MAT311, Cours 5 : Théorie de la mesure [1]

§1. Des intégrales classiques aux tribus

1.1. Familles sommables

1.2. Ensemble $\overline{\mathbf{R}}_+$ et sommes de familles de nombres $\leqslant +\infty$

On note $\overline{\mathbf{R}}_+$ l'ensemble des nombre réels $\geqslant 0$ auxquels on a ajouté le symbole $+\infty$. On met sur cet ensemble les règles de calcul suivantes.

pour tout $a \in \overline{\mathbf{R}}_+$, on pose que : $(+\infty) + a = a + \infty = +\infty$;

pour tout $a \in \overline{\mathbf{R}}_+$ non nul, on pose que : $(+\infty) \times a = a \times (+\infty) = +\infty$;

enfin, on pose que : $(+\infty) \times 0 = 0 \times (+\infty) = 0$.

Définition 1.1. *Soit $\{u_i\}_{i\in I}$ une famille d'éléments de $\overline{\mathbf{R}}_+$ indexée par un ensemble d'indices I. On pose :*

$$\sum_{i\in I} u_i = \sup_{J\in\mathcal{P}_f(I)} \Big(\sum_{j\in J} u_j\Big) \leqslant +\infty,$$

où $\mathcal{P}_f(I)$ désigne l'ensemble des parties finies de l'ensemble (quelconque) I.

1.3. Propriétés de la somme Soient $\{u_i\}_{i\in I}$ et $\{v_i\}_{i\in I}$ des familles comme ci-dessus (même ensemble d'indices I). On a les propriétés suivantes :

1. **Croissance :** si $u_i \leqslant v_i$ pour tout $i \in I$, alors $\sum_{i\in I} u_i \leqslant \sum_{i\in I} v_i$.
2. **Homogénéité :** si $\lambda \in \overline{\mathbf{R}}_+$, alors $\sum_{i\in I} \lambda u_i = \lambda \sum_{i\in I} \lambda u_i$.
3. **Additivité :** on a : $\sum_{i\in I}(u_i + v_i) = \sum_{i\in I} u_i + \sum_{i\in I} v_i$.

Les preuves de ces énoncés relèvent de l'usage standard de la borne supérieure et permettent de prouver un théorème «à la Fubini» .

1.4. Petit Fubini et sommes par paquets

Proposition 1.1 («petit Fubini»)**.** *Soient I et J des ensembles et soit $\{u_{ij}\}_{i\in I\times J}$ une famille d'éléments de $\overline{\mathbf{R}}_+$ indexée par le produit $I \times J$. Alors on a :*

$$\sum_{i\in I}\Big(\sum_{j\in J} u_{ij}\Big) = \sum_{j\in J}\Big(\sum_{i\in i} u_{ij}\Big) = \sum_{(i,j)\in I\times J} u_{ij} \leqslant +\infty.$$

Corollaire 1.1. *Soit $\{I_\lambda\}_{\lambda\in\Lambda}$ une partition d'un ensemble I et soit $\{u_i\}_{i\in I}$ une famille d'éléments de $\overline{\mathbf{R}}_+$ indexée par I. Alors, on a :*

$$\sum_{i\in I} u_i = \sum_{\lambda\in\Lambda}\Big(\sum_{i\in I_\lambda} u_i\Big) \leqslant +\infty.$$

1.5. Intégrales classiques

[1] Bertrand Rémy (École polytechnique). Palaiseau, 31 mai 2017.

1.6. Deux points de vue sur l'intégration Revenons un instant sur la question de l'intégration des fonctions. Il y a essentiellement deux points de vue :

- celui des mesures (adapté par exemple aux probabilités, systèmes dynamiques etc.) ;
- celui des formes linéaires sur les espaces de fonctions (adapté par exemple aux équations aux dérivées partielles etc.).

Dans le second, on voit l'opération $f \mapsto \int_\Omega f(x)\,\mathrm{d}x$ d'intégrer sur un ouvert Ω de $\mathbf{R}^N$ comme une forme linéaire sur un espace de fonctions (convenables) sur Ω. Dans cette approche, on travaille sur des espaces de fonctions de plus en plus gros et on généralise l'intégrale par passages à la limite successifs. Les espaces fonctionnels ont en commun de contenir l'ensemble $\mathcal{C}_c(\Omega)$ des fonctions continues, à valeurs réelles, et nulles en-dehors d'un compact (non fixé) de Ω. En particulier, on impose que le calcul intégral sur $\mathcal{C}_c(\Omega)$ soit celui qu'on connaît déjà.

