Chapitre 0

Rappels d'intégration et de topologie

0.1 Généralités sur la théorie de l'intégration

On rappelle que $\mathcal{S} \subset \mathcal{P}(X)$ est une *σ-algèbre* sur X si

1. $X \in \mathcal{S}$;
2. $V \in \mathcal{S} \Longrightarrow X \setminus V \in \mathcal{S}$;
3. si $(E_i)_{i\in\mathbb{N}}$, $E_i \in \mathcal{S}$, alors $\displaystyle\bigcup_{i\in\mathbb{N}} E_i \in \mathcal{S}$.

On dit alors que $(X, \mathcal{S})$ est un *espace mesurable*. Une mesure (positive) sur $(X, \mathcal{S})$ est une application $\mu : \mathcal{S} \to [0, +\infty]$ avec $\mu(\emptyset) = 0$ et qui est σ-additive.

À partir du moment où l'on connait $\mu(A)$ pour les parties mesurables $A \subset \mathcal{S}$, on peut définir l'intégrale comme suit : d'abord on pose $\int 1_A \, d\mu = \mu(A)$. On étend cette définition aux fonctions étagées mesurables positives (par linéarité), puis aux fonctions mesurables positives quelconque en approchant celles-ci de manière croissante par des fonctions étagées (lemme d'approximation). On sait donc intégrer toutes les fonctions positives mesurables$f$ (et on a $\int f \, d\mu \in [0, +\infty]$) et on dispose pour celles-ci du théorème de convergence monotone. On ne peut cependant pas étendre la définition à une fonction réelle ou complexe mesurable $f : X \to \mathbb{C}$ quelconque. Mais dans le cas où elle est μ-intégrable, c'est-à-dire où la fonction positive $|f|$ vérifie $\int |f| \, d\mu < +\infty$, alors on peut. On établit ensuite le théorème de convergence dominée.

Pour résumer : quand est-ce que l'expression $\int f \, d\mu$ a un sens ? Vous devez vous poser systématiquement cette question, et y répondre, <u>avant</u> de pouvoir écrire cette expression. Il y a deux cas où l'on peut écrire $\int f \, d\mu$, qu'il vous faut donc identifier :

- Soit la fonction f est à valeurs réelles positives, et on peut toujours écrire $\int f \, d\mu$, qui peut éventuellement être égal à $+\infty$.
- Soit la fonction mesurable positive $|f|$ a une intégrale finie, $\int |f| \, d\mu < +\infty$, c'est-à-dire f est (X, μ)-intégrable. Dans ce cas, $\int f \, d\mu$ est un nombre fini, appartenant à $\mathbb{R}$ ou $\mathbb{C}$.

Avant de pouvoir écrire $\int f \, d\mu$, il faut justifier que l'on se trouve dans l'une des deux situations ci-dessus. De même, il est inexact de dire que l'intégrale est linéaire. Elle ne l'est pas en général, mais elle le devient si on est dans un des cas ci-dessus (dans le premier cas, on ne peut faire que des combinaisons positives de fonctions positives, bien sûr).

Remarque 0.1. *Étant donnée une fonction réelle mesurable $f : X \to \mathbb{R}$, il y a un troisième cas un peu plus général qui permet de définir $\int f \, d\mu$, c'est quand l'une des deux intégrales de fonctions positives $\int f_+ \, d\mu$ ou $\int f_- \, d\mu$ est finie (dans la discussion précédente, on demandait que les deux soient finies, car cela équivaut à $\int |f| \, d\mu$ finie). Dans ce cas, on dit que f est semi-intégrable, et on peut encore poser $\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu$ qui appartient à $[-\infty, +\infty]$.*

On retrouve tout le temps, en théorie de l'intégration, cette dichotomie f positive ou/et intégrable. Par exemple, si on travaille avec la mesure de Lebesgue sur $\mathbb{R}^+$ et avec une fonction borélienne f qui est localement intégrable sur $\mathbb{R}^+$ (c'est-à-dire $\int_{[a,b]} |f| < +\infty$ pour tout segment $[a, b] \subset \mathbb{R}^+$) on a .

— Soit la fonction f est à valeurs réelles positives, et alors $\int_{[0,M]} f \to \int_{\mathbb{R}^+} f$ *par convergence monotone.*

— Soit f est intégrable sur $\mathbb{R}^+$, et alors $\int_{[0,M]} f \to \int_{\mathbb{R}^+} f$ *par convergence dominée.*

Lorsqu'on souhaite faire des inégalités sur les intégrales, on se ramène presque toujours au cas de fonction positives à l'aide de l'inégalité triangulaire.

Proposition 0.2. *Soit (X, μ) un espace mesuré et $f : X \to \mathbb{C}$ une fonction μ-intégrable. On a alors*

$$\Big| \int f \, d\mu \Big| \leq \int |f| \, d\mu,$$

avec égalité si et seulement si il existe $\alpha_0 \in \mathbb{C}$ avec $|\alpha_0| = 1$ tel que $f = \alpha_0 |f|$ μ-presque-partout.

Dans le cas d'une fonction réelle, il y a donc égalité si et seulement si f est presque-partout positive, ou bien presque-partout négative.

Démonstration. Il y a plusieurs preuves possibles. On peut raisonner par dualité comme suit. Pour $w \in \mathbb{C}^n$, on a

$$|w| = \max\{\Re(wz) \; ; \; z \in \mathbb{C}, \; |z| \leq 1\}.$$

Ainsi, il existe $z_0 \in \mathbb{C}$ tel que

$$\Big| \int f \, d\mu \Big| = \Re\Big[\Big(\int f(x) \, d\mu(x) \Big) z_0 \Big] = \int \Re(f(x) z_0) \, d\mu(x),$$

où l'on a utilisé la linéarité de l'intégrale. En réutilisant la formule ci-dessous (cette fois juste l'inégalité), on conclut en écrivant que $\Re(f(x)z_0) \leq |f(x)|$.

En ce qui concerne la caractérisation des cas d'égalité, on va traiter le cas d'une fonction à valeurs réelles ; (faites le cas complexe à titre d'exercice). Soit $f : X \to \mathbb{R}$ est une fonction μ-intégrable telle que $\Big| \int f \, d\mu \Big| = \int |f| \, d\mu$. Quitte à changer f en $-f$, on peut supposer que $\int f \, d\mu \geq 0$, ce qui implique que

$$\int \big(|f| - f\big) \, d\mu = 0.$$

Or la fonction $|f| - f$ est positive, et donc elle doit être nulle μ-presque-partout.

□

Remarque 0.3 (Valeur $+\infty$)**.** *En théorie de l'intégration, on autorise les fonction à prendre la valeur $+\infty$. Mais on peut toujours se ramener au cas où les fonctions ne prennent que des valeurs finies. En effet, si f est une fonction mesurable de X dans $[0, +\infty]$ alors soit $\int f \, d\mu = +\infty$, et alors il n'y a en général plus rien à dire, ou bien $\int f \, d\mu < +\infty$, et dans ce cas,*

$$\mu(\{f = +\infty\}) = 0.$$

Ainsi, lorsque $\int f \, d\mu < +\infty$, alors pour toute fonction intégrable H positive ou intégrable (et en particulier pour $H = f$) on a

$$\int_X H \, d\mu = \int_{\{f < +\infty\}} H \, d\mu.$$

Cela découle de propriétés des ensembles de mesure nulle.

Remarque 0.4 (Ensembles de mesure nulle)**.** *Si A est une partie mesurable de mesure nulle, i.e. $\mu(A) = 0$, alors si f est une fonction positive ou intégrable, on a*

$$\int_A f \, d\mu = 0.$$

Cela est vrai même si f prend la valeur $+\infty$. On pourrait dire que ça découle de la convention $0 \cdot +\infty = 0$, ou mieux, de la convergence monotone ou dominée.

Remarque 0.5. *En théorie de l'intégration, on manipule souvent, pour $g : X \to \mathbb{R}$ des ensembles du type*

$$\{g \in B\} := \{x \in X \ ; \ g(x) \in B\} = g^{-1}(B)$$

associé à un ensemble $B \subset \mathbb{R}$, typiquement un borélien et souvent un intervalle, comme

$$\{g \geq a\} = \{x \in X \ ; \ g(x) \geq a\} = g^{-1}([a, +\infty[).$$

0.2 Inégalités

Dans toute cette section , $(X, \mathcal{S}, \mu)$ désigne un espace mesuré quelconque. Lorsque des hypothèses supplémentaires seront nécessaires, elles seront précisées.

Inégalité arithmético-géométrique

Si $x, y \in \mathbb{R}$, on a $|x - y|^2/2 \geq 0$, ce qui s'écrit aussi $xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2$. On peut aussi réecrire cela en disant que

$$\forall a, b \geq 0, \qquad \sqrt{ab} \leq \frac{1}{2}a + \frac{1}{2}b \tag{1}$$

L'inégalité suivante, qui est à la base de la plupart des inégalités de ce chapitre, est une version pour une combinaison plus générale. Elle exprime que le logarithme est concave.
Il peut être pratique d'autoriser la valeur $+\infty$. Si besoin, on conviendra que $0 \cdot (+\infty) = 0$.

