixé, $x \mapsto f(x,t)$ est continue ; et $|h_n(t)| \leq g(t)$ pour tout $t \in I$. Le théorème de convergence dominée nous donne alors

$$\lim_{n \to +\infty} \int_a^b h_n(t)\,dt = \int_a^b f(x,t)\,dt \ .$$

Autrement dit, $\lim(F(x_n)) = F(x)$ pour toute suite x_n qui converge vers $x \in J$: on vient de prouver que F est continue sur J. □

On pourrait donner une preuve d'un cas un peu plus élémentaire de ce théorème (en supposant f continue sur $J \times I$) « à la main », en n'utilisant pas le théorème de convergence dominée.

Dans le cas où I est un segment (et où on ne considère donc pas d'intégrales généralisées) et où f est continue, on peut se passer de l'hypothèse de domination (qui est en fait automatiquement satisfaite).

Corollaire 4.19. Soit $I = [a,b]$ un segment de $\mathbb{R}$, J un intervalle de $\mathbb{R}$ et $f \colon J \times I \to \mathbb{C}$ une fonction continue sur $J \times I$. Alors la fonction $F \colon x \mapsto \int_a^b f(x,t)\,dt$ est continue sur J.

Démonstration. Commençons par le cas où J est aussi un segment. Alors la fonction f, étant continue sur $J \times I$ qui est compact (produit de deux segments : c'est un rectangle dans le plan), est bornée sur $J \times I$. Par conséquent, il existe M tel que $|f(x,t)| \leq M$ pour tout $(x,t) \in J \times I$, et on peut appliquer le théorème précédent avec la fonction $g \colon t \mapsto M$, qui est intégrable sur I, pour conclure que F est continue.

Le raisonnement ci-dessus montre que, pour tout segment $J' \subseteq J$, la restriction de f à J' est continue ; comme J est une réunion de segments, cela prouve que f est continue sur J. □

On peut, de manière analogue, établir un théorème de dérivabilité des intégrales à paramètre. Rappelons avant cela que, si f est une fonction de deux variables, la notation $\frac{\partial f}{\partial x}(x_0,t_0)$ désigne la dérivée de la fonction $x \mapsto f(x,t_0)$ au point x_0.

Théorème 4.20 (Théorème de dérivabilité des intégrales à paramètre)**.** *Soit I un intervalle de $\mathbb{R}$ d'extrémités a,b, J un intervalle de $\mathbb{R}$ et $f \colon J \times I \to \mathbb{C}$ une fonction telle que :*

1. *Pour tout $x \in J$ la fonction $t \mapsto f(x,t)$ est continue par morceaux sur I.*
2. *Il existe $x_0 \in J$ tel que $\int_a^b f(x_0,t)\,dt$ converge.*
3. *$\frac{\partial f}{\partial x}$ existe sur $J \times I$.*
4. *Pour tout x dans J, $t \mapsto \frac{\partial f}{\partial x}(x,t)$ est continue par morceaux sur I.*
5. *Il existe g intégrable et à valeurs positives telle que $\left|\frac{\partial f}{\partial x}(x,t)\right| \leq g(t)$ pour tout $(x,t) \in J \times I$.*

Alors la fonction $F \colon x \mapsto \int_a^b f(x,t)\,dt$ est bien définie sur J, dérivable, et pour tout $x \in J$ on a

$$F'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)\,dt \ .$$

Remarque 4.21. Il est notable que l'on n'ait besoin que de supposer la convergence de F en un point pour obtenir la convergence de F sur J tout entier. Par ailleurs, notons que la quatrième hypothèse ci-dessus est vérifiée si jamais $\frac{\partial f}{\partial x}$ existe et est continue sur $J \times I$, ce qui sera le cas dans la plupart de nos exemples. Dans ce cas, la fonction F est de classe C^1, puisque le théorème de continuité des intégrales à paramètre s'applique pour prouver que sa dérivée est continue sur J.

Par ailleurs, ici encore, on pourrait affaiblir l'hypothèse de domination de $\frac{\partial f}{\partial x}$ en demandant simplement que, pour tout segment J' contenu dans J, il existe g positive et intégrable sur I telle que $|\frac{\partial f}{\partial x}(x,t)| \leq g(t)$ pour tout $(x,t) \in J' \times I$, et obtenir la même conclusion.

Démonstration. Il nous faut commencer par montrer que $F(x)$ est bien définie sur J. Pour $t \in I$ fixé, l'inégalité des accroissements finis, appliquée à la fonction $x \mapsto f(x,t)$, montre que l'on a

$$|f(x,t) - f(x_0,t)| \le |x - x_0| \sup_{x \in J} \left(\left| \frac{\partial f}{\partial x}(x,t) \right| \right) \le |x - x_0| g(t) \ .$$

Fixons $x \in J$. L'inégalité ci-dessus montre que $\int_a^b (f(x,t) - f(x_0,t))\, dt$ converge, par conséquent $\int_a^b f(x,t)\, dt$ converge puisque c'est la somme de deux intégrales convergentes.

Ensuite, toujours pour $x \in J$ fixé, prenons une suite (x_n) qui tend vers x (avec $x_n \neq x$) et considérons la suite de fonctions

$$h_n \colon t \mapsto \frac{f(x,t) - f(x_n,t)}{x - x_n} \ .$$

Pour un $t \in I$ fixé, l'inégalité des accroissements finis applliquée comme ci-dessus montre que $|h_n(t)| \le g(t)$. De plus, quand n tend vers $+\infty$, $h_n(t)$ converge vers $\frac{\partial f}{\partial x}(x,t)$ (c'est un taux d'accroissement).

Autrement dit, (h_n) converge simplement vers $t \mapsto \frac{\partial f}{\partial x}(x,t)$; on peut appliquer le théorème de convergence dominée, et obtenir que

$$\lim_{n \to +\infty} \int_a^b h_n(t)\, dt = \int_a^b \frac{\partial f}{\partial x}(x,t) dt \ .$$

On vient de montrer que, pour toute suite x_n convergeant vers x (avec $x_n \neq x$ pour tout n), on a

$$\lim_{n \to +\infty} \frac{F(x) - F(x_n)}{x - x_n} = \int_a^b \frac{\partial f}{\partial x}(x,t) dt \ .$$

Cela revient à dire que F est dérivable et que

$$F'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)\, dt \ .$$

Notons encore ici que, si on avait supposé $x \mapsto \frac{\partial f}{\partial x}(x,t)$ continue pour tout $t \in I$, on obtiendrait que F est de classe C^1 puisque le théorème de continuité des intégrales à paramètre, appliqué à la formule ci-dessus pour F', montrerait que F' est continue. □

Pour clore ce chapitre, notons que, si I est un segment et $\frac{\partial f}{\partial x}$ est continue sur $J \times I$, on peut se passer de l'hypothèse de domination sur $\frac{\partial f}{\partial x}$ (qui est en fait automatiquement satisfaite sur tout segment de J, exactement comme pour le théorème de continuité des intégrales à paramètre)

Corollaire 4.22. Soit I un segment de $\mathbb{R}$ d'extrémités a, b, J un intervalle de $\mathbb{R}$ et $f \colon J \times I \to \mathbb{C}$ une fonction telle que :

1. Pour tout $x \in J$ la fonction $t \mapsto f(x,t)$ est continue par morceaux sur I.
2. Il existe $x_0 \in J$ tel que $\int_a^b f(x_0,t)\, dt$ converge.
3. $\frac{\partial f}{\partial x}$ existe et est continue sur $J \times I$.

Alors la fonction $F \colon x \mapsto \int_a^b f(x,t)\, dt$ est bien définie sur J, de classe C^1, et pour tout $x \in J$ on a

$$F'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)\, dt \ .$$

Démonstration. Exercice (recommandé). □

Concluons ce chapitre par un exemple : pour $x \ge 0$, on pose

$$F(x) = \int_0^1 \frac{e^{-x^2(1+t^2)}}{t^2+1}\, dt \quad \text{et} \quad G(x) = \left(\int_0^x e^{-t^2}\, dt \right)^2 \ .$$

A x fixé, $G(x)$ est le carré de l'intégrale d'une fonction continue sur un segment, donc G est bien définie ; de plus, le théorème fondamental de l'analyse, et la formule de dérivation d'un produit, donnent que G est dérivable sur $\mathbb{R}$ et

$$G'(x) = 2e^{-x^2} \int_0^x e^{-t^2}\, dt = 2 \int_0^x e^{-x^2-t^2}\, dt.$$

Comme la fonction $(x,t) \mapsto f(x,t) = \frac{e^{-x^2(1+t^2)}}{t^2+1}$ est continue sur $\mathbb{R} \times [0,1]$, on peut affirmer grâce au théorème de continuité des intégrales à paramètre que la fonction F est bien définie et continue. De plus, on a

$$\frac{\partial f}{\partial x}(x,t) = -2x(1+t^2)\frac{e^{-x^2(1+t^2)}}{t^2+1} = -2xe^{-x^2(1+t^2)} \ .$$

Cette fonction est continue sur $\mathbb{R} \times [0,1]$, donc on peut appliquer le théorème de dérivation des intégrales à paramètre et obtenir que F est dérivable et

$$F'(x) = \int_0^1 -2xe^{-x^2(1+t^2)}\, dt$$

Pour $x > 0$, on peut faire le changement de variables $u = xt$ dans cette intégrale, ce qui donne

$$F'(x) = -2\int_0^x e^{-x^2-u^2}\, du = -G'(x) \ .$$

Par conséquent, la fonction $F+G$ est de dérivée nulle sur $]0,+\infty[$, donc constante sur $[0,+\infty[$. Comme $F(0) = \frac{\pi}{4}$ et $G(0) = 0$, on a $F(x) + G(x) = \frac{\pi}{4}$ pour tout $x \in [0,+\infty[$.

L'intégrale $I = \int_0^{+\infty} e^{-t^2}\, dt$ converge, donc la limite de $G(x)$ quand x tend vers $+\infty$ existe et vaut I^2. Pour déterminer la limite de F en $+\infty$, considérons une suite x_n qui tend vers $+\infty$. On a

$$F(x) = \int_0^1 \frac{e^{-x_n^2(1+t^2)}}{t^2+1}\, dt \ .$$

La suite de fonctions $f_n\colon t \mapsto \frac{e^{-x_n^2(1+t^2)}}{t^2+1}$ converge simplement vers 0 quand n tend vers $+\infty$; de plus, on a $0 \leq f_n(t) \leq \frac{1}{1+t^2}$, qui est intégrable sur $[0,1]$. Par conséquent on peut appliquer le théorème de convergence dominée et obtenir que $\lim_{n\to+\infty} \int_0^1 f_n(t)\, dt = 0$. Autrement dit, pour toute suite (x_n) qui tend vers $+\infty$, on a $\lim_{n\to+\infty} F(x_n) = 0$, ce qui revient à dire que $\lim_{x\to+\infty} F(x) = 0$.

De tout cela, on déduit que $I^2 = \frac{\pi}{4}$, ou encore

$$\int_0^{+\infty} e^{-t^2}\, dt = \frac{\sqrt{\pi}}{2} \ .$$

Cette intégrale est appelée « intégrale de Gauss »et est importante en probabilités [ii].

ii. voir par exemple `http://fr.wikipedia.org/wiki/Intégrale_de_Gauss`

Chapitre 5

Fonctions de plusieurs variables

On va maintenant considérer des fonctions de plusieurs variables réelles, à valeurs dans un espace vectoriel réel de dimension finie. Autrement dit, on va étudier des fonctions définies sur une partie U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^m$. On va commencer par rappeler quelques notions élémentaires de topologie en dimension finie ; toutes ces notions ont été vues dans l'UE « Topologie élémentaire », donc il s'agit simplement de révisions que nous ne ferons pas (ou peu) en cours. De même, on passera vite en cours sur les propriétés des fonctions continues de plusieurs variables, ainsi que les définitions élemntaires liées à la différentiabilité.

5.1 Rappels de topologie en dimension finie

Définition 5.1. Soit n un entier. Une *norme* sur $\mathbb{R}^n$ est une application N telle que :

1. $\forall x \in \mathbb{R}^n \; N(x) \geq 0$. (Positivité de la norme)
2. $\forall x \in \mathbb{R}^n \; N(x) = 0 \Leftrightarrow x = 0$. (Axiome de séparation)
3. $\forall x \in \mathbb{R}^n \; \forall \lambda \in \mathbb{R} \; N(\lambda x) = |\lambda| N(x)$. (Homogénéité de la norme)
4. $\forall x, y \in \mathbb{R}^n \; N(x+y) \leq N(x) + N(y)$. (Inégalité triangulaire)

Intuitivement, une norme sert à mesurer la longueur des vecteurs : une longueur est positive, le seul vecteur dont la longueur est nulle est le vecteur nul, la longueur de λx est $|\lambda|$ fois la longueur de x, et la longueur de la somme de deux vecteurs est plus petite que la somme de leurs longueurs.

On ne va pas manipuler de normes ésotériques dans ce cours, mais il est important de bien comprendre cette notion abstraite. Rappelons quelques exemples.

Exemple.

– La *norme euclidienne* sur $\mathbb{R}^n$ est définie par la formule suivante :

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \; .$$

– La norme 1 est définie par

$$\|(x_1, \ldots, x_n)\|_1 = \sum_{i=1}^{n} |x_i| \; .$$

– Enfin, la norme infinie est définie par

$$\|(x_1, \ldots, x_n)\|_\infty = \max_{i=1,\ldots,n} |x_i| \; .$$

Théorème 5.2 (Théorème de Bolzano-Weierstrass pour $\|\cdot\|_\infty$)**.** *Soit $x_i = (x_1^i, \ldots, x_n^i)$ une suite d'éléments de $\mathbb{R}^n$ telle que pour tout $k \in \{1, \ldots, n\}$ la suite (x_k^i) soit bornée. Alors il existe une application strictement croissante $\psi \colon \mathbb{N} \to \mathbb{N}$ et $x \in \mathbb{R}^n$ tel que $\|x - x_{\psi(i)}\|_\infty$ tende vers 0.*

Démonstration. Pour tout $k \in \{1, \dots, n\}$ fixé, la suite x_k^i est une suite bornée de réels, et on peut lui appliquer le théorème de Bolzano-Weierstrass pour les suites réelles. On commence par trouver $\varphi_1 \colon \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $(x_1^{\varphi_1(i)})$ converge vers x_1, puis φ_2 tel que $(x_2^{\varphi_1(\varphi_2(i))})$ converge vers x_2, etc., jusqu'à φ_n telle que $(x_n^{\varphi_1(\varphi_2 \dots \varphi_n((i)))})$ converge vers x_n. Ensuite, on pose $\psi(i) = \varphi_1(\varphi_2(\dots(\varphi_n(i))))$. La fonction $\psi \colon \mathbb{N} \to \mathbb{N}$ est strictement croissante, et par construction on a que $x_k^{\psi(i)}$ converge vers x_k pour tout $k \in \{1, \dots, n\}$. En posant $x = (x_1, \dots, x_n)$, ceci est équivalent à dire que $\|x - x_{\psi(i)}\|_\infty$ tend vers 0. □

Dans la suite, on utilisera principalement la norme infinie. Pour ce que nous souhaitons faire, le choix de norme ne sera pas fondamental, à cause du théorème suivant.