Une fois l'intégrale de Lebesgue construite par cette approche, on définit la mesure d'une partie (convenable) A de Ω comme étant $\int_\Omega \mathbf{1}_A(x)\,\mathrm{d}x$, où $\mathbf{1}_A$ est la fonction caractéristique de A.

1.7. Intégrale des fonctions réglées Dans cette séance, on choisit de privilégier (au départ) le point de vue des mesures, c'est-à-dire de commencer par le problème de définir une notion de volume pour des parties (convenables) de l'ensemble sur lequel on veut intégrer des fonctions (convenables). Nous allons motiver ce point de vue en revenant à la définition de deux intégrales classiques.

Rappelons qu'une fonction sur un intervalle $[a;b]$ est dite *réglée* si elle admet une limite à droite en tout point de $[a;b[$ et une limite à gauche en tout point de $]a;b]$. Un de résultats de base sur cette classe de fonctions est qu'elle coïncide avec l'ensemble des limites uniformes de fonctions en escalier sur $[a;b]$. La définition initiale permet vérifier concrètement qu'une fonction donnée est réglée. Quant au critère de limite uniforme, il permet justement de définir l'intégrale des fonctions réglées à partir de celle (évidente) des fonctions en escalier : si f est une telle fonction, on l'écrit comme limite uniforme $f = \lim_{n\to\infty} f_n$ et on pose :

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

Enfin, on vérifie que cette limite d'intégrales ne dépend pas du choix de la suite $(f_n)_{n\geqslant 0}$.

1.8. Intégrale de Riemann Rappelons que si f est une fonction $[a;b] \to \mathbf{R}$ bornée, l'*intégrale supérieure* (resp. *inférieure*) de f est la borne inférieure (resp. supérieure) des fonctions en escalier $\geqslant f$ (resp. $\leqslant f$). L'intégrale supérieure majore toujours l'intégrale inférieure, et f est dite *intégrable au sens de Riemann* (ou *Riemann-intégrable*) quand il y a égalité.

L'intégrale de Riemann est une généralisation de l'intégrale précédente dans le sens où toute fonction réglée est Riemann-intégrable, et où les deux calculs d'intégrale coïncident.

C'est une généralisation stricte dans le sens où il existe des fonctions non réglées mais Riemann-intégrables, comme par exemple la fonction $f:[0;1]\to\mathbf{R}$ qui vaut $\sin(\frac{1}{x})$ pour $x\in]0;1]$.

Les fonctions Riemann-intégrables sur un segment $[a;b]$ donné forment un espace vectoriel et l'application $f\mapsto\int_a^b f(x)\,\mathrm{d}x$ est une forme linéaire positive. L'application $f\mapsto\int_a^b |f(x)|\,\mathrm{d}x$ est une semi-norme : elle vérifie les axiomes d'une norme, sauf celui de séparation.

1.9. Problèmes avec l'intégrale de Riemann L'intégrale de Riemann présente un certain nombre d'inconvénients.

L'argument classique est que la fonction caractéristique $\mathbf{1}_{[0;1]\cap\mathbf{Q}}$ n'est pas Riemann-intégrable alors qu'elle est nulle sur une partie de $[0;1]$ de complémentaire dénombrable. Pire : en numérotant les nombres rationnels de $[0;1]\cap\mathbf{Q}$ on produit un suite $(f_n)_{n\geqslant 0}$ de fonctions Riemann-intégrables : f_n est la fonction caractéristique des n 《premiers》 nombres rationnels entre 0 et 1. Ces fonctions sont uniformément bornées, chacune est d'intégrale nulle et $(f_n)_{n\geqslant 0}$ converge simplement ; cependant, la limite n'est pas Riemann-intégrable.

Le point de départ de la théorie de l'intégration de Lebesgue est d'observer que toute fonction bornée est limite uniforme d'une suite de fonctions ne prenant qu'un nombre fini de valeurs. Ces dernières fonctions sont de la forme $\sum_{i=1}^{k}\lambda_k\mathbf{1}_{\{x\in[a;b]:f(x)=\lambda_k\}}$. Sous ce point de vue, par linéarité il s'agit de savoir calculer en premier lieu l'intégrale d'une fonction caractéristique d'une partie E de $[a;b]$; cette intégrale peut être vue comme une 《mesure》 de E.

1.10. Tribus de parties d'un ensemble

1.11. Les axiomes d'une tribu Le problème est qu'un segment réel admet beaucoup de parties. Pour pouvoir définir une notion cohérente de mesure de parties, on a recours à la notion de tribu qui sélectionne des parties mesurables dans un ensemble donné quelconque (pas nécessairement un intervalle réel).