Proposition 0.6 (Inégalité arithmético-géométrique)**.** *Soit $t \in]0, 1[$ et $a, b \in [0, +\infty]$. Alors*

$$a^{1-t}b^t \leq (1 - t)a + tb \tag{2}$$

avec, lorsque a et b sont finis, égalité si et seulement si $b = a$.

L'inégalité est aussi vraie, mais triviale, lorsque $t = 0$ ou $t = 1$.

Démonstration. Si $a = 0$ ou $b = 0$, l'inégalité est vraie, et il y a égalité si et seulement si $b = 0$. Si $a = +\infty$ ou $b = \infty$, l'inégalité est trivialement vraie. On peut donc supposer $0 < a, b < +\infty$.

L'inégalité exprime exactement le fait que le logarithme est concave.

On peut aussi en donner une démonstration sans utiliser le logarithme. Introduisons la fonction C^1 sur $]0, +\infty[$ définie par

$$\forall a > 0, \quad f(a) = a^{1-t}b^t - (1 - t)a - tb. \tag{3}$$

On a $f'(a) = (1 - t)a^{-t}b^t - (1 - t) = (1 - t)\left[\left(\frac{b}{a}\right)^t - 1\right]$. Ainsi, f est (strictement) croissante sur $]0, b]$ et (strictement) décroissante sur $[b, +\infty[$, ce qui veut dire qu'elle atteint son maximum en $a = b$. Or $f(b) = 0$ et donc $f \leq 0$. La stricte monotonie assure que f ne s'annule que en b. $\square$

Pour $\lambda > 0$, on peut multiplier a par $\lambda^{1/(1-t)}$ et b par $1/\lambda^{1/t}$ et on obtient le

Lemme 0.7 (Inégalité arithmético-géométrique bis)**.** *Soit* $t \in]0,1[$. *Alors, pour* $a, b \in [0,+\infty]$ *et* $\lambda > 0$ *on a*

$$a^{1-t}b^t \leq (1-t)\lambda^{\frac{1}{1-t}}a + t\frac{b}{\lambda^{\frac{1}{t}}} \tag{4}$$

et, lorsque a *et* b *sont finis, il y a égalité si et seulement si* $\lambda^{\frac{1}{1-t}}a = \frac{b}{\lambda^{\frac{1}{t}}}$.
En particulier, si $0 < a, b < +\infty$, *il y a égalité pour un certain* $\lambda > 0$, *ce qui peut se résumer par :*

$$\inf_{\lambda>0}\left((1-t)\lambda^{\frac{1}{1-t}}a + t\frac{b}{\lambda^{\frac{1}{t}}}\right) = \min_{\lambda>0}\left((1-t)\lambda^{\frac{1}{1-t}}a + t\frac{b}{\lambda^{\frac{1}{t}}}\right) \;=\; a^{1-t}b^t. \tag{5}$$

Démonstration. Il suffit donc de prendre $\lambda := \left(\frac{b}{a}\right)^{t(1-t)}$. □

On préfère parfois introduire $p = \frac{1}{1-t} \in]1,+\infty[$ et $q = \frac{1}{t} \in]1,+\infty[$, qui vérifient $\frac{1}{p} + \frac{1}{q} = 1$ –on dit que ces nombres sont *conjugués*[1]–, et remplacer a par a^p et b par b^q, de sorte qu'on trouve le

Lemme 0.8 (Inégalité arithmético-géométrique bis bis)**.** *Soit* $p, q > 1$ *avec* $\frac{1}{p} + \frac{1}{q} = 1$. *Alors, pour* $a, b \in [0,+\infty]$ *et* $\lambda > 0$ *on a*

$$ab \leq \lambda^p\, a^p/p + \lambda^{-q}\, b^q/q \tag{6}$$

et, lorsque a *et* b *sont finis, il y a égalité si et seulement si* $\lambda^p a^p = \lambda^{-q} b^q$.
En particulier, si $0 < a, b < +\infty$, *il y a égalité pour un certain* $\lambda > 0$, *ce qui peut se résumer par :*

$$\inf_{\lambda>0}\left(\lambda^p a^p/p + t\lambda^{-q}b^q/q\right) = \min_{\lambda>0}\left(\lambda^p a^p/p + t\lambda^{-q}b^q/q\right) \;=\; ab. \tag{7}$$

Sous cette dernière forme, l'inégalité arithmético-géométrique s'appelle aussi *inégalité d'Young*. L'inégalité la plus simple et la plus classique (liée à Cauchy-Schwartz) correspond à $t = 1/2$, c'est-à-dire $p = q = 2$.

Inégalités de Hölder

Proposition 0.9 (Inégalité de Hölder)**.** *Soit* $f, g : X \to [0,+\infty]$ *deux fonctions mesurables positives et* $t \in]0,1[$. *Alors on a*

$$\int f^{1-t}g^t\, d\mu \leq \Big(\int f\, d\mu\Big)^{1-t}\Big(\int g\, d\mu\Big)^t.$$

De plus, si $0 < \int f\, d\mu, \int g\, d\mu < +\infty$, *il y a égalité si et seulement si il existe* $\lambda > 0$ *tel que* $g = \lambda f$ μ*-pp.*

Remarque 0.10. *Vous noterez que l'inégalité est bien homogène en* f *et* g.

Remarque 0.11. *Pour une fonction positive* f *et* $r > 0$ *on a*

$$\int f^r\, d\mu > 0 \Longleftrightarrow \mu(\{f \neq 0\}) = \mu(\{x \in X\ ;\ f(x) \neq 0\}) \neq 0,$$

ce qui veut dire qu'il existe $A \in \mathcal{S}$, *avec* $\mu(A) \neq 0$ *tel que* $f > 0$ *sur* A. *On écrit aussi "*$f \not\equiv 0$ μ*-pp" qu'il faut comprendre comme "*f *n'est pas égale à une fonction nulle* μ*-pp".*

1. par exemple $p = q = 2$ ou $p = 1$ et $q = +\infty$ (ici non considéré)

Démonstration. D'abord, on remarque que l'inégalité devient une égalité lorsque $f = g$.
Si $\int f\,d\mu = +\infty$, il n'y a rien à montrer. De même, si $\int f\,d\mu = 0$ alors $f = 0$ μ-pp et l'inégalité est triviale. Idem avec g. On supposera donc que $0 < \int f\,d\mu, \int g\,d\mu < +\infty$.
On va utiliser deux fois le Lemme 0.7. Tout d'abord, pour tout $\lambda > 0$ et tout $x \in X$ on a

$$f(x)^{1-t}g(x)^t \leq (1-t)\lambda^{\frac{1}{1-t}} f(x) + t\lambda^{-\frac{1}{t}} g(x)$$

et donc en intégrant on trouve que pour tout $\lambda > 0$,

$$\int f^{1-t}g^t\,d\mu \leq (1-t)\lambda^{\frac{1}{1-t}} \int f\,d\mu + t\lambda^{-\frac{1}{t}} \int g\,d\mu.$$

En prenant l'infimum sur les λ, ou trouve donc bien $\int f^{1-t}g^t\,d\mu \leq \Big(\int f\,d\mu\Big)^{1-t}\Big(\int g\,d\mu\Big)^t$.
Pour la réciproque, on suppose donc que les intégrales sont non-nulles et finies. On reprend la démonstration ci-dessous mais au lieu de prendre l'infimum sur les λ, on suppose qu'on a pris, dès le début le $\lambda = \lambda_0 > 0$ optimal pour lequel l'infimum est atteint (la valeur valeur exacte, dont on n'a pas besoin, est $\lambda_0 = \Big(\frac{\int g\,d\mu}{\int f\,d\mu}\Big)^{t(1-t)} > 0$). Alors on a

$$\Big(\int f\,d\mu\Big)^{1-t}\Big(\int g\,d\mu\Big)^t - \int f^{1-t}g^t\,d\mu = \int \Big[(1-t)\lambda_0^{\frac{1}{1-t}} f(x) + t\lambda_0^{-\frac{1}{t}} g(x)^t - f(x)^{1-t}g(x)^t\Big]\,d\mu.$$

Le terme sous l'intégrale est positif, et donc pour que son intégrale soit nulle, il faut qu'il soit nul μ-pp. Mais par l'étude des cas d'égalité dans l'inégalité arithmético-géométrique, cela implique que $\lambda_0^{\frac{1}{1-t}} f = \lambda_0^{-\frac{1}{t}} g$ μ-pp. ☐

Il y a beaucoup de formulations équivalentes de l'inégalité de Hölder. En voici une pour ceux qui préfèrent les p, q. On rappelle que $p/q = p - 1$.