Théorème 5.3 (Théorème d'équivalence des normes en dimension finie)**.** *Soit $n \geq 1$ un entier, et N_1, N_2 deux normes sur $\mathbb{R}^n$. Alors N_1 et N_2 sont* équivalentes, *c'est-à-dire qu'il existe deux réels strictement positifs m, M tels que*

$$\forall x \in \mathbb{R}^n \; mN_1(x) \leq N_2(x) \leq MN_1(x) \; .$$

Démonstration. Soit N une norme sur R^n. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, l'inégalité triangulaire appliquée à N nous donne

$$N(x) = N\left(\sum_{k=1}^{n} x_k e_k\right) \leq \sum_{k=1}^{n} |x_k| N(e_k) \leq \|x\|_\infty \sum_{k=1}^{n} N(e_k) \; .$$

En posant $M = \sum_{k=1}^{n} N(e_k)$, on vient de montrer que $N(x) \leq M\|x\|_\infty$ pour tout $x \in \mathbb{R}^n$.

Soit maintenant $A = \{x \in \mathbb{R}^n \colon \|x\|_\infty = 1\}$, et $m = \inf\{N(x) \colon x \in A\}$. Il existe une suite (x_i) d'eléments de A tels que $N(x_i)$ tende vers M. Par le théorème de Bolzano-Weierstrass pour $\|\cdot\|_\infty$, on peut trouver une suite extraite $(x_{\varphi(i)})$ telle que $\|x_{\varphi(i)} - x\|$ tende vers 0. Il est facile de vérifier que $x \in A$ (en particulier x est non nul), et de plus on a $N(x_{\varphi(i)} - x) \leq M\|x_i - x\|_\infty|$donc $N(x_{\varphi(i)} - x)$ tend vers 0 quand i tend vers $+\infty$.

Par l'inégalité triangulaire pour N, $|N(x_{\varphi(i)}) - N(x)| \leq N(x_{\varphi(i)} - x)$, donc $N(x_{\varphi(i)})$ converge vers $N(x)$, par conséquent $N(x) = m$ est non nul puisque x est non nul. Ce qui nous intéresse est que $m > 0$, et que, par définition,

$$\forall x \in \mathbb{R}^n \; \|x\|_\infty = 1 \Rightarrow N(x) \geq m \; .$$

Alors, si $y \in \mathbb{R}^n$ est un vecteur non nul de $\mathbb{R}^n$, l'inégalité ci-dessus appliquée au vecteur $\frac{y}{\|y\|_\infty}$ nous donne

$$\forall y \in \mathbb{R}^n \setminus \{0\} \; N\left(\frac{y}{\|y\|_\infty}\right) \geq m \, , \text{ ou encore } \frac{N(y)}{\|y\|_\infty} \geq m \; .$$

On vient donc de démontrer que $N(y) \geq m\|y\|_\infty$ pour tout vecteur de $\mathbb{R}^n$ différent de 0. Cette inégalité est bien sûr vraie aussi en 0, et finalement on a trouvé m, M strictement positifs tels que

$$\forall y \in \mathbb{R}^n \; m\|y\|_\infty \leq N(y) \leq m = M\|y\|_\infty \; .$$

Ceci prouve que N est équivalente à $\|\cdot\|_\infty$; toutes les normes sont donc équivalentes à $\|\cdot\|_\infty$, donc elles sont toutes équivalentes entre elles. □

Définition 5.4. Soit N une norme sur $\mathbb{R}^n$, $x \in \mathbb{R}^n$ et $R > 0$ un réel. La *boule ouverte de centre x et de rayon R* est l'ensemble

$$B(x, R[= \{y \in \mathbb{R}^n \colon N(x - y) < R\} \; .$$

En dimension 2, si N est la norme euclidienne, la boule ouverte de rayon R et de centre 0 est le disque ouvert de rayon R. Si par exemple N est la norme infinie, la boule ouverte de centre 0 et de rayon R est l'intérieur du carré $]-N, N[\times]-N, N[$.

On peut maintenant définir les ouverts et les fermés de $\mathbb{R}^n$.

Définition 5.5. Soit n un entier ≥ 1, $U \subseteq \mathbb{R}^n$ et N une norme sur $\mathbb{R}^n$. On dit que U est *ouvert* si pour tout $x \in U$ il existe $r > 0$ tel que la boule ouverte de centre x et de rayon r soit contenue dans U.

On dit que $F \subseteq \mathbb{R}^n$ est *fermé* si son complémentaire est ouvert.

Remarquons que, à cause du théorème d'équivalence des normes en dimension finie, la notion d'ensemble ouvert ne dépend pas du choix de la norme dans la définition. Il en va de même pour la notion d'ensemble fermé.

Exercice 5.6. Soit N une norme sur $\mathbb{R}^n$. En utilisant l'inégalité triangulaire, montrer que toutes les boules ouvertes pour N sont des ouverts.

Définition 5.7. Soit (x_i) une suite d'élements de $\mathbb{R}^n$, N une norme sur $\mathbb{R}^n$ et $x \in \mathbb{R}^n$. On dit que x_n *converge vers* x si $N(x - x_i)$ tend vers 0 quand i tend vers $+\infty$.

Encore une fois, la notion de suite convergente ne dépend pas de la norme choisie, ce qui nous donne la caractérisation suivante.

Proposition 5.8. *Soit (x_i) une suite d'élements de $\mathbb{R}^n$, N une norme sur $\mathbb{R}^n$ et $x \in \mathbb{R}^n$. Notons $x_i = (x_1^i, \ldots, x_n^i)$ et $x = (x_1, \ldots, x_n)$. Alors (x_i) converge vers x si, et seulement si, (x_k^i) converge vers x_k pour tout $k \in \{1, \ldots, n\}$.*

Démonstration. La condition de l'énoncé exprime simplement le fait que $\|x_i - x\|_\infty$ tend vers 0. □

Proposition 5.9 (Caractérisation des fermés par les suites)**.** *Soit $n \geq 1$ un entier et F une partie de $\mathbb{R}^n$. Alors F est fermé si, et seulement si, pour toute suite (x_i) d'élements de F qui converge on a $\lim(x_i) \in F$.*

Démonstration. Fixons une norme N sur $\mathbb{R}^n$.

Supposons d'abord que F est fermé, i.e. que son complémentaire est ouvert, et qu'il existe une suite (x_i) d'élements de F qui converge vers $x \notin F$. Alors, il existe $r > 0$ tel que $B(x, r[\cap F = \emptyset$ et, comme $N(x_i - x)$ tend vers 0, on voit que $N(x_i - x) < r$ pour i suffisamment grand, donc $x_i \in B(x, r[$ pour i grand, ce qui contredit l'hypothèse selon laquelle $x_i \in F$. Donc, si F est fermé, pour toute suite (x_i) d'élements de F qui converge on a bien $\lim(x_i) \in F$.

Ensuite, supposons que F n'est pas fermé, i.e. que son complémentaire n'est pas ouvert. Cela signifie qu'il existe $x \notin F$ tel que, pour tout $r > 0$, la boule ouverte de centre x et de rayon r ne soit pas contenue dans le complémentaire de F. En particulier, pour tout $i > 0$ il doit exister $x_i \in F$ tel que $N(x - x_i) < \frac{1}{i}$. La suite x_i est une suite d'éléments de F qui converge vers x, qui n'appartient pas à F. □

Définition 5.10. Soit F une partie de $\mathbb{R}^n$. On dit que F est *compacte* si de toute suite (x_i) d'éléments de F on peut extraire une sous-suite qui converge vers $x \in F$.

Proposition 5.11. *Une partie compacte de $\mathbb{R}^n$ est nécessairement fermée.*

Démonstration. Soit F une partie compacte de $\mathbb{R}^n$, et (x_i) une suite d'éléments de F qui converge vers $x \in \mathbb{R}^n$. Par compacité de F, on peut extraire une sous-suite $(x_{\varphi(i)})$ qui converge vers $x' \in F$. Comme (x_i) tend vers x, on a aussi $\lim(x_{\varphi(i)}) = x$, donc $x = x' \in F$. □

En fait, **en dimension finie**, on a une caractérisation simple des parties compactes.

Définition 5.12. Soit A une partie de $\mathbb{R}^n$. On dit que A est *bornée* s'il existe une norme N sur $\mathbb{R}^n$, et une constante M, telles que

$$\forall x \in A \; N(x) \leq M \; .$$

Encore une fois, à cause du théorème d'équivalence des normes, s'il existe une norme N et une constante M telles que $\forall x \in A \; N(x) \leq M$, alors *pour toute* norme N' il existe une constante M' telle que $\forall x \in A \; N'(x) \leq M'$.

Exercice 5.13. Soit (x_i) une suite d'éléments de $\mathbb{R}^n$ qui converge vers $x \in \mathbb{R}^n$. Montrer que la suite (x_i) est bornée, i.e $\{x_i \colon i \in \mathbb{N}\}$ est un sous-ensemble borné de $\mathbb{R}^n$.

Théorème 5.14 (Théorème de Bolzano-Weierstrass)**.** *Soit (x_i) une suite bornée dans $\mathbb{R}^n$. Alors on peut en extraire une sous-suite convergente.*

Démonstration. C'est une conséquence immédiate du théorème de Bolzano-Weierstrass pour $\|\cdot\|_\infty$ et de l'équivalence des normes en dimension finie. □

Théorème 5.15. *Soit F une partie de $\mathbb{R}^n$. Alors F est compacte si, et seulement si, F est* **à la fois** *fermée et bornée.*

Démonstration. Supposons d'abord F compacte. On a déjà montré que F est fermée. Pour voir que F est bornée, raisonnons par l'absurde : si ce n'est pas le cas, il existe une suite (x_i) d'éléments de F telle que $\|x_i\|_\infty$ tende vers $+\infty$. Alors (x_i) ne peut pas avoir de sous-suite convergente, puisque aucune sous-suite de (x_i) n'est bornée et une suite convergente est nécessairement bornée. Par conséquent F n'est pas compacte, contredisant notre hypothèse de départ.

Supposons maintenant que F soit fermée et bornée, et soit $x_i = (x_1^i, \dots, x_n^i)$ une suite d'éléments de F. D'après le théorème de Bolzano-Weierstras, il existe $x \in \mathbb{R}^n$ et $\psi \colon \mathbb{N} \to \mathbb{N}$ strictement croissante tels que $(x_{\psi(i)})$ converge vers $x = (x_1, \dots, x_n)$; comme F est fermé, $x \in F$, et on vient de trouver une sous-suite de (x_i) qui converge vers $x \in F$. Autrement dit, on vient de prouver que F est compacte. □

Exercice 5.16. Soient $A_1, \dots, A_n$ des parties fermées et bornées de $\mathbb{R}$. Montrer que $A_1 \times A_2 \times \dots \times A_n$ est un compact de $\mathbb{R}^n$.

Ce résultat est souvent utilisé dans le cas où chaque A_i est un segment.

5.2 Fonctions continues de plusieurs variables

Définition 5.17. Soit A une partie de $\mathbb{R}^n$, et $f \colon A \to \mathbb{R}^m$ une fonction. On dit que f est *continue* en $a \in A$ si, pour toute suite (a_i) d'éléments de A qui converge vers a, on a $\lim f(a_i) = f(a)$. On dit que f est continue sur A si f est continue en a pour tout $a \in A$.

Exercice 5.18. Soit A une partie de $\mathbb{R}^n$, et $f \colon A \to \mathbb{R}^m$ une fonction. Soit N_1 une norme sur $\mathbb{R}^n$ et N_2 une norme sur $\mathbb{R}^m$. Montrer que f est continue en $x \in A$ si, et seulement si, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que

$$\forall y \in A \; N_1(x-y) \le \delta \Rightarrow N_2(f(x) - f(y)) \le \varepsilon \; .$$

Proposition 5.19. *Soit A une partie de $\mathbb{R}^n$, et $f \colon A \to \mathbb{R}^m$ une fonction ; on note $f(x) = (f_1(x), \dots, f_m(x))$. Alors f est continue sur A si, et seulement si, chaque fonction f_k est continue sur A.*

Démonstration. Exercice, en utilisant la norme infinie. □

Pour cette raison, les fonctions de plusieurs variables réelles et à valeurs réelles jouent un rôle particulièrement important.

Exercice 5.20. Soit $f \colon \mathbb{R}^2 \to \mathbb{R}$ une fonction. Montrer que si f est continue alors pour tout $x \in \mathbb{R}$ l'application $y \mapsto f(x, y)$ est continue sur $\mathbb{R}$, et pour tout $y \in \mathbb{R}$ l'application $x \mapsto f(x, y)$ est continue sur $\mathbb{R}$.

On considère la fonction définie f sur $\mathbb{R}^2$ par

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

Montrer que f n'est pas continue en 0, et que pourtant pour tout $x \in \mathbb{R}$ l'application $y \mapsto f(x, y)$ est continue sur $\mathbb{R}$, et pour tout $y \in \mathbb{R}$ l'application $x \mapsto f(x, y)$ est continue sur $\mathbb{R}$.

Théorème 5.21. *Soit A une partie compacte non vide de $\mathbb{R}^n$, et f une fonction continue de A dans $\mathbb{R}$. Alors f est bornée sur A et atteint ses bornes.*

Démonstration. La preuve est essentiellement la même que dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Soit M la borne supérieure de $f(A)$ (à ce stade de la preuve, il est possible que $M = +\infty$). Il existe une suite (x_i) d'éléments de A telle que $\lim f(x_i) = M$. Comme A est compacte, (x_i) admet une sous-suite $(x_{\varphi(i)})$qui converge vers $x \in A$. Par continuité de f, on obtient $f(x) = \lim f(x_{\varphi(i)}) = M$, ce qui montre à la fois que M est finie et que M est atteinte.

