Analyse et Probabilités en Deug Mass 2, semestre 2

Marc et Francine DIENER

25 mai 2001

Table des matières

Chapitre 1

Séries numériques

Ce chapitre traite d'un exemple de suites particulièrement fréquent dans la pratique. Il ne nécessite que d'avoir compris la notion de suite réelle convergente ou divergente.

1.1 Introduction et remarques élémentaires

Une manière de calculer e avec une précision arbitrairement bonne est d'utiliser la formule

$$\begin{aligned} e &= 1+1+\frac{1}{2}+\frac{1}{6}+\frac{1}{24}+\frac{1}{120}+\ldots && (1.1)\\ &= \frac{1}{0\,!}+\frac{1}{1\,!}+\frac{1}{2\,!}+\frac{1}{3\,!}+\frac{1}{4\,!}+\frac{1}{5\,!}+\ldots+\frac{1}{n\,!}+\ldots && (1.2)\end{aligned}$$

c'est-à-dire en calculant l'approximation $s_N=\sum_{n=0}^{N}\frac{1}{n\,!}$. En d'autres termes, la suite des $(s_N)_{N\geq 0}$ tend vers le nombre e, nombre que nous ne connaitrons jamais mieux que par l'information que la suite $(s_N)_{N\geq 0}$ tend vers e. En fait, souvent des nombres irrationnels utiles nous sont accessibles par le fait qu'ils sont limites de suites $(s_N)_{N\geq 0}$ se présentant sous la forme $s_N=\sum_{n=0}^{N}u_n$, d'où l'intérêt d'étudier des *critères de convergence* de suites se présentant précisément ainsi dites *séries associées à* la suite $(u_n)_{n\geq 0}$, et permettant de préciser la *nature* (convergente ou divergente) de la série.

Notons que toute suite $(s_N)_{N\geq 0}$ peut être vue comme la série associée à une suite $(u_n)_{n\geq 0}$: il suffit de poser $u_0=s_0$, puis $u_n=s_n-s_{n-1}$; il se trouve simplement que pour beaucoup d'exemples importants, l'écriture sous forme de série est celle qui s'introduit naturellement (voir l'études des *séries entières*, plus loin).

Il peut biensûr se produire que la série se présente sous la forme $s'_n=\sum_{n=n_0}^{N}u_n$, où n_0 est différent de 0, par exemple $n_0=1$; la différence entre s_N et s'_N est une fois pour toute la somme $d=\sum_{n=0}^{n_0-1}u_n$, et cette différence n'affecte pas l'étude de la convergence ; si s et s' désignent les limites respectives, on aura simplement $s=d+s'$.

Pour l'étudiant(e), souvent l'étude des séries est la première occasion de faire personnellement un usage intensif de la notation $\sum_{n=n_0}^{N}u_n$ au lieu de la notation plus évocatrice mais moins concise $u_{n_0}+u_{n_0+1}+\ldots+u_N$; si tel est le cas, il n'est pas inutile de méditer aux identités suivantes, où $[p]$ désigne la "partie entière" du réel p, c'est-à-dire le plus grand entier inférieur ou égal à p :

$$\sum_{n=n_0}^{N}\lambda u_n=\lambda\sum_{n=n_0}^{N}u_n=\lambda\sum_{p=n_0}^{N}u_p=\lambda\sum_{i=n_0}^{N}u_i=\ldots\quad(\lambda\text{ nombre quelconque})$$

$$\sum_{n=n_0}^{N}(u_n+v_n)=\left(\sum_{n=n_0}^{N}u_n\right)+\left(\sum_{n=n_0}^{N}v_n\right)\quad;\quad\sum_{n=0}^{N}u_n=\sum_{p=0}^{[N/2]}u_{2p}+\sum_{p=0}^{[(N+1)/2]-1}u_{2p+1}$$

$$\left(\sum_{m=m_0}^{M}u_m\right)\left(\sum_{n=n_0}^{N}v_n\right)=\sum_{m=m_0}^{M}\left(\sum_{n=n_0}^{N}(u_mv_n)\right)=\sum_{n=n_0}^{N}\left(\sum_{m=m_0}^{M}(u_mv_n)\right).$$

Si l'on conçoit la suite $(s_N)_{N\geq 0}$ comme une suite d'approximations de sa limite s, on voit que $s_0(=u_0)$ est une "première approximation" de s, puis u_1 est une "correction" à ajouter à l'approximation s_0, etc... ; nous voyons donc que pour que la série s_N converge ; *il faut que la suite (des corrections)* $(u_n)_{n\geq 0}$ *tende*

Figure 1.1: Majoration et minorations de la série de Riemann ; l'aire de chacun des rectangles est égale à $u_n := \frac{1}{n^\alpha} =: f_\alpha(n)$.

vers 0 (nous verrons que cette condition nécessaire à la convergence de la série ne suffit pas à assurer la convergence de la série.)

Souvent nous désignerons la suite $(s_N)_{N\geq 0}$ par le symbole $\sum_{n\geq 0} u_n$, qui ne préjuge pas de la convergence de cette série. Quand elle converge, on notera $\sum_{n=0}^{\infty} u_n$ la limite de cette série, qui est souvent appelée *somme de la série* considérée. Attention, ce nombre est une limite et n'est pas, comme s_N, une somme finie de termes u_n, et cette écriture n'autorise plus, à priori, les manipulations élémentaires que nous avons rappelées ; c'est pourquoi, lorsque la suite $(u_n)_{n\geq 0}$ (, et donc le nombre $\sum_{n=0}^{\infty} u_n$,) sont standard, nous choisirons souvent $N = \omega$ i-grand, et ainsi $\sum_{n=0}^{\omega} u_n$ sera une somme finie i-voisine du nombre $\sum_{n=0}^{\infty} u_n$.

Néanmoins, si $\sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty} v_n$ ont bien un sens, il en est de même de $\sum_{n=0}^{\infty} \lambda u_n$ et $\sum_{n=0}^{\infty}(u_n+v_n)$, et on a $\sum_{n=0}^{\infty} \lambda u_n = \lambda \sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty}(u_n+v_n) = (\sum_{n=0}^{\infty} u_n) + (\sum_{n=0}^{\infty} v_n)$ (pourquoi ?) ; en revanche, par exemple si $u_n = \frac{(-1)^n}{n}$, on n'a pas $\sum_{n=0}^{\infty} u_n = \sum_{p=0}^{\infty} u_{2p} + \sum_{p=0}^{\infty} u_{2p+1}$: cette bizarerie deviendra compréhensible dès que nous aurons étudiés deux familles d'exemples essentielles auxquelles est consacrée la prochaine section.

1.2 Exemples et critère fondamentaux

1.2.1 Séries géométriques

Pour $q \in \mathbb{R}$ quelconque, soit $u_n = q^n$; on appelle *série géométrique* la série $\sum_{n\geq 0} q^n$. On sait (et il est facile de le vérifier par une preuve par récurrence) que $s_N := \sum_{n=0}^{N} q^n = \frac{1-q^{N+1}}{1-q}$ si $q \neq 1$ et $s_N = N+1$ si $q = 1$, d'où l'on déduit immédiatement que $\sum_{n\geq 0} q^n$ converge (vers $1/(1-q)$) si $|q| < 1$, et ne converge pas sinon.

Ce premier exemple, capital, a pu être traité facilement car on disposait d'une expression commode pour réécrire s_N. En fait, ceci est exceptionnel ; l'exemple suivant, également capital, demandera un peu plus de “bricolage”.

1.2.2 Séries de Riemann

On appelle *fonction zéta de Riemann*, notée $\alpha \mapsto \zeta(\alpha)$, la somme, lorsqu'elle existe, de la série $\sum_{n\geq 1} \frac{1}{n^\alpha}$, dite *série de Riemann*. L'étude de la convergence des séries de Riemann ne s'effectue plus par un calcul explicite, mais par un bel encadrement par des intégrales qui, elles, se calculent. Notons $f_\alpha(x) := \frac{1}{x^\alpha}$; on a donc $u_n := \frac{1}{n^\alpha} \leq \frac{1}{x^\alpha} =: f_\alpha(x)$ si et seulement si $x \leq n$, et donc $\int_n^{n+1} f_\alpha(x)dx \leq u_n \leq \int_{n-1}^{n} f_\alpha(x)dx$, d'où

$$\begin{aligned}
\int_1^{n+1} f_\alpha(x)dx &= \int_1^2 f_\alpha(x)dx + \int_2^3 f_\alpha(x)dx + \ldots + \int_n^{n+1} f_\alpha(x)dx && (1.3)\\
&\leq u_1 + u_2 + \ldots + u_n =: s_n && (1.4)\\
&\leq 1 + \int_1^2 f_\alpha(x)dx + \int_2^3 f_\alpha(x)dx + \ldots + \int_{n-1}^{n} f_\alpha(x)dx && (1.5)\\
&= 1 + \int_1^n f_\alpha(x)dx. && (1.6)
\end{aligned}$$

L'avantage ici de l'intégrale sur la série, est qu'on connait une primitive F_α de f_α : il suffit de poser $F_\alpha(x) := \frac{1}{(1-\alpha)x^{\alpha-1}}$ si $\alpha \neq 1$ et $F_\alpha(x) := \ln(x)$ si $\alpha = 1$. En réécrivant les majorations (1.3-1.6) nous obtenons

$$F_\alpha(N+1) - F_\alpha(1) \leq \sum_{n=1}^{N} \frac{1}{n^\alpha} \leq 1 + F_\alpha(N) - F_\alpha(1),$$

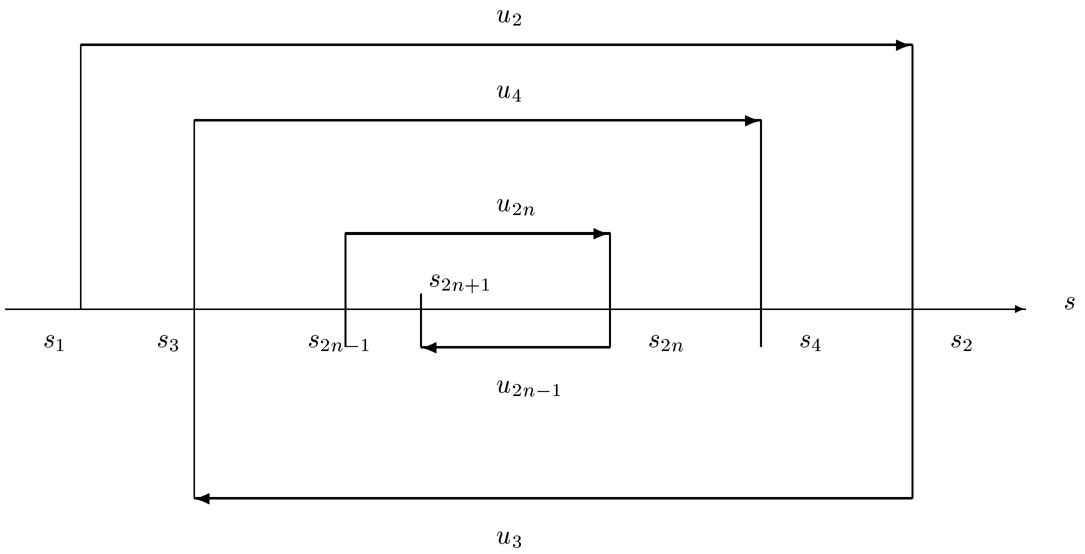
d'où l'on déduit que la série $\sum_{n\geq 1} \frac{1}{n^\alpha}$ tend vers l'infini si $\alpha \leq 1$ (utiliser l'inégalité de gauche), et que si $\alpha > 1$, la série $\sum_{n\geq 1} \frac{1}{n^\alpha}$ est une suite majorée (utiliser l'inégalité de droite), croissante puisque les "corrections" $u_n := \frac{1}{n^\alpha}$ sont positives, et donc convergente.

Résumons cette étude d'exemples fondamentaux.

Théorème 1.1 *La série géométrique $\sum_{n\geq 0} q^n$ converge si et seulement si $|q| < 1$*
La série de Riemann $\sum_{n\geq 1} \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

1.2.3 Série harmonique alternée

Comme nous l'avons vu lors de l'étude des séries de Riemann, la série $\sum_{n\geq 1} 1/n$, dite *série harmonique*, n'est pas convergente. Considérons à présent la série $\sum_{n\geq 1}(-1)^n/n$, dite *série harmonique alternée.* Cette série, elle, est convergente. En effet, en calculant les premiers temes de la suite $s_N := \sum_{n=1}^{N} \frac{(-1)^n}{n}$, on obtient $s_1 = -1$, $s_2 = s_1 + \frac{1}{2}$, $s_3 = s_2 + \frac{-1}{3} = s_1 + \frac{-1}{3} + \frac{+1}{2} = s_1 + \frac{1}{6}$, $s_4 = s_3 + \frac{+1}{4} = s_2 + \frac{1}{4} + \frac{-1}{3} = s_2 + \frac{-1}{12}$, et on comprend que la sous-suite $s'_K := s_{2K+1}$ des termes d'indices impairs constituent une suite croissante, et que la sous-suite $s''_K := s_{2K}$ des termes d'indice pair constitue une suite décroissante ; enfin $s'_K < s''_K$, puisque $s''_K = s_{2K} = s_{2K-1} + \frac{(-1)^{2K}}{2K} > s_{2K-1} = s'_K$.



On en déduit que si $q > p > 2K$, on a à la fois $s'_K < s_q < s''_K$ et $s'_K < s_p < s''_K$, et donc $|s_q - s_p| \leq s''_K - s'_K = s_{2K} - s_{2K-1} = \frac{(-1)^K}{2K} = \frac{1}{2K}$. A présent, si p et q sont i-grands, nous pouvons choisir K i-grand, et donc $|s_q - s_p| \simeq 0$, ce qui montre que la suite $(s_N) := \sum_{n=1}^{N} \frac{(-1)^n}{n}$ est de Cauchy, donc convergente.

On aura compris qu'on a, plus généralement, la proposition (1.2) suivante, que l'on peut également prouver à l'aide du critère d'Abel (1.9) que nous étudierons plus loin.

Proposition 1.2 *Pour toute série* alternée $\sum_{n\geq 0} u_n$*, c'est-à-dire tel que u_{n+1} est de signe opposé à celui de u_n, si la suite $(|u_n|)_{n\geq 0}$ décroit vers zéro, alors la série $\sum_{n\geq 0} u_n$ est convergente.*

1.2.4 Le critère de Cauchy pour les séries

Formulons, en termes de séries, le critère de Cauchy de convergence fondé sur la convergence des suites de Cauchy, tel qu'il peut être formulé pour les séries standard. Rappelons qu'une série $(s_N)_{N\geq 0}$ est de Cauchy si et seulement si "$\lim_{p,q\to\infty}$"$s_p - s_q = 0$, ce qui s'exprime, si la suite est standard, par le fait que $s_p - s_q \simeq 0$ dès que p et q sont i-grands. Une propriété essentiel de $\mathbb{R}$ (que n'a pas $\mathbb{Q}$, par exemple), est que toute suite de Cauchy est convergente (de limite égale à ${}^o s_\omega$ pour ω i-grand quelconque, toujours si l'on suppose la suite standard).

Théorème 1.3 (Critère de Cauchy pour la convergence des séries) *Soit $\sum_{n\geq 0} u_n$ une série standard quelconque. Cette série converge si et seulement si pour tous p et q i-grands, $\sum_{n=p+1}^{q} u_n$ est i-petit.*

Preuve : Comme nous venons de le rappeler, une suite réelle est convergente si (et bien entendu seulement si) elle est de Cauchy. La série considérée $(s_N)_{N\geq 0} := \sum_{n\geq 0} u_n$ est donc convergente si et seulement si $s_p - s_q \simeq 0$ dès que p et q sont i-grands ; mais précisément $s_p - s_q = \sum_{n=p+1}^{q} u_n$, d'où le résultat annoncé. □

1.3 Séries absolument convergentes

Définition : On dit que la série $\sum_{n\geq 0} u_n$ est *absolument convergente* si et seulement si la série $\sum_{n\geq 0} |u_n|$ est convergente.

Proposition 1.4 *Toute série absolument convergente est convergente.*

Preuve : Par transfert, nous démontrons le résultat sous l'hypothèse que la série est standard. Par le critère de Cauchy pour les séries (1.3), si la série est absolument convergente, pour tout p et q i-grands, on a $\sum_{n=p+1}^{q} |u_n| \simeq 0$; donc par l'inégalité triangulaire, $0 \leq \left|\sum_{n=p+1}^{q} u_n\right| \leq \sum_{n=p+1}^{q} |u_n| \simeq 0$; on conclut par le critère de Cauchy (1.3), appliqué à la série $\sum_{n\geq 0}^{q} u_n$ cette fois. □

La réciproque de ce résultat est fausse comme le montre l'exemple de la série harmonique alternée.

Proposition 1.5 *Si $0 \leq u_n \leq v_n$ pour tout $n \geq 0$, et si $\sum_{n\geq 0} v_n$ converge, alors $\sum_{n\geq 0} u_n$ converge.*

Preuve : Il suffit de remarquer que $0 \leq \sum_{n=p+1}^{q} u_n \leq \sum_{n=p+1}^{q} v_n$, de supposer $q > p$ i-grands, et d'appliquer le critère de Cauchysuccessivement à la série $\sum_{n\geq 0} v_n$ puis à la série $\sum_{n\geq 0} u_n$.

□

Définition : Soient $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$ deux suites quelconques, non nulles pour n suffisamment grand. On dit que la suite $(u_n)_{n\geq 0}$ est *asymptotique à* la suite $(v_n)_{n\geq 0}$ si et seulement si $\lim_n u_n/v_n = 1$.

Il est facile de voir que la relation “être asymptotique à” est une relation d'équivalence (standard) sur les suites ; d'ailleurs on utilise souvent la terminologie $(u_n)_{n\geq 0}$ est é*quivalente à* la suite $(v_n)_{n\geq 0}$ pour exprimer la même propriété, et on note $(u_n)_{n\geq 0} \sim (v_n)_{n\geq 0}$, ou plus abusivement $u_n \sim v_n$ qui n'a à proprement parler pas de sens (u_n et v_n sont des nombres, pas des suite), voire (lors de l'application de la proposition suivante) $\sum_{n\geq 0} u_n \sim \sum_{n\geq 0} v_n$, ce qui a un sens franchement différent et constitue de ce fait un abus à proscrire.

Proposition 1.6 *Si les suites $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$ sont positives, et si $(u_n)_{n\geq 0} \sim (v_n)_{n\geq 0}$, alors $\sum_{n\geq 0} u_n$ converge si et seulement si $\sum_{n\geq 0} v_n$ converge.*

Preuve : La relation $\sim$ étant symétrique, quitte à échanger les suites $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$, il suffit de montrer que si $\sum_{n\geq 0} v_n$ converge, il en est de même de $\sum_{n\geq 0} u_n$. Par transfert, nous pouvons supposer les suite $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$ standard, et donc l'hypothèse équivaut au fait que $u_n/v_n = 1 + \varepsilon_n$ avec $\varepsilon_n \simeq 0$ dès que n est i-grand. En vue d'appliquer le critère de Cauchy, soient $p < q$ i-grands quelconques. Nous avons

$$\sum_{n=p+1}^{q} u_n = \sum_{n=p+1}^{q} (1+\varepsilon_n)v_n = \sum_{n=p+1}^{q} v_n + \sum_{n=p+1}^{q} \varepsilon_n v_n$$

Il suffit donc de montrer que $\sum_{n=p+1}^{q} \varepsilon_n v_n \simeq 0$. Soit $\varepsilon := \mathrm{Max}\,\{|\varepsilon_{p+1}|, \ldots, |\varepsilon_q|\}$ qui est l'un des éléments, tous i-petits, de cet ensemble fini, et est donc i-petit. On a

$$0 \leq |\sum_{n=p+1}^{q} \varepsilon_n v_n| \leq \sum_{n=p+1}^{q} |\varepsilon_n v_n| \leq \sum_{n=p+1}^{q} \varepsilon v_n = \varepsilon \sum_{n=p+1}^{q} v_n \simeq 0,$$

puisque $\varepsilon \simeq 0$, et $\sum_{n=p+1}^{q} v_n \simeq 0$ par le critère de Cauchy. Par ce même critère de Cauchy, la série $\sum_{n\geq 0} u_n$ est donc également convergente. □

Exercice : Soient $\mathcal{A} := \{a_1, a_2, \ldots, a_M\}$ une famille finie de réels *de même signe*, et $E := \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M\}$ une famille finie de nombres, réels ou complexes. Il exite dans l'enveloppe convexe de E un ε tel que

$$\sum_{n=1}^{M} \varepsilon_n a_n = \varepsilon \sum_{n=1}^{M} a_n$$

En particulier, si les ε_n sont i-petits, il en est de même de ε. Et si les ε_n sont limités, ε est également limité. En déduire que $\sum_{n=1}^{M} a_n(1+\varepsilon_n) = (1+\varepsilon)\sum_{n=1}^{M} a_n$.

Solution : Quitte à éventuellement multiplier les deux membres de l'égalité par -1 et écarter les éléments nuls de $\mathcal{A}$, nous pouvons supposer que les a_n sont strictement positifs. Posons $\sigma := \sum_{m=0}^{M} a_m$ et $\lambda_n = a_n/\sigma$. Nous avons donc $0 < \lambda_n \leq 1$ et $\sum_{n=0}^{M} \lambda_n = 1$. Nous voyons qu'il convient de choisir $\varepsilon = \sum_{n=1}^{M} \varepsilon_n a_n / \sum_{n=1}^{M} a_n$, ce qui se réécrit,

$$\varepsilon = \frac{1}{\sigma}\sum_{n=1}^{M} \varepsilon_n a_n = \sum_{n=1}^{M} \varepsilon_n \frac{a_n}{\sigma} = \sum_{n=1}^{M} \varepsilon_n \lambda_n,$$

qui est une combinaison barycentrique des éléments de E, donc un élément de l'enveloppe convexe de E.

Exercice : Séries de Bertrand : Montrer de façon analogue aux séries de Riemann que $\sum_{n\geq 1} \frac{1}{n^\alpha (\ln n)^\beta}$ converge si et seulement si $\alpha > 1$, ou si $\alpha = 1$ et $\beta > 1$.

A la section 1.5 nous verrons que seules les séries absolument convergentes sont les séries qui convergent indépendamment de l'ordre dans lequel on additionne les termes de la série.

1.3.1 Séries à convergence géométrique

Soit $\sum_{n\geq 0} u_n$ une série à termes u_n tous positif. Si la série $\sum_{n\geq 0} u_n$ converge ou diverge aussi rapidement qu'une série géométrique, voici deux tests qui permettent souvent de s'en assurer.

Corollaire 1.7 (Test du rapport, dit "de d'Alembert") *Supposons que* $\lim_n u_{n+1}/u_n = l \neq 1$. *La série* $\sum_{n\geq 0} u_n$ *est convergente si et seulement si* $l < 1$.

Corollaire 1.8 (Test de la racine, dit "de Cauchy") *Supposons que* $\lim_n \sqrt[n]{u_n} = l \neq 1$. *La série* $\sum_{n\geq 0} u_n$ *est convergente si et seulement si* $l < 1$.

Preuve : Par transfert, nous pouvons supposer la suite $(u_n)_{n\geq 0}$ standard, et donc aussi la limite l. Nous savons que la convergence d'une série ne dépends pas des premiers termes de la série ; ainsi, par la proposition 1.5, pour montrer la convergence (resp. divergence) de $\sum_{n\geq 0} u_n$, il nous suffit de touver une série convergente (resp. divergente) $(v_n)_{n\geq 0}$ et un indice n_0 tel que pour tout $n \geq n_0$ on ait $u_n \leq v_n$ (resp. $0 \leq v_n \leq u_n$). Ici, il suffira de choisir n_0 i-grand quelconque ; voici comment procéder pour le choix de $(v_n)_{n\geq n_0}$: on choisira tout d'abord q standard strictement compris entre l et 1 ; par exemple la moyenne $q := \frac{l+1}{2}$. Comme $l \neq 1$, nous avons donc ou bien $l \lneqq q \lneqq 1$ ou bien $1 \lneqq q \lneqq l$.

Si $\lim_n \sqrt[n]{u_n} = l$, et si $l < 1$ (et donc $q < 1$), pour $n \geq n_0$ on a $\sqrt[n]{u_n} \simeq l$ et donc $\sqrt[n]{u_n} < q$, d'où $u_n < q^n$: il suffit de poser $v_n := q^n$. Si $l > 1$ (et donc $q > 1$), pour $n \geq n_0$ on a $\sqrt[n]{u_n} \simeq l$ et donc $\sqrt[n]{u_n} > q$, d'où $u_n > q^n$: il suffit là encore de poser $v_n := q^n$.

Si $\lim_n u_{n+1}/u_n = l$, et si $l < 1$ (et donc $q < 1$), pour $n \geq n_0$ on a $u_{n+1}/u_n \simeq l$, d'où $u_{n+1}/u_n < q$, et plus généralement

$$\frac{u_n}{u_{n-1}} < q \ ; \ \frac{u_{n-1}}{u_{n-2}} < q \ ; \ \ldots \ ; \ \frac{u_{n_0+1}}{u_{n_0}} < q$$

$$\text{d'où} \quad \frac{u_n}{u_{n_0}} = \frac{u_n}{u_{n-1}}\frac{u_{n-1}}{u_{n-2}} \cdots \frac{u_{n_0+1}}{u_{n_0}} < q^{n-n_0}$$

ou encore $u_n < aq^n$, où on a posé $a := u_{n_0}/q^{n_0}$: il suffit de poser $v_n := aq^n$.

De même, si $\lim_n u_{n+1}/u_n = l > q > 1$, pour $n \geq n_0$ on a cette fois $u_n > v_n$ pour ce même choix $v_n := aq^n$, qui est maintenant le terme général d'une série divergente. □

1.4 Théorème d'Abel de semi-convergence

Comme nous le verrons dans la section suivante, une série qui est convergente mais pas absolument convergente a cette propriété que si l'on modifie convenablement l'ordre de ses termes, la limite de ses sommes partielles peut être affectée par ce changement d'ordre, voire ne pas exister du tout. C'est pourquoi il y a une tradition d'appeler *semi-convergente* toute série convergente qui ne serait pas absolument convergente. Voici un théorème, dû à Niels H. Abel (1802-1829) assurant la semi-convergence de certaines séries.

Théorème 1.9 (critère d'Abel) *Si $(a_n)_{n\geq 0}$ est une suite telle que la suite des $A_N := \sum_{n=0}^N a_n$ soit bornée, et si $(b_n)_{n\geq 0}$ est une suite de réels positifs décroissant vers 0, alors la série $\sum_{n\geq 0} a_n b_n$ est (semi-)convergente.*

Preuve : La preuve sera facile moyennant le lemme combinatoire suivant :

Lemme 1.10 (sommation par partie discrète) *Pour toutes suites $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 0}$ de nombres, et tous indices $q > p$, on a*

$$\sum_{n=p+1}^{q} a_n b_n = A_q b_q - A_p b_{p+1} + \sum_{n=p+1}^{q-1} A_n (b_n - b_{n+1})$$

où l'on a posé $A_N := \sum_{n=0}^N a_n$.

Preuve : On a

$$\begin{aligned}
\sum_{n=p+1}^{q} a_n b_n &= \sum_{n=p+1}^{q} (A_n - A_{n-1}) b_n = \sum_{n=p+1}^{q} A_n b_n - \sum_{n=p+1}^{q} A_{n-1} b_n \\
&= \sum_{n=p+1}^{q} A_n b_n - \sum_{n=p}^{q-1} A_n b_{n+1} \\
&= \sum_{n=p+1}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_p b_{p+1}
\end{aligned}$$

□

A présent nous pouvons prouver le critère d'Abel au moyen du critère de Cauchy : Par transfert, nous pouvons supposer les suites $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 0}$ standard, et partant, aussi la suite $(A_n)_{n\geq 0}$ ainsi que le majorant commun M de tous les $|A_n|$. Soient $q > p$ i-grands quelconques. Observons que, comme les b_n sont décroisants, on a $|b_n - b_{n+1}| = b_n - b_{n+1}$, et donc, en appliquant l'inégalité triangulaire à la formule du lemme, nous avons

$$\begin{aligned}
\left|\sum_{n=p+1}^{q} a_n b_n\right| &\leq |A_q| b_q + |A_p| b_{p+1} + \sum_{n=p+1}^{q-1} |A_n| (b_n - b_{n+1}) \\
&\leq M b_q + M b_{p+1} + M \sum_{n=p+1}^{q-1} (b_n - b_{n+1}) = M(b_q + b_{p+1} + b_{p+1} - b_q) \simeq 0
\end{aligned}$$

□

Exercice : Déduire une nouvelle preuve de la proposition 1.2 du critère d'Abel 1.9.

Après avoir vérifié que la preuve du critère d'Abel 1.9 s'applique également au cas où les a_n sont des nombres complexes, formuler et montrer que si $(b_n)_{n\geq 0}$ est une suite de réels décroissant vers 0, alors la séries trigonométriques $\sum_{n\geq 0} b_n e^{in\theta} := \sum_{n\geq 0} b_n \cos n\theta + i \sin n\theta$ converge pourvu que $\theta \not\equiv 0 (\mathrm{mod} 2\pi)$. (Indication : montrer que $A_n = (1 - e^{in\theta})/(1 - e^{i\theta})$).

1.5 Suites sommables

Lorsque l'on doit additionner les éléments d'une famille finie de nombres réels, l'associativité et la commutativité de l'addition dans $\mathbb{R}$ (ou $\mathbb{C}$) font qu'on est libre d'effectuer les additions dans n'importe quel ordre. Qu'en est-il si l'on a un ensemble infini $(x_i)_{i\in I}$ de nombres ? Lorsque $I=\mathbb{N}$, nous disposons d'un ordre privilégié sur l'ensemble I des indices, ce qui permet d'associer à ce problème une unique suite $(s_N)_{N\in\mathbb{N}}$, avec $s_N := \sum_{n=0}^N x_n$, et de convenir que la somme des $(x_n)_{n\in\mathbb{N}}$ est la limite des s_N, si elle existe. Mais qu'en est-il si l'on ne veut pas supposer qu'il y a un ordre "naturel" sur l'ensemble des indices ?

Envisageons ici simplement le cas où I est dénombrable, c'est-à-dire qu'il existe une bijection $n\mapsto i(n)$ de $\mathbb{N}$ sur I [1] ; dans ce cas on peut considérer la suite des $u_n := x_{i(n)}$ et étudier la convergence de la suite des

$$s_N := \sum_{n=0}^N u_n = \sum_{n=0}^N x_{i(n)} = \sum_{i\in i(\{0,\dots,N\})} x_i.$$

Si la bijection i est naturellement associée à la question, ceci ramène la question à celle de la convergence de la série $\sum_{n\geq 0} u_n.$; mais si ce n'est pas le cas, il existe bien d'autres bijections j de $\mathbb{N}$ sur I (en fait n'importe quelle $j := i\circ\sigma$ où σ est une bijections quelconque de $\mathbb{N}$ sur lui-même), et on peut légitimement souhaiter que la suite des

$$t_N := \sum_{n=0}^N v_n := \sum_{n=0}^N x_{j(n)} = \sum_{j\in j(\{0,\dots,N\})} x_j$$

converge si et seulement si la suite des

$$s_N := \sum_{n=0}^N u_n = \sum_{n=0}^N x_{i(n)} = \sum_{i\in i(\{0,\dots,N\})} x_i$$

converge et qu'elle conduit à la même somme, d'où la définition :

Définition : On dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ est *sommable* si et seulement si pour toute bijection σ des entiers, la série $\sum_{n\geq 0} v_n := \sum_{n\geq 0} u_{\sigma(n)}$ est convergente. On dit aussi que la série $\sum_{n\geq 0} u_n$ est *commutativement sommable.*

Visiblement, la notion de sommabilité d'une suite de nombre est une notion standard ; afin de trouver une caractérisation plus commode, nous pouvons donc, sans perte de généralité, commencer par l'étudier dans le cas des suites standard, et pour les bijections σ standard. Observons que dans ce cas, comme l'image d'un standard par une application standard est standard, que pour tout n standard, $\sigma(n)$ et aussi $\sigma^{-1}(n)$ sont standard, puisque la bijection inverse σ^{-1} d'une bijection standard est unique, donc standard. En raisonnant par l'absurde nous en déduisons que pour une bijection standard, l'image $\sigma(\omega)$ de tout non-standard ω est non-standard. En d'autres termes, si ω est i-grand, il en est de même de $\sigma(\omega)$.

A présent, pour que la suite $(u_n)_{n\geq 0}$ soit sommable, nous voyons qu'il faut que pour tout N i-grand, on ait $s_N \simeq t_N$, c'est-à-dire

$$\sum_{n\in\{0,\dots,N\}} u_n - \sum_{p\in\sigma(\{0,\dots,N\})} u_p \simeq 0.$$

Observons que les deux ensembles d'indices $\{0,\dots,N\}$ et $\sigma(\{0,\dots,N\}) = \{\sigma(0),\dots,\sigma(N)\}$ ont tous les entiers standard en commun ; en d'autres termes, dans la différence ci-dessus, tous les u_n d'indice n standard de la première somme figurent également dans la seconde somme et s'annulent donc deux à deux. Il ne subsite donc, dans la différence, que des u_ω d'indice ω i-grand. On est donc conduit à souhaiter que pour une suite sommable, et pour tout ensemble fini F d'indices tous i-grand, on ait $\sum_{\omega\in F} u_\omega \simeq 0$; d'autant que, par le critère de Cauchy, nous savons déjà que pour que $\sum_{n\geq 0} u_n$ converge, il faut et il suffit que $\sum_{\omega\in F} u_\omega \simeq 0$ pour tout tel ensemble d'indices F de la forme $F=\{p+1,p+2,\dots,q\}$, p et q i-grands. Et c'est la bonne idée ; en effet, on a

Lemme 1.11 (critère de sommabilité) *Soit $(u_n)_{n\geq 0}$ une suite* standard. *Cette suite est sommable si et seulement si pour tout ensemble fini F d'indices i-grands, on a*

$$\sum_{\omega\in F} u_\omega \simeq 0.$$

[1]En fait ceci est le seul cas intéressant : on montre qu'il existe une notion naturelle de somme d'une famille infinie seulement dans le cas où l'ensemble des indices pour lesquels x_i est non nul est dénombrable

Preuve : La preuve de ce lemme ne demande pas de connaissance nouvelle, mais elle met en oeuvre une technique non-standard dite du “bootstrapping” sans doute neuve pour le lecteur ; celui-ci est donc invité à y revenir ultérieurement, et à s'intéresser en priorité au théorème suivant et à sa preuve qui découle facilement du lemme : ceci peut constituer une motivation à l'étude de la preuve du lemme !

$\boxed{\Leftarrow}$ Supposons que pour tout ensemble fini F d'indices i-grands on a $\sum_{\omega\in F} u_\omega \simeq 0$. Soit σ une bijection standard de $\mathbb{N}$; soit $(v_n)_{n\geq 0}$: $=(u_{\sigma(n)})_{n\geq 0}$ la suite $(u_n)_{n\geq 0}$ ainsi renumérotée. Nous devons montrer que $\sum_{n\geq 0} v_n$ converge, et vers la même limite $s := \sum_{n=0}^{+\infty} u_n$ que la série $\sum_{n\geq 0} u_n$ (qui correspond au cas où σ est l'identité : $\sigma(n) := n$).

Soit $p < q$ i-grands quelconques ; on a

$$\sum_{n=p+1}^{q} v_n = \sum_{n\in\{p+1,\ldots,q\}} u_{\sigma(n)} = \sum_{m\in\sigma(\{p+1,\ldots,q\})} u_m = \sum_{m\in F} u_m,$$

où $F := \sigma(\{p+1,\ldots,q\})$ qui est un ensemble fini d'entiers tous non standard, puisque σ est une bijection standard, et que les indices $p+1,\ldots,q$ sont tous non standard (voir l'introduction heuristique au lemme). Donc, par le critère de Cauchy, $\sum_{n\geq 0} v_n$ converge. Il reste à vérifier que sa limite $t := \sum_{n=0}^{+\infty} v_n$ est la même que la limite $s := \sum_{n=0}^{+\infty} u_n$.

Soit ω i-grand quelconque ; on a $t \simeq \sum_{n=0}^{\omega} v_n$ et $s \simeq \sum_{n=0}^{\omega} u_n$. Comme s et t sont standard, il suffit de montrer que $s \simeq t$, et, pour cela, il suffit de montrer que

$$(t \simeq) \sum_{n=0}^{\omega} v_n \simeq \sum_{n=0}^{\omega} u_n (\simeq s),$$

ce qui s'écrit encore

$$(\sum_{n\in D_+} u_n - \sum_{n\in D_-} u_n =) \sum_{n\in G} u_n - \sum_{n\in\Omega} u_n \simeq 0,$$

où l'on a posé $\Omega := \{0,\ldots,\omega\}$, $G := \{\sigma(0),\ldots,\sigma(\omega)\} = \sigma(\{0,\ldots,\omega\})$, et où D_+ est l'ensembles des indices n dans G qui ne sont pas dans Ω, et D_- est l'ensembles des indices n dans Ω qui ne sont pas dans G : seuls les u_n tels que $n \in D_+ \cup D_-$ ne s'annulent pas deux à deux dans la différence. Les ensembles D_- et D_+ sont finis, et ne comporte aucun élément standard, comme cela a été expliqué dans l'introduction. Par hypothèse, on a donc $\sum_{n\in D_+} u_n \simeq 0$ et $\sum_{n\in D_-} u_n \simeq 0$, ce qui permet de conclure.

$\boxed{\Rightarrow}$ Nous devons montrer que si la série (standard) $\sum_{n\geq 0} u_n$ est sommable, alors pour tout ensemble fini F d'entiers non standard, on a $\sum_{n\in F} u_n \simeq 0$. Nous allons montrer la contrapposée : s'il existe un ensemble fini F d'entiers non standard tel que $\sum_{n\in F} u_n$ n'est pas i-petit, alors il existe une bijection standard σ de $\mathbb{N}$ telle que la série $\sum_{n\geq 0} v_n := \sum_{n\geq 0} u_{\sigma(n)}$ ne soit pas convergente.

Comme par hypothèse $\sum_{n\in F} u_n \not\simeq 0$, il existe un standard $a > 0$ tel que $|\sum_{n\in F} u_n| > a$. Nous avons donc que pour tout *standard* N

$$\text{il existe un ensemble fini } F \text{ d'entiers tous supérieurs a } N+1 \text{ tel que} \left|\sum_{n\in F} u_n\right| > a \qquad (1.7)$$

En effet, l'ensemble F dont nous avons supposé l'existence convient bien pour chaque N standard. Mais la formule (1.7) étant standard (puisque la suite $(u_n)_{n\geq 0}$ et le réel a le sont), nous avons donc le droit d'applique le *transfert*, et la formule (1.7) est donc vrai pour tout N, standard ou non. Pour chaque N fixons un choix $F(N)$ de tel ensemble vérifiant (1.7) ; nous venons de montrer l'existence d'une fonction $N \mapsto F(N)$ qui a chaque entier associe un ensemble fini d'entiers ayant les propriétés exprimées dans la formule (1.7). Par transfert, nous pouvons supposer en outre qu'il existe en outre une telle fonction qui est standard.[2]

La figure 1.2 montre comment confectionner une renumérotation $v_n := u_{\sigma(n)}$ des éléments de la suite $(u_n)_{n\geq 0}$ telle que la série $\sum_{n\geq 0} v_n$ ne soit pas convergente. L'idée est simple on s'arrange pour que pour la série renumérotée $(v_n)_{n\geq 0}$, pour une suite de N, les éléments u_n pour $n \in F(N_m)$ soient consécutifs,

[2] Avec la preuve de l'existence de cette fonction standard, nous venons de voir un exemple de “bootstrapping”, expression anglo-saxonne dérivant littéralement de “lacet de chaussure”, et évoquant l'opération (périlleuse et pour le moins extrèmement difficile) consistant à s'auto-soulever en tirant (fort) sur ses lacets. En effet, au moyen de l'existence d'un seul ensemble F d'entiers i-grandsayant une propriété standard donnée, nous avons démontré l'existence de toute une suite standard $(F(N))_{N\geq 0}$ d'ensembles ayant cette propriété. Le bootstrapping est un exemple de la force de l'axiomatique non-standard.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	$\mathbb{N}$
			x	x		x			y		y	y				z	z			z					
			1	2		3=$q(1)$																			
4	5	6			7																				
									8		9	10=$q(2)$													
							11	12		13															
																14	15			16=$q(2)$					
													17	18	19			20	21						

x : éléments de $F(1)$

y : éléments de $F(\sigma^{-1}(q(1))+1)$

z : éléments de $F(\sigma^{-1}(q(2))+1)$

Figure 1.2: Confections d'une renumérotation $v_n = u_{\sigma(n)}$ (bijection σ de $\mathbb{N}$) telle que pour tout N il existe $q(N) > p(N) \geq N$ tel que $|\sum_{n=p(N)+1}^{q(N)} u_{\sigma(n)}| \geq a > 0$; pour une telle renumérotation σ la série $\sum_{n\geq 0} v_n := \sum_{n\geq 0} u_{\sigma(n)}$ ne peut être convergente, d'après le critère de Cauchy. Par hypothèse pour tout N il existe un ensemble fini $F(N)$ d'entiers supérieurs à N (on notera $F(N) \geq N$) tel que $|\sum_{n\in F(N)}| \geq a$.

numérotés de $p(m)+1$ à $q(m)$, de telle manière à ce que pour chaque m on ait $|\sum_{n=p(m)+1}^{q(N_m)} v_n| = |\sum_{n\in F(N_m)} u_n| > a > 0$, ce qui montre, par le critère de Cauchy, que $\sum_{n\geq 0} v_n$ ne peut converger.

