**Université Paris-Dauphine** **2020-2021**
Cours “intégrale de Lebesgue et probabilités” Juillet 2020

# Présentation du cours

## Table des matières



Sont écrites en rouge les parties qui peuvent être omises en première lecture. Bien que ne comportant pas de difficulté technique, les notes historiques en violet peuvent également être omises.

# 1 Calcul intégral et intégrale de Riemann

Le calcul intégral est apparu historiquement pour le calcul des longueurs, aires et volumes. Pour une surface $S$ délimitée par le graphe d'une fonction $f : [a, b] \to \mathbb{R}$, $a, b \in \mathbb{R}$, $a \le b$, on définit

$$\text{aire}(S) = \text{aire}(S_+) - \text{aire}(S_-) = \int_a^b f(x)\,dx, \tag{1.1}$$

où $S_\pm := \{(x, y) \in [a, b] \times \mathbb{R},\ 0 \le y \le f_\pm(x)\}$, $f_\pm(x) := \max(0, \pm f(x))$. Lorsque $S$ est le rectangle $S := [a, b] \times [0, 1]$, ou de manière équivalente, lorsque $f$ est la fonction indicatrice de l'intervalle $[a, b]$, cette aire/intégrale est par définition

$$\text{aire}(S) = \int_a^b \mathbf{1}\,dx = b - a. \tag{1.2}$$

Lorsque $S$ est une réunion finie de rectangles, ou de manière équivalente, lorsque $f$ est une fonction en escalier, on note $f \in \mathscr{E}sc$, c'est-à-dire, si

$$f(x) := \sum_{i=1}^n \alpha_i \mathbf{1}_{[x_{i-1}, x_i[}, \quad a = x_0 < x_1 < \ldots < x_n = b,\ \alpha_i \in \mathbb{R},$$

cette aire/intégrale est également facile à définir (et calculer), et l'on a

$$\int_a^b f(x)\,dx = \sum_{i=1}^n \alpha_i (x_i - x_{i-1}).$$

Pour une surface plus générale, l'intégrale est définie par approximation (méthodes d'exhaustion)

$$\int_a^b f(x)\,dx = \text{«}\ \sum f(x_i)\delta x_i\ \text{»}. \tag{1.3}$$

Plus précisément, lorsque $f$ est une fonction continue, on peut définir l'intégrale comme limite des sommes (de Riemann)

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} \sum_{i=1}^n f(x_i)(x_i - x_{i-1}), \tag{1.4}$$

où la famille $(x_i)_{i=0,\dots,n}$ est une subdivision de $[a,b]$ dont la longueur maximale $\delta_n := \max(x_i - x_{i-1})$ tend vers 0 lorsque $n$ tend vers $\infty$. De manière équivalente, l'intégrale est définie comme limite (borne inférieure et borne supérieure) des sommes (de Darboux)

$$\int_a^b f(x)\,dx = \sup_{g\in\mathscr{E}sc,\, g\le f} \int_a^b g(x)\,dx = \sup_{h\in\mathscr{E}sc, f\le h} \int_a^b h(x)\,dx, \tag{1.5}$$

c'est-à-dire comme limite des approximations par défaut et par excès par des fonctions en escalier. Plus généralement, on dira qu'une fonction $f : [a,b] \to \mathbb{R}$ est Riemann intégrable si les deux derniers termes ci-dessus sont finis et égaux, et la formule précédente servira alors de définition à l'intégrale de Riemann de $f$. En particulier, $f$ est Riemann intégrable si elle est rectifiable, c'est-à-dire, si $f$ est la différence de deux fonctions croissantes, ou de manière équivalente, si elle est limite uniforme d'une suite de fonctions en escalier. Cette intégrale connait différentes généralisations : aux intervalles $]a,b[$, $a,b \in [-\infty,+\infty]$, aux fonctions "à variations bornées".

On dispose alors de différents résultats célèbres :

- La relation fondamentale du calcul intégral liant calcul de primitive et calcul de dérivée

$$\int_a^b f(x)\,dx = F(b) - F(a), \tag{1.6}$$

si $F : [a,b] \to \mathbb{R}$ est une fonction de classe $C^1$ telle que $F' = f$, et sa variante, la formule d'intégration par parties.

- La formule de changement de variables

$$\int_a^b f(y)\,dy = \int_{\phi(a)}^{\phi(b)} f(\phi(x))\,\phi'(x)\,dx, \tag{1.7}$$

si $\phi : [a,b] \to \mathbb{R}$ est une fonction de classe $C^1$.

- Pour les intégrales multiples : la méthode des indivisibles (ou principe de Cavalieri), les formules de changement de variables et d'inversion de l'ordre d'intégration.

- Le formule de Dirichlet d'inversion des séries de Fourier

$$\frac{1}{2}(f(x+0) + f(x-0)) = \sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad c_n := \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt, \tag{1.8}$$

si $f$ est une fonction $2\pi$-périodique de classe $C^1$ par morceaux.

*Notes historiques.*

Le calcul intégral est développé depuis Archimède et Eudoxe de Cnide ($\sim -250$) et est formalisé dans sa forme quasiment actuelle par Newton et Leibniz ($\sim 1670 - 1690$) qui introduisent la notation moderne de l'intégrale et établissent les formules (1.3), (1.6) et (1.7).

Le besoin de fonder rigoureusement le calcul intégral se fait sentir au XIXᵉ siècle. En 1821, dans la leçon 21 de son cours d'analyse à l'école polytechnique, Cauchy démontre (1.4) en utilisant la propriété d'uniforme continuité de $f$ et son (désormais fameux) critère de Cauchy sur la suite des séries de Riemann.