Définition 1.2. *Une* tribu *(ou* σ-algèbre *) sur un ensemble X est un ensemble, disons $\mathcal{T}$, de parties de X satisfaisant les axiomes suivants :*

(i) *la partie vide $\varnothing$ est dans $\mathcal{T}$;*

(ii) *une réunion dénombrable de parties de X appartenant à $\mathcal{T}$ est encore une partie de X appartenant à $\mathcal{T}$;*

(iii) *le complémentaire d'une partie de X appartenant à $\mathcal{T}$ est encore une partie de X appartenant à $\mathcal{T}$.*

Pour faire court, : une tribu de parties de X est un ensemble de parties de X contenant le vide et X lui-même, et stable par réunion dénombrable et par passage au complémentaire.

1.12. Propriétés élémentaires, exemples et non exemples de tribus **Propriétés élémentaires :** il découle des axiomes qu'une tribu est stable par deux opérations ensemblistes importantes, la différence et la différence symétrique. Rappelons que si A et B sont des parties de X, la *différence* de A et de B est $A\setminus B=\{x\in X : x\in A \text{ mais } x\notin B\}$, et que leur *différence symétrique* est $A\Delta B=(A\cup B)\setminus(A\cap B)=(A\setminus B)\cup(B\setminus A)$.

Une tribu est stable par des opérations ensemblistes plus subtiles : plus grande et plus petite limite d'une suite de parties. Si $\{A_n\}_{n\geqslant 0}$ est une suite de parties de X, la plus grande limite $\overline{\lim}_{n\to\infty}A_n$ de $\{A_n\}_{n\geqslant 0}$ est l'ensemble des points qui appartiennent à A_n pour une infinité d'indices n, et la plus petite limite $\underline{\lim}_{n\to\infty}A_n$ de $\{A_n\}_{n\geqslant 0}$ est l'ensemble de ceux qui appartiennent à tous les A_n à partir d'un certain rang.

Exemple : la paire $\{\varnothing;X\}$, ainsi que l'ensemble $\mathcal{P}(X)$ de toutes les parties de X, sont des tribus sur X.

Non-exemple : si (X,d) est un espace métrique (par exemple, la droite réelle munie de la valeur absolue) la topologie de X, c'est-à-dire l'ensemble de ses ouverts, n'est en général pas une tribu sur X.

1.13. Tribus engendrées Il y a une relation d'ordre pour les tribus sur un ensemble X fixé : on dit qu'une tribu $\mathcal{A}$ *contient* une tribu $\mathcal{B}$ si toute partie de X qui est dans $\mathcal{B}$ figure aussi dans $\mathcal{A}$. Pour un ensemble *quelconque* de parties de X, on veut alors définir la plus petite tribu sur X contenant $\mathcal{E}$, ou tribu *engendrée* par $\mathcal{E}$.

Ce qui permet de le faire est de voir que les axiomes impliquent qu'*une intersection quelconque de tribus sur X reste une tribu.* Soit en effet $\{\mathcal{T}_i\}_{i\in I}$ une famille de tribus sur X. Le vide étant dans toutes les tribus $\mathcal{T}_i$, il est dans $\bigcap_{i\in I}\mathcal{T}_i$; si $A\in\bigcap_{i\in I}\mathcal{T}_i$ alors par stabilité par passage au complémentaire des tribus, on a $X\setminus A\in\mathcal{T}_i$ pour tout $i\in I$ soit $X\setminus A\in\bigcap_{i\in I}\mathcal{T}_i$, et la stabilité de $\bigcap_{i\in I}\mathcal{T}_i$ par réunion dénombrable se voit de façon similaire.

Définition 1.3. *La* tribu engendrée *par $\mathcal{E}$ est l'intersection de toutes les tribus contenant $\mathcal{E}$.*

Remarquons qu'il existe toujours au moins une tribu contenant $\mathcal{E}$, à savoir $\mathcal{P}(X)$.

1.14. Ensembles boréliens d'un espace métrique Considérons maintenant le cas particulier où l'ensemble X est muni d'une distance d, c'est-à-dire où X est un espace métrique. Dans ce cas, on a vu que la distance d, via les boules ouvertes associées, permettait de définir la collection des ouverts de X : un ouvert est une réunion de boules ouvertes. La collection des ouverts de X s'appelle aussi la *topologie* de X.

Définition 1.4. *Soit (X,d) un espace métrique. La* tribu borélienne *de X est la tribu engendrée par la topologie de X, c'est-à-dire par la collection des ouverts de X. Un ensemble* borélien *de X est un élément de la tribu borélienne de X.*

Par définition, les ouverts de X sont des boréliens, et par stabilité d'une tribu par passage au complémentaire, les fermés de X sont aussi des boréliens.