Le cas de la borne inférieure se traite de la même façon (ou, si on ne veut pas répéter la même preuve, on applique le cas précédent à la fonction $g = -f$). □

Ce résultat est ce dont nous aurons besoin dans la suite ; notons qu'il se généralise facilement au cas des fonctions à valeurs dans $\mathbb{R}^n$.

Théorème 5.22. *Soit A une partie compacte non vide de $\mathbb{R}^n$, et f une fonction continue de A dans $\mathbb{R}^m$. Alors $f(A)$ est compacte.*

Démonstration. Pour voir que $f(A)$ est compacte, prenons une suite (y_i) d'élements de $f(A)$. Alors il existe une suite (a_i) d'éléments de A tels que $f(a_i) = y_i$. Par compacité de A, il existe une sous-suite $(a_{\varphi(i)})$ qui converge vers $a \in A$. Alors $y_{\varphi(i)} = f(a_{\varphi(i)})$ converge vers $f(a)$ puisque f est continue. On vient de construire une suite extraite de (y_i) qui converge vers un élément de $f(A)$, ce qui achève la démonstration. □

Définition 5.23. Soit A une partie de $\mathbb{R}^n$, et $f\colon A \to \mathbb{R}^m$ une fonction. Soit N_1 une norme sur $\mathbb{R}^n$ et N_2 une norme sur $\mathbb{R}^m$. On dit que f *est uniformément continue* si pour tout $\varepsilon >$0 il existe $\delta > 0$ tel que

$$\forall x, y \in A \; N_1(x - y) \leq \delta \Rightarrow N_2(f(x) - f(y)) \leq \varepsilon \; .$$

Bien sûr, toute fonction uniformément continue sur A est continue ; la réciproque, qui est fausse en général (elle est déjà fausse pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$!) est vraie sur les compacts. A cause de l'équivalence des normes en dimension finie, la notion de fonction uniformément continue ne dépend pas du choix de N_1,N_2.

Théorème 5.24 (Théorème de Heine-Borel)**.** *Soit A une partie compacte de $\mathbb{R}^n$, et $f\colon A \to \mathbb{R}^m$ une application continue. Alors f est uniformément continue sur A.*

Démonstration. En exercice : c'est essentiellement la même preuve que celle du théorème 1.5 (qui est un cas particulier du théorème ci-dessus). □

5.3 Différentiabilité des fonctions de plusieurs variables

Définition 5.25. Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f\colon U \to \mathbb{R}^m$ une fonction. Soit N une norme sur $\mathbb{R}^n$. On dit que f est *différentiable* en $x \in U$ s'il existe une application linéaire $D\colon \mathbb{R}^n \to \mathbb{R}^m$ et une fonction $\varepsilon\colon U \to \mathbb{R}^m$ telles que l'on ait, pour tout $y \in U$,

$$f(y) = f(x) + D(y - x) + N(y - x)\varepsilon(y) \, , \text{ avec } \lim_{y \to x} \varepsilon(y) = 0 \; .$$

On dit que f est différentiable sur U si f est différentiable en x pour tout $x \in U$.

Cette définition ne dépend pas du choix de la norme N. Si $n = m = 1$, on retrouve la définition usuelle de la dérivée : l'application D est alors la multiplication par $f'(x)$. Remarquons aussi qu'une application différentiable en x est nécessairement continue en x.

Le premier exemple d'application différentiable est donné par les applications linéaires : si f est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, alors pour tout $x, y \in \mathbb{R}^n$ on a $f(y) = f(x) + f(y - x) + 0$ par linéarité, donc on voit que f est différentiable et $d_x(f) = f$ pour tout $x \in \mathbb{R}^n$. L'idée de la différentiabilité est en fait de s'intéresser aux fonctions qu'on peut bien approcher localement par une fonction linéaire - en espérant montrer des théorèmes valables pour toutes les fonctions différentiables en commençant par les montrer pour les applications linéaires. Souvent on a besoin d'une notion plus forte que la différentiabilité (fonctions de classe C^1) qu'on verra plus loin dans ces notes.

Lemme 5.26. *Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f\colon U \to \mathbb{R}^m$ une fonction. Si f est différentiable en $x \in U$, alors l'application D définie ci-dessus est unique ; on l'appelle* différentielle de f au point x *et on la note df_x ou $df(x)$.*

Démonstration. Fixons une norme N sur $\mathbb{R}^n$. Soient D_1, D_2 deux applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et $\varepsilon_1, \varepsilon_2$ deux fonctions de U dans $\mathbb{R}^m$ telles que, pour tout $y \in U$, on ait

$$\begin{aligned} f(y) &= f(x) + D_1(y - x) + N(y - x)\varepsilon_1(y) \text{ avec } \lim_{y \to x} \varepsilon(y) = 0 \text{ et} \\ f(y) &= f(x) + D_2(y - x) + N(y - x)\varepsilon_2(y) \text{ avec } \lim_{y \to x} \varepsilon(y) = 0 \; . \end{aligned}$$

Alors, pour tout $y \in U$, on a

$$D_1(y - x) - D_2(y - x) = N(y - x)(\varepsilon_2(y) - \varepsilon_1(y)) \; .$$

Comme U est ouvert, il existe $r > 0$ tel que $N(y - x) < r \Rightarrow y \in U$. Posons $u = y - x$; l'équation ci-dessus devient, pour tout u tel que $N(u) < r$:

$$D_1(u) - D_2(u) = N(u)(\varepsilon_2(u + x) - \varepsilon_1(u + x)) \; .$$

Pour $\lambda \in]0,1]$, on a alors, pour tout u tel que $0 < N(u) < r$:

$$\frac{D_1(\lambda u) - D_2(\lambda u)}{N(\lambda u)} = \varepsilon_2(\lambda u + x) - \varepsilon_1(\lambda u + x) \ .$$

Le vecteur de droite tend vers 0 quand λ tend vers 0 tandis que, par linéarité de D_1, D_2, le vecteur de gauche est constant égal à

$$\frac{D_1(u) - D_2(u)}{N(u)} \ .$$

On en déduit que, pour tout u tel que $0 < N(u) < r$ on a $D_1(u) = D_2(u)$. Encore par linéarité, cela entraîne que $D_1 = D_2$: pour tout $v \in \mathbb{R}^n$ on a $N(\lambda v) < r$ pour λ assez petit, et donc $D_1(\lambda v) = D_2(\lambda v)$; par linéarité on a donc $\lambda D_1(v) = \lambda D_2(v)$ pour tout λ assez petit, autrement dit $D_1(v) = D_2(v)$. □

Comme pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$, la différentiabilité se comporte bien par rapport aux combinaisons linéaires et à la composition.

Proposition 5.27. *Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$, $f, g \colon U \to \mathbb{R}^m$ deux fonctions différentiables et α, β deux réels. Alors $\alpha f + \beta g$ est différentiable sur U, et $d(\alpha f + \beta g) = \alpha df + \beta dg$.*

Démonstration. C'est une conséquence immédiate des définitions (donc un bon exercice si jamais ce résultat n'est pas clair pour vous). □

Théorème 5.28 (Règle de la chaîne)**.** *Soient $n, m, p \geq 1$ trois entiers, U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^m$, $f \colon U \to V$ et $g \colon V \to \mathbb{R}^p$ deux fonctions. Supposons que f soit différentiable en $x \in U$ et que g soit différentiable en $f(x)$. Alors $g \circ f$ est différentiable en x, et on a*

$$d(g \circ f)(x) = dg(f(x)) \circ df(x) \ .$$

Avant de donner cette preuve, rappelons que, si $f \colon R^n \to \mathbb{R}^m$ est une application linéaire, que $\mathbb{R}^n$ est muni d'une norme N_1 et $\mathbb{R}^m$ est muni d'une norme N_2 alors on définit la *norme de f* relativement à N_1, N_2 par

$$\|f\|_{N_1,N_2} = \sup_{N_1(x) \leq 1} N_2(f(x)) \ .$$

La propriété fondamentale de cette norme [i] est qu'on a, pour tout $x \in \mathbb{R}^n$,

$$N_2(f(x)) \leq \|f\|_{N_1,N_2} N_1(x) \ .$$

Preuve de la règle de la chaîne. Fixons une norme N_1 sur $\mathbb{R}^n$ et une norme N_2 sur $\mathbb{R}^m$. On a :

$$\begin{aligned} \forall y \in U \ f(y) &= f(x) + df_x(y-x) + N_1(y-x)\varepsilon_1(y) \text{ avec } \lim_{y \to x} \varepsilon_1(y) = 0 \text{ et} \\ \forall z \in V \ g(z) &= g(f(x)) + dg_{f(x)}(z - f(x)) + N_2(z - f(x))\varepsilon_2(z) \text{ avec } \lim_{z \to f(x)} \varepsilon(z) = 0 \ . \end{aligned}$$

On a donc, pour tout $y \in U$:

$$\begin{aligned} g(f(y)) &= g\big(f(x) + df_x(y-x) + N_1(y-x)\varepsilon_1(y)\big) \\ &= g(f(x)) + dg_{f(x)}(df_x(y-x)) + dg_{f(x)}(N_1(y-x)\varepsilon_1(y)) + N_2(d_x(f)(y-x) + N_1(y-x)\varepsilon_1(y))\varepsilon_2(f(y)) \end{aligned}$$

Par linéarité de $dg_{f(x)}$, on a

$$dg_{f(x)}(N_1(y-x)\varepsilon_1(y)) = N_1(y-x) dg_{f(x)}(\varepsilon_1(y)) = N_1(y-x)\varepsilon_3(y) \,,$$

où $\varepsilon_3(y)$ tend vers 0 quand y tend vers x. Ensuite, on a, en notant M la norme de $d_x(f)$ relativement à N_1, N_2 :

$$\begin{aligned} N_2(d_x(f)(y-x) + N_1(y-x)\varepsilon_1(y)) &\leq N_2(d_x(f)(y-x)) + N_2(N_1(y-x)\varepsilon_1(y)) \\ &\leq M N_1(y-x) + N_1(y-x) N_2(\varepsilon_1(y)) \ . \end{aligned}$$

On a donc

$$N_2(d_x(f)(y-x) + N_1(y-x)\varepsilon_1(y))\varepsilon_2(f(y)) = N_1(y-x)\varepsilon_4(y) \,,$$

avec $\varepsilon_4(y)$ qui tend vers 0 quand y tend vers x, et on a finalement obtenu :

$$g(f(y)) = g(f(x)) + dg_{f(x)}(d_x(f)(y-x)) + N_1(y-x)(\varepsilon_3(y) + \varepsilon_4(y)) \ .$$

Ceci montre que $g \circ f$ est différentiable en x et que $d(g \circ f)(x) = dg(f(x)) \circ df(x)$. □

i. qui mérite son nom de norme : c'est une norme sur l'espace vectoriel formé par les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$.

Définition 5.29. Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f \colon U \to \mathbb{R}^m$ une fonction différentiable en $x \in U$. On appelle *matrice jacobienne de f en $x \in U$* la matrice de la différentielle de f en x relativement aux bases canoniques de $\mathbb{R}^n$, $\mathbb{R}^m$. On la note $M(f)(x)$ ou $M_x(f)$.

Pour l'instant, tout ce qu'on a fait est très abstrait - par exemple, comment calculer la différentielle en un point x de l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par $f(r,\theta) = (r\cos(\theta), r\sin(\theta))$? Comme pour la continuité, il est utile de se ramener au cas des fonctions à valeurs réelles ; dans ce cas, la différentielle en un point est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$, donc sa matrice est un vecteur ligne avec n entrées.

Lemme 5.30. *Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f \colon U \to \mathbb{R}^m$ une fonction. On note $f = (f_1, \ldots, f_m)$. Alors f est différentiable en $x \in U$ si, et seulement si, chaque f_i est différentiable en x, et on a*

$$M(f)(x) = \begin{pmatrix} M(f_1)(x) \\ \ldots \\ M(f_n)(x) \end{pmatrix} .$$

La matrice ci-dessus est la matrice à m lignes et n colonnes dont la i-ième ligne est égale à $M(f_i)$. Dans le cas d'une application définie sur un intevralle ouvert de $\mathbb{R}$ et à valeurs dans $\mathbb{R}^n$, notée $f(x) = (f_1(x), \ldots, f_n(x))$, le résultat ci-dessus dit que f est différentiable en x si, et seulement si, chaque f_i est dérivable en x, et la matrice de la différentielle de f est le vecteur de $\mathbb{R}^n$ de coordonnées $(f_1'(x), \ldots, f_n'(x))$, que l'on note $f'(x)$.

Démonstration. C'est une conséquence immédiate de la définition, et donc un bon exercice pour vérifier que celle-ci est bien comprise... $\square$

Définition 5.31. Soit $n \geq 1$ un entier, $j \in \{1, \ldots, n\}$, U un ouvert de $\mathbb{R}^n$ et $f \colon U \to \mathbb{R}$ une fonction. On dit que f *admet une dérivée partielle par rapport à la variable x_j* en $x = (x_1, \ldots, x_n) \in U$ si l'application

$$g_j \colon t \mapsto f(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_n)$$

admet une dérivée en $t = x_j$. Dans ce cas, on pose $\frac{\partial f}{\partial x_j}(f)(x) = g_j'(x_j)$.

Explicitement, ce que cette définition signifie est : on regarde l'application obtenue en ne faisant varier que la j-ième variable ; si cette application a une dérivée au point où on s'est placé, cette dérivée est la dérivée partielle de f par rapport à la variable x_j.

Exemple. Soit $f \colon (r,\theta) \mapsto r\cos(\theta)$. Cette application a des dérivées partielles selon la variable r et la variable θ en tout point de $\mathbb{R}^2$, et on a

$$\frac{\partial f}{\partial r}(r,\theta) = \cos(\theta) \quad et \quad \frac{\partial f}{\partial \theta}(f) = -r\sin(\theta) .$$

Théorème 5.32. *Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f \colon U \to \mathbb{R}^m$ une fonction. On note $f = (f_1, \ldots, f_m)$. Alors, si f est différentiable en $x \in U$, chaque f_i admet des dérivées partielles en x par rapport aux variables $x_1, \ldots, x_n$ et on a*

$$M(f)(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \ldots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

La matrice ci-dessus est la matrice dont le coefficient sur la i-ièüe ligne et la j-ième colonne est donné par la dérivée partielle de f_i par rapport à la variable x_j.

Démonstration. Le lemme 5.30 permet de supposer que $m = 1$, ce qu'on fait dans la suite. Supposons que f soit différentiable en $x \in U$; comme $m = 1$ la matrice de df_x est un vecteur ligne de longueur n. Fixons $i \in \{1, \ldots, n\}$, notons $x = (x_1, \ldots, x_m)$ et considérons l'application g définie au voisinage de x_i par $t \mapsto f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$.