Au cours du XIXᵉ siècle, les notions de fonction et de nombres réels sont précisées par Bolzano, Dirichlet, Wieirestrass, Heine, Dedekind, Cantor conjointement au développement de la théorie de l'intégration par Riemann, Darboux, Jordan, Stieltjes qui établissent (1.4) et (1.5) avec une grande généralité. A noter ici le rôle important joué par les séries de Fourier qui donne un accès naturel à des fonctions "peu régulières" et à la formule (1.8) démontrée par Dirichlet y compris pour la fonction (de Dirichlet) $f = \mathbf{1}_{[0,1]}$.

## 2 L'intégrale de Lebesgue sur $[0,1]$

Au tournant du siècle la question est alors de donner un sens à la longueur de sous-ensembles de $\mathbb{R}$ les plus généraux, à la surface de sous-ensembles de $\mathbb{R}^2$ les plus généraux et à l'intégrale de fonctions $f:[a,b]\to\mathbb{R}$ les plus générales.

L'intégrale de Lebesgue opère un changement radical de point de vue et repose sur deux idées que nous exposons brièvement dans le cas de l'intégrale sur le segment $[0,1]$ :

- La première question est celle de savoir définir la mesure (ou longueur) des parties de $[0,1]$. Il est effectivement possible de définir une classe de parties de $[0,1]$, qui contient les intervalles mais qui est bien plus générale, appelée tribu de Lebesgue et notée $\mathscr{L}([0,1])$, telle que l'on soit capable de définir la longueur (ou mesure de Lebesgue) $\lambda(A)\in[0,1]$ pour tout $A\in\mathscr{L}([0,1])$ et telle que
(i) la longueur $\lambda(A)$ coïncide avec la longueur usuelle lorsque $A$ est un intervalle ;
(ii) la longueur d'un ensemble est invariante par translations ($\lambda(a+A)=\lambda(A)$ si $a\in\mathbb{R}$) ;
(iii) la fonction d'ensembles $\lambda:\mathscr{L}([0,1])\to[0,1]$ se comporte convenablement vis à vis des réunions dénombrables disjointes d'ensembles.
Il est toutefois impossible de définir la longueur de toutes les parties de $[0,1]$ de sorte que ces deux dernières propriétés soient satisfaites, en particulier $\mathscr{L}([0,1])\neq\mathscr{P}([0,1])$.

De manière plus précise, on définit successivement, la mesure d'un intervalle ouvert $I=]a,b[\subset[0,1]$ par

$$\lambda(I)=b-a, \tag{2.9}$$

et la mesure d'un ouvert $O\subset[0,1]$ par

$$\lambda(O)=\sum_{n=1}^{\infty}\lambda(I_n),\quad \text{si } O=\bigcup_n I_n,\ (I_n)\ \underline{\text{suite}}\ \text{intervalles disjoints.} \tag{2.10}$$

Pour un ensemble quelconque $A\subset[0,1]$, on définit alors sa mesure extérieur $\lambda^*(A)$ par

$$\lambda^*(A):=\inf\{\lambda(O);\ A\subset O,\ O\ \text{ouvert}\}\in[0,1].$$

On dit qu'un ensemble $A\subset[0,1]$ est mesurable au sens de Lebesgue, on note $A\in\mathscr{L}([0,1])$, si $\lambda^*(A)+\lambda^*([0,1]\backslash A)=1$, et on définit sa mesure par

$$\lambda(A):=\lambda^*(A).$$

On peut alors montrer que l'ensemble de parties $\mathscr{L}([0,1])$ est une tribu, au sens où
(i) $\emptyset\in\mathscr{L}([0,1])$ ;
(ii) $A\in\mathscr{L}([0,1])$ implique $[0,1]\backslash A\in\mathscr{L}([0,1])$ ;
(iii) $A_n\in\mathscr{L}([0,1])$ pour tout $n\in\mathbb{N}$ implique $\bigcup A_n\in\mathscr{L}([0,1])$,
et que $\lambda$ est une fonction $\sigma$-additive d'ensembles, au sens où

$$\lambda(\emptyset)=0,\quad \lambda\Big(\bigcup_n A_n\Big)=\sum_n\lambda(A_n),$$

si $(A_n)$ est une suite d'ensembles de $\mathscr{L}([0,1])$ deux à deux disjoints. [1]

1. Une définition alternative et équivalente de la mesure de Lebesgue consiste à définir la longueur d'un compact $K\subset[0,1]$ par

$$\lambda(K):=\inf\{\lambda(O);\ K\subset O,\ O\ \text{ouvert}\}, \tag{2.11}$$

puis pour un ensemble quelconque $A\subset[0,1]$, de définir sa mesure extérieur $\lambda^*(A)$ et sa mesure intérieur $\lambda_*(A)$ par

$$\begin{aligned}\lambda^*(A) &:= \inf\{\lambda(O);\ A\subset O,\ O\ \text{ouvert}\},\\ \lambda_*(A) &:= \sup\{\lambda(K),\ K\subset A,\ K\ \text{compact}\}.\end{aligned}$$

On dit que $A\subset[0,1]$ est mesurable au sens de Lebesgue, on note $A\in\mathscr{L}([0,1])$, si sa mesure intérieure et sa mesure extérieure coïncident, et on définit sa mesure par

$$\lambda(A):=\lambda_*(A)=\lambda^*(A).$$

- La deuxième question est celle de définir l'intégrale. Pour une fonction indicatrice d'ensemble $\mathbf{1}_A$ avec $A \in \mathscr{L}([0,1])$, on définit

$$\int_{[0,1]} \mathbf{1}_A d\lambda := \lambda(A), \tag{2.12}$$

puis on procède de manière similaire à la définition de l'intégrale classique. Pour une fonction $f$ qui est combinaison linéaire de fonctions indicatrices d'ensemble mesurable, on dit que $f$ est étagée et on note $f \in \mathcal{E}$, l'intégrale de $f$ est la combinaison linéaire des intégrales des fonctions indicatrices :

$$\int_{[0,1]} f d\lambda := \sum_{i=1}^{n} \alpha_i \lambda(A_i), \quad \text{si } f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i}.$$

Pour une fonction positive $f$ qui est limite croissante d'une suite $(f_n)$ de fonctions étagées positives, l'intégrale de $f$ est la limite des intégrales des fonctions étagées :

$$\int_{[0,1]} f d\lambda := \lim \int_{[0,1]} f_n d\lambda, \quad \text{si } f = \lim f_n,\ f_n \uparrow,\ f_n \text{ étagée positive.}$$

Lorsque $f$ est la différence de deux fonctions positives d'intégrales finies, l'intégrale de $f$ est la différence des intégrales de ces deux fonctions.

