

# Chapitre 1

## Rappels d'intégration

### 1.1 Généralités

#### 1.1.1 Remarque liminaire

Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré. Si  $f$  est une fonction de  $X$  dans  $[0, +\infty]$  qui est mesurable (à l'arrivée, on prendra toujours la tribu des boréliens ou la tribu de Lebesgue sur  $\mathbb{R}$  ou  $\mathbb{C} \simeq \mathbb{R}^2$ ), alors la quantité

$$\int_X f d\mu \in [0, +\infty]$$

est bien définie (vous souvenez-vous de comment on la construit ? Voir plus loin).

Par contre, si  $f$  est à valeurs réelles ou complexe, la quantité  $\int_X f d\mu$  n'a *a priori* pas de sens. Elle a un sens lorsque  $\int_X |f| d\mu < +\infty$  (on dit que  $f$  est  $\mu$ -intégrable), ce dont il faut s'assurer avant d'écrire l'intégrale.

De même, il est erroné de croire que pour deux fonctions mesurables  $f_1, f_2$  on a

$$\int (f_1 + f_2) d\mu = \int f_1 d\mu + \int f_2 d\mu.$$

Cependant, cela est vrai lorsque les fonctions sont positives, ou lorsqu'elles sont intégrables. Si seule l'une d'entre elles est intégrable, cela vous irait-il ?

#### 1.1.2 Rappels sur la construction de l'intégrale

Sur un espace mesuré  $(X, \mathcal{A}, \mu)$ , on commence par construire l'intégrale d'une fonction mesurable étagée (c'est-à-dire ne prenant qu'un nombre fini de valeurs) positive

$$f = \sum_{i=1}^k c_i 1_{A_i}$$

en posant

$$\int_X f d\mu = \sum_{i=1}^k c_i \mu(A_i).$$

Ici on autorise les  $c_i$  à prendre la valeur  $+\infty$ , i.e. la fonction  $f$  est à valeurs dans  $[0, +\infty]$ . On notera que par définition

$$\mu(A) = \int_X 1_A d\mu$$

pour toute partie mesurable  $A \subset X$ .

On passe ensuite aux fonctions positives (à valeurs dans  $[0, +\infty]$ ) mesurables à l'aide du résultat suivant.

**Proposition 1** (lemme fondamental d'approximation). *Pour toute  $f : X \rightarrow \mathbb{C}$  ou  $[-\infty, +\infty]$  mesurable, il existe une suite  $(f_n)$  de fonctions étagées convergeant simplement vers  $f$ . De plus,*

- a) *si  $f$  est positive, on peut choisir la suite  $(f_n)$  positive et croissante.*
- b) *si  $f$  est bornée, on peut choisir  $(f_n)$  de sorte que la convergence soit uniforme sur  $X$ .*

En fait, on montre la version plus précise suivante :

**Lemme 2.** *Soit  $f : X \rightarrow [0, 1]$  une fonction mesurable. Alors il existe une suite  $(A_n)$  d'ensembles mesurables tel que*

$$f = \sum_{n=1}^{+\infty} 2^{-n} 1_{A_n}.$$

*Démonstration.* Voit TD. □

Ainsi, si  $f$  est une fonction à valeurs dans  $[0, +\infty]$  on "sait" (il y a plusieurs points et propriétés à vérifier tout de même...) définir

$$\int_X f d\mu \in [0, +\infty]$$

Ensuite, on souhaite passer à une fonction réelle  $f$ , à valeurs dans  $[-\infty, +\infty]$ . Pour continuer, on fait l'hypothèse que la fonction positive  $|f|$  a une intégrale (comme définie précédemment) finie,

$$\int_X |f| d\mu < +\infty.$$

On décompose alors  $f$  en parties positive et négative,  $f = f_+ - f_-$  et on pose

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu.$$

Cela est possible car notre hypothèse garanti que les quantités  $\int f_+$  et  $\int f_-$  sont finies.

**Remarque 3** (Semi-intégrabilité). Si  $\int f_- d\mu < +\infty$ , on peut encore poser

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu \in ]-\infty, +\infty].$$

On dit que la fonction est semi-intégrable. Idem si  $\int f_+ < +\infty$ .

On passe ensuite de même à une fonction  $f$  valeurs complexes, lorsque le module de  $f$ ,  $|f|$ , est une fonction positive d'intégrale finie.

Bref, lorsque  $f$  est positive ou intégrable, on sait définir  $\int_X f d\mu$ .

### 1.1.3 Ce que vous devez connaître

- Les propriétés de base de l'intégrale.
- Les grands théorèmes de convergences : convergence monotone, convergence dominée, Lemme de Fatou.
- La définition du produit de deux tribus, de la mesure produit, et le théorème de Fubini (avec précision).