Pour $u \in \mathbb{R}$, appelons u_i le vecteur de $\mathbb{R}^n$ dont toutes les coordonnées sont nulles, sauf la i-ième coordonnée qui vaut u. Alors, pour tout $u \in \mathbb{R} \setminus \{0\}$ suffisamment petit, $x + u_i \in U$ et on a

$$\begin{aligned} g(x_i + u) &= f(x + u_i) \\ &= f(x) + df_x(u_i) + \|u_i\|_\infty \varepsilon(u_i) \\ &= g(x_i) + df_x(u_i) + |u|\varepsilon'(u) \end{aligned}$$

où ε' est une fonction qui tend vers 0 quand u tend vers 0. On vient de montrer que g est dérivable en x_i, et que

$$g'(x_i) = \lim_{u\to 0} \frac{df_x(u_i)}{u} = \underbrace{df_x(0,\ldots,0,1,0,\ldots,0)}_{\text{toutes les coordonnées nulles sauf la } i\text{-ième}} .$$

Ceci revient à dire que $\frac{\partial f}{\partial x_i}(x)$ existe et est égal à la i-ième coordonnée de la matrice de df_x, ce qu'on voulait démontrer. □

Remarque 5.33. Il est très tentant de supposer que, si toutes les dérivées partielles de f existent en un point x, alors f est différentiable en ce point. Ce n'est pas le cas en général ! Pour le voir, considérez par exemple la fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$ définie par

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

Alors f admet des dérivées partielles en $(0,0)$, qui sont nulles. Donc si f était différentiable en 0 sa différentielle serait la fonction nulle, et on aurait pour tout $u \in \mathbb{R}^2$ $f(u) = \|u\|_\infty \varepsilon(u)$, où ε tend vers 0 quand u tend vers 0. On devrait donc avoir

$$\lim_{u\to 0} \frac{f(u)}{\|u\|_\infty} = 0 .$$

Puisque $f(\frac{1}{n},\frac{1}{n}) = \frac{1}{n}$, on voit qu'on n'a pas $\lim_{u\to 0} \frac{f(u)}{\|u\|_\infty} = 0$, donc f n'est pas différentiable en 0.

Le problème dans l'exemple ci-dessus est que les dérivées partielles de f ne sont pas continues en 0.

Théorème 5.34. *Soient $n,m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f\colon U \to \mathbb{R}^m$ une fonction. On note $f = (f_1,\ldots,f_m)$. Si les dérivées partielles $\frac{\partial f_i}{\partial x_j}$ existent sur un voisinage de $x \in U$* **et sont continues** *en x alors f est différentiable en $x \in U$.*

Démonstration. Comme précédemment, il suffit de traiter le cas des fonctions à valeurs dans $\mathbb{R}$. Dans l'espoir que l'idée de la preuve soit claire, on va se contenter de la donner pour des fonctions de deux variables. Soit donc $f\colon U \to \mathbb{R}$ une fonction de deux variables dont les dérivées partielles existent et sont continues au voisinage d'un point $(x_0,y_0) \in U$. Pour tout (s,t) suffisamment petit, (x_0+s, y_0+t) appartient à ce voisinage et on a

$$\begin{aligned} f(x_0+s,y_0+t) &= f(x_0+s,y_0) + (f(x_0+s,y_0+t) - f(x_0+s,y_0)) \\ &= f(x_0,y_0) + (f(x_0+s,y_0) - f(x_0,y_0)) + (f(x_0+s,y_0+t) - f(x_0+s,y_0)) \\ &= f(x_0,y_0) + s\frac{\partial f}{\partial x}(x_0+cs,y_0) + t\frac{\partial f}{\partial y}(x_0+s,y_0+dt) \end{aligned}$$

pour une certaine paire (c,d) d'éléments de $]0,1[$ (la dernière égalité résulte du théorème des accroissements finis, appliqué à des fonctions d'une variable réelle à valeurs réelles). Par continuité de $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$, on a

$$\frac{\partial f}{\partial x}(x_0+cs,y_0) = \frac{\partial f}{\partial x}(x_0,y_0) + \varepsilon_1(s), \text{ où } \lim_{s\to 0} \varepsilon_1(s) = 0 .$$

De même, on a

$$\frac{\partial f}{\partial y}(x_0+cs,y_0+dt) = \frac{\partial f}{\partial y}(x_0,y_0) + \varepsilon_2(s,t), \text{ où } \lim_{(s,t)\to 0} \varepsilon_2(s,t) = 0 .$$

Tout ceci nous donne finalement

$$f(x_0+s,y_0+t) = f(x_0,y_0) + s\frac{\partial f}{\partial x}(x_0,y_0) + t\frac{\partial f}{\partial y}(x_0,y_0) + s\varepsilon_1(s) + t\varepsilon_2(s,t) .$$

L'application $(s,t) \mapsto s\frac{\partial f}{\partial x}(x_0,y_0) + t\frac{\partial f}{\partial y}(x_0,y_0)$ est linéaire, et on a

$$\lim_{(s,t)\to(0,0)} \frac{s\varepsilon_1(s) + t\varepsilon_2(s,t)}{\|(s,t)\|_\infty} = 0 .$$

On vient de démontrer que f est différentiable en (x_0,y_0), et que sa différentielle est l'application $(s,t) \mapsto s\frac{\partial f}{\partial x}(x_0,y_0) + t\frac{\partial f}{\partial y}(x_0,y_0)$. □

Exercice 5.35. Prouver le résultat précédent pour des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ avec $n \geq 2$ un entier quelconque.

Définition 5.36. Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$ et $f = (f_1, \ldots, f_m) \colon U \to \mathbb{R}^m$ une fonction. Si toutes les dérivées partielles de chaque f_i existent et sont continues sur U alors on dit que f est *de classe* C^1 sur U.

Plus généralement, on définit par récurrence les applications de classe C^k : f est de classe C^{k+1} sur U si toutes les dérivées partielles des f_i existent sur U et sont de classe C^k.

Grâce à la règle de la chaîne, on vérifie facilement par récurrence qu'une composée d'applications de classe C^p est encore de classe C^p.

Remarque 5.37. La définition d'une fonction de classe C^1 ci-dessus est équivalente au fait de dire que $df(x)$ existe en tout point de U et que l'application $x \mapsto df(x)$ est continue de U dans l'espace vectoriel formé par les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$ (modulo l'identification d'une application linéaire à sa matrice dans les bases canoniques, cet espace n'est autre que l'espace $M_{m,n}(\mathbb{R})$ des matrices à n colonnes et m lignes, qui est un espace vectoriel de dimension nm).

Théorème 5.38 (Théorème de Schwarz)**.** *Soient n un entier, U un ouvert de $\mathbb{R}^n$ et $f \colon U \to \mathbb{R}$ une fonction de classe C^2 sur U. Alors on a, pour tout $i, j \in \{1, \ldots, n\}$ et tout $x \in U$:*

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

En d'autres termes, pour des fonctions de classe C^2, l'ordre dans lequel on effectue les dérivations n'a pas d'influence sur le résultat. On note alors $\frac{\partial^2 f}{\partial x_i \partial x_j}$ la fonction obtenue en dérivant une fois par rapport à x_i et une fois par rapport à x_j ; dans le cas où $i = j$, on note $\frac{\partial^2 f}{\partial x_i^2}$ la fonction obtenue en dérivant deux fois par rapport à x_i.

Démonstration. On va donner la preuve pour une fonction définie sur un ouvert U de $\mathbb{R}^2$(le cas général s'en déduit facilement : si on considère deux dérivées partielles en un point, il n'y a que deux variables qui ne sont pas fixées !). Fixons $(x_0, y_0) \in U$. Pour (s, t) proche de 0, $(x_0 + s, y_0 + t) \in U$, et on pose

$$F(s,t) = f(x_0 + s, y_0 + t) - f(x_0 + s, y_0) + f(x_0, y_0) - f(x_0, y_0 + t) \ .$$

A s, t fixés, on pose $\varphi(x) = f(x, y_0 + t) - f(x, y_0)$. On a a lors $F(s,t) = \varphi(x_0 + s) - \varphi(x_0)$. Alors, φ est dérivable, de dérivée

$$\varphi'(x) = \frac{\partial f}{\partial x}(x, y_0 + t) - \frac{\partial f}{\partial x}(x, y_0) \ .$$

On peut appliquer le théorème des accroissements finis (pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$) à φ sur le segment d'extrémités x_0, $x_0 + s$, et obtenir $c_1 \in]0, 1[$ tel que

$$\varphi(x_0 + s) - \varphi(x_0) = s\varphi'(x_0 + c_1 s) \ ,$$

c'est-à-dire

$$F(s,t) = s\Big(\frac{\partial f}{\partial x}(x_0 + c_1 s, y_0 + t) - \frac{\partial f}{\partial x}(x_0 + c_1 s, y_0)\Big) \ .$$

Le théorème des accroissements finis, appliqué cette fois à $y \mapsto \frac{\partial f}{\partial x}(x_0 + c_1 s, y)$ sur le segment d'extrémités y_0, $y_0 + t$, nous donne un $d_1 \in]0, 1[$ tel que

$$F(s,t) = st\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0 + c_1 s, y_0 + d_1 t) \ .$$

En appliquant exactement le même raisonnement avec la fonction $\psi \colon y \mapsto f(x_0 + s, y) - f(x_0, y)$, on obtient l'existence de $c_2, d_2 \in]0, 1[$ tels que

$$F(s,t) = st\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0 + c_2 s, y_0 + d_2 t) \ .$$

On a donc (dès que s et t sont tous les deux non nuls et suffisamment proches de 0) :

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0 + c_1 s, y_0 + d_1 t) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0 + c_2 s, y_0 + d_2 t) \ .$$

En faisant tendre (s,t) vers 0 et en utilisant la continuité de $\frac{\partial}{\partial y}\frac{\partial f}{\partial x}$ et de $\frac{\partial}{\partial x}\frac{\partial f}{\partial y}$ (qui fait partie des hypothèses du théorème !), on obtient finalement

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0,y_0) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0,y_0)$$

□

5.4 Inégalité des accroissements finis

Lemme 5.39 (Inégalité des accroissements finis pour les fonctions d'une variable réelle)**.** *Soit $I=[a,b]$ un segment de $\mathbb{R}$, $m\geq 1$ un entier, $f\colon I\to\mathbb{R}^m$ et $g\colon I\to\mathbb{R}$ des fonctions continues sur $[a,b]$ et dérivables sur $]a,b[$. Soit $\|\cdot\|$ une norme sur $\mathbb{R}^m$, et supposons que $\|f'(x)\|\leq g'(x)$ pour tout $x\in]a,b[$.*

Alors on a $\|f(b)-f(a)\|\leq g(b)-g(a)$. En particulier,on a

$$\|f(b)-f(a)\|\leq (\sup_{x\in]a,b[}\|f'(x)\|\,)|b-a|\ .$$

Démonstration. Pour $\varepsilon>0$, on définit $\varphi_\varepsilon\colon[a,b]\to\mathbb{R}$ par

$$\varphi_\varepsilon(t)=\|f(t)-f(a)\|-g(t)-\varepsilon t\ .$$

Pour tout $x\in]a,b[$, φ_ε est continue sur $[x,b]$ donc y admet un minimum. Montrons que ce minimum est atteint en $x=b$; pour cela, il suffit de montrer que φ_ε n'atteint pas son minimum en un $t\in[x,b[$. Si on prend $t\in[x,b[$, alors on a, pour $s\in]t,b[$ suffisamment proche de t :

$$\left\|\frac{f(s)-f(t)}{s-t}\right\|-\frac{\varepsilon}{2}\leq\|f'(t)\|\leq g'(t)<\frac{g(s)-g(t)}{s-t}+\frac{\varepsilon}{2}\ .$$

Par conséquent, pour $s\in[t,b[$ suffisamment proche de t, on a $\|f(s)-f(t)\|<g(s)-g(t)+\varepsilon(s-t)$, donc

$$\begin{aligned}\varphi_\varepsilon(s)-\varphi_\varepsilon(t) &= g(s)-g(t)+\varepsilon(s-t)+\|f(s)-f(a)\|-\|f(t)-f(a)\|\\ &< g(s)-g(t)+\varepsilon(s-t)+\|f(s)-f(t)\|\\ &< 0.\end{aligned}$$

Ceci montre que φ_ε n'atteint pas son minimum en t. Donc ce minimum est nécessairement atteint en b, ce dont on déduit que, pour tout $x\in]a,b]$ et tout $\varepsilon>0$, on a

$$\|f(x)-f(a)\|-g(x)-\varepsilon t\geq\|f(b)-f(a)\|-g(b)-\varepsilon b\ .$$

En faisant tendre ε vers 0 (à x fixé), ceci donne

$$\|f(x)-f(a)\|-g(x)\geq\|f(b)-f(a)\|-g(b)\ .$$

Finalement, en faisant tendre x vers a, on obtient

$$-g(a)\geq\|f(b)-f(a)\|-g(b)\ .$$

C'est l'inégalité qu'on souhaitait démontrer. □

Définition 5.40. Soit $n\geq 1$ un entier et x,y deux éléments de $\mathbb{R}^n$. Le *segment* reliant a et b, noté $[x,y]$, est défini par

$$[x,y]=\{tx+(1-t)y\colon t\in[0,1]\}\ .$$

Théorème 5.41 (Inégalité des accroissements finis pour des fonctions de plusieurs variables)**.** *Soient $n,m\geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n$, et f une fonction de U dans $\mathbb{R}^m$. On fixe une norme N_1 sur $\mathbb{R}^n$, une norme N_2 sur $\mathbb{R}^m$, et on note $\|f\|$ la norme d'une application linéaire f de $\mathbb{R}^n$ dans $\mathbb{R}^m$ subordonnée aux normes N_1, N_2.*

Supposons que f soit différentiable sur U. Alors, pour tout $x,y\in U$ tels que $[x,y]\subseteq U$, on a

$$N_2(f(x)-f(y))\leq(\sup_{t\in]x,y[}\|df(t)\|\,)N_1(x-y)\ .$$

Démonstration. Ce théorème découle immédiatement de l'inégalité des accroissements finis pour les fonctions d'une variable réelle : définissons une fonction $\varphi\colon [0,1] \to \mathbb{R}^n$ en posant $\varphi(t) = f(tx+(1-t)y)$. Alors $\varphi(0) = f(y)$, $\varphi(1) = f(x)$, φ est continue sur $[x,y]$ et dérivable sur $]x,y[$, et la règle de la chaîne donne

$$\varphi'(t) = df(tx + (1-t)y)(x-y)$$

En particulier, on a pour tout $t \in]0,1[$

$$N_2(\varphi'(t)) \leq (\sup_{t\in]x,y[} \|df(t)\|)N_1(x-y) \ .$$

En appliquant l'inégalité des accroissements finis à φ sur $[0,1]$, on obtient

$$N_2(\varphi(1) - \varphi(0)) \leq (\sup_{t\in]x,y[} \|df(t)\|)N_1(x-y) \ .$$

C'est ce qu'on voulait démontrer. □

5.5 Gradient, hessienne et extrema

Définition 5.42. Soit $n \geq 1$ un entier, U un ouvert de $\mathbb{R}^n$ et $f\colon U \to \mathbb{R}$ une fonction différentiable en $x \in U$. Le *gradient* de f en x est le vecteur

$$\mathrm{grad}(f)(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right) \ .$$

Remarquons que $\mathrm{grad}(f)(x)$ est simplement la matrice jacobienne de f au point x, vue comme un vecteur de $\mathbb{R}^n$; pour tout $h \in \mathbb{R}^n$ et tout $x \in U$ on a

$$d_x(f)(h) = \langle \mathrm{grad}(f)(x), h\rangle \ .$$

(On utilise la notation $\langle \cdot, \cdot \rangle$ pour désigner le produit scalaire usuel sur $\mathbb{R}^n$).