A ce stade, il n'est pas inutile de faire plusieurs remarques.

- On voit que les fonctions élémentaires que l'on peut intégrer grâce à la théorie de Lebesgue, à savoir les fonctions indicatrices d'ensemble mesurable, sont beaucoup plus générales que les fonctions élémentaires que l'on peut intégrer grâce à la théorie de Riemann, qui sont les fonctions indicatrices d'intervalle. On voit également que l'intégrale de Riemann correspond à la limite d'une approximation par un découpage par bandes verticales alors que l'intégrale de Lebesgue correspond à la limite d'une approximation par un découpage par tranches horizontales. On peut enfin montrer que les fonctions Riemann intégrables sont Lebesgue intégrables et que les fonctions Riemann intégrables sont les fonctions continues sauf éventuellement sur un ensemble de points de mesure de Lebesgue nul. Cette construction a deux conséquences fondamentales et recherchées.

(i) Pour une suite uniformément bornée de fonctions $(f_n)$, si $f_n$ est intégrable et si $f_n(x) \to f(x)$ pour tout $x \in [0,1]$, alors $f$ est intégrable. Cela contraste avec l'intégrale de Riemann, puisque si $D_n$ est la suite des $n$ « premiers » rationnels de $D := \mathbb{Q} \cap [0,1]$, les fonctions $\mathbf{1}_{D_n}$ sont Riemann intégrables (d'intégrale nulle) mais $\mathbf{1}_D = \lim \mathbf{1}_{D_n}$ ne l'est pas (puisque le meilleur encadrement par des fonctions en escalier est $g := 0 \le \mathbf{1}_D \le h := 1$ et que $g$ et $h$ n'ont pas la même intégrale). On peut également écrire

$$\mathbf{1}_D(x) = \lim_{n\to\infty} \mathbf{1}_{D_n}(x), \quad \mathbf{1}_{D_n}(x) = \lim_{m\to\infty} [\cos(2\pi n! x)]^{2m} \mathbf{1}_{[0,1]}(x),$$

où les fonctions $x \mapsto [\cos(2\pi n! x)]^{2m} \mathbf{1}_{[0,1]}(x)$ sont régulières (donc Riemann intégrables).

(ii) La relation fondamentale du calcul intégral (1.6) est vraie au sens de l'intégrale de Lebesgue pour toute fonction $F$ continue sur $[a,b]$, dérivable sur $]a,b[$ et de dérivée bornée. En ce sens l'opération d'intégration au sens de Lebesgue est exactement l'opération inverse à celle de la dérivation. Cela contraste également avec l'intégrale de Riemann puisque par exemple Volterra a construit "en recollant" des copies de la fonction assez oscillante $x \mapsto x^2 \sin(1/x)$ une fonction dérivable sur $\mathbb{R}$ et de dérivée bornée, qui n'est Riemann intégrable sur aucun intervalle de $\mathbb{R}$.

- Il convient enfin de souligner encore une fois que toutes les parties de $[0,1]$ ne sont pas mesurables au sens de Lebesgue, soit donc $\mathscr{L}([0,1]) \neq \mathscr{P}([0,1])$, mais qu'il est nécessaire d'avoir recours à l'axiome du choix pour établir ce résultat. Les questions autours des limites de la théorie de Borel et de Lebesgue nous emmènent immédiatement à s'intéresser à des problèmes subtils à la frontière entre mathématiques et logique axiomatique. Le résultat le plus emblématique dans cette direction est probablement le paradoxe de Banach-Tarski qui affirme qu'il est possible (à l'aide de l'axiome du choix) de trouver une partition en $2n$ ensembles $A_1, \ldots, A_{2n}$ de la boule unité $B$ de $\mathbb{R}^3$ et des isométries $g_1, \ldots, g_{2n}$ de $\mathbb{R}^3$ à partir desquels on peut reconstituer deux fois la boule unité, soit donc $B = \cup_{i=1}^n g_i A_i = \cup_{i=1}^n g_{i+n} A_{i+n}$. Cela implique que les ensembles $A_i$ ne sont pas (tous) mesurables, puisque sinon nous arriverions à la contradiction $\mathrm{vol}(B) = 2\mathrm{vol}(B)$ ! Ces questions ne seront pas abordées ici car d'une part elles ne sont pas au cœur de la théorie de Lebesgue (on peut développer et utiliser cette théorie sans répondre à ces questions) et que d'autre part elles mettent en jeu de manière fondamentale l'axiome du choix, axiome auquel nous souhaitons éviter d'avoir recours.

*Notes historiques.*

- Dans les années 1880, Jordan formalise la notion d'ensemble du plan quarrable : c'est un ensemble $A$ que l'on peut approcher par l'intérieur et l'extérieur par des <u>réunions finies</u> de rectangles de sorte que les surfaces ainsi couvertes sont arbitrairement proches. L'aire $\mu(A)$ de $A$ est alors la limite commune des ces approximations par défaut et par excès. Jordan montre que cette fonction d'ensemble est une mesure additive au sens où $\mu(\emptyset) = 0$ et

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i),$$

si les $(A_i)$ sont des sous-ensembles quarrables et disjoints du plan.