Voici comment on numérote $\sigma(n)$ chaque élément n de $\mathbb{N}$:

1.- On numérote (en croissant) les éléments de $F(1)$; soit $p(1)$ tel que $p(1)+1$ soit le numéro du premier élément de $F(1)$ numéroté, et $q(1)$ le numéro du dernier entier $(\sigma^{-1}(q(1)))$ numéroté.

2.- S'il en existe, on numérote ensuite les entiers inférieurs à $\sigma^{-1}(q(1))$ qui ne sont pas encore numérotés. On atteint ainsi le numéro $\sigma^{-1}(q(1))+1$.

3.- On numérote (en croissant) les éléments de $F(\sigma^{-1}(q(1))+1)(\geq \sigma^{-1}(q(1))+1)$; soit $p(2)$ tel que $p(2)+1$ soit le numéro du premier élément de $F(\sigma^{-1}(q(1))+1)$ numéroté, et $q(2)$ le numéro du dernier entier $(\sigma^{-1}(q(2)))$ numéroté.

4.- S'il en existe, on numérote ensuite les entiers inférieurs à $\sigma^{-1}(q(2))$ qui ne sont pas encore numérotés. On atteint ainsi le numéro $\sigma^{-1}(q(2))+1$.

5.- On numérote (en croissant) les éléments de $F(\sigma^{-1}(q(2))+1)$; ...

...

2m-1.- On numérote (en croissant) les éléments de $F(\sigma^{-1}(q(m-1))+1)(\geq \sigma^{-1}(q(m-1))+1)$; soit $p(m)$ tel que $p(m)+1$ soit le numéro du premier élément de $F(\sigma^{-1}(q(m-1))+1)$ numéroté, et $q(m)$ le numéro du dernier entier $(\sigma^{-1}(q(m)))$ numéroté.

2m.- S'il en existe, on numérote ensuite les entiers inférieurs à $\sigma^{-1}(q(m))$ qui ne sont pas encore numérotés. On atteint ainsi le numéro $\sigma^{-1}(q(m))+1$.

...

□

Théorème 1.12 *Soit $(u_n)_{n\geq 0}$ une suite quelconque. La suite $(u_n)_{n\geq 0}$ est sommable si et seulement si la série $\sum_{n\geq 0} u_n$ est absolument convergente.*

Preuve : Par transfert, il suffit de prouver le théorème dans le cas où la suite $(u_n)_{n\geq 0}$ est standard, pour laquelle le critère de sommabilité 1.11 s'applique.

$\boxed{\Rightarrow}$ Supposons $(u_n)_{n\geq 0}$ sommable, et montrons que $\sum_{n\geq 0} |u_n|$ converge. Pour cela, en vue d'appliquer le critère de Cauchy, soient $p < q$ i-grandsquelconques. Nous devons montrer que $\sum_{n\in F} |u_n|$ est i-petit, où F désigne l'ensemble $\{p+1, \ldots, q\}$ des indices compris entre $p+1$ et q. Décomposons F en l'ensemble F_+ des indices où $|u_n| = u_n > 0$ et l'ensemble F_- des indices où $u_n \leq 0$, c'est-à-dire $|u_n| = -u_n$. Nous avons donc

$$\sum_{n=p+1}^{q} |u_n| = \sum_{n\in F} |u_n| = \sum_{n\in F_+} |u_n| + \sum_{n\in F_-} |u_n| = \sum_{n\in F_+} u_n - \sum_{n\in F_-} u_n.$$

Il ne reste plus qu'à observer que F_+ et F_- sont des ensembles finis d'indices i-grands, et donc, d'après le critère de sommabilité, les deux dernières sommes sont i-petites, d'où $\sum_{n=p+1}^{q} u_n \simeq 0$.

$\boxed{\Leftarrow}$ Supposons $\sum_{n\geq 0} |u_n|$ convergente, et montrons que $(u_n)_{n\geq 0}$ est sommable. En vue d'appliquer le critère de sommabilité, soit F un ensemble quelconque d'indices tous i-grands. Posons $p := \mathrm{Min}\, F - 1$ et $q := \mathrm{Max}\, F$. Nous avons donc

$$|\sum_{n\in F} u_n| \leq \sum_{n\in F} |u_n| \leq \sum_{n=p+1}^{q} |u_n|,$$

et cette dernière somme est i-petite, puisque la série $\sum_{n\geq 0} |u_n|$ est convergente. D'où $\sum_{n\in F} u_n \simeq 0$. On conclut en appliquant le critère de sommabilité.

$\square$

Chapitre 2

Intégrale généralisée

2.1 Définition et critère de Cauchy

Ce chapitre traite de la question du sens à donner à des expressions telles que $\int_0^{+\infty} e^{-t}\sin t dt$ ou $\int_0^{\frac{\pi}{2}} \frac{dt}{\sin\sqrt{t}}$. Nous y verrons dans quel sens certaines régions non bornées du plan sont d'aire finie. Soient $-\infty \leq a < b \leq +\infty$, et $f :]a,b[\longrightarrow \mathbb{R}$ une fonction que l'on suppose intégrable au sens Riemann sur tout intervalle $[c,d] \subset]a,b[$, comme dans les exemples ci-dessus. Pour tout $x \in]a,b[$ nous pouvons donc poser $F_c(x) := \int_c^x f(t)dt$; si c_1 et c_2 appratiennent l'un et l'autre à $]a,b[$, les fonctions F_{c_1} et F_{c_2} diffèrent par une constante C, et plus précisément $F_{c_2}(x) = F_{c_1}(x) + C$, avec $C := \int_{c_2}^{c_1} f(t)dt$. L'existence des limites que nous allons examiner maintenant ne dépend donc pas du choix de c dans $]a,b[$.

Définition : Si les limites $\lim_{x\to a} F_c(x)$ et $\lim_{x\to b} F_c(x)$ existent, on appelle *intégrale généralisée* de f sur $]a,b[$ le nombre, indépendant du choix de c,

$$\int_a^b f(t)dt := \lim_{\alpha\to a}\int_\alpha^c f(t)dt + \lim_{\beta\to b}\int_c^\beta f(t)dt,$$

et on dit que l'*intégrale* $\int_a^b f(t)dt$ *converge*. Nous dirons que f est intégrable sur $]a,b[$.

Exemples fondamentaux

Du fait qu'il est possible d'expliciter une primitive des fonctions e^x et $1/x^\alpha$, il est immédiat de vérifier que

- $\int_0^{+\infty} e^t dt$ diverge alors que $\int_0^{+\infty} e^{-t}dt$ converge.
- $\int_1^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$
- $\int_0^1 \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha < 1$

Exercice : Montrer que si f est intégrable sur $[a,b]$ (au sens usuel, de Riemann), alors les limites $\lim_{x\to a} F_c(x)$ et $\lim_{x\to b} F_c(x)$ existent, et les deux définitions du symbole $\int_a^b f(t)dt$ correspondent au même nombre.

Observons que si f est intégrable au sens de Riemann sur l'intervalle $[c,d]$, cet intervalle est nécessairement borné et la fonction f est nécessairement bornée sur $[c,d]$. Ce qui peut donc empècher que f soit intégrable sur $]a,b[$ alors qu'elle est intégrable sur tout intervalle $[c,d] \subset]a,b[$ est, ou bien que l'intervalle $]a,b[$ n'est pas borné (c'est-à-dire que $a = -\infty$ ou $b = +\infty$), ou que f n'est pas bornée : c'est ce qui se produit dans les deux exemples liminaires, où toutefois les deux intégrales sont bel et bien convergentes.

Exercice : Montrer que si l'intevalle $]a,b[$ est borné, et si la fonction f est bornée sur $]a,b[$, alors f est intégrable sur $]a,b[$. (Indication : appliquer le critère de Cauchy ci-dessous.

Pour les critères d'intégrabilité auxquels sera consacré la fin de ce chapitre, nous n'envisagerons plus que le problème de l'existence de $\lim_{\beta\to b}\int_c^\beta f(t)dt$, le cas de $\lim_{\alpha\to a}\int_\alpha^c f(t)dt$ se traitant au moyen de la remarque que $\int_\alpha^c f(t)dt = \int_{-c}^{-\alpha} f(-t)dt$. En pratique, nous ne considèrerons dorénavant que des fonctions intégrables sur tout intervalle $[a,\beta] \subset [a,b[$ et n'étudierons que la question de l'intégrabilité sur $[a,b[$, c'est-à-dire l'existence de $\lim_{\beta\to b}\int_a^\beta f(t)dt$.

Notations : Nous dirons que x est i-voisin de $+\infty$ et noterons $x \simeq +\infty$ si et seulement si x est i-grand. En revanche nous proscrirons l'écriture $+\infty \simeq x$ qui encouragerait (par transitivité abusive sur $x' \simeq +\infty \simeq x$) l'erreur de déduire que deux i-grands x' et x sont toujours i-voisins.

Théorème 2.1 (Critère de Cauchy) *Soient $a < b \leq +\infty$ et $f : [a,b[\longrightarrow \mathbb{R}$ une fonction* standard *intégrable sur tout intervalle $[a,\beta] \subset [a,b[$. La fonction f est intégrable sur $[a,b[$ si et seulement si pour tous $\beta' \simeq b$ et $\beta'' \simeq b$, on a $\int_{\beta'}^{\beta''} f(t)dt \simeq 0$.*

Preuve : Par transfert, nous pouvons supposer que f, a, et b sont standard, et donc aussi F_a. Nous pouvons appliquer la caractérisation non standard de la limite, c'est-à-dire que $\lim_{x\to b} F_a(x)$ existe et vaut I si et seulement si I est standard et pour tout $\beta \simeq b$ $F_a(\beta) = I$. Voici la preuve de chacune des implications constituant l'équivalence affirmée :

$\boxed{\Rightarrow}$ Pour tous $\beta' \simeq b$ et $\beta'' \simeq b$, nous avons $F_a(\beta') \simeq I$ et $F_a(\beta'') \simeq I$, donc

$$\int_{\beta'}^{\beta''} f(t)dt = \int_{\beta'}^{a} f(t)dt + \int_{a}^{\beta''} f(t)dt = \int_{a}^{\beta''} f(t)dt - \int_{a}^{\beta'} f(t)dt = F_a(\beta'') - F_a(\beta') \simeq I - I = 0$$

$\boxed{\Leftarrow}$ Soit $\beta_0 \simeq b$. Nous souhaitons poser $I := {}^o\!(F_a(\beta_0))$, mais pour cela nous devons nous assurer que $F_a(\beta_0)$ est bien presque-standard, c'est à dire limité, puisque tout réel limité est presque standard. Considérons l'ensemble *interneE* suivant : $E := \{x \in [a,b[\ | \ |F_a(\beta_0) - F_a(x)| \leq 1\}$. Si on pose $\beta'' := \beta_0$ et $\beta' := x$, on voit que l'ensemble E contient tous les réels i-voisins de β_0. Mais les réels i-voisins de β_0 ne formant pas un ensemble interne (pourquoi ?), E ne peut pas se réduire à ces réels là, et il doit donc aussi comporter un (en fait des) c qui ne soit pas i-voisin de b. Il existe donc aussi un standard $c_0 \in [a,b[$, $c \simeq c_0$ (la *partie standard* de c). A présent $F_a(c_0)$ est standard, $F_a(c) \simeq F_a(c_0)$ est donc limité, et donc $F_a(\beta_0)$ est également limité, puisque la distance $|F_a(\beta_0) - F_a(c)|$ de $F_a(c)$ à $F_a(\beta_0)$ est inférieure à 1. Nous sommes donc bien assurés de l'existence de $I := {}^o\!(F_a(\beta_0))$, et nous avons bien, pour tout $\beta \simeq b$,en posant cette fois $\beta' := \beta$, $F_a(\beta) \simeq F_a(\beta_0) \simeq {}^o\!(F_a(\beta_0)) = I$, ce qui montre que $I = \lim_{x\to b} F_a(x)$.

□

2.2 Critères d'intégrabilité des fonctions positives

Dans toute cette section f et g désignerons des fonctions réelles définies sur $[a,b,[$, intégrable au sens de Riemann sur tout intervalle $[\alpha,\beta]$ contenu dans $[a,b,[$. On notera $F(x) := \int_a^x f(t)dt$ et $G(x) := \int_a^x g(t)dt$. Le résultat suivant indique une raison de s'intéresser à l'intégrabilité de fonctions positives.

Proposition 2.2 *Si $\int_a^b |f(t)|dt$ converge, alors $\int_a^b f(t)dt$ converge.*

Preuve : Par transfert, nous pouvons supposer que f, a, et b sont standard ; $|f|$ est donc également standard. Soient $\beta' < \beta''$ dans $[a,b[$ i-voisins de b. On a $0 \leq |\int_{\beta'}^{\beta''} f(t)dt| \leq \int_{\beta'}^{\beta''} |f(t)|dt$, or ce dernier nombre est i-petit puisque $\int_a^b |f(t)|dt$ converge (critère de Cauchy 2.1), donc $\int_{\beta'}^{\beta''} f(t)dt$ est également i-petit, ce qui implique que $\int_a^b f(t)dt$ converge, par ce même critère de Cauchy. □

Proposition 2.3 *On suppose qu'il existe $t_0 < b$ tel que pour tout $t \in [t_0,b[$ on a $0 \leq f(t) \leq g(t)$. Si $\int_a^b g(t)dt$ converge, alors $\int_a^b f(t)dt$ converge.*

Preuve : Par transfert, nous pouvons supposer que f, g, a, et b sont standard. Soient $\beta' < \beta''$ dans $[a,b[$ i-voisins de b. On a $0 \leq \int_{\beta'}^{\beta''} f(t)dt \leq \int_{\beta'}^{\beta''} g(t)dt$, or ce dernier nombre est i-petit puisque $\int_a^b g(t)dt$ converge (critère de Cauchy 2.1), donc $\int_{\beta'}^{\beta''} f(t)dt$ est également i-petit, ce qui implique que $\int_a^b f(t)dt$ converge, par ce même critère de Cauchy. □

Définition : On dit f est é*quivalente* à g au point b, et on note $f \sim g$ si et seulement si $\lim_{t\to b} f(t)/g(t) = 1$.

La relation $\sim$ est une relation d'équivalence sur l'ensemble des fonctions définies sur $[a,b[$.

Proposition 2.4 *On suppose qu'il existe $t_0 < b$ tel que $g(t) > 0$ pour tout $t \in [t_0,b[$. Si $f \sim g$ au point b, alors $\int_a^b f(t)dt$ converge si et seulement si $\int_a^b g(t)dt$ converge.*

Preuve : Par transfert, nous pouvons supposer que f, g, a, b, et t_0 sont standard. Observons que comme $f(\beta)/g(\beta) \simeq 1$ pour tout $\beta \simeq b$, il suffit de choisir un $t_1 \simeq b$ pour s'assurer qu'il existe un $t_1 \in [t_0, b[$ tel que pour tout $t \in [t_1, b[$ on a $0 < \frac{1}{2}g(t) \leq f(t) \leq 2g(t)$. Appliquons le *transfert* a cette propriété *standard :* nous en déduisons qu'il existe un tel t_1 qui est standard. Partant, pour tout $t \simeq b$, nous avons $t_1 \leq t < b$ puisque $t_1 \lneqq b$, et donc $0 < f(t) \leq 2g(t)$, d'où, pour tout $\beta' < \beta''$ dans $[a, b[$ i-voisins de b on a $0 \leq \int_{\beta'}^{\beta''} f(t)dt \leq 2\int_{\beta'}^{\beta''} g(t)dt$, or ce dernier nombre est i-petit puisque $\int_a^b g(t)dt$ converge (critère de Cauchy 2.1), donc $\int_{\beta'}^{\beta''} f(t)dt$ est également i-petit, ce qui implique que $\int_a^b f(t)dt$ converge, par ce même critère de Cauchy. □

2.3 Le critère d'Abel

Proposition 2.5 (Critère d'Abel) *Si $F(x) := \int_a^x f(t)dt$ est une fonction bornée sur $[a, b[$, et si g est dérivable et décroissante vers 0, alors $\int_a^b f(t)g(t)dt$ converge.*

Preuve : L'idée de la preuve tient dans une intégration par partie, qui donne

$$\int_{\beta'}^{\beta''} f(t)g(t)dt = [F_a(t)g(t)]_{\beta'}^{\beta''} - \int_{\beta'}^{\beta''} F_a(t)g'(t)dt. \tag{2.1}$$

Pour le reste, on raisonne comme à l'accoutumée : Par transfert, nous pouvons supposer que f, g, a, et b sont standard. Il en est donc de même pour $F = F_a$ qui est donc bornée par un standard M, et donc à valeurs limitées. Par ailleurs $g(\beta) \simeq 0$ pour tout $\beta \simeq b$; Donc, si $\beta' \simeq b$ et $\beta'' \simeq b$, nous voyons que le terme $[F_a(t)g(t)]_{\beta'}^{\beta''} = F_a(\beta'')g(\beta'') - F_a(\beta')g(\beta')$ est i-petit. Par ailleurs, pour $\beta' < \beta''$, et comme $-g'$ est positive puisque g est supposée décroissante

$$\left| - \int_{\beta'}^{\beta''} F_a(t)g'(t)dt \right| \leq \int_{\beta'}^{\beta''} |F_a(t)(-g'(t))|dt \leq M \int_{\beta'}^{\beta''} |(-g'(t))|dt = M \int_{\beta'}^{\beta''} -g'(t))dt = M(-g(\beta'')+g(\beta')) \simeq 0.$$

□

Chapitre 3

Intégrale double

3.1 Intégrale itérée et théorème de Fubini

On considère un domaine $\mathcal{D}$, fermé-borné de $\mathbb{R}^2$, limité par les lignes $x = a$, $x = b$, $y = \overline{\alpha}(x)$, et $y = \overline{\beta}(x)$, où $\overline{\alpha} : [a, b] \longrightarrow \mathbb{R}$ et $\overline{\beta} : [a, b] \longrightarrow \mathbb{R}$ sont des fonctions continues, telles que $\overline{\alpha}(x) \leq \overline{\beta}(x)$ pour tout $x \in [a, b]$. Précisément

$$\mathcal{D} := \{(x, y) \in \mathbb{R}^2 \ \mid \ x \in [a, b] \text{ et } y \in [\overline{\alpha}(x), \overline{\beta}(x)]\}.$$

En considérant x comme un paramètre, on peut intégrer $f(x, y)$ par rapport à y sur l'intervalle $[\overline{\alpha}(x), \overline{\beta}(x)]$, ce qui donne un nombre dépendant du choix de x, c'est-à-dire une fonction

$$\varphi : [a, b] \longrightarrow \mathbb{R} \ , \ \ \varphi(x) := \int_{\overline{\alpha}(x)}^{\overline{\beta}(x)} f(x, y)dy.$$

Cette fonction est continue ; on peut l'intégrer sur $[a, b]_x$ et obtenir le nombre $\int_a^b \varphi(x)dx =: I_{1,2}[f]$, c'est à dire *l'intégrale itérée* suivante :

$$I_{1,2}[f] := \int_a^b \left(\int_{\overline{\alpha}(x)}^{\overline{\beta}(x)} f(x, y)dy \right) dx.$$

Exercice : Former $I_{1,2}[f]$ dans le cas $[a, b] = [1, 2]$, $\overline{\alpha}(x) = \ln(x)$, $\overline{\beta}(x) = \ln(3x)$, et $f(x, y) = e^{x^2+y}$. On obtient $I_{1,2}[e^{x^2+y}] = e^4 - e$.

Il peut aussi se trouver que le domaine $\mathcal{D}$ se présente sous la forme

$$\mathcal{D} := \{(x, y) \in \mathbb{R}^2 \ \mid \ y \in [\alpha, \beta] \text{ et } x \in [\overline{a}(y), \overline{b}(y)]\},$$

où les fonctions $\overline{a}$ et $\overline{b}$ sont des fonctions continues, $\overline{a}(y) \leq \overline{b}(y)$ pour $y \in [\alpha, \beta]$. Dans ce cas on peut former la fonction (continue)

$$\psi(y) := \int_{\overline{a}(y)}^{\overline{b}(y)} f(x, y)dx,$$

qu'on peut intégrer sur $[\alpha, \beta]$, et obtenir le nombre $\int_\alpha^\beta \psi(y)dy =: I_{2,1}[f]$, c'est à dire *l'intégrale itérée* suivante :

$$I_{2,1}[f] := \int_\alpha^\beta \left(\int_{\overline{a}(y)}^{\overline{b}(y)} f(x, y)dx \right) dy.$$

Exercice : En s'aidant d'une représentation graphique du domaine $\mathcal{D}$ de l'exercice précèdant, vérifier qu'on peut ici également considérer $I_{2,1}[e^{x^2+y}]$. La remarque que $y = \ln(x) \Leftrightarrow x = e^y$ et $y = \ln(3x) \Leftrightarrow x = e^y/3$ et la représentation graphique de $\mathcal{D}$ permettent de trouver que $[\alpha, \beta] = [0, \ln 6]$, que $\overline{\alpha}(y) = 1$ si $y \leq \ln 3$[1], et $\overline{\alpha}(y) = e^y/3$ si $y \geq \ln 3$, et enfin que $\overline{\beta}(y) = e^y$ si $y \leq \ln 2$, et $\overline{\beta}(y) = 2$ si $y \geq \ln 2$. On notera que l'expression $I_{2,1}[e^{x^2+y}]$ obtenue ne peut s'exprimer aussi simplement que $I_{1,2}[e^{x^2+y}]$, faute de disposer d'une primitive de e^{x^2}. Pourtant :

[1] Hommage à Homère !

Théorème 3.1 (Fubini) *Soit f définie et continue sur un ouvert U contenant le domaine $\mathcal{D}$. Si le domaine $\mathcal{D}$ est tel que $I_{1,2}[f]$ et $I_{2,1}[f]$ sont tous les deux définis, alors $I_{1,2}[f] = I_{2,1}[f]$. La valeur commune est, par définition, égale à $\iint_{\mathcal{D}} f$.*

Ce théorème est une "évidence" si l'on observe que $I_{1,2}[f]$ comme $I_{2,1}[f]$ représentent l'un et l'autre le volume de la portions du "cylindre" $\mathcal{D} \times \mathbb{R}_z$ comprise entre le plan $z = 0$ et le graphe de f, volume algébrique (i.e. éventuellement négatif), compté positivement là où $f(x, y) \geq 0$ et négativement là où $f(x, y) \leq 0$. La difficulté réside dans le fait que la notion physique de volume n'est pas définie mathématiquement à ce stade, sauf pour le cas d'un parallélépipède $\Pi(v_1, v_2, v_3) = \{\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 | \lambda_i \in [0, 1]\}$ dont le volume est défini par

$$\mathrm{vol}\, \Pi(v_1, v_2, v_3) := |\det_{\mathcal{B}}(v_1, v_2, v_3)|,$$

où $\det_{\mathcal{B}}(v_1, v_2, v_3)$ est le déterminant de la matrice formée des vecteurs v_1, v_2, v_3 exprimés dans la base $\mathcal{B}$, dont le choix equivaut au choix de l'unité de volume.

3.2 Intégrale double

A l'analogue de la notion d'aire, voici comment définir la notion de volume (algébrique) à partir de la notion de volume d'un parallélépipède.

3.2.1 Volume algébrique de la région $\mathcal{V}$ limitée par le graphe de f et $z = 0$

Nous considérons un domaine *standard* quelconque $\mathcal{D}$, contenu dans un "pavé fermé" $[a, b] \times [\alpha, \beta]$ et une fonction f *standard* continue définie sur $\mathcal{D}$. Soient $(x_i)_{i=i_-..i_+}$ et $(y_j)_{j=j_-..j_+}$ des découpages fins de $[a, b]$ et $[\alpha, \beta]$ respectivement, ce qui signifie que

$$a = x_{i_-} < x_{i_-+1} < \ldots < x_{i_+-1} < x_{i_+} = b \text{ , et}$$

$$\beta = y_{j_-} < y_{j_-+1} < \ldots < y_{j_+-1} < y_{j_+} = \beta \text{ ,}$$

avec $x_{i+1} \simeq x_i$ et $y_{j+1} \simeq y_j$ pour tous i et j. Dans chaque segment $[x_i, x_{i+1}[$ choisissons un point ξ_i et de même, choisissons un $\eta_j \in [y_j, y_{j+1}[$ dans chacun de ces segments. Ainsi, chaque point (ξ_i, η_j) appartient au *pavé* $\mathcal{P}_{ij} := [x_i, x_{i+1}[\times[y_j, y_{j+1}[$. Posons $\overline{\mathcal{P}}_{ij} := [x_i, x_{i+1}] \times [y_j, y_{j+1}]$. Le nombre positif $\Delta_{ij} := (x_{i+1} - x_i)(y_{j+1} - y_j)$ représente aussi bien l'aire (arithmétique, i.e. positive) du pavé $\mathcal{P}_{ij}$ que du pavé fermé $\overline{\mathcal{P}}_{ij}$. Si $(\xi_i, \eta_j) \in \mathcal{D}$, $f(\xi_i, \eta_j)$ est défini, et nous pouvons condidérer la somme S suivante, appelée *somme de Riemann double* associée au pavage pointé considéré du domaine $\mathcal{D}$,

$$S := \sum_{(\xi_i, \eta_j) \in \mathcal{D}} f(\xi_i, \eta_j) \Delta_{ij}. \tag{3.1}$$

Le nombre $f(x_i, y_j)\Delta_{ij}$ représente le volume du parallélépipède de base $\mathcal{P}_{ij}$ et de hauteur égale à la valeur de f au point (ξ_i, η_j) choisi dans $\mathcal{P}_{ij}$; cette intuition géométrique demande de concevoir $\mathbb{R}^2_{xy}$ comme le plan $z = 0$ de $\mathbb{R}^3_{xyz}$, et son intérêt apparaît lorsqu'on conçoit f comme étant son graphe $z = f(x, y)$. Comme la fonction f est continue, $f(\xi_i, \eta_j) \simeq f(x, y)$ pour tout $(x, y) \in \mathcal{P}_{ij}$. Comme f et $\mathcal{D}$ sont standard, il est naturel de *définir* le volume algébrique de la région limitée par le graphe de f et $z = 0$ comme étant la partie standard de S, que nous notons $\iint_{\mathcal{D}} f$ ou encore $\iint_{\mathcal{D}} f(x, y) dx dy$:

$$\iint_{\mathcal{D}} f := {}^o\!\left(\sum_{(\xi_i, \eta_j) \in \mathcal{D}} f(\xi_i, \eta_j) \Delta_{ij} \right) =: {}^oS. \tag{3.2}$$

Observons qu'ainsi nous avons défini une opération $f \mapsto \iint_{\mathcal{D}} f$ que pour des f et $\mathcal{D}$ *standard.* Nous *postulons* en outre que cette opération s'étend à une opération standard définie pour tous $\mathcal{D}$ et f, standard ou non. La justification de ce postulat est purement logique : c'est un cas particulier de l'axiome de standardisation, une des trois axiomes fondant l'analyse non standard.

3.2.2 Réduction aux fonctions positives

Si $f : \mathcal{U} :\longrightarrow \mathbb{R}$ et $g : \mathcal{U} :\longrightarrow \mathbb{R}$ sont deux fonctions, posons

$$\mathrm{Max}_{f,g}(u) := \mathrm{Max}\,\{f(u), g(u)\} \text{ et } \mathrm{Min}_{f,g}(u) := \mathrm{Min}\,\{f(u), g(u)\}.$$

On voit immédiatement que $\mathrm{Min}_{f,g} = -\mathrm{Max}_{-f,-g}$.

Proposition 3.2 *Si f et g sont continues, il en est de même de $Max_{f,g}$ et $Min_{f,g}$.*

Preuve : Comme $\mathrm{Min}_{f,g} = -\mathrm{Max}_{-f,-g}$, il suffit de montrer la continuité de $\mathrm{Max}_{f,g}$. Par transfert, on peut supposer que f, g, et $\mathcal{U}$ sont standard. Soit u_0 standard, et $u \simeq u_0$. Supposons, pour fixer les idées, que $\mathrm{Max}_{f,g}(u) = f(u)$, donc

$$f(u) \geq g(u),$$

d'où, par continuité de f et g,

$$f(u_0) = {}^o\!(f(u) \geq {}^o\!(g(u)) = g(u_0),$$

et donc $\mathrm{Max}_{f,g}(u_0) = f(u_0)$, d'où, par continuité de f,

$$\mathrm{Max}_{f,g}(u) = f(u) \simeq f(u_0) = \mathrm{Max}_{f,g}(u_0).$$

□

Pour toute fonction f, posons $f_+ = \mathrm{Max}_{f,0}$ et $f_- = \mathrm{Max}_{-f,0}$. C'est deux fonctions sont manifestement positives, et il est facile de voir que $f = f_+ - f_-$. On en déduit que toute fonction f est la différence de deux fonctions positives, et que si f est continue, ces deux fonctions positives sont continues. Enfin

$$\begin{aligned}\iint_{\mathcal{D}} f &= {}^o\!\left(\sum_{(\xi_i,\eta_j)\in\mathcal{D}} (f_+ - f_-)(\xi_i,\eta_j)\Delta_{ij}\right) = {}^o\!\left(\sum_{(\xi_i,\eta_j)\in\mathcal{D}} f_+(\xi_i,\eta_j)\Delta_{ij}\right) - {}^o\!\left(\sum_{(\xi_i,\eta_j)\in\mathcal{D}} f_-(\xi_i,\eta_j)\Delta_{ij}\right)\\ &= \iint_{\mathcal{D}} f_+ \ - \ \iint_{\mathcal{D}} f_-.\end{aligned}$$

Nous pourrons donc dorénavant nous borner à l'étude de l'intégrale de fonctions continues *positives.*

3.3 Preuve du théorème de Fubini

Nous supposons dorénavant que f est positive. Observons que la somme S est le volume d'une réunion de parallélépipèdes, et cette réunion n'est pas tout-à-fait égale à la région entre $\mathcal{D}$ et le graphe de f : le "haut" des parallélépipèdes n'est qu'i-voisin du graphe de f, et les parallélépipèdes "au-dessus" des pavés $\mathcal{P}_{ij}$ à cheval sur le bord $y = \overline{\alpha}(x,y)$ et le bord $y = \overline{\beta}(x,y)$ sont tantôt pris en entier (si $(\xi_i,\eta_j) \in \mathcal{D}$), tantôt totalement exclus (si $(\xi_i,\eta_j) \notin \mathcal{D}$). En fait, nous voyons que

$$0 \leq S_- := \sum_{\mathcal{P}_{ij}\subset\mathcal{D}} f(\xi_i,\eta_j)\Delta_{ij} \leq S := \sum_{(\xi_i,\eta_j)\in\mathcal{D}} f(\xi_i,\eta_j)\Delta_{ij} \leq \sum_{\mathcal{P}_{ij}\cap\mathcal{D}\neq\emptyset} f(\xi_i,\eta_j)\Delta_{ij} =: S_+ \tag{3.3}$$

puique ces sommes de nombres tous positifs sont étendues à des familles de plus en plus larges : tout d'abord les pavés tout entier contenus dans $\mathcal{D}$, puis les pavés dont le point (ξ_i,η_j) est contenu dans $\mathcal{D}$, enfin ceux qui rencontrent $\mathcal{D}$ en un point au moins. Posons quelques notations : soient

$$m := \mathrm{Min}\, f(\mathcal{D}) \ , \ m_{ij} := \mathrm{Min}\, f(\overline{\mathcal{P}}_{ij}) \ , \ M_{ij} := \mathrm{Max}\, f(\overline{\mathcal{P}}_{ij}) \ , \text{ et } M := \mathrm{Max}\, f(\mathcal{D}).$$

Ces nombres sont bien définis puisque f est continue, et $\mathcal{D}$ et les $\overline{\mathcal{P}}_{ij}$ fermés et bornés, et $m_{ij} \simeq f(\xi_i,\eta_j) \simeq M_{ij}$, puisque la fonction f est en outre standard. On a $m_{ij} \leq M_{ij}$ et $m_{ij} \simeq M_{ij}$, d'où $0 \leq \delta_{ij} := M_{ij} - m_{ij} \simeq 0$. Soit δ le plus grand des δ_{ij} de sorte que

$$0 \leq M_{ij} - m_{ij} = \delta_{ij} \leq \delta \simeq 0. \tag{3.4}$$

3.3.1 Aire de $\mathcal{D}$

Attachons notre attention à l'aire de la réunions des pavés. Rappelons que le pavé $\mathcal{P}_{ij}$ a pour aire $\Delta_{ij} = (x_{i+1} - x_i)(y_{j+1} - y_j)$; posons

$$d_- := \sum_{\mathcal{P}_{ij} \subset \mathcal{D}} \Delta_{ij} \ , \quad d_0 := \sum_{(\xi_i,\eta_j) \in \mathcal{D}} \Delta_{ij} \ , \text{ et } \ d_+ := \sum_{\mathcal{P}_{ij} \cap \mathcal{D} \neq \emptyset} \Delta_{ij}. \tag{3.5}$$

Comme $\mathcal{D}$ contient tous les pavés (disjoints) pris en compte dans la premières sommes, et est contenu dans la réunion des pavés (disjoints) pris en compte dans la dernière, son aire

$$\partial := \int_a^b \overline{\beta} - \int_a^b \overline{\alpha} \quad \left(= \int_a^b (\overline{\beta} - \overline{\alpha}) \right)$$

satisfait à $d_- \leq \partial \leq d_+$, et comme $\overline{\alpha}$ et $\overline{\beta}$ sont standard et intégrables, $d_- \simeq d_+$, et donc

$$d_- \simeq \int_a^b (\overline{\beta} - \overline{\alpha}) \simeq d_+. \tag{3.6}$$

Aire de la "tranche" $x = \xi_i$

Posons

$$A(\xi) = \int_{\overline{\alpha}(\xi)}^{\overline{\beta}(\xi)} f(\xi, y) dy \ ;$$

$A(\xi)$ représente l'aire de l'intersection de la région $\mathcal{V}$ dont nous étudions le volume avec le plan $x = \xi$. Pour $\xi = \xi_{i_0}$, comme $0 \leq m_{i_0 j} \leq f(\xi_{i_0}, y) \leq M_{i_0 j}$ pour tout $y \in [y_j, y_{j+1}]$, on a

$$\sum_{\mathcal{P}_{i_0 j} \subset \mathcal{D}} m_{i_0 j}(y_{j+1} - y_j) \leq A(\xi_{i_0}) \leq \sum_{\mathcal{P}_{i_0 j} \cap \mathcal{D} \neq \emptyset} M_{i_0 j}(y_{j+1} - y_j). \tag{3.7}$$

Revenons à présent aux sommes des volumes des parallélépipèdes; par (3.3) et définition de m_{ij} et M_{ij}, nous avons

$$0 \leq \sum_{\mathcal{P}_{ij} \subset \mathcal{D}} m_{ij} \Delta_{ij} \leq \sum_{(\xi_i,\eta_j) \in \mathcal{D}} f(\xi_i, \eta_j) \Delta_{ij} \leq \sum_{\mathcal{P}_{ij} \cap \mathcal{D} \neq \emptyset} M_{ij} \Delta_{ij} \leq M d_+. \tag{3.8}$$

De façon analogue, si nous ne considérons que les $\mathcal{P}_{ij}$ tels que $i = i_0$, et mettons en facteur le terme $(x_{i_0+1} - x_{i_0})$ des $\Delta_{i_0 j} := (x_{i_0+1} - x_{i_0})(y_{j+1} - y_j)$ nous obtenons

$$\begin{aligned}
0 \ &\leq \ (x_{i_0+1} - x_{i_0}) \sum_{\mathcal{P}_{i_0 j} \subset \mathcal{D}} m_{i_0 j}(y_{j+1} - y_j) \\
&\leq \ (x_{i_0+1} - x_{i_0}) \sum_{(\xi_{i_0},\eta_j) \in \mathcal{D}} f(\xi_{i_0}, \eta_j)(y_{j+1} - y_j) \\
&\leq \ (x_{i_0+1} - x_{i_0}) \sum_{\mathcal{P}_{i_0 j} \cap \mathcal{D} \neq \emptyset} M_{i_0 j}(y_{j+1} - y_j),
\end{aligned}$$

et par (3.7) on a

$$0 \leq (x_{i_0+1} - x_{i_0}) \sum_{\mathcal{P}_{i_0 j} \subset \mathcal{D}} m_{i_0 j}(y_{j+1} - y_j) \leq (x_{i_0+1} - x_{i_0}) A(\xi_{i_0}) \leq (x_{i_0+1} - x_{i_0}) \sum_{\mathcal{P}_{i_0 j} \cap \mathcal{D} \neq \emptyset} M_{i_0 j}(y_{j+1} - y_j).$$

Regroupement des "tranches" $x = \xi_i$

Etendons à nouveau la somme à tous les indices i. Nous obtenons

$$\sum_{\mathcal{P}_{ij} \subset \mathcal{D}} m_{ij} \Delta_{ij} \leq \sum_{[x_i, x_{i+1}[\subset [a,b]} (x_{i+1} - x_i) A(\xi_i) \leq \sum_{\mathcal{P}_{ij} \cap \mathcal{D} \neq \emptyset} M_{ij} \Delta_{ij}. \tag{3.9}$$

Or, par définition de $\int_a^b A(x) dx$, nous avons

$$\sum_{[x_i, x_{i+1}[\subset [a,b]} (x_{i+1} - x_i) A(\xi_i) \simeq \int_a^b A(x) dx. \tag{3.10}$$

Par ailleurs, les trois termes de l'inégalité (3.9) sont tous i-voisins les uns des autres : ceci découle du lemme suivant :

Lemme 3.3

$$\Delta V := \sum_{\mathcal{P}_{ij}\cap\mathcal{D}\neq\emptyset} M_{ij}\Delta_{ij} - \sum_{\mathcal{P}_{ij}\subset\mathcal{D}} m_{ij}\Delta_{ij} \simeq 0.$$

Mais achevons tout d'abord la preuve du théorème en admettant provisoirement ce résultat. En appliquant le lemme 3.3 à (3.8) et (3.9), nous voyons que les termes médians de ces inégalités sont i-voisins :

$$\sum_{(\xi_i,\eta_j)\in\mathcal{D}} f(\xi_i,\eta_j)\Delta_{ij} \simeq \sum_{[x_i,x_{i+1}[\subset[a,b]} (x_{i+1}-x_i)A(\xi_i),$$

d'où, par (3.10), en utilisant la définition de $A(x)$, nous avons finalement

$$\int_a^b \left(\int_{\overline{\alpha}(x)}^{\overline{\beta}(x)} f(x,y)dy\right) dx \simeq \sum_{(\xi_i,\eta_j)\in\mathcal{D}} f(\xi_i,\eta_j)\Delta_{ij}, \tag{3.11}$$

d'où, comme $\int_a^b \left(\int_{\overline{\alpha}(x)}^{\overline{\beta}(x)} f(x,y)dy\right) dx$ est standard,

$$\int_a^b \left(\int_{\overline{\alpha}(x)}^{\overline{\beta}(x)} f(x,y)dy\right) dx = {}^o\!\left(\sum_{(\xi_i,\eta_j)\in\mathcal{D}} f(\xi_i,\eta_j)\Delta_{ij}\right) =: \iint_{\mathcal{D}} f.$$

On montrerait de même que

$$\int_\alpha^\beta \left(\int_{\overline{a}(y)}^{\overline{b}(y)} f(x,y)dx\right) dy = \iint_{\mathcal{D}} f$$

ce qui achève la preuve du théorème de Fubini 3.1, à la preuve du lemme près. □

Preuve du lemme 3.3 :

Appelons "pavé à cheval" tout $\mathcal{P}_{ij}$ tel qu'à la fois $\mathcal{P}_{ij}\cap\mathcal{D}\neq\emptyset$ et $\mathcal{P}_{ij}\not\subset\mathcal{D}$. En (3.6) nous avons vu que l'aire totale $d_+ - d_-$ des pavés à cheval est i-petite. Par ailleurs

$$\sum_{\mathcal{P}_{ij}\cap\mathcal{D}\neq\emptyset} M_{ij}\Delta_{ij} = \sum_{\mathcal{P}_{ij}\subset\mathcal{D}} M_{ij}\Delta_{ij} + \sum_{\mathcal{P}_{ij}\text{ à cheval}} M_{ij}\Delta_{ij}.$$

d'où

$$\Delta V = \sum_{\mathcal{P}_{ij}\subset\mathcal{D}} (M_{ij}-m_{ij})\,\Delta_{ij} + \sum_{\mathcal{P}_{ij}\text{ à cheval}} M_{ij}\Delta_{ij}.$$

Or tous les $M_{ij}-m_{ij}(=M_{ij}-m_{ij}\geq 0)$ sont i-petits ; soit δ le plus grand de ces nombres. Par ailleurs $0=m\leq M_{ij}\leq M=M-m$ pour tout i et j, et comme $M-m$ est standard, il est limité. Finalement, par (3.4),

$$\begin{aligned} 0 \;\leq\; \Delta V \;&\leq\; \sum_{\mathcal{P}_{ij}\subset\mathcal{D}} (M_{ij}-m_{ij})\,\Delta_{ij} \;+\; \sum_{\mathcal{P}_{ij}\text{ à cheval}} M_{ij}\Delta_{ij} \\ &\leq\; \delta\sum_{\mathcal{P}_{ij}\subset\mathcal{D}}\Delta_{ij} \;+\; M\left(\sum_{\mathcal{P}_{ij}\text{ à cheval}}\Delta_{ii}\right) \\ &=\; \delta d_- \;+\; M(d_+-d_-) \;\simeq\; 0. \end{aligned}$$

□

Chapitre 4

Espaces normés

Les normes sont des généralisations de la notion de longueur de la géométrie euclidienne. Elles permettent de définir la notion de fonctions continues et surtout de fonctions différentiables entre certaines parties - les ouverts - d'espaces "normés", c'est-à-dire munis d'une norme.