Proposition 0.12 (Inégalité de Hölder)**.** *Soit $p, q \in]1, +\infty[$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$ et $f, g : X \to [0, +\infty]$ deux fonctions mesurables positives sur X. Alors,*

$$\int fg\,d\mu \leq \Big(\int f^p\,d\mu\Big)^{1/p}\Big(\int g^q\,d\mu\Big)^{1/q}.$$

avec égalité lorsque $g = f^{p-1}$. De plus, lorsque $0 < \int f^p\,d\mu, \int g^q\,d\mu < +\infty$ il y a égalité si et seulement si il existe $\lambda > 0$ tel que $g^q = \lambda f^p$ (i.e. $g = \tilde{\lambda}\, f^{p-1}$ pour un $\tilde{\lambda} \geq 0$).

Démonstration. On applique la proposition précédente en remplaçant $(1-t)$ par $\frac{1}{p}$, et donc t par $\frac{1}{q}$, et en l'appliquant à f^p à la place de f et g^q à la place de g. ☐

Le cas le plus rencontré est le cas $p = q = 2$, et l'inégalité s'appelle alors *inégalité de Cauchy-Schwartz.*

Remarque 0.13. *Le cas du couple $p = 1$ et $q = \infty$ est trivial et s'énonce comme suit : si f, g sont deux fonctions mesurables positives sur X, alors*

$$\int fg\,d\mu \leq \big(\sup g\big) \int f\,d\mu.$$

Remarque 0.14. *Une conséquence de l'inégalité de Hölder est que si on se donne* p *et* q *dans* $]1, +\infty[$ *avec* $\dfrac{1}{p} + \dfrac{1}{q} = 1$ *et* $f : X \to [0, +\infty]$ *une fonction mesurable positive avec* $\int f^p \, d\mu < +\infty$*, alors*

$$\Big(\int f^p \, d\mu \Big)^{1/p} = \sup_{g \geq 0} \frac{\displaystyle\int fg \, d\mu}{\Big(\displaystyle\int g^q \, d\mu \Big)^{1/q}} = \sup_{g \geq 0, \, \int g^q \, d\mu \leq 1} \int fg \, d\mu.$$

où les sup sont pris sur les fonction mesurables positives telles que $0 < \int g^q \, d\mu < +\infty$*). De plus ce sup est atteint. En effet, l'inégalité de Hölder montre que* $\dfrac{\displaystyle\int fg \, d\mu}{\Big(\displaystyle\int g^q \, d\mu \Big)^{1/q}} \leq \Big(\int f^p \, d\mu \Big)^{1/p}$*, et donc* idem *pour le sup sur* g*. On voit par ailleurs qu'il y a égalité si* $g = f^{p-1}$ *par exemple (si* f *est non nulle ; si* f *est nulle* μ*-pp, on prend n'importe que* g*), ce qui donne à la fois l'égalité voulue, et le fait que le sup est atteint. La deuxième inégalité découle par homogénéité.*
On retrouvera une formule similaire lors de l'étude de la dualité $L^p - L^q$*.*

Inégalité de Minkowski

Proposition 0.15. *Soit* $p \in]1, +\infty[$*. Si* $f, g : X \to [0, +\infty]$ *sont deux fonctions mesurables positives sur* X*, alors*

$$\Big(\int (f+g)^p \, d\mu \Big)^{1/p} \leq \Big(\int f^p \, d\mu \Big)^{1/p} + \Big(\int g^p \, d\mu \Big)^{1/p}.$$

Si $0 < \int f^p \, d\mu, \int g^p \, d\mu < +\infty$*, alors il y a égalité si et seulement si il existe* $\lambda > 0$ *tel que* $g = \lambda f$ μ*-pp.*

Démonstration. Soit $q > 1$ tel que $\frac{1}{p} + \frac{1}{q} = 1$. On rappelle que $q(p-1) = p$. Si f ou g est nulle μ-pp, il n'y a rien à montrer ; on supposera donc que ce n'est pas le cas. Idem si l'une des intégrales de droite vaut $+\infty$. On suppose donc les intégrales du terme de droite sont finies et que l'intégrale du terme de gauche est non-nulle.
On a

$$\forall a, b \in [0, +\infty], \qquad (a+b)^p \leq 2^{p-1}(a^p + b^p). \tag{8}$$

On montrera cette inégalité plus loin. Cela permet de voir, en l'appliquant à $a = f(x)$ et $b = g(x)$ et en intégrant sur X par rapport à $d\mu(x)$ que si les intégrales de droites sont finies, l'intégrale de gauche aussi. Alors, par le cas d'égalité (trivial) dans l'inégalité de Hölder, on sait qu'il existe $H \geq 0$ tel que

$$\Big(\int (f+g)^p \, d\mu \Big)^{1/p} = \int (f+g) H \, d\mu \quad \text{et} \quad \int H^q \, d\mu = 1.$$

De façon explicite $H = \dfrac{1}{\big(\int (f+g)^p \, d\mu \big)^{1/q}} (f+g)^{p-1}$ sur X. On a donc,

$$\Big(\int (f+g)^p \, d\mu \Big)^{1/p} = \int fH \, d\mu + \int gH \, d\mu \leq 1 \times \Big(\int f^p \, d\mu \Big)^{1/p} + 1 \times \Big(\int g^p \, d\mu \Big)^{1/p},$$

où l'on a utilisé deux fois l'inégalité de Hölder. Cela montre l'inégalité voulue.
On voit qu'il y a égalité si $g = \lambda f$ μ-pp. Réciproquement, pour qu'il y ait égalité, il faut que dans la preuve ci-dessus, il y ait égalité dans les deux inégalités de Hölder utilisées. Et donc, il faut que μ-pp $f = \lambda_1 H^{p-1}$ et $g = \lambda_2 H^{p-1}$, avec $\lambda_1, \lambda_2 > 0$. Donc il faut que $g = \lambda f$ pour un certain $\lambda > 0$. $\square$

Convexité et inégalité de Jensen

Soit I un intervalle de $\mathbb{R}$. Le cas le plus important, de loin, est celui où $I = \mathbb{R}$.

On dit qu'une fonction $f : I \to \mathbb{R}$ est convexe si pour tout $x, y \in I$ et $t \in [0,1]$ on a

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

Par associativité du barycentre, cette propriété est équivalente à la forme plus générale suivante : pour $m \geq 1$, $x_1, \ldots, x_m \in I$ et $\alpha_1, \ldots, \alpha_m \in [0,1]$ tels que $\alpha_1 + \ldots \alpha_m = 1$,

$$f\big(\alpha_1 x_1 + \ldots + \alpha_m x_m\big) \leq \alpha_1 f(x_1) + \ldots + \alpha_m f(x_m).$$

Une fonction f est dite concave sur I si les inégalités précédentes ont lieu dans l'autre sens, c'est-à-dire si $-f$ est convexe sur I.

Si f est convexe sur I, alors on voit que pour tout $s, t, u \in I$

$$s < t < u \Longrightarrow \frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(t)}{u - t}. \tag{9}$$

En fait, la propriété (9) est équivalente à la convexité de f sur I. En effet, pour $s, u \in I$, $s < u$, et $t \in]s, u[$, introduisons $r \in [0,1]$ tel que

$$t = (1-r)s + ru,$$

à savoir $r = \frac{t-s}{u-s}$ et donc $1 - r = \frac{u-t}{u-s}$. On a alors

$$\begin{aligned} \frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(t)}{u - t} \quad &\Leftrightarrow \quad [(u-t) + (t-s)]f(t) \leq (u-t)f(s) + (t-s)f(u) \\ &\Leftrightarrow \quad f(t) \leq (1-r)f(s) + rf(u) \end{aligned}$$

On peut tirer plusieurs propriétés intéressantes de l'inégalité (9).

Proposition 0.16. *Soit f une fonction dérivable (en pratique C^1) sur un intervalle ouvert I. Alors f est convexe sur I si et seulement si f' est croissante sur I.*

En particulier, une fonction deux fois dérivable sur I est convexe sur I si et seulement si $f'' \geq 0$ sur I.

Démonstration. Supposons d'abord que f convexe. Soit $s, v \in I$ avec $s < v$. Alors pour tout $t, u \in I$ tels que $s < t < u < v$ on a, en appliquant deux fois l'inégalité (9),

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(t)}{u - t} \leq \frac{f(v) - f(u)}{v - u}. \tag{10}$$

En faisant $t \to s^+$ et $u \to v^-$, on trouve donc que $f'(s) \leq f'(v)$.

