cours 3, le jeudi 4 février 2010

II. Mesure et intégrale (de Lebesgue)

II.1. Mesures

II.1.1. *Opérations ensemblistes finies, « mesure de Riemann »*

On va travailler dans $X = [0, 1]$, intervalle fixé qui sera notre espace ambiant dans cette section. On pose pour $A \subset [0, 1]$ tel que l'indicatrice $\mathbf{1}_A$ soit R-intégrable

$$\mu(A) = \int_0^1 \mathbf{1}_A(t)\,dt \ge 0.$$

Désignons par $\mathcal{R}$ la classe des ensembles $A \subset [0, 1]$ tels que $\mathbf{1}_A$ soit R-intégrable.

Propriétés de la classe $\mathcal{R}$

Si $A \subset [0, 1]$, son complémentaire $[0, 1] \setminus A$ (dans $X = [0, 1]$, l'espace ambiant) sera noté A^c tant qu'il n'y aura aucune ambiguïté sur l'espace ambiant. Si $A \in \mathcal{R}$, alors A^c est aussi dans $\mathcal{R}$, puisqu'on a

$$\mathbf{1}_{A^c} = \mathbf{1} - \mathbf{1}_A,$$

et on en déduit que $\mu(A^c) = 1 - \mu(A) = \mu(X) - \mu(A)$. Plus généralement, pour les mêmes raisons, si $A, B \in \mathcal{R}$ et $A \subset B$, alors $B \setminus A \in \mathcal{R}$ et $\mu(B \setminus A) = \mu(B) - \mu(A)$.

Si A, B sont deux ensembles de la classe $\mathcal{R}$, l'intersection $A \cap B$ est aussi dans $\mathcal{R}$, puisque

$$\mathbf{1}_{A\cap B} = \mathbf{1}_A \mathbf{1}_B,$$

qui est intégrable comme produit de deux fonctions intégrables. Si $A, B \in \mathcal{R}$ et sont disjoints, l'indicatrice $\mathbf{1}_{A\cup B}$ est la somme des deux fonctions indicatrices, donc $A \cup B \in \mathcal{R}$ et

$$\mu(A \cup B) = \int_0^1 \mathbf{1}_{A\cup B}(t)\,dt = \int_0^1 \big(\mathbf{1}_A(t) + \mathbf{1}_B(t)\big)\,dt = \mu(A) + \mu(B)$$

dans ce cas.

Si $A, B \in \mathcal{R}$, alors $A \cup B \in \mathcal{R}$, puisque son complémentaire $A^c \cap B^c$ est dans $\mathcal{R}$. Pour calculer la mesure de la réunion en général, on découpe $A \cup B$ en trois parties disjointes $A \setminus B$, $A \cap B$ et $B \setminus A$. On obtient

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

On en déduit l'inégalité

$$\mu(A \cup B) \le \mu(A) + \mu(B).$$

Définition. Une algèbre de parties de X est une famille $\mathcal{A} \subset \mathcal{P}(X)$, contenant X, et telle que si $A \in \mathcal{A}$, alors $A^c \in A$ et si $A, B \in \mathcal{A}$, alors $A \cup B \in \mathcal{A}$. On dit que $\mathcal{A}$ est *stable par complémentaire et par réunion (finie)*. En conséquence, $\mathcal{A}$ est aussi stable par intersection finie.

Une mesure μ (finiment) additive finie et ≥ 0 sur $\mathcal{A}$ est une fonction de $\mathcal{A}$ dans $[0, +\infty[$ telle que

$$\mu(\mathrm{A} \cup \mathrm{B}) = \mu(\mathrm{A}) + \mu(\mathrm{B})$$

quand $\mathrm{A}, \mathrm{B} \in \mathcal{A}$ sont disjoints.