4.1 Normes et boules

Définition : Une norme sur un espace vectoriel E sur $\mathbb{R}$ ou $\mathbb{C}$, généralement notée $\|\quad\|$, est une fonction de E dans $\mathbb{R}$, à valeurs positives, telle que, pour tous v et w dans E et tous scalaires (nombres) λ et μ, les trois propriétés suivantes sont satisfaites :

1. $\|v\| = 0$ si et seulement si $v = 0$ (séparation)
2. $\|\lambda v\| = |\lambda|\|v\|$ (homogénéité)
3. $\|v + w\| \leq \|v\| + \|w\|$ (inégalité triangulaire)

Selon que λ est réel ou complexe, $|\lambda|$ désigne sa valeur absolue ou son module.

Observons que par l'homogénéité, pour tout v, on a $\|v\| = \|-v\|$, et si $v \neq 0$, le vecteur $u := v/\|v\|$ est tel que $\|u\| = 1$: on dit que u est le *normalisé* du vecteur v.

Dans le plan affine, si l'on choisi une origine O, alors pour tout couple de point M et N, on peut poser $v := \vec{OM}$ et $w := \vec{ON}$; dans ce cas, la distance de M à N est la longueur du vecteur $\vec{MN} = \vec{ON} - \vec{OM} = w - v$; c'est cette idée que l'on généralise dans un espace normé quelconque de la manière suivante : on appelle *distance* de v à w le nombre $d(v, w) := \|w - v\|$; la distance de v à w est donc nulle si et seulement si $v = w$. La *boule ouverte* centrée en v et de rayon r est l'ensemble, noté $B(v, r)$, de tous les w tels que $d(v, w) < r$. La *boule-fermée* centrée en v et de rayon r est l'ensemble, noté $\overline{B}(v, r)$, de tous les w tels que $d(v, w) \leq r$.

Figure 4.1: a) les boules $B((1, 2), 1)$ et $\overline{B}((2, 2), 0.5)$ dans $(\mathbb{R}^2, \|\quad\|_1)$; b) $\overline{B}((0, 0), 1)$ dans $(\mathbb{R}^2, \|\quad\|_\infty)$.

La deuxième inégalité triangulaire

Si v et w sont deux vecteurs quelconques, on a $\|v\| = \|v-w+w\| \le \|v-w\|+\|w\|$, d'où $\|v\|-\|w\| \le \|v-w\|$. En échageant les rôles de v et w, on a de même $\|w\| - \|v\| \le \|w - v\| = \|v - w\|$. Donc, que $\|v\| - \|w\|$ soit positif ou négatif, on a

$$|\|v\| - \|w\|| \le \|v - w\| (\text{deuxième inégalité triangulaire})$$

Une conséquence de cette "deuxième inégalité triangulaire" est que si $v \simeq w$, alors $\|v\| \simeq \|w\|$: deux vecteurs i-voisins sont de longueurs i-voisines. Ceci montre également que la fonction $\| \ \| : (E, \| \ \|) \longrightarrow (\mathbb{R}, | \ |)$, $v \mapsto \|v\|$, est une fonction continue.

4.1.1 Quelques normes usuelles

La norme euclidienne

Il est bien connu que la "norme euclidienne" $\|v\|_2 = \|(v_1, \dots, v_d)\|_2 := (v_1^2 + \dots + v_d^2)^{\frac{1}{2}}$ est une norme sur $\mathbb{R}^d$. Elle est très importante et est étudiée et généralisée lors de l'étude des *produits scalaire* ($\|v\|_2 := (v \cdot v)^{\frac{1}{2}}$).

La norme $\| \ \|_1$ sur $\mathbb{R}^d$

Elle est défini par $\|v\|_1 = \|(v_1, ..., v_d)\|_1 := |v_1| + ... + |v_d|$. C'est une norme ; en effet $0 = \|v\|_1 = |v_1| + ... + |v_d|$ si et seulement si $|v_i| = 0$ pour tout i, c'est-à-dire $v = 0$, d'où la propriété de séparation. D'autre part $\|\lambda v\|_1 = |\lambda v_1| + ... + |\lambda v_d| = |\lambda|(|v_1| + ... + |v_d|) = |\lambda|\|v\|_1$, d'où l'homogénéité. Enfin $\|v + w\|_1 = |v_1 + w_1| + ... + |v_d + +w_d| \le |v_1| + |w_1| + ... + |v_d| + |w_d| = |v_1| + ... + |v_d| + |w_1| + ... + |w_d| = \|v\|_1 + \|w\|_1$, d'où l'inégalité triangulaire.

La norme du max

Rappelons que pour $\mathcal{E} \subset \mathbb{R}$, $\mathrm{Max}\,\mathcal{E}$ désigne le plus grand élément de $\mathcal{E}$, s'il existe, ce qui est toujours le cas si $\mathcal{E}$ est fini, ou par exemple si $\mathcal{E} =]0, 1]$; en revanche $\mathrm{Max}\,[0, 1[$ n'existe pas (alors que $\mathrm{Sup}\,[0, 1[$ existe et vaut 1).

Pour $v = (v_1, ..., v_d)$, on pose $\|v\|_\infty := \mathrm{Max}\,\{|v_i|, i = 1..d\}$; c'est $|v_{i_0}|$ pour un (ou plusieurs) indice(s) i_0. On a $0 = \|v\|_\infty = \mathrm{Max}\,\{|v_i|\}$ si et seulement si $|v_i| = 0$ pour tout i, ce qui équivaut à $v = (0, \dots, 0) = 0$, d'où la propriété de séparation. La propriété d'homogénéité est évidente si $\lambda = 0$; si $|\lambda| > 0$, comme $|v_i| \le |v_{i_0}|$ pour tout i si et seulement si $|\lambda||v_i| \le |\lambda||v_{i_0}|$ pour tout i, on a

$$|\lambda|\|v\|_\infty = |\lambda|\mathrm{Max}\,\{|v_i|\} = |\lambda||v_{i_0}| = \mathrm{Max}\,\{|\lambda||v_i|\} = \mathrm{Max}\,\{|\lambda v_i|\} = \|\lambda v\|_\infty,$$

d'où l'homogénéité. Enfin, soient i_0 et i_1 tels que $\|v\|_\infty = |v_{i_0}|$ et $\|w\|_\infty = |w_{i_1}|$. Donc pour tout indice i on a $|v_i + w_i| \le \ (|v_i| + |w_i|) \le |v_{i_0}| + |w_{i_1}| =) \ \|v\|_\infty + \|w\|_\infty$, et donc $\|v + w\|_\infty = \mathrm{Max}\,\{|v_i + w_i|\} \le \|v\|_\infty + \|w\|_\infty$, d'où l'inégalité triangulaire.

L'exemple $v = (100, 1, -99)$ et $w = (-99, 1, 100)$, d'où $v+w = (1, 2, 1)$ montre que l'inégalité peut être stricte et surtout que ce sont des composantes d'indices différents qui peuvent réaliser les trois nombres $\|v\|_\infty$, $\|w\|_\infty$, et $\|v + w\|_\infty$.

Un exemple de norme non-standard

Soit $\varepsilon > 0$ i-petitet $E = \mathbb{R}^2$. Posons $\|(x, y)\| := \varepsilon|x| + |y|$. Nous voyons que le vecteur standard $(1, 0)$ a pour norme $\|(1, 0)\| = \varepsilon$ qui est non-standard ; cette norme est donc non-standard. Observons aussi que le vecteur $(1/\varepsilon, 0)$ a pour norme $\|(1/\varepsilon, 0)\| = 1$ bien que l'une de ses composantes soit i-grand.

Exercice : Dessiner la boule $B((0, 0), 1)$ pour cette norme.

Norme sur un produit cartésien $E' \times E''$

Rappelons que le produit cartésien de deux espaces vectoriels E' et E'' est l'espace vectoriel des couples (v', v'') ; si les deux espaces sont munis respectivement des normes $\| \ \|'$ et $\| \ \|''$, on défini une norme sur $E' \times E''$, en posant, pour $v := (v', v'')$, $\|v\| := \mathrm{Max}\,\{\|v'\|', \|v''\|''\}$. La vérification est identique au cas précédant ; nous la laissons en exercice.

4.1.2 Convexité des boules

Si E est un espace vectoriel et si v et w sont des éléments de E, on appelle segment fermé joignant v et w et on note $[v,w]$ le sous-ensemble de E suivant :

$$[v,w] := \{(1-\lambda)v + \lambda w \mid \lambda \in [0,1]\}$$

Rappelons qu'un sous-ensemble $C \subseteq E$ est dit *convexe* si et seulement si pour tout $v \in C$ et tout $w \in C$ on a $[v,w] \subseteq C$.

Proposition 4.1 *Soit $(E, \| \ \|)$ un espace normé. Toute boule-ouverte $B(u,r)$ et toute boule-fermée $\overline{B}(u,r)$ est convexe.*

Preuve : Calculons la distance du point $(1-\lambda)v + \lambda w$ du segment $[v,w]$ au centre u de la boule. On a, en écrivant $1 = 1-\lambda+\lambda$,

$$\begin{aligned} \|(1-\lambda)v + \lambda w - u\| &= \|(1-\lambda)v + \lambda w - ((1-\lambda)u + \lambda u)\| && (4.1)\\ &= \|(1-\lambda)(v-u) + \lambda(w-u)\| && (4.2)\\ &\leq |1-\lambda|\|v-u\| + |\lambda|\|w-u\|. && (4.3) \end{aligned}$$

A présent, comme $\lambda \in [0,1]$, $|1-\lambda| = 1-\lambda$ et $|\lambda| = \lambda$, et d'ailleurs un de ces deux nombres est non nul. Donc si $\|v-u\| \leq r$ et $\|w-u\| \leq r$, alors $|1-\lambda|\|v-u\| + |\lambda|\|w-u\| \leq (1-\lambda)r + \lambda r = r$, d'où on déduit que toute boule-fermée $\overline{B}(u,r)$ est convexe, et, pour $r > 0$, si $\|v-u\| < r$ et $\|w-u\| < r$, alors $|1-\lambda|\|v-u\| + |\lambda|\|w-u\| < (1-\lambda)r + \lambda r = r$, d'où on déduit que toute boule-ouverte $B(u,r)$ est convexe. □

4.2 Continuité

4.2.1 Ordres de grandeurs dans un normé standard

Dire que $(E, \| \ \|)$ est standard signifie à la fois que l'espace vectorie E est standard et que la fonction $\| \ \| : E \longrightarrow \mathbb{R}$ est standard. Ainsi, si $E = \mathbb{R}^d$, ceci implique que d est standard, et si $v \in E$ est un vecteur standard, alors il en est de même de sa norme $\|v\|$. Dans $\mathbb{R}$ on sait définir la notion de nombre i-petit, limité, appréciable, et i-grand. Cette notion s'étend immédiatement aux vecteurs d'un espace normé : $v \in E$ est i-petit, limité, appréciable, ou i-grand si tel est le cas de $\|v\|$. On dit que les vecteurs v et w sont i-voisins si et seulement si $v - w$ est i-petit.

La proposition suivante est une simple traduction géométrique de $\|v-w\| \simeq 0$:

Proposition 4.2 *On a $v \simeq w$ si et seulement si v appartient a toute boule-ouverte non vide de rayon standard centrée en w.*

Preuve : Exercice. □

4.2.2 Normes équivalentes

La notion de vecteur i-petit dépend bien entendu du choix de la norme ; par exemple, si $\varepsilon > 0$ est i-petit, $\|v\| := \varepsilon\|v\|_2$ détermine une norme sur $\mathbb{R}^2$ pour laquelle le vecteur $(1,1)$ est i-petit. Toutefois, cette nouvelle norme n'est pas standard (alors que $\| \ \|_2$ l'est).

La notion de normes équivalentes est une notion standard. Deux normes standard $\| \ \|'$ et $\| \ \|''$ sur E sont dites équivalentes, si et seulement si elles déterminent les mêmes vecteurs i-petits, c'est-à-dire que $\|v\|'$ est i-petit si et seulement si $\|v\|''$ est i-petit. En fait, si deux normes déterminent les mêmes i-petits, elles déterminent les mêmes ordres de grandeurs. Plus précisément :

Proposition 4.3 *Pour deux normes standard quelconques $\| \ \|'$ et $\| \ \|''$ sur un même espace standard, les propriétés suivantes sont équivalentes*

1. *Les deux normes déterminent les mêmes vecteurs i-petits.*
2. *Les deux normes déterminent les mêmes vecteurs limités.*
3. *Les deux normes déterminent les mêmes vecteurs appréciables.*
4. *Les deux normes déterminent les mêmes vecteurs i-grands.*
5. *Il existe des $M \geq m > 0$ tels que pour tout $v \in E$ $m\|v\|' \leq \|v\|'' \leq M\|v\|'$.*

Preuve : Montrons tout d'abord que **(i)⇒(ii)⇒(iii)⇒(iv)⇒(i)** ce qui montrera l'équivalence des quatre premières notions.

(i)⇒(ii) Soit v un vecteur dont l'une des normes, par exemple $\|v\|'$, soit limitée. Posons $\nu := \|v\|''$ et $u := v/\nu$. On a donc $\|u\|'' = 1$. L'hypothèse **(i)** implique que $\|u\|'$ est non i-petit comme $\|u\|''$. Or $\|u\|' = \|v/\nu\|' = \|v\|'/\nu$. Donc ν ne peut être i-grand. Donc la norme $\|v\|''$ $(= \nu)$ est, elle aussi, limitée.

(ii)⇒(iii) Soit v un vecteur dont l'une des normes, par exemple $\|v\|'$, soit appréciable. Posons $\nu := \|v\|''$ et $u := v/\nu$. On a donc $\|u\|'' = 1$. L'hypothèse **(ii)** implique que $\|v\|''$ est limité comme $\|v\|'$. Donc ν ne peut être i-grand. L'hypothèse **(ii)** implique que $\|u\|'$ est limité comme $\|u\|''$. Or $\|u\|' = \|v\|'/\nu$. Donc ν ne peut être i-petit. Donc la norme $\|v\|''$ $(= \nu)$ est, elle aussi, appréciable.

(iii)⇒(iv) Soit v un vecteur dont l'une des normes, par exemple $\|v\|'$, soit i-grande. Posons $\nu := \|v\|''$ et $u := v/\nu$. On a donc $\|u\|'' = 1$. L'hypothèse **(iii)** implique que $\|u\|'$ est appréciable comme $\|u\|''$. Or $\|u\|' = \|v\|'/\nu$. Donc ν ne peut être limité. Donc la norme $\|v\|''$ $(= \nu)$ est, elle aussi, i-grande.

(iv)⇒(i) Soit v un vecteur dont l'une des normes, par exemple $\|v\|'$, soit i-petite. Posons $\mu := \|v\|' \simeq 0$ et $w := v/\mu$. On a donc $\|w\|' = 1$. L'hypothèse **(iv)** implique que $\|w\|''$ est non i-grand comme $\|w\|'$. Or $\|w\|'' = \|v\|''/\mu$. Donc $\|v\|''$ doit être i-petit puisque μ l'est. Donc la norme $\|v\|''$ $(= \nu)$ est, elle aussi, i-petite.

Montrons à présent que **(iii)⇔(v)** ce qui achèvera la preuve de la proposition. Observons que la relation 4.3-5 est toujours vrai si $v = 0$. Si $v \neq 0$, la relation 4.3-5 équivaut à $m \leq \|v\|''/\|v\|' \leq M$. Or $\|v\|''/\|v\|' = \|v/\|v\|'\|'' = \|u\|''$, où $u = v/\|v\|'$. Dire que la relation 4.3-5est vraie pour tout $v \in E$ équivaut donc à dire que

$$\text{Pour tout } u \in E \text{ tel que } \|u\|' = 1 \text{ on a} m \leq \|u\|'' \leq M \tag{4.4}$$

(iii)⇒(v) L'hypothèse **(iii)** implique que $\|u\|''$ est appréciable dès que $\|u\|' = 1$. Donc 4.4 est vrai pour tout m i-petit est tout M i-grand. Nous pouvons donc considérer $m_0 := \mathrm{Inf}\{\|u\|'' \mid u \in E, \|u\|' = 1\}$ et $M_0 := \mathrm{Sup}\{\|u\|'' \mid u \in E, \|u\|' = 1\}$. Nous avons vu que m_0 ne peut être i-petit et M_0 ne peut être i-grand. Ils sont donc appréciables l'un et l'autre. Il suffit donc de choisir des standard m et M tels que $0 < m \leq m_0 \leq M_0 \leq M$, par exemple $m := {}^o\!(m_0)/2$ et $M := 2^o\!(M_0)$.

(v)⇒(iii) Il suffit de remarquer que si $M \geq m > 0$ sont standard, la relation 4.3-5 et la relation $\|v\|''/M \leq \|v\|' \leq \|v\|''/m$ qui en découle montrent immédiatement que $\|v\|'$ est appréciable si et seulement si $\|v\|''$ est appréciable.

□

La propriété 4.3-5 est manifestement standard : c'est elle qu'on a retenu pour *définir* la notion de normes équivalentes dans le cas général, c'est-à-dire que ces normes soient *standard ou non.*

Définition : Deux normes $\|\ \|'$ et $\|\ \|''$ sur E sont é*quivalentes* si et seulement si il existe des réels $M \geq m > 0$ tels que pour tout $v \in E$

$$m\|v\|' \leq \|v\|'' \leq M\|v\|'$$

4.2.3 Fonctions continues

La notion, standard, de fonctions continues s'étant également aux fonctions $f : A \longrightarrow E''$, où A est une partie d'un espace normé E', et E'' est également un espace normé : si f est une fonction standard, elle est *continue* si et seulement si pour tout standard $v_0 \in A$ et tout $v \in A$, si v est i-voisin de v_0, alors $f(v)$ est i-voisin de $f(v_0)$.

Définition : $f : A \longrightarrow E''$ est continue si et seulement si pour tout $v_0 \in A$

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall v \in A\ [\ \|v - v_0\|' \leq \delta \ \Rightarrow\ \|f(v) - f(v_0)\|''\ ].$$

Comme pour le cas où $E' = \mathbb{R} = E''$, on montre immédiatement la proposition suivante :

Proposition 4.4 *Si $f : A \longrightarrow E''$ et $g : A \longrightarrow E''$ sont continue alors $f + g$ est continue. Si $E'' = \mathbb{R}$, alors fg est continue ; si, de plus, g ne s'annule pas, alors f/g est également continue.*

Preuve : Exercice. □

Notations : Soit $l \in E''$; on dit que *l est la limite de f quand v tends vers* v_0 et on écrit $l = \lim_{v \to v_0} f(v)$ si et seulement si

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall v \in A \ \ [\ \|v - v_0\|' \leq \delta \ \Rightarrow \ \|f(v) - l\|'' \ ].$$

Avec cette notation, f est continue au point v_0 si et seulement si $\lim_{v \to v_0} f(v) = f(v_0)$. Observons que la locution "v tend vers v_0" n'a pas de sens en elle-même, alors que $v \simeq v_0$ en a un.

Proposition 4.5 (Fonction à valeurs dans un produit) *Soient E, E', et E'' des espaces normés et $A \subseteq E$; on munit $E' \times E''$ de la norme-produit. Toute fonction $f : A \longrightarrow E' \times E''$, $v \mapsto f(v) =: (f_1(v), f_2(v))$ est continue si et seulement si les fonctions-composantes $f_1 : A \longrightarrow E'$ et $f_2 : A \longrightarrow E''$ sont continues.*

Preuve : Exercice. □

La section suivante nous donnera une définition plus géométrique des fonctions continues.

4.3 Ouverts et fermés

Définition : On dit que le sous-ensemble $U \subseteq E$ est *ouvert*, si et seulement si pour tout $u \in U$, il existe un rayon $r > 0$ tel que $B(u, r) \subseteq U$.

Proposition 4.6 *Si le sous-ensemble U et l'espace normé E sont standard, $U \subseteq E$ est ouvert si et seulement si pour tout standard $u_0 \in U$ et tout $u \simeq u_0$, u appartient lui aussi à U.*

Preuve : Exercice. Indication : penser au boules de rayon i-petit et au transfert ! □

Exemple : Tout intervalle ouvert $]a, b[$ et toute demi-droite ouverte $]-\infty, b[$ ou $]a, +\infty[$ est un ouvert de $(\mathbb{R}, |\ |)$ On a le très joli théorème suivant, que nous ne démontrons pas :

Théorème 4.7 *Pour tout ouvert $U \subseteq \mathbb{R}$ il existe une famille finie ou dénombrable d'intervalles ouverts* $]a_i, b_i[$ disjoints deux-à-deux *tels que $U = \bigcup_i]a_i, b_i[$.*

Exercice : Montrer que toute boule-ouverte $B(u, r)$ est ouverte.

Définition : On dit qu'un sous-ensemble $F \subseteq E$ est *fermé* si et seulement si le complémantaire $F^c := E \setminus F$ est ouvert.

Exemple : Tout intervalle-fermé $[a, b]$ et toute demi-droite-fermée $]-\infty, b[$ ou $[a, +\infty[$ est un fermé de $(\mathbb{R}, |\ |)$

Proposition 4.8 *Si le sous-ensemble F et l'espace normé sont standard, F est fermé si et seulement si pour tout $u \in U$, s'il existe un standard $u_0 \simeq u$, alors ce standard u_0 appartient lui aussi a F.*

Preuve : D'après la proposition 4.6, F^c est ouvert si et seulement si $u_0 \in F^c$ implique que $u \in F^c$; on en déduit immédiatement les deux implications annoncées : si F est fermé, u_0 ne peut pas appartenir F^c qui est ouvert, ce qui impliquerait que $u \in F^c$: il est donc dans F ; réciproquement, si $v_0 \in F^c$ et si $v \simeq v_0$, alors v ne peut appartenir F (sans quoi on aurait aussi $v_0 \in F$), et donc $v \in F^c$, ce qui montre que F^c est ouvert. □

Exercice : Montrer que toute boule-fermée $\overline{B}(u,r)$ est fermée.

Rappelons que si $f : A \longrightarrow B$ est une fonction quelconque, pour tout sous-ensemble $V \subseteq B$, on appelle *image réciproque* de V par f et on note $f^{-1}(V)$ le sous-ensemble suivant de A :

$$f^{-1}(V) := \{x \in A \ | \ f(x) \in V\}$$

Cette notion a donc un sens et définit une fonction de $\mathcal{P}(B)$ vers $\mathcal{P}(A)$, que f soit bijective ou non. Biensûr, si de plus f est bijective, alors elle admet une *fonction réciproque* $f^{-1} : B \longrightarrow A$, et $y = f(x)$ *ssi* $x = f^{-1}(y)$, mais aussi $\{x\} = f^{-1}(\{y\})$, au sens de l'image réciproque. On voit facilement que pour tous V', V'', et V contenus dans B, on a

$$\begin{aligned} f^{-1}(V' \cup V'') &= f^{-1}(V') \cup f^{-1}(V'') \\ f^{-1}(V' \cap V'') &= f^{-1}(V') \cap f^{-1}(V'') \\ f^{-1}(B \setminus V) &= A \setminus f^{-1}(V). \end{aligned}$$

Théorème 4.9 *Soit $f : E' \to E''$ une application entre espaces normés $(E', \|\ \|_{E'})$ et $(E'', \|\ \|_{E''})$. Il y a équivalence entre les trois propriétés suivantes :*

(i) *f est continue*

(ii) *pour tout ouvert $U'' \subseteq E''$ $f^{-1}(U'')$ est un ouvert $U' \subseteq E'$*

(iii) *pour tout fermé $F'' \subseteq E''$ $f^{-1}(F'')$ est un fermé $F' \subseteq E'$.*

Preuve : Par transfert nous pouvons supposer E', E'', f, U'', et F'' standard, ce qui implique que $U' := f^{-1}(U'')$ et $F' := f^{-1}(F'')$ sont standard puisque déterminés entièrement par des ensembles et fonction standard. Nous pouvons donc utiliser la propriété caractéristique non-standard d'ouvert et fermé.

(ii)⇔(iii) Comme pour tout sous-ensemble $A \subseteq E''$ on a la relation ensembliste $f^{-1}(E'' \setminus A) = E' \setminus f^{-1}(A)$, il est immédiat de voir que les propriétés (ii) et (iii) sont équivalentes.

(i)⇒(ii) Soit $u_0 \in U'$ un élément standard quelconque de $U' = f^{-1}(U'')$, et $u \simeq u_0$. Comme f est standard continue, $f(u) \simeq f(u_0)$, et comme $f(u_0)$ est un élément standard de U'' qui est ouvert, $f(u) \in U''$: ceci montre que $u \in U'$ lui aussi, et donc U' est bien ouvert.

(ii)⇒(i) Soit u_0 un élément standard de E' et $u \simeq u_0$. Nous devons montrer que $f(u) \simeq f(u_0)$, c'est-à-dire, d'après la caractérisation géométrique d'être i-voisins 4.2, que $f(u)$ appartient a toute boule de rayon standard centrée en $f(u_0)$. Or chacune de ces boules est en particulier un ouvert standard U'' contenant $f(u_0)$; par hypothèse $U' := f^{-1}(U'')$ est un ouvert standard, et $u_0 \in U'$ puisque $f(u_0) \in U''$; par la caractérisation non standard des ouverts standards 4.6 on a donc $u \in U'$ puisque $u \simeq u_0$, et donc $f(u) \in U''$

□

Corollaire 4.10 *La composée $g \circ f$ de fonctions continues est continue.*

Preuve : Soit U un ouvert quelconque du but de l'application g, qui est aussi le but de $g \circ f$. Comme g est continue $V := g^{-1}(U)$ est un ouvert du domaine de g, qui est aussi le but de f ; comme f est continue, $f^{-1}(V)$ est un ouvert. Donc $(g \circ f)^{-1}(U) = g^{-1}(f^{-1}(U)) = g^{-1}(V)$ est bien ouvert. □

Exemples : Soient $F := \{(x,y) \ | \ (x-2y)y \leq 0\}$, $U := \{(x,y) \ | \ (x-2y)y < 0\}$, et $A := \{(x,y) \ | \ \frac{x}{2y} \leq 2\}$. Posons $f(x,y) := (x-2y)y$ qui est somme et produit de fonctions continues, donc continu. Comme $F = f^{-1}(]-\infty, 0]$, et que $]-\infty,0]$ est fermé, F est donc fermé. Comme $U = f^{-1}(]-\infty,0[$, et que $]-\infty,0[$ est ouvert, U est donc ouvert. Soit $g(x,y) := \frac{x}{y}$, qui est le rapport de deux fonctions continues, et est donc continue *sur son domaine de définition*, ici les (x,y) tels que $y \neq 0$. On a bien $A = g^{-1}(]-\infty,2]$ et $]-\infty,2]$ est bien fermé, mais A n'est pas fermé dans $\mathbb{R}^2$, car g n'est pas défini sur tout $\mathbb{R}^2$. Pourtant la relation $\frac{x}{y} \leq 2$ équivaut à $(x-2y)y \leq 0$, dès que $y \neq 0$. En fait $A = F \cap \{(x,y) \ | \ y \neq 0\}$, et F est le plus petit fermé contenant A, ou encore F est l'adhérence de A, dans le sens suivant :

Définition : On dit que $u \in E$ est *adhérent* à $A \subseteq E$ si et seulement si toute boule de rayon strictement positif centrée au point u rencontre A ; on appelle *adhérence* de A et on note $\overline{A}$ ou $\mathrm{Adh}\,(A)$ l'ensemble des points adhérents à A. En d'autres termes

$$u \in \mathrm{Adh}\,(A) \Leftrightarrow \forall r > 0 \;\; B(u,r) \cap A \neq \emptyset.$$

Proposition 4.11 *Pour tout $A \subseteq E$, $\mathrm{Adh}\, A$ est fermé ; c'est le plus petit fermé (au sens de l'inclusion) contenant A.*

Preuve : exercice. □

Proposition 4.12 *Soit E un espace normé standard. Si $u_0 \in E$ et A sont standard, alors $\mathrm{Adh}\,(A)$ est standard ; $u_0 \in \mathrm{Adh}\,(A)$ si et seulement si il existe $a \in A$ tel que $a \simeq u_0$.*

Preuve : $\mathrm{Adh}\,(A)$ est déterminé de manière unique à partir de A et de la norme sur E qui sont standard par hypothèse ; $\mathrm{Adh}\,(A)$ est donc standard ; à présent

[⇒] Si $u \in \mathrm{Adh}\,(A)$ on a $B(u,r) \cap A \neq \emptyset$ pour tout $r > 0$; il suffit d'appliquer cela pour un tel r qui soit infinitésimal.

[⇐]Par transfert (licite car A et u sont standard), il suffit de montrer que $B(u_0,r) \cap A \neq \emptyset$ dans le cas où $r > 0$ est *standard.* Mais alors $a \in B(u_0,r)$ puisque $a \simeq u_0$, et donc $\|a - u_0\| < r$, le premier étant i-petit et le second étant appréciable. Donc $a \in B(u_0,r) \cap A \neq \emptyset$.

□

4.4 Sous ensembles compacts d'un normé

Rappelons qu'un sous-ensemble standard $F \subseteq E$ est fermé si et seulement si pour tout $u \in F$, si $u_0 \in E$ est tel que $u \simeq u_0$, alors $u_0 \in F$ lui aussi. Mais, même si F est fermé, ceci n'assure pas qu'il existe effectivement un standard u_0 tel que $u_0 \simeq u$; ainsi la demi-droite fermée $[1, +\infty[$ est bien fermé, mais elle comporte des u i-grands et ceux-ci ne sont i-voisins d'aucun standard. Cette situation ne pourra pas se présenter dans le cas d'un sous-ensemble compact.

La notion de sous-ensemble compact $K \subseteq E$ est une notion standard. Nous ne la caractériserons pas en toute généralité[1], mais seulement dans le cas où K et l'espace normé sont standard. Nous en donnerons au chapitre une caractérisation générale dans le cas d'un espace de dimension finie.

Définition : Soit K un sous-ensemble standard d'un espace normé standard. On dit que K est *compact* si et seulement si pour tout $u \in K$ il existe un standard $u_0 \in K$ tel que $u \simeq u_0$.

Observons que deux standard u_0' et u_0'' i-voisins sont nécessairement égaux, puisque $\|u_0' - u_0''\| \simeq 0$ implique que $\|u_0' - u_0''\| = 0$ car $\|u_0' - u_0''\|$ est standard, et donc $u_0' - u_0'' = 0$, c'est-à-dire $u_0' = u_0''$. C'est pourquoi u_0 est appelé *la* partie-standard de u.

Comme évoqué dans l'introduction de cette section, on a :

Proposition 4.13 *Si $K \subseteq E$ est compact, alors K est fermé.*

Preuve : Les notions de compact et de fermé étant des notions standard, on peut appliquer le transfert, et supposer E et K standard. Nous allons appliquer la caractérisation 4.8 des fermés standard : soit $u \in K$ et u_0 un standard de E tel que $u \simeq u_0$. Donc u_0 est la partie standard de u, et par compacité de K, cette partie standard d'un élément de K appartient K. □

Proposition 4.14 *Pour tout fermé $F \subseteq E$, si F est contenu dans un compact K, ce fermé est lui-même compact.*

[1] Voici néanmoins la définition générale, que nous n'utiliserons pas : K est compact si et seulement si pour tout famille $\mathcal{U} := (U_i)_{i \in I}$ d'ouverts telle que $K \subseteq \bigcup_{i \in I} U_i$, il existe une sous-famille *finie* $\mathcal{U}_0 := (U_i)_{i \in I_0}$ telle que $K \subseteq \bigcup_{i \in I_0} U_i$, ce qui se prononce "de tout recouvrement ouvert de K, on peut extraire un recouvrement fini".

Preuve : Par transfert on peut supposer F et K compact. Soit $u \in F \subseteq K$; donc u admet une partie standard $u_0 \in K$, mais, comme F est fermé, $u_0 \in F$, et donc tout u admet une partie standard $u_0 \in F$. □

Voici un exemple utile de compact :

Proposition 4.15 *Toute boule-fermée $\overline{B}_\infty(v_0, r)$ de $(\mathbb{R}^d, \| \ \|_\infty)$ est compacte.*

Preuve : Par transfert, on peut supposer que d, v_0, et r sont standard. Soit $u = (u^1, \ldots, u^d) \in \overline{B}_\infty(v_0, r)$; pour tout $i = 1..d$

$$|u^i| \leq \|u\|_\infty \leq \|u - v_0\|_\infty + \|v_0\|_\infty \leq r + \|v_0\|_\infty,$$

ce dernier nombre étant standard ; donc u^i est limité, donc i-proche d'un standard u_0^i. Soit $u_0 := (u_0^1, \ldots, u_0^d)$ qui est standard car composé que d'éléments standard ; on a $\|u - u_0\|_\infty = \text{Max}\,\{u^1 - u_0^1, \ldots, u^d - u_0^d\} \simeq 0$ puisque $u^i \simeq u_0^i$ pour tout $i = 1..d$, et donc $u \simeq u_0$. □

Rappelons qu'on dit que A est *borné* s'il existe $R \geq 0$ tel que $A \subseteq B(0, R)$.

Proposition 4.16 *Tout compact est borné.*

Preuve : Par transfert, on peut supposer l'espace, E, la norme, $\| \ \|$, et le compact, K standard. Choisissons R i-grand ; la norme $\|u_0\|$ de tout standard $u_0 \in E$ est standard, et si $u \simeq u_0$, sa norme $\|u\|$ est i-voisine de $\|u_0\|$, donc limitée, d'où $u \in B(0, R)$. Comme K est compact, chaque $u \in K$ est i-voisin d'un standard, et donc $K \subseteq B(0, R)$. □

On déduit des propositions 4.13 et 4.16 le résultat suivant relatif aux compacts de $(\mathbb{R}, | \ |)$:

Corollaire 4.17 *Si $K \subseteq \mathbb{R}$ est compact, $\text{Sup}\,K$ et $\text{Inf}\,K$ existent et appartiennent à K.*

Preuve : Par transfert, on peut supposer K standard. Comme K est borné, il existe R tel que $K \subseteq\ ]-R, +R[\ \ (= B(O, R)$ dans $\mathbb{R})$. Donc K est majoré et minoré ; soient $k_- = \text{Inf}\,K$ et $k_+ = \text{Sup}\,K$. Comme ces nombre sont uniques et déterminés par K qui est standard, ils sont standard. Montrons qui'ils sont dans K. Comme $k_+ (= \text{Sup}\,K)$ est le plus petit minorant de K, pour tout $\varepsilon > 0$ on a $K \cap [k_+ - \varepsilon, k_+] \neq \emptyset$. Si on choisit ε i-petit on voit qu'il existe $u \in K$ tel que $u \simeq k_+$. Comme K est fermé et k_+ standard, ceci implique que $k_+ \in K$. On raisonne de façon analogue pour k_- □

Proposition 4.18 *Soient $K \subseteq E'$ et $f : K \longrightarrow E''$. Si K est compact et f continue, alors $f(K)$ est compact.*

En d'autres termes, *l'image continue de tout compact et compact.*

Preuve : Par transfert on peut supposer E', E'', K, et f standard. Soit $v \in f(K)$; il existe donc $u \in K$ tel que $v = f(u)$. Comme K est compact, il existe un standard $u_0 \in K$ tel que $u \simeq u_0$. Comme f est continue $v_0 := f(u_0) \simeq f(u) = v$. Comme f et u_0 sont standard, $v_0 := f(u_0)$ est standard. Il existe donc bien un standard $v_0 \in f(K)$ tel que $v \simeq v_0$; K est donc bien compact. □

Théorème 4.19 *Soit $K \subseteq E$ et $f : K \longrightarrow \mathbb{R}$. Si K est compact et f continue, il existe $u_- \in K$ et $u_+ \in K$ tels que $f(u_-) \leq f(u) \leq f(u_+)$ pour tout $u \in K$.*

On dit que u_- est *un* minimum de f et que $f(u_-)$ est *la* valeur minimale de f ; de façon analogue, u_+ est *un* maximum de f et $f(u_+)$ est *la* valeur maximale. Ce théorème s'ennonce souvent *toute fonction continue sur un compact atteind ses valeurs minimale et maximale sur ce compact.*

Preuve : Par la proposition 4.18 $f(K)$ est compact, et par la proposition 4.17, il existe k_- et k_+ dans ce compact $f(K)$ tel que $k_- = \text{Inf}\,f(K)$ et $k_+ = \text{Sup}\,f(K)$; $k_\pm \in f(K)$ signifie précisément qu'il existe $u_- \in K$ et $u_+ \in K$ tels que $f(u_-) = k_- = \text{Inf}\,f(K)$ et $f(u_+) = k_+ = \text{Sup}\,f(K)$. □

Chapitre 5

Espaces normés de dimension finie

Nous dégageons dans ce chapitre quelques résultats spécifiques aux espaces de dimension finie. Si E est un $\mathbb{R}$-espace vectoriel de dimension finie d, il existe une base $\mathcal{B}$ comportant d éléments :

$$\mathcal{B} = \{e_1, \ldots, e_d\}.$$

De façon analogue à la norme $\|\ \|_\infty$ sur $\mathbb{R}^r$, nous pouvons considérer la norme $\|\ \|_\mathcal{B}$ sur E définie par

$$\|v\|_\mathcal{B} := \mathrm{Max}\,\{|v^1|, \ldots, |v^d|\}$$

si $v = v^1 e_1 + \ldots + v^d e_d$. Si E est supposé standard, il en est de même de d (qui est uniquement déterminé par E), et il existe une base *standard* $\mathcal{B}$; comme cet ensemble est fini, ses éléments e_i sont tous standard, et la norme $\|\ \|_\mathcal{B}$ est elle aussi standard, puisqu'uniquement déterminée par E et $\mathcal{B}$.

Dans la suite E désignera un espace vectoriel normé sur $\mathbb{R}$ ou $\mathbb{C}$ de dimension finie d.

5.1 Normes equivalentes

Théorème 5.1 *Si E est de dimension finie, toutes les normes sur E sont équivalentes.*

Preuve : Il suffit de montrer que toute norme $\|\ \|$ est équivalente à la norme $\|\ \|_\mathcal{B}$. Par ailleurs, le théorème étant standard, nous pouvons supposer que E, $\|\ \|$, et $\mathcal{B}$ sont standard, ce qui implique que $\|\ \|_\mathcal{B}$ est standard. En vertu de la proposition 4.3 donnant des propriétés caractéristiques des normes standard équivalentes, il suffit que nous montrions que pour tout vecteur $v = v^1 e_1 + \ldots + v^d e_d \in E$ non nul, $\|v\| \simeq 0$ si et seulement si $\|v\|_\mathcal{B} \simeq 0$.

Supposons tout d'abord que $\|v\|_\mathcal{B} \simeq 0$; donc $\|v\|_\mathcal{B} := \mathrm{Max}\,\{|v^1|, \ldots, |v^d|\} \simeq 0$, d'où $v^1 \simeq 0, \ldots, v^d \simeq 0$. A présent, comme les e_i sont standard, leur norme $\|e_i\|$ est standard donc limitée, et comme d est standard, nous avons

$$\|v\| = \|v^1 e_1 + \ldots + v^d e_d\| \leq |v^1|\|e_1\| + \ldots + |v^d|\|e_d\| \simeq 0.$$

puisque somme de d nombres i-petits, avec d standard.

Réciproquement, supposons que $\|v\| \simeq 0$. Posons $u := v/\|v\|_\mathcal{B} =: u^1 e_1 + \ldots + u^d e_d$. Donc $\|u\|_\mathcal{B} = 1$, d'où $|u_i| \leq 1$ pour tout $i = 1 \ldots d$. Les nombres u_i sont donc limités et donc presque-standard : pour chaque $i = 1 \ldots d$, soit u_0^i l'unique standard i-voisin de u^i, et $u_0 := u_0^1 e_1 + \ldots + u_0^d e_d$, qui est standard puisque combinaison linéaire à coefficients standard d'une famille finie standard ($\mathcal{B}$) de vecteurs.

Notons que $u_0 \neq 0$, puisque, comme $\|u\|_\mathcal{B} = 1$, il existe au moins un indice i_0 tel que $|u^{i_0}| = 1$, et donc $|u_0^{i_0}| = 1$. Biensûr $\|u - u_0\| = \mathrm{Max}\,\{|u^1 - u_0^1|, \ldots, |u^d - u_0^d|\} \simeq 0$.