L'égalité ci-dessus montre aussi que, pour h de norme fixée, $df_x(h)$ est maximal quand h est colinéaire et de même sens que $\mathrm{grad}(f)(x)$: $\mathrm{grad}(f)(x)$ donne la direction dans laquelle f « augmente le plus vite ».

Définition 5.43. Soit $n \geq 1$ un entier, A une partie de $\mathbb{R}^n$ et $f\colon A \to \mathbb{R}$ une fonction. On dit que f a un *extremum* en $x \in A$ si $f(x)$ est le maximum, ou le minimum, de f sur A. Si pour tout $y \in A \setminus \{x\}$ on a $f(y) < f(x)$ alors on dit que x est un maximum *strict* de f sur A. On définit de même la notion de minimum strict.

S'il existe un ouvert U contenant x et tel que $f_{|U\cap A}$ admette un extremum en x, on dit que x est un *extremum local* de f. On définit de même la notion d'*extremum local strict*.

Proposition 5.44. *Soit $n \geq 1$ un entier, U un ouvert de $\mathbb{R}^n$ et $f\colon U \to \mathbb{R}$ une fonction différentiable en $x \in U$. Si f admet un extremum en $x \in U$, alors $\mathrm{grad}(f)(x) = 0$. On dit alors que x est un* point critique *de f sur U.*

Démonstration. Pour tout vecteur $u \in \mathbb{R}^n$, considérons la fonction $f_u\colon t \mapsto f(x+tu)$. Cette fonction est définie sur un intervalle ouvert contenant 0, et admet un extremum local en 0. Par conséquent, on doit avoir $f_u'(0) = 0$. Par la règle de la chaîne, on a

$$f_u'(0) = df_x(u) = \langle \mathrm{grad}(f)(x), u\rangle \ .$$

Par conséquent, on a $\langle \mathrm{grad}(f)(x), u\rangle = 0$ pour tout $u \in \mathbb{R}^n$, et ceci n'est possible que si $\mathrm{grad}(f)(x) = 0$. □

La réciproque n'est pas vraie : on peut avoir $\mathrm{grad}(f)(x) = 0$ sans que x soit un extremum local pour f ; c'était déjà le cas pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, considérez par exemple $f\colon x \mapsto x^3$. Alors $f'(0) = 0$ mais 0 n'est pas un extremum local pour f.

Si on cherche les extrema d'une fonction f sur un ouvert U, on peut donc commencer par chercher les éléments x tels que $\mathrm{grad}(f)(x) = 0$. Puisque les différentielles ne donnent qu'une information locale (elles ne disent rien sur ce que fait f « loin »de x !), on ne peut de toute façon pas espérer qu'elles nous suffisent à décider si un point est un extremum sur U. Par contre, on peut essayer d'utiliser des dérivées pour savoir si x est un extremum local pour f ; pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on utiliserait un développement limité à l'ordre 2 : si x est tel que $f'(x)=0$ et f est deux fois dérivable en x, alors on a, par la formule de Taylor-Young,

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

Ainsi, pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, si $f'(x)=0$ et $f''(x)>0$, alors f admet un minimum local strict en x ; si $f'(x)=0$ et $f''(x)<0$, alors f admet un maximum local strict en x. Si $f'(x)=0$ et $f''(x)=0$, alors le développement limité à l'ordre 2 ne nous permet pas de conclure.

La situation est similaire pour les fonctions de plusieurs variables, mais la notion de dérivée seconde est plus compliquée : il y a beaucoup de dérivées secondes possibles, puisqu'on peut d'abord dériver par rapport à la variable x_i, puis dériver une nouvelle fois par rapport à la variable x_j... La bonne approche consiste à regrouper toutes ces dérivées dans une matrice, et à étudier les propriétés de cette matrice.

Définition 5.45. Soit $n\geq 1$ un entier, U un ouvert de $\mathbb{R}^n$ et $f\colon U\to\mathbb{R}$ une fonction de classe C^2 sur U. On définit la *matrice hessienne* $H(f)(x)$ de f en $x\in U$ comme étant la matrice dont le coefficient sur la i-ième ligne et la j-ième colonne est égal à $\frac{\partial^2 f}{\partial x_i\partial x_j}(x)$:

$$H(f)(x)=\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \dots & \frac{\partial^2 f}{\partial x_1\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_i\partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_i\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_i\partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

C'est une matrice carrée $n\times n$. Evidemment, quand $n=1$, la matrice hessienne de f en x est une matrice 1×1, dont le coefficient vaut $f''(x)$, donc on ne fait que se compliquer la vie en y pensant comme étant une matrice - mais en dimension supérieure, il faut bien prendre en compte toutes les dérivées secondes possibles. Grâce au théorème de Schwarz, on sait que la matrice hessienne est symétrique ; par conséquent, elle est diagonalisable sur $\mathbb{R}$ et est la matrice d'une forme bilinéaire symétrique. Ce sont les propriétés de la forme quadratique associée qui jouent un rôle dans l'étude des extrema de f.

Théorème 5.46. *Soit $n\geq 1$ un entier, U un ouvert de $\mathbb{R}^n$, $f\colon U\to\mathbb{R}$ une fonction de classe C^2 sur U, et $x\in U$ tel que grad$(f)(x)=0$. Alors :*

1. *Si la matrice hessienne de f en x est définie positive (i.e. si toutes ses valeurs propres sont strictement positives) alors x est un minimum local de f.*
2. *Si la matrice hessienne de f en x est définie négative (i.e. si toutes ses valeurs propres sont strictement négatives) alors x est un maximum local de f.*
3. *Si la matrice hessienne de f en x a une valeur propre strictement positive et une valeur propre strictement négative, alors x n'est pas un extremum local de f ; on dit alors que x est un* point selle *de f.*
4. *Si l'on n'est pas dans un des cas précédents, alors on ne peut pas savoir si x est, ou non, un extremum local de f.*

Réciproquement, si x est un minimum local de f alors $H(f)(x)$ doit être positive (mais pas forcément définie positive), et si x est un maximum local de f alors $H(f)(x)$ doit être négative (mais pas forcément définie négative).

Démonstration. Prenons notre courage à deux mains et appliquons une stratégie déjà utilisée plus haut : fixons $u=(u_1,\dots,u_n)\in\mathbb{R}^n\setminus\{0\}$, et considérons l'application $\varphi_u\colon t\mapsto f(x+tu)$, qui est définie sur un voisinage de 0. Alors on a

$$\varphi_u'(t)=\langle \mathrm{grad}(f)(x+tu),u\rangle=\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x+tu)u_i\ .$$

Le gradient de f en x est nul, donc on a $\varphi_u'(0)=0$; comme f est de classe C^2, on peut dériver une fois de plus et obtenir

$$\varphi_u''(t)=\sum_{i=1}^n\left(\sum_{j=1}^n \frac{\partial^2 f}{\partial x_j\partial x_i}(x+tu)u_j\right)u_i\ .$$

En 0, on a donc
$$\varphi_u''(0) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x) u_j \right) u_i = \langle H(f)u, u \rangle \ .$$
En se souvenant du cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ décrit plus haut, on voit que

(a) Si $\langle H(f)u, u \rangle > 0$, alors x est un minimum local pour la fonction φ_u.

(b) Si $\langle H(f)u, u \rangle < 0$, alors x est un minimum local pour la fonction φ_u.

On peut maintenant conclure : si la condition (1) du théorème est vérifiée, alors pour tout u on est dans le cas (a) ci-dessus, donc pour tout u x est un minimum local de φ_u, ce dont on déduit que x est un minimum local pour f. De même, si la condition 2 du théorème est vérifiée alors x est un maximum local de f.

Si par contre on est dans le cas (3), alors il existe u_1 (un vecteur propre pour une valeur propre strictement positive) pour lequel x est un minimum local strict pour φ_{u_1}, et u_2 (un vecteur propre pour une valeur propre strictement négative) pour lequel x est un maximum local strict pour φ_{u_2}. Ceci montre que x ne peut être ni un minimum local, ni un maximum local, pour f, donc x n'est pas un extremum local de f. $\square$

Dans le cas des applications de $\mathbb{R}^2$ dans $\mathbb{R}$, nul besoin de calculer les valeurs propres pour savoir si une matrice symétrique est définie positive (négative), comme le rappelle l'exercice suivant.

Exercice 5.47. Soit $A \in M_2(\mathbb{R})$ une matrice symétrique, de valeurs propres λ_1, λ_2. Montrer que $\det(A) = \lambda_1\lambda_2$ et $\mathrm{tr}(A) = \lambda_1 + \lambda_2$. En déduire que :

1. A est définie positive si, et seulement si, $\det(A) > 0$ et $\mathrm{tr}(A) > 0$.
2. A est définie négative si, et seulement si, $\det(A) > 0$ et $\mathrm{tr}(A) < 0$.
3. A est positive si, et seulement si, $\det(A) \geq 0$ et $\mathrm{tr}(A) \geq 0$.
4. A est négative si, et seulement si, $\det(A) \leq 0$ et $\mathrm{tr}(A) \leq 0$.
5. A a une valeur propre strictement positive et une valeur propre strictement négative si, et seulement si, $\det(A) < 0$.

Exercice 5.48. Pour les applications de $\mathbb{R}^2$ dans $\mathbb{R}$, reformuler les conditions du Théorème 5.46 en utilisant le déterminant et la trace de la différentielle de f au point x.

5.6 Difféomorphismes de classe C^1

Définition 5.49. Soient U, V deux ouverts de $\mathbb{R}^n$ et $f \colon U \to \mathbb{R}^n$ une fonction. On dit que f est un *difféomorphisme de classe* C^1 de U sur V si :

1. f est une bijection de U sur V (i.e. f est injective sur U, et $f(U) = V$)
2. f est de classe C^1 sur U.
3. f^{-1} est différentiable sur U.

Le premier exemple de difféomorphisme de classe C^1 est fourni par les applications linéaires inversibles, qui sont des difféomorphismes de classe C^1 de $\mathbb{R}^n$ sur lui-même.

Proposition 5.50. *Soient U, V deux ouverts de $\mathbb{R}^n$,$k \geq 1$ et $f \colon U \to \mathbb{R}^n$ un difféomorphisme de classe C^1 de U sur V. Alors l'application inverse f^{-1} de f est un difféomorphisme de classe C^k de V sur U, la différentielle de f est inversible en tout $x \in U$, et on a, pour tout $x \in U$:*

$$df^{-1}(f(x) = (df(x))^{-1} \ .$$

Démonstration. La formule permettant de calculer la différentielle de f^{-1} est une conséquence immédiate de la règle de la chaîne : pour tout $x \in U$ on a, par définition de f^{-1}, que $f^{-1} \circ f(x) = x$. En différentiant cette égalité, et en utilisant le fait que la différentielle de l'application $x \mapsto x$ est l'application identité, on obtient, pour tout $x \in U$:

$$df^{-1}(f(x)) \circ df(x) = I \ .$$

On en déduit donc que $df(x)$ est inversible, d'inverse $df^{-1}(f(x)$, ce qui montre que $df^{-1}(f(x)) = (df(x))^{-1}$.

Reste à vérifier que f^{-1} est de classe C^1. Pour cela, on rappelle que la *comatrice* $c(A)$ d'une matrice carrée A est la matrice dont le coefficient sur la i-ième ligne et la j-ième colonne est le déterminant de la matrice obtenue en enlevant de A sa i-ième ligne et sa j-ième colonne ; si A est inversible, alors l'inverse de A est égale à $\frac{1}{det(A)} c(A)^T$. Les coefficients apparaissant dans la matrice jacobienne de $df^{-1}(y)$ sont donc des quotients de fonctions continues dont le dénominateur ne s'annule pas ; ce sont donc des fonctions continues, ce qui montre que f^{-1} est de classe C^1.

Une autre façon de montrer que f^{-1} est de classe C^1 : si A est une matrice inversible à n lignes et n colonnes, le théorème de Cayley–Hamilton nous dit qu'on a $P(A) = 0$, pù P est le polynôme caractéristique de A, dont le coefficient constant vaut $(-1)^n \det(A)$ et le coefficient dominant vaut 1. Par conséquent, on a une égalité de la forme $A^n + a_{n-1}A^{n-1} + \ldots + a_1 A = (-1)^{n+1}\det(A) I_n$; en mettant A en facteur à gauche, on voit que A^{n-1} vaut $A^{n-1} + a_{n-1} + A^{n-2} + \ldots + a_1 I_n$; donc les coefficients de A^{n-1} sont des fonctions continues des coefficients de A, ce qu'on voulait démontrer.

□

Remarque 5.51. Si U est un ouvert non vide de $\mathbb{R}^n$, et V est un ouvert de $\mathbb{R}^m$, alors il ne peut exister de bijection différentiable et d'inverse différentiable de U sur V que si $n = m$: en effet, la différentielle de f en un point $x \in U$ quelconque devrait être une bijection linéaire de $\mathbb{R}^n$ sur $\mathbb{R}^m$, et une telle application ne peut exister que si $n = m$.