Réciproquement, supposons f' est croissante, et soit $x, y \in I$, $x < y$, et $z \in]x, y[$. Par le théorème des accroissement finis, il existe $\alpha \in [x, z]$ tel que $f(z) - f(x) = (z - x)f'(\alpha)$ et $\beta \in [z, y]$ tel que $f(y) - f(z) = (y - z)f'(\beta)$. Comme $\alpha \leq \beta$ on a $f'(\alpha) \leq f'(\beta)$, ce qui donne la propriété (9) pour le triplet $x < z < y$. $\square$

Remarque 0.17. *La propriété (9) sous la forme (10) montre qu'en tout point, f admet une dérivée à gauche et une dérivée à droite, que ces dérivées sont croissantes, et que $f'_g \leq f'_d$ en tout point.*

Proposition 0.18. *Soit f une fonction convexe sur un intervalle I. Alors en tout point x_0 de l'intérieur de I, il existe $\beta_0 \in \mathbb{R}$ tel que*

$$\forall x \in I, \qquad f(x) \geq f(x_0) + \beta_0(x - x_0).$$

Démonstration. Pour tout $s, u \in I$ tel que $s < x_0 < u$, on a

$$\frac{f(x_0) - f(s)}{x_0 - s} \leq \frac{f(u) - f(x_0)}{u - x_0}.$$

Notons $\beta = \sup_{s \in I, s < x_0} \frac{f(x_0) - f(s)}{x_0 - s}$. C'est un nombre fini, car la quantité est majorée par n'importe quel $\frac{f(u) - f(x_0)}{u - x_0}$ avec $u > x_0$.

Pour $x \in I$, on a, lorsque $x < x_0$, $f(x_0) - f(x) \leq \beta(x_0 - x)$ par définition de β, et lorsque $x > x_0$, $\beta(x - x_0) \leq f(x) - f(x_0)$ d'après l'inégalité ci-dessus (avec $u = x$). □

Géométriquement, on a donc qu'en tout point x_0 de l'intérieur de I, on peut trouver une droite tangente au graphe de I en x_0, telle que f reste au dessus de cette droite. Lorsque f est dérivable, on peut (et on doit) prendre $\beta = f'(x_0)$. Dans ce cas on a,

$$\forall x \in I, \qquad f(x) \geq f(x_0) + f'(x_0)(x - x_0).$$

Exemple 0.19. La fonction log est concave sur $]0, +\infty[$, en effet sa dérivée seconde vaut $\log''(x) = -1/x^2 \leq 0$ pour tout $x > 0$. On a donc, pour tout $a, b > 0$,

$$\log((1 - t)a + tb) \geq (1 - t)\log(a) + t\log(b) = \log(a^{1-t}b^t)$$

ce qui s'écrit encore, en prenant l'exponentielle (qui est croissante) : $(1 - t)a + tb \geq a^{1-t}b^t$.

Exemples de fonctions convexes sur $\mathbb{R}$: $x \to |x|$ (inégalité triangulaire sur $\mathbb{R}$), $x \to e^x$, $x \to x^2$, et plus généralement $x \to |x|^p$, avec $p > 1$. Cet exemple permet de voir que pour $a, b \geq 0$ on a $\left(\frac{a+b}{2}\right)^p \leq \frac{1}{2}a^p + \frac{1}{2}b^p$, ce qui donne (8).

Théorème 0.1 (Inégalité de Jensen)**.** *Soit $\varphi : I \to \mathbb{R}$ une fonction convexe sur un intervalle I de $\mathbb{R}$ (en général $I = \mathbb{R}$ ou $\mathbb{R}^+$). Si μ est une* probabilité *sur un espace mesurable $(X, \mathcal{A})$ et f une fonction μ-intégrable à valeurs dans un intervalle I, alors $\int_X f\,d\mu \in I$, $\int_X \varphi(f)\,d\mu$ existe dans $]-\infty, +\infty]$ et*

$$\int_X \varphi(f)\,d\mu \geq \varphi\left(\int_X f\,d\mu\right).$$

avec égalité si f est constante (i.e. si $\exists c \in I$ tel que $f \equiv c$ μ-pp)

Démonstration. Soient $a := \inf(I) \in \mathbb{R} \cup \{-\infty\}$ et $b := \sup(I) \in \mathbb{R} \cup \{+\infty\}$, de sorte que l'intérieur de I est $]a, b[$. Ayant $a \leq f \leq b$, comme μ est une probabilité,

$$a = \int_X a\,d\mu \leq \int_X f\,d\mu \leq \int_X b\,d\mu = b,$$

donc $m := \int_X f\,d\mu \in I$. De plus, dans l'éventualité où $a \in I$, si $m = a$, on doit avoir $f = a$ μ-pp, puisque $f \geq a$ et $\int (f - a)\,d\mu = 0$, et alors l'inégalité à montrer devient triviale. Même chose si $m = b$. On supposera donc dans la suite que m appartient à l'intérieur de I.

Comme φ est convexe, il existe au moins une droite située en-dessous du graphe de φ et passant par $(m, \varphi(m))$, d'équation $y = \beta(x - m) + \varphi(m)$. Ceci se traduit par

$$\varphi(u) \geq \beta(u - m) + \varphi(m) \qquad u \in I,$$

et donc pour tout $x \in X$,

$$\varphi \circ f(x) \geq \beta(f(x) - m) + \varphi(m).$$

La fonction $\varphi \circ f$ étant minorée par une fonction intégrable, elle admet une intégrale (qui ne peut être égale à $-\infty$) et

$$\int_X \varphi \circ f \, d\mu \geq \beta \int_X (f - m) \, d\mu + \int_X \varphi(m) \, d\mu = 0 + \varphi(m),$$

par linéarité (pour les fonctions μ-intégrables), et parce que μ est une probabilité. □

Remarque 0.20. *En général, il peut y avoir d'autres cas d'égalité que les fonctions constantes.Par exemple, si φ est constante ou plus généralement, affine, il y a égalité pour toute fonction f. Par contre, les fonctions constantes sont les seuls cas d'égalité si φ est strictement convexe. La fonction φ est dite* strictement convexe *sur I si pour $s, t \in I$, $s \neq t$ et $\lambda \in]0, 1[$, on a*

$$\varphi((1 - \lambda)s + \lambda t) < (1 - \lambda)\varphi(s) + \lambda\varphi(t).$$

Cela équivaut à dire que le graphe de φ ne contient pas de segment. En particulier, toute tangente ne touche le graphe qu'en un seul point.

0.3 Espace $\mathcal{L}^p$ et espace L^p

Si f est une fonction mesurable à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, on note, pour $p \in [1, +\infty[$

$$\| f \|_p := \left(\int_E | f |^p \, d\mu \right)^{\frac{1}{p}},$$

qu'on appelle[2] *norme* $\mathcal{L}^p$ de f, et lorsque $p = \infty$,

$$\| f \|_\infty := \inf\{a > 0 \; ; \; \mu(\{| f | \geq a\}) = 0\},$$

qu'on appelle[3] *supremum essentiel* de f.

Définition 0.21 ($p \in [1, +\infty[$)**.** On note $\mathcal{L}^p(E, \mathcal{S}, \mu)$, ou $\mathcal{L}^p(\mu)$, l'ensemble de toutes les fonctions mesurables à valeurs dans ans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ telles que $| f |^p$ est μ-intégrable, i.e. telles que $\|f\|_p < +\infty$.

Définition 0.22 ($p = \infty$)**.** On note $\mathcal{L}^\infty(E, \mathcal{S}, \mu)$, ou $\mathcal{L}^\infty(\mu)$, l'ensemble de toutes les fonctions mesurables f à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ qui sont *μ-essentiellement bornées*, c'est-à-dire telles qu'il existe $a > 0$ pour lequel $\mu(\{| f | \geq a\}) = 0$, soit encore telles que $\|f\|_\infty < +\infty$.

On remarquera que pour $f \in \mathcal{L}^\infty(\mu)$ on a

$$\mu(\{|f| > \|f\|_\infty\}) = 0.$$

En effet, on a $\mu(\{|f| > \|f\|_\infty\}) = \mu\Big(\bigcup_{n \geq 1} \Big\{|f| \geq \|f\|_\infty + \dfrac{1}{n}\Big\}\Big)$ et on conclut par convergence monotone. En particulier, pour toute partie mesurable A on a $\mu(A) = \mu(A \cap \{|f| \leq \|f\|_\infty\})$.