- En 1898, Borel définit une classe d'ensembles, appelés aujourd'hui boréliens ou tribu de Borel, qui est constituée des ensembles obtenus à partir des intervalles ouverts de $[0, 1]$ par union dénombrable et passage au complémentaire, puis en recommençant éventuellement indéfiniment ces deux opérations. De manière naïve, il définit également la mesure des premiers boréliens à l'aide des formules (2.9), (2.10) et (2.11), sans toutefois démontrer la validité de la mesure pour tous les boréliens.

- Il revient à Lebesgue en 1901 de fonder la théorie de la mesure et de l'intégrale comme indiqué ci-dessus en complétant, généralisant et rendant rigoureuses certaines intuitions de Jordan et Borel. Il est important de souligner que la tribu de Lebesgue (des ensembles mesurables) est bien plus vaste que la tribu de Borel, et que l'on ne sait mesurer les boréliens que parce que l'on sait qu'ils appartiennent à cette tribu de Lebesgue. Lebesgue développera ensuite sa théorie dans sa thèse (1902) et deux cours Peccot au Collège de France (1902-1903 et 1904-1905) dans lesquels il généralisera sa construction et démontrera les principaux théorèmes de convergence qui portent aujourd'hui son nom.

- Mentionnons enfin que Lebesgue et les analystes français de l'époque (Baire, Fréchet, Fatou) ne s'intéresseront pas ou rapidement plus à cette théorie de l'intégrale qui ne fut d'ailleurs pas enseignée à l'université en France jusque dans les années 1950. Denjoy développera même une théorie alternative (dont l'utilité n'a jamais vraiment convaincu !) qui généralise la théorie de l'intégrale de Riemann. Toutefois Lévy et Leray s'appuieront sur la théorie de l'intégrale de Lebesgue dans leurs travaux respectifs en théorie des probabilités et des équations aux dérivées partielles. C'est plus à l'Est que la théorie de Lebesgue est reprise et développée, notamment :

- en Italie, avec Beppo Levi, Vitali, Fubini et Tonelli qui s'intéressent à la théorie de la mesure (premier exemple d'ensemble non mesurable) et à la théorie de l'intégrale (en particulier, à l'intégrale multiple) ;

- en Autriche-Hongrie, avec Fischer, Hahn, Riesz, Fèjer et Radon qui s'intéressent à l'analyse fonctionnelle et aux séries de Fourier ;

- en Allemagne avec Carathéodory, Perron et Hausdorff qui s'intéressent à la théorie de la mesure et à la théorie de l'intégration ;

- en Pologne avec Steinhaus, Banach et Nikodym qui fondent l'analyse fonctionelle en introduisant et étudiant les espaces vectoriels normés complets (espaces de Banach) et s'intéressent à la théorie de l'intégrale de Lebesgue ;

- en Russie avec Egorov puis plus tard avec Sobolev et Kolmogorov qui étudient la théorie de l'intégration et l'analyse fonctionelle. Kolmogorov fonde en 1929 la théorie axiomatique des probabilités et Sobolev introduira de nouveaux espaces fonctionnels issus de la théorie de l'intégrale de Lebesgue et incontournables en théorie des équations aux dérivées partielles.

## 3 L'intégrale de Lebesgue sur un ensemble mesuré

L'objectif de ce cours est de construire l'intégrale de Lebesgue

$$\int_E f \, d\mu$$

pour des ensembles $E$ quelconques, des fonctions « mesurables » $f$ et des « mesures positives » $\mu$, objets que nous définirons précisément dans les deux premiers chapitres du cours.

Cette intégrale jouit de deux propriétés essentielles :

- Les normes $\| \cdot \|_p$ usuelles permettent de définit des espaces de Banach, ce qui est nécessaire pour savoir démontrer certains théorèmes d'existence. Son utilisation dans le domaine de l'analyse fonctionnelle a été cruciale dans le développement de cette branche des mathématiques au cours du XXᵉ siècle et reste fondamental dans ce domaine et dans le domaine de l'analyse en général.

- L'ensemble $E$ peut-être pris "très gros" comme par exemple $E = \mathbb{R}^{\mathbb{N}}$ ou même $E = C([0,\infty);\mathbb{R})$, ce qui constituent les ensembles naturels pour pouvoir définir des "suites de variables aléatoires" ou des "processus" en théorie des probabilités. La théorie de Lebesgue s'est définitivement imposée part le fait qu'elle permet de fonder la théorie des probabilités (ce que l'on doit à Kolmogorov).

Une fois l'intégrale de Lebesgue construite, nous disposerons de théorèmes assez faciles à utiliser pour calculer des intégrales multiples, passer à la limite sous le signe intégrale, ..., et une théorie qui unifie la théorie de l'intégration usuelle et la théorie de sommation (séries numériques ou séries de fonctions). Le prix à payer est une théorie sensiblement plus abstraite et subtile que celle de Cauchy et Riemann.

Comme nous l'avons déjà dit, une des motivations historiques était d'apporter une solution au problème des primitives, c'est-à-dire, d'établir (1.6) pour une large classe de fonctions pour lesquelles l'opération de dérivation et d'intégration sont des opérations inverses l'une de l'autre. Une autre motivation était de classer les ensembles de $\mathbb{R}$ ou $\mathbb{R}^d$ suivants leur taille. La théorie de Lebesgue permet également d'aborder avec succès cette question et a été à la base des développements récents de la théorie géométrique de la mesure et de l'analyse fine (analyse harmonique, inégalités fonctionnelles) contemporaine.