Plus généralement, il est impossible qu'un ouvert U de $\mathbb{R}^n$ non vide soit *homéomorphe* à un ouvert V de $\mathbb{R}^m$ si $n \neq m$. C'est le *théorème d'invariance du domaine*, beaucoup plus général que celui qu'on vient d'énoncer, et hors de portée dans ce cours.

La section suivante est hors programme cette année.

5.7 Théorème d'inversion locale

Théorème 5.52 (Théorème d'inversion locale)**.** *Soit U un ouvert de $\mathbb{R}^n$, et $f \colon U \to \mathbb{R}^n$ une fonction de classe C^1. Supposons que $x \in U$ soit tel que $df(x)$ soit inversible. Alors il existe deux ouverts U_1, V_1 tels que $x \in U_1$, $f(x) \in V_1$, et $f_{|U_1}$ soit un difféomorphisme de classe C^1 de U_1 sur V_1.*

Démonstration. On va utiliser le théorème du point fixe de Picard. Pour cela, on se fixe une norme $\|\cdot\|$ sur $\mathbb{R}^n$, et on note aussi $\|\cdot\|$ la norme sur les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^n$ subordonnée à $\|\cdot\|$.

Puisque la translation $x \mapsto x + x_0$ et $df(x_0)$ sont des difféomorphismes on peut, quitte à composer, supposer que $x_0 = 0$ et $df(x_0) = I$, ce qu'on fait dans la suite.

Définissons une application g sur U en posant $g(x) = f(x) - x$. Alors g est de classe C^1 sur U, et $dg(0) = df(0) - I = 0$. Par continuité de df en 0, il existe une boule ouverte $B \subseteq U$ contenant 0 et telle que

$$\forall x \in B \ \|dg(x)\| \leq \frac{1}{2} .$$

Alors, en appliquant l'inégalité des accroissements finis, on voit que

$$\forall x, y \in B \ \|g(x) - g(y)\| \leq \frac{1}{2}\|x - y\| .$$

Maintenant, on a, pour tout $x, y \in B$:

$$\|x - y\| = \|g(y) - g(x) + f(x) - f(y)\| \leq \frac{1}{2}\|x - y\| + \|f(x) - f(y)\| .$$

Donc $\frac{1}{2}\|x - y\| \leq \|f(x) - f(y)\|$: ceci montre que f est injective sur B. Montrons maintenant que $f(B)$ est ouvert. Fixons $z_0 = f(x_0) \in f(B)$ et $r > 0$ tel que $B(x, r]$ soit contenu dans B. Pour un z fixé, définissons une fonction auxiliaire g_z en posant $g_z(x) = z - g(x)$. On cherche à montrer que, pour z suffisamment proche de z_0, il existe $x \in B$ tel que $f(x) = z$, autrement dit on cherche $x \in B$ tel que $g_z(x) = x$, i.e. un point fixe de g_z.

Notons que, si $x, y \in B$ on a

$$\|g_z(x) - g_z(y)\| = \|g(x) - g(y)\| \leq \frac{1}{2}\|x - y\| .$$

En appliquant cela pour $z = z_0$, on voit, puisque $g_{z_0}(x_0) = x_0$, que $g_{z_0}(B(x_0, r]) \subseteq B(x_0, \frac{r}{2}]$. Puisque $g_z = g_{z_0} + z - z_0$ on a aussi

$$g_z(B(x_0, r]) = g_{z_0}(B(x_0, r]) + z - z_0 \subseteq B(x_0, \frac{r}{2} + \|z - z_0\|] .$$

Tout ceci montre que, si $\|z - z_0\| \leq \frac{r}{2}$, g_z est une contraction (de rapport 1/2) de $B(x, r]$ dans lui-même, donc admet un unique point fixe (c'est ici qu'on utilise le théorème du point fixe de Picard) ; par conséquent, pour tout z tel que $\|z - z_0\| \leq \frac{r}{2}$ il existe $x \in B([x_0, r] \subseteq B$ tel que $f(x) = r$. Autrement dit, la boule fermée de centre z_0 et de rayon $r/2$ est contenue dans $f(B)$, ce qui montre que $f(B)$ est ouvert.

Pour l'instant, on a montré que f est injective de B sur $f(B)$ et que $f(B)$ est ouvert ; pour finir la preuve du théorème, il nous reste à prouver que f^{-1} est différentiable sur $f(B)$. Rappelons que $\frac{1}{2}\|x - x'\| \leq \|f(x) - f(x')\|$ pour tout x, x' de B, et donc

$$\forall y, y' \in f(B) \ \frac{1}{2}\|f^{-1}(y) - f^{-1}(y')\| \leq \|y - y'\| .$$

Fixons $y \in f(B)$, et posons $x = f^{-1}(y)$. Pour $y' \in B$, notons aussi $x' = f^{-1}(y')$; on a

$$y' - y = f(x') - f(x) = df(x)(x' - x) + \|x' - x\|\varepsilon(x') ,$$

où $\varepsilon(x')$ tend vers 0 quand x' tend vers x. Puisqu'on sait que $\|x - x'\| \leq 2\|y - y'\|$, on peut réécrire cela sous la forme :

$$y' - y = df(x)(x' - x) + \|y' - y\|\varepsilon(f(y')) .$$

En appliquant $(df(x))^{-1}$ à cette égalité, on obtient, en utilisant la linéarité $(df(x))^{-1}$,

$$(df(x))^{-1}(y - y') = x' - x + \|y' - y\|(df(x))^{-1}(\varepsilon(f(y'))) .$$

Puisque $x' = f^{-1}(y')$, $x = f^{-1}(y)$, ceci s'écrit aussi sous la forme :

$$f^{-1}(y') = f^{-1}(y) + (df(x))^{-1}(y - y') + \|y' - y\|(df(x))^{-1}(\varepsilon(f(y'))) .$$

Comme $(df(x))^{-1}(\varepsilon(f(y')))$ tend vers 0 quand y' tend vers y, on vient de montrer que f^{-1} est différentiable en y, de différentielle égale à $(df(x))^{-1}$. Ceci conclut la démonstration.

$\square$

5.8 Théorème d'inversion globale

Corollaire 5.53 (Théorème d'inversion globale)**.** Soit U un ouvert de $\mathbb{R}^n$, et $f\colon U \to \mathbb{R}^n$ une fonction de classe C^1 et injective. Si la différentielle de f est inversible en tout point de U, alors $f(U)$ est ouvert et f est un difféomorphisme de classe C^1 de U sur $f(U)$.

Cet énoncé est au programme de l'U.E, mais sa démonstration, qui utilise le théorème d'inversion locale, ne l'est pas.

Remarque 5.54. L'hypothèse selon laquelle la différentielle de f est inversible en tout point de U est essentielle : sans cela le théorème est faux. Par exemple, l'application $x \mapsto x^3$ est une bijection de classe C^1 de $\mathbb{R}$ sur $\mathbb{R}$, mais sa fonction réciproque $x \mapsto x^{1/3}$ n'est pas dérivable en 0.

Démonstration. Montrons d'abord que $f(U)$ est ouvert : si $y \in f(U)$, alors il existe $x \in U$ tel que $f(x) = y$. Par le théorème d'inversion locale, il existe un ouvert $U_1 \ni x$ et un ouvert $V_1 \ni y$ tels que $f_{|U_1}$ soit un difféomorphisme de U_1 sur V_1. Alors V_1 est ouvert, contient y, et est contenu dans $f(U)$. Ceci prouve que $f(U)$ est ouvert.

Ensuite, le théorème d'inversion locale nous assure que f^{-1} est différentiable en $f(x)$ pour tout $x \in U$; autrement dit f^{-1} est différentiable en y pour tout $y \in f(U)$, et toutes les hypothèses définissant un difféomorphisme sont vérifiées. □

5.9 Fonctions implicites

Théorème 5.55 (Théorème des fonctions implicites)**.** *Soient $n, m \geq 1$ deux entiers, U un ouvert de $\mathbb{R}^n \times \mathbb{R}^m$ et $f\colon U \to \mathbb{R}^m$ une fonction de classe C^1. Soit $(x_0, y_0) \in U$ tel que $f(x_0, y_0) = 0$, et la différentielle de l'application $y \mapsto f(x_0, y)$ soit inversible en y_0. Alors il existe un ouvert O contenant x_0, un ouvert W contenant (x_0, y_0), et une application $\varphi\colon U_0 \to \mathbb{R}^m$ de classe C^1 tels que :*

$$\forall (x, y) \in W \; f(x, y) = 0 \Leftrightarrow x \in O \text{ et } \varphi(x) = y \; .$$

(en particulier $\varphi(x_0) = y_0$)

On dit alors que l'équation $f(x, y) = 0$ définit implicitement y en fonction de x *au voisinage de* (x_0, y_0).

Ce théorème se généralise aux fonctions de classe C^p, c'est-à-dire qu'on peutremplacer « C^1 »par « C^p »dans l'énoncé ci-dessus.

Rssayons d'expliquer ce que signifie ce théorème : les deux premières lignes signifient que, au voisinage de (x_0, y_0), l'équation $f(x, y) = \varphi(x_0, y_0)$ définit y comme une fonction de x (c'est la « fonction implicite »donnant sont nom au théorème). En utilisant la règle de la chaîne, on peut calculer les différentielles successives de φ en x_0, voir les exemples donnés après la preuve du théorème.

Si l'on considère l'exemple de l'application $f\colon (x, y) \mapsto x^2 + y^2 - 1$, on voit que, au voisinage de $(1, 0)$, cette équation ne peut pas définir y comme une fonction de x : il y a deux solutions pour y. La raison pour laquelle le théorème ne s'applique pas est que la dérivée partielle de f par rapport à y en ce point vaut 0...

La preuve, qui utilise le théorème d'inversion locale, est hors programme cette année.

Preuve du théorème des fonctions implicites. Appelons g l'application $y \mapsto f(x_0, y)$.

Il existe un ouvert $U_0 \ni x_0$ de $\mathbb{R}^n$, et un ouvert $V_0 \ni y_0$ de $\mathbb{R}^m$, tels que $U_0 \times V_0 \subseteq U$. Sur $U_0 \times V_0$, on définit $F(x, y) = (x, f(x, y))$. Alors F est de classe C^1, et sa matrice jacobienne en (x_0, y_0) vaut

$$M(f)(x_0, y_0) = \begin{pmatrix} I_n & 0 \\ * & M(g)(y_0) \end{pmatrix} \; .$$

(Ci-dessus, on a écrit la matrice par blocs, I_n étant la matrice identité de taille $n \times n$). Donc $M(f)(x_0, y_0)$ est inversible (son déterminant est égal à celui de $M(g)(y_0)$), et le théorème d'inversion locale assure l'existence d'un ouvert W contenant (x_0, y_0) et tel que la restriction de F à cet ouvert soit un difféomorphisme de classe C^1 sur l'ouvert $F(W)$.

Notons que pour tout $(x, y) \in F(W)$ on a $F^{-1}(x, y) = (x, \psi(x, y))$, et ψ est de classe C^1 sur $F(W)$. Appelons O l'ensemble (ouvert) formé par tous les x tels que $(x, 0) \in F(W)$ et définissons une fonction φ de classe C^1 sur O par $\varphi(x) = \psi(x, 0)$. Alors, pour tout (x, y) dans W, on a

$$f(x, y) = 0 \Leftrightarrow F(x, y) = (x, 0) \Leftrightarrow x \in O \text{ et } (x, y) = F^{-1}(x, 0) \Leftrightarrow x \in O \text{ et } y = \varphi(x) \; .$$

Ce qu'on vient d'écrire revient à

$$\forall (x,y) \in W\ f(x,y) = 0 \Leftrightarrow x \in O \text{ et } \varphi(x) = y\ .$$

$\square$

Discutons maintenant quelques exemples. Commençons par considérer la fonction $f\colon (x,y) \mapsto xy + \ln(xy) - 1$, définie sur $U = \{(x,y)\colon x > 0, y > 0\}$. Alors f est de classe C^∞ sur U, on a $f(1,1) = 0$, et $\frac{\partial f}{\partial y}(x,y) = x + \frac{1}{y}$ donc $\frac{\partial f}{\partial y}(1,1) = 2 \neq 0$. Par conséquent, le théorème des fonctions implicites nous permet d'affirmer que l'équation $xy + \ln(xy) = 1$ définit implicitement y comme une fonction φ de x au voisinage de $(1,1)$. Si on doit calculer la dérivée de φ en 1, on écrit : $x\varphi(x) + \ln(x\varphi(x)) = 1$, ce qui se dérive en

$$\varphi(x) + x\varphi'(x) + \frac{1}{x} + \frac{\varphi'(x)}{\varphi(x)} = 0\ .$$

Comme $\varphi(1) = 1$, l'équation ci-dessus donne $2 + 2\varphi'(1) = 0$, donc $\varphi'(1) = -1$. Ceci nous permettrait par exemple de trouver l'équation de la tangente à la courbe définie par l'équation $xy + \ln(xy) = 1$ au voisinage de $(1,1)$.

Considérons maintenant l'équation $2xy - z + 2xz^3 = 5$. Cette équation définit-elle implicitement z comme une fonction de (x,y) au voisinage de $(1,2,1)$? Pour le savoir, on pose $f(x,y,z) = 2xy - z + 2xz^3 - 5$, et on calcule la matrice jacobienne de f, qui vaut $\begin{pmatrix} 2y + 2z^3 & 2x & -1 + 6xz^2 \end{pmatrix}$. En $(1,2,1)$, cela donne $\begin{pmatrix} 4 & 2 & 5 \end{pmatrix}$. Puisque $5 \neq 0$, on voit que l'équation définit bien implicitement z comme fonction de (x,y) au voisinage de $(1,2,1)$. Notons $z = \varphi(x,y)$ et essayons de calculer $\frac{\partial^2 \varphi}{\partial y^2}(1,2)$. Pour cela on doit d'abord calculer les dérivées partielles de φ en $(1,2)$, ce qu'on fait en dérivant l'équation $f(x,y,\varphi(x,y)) = 0$, qui donne par la règle de la chaîne :

$$\begin{pmatrix} 2y + 2\varphi^3(x,y) & 2x & -1 + 6x\varphi^2(x,y) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial \varphi}{\partial x}(x,y) & \frac{\partial \varphi}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}\ .$$

On a donc le système suivant :

$$\begin{cases} 2y + 2\varphi^3(x) + (6x\varphi^2(x) - 1)\frac{\partial \varphi}{\partial x}(x,y) & = 0 \\ 2x + (6x\varphi^2(x) - 1)\frac{\partial \varphi}{\partial y}(x,y) & = 0 \end{cases}$$

On en déduit les formules suivantes :

$$\frac{\partial \varphi}{\partial x}(x,y) = \frac{2y + 2\varphi^3(x)}{1 - 6x\varphi^2(x)} \quad \text{et} \quad \frac{\partial \varphi}{\partial y}(x,y) = \frac{2x}{1 - 6x\varphi^2(x)}\ .$$

Ces deux équations nous donnent $\frac{\partial \varphi}{\partial x}(1,2) = -\frac{6}{5}$ et $\frac{\partial \varphi}{\partial y}(1,2) = -\frac{2}{5}$.