Remarque 0.23. *Si on éprouve le besoin de préciser que l'on travaille avec des fonctions réelles ou des fonctions complexes, on peut ajouter "espace $\mathcal{L}^p$-réel" ou "espace $\mathcal{L}^p$-complexe".*

2. mais dont nous verrons qu'il ne s'agit en fait que d'une semi-norme
3. mais on devrait dire μ-supremum essentiel

Remarque 0.24. *Dans les deux définitions ci-dessus, on peut autoriser la fonction $|f|$ à prendre la valeurs $+\infty$ (en particulier on peut considérer des fonctions à valeurs dans $\mathbb{R} \cup \{\pm\infty\}$). Cela ne change rien du point de vue de l'intégration par rapport à μ, car pour une fonction dans $\mathcal{L}^p$, cela ne peut avoir lieu que sur un ensemble de μ-mesure nulle. En effet, si p est fini, $|f|^p$ μ-intégrable entraîne que $\mu(\{|f| = +\infty\}) = 0$. Pour $p = +\infty$, si $\|f\|_\infty < +\infty$, cela veut dire qu'il existe $a > 0$ fini tel que $\mu(\{|f| \geq a\}) = 0$, et $\mu(\{|f| = +\infty\}) \leq \mu(\{|f| \geq a\})$.*

Proposition 0.25. *Pour tout $p \in [1, +\infty]$, on a, pour $\lambda \in \mathbb{K}$ et $f, g \in \mathcal{L}^p(\mu)$, on a*

1. $\|\lambda f\|_p = |\lambda| \|f\|_p$, *et*
2. $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

En particulier, $\mathcal{L}^p(\mu)$ est un espace vectoriel.

Démonstration. Le premier point est évident par linéarité de l'intégrale si $p < \infty$. Si $p = +\infty$,

$$\begin{aligned}\| af \|_\infty &= \inf\{m > 0 : \mu(\{| af | \geq m\}) = 0\} = | a | \inf\{m' > 0 : \mu(\{| af | \geq | a | m'\}) = 0\} \\ &= | a | \, inf\{m' > 0 : \mu(\{| f | \geq m'\}) = 0\} = | a | \| f \|_\infty.\end{aligned}$$

Le deuxième point est évident pour $p = 1$ à partir de l'inégalité $|f + g| \leq |f| + |g|$. Pour $p \in]1, +\infty[$, on combine cela avec l'inégalité de Minkowski. Pour le cas $p = \infty$, on remarque que si $a > \|f\|_\infty + \|g\|_\infty$ on a,

$$\mu(\{|f + g| \geq a\}) \leq \mu(\{|f| + |g| \geq a\}) = \mu\Big(\{|f| + |g| \geq a\} \cap \{|f| \leq \|f\|_\infty\} \cap \{|g| \leq \|g\|_\infty\}\Big) = \mu(\emptyset) = 0,$$

et donc $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. □

Par ailleurs, si f est la fonction nulle, on a $\|f\|_p = 0$. Alors que manque-t-il à $\| \cdot \|_p$ pour être une norme sur $\mathcal{L}^p$? Pas grand chose, mais le problème vient du fait que pour $f \in \mathcal{L}^p$ on a

$$\|f\|_p = 0 \Longleftrightarrow f = 0 \quad \mu\text{-p.p.}$$

Cela est clair pour $p < +\infty$, puisque dire que la fonction positive $|f|^p$ a une intégrale nulle, cela veut dire qu'elle est nulle μ-pp. Pour $p = \infty$, si $\|f\|_\infty = 0$, alors $\mu(\{|f| > 0\}) = \mu(\bigcap_{n\geq 1} \{|f| \geq \frac{1}{n}\}) = 0$, par convergence monotone.

Ainsi, on veut construite un espace tel que f nulle μ-presque partout veut dire que f est le vecteur nul. Pour cela, on fait le quotient de $\mathcal{L}^p$ par la relation d'équivalence suivante

$$f \sim g \Longleftrightarrow f = g \quad \mu\text{-p.p.} \Longleftrightarrow \|f - g\|_p = 0.$$

Ainsi, on considère l'ensemble quotient $\mathcal{L}^p(\mu)/\sim$ (que l'on notera $L^p(\mu)$) formé par les classes d'équivalences modulo $\sim$. Notez que la relation d'équivalence associée à chaque $\| \cdot \|_p$ ne dépend pas de p et est la même pour tous les espaces $\mathcal{L}^p$: la classe d'une fonction f est constituée par les fonctions qui coïncident avec f μ-presque partout.

Si on note $\mathcal{N}$ l'ensemble des fonctions mesurables nulles μ-presque partout, on peut aussi écrire

$$f \sim g \Longleftrightarrow f - g \in \mathcal{N}.$$

Or $\mathcal{N}$ est un espace vectoriel (et un sous-espace vectoriel de tout $\mathcal{L}^p(\mu)$), et il est classique de voir que les structures d'espace vectoriel passent au quotient. En résumé, on obtient la

Définition 0.26. Pour $p \in [1, +\infty]$, on note $L^p(E, \mathcal{S}, \mu)$, ou $L^p(\mu)$, l'ensemble des classes d'équivalence des éléments de $\mathcal{L}^p(\mu)$ par la relation d'équivalence définie par l'égalité μ-p.p.
Soit $\tilde{f} := \{g \ ; \ g = f \ \mu\text{-p.p.}\}$ la classe d'équivalence de f. Les opérations classiques s'étendent aux classes d'équivalence, avec $\widetilde{af} = a\tilde{f}$ et $\widetilde{f+g} = \overline{f} + \overline{g}$.
On peut également définir $\| \cdot \|_p$ sur $L^p(\mu)$ par $\| \tilde{f} \|_p = \| f \|_p$, qui ne dépend pas du représentant choisi, car $f = g$ μ-p.p. implique $\| f \|_p = \| g \|_p$.

Remarque 0.27. *On fera systématiquement l'abus de notation qui consiste à ne pas différencier fonctions et classes d'équivalences, c'est-à-dire à utiliser le même symbole pour une fonction f et pour sa classe d'équivalence $\overline{f}$.*
C'est une question d'habitude. La seule manière de comprendre L^p, c'est de l'utiliser. En fait, sur $L^p(\mu)$ on pense plutôt "$\mathcal{L}^p(\mu)$", c'est-à-dire à des fonctions, plutôt qu'à des classes d'équivalences, mais on se souvient que les objets ne sont définis que μ-pp. Ainsi, par exemple, on a coutume de dire que deux fonctions f et g sont égales dans L^p si elles coïncident μ-pp (même si on devrait simplement dire qu'elle définissent la même classe d'équivalence dans $L^p(\mu)$).

On a alors immédiatement ce que l'on cherchait.

Théorème 0.28. *L'ensemble $(L^p(\mu), \| \cdot \|_p)$ est un espace vectoriel normé.*

Exemple 0.29. On note $\ell_p(\mathbb{N})$ ou simplement ℓ_p l'espace $\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), m)$, où m est la mesure de comptage. On distingue parfois les espaces réels $\ell_p^{\mathbb{R}}(\mathbb{N})$et complexes $\ell_p^{\mathbb{C}}(\mathbb{N})$.
Soit $u \in \ell^p$. Si $p < \infty$, alors

$$\| u \|_p = \left(\sum_n | u_n |^p \right)^{\frac{1}{p}},$$

tandis que si $p = +\infty$,

$$\| u \|_\infty = \sup_n | u_n |.$$

Il n'est pas besoin ici de quotienter $\mathcal{L}^p$ car $\| u \|_p = 0$ implique $u = 0$.
On a la même chose pour $\ell^p(\mathbb{Z}) = \mathcal{L}^p(\mathbb{Z}, \mathcal{P}(\mathbb{Z}), m)$.

0.4 Convergence dans L^p et convergence simple

Rappelons que la topologie usuelle d'un espace vectoriel normé est la topologie relative à la distance $d(f, g) = \| f - g \|$. Ainsi on dira que la suite (f_n) converge (vers f) dans L^p si

a) pour tout $n \in \mathbb{N}$, $f_n \in L^p$ et $f \in L^p$;

b) $\lim_n \| f - f_n \|_p = 0$.

On rappelle que la suite (f_n) converge simplement vers f si $\lim_n f_n(x) = f(x)$ pour μ-presque tout x.
On remarque que si (f_n) converge dans $L^1(\mu)$ vers f, alors $\int f_n \, d\mu \to \int f \, d\mu$. La réciproque est fausse en général.
Le théorème de convergence dominée est généralement énoncé en terme de fonction intégrable, mais on peut aussi en donner une version (équivalente) L^p.

Proposition 0.30 (Convergence L^p-dominée)**.** *Soit $p \in [1, +\infty[$. Si $f_n \to f$ μ-p.p. et qu'il existe $g \in L^p$ tel que $| f_n | \leq g$ pour tout entier n, alors $f_n \overset{L^p}{\to} f$.*

Démonstration. On applique le théorème de convergence dominée. En effet, $| f_n - f |^p \leq (| f_n | + | f |)^p \leq 2^p | g |^p$ μ-p.p., et par hypothèse $| g |^p$ est intégrable, donc comme $| f_n - f |^p \to 0$, μ-p.p., on a la convergence vers 0 de $\int | f_n - f |^p \, d\mu$. $\square$

Proposition 0.31 (Extraction d'une sous-suite convergeant simplement)**.** *Soit $p \in [1, +\infty]$. Si $f_n \overset{L^p}{\to} f$, alors il existe une suite extraite de (f_n) qui converge vers f μ-p.p.*
Dans le cas $p = +\infty$, on a bien sûr beaucoup mieux : $f_n \to f$ uniformément en dehors d'un ensemble négligeable (en particulier, $f_n \to f$ μ-p.p., pas besoin de sous-suite).