Remarque. On peut prouver qu'il y a beaucoup plus de parties de $\mathbf{R}$ que de boréliens, mais il la question d'exhiber une partie non borélienne de $\mathbf{R}$ est délicate.

1.15. Comportement vis-à-vis des flèches Puisqu'on a défini une structure (les tribus), il s'agit se savoir comment celle-ci se comporte vis-à-vis des applications. On va introduire à la fin de ce cours la notion d'application mesurable. Pour l'instant, on considère la notion de tribu image réciproque et de tribu induite.

Proposition 1.2. *Soit $f : X \to Y$ une application d'un ensemble X vers un ensemble Y.*

(i) *L'image réciproque par f de toute tribu de Y est une tribu sur X.*

(ii) *Si $\mathcal{A}$ est une tribu sur X, alors l'ensemble $\mathcal{B}$ des parties de Y dont l'image réciproque est dans $\mathcal{A}$ est une tribu de Y.*

Preuve. On prouve (ii), la démonstration étant similaire pour (i). La partie vide de Y a pour image réciproque par f la partie vide de X, qui est dans $\mathcal{A}$; ainsi $\varnothing \in \mathcal{B}$. Soit $B \in \mathcal{B}$; alors le complémentaire $Y \setminus B$ vérifie $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$. Or $f^{-1}(B) \in \mathcal{A}$ par définition de $\mathcal{B}$, et donc $X \setminus f^{-1}(B) \in \mathcal{A}$ par stabilité d'une tribu par passage au complémentaire ; ainsi $Y \setminus B \in \mathcal{B}$. La stabilité de $\mathcal{B}$ par réunion dénombrable vient de ce que l'image réciproque d'une réunion est la réunion des images réciproques. □

§2. Mesures positives sur les tribus

2.1. Notion de mesure et propriétés élémentaires

2.2. Dénombrabilité et limite inférieure La notion de dénombrabilité va s'avérer très importante dans ce cours : rappelons qu'un ensemble X est dit *dénombrable* s'il est fini ou en bijection avec $\mathbf{N}$. Pour cela, il suffit en fait qu'il existe une application surjective $\mathbf{N} \to X$ ou injective $X \to \mathbf{N}$. Cette notion formalise le fait d'être un《petit》 ensemble, elle est donc propre à définir les ensembles négligeables de la théorie de l'intégration. Elle est raisonnablement robuste dans le sens où elle est stable par produit fini et réunion dénombrable.

Nous aurons aussi besoin de la notion de *limite inférieure* (éventuellement infinie) d'une suite numérique. La droite numérique achevée $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$ est munie d'une distance qui induit la topologie usuelle sur $\mathbf{R}$ et assure qu'une suite réelle $(x_n)_{n\geqslant 0}$ tend vers $\pm\infty$ au sens de la distance de $\overline{\mathbf{R}}$ si, et seulement si, $\lim\limits_{n\to+\infty} x_n = \pm\infty$ au sens usuel. Une telle distance est par exemple donnée par $d(x,y) = |f(x) - f(y)|$ pour $f(x) = \frac{x}{1+|x|}$ et $f(\pm\infty) = \pm 1$. Elle fait de $\overline{\mathbf{R}}$ un espace compact, si bien que toute suite numérique $(x_n)_{n\geqslant 0}$ a des valeurs d'adhérence dans $\overline{\mathbf{R}}$, notamment une plus petite : celle-ci est notée $\varliminf\limits_{n\to\infty} x_n$ et est aussi égale à la limite de la suite croissante $(z_n)_{n\geqslant 0}$ des $z_n = \inf\{x_k : k \geqslant n\}$.

2.3. Notion de mesure

Définition 2.1. *Soit $(X,\mathcal{A})$ un* espace mesurable *, c'est-à-dire un ensemble X muni d'une tribu de parties – notée ici $\mathcal{A}$. On appelle* mesure (positive) *sur $(X,\mathcal{A})$ une application $\mu : \mathcal{A} \to \overline{\mathbf{R}}_+$ satisfaisant les axiomes suivants.*

(i) *On a $\mu(\varnothing) = 0$.*

(ii) *Si $\{A_n\}_{n\geqslant 0}$ est une suite (finie ou dénombrable) d'éléments de $\mathcal{A}$ deux à deux disjoints, on a :*

$$\mu\Big(\bigcup_{n\geqslant 0} A_n\Big) = \sum_{n\geqslant 0} \mu(A_n).$$

Un espace mesurable muni d'une mesure est appelé un espace mesuré *; c'est un triplet $(X,\mathcal{A},\mu)$ pour X, $\mathcal{A}$ et μ comme ci-dessus.*

Le second axiome s'appelle *additivité complète*, ou σ-*additivité*, et le premier axiome peut-être remplacé par la condition suivant laquelle μ n'est pas constante égale à $+\infty$.