A présent, par le deuxième inégalité triangulaire, $|\|u\| - \|u_0\|| \leq \|u - u_0\| \simeq 0$, d'où $\|u\| \simeq \|u_0\|$, et comme ce dernier nombre est standard et non nul, nous voyons que $\|u\| =: a$ est limité. Résumons :

$$a = \|u\| = \|v/\|v\|_\mathcal{B}\| = \|v\|/\|v\|_\mathcal{B},$$

d'où $\|v\|_\mathcal{B} = a\|v\| \simeq 0$

□

5.2 Convergence des suites de Cauchy

Rappelons que la notion de suite de Cauchy de E est une notion standard, et si $(u_n)_{n\geq 0}$ est une suite standard, elle est de Cauchy si et seulement si pour tous $q \geq p$ i-grands, $u_p \simeq u_q$. Si $(u_n)_{n\geq 0}$ converge, elle est de Cauchy, puisque si $u := \lim_n u_n$, on a $u_p \simeq u \simeq u_q$, et donc $u_p \simeq u_q$. La réciproque est vraie également :

Théorème 5.2 *Si E est de dimension finie, toute suite de Cauchy de E converge dans E.*

Preuve : Soit $(u_n)_{n\geq 0}$ une suite de Cauchy de E quelconque et montrons que $(u_n)_{n\geq 0}$ est convergente. Par transfert nous pouvons supposer que E, $\|\ \|$, et $(u_n)_{n\geq 0}$ sont standard. Soit $\mathcal{B}$ une base standard de E ; par le théorème précédant 5.1, les normes $\|\ \|$ et $\|\ \|_{\mathcal{B}}$ sont équivalentes, et comme elles sont standard, elle déterminent les mêmes vecteurs i-petits. En d'autres termes, comme $(u_n)_{n\geq 0}$ une suite de Cauchy (pour $\|\ \|$), pour tout $q \geq p$ i-grands, $\|u_p - u_q\| \simeq 0$, et donc $\|u_p - u_q\|_{\mathcal{B}} \simeq 0$. Nous en déduisons

$$\exists N \ \ \forall n \geq N \ \ \|u_n - u_N\|_{\mathcal{B}} \leq 1,$$

puisque n'importe quel N i-grand convient. Mais cette propriété est standard, et ne fait intervenir que des objets standard ($(u_n)_{n\geq 0}$, $\|\ \|_{\mathcal{B}}$, $-$, $\leq$, et 1). Par transfert

$$\exists^{st} N \ \ \forall n \geq N \ \ \|u_n - u_N\|_{\mathcal{B}} \leq 1,$$

Désignons par N un tel standard tel que $\|u_n - u_N\|_{\mathcal{B}} \leq 1$ pour tout $n \geq N$. Nous pouvons à présent confectionner une (en fait *la*) limite de $(u_n)_{n\geq 0}$; voici comment :

Soit ω i-grand fixé ; comme $\omega \geq N$ nous avons $|u^i_\omega - u^i_N| \leq 1$ pour tout $i = 1 \ldots d$, et donc $|u^i_\omega| \leq 1 + |u^i_N|$ et ce dernier nombre est standard, donc u^i_ω est limité ; soit $\overline{u}^i$ l'unique standard i-voisin de u^i_ω, et $\overline{u} := \overline{u}^1 e_1 + \ldots + \overline{u}^d e_d$, qui est standard puisque combinaison linéaire à coefficients standard d'une famille finie standard ($\mathcal{B}$) de vecteurs, et biensûr $\|\overline{u} - u_\omega\|_{\mathcal{B}} = \mathrm{Max}\,\{|\overline{u}^1 - u^1_\omega|, \ldots, |\overline{u}^d - u^d_\omega|\} \simeq 0$. Mais comme les normes, standard, $\|\ \|_{\mathcal{B}}$ et $\|\ \|$ sont équivalentes, on a aussi, $\|\overline{u} - u_\omega\| \simeq 0$. A présent, pour tout n i-grand, comme ω est également i-grand et que $(u_n)_{n\geq 0}$ est de Cauchy, nous avons bien

$$\|\overline{u} - u_n\| \leq \|\overline{u} - u_\omega\| + \|u_\omega - u_n\| \simeq 0,$$

et donc $\overline{u} = \lim_n u_n$. □

De façon générale, on appelle *espace de Banach* tout espace normé dont toutes les suites de Cauchy sont convergentes. Le théorème 5.2 montre donc que tout espace normé de dimension finie est un espace de Banach.

5.3 Compacts

Théorème 5.3 *Soit K un sous-ensemble de E. Si E est de dimension finie, alors K est compact si et seulement si K est à la fois fermé et borné.*

Preuve : Par les propositions 4.13 et 4.16 et sans avoir à se restreindre aux espaces de dimension finie, nous savons que tout compact est à la fois fermé et borné.

Nous n'avons donc que la réciproque à démontrer ; supposons donc $K \subseteq E$ fermé et borné, c'est-à-dire qu'il existe $R \geq 0$ tel que $K \subseteq B(0, R)$. Par transfert nous pouvons supposer E, d, $\|\ \|$, K, et R standard. Nous devons montrer que tout $u \in K$ est i-voisin d'un standard $u_0 \in K$. Choisissons une base standard $\mathcal{B} = \{e_1, \ldots, e_d\}$ de E et considérons la norme $\|\ \|_{\mathcal{B}}$ associée. Décomposons u sur $\mathcal{B}$; on a $u = u^1 e_d + \ldots + u^d e_d$ et $\|u\|_{\mathcal{B}} = \mathrm{Max}\,\{|u^i|\}$; par le théorème 5.1 les normes $\|\ \|$ et $\|\ \|_{\mathcal{B}}$ sont équivalentes, et comme elles sont standard, par la proposition 4.3, elles détermines les mêmes limités ; comme $\|u\| \leq R$ est limité, $\|u\|_{\mathcal{B}}$ est limité, c'est-à-dire $\mathrm{Max}\,\{|u^i|\}(= \|u\|_{\mathcal{B}})$ est limité. Donc toutes les composantes u^i de u sur la base $\mathcal{B}$ sont limitées, donc chacune i-voisine d'un standard u^i_0. Il suffit de poser $u_0 = u^1_0 e_1 + \ldots + u^d_0 e_d$, qui est standard puisque combinaison linéaire à coefficients standard d'une famille finie standard ($\mathcal{B}$) de vecteurs, et biensûr $\|u_0 - u\|_{\mathcal{B}} = \mathrm{Max}\,\{|u^1_0 - u^1|, \ldots, |u^d_0 - u^d|\} \simeq 0$. Mais comme les normes $\|\ \|$ et $\|\ \|_{\mathcal{B}}$ sont équivalentes, ceci implique que $\|u_0 - u\| \simeq 0$: on a bien $u_0 \simeq u$, et comme $u \in K$ et que K est fermé, $u_0 \in K$. □

Rappelons que si K est compact et que si $f : K \longrightarrow \mathbb{R}$ est continue, alors f admet sur K (au moins) un minimum et un maximum (4.17). On a donc, dans tout espace de dimension finie E, si K est fermé et borné il existe des éléments $k_- \in K$ et $k_+ \in K$ tels que pour tout $k \in K$

$$\mathrm{Min}_{v\in K} f(v) := f(k_-) \leq f(k) \leq f(k_+) =: \mathrm{Max}_{v\in K} f(v).$$

Observons également que dans un espace normé de dimension finie, toute boule fermée $\overline{B}(u, R)$ et toute sphère $S(u, R)$ est donc compacte.

5.4 Applications linéaires continues

Théorème 5.4 *Soit $L : E \longrightarrow F$ une application linéaire. Si E est de dimension finie, L est nécessairement continue.*

Preuve : Par transfert, on peut supposer que L, E, $\|\ \|_E$, F, et $\|\ \|_F$ sont standard ; soit $\mathcal{B}$ une base standard de E, et $\|\ \|_{\mathcal{B}}$ la norme associée à cette base. Soit $u_0 = u_0^1 e_d + \ldots + u_0^d e_d$ un élément standard de E et $u = u^1 e_d + \ldots + u^d e_d \simeq u_0$; ceci signifie que $\|u - u_0\| \simeq 0$, et implique que $\|u - u_0\|_{\mathcal{B}} \simeq 0$ puisque ces normes sont équivalentes. Donc $u^1 \simeq u_0^1, \ldots, u^d \simeq u_0^d$. Nous devons montrer que $L[u] \simeq L[u_0]$, on encore, compte tenu de la linéarité de L, $L[u - u_0] = L[u] - L[u_0] \simeq 0$. Or, toujours par linéarité de L,

$$L[u - u_0] = (u^1 - u_0^1)L[e_1] + \ldots + (u^d - u_0^d)L[e_d] \text{ , donc}$$

$$\|L[u - u_0]\| \leq |u^1 - u_0^1|\|L[e_1]\| + \ldots + |u^d - u_0^d|\|L[e_d]\| \text{ ;}$$

or les $\|L[e_i]\|$ sont standard donc limités, et les $|u^i - u_0^i|$ sont i-petits. Donc $\|L[u - u_0]\|$ est majoré par la somme de d nombres i-petits, avec d standard, donc $\|L[u - u_0]\| \simeq 0$. □

Le théorème 5.3 montre en particulier que si E est de dimension finie, la boule unité centrée en 0, consituée des $u \in E$ tels que $\|u\| \leq 1$, est compacte. Si $L : E \longrightarrow F$ est linéaire, et si E est de dimension finie, la fonction $l : E \longrightarrow \mathbb{R}$ définie par $l(u) = \|L[u]\|$ est continue. On peut donc considérer le nombre $N(L) := \mathrm{Max}_{\|u\|\leq 1} l(u)$. L'application N qui a toute application linéaire L de E vers F associe le nombre $N(L)$ défini ci-dessus est une norme sur l'espace vectoriel des applications linéaires de E vers F. Le nombre $N(L)$ est de ce fait généralement noté $\|L\|$; en d'autres termes[1] $\|L\| := \mathrm{Max}_{\|u\|\leq 1} \|L[u]\|$.

Proposition 5.5 *Si E est de dimension finie, la norme sur $\mathcal{L}(E, F)$ définie par*

$$\|L\| := Max_{\|u\|\leq 1} \|L[u]\| \tag{5.1}$$

vérifie les propriétés suivantes : Pour tout $v \in E$, on a

$$\|L[v]\| \leq \|L\| \cdot \|v\| \tag{5.2}$$

et si $M : F \longrightarrow G$ est également une application linéaire et si F est de dimension finie, alors $\|M\|$ et $\|M \circ L\|$ sont également définies, et on a

$$\|M \circ L\| \leq \|M\| \cdot \|L\|. \tag{5.3}$$

Preuve : Pour $v = 0$ l'inégalité 5.2 résulte de $L[0] = 0$. Si $v \neq 0$, considérons $u := v/\|v\|$, d'où $\|u\| = 1$, et par la définition 5.1,

$$\|L[u]\| \leq \|L\| \text{ ;}$$

or

$$(\|L\| \geq)\ \ \|L[u]\| = \|L[v/\|v\|]\| = \|L[v]/\|v\|\| = \|L[v]\|/\|v\|$$

d'où 5.2, en multipliant les deux membres de cette dernière inégalité par $\|v\|$.

Soit $u \in E$ tel que $\|u\| \leq 1$. En appliquant 5.2 successivement à $(M, L[u])$ et (L, u) on a

$$\|M[L[u]]\| \leq \|M\| \cdot \|L[u]\| \leq \|M\| \cdot \|L\| \cdot \|u\| \leq \|M\| \cdot \|L\|$$

pour tout $u \in E$ tel que $\|u\| \leq 1$. Comme $\|M \circ L\| := \mathrm{Max}_{\|u\|\leq 1} \|M \circ L\ \ [u]\|$ désigne le plus petit des majorants de $\|M[L[u]]\|$ pour tout $u \in E$ tel que $\|u\| \leq 1$, on a bien $\|M \circ L\| \leq \|M\| \cdot \|L\|$, c'est-à-dire l'inégalité 5.3. □

[1] Notez que la notation $\|\ \|$ est utilisée ici dans trois sens différents : dans $\|u\|$ elle désigne la norme dans E, dans $\|L[u]\|$ elle désigne la norme dans F, et dans $\|L\|$ elle désigne la norme dans l'espace vectoriel $\mathcal{L}(E, F)$ des applications linéaires (continues) de E dans F.

Chapitre 6

Matrice jacobienne d'une fonction

Nous considérons dans ce chapitre une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $v \mapsto w = f(v)$. Plus précisément, on supposera que le domaine de définition $\mathcal{D}_f$ de la fonction f est un ouvert $\mathcal{U}$ de $\mathbb{R}^n$. Ainsi, pour $v_0 \in \mathcal{D}_f$, $f(v_0 + h)$ sera défini pour tout h i-petit relativement[1] à f et v_0. On notera indifféremment par $\| \ \|$ la norme sur $\mathbb{R}^n$ et celle sur $\mathbb{R}^p$.

6.1 Différentielle

Définition : On dit que f est *différentiable au point* $v_0 \in \mathcal{D}_f$ si et seulement si il existe une fonction linéaire continue L telle que

$$f(v_0 + h) = f(v_0) + L[h] + \|h\|\delta(h) \tag{6.1}$$

où la fonction δ déterminée par cette relation tend vers 0 avec h. La fonction linéaire L est dite *tangente à f au point v_0*.

Si f, v_0, et L sont standard, il en sera de même de la fonction δ qui est déterminée, pour $h \neq 0$, par (6.1) et ces trois objets, puisque

$$\delta(h) = \frac{f(v_0 + h) - f(v_0)}{\|h\|}.$$

Dans ce cas l'hypothèse que $\lim_{h \to 0} \delta(h) = 0$ se réduit à une notion d'ordre de grandeur :

si $h \neq 0$ est i-petit, alors $\delta(h)$ est i-petit.

L'intérêt de cette approche est qu'elle est statique (on ne s'intéresse qu'à **un** $h \simeq 0$), alors que l'approche par $\lim_{h \to 0} \delta(h)$ évoque l'évolution de $\delta(h)$ lorsque $\|h\|$ devient de plus en plus petit : c'est une idée dynamique, plus difficile à manier dès que d'autres grandeurs doivent également évoluer. Le prix à payer pour pouvoir sans perte de généralité fixer ainsi h à une valeur infinitésimale est à peu près toujours la même : montrer au préalable qu'il suffit de ne considérer le problème que dans le cas où les "paramètres libres" (ici f, v_0, et L) sont tous standard ; ceci est rendu possible très souvent par application d'une des deux versions du principe de transfert. Si tel est le cas, cette phase préparatoire se déroule toujours à la manière de l'incantation "par transfert, on peut supposer que ... sont standard". Une fois que cette démarche est comprise dans un exemple, elle devient naturelle dans toute autre situation, totalement indépendamment des situations géométriques envisagées.

Proposition 6.1 *Il existe au plus une fonction linéaire tangente en un point v_0 à une fonction f. Si une telle fonction linéaire existe, on l'appelle la* différentielle *de f au point v_0, et on la note $D_{v_0}f$. En d'autres termes, on écrira $D_{v_0}f[h]$ pour $L[h]$ et on aura $f(v_0 + h) = f(v_0) + D_{v_0}f[h] + \|h\|\delta(h)$.*

Preuve : Par transfert, on peut supposer f et v_0 standard. Supposons qu'il existe deux applications linéaires L_1 et L_2 satisfaisant (6.1). Par transfert, il existe alors aussi deux applications linénaires *standard* L_1 et L_2 satisfaisant (6.1). On a

$$(L_1 - L_2)[h] = L_1[h] - L_2[h] = \|h\|\delta_1(h) - \|h\|\delta_2(h),$$

[1] c'est-à-dire si f et v_0 sont supposés standard

Figure 6.1: Graphe de la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}^2$, $v \mapsto w = f(v) = (\cos(v), \sin(v))$ et son image par une loupe

et donc

$$(L_1 - L_2)\left[\frac{h}{||h||}\right] = \delta_1(h) - \delta_2(h) =: \delta(h).$$

Lemme 6.2 *Soit $\mathcal{E} \neq \emptyset$ un ensemble (interne) de nombre tous i-petits, alors Sup$\mathcal{E} \simeq 0 \simeq$ Inf$\mathcal{E}$.*

Preuve : [du lemme] Soit $\varepsilon_0 \in \mathcal{E}$; $\operatorname{Sup}\mathcal{E} \geq \varepsilon_0$ donc $S := \operatorname{Sup}\mathcal{E}$ est soit i-petit soit appréciable positif. Si S était appréciable positif, il en serait de même de $S/2$ et donc, pour tout $\varepsilon \in \mathcal{E}$, comme ε est i-petit, on aurait $\varepsilon < S/2 < S$ d'où $S = \operatorname{Sup}\mathcal{E} \leq S/2 < S$ ce qui est absurde. Le preuve de $\inf \mathcal{E} \simeq 0$ est analogue, d'où le lemme. □

Soient $u \in E$ quelconque tel que $||u|| = 1$, et $h_u = h := \alpha u$, avec $\alpha \simeq 0$ non nul. On a $h/||h|| = \alpha u/||\alpha u|| = \alpha u/\alpha||u|| = u$, donc, en appliquant le lemme 6.2,

$$||L_1 - L_2|| = \operatorname{Sup}_{||u||=1}||(L_1 - L_2)[u]|| = \operatorname{Sup}_{||u||=1}||(L_1 - L_2)[h_u/||h_u||]||\operatorname{Sup}_{||u||=1}||\delta(h_u)|| \simeq 0.$$

Mais $L_1 - L_2$ est standard, donc $||L_1 - L_2||$ est standard ; donc $||L_1 - L_2|| = 0$, puisque 0 est le seul standard i-petit. Finalement $L_1 - L_2 = 0$, c'est-à-dire $L_1 = L_2$. □

Remarques :

Loupe Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction standard, et $v_0 \in \mathbb{R}$ un point standard. En première année, vous avez vu que f est dérivable au point v_0 si et seulement si l'image sous la loupe centrée au point $(v_0, f(v_0))$ de son graphe est une droite de pente p, où p est la *dérivée* au point considéré ; la situation ici est parfaitement identique, moyennant le fait de généraliser le nom "droite" par "graphe de fonction linéaire L". En effet, rappelons qu'on entend par loupe centrée au point standard[2] $(v_0, f(v_0))$, le fait de considérer les valeurs limitées des variables V et W liées à v et w par les relations

$$v = v_0 + \varepsilon V \text{ et } w = f(v_0) + \varepsilon W$$

pour $\varepsilon > 0$ infinitésimal quelconque. Sous la loupe on obtient le graphe de la fonction F telle que $W = F(V)$, c'est-à-dire, d'après $w = f(v)$, on a

$$f(v_0) + \varepsilon W = w_0 + \varepsilon W = w = f(v) = f(v_0 + \varepsilon V),$$

d'où

$$F(V) = W = \frac{1}{\varepsilon}(f(v_0 + \varepsilon V) - f(v_0))$$

Choisissons $h = \varepsilon V$; En utilisant la linéarité de L nous voyons que la relation (6.1) équivaut à

$$\delta(h) = \frac{f(v_0 + h) - f(v_0)}{||h||} - L\left[\frac{h}{||h||}\right] = \frac{f(v_0 + \varepsilon V) - f(v_0)}{||\varepsilon V||} - L\left[\frac{\varepsilon V}{||\varepsilon V||}\right] = \frac{1}{||V||}(F(V) - L[V]).$$

Nous voyons donc que pour V appréciable, $\delta(h) \simeq 0$ équivaut bien a $F(V) \simeq L[V]$.

normes équivalentes Comme nous venons de le voir, la relation (6.1) détermine la valeur de $\delta(h)$ et le fait $L = D_{v_0}f$ équivaut à ce que $\delta(h)$ soit i-petit dès que h est i-petit. A présent, si on remplace les normes sur $\mathbb{R}^n$ et $\mathbb{R}^p$ par des normes équivalentes, on ne change pas les i-petits : on ne change donc pas non plus la notion de différentiabilité. En particulier, en dimension finie, on pourra choisir la norme qui nous arrange pour la fonction considérée.

[2] Si on supposait ceci aussi pour les points *presque*-standard, c'est-à-dire i-voisins d'un standard du domaine de f, on déboucherait directement sur la notion de fonction de classe $\mathcal{C}^1$.

6.2 Différentiabilité, matrice jacobienne

6.2.1 Quelques règles de différentiabilité

Proposition 6.3 *Soit g également définie sur un ouvert de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^p$. Soient λ et μ des nombres. Si f et g sont différentiables au point v, il en est de même de la fonction $\lambda f + \mu g$, et $D_v(\lambda f + \mu g) = \lambda D_v f + \mu D_v g$.*

Preuve : Comme f et g sont différentiables au point v_0, il existe des fonctions $h \mapsto \delta(h)$ et $h \mapsto \varepsilon(h)$ avec $\lim_{h\to 0} \delta(h) = 0 = \lim_{h\to 0} \varepsilon(h)$ telles que

$$f(v_0 + h) = f(v_0) + D_{v_0} f[h] + \|h\|\delta(h)$$
$$g(v_0 + h) = g(v_0) + D_{v_0} g[h] + \|h\|\varepsilon(h)$$

d'où

$$\begin{aligned}
(\lambda f + \mu g)(v_0 + h) &:= \lambda f(v_0 + h) + \mu g(v_0 + h) \\
&= \lambda(f(v_0) + D_{v_0} f[h] + \|h\|\delta(h)) + \mu(g(v_0) + D_{v_0} g[h] + \|h\|\varepsilon(h)) \\
&= \lambda f(v_0) + \mu g(v_0) + \lambda D_{v_0} f[h] + \mu D_{v_0} g[h] + \|h\|\lambda\delta(h) + \|h\|\mu\varepsilon(h) \\
&=: (\lambda f + \mu g)(v_0) + (\lambda D_{v_0} f + \mu D_{v_0} g)[h] + \|h\|(\lambda\delta + \mu\varepsilon)(h)
\end{aligned}$$

Comme $\lim_{h\to 0}(\lambda\delta + \mu\varepsilon)(h) = \lim_{h\to 0}(\lambda\delta(h) + \mu\varepsilon(h)) = \lambda \lim_{h\to 0} \delta(h) + \mu \lim_{h\to 0} \varepsilon(h) = 0$, ceci montre que $\lambda f + \mu g$ est différentiable au point v_0 avec

$$D_{v_0}(\lambda f + \mu g) = \lambda D_{v_0} f + \mu D_{v_0} g.$$

□

Proposition 6.4

1. *Toute fonction constante $f(v) = c$ est différentiable en chaque point, et sa différentielle en chaque $v \in \mathbb{R}^n$ est la fonction nulle.*
2. *Toute fonction linéaire (continue) $L : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ est différentiable en chaque point $v \in \mathbb{R}^n$, et $D_v L[h] = L[h]$; $D_v L$ est donc indépendante du choix de $v \in \mathbb{R}^n$.*
3. *Si $n = 1$, f est dérivable au point v de dérivée $f'(v)$ si et seulement si f est différentiable au point v, et $D_v f[h] = hf'(v)$.*
4. *Posons $E_1 = \mathbb{R}^{n_1}$ et $E_2 = \mathbb{R}^{n-2}$; donc $E_1 \times E_2 = \mathbb{R}^{n_1+n_2}$. Si $b : E_1 \times E_2 \longrightarrow F$ est bilinéaire continue, elle est différentiable en tout point $(v_1, v_2) \in E_1 \times E_2$, et $D_{(v_1,v_2)} b[(h_1, h_2)] = b[h_1, v_2] + b[v_1, h_2]$.*

Preuve :

1. On a $f(v + h) = c = f(v) = f(v) + 0[h] + \|h\|0$, d'où $D_v f[h] = 0$.
2. On a $L[v + h] = L[v] + L[h] = L[v] + L[h] + \|h\|0$, d'où $D_v L[h] = L[h]$.
3. Supposons f dérivable au point v, de dérivée $f'(v)$, c'est-à-dire

   $$f'(v) = \lim_{h\to 0} \frac{f(v + h) - f(v)}{h} \in \mathbb{R}^p$$

   (pour $h \neq 0$, $\frac{1}{h}$ a bien un sens puisque $h \in \mathbb{R}$). Posons $\hat{\delta}(h) := \frac{f(v+h)-f(v)}{h} - f'(v)$, d'où $\lim_{h\to 0} \hat{\delta}(h) = 0$, et $\delta(h) := \frac{h}{|h|}\hat{\delta}(h)$, d'où $\lim_{h\to 0} \delta(h) = \pm \lim_{h\to 0} \hat{\delta}(h) = 0$. On a donc

   $$f(v + h) = f(v) + hf'(v) + h\hat{\delta}(h) = f(v) + hf'(v) + |h|\delta(h)$$

   ce qui montre que $D_v f[h] = hf'(v)$.
   Réciproquement, si $f(v + h) = f(v) + D_v f[h] + |h|\delta(h)$, avec $\lim_{h\to 0} \delta(h) = 0$, alors

   $$\begin{aligned}
   \lim_{h\to 0} \frac{f(v + h) - f(v)}{h} &= \lim_{h\to 0}(D_v f\left[\frac{h}{h}\right] + \frac{|h|}{h}\delta(h)) \\
   &= D_v f[1] + \lim_{h\to 0} \pm\delta(h) = D_v f[1]
   \end{aligned}$$

   ce qui montre que $f'(v) = D_v f[1]$, et donc $hf'(v) = hD_v f[1] = D_v f[h]$.

4. Par transfert, on peut, sans perte de généralité, supposer n_1, n_2, p, $\| \ \|$, b, v_1 et v_2 standard. En appliquant la linéarité de b par rapport à chacun de ses deux arguments, on a :

$$b(v_1+h_1, v_2+h_2) = b(v_1, v_2+h_2) + b(h_1, v_2+h_2) = b(v_1,v_2) + b(v_1,h_2) + b(h_1,v_2) + b(h_1,h_2)$$

Nous avons $\|h\| = \|(h_1,h_2)\| = \text{Max}\{\|h_1\|, \|h_2\|\}$ si nous choisissons la norme $\| \ \|_\infty$ sur $\mathbb{R}^{n_1}$, $\mathbb{R}^{n_2}$, et $\mathbb{R}^{n_1+n_2}$. Observons que

$$\left\| \frac{h_1}{\sqrt{\|h\|}} \right\| = \frac{\|h_1\|}{\sqrt{\|h\|}} = \sqrt{\frac{\|h_1\|^2}{\|h\|}} \leq \sqrt{\frac{\|h\|^2}{\|h\|}} = \sqrt{\|h\|} \simeq 0 \text{ si } h \simeq 0,$$

et de même $\left\| \frac{h_2}{\sqrt{\|h\|}} \right\| \simeq 0$ si $h \simeq 0$. Définissons $\delta(h)$ par $b(h_1,h_2) = \|h\|\delta(h)$. On a $\delta(h) = \frac{1}{\sqrt{\|h\|}\sqrt{\|h\|}} b(h_1,h_2) = b\left(\frac{h_1}{\sqrt{\|h\|}}, \frac{h_2}{\sqrt{\|h\|}}\right) \simeq 0$ puisque b est continue, que $b(0,0)=0$ et que $\left(\frac{h_1}{\sqrt{\|h\|}}, \frac{h_2}{\sqrt{\|h\|}}\right) \simeq 0$. D'où $D_{(v_1,v_2)}b[(h_1,h_2)] = b[h_1,v_2] + b[v_1,h_2]$.

□

Théorème 6.5 *Soit f différentiable au point v et soit g différentiable au point $w = f(v)$; alors $g \circ f$ est différentiable au point v, et $D_v(g \circ f) = D_w g \circ D_v f$. En particulier*

$$Jac\,(g \circ f)(v) = Jac\,g(f(v)) \times Jac\,f(v).$$

Preuve : On peut, sans perte de généralité, supposer f, g et v standard ; donc $w = f(v)$, $D_v f$ et $D_w g$ sont standard, puisqu'ils sont déterminés uniquement par f, g et w. Par hypothèse, pour tout $h \simeq 0$ et $k \simeq 0$, il existe $\delta \simeq 0$ et $\varepsilon \simeq 0$ tels que :

$$f(v+h) = f(v) + D_v f[h] + \|h\|\delta$$

$$g(w+k) = g(w) + D_w g[k] + \|k\|\varepsilon$$

Choisissons $k = D_v f[h] + \|h\|\delta$. On a $k \simeq 0$ puisque $D_v f[h] \simeq 0$. Nous avons

$$\begin{aligned} g \circ f(v+h) &= g(f(v) + D_v f[h] + \|h\|\delta) \\ &= g(f(v)) + D_w g[D_v f[h] + \|h\|\delta] + \|k\|\varepsilon \\ &= g \circ f(v) + D_w g \circ D_v f[h] + D_w g[\|h\|\delta] + \|k\|\varepsilon. \end{aligned}$$

Nous devons vérifier que $\frac{1}{\|h\|} D_w g[\|h\|\delta] + \|k\|\varepsilon \simeq 0$. Or $\|k\| = \|D_v f\|[h] + \|h\|\delta\| \leq D_v f\| \cdot \|h\| + \|h\| \cdot \|\delta\| = \|h\|l$ avec $l := \|D_v f\| + \|\delta\|$ donc limité puisque $\|D_w f\|$ est standard et $\delta \simeq 0$.

On a donc

$$\begin{aligned} \left\| \frac{1}{\|h\|}(D_w g[\|h\|\delta] + \|k\|\varepsilon) \right\| &= \frac{1}{\|h\|}\|D_w g[\|h\|\delta] + \|k\|\varepsilon\| \\ &\leq \frac{1}{\|h\|}(\|D_w g[\|h\|\delta]\| + \|h\|l\|\varepsilon\|) \\ &\leq \frac{1}{\|h\|}(\|h\|\|D_w g[\delta]\| + \|h\|l\|\varepsilon\|) \\ &\leq \|D_w g[\delta]\| + l\|\varepsilon\| \end{aligned}$$

Or cette dernière quantité est i-petite puisque $\delta \simeq 0$ et donc $D_w g[\delta] \simeq 0$, $\varepsilon \simeq 0$ et l est limité. □

6.2.2 Fonctions-composantes, dérivées partielles, jacobienne

Comme $f(v) \in \mathbb{R}^p$, $f(v)$ est un vecteur $(f_1(v), .., f_p(v))$; cette écriture revient à poser $f_i = pr_i \circ f$ où pr_i désigne la projection de $\mathbb{R}^p$ sur la i-ème composante, c'est-à-dire $pr_i[w] = pr_i[(w_1, .., w_p)] := w_i$. Comme un ensemble fini est standard si et seulement si tous ses éléments sont standard, la fonction f est standard si et seulement si toutes ses composantes $f_1, f_2, \ldots, f_p$ le sont. Rappelons que nous savons déjà que la fonction f est continue si et seulement si f_i est continue, pour tout $i = 1..p$.

Proposition 6.6 *Avec les notations qui précèdent, la fonction f est différentiable au point v si et seulement si toutes les fonctions f_i le sont.*

Preuve : Par transfert, on peut supposer f et v standard, ce qui implique que les f_i le sont aussi. Comme $f_i = pr_i \circ f$, la différentiabilité de f implique celle des f_i puisque pr_i est linéaire donc différentiable (proposition 6.4) et que la composée de fonctions différentiables est différentiable (théorème 6.5). Réciproquement, supposons les f_i différentiables et posons $L[h] = (D_v f_1[h] + \ldots + D_v f_p[h])$. Comme par hypothèse

$$f_i(v+h) = f_i(v) + D_v f_i[h] + \|h\|\delta_i(h)$$

avec $\delta_i(h) \simeq 0$ dès que $h \simeq 0$, on a donc

$$f(v+h) = f(v) + L[h] + \|h\|\delta(h)$$

avec $\delta(h) = (\delta_1(h), \ldots, \delta_p(h))$. Si $h \simeq 0$, $\|\delta(h)\|_\infty = \mathrm{Max}\{|\delta_1(h)|, \ldots, |\delta_p(h)|\} \simeq 0$. Comme δ est standard puisque tel est le cas de chacune de ses composantes, ceci implique que $\lim_{h\to 0} \delta(h) = 0$.

□

La notion de dérivée partielle de $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ a été introduite en première année, et il a été démontré que si f_i est différentiable au point v, alors toutes les dérivées partielles $\partial f_i/\partial v_j(v)$ existent ($j = 1..n$), et la matrice de la différentielle $D_v f_i$ dans la base canonique de $\mathbb{R}^n$ est égale à $\mathrm{Grad}\, f := (\partial f_i/\partial v_1(v) \quad .. \quad \partial f_i/\partial v_n(v))$; ce vecteur ligne est appelé *vecteur gradient* de f_i au point v.

Corollaire 6.7 *Si f est différentiable au point v, alors la matrice de $D_v f$ dans la base canonique est la matrice*

$$Jac\, f(v) := \begin{pmatrix} \frac{\partial f_1}{\partial v_1}(v) & \cdots & \frac{\partial f_1}{\partial v_j}(v) & \cdots & \frac{\partial f_1}{\partial v_n}(v) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_i}{\partial v_1}(v) & \cdots & \frac{\partial f_i}{\partial v_j}(v) & \cdots & \frac{\partial f_i}{\partial v_n}(v) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_p}{\partial v_1}(v) & \cdots & \frac{\partial f_p}{\partial v_j}(v) & \cdots & \frac{\partial f_p}{\partial v_n}(v) \end{pmatrix}$$

appelée matrice Jacobienne[3] *de f au point v.*

Théorème 6.8 *Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^n$. Si pour chaque $i = 1..p$ et chaque $j = 1..n$ la fonction $v \mapsto \frac{\partial f_i}{\partial x_j}(v)$ est continue pour $v \in \mathcal{U}$, alors f est différentiable en tout point $v \in \mathcal{U}$, et $D_v f[V] = Jac\, f(v) \cdot V$.*

Preuve : Ce théorème vous a été démontré en première année dans le cas $p = 1$. Pour le cas général, la proposition 6.6 montre que f est différentiable *ssi* chaque fonction f_i l'est. Or les fonctions f_i sont à valeurs dans $\mathbb{R}$ et on peut donc leur appliquer le théorème connu du cas $p = 1$.

□

Définition : On dit que f est continûment différentiable, ou encore f est $\mathcal{C}^1$, si et seulement si les hypothèses du théorème 6.8 sont satisfaites.On désigne par $\mathcal{C}^1(\mathcal{U}, \mathbb{R}^p)$ l'ensemble des fonctions qui sont continûment différentiables sur $\mathcal{U}$ et à valeurs dans $\mathbb{R}^p$.

6.2.3 Le théorème des accroissements finis

Rappelons que pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$ suffisamment régulière le théorème des accroissements finis assure l'existence d'un $c \in]a, a+H[$ tel que $f(a+H) - f(a) (= Hf'(c)) = D_c f[H]$. Sous cette forme, le théorème ne se généralise pas à une fonction à valeurs dans $\mathbb{R}^p$ de dimension plus grande. Ceci peut se voir sur l'exemple de la fonction $f : \mathbb{R} \to \mathbb{R}^2$ définie par $f(t) := (\cos t, \sin t)$, pour laquelle on a $\mathrm{Jac}\, f(t) = (-\sin t, \cos t)$, et donc, pour $H \neq 0$, l'"accroissement linéaire" $\mathrm{Jac}\, f(c) \times H$ est non nul pour n'importe quel c. En revanche, pour $H = 2\pi$, l'accroissement "exact" $f(a+2\pi) - f(a)$ est toujours nul.

Il est toutefois possible de *majorer* l'accroissement, moyennant une majoration de la différentielle. Commençons par une majoration dans le cas particulier d'une fonction $\varphi : \mathbb{R} \longrightarrow \mathbb{R}^p$.

Proposition 6.9 *Soit $\varphi : \mathbb{R} \supseteq [a,b] \longrightarrow \mathbb{R}^p$, continue sur $[a,b]$ et dérivable sur $]a,b[$. Si $\|\varphi'(t)\|_\infty \leq M$ pour tout $t \in]a,b[$, alors $\|\varphi(b) - \varphi(a)\|_\infty \leq |b-a|M$.*

[3] Carl Gustav Jacob Jacobi (1804-1851)

Preuve : Posons $\varphi(t) = (\varphi_1(t), \dots, \varphi_p(t))$; φ est continue si et seulement si les φ_i le sont, et dérivable si et seulement si les φ_i le sont. Le théorème des accroissements finis pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$ assure l'existence, pour chaque[4] $i = 1 \dots p$ d'un $c_i \in]a, b[$ tel que

$$\varphi_i(b) - \varphi_i(a) = (b-a)\varphi_i'(c_i).$$

Finalement,

$$\|\varphi(b) - \varphi(a)\|_\infty = \|(b-a)(\varphi_1'(c_1), \dots, \varphi_p'(c_p))\|_\infty = |b-a| \|(\varphi_1'(c_1), \dots, \varphi_p'(c_p))\|_\infty \leq |b-a|M$$

□

Théorème 6.10 *On munit $\mathbb{R}^p$ de la norme $\| \ \|_\infty$. Soit $f : \mathbb{R}^n \supseteq [v, w] \longrightarrow \mathbb{R}^p$, continue sur $[v, w]$ et différentiable sur $]v, w[$, où*

$$[v, w] := \{v + t(w-v) \in \mathbb{R}^n \ \mid \ t \in [0, 1]\} \ \text{et} \ ]v, w[:= \{v + t(w-v) \in \mathbb{R}^n \ \mid \ t \in]0, 1[\}.$$

Si $\|D_u f\| \leq M$ pour tout $u \in]v, w[$, alors

$$\|f(w) - f(v)\|_\infty \leq \|w - v\| M.$$

Preuve : Soit $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^n$, $\gamma(t) = v + t(w-v)$, et $\varphi : [0, 1] \longrightarrow \mathbb{R}^p$, $\varphi = f \circ \gamma$. Par le théorème de diférentiabilité des fonctions composée 6.5, $\varphi = f \circ \gamma$, qui est continue sur $[0, 1]$, est aussi différentiable sur $]0, 1[$, et

$$\varphi'(t) = D_t\varphi[1] = D_{\gamma(t)}f \circ D_t\gamma[1] = D_u f[\gamma'(t)] = D_u f[w - v],$$

avec $u := \gamma(t)$. Pour chaque $t \in]0, 1[$ on a donc

$$\|\varphi(t)\|_\infty = \|D_u f[w-v]\|_\infty = \|w - v\| \|D_u f\left[\frac{w-v}{\|w-v\|_\infty}\right]\|_\infty =: \|w-v\| \|D_u f[U]\|_\infty,$$

avec $U := (w - v)/\|w - v\|$ et donc $\|U\| = 1$, d'où

$$\|D_u f[U]\|_\infty \leq \operatorname{Sup}_{\|U\|} \|D_u f[U]\|_\infty = \|D_u f\| \leq M.$$

Finalement, pour tout $t \in]0, 1[$,

$$\|\varphi'(t)\|_\infty = \|D_u f[w - v]\|_\infty \leq \|w - v\| M.$$

A présent, comme $\varphi(0) = f(\gamma(0)) = f(v)$ et $\varphi(1) = f(\gamma(1)) = f(w)$, on a

$$\|f(w) - f(v)\|_\infty = \|\varphi(1) - \varphi(0)\|_\infty \leq (1 - 0)\|w - v\| M = \|w - v\| M.$$

□

6.3 Difféomorphismes

Définition : Soient $\mathcal{V}$ et $\mathcal{W}$ deux ouverts de $\mathbb{R}^n$. On appelle $\mathcal{C}^1$-difféomorphisme de $\mathcal{V}$ sur $\mathcal{W}$ toute application $f \in \mathcal{C}^1(\mathcal{V}, \mathbb{R}^n)$, bijective de $\mathcal{V}$ sur $\mathcal{W}$, telle que pour tout $v \in \mathcal{V}$, $\operatorname{Jac} f(v)$ soit inversible.

Remarque : On pourrait penser que le fait que f soit bijective implique que $\operatorname{Jac} f(v)$ le soit automatiquement. L'exemple $f(v) := v^3$, où $n = 1$ et $\mathcal{V} = \mathbb{R} = \mathcal{W}$, montre qu'il n'en est rien : f est bien bijective, avec $f^{-1}(w) = \sqrt[3]{w}$, et $\operatorname{Jac} f(v) = (3v^2)$ qui est nulle si $v = 0$. De fait, s'il est bien clair que l'image sous la loupe d'une application bijective est bijective, il ne faut pas oublier que la différentielle n'est qu'une approximation de l'application vue sous la loupe. Ainsi, dans notre exemple, sous la loupe centrée au point $(0, f(0))$, l'application $w = f(v)$ devient $W = F(V) := \varepsilon^2 V$, et $\varepsilon^2 V \simeq 0$ tant que V est limité.

A l'inverse, si $\operatorname{Jac} f(v_0)$ est inversible, alors f est un $\mathcal{C}^1$ difféomorphisme au voisinage du point v_0 : c'est précisément ce qu'assure le théorème d'inversion locale que nous examinons à présent.

[4] à noter que la valeur de c_i dépend de la composante φ_i considérée ; il n'y a pas de raison qu'en général les c_i soient égaux.

6.3.1 Le théorème d'inversion locale

Nous admettrons le résultat suivant :

Théorème 6.11 (d'inversion locale) *Ici $p = n$. Soient $\mathcal{U}$ un ouvert de $\mathbb{R}^n$, $v \in \mathcal{U}$ et $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^n)$. Si la matrice $Jac\, f(v)$ est inversible alors il existe un ouvert $\mathcal{V}$, $v \in \mathcal{V} \subseteq \mathcal{U} \subseteq \mathbb{R}^n$, et un ouvert $\mathcal{W}$, $f(v) \in \mathcal{W} \subseteq f(\mathcal{U}) \subseteq \mathbb{R}^n$ tels que f soit une bijection de $\mathcal{V}$ sur $\mathcal{W}$. Soit $g : \mathcal{W} \to \mathcal{V}$ la bijection réciproque ; on a $g \in \mathcal{C}^1(\mathcal{W}, \mathbb{R}^n)$ et $Jac\, g(f(v)) = [Jac\, f(v)]^{-1}$ pour tout $v \in \mathcal{V}$.*

Preuve : Nous ne démontrons pas ce théorème qui est un point culminant du cours de Calcul Différentiel de Licence, et nous nous bornerons à quelques observations simples.