Démonstration. On traite d'abord le cas $1 \leq p < +\infty$. Si (f_n) converge vers f dans $L^p(\mu)$, on peut trouver une sous-suite $(f_{n_k})_{k\geq 0}$ tel que pour tout $k \geq 1$,

$$\|f_{n_k} - f\|_p \leq 2^{-k}.$$

Introduisons la suite de fonctions positives $u_k = |f_{n_k} - f|^p$. Par le théorème de Beppo-Levi (convergence monotone) on a

$$\int \sum_{k\geq 0} u_k(x)\, d\mu(x) = \sum_{k\geq 0} \int u_k(x)\, d\mu(x) \leq \sum_{k\geq 0} (2^{1/p})^{-k} < +\infty.$$

Par conséquent, il existe un ensemble de mesure nulle $\mathcal{N}$ tel que $\forall x \in X \setminus \mathcal{N}$, $\sum_{k\geq 0} u_k(x) < +\infty$. Donc, pour $x \in X \setminus \mathcal{N}$ la série réelle $\sum u_k(x)$ est convergente, et donc son terme général tend vers zéro, c'est à dire $f_{n_k}(x) \to f(x)$.
Pour $p = +\infty$, c'est la définition de la convergence dans $L^\infty(\mu)$. En effet, si on introduit $A_n = \{|f - f_n| > \|f - f_n\|_\infty\}$ et $A = \cup A_n$, alors A est un ensemble de mesure nulle et

$$\forall x \in X \setminus A, \qquad |f(x) - f_n(x)| \leq \|f - f_n\|_\infty \to 0,$$

ce qui traduit le fait que f_n converge uniformément vers f sur $X \setminus A$. □

Exemple 0.32. Dans le cas de l'espace ℓ^p (pour $p < \infty$), une suite (de fonctions, aussi appelées suites ici...) $(u^{(n)})$ converge vers la fonction $u \in \ell^p$ si $\sum_k |\, u_k^{(n)} \,|^p < \infty$, si $\sum_k |\, u_k \,|^p < \infty$ et si

$$\lim_n \sum_k |\, u_k^{(n)} - u_k \,|^p = 0.$$

Ceci implique en particulier que $u_k^{(n)} \longrightarrow u_k$ lorsque $n \to \infty$. En conclusion, dans l'espace ℓ^p (vrai aussi si $p = +\infty$ par b)ii)),

$$f_n \overset{\ell^p}{\to} f \qquad \Longrightarrow \qquad f_n \to f \quad \text{simplement (partout).}$$

Évidemment, on n'a pas la réciproque, comme on peut le voir sur le contre-exemple $u^{(n)} = \mathbf{1}_{\{n\}}$. Alors la suite $(u^{(n)})$ converge simplement vers la fonction nulle car $u_k^{(n)} = 0$ pour tout $k > n$. Néanmoins pour tout n, la fonction $u^{(n)}$ est à distance 1 de la fonction nulle : $\|\, u^{(n)} - 0 \,\|_p = (\sum_k |\, u_k^{(n)} \,|^p)^{1/p} = 1$ pour tout p (même $p = \infty$), et donc ne converge pas vers la suite nulle dans ℓ^p. En effet, ici la plus petite fonction dominant la suite $(u^{(n)})$ est la fonction v constante à 1. Pour $p < \infty$, cette fonction n'est pas dans ℓ^p, donc on ne peut pas appliquer a). De plus, $v \in \ell^\infty$, ce qui montre aussi que la Proposition 0.30 n'est pas valide en général pour $p = \infty$.

Corollaire 0.33. *Soit $p \in [1, +\infty]$. Si l'on a la convergence de la suite (f_n) vers f dans L^p et vers g μ-p.p. alors f et g sont égales μ-p.p.*

Démonstration. On sait qu'il existe une suite extraite $(f_{\varphi(n)})$ qui converge μ-p.p. vers f. Or la suite (f_n) converge μ-p.p. vers g, donc la sous-suite $(f_{\varphi(n)})$ également. Ainsi $f = g$ μ-p.p. □

0.5 Complétude des espaces L^p

Théorème 0.34 (Théorème de Riesz–Fisher)**.** *Pour tout $p \in [1, +\infty]$, $L^p(\mu)$ est un espace de Banach.*

Démonstration. Soit (f_n) une suite de Cauchy de $L^p(\mu)$. On veut montrer qu'elle converge dans L^p. Remarquons qu'il suffit de montrer qu'une sous-suite $(f_{n_k})_k$ converge. En effet, si f est la limite de cette sous-suite on a alors pour tout n, k,

$$\|f_n - f\|_p \leq \|f - f_{n_k}\|_p + \|f_{n_k} - f_n\|_p,$$

et chaque terme peut être rendu petit, le premier en prenant k assez grand (par définition de la limite), et le deuxième en prenant k (puisque $n_k \geq k$) et n assez grands, par le caractère de Cauchy.
Le caractère de Cauchy nous permet de trouver une sous-suite (f_{n_k}) tel que

$$\forall k \geq 0, \qquad \|f_{n_{k+1}} - f_{n_k}\|_p \leq 2^{-k}.$$

Posons alors

$$u_0 = f_{n_0}, \qquad \text{et pour } k \geq 1 \;\; u_k = f_{n_k} - f_{n_{k-1}},$$

de sorte que pour $N \geq 0$, la somme partielle vérifie

$$U_N := u_0 + u_1 + \ldots + u_N = f_{n_N}.$$

On se demande donc si la série $\sum u_k$ converge dans $L^p(\mu)$.
Posons, pour (presque tout) $x \in E$, et $N \geq 0$,

$$V_N(x) = \sum_{k=0}^{N} |u_k(x)|.$$

Pour x fixé, c'est une suite croissante qui converge vers une limite que l'on note $V(x)$:

$$V(x) := \sum_{k=0}^{+\infty} |u_k(x)| \in [0, +\infty].$$

La suite croissante $V_N(x)^p$ converge elle vers $V(x)^p$ lorsque $N \to +\infty$ (en convenant que $(+\infty)^p = +\infty$) et comme

$$\int V_N(x)^p \, d\mu(x) = \Big\| \sum_{k=0}^{N} |u_j| \Big\|_p^p \leq \Big(\sum_{k=0}^{N} \|u_j\|_p \Big)^p \leq \Big(\|f_{n_0}\|_p + \sum_{k=0}^{+\infty} 2^{-k} \Big)^p =: M < +\infty,$$

on a par convergence monotone

$$\int V(x)^p \, d\mu(x) \leq M < +\infty.$$

Cela force l'ensemble $\mathcal{N} := \{V^p = +\infty\} = \{V = +\infty\}$ à être de mesure nulle. Pour $x \notin \mathcal{N}$, on a

$$V(x) = \sum_{k=0}^{+\infty} |u_k(x)| < +\infty,$$

ce que veut dire que la série dans $\mathbb{K}$, $\sum u_k(x)$ est elle aussi convergente, car absolument convergente ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est complet). Pour $x \notin \mathcal{N}$, on pose

$$f(x) = \sum_{k=0}^{+\infty} u_j(x) = \lim_{N \to +\infty} U_N(x) = \lim_{N \to +\infty} f_{n_N}(x)$$

la somme de cette série convergente. On peut poser $f(x) = 0$ pour $x \in \mathcal{N}$, si on veut, mais ce n'est pas nécessaire si on raisonne μ-pp. Notez que f est une fonction mesurable comme limite simple (presque partout) d'une suite de fonctions mesurables. Par ailleurs, on a μ-pp, par convergence simple,

$$|f|^p \leq \Big(\sum_{k=0}^{+\infty} |u_k|\Big)^p = V^p$$

et donc $f \in L^p(\mu)$. Il reste à montrer la convergence dans $L^p(\mu)$. On a que U_N converge simplement vers f et $|U_N| \leq V_N \leq V$. Comme $V \in L^p(\mu)$, on peut conclure par convergence dominée dans $L^p(\mu)$ que U_N converge vers f dans $L^p(\mu)$ dans le cas $p < +\infty$. On a donc bien montré que la sous-suite (f_{n_k}) convergeait dans $L^p(\mu)$. □

0.6 Quelques rappels de topologie

Définition 0.35. Soit X un ensemble. L'ensemble $\mathcal{T} \subset \mathcal{P}(X)$ est une *topologie* sur X

1. $\emptyset \in \mathcal{T}$ et $X \in \mathcal{T}$;
2. si $(V_i)_{i\in I}$, $V_i \in \mathcal{T}$, alors $\bigcup_{i\in I} V_i \in \mathcal{T}$;
3. si $n \geq 1$ et $(V_i)_{1\leq i\leq n}$, $V_i \in \mathcal{T}$, alors $\bigcap_{1\leq i\leq n} V_i \in \mathcal{T}$.