2.4. Exemples de mesures **Exemples.** Voici une liste d'exemples de mesures, qui illustre la grande variété de possibilités.

1. La *mesure grossière* sur une tribu $\mathcal{A}$ est la mesure qui vaut $+\infty$ sur toutes les parties de $\mathcal{A}$ différentes de $\varnothing$.

2. La *mesure nulle* est celle qui vaut 0 tout le temps.

3. La *mesure de Dirac* en un point $a \in X$ est la mesure qui vaut 1 sur tout $A \in \mathcal{A}$ tel que $a \in A$, et 0 autrement.

4. La *mesure de décompte* sur X est la mesure – définie sur $\mathcal{P}(X)$ tout entier – qui vaut le cardinal de $A \subset X$ si A est fini, et $+\infty$ sinon.

On peut englober tous ces exemples sous la construction suivante. On se donne $m : X \to \overline{\mathbf{R}}_+$ une fonction et on pose : $\mu_m(A) = \sum\limits_{a\in A} m(a)$.

L'exemple qui nous intéressera le plus par la suite est celui qui conduit à la bonne généralisation de l'intégrale de Riemann : il s'agit de la *mesure de Lebesgue*.

2.5. Mesure de Lebesgue, première approche La *mesure de Lebesgue* est l'unique mesure définie sur la tribu borélienne de $\mathbf{R}^n$ et possédant les propriétés suivantes.

(i) Elle est invariante par toute translation de $\mathbf{R}^n$.

(ii) Elle attribue au cube unité la mesure 1.

Le point (i) signifie que deux boréliens déduits l'un de l'autre par translation ont même mesure de Lebesgue, et le point (ii) est une normalisation naturelle. À ce stade, l'existence et l'unicité de cette mesure, qu'on note λ_n ou λ tout court s'il n'y a pas de risque de confusion, ne sont pas du tout évidentes. Voici quelques propriétés pratiques avec lesquelles il faut se familiariser rapidement (elles sont assez naturelles une fois qu'on a adopté le changement de point de vue initial intégration → mesures).

1. Pour tout borélien A de $\mathbf{R}^n$, on a $\lambda_n(A) = \inf_O \lambda_n(O)$ où O parcourt l'ensemble des ouverts contenant A.

2. Pour tout borélien A de $\mathbf{R}^n$, on a $\lambda_n(A) = \sup_K \lambda_n(K)$ où K parcourt l'ensemble des compacts contenus dans A.

3. Pour des réels $a < b$, on a $\lambda_1([a;b]) = b - a$ et on peut ouvrir l'intervalle en a et/ou en b.

2.6. Croissance des mesures Nous passons désormais aux propriétés de bases des mesures générales, c'est-à-dire les propriétés qui se déduisent directement des axiomes.

Dans les énoncés qui suivent, $\mu : \mathcal{A} \to \overline{\mathbf{R}}_+$ désigne une mesure sur un espace mesuré $(X, \mathcal{A})$.

Lemme 2.1 (croissance des mesures)**.** *Une mesure $\mu : \mathcal{A} \to \overline{\mathbf{R}}_+$ est une fonction croissante pour la relation d'ordre donnée par l'inclusion sur les parties de la tribu $\mathcal{A}$; autrement dit, si A et B dans $\mathcal{A}$ sont tels que $A \subset B$, alors $\mu(A) \leqslant \mu(B)$.*

Preuve. On écrit A comme réunion disjointe de A et $B \setminus A$ (en observant que $B \setminus A$ est encore dans $\mathcal{A}$) et on applique l'axiome (ii) des mesures à la suite finie $\{A; B \setminus A\}$ d'éléments de $\mathcal{A}$. □

Remarquons que les conventions de calculs dans $\overline{\mathbf{R}}_+$ permettent d'écrire que

$$\mu(B \setminus A) = \mu(B) - \mu(A) \text{ si } \mu(A) < +\infty,$$

mais qu'on ne peut rien conclure de tel si A est de μ-mesure infinie.

2.7. Forte additivité des mesures Dans un contexte (comme le nôtre) où la tribu $\mathcal{A}$ sur un ensemble X est fixée, on dit – pour faire court – qu'un ensemble est *mesurable* s'il appartient à $\mathcal{A}$.