On pourra consulter le livre de Rudin pour ces points.

## 1.2 Entracte

Par construction, l'intégrale de Lebesgue d'une fonction  $f$  est intimement liée aux mesures des sous-ensembles de niveau  $\{f > t\}$  où

$$\{f > t\} := \{x \in X ; f(x) > t\}.$$

Remarquons tout d'abord que, pour une fonction positive  $f$ , la fonction  $t \rightarrow \mu(\{f > t\})$  est une fonction décroissante sur  $\mathbb{R}^+$ , qui est continue à droite (pourquoi?).

Si  $f$  est une fonction mesurable positive sur un espace mesuré  $(X, \mathcal{A}, \mu)$ , alors

$$\int_X f d\mu = \int_0^{+\infty} \mu(\{f > t\}) dt \tag{1.1}$$

où  $dt$  est la mesure de Lebesgue sur  $\mathbb{R}$ . Pour le voir, on peut commencer par étudier le cas d'une fonction étagée positive, et passer au cas général par convergence monotone (dans les deux intégrales).

Faisons ici une parenthèse sur les fonctions indicatrices. On a donc noté  $1_A$  la fonction indicatrice d'un ensemble  $A$ . Cette fonction est une fonction *booléenne* : elle ne prend que deux valeurs, 0 et 1. Réciproquement, toute fonction booléenne  $H$  sur un ensemble  $X$  s'écrit comme une fonction indicatrice,  $H = 1_A$  avec  $A = \{H = 1\}$ . Lorsqu'on est sur un espace mesurable, la mesurabilité de  $1_A$  équivaut à la mesurabilité de  $A$ . Le point de vue booléen est intéressant lorsqu'on commence à combiner plusieurs variables/paramètres. Par exemple, si on a une fonction réelle positive  $f$  sur un ensemble  $X$  et  $t \in \mathbb{R}^+$ , on peut

considérer  $1_{\{f>t\}}$  comme une fonction sur  $X$ , mais on peut aussi voir  $t$  comme une variable et considérer la fonction sur  $X \times \mathbb{R}^+$

$$(x, t) \rightarrow 1_{\{f>t\}}(x).$$

Cette fonction est une fonction booléenne sur  $X \times \mathbb{R}^+$  : elle vaut 1 si et seulement si la condition  $f(x) > t$  est vérifiée. On a plusieurs écriture possible, suivant le point de vue privilégié :

$$1_{\{f>t\}}(x) = 1_{\{(y,s) \in X \times \mathbb{R}^+ ; f(y) > s\}}(x, t) = 1_{\{s \in \mathbb{R}^+ ; s < f(x)\}}(t).$$

Revenons à l'expression de l'intégrale ci-dessus. Pour tout fonction  $f$  positive, sur un ensemble  $X$  on peut écrire :

$$\forall x \in X, \quad f(x) = \int_0^{+\infty} 1_{\{f>t\}}(x) dt. \quad (1.2)$$

Le résultat précédent (1.2) s'ensuit alors par le théorème de Fubini.

Voici une version générale.

**Proposition 4.** *Soit  $\nu$  une mesure sur les boréliens de  $\mathbb{R}^+$  tel que la fonction*

$$\phi(t) := \nu([0, t]) \quad (1.3)$$

*soit finie pour tout  $t > 0$ . On remarquera que  $\phi(0) = 0$  et que  $\phi$  est croissante.*

*Alors, si  $f$  est une fonction positive mesurable sur un espace mesuré  $(X, \mathcal{A}, \mu)$  on a*

$$\int_X \phi(f(x)) d\mu(x) = \int_0^{+\infty} \mu(\{f > t\}) d\nu(t).$$

*En particulier, en prenant  $d\nu(t) = pt^{p-1} dt$  avec  $p > 0$ , on a*

$$\int_X f^p d\mu = p \int_0^{p-1} \mu(\{f > t\}) t^{p-1} dt.$$

*Démonstration.* [Lieb and Loss] p. 26-27 □

On remarque que si on prend pour  $\mu$  la mesure de Dirac en  $x \in X$ , alors on retrouve la formule (1.2).

Si on se donne une fonction  $\phi$  croissante continue sur  $[0, +\infty[$ , avec  $\phi(0) = 0$  et telle que  $\phi$  soit de classe  $C^1$  sur  $]0, +\infty[$ , alors la mesure  $d\nu(t) = \phi'(t) dt$  vérifie bien les hypothèse et on a

$$\int_X \phi(f(x)) d\mu(x) = \int_0^{+\infty} \mu(\{f > t\}) \phi'(t) dt.$$

L'exemple principal reste  $\phi(t) = t^p$  avec  $p > 0$ .

Voici un autre exemple d'utilisation des ensembles de niveaux, pour un problème variationnel.