On dit alors que $(X, \mathcal{T})$ est un *espace topologique*. Les éléments de $\mathcal{T}$ sont appelés les *ouverts* de la topologie.

L'*intérieur* d'un ensemble E, notée $\overset{\circ}{E}$, est le plus grand ouvert contenu dans cet ensemble ; c'est aussi la réunion de tous les ouverts contenu dans cet ensemble. Les *fermés* sont les complémentaires des ouverts. L'*adhérence* d'un ensemble E, notée $\overline{E}$, est le plus petit fermé contenant cet ensemble ; c'est aussi l'intersection de tous les fermés contenant cet ensemble (par la définition des fermés et le 3 de la définition 0.35). On dit que V est un *voisinage* du point x s'il existe un ouvert U tel que $x \in U \subset V$.

Définition 0.36. Un ensemble $K \subset X$ est *compact* si, de tout recouvrement ouvert de K (i.e. une famille d'ouverts $(O_j)_{j\in J}$ tels que $K \subset \bigcup_{j\in J} O_j$), on peut extraire un sous recouvrement fini (i.e. il existe $j_1, \cdots, j_n$ dans J tels que $K \subset O_{j_1} \cup \cdots \cup O_{j_n}$).
Un ensemble est dit *relativement* compact s'il est inclus dans un compact.

Exercice 0.37. Montrer qu'un fermé contenu dans un compact est compact.

Remarque 0.38. *Il est important de noter que la compacité est une notion intrinsèque qui ne dépend que de la topologie trace. On rappelle que pour un sous-ensemble $F \subset X$ d'un espace topologique $(X, \mathcal{T})$, la topologie trace est donnée par $\mathcal{T}_F := \{V \cap F \; ; \; V \in \mathcal{T}\}$; ainsi $(F, \mathcal{T}_F)$ est un espace topologique. Etant donné un ensemble $K \subset X$ on a*

$$K \text{ compact dans } (X, \mathcal{T}) \Longleftrightarrow K \text{ compact dans } (K, \mathcal{T}_K).$$

En effet, étant donné une famille d'ouverts (V_i) de X, dire que les V_i recouvrent K équivaut à dire que les $V_i \cap K$ recouvrent K.

Cela justifie la pertinence de la notion (topologique) d'espace compact : on dit que K est un (espace) compact, si c'est un espace topologique pour lequel il est compact.

Définition 0.39. $(X, \mathcal{T})$ un espace topologique est dit séparé (ou de Hausdorff) si et seulement si pour tous $(x, y) \in X^2$, il existe U voisinage de x et V voisinage de y tels que $U \cap V = \emptyset$.

Théorème 0.40. *Soit X un espace de Hausdorff, $K \subset X$ un compact et $x \in X \setminus K$. Alors il existe U et V ouverts de X tels que $K \subset U$, $x \in V$ et $U \cap V = \emptyset$.*

On peut aussi dire que si K est compact et $x \notin K$, alors il existe un ouvert U tel que $K \subset U$ et $x \notin \overline{U}$.

Démonstration. Comme X est un espace de Hausdorff, pour chaque $y \in K$ il existe U_y voisinage de y et V_y voisinage de x tel que $U_y \cap V_y = \emptyset$. Alors $(U_y)_{y \in K}$ est un recouvrement ouvert de K dont on peut extraire le recouvrement fini $(U_{y_j})_{1 \leq j \leq J}$ (car K est compact). Ainsi $U = \bigcup_{1 \leq j \leq J} U_{y_j}$ est un ouvert contenant K et $V := \bigcap_{1 \leq j \leq J} V_{y_j}$ est un ouvert contenant x. On a clairement $V \cap U = \emptyset$. □

Le complémentaire d'un compact dans un espace de Hausdorff est ouvert comme réunion des voisinages ouverts de chacun de ses points construits par le théorème 0.40. On obtient ainsi

Corollaire 0.41. *Un sous ensemble compact d'un espace de Hausdorff est fermé.*

Exercice 0.42. Montrer que, dans un espace de Hausdorff, l'adhérence d'un sous-ensemble d'un compact est compacte.
Montrer que, dans un espace de Hausdorff, un ensemble est relativement compact (i.e. contenu dans un compact) si et seulement si son adhérence est compacte.

Théorème 0.43. *Soient $(K_i)_{i \in I}$ des compacts de X, un espace de Hausdorff tels que $\bigcap_{i \in I} K_i = \emptyset$. Alors il existe $\{i_1, \cdots, i_n\} \subset I$ tel que $\bigcap_{1 \leq j \leq n} K_{i_j} = \emptyset$.*

Démonstration. On définit l'ouvert $V_i = (K_i)^c$ et on choisit $i_0 \in I$. Comme K_{i_0} ne rencontre pas tous les $(K_i)_{i \neq i_0}$, $(V_i)_{i \neq i_0}$ est un recouvrement ouvert de K_{i_0}, on peut donc en extraire un sous-recouvrement fini $K_{i_0} \subset \bigcup_{1 \leq j \leq n} V_{i_j}$. Ainsi $\bigcap_{0 \leq j \leq n} K_{i_j} = K_{i_0} \cap \bigcap_{1 \leq j \leq n} K_{i_j} = \emptyset$. □

Définition 0.44. Un espace topologique $(X, \mathcal{T})$ est dit *localement compact* si tout point admet un voisinage compact, ce qui revient à dire qu'en tout point on peut trouver un voisinage ouvert relativement compact.

Théorème 0.45. *Soit U un ouvert de X, un espace de Hausdorff localement compact et $K \subset U$ un compact. Alors il existe un ouvert V relativement compact tel que $K \subset V \subset \overline{V} \subset U$.*

Démonstration. Comme tout point de K admet un voisinage compact, on peut trouver un recouvrement fini de K par des ouverts relativement compacts. La réunion de ces ouverts, disons, O, est ouverte relativement compacte. Si $U = X$, on pose $V := O$ et la preuve est achevée.
Si $U \neq X$, soit C le complémentaire de U. Par le théorème 0.40, pour tout $c \in C$, il existe V_c un ouvert tel que $K \subset V_c$ et $c \notin \overline{V_c}$. Ainsi si on pose $K_c := C \cap \overline{O} \cap \overline{V_c}$, K_c est compact et l'intersection $\bigcap_{c \in C} K_c$ est vide. Par le théorème 0.43, il existe $c_1, \cdots, c_n$, des points de C, tels que $C \cap \overline{O} \cap \overline{V_{c_1}} \cap \cdots \cap \overline{V_{c_n}} = \emptyset$. On pose alors $V := O \cap V_{c_1} \cap \cdots \cap V_{c_n}$ qui a les propriétés annoncées comme $\overline{V} \subset \overline{O} \cap \overline{V_{c_1}} \cap \cdots \cap \overline{V_{c_n}}$. □

Définition 0.46. Soit X un espace topologique et f une fonction de X à valeurs dans $\overline{\mathbb{R}}$ (où $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$).
- f est semi-continue inférieurement si, pour tout $\alpha \in \mathbb{R}$, $\{x;\ f(x) > \alpha\}$ est ouvert ;
- f est semi-continue supérieurement si, pour tout $\alpha \in \mathbb{R}$, $\{x;\ f(x) < \alpha\}$ est ouvert.

On vérifie alors que

1. une fonction à valeurs réelles est continue si et seulement si elle est à la fois semi-continue inférieurement et semi-continue supérieurement.
2. Les fonctions indicatrices d'ouverts sont semi-continues inférieurement, celles de fermés semi-continues supérieurement.
3. L'infimum d'une famille de fonctions semi-continues supérieurement est semi-continue supérieurement.
4. Le supremum d'une famille de fonctions semi-continues inférieurement est semi-continue inférieurement.

Exercice 0.47. Montrer que l'opposé d'une fonction semi-continue inférieurement (resp. supérieurement) est semi-continue supérieurement (resp. inférieurement).
Montrer qu'une somme de fonctions semi-continues inférieurement (resp. supérieurement) est semi-continue inférieurement (resp. supérieurement).

Définition 0.48. Soit $f : X \to \mathbb{C}$. Le support de f, noté supp$f$, est l'adhérence de $\{x \in X;\ f(x) \neq 0\}$.

On notera $\mathcal{C}_c(X)$ l'ensemble des fonctions $f : X \to \mathbb{C}$ continues à support compact. En pratique, on n'a souvent pas besoin de connaitre de manière exacte le support. On utilise plutôt :

Exercice 0.49. Soit X un espace de Hausdorff et $f : X \to \mathbb{C}$. Montrer que $f \in \mathcal{C}_c(X)$ si et seulement si f est continue sur X et il existe un compact $K \subset X$ tel que f est identiquement nulle en dehors de K.