Notons tout d'abord que, même si sa preuve est assez fine, ce résultat est naturel puisque, dans le cas où f et v sont standard, l'hypothèse implique que, vu sous toute loupe centrée en v, son graphe est i-voisin du graphe d'une fonction linéaire inversible (de matrice $\mathrm{Jac}\, f(v)$) : il convient toutefois de s'assurer que ce graphe, “vu depuis les $w(= f(v))$”, ne présente pas de trou (non surjectivité) ni de pli (non injectivité) infinitésimaux ! En revanche, le fait que $\mathrm{Jac}\, g(f(v)) = [\mathrm{Jac}\, f(v)]^{-1}$ découle directement de la règle de dérivation des fonction composées (théorème 6.5) dès que l'on est assuré de l'existence de g ; en effet $g \circ f = \mathrm{Id}_{\mathbb{R}^n} = f \circ g$ et donc

$$\mathrm{Jac}\, g(f(v)).\mathrm{Jac}\, f(v) = \mathbb{I}_{n\times n} = \mathrm{Jac}\, f(v).\mathrm{Jac}\, g(f(v))$$

□

Exemple : La fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $(x, y) \mapsto (x^2 - y^2, 2xy)$ est de classe $\mathcal{C}^1$. Sa jacobienne

$$\mathrm{Jac}\, f(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

est inversible pour tout $(x, y) \neq (0, 0)$, puisque son déterminant[5] $\mathrm{jac}\, f(x, y) := \det \mathrm{Jac}\, f(x, y)$ est égal à $4(x^2 + y^2)$ et est non nul en dehors de $(0, 0)$. Le théorème d'inversion locale s'applique donc en tout point $(x, y) \neq (0, 0)$. Pourtant f n'est pas une bijection de $\mathbb{R}^2_* := \mathbb{R}^2 - \{(0, 0)\}$ sur $f(\mathbb{R}^2_*)$, puisque (x, y) et $(-x, -y)$ ont même image. Il convient donc de bien veiller au caractère local de ce théorème.

6.3.2 Changements de variables

La notion de $\mathcal{C}^1$-difféomorphisme peut paraître un peu abstraite au premier abord ; elle devient beaucoup plus concrète lorsque qu'elle décrit un *changement de variable*. Bornons-nous au cas n=2 ; notons plutôt $\mathcal{X}$ pour l'ouvert-source (plus haut nous écrivions $\mathcal{V}$) et désignons par (x, y) la variable de $\mathcal{X}$ (plus haut nous écrivions v au lieu de (x, y)). Notons toujours $\mathcal{W}$ pour l'ouvert-but, mais notons (u, v) la variable de $\mathcal{W}$ (nous écrivions w au lieu de (u, v)). Nous concevrons notre $\mathcal{C}^1$-difféomorphisme comme un changement de variable et nous l'appelons de ce fait $c : \mathcal{X} \longrightarrow \mathcal{W}$ (au lieu de $f : \mathcal{V} \longrightarrow \mathcal{W}$). Nous avons donc $c(x, y) = (u, v)$, ou encore

$$\left[\begin{array}{ccc} u & = & \overline{u}(x, y) \\ v & = & \overline{v}(x, y) \end{array} \right. \tag{6.2}$$

où $c(x, y) = (\overline{u}(x, y), \overline{v}(x, y))$, les fonctions $\overline{u}$ et $\overline{v}$ étant $\mathcal{C}^1$.

Exemple : *Coordonnées polaires :* en posant $u = x \cos y$ et $v = x \sin y$, x et y sont simplement le module et l'angle polaire du point (u, v) ; $c(x, y) := (x \cos y, x \sin y)$ est un $\mathcal{C}^1$-difféomorphisme sur tout ouvert $\mathcal{X} := \mathbb{R}^+_* \times]y_-, y_+[$ dès que $0 < y_+ - y_- < 2\pi$.

Voici une propriété essentielle des changements de variables standard.

6.3.3 Changements de variables standard

Pour des raisons de clarté d'exposition nous nous restraignons toujours au cas $n = 2$, mais ce qui suit se généralise facilement au cas n quelconque (standard). Soit $\mathcal{W}$ un ouvert standard de $\mathbb{R}^2$.

Soit $\mathcal{C}$ le carré de côté 1 centré en $(0, 0)$, à côtés parallèles aux axes :

$$\mathcal{C} := \{xe_1 + ye_2 \mid x \in \left[\frac{-1}{2}, \frac{+1}{2}\right], y \in \left[\frac{-1}{2}, \frac{+1}{2}\right]\}.$$

[5] $\mathrm{jac}\, f = \det \mathrm{Jac}\, f$: ce sont ces déterminants fonctionnels que Jacobi a abondamment étudiés

Soit L une application linéaire inversible, et $\mathcal{K} := L[\mathcal{C}]$, donc

$$\mathcal{K} := \{xL[e_1] + yL[e_2] \mid x \in \left[\frac{-1}{2}, \frac{+1}{2}\right], y \in \left[\frac{-1}{2}, \frac{+1}{2}\right]\}.$$

c'est-à-dire un parallélogramme (plein), centré en $(0,0)$, à coté parallèles à $L[e_1]$ et $L[e_2]$ et de même longueur que ces vecteurs, qui sont indépendants puisque L est inversible. Son aire est égale à $|\det L|$.

Soit (u', v') un point de $\mathcal{W}$ et $\varepsilon > 0$ un nombre i-petit. Notons

$$\kappa := (u', v') + \varepsilon\mathcal{K} := \{(u', v') + \varepsilon(U, V) \mid (U, V) \in \mathcal{K}\}.$$

Si L est standard ou i-voisine d'une application standard, alors κ est un parallélogramme infinitésimal centré au point (u', v'). Son aire est égale à $\varepsilon^2 \operatorname{Aire} \mathcal{K} = \varepsilon^2 |\det L|$.

Définition : Nous dirons que $\tilde{A} \subseteq \mathbb{R}^2$ est *asymptotiquement proche* de $A \subseteq \mathbb{R}^2$ s'il existe un point $w \in A$ et des infinitésimaux $\delta_- > 0$ et $\delta_+ > 0$ tels que

$$A_- \subseteq \tilde{A} \subseteq A_+,$$

où A_- et A_+ désignent les homothétiques de A par les homothéties de centre w et de rapports respectifs $1 - \delta_-$ et $1 + \delta_+$.

Il est facile de voir que si l'aire de $\tilde{A}$ et de A sont définies, alors

$$\frac{\operatorname{Aire} \tilde{A}}{\operatorname{Aire} A} \simeq 1.$$

Proposition 6.12 *Soit $c : \mathcal{X} \longrightarrow \mathcal{W}$ un changement de variable standard, et $(x_0, y_0) \in \mathcal{X}$ un point standard. Soit $(x', y') \simeq (x_0, y_0)$ et $\varepsilon > 0$ i-petit. Soit*

$$\mathcal{P} := (x', y') + \varepsilon\mathcal{C}.$$

Alors $c(\mathcal{P})$ est asymptotiquement proche du parallélogramme $\kappa := c(x', y') + \varepsilon D_{(x',y')} f[\mathcal{C}]$ dont l'aire est égale à $|\operatorname{jac} f(x', y')|$.

Preuve : ; posons $L := D_{(x',y')} f$ □

Chapitre 7

Intégrales doubles : Changements de variables

Définition : Soient $\mathcal{U}$ et $\mathcal{V}$ des ouverts de $\mathbb{R}^n$, et $c : \mathcal{U} \longrightarrow \mathcal{V}$ une application. On dit c est un $\mathcal{C}^1$-*difféomorphisme* de $\mathcal{U}$ sur $\mathcal{V}$ si et seulement si

1. c est une bijection de $\mathcal{U}$ sur $\mathcal{V}$;
2. c et c^{-1} sont des applications continûment différentiables.

Plutôt que $\mathcal{C}^1$-difféomorphisme, nous dirons ici *changement de variable*. Par le théorème de dérivation des fonctions composées appliqué à $c \circ c^{-1} = Id_{\mathbb{R}^n}$, nous savons que pour $w \in \mathcal{V} \subseteq \mathbb{R}^n$, $\mathrm{Jac}\, c^{-1}(w) = [\mathrm{Jac}\, c(c^{-1}(w))]^{-1}$. Nous noterons $\mathrm{jac}\, c := \det(\mathrm{Jac}\, c)$.

Nous dirons qu'un domaine $\mathcal{D}$ est *mesurable* s'il est du type de ceux où nous avons défini l'intégrale itérée, c'est-à-dire tel qu'ou bien il existe $\overline{\alpha} : [a,b] \longrightarrow \mathbb{R}$ et $\overline{\beta} : [a,b] \longrightarrow \mathbb{R}$, $\overline{\alpha} \leq \overline{\beta}$ continues, avec

$$\mathcal{D} := \{(x,y) \in \mathbb{R}^2 \;\mid\; x \in [a,b] \text{ et } y \in [\overline{\alpha}(x), \overline{\beta}(x)]\}.$$

ou bien il existe $\overline{a} : [\alpha,\beta] \longrightarrow \mathbb{R}$ et $\overline{b} : [\alpha,\beta] \longrightarrow \mathbb{R}$, $\overline{a} \leq \overline{b}$ continues, avec

$$\mathcal{D} := \{(x,y) \in \mathbb{R}^2 \;\mid\; y \in [\alpha,\beta] \text{ et } x \in [\overline{a}(y), \overline{b}(y)]\},$$

Théorème 7.1 *Soit $\mathcal{D}$ un domaine mesurable de $\mathbb{R}^2$ contenu dans un ouvert $\mathcal{U}$. Soit $c : \mathcal{U} \longrightarrow \mathcal{V}$ un changement de variables. On suppose également $c(\mathcal{D})$ mesurable. Soit $g : \mathcal{V} \longrightarrow \mathbb{R}$ une fonction continue. Alors*

$$\iint_{c(\mathcal{D})} g(u,v)dudv = \iint_{\mathcal{D}} g(c(x,y))|jac\, c(x,y)|dxdy. \tag{7.1}$$

Preuve : Posons $f(x,y) := g(c(x,y))|\mathrm{jac}\, c(x,y)|$; nous avons donc à montrer que

$$\iint_{c(\mathcal{D})} g(u,v)dudv = \iint_{\mathcal{D}} f(x,y)dxdy.$$

Par transfert, nous pouvons supposer que $\mathcal{U}$, $\mathcal{D}$, g, et c sont standard ; f et $c(\mathcal{D})$ sont donc standard également. Comme toute fonction continue est la différence de deux fonctions continues positives, *nous pouvons nous borner à étudier le cas où g est positive ; f sera donc également positive.*

Rappelons comment nous avons défini $\iint_{\mathcal{D}} f$:

Somme de Riemann

Par hypothèse $\mathcal{D} \subseteq [a,b] \times [\alpha,\beta]$; on découpe le rectangle $[a,b] \times [\alpha,\beta]$ en pavés i-petits pointés :

$$\mathcal{P}_{ij} := [x_i, x_{i+1}[\times[y_j, y_{j+1}[\ni (\xi_i, \eta_j).$$

Nous supposerons, pour simplifier les notations, que tous ces intervalles ont même longueur : $x_{i+1} - x_i = \varepsilon = y_{j+1} - y_j$, et que le *point-base* (ξ_i, η_j) est le milieu de ce carré.

On considère alors trois familles emboitées de tels pavés :

- la famille des pavés contenus entièrement dans $\mathcal{D}$: ceux pour lesquels on a $\mathcal{P}_{ij} \subseteq \mathcal{D}$.
- la famille des pavés dont le point-base appartient à $\mathcal{D}$: ceux pour lesquels on a $(\xi_i, \eta_j) \in \mathcal{D}$.
- la famille des pavés dont un point au-moins appartient à $\mathcal{D}$: ceux pour lesquels on a $\mathcal{P}_{ij} \cap \mathcal{D} \neq \emptyset$.

L'aire Δ_{ij} de $\mathcal{P}_{ij}$ est égale à $(x_{i+1}-x_i)(y_{j+1}-y_j)$. On appelle somme de Riemann associée à ce pavage pointé du domaine de f la somme $R := \sum_{(\xi_i,\eta_j)\in\mathcal{D}} f(\xi_i,\eta_j)\Delta_{ij}$, c'est-à-dire la somme des volumes des parallélépipèdes construits sur les pavés de la deuxième famille dont la hauteur est la valeur de f au point-base, et nous posions $\iint_{\mathcal{D}} f = {}^oR$.

Sommes de Darboux

Deux autres sommes sont naturelles dans cette question : R_{int} et R_{ext} dont voici la définition. Soit $\overline{\mathcal{P}}_{ij}$ le pavé fermé-borné infinitésimal

$$\overline{\mathcal{P}}_{ij} := [x_i, x_{i+1}] \times [y_j, y_{j+1}].$$

Comme f est continue, il existe des points $p_{ij}^+ \in \overline{\mathcal{P}}_{ij}$ et $p_{ij}^- \in \overline{\mathcal{P}}_{ij}$ tels que

$$f(p_{ij}^-) = \operatorname{Inf} f(\overline{\mathcal{P}}_{ij}) =: M_{ij}^- \quad , \text{et} \quad f(p_{ij}^+) = \operatorname{Inf} f(\overline{\mathcal{P}}_{ij}) =: M_{ij}^+ \ .$$

Comme f est standard, continue, et les points (ξ_i, η_j), p_{ij}^-, et p_{ij}^+ i-voisins d'un même standard de $\mathcal{D}$, on a, en ordre croissant,

$$f(p_{ij}^-) \simeq f(\xi_i, \eta_j) \simeq f(p_{ij}^+).$$

Au-dessus de chaque pavé $\mathcal{P}_{ij}$ nous introduisons trois parallélépipèdes :

$$\mathcal{R}_{ij} = \mathcal{P}_{ij} \times [0, f(\xi_i, \eta_j)] \quad , \quad \mathcal{R}_{ij}^- = \mathcal{P}_{ij} \times [0, M_{ij}^-] \quad , \text{et} \quad \mathcal{R}_{ij}^+ = \mathcal{P}_{ij} \times [0, M_{ij}^+].$$

Soit $\mathcal{R}$ la région de $\mathbb{R}^3_{xyz}$ comprise entre $z = 0$ et $z = f(x, y)$ pour $(x, y) \in \mathcal{D}$, et soit

$$\mathcal{R}^- := \bigcup_{\mathcal{P}_{ij}\subseteq\mathcal{D}} \mathcal{R}_{ij}^- = \bigcup_{\mathcal{P}_{ij}\subseteq\mathcal{D}} \mathcal{P}_{ij} \times [0, M_{ij}^-] \quad , \text{et} \quad \mathcal{R}^+ := \bigcup_{\mathcal{P}_{ij}\subseteq\mathcal{D}} \mathcal{R}_{ij}^+ = \bigcup_{\mathcal{P}_{ij}\subseteq\mathcal{D}} \mathcal{P}_{ij} \times [0, M_{ij}^+].$$

Nous avons donc

$$\mathcal{R}^- \subseteq \mathcal{R} \subseteq \mathcal{R}^+.$$

Définissons R_{int} comme étant le volume de $\mathcal{R}^-$, R_{ext} comme étant le volume de $\mathcal{R}^+$. Par définition $\iint_{\mathcal{D}} f$ est le volume de $\mathcal{R}$, et comme $\mathcal{R}^- \subseteq \mathcal{R} \subseteq \mathcal{R}^+$, on a[1]

$$R_{\text{int}} := \sum_{\mathcal{P}_{ij}\subseteq\mathcal{D}} M_{ij}^- \Delta_{ij} \quad \leq \iint_{\mathcal{D}} f \leq \sum_{\mathcal{P}_{ij}\cap\mathcal{D}\neq\emptyset} M_{ij}^+ \Delta_{ij} =: R_{\text{ext}}.$$

Nous avons montré (lemme 3.3) que

$$R_{\text{int}} \simeq R_{\text{ext}}.$$

Pour établir le théorème, nous allons associer à g des régions $\mathcal{S}^-$ et $\mathcal{S}^+$ de volume respectif S_{int} et S_{ext} telles que

$$S_{\text{int}} \leq \iint_{c(\mathcal{D})} g \leq S_{\text{ext}}.$$

Nous montrerons par ailleurs que $S_{\text{int}} \simeq R_{\text{int}}$ et $S_{\text{ext}} \simeq R_{\text{ext}}$, ce qui montrera que les nombres standard $\iint_{\mathcal{D}} f$ et $\iint_{c(\mathcal{D})} g$ sont i-voisins donc égaux.

Voici l'observation fondamentale :

Comme le changement de variables c est $\mathcal{C}^1$, par la proposition 6.12, pour chaque pavé $\mathcal{P}_{ij}$, son image $c(\mathcal{P}_{ij})$ est asymptotique à un parallélogramme κ_{ij} d'aire $|\operatorname{jac} f(\xi_i, \eta_j)|\Delta_{ij}$; en d'autres termes, il existe des paralélogrammes κ_{ij}^- et κ_{ij}^+ tels que

$$\kappa_{ij}^- \subseteq c(\mathcal{P}_{ij}) \subseteq \kappa_{ij}^+$$

et

$$\operatorname{Aire} \kappa_{ij}^{\pm} = \operatorname{Aire} \kappa_{ij}(1 + \phi) = |\operatorname{jac} f(\xi_i, \eta_j)|\Delta_{ij}(1 + \phi). \tag{7.2}$$

L'intégrale $\iint_{c(\mathcal{D})} g$ est le volume d'une région $\mathcal{S}$ dont nous pouvons à présent encadrer le volume.

[1]ici, en toute rigueur, il conviendrait d'abord de montrer que si $\varphi \leq \psi$, alors $\iint_{\mathcal{D}} \varphi \leq \iint_{\mathcal{D}} \psi$

Approximation de $\mathcal{S}$

Soit $\mathcal{S}$ la région de $\mathbb{R}^3_{uvz}$ comprise entre $z = 0$ et $z = g(u,v)$, pour $(u,v) \in c(\mathcal{D})$. Par définition, $\iint_{c(\mathcal{D})} g$ est le volume de $\mathcal{S}$. Nous allons encadrer ce volume de façon analogue à ce que nous avons fait pour f, mais nous allons utiliser cette fois des parallélogrammes κ^-_{ij} et κ^+_{ij} à la place des pavés $\mathcal{P}_{ij}$ que nous avons utilisés pour $\iint_{\mathcal{D}} f$.

Comme g est continue et que les parallélogrammes κ^+_{ij} sont fermés-bornés, il existe des points $q^-_{ij} \in \kappa^+_{ij}$ et $q^+_{ij} \in \kappa^+_{ij}$ tels que

$$g(q^-_{ij}) = \operatorname{Inf} g(\kappa^+_{ij}) =: N^-_{ij},$$

$$g(q^+_{ij}) = \operatorname{Sup} g(\kappa^+_{ij}) =: N^+_{ij}.$$

Comme g est standard et continue, et que les points $c(\xi_i, \eta_j)$, q^-_{ij}, et q^+_{ij} sont i-voisins d'un même standard de $c(\mathcal{D})$, on a

$$g(c(\xi_i, \eta_j)) \simeq g(q^-_{ij}) \simeq g(q^+_{ij}). \tag{7.3}$$

Au-dessus de chaque parallélogramme κ^-_{ij} et κ^+_{ij} nous considérons les "prismes"

$$\mathcal{S}^- = \kappa^-_{ij} \times [0, N^-_{ij}] \quad , \text{et} \quad \mathcal{S}^+ = \kappa^+_{ij} \times [0, N^+_{ij}].$$

Comme $\kappa^-_{ij} \subseteq \kappa^+_{ij}$ et $N^-_{ij} \leq N^+_{ij}$, nous avons $\mathcal{S}^-_{ij} \subseteq \mathcal{S}^+_{ij}$. Posons

$$\mathcal{S}^- := \bigcup_{c(\mathcal{P}_{ij}) \subseteq c(\mathcal{D})} \mathcal{S}^-_{ij} = \bigcup_{c(\mathcal{P}_{ij}) \subseteq c(\mathcal{D})} \kappa_{ij} \times [0, N^-_{ij}] = \bigcup_{\mathcal{P}_{ij} \subseteq \mathcal{D}} \kappa_{ij} \times [0, N^-_{ij}],$$

puisqu'on a $c(\mathcal{P}_{ij}) \subseteq c(\mathcal{D})$ si et seulement si $\mathcal{P}_{ij} \subseteq \mathcal{D}$. De façon analogue,

$$\mathcal{S}^+ := \bigcup_{c(\mathcal{P}_{ij}) \subseteq c(\mathcal{D})} \mathcal{S}^+_{ij} = \bigcup_{c(\mathcal{P}_{ij}) \subseteq c(\mathcal{D})} \kappa_{ij} \times [0, N^+_{ij}] = \bigcup_{\mathcal{P}_{ij} \subseteq \mathcal{D}} \kappa_{ij} \times [0, N^+_{ij}].$$

Comme $\kappa^-_{ij} \subseteq c(\mathcal{P}_{ij}) \subseteq \kappa^+_{ij}$, et $N^-_{ij} \leq g(u,v) \leq N^+_{ij}$ pour tout $(u,v) \in \kappa^+_{ij}$, nous voyons que

$$\mathcal{S}^- \subseteq \mathcal{S} \subseteq \mathcal{S}^+.$$

Observons que les prismes $\mathcal{S}^-_{ij}$ sont disjoints, alors que les $\mathcal{S}^+_{ij}$ consécutifs se chevauchent. Le volume de $\mathcal{S}^-$ est donc la somme des volumes des prismes qui le compose, alors que le volume de $\mathcal{S}^+$ est inférieur à la somme des volumes des primes qui le composent. En pratique, cela donne

$$S_{\text{int}} := \sum_{\mathcal{P}_{ij} \subseteq \mathcal{D}} N^-_{ij} \ \operatorname{Aire} \kappa^-_{ij} \leq \iint_{c(\mathcal{D})} g \leq \sum_{\mathcal{P}_{ij} \cap \mathcal{D}} N^+_{ij} \operatorname{Aire} \kappa^+_{ij} =: S_{\text{ext}}. \tag{7.4}$$

Calcul asymptotique final : $R_{\text{int}} \simeq S_{\text{int}}$ et $R_{\text{ext}} \simeq S_{\text{ext}}$

Par (7.2) et (7.3), et puisque $(1+\phi)(1+\phi) = 1+\phi$, nous avons

$$\begin{aligned}
N^\pm_{ij} \ \operatorname{Aire} \kappa^\pm_{ij} &= g(q^\pm_{ij}) \ \operatorname{jac}_{ij} \Delta_{ij} (1+\phi) \\
&= (g(c(\xi_i, \eta_j)) + \phi) \ \operatorname{jac}_{ij} \Delta_{ij} (1+\phi) \\
&= g(c(\xi_i, \eta_j)) \ \operatorname{jac}_{ij} \Delta_{ij} + g(c(\xi_i, \eta_j)) \ \operatorname{jac}_{ij} \Delta_{ij} \phi + \phi \ \operatorname{jac}_{ij} \Delta_{ij} (1+\phi) && (7.5) \\
&= g(c(\xi_i, \eta_j)) \ |\operatorname{jac}(\xi_i, \eta_j)| \Delta_{ij} + \phi \Delta_{ij} && (7.6) \\
&= f(\xi_i, \eta_j) \Delta_{ij} + \phi \Delta_{ij} && (7.7)
\end{aligned}$$

où (7.6) se déduit de (7.5) en remarquant que

$$g(c(\xi_i, \eta_j)) \ \operatorname{jac}_{ij} \phi + \phi \ \operatorname{jac}_{ij} (1+\phi) = \phi,$$

puisque $g(c(\xi_i, \eta_j))$ et jac_{ij} sont limités, et (7.7) se déduit de (7.6) par définition de f. On a donc

$$\begin{aligned}
S_{\text{int}} := \sum_{\mathcal{P}_{ij} \subseteq \mathcal{D}} N^-_{ij} \ \operatorname{Aire} \kappa^-_{ij} &= \sum_{\mathcal{P}_{ij} \subseteq \mathcal{D}} (f(\xi_i, \eta_j) \Delta_{ij} + \phi \Delta_{ij}) \\
&= \sum_{\mathcal{P}_{ij} \subseteq \mathcal{D}} f(\xi_i, \eta_j) \Delta_{ij} + \sum_{\mathcal{P}_{ij} \subseteq \mathcal{D}} \phi \Delta_{ij} && (7.8)
\end{aligned}$$

$$\begin{aligned} &= R_{\text{int}} + \not{0} \sum_{\mathcal{P}_{ij} \subseteq \mathcal{D}} \Delta_{ij} \qquad (7.9) \\ &= R_{\text{int}} + \not{0} d_- \\ &\simeq R_{\text{int}} \ . \end{aligned}$$

où (7.9) résulte de (7.8) par application de la proposition 7.2. De façon analogue,

$$S_{\text{ext}} := \sum_{\mathcal{P}_{ij} \cap \mathcal{D} \neq \emptyset} N_{ij}^{+} \ \text{Aire}\, \kappa_{ij}^{+} \quad \simeq \quad R_{\text{ext}} \ .$$

Comme le lemme 3.3 montre que $R_{\text{int}} \simeq R_{\text{ext}}$, on en déduit que $S_{\text{int}} \simeq S_{\text{ext}}$, et comme $S_{\text{int}} \leq \iint_{c(\mathcal{D})} g \leq S_{\text{ext}}$, nous avons

$$\iint_{c(\mathcal{D})} g \simeq S_{\text{ext}} \simeq R_{\text{ext}} \simeq \iint_{\mathcal{D}} f.$$

Comme $\iint_{c(\mathcal{D})} g$ et $\iint_{\mathcal{D}} f$ sont des nombres standard, i-voisins ils ne peuvent être qu'égaux. □

Le “zérobarre” de Van den Berg

Nous avons adopté ici la notation $\not{0}$ (prononcer “zérobarre”) introduite par I. van den Berg, pour désigne un i-petit “générique” : deux occurences distinctes de $\not{0}$ ne désigne pas nécessairement le même i-petit. Ce zérobarre constitue un des éléments du *calcul asymptotique* de Van den Berg. Voici un exemple de joli résultat de calcul asymptotique que nous avons utilisé :

Proposition 7.2 *Soit $(p_i)_{i \in I}$ un ensemble fini de nombres* positifs*. On a*

$$\sum_{i \in I} \not{0} p_i = \not{0} \sum_{i \in I} p_i$$

Preuve : Le terme $\sum_{i \in I} \not{0} p_i$ représente la somme $\sum_{i \in I} \varepsilon_i p_i$ où les ε_i sont les infinitésimaux à priori *distincts* représentés par cette occurence de $\not{0}$. Posons, pour tout $j \in I$, $\lambda_j := p_j / \left(\sum_{i \in I} p_i\right) \geq 0$; on a

$$\sum_{j \in I} \lambda_j = \sum_{j \in I} \frac{p_j}{\sum_{i \in I} p_i} = \frac{\sum_{j \in I} p_j}{\sum_{i \in I} p_i} = 1.$$

Posons à présent $\varepsilon := \sum_{j \in I} \varepsilon_j \lambda_j$ qui est i-petit, puisque compris entre le plus petit est plus grand des éléments, tous i-petits, de $(\varepsilon_i)_{i \in I}$. Nous avons maintenant

$$\varepsilon \sum_{i \in I} p_i = \left(\sum_{j \in I} \varepsilon_j \lambda_j \right) \left(\sum_{i \in I} p_i \right) = \sum_{j \in I} \varepsilon_j \left(\lambda_j \sum_{i \in I} p_i \right) = \sum_{j \in I} \varepsilon_j \frac{p_j}{\sum_{i \in I} p_i} \sum_{i \in I} p_i = \sum_{j \in I} \varepsilon_j p_j \quad .$$

Nous avons finalement bien $\sum_{j \in I} \varepsilon_j p_j = \varepsilon \sum_{i \in I} p_i = \not{0} \sum_{i \in I} p_i$, puisque ε est i-petit, d'où le résultat annoncé. □

Chapitre 8

Le théorème de la fonction implicite

L'objet de cette leçon est l'étude de cas où l'ensemble des (x, y) tels que $f(x, y) = 0$ est, "localement", un graphe de fonction $y = \overline{y}(x)$. E désignera un espace normé de dimension finie (de manière à nous permettre d'utiliser le théorème d'inversion locale) ; en pratique, dans les exercices, les cas où $E = \mathbb{R}$ ou $E = \mathbb{R}^2$ seront tout-à-fait intéressants.

8.1 Le théorème

Théorème 8.1 *Soit $f : E \times \mathbb{R} \longrightarrow \mathbb{R}$ une fonction $\mathcal{C}^1$ sur un ouvert contenant le point (x^0, y^0). Si $f(x^0, y^0) = 0$ et $f'_y(x^0, y^0) \neq 0$, il existe un ouvert $V \ni x^0$ de E, un ouvert $W \ni (x^0, y^0)$ de $E \times \mathbb{R}$, et une fonction $\mathcal{C}^1$ $\overline{y} : V \longrightarrow \mathbb{R}$ telle que*

$$(x, y) \in W \text{ et } f(x, y) = 0 \quad \Leftrightarrow \quad x \in V \text{ et } y = \overline{y}(x). \tag{8.1}$$

Remarques :

- La fonction $\overline{y}$ est la (ou plutôt *une*) *fonction implicite* définie par la relation $f(x, y) = 0$.
- Le point (x^0, y^0) est une *condition initiale* contribuant a choisir $\overline{y}$, puisque $\overline{y}(x^0) = y^0$.
- L'exemple $f(x, y) := (y - x^2)^2$ (et $E = \mathbb{R}$) est tel que $f'_y(x, y) = 0$ dès que $f(x, y) = 0$. Il montre que la condition $f'_y((x^0, y^0)) \neq 0$ n'est nullement nécessaire pour que l'équivalence 8.1 soit satisfaite ; toutefois cette situation est exceptionelle.
- L'exemple $f(x, y) := y^2 - x^2 = (y - x)(y + x)$ et $(x^0, y^0) := (0, 0)$ donne une situation telle que $f'_y(x^0, y^0) \neq 0$ est en défaut, et où c'est l'unicité de $\overline{y}$ qui n'est pas assurée ($\overline{y}_1(x) = +x$ et $\overline{y}_2(x) = -x$ conviennent).
- L'exemple $f(x, y) := x^2 + y^2$ et $(x^0, y^0) := (0, 0)$ donne une situation telle que $f'_y(x^0, y^0) \neq 0$ est en défaut, et où c'est l'existence qui n'est pas assurée : l'ensemble des points où $f(x, y) = 0$ se réduite au point $(0, 0)$ qui ne saurait être le graphe d'une fonction définie sur un ouvert.
- L'exemple $f(x, y) := y^2 - x$ commule d'une certaine manière les deux difficultés : deux solutions $y = \pm\sqrt{x}$ pour $x > 0$, et aucune pour $x < 0$.
- L'exemple $f(x, y) := xy - y - 1$, pour lequel on a bien $f'_y(x, y) \neq 0$ dès que $f(x, y) = 0$ montre que cette dernière condition ne suffit pas pour que V soit égal à E tout entier. Ici $\overline{y}(x) = 1/(x - 1)$ et donc $]-\infty, 1[\cup]1, +\infty[$ est le plus grand V possible.
- L'exemple $f(x, y) = f(x_1, x_2, y) := x_2 \cos(y) - x_1 \sin(y)$, où $E = \mathbb{R}^2 = \mathbb{R}_{x_1} \times \mathbb{R}_{x_2}$ est une autre exemple de ce phénomène, un peu plus subtile. En dehors de l'axe des y, d'équation $x = 0$, on a $f'_y(x, y) \neq 0$ dès que $f(x, y) = 0$. Pourtant V ne peut être choisi égal à $\mathbb{R}^2 \setminus \{(0, 0)\}$: la fonction $\overline{y}$, continue puisque $\mathcal{C}^1$, ne peut être "uniforme", puisque $f(x, y) = 0$ si et seulement si $y = \theta + k\pi$, où θ est "l'angle polaire de x", c'est-à-dire que $x = (r\cos\theta, r\sin\theta)$, et $k \in \mathbb{Z}$ est un entier relatif quelconque. L'ensemble des (x, y) tels que $f(x, y) = 0$ est un double colimaçon.

Figure 8.1: Le difféomorphisme local F de $W := F^{-1}(V \times I)$ sur $V \times I$.

Preuve : 8.1 Posons $F(x,y) = (F_1(x,y), F_2(x,y)) := (x, f(x,y))$; le lecteur qui souhaite comprendre ce choix (qui constitue l'idée fondamentale de la preuve) est invité à considérer les courbes et hypersurfaces $F_1(x,y) = \text{Cste}$ et $F_2(x,y) = \text{Cste}$ (voir figure 8.1). Nous avons $F(x^0, y^0) = (x^0, f(x^0, y^0)) = (x^0, 0)$ et $\operatorname{jac} F(x,y) = f'_y(x,y)$ (l'application $(x,y) \mapsto F_1(x,y) = x$ est la projection sur la première composante ; la différentielle de cette application linéaire est donc, en tout point (x,y), cette première projection ; ceci suffit à calculer facilement la matrice jacobienne de F, puis son déterminant).

En particulier $\operatorname{jac} F(x^0, y^0) = f'_y(x^0, y^0) \neq 0$, et donc, par le théorème d'inversion locale 6.11, F est un difféomorphisme d'un voisinage $\mathcal{V}$ de $v_0 := (x^0, y^0)$ sur un voisinage $\mathcal{W}$ de $w_0 := (x_0, 0)$; notons $G = (G_1, G_2) : \mathcal{W} \longrightarrow \mathcal{V}$ l'inverse local de F. Soient V et I des boules ouvertes centrées en x_0 et 0 suffisamment petites pour qu'on ait $\mathcal{W} \supseteq V \times I (\ni (x_0, 0))$; I est en fait simplement un intervalle. Soit $W := F^{-1}(V \times I)$, et posons $\overline{y}(x) := G_2(x, 0)$, pour $x \in V$. Avec l'aide de la figure 8.1, il est évident que ce choix de V, W, et $\overline{y} : V \longrightarrow \mathbb{R}$ convient ; en voici néanmoins une preuve algébrique.

Observons tout d'abord que pour $(X,Y) \in V \times I (\subset \mathcal{W})$ nous avons par définition de F

$$(X,Y) = F \circ G(X,Y) = F(G_1(X,Y), G_2(X,Y)) = (G_1(X,Y), *)$$

où l'étoile $*$ désigne une expression dont la valeur précise ne nous sera pas utile. Donc $G_1(X,Y) = X$ pour tout $(X,Y) \in V \times I$.

Notons ensuite que, comme $X = G_1(X, 0)$, nous avons

$$(x, f(x, \overline{y}(x))) = F(x, \overline{y}(x)) = F(G_1(x,0), G_2(x,0)) = F \circ G(x,0) = (x,0)$$

et donc $f(x, \overline{y}(x)) = 0$.

Finalement, comme F est bijective de W sur $F(W) = V \times I$,

$$(x,y) \in W \text{ et } f(x,y) = 0$$

équivaut à

$$(x, f(x,y)) = F(x,y) \in F(W) = V \times I \text{ et } (x, f(x,y)) = F(x,y) = (x,0) = F(x, \overline{y}(x))$$

qui équivaut, puisque F est injective, à

$$x \in V \text{ et } y = \overline{y}(x).$$

$\square$

Proposition 8.2 *Soit $E = \mathbb{R}^n$; si $f(x, \overline{y}(x)) = 0$ pour tout x, alors, pour tout $i = 1 \ldots n$, on a*

$$f'_{x_i}(x, \overline{y}(x)) + f'_y(x, \overline{y}(x)) \cdot \overline{y}'_{x_i}(x) = 0 \quad (x \in V).$$

En particulier

$$\overline{y}'_{x_i} = -\frac{f'_{x_i}(x^0, y^0)}{f'_y(x^0, y^0)}.$$

Preuve : Il suffit de calculer la dérivée partielle par rapport à chaque coordonnée x_i de la fonction composée $x \mapsto f(x, \overline{y}(x))$ dont on sait, par ailleurs, qu'il s'agit de la fonction identiquement nulle, donc constante, donc à dérivées partielles toutes nulles. $\square$

8.2 Application : sous-variétés de codimension un.

Soit G un espace vectoriel de dimension n. Rappelons qu'on dit que l'espace G est somme directe de ses deux sous-espaces E et F, et on note $G = E \oplus F$, si et seulement si tout vecteur $v \in G$ s'écrit de façon unique $v = x + y$, avec $x \in E$ et $y \in F$. Nous écrirons, dans ce cas, $v \overset{E\oplus F}{=} x + y$. Notons que si G est normé, il en est de même de de E et F, en considérant simplement la norme induite par restriction à E ou F.

Définition : Soit $A \subset G$ un sous-ensemble A d'un espace normé G. On dit que A est *lisse de dimension* r au point $a \in A$ si et seulement si G admet une décomposition en somme directe $G = E \oplus F$ ayant la propriété suivante :

1. il existe un voisinage W de $a \overset{E\oplus F}{=} x_0 + y_0$ dans G,
2. il existe un voisinage V de x_0 dans E,
3. il existe une fonction $\mathcal{C}^1$ $\overline{y}$ de E dans F définie sur V,

tels que $w \overset{E\oplus F}{=} x + y \in A \cap W$ si et seulement si $x \in V$ et $y = \overline{y}(x)$.

On dit que A est une sous-variété de dimension r si et seulement si A est lisse de dimension r au voisinage de chacun de ses points $a \in A$.

Corollaire 8.3 *Soient $f : G \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ et $a \in G$ tel que $f(a) = 0$; posons $A := f^{-1}(0)$. Pour que A soit lisse de dimension $n-1$ au point a, il suffit que Grad $f(a) \neq 0$. Dans ce cas, l'espace tangeant au point a à l'hypersurface d'équation $f(v) = 0$ a pour équation Grad $f(a)V = 0$.*

Chapitre 9

Convergence uniforme

Avec ce cours, nous abordons, de façon aussi opérationelle et concrète que possible, une perspective sur un apport important du XXème siècle aux Mathématiques : l'analyse fonctionelle. Alors que l'algèbre se propose de dégager les structures des solutions d'équations sur des ensembles de nombres (entiers, rationnels, complexes), l'analyse fonctionelle se propose d'étudier ce même type de questions sur des ensembles de fonctions, sur lesquels sont définis par exemple les équations différentielles[1] ou intégrales. Avouons que nous avons déjà abordé cette question sans la mentionner lorsque nous avons traité des espaces normés sans toujours nous restreindre à la dimension finie ; nous verrons, à la dernière section de ce chapitre, comment la notion de la convergence uniforme est liée à certains espaces normés.

9.1 Fonctions i-voisines

On peut considérer que ce chapitre est consacré à la question de savoir quelles propriétés une fonction g peut hériter d'une fonction f dès lors que l'on est assuré que les deux fonctions sont i-voisines. Il nous faut donc tout d'abord une définition, guère surprenante :

Définition : Deux fonctions $f : \mathcal{D} \longrightarrow \mathbb{R}$ et $g : \mathcal{D} \longrightarrow \mathbb{R}$, définies sur un même domaine $\mathcal{D}$, sont dites i-voisines et on note $f \simeq g$ si et seulement si $f(x) \simeq g(x)$ pour tout $x \in \mathcal{D}$.

Nous allons vouloir intégrer ou dériver f et g ; l'exemple le plus simple où cela est possible, c'est lorsque $\mathcal{D}$ est un intervalle compact $[a, b]$ lorsqu'on veut intégrer, ou un intervalle ouvert $]a, b[$ lorsqu'on veut dériver : c'est à de tels cas que nous nous restreindrons, mais le lecteur curieux constatera sans doute que les preuves s'étendent facilement à des compacts ou ouverts plus généraux, ce qui montre, s'il était nécessaire, que les résultats éventuellement sophistiqués, peuvent être découverts et compris sur des cas simples.

Cette définition nous conduit tout naturellement à définir la convergence d'une suite standard $(f_n)_{n\geq 0}$ de fonctions vers une fonctions standard $\overline{f}$: on dit que $\overline{f}$ est la limite de $(f_n)_{n\geq 0}$ si et seulement si $f_\omega \simeq \overline{f}$ pour tout indice ω i-grand. En reprenant ce qui a été fait pour les suites numériques en première année, on trouve la définition suivante dans le cas général, que $(f_n)_{n\geq 0}$ et $\overline{f}$ soient standard ou non :

Définition : [convergence uniforme] On dit que la suite $(f_n)_{n\geq 0}$ converge vers $\overline{f}$ si et seulement si pour tout $\varepsilon > 0$, on peut trouver un N tel que $|f_n(x) - \overline{f}(x)| \leq \varepsilon$ pour tout $n \geq N$ et tout $x \in \mathcal{D}$, ce qui ce formule ainsi :

$$\forall \varepsilon > 0 \;\; \exists N \;\; \forall n \;\; n \geq N \Rightarrow [\forall x \in \mathcal{D} \;\; |f_n(x) - \overline{f}(x)| \leq \varepsilon]$$

Exemples :

- Pour chaque indice $n > 0$, posons $f_n(x) := \frac{1}{n} \sin n^2 x$. Nous voyons que pour ω i-grand $|f_\omega(x) - 0| \leq \frac{1}{\omega}$ est i-petit pour tout $x \in \mathbb{R}$; la suite de fonctions $(f_n)_{n>0}$ définies sur $\mathcal{D} := \mathbb{R}$ est une suite standard, et celle-ci converge vers la fonction (standard) nulle.