Lemme 2.2 (forte additivité des mesures)**.** *Si A et B sont des ensembles mesurables quelconques, alors on a :*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Preuve. On écrit $A \cup B$ comme réunion disjointe des trois ensembles (mesurables) $A \setminus B$, $B \setminus A$ et $A \cap B$. En appliquant l'axiome (ii) des mesures, on obtient :

$$\mu(A \cup B) = \mu(A \setminus B) + \mu(B \setminus A) + \mu(A \cap B).$$

Il reste à ajouter aux deux membres de l'égalité le terme $\mu(A \cap B)$ et à utiliser, comme dans la preuve précédente, que $\mu(A \setminus B) + \mu(A \cap B) = \mu(A)$ et $\mu(B \setminus A) + \mu(A \cap B) = \mu(B)$. □

2.8. Sous-additivité dénombrable En théorie de la mesure la lettre σ est souvent utilisée comme référence à la dénombrabilité, comme dans le résultat qui suit. La différence avec le second axiome des mesures est qu'on ne suppose pas que les A_n sont deux à deux disjoints, mais le résultat est une inégalité.

Lemme 2.3 (σ-sous-additivité des mesures)**.** *Pour toute suite (au plus) dénombrable d'ensembles mesurables $(A_n)_{n \geqslant 0}$ on a :*

$$\mu\Big(\bigcup_{n \geqslant 0} A_n\Big) \leqslant \sum_{n \geqslant 0} \mu(A_n).$$

Preuve. C'est évident dès qu'un des termes A_n est de μ-mesure infinie, ce qu'on exclut désormais. On fait alors la manipulation naturelle pour produire, à partir de $(A_n)_{n \geqslant 0}$, une suite $(B_n)_{n \geqslant 0}$ satisfaisant la condition de disjonction du second axiome. Précisément, on pose $B_1 = A_1$ et $B_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k$ pour $n \geqslant 2$. Cela founit :

$$\mu\Big(\bigcup_{n \geqslant 0} A_n\Big) = \mu\Big(\bigcup_{n \geqslant 0} B_n\Big) = \sum_{n \geqslant 0} \mu(B_n) \leqslant \sum_{n \geqslant 0} \mu(A_n). \qquad \square$$

2.9. Continuité à gauche

Lemme 2.4 (continuité à gauche des mesures)**.** *Pour toute suite croissante d'ensembles mesurables $(A_n)_{n \geqslant 0}$ et de limite (i.e. réunion) A, on a :*

$$\mu(A) = \lim_{n \to \infty} \mu(A_n) \leqslant +\infty.$$

Preuve. C'est la même manipulation. On pose $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geqslant 2$. La suite $(B_n)_{n \geqslant 0}$ est formée d'ensembles mesurables et deux à deux disjoints, donc on a :

$$\lim_{n \to \infty} \mu(A_n) = \sum_{m=1}^{\infty} \mu(B_m) = \mu\Big(\bigcup_{m=1}^{\infty} B_m\Big) = \mu(A).$$

□

2.10. Passage au complémenaire dans la continuité à gauche

Lemme 2.5 (suites décroissantes d'ensembles mesurables)**.** *Si $(A_n)_{n \geqslant 0}$ est une suite décroissante d'ensembles mesurables de limite (i.e. intersection) A, et si l'un de ses termes est de μ-mesure finie, on a :*

$$\mu(A) = \lim_{n \to \infty} \mu(A_n) \leqslant +\infty.$$

Preuve. Par hypothèse, pour un indice p, on a $\mu(A_p) < +\infty$. Le lemme précédent pour la suite croissante $(A_p \setminus A_n)_{n \geqslant p}$ fournit alors

$$\mu(A_p) - \mu(A) = \mu(A_p \setminus A) = \lim_{n \to \infty} \mu(A_p \setminus A_n) = \mu(A_p) - \lim_{n \to \infty} \mu(A_n),$$

et l'on peut conclure précisément parce qu'on peut simplifier par $\mu(A_p) < +\infty$. □

Remarque : la condition de finitude est indispensable (prendre pour μ la mesure de décompte sur $\mathbf{N}$ et la suite donnée par $A_n = \{\text{entiers} \geqslant n\}$; dans ce cas, $A = \bigcap_n A_n = \varnothing$, et donc $\mu(A) = 0$, alors que $\mu(A_n) = +\infty$ pour tout $n \geqslant 0$).

2.11. Parties négligeables et espaces mesurés complets

2.12. Parties négligeables Si l'on garde en tête que l'on veut à terme calculer des intégrales, on va maintenant s'intéresser aux parties sur lesquelles les intégrales seront nulles.