Comme les valeurs de μ sont finies, il en résulte en prenant $\mathrm{A} = \emptyset$ et $\mathrm{B} \in \mathcal{A}$ quelconque que $\mu(\emptyset) = 0$. On déduit pour toute $\mu \geq 0$ finie, finiment additive, les propriétés données ci-dessus pour la mesure de Riemann,

$$\mu(\mathrm{A}^c) = \mu(\mathrm{X}) - \mu(\mathrm{A}),$$

$$\mu(\mathrm{A} \cup \mathrm{B}) + \mu(\mathrm{A} \cap \mathrm{B}) = \mu(\mathrm{A}) + \mu(\mathrm{B}), \quad \mu(\mathrm{A} \cup \mathrm{B}) \leq \mu(\mathrm{A}) + \mu(\mathrm{B}).$$

Vers la théorie moderne

Bien avant la création de la théorie de la mesure moderne, on a eu envie de définir la « longueur » des ouverts en posant

$$|\mathrm{U}| = \sum_{n \geqslant 0} \mu(\mathrm{I}_n)$$

quand U est un ouvert de $[0, 1]$ obtenu comme réunion dénombrable d'intervalles disjoints $(\mathrm{I}_n)_{n \geqslant 0}$.

Donnons un exemple de construction qui était connu trente ans avant la création de la théorie de la mesure. L'ensemble Q des rationnels de $[0, 1]$ est dénombrable, on l'énumère sous la forme d'une suite $(r_n)_{n \geqslant 0}$ et on choisit une suite de nombres $\varepsilon_n > 0$ telle que

$$\sum_{n \geqslant 0} \varepsilon_n < \varepsilon,$$

où $\varepsilon > 0$ est petit et donné à l'avance ; on construit $\mathrm{I}_0 = (r_0 - \alpha_0, r_0 + \alpha_0)$, en choisissant $\alpha_0 < \varepsilon_0$ irrationnel > 0, pour garantir que les extrémités de I_0 ne soient pas rationnelles ; on désigne ensuite par récurrence n_{k+1} comme le plus petit indice tel que $r_{n_{k+1}}$ ne soit pas dans I_k, on choisit $\alpha_{k+1} < \varepsilon_{k+1}$ irrationnel > 0 tel que le nouvel intervalle $\mathrm{I}_{k+1} = (r_{n_{k+1}} - \alpha_{k+1}, r_{n_{k+1}} + \alpha_{k+1})$ soit disjoint de la réunion des précédents. On construit ainsi une suite d'intervalles ouverts disjoints ; la réunion est un ouvert V de « longueur » plus petite que ε, et qui contient tous les rationnels de $[0, 1]$.

Notons que l'indicatrice de V ne peut pas être intégrée par la méthode de Riemann ; en effet,

$$\mathbf{1}_{\mathrm{Q}} \leq \mathbf{1}_{\mathrm{V}}$$

entraîne que toute fonction en escalier $\varphi_2 \geq \mathbf{1}_{\mathrm{V}}$ est d'intégrale ≥ 1 ; mais dans la minoration $\varphi_1 \leq \mathbf{1}_{\mathrm{V}}$ on aura des intégrales $\int_0^1 \varphi_1 < \varepsilon$. Le problème de cet exemple, c'est qu'on ne savait pas (en 1880) quelle décision prendre : l'ensemble Q est contenu dans des ouverts de longueur arbitrairement petite, mais son adhérence est $[0, 1]$; faut-il décider que Q est petit, ou bien qu'il est « gros » ?

Le complémentaire $\mathrm{K} = [0, 1] \setminus \mathrm{V}$ est un ensemble fermé très compliqué, un *ensemble de Cantor*, ici un exemple de « mesure non nulle », contrairement à l'exemple plus courant donné par *l'ensemble triadique*. Tout comme pour le complémentaire V, la fonction indicatrice de K n'est pas R-intégrable.

On disposait tout de même d'une notion avec un parfum d'intégrale de Riemann. Le *contenu* d'un fermé F de $[0, 1]$ (une notion de l'époque de Cantor) est l'inf des mesures de Riemann des unions *finies* d'intervalles contenant F ; si U est l'ouvert complémentaire, Cantor savait que le contenu de F est $1 - |\mathrm{U}|$, où la longueur de l'ouvert est définie par la série ci-dessus.

L'exemple précédent est un ensemble de Cantor de contenu $> 1 - \varepsilon$. Toutes ces considérations deviendront très simples dans la théorie générale, qui englobera la longueur des ouverts et le contenu des fermés.