En résumé, l'intégrale de Lebesgue doit son succès à sa simplicité d'utilisation et parce qu'elle a été à l'origine du développement de domaines importants des mathématiques au début du XXᵉ siècle, à savoir notamment :

▷ L'analyse fonctionelle. En 1907, Riesz et Fréchet montrent et utilisent le caractère complet de l'espace $L^2$, puis $L^p$, ce qui sera formalisé par les fondateurs de l'analyse fonctionnelle moderne à Lwów (Pologne) par Banach, Steinhaus et Nikodým à partir de 1920.

▷ La théorie des probabilités. Au cours de la décennie 1920-1930, Wiener et Lévy construisent une mesure qui permet de définir le mouvement Brownien, Steinhaus et Kolmogorov autour des années 1930 fondent définitivement les probabilités et introduisent le fameux triplet $(\Omega, \mathscr{A}, P)$.

▷ La théorie géométrique de la mesure. Un résultat géométrique emblématique est le *théorème isopérimétrique* qui postule que pour tout compact $K \subset \mathbb{R}^n$, on a

$$\frac{\mathrm{vol}(K)}{(\mathrm{vol}(\partial K))^{1-1/n}} \geq \frac{\mathrm{vol}(B)}{(\mathrm{vol}(\partial B))^{1-1/n}},$$

résultat qui ne doit toujours pas être démontré en toute généralité. Les premières versions et preuves remontent à Zénodore ($\sim -200$) et des contributions importantes sont dues à Minkowski

(1864-1909). Le lien avec la théorie de la mesure est fait dans l'article fondateur de Hausdorff en 1919, travail qui sera repris et poursuivit par Alexandrov, Fenchel, Steiner, ....

Venons-en à une présentation succincte de ce cours d'intégrale de Lebesgue et probabilités. Ce domaine des mathématiques étant extrêmement vaste, ce cours ne doit être considéré que comme une introduction aux différents thèmes que sont la théorie de la mesure (ensembliste), la construction de l'intégrale de Lebesgue et ses propriétés fondamentales, les inégalités fonctionnelles, l'analyse fonctionnelle et l'analyse de Fourier, les fondements axiomatiques des probabilités et les premières propriétés des variables aléatoires.

Dans le choix des thèmes abordés et l'importance relative de chacun d'entre eux, nous avons essayer d'emprunter une voie médiane entre une présentation « pour l'analyse » et une présentation « pour les probabilités ». Plus précisément :

- Nous avons choisi de laisser de côté les aspects les plus techniques de la théorie de la mesure, notamment la construction de la mesure de Lebesgue, et des mesures en général, qui ne sera abordé (si nous avons le temps) qu'au dernier chapitre. Du même coup, les aspects topologiques sont réduits à la portion congrue, et c'est le cadre général ensembliste qui est présenté.

- Nous avons toutefois choisi de ne rien laissé de côté et tous les résultats seront donc démontrés dans les notes (éventuellement certaines preuves ne seront pas exposées en cours et nous renverrons les étudiants au cours d'analyse fonctionnelle du second semestre pour leur présentation).

- Notre objectif étant d'exposer les outils efficaces du calcul intégral et du calcul des probabilités, nous avons laissé de côté les discussions subtiles concernant les contre-exemples et les comparaisons entre les différentes intégrales (au sens de Riemann, au sens de Borel, au sens de Lebesgue, sur un espace métrique ou non, sur une tribu complète ou non).

- Nous avons choisi de travailler uniquement dans le cadre des fonctions boréliennes et de la restriction de la mesure de Lebesgue à la tribu de Borel et non pas dans le cadre (apparemment plus général) des fonctions mesurables au sens de Lebesgue et de la mesure de Lebesgue sur la tribu de Lebesgue. Il s'avère que ce cadre est suffisant pour traiter la plupart des questions (pour ne pas dire toutes !) que l'on peut rencontrer en analyse et en probabilités. A noter en particulier, que les espaces $L^p$ de Lebesgue (de classes de fonctions) sont identiques lorsqu'ils sont construits à partir des fonctions Borel mesurables et des fonctions Lebesgue mesurables. Encore une fois, cela permet de simplifier l'exposé et d'éviter de nombreuses discussions qui risqueraient plus d'embrouiller que d'éclairer le lecteur.

- Les espaces de probabilités et l'étude des variables aléatoires ne seront traités que dans les derniers chapitres lorsque l'ensemble des outils ensemblistes et analytiques auront été présentés, cela permettra d'obtenir en peu de pages des résultats déjà très aboutis.

- Enfin, nous ne reviendrons pas sur la construction et les propriétés de l'intégrale de Riemann déjà abordées dans des cours plus élémentaires d'analyse en première et deuxième année de Licence. L'intégrale au sens de Lebesgue étant un prolongement et un raffinement de l'intégration au sens de Riemann, le lecteur est toutefois invité à « relire ses classiques ». Faute de temps, nous ne nous étendrons pas non plus sur les bases de la théorie des ensembles qui ont probablement déjà été entrevues dans des cours de probabilités en première et deuxième année de Licence, et dont nous nous servirons lorsque cela nous sera utile. Nous invitons cependant le lecteur curieux à se cultiver sur ce sujet, ce dernier pouvant parfois éclairer quelques remarques que nous formulerons sur la théorie de l'intégrale de Lebesgue.

Terminons enfin par une description rapide des différents chapitres.

- Dans un premier chapitre, on introduira les classes d'ensembles et de fonctions qui peuvent apparaitre dans l'intégrale.

- Le second chapitre sera consacré à la construction de l'intégrale proprement dite. On commencera par définir

$$\int_E \mathbf{1}_A \, d\mu =: \mu(A)$$

pour la classe des ensembles $A$ exhibée au premier chapitre et pour une mesure $\mu$ convenablement définie. Nous généraliserons ensuite l'intégration à la classe de fonctions du premier chapitre grâce à un procédé d'approximation.