En redérivant par rapport à y l'équation donnant $\frac{\partial \varphi}{\partial y}(x,y)$, on obtient :

$$\frac{\partial^2 \varphi}{\partial y^2}(x,y) = 2x\frac{12x\frac{\partial \varphi}{\partial y}(x,y)}{(1 - 6x\varphi^2(x))^2}\ .$$

En $(1,2)$, on obtient $\frac{\partial^2 \varphi}{\partial y^2}(1,2) = -\frac{48}{125}$.

Un dernier exemple, pour une fonction f de trois variables, à valeurs dans $\mathbb{R}^2$. On considère le système d'équations suivant :

$$\begin{cases} 4xy + 2xz + y + 4y^2 & = 0 \\ x^3y + xz + yz - z & = 0 \end{cases}$$

Essaons de voir si ce système définit (y,z) comme une fonction de x au voisinage de $(0,0,0)$. Pour cela, on considère l'application $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ définie par

$$f(x,y,z) = (4xy + 2xz + y + 4y^2, x^3y + xz + yz - z)$$

Cette fonction est de classe C^∞, et sa matrice jacobienne en (x,y,z) vaut $\begin{pmatrix} 4y+2z & 4x+8y+1 & 2x \\ 3x^2y+z & x^3+z & x+y-1 \end{pmatrix}$.
En $(0,0,0)$, cette matrice vaut $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Pour décider si cette équation définit implicitement (y,z) comme fonction de x au voisinage de $(0,0,0)$, il nous faut donc décider si la matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est inversible ; son déterminant vaut -1, donc c'est bien le cas. On peut donc écrire $(y,z) = \varphi(x)$ au voisinage de $(0,0,0)$.

Essayons maintenant de calculer la dérivée de φ en 0 : si on note $\varphi = (\varphi_1, \varphi_2)$; le fait que $f(x, \varphi_1(x), \varphi_2(x)) = 0$ donne, par la règle de la chaîne, $df_{(x,\varphi_1(x),\varphi_2(x))}(1, \varphi_1'(x), \varphi_2'(x)) = 0$, d'où le système suivant :

$$\begin{cases} \varphi_1'(0) & = 0 \\ \varphi_2'(0) & = 0 \end{cases}$$

Autrement dit, $\varphi_1'(0) = \varphi_2'(0) = 0$. S'il avait fallu calculer les dérivées de φ en 0 à un ordre supérieur, alors on aurait dû écrire le système suivant (toujours donné par la règle de la chaîne) :

$$\begin{cases} 4\varphi_1(x) + 2\varphi_2(x) + (4x + 8\varphi_1(x) + 1)\varphi_1'(x) + 2x\varphi_2'(x) & = 0 \\ 3x^2\varphi_1(x) + \varphi_2(x) + (x^3 + \varphi_2(x))\varphi_1'(x) + (x + \varphi_1(x) - 1)\varphi_2'(x) & = 0 \end{cases}$$

Pour calculer $\varphi''(0)$, par exemple, il aurait fallu dériver ce système, puis l'écrire en 0 en y substituant le fait que $\varphi_1'(0) = \varphi_2'(0)$, pour obtenir un nouveau système de deux équations à deux inconnues $\varphi_1''(0)$, $\varphi_2''(0)$.

Plutôt que de continuer ces calculs, une dernière question : est-ce que l'équation $f(x,y,z) = (0,0)$ définit implicitement (x,y) en fonction de z au voisinage de $(0,0,0)$? Pour déterminer cela, il nous faut décider si la matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est inversible. Ce n'est clairement pas le cas (une colonne ne contient que des 0), donc l'équation ne définit pas implicitement (x,y) comme fonction de z au voisinage de $(0,0,0)$.

Chapitre 6

Intégrale double

6.1 Intégration sur un domaine compact du plan

Dans tout ce chapitre, on ne considèrera que des fonctions continues de deux variables. Commençons par rappeler comment on intègre une fonction continue sur un rectangle de côtés parallèles aux axes de coordonnées.

Théorème 6.1. *Soit $I = [a, b]$, $J = [c, d]$ deux segments de $\mathbb{R}$, et $R = I \times J$. Soit $f : R \to \mathbb{C}$ une fonction continue. On a*

$$\int_a^b \left(\int_c^d f(x, y)\, dy \right) dx = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy \ .$$

On appelle cette valeur commune l'intégrale de f sur R *et on la note* $\iint_R f(x, y)\, dxdy$.

La formule figurant dans le théorème ci-dessus est un cas particulier du *théorème de Fubini*, qu'on reverra plus loin.

Démonstration. Pour $x \in [a, b]$, $t \in [c, d]$, on définit $g(x, t) = \int_c^t f(x, y)\, dy$, et $G(t) = \int_a^b g(x, t)\, dx$. Comme f est continue sur R, on peut vérifier que g est elle aussi continue sur R (en utilisant le fait que f doit être uniformément continue sur R) ; de plus, le théorème fondamental de l'analyse appliqué à la fonction $t \mapsto g(x, t)$ (à x fixé) permet de voir que

$$\frac{\partial g}{\partial t}(x, t) = f(x, t) \ .$$

Ainsi, $\frac{\partial g}{\partial t}$ existe et et continue sur R. On peut appliquer le théorème de dérivabilité des intégrales à paramètre (fonction de deux variables continue et à dérivée partielle continue sur un produit de segments), et on obtient que G est dérivable et

$$G'(t) = \int_a^b \frac{\partial g}{\partial t}(x, t)\, dx = \int_a^b f(x, t)\, dx \ .$$

De même, pour $t \in [c, d]$, on peut définir $H(t) = \int_c^t \left(\int_a^b f(x, y)\, dx \right) dy$. On commence par appliquer le théorème de continuité des intégrales à paramètre pour voir que $y \mapsto \int_a^b f(x, y)\, dx$ est une fonction continue sur $[c, d]$, à laquelle on peut donc appliquer le théorème fondamental de l'analyse et obtenir

$$H'(t) = \int_a^b f(x, t)\, dx$$

On voit donc que $G'(t) = H'(t)$ pour tout $t \in [c, d]$; comme de plus on a $G(c) = H(c) = 0$, on en conclut que $G(t) = H(t)$ pour tout $t \in [c, d]$; en particulier $G(d) = H(d)$, et c'est l'égalité qu'on souhaitait démontrer. □

On sait maintenant comment intégrer des fonctions continues sur un rectangle $[a, b] \times [c, d]$; on retrouve les propriétés usuelles de l'intégrale, en particulier la linéarité et la positivité.

Comment faire pour intégrer sur un domaine plus général ? On est assez restreint par notre définition de l'intégrale ; on ne va considérer que des domaines d'intégration très particuliers.

Définition 6.2. On dit que $A \subseteq \mathbb{R}^2$ est une *partie élémentaire compacte* s'il existe deux segments $[a,b],[c,d] \subset \mathbb{R}$ et des fonctions continues $\phi_1, \phi_2 \colon [a,b] \to \mathbb{R}$, $\psi_1, \psi_2 \colon [c,d] \to \mathbb{R}$ telles que l'on ait *à la fois*

- $\forall x \in]a,b[\ \phi_1(x) < \phi_2(x)$ et $A = \{(x,y) \colon a \leq x \leq b \text{ et } \phi_1(x) \leq y \leq \phi_2(x)\}$.
- $\forall y \in]c,d[\ \psi_1(y) < \psi_2(y)$ et $A = \{(x,y) \colon c \leq y \leq d \text{ et } \psi_1(y) \leq c \leq \psi_2(y)\}$.

Remarquons qu'alors A est contenu dans le rectangle $[a,b] \times [c,d]$; de plus A est fermé, donc A est à la fois fermé et borné, c'est-à-dire compact.

Intuitivement, la définition signifie que A est obtenu en traçant une courbe fermée dans le plan, qui ne se recoupe pas, de telle façon qu'une ligne horizontale passant par un point à l'intérieur de la courbe rencontre la courbe en exactement deux points, et de même pour toute ligne verticale (faites un dessin !).

On peut assez facilement étendre la définition des intégrales de fonctions continues sur un rectangle aux parties élémentaires compactes.

Théorème 6.3. *Soit $A \subseteq \mathbb{R}^2$ une partie élémentaire compacte (on reprend les notations de la définition ci-dessus), et $f \colon A \to \mathbb{C}$ une fonction continue sur A.*

Soit $\hat{f}$ la fonction définie sur $[a,b] \times [c,d]$ par $\hat{f}(x,y) = \begin{cases} f(x,y) & \textit{si } (x,y) \in A \\ 0 & \textit{sinon} \end{cases}$

*Alors les intégrales $\int_a^b \left(\int_c^d \hat{f}(x,y)\,dy \right) dx$ et $\int_c^d \left(\int_a^b \hat{f}(x,y)\,dx \right) dy$ existent et sont égales ; on appelle cette valeur commune l'*intégrale de f sur A, *et on la note $\iint_A f(x,y)\,dxdy$.*

Une reformulation importante de l'égalité ci-dessus est la formule suivante, dite *formule de Fubini* : pour une fonction f continue sur une partie élémentaire compacte A, toujours en utilisant les mêmes notations que ci-dessus, on a

$$\iint\limits_A f(x,y)\,dxdy = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dy \right) dx = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx \right) dy$$

Cette formule est très importante en pratique, puisqu'elle permet de ramener le calcul d'une intégrale double à deux intégrales simples successives.

Ebauche de preuve. On commence par supposer que f est à valeurs réelles positives (le cas général se déduit facilement de ce cas particulier). On peut étendre f à une fonction continue $g \colon [a,b] \times [c,d] \to \mathbb{R}^+$ en posant, pour $(x,y) \in [a,b] \times [c,d]$:

$$g(x,y) = \begin{cases} f(x,\phi_2(x)) & \text{si } y > \phi_2(x) \\ f(x,y) & \text{si } \phi_1(x) \leq y \leq \phi_2(x) \\ f(x,\phi_1(x)) & \text{si } y < \phi_1(x) \end{cases}$$

Ensuite, on fixe $\varepsilon > 0$. On peut trouver deux fonctions continues α, β sur $\mathbb{R}^2$ qui aient les propriétés suivantes :

- $\forall (x,y)\ \alpha(x,y) \leq \mathbf{1}_A(x,y) \leq \beta(x,y)$ (où $\mathbf{1}_A$ désigne la fonction caractéristique de A).
- $\iint\limits_{a,b] \times [c,d]} (\beta - \alpha)(x,y)\,dxdy \leq \varepsilon$.

(Prouver l'existence de ces deux fonctions n'est pas évident, c'est en cela en particulier que ce qui est présenté ici n'est qu'une ébauche de preuve)

Puisque $\hat{f} = g \times \mathbf{1}_A$, nos définitions entraînent que, pour tout $(x,y) \in [a,b] \times [c,d]$, on a $\alpha(x,y)g(x,y) \leq \hat{f}(x,y) \leq \beta(x,y)g(x,y)$. En utilisant la positivité de l'intégrale simple, on en déduit les inégalités suivantes :

$$\begin{aligned} \int_a^b \left(\int_c^d \alpha(x,y)g(x,y)\,dy \right) dx \quad &\leq \int_a^b \left(\int_c^d \hat{f}(x,y)\,dy \right) dx \leq \int_a^b \left(\int_c^d \beta(x,y)g(x,y)\,dy \right) dx \\ \int_c^d \left(\int_a^b \alpha(x,y)g(x,y)\,dx \right) dy \quad &\leq \int_c^d \left(\int_a^b \hat{f}(x,y)\,dx \right) dy \leq \int_c^d \left(\int_a^b \beta(x,y)g(x,y)\,dx \right) dy \end{aligned}$$

Comme αg et βg sont continues, les termes de gauche de chaque ligne sont égaux à $\iint_{[a,b]\times[c,d]} \alpha(x,y)g(x,y)\,dxdy$,et ceux de droite sont égaux à $\iint_{[a,b]\times[c,d]} \beta(x,y)g(x,y)\,dxdy$. On en déduit l'inégalité

$$\left| \int_a^b \left(\int_c^d \hat{f}(x,y)\,dy \right) dx - \int_c^d \left(\int_a^b \hat{f}(x,y)\,dx \right) dy \right| \leq \iint\limits_{[a,b]\times[c,d]} (\beta - \alpha)(x,y)g(x,y)\,dxdy\ .$$

Si l'on appelle M le maximum de la fonction continue g sur le compact $[a,b]\times[c,d]$, on obtient, grâce à la positivité de l'intégrale des fonctions continues sur un rectangle, que

$$\left|\int_a^b \left(\int_c^d \hat{f}(x,y)\,dy\right)dx - \int_c^d \left(\int_a^b \hat{f}(x,y)\,dx\right)dy\right| \leq M\cdot \iint\limits_{[a,b]\times[c,d]} (\beta-\alpha)(x,y)\,dxdy \leq M\varepsilon\ .$$

Ceci étant vrai pour tout $\varepsilon > 0$, les deux intégrales itérées doivent être égales.

Le cas d'une fonction à valeurs réelles se déduit du résultat qu'on vient d'obtenir en posant $f^+ = \max(f,0)$, $f^- = \max(-f,0)$; alors f^+ et f^- sont continues, à valeurs positives, et $f = f^+ - f^-$. Ensuite, le cas d'une fonction à valeurs complexes s'obtient en décomposant en partie réelle et partie imaginaire. □

Les parties sur lesquelles on voudrait pouvoir calculer des intégrales doubles ne sont pas toujours des parties élementaires ; ceci motive la définition suivante.

Définition 6.4. $A \subseteq \mathbb{R}^2$ est une *partie simple compacte* s'il existe des parties élémentaires compactes $A_1,\ldots,A_n$ d'intérieurs deux à deux disjoints et telles que $A = \bigcup\limits_{i=1}^n A_i$.