Le programme de Borel (1898)

Dans un livre de 1898, Émile Borel introduit sur $[0, 1]$ une classe d'ensembles qu'il nomme mesurables et il définit en même temps ces ensembles et leur mesure *par récurrence*, au moyen des trois principes suivants :

— un intervalle contenu dans $[0, 1]$ est mesurable et sa mesure est égale à sa longueur ;

— si (E_n) est une suite d'ensembles mesurables disjoints, leur réunion est mesurable et la mesure de la réunion est la somme de la série des mesures ;

— si un ensemble mesurable en contient un autre, leur différence est mesurable et a pour mesure la différence des mesures.

Borel n'est pas très clair sur la possibilité de cette « récurrence ». Il se contente d'indiquer en note de bas de page : « *Le théorème fondamental démontré pages 41-43 nous assure que ces définitions ne seront jamais contradictoires entre elles* ». Le théorème en question est le théorème de recouvrement connu aujourd'hui comme *théorème de recouvrement de Borel-Lebesgue*.

II.1.2 *Sommation de réels positifs*

On considère l'ensemble ordonné $[0, +\infty]$, qui a $+\infty$ comme plus grand élément. Tout sous-ensemble de $[0, +\infty]$ possède un plus petit majorant, autrement dit, une borne supérieure (on convient que $\emptyset$ admet 0, le plus petit élément, pour borne supérieure). On dit qu'une suite (x_n) dans cet ensemble tend vers $+\infty$ si elle dépasse tout nombre $c < +\infty$ à partir d'un certain rang. Toute suite croissante dans cet ensemble converge vers son sup.

Proposition. *Si $\sum u_n$ est une série, convergente ou non, à termes positifs, on a*

$$\sum_{n=0}^{+\infty} u_n = \sup\Big\{\sum_{j\in\mathrm{J}} u_j : \mathrm{J} \text{ fini, } \mathrm{J} \subset \mathbb{N}\Big\}.$$

Preuve. — On pose

$$\mathrm{S}_\infty = \sum_{n=0}^{+\infty} u_n,$$

qui est $+\infty$ si la série diverge. Si $\mathrm{J} \subset \mathbb{N}$ est un ensemble fini, il a un plus grand élément N, donc $\mathrm{J} \subset \{0, 1, \ldots, \mathrm{N}\}$ et

$$\Sigma_\mathrm{J} := \sum_{j\in\mathrm{J}} u_j \le \sum_{n=0}^{\mathrm{N}} u_n \le \mathrm{S}_\infty$$

donc S_∞ est un majorant pour l'ensemble des sommes finies $(\Sigma_\mathrm{J})_\mathrm{J}$, et par conséquent on a $\Sigma_\infty \le \mathrm{S}_\infty$, où Σ_∞ désigne le sup des sommes finies. L'inégalité inverse est évidente : si $\mathrm{J}_n = \{0, 1, \ldots, n\}$, on a

$$\mathrm{S}_n = \Sigma_{\mathrm{J}_n} \le \Sigma_\infty$$

et à la limite on obtient $\mathrm{S}_\infty \le \Sigma_\infty$.

Remarque. On a retrouvé le fait que la somme de la série (à termes positifs) ne dépend pas de l'ordre des termes.

Arithmétique de $[0, +\infty]$

On introduit une extension à $[0, +\infty]$ des opérations arithmétiques d'addition et de multiplication (il n'y a pas de bonne extension de la soustraction ni de la division) : on pose pour tout $a \in [0, +\infty]$

$$a + \infty = \infty + a = \infty\,;$$

si $a > 0$ on pose $a \times \infty = \infty \times a = \infty$ mais pour $a = 0$ on pose la convention importante

$$0 \times \infty = \infty \times 0 = 0.$$

On note la croissance des opérations : si $a_1 \le a_2$, $b_1 \le b_2$, on a $a_1 + a_2 \le b_1 + b_2$, $a_1 a_2 \le b_1 b_2$.

Si on a $a = \lim a_n$ et $b = \lim b_n$, où les limites sont prises dans $[0, +\infty]$ et où les suites (a_n), (b_n) sont croissantes, alors

$$a + b = \lim_n (a_n + b_n), \quad ab = \lim_n a_n b_n\,;$$

on a donc une sorte de « continuité » des opérations pour les limites de suites croissantes.