- Pour chaque indice $n > 0$, posons

$$f_n(x) := 1 - \frac{x^2}{2\,!} + \frac{x^4}{4\,!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)\,!}$$

[1] c'est-à-dire équations différentielles ordinaires ou équations aux dérivées partielles

La formule de Taylor-MacLaurin[2] montre que pour chaque $x \in \mathbb{R}$, $\lim_{n\to\infty} f_n(x) = \cos(x)$, ce qui montre que si la suite $(f_n)_{n\geq 0}$ a une limite, ce ne peut qu'être la fonction cos. Pourtant, comme f_n est un polynôme, $|f_n(x)|$ peut être arbitrairement grand, pourvu que $|x|$ soit suffisamment grand, alors que $|\cos(x)| \leq 1$, donc $|f_n(x) - \cos(x)| \geq |f_n(x)| - 1$ ne saurait être petit pour tout x, donc la suite de fonctions $(f_n)_{n>0}$ définies sur $\mathcal{D} := \mathbb{R}$ ne converge pas vers la fonction cos.

- Formons un troisième exemple à partir du précédant, en considérant cette fois les restrictions $\hat{f}_n$ et côs des fonctions f_n et cos à un intervalle $[-a, +a]$, pour un $a \geq 0$ quelconque fixé. En choisissant un entier $M \geq a$ pour la majoration de $|x^{2n+1}/(2n+1)\,!|$, nous voyons que la suite $(\hat{f}_n)_{n\geq 0}$ converge vers côs. On dit aussi que pour tout $a \geq 0$, *la suite* $(f_n)_{n\geq 0}$ *converge uniformément vers* cos *sur l'intervalle* $[-a, a]$.

9.2 Convergence uniforme et convergence simple

Il est d'usage, pour une suite numérique $(u_n)_{n\geq 0}$ convergeant vers $\overline{u}$, de noter cette propriété $u_u \to \overline{u}$; nous généralisons cet usage en notant $f_n \to \overline{f}$ si $\overline{f}$ est la limite des f_n au sens de la limite uniforme que nous venons de définir. Avec les polynômes de Taylor f_n de la fonction cos, nous venons de voir qu'il se peut que pour chaque x dans $\mathcal{D}$ (ici, on avait $\mathcal{D} = \mathbb{R}$) on ait $f_n(x) \to \overline{f}(x)$ sans que pour autant on ait $f_n \to \overline{f}$ (ce qu'on pourra dénoter par $f_n \not\to \overline{f}$) ; dans ce cas, on dit que $\overline{f}$ est la *limite simple* des f_n, et qu'elle *n'est pas la limite uniforme* des f_n. Comme il est souvent commode, pour les calculs, de faire apparaître une notations pour la "variable" (ici x) d'une fonction, comme dans $\frac{1}{n}\sin n^2x$, il peut y avoir confusion entre convergence simple et convergence uniforme ; on notera donc aussi $f_n(x) \overset{\to}{\to} \overline{f}(x)$ pour $f_n \to \overline{f}$ (ce qui ne laisse qu'une éventuelle ambiguité sur le choix domaine $\mathcal{D}_x$, qu'il convient de fixer par le discours ou par une déclaration telle que $\mathcal{D}_x := \mathbb{R}$ ou $\mathcal{D}_\xi := [-a, +a]$). Ainsi $\frac{1}{n}\sin n^2x \overset{\to}{\to} 0$ et $\sum_{p=0}^n (-1)^p x^{2p}/(2p)\,! \overset{\to}{\not\to} \cos x$ pour $\mathcal{D}_x := \mathbb{R}$, alors que $\sum_{p=0}^n (-1)^p \xi^{2p}/(2p)\,! \overset{\to}{\to} \cos\xi$ pour $\mathcal{D}_\xi := [-a, +a]$.

A noter que si $f_n(x) \overset{\to}{\to} \overline{f}(x)$, alors $f_n(x) \to \overline{f}(x)$ pour tout $x \in \mathcal{D}_x$; voyez-vous pourquoi?

Dans l'exemple de la non uniforme convergence des polynômes de Taylor f_n de cos vers cos, c'étaient les grandes valeurs de x qui étaient en cause. Voici un exemple où la difficulté provient des petites valeurs de x.

Posons $f_n(x) := nx(1-x)^n$ pour $\mathcal{D}_x := [0, 1]$, et supposons par l'absurde que la suite ainsi définie admette une limite uniforme $\overline{f}$. Pour $x = 0$ et $x = 1$ on a $f_n(x) = 0$, donc $\overline{f}(0) = 0$ et $\overline{f}(1) = 0$. Par ailleurs, pour chaque $x \in]0, 1[$, on a $0 < q := (1 - x^2) < 1$, et donc $0 \leq f_n(x) \leq nxq^n$, d'où $0 \leq \lim_{n\longrightarrow\infty} f_n(x) \leq \lim_{n\longrightarrow\infty} nxq^n = 0$ (exercice de type DEUG 1) ; d'où $\overline{f}(x) = 0$ pour tout $x \in \mathcal{D}_x = [0, 1]$. Et pourtant f_n ne tend pas vers f. En effet, le plus grand écart entre f_ω et $0 = \overline{f}$ se trouve au point x_ω où la fonction f_ω est maximale ; la recherche des zéros de $\overline{f}'_\omega$ donne $x_\omega = -\omega + \sqrt{\omega^2+1} = \omega(\sqrt{1 + \frac{1}{\omega^2}} - 1) = \omega(1 + \frac{1+\delta}{2\omega^2} - 1) = \frac{1+\delta}{2\omega}$, avec $\delta \simeq 0$ dès que ω est i-grand, ce qui suggère d'évaluer la fonction f_ω au point $\tilde{x}_\omega := \frac{1}{2\omega}$, avec $\delta \simeq 0$ dès que ω est i-grand ; on obtient $f_\omega(\tilde{x}_\omega) = \frac{1}{2}(1 - \frac{1}{2\omega})^\omega = \frac{1}{2}e^{-\frac{1}{2}}(1+\eta)$, avec $\eta \simeq 0$ dès que ω est i-grand, et donc $f_\omega(\tilde{x}_\omega) \not\simeq 0 = \overline{f}(\tilde{x}_\omega)$. La suite $(f_n(x))_{n\geq 0}$ n'a donc pas de limite (uniforme).

---

[2]) $\cos(x) - f_n(x) = \frac{x^{2n+1}}{(2n+1)\,!}\cos^{(n+1)}(\xi)$ pour un ξ entre 0 et x ; en choisissant un entier $M \geq |x|$, on voit que

$$\left|\frac{x^{2n+1}}{(2n+1)\,!}\right| \leq \frac{M}{1}\frac{M}{2}\frac{M}{3}\cdots\frac{M}{M-1}\frac{M}{M}\frac{M}{M+1}\frac{M}{M+2}\cdots\frac{M}{2n}\frac{M}{2n+1}.$$

Passés les $M-1$ premiers termes, dont on dénote par P le produit, tous les suivants sont inférieurs à 1, et en particulier $\left|\frac{x^{2n+1}}{(2n+1)\,!}\right| \leq \frac{P}{2n+1}$, qui tend bien vers 0 lorsque n tends vers l'infini. A noter que C. MacLaurin ne mérite pas seulement sa célébrité pour avoir attiré notre attention sur le fait qu'il suffit de mémoriser la formule de Taylor pour $x_0 = 0$, selon un principe d'économie fécond en Mathématiques. C'est également lui, en admirateur de l'anglais Newton, qui devant le scepticisme affiché par G. Berkley, philosophe évèque dans le sud de l'Irlande, envers les infiniment petits de Newton, a proposé une explicitation de la méthode des quantités "évanouissante" de Newton, dans son Traité des Fluxions, publié à Edinburg en 1742 (pour plus de détails, voir par exemple E.F. Robertson et R.A.A. Devéria, *Colin MacLaurin, Argyllshire's mathematician*, ISBN 1-902847-10-5). A noter qu'il semble bien que, dans sa critique, Georges Berkley était moins préoccupé d'une quète de fondements pour le calcul de Newton que d'un argument en faveur de ses réflexions théologiques destiné a ses lecteurs rationalistes.

Le non-standard permet de préciser, pour une suite $(f_n)_{n\geq 0}$ standard, cette question des valeurs de la variable qui "empèchent la convergence uniforme". En effet, la convergence simple se formule

$$\forall x \quad \lim_n f_n(x) = \overline{f}(x)$$

Mais comme $(f_n)_{n\geq 0}$ et $\overline{f}$ sont standard, ceci equivaut, par transfert, à

$$\forall^{st} x \quad \lim_n f_n(x) = \overline{f}(x)$$

c'est à dire que, pour x standard, la suite standard $(u_n)_{n\geq 0}$ définie par $u_n = f_n(x)$ converge vers la valeur standard $\overline{u} := \overline{f}(x)$, ou encore que $u_\omega \simeq \overline{u}$ dès que ω est i-grand. Nous obtenons donc une troisième forme pour la convergence simple des $f_n(x)$:

$$\forall^{st} x \ \forall^{\text{i-grand}} \omega \quad f_\omega(x) \simeq f(x).$$

En d'autres termes, pour ω i-grand, en cas de convergence simple, on a $f_\omega(x) \simeq \overline{f}(x)$ pour les x standard, alors que pour la convergence uniforme, nous avons $f_\omega(x) \simeq \overline{f}(x)$ pour tous les x, standard ou non.

9.3 Convergence uniforme et opérations fonctionnelles

Pour connaître la valeur de la dérivée $f'(x)$, il ne suffit pas de connaître $f(x)$: il convient, par exemple, de connaître f sur tout un voisinage ouvert $B(x, r)$. Par connaître $\int f$, il est même nécessaire de connaître f (presque-)[3] partout. C'est en ce sens que dérivation et intégration sont des "opérations fonctionnelles", par opposition à la somme et au produit des fonctions.

Le résultat suivant a peu de contenu opérationnel, mais il est capital du point de vue de l'Analyse :

Théorème 9.1 *La limite uniforme de toute suite de fonctions continues est une fonction continue.*

Le théorème 9.7 ci-dessous est un peu plus fort. Nous démontrerons ce théorème.

Théorème 9.2 *Soit* $(f_n)_{n\geq 0}$ *une suite de fonction continues sur* $[a, b]$ *convergeant uniformément vers* $\overline{f}$. *Alors* $\lim_{n\longrightarrow\infty} \int_a^b f_n(x)dx = \int_a^b \overline{f}(x)dx$.

Ce théorème s'écrit aussi, de façon opérationnelle

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

Notez bien les parenthèses : c'est sur elles seules que reposent de rappeler que ce théorème demande la convergence (uniforme !) des f_n. Il est bien clair que dans les calculs explicites, on les néglige généralement, ce qui donne

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

... avec le risque d'erreur que cela entraine, puisque $\lim_n f_n(x)$ suggère la seule limite simple. Ainsi voici une "preuve" de $\frac{1}{2} = 0$: nous avons

$$\int_0^1 nx(1-x^2)^n dx = \left[\frac{-1}{2}\frac{n}{n+1}(1-x^2)^{n+1}\right]_0^1 = \frac{1}{2}\frac{n}{n+1}$$

d'où

$$\frac{1}{2} = \lim_n \frac{1}{2}\frac{n}{n+1} = \lim_n \int_0^1 nx(1-x^2)^n dx \quad \text{“=”} \quad \int_0^1 \lim_n nx(1-x^2)^n dx = \int_0^1 0 dx = 0.$$

Cette erreur est pourtant moins naïve qu'il n'y paraît. Souvent, la seule méthode à notre disposition pour *calculer* la limite (uniforme, si elle existe) $\overline{f}$ des f_n est de calculer la limite pour chaque x, c'est-à-dire la limite simple. Le théorème 9.2 n'est pas le seul autorisant le "calcul simple" d'une limite d'intégrales : on doit à Lebesgue un cadre beaucoup plus général (son "théorème de convergence dominée") que vous apprendrez prochainement. Biensûr, ce théorème ne s'applique pas non plus à notre exemple.

La preuve du théorème 9.2 se déduit facilement du lemme suivant :

Lemme 9.3 *Soit* $[a, b]$ *un intervalle de longueur* $(b-a)$ *limitée et* $\delta : [a, b] \longrightarrow \mathbb{R}$ *une fonction intégrable i-petite en tout point* $x \in [a, b]$. *Alors* $\int_a^b \delta(x)dx \simeq 0$.

[3] Ce "presque" est placé là pour aiguiser votre curiosité sur une théorie que vous verrez vraissemblablement en Licence : celle de l'intégrale de Lebesgue

Preuve : Si δ est supposée en outre continue, il en est de même de $|\delta|$, et comme $[a,b]$ est compacte, cette fonction admet un maximum $x_0 \in [a,b]$, d'où, pour tout $x \in [a,b]$, $|\delta(x)| \leq |\delta(x_0)| =: M$, ce dernier nombre étant i-petit par hypothèse sur δ. On a donc

$$\left|\int_a^b \delta(x)dx\right| \leq \int_a^b |\delta(x)|dx \leq \int_a^b M dx = (b-a)M \simeq 0.$$

Si l'on ne suppose plus δ continue, il suffit de choisir $M := \mathrm{Sup}_{x\in[a,b]}|\delta(x)|$ et de remarquer que le Sup d'un ensemble (interne) de nombre tous i-petits qui existe puisque l'ensemble est majoré (par 1 par exemple), ne peux qu'être i-petit lui-même. □

Théorème 9.4 *Soit $(f_n)_{n\geq 0}$ une suite de fonctions dérivables sur $]a,b[$ dont la suite des dérivées $(f_n')_{n\geq 0}$ converge uniformément vers g. Supposons en outre qu'il existe $x_0 \in]a,b[$ et y_0 tel que $\lim_{n\longrightarrow\infty} f_n(x_0) = y_0$. Alors $(f_n)_{n\geq 0}$ converge uniformément vers $y_0 + \int_{x_0}^x g(t)dt$.*

Théorème 9.5 *Soit $f(x,y)$ une fonction de classe $\mathcal{C}^1$ sur un ouvert contenant le pavé $[a,b]_x \times]c,d[_y$. Soit g la fonction définie sur $]c,d[_y$ par $g(y) := \int_a^b f(x,y)dx$. Alors g est de classe $\mathcal{C}^1$ et $g'(y) = \int_a^b f_y'(x,y)dx$*

Preuve : Par transfert, on peut supposer f, a, b, et y standard. Nous devons montrer que $\lim_{h\to 0} \frac{1}{h}(g(y+h)-g(y)) = \int_a^b f_y'(x,y)dx$. Comme $\int_a^b f_y'(x,y)dx$ est standard, ceci revient à montrer que pour tout $h \simeq 0$ non nul, on a $\frac{1}{h}(g(y+h)-g(y)) \simeq \int_a^b f_y'(x,y)dx$. Or

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

Or, par le théorème des accroissement finis appliqué à la fonction $\varphi(y) := f(x,y)$ (x fixé), il existe $\eta \in]y, y+h[$ tel que $f(x,y+h) - f(x,y) = hf_y'(x,\eta)$. Comme $[a,b]$ est compact standard, pour chaque $x \in [a,b]$, il existe un standard $x_0 \in [a,b]$ tel que $x \simeq x_0$ (la "partie standard" de x), et comme f_y' est standard continue, et que $\eta \simeq y$ puisque h est i-petit, on a $f_y'(x,\eta) \simeq f_y'(x_0,y) \simeq f_y'(x,y)$, ou encore $f_y'(x,\eta) = f_y'(x,y) + \delta(x)$, avec $\delta(x) \simeq 0$. Finalement, en appliquant le lemme 9.3, on a

$$\frac{1}{h}(g(y+h)-g(y)) = \frac{1}{h}\left(\int_a^b (f(x,y+h)-f(x,y))dx\right)$$

$$= \frac{1}{h}\left(\int_a^b h(f_y'(x,y)+\delta(x))dx\right) = \int_a^b f_y'(x,y) + \int_a^b \delta(x))dx \simeq \int_a^b f_y'(x,y).$$

□

9.4 La norme de la convergence uniforme

Considérons l'ensemble E des fonctions continues $f : \mathcal{D} - \mathbb{R}$ qui sont bornées, c'est-à-dire les fonctions telles qu'il existe, pour chaque d'elles, un $M \geq 0$ ayant la propriété que $|f(x)| \leq M$ pour tout $x \in \mathcal{D}$. Du fait que, pour tout f_1 et f_2 dans E et tout λ_1 et λ_2 dans $\mathbb{R}$, si M_1 et M_2 sont tels que $|f_i(x)| \leq M_i$, $i = 1,2$ et $x \in \mathcal{D}$, alors $|(\lambda_1 f_1 + \lambda_2 f_2)(x)| \leq \lambda_1 M_1 + \lambda_2 M_2 =: M$, on voit que E est un sous-espace vectoriel de l'espace vectoriel des fonctions de $\mathcal{D}$ à valeurs dans $\mathbb{R}$.

Proposition 9.6 *La fonction $\|\ \|_\infty : E \longrightarrow \mathbb{R}$, définie par*

$$\|f\|_\infty := Sup_{x\in\mathcal{D}}|f(x)|$$

est une norme sur l'espace vectoriel E des fonctions continues sur $\mathcal{D}$. On l'appelle la *norme de la convergence uniforme* sur $\mathcal{D}$.

Preuve : Détaillons la preuve dans le cas où $\mathcal{D}$ est compact. Dans ce cas, comme la fonction $|f|$ est continue, elle atteint son maximum sur $\mathcal{D}$, c'est-à-dire qu'il existe $x_0 \in \mathcal{D}$ tel que $|f(x_0)| = \text{Sup}_{x\in\mathcal{D}}|f(x)| = \|f\|_\infty$. Soient donc f, f_1, et f_2 dans E et λ_1, et λ_2 dans $\mathbb{R}$. Notons x', et x'' des maxima respectifs de $|f|$, et $|f_1+f_2|$; nous vérifions les trois propriétés caractéristiques d'une norme :

- On a $0 \leq |f(x)| \leq \text{Sup}_{x\in\mathcal{D}}|f(x)| = \|f\|_\infty$; donc $f(x)=0$ pour tout $x \in \mathcal{D}$ si et seulement si $\|f\|_\infty = 0$, et donc $\|f\|_\infty = 0 \Leftrightarrow f = 0$.

- On a, pour tout $x \in \mathcal{D}$,
$$|\lambda f(x)| = |\lambda||f(x)| \leq |\lambda|\|f\|_\infty = |\lambda||f(x')| = |\lambda f(x')| \leq \text{Sup}_{x\in\mathcal{D}}|\lambda f(x)| = \|\lambda f\|_\infty,$$
et donc $\|\lambda f\|_\infty = \text{Sup}_{x\in\mathcal{D}}|\lambda f(x)| \leq |\lambda|\|f\|_\infty \leq \|\lambda f\|_\infty$, d'où $\|\lambda f\|_\infty = |\lambda|\|f\|_\infty$.

- On a $\|f_1+f_2\|_\infty = |(f_1+f_2)(x'')| \leq |f_1(x'')| + |f_2(x'')| \leq \text{Sup}_{x\in\mathcal{D}}|f_1(x)| + \text{Sup}_{x\in\mathcal{D}}|f_2(x)| = \|f_1\|_\infty + \|f_2\|_\infty$, d'où $\|f_1+f_2\|_\infty \leq \|f_1\|_\infty + \|f_2\|_\infty$.

Indications pour le cas général ($\mathcal{D}$ non nécessairement compact). Dans ce cas, nous ne sommes assuré que de l'existence, pour chaque $\varepsilon > 0$, d'un x_ε tel que $|f(x_\varepsilon)| \leq \text{Sup}_{x\in\mathcal{D}}|f(x)| = \|f\|_\infty \leq |f(x_\varepsilon)| + \varepsilon$. En adaptant en conséquence les notations et calculs ci-dessus, nous obtenons

- Pour tout $x \in \mathcal{D}$ et tout $\varepsilon > 0$
$$\begin{aligned} |\lambda f(x)| &= |\lambda||f(x)| \leq |\lambda|\|f\|_\infty \leq |\lambda|(|f(x'_\varepsilon)| + \varepsilon) = |\lambda f(x'_\varepsilon)| + |\lambda|\varepsilon \\ &\leq \text{Sup}_{x\in\mathcal{D}}|\lambda f(x)| + |\lambda|\varepsilon = \|\lambda f\|_\infty + |\lambda|\varepsilon. \end{aligned}$$
On a donc $\|\lambda f\|_\infty = \text{Sup}_{x\in\mathcal{D}}|\lambda f(x)| \leq |\lambda|\|f\|_\infty \leq \|\lambda f\|_\infty + |\lambda|\varepsilon$. En passant à la limite pour $\varepsilon \to 0^+$, on retrouve $\|\lambda f\|_\infty = \text{Sup}_{x\in\mathcal{D}}|\lambda f(x)| \leq |\lambda|\|f\|_\infty \leq \|\lambda f\|_\infty$ d'où $\|\lambda f\|_\infty = |\lambda|\|f\|_\infty$.

- On a $\|f_1+f_2\|_\infty \leq |(f_1+f_2)(x''_\varepsilon)|+\varepsilon \leq |f_1(x''_\varepsilon)|+|f_2(x''_\varepsilon)|+\varepsilon \leq \text{Sup}_{x\in\mathcal{D}}|f_1(x)|+\text{Sup}_{x\in\mathcal{D}}|f_2(x)|+\varepsilon = \|f_1\|_\infty + \|f_2\|_\infty + \varepsilon$, d'où
$$\|f_1+f_2\|_\infty \leq \|f_1\|_\infty + \|f_2\|_\infty + \varepsilon.$$
En passant à la limite pour $\varepsilon \to 0^+$, on obtient le résultat souhaité.

□

Théorème 9.7 *L'espace normé $(E, \|\ \|_\infty)$ est complet, c'est-à-dire, par définition, que, pour toute suite $(f_n) \subset E$, (f_n) converge si et seulement si (f_n) est de Cauchy. Si f dénote la limite des f_n, on a $f(x) = \lim_n f_n(x)$.*

Preuve : Nous savons que toute suite convergente est de Cauchy. Réciproquement, soit (f_n) une suite de Cauchy. Pour chaque $x \in \mathcal{D}$, comme
$$|f_p(x) - f_q(x)| \leq \text{Sup}_{\xi\in\mathcal{D}}|f_p(\xi) - f_q(\xi)| = \|f_p - f_q\|_\infty.$$
Donc $(f_n(x))$ est de Cauchy, et comme $\mathbb{R}$ est complet, la suite $(f_n(x))$ a une limite. On pose $f(x) := \lim_n f_n(x)$. Il nous faut nous assurer que la fonction f, ainsi définie pour chaque $x \in \mathcal{D}$ est bien dans E, c'est-à-dire continue et bornée, et que $f = \lim f_n$ *au sens de la norme*$\|\ \|_\infty$ (voir l'exemple $f_n(x) = nx(1-n)^n$ ci-dessus, qui conduit à poser $f(x) = 0$, et cette fonction n'est pas la limite au sens de $\|\ \|_\infty$ de cette suite, *qui n'est donc pas convergente*).

Voyons d'abord la continuité de f en tout point $x_0 \in E$: soit $\varepsilon > 0$ quelconque ; comme $f(x_0) = \lim_n f_n(x_0)$
$$\exists N_1 \ \ \forall n \geq N_1 \ \ |f_n(x_0) - f(x_0)| \leq \frac{\varepsilon}{3}. \tag{9.1}$$
Mais comme (f_n) est de Cauchy, il existe aussi N_2 tel que, pour tous $q \geq p \geq N_2$, on a $\|f_p - f_q\|_\infty \leq \frac{\varepsilon}{3}$; en d'autres termes
$$\forall x \in \mathcal{D} \ \ \forall q \geq p \geq N_2 \ \ |f_p(x) - f_q(x)| \leq \|f_p - f_q\|_\infty \leq \frac{\varepsilon}{3}. \tag{9.2}$$
Posons maintenant $N = \text{Max}\,\{N_1, N_2\}$. Comme f_N est continue au point x_0, il existe $\eta > 0$ tel que pour tout $x \in \mathcal{D}$,
$$|x - x_0| \leq \eta \Rightarrow |f_N(x) - f_N(x_0)| \leq \frac{\varepsilon}{3}. \tag{9.3}$$

A présent, pour tout $n \geq N = \text{Max}\{N_1, N_2\}$ et tout $x \in \mathcal{D}$ tel que $|x - x_0| \leq \eta$, nous avons

$$\begin{aligned} |f_n(x) - f(x_0)| &\leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

la majoration de chacun des trois termes résultant respectivement de (9.2) puisque $n \geq N \geq N_2$, de (9.3) puisque $|x - x_0| \leq \eta$, et de (9.1) puisque $N \geq N_1$.

En pratique, pour tout $n \geq N = \text{Max}\{N_1, N_2\}$, nous avons donc

$$|f_n(x) - f(x_0)| \leq \varepsilon \text{ pour tout } n \geq N \text{ et tout } x \in \mathcal{D} \text{ tel que } |x - x_0| \leq \eta$$

et donc, comme $|\ \ | : \mathbb{R} \longrightarrow \mathbb{R}$ est continue, on a pour tout $x \in \mathcal{D}$ tel que $|x - x_0| \leq \eta$

$$|f(x) - f(x_0)| = |\lim_n f_n(x) - f(x_0)| = \lim_n |f_n(x) - f(x_0)| \leq \lim_n \varepsilon = \varepsilon,$$

ce qui montre que f est bien continue.

Enfin, en reprenant (9.2), on a pour tout $x \in \mathcal{D}$ et tout $q \geq n \geq N_2$,

$$|f_n(x) - f(x)| = |f_n(x) - \lim_q f_q(x)| = \lim_q |f_n(x) - f_q(x)| \leq \lim_q \frac{\varepsilon}{3} = \frac{\varepsilon}{3} \leq \varepsilon.$$

En d'autres termes, $|f_n(x) - f(x)| \leq \varepsilon$ pour tout $x \in \mathcal{D}$ et tout $n \geq N_2$, ce qui montre d'une part que f est bornée (par $\|f_n\|_\infty + \varepsilon$), et d'autre part que $\|f_n - f\|_\infty = \text{Sup}_{x \in \mathcal{D}} |f_n(x) - f(x)| \leq \varepsilon$, ce qui montre que f est bien la limite, au sens de l'espace normé $(E, \|\ \ \|_\infty)$, de la suite (f_n). □

Exercice : Déduire le théorème 9.1 de ce théorème 9.7.

Corollaire 9.8 *Soit (u_n) une suite de fonctions continues sur $\mathcal{D}$. Supposons qu'il existe une suite numérique (α_n) telles que*

1. *on a $|u_n(x)| \leq \alpha_n$ pour tout $x \in \mathcal{D}$,*
2. *la série $\sum_{n \geq 0} \alpha_n$ converge.*

Posons $f_N(x) := \sum_{n=0}^N u_n(x)$. Alors, la suite de fonctions f_N ainsi définie converge uniformément vers une limite f. On note $\sum_{n=0}^\infty u_n$ cette limite ; elle est continue et a pour valeurs $f(x) = \sum_{n=0}^\infty u_n(x)$.

Preuve : Comme la série numérique $\sum_{n \geq 0} \alpha_n$ converge, la suite $s_N := \sum_{n=0}^N \alpha_n$ est convergente ; c'est donc une suite de Cauchy. On en déduit que la suite de fonctions (f_N) est une suite de Cauchy de l'espace normé $(E, \|\ \ \|_\infty)$, puisque, pour $q \geq p$,

$$\begin{aligned} \|f_q - f_p\|_\infty &\leq \|f_q - f_{q-1}\|_\infty + \|f_{q-1} - f_{q-2}\|_\infty + \ldots + \|f_{p+1} - f_p\|_\infty \\ &= \|u_q\|_\infty + \|u_{q-1}\|_\infty + \ldots + \|u_{p+1}\| \\ &\leq \sum_{n=p+1}^{q} \alpha_n = s_q - s_p = |s_q - s_p|. \end{aligned}$$

Donc la série $\sum_{n \geq 0} u_n$ converge dans l'espace normé $(E, \|\ \ \|_\infty)$ en vertu du théorème 9.7. □

Définition : Soit (u_n) une suite de fonctions continues et bornées définies sur $\mathcal{D}$. Posons $a_n := \|u_n\|_\infty$. On dit que la série de fonctions $\sum_{n \geq 0} u_n(x)$ est *normalement convergente*, si la série numérique $\sum_{n \geq 0} a_n$ converge.

Corollaire 9.9 *Toute série normalement convergente converge uniformément vers sa limite, et cette limite est continue.*

Preuve : Il suffit, dans le corollaire 9.8, de poser $\alpha_n := \|u_n\|_\infty$. □

Exemple : Revenons sur le troisième exemple du début du chapitre. On a $\mathcal{D} = [-a, +a]$ et $u_n(x) = x^{2n}/(2n)\,!$. On a $a_n = \|u_n\|_\infty = a^{2n}/(2n)\,!$. La série $\sum_{n \geq 0} a_n$ est convergente (appliquer par exemple le critère de d'Alembert). La série $\sum_{n \geq 0} x^{2n}/(2n)\,!$ est donc normalement convergente. Sur $[-a, +a]$ la série $\sum_{n \geq 0} x^{2n}/(2n)\,!$ est donc uniformément convergente vers sa limite, $\cos(x)$. On retrouve ainsi que la fonction cos est continue.

Chapitre 10

Fonctions holomorphes et séries entières

10.1 Nombres complexes

Comme vous le savez, les nombres complexes sont les expressions de la forme $z = x + iy$, avec x et y dans $\mathbb{R}$, que l'on traite tout d'abord comme des polynômes de degré un en l'indéterminée i, ce quindéfini $z' + z''$ pour $z' = x' + iy'$ et $z'' = x'' + iy''$. Par bilinéarité du produit de polynômes, on a $z' \cdot z'' = (x' + iy') \cdot (x'' + iy'') = x'x'' + i(x'y'' + x''y') + i^2 y'y''$. Puis on convient que $i^2 = -1$, ce qui fait de i une solution de $1 + z^2 = 0$, et conduit à

$$z' \cdot z'' = (x' + iy') \cdot (x'' + iy'') = x'x'' - y'y'' + i(x'y'' + x''y').$$

Selon une idée de Jean Robert Argand apparue à la fin du 18ème siècle, donc après l'introduction des nombres complexes, on représente l'expression $x + iy$ par le couple de réels (x, y), donc $1 = 1 + i0$ par $(1, 0)$ et $i = 0 + i1$ par $(0, 1)$. Poser $i^2 = -1$ revient à poser $(0, 1) \cdot (0, 1) = (1, 0)$. Avant la représentation géométrique des complexes, la relation $i^2 = -1$ paraissait tellement peu naturelle, que la composante y de $z = x + iy$ a été appelée la partie *imaginaire* de z, alors que x est sa partie *réelle*. Nous noterons $x = \operatorname{Re} z$ et $y = \operatorname{Im} z$. Vous avez vu en terminale que cette manière de faire étend le calcul du corps de réels $x \in \mathbb{R}$ à un calcul sur le corps des complexes $z \in \mathbb{C}$. L'opération $x + iy = z \mapsto \overline{z} = x - iy$ (c'est-à-dire la symétrie $(x, y) \mapsto (x, -y)$ dans le plan d'Argand) est souvent utile car pour tout $z \in \mathbb{C}$

$$z \cdot \overline{z} = (x + iy) \cdot (x - iy) = x^2 + y^2 \ (= \|(x, y)\|_2^2 \text{ dans le plan d'Argand}),$$

c'est-à-dire un réel, positif de surcroît. On pose $|z| = \sqrt{z \cdot \overline{z}}$, c'est-à-dire $\|(x, y)\|_2$ du plan d'Argand, d'où $|z' + z''| \leq |z'| + |z''|$; on vérifie facilement les relations

$$\overline{z' + z''} = \overline{z'} + \overline{z''} \ , \ \overline{z' \cdot z''} = \overline{z'} \cdot \overline{z''} \ \ ; \ \ |z' \cdot z''| = |z'| \cdot |z''| \ , \ \frac{\overline{z}}{|z|^2} = \frac{\overline{z}}{z \cdot \overline{z}} = \frac{1}{z}.$$

Si on note $r := |z|$ et $\vartheta = \operatorname{angle}(Ox, Oz)$ dans le plan d'Argand, et $[r, \vartheta] := z$, on vérifie facilement que $z' \cdot z'' = [r', \vartheta'] \cdot [r'', \vartheta''] = [r'r'', \vartheta' + \vartheta'']$ et, partant, $\frac{z'}{z''} = \frac{[r', \vartheta']}{[r'', \vartheta'']} = [r'/r'', \vartheta' - \vartheta'']$, d'où en particulier $\frac{1}{z} = \frac{[1,0]}{[r,\vartheta]} = [1/r, -\vartheta]$.

10.2 Dérivation complexe

Le fait que $\mathbb{C}$ soit un corps permet de considérer, pour toute fonction $z \mapsto f(z)$, la question de l'existence d'une limite pour

$$\frac{f(z + h) - f(z)}{h}$$

lorsque $h \to 0$, $h \in \mathbb{C} \div \{0\}$. A titre de comparaison, souvenez-vous des difficultés que nous avons eu pour traité de la différentielle Dg pour une fonction $(x, y) \mapsto g(x, y) = (g_1(x, y), g_2(x, y))$ qui, elle, se représente par une matrice 2×2 et non un nombre.

Définition : Soit U un ouvert de $\mathbb{C}$ (vu comme le plan d'Argand). On dit que $f : U \longrightarrow \mathbb{C}$ admet l pour dérivée au point z si et seulement si

$$l = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}.$$

Si f admet une dérivée en chaque point z de U, on dit que f est *holomorphe*[1], et on pose

$$f'(z) := \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}.$$

En copiant sur le cas réel , on montre

Proposition 10.1 *Si f et g sont holomorphes sur U, il en est de même de $f+g$, fg et f/g, et $(f+g)' = f' + g'$, $(fg)' = f'g + fg'$, $(f/g)' = (f'g - fg')/g^2$.*

En particulier $(az^n)' = naz^{n-1}$, tout comme pour un monôme réel. A noter que la fonction $z \mapsto \overline{z}$ n'est nulle part dérivable. En effet

$$\frac{\overline{(z+h)} - \overline{z}}{h} = \frac{\overline{z} + \overline{h} - \overline{z}}{h} = \frac{\overline{h}}{h}$$

vaut 1 si $h = h_1$ est réel et vaut -1 si $h = ih_2$ est imaginaire pur, aussi petits soient-ils. Il n'y a donc pas de limite au rapport de dérivations, en quelque point z que ce fût.

Toujours en copiant sur le cas réel, on définit de manière analogue la dérivée seconde $f''(z) := (f')'(z)$, et, par récurrence, $f^{(n)}(z) = (f^{(n-1)})'(z)$. Ici se produit un miracle (dont la preuve naturelle est très détournée) : si f est *holomorphe* sur U, alors elle admet des dérivées de tous les ordres ! En d'autres termes, pour la dérivation complexe, $\mathcal{C}^1$ implique $\mathcal{C}^\infty$, et il y a mieux : on montre (bien-entendu) une formule de Taylor

$$f(z) = f(z_0) + (z - z_0) f'(z_0) + \ldots + \frac{(z-z_0)^n}{n\,!} f^{(n)}(z_0) + \frac{(z-z_0)^{n+1}}{(n+1)\,!} \phi_{n+1}(z),$$

où $\phi_{n+1}(z)$ est une quantité qui reste bornée lorsque $z - z_0$ devient petit. Ceci conduit à des séries de la forme $\sum_{n \geq 0} a_n (z - z_0)^n$ et nous verons que ces séries convergent vers $f(z)$.

Rappelons que ceci n'est pas toujours vrai pour les fonctions $\mathcal{C}^\infty$ réelles : il suffit de considérer la fonction, nulle en 0, et définie par $f(x) = e^{-\frac{1}{x^2}}$ pour $x \neq 0$ dans $\mathbb{R}$; on montre que pour tout n on a $f^{(n)}(0) = 0$, et la série ci-dessus, pour $z_0 = x_0 = 0$, est donc nulle, et ne converge donc pas vers $f(x)$, qui elle, ne s'annule pas en dehors de $x_0 = 0$. Une telle pathologie n'est, de fait, pas possible pour une fonction holomorphe.

10.3 Séries entières

Nous avons étudié la convergence de séries de fonctions telles que $\sum_{n \geq 0} u_n(x)$; nous considèrerons ici le cas où $u_n(x) = a_n x^n$, faisant intervenir des puissances *entières* de la variable x, d'où le nom de ces séries. Nous étendons notre étude au cas où la variable est complexe : ce cas n'est pas plus difficile, et donne des résultats souvent plus naturels.

Définition : On appelle *série entière* toute série de la forme $\sum_{n \geq 0} a_n z^n$, $a_n \in \mathbb{C}$ et $z \in \mathbb{C}$

10.3.1 Disque et rayon de convergence

La question naturelle est, bien-entendu, de savoir quand une telle série converge. Nous savons que pour qu'une telle série converge, il faut que $|a_n z^n|$ tende vers 0. Cette condition est, en fait, "presque" suffisante.

Théorème 10.2 *Il existe un réel R, avec $0 \leq R \leq +\infty$, tel que*

1. *Si $r \in [0, R[$, $\sum_{n \geq 0} a_n z^n$ converge normalement sur $\{|z| \leq r\}$.*
2. *Si $r > R$, la suite $(a_n z^n)$ est non bornée et $\sum_{n \geq 0} a_n z^n$ diverge.*

[1] du grec *holos*, "entier", et *morphê*, "forme", c'est-à-dire admettant une représentation sous forme de série entière.

Ce nombre R s'appelle le rayon de convergence *de la série entière* $\sum_{n\geq 0} a_n z^n$ *et vérifie*

$$R = Sup\{r \geq 0 \mid \text{la suite } (a_n r^n)_{n\geq 0} \text{ est bornée}\} = Sup\{r \geq 0 \mid \text{la suite } (a_n r^n)_{n\geq 0} \text{ tend vers } 0.\}$$

La preuve tient, outre les théorèmes généraux de convergence, au résultat suivant.

Proposition 10.3 *Soit $\alpha = (\alpha_n)_{n\geq 0}$ une suite de réels positifs. Soit $R = Sup \mathcal{B}_\alpha$, où*

$$\mathcal{B}_\alpha = \{r_0 \geq 0 \mid \text{la suite } (\alpha_n r_0^n)_{n\geq 0} \text{ est bornée}\}.$$

Si $0 \leq r < R$, $\sum_{n\geq 0} \alpha_n r^n$ converge ; si $r > R$, la suite $(\alpha_n r^n)$ est non bornée et $\sum_{n\geq 0} \alpha_n r^n$ diverge.

Preuve : On remarque que si $r_0 \in \mathcal{B}_\alpha$, alors $[0, r_0] \subseteq \mathcal{B}_\alpha$; en effet, si $0 \leq r < r_0$, on a $\alpha_n r^n = \alpha_n r_0^n \left(\frac{r}{r_0}\right)^n$, qui tend vers 0 car $0 \leq q := \frac{r}{r_0} < 1$, et est donc bornée. Donc

$$\mathcal{B}_\alpha = \bigcup_{r_0 \in \mathcal{B}_\alpha} [0, r_0].$$

En d'autres termes $\mathcal{B}_\alpha$ est une réunion d'intervalles ayant un point (0) en commun. Donc $\mathcal{B}_\alpha$ est une intervalle, ici $\mathcal{B}_\alpha = [0, b]$ ou $\mathcal{B}_\alpha = [0, b[$, et par définition $R = \operatorname{Sup} \mathcal{B}_\alpha = b$.

A présent, si $0 \leq r < R$, choisissont un $r_0 \in]r, R[$. Donc $0 < q := \frac{r}{r_0} < 1$, et $0 \leq \alpha_n r^n = \alpha_n r_0^n \left(\frac{r}{r_0}\right)^n = \alpha_n r_0^n q^n \leq \beta q^n$, où β est un majorant de la suite $(\alpha_n r_0^n)$ qui est bornée puisque $r_0 \in \mathcal{B}_\alpha$. Donc, comme $\sum_{n\geq 0} \beta q^n$ converge, il en est de même de $\sum_{n\geq 0} \alpha_n r^n$.

Par ailleurs, si $R < r$, la suite $(\alpha_n r^n)_{n\geq 0}$ est non-bornée par définition de $R = \operatorname{Sup} \mathcal{B}_\alpha$, et donc $\sum_{n\geq 0} \alpha_n r^n$ diverge. □

Preuve : (du théorème 10.2) Posons $\alpha_n := |a_n|$ et choisissons $R = \operatorname{Sup} \mathcal{B}_\alpha$ comme dans la proposition 10.3.

Si $|z| =: r > R$, nous voyons que la suite $|a_n z^n| = \alpha_n r^n$ n'est pas pas bornée, donc la suite des $a_n z^n$ n'est pas bornée, et ne tends à *forciori* pas vers 0. Donc la série $\sum_{n\geq 0} a_n z^n$ diverge.