- Dans le troisième chapitre nous exposerons différents "théorèmes limites" qui font que la théorie de Lebesgue est particulièrement agréable et simple d'utilisation et nous présenterons quelques inégalités fonctionnelles élémentaires.

- Le quatrième chapitre sera consacré au théorème de Fubini sur les intégrales multiples, au transport de mesures et au théorème de changement de variables.

- Le cinquième chapitre sera consacré aux espaces $L^p$ de Lebesgue sur un espace mesuré général, au moins dans le cas $p = 1, 2, \infty$.

- Le sixième chapitre sera consacré au cas particulier de $E = \mathbb{R}^d$ et à plusieurs objets/notions qui peuvent être définis dans ce cadre (convolution, densité, transformation de Fourier).

- Le septième chapitre sera consacré aux fondements des Probabilités (variable aléatoire, indépendance, convergence en loi, autres convergences, marche aléatoire).

- Les huitième chapitre sera consacré aux vecteurs gaussiens et au conditionnement.

- Le neuvième (et dernier) chapitre sera consacré à des compléments hors programme : construction de la mesure de Lebesgue, tribu de Lebesgue, mesures signées, dualité, ...

Le cours est inspiré de

- G. Ben Arous, "Intégration" (notes personnelles)
- J. Deny (notes personnelles)
- C. Villani, "cours d'intégration" (probablement sur le net)
- J.-F. Le Gall, "Intégration, Probabilités et Processus Aléatoires" (pobablement sur le net)

On trouve une bibliographie abondante sur le net :

- F. Bolley, R. Dujardin, A. Lambert, "Théorie de la Mesure et Intégration" (cours de Paris 6) ;
- J. Féjoz, "Chapitres d'intégration et de probabilités" (cours UPD 2012-2015) ;
- H. Doss, "Probabilités" (cours UPD 2016-2018) ;
- C. Suquet, "Intégration, Analyse de Fourier, Probabilités" ;
- F. Hérau, "Cours Intégration" ;
- A. Yger, "Théorie de l'intégration" ;
- R. Danchin, "Cours Intégration" ;
- O. Garet, "Intégration et Probabilités".

## 4 Compléments : Riemann, Borel et Lebesgue

- Des exemples de fonctions non Riemann intégrables mais intégrables au sens de Riemann généralisé sont $f(x) = x^{-1/2}$ et $f(x) = \sin(1/x)$ sur $[0,1]$ (puisque ces fonctions sont continues sur $[\varepsilon, 1]$, $\varepsilon > 0$, et les intégrales sur $[\varepsilon, 1]$ convergent lorsque $\varepsilon \to 0$).

- Notons $\mathscr{B}([0,1])$ l'ensemble des fonctions élémentaires et boréliennes, c'est-à-dire de la forme $f := \mathbf{1}_A$, avec $A \subset [0,1]$ borélien. On peut montrer que l'ensemble des boréliens $\mathcal{B}$ est équipotent à $\mathbb{R}$, ce qui montre du même coup que $\sharp\mathscr{B}([0,1]) = \sharp\mathbb{R}$. Pour s'en convaincre, on observera que l'ensemble $\mathcal{O}_\mathbb{R}$ des ouverts de $[0,1]$ peut être obtenu à partir de l'ensemble $\mathcal{O}_\mathbb{Q}$ des ouverts de $[0,1]$ dont les extrémités sont dans $\mathbb{Q}$ : chaque élément de $\mathcal{O}_\mathbb{R}$ est une union dénombrable d'éléments de $\mathcal{O}_\mathbb{Q}$. Cela implique $\sharp\mathcal{O}_\mathbb{R} = \sharp\mathcal{O}_\mathbb{Q}$. Par ailleur, dans $\mathbb{Q}$, la tribu $\mathcal{T}_\mathbb{Q}$ engendrée par $\mathcal{O}_\mathbb{Q}$ contient moins d'éléments que $\mathscr{P}(\mathbb{Q})$, on en déduit donc que $\sharp\mathcal{T}_\mathbb{Q} = \sharp\mathscr{P}(\mathbb{Q}) = \sharp\mathbb{R}$. Malheureusement, $\mathcal{T}_\mathbb{Q} \neq \mathcal{B}$, et cela ne permet donc pas de conclure. La preuve rigoureuse utilise la description de Borel de la tribu de Borel et un « argument de récurrence transfinie » (de logique / théorie descriptive des ensembles).

- On peut montrer qu'il existe des boréliens non dénombrables (équipotents à $\mathbb{R}$) et de mesure de Lebesgue nulle. Le plus connu est l'ensemble « triadique » de Cantor $K \subset [0,1]$. Il s'ensuit que pour tout $A \in \mathscr{P}(K)$, on a $\mathbf{1}_A$ est Riemann et Lebesgue intégrable, et donc en notant $\mathscr{R}([0,1])$ l'ensemble des fonctions indicatrices d'ensemble Riemann intégrable et $\mathscr{L}([0,1])$ l'ensemble des fonctions indicatrices d'ensemble Lebesgue intégrable, on a $\mathscr{P}(K) \subset \mathscr{R}([0,1]) \subset \mathscr{L}([0,1])$. On en déduit

$$\sharp\mathscr{P}([0,1]) = \sharp\mathscr{P}(K) \leq \sharp\mathscr{R}([0,1]) \leq \sharp\mathscr{L}([0,1]) \leq \sharp\mathscr{P}([0,1]),$$

soit donc $\sharp\mathscr{R}([0,1]) = \sharp\mathscr{L}([0,1]) > \sharp\mathscr{B}([0,1])$.

- On peut montrer que la $\mathscr{B}([0,1]) \otimes \mathscr{B}([0,1]) = \mathscr{B}([0,1] \times [0,1])$ et que $\mathscr{L}([0,1]) \otimes \mathscr{L}([0,1]) \subset \mathscr{L}([0,1] \times [0,1])$ (sans égalité).