Remarque en passant : la définition d'une somme finie $\sum_{j \in \mathrm{J}} u_j$ se fait proprement par récurrence sur le nombre d'éléments de J, en utilisant l'associativité. On montre par récurrence que $\sum_{j \in \mathrm{J}} u_j$ possède la propriété suivante : si j_0 est un élément de J, on a

$$\sum_{j \in \mathrm{J}} u_j = \Bigl(\sum_{j \in \mathrm{J} \setminus j_0} u_j \Bigr) + u_{j_0}.$$

Preuve. — On commence par dire que la somme sur un ensemble vide d'indices est nulle, puis pour un singleton $\mathrm{J} = \{j_0\}$ on pose

$$\sum_{j \in \{j_0\}} u_j = u_{j_0}.$$

Pour le pas de récurrence, si J est de cardinal > 1, on pose

$$\sum_{j \in \mathrm{J}} u_j = \Bigl(\sum_{j \in \mathrm{J} \setminus j_0} u_j \Bigr) + u_{j_0},$$

où j_0 est un élément de J, et on vérifie grâce à l'associativité que le résultat ne dépend pas du j_0 choisi. Si j_1 est un autre élément de J, distinct de j_0, on écrit

$$\Bigl(\sum_{j \in \mathrm{J} \setminus j_0} u_j \Bigr) + u_{j_0} = \Bigl(\Bigl(\sum_{j \in \mathrm{J} \setminus \{j_0, j_1\}} u_j \Bigr) + u_{j_1} \Bigr) + u_{j_0} = \Bigl(\Bigl(\sum_{j \in \mathrm{J} \setminus \{j_0, j_1\}} u_j \Bigr) + u_{j_0} \Bigr) + u_{j_1}.$$

Si on fait une partition d'un ensemble fini J en sous-ensembles $(\mathrm{J}_\beta)_{\beta\in\mathrm{B}}$, et si on a donné des nombres $u_j \in [0,+\infty]$, alors

$$\sum_{j\in\mathrm{J}} u_j = \sum_{\beta\in\mathrm{B}} \Bigl(\sum_{j\in\mathrm{J}_\beta} u_j\Bigr).$$

Définition. Si $(u_i)_{i\in\mathrm{I}}$ est une famille d'éléments $u_i \in [0,+\infty]$, on pose

$$\sum_{i\in\mathrm{I}} u_i = \sup\{\sum_{j\in\mathrm{J}} u_j : \mathrm{J} \text{ fini } \subset \mathrm{I}\}.$$

Remarque. D'après la proposition précédente, cette définition est cohérente avec la définition de la somme d'une série à termes positifs.

Il est clair que $\mathrm{I}_1 \subset \mathrm{I}_2$ implique $\sum_{i\in\mathrm{I}_1} u_i \le \sum_{i\in\mathrm{I}_2} u_i$, puisque pour I_2 le sup porte sur plus de possibilités de sous-ensembles finis $\mathrm{J} \subset \mathrm{I}_2$.

Découpages arbitraires

Proposition. *Si* $\mathrm{I} = \bigcup_{\alpha\in\mathrm{A}} \mathrm{I}_\alpha$, *partition quelconque de* I *en sous-ensembles disjoints, on a*

$$\sum_{i\in\mathrm{I}} u_i = \sum_{\alpha\in\mathrm{A}} \Bigl(\sum_{i\in\mathrm{I}_\alpha} u_i\Bigr).$$