Si z=: r¿R, pour montrer que la série $\sum_{n\geq 0} a_n z^n$ converge, il faut et il suffit que la série $\operatorname{Re}(\sum_{n\geq 0} a_n z^n) = \sum_{n\geq 0} \operatorname{Re}(a_n z^n)$ converge et que la série $\operatorname{Im}(\sum_{n\geq 0} a_n z^n) = \sum_{n\geq 0} \operatorname{Im}(a_n z^n)$ converge. Choisissons r_0 tel que $r < r_0 < R$ et posons $q = \frac{r}{r_0} < 1$. Comme $r_0 \in \mathcal{B}_\alpha$, il existe M tel que $\alpha_n r_0^n \leq M$ pour tout n. Nous avons

$$|\operatorname{Re}(a_n z^n| \leq |a_n z^n| = |a_n||z|^n = \alpha_n r^n = \alpha_n r_0^n \left(\frac{r}{r_0}\right)^n = \alpha_n r_0^n q^n \leq M q^n,$$

et ce dernier nombre est le terme général d'une série convergente. Donc $\sum_{n\geq 0} \operatorname{Re}(a_n z^n)$ est absolument convergente. On raisonnerait de manière analogue pour montrer que $\operatorname{Im}(\sum_{n\geq 0} a_n z^n)$ est convergente. □

Proposition 10.4 *Soit $\sum_{n\geq 0} a_n r^n$ une série entière, et R son rayon de convergence.*

1. *Si $\lim_{n\to\infty} \sqrt[n]{|a_n|} = l$ existe, alors $R = 1/l$.*
2. *Si $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = l$ existe, alors $R = 1/l$.*

Preuve : Par définition de R, et en choisissant $z = r = |z|$, nous savons que $\sum_{n\geq 0} |a_n| r^n$ converge si $r < R$ et diverge si $r > R$. Soit $L := \lim_n \sqrt[n]{|a_n| r^n} = \lim_n r \sqrt[n]{|a_n|} = r \lim_n \sqrt[n]{|a_n|} = rl$. Par le critère de la racine, nous savons aussi que $\sum_{n\geq 0} |a_n| r^n$ converge si $(rl =)L < 1$, c'est-à-dire $r < \frac{1}{l}$, et diverge si $(rl =)L > 1$, c'est-à-dire $r > \frac{1}{l}$. Ceci montre que $R = \frac{1}{l}$. □

Exemples :

- Pour la série $\sum_{n\geq 0} z^n$ on a $a_n = 1$, et donc $\lim \sqrt[n]{a_n} = 1 = \lim \frac{a_{n+1}}{a_n}$, d'où $R = 1/l = 1$.
- Pour la série $\sum_{m\geq 0} z^{2m}$ on a $a_n = 1$ si $n = 2m$ et $a_n = 0$ si n est impair. La suite $\frac{a_{n+1}}{a_n}$ n'est donc pas définie pour les n impairs ; par ailleurs $\sqrt[n]{a_n}$ est égal à 1 ou 0 selon que n est pair ou impair ; la suite $(\sqrt[n]{a_n})_{n\geq 0}$ n'a donc pas de limite. L'issue, ici, est de poser $u = z^2$ et d'étudier la série $\sum_{m\geq 0} u^m = \sum_{m\geq 0} b_m u^m$, avec $b_m = 1$; comme nous venons de le voir, $R_u = 1$, et donc $\sum_{m\geq 0} z^{2m}$ converge pour $|z|^2 = |z^2| = |u| < 1$ et diverge si $|z|^2 = |z^2| = |u| > 1$, d'où $R_z = \sqrt{1} = 1$.
- Pour la série $\sum_{n\geq 0} \frac{z^n}{n}$ on a $a_n = \frac{1}{n}$, et donc $\lim \frac{a_{n+1}}{a_n} = \lim \frac{n}{n+1} = 1$, d'où $R = 1/l = 1$.
- Pour la série $\sum_{n\geq 0} \frac{z^n}{n\,!}$ on a $a_n = \frac{1}{n\,!}$, et donc $\lim \frac{a_{n+1}}{a_n} = \lim \frac{n\,!}{(n+1)\,!} = \frac{1}{n+1} = 0$, d'où $R = 1/0 = \infty$.

10.3.2 Intégration et dérivation des séries entières

Considérons une série entière $\sum_{n\geq 0} a_n z^n$; posons $f_N(z) = \sum_{n=0}^{N} a_n z^n$. On a

$$f'_N(z) = \sum_{n=1}^{N} n a_n z^{n-1} = \sum_{n=0}^{N-1} d_n z^n, \text{ pour } d_n := (n+1)a_{n+1}, \text{ et}$$

$$F_N(z) := A_0 + \int_0^z f_N(\zeta)d\zeta = A_0 + \sum_{n=0}^{N} \frac{a_n}{n+1} z^{n+1} = \sum_{n=0}^{N+1} A_n z^n, \text{ avec } A_n = \tfrac{a_{n-1}}{n} \text{ pour } n \geq 1.$$

Proposition 10.5 *Les séries entières $\sum_{n\geq 0} a_n z^n$, $\sum_{n\geq 0} d_n z^n$, et $\sum_{n\geq 0} A_n z^n$ ainsi définies ont même rayon de convergence.*

Preuve : Soit R le rayon de convergence de $\sum_{n\geq 0} a_n z^n$. Donc la suite $(a_n r^n)_{n\geq 0}$ tend vers 0 si $r < R$ et est non bornée si $r > R$. Pour $r \neq R$ choisissons r_0 strictement compris entre r et R et posons $q = \frac{r}{r_0}$. Considérons le cas de la série-dérivée $\sum_{n\geq 0} d_n z^n = \sum_{n\geq 0}(n+1)a_{n+1} z^n$. Notons R_d son rayon de convergence. Si $r < R$ (et donc $r < r_0 < R$ et $q < 1$), on a $d_n r^n = (n+1)a_{n+1} r_0^n q^n = a_{n+1} r_0^{n+1} \frac{n+1}{r_0} q^n$. Or la suite $(a_{n+1} r_0^{n+1})_{n\geq 1}$ tend vers 0 puisque $r_0 < R$, et la suite $(\frac{n+1}{r_0} q^n)_{n\geq 1}$ tend vers 0 car $q < 1$. Donc $r \leq R_d$ et cela pour tout $r < R$, donc $R \leq R_d$.

Si $r > R$ (et donc $r > r_0 > R$ et $q > 1$), on a $d_n r^n = (n+1)a_{n+1} r_0^n q^n = a_{n+1} r_0^{n+1} \frac{n+1}{r_0} q^n$. Or la suite $(a_{n+1} r_0^{n+1})_{n\geq 1}$ est non bornée puisque $r_0 > R$, et la suite $(\frac{n+1}{r_0} q^n)_{n\geq 1}$ tend vers $+\infty$ car $q > 1$. Donc $r \geq R_d$ et cela pour tout $r > R$, donc $R \geq R_d$, d'où finalement $R = R_d$.

On procèderait de manière similaire pour la série-primitive $\sum_{n\geq 0} A_n z^n = \sum_{n\geq 0} \frac{a_{n-1}}{n} z^n$. □

Pour $|z| < R$, soient $f(z) = \lim_{n\to\infty} f_N(z) = \sum_{n=0}^{\infty} a_n z^n$, $F(z) = \lim_{n\to\infty} F_N(z) = \sum_{n=0}^{\infty} A_n z^n$, et $g(z) = \lim_{n\to\infty} f'_N(z) = \sum_{n=0}^{\infty} d_n z^n$

Théorème 10.6 *Les fonction f, F, et g ainsi définies sont holomorphes sur le disque ouvert $U = \{z \in \mathbb{C} \mid |z| < R\}$; on a F'(z)=f(z) et $f'(z) = g(z)$. Enfin $a_n = \frac{f^{(n)}(0)}{n\,!}$.*

Preuve : Pour $z = x \in \mathbb{R}$ et $' = \frac{d}{dx}$ (dérivation au sens réel, c'est-à-dire pour des accroissements h réels) ceci est un simple corollaire du théorème 9.4, puisque la proposition 10.3 assure la convergence uniforme sur tout intervalle $[-r, r]$ pour $0 \leq r < R$. Le résultat est toutefois vrai dans la généralité énoncée ici mais sa preuve déborde des ambitions raisonables pour un cours de Deug. □

10.3.3 Séries entières fondamentales

- $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$ avec $R = 1$.
- $\sum_{n=0}^{\infty}(-1)^n z^{2n} = \frac{1}{1+z^2}$ avec $R = 1$; (voir ci-dessus, pour $u = -z^2$.)
- $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} z^n =: \ln(1+z)$ avec $R = 1$; (intégrer $\frac{1}{1+z} = \sum_{n=0}^{\infty}(-1)^n z^n$.)
- $\sum_{n=0}^{\infty} \frac{1}{n\,!} z^n =: \exp(z) = e^z$; (définition de l'exponentielle complexe.)
- $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)\,!} z^{2n+1} =: \sin(z)$; (définition du sinus complexe.)
- $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)\,!} z^{2n} =: \cos(z)$; (définition du cosinus complexe.)

Exercice : Vérifier que $\cos(z) = \frac{e^{iz}+e^{-iz}}{2}$, $\sin(z) = \frac{e^{iz}-e^{-iz}}{2i}$, et donc $e^{i\vartheta} = \cos(\vartheta) + i\sin(\vartheta)$.

Chapitre 11

Fonctions génératrices

11.1 Fonction génératrice des probabilité d'une v.a. entière

Soit $X \rightsquigarrow \mathcal{B}(n,p)$ une v.a. binômiale ; on a $p_X(k) := \mathbb{P}\{X = k\} = \begin{pmatrix} n \\ p \end{pmatrix} p^k(1-p)^{n-k}$. On observe que les nombres $p_X(k)$ sont précisément les coefficients du polynôme

$$G_X(s) := ((1-p)+ps)^n = \sum_{k=0}^{n} \begin{pmatrix} n \\ p \end{pmatrix} p^k(1-p)^{n-k}s^k = \sum_{k=0}^{n} p_X(k)s^k = \mathbb{E}(s^X).$$

Cette dernière expression, $\mathbb{E}(s^X)$, se prète à la généralisation :

Définition : Soit X une v.a. entière. On appelle *fonction génératrice des probabilités* (fgp) de X, et on note G_X, la fonction

$$G_X(s) := \mathbb{E}(s^X) = \sum_{k=0}^{\infty} p_X(k)s^k.$$

Remarques :

- Comme $\sum_{k=0}^{\infty} p_X(k) = 1$, le rayon de convergence R_G de la série entière $\sum_{k=0}^{\infty} p_X(k)s^k$ est au moins égal à 1.
- Si X et Y ont même fgp, elles ont même loi, puisque

$$\mathbb{P}\{X = k\} = p_X(k) = \frac{G_X^{(k)}(0)}{k\ !} = \frac{G_Y^{(k)}(0)}{k\ !} = p_Y(k) = \mathbb{P}\{y = k\}.$$

Théorème 11.1 *Si* $\mathbb{E}X^r$*, le moment d'ordre* r *de la v.a.* X *existe, et la dérivé* $G_X^{(r)}$ *de la fgp de* X *vérifie*

$$G_X^{(r)}(1) = \mathbb{E}[X(X-1)\ldots(X-r+1)], \tag{11.1}$$

où, lorsque $R = 1$*,* $G^{(r)}(1)$ *désigne* $\lim_{s\to 1^-} G^{(r)}(s) = \sum_{k=0}^{\infty} \frac{d^r}{ds^s}(p(k)s^k)$.

Comme on sait seulement que $R_G \geq 1$, on est assuré de la dérivabilité de G_X que pour $|s| < 1$. Comme $G_X(1)$ est toutefois bien défini, le résultat suivant s'applique ; il servira dans la preuve du théorème 11.1.

Lemme 11.2 (Abel) *Si la série entière de* $f(x) := \sum_{n=0}^{\infty} a_n x^n$ *converge pour tout* $|x| < 1$*, et si* $\sum_{n=0}^{\infty} a_n$ *converge, alors* $\lim_{x\to 1^-} f(x) = \sum_{n=0}^{\infty} a_n = f(1)$

Preuve : Comme les dérivées terme à terme d'une série entière ont le même rayon de convergence que cette série entière, le rayon de convergence de la série de $G^{(r)}(s)$ pour $|s| < R$ est lui aussi, au moins égal à 1. Comme $\mathbb{E}X^r$ existe, il en est de même de $\mathbb{E}X^s$ pour $0 \leq s \leq r$, et donc $\mathbb{E}[X(X-1)\ldots(X-r+1)]$ existe bien. Montrons le théorème pour $r = 1$; le cas général se montrerait de même par récurrence. Pour $|s| < 1$, on a, en dérivant terme à terme,

$$G_X'(s) = \sum_{k=0}^{\infty} \frac{d}{ds}(p(k)s^k) = \sum_{k=0}^{\infty} kp(k)s^{k-1},$$

d'où, par le lemme d'Abel 11.2, $\lim_{s\to 1^-} G_X'(s) = \sum_{k=0}^{\infty} kp(k) = \mathbb{E}X$. $\square$

Exemples :

- *Fonction génératrice des probablités d'une v.a. de Poisson $X \rightsquigarrow \mathcal{P}(\lambda)$:*
  On a $\mathbb{P}\{X=k\} = e^{-\lambda}\frac{\lambda^k}{k\,!}$, donc
  $$G_X(s) = \mathbb{E}s^X = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k\,!}s^k = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda s)^k}{k\,!} = e^{-\lambda}e^{\lambda s} = e^{\lambda(s-1)}.$$
  D'où $G'_X(s) = \lambda e^{-\lambda(s-1)}$, et donc $\mathbb{E}X = G'_X(1) = \lambda$. Plus généralement, $G_X^{(r)}(s) = \lambda^r e^{-\lambda(s-1)}$, et donc $\mathbb{E}[X(X-1)\dots(X-r+1)] = G_X^{(r)}(1) = \lambda^r$. En particulier $\lambda^2 = \mathbb{E}[X(X-1)] = \mathbb{E}X^2 - \mathbb{E}X$, donc $\mathbb{E}X^2 = \lambda^2+\lambda$, d'où $\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \lambda^2+\lambda-\lambda^2 = \lambda$.

- *Fonction génératrice des probablités d'une v.a. géométrique $X \rightsquigarrow \mathcal{G}(\alpha)$:*
  On a $\mathbb{P}\{X=k\} = (1-\alpha)\alpha^{k-1}$, $k=1,2,\dots$, donc
  $$G_X(s) = \mathbb{E}s^X = \sum_{k=1}^{\infty}(1-\alpha)\alpha^{k-1}s^k = \frac{1-\alpha}{\alpha} = \sum_{k=1}^{\infty}(\alpha s)^k = (1-\alpha)\frac{s}{1-\alpha s},$$
  donc $G'_X(s) = \frac{1-\alpha}{(1-\alpha s)^2}$, et $\mathbb{E}X = G'_X(1) = \frac{1-\alpha}{(1-\alpha)^2} = \frac{1}{1-\alpha}$; par ailleurs
  $$\mathbb{E}X^2 - \mathbb{E}X = G''_X(s) = (1-\alpha)\frac{+2\alpha}{(1-\alpha)^3} = \frac{2\alpha}{(1-\alpha)^2},$$
  d'où $\mathbb{E}X^2 = \frac{2\alpha}{(1-\alpha)^2} + \frac{1}{1-\alpha} = \frac{1+\alpha}{(1-\alpha)^2}$, et
  $$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1+\alpha}{(1-\alpha)^2} - \frac{1}{(1-\alpha)^2} = \frac{\alpha}{(1-\alpha)^2}.$$

Proposition 11.3 *Si X et Y sont deux v.a. indépendantes, $G_{X+Y}(s) = G_X(s)G_Y(s)$.*

Preuve : Comme X et Y sont indépendantes, il en est de même pour les v.a. s^X et s^Y, d'où
$$G_{X+Y}(s) = \mathbb{E}[s^{X+Y}] = \mathbb{E}[s^X s^Y] = \mathbb{E}[s^X]\mathbb{E}[s^Y] = G_X(s)G_Y(s).$$
□

Exemples :

- *V.a. de Bernoulli :* on sait que la loi binômiale est la loi de la somme de n v.a de Bernoulli indépendantes : ceci explique que la fgp d'une v.a. de Bernouli, $G(s) = ((1-p)+ps)^n$ est la puissance n-ième de la fgp $(1-p)+ps$ d'une v.a. de Bernoulli.

- *Somme aléatoire de v.a. i.i.d. :* Soit N, X_1, ..., X_k, ...des v.a. entières indépendantes, les X_k étant identiquement distribuées. Notons respectivement G_N et G_X les fgp de N et des X_k. Soit $T := \sum_{k=1}^N X_k$ la somme aléatoire des X_k (un modèle possible pour le coût annuel des sinistres pour une compagnie d'assurance : N désigne le nombre de sinistres par an, et X_k représente le coût du k-ième). Alors $G_T = G_N \circ G_X$. En effet, $G_T(s) = \mathbb{E}(\mathbb{E}(s^T|N)) = \mathbb{E}g(N)$, avec
  $$g(n) := \mathbb{E}(s^T|N=n) = \mathbb{E}\left(s^{\sum_{k=1}^n X_k}|N=n\right) \overset{\sharp}{=} \mathbb{E}\left(s^{\sum_{k=1}^n X_k}\right) \overset{\natural}{=} \prod_{k=1}^n \mathbb{E}(s^{X_k}) = (G_X(s))^n,$$
  où l'égalité $\sharp$ résulte de l'indépendance de $X_1,\dots,X_n$ de N, et l'égalité $\natural$ résulte de l'indépendance mutuelle des $X_1,\dots,X_n$. En posant $y := G_X(s)$, on a donc $G_T(s) = \mathbb{E}y^N = G_N(y) = G_N \circ G_X(s)$.

Théorème 11.4 (admis) *Soient X, X_1, X_2, ... une suite de v.a.. Si, pour tout $|s| \le 1$, $\lim_{n\to\infty} G_{X_n}(s) = G_X(s)$, alors la suite des X_n* tend en loi *vers X, c'est-à-dire que*
$$\lim_{n\to\infty}\mathbb{P}\{X_n = k\} = \mathbb{P}\{X_n = k\} \textit{ pour tout } k.$$

Dans l'exemple $Y_n \rightsquigarrow \mathcal{U}(1..n)$ (loi *discrète* uniforme sur $1..n$) on devine que si X est la limite des $X_n = \frac{1}{n}Y_n$, X suivra une loi *continue* uniforme $\mathcal{U}[0,1]$. Nous souhaitons donc une notion de fonction génératrice qui s'applique aussi au cas de v.a. absolument continues.

11.2 Fonction génératrice des moments (v.a. continue ou discrète)

Soit $X \rightsquigarrow \mathcal{E}(\lambda)$ une v.a. suivant une loi exponentielle. Son k-ième moment $\mu_X(k) = \mathbb{E}X^k$, se calcule facilement au moyen de k intégrations par parties et est donc égal à

$$\mu_X(k) = \mathbb{E}X^k = \int_0^\infty x^k \lambda e^{-\lambda x} dx = \frac{k\,!}{\lambda^k}.$$

On observe que les nombres $\frac{1}{k\,!}\mu_X(k)$ sont précisément les coefficients du développement en série entière de la fonction

$$\begin{aligned} M_X(t) &= \frac{\lambda}{\lambda - t} = \frac{1}{1-(\frac{t}{\lambda})} = \sum_{k=0}^\infty \frac{t^k}{\lambda^k} = \sum_{k=0}^\infty \frac{\mu_X(k)}{k\,!} t^k \\ &= \sum_{k=0}^\infty \frac{\mathbb{E}X^r}{k\,!} t^k = \sum_{k=0}^\infty \mathbb{E}\left[\frac{X^r}{k\,!}t^k\right] \stackrel{\flat}{=} \mathbb{E}\left[\sum_{k=0}^\infty \frac{X^r}{k\,!}t^k\right] = \mathbb{E}e^{tX}, \end{aligned}$$

où l'égalité surmontée d'un bémol $\flat$ impose la circonspection usuelle vis-à-vis du caractère infini de la somme considérée. La dernière expression, $\mathbb{E}e^{tX}$, se prête à la généralisation :

Définition : On appelle *fonction génératrice des moments* de X la fonction M_X définie par

$$M_X(t) := \mathbb{E}e^{tX}$$

Théorème 11.5 *Supposons que tous les moments $\mu_X(k) := \mathbb{E}X^k$ d'une v.a. X soient définis, et que la série $\sum_{k=0}^\infty \frac{\mu_X(k)}{k\,!}t^k$ converge, avec un rayon de convergence R non nul. Alors, pour $|t| < R$, $\mathbb{E}e^{tX}$ existe et est égal à $\sum_{k=0}^\infty \frac{\mathbb{E}X^r}{k\,!}t^k$, et*

$$M_X(t) = \mathbb{E}e^{tX} = \sum_{k=0}^\infty \frac{\mu_X(k)}{k\,!}t^k \ , \ (|t| < R).$$

Remarques :

- Nous admettons ce théorème et retenons sa conclusion sous la forme que $\mathbb{E}e^{tX} =: M_X(t) = \sum_{k=0}^\infty \frac{\mathbb{E}X^r}{k\,!}t^k$ pour $|t| < R$. Sa preuve est simple dès lors qu'on dispose de la théorie de Lebesgue de l'intégrale.
- Si X est à valeurs entières positives, en posant $s = e^t$, nous voyons que

  $$M_X(t) = \mathbb{E}e^{tX} = \mathbb{E}(e^t)^X = \mathbb{E}s^X = G_X(s) = G_X(e^t).$$

  Cette dernière expression est définie pour $|s| \le 1$, donc $M_X(t)$ est définie pour $t \le 0$ (au moins) ; les hypothèses sur la convergence de la série formée à partir des moments n'est utile que pour assurer une représentation de M_X par une série en puissances entières de t.
- Si X est absolument continue, de fonction de densité de probabilité (fdp) f_X, nous avons

  $$M_X(t) = \mathbb{E}e^{tX} = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx.$$

  L'opération fonctionnelle $f_X \mapsto M_X$ associant à une fonction $(x \mapsto f(x))$ la fonction $(t \mapsto M(t) = \int_{-\infty}^{+\infty} f(t)e^{tx}dx)$ s'appelle la *transformation de Laplace (bilatère)* et est utilisée dans biens des branches des Mathématiques, allant de méthodes pour l'ingénieur aux théories les plus récentes de sommation des séries divergentes qui apparaissent notamment en physique théorique. Nous l'étudierons plus en détails plus loin.
- Par la formule de Taylor nous savons que le coefficient de t^k dans la série entière de M_X n'est autre que la valeur de $\frac{M_X^{(k)}(0)}{k\,!}$, donc $\mu_X(k) = M^{(k)}(0)$, d'où

  $$\mathbb{E}X^k = M_X^{(k)}(0). \tag{11.2}$$

Exemples :

- *Fonction génératrice des moments d'une v.a. uniforme* continue $X \rightsquigarrow \mathcal{U}[a,b]$:

  On a $f_X(x) = \frac{1}{b-a}\mathbb{I}_{[a,b]}$, d'où $\mu_X(k) = \int_a^b \frac{x^k}{b-a}dx = \frac{b^{k+1}-a^{k+1}}{b-a}$, une expression guère commode. De fait

$$M_X(t) = \mathbb{E}e^{tX} = \int e^{tx}\frac{1}{b-a}\mathbb{I}_{[a,b]}(x)dx = \frac{1}{b-a}\int_a^b e^{tx}dx = \frac{1}{b-a}\left[\frac{1}{t}e^{tx}\right]_{x=a}^{b} = \frac{e^{bt}-e^{at}}{(b-a)t}.$$

- *Fonction génératrice des moments d'une v.a. uniforme* discrete $X_n \rightsquigarrow \mathcal{U}]0..1]_{\frac{1}{n}} = \frac{1}{n}Y_n$ avec $Y_n \rightsquigarrow \mathcal{U}]1..n]$:

  On a $p_{X_n}(\frac{i}{n}) = \frac{1}{n}$, pour $i = 1..n$, d'où

$$M_{X_n}(t) = \mathbb{E}e^{tX} = \sum_{i=1}^{n}\frac{1}{n}e^{t\frac{i}{n}} = \frac{1}{n}\sum_{i=1}^{n}\left(e^{\frac{t}{n}}\right)^i = \frac{1}{n}e^{\frac{t}{n}}\frac{\left(e^{\frac{t}{n}}\right)^n - 1}{e^{\frac{t}{n}}-1} = \frac{e^t-1}{n(1-e^{-\frac{t}{n}})}.$$

- *Fonction génératrice des moments d'une v.a. exponentielle* $X \rightsquigarrow \mathcal{E}(\lambda)$:

  On a $f_X(x) = \lambda e^{-\lambda x}\mathbb{I}_{[0,+\infty[}(x)$. Là encore, inutile de calculer les moments. Pour $t < \lambda$ on a

$$M_X(t) = \mathbb{E}e^{tX} = \int e^{tx}\lambda e^{-\lambda x}\mathbb{I}_{[0,+\infty[}(x) = \int_0^{+\infty}\lambda e^{(t-\lambda)x}dx = \frac{\lambda}{t-\lambda}\left[e^{(t-\lambda)x}\right]_{x=0}^{+\infty} = \frac{\lambda}{\lambda - t}.$$

- *Fonction génératrice des moments d'une v.a. normale* $X \rightsquigarrow \mathcal{N}(0,1)$:

  On a $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, d'où

$$M_X(t) = \mathbb{E}e^{tX} = \int_{-\infty}^{+\infty} e^{tx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}+tx}dx.$$

  C'est une intégrale gaussienne ; il convient de mettre l'exposant de l'intégrand sous la forme $-\frac{v^2}{2}+c$. On a

$$-\frac{x^2}{2}+tx = -\frac{1}{2}(x^2-2tx) = -\frac{1}{2}[(x-t)^2-t^2] = -\frac{1}{2}(x-t)^2+\frac{t^2}{2} = -\frac{v^2}{2}+\frac{t^2}{2}$$

  pour $v = x-t$, et finalement

$$M_X(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}+tx}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{v^2}{2}+\frac{t^2}{2}}dv = e^{+\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{v^2}{2}}dv = e^{+\frac{t^2}{2}}.$$

  On pourrait procéder de manière similaire pour $X \rightsquigarrow \mathcal{N}(\mu,\sigma)$ et obtenir

$$M_X(t) = e^{\mu t}e^{\frac{1}{2}\sigma^2 t^2}.$$

  Les calculs sont toutefois plus laborieux, et seront bien plus aisés avec le recul des propriétés de la transformation de Laplace. Notons qu'on trouve $M_X'(t) = (\mu+\sigma^2 t)M_X(t)$ et $M_X''(t) = (\sigma^2+(\mu+\sigma^2 t)^2)M_X(t)$, d'où $M_X'(0) = \mu$ et $M_X''(0) = \sigma^2+\mu^2$, d'où $\mathbb{E}X = M_X'(0) = \mu$, $\mathbb{E}X^2 = M_X''(0) = \sigma^2+\mu^2$, et $\operatorname{Var} X = \sigma^2+\mu^2-\mu^2 = \sigma^2$. On retrouve bien l'espérance et la variance d'une v.a. suivant une loi $\mathcal{N}(\mu,\sigma)$.

En raisonnant comme dans le cas de la fonction génératrice des probabilités on voit qu'on a

Proposition 11.6 *Si X et Y sont deux v.a. indépendantes, $M_{X+Y}(t) = M_X(t)M_Y(t)$.*

Théorème 11.7 (admis) *Soient $X, X_1, X_2, \ldots$ une suite de v.a.. Si, pour tout $|t| \leq R$, $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$, alors la suite des X_n* tend en loi *vers X, c'est-à-dire que*

$$\lim_{n\to\infty} M_{X_n}(t) = M_X(t) \textit{pour tout } t \textit{ implique que } \lim_{n\to\infty} F_{X_n}(x) = F_X(x) \textit{ pour tout } x,$$

où $F_X(x) := \mathbb{P}\{X \leq x\}$ et $F_{X_n}(x) := \mathbb{P}\{X_n \leq x\}$ désignent les fonctions caractérisques des v.a. considérées. En particulier, si pour deux v.a. X et Y on a $M_X = M_Y$, alors X et Y ont même loi.

Reprenons l'exemple $Y_n \rightsquigarrow \mathcal{U}[1..n]$ (loi *discrète* uniforme sur $1..n$) et supposons que X est la limite des $X_n = \frac{1}{n}Y_n$. Nous avons vu que $M_{X_n}(t) = \frac{e^t-1}{n(1-e^{-\frac{t}{n}})}$. Un développement limité à l'ordre 1 du dénominateur $n(1-e^{-\frac{t}{n}}) = n(1-(1-\frac{t}{n}+\ldots))$ montre que $\lim_n M_{X_n}(t) = \frac{e^t-1}{t}$, la fonction génératrice des moments d'une v.a. continue uniforme sur $[0,1]$, ce qui montre que $X \rightsquigarrow \mathcal{U}[0,1]$.

Chapitre 12

Transformation de Laplace

12.1 Transformation de Laplace bilataire

Pour $p = \alpha + i\omega \in \mathbb{C}$, $x \in \mathbb{R}$, et $f(x) \in \mathbb{R}$, f continue par morceaux, on a

$$g(p\,;x) := e^{-px}f(x) = e^{-\alpha x}e^{-i\omega x}f(x) = e^{-\alpha x}f(x)(\cos x - i\sin x) =: g_1(p\,;x) + ig_2(p\,;x),$$

où les $g_k(p\,;x)$, $k = 1, 2$, sont des fonctions réelles. On peut donc définir

$$\int_{-\infty}^{+\infty} g(p\,;x)dx := \int_{-\infty}^{+\infty} g_1(p\,;x)dx + i\int_{-\infty}^{+\infty} g_2(p\,;x)dx.$$

Bien entendu, une question d'intégrabilité des fonctions g_k se pose ; pour cela, observons que pour $k = 1, 2$, on a

$$|g_k(p\,;x)| \leq |g(p\,;x)| = |e^{-px}f(x)| = e^{-\alpha x}|f(x)| \text{ , puisque } |e^{-i\omega x}| = |\cos(\omega x) - i\sin(\omega x)| = 1.$$

Supposons que f soit de *type exponentiel* tant en $x = +\infty$ que $x = -\infty$, c'est-à-dire qu'il existe $C \geq 1$ et $\alpha_- < \alpha_+$ tels que

$$\begin{aligned} |f(x)| &\leq Ce^{\alpha_- x} \text{ pour } x > 0 \text{ (type exponentiel en } x = +\infty\text{)},\\ |f(x)| &\leq Ce^{\alpha_+ x} \text{ pour } x < 0 \text{ (type exponentiel en } x = -\infty\text{) ;} \end{aligned}$$

observez que α_- va avec $x > 0$, et α_+ va avec $x < 0$. On a donc $|g_k(p\,;x)| \leq Ce^{(\alpha_- - \alpha)x}$ pour $x > 0$, et $|g_k(p\,;x) \leq Ce^{(\alpha_+ - \alpha)x}$ pour $x < 0$; pour assurer l'intégrabilité des g_k en $x = +\infty$ il suffit donc d'avoir $\alpha_- - \alpha > 0$, et pour l'intégrabilité des g_k en $x = -\infty$ il suffit d'avoir $\alpha_+ - \alpha > 0$. Les deux intégrales sont donc bien définies pourvu que

$$\alpha_- < \alpha := \mathrm{Re}\,(p) < \alpha_+.$$

Définition : Pour f continue par morceaux et de type exponentiel, on appelle *transformée de Laplace bilataire* de f et on note $F := L[f]$ la fonction

$$p \mapsto F(p) = L[f](p) := \int_{-\infty}^{+\infty} e^{-px}f(x)dx$$

définie pour $\alpha_- < \mathrm{Re}\,(p) < \alpha_+$. Pour exprimer que F est la transformée de Laplace bilataire de f nous écrirons

$$f(x) \sqsupset F(p).$$

Exemple : Si f_X est la fonction de densité d'un v.a. X, et M_X sa fgm, on a

$$M_X(t) = L[f](-t).$$

12.1.1 Quelques propriétés de la transformation de Laplace bilataire

On notera systématiquement $F = L[f]$.

- linéarité : comme $\int_{-\infty}^{+\infty} e^{-px}(\lambda f(x) + \mu g(x))dx = \lambda \int_{-\infty}^{+\infty} e^{-px} f(x)dx + \mu \int_{-\infty}^{+\infty} g(x))dx$, on a
$$L[\lambda f + \mu g] = \lambda L[f] + \mu L[g].$$

- $\boxed{f(ax) \sqsupset \frac{1}{|a|} F\left(\frac{p}{a}\right), \text{ pour } a \neq 0 \text{ réel.}}$
En effet, en posant $y = ax$ dans l'intégrale ci-dessous, on a, si $a > 0$,
$$\int_{-\infty}^{+\infty} e^{-px} f(ax)dx = \int_{-\infty}^{+\infty} e^{-\frac{py}{a}} \frac{1}{a} dy = \frac{1}{a} \int_{-\infty}^{+\infty} e^{-\left(\frac{p}{a}\right)y} dy = \frac{1}{|a|} F\left(\frac{p}{a}\right).$$
Si $a < 0$, le même changement de variable $x = ay$ interverti les bornes d'intégration, et en observant que $-a = |a|$, on a ici
$$\int_{-\infty}^{+\infty} e^{-px} f(ax)dx = \int_{+\infty}^{-\infty} e^{-\frac{py}{a}} \frac{1}{a} dy = \frac{1}{-a} \int_{-\infty}^{+\infty} e^{-\left(\frac{p}{a}\right)y} dy = \frac{1}{|a|} F\left(\frac{p}{a}\right).$$

- $\boxed{f(x-a) \sqsupset e^{-ap} F(p).}$
En effet, il suffit cette fois d'effectuer le changement de variable $y = x - a$:
$$\int_{-\infty}^{+\infty} e^{-px} f(x-a)dx = \int_{-\infty}^{+\infty} e^{-p(y+a)} f(y)dy = \int_{-\infty}^{+\infty} e^{-py-pa} f(y)dy = e^{-pa} \int_{-\infty}^{+\infty} e^{-py} f(y)dy = e^{-pa} F(p)$$

- $\boxed{e^{-ax} f(x) \sqsupset F(p+a).}$
En effet,
$$\int_{-\infty}^{+\infty} e^{-px} e^{-ax} f(x)dx = \int_{-\infty}^{+\infty} e^{-(p+a)x} f(x)dx = F(p+a).$$

- $\boxed{\mathcal{H}(x) \sqsupset \frac{1}{p}, \text{ pour } \mathrm{Re}\,(p) > 0,}$ où $x \mapsto \mathcal{H}(x)$ est la *fonction de Heaviside*, définie par $\mathcal{H}(x) = 1$ si $x > 0$ et $\mathcal{H}(x) = 0$ si $x < 0$. En effet
$$\int_{-\infty}^{+\infty} e^{-px} \mathcal{H}(x)dx = \int_{0}^{+\infty} e^{-px} dx = \left[\frac{-1}{p} e^{-px}\right]_{x=0}^{x=+\infty} = -\frac{-1}{p} = \frac{1}{p},$$
puisque $\lim_{x\to+\infty} |e^{-px}| = \lim_{x\to+\infty} |e^{-\alpha x}||e^{-i\omega x}| = \lim_{x\to+\infty} e^{-\alpha x} = 0$, si $\alpha := \mathrm{Re}\,(p) > 0$, puisque $|e^{-i\omega x}| = 1$.

12.1.2 Application : calcul de la fgm d'une v.a. X normale, $X \rightsquigarrow \mathcal{N}(\mu, \sigma)$.

Soit $f(x) := f_X(x)$. On a $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{\sigma} - \frac{\mu}{\sigma}\right)^2} = \frac{1}{\sigma} g_{\frac{\mu}{\sigma}}\left(\frac{x}{\sigma}\right)$, où
$$g_\lambda(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-\lambda)^2}{2}} = h(y-\lambda) \text{ , avec } h(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$
Nous avons calculé au chapitre 11 la fgm d'une v.a. normale, centrée, et réduite, et donc
$$h(x) \sqsupset H(p) = e^{+\frac{(-p)^2}{2}} = e^{+\frac{p^2}{2}} \text{ , d'où}$$
$$g_\lambda(y) = h(y-\lambda) \sqsupset G_\lambda(p) = e^{-\lambda p} H(p) = e^{-\lambda p} e^{+\frac{p^2}{2}}.$$
Posons $\lambda := \frac{\mu}{\sigma}$, et $a = \frac{1}{\sigma} > 0$; on a
$$f(x) = \frac{1}{\sigma} g_\lambda\left(\frac{x}{\sigma}\right) \sqsupset F(p) = \frac{1}{\sigma}\frac{1}{a} G_\lambda\left(\frac{p}{a}\right) = e^{-\lambda\frac{p}{a}} e^{\frac{1}{2}\left(\frac{p}{a}\right)^2} = e^{-\mu p} e^{\frac{\sigma^2 p^2}{2}}.$$
Finalement, $M_X(t) = L[f](-t) = e^{\mu t} e^{\frac{\sigma^2 t^2}{2}}$, comme annoncé au chapitre 11.

12.2 Transformation de Laplace (unilataire)

Le fait d'avoir à demander que f soit de type exponentiel aussi bien en $x = +\infty$ qu'en $x = -\infty$ est très contraignant. Ainsi, si l'on considère $f(x) = x$, il est nécessaire, pour $x = +\infty$, de choisir $\alpha_- > 0$, et pour $x = -\infty$ il faut choisir $\alpha_+ < 0$, et la condition $\alpha_- < \alpha_+$ ne peut donc plus être satisfaite. Si on peut ne s'intéresser à $f(x)$ que pour $x \geq 0$, on a le recours suivant : considérer la fonction $f_+(x) := \mathcal{H}(x)f(x)$, qui coïncide avec f pour $x \geq 0$, et est nulle pour $x < 0$, ce qui permet de choisir $\alpha_+ = +\infty$ (ce qui signifie une valeur de α_+ aussi grande que souhaitée), puisque $0 \leq e^{\alpha_+ x}$, pour n'importe quel α_+.

Définition : Pour toute fonction f continue par morceaux, de type exponentiel en $x = +\infty$, on appelle *transformée de Laplace* (unilataire) de f et on note $\mathcal{L}[f]$ la transformée de Laplace bilataire de $\mathcal{H}(x)f(x)$, c'est-à-dire la fonction

$$p \mapsto \mathcal{L}[f](p) = \int_0^{+\infty} e^{-px} f(x)dx.$$

Exemple : Si f_X est la fonction de densité de probabilité d'une v.a. *positive* X, on a $L[f] = \mathcal{L}[f]$, et

$$M_X(t) = \mathcal{L}[f](-t).$$

La transformation de Laplace unilataire est définie pour la plupart des fonctions usuelles ; c'est également pour cette version de la transformation de Laplace que nous verrons l'un de ses domaines d'application : remplacer les opérations de dérivation et intégration par des opérations algébriques.

12.2.1 Quelques relations relatives de la transformation de Laplace

La plupart des résultats suivants s'obtiennent par une intégration par parties d'expressions du type $\int u'v$ avec e^{-px} tantôt égal à u' tantôt à v.

- $\boxed{\mathcal{L}(x \mapsto x)(p) = \frac{1}{p^2}, \text{ pour } \mathrm{Re}\,(p) > 0.}$

  En effet, en pratiquant une intégration par parties, en posant $u' = e^{-px}$, on a

$$\int_0^{+\infty} e^{-px} x dx = \left[\frac{-1}{p} e^{-px} x\right]_{x=0}^{+\infty} - \int_0^{+\infty} \frac{-1}{p} e^{-px} x dx = 0 - 0 + \frac{1}{p}\left[\frac{-1}{p} e^{-px} x\right]_{x=0}^{+\infty} = \frac{1}{p^2}.$$

- $\boxed{\mathcal{L}(x \mapsto x^n)(p) = \frac{n\,!}{p^{n+1}}, \text{ pour } \mathrm{Re}\,(p) > 0.}$

  En effet, en pratiquant une intégration par parties, en posant une nouvelle fois $u' = e^{-px}$, on a

$$\mathcal{L}(x \mapsto x^n)(p) = \int_0^{+\infty} e^{-px} x^n dx = \left[\frac{-1}{p} e^{-px} x^n\right]_{x=0}^{+\infty} - \int_0^{+\infty} \frac{-n}{p} e^{-px} x^{n-1} dx = 0 - 0 + \frac{n}{p}\mathcal{L}(x \mapsto x^{n-1})(p)\,;$$

  le résultat annoncé s'en déduit immédiatement par récurrence.

- $\boxed{\mathcal{L}(x \mapsto e^{-ax})(p) = \frac{1}{p+a}, \text{ pour } \mathrm{Re}\,(p) > -a.}$

  En effet, on a

$$\int_0^{+\infty} e^{-px} e^{-ax} dx = \int_0^{+\infty} e^{-(p+a)x} dx = \left[\frac{-1}{p+a} e^{-(p+a)x}\right]_{x=0}^{+\infty} = 0 - \frac{-1}{p+a} = \frac{1}{p+a}.$$

- $\boxed{\mathcal{L}(x \mapsto \cos(\omega x))(p) = \frac{p}{p^2+\omega^2}, \text{ pour } \mathrm{Re}\,(p) > 0.}$

- $\boxed{\mathcal{L}(x \mapsto \sin(\omega x))(p) = \frac{\omega}{p^2+\omega^2}, \text{ pour } \mathrm{Re}\,(p) > 0.}$

  Ces deux résultats se démontrent de pair, par une intégrations par partie, en posant $v = e^{-px}$. Nous utiliserons les notations $C(p) := \mathcal{L}(x \mapsto \cos(\omega x))(p)$ et $S(p) := \mathcal{L}(x \mapsto \sin(\omega x))(p)$. On a

$$C(p)=\int_0^{+\infty} e^{-px}\cos(\omega x)dx=\left[e^{-px}\frac{1}{\omega}\sin(\omega x)\right]_{x=0}^{+\infty}-\int_0^{+\infty}-pe^{-px}\frac{1}{\omega}\sin(\omega x)dx=+\frac{p}{\omega}S(p)\ \text{, et}$$

$$S(p)=\int_0^{+\infty} e^{-px}\sin(\omega x)dx=\left[e^{-px}\frac{-1}{\omega}\cos(\omega x)\right]_{x=0}^{+\infty}-\int_0^{+\infty}-pe^{-px}\frac{-1}{\omega}\cos(\omega x)dx=\frac{+1}{\omega}-\frac{p}{\omega}C(p).$$

En d'autres termes, $C(p)$ et $S(p)$ forment la solution du système linéaire

$$\left\{\begin{array}{lcl} C(p)-\frac{p}{\omega}S(p) & = & 0 \\ \frac{p}{\omega}C(p)+S(p) & = & \frac{1}{\omega}, \end{array}\right. \tag{12.1}$$

qui se résoud facilement par combinaison, et on obtient le résultat annoncé.