**En conclusion.** Pour les fonctions bornées, nous disposons de trois théories :

- les fonctions Riemann intégrables : ce sont les fonctions qui sont encadrées par des fonctions en escalier proches en norme $L^1$. Les plus : c'est l'approche historique. Les moins : beaucoup de fonctions même assez régulières ne sont pas Riemann intégrables, la classe n'est pas stable par limite simple et les espaces de fonctions associés ne sont pas complets.

- les fonctions mesurables au sens de Borel : ce sont celles qui sont en tout point limite d'une suite de fonctions régulières et bornées par une même fonction bornée. Les plus : une fois construite la théorie est particulièrement robuste et se comporte bien vis à vis des limites simples, des espaces produits et les espaces de fonctions associés sont complets. Les moins : toutes les fonctions (même Riemann intégrables) ne sont pas mesurables. En particulier, la somme d'une fonction réglée et d'une fonction nulle sur le complémentaire d'un ensemble négligeable non borelien (par exemple beaucoup de sous-ensembles de l'ensemble de Cantor) est Riemann intégrable mais pas borélienne.

- les fonctions mesurables au sens de Lebesgue : ce sont celles qui sont presqu'en tout point limite d'une suite de fonctions régulières bornées par une même fonction bornée. Ce sont aussi les fonctions proches en norme $L^1$ des fonctions en escalier. Les plus : les fonctions Riemann intégrables et les fonctions Borel mesurables sont mesurables au sens de Lebesgue. La théorie se comporte bien vis à vis des limites simples et les espaces de fonctions associés sont complets. Les moins : moins explicite que la théorie de Borel et ne se comporte pas bien vis à vis des espaces produits.

Cependant, les espaces $L^p$ construits à partir des espaces $\mathcal{L}^p$ des fonctions mesurables au sens de Lebesgue et de Borel sont identiques ! En théorie des probabilités non plus, la complétion des tribus est une opération non nécessaire, voire dangereuse ! Il y a donc plus d'inconvénients que de bénéfices à considérer la tribu de Lebesgue, et nous nous restreignons donc (lorsqu'il est nécessaire de le préciser) aux tribus boréliennes dans la suite du cours.

# 5 Chronologie

• $\sim -250$. Le calcul effectif de la longueur d'une courbe du plan et de l'aire d'une forme géométrique du plan (ou du volume d'un solide de l'espace) sont des problèmes mathématiques connus et traités depuis l'antiquité. On doit en particulier à Archimède et Eudoxe de Cnide ($\sim -250$) le développement des méthodes d'exhaustion : rectification d'une courbe en l'approchant par une courbe affine par morceaux, quadrature d'une surface en l'approchant par des carrés ou des rectangles. On peut ainsi définir

$$\operatorname{aire}(S) := \sum_i \operatorname{aire}(Q_i), \tag{5.13}$$

pour une famille de carrés $(Q_i)$ contenus dans $S$, disjoints et dont l'union est $S$. Le calcul de l'aire de chaque $Q_i$ est aisé et est par définition le carré de la longueur d'un de ses côtés.

• XVI$^{\text{e}}$ siècle : Les méthodes des indivisibles (précurseurs des méthodes de changement de variables) sont développées par Cavalieri, Pascal (1659), Barrow (1660), Torricelli, ....

• 1669-1686 : On doit à Newton (1669) et à Leibniz (1684-86) le fondement du calcul infinitésimal (calcul différentiel et calcul intégral) : définition de la dérivée (ou fluxion chez Newton), formules de changement de variables dans l'intégrale, définition naïve de l'intégrale comme limite de la somme des surfaces de fines bandelettes sous-tendue par le graphe de la fonction lorsque l'épaisseur des bandelettes tend vers 0, la relation fondamental du calcul intégral établissant l'intégration comme l'opération inverse à la dérivation.

• 1707-1783 : Euler apporte la notation moderne $f(x)$ pour les fonctions. On lui doit également l'étude de problèmes liés à la sommation et aux intégrales (fonction $\Gamma$, résolution du problème de Bâle, ...). Bien que les fonctions étudiées soient des séries entières, il en a probablement déjà l'intuition d'une définition non explicite.

• 1812-1824 - Fourier établit l'équation de la chaleur et la résout grâce aux séries de Fourier (sommes infinies de fonctions trigonométriques, déjà présentes dans les travaux de Bernouilli sur l'équation les ondes). La question de la convergence des séries de Fourier sera une motivation puissante du développement de l'analyse (régularité, intégration) pendant plus d'un siècle. Fourier considère que toute fonction est développable en série de Fourier (...).

• 1821 - Cauchy démontre que l'intégrale d'une fonction (uniformément) continue $f : [a,b] \to \mathbb{R}$ est la limite des sommes (de Riemann) dans la leçon 21 de son cours d'analyse à l'école polytechnique. Sa preuve repose sur l'observation que la suite $(S_n)$ des sommes de Riemann vérifie le critère de Cauchy et utilise le fait que toute fonction continue sur $[a,b]$ est uniformément continue et que toute suite réelle vérifiant le critère de Cauchy est convergente (deux résultats qui ne seront vraiment compris et démontrés que quelques années plus tard par Heine et Cantor).

• 1828 - Dirichlet montre qu'une fonction discontinue ($\mathbf{1}_{\mathbb{Q}}$) peut être limite d'une suite de fonctions régulières (série de Fourier). En 1837, Dirichlet propose de manière claire la définition moderne d'une fonction numérique (dont l'intuition se trouve déjà dans les travaux d'Euler en 1755 ou de Fourier en 1822). Il utilise aussi la définition de la continuité pour obtenir le théorème (de Dirichlet) de convergence en tout point des séries de Fourier vers la fonction initiale, et plus généralement il établit la formule (1.8).