Pour toute fonction continue f sur A, on pose alors

$$\iint\limits_A f(x,y)\,dxdy = \sum_{i=1}^n \iint\limits_{A_i} f(x,y)\,dxdy\ .$$

Remarquons qu'il n'est pas clair a priori que la définition ci-dessus soit indépendante de la décomposition de A en réunion de parties élémentaires compactes ; on admet que c'est le cas. L'idée est que, du moment qu'on peut découper A en une réunion finie de parties sur lesquelles on sait définir l'intégrale d'une fonction continue, alors on sait aussi définir l'intégrale d'une fonction continue sur A.

On retrouve les propriétés habituelles de l'intégrale :

- Pour toute partie simple compacte A, et toute fonction f continue et à valeurs positives sur A, on a $\iint_A f(x,y)\,dxdy \geq 0$.. (positivité)
- Si A est une partie simple compacte, $f,g\colon A\to\mathbb{C}$ sont des fonctions continues et $\alpha,\beta\in\mathbb{C}$ alors
$$\iint\limits_A (\alpha f+\beta g)(x,y)\,dxdy = \alpha\iint\limits_A f(x,y)\,dxdy + \beta\iint\limits_A f(x,y)\,dxdy \qquad \text{(linéarité)}$$
En particulier, pour toute fonction continue $f\colon A\to\mathbb{C}$ on a $\iint_A f(x,y)\,dxdy = \iint_A \mathrm{Ré}(f)(x,y)\,dxdy + i\iint_A \mathrm{Im}(f)(x,y)\,dxdy$.
- Pour toute partie simple compacte A et toute fonction f continue sur A et à valeurs complexes, on a $\left|\iint_A f(x,y)\,dxdy\right| \leq \iint_A |f(x,y)|\,dxdy$. (inégalité triangulaire)

Notons que, comme dans le cas des fonctions d'une variable, l'inégalité triangulaire est une conséquence de la positivité et de la linéarité de l'intégrale. Notons également la *croissance de l'intégrale par rapport au domaine* : si f est une fonction continue à valeurs positives sur un domaine simple compact A_1, et A_2 est un domaine simple compact contenu dans A_1, alors $\iint_{A_1} f(x,y)\,dxdy \leq \iint_{A_2} f(x,y)\,dxdy$.

Définition 6.5. Pour toute partie simple compacte A, on définit l'*aire* de A par la formule

$$\mathrm{aire}(A) = \iint\limits_A dxdy\ .$$

Ceci donne une définition formelle de l'aire, qui étend celle qui était déjà connue pour les triangles, les disques, les parallélogrammes, etc. Mais est-on capable de retrouver les formules connues pour les aires de ces figures à partir de notre définition de l'intégrale ? Pour l'instant, le seul moyen qu'on connaît pour calculer l'aire d'un disque de rayon 1 est de calculer l'intégrale

$$\int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy\,dx = \int_{-1}^1 2\sqrt{1-x^2}dx\ .$$

Cette intégrale n'est à première vue pas évidente à calculer (c'est d'ailleurs un bon exercice ; un changement de variables serait du meilleur effet) ; de même, il n'est pas aisé de calculer, par exemple, l'aire d'un parallélogramme en utilisant des intégrales itérées. C'est pourquoi on utilise fréquemment des changements de variables, qui sont une technique fondamentale : fréquemment, on est amené à faire un changement de variables pour ramener un domaine d'allure compliquée à un domaine plus simple, typiquement un rectangle.

Théorème 6.6 (Théorème de changement de variables)**.** *Soit D_1, D_2 deux parties simples compactes, et $\varphi\colon D_1 \to D_2$ une bijection continue telle que φ soit un difféomorphisme de classe C^1 de l'intérieur de D_1 sur l'intérieur de D_2. Alors, pour toute fonction continue $f\colon D_2 \to \mathbb{C}$, on a*

$$\iint\limits_{D_2} f(x,y)\,dxdy = \iint\limits_{D_1} f\circ\varphi(x,y)\,|det(Jac(\varphi)(x,y))|\,dxdy\ .$$

Démonstration. Admis. □

Remarque 6.7.

- Les hypothèses sont en particulier vérifiées quand $\varphi\colon U_1 \to U_2$ est un C^1-difféomorphisme défini sur un ouvert contenant D_1 et $\varphi(D_1) = D_2$; un cas particulier important est le cas où φ est une application linéaire bijective.
- Le déterminant de la matrice jacobienne de φ s'appelle *déterminant jacobien* de φ ; on le notera $J\varphi(x,y)$ dans la suite.
- Il est important de ne pas oublier la valeur absolue dans la formule !

Un exemple important : le *changement de variables en coordonnées polaires*. On définit

$$\varphi\colon \begin{cases} [0,+\infty[\times[0,2\pi[\to \mathbb{R}^2 \\ \varphi(r,\theta) = (r\cos(\theta), r\sin(\theta)) \end{cases}$$

Alors φ est une surjection de classe C^∞ de $[0,+\infty[\times[0,2\pi[$ sur $\mathbb{R}^2$, et φ est un C^∞-difféomorphisme de $]0,+\infty[\times]0,2\pi[$ sur $\mathbb{R}^2\setminus\{0\}$. Son déterminant jacobien en (r,θ) est le déterminant de la matrice $\begin{pmatrix}\cos(\theta) & -r\sin(\theta)\\ \sin(\theta) & r\cos(\theta)\end{pmatrix}$, c'est-à-dire r.

Utilisons φ pour calculer $\iint_D (x+y)^2\,dxdy$, où D est le disque de centre 0 et de rayon 1. Pour être rigoureux lors de notre première application d'un changement de variables, écrivons $D = D^+ \cup D^-$, où D est la partie du disque au-dessus de l'axe des abscisses et D^- la partie en-dessous de cet axe. Alors φ est un C^1-difféomorphisme de $]0,1[\times]0,\pi[$ sur l'intérieur de D^+, et un C^1-difféomorphisme de $]0,1[\times]\pi,2\pi[$ sur l'intérieur de D^-. On obtient donc :

$$\begin{aligned}
\iint\limits_D (x+y)^2\,dxdy &= \iint\limits_{D^+}(x+y)^2\,dxdy + \iint\limits_{D^-}(x+y)^2\,dxdy \\
&= \iint\limits_{[0,1]\times[0,\pi]} (r\cos(\theta)+r\sin(\theta))^2 r\,drd\theta + \iint\limits_{[0,1]\times[\pi,2\pi]} (r\cos(\theta)+r\sin(\theta))^2 r\,drd\theta \\
&= \iint\limits_{[0,1]\times[0,2\pi]} (r\cos(\theta)+r\sin(\theta))^2 r\,drd\theta \\
&= \int_0^1 \left(\int_0^{2\pi} r^3(\cos^2(\theta)+\sin^2(\theta)+2\sin(\theta)\cos(\theta))\,d\theta\right) dr \\
&= \int_0^1 r^3\left(\int_0^{2\pi}(1+\sin(2\theta))d\theta\right) dr \\
&= \int_0^1 2\pi r^3\,dr \\
&= \frac{\pi}{2}\ .
\end{aligned}$$

On n'a pas mis de valeur absolue autour de r dans le calcul ci-dessus (passage de $dxdy$ à $rdrd\theta$), parce que r est positif et donc $|r| = r$.

Le calcul ci-dessus est très détaillé, trop sans doute ; en pratique, lors d'un changement de variable en coordonnées polaires, on pourra aller plus vite, en ne mentionnant pas l'étape de découpage de D en D^+, D^- par exemple. Le même calcul, rédigé de manière plus rapide mais suffisamment détaillée, donnerait :

$$\begin{aligned}
\iint_D (x+y)^2\,dxdy &= \iint_{[0,1]\times[0,2\pi]} (r\cos(\theta)+r\sin(\theta))^2 r\,drd\theta \text{ (passage en coordonnées polaires)}\\
&= \int_0^1 \left(\int_0^{2\pi} r^3(\cos^2(\theta)+\sin^2(\theta)+2\sin(\theta)\cos(\theta))\,d\theta\right) dr\\
&= \int_0^1 r^3 \left(\int_0^{2\pi} (1+\sin(2\theta))d\theta\right) dr\\
&= \int_0^1 2\pi r^3\,dr\\
&= \frac{\pi}{2}\ .
\end{aligned}$$

Exercice 6.8. 1. Utiliser un changement de variables en coordonnées polaires pour calculer l'aire du disque de centre 0 et de rayon R.

2. Soit $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ une application linéaire inversible de matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et A le parallélogramme $f([0,1]^2)$. Montrer que l'aire de A est égale à $|\det(M)|$. Que pensez-vous de cette formule dans le cas où f n'est pas inversible ?

6.2 Intégrales doubles sur des ouverts du plan

On va finir ce cours en définissant, sans démonstrations, une notion d'intégrale double pour des fonctions continues sur un ouvert du plan. Ces intégrales sont une extension au plan des intégrales généralisées. A cause de l'absence d'une relation d'ordre « naturelle »sur $\mathbb{R}^2$, on ne peut pas avoir de théorie satisfaisante des intégrales semi-convergentes (i.e. convergentes mais pas absolument convergentes), et on a seulement un analogue des intégrales absolument convergentes. Ces intégrales doubles sont en particulier utiles pour calculer des intégrales généralisées sur un intervalle de $\mathbb{R}$ (horresco referens).

Commençons par observer que tout ouvert O du plan peut s'écrire sous la forme $O = \bigcup_{n\in\mathbb{N}} R_n$, où les R_n sont des rectangles fermés (d'intérieurs deux à deux disjoints si on veut). Commençons par une définition théorique qui ne nous servira pas en pratique.

Définition 6.9. Soit O un ouvert de $\mathbb{R}^2$, et $f\colon O \to \mathbb{R}^+$ une fonction continue et à valeurs positives. On dit que f est *intégrable* si

$$M = \sup\{\iint_A f(x,y)\,dxdy\colon A \subseteq O \text{ est une réunion finie de rectangles fermés}\} < \infty\ .$$

On pose alors $\iint_O f(x,y)\,dxdy = M$.

Si $f\colon O \to \mathbb{C}$ est une fonction continue, on dit que f est intégrable si $|f|$ est intégrable.

Remarquons tout de suite que cela ne nous dit pas du tout comment calculer $\iint_O f(x,y)\,dxdy$!

Proposition 6.10. *Si f est continue, intégrable et à valeurs positives sur un ouvert O de $\mathbb{R}^2$, alors pour toute suite croissante (A_n) de parties simples compactes telles que $\cup A_n = O$, la suite $\iint_{A_n} f(x,y)\,dxdy$ converge vers $\iint_O f(x,y)\,dxdy$.*

Si f est continue et intégrable sur un ouvert O de $\mathbb{R}^2$, alors pour toute suite (A_n) de parties simples compactes telles que $\cup A_n = 0$, la suite $\iint_{A_n} f(x,y)\,dxdy$ converge ; cette limite ne dépend pas de la suite (A_n), et on pose

$$\iint_O f(x,y)\,dxdy = \lim_n \iint_{A_n} f(x,y)\,dxdy\ .$$

En particulier, dans le cas d'une fonction f continue et intégrable sur $\mathbb{R}^2$ tout entier, on a par exemple

$$\iint_{\mathbb{R}^2} f(x,y)\,dxdy = \lim_{n\to+\infty} \iint_{[-n,n]^2} f(x,y)\,dxdy\ .$$

Notons que les propriétés habituelles de l'intégrale (positivité, linéarité, croissance par rapport au domaine) sont toujours vérifiées. On a maintenant un premier moyen de calculer des intégrales doubles sur des ouverts du plan, mais ce critère ne s'applique que si l'on a d'abord vérifié que f est intégrable, ce qui peut se faire en commençant par étudier $|f|$.

Proposition 6.11. *Si O est un ouvert de $\mathbb{R}^2$, g est une fonction continue à valeurs positives intégrable sur O, et $f\colon O \to \mathbb{C}$ est une fonction continue telle que $|f(x)| \leq g(x)$ pour tout $x \in O$, alors f est intégrable sur O.*

Un cas particulier est particulièrement important : celui où O est une « bande », i.e. un produit de deux intervalles ouverts. Alors on peut appliquer la formule de Fubini.

Théorème 6.12 (Théorème de Fubini pour les produits d'intervalles ouverts)**.** *Soit I, J deux intervalles ouverts du plan.*

- *Soit $f\colon I \times J \to \mathbb{C}$ une fonction continue. Si, pour tout $x \in I$, la fonction $y \mapsto g(x, y)$ est d'intégrale absolument convergente sur J, et si la fonction $g\colon x \mapsto \int_J f(x, y)\, dy$ est continue par morceaux et d'intégrale absolument convergente sur I, alors f est intégrable sur $I \times J$ et on a*

$$\iint_{I\times J} f(x,y)\, dxdy = \int_I \left(\int_J f(x,y)\, dy \right) dx \ .$$

- *En particulier, si f est à valeurs positives, si pour tout $x \in I$ l'intégrale $\int_J f(x, y)\, dy$ converge et si l'intégrale $\int_I \left(\int_J f(x, y)\, dy\right) dx$ converge, alors f est intégrable et*

$$\iint_{I\times J} f(x,y)\, dxdy = \int_I \left(\int_J f(x,y)\, dy \right) dx \ .$$

On obtiendrait un énoncé similaire en échangeant les rôles de x et y ci-dessus (c'est-à-dire en intégrant d'abord par rapport à x puis par rapport à y). En particulier, si f est intégrable sur $I \times J$, et toutes les intégrales apparaissant dans la formule ci-dessous sont absolument convergentes, alors on a

$$\int_I \left(\int_J f(x,y)\, dy \right) dx = \int_J \left(\int_I f(x,y)\, dx \right) dy \ .$$

Bien sûr, tous les ouverts ne sont pas des produits d'intervalles ouverts : dans le cas d'un ouvert plus général, on peut soit essayer de l'écrire comme une réunion croissante de domaines simples compacts, soit utiliser le théorème de changement de variables, sous la forme suivante.

Théorème 6.13 (Théorème de changement de variables pour des ouverts de $\mathbb{R}^2$)**.** *Soit U, V deux ouverts de $\mathbb{R}^2$, $\varphi\colon U \to V$ un difféomorphisme de classe C^1 et $f\colon V \to \mathbb{C}$ une fonction continue. Alors f est intégrable sur V si, et seulement si, $(x, y) \mapsto f \circ \varphi(x, y)|J\varphi(x, y)|$ est intégrable sur U, et on a l'égalité*

$$\iint_V f(x,y)\, dxdy = \iint_U f(\varphi(x,y))|J\varphi(x,y)|\, dxdy \ .$$

(On rappelle que $J\varphi(x, y)$ désigne le déterminant de la matrice jacobienne de φ en (x, y) ; remarquons que ce déterminant est toujours non nul puisque φ est un difféomorphisme).

Index