- $\boxed{\mathcal{L}[f'](p)=p\mathcal{L}[f]-f(0^+).}$

  En effet, en posant $v=e^{-px}$, on a

$$\int_0^{+\infty} e^{-px}f'(x)dx=[e^{-px}f(x)]_{x=0}^{+\infty}-\int_0^{+\infty}-pe^{-px}f'(x)dx=f(0)+p\mathcal{L}[f](p).$$

- $\boxed{\mathcal{L}[\int_0^x f(\xi)d\xi](p)=\frac{1}{p}\mathcal{L}[f](p),}$ pourvu que $\lim_{x\to+\infty} e^{-\alpha x}\int_0^x f(\xi)d\xi=0$, avec $\alpha=\mathrm{Re}\,(p)$.

  En effet, en posant cette fois $u'=e^{-px}$, on a

$$\int_0^{+\infty} e^{-px}\int_0^x f'(\xi)d\xi=[\frac{-1}{p}e^{-px}\int_0^x f'(\xi)d\xi]_{x=0}^{+\infty}-\int_0^{+\infty}\frac{-1}{p}e^{-px}f(x)dx=+\frac{1}{p}\mathcal{L}[f](p).$$

- $\boxed{\mathcal{L}[f*g]=\mathcal{L}[f]\mathcal{L}[g],}$ où $f*g(x):=\int_0^{+\infty} f(t)g(x-t)dt$; nous reviendrons plus tard sur le *produit de convolution* $f*g$, et prouverons alors cette identité.

Le résultat suivant est à la base de la résolution de problèmes par transformation de Laplace ; il permet la démarche générale suivante :

1. chercher la relation satisfaite par la transformée de Laplace F de la solution f du problème posé,
2. déterminer la solution F du problème posé par cette nouvelle relation,
3. trouver la fonction f ayant pour transformée de Laplace la fonction F.

Dans les cas élémentaires, ces trois étapes se réduisent à de simples calculs algébriques.

Théorème 12.1 (admis) *Soient f_1 et f_2 continues par morceaux, de type exponentiel en $+\infty$, et si $\mathcal{L}[f_1]=\mathcal{L}[f_2]$, alors $f_1(x)=f_2(x)$ en tout point $x\geq 0$ où f_1 et f_2 sont continues.*

Exemple : Soit à trouver la solution de $y'(x)=ay(x)$, une équation différentielle ordinaire que vous connaissez bien. Posons $y(x)\sqsupset Y(p)$. On a donc

$$pY(p)-y(0)=\mathcal{L}[y'](p)=\mathcal{L}[ay(x)](p)=aY(p),$$

d'où $Y(p)=\frac{y(0)}{p-a}$. Or $y(0)e^{ax}\sqsupset\frac{y(0)}{p-a}$; donc $y(x)=y(0)e^{ax}$ pour $x\geq 0$, puisque le théorème nous assure que ces deux fonctions continues, qui ont même transformées de Laplace, sont nécessairement égales. Vous savez que la restriction à $x\geq 0$ est inutile ; elle résulte ici de l'usage de la transformée de Laplace unilataire qui, en "écrasant" les valeurs de y pour $x<0$ empèche un théorème dont la conclusion s'applique aussi pour $x<0$.

Chapitre 13

Le théorème central-limit

13.1 Une motivation issue de la finance

Chapitre 14

Séries de Fourier

Ce cours est consacré à une représentation particulièrment commode des fonctions définies sur $\mathbb{R}$ à valeurs réelles ou complexes, et 2π-périodiques. Elle permet de coder ou résumer une telle fonction par deux suites de nombres $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 1}$, voire en une seule suite $(c_n)_{n\in\mathbb{Z}}$. Le calcul avec ces codes est simple et la connaissance des premiers termes des suites permet une bonne approximation des fonctions qu'ils résument. Ces représentations ont été introduites par Joseph Fourier, notamment pour la résolution de problèmes de Physique (l'équation de la chaleur), et il n'est guère aujourd'hui de théorie physique qui n'y fasse appel. Elles ont une nouvelle fois été popularisées au mileu du XXème siècle par Norbert Wiener pour la théorie du signal où elles se sont mèlées au calcul de probabilités ; c'est par ce canal qu'elles sont devenues un outil précieux pour l'économètre, avec les travaux de Box et Jenkins, dans les années 70, sur les séries temporelles. C'est à l'initiation aux méthodes issues du mariage du calcul de Fourier et du calcul des probabilités que sera consacré la fin du cours d'Analyse et Probablités du Deug Mass.[1]

Nous commençons par la notion de séries trigonométriques, puis nous découvrirons progressivement pourquoi ces séries particulières sont en fait toutes les fonctions 2π-périodiques quelque peu régulières. Nous terminons par la version complexe de la théorie, qui est la plus commode à manipuler.

14.1 Séries trigonométriques

Définition : Soient $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 1}$ deux suites numériques. On appelle *série trigonométrique* toute expression de la forme $\sum_{n\geq 0} u_n(x)$, où $u_n(x) = a_n\cos(nx) + b_n\sin(nx)$, et $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 1}$ sont deux suites numériques.

Proposition 14.1 *Si les séries numériques $\sum_{n\geq 0}|a_n|$ et $\sum_{n\geq 1}|b_n|$ convergent, alors la série de fonctions $\sum_{n\geq 0} u_n$ converge normalement et uniformément vers sa limite qui est continue. Si g est une fonction continue sur $[0,2\pi]$, on a*

$$\int_0^{2\pi}\sum_{n=0}^{\infty} u_n g = \sum_{n=0}^{\infty}\int_0^{2\pi} u_n g.$$

Preuve : Pour tout x on a

$$|u_n(x)| = |a_n\cos(nx) + b_n\sin(nx)| \leq |a_n||\cos(nx)| + |b_n|\sin(nx)| \leq |a_n| + |b_n|.$$

Donc $\|u_n\|_\infty = \mathrm{Sup}_{x\in[0,2\pi]}|u_n(x)|$, qui est le plus petit majorant des $|u_n(x)|$, vérifie $\|u_n\|_\infty \leq |a_n| + |b_n|$, qui est le terme général d'une série convergente, ce qui montre que $\sum_{n\geq 0}|u_n|$ converge normalement.

Soit M un majorant des $|g(x)|$ sur le compact $[0,2\pi]$. On a donc $\|u_n g\|_\infty \leq (|a_n| + |b_n|)M$, donc $\sum_{n\geq 0} u_n g$ converge, elle aussi, normalement et uniformément vers sa limite qui est continue. On a donc

$$\int_0^{2\pi}\sum_{n=0}^{\infty} u_n g = \int_0^{2\pi}\lim_{N\to\infty}\sum_{n=0}^{N} u_n g = \lim_{N\to\infty}\int_0^{2\pi}\sum_{n=0}^{N} u_n g = \lim_{N\to\infty}\sum_{n=0}^{N}\int_0^{2\pi} u_n g = \sum_{n=0}^{\infty}\int_0^{2\pi} u_n g.$$

□

[1]Pour la préparation de ce cours nous nous sommes aidés des livres suivants, dans lesquels le lecteur trouvera bien des développements intéressants auquels nous avons du renoncer : W. Rudin, *Principes d'analyse mathématique ;* P. Thuillier et J.-C. Belloc, *Mathématiques, tome 3 : analyse*, Masson (1975).

14.2 Série de Fourier d'une fonction

Soit f une fonction 2π-périodique. Nous étudions ici comment écrire f sous forme d'une série trigonométrique, appelée série de Fourier de la fonction f. Le calcul peut paraître un peu magique au premier abord. Nous verrons, à la section suivante, que cette magie a une interprétation géométrique très naturelle. Rappelons tout d'abord trois formules de trigonométrie *qu'il est maintenant indispensable d'apprendre par coeur.*

$$\begin{aligned}
\cos(a)\cos(b) &= \frac{1}{2}(\cos(a-b)+\cos(a+b)) \\
\sin(a)\sin(b) &= \frac{1}{2}(\cos(a-b)-\cos(a+b)) \\
\cos(a)\sin(b) &= \frac{1}{2}(\sin(a-b)+\sin(a+b))
\end{aligned}$$

On appelle *coefficient de Fourier* de f les nombres

$$\begin{aligned}
a_0 &= \frac{1}{2\pi}\int_0^{2\pi} f(x)dx \\
a_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)dx \quad (n\geq 1) \\
b_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)dx \quad (n\geq 1).
\end{aligned}$$

On appelle *série de Fourier* de f la série trigonométrique $\sum_{n\geq 0} u_n$, avec $u_n(x) = a_n\cos(nx)+b_n\sin(nx)$.

Pour $r = 0, 1, \ldots$ nous dirons que f est $\mathcal{C}^r$ *par morceaux* sur $[0, 2\pi]$ et écrirons $f \in \mathcal{C}^r_{\text{mcx}}[0, 2\pi]$ s'il existe un ensemble fini de points $0 \leq x_1 \leq \ldots \leq x_m \leq 2\pi$ tels que sur $[x_i, x_{i+1}[$ la fonction f est la restriction d'une fonction définie et $\mathcal{C}^r$ sur $[x_i, x_{i+1}]$. Nous posons $f(x_i^-) = f_{i-1}(x_i)$ et $f(x_i^+) = f_i(x_i)(= f(x_i))$. Nous admettrons le théorème capital suivant et donnerons une preuve d'un résultat analogue sur les séries de Fourier complexes, sous des hypothèses un peu plus restrictives.

Théorème 14.2 (Dirichlet) *Soit $f \in \mathcal{C}^1_{mcx}[0, 2\pi]$.*

- *Si f est continue au point x, alors $\sum_{n=0}^{\infty} u_n(x) = f(x)$.*
- *Si x_i est un point de discontinuité de f, alors $\sum_{n=0}^{\infty} u_n(x) = \frac{1}{2}(f(x_i^+)+f(x_i^-))$.*

Exemple : Soit f la fonction 2π-périodique telle que $f(x) = x$ pour $x \in [-\pi, +\pi[$. Observons que pour toute fonction $x \mapsto g(x)$ 2π-périodique, on a

$$\begin{aligned}
\int_0^{2\pi} fg &= \int_0^{\pi} fg + \int_{\pi}^{2\pi} f(x)g(x)dx = \int_0^{\pi} fg + \int_{-\pi}^{0} f(y+2\pi)g(y+2\pi)dy \quad (\text{poser } x = y+2\pi) \\
&= \int_0^{\pi} fg + \int_{-\pi}^{0} f(y)g(y)dy \quad (\text{périodicité}) \\
&= \int_0^{\pi} fg + \int_{-\pi}^{0} fg = \int_{-\pi}^{\pi} fg.
\end{aligned}$$

A présent, comme $x \mapsto f(x)\cos(nx)$ est impaire sur $]-\pi, +\pi[$, on voit que $a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx = 0$ et $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx = 0$. Par ailleurs,

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)dx = \frac{1}{\pi}\left[\frac{-1}{n}x\cos(nx)\right]_{-\pi}^{+\pi} - \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{-1}{n}\cos(nx)dx = \frac{2\pi}{\pi n} - 0 = \frac{2}{n}.$$

Donc $\sum_{n=0}^{\infty}\frac{2}{n}\sin(nx)(= f(x)) = x$ pour $x \in]-\pi, +\pi[$ puisque f y est continue. Enfin, pour les points de discontinuité $x = \pi + k\pi$, on a bien

$$\sin(nx) = \sin(n\pi + 2nk\pi) = \sin(n(1+2k)\pi) = 0 = \frac{1}{2}(-\pi+\pi) = \frac{1}{2}(f(\pi^+)+f(\pi^-)).$$

Remarques : Cet exemple suggère les remarques générales suivantes :

- Si f est une fonction impaire, alors $a_n = 0$ pour tout $n \geq 0$, et si f est une fonction paire, alors $b_n = 0$ pour tout $n \geq 1$.
- Si f est une fonction T-périodique, la théorie de Fourier s'adapte immédiatement : il suffit de poser $\omega := \frac{2\pi}{T}$ et $u_n(x) := a_n \cos(nx) + b_n \sin(nx)$, avec $a_0 := \frac{1}{2T}\int_{\mathcal{T}} f(x)dx$, $a_n := \frac{1}{T}\int_{\mathcal{T}} f(x)\cos(nx)dx \quad (n \geq 1)$, et $b_n := \frac{1}{T}\int_{\mathcal{T}} f(x)\sin(nx)dx \quad (n \geq 1)$, où $\mathcal{T} := [t_0, t_0 + T]$, avec t_0 arbitraire.
- Si f n'est pas continue la convergence de $\sum u_n$ ne saurait être uniforme, puisque la limite uniforme de fonctions continues (ici $f_N = \sum_{n=0}^{N} u_n$) est continue. En particulier $\sum_{n\geq 0}|a_n|$ ou $\sum_{n\geq 0}|b_n|$ est divergente : c'est bien ce qui arrive dans le cas de l'exemple, qui fait apparaître une série harmonique.

14.3 Interprétation géométrique

Rappelons que si $(E, < \ , \ >)$ est un espace euclidien et $\mathcal{E} = \{e_1, e_2, \ldots\}$ est une base orthonormée de E, alors tout vecteur $f \in E$ se décompose sur la base $\mathcal{B}$ en $f = \sum_n \lambda_n e_n$, la valeur de la composante λ_n de f dans la direction e_n s'obtenant par *projection orthogonale* sur e_n :

$$\lambda_n = < f, e_n > .$$

Orthogonalité et orthonormalité : Soulignons ce petit miracle : λ_n ne dépends que de e_n et pas du choix des autres vecteurs e_m $m \neq n$ choisis pour complèter la base orthonormée. En fait, il suffit que la base soit orthogonale, moyennant une légère adaptation : soit $\mathcal{F} = \{f_1, f_2, \ldots\}$ une base orthogonale ; pour obtenir une base orthonormée, il suffit de poser $e_n = \frac{f_n}{\|f_n\|}$ pour former une base orthonormée. Posons $\mu_n := < f, f_n >$. On a donc $\mu_n = < f, \|f_n\| e_n > = \|f_n\| < f, e_n > = \|f_n\|\lambda_n$, d'où

$$f = \sum_n \lambda_n e_n = \sum_n \frac{\mu_n}{\|f_n\|}\frac{f_n}{\|f_n\|} = \sum_n \frac{\mu_n}{\|f_n\|^2} f_n =: \sum_n c_n f_n \text{ , avec } c_n := \frac{< f, f_n >}{< f_n, f_n >}.$$

La présence du dénominateur $< f_n, f_n >$ explique, dans les coeficients de Fourier, la présence du facteur $\frac{1}{2\pi}$ dans a_0 et la présence du facteur $\frac{1}{\pi}$ dans a_n et b_n, $n \geq 1$. En effet, la base en cause ici est constituée des fonction $x \mapsto \cos(nx)$, $n \geq 0$, et $x \mapsto \sin(nx)$, $n \geq 1$; le produit scalaire est $< f, g > := \int_0^{2\pi} f(x)g(x)dx$.

Exercice : Monter que les $x \mapsto \cos(nx)$, $n \geq 0$, et $x \mapsto \sin(nx)$, $n \geq 1$ sont tous orthogonaux deux à deux, et que

$$< \cos(nx), \cos(nx) > = \pi = < \sin(nx), \sin(nx) > \text{ et } < 1, 1 > = 2\pi.$$

14.4 Série de Fourier complexe

Lors du cours sur les fonctions analytiques nous avons vu que la fonction exponentielle se définit, par sa séries entière, non seulement pour des valeurs réelles de la variable, mais aussi pour tout $z \in \mathbb{C}$. De ce fait, elle concerve d'ailleurs ses propriétés formelles telles que $e^{z_1+z_2} = e^{z_1}e^{z_2}$ et $\frac{d}{dz}e^z = e^z$. Elle n'est plus toujours positive ; en revanche elle est $2i\pi$-périodique, comme il ressort de la *formule d'Euler*

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}) \text{ et } \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}) \text{ (avez-vous notez de "}i\text{" au dénominateur?}$$

Ainsi sommes-nous naturellement ammenés à étendre à la théorie de Fourier au cas de fonctions à valeurs complexes, telle que $e^{ix} = \cos x + i \sin x$, qui, pour $x \in \mathbb{R}$ comme nous le supposerons dans tout ce chapitre, a pour conjugué

$$\overline{e^{ix}} = \cos x - i \sin x = \cos(-x) + i \sin(-x) = e^{-ix}.$$

On sait que, dans le cas d'un espace vectoriel complexe, pour assurer que $< f, f >$ reste réel positif, il convient de définir $< f, g >$ comme produit hermitien[2]

$$< f, g > = \int_0^{2\pi} f(x)\overline{g(x)}dx.$$

[2]Mêmes propriété qu'en produit scalaire, mais défini sur un espace vectoriel complexe et à valeurs complexes ; de plus, pour λ et μ complexes, $< \lambda f, \mu g > = \lambda\overline{\mu} < f, g >$.

Exemple : Soient $m \in \mathbb{Z}$ et $n \in \mathbb{Z}$, c'est-à-dire des entiers *relatifs*. On a

$$< e^{imx}, e^{inx} >= \int_0^{2\pi} e^{imx} e^{-inx} dx = \int_0^{2\pi} e^{i(m-n)x} dx.$$

Nous voyons que, si $m \neq n$, $e^{i(m-n)z}$ a pour primitive $\frac{1}{i(m-n)} e^{i(m-n)z} = \frac{1}{i(m-n)}(\cos(m-n)z + i\sin(m-n)z$ qui est 2π-périodique. Si $m = n$, $e^{i(m-n)z} = e^0 = 1$; donc

$$\begin{aligned}
\text{si } m \neq n,\ < e^{imx}, e^{inx} > &= \frac{1}{i(m-n)} \left[e^{i(m-n)x} \right]_0^{2\pi} = 0 \text{ par } 2\pi\text{-périodicité, et} \\
\text{si } m = n,\ < e^{imx}, e^{inx} > &= [\ 1\ ]_0^{2\pi} = 2\pi.
\end{aligned}$$

Nous venons donc de voir que les fonctions e^{inx} forment une famille de vecteurs orthogonaux, tous de même norme $\|e^{inx}\| = \sqrt{e^{inx}, e^{inx}} = \sqrt{2\pi}$.

Expression de $\sum u_n$ au moyen des fonctions e^{inx} : par le formule d'Euler on a, si $n > 0$,

$$\begin{aligned}
u_n(x) = a_n \cos(nx) + b_n \sin(nx) &= a_n \frac{1}{2}(e^{inx} + e^{-inx}) + b_n \frac{1}{2}(e^{inx} - e^{-inx}) \\
&= \frac{1}{2}(a_n - ib_n)e^{inx} + \frac{1}{2}(a_n + ib_n)e^{-inx} \text{ , puisque } \frac{1}{i} = \frac{-i^2}{i} = -i \\
&= c_n e^{inx} + c_{-n} e^{-inx} \text{ en posant}
\end{aligned}$$

$$\begin{aligned}
c_n = \frac{1}{2}(a_n - ib_n) &= \frac{1}{2}\left(\frac{1}{\pi} \int_0^{2\pi} f(x)\cos(nx)dx - \frac{i}{\pi} \int_0^{2\pi} f(x)\sin(nx)dx \right) \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(x)(\cos(nx) - i\sin(nx))dx = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-inx}dx \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{e^{+inx}}dx = \frac{1}{2\pi} < f, e^{+inx} > \text{ et} \\
c_{-n} = \frac{1}{2}(a_n + ib_n) = \overline{c_n} &= \frac{1}{2\pi}\overline{< f, e^{+inx} >} = \frac{1}{2\pi} < \overline{f}, \overline{e^{+inx}} >= \frac{1}{2\pi} < f, e^{-inx} > .
\end{aligned}$$

Par ailleurs, pour $n = 0$, on a

$$u_0(x) = a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x)dx = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{e^{i0x}}dx = \frac{1}{2\pi} < f, e^{i0x} >=: c_0 = c_0 e^{i0x}.$$

En d'autres termes, $\sum_{n=0}^{N} u_n = \sum_{n=-N}^{N} c_n e^{inx}$, avec la définition (unifiée !) des coefficients c_n, dit *coefficients de Fourier complexes* de f :

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-inx}dx = \frac{< f, e^{inx} >}{< e^{inx}, e^{inx} >}, \tag{14.1}$$

et, en passant à la limite sur N,

$$\sum_{n=0}^{\infty} u_n = \lim_{N\to\infty} \sum_{n=0}^{N} u_n = \lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-\infty}^{\infty} c_n e^{inx}. \tag{14.2}$$

Bien entendu, le théorème de Dirichlet 14.2 reste vrai ; la relation (14.2) donne simplement une nouvelle expression de $\sum_{n=0}^{N} u_n(x)$.

Exercice : Montrer directement que $c_{-n} = \overline{c_n}$.

Nous admettons le théorème suivant qui montre que les coefficients de Fourier d'une fonction continue constituent un bon codage :

Théorème 14.3 *Soient f et g deux fonctions 2π-périodique continues. Si les coeficients de Fourier réels ou complexes de f et g sont égaux, alors $f = g$.*

14.5 Inégalité de Bessel et égalité de Parseval

En reprenant la preuve de la proposition 14.1 on voit que si $\sum_{k=-\infty}^{k=+\infty} c_k e^{ikx}$ converge uniformément et si g est continue, on a

$$< \sum_{k=-\infty}^{k=+\infty} c_k e^{ikx}, g >= \sum_{k=-\infty}^{k=+\infty} < c_k e^{ikx}, g > . \tag{14.3}$$

Si f est continue nous pouvons donc calculer la norme de la projection orthogonale de f sur le sous-espace engendré par les $x \mapsto e^{inx}$, n=-N..N :

$$\begin{aligned}
< f, \sum_{n=-N}^{+N} c_n e^{inx} > &= < \sum_{k=-\infty}^{k=+\infty} c_k e^{ikx}, \sum_{n=-N}^{n=+N} c_n e^{inx} > \text{ (par le théorème de Dirichlet 14.2)} \\
&= \sum_{k=-\infty}^{k=+\infty} < c_k e^{ikx}, \sum_{n=-N}^{n=+N} c_n e^{inx} > \text{ par (14.3)} \\
&= \sum_{k=-\infty}^{k=+\infty} \sum_{n=-N}^{n=+N} c_n \overline{c_k} < e^{ikx}, e^{inx} > \text{ où seul } k = n \text{ a un apport non nul, d'où} \\
&= \sum_{n=-N}^{n=+N} |c_n|^2 < e^{inx}, e^{inx} >= 2\pi \sum_{n=-N}^{n=+N} |c_n|^2.
\end{aligned}$$

En passant à la limite sur $N \to \infty$, on en déduit

Théorème 14.4 (identité de Bessel-Parseval) *Si f est continue, on a*

$$\int_0^{2\pi} f^2(x)dx = 2\pi \sum_{n=-\infty}^{n=+\infty} |c_n|^2. \tag{14.4}$$

Comme $(\sum_{n=-N}^{n=+N} |c_n|^2)_N$ est une suite croissante, on a $\sum_{n=-N}^{n=+N} |c_n|^2 \leq \sum_{n=-\infty}^{n=+\infty} |c_n|^2$, et nous avons donc montré ainsi l'*inégalité de Bessel*

$$\frac{1}{2\pi} \int_0^{2\pi} f^2(x)dx \geq \sum_{n=-N}^{n=+N} |c_n|^2$$

sous l'hypothèse que f est continue. A noter que cette inéalité est vraie sous la seule hypothèse, bien plus générale, que f^2 est intégrable.

Chapitre 15

Fonctions caractéristiques

15.1 Espérance d'une variable aléatoire complexe

Si V et W sont deux v.a. réelles, il est commode de noter $U := V + iW$ le vecteur aléatoire (V, W), tout comme on peut noter $u = w + iw$ un point (v, w) du plan, conformément à l'interprétation géométrique des nombres complexes. On pose

$$\mathbb{E}(V + iW) := \mathbb{E}V + i\mathbb{E}W.$$

Exercice : Montrer que la fonction $U \mapsto \mathbb{E}U$ ainsi définie est linéaire (à valeurs complexes) et, en particulier, si $\lambda = \alpha + i\omega$ et $U = V + iW$, on a bien $\mathbb{E}(\lambda U) = \lambda \mathbb{E}U$.

Soit $t \in \mathbb{R}$ et X est une v.a. *réelle*, alors $e^{itX} = \cos(tX) + i\sin(tX)$ est une v.a. complexe et, avec la définition ci-dessus

$$\mathbb{E}e^{itX} = \mathbb{E}\cos(tX) + i\mathbb{E}\sin(tX). \tag{15.1}$$

Notons que les v.a. $\cos(tX)$ et $\sin(tX)$ étant bornées, elles ont bien un moment d'ordre un et donc (15.1) est bien défini, ce qui permet de poser la définition suivante :

Définition : On appelle *fonction caractéristique* de la v.a. réelle X la fonction $\varphi_X : \mathbb{R} \longrightarrow \mathbb{C}$ définie par

$$\varphi_X(t) = \mathbb{E}e^{itX}.$$

Les fonctions caractéristiques sont des outils précieux dans l'étude des lois limites du fait du théorème suivant dû à Paul Lévy, que nous admettrons :

Théorème 15.1 (P. Lévy) *Soit (X_n) une suite de v.a., de fonctions caractéristiques φ_{X_n}. S'il existe une fonction φ, continue en 0, telle que $\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi(t)$ pout tout $t \in \mathbb{R}$, alors il existe une loi, de fonction de répartition F_X, tel que la suite (X_n) tend en loi vers X, c'est-à-dire que pour tout x où F_X est continue, on a $\lim_{n\to\infty} F_{X_n}(x) = F_X(t)$.*

En particulier, si φ_X et φ_Y désignent les fonctions caractéristique des v.a. réelles X et Y, si $\varphi_X = \varphi_Y$, alors X et Y ont même loi : ceci légitime la terminologie "caractéristique" adoptée, un résultat qui d'ailleurs admet un preuve plus directe que le passage par ce théorème sur les limites de suites de variables aléatoires.

Nous voyons que la fonction caractéristique d'une v.a. joue un rôle analogue à la fonctions génératrice des moments M_X que nous avons déja étudiée ; un avantage de la fonction caractéristique est qu'elle est définie pout tout t réel. Remarquons que

$$\varphi_X(t) = \mathbb{E}e^{itX} = \mathbb{E}e^{uX} = M_X(u) \text{ , pour } u = it.$$

La fonction génératrice $M_X(u)$ est donc bien définie pour tout u imaginaire pur : c'est son *prolongement* (analytique) à u réel qui peut faire défaut pour certaines lois telles que la loi de Cauchy.

Exemple : Soit $X \rightsquigarrow \mathcal{U}[0, 1]$ une v.a. uniforme sur $[0, 1]$. On a

$$\begin{aligned}
\varphi_X(t) &= \int_0^1 e^{itx}dx = \int_0^1 \cos(tx)dx + i\int_0^1 \sin(tx)dx = \left[\frac{1}{t}\sin(tx)\right]_{x=0}^{x=1} + i\left[\frac{-1}{t}\sin(tx)\right]_{x=0}^{x=1} \\
&= \frac{1}{t}\sin t + \frac{i}{i^2 t}(\cos t - 1) = \frac{1}{it}(i\sin t + \cos t - 1) = \frac{1}{it}(e^{it} - 1).
\end{aligned}$$

A noter que nous aurions pu effectuer ce calcul plus rapidement en remarquant que la dérivée réelle $\frac{d}{dx}$ est un cas particulier de la dérivée au sens complexe $\frac{d}{dz}$ lorsqu'elle existe, que $\frac{d}{dx}(\operatorname{Re} F(x)) = \operatorname{Re}(\frac{d}{dx}F(x))$, et que $\frac{d}{dx}(\operatorname{Im} F(x)) = \operatorname{Im}(\frac{d}{dx}F(x))$; on a donc directement

$$\varphi_X(t) = \int_0^1 e^{itx}dx = \int_0^1 \frac{d}{dx}\left(\frac{1}{it}e^{itx}\right)dx = \left[\frac{1}{it}e^{itx}\right]_{x=0}^{x=1} = \frac{1}{it}(e^{it}-1).$$

Proposition 15.2 *Pour toute v.a. réelle X la fonction caractéristique $\varphi_X : \mathbb{R} \longrightarrow \mathbb{C}$ est définie, continue, bornée ; pour tous t, a, et b réels on a*

1. $|\varphi_X(t)| \leq \varphi_X(0) = 1$,
2. $\varphi_X(-t) = \overline{\varphi_X(t)}$,
3. $\varphi_{aX+b}(t) = e^{ibt}\varphi_X(at)$.

Exercice : Montrer que si X est symétrique, c'est-à-dire que $-X$ a même loi que X, alors $\varphi(t)$ est réel, et que φ_X est pair ; (indication : montrer que $\overline{\varphi_X(t)} = \varphi_X(t)$.)

Proposition 15.3 *Si X et Y sont des v.a. indépendantes, on a*

$$\varphi_{X+Y} = \varphi_X\varphi_Y.$$

Preuve : Comme X et Y sont indépendantes, il en est de même de e^{itX} et e^{itY} ; on en déduit que pour tout $t \in \mathbb{R}$

$$\varphi_{X+Y}(t) = \mathbb{E}e^{it(X+Y)} = \mathbb{E}(e^{itX}e^{itY}) = \mathbb{E}e^{itX}\mathbb{E}e^{itY} = \varphi_X(t)\varphi_Y(t).$$

□

15.2 Fonctions caractéristiques des principales lois.

Loi binomiale : Supposons que $X \rightsquigarrow \mathcal{B}(n,p)$; donc, pour $k = 0..n$, $\mathbb{P}\{X = k\} = \binom{n}{k} p^k(1-p)^{n-k}$; et

$$\varphi_X(t) = (1-p+pe^{it})^n.$$

Loi de Poisson : Supposons que $X \rightsquigarrow \mathcal{P}(\lambda)$; donc pour $k = 0, \ldots$ on a $\mathbb{P}\{X = k\} = e^{-\lambda}\frac{\lambda^k}{k!}$; et

$$\varphi_X(t) = e^{\lambda(e^{it}-1)}.$$

Loi géométrique : Supposons que $X \rightsquigarrow G(p)$; donc, pour $k = 1, \ldots$, $\mathbb{P}\{X = k\} = (1-p)^{k-1}p$; et

$$\varphi_X(t) = \frac{pe^{it}}{1-(1-p)e^{it}}.$$

Loi uniforme : Supposons que $X \rightsquigarrow \mathcal{U}[a,b]$; donc $f_X(x) = \frac{1}{b-a}\mathbb{I}_{[a,b]}(x)$; et

$$\varphi_X(t) = \frac{1}{i(b-a)t}(e^{ibt} - e^{iat}).$$

Loi exponentielle : Supposons que $X \rightsquigarrow \mathcal{E}(\lambda)$; donc $f_X(x) = \lambda e^{-\lambda x}\mathbb{I}_{[0,\infty[}(x)$; et

$$\varphi_X(t) = \frac{\lambda}{\lambda - it}.$$

Première loi de Laplace : Supposons que X suive la première loi de Laplace, donc $f_X(x) = \frac{1}{2}e^{-|x|}$; alors

$$\varphi_X(t) = \frac{1}{1+t^2}.$$

Loi normale : Supposons que $X \rightsquigarrow \mathcal{N}(\mu,\sigma)$; donc $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{(x-\mu)}{\sigma})^2}$; et

$$\varphi_X(t) = e^{it\mu}e^{-\frac{t^2\sigma^2}{2}}.$$

Loi de Cauchy : Supposons que $X \rightsquigarrow \mathcal{C}(0)$; donc $f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}$; et

$$\varphi_X(t) = e^{-|t|}.$$

Chapitre 16

Transformation de Fourier

Dans le cours de probabilités sur la fonction caractéristique φ_X de la loi d'une v.a. X, dans le cas où X est absolument continue de densité f_X, nous avions

$$\varphi_X(t) = \int_{x=-\infty}^{x=+\infty} e^{+itx} f_X(x)dx,$$

définissant ainsi une application linéaire $f_X \mapsto \varphi_X$ (on dit aussi "transformation" ou "opérateur", pour souligner que les vecteurs considérés ici sont des fonctions). Cette application, appelée transformation de Fourier, est définie et a des propriétés remarquables au delà du cas particulier d'une densité (pour laquelle on a toujours $f_X(x) \geq 0$ et $\int f_X = 1$) et demande seulement que f soit "intégrable" sur $\mathbb{R}$ (ce qu'on note par $f \in L^1(\mathbb{R})$ ou simplement $f \in L^1$), c'est-à-dire, essentiellement, que $\int_{-\infty}^{+\infty} |f(x)|dx$ soit fini. Notons qu'il y a de petites différences d'un auteur à l'autre dans la définition de la transformation de Fourier. Nous avons adopté ici la définition choisie par J. Jacod et Ph. Protter ; d'autres auteurs considèrent $\int_{x=-\infty}^{x=+\infty} e^{-itx} f_X(x)dx$ (A. Kolmogorov et S. Fomine), $\int_{x=-\infty}^{x=+\infty} e^{-2i\pi tx} f_X(x)dx$ (F. Rodier), ou $\frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{x=+\infty} e^{-itx} f_X(x)dx$ (W. Rudin). Elles débouchent évidement sur des théories similaires mais comportent des différences dans les formules, ce qui impose au lecteur une certaine vigilance.

Définition : Soit $f \in L^1$. On appelle *transformée de Fourier* de f et on note $\mathcal{F}(f) := \hat{f}$ la fonction

$$t \mapsto \hat{f}(t) := \int_{x=-\infty}^{x=+\infty} e^{+itx} f(x)dx.$$

Remarquons que, comme $|e^{+itx} f(x)| = |f(x)|$ et $f \in L^1$, $\hat{f}(t)$ est bien défini pour tout $t \in \mathbb{R}$.

Proposition 16.1 *Soient $f \in L^1$, $\alpha \in \mathbb{R}$, et $\lambda \in \mathbb{R}$; alors $f \mapsto \hat{f}$ est linéaire et*

1. *Si $g(x) := f(x)e^{i\alpha x}$, alors $\hat{g}(t) = \hat{f}(t+\alpha)$.*
2. *Si $g(x) := f(x-\alpha)$, alors $\hat{g}(t) = e^{it\alpha}\hat{f}(t)$.*
3. *Si $g(x) := \overline{f(-x)}$, alors $\hat{g}(t) = \overline{\hat{f}(x)}$.*
4. *Si $g(x) := f(x/\lambda)$ et si $\lambda > 0$, alors $\hat{g}(t) = \lambda\hat{f}(\lambda t)$.*

La preuve de cette proposition ne présente aucune difficulté et est laissée en exercice. La situtation est un peu différente pour les résultats suivants : leur preuve comporte à un moment un problème de justification d'interversion d'ordre dans un passage à la limite. Ce passage est marqué d'une étoile $*$. Dans ce cours de Deug il vous a été donné des théorèmes fournissant certaines hypothèses sous lesquelles cette interversion donne bien des résultats égaux. On attend donc de vous que vous soyez sensibles au fait qu'une justification est nécessaire. Toutefois ces théorèmes élémentaires ne s'appliquent pas ici, essentiellement du fait que le domaine d'intégration, $\mathbb{R}$, n'est pas compact. Les théorème convenant à la justification du cas considéré ici vous seront donnés dans le cours d'intégration de licence.

Proposition 16.2 (produit de convolution) *Soient $f \in L^1$ et $g \in L^1$; la fonction $x \mapsto (f*g)(x)$ définie par*

$$(f*g)(x) = \int_{y=-\infty}^{y=+\infty} f(x-y)g(y)dy$$

est définie pour presque-tout $x \in \mathbb{R}$; si $h := f*g$, alors $\hat{h} = \hat{f}\hat{g}$.

Preuve : La clé de ce résultat est la formule de Fubini ; pour assurer qu'elle s'applique bien, il conviendrait de vérifier que la fonction $F(x,y) := f(x-y)g(y)$ est bien intégrable sur $\mathbb{R}^2$; au delà, la preuve est élémentaire :

$$\begin{aligned}
\hat{h}(t) &= \int_{x=-\infty}^{x=+\infty} e^{itx}\left(\int_{y=-\infty}^{y=+\infty} f(x-y)g(y)dy\right)dx \\
&= \int_{x=-\infty}^{x=+\infty}\left(\int_{y=-\infty}^{y=+\infty} e^{itx}f(x-y)g(y)dy\right)dx \overset{*}{=} \int_{y=-\infty}^{y=+\infty}\left(\int_{x=-\infty}^{x=+\infty} e^{itx}f(x-y)g(y)dx\right)dy \quad \text{(Fubini)} \\
&= \int_{y=-\infty}^{y=+\infty} g(y)\left(\int_{u=-\infty}^{u=+\infty} e^{ity}e^{itu}f(u)du\right)dy \quad \text{(changement de variable } x\mapsto u=x-y) \\
&= \int_{y=-\infty}^{y=+\infty} g(y)e^{ity}\hat{f}(t)dy = \hat{f}(t)\int_{y=-\infty}^{y=+\infty} g(y)e^{ity}dy = \hat{f}(t)\hat{g}(t).
\end{aligned}$$

□

Proposition 16.3 *Soit $g(x) = ixf(x)$. Si $g \in L^1$ alors $\hat{f}$ est dérivable et*

$$(\hat{f})' = \hat{g}.$$

Preuve : Soit $h \neq 0$; un développement de e^{ihx} à l'ordre 1 montre que $\lim_{h\to 0}\frac{e^{ihx}-1}{h} = ix$, et donc

$$\begin{aligned}
\hat{f}'(t) = \lim_{h\to 0}\frac{1}{h}(\hat{f}(t+h)-\hat{f}(t)) &= \lim_{h\to 0}\int_{x=-\infty}^{x=+\infty}\frac{e^{ihx}-1}{h}e^{itx}f(x)dx \\
&\overset{*}{=} \int_{x=-\infty}^{x=+\infty}\lim_{h\to 0}\frac{e^{ihx}-1}{h}e^{itx}f(x)dx = \int_{x=-\infty}^{x=+\infty} ixe^{itx}f(x)dx = \hat{g}(t).
\end{aligned}$$

□

Proposition 16.4 *Soit $f \in L^1$ dérivable ; supposons que $f' \in L^1$ et que $f(x) = \int_{-\infty}^{x} f'(\xi)d\xi$. Soit $g = f'$; alors*

$$\hat{g}(t) = -it\hat{f}(t).$$

Preuve : Posons $\gamma_n(t) = \int_{x=-n}^{x=+n} e^{itx}f'(x)dx$; par intégration par parties on a

$$\gamma_n(t) = \int_{x=-n}^{x=+n} e^{itx}f'(x)dx = [e^{itx}f(x)]_{x=-n}^{x=+n} - \int_{x=-n}^{x=+n} ite^{itx}f(x)dx.$$

Par définition $\hat{g}(t) = \lim_{n\to\infty}\gamma_n(t)$, et $\lim_{n\to\infty}\int_{x=-n}^{x=+n} e^{itx}f(x)dx = \hat{f}(t)$. Il reste à s'assurer que le crochet est nul. Par hypothèse $\lim_{x\to-\infty} f(x) = 0$ et, comme $f' \in L^1$, $\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty}\int_{-\infty}^{x} f'(\xi)d\xi = f_+$ existe. Comme $f \in L^1$, nécessairement $f_+ = 0$. On en déduit que $\lim_{n\to\infty}[e^{itx}f(x)]_{x=-n}^{x=+n} = 0$, d'où le résultat annoncé. □

Le théorème fondamental suivant (que nous admettons) montre que la transformation de Fourier $f \mapsto \hat{f}$ admet une transformation "inverse" $\hat{f} \mapsto \tilde{f}$, avec $\tilde{f}(x) = f(x)$ pour presque-tout x. La définition de "presque-tout" relève du cours de mesure et intégration ; cela signifie que l'ensemble des x pour lesquels la relation est fausse est de mesure nulle.

Théorème 16.5 *Si $f \in L^1$, alors $\hat{f}$ est continue et $\lim_{x\to\pm\infty}\hat{f}(x) = 0$. De plus $\|\hat{f}\|_\infty \le \|f\|_1$. Si $f \in L^1$ et si $\hat{f} \in L^1$, alors pour presque-tout $x \in \mathbb{R}$ on a $f(x) = \tilde{f}(x)$, où*

$$\tilde{f}(x) := \frac{1}{2\pi}\int_{t=-\infty}^{t=+\infty} e^{-ixt}\hat{f}(t)dt \quad (= \overline{\mathcal{F}}(\hat{f})(x),\ \textit{voir ci-dessous}).$$

On est donc amené à poser la définition suivante :

Définition : On appelle *transformation de Fourier inverse* l'opérateur $\varphi \mapsto \overline{\mathcal{F}}(\varphi)$ défini pour tout $\varphi \in L^1$ par

$$\overline{\mathcal{F}}(\varphi)(x) = \frac{1}{2\pi}\int_{t=-\infty}^{t=+\infty} e^{-ixt}\varphi(t)dt.$$