Définition 2.2. *Un* espace mesuré *est un triplet $(X, \mathcal{A}, \mu)$ où X est un ensemble, $\mathcal{A}$ est une tribu de parties de X et $\mu : \mathcal{A} \to \overline{\mathbf{R}}_+$ est une mesure sur l'espace mesuré $(X, \mathcal{A})$.*

Définition 2.3. *Pour un espace mesuré $(X, \mathcal{A}, \mu)$, une partie de X est dite* négligeable *(ou* μ-négligeable *) si elle est contenue dans une partie mesurable de mesure nulle.*

On note $\mathcal{N}$ l'ensemble des parties négligeables de $(X, \mathcal{A}, \mu)$.

Définition 2.4. *On dit que $(X, \mathcal{A}, \mu)$ est* complet **en tant qu'espace mesuré** *si toute partie négligeable est mesurable ; autrement dit, si $\mathcal{N} \subset \mathcal{A}$.*

2.13. Complétion et propriétés vraies presque partout On peut montrer que la mesure de Lebesgue sur la tribu des boréliens de la droite réelle $\mathbf{R}$ n'est pas complète, i.e. qu'il existe des parties négligeables pour la mesure de Lebesgue, mais non mesurables (c'est-à-dire non contenues dans la tribu borélienne). Ces problèmes peuvent être résolus – en théorie – car on peut toujours associer à un espace mesuré $(X,\mathcal{A},\mu)$ une plus petite tribu $\mathcal{A}'$ contenant $\mathcal{A}$ et $\mathcal{N}$ sur laquelle μ se prolonge en une mesure μ', faisant de $(X,\mathcal{A}',\mu')$ un espace mesuré. Nous ne détaillerons pas ces considérations dans ce cours.

En revanche, la notion suivante est importante pour la suite.

Définition 2.5. *Soit $(X,\mathcal{A},\mu)$ un espace mesuré. On dit qu'une propriété portant sur les points de X a lieu* presque partout *(ou μ-*presque partout *) si l'ensemble des $x \in X$ où elle n'est pas vraie est une partie négligeable de X.*

Remarque. On dit que μ est une mesure de probabilité si $\mu(X) = 1$. Dans ce cas, on dit d'une propriété qui est vraie μ-presque partout, qu'elle a lieu **presque sûrement**.

2.14. Mesure de Lebesgue

2.15. Retour à la mesure de Lebesgue Le cours suivant aura pour objet de définir l'intégrale des fonctions pour un espace mesuré $(X,\mathcal{A},\mu)$ général.

Pour rendre plus concrets les débuts de théorie de la mesure, mentionnons le polycopié qui contient la construction – au final **complètement équivalente** – de l'intégrale et de la mesure de Lebesgue sur un ouvert Ω d'un espace vectoriel $\mathbf{R}^n$, en privilégiant le point de vue de l'intégration des fonctions en premier lieu. La lecture (facultative) de cette construction est l'occasion :

1. de revoir les notions de convergence de suites de fonctions de la fin du cours de topologie ;
2. de comprendre un peu plus en détail le passage d'un point de vue à l'autre (mesures $\leftrightarrow$ intégrales).

2.16. Propriétés élémentaires de l'intégrale de Lebesgue Voici quelques faits importants sur l'intégrale de Lebesgue des fonctions à valeurs réelles.

1. L'ensemble $\mathcal{L}^1(\Omega)$ est un $\mathbf{R}$-espace vectoriel.
2. L'application $f \mapsto \displaystyle\int_\Omega f(x)\,\mathrm{d}x$ est une forme $\mathbf{R}$-linéaire sur $\mathcal{L}^1(\Omega)$.
3. On a l'inclusion : $\mathcal{C}_c(\Omega) \subset \mathcal{L}^1(\Omega)$ et l'intégrale de Lebesgue coïncide avec l'intégrale usuelle sur $\mathcal{C}_c(\Omega)$.
4. Si $f,g \in \mathcal{L}^1(\Omega)$ et $f \leqslant g$ p.p. sur Ω, alors
$$\int_\Omega f(x)\,\mathrm{d}x \quad\leqslant\quad \int_\Omega g(x)\,\mathrm{d}x.$$
5. Pour toutes $f,g \in \mathcal{L}^1(\Omega)$, on a :
$$|f|, \max(f,g), \min(f,g) \in \mathcal{L}^1(\Omega)$$
et
$$\left|\int_\Omega f(x)\,\mathrm{d}x\right| \leqslant \int_\Omega |f(x)|\,\mathrm{d}x.$$

2.17. De l'intégrale à la mesure de Lebesgue Pour tout $A \subset \mathbf{R}^N$ mesurable, la mesure de Lebesgue de A peut alors être définie par
$$|A| = \int_{\mathbf{R}^N} \mathbf{1}_A(x)\,\mathrm{d}x.$$
Grâce aux propriétés de l'intégrale de Lebesgue, on vérifie qu'on définit bien une mesure de cette façon, et cette approche a l'avantage de permettre d'utiliser toutes les techniques de calcul intégral (convergence monotone, dominée, Fubini, changement de variable etc.) pour calculer ou estimer des mesures d'ensembles. Par exemple, on a :