**Proposition 5.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $f$  une fonction positive mesurable sur  $X$ . On se donne  $s_0 \in ]0, \mu(X)[$  et on suppose qu'il existe  $t_0 > 0$  tel que  $\mu(\{f > t_0\}) = s_0$ . Alors, parmi tous les ensembles  $A \subset X$  de mesure fixée  $\mu(A) = s_0$ , la quantité

$$\int_A f(x) d\mu(x)$$

est maximale pour

$$A_0 = \{f > t_0\}$$

*Démonstration.* On a

$$\int (1_A - 1_{A_0})f d\mu = \int (1_A - 1_{A_0})(f - t_0) d\mu$$

et

$$\forall x \in X, \quad (1_A(x) - 1_{A_0}(x))(f(x) - t_0) \leq 0.$$

□

### 1.3 L'inégalité de Jensen

Soit  $\varphi$  une fonction convexe sur  $\mathbb{R}$ . La convexité de  $\varphi$  est équivalente à la propriété suivante pour les taux d'accroissements : pour tout  $s, t, u \in \mathbb{R}$

$$s < t < u \implies \frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(u) - \varphi(t)}{u - t}. \quad (1.4)$$

En utilisant deux fois cette inégalité, on voit que pour  $s < t < u < v$ ,

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(u) - \varphi(t)}{u - t} \leq \frac{\varphi(v) - \varphi(u)}{v - u}. \quad (1.5)$$

De cela, on déduit qu'en tout point,  $\varphi$  admet une dérivée à gauche et une dérivée à droite (en particulier  $\varphi$  est continue), que ces dérivées sont croissantes, et que  $\varphi'_g \leq \varphi'_d$  en tout point. On a pour, pour tout  $t_0 \in \mathbb{R}$  et  $\ell \in [\varphi'_g(t_0), \varphi'_d(t_0)]$ ,

$$\forall s \in \mathbb{R}, \quad \varphi(s) \geq \varphi(t_0) + \ell(s - t_0).$$

Pour le voir, on utilise (1.5). En résumé : en tout point, le graphe d'une fonction convexe sur  $\mathbb{R}$  admet une tangente, et le graphe reste au dessus de ses tangentes.

La fonction  $\varphi$  est dite *strictement convexe* sur  $\mathbb{R}$  si pour  $s, t \in \mathbb{R}$ ,  $s \neq t$  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  $\varphi$  ne contient pas de segment. En particulier, toute tangente ne touche le graphe qu'en un seul point.

**Théorème 6.** Soit  $\varphi$  une fonction convexe sur  $\mathbb{R}$ . Si  $\mu$  est une probabilité sur un espace mesurable  $(X, \mathcal{A})$  et  $f$  une fonction  $\mu$ -intégrable à valeurs  $\mathbb{R}$ , alors  $[\varphi \circ f]_-$  est intégrable, et ainsi  $\varphi \circ f$  est semi-intégrable, et on a

$$\varphi\left(\int_X f d\mu\right) \leq \int_X \varphi(f(x)) d\mu(x), \quad (1.6)$$

avec égalité si  $f$  est une fonction constante.

Lorsque  $\varphi$  est strictement convexe, il y a égalité dans (1.6) si et seulement si  $f$  est une fonction "constante" (plus précisément, égale  $\mu$ -pp à une constante, i.e. il existe  $a \in \mathbb{R}$  tel que  $f = a$   $\mu$ -pp).

*Démonstration.* On pose  $m = \int_X f d\mu \in \mathbb{R}$ . Comme  $\varphi$  est convexe, il existe au moins une droite située en-dessous du graphe de  $\varphi$  et passant par  $(m, \varphi(m))$ , d'équation  $y = \beta(x - m) + \varphi(m)$ . Ceci se traduit par

$$\varphi(u) \geq \beta(u - m) + \varphi(m) \quad u \in \mathbb{R}$$

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  $\mu$ -intégrables), et parce que  $\mu$  est une probabilité.

Si  $f$  n'est pas constante, il existe un ensemble  $A \in \mathcal{A}$  de mesure non-nulle tel que  $f(x) \neq m$  pour tout  $x \in A$ . Si  $\varphi$  est strictement convexe, on a alors

$$\forall x \in A, \quad F(x) := \varphi \circ f(x) - \beta(f(x) - m) + \varphi(m) > 0$$

et  $\int_X F d\mu \geq \int_A F d\mu > 0$ . L'inégalité sera donc stricte. □

## 1.4 La mesure de Lebesgue

On admettra l'existence et les propriétés fondamentales (régularité) de la mesure de Lebesgue, qui sont résumées dans le théorème suivant. Voir le livre de Rudin pour la construction et les preuves.