• 1852 - Riemann étend la définition de l'intégrale de Cauchy à une large classe de fonctions discontinues, mais surtout change de point de vue : une fonction est intégrable au sens de Riemann si elle est "proche" d'une fonction en escalier.

• 1858 & + : construction des nombres réels à partir des nombres rationnels par Dedekind (1858) grâce aux coupures, puis par Méray (1869) et Cantor (1872) grâce aux suites de Cauchy

• 1872 - Heine démontre que sur un intervalle fermé borné toute fonction continue est uniformément continue.

• 1872 – Weierstrass construit une fonction continue qui n'est dérivable en aucun point. Quelques années plus tôt il avait donné la définition moderne (avec $\varepsilon$) des fonctions continues.

• 1875 - Darboux donne une définition alternative et équivalente à Riemann des fonctions intégrables : ce sont les fonctions telles que les séries de Darboux par défaut et par excès convergent vers la même limite.

• 1881 – Jordan introduit les fonctions dites à variation bornée, et plus tard en 1887, il montre qu'elles sont naturellement reliées à la rectifiabilité des courbes. Il définit également une notion de mesure additive (mais pas encore $\sigma$-additive) sur $\mathbb{R}^d$ à partir des notions de mesure additive extérieure et mesure additive intérieure. Lorsque ces deux quantités coïncident on dit qu'un ensemble est quarrable.

• 1883 – Cantor introduit l'ensemble ternaire qui porte son nom, source de nombreux contre-exemples pathologiques mais intéressants. La théorie des ensembles est initiée par Cantor à partir des années 1870 : théorème de Cantor 1891, théorème de Cantor-Bernstein par Dedekind 1887, Bernstein 1897, Zermelo 1901.

• 1890 – Peano construit une courbe continue qui parcourt tous les points d'un carré. Peano introduit la notation des quantificateurs.

• 1894 – Stieltjes étend l'intégrale de Riemann au cas où « la mesure uniforme » est remplacée par « une mesure associée à une fonction à variation bornée » sur la droite réelle.

• 1898 – Borel introduit les ensembles mesurables pour lesquels il est possible de définir leur longueur (avec une définition naïve qui n'est pas complètement satisfaisante/rigoureuse) : ce sont les intervalles, leurs unions dénombrables et les complémentaires de ceux-ci, les unions dénombrables et les complémentaires des ensembles ainsi formés, et ainsi de suite . . ..

• 1899 – Baire défend sa thèse sur l'étude de la « *Continuité des fonctions limites de fonctions continues* » et poursuivra ensuite ses recherches sur les thèmes des liens entre théorie des ensembles et l'analyse et sur les fondements de la topologie et la continuité.

• 1902 – Lebesgue développe la théorie de la mesure et la théorie de l'intégration (essentiellement pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$). L'intégrale de Lebesgue naît en 1901 (dans une note de quelques pages aux comptes rendus de l'académie des sciences intitulée « Sur une généralisation de l'intégrale définie » qui s'intéresse à la définition de l'intégrale d'une fonction réelle bornée $f$ définie sur un intervalle $[a, b] \subset \mathbb{R}$). La construction repose sur deux idées :
- découper l'ensemble d'arrivée plutôt que l'ensemble de départ (contrairement à l'intégrale de Riemann) afin de se ramener à intégrer des fonctions caractéristiques d'ensemble ;
- définir l'intégrale d'une fonction caractéristique d'ensemble comme étant la longueur ou mesure de celui-ci, et il définit de manière convenable les ensembles « Lebesgue mesurables » et la mesure de ces derniers.
Lebesgue développera ensuite sa théorie en trois étape :
- sa thèse (1902) intitulée « Intégrale, longueur, aire » ;
- un premier cours Peccot au Collège de France (1902-1903) intitulé « Leçons sur l'intégration et la recherche des fonctions primitives » ;
- un second cours Peccot au Collège de France (1904-1905) intitulé « Leçons sur les séries trigonométriques ».

• 1905 – Vitali construit un ensemble non-mesurable en utilisant l'axiome du choix.

• 1906 – Fatou applique la théorie de Lebesgue à l'Analyse Complexe.

• 1906 – Fréchet participe aux fondements de l'analyse fonctionnelle et à la topologie (il définit les espaces métriques). On lui doit également la définition ensembliste moderne d'une fonction (entre ensembles abstraits).

• 1914 – Carathéodory démontre un « théorème d'extension » qui permet de construire de façon générale des mesures (sur des tribus).

• Terminons par une liste (qui ne se veut pas exhaustive) de quelques autres mathématiciens qui ont participé au développement de la théorie de la mesure, de l'intégration, de l'analyse fonctionnelle et des probabilités :
- en France : Lévy (thèse en 1912), Fatou (thèse en 1906), Denjoy ;

- en Italie : Beppo Levi (thèse à Turin en 1896), Vitali (thèse à Pise en 1899), Fubini (thèse à Pise en 1900), et Tonelli (Bologne 1906) ;
- en Autriche-Hongrie : Fischer (thèse à Vienne en 1899), Hahn (thèse à Vienne en 1902), Riesz (thèse à Budapest, 1902), Fèjer (hongrois, thèse à Berlin en 1902), Radon (thèse à Vienne en 1910) ;
- en Allemagne : Perron (thèse à Munich en 1902) et Hausdorff (thèse de physique à Leipzig en 1891) ;
- en Pologne : Steinhaus (thèse à Göttingen en 1911), Banach (thèse à Lvov en 1919) et Nikodym (thèse à Cracovie en 1924) ;
- en Russie : Egorov (thèse à Moscou en 1901), Sobolev (thèse à Leningrad en 1929) et Kolmogorov (thèse à Moscou en 1929) .
- enfin, aux Etats-Unis, Wiener et Doob participent au développement de la théorie des probabilités.