- La **mesure de Lebesgue d'un pavé** de $\mathbf{R}^N$ est égale à son volume $|\lfloor a_1,b_1\rceil \times \ldots \times \lfloor a_N,b_N\rceil| = \prod_{k=1}^N (b_k - a_k)$, où ici $\lfloor a,b\rceil$ est un intervalle de la forme $[a,b]$, $[a,b[$, $]a,b[$ ou $]a,b[$.
- Pour tout sous-ensemble mesurable $A \subset \mathbf{R}^N$ et **pour toute isométrie affine** T de $\mathbf{R}^N$, on a : $|T(A)| = |A|$.
- **Mesure de Lebesgue d'un ouvert de $\mathbf{R}$.** Pour tout ouvert Ω de $\mathbf{R}$, on a : $|\Omega| = \sum_{n\in\mathbf{N}}(b_n - a_n)$, où $\Omega = \bigcup_{n\in\mathbf{N}}]a_n,b_n[$ est la décomposition de Ω en intervalles deux à deux disjoints.

§3. Fonctions mesurables

3.1. Fonctions mesurables Voici une définition de fonction adaptée à la notion de tribu, qui se calque sur la caractérisation topologique des fonctions continues (usage des images réciproques).

Définition 3.1. *Si $(X,\mathcal{A})$ et $(Y,\mathcal{B})$ sont des espaces mesurés, on dit qu'une application $f : X \to Y$ est* mesurable *si pour tout $B \in \mathcal{B}$ on a $f^{-1}(B) \in \mathcal{A}$.*

Dans le cas de la mesure de Lebesgue, on peut voir qu'une fonction $f : \Omega \to [-\infty,+\infty]$ est mesurable (pour la tribu complétée) si et seulement s'il existe une suite $(f_n)_{n\geqslant 0}$, de fonctions continues à support compact qui sont définies sur Ω, qui converge vers f p.p. sur Ω. Si $f,g : \Omega \to \mathbf{R}$ sont mesurables et si $\lambda,\mu \in \mathbf{R}$, alors
$$\lambda f + \mu g, \qquad fg, \qquad \max(f,g) \qquad \text{et} \qquad \min(f,g)$$
sont des fonctions mesurables. En particulier
$$f^+ := \max(f,0), \qquad f^- := -\min(f,0) \qquad \text{et} \qquad |f|$$
sont mesurables.

3.2. Exemples de fonctions mesurables Voici quelques exemples ou énoncés de stabilité pour la mesurabilité des applications.

1. Si $f_1,\ldots,f_N$, définies sur Ω, sont mesurables et si $\Phi : \mathbf{R}^N \to \mathbf{R}$ est continue, alors
$$\Phi(f_1,\ldots,f_N)(x) := \Phi\big(f_1(x),\ldots,f_N(x)\big)$$
est mesurable.
2. Si $f(x) \neq 0$ p.p. sur Ω, alors $x \mapsto 1/f(x)$ est mesurable.
3. Toute fonction continue sur Ω est mesurable.
4. Toute fonction continue par morceaux sur un intervalle $I \subset \mathbf{R}$ est mesurable.
5. Toute fonction appartenant à $\mathcal{L}^1(\Omega)$ est mesurable.
6. Soit $(f_n)_{n\geqslant 0}$ une suite de fonctions mesurables, définies sur Ω et f telles que $f_n \to f$ p.p. sur Ω. Alors, f est mesurable.

Remarque : Il existe des fonctions non mesurables (leur construction repose toujours sur l'axiome du choix).

3.3. Intégration et fonctions mesurables Étant donnée une fonction f définie p.p. sur Ω, comment vérifie-t-on que $f \in \mathcal{L}^1(\Omega)$?

Théorème 3.1. *Soit f une fonction mesurable sur Ω. Supposons qu'il existe $g \in \mathcal{L}^+(\Omega)$ telle que $|f| \leqslant g$ p.p. sur Ω. Alors $f \in \mathcal{L}^1(\Omega)$.*

Preuve. Soit $f_n \in \mathcal{C}_c(\Omega)$ telles que $f_n \to f$ p.p. sur Ω. On note

$$h_n := \max\left(\min(f_n, g), -g\right),$$

Par construction $h_n \in \mathcal{L}^1(\Omega)$. De plus, $h_n \to f$ et $|h_n| \leqslant g$ p.p. sur Ω. Donc par convergence dominée) $f \in \mathcal{L}^1(\Omega)$. $\square$