**Théorème 7.** Il existe une mesure  $\lambda$  définie sur une tribu  $\mathcal{L}$  de  $\mathbb{R}^n$  qui vérifie les propriétés suivantes :

1.  $\mathcal{L}$  contient les boréliens de  $\mathbb{R}^d$ , et est complète (càd si  $A \in \mathcal{L}$  est tel que  $\lambda(A) = 0$ , alors  $N \subset A \Rightarrow N \in \mathcal{L}$ ).

$$2. \lambda\left(\prod_{i=1}^n [a_i, b_i]\right) = \prod_{i=1}^k |b_i - a_i|.$$

3.  $\lambda$  est invariant par translation : si  $E \in \mathcal{L}$  et  $x_0 \in \mathbb{R}$ ,  $\lambda(E + x_0) = \lambda(E)$ .

De plus, elle vérifie aussi les propriétés suivantes :

a) Pour tout  $A \in \mathcal{L}$ ,

$$\lambda(A) = \inf\{\lambda(U) ; U \supset A, U \text{ ouvert}\}.$$

b) Pour tout  $A \in \mathcal{L}$ , avec  $\lambda(A) < +\infty$ , ou  $A$  ouvert, on a

$$\lambda(A) = \sup\{\lambda(K) ; K \subset A, K \text{ compact}\}.$$

En conséquence, pour  $A \in \mathcal{L}$ , on peut trouver  $F$  une réunion dénombrable de compacts et  $V$  intersection dénombrable d'ouverts tel que  $F \subset A \subset V$  avec  $\lambda(V \setminus F) = 0$ .

Pour la mesure de Lebesgue sur  $\mathbb{R}^n$ , on parle aussi de volume.

Je suppose aussi que vous connaissez la formule de changement de variable dans  $\mathbb{R}^n$ .

La régularité de la mesure de Lebesgue, à savoir les points a) et b) du théorème précédent, est essentielle pour étudier les approximations par des fonctions régulières. À titre d'exemple, voici un joli résultat dû à Lusin.

**Théorème 8 (Lusin).** Soit  $f$  une fonction mesurable sur  $\mathbb{R}^n$ , à valeurs complexes, nulle en dehors d'un ensemble de mesure (volume) fini. Pour tout  $\varepsilon > 0$ , il existe une fonction  $g$  continue sur  $\mathbb{R}^n$  à support compact tel que

$$\lambda(\{f \neq g\}) \leq \varepsilon.$$

De plus, on peut s'arranger pour que  $\sup |g| \leq \sup |f|$ .

Nous n'utiliserons pas explicitement ce résultat ; on pourra consulter [Rudin, p. 55] pour la démonstration. Nous utiliserons cependant la brique suivante, qui est un résultat purement topologique.

**Lemme 9 (Lemme d'Urysohn).** Soit  $K$  un compact et  $U$  un ouvert de  $\mathbb{R}^n$  tels que  $K \subset U$ . Alors, il existe une fonction  $f$  continue sur  $\mathbb{R}^n$ , qui vérifie

1.  $f$  est à valeurs dans  $[0, 1]$  et à support compact,
2.  $f(x) = 1$  si  $x \in K$ ,
3.  $f(x) = 0$  si  $x \notin U$ .

*Démonstration.* On rappelle que la distance à un ensemble  $A \subset \mathbb{R}^n$  est définie par  $d(x, A) = \inf\{|y - x| ; y \in A\}$ . Pour tout ensemble  $A$  (non vide), cette fonction est 1-lipschitzienne, donc continue. On a  $d(x, A) = 0$  si et seulement si  $x$  appartient à l'adhérence de  $A$ .

Pour  $\varepsilon > 0$  on considère

$$K_\varepsilon := \{x \in \mathbb{R}^n, d(x, K) < \varepsilon\}$$

C'est un ouvert borné, contenant  $K$ , et qui pour un certain  $\varepsilon_0$  assez petit sera dans  $U$ . En effet, si ce n'est pas le cas, on pourrait construire une suite  $x_n \notin U$  avec  $d(x_n, K) < \frac{1}{n}$ . Comme  $K$  est borné, cette suite est bornée et on peut donc en extraire une sous-suite convergente. Sa limite  $x_\infty$  serait dans  $K$ , puisque  $d(x_\infty, K) = 0$  et  $K$  est fermé, et ne serait pas dans  $U$  car le complémentaire de  $U$  est fermé. Cela contredirait  $K \subset U$ .

On a donc  $K \subset K_{\varepsilon_0} \subset U$ . On définit sur  $\mathbb{R}^n$ ,

$$f(x) = \frac{1}{\varepsilon_0} \left( \varepsilon_0 - d(x, K) \right)_+ = \left( 1 - \frac{1}{\varepsilon_0} d(x, K) \right)_+.$$

Cette fonction vérifie les conditions demandées. □