L'ensemble $\mathcal{C}_c(X)$ un espace vectoriel (sur $\mathbb{C}$) pour l'addition usuelle des fonctions (et la multiplication par un scalaire). Mais c'est même une algèbre si on la munit du produit des fonctions.

Remarque 0.50. *En théorie de l'intégration, on travaille souvent avec l'image réciproque. Si $f : X \to Y$ et $B \subset Y$, on définit $f^{-1}(B) = \{x \in X\ ;\ f(x) \in B\}$. Si la notation peut prêter à confusion, il faut retenir que cet objet est très simple. Il est facile (d'un point de vue calculatoire/algoritmique) de tester si $x \in f^{-1}(B)$: il suffit de vérifier si $f(x) \in B$. Si vous voulez, on peut écrire*

$$1_{f^{-1}(A)}(x) = 1_{f(x)\in B}$$

où cette dernière expression doit se comprendre comme une fonction booléenne valant 1 si la condition est vérifiée et zéro sinon. Travailler avec les images réciproques est facile, et d'ailleurs toutes les propriétés imaginables sont vraies avec f^{-1} (genre image réciprique d'une union, d'une réunion, complémentaire, etc).

Il en va tout autrement de l'image directe d'un ensemble $A \subset X$,

$$f(A) = \{f(a)\ ;\ a \in A\} = \{y \in Y\ ;\ \exists x \in A,\ f(x) = y\}.$$

Il est difficile, voire impossible, de tester si un élément est dans l'image directe. L'image directe est un objet compliqué, et il faut faire attention que des propriétés élémentaires sont fausses pour l'image directe.

Notons que si $f : X \to \mathbb{R}$ est une application borélienne sur un espace mesurable $(X, \mathcal{A})$, l'image réciproque d'un borélien est par définition mesurable, alors que l'image d'une partie mesurable n'est pas nécessairement borélienne. Même pour une application continue de $\mathbb{R}$ dans $\mathbb{R}$, l'image d'un borélien n'est pas nécessairement un borélien.

En général, on ne peut rien dire de l'image (directe) d'un ouvert ou d'un fermé, même par une application continue. Cependant :

Proposition 0.51. *Soient X et Y des espaces topologiques et $f :\ X \to Y$ continue. Alors si $K \subset X$ est compact, $f(K)$ est compact dans Y.*

Exercice 0.52. Démontrer la proposition 0.51.

Notations. La notation $K \prec f$ désigne un compact K dans X et un fonction f dans $\mathcal{C}_c(X)$ tels que $0 \leq f \leq 1$ et $f_{|K} = 1$.
La notation $f \prec V$ désigne un ouvert V dans X et un fonction f dans $\mathcal{C}(X)$ tels que $0 \leq f \leq 1$ et $\operatorname{supp} f \subset V$.
On notera $K \prec f \prec V$ quand, à la fois, on a $K \prec f$ et $f \prec V$.

Théorème 0.53. *(Lemme d'Urysohn) Soit X un espace de Hausdorff localement compact, V un ouvert de X et $K \subset V$ un compact de X. Alors il existe $f \in \mathcal{C}_c(X)$ telle que $K \prec f \prec V$.*

Démonstration. On pose $r_1 = 0$ et $r_2 = 1$. Soit $\{r_i; i \geq 3\} =]0,1[\cap\mathbb{Q}$ (où $r_i \neq r_j$ si $i \neq j$). Grâce au théorème 0.45, on construit V_0 puis V_1 ouverts relativement compacts tels que

$$K \subset V_1 \subset \overline{V_1} \subset V_0 \subset \overline{V_0} \subset V. \tag{11}$$

Pour $n \geq 2$, supposons que l'on a construit $V_{r_1}, \cdots, V_{r_n}$ ouverts relativement compacts de telle façon que si $r_i < r_j$ alors $\overline{V_{r_j}} \subset V_{r_i}$. Soit $r_i = \max\{r_k; r_k < r_{n+1},\ 1 \leq k \leq n\}$ et $r_j = \min\{r_k; r_k > r_{n+1},\ 1 \leq k \leq n\}$. Par le théorème 0.45, on construit $V_{r_{n+1}}$ tel que

$$\overline{V_{r_j}} \subset V_{r_{n+1}} \subset \overline{V_{r_{n+1}}} \subset V_{r_i}.$$

Par récurrence, on construit ainsi une famille d'ouverts relativement compacts $(V_r)_{r\in]0,1[\cap\mathbb{Q}}$ tels que $K \subset V_1$, $\overline{V_0} \subset V$ et

$$s > r \Longrightarrow \overline{V_s} \subset V_r. \tag{12}$$

Pour $r \in]0,1[\cap\mathbb{Q}$ et $x \in X$, posons

$$f_r(x) = r\mathbf{1}_{V_r} = \begin{cases} r \text{ si } x \in V_r, \\ 0 \text{ sinon} \end{cases} \qquad \text{et} \quad g_r(x) = r + (1-r)\mathbf{1}_{\overline{V_r}} = \begin{cases} 1 \text{ si } x \in \overline{V_r}, \\ r \text{ sinon} \end{cases}. \tag{13}$$

Les remarques qui suivent la définition 0.46 nous disent que pour tout r, f_r est semi-continue inférieurement et g_r semi-continue supérieurement.
On pose alors

$$f = \sup_{r\in]0,1[\cap\mathbb{Q}} f_r \quad \text{et} \quad g = \inf_{r\in]0,1[\cap\mathbb{Q}} g_r. \tag{14}$$

Ainsi f est semi-continue inférieurement et g semi-continue supérieurement ; elles prennent clairement leur valeurs dans $[0,1]$. De plus, f est constante égale à 1 sur K et son support est contenu dans $\overline{V_0}$. Pour achever la preuve du lemme d'Urysohn, il suffit de démontrer que $f = g$.
On a $f \leq g$. En effet, on ne peut avoir $f_r(x) > g_s(x)$ que si $r > s$, $x \in V_r$ et $x \notin \overline{V_s}$. Or, $r > s$ implique $V_s \subset V_r$. Ainsi, on voit que pour tout (r,s), on a $f_r \leq g_s$. On en déduit que $f \leq g$.
Supposons maintenant qu'il existe x tel que $f(x) < g(x)$. Il existe donc des rationnels r et s tels que $f(x) < r < s < g(x)$. Comme $f(x) < r$, on a $x \notin V_r$; mais comme $g(x) > s$, on a $x \in \overline{V_s}$ ce qui contredit (12). On a donc $f = g$. Ceci achève la preuve du lemme d'Urysohn. □

On va maintenant utiliser le lemme d'Urysohn pour construire une partition continue de l'unité.

Théorème 0.54. *Soient $V_1, \cdots, V_n$ des ouverts de X un espace de Hausdorff localement compact et K un compact tel que $K \subset V_1 \cup \cdots \cup V_n$. Alors, pour $1 \leq i \leq n$, il existe $f_i \prec V_i$ telles que $K \prec f_1 + f_2 + \cdots + f_n$.*

Démonstration. Par le théorème 0.45, tout point $x \in K$ est contenu dans un ouvert W_x relativement compact d'adhérence contenue dans l'un des $(V_i)_{1 \leq i \leq n}$. On peut donc recouvrir K par un nombre fini de ces ouverts $K \subset W_{x_1} \cup \cdots \cup W_{x_m}$. Pour $1 \leq i \leq n$, on pose $F_i = \bigcup_{\substack{1 \leq j \leq m \\ \overline{W_{x_j}} \subset V_i}} \overline{W_{x_j}}$ qui est compact. Le lemme d'Urysohn nous donne alors $g_i \in \mathcal{C}_c(X)$ telle que $\mathbf{1}_{F_i} \leq g_i \leq \mathbf{1}_{F_i'}$ où F_i' est un compact de V_i. On pose

$$\begin{aligned} f_1 &= g_1, \\ f_2 &= (1 - g_1) g_2, \\ \vdots \quad & \quad \vdots \\ f_n &= (1 - g_1)(1 - g_2) \cdots (1 - g_{n-1}) g_n. \end{aligned} \tag{15}$$

Pour $1 \leq i \leq n$, on a $f_i \in \mathcal{C}_c(X)$ et $0 \leq f_i \leq \mathbf{1}_{F_i'}$. D'autre part,

$$f_1 + f_2 + \cdots + f_n = 1 - (1 - g_1)(1 - g_2) \cdots (1 - g_n).$$

Or comme $K \subset F_1 \cup \cdots \cup F_n$, $(1 - g_1)(1 - g_2) \cdots (1 - g_n)$ s'annule sur K ; ainsi $\mathbf{1}_K \leq f_1 + f_2 + \cdots .f_n$. Ceci complète la preuve du théorème 0.54. $\square$