Preuve. — Si $\mathrm{J} \subset \mathrm{I}$ est un sous-ensemble fini, il n'y a qu'un ensemble fini $\mathrm{B} \subset \mathrm{A}$ d'indices $\alpha \in \mathrm{A}$ tel que J intersecte I_α selon un ensemble *non vide* $\mathrm{J}_\alpha = \mathrm{J} \cap \mathrm{I}_\alpha$, et J est la réunion disjointe de ces $(\mathrm{J}_\alpha)_{\alpha\in\mathrm{B}}$. Alors

$$\sum_{j\in\mathrm{J}} u_j = \sum_{\beta\in\mathrm{B}} \Bigl(\sum_{j\in\mathrm{J}_\beta} u_j\Bigr) \le \sum_{\alpha\in\mathrm{A}} \Bigl(\sum_{i\in\mathrm{I}_\alpha} u_i\Bigr),$$

d'où une première inégalité en prenant le sup en J. L'inégalité inverse est claire si l'une des sommes $\sum_{i\in\mathrm{I}_\alpha} u_i$ est infinie : dans ce cas les deux côtés sont égaux à $+\infty$. On les supposera donc toutes finies ; on prend un ensemble fini $\mathrm{B} \subset \mathrm{A}$, de cardinal noté $|\mathrm{B}|$, un $\alpha > 0$ tel que $\alpha|\mathrm{B}| < 1$ et pour chaque $\beta \in \mathrm{B}$ un sous-ensemble fini $\mathrm{J}_\beta \subset \mathrm{I}_\beta$ tel que

$$\sum_{j\in\mathrm{J}_\beta} u_j > \sum_{i\in\mathrm{I}_\alpha} u_i - \varepsilon\alpha.$$

On a alors

$$\sum_{i\in\mathrm{I}} u_i \ge \sum_{\beta\in\mathrm{B}} \Bigl(\sum_{j\in\mathrm{J}_\beta} u_j\Bigr) > \sum_{\beta\in\mathrm{B}} \Bigl(\sum_{i\in\mathrm{I}_\alpha} u_i\Bigr) - \varepsilon\alpha|\mathrm{B}|.$$

Il en résulte que $\sum_{i\in\mathrm{I}} u_i$ est un majorant de l'ensemble des sommes finies $\sum_{\beta\in\mathrm{B}}$, ce qui donne l'inégalité inverse, et termine la preuve.

Exemple traité. En découpant la somme des

$$u_{m,n} = \frac{x^m}{m!}\,\frac{y^n}{n!}$$

de deux façons, on retrouve $\mathrm{e}^{x+y} = \mathrm{e}^x\,\mathrm{e}^y$ (avec $x, y \ge 0$) par une méthode plus souple que celle qu'on pratique habituellement en L1-L2.

Ici, on travaille avec l'ensemble $\mathrm{I} = \mathbb{N} \times \mathbb{N}$; pour chaque indice $i = (m, n)$ on a défini $u_{m,n}$. Si on découpe I sous la forme $\mathrm{I} = \bigcup_{m \in \mathbb{N}} \{m\} \times \mathbb{N}$, la sommation groupée de cette façon donne

$$\sum_i u_i = \sum_{m \geqslant 0} \Bigl(\sum_{n \geqslant 0} \frac{x^m}{m!} \frac{y^n}{n!} \Bigr) = \sum_{m \geqslant 0} \frac{x^m}{m!} \,\mathrm{e}^y = \mathrm{e}^x \,\mathrm{e}^y \,.$$

L'autre découpage est obtenu en posant pour tout $k \geq 0$

$$\mathrm{I}_k = \{(m, n) \in \mathrm{I} : m + n = k\}$$

et avec ce découpage on obtient

$$\sum_{(m,n) \in \mathrm{I}_k} u_{m,n} = \frac{(x + y)^k}{k!}$$

d'où le résultat

$$\sum_i u_i = \sum_{k \geqslant 0} \Bigl(\sum_{i \in \mathrm{I}_k} u_i \Bigr) = \sum_{k \geqslant 0} \frac{(x + y)^k}{k!} = \mathrm{e}^{x+y} \,.$$

II.1.3. *Mesure σ-additive et σ-algèbre (ou tribu)*

Définition. Une σ-algèbre (ou tribu) $\mathcal{F}$ de parties de X est une algèbre de parties de X qui possède la propriété additionnelle que si $(\mathrm{A}_n)_{n \geqslant 0}$ est une suite de parties de $\mathcal{F}$, alors

$$\bigcup_{n \geqslant 0} \mathrm{A}_n \in \mathcal{F}.$$

Une σ-algèbre est donc une algèbre qui est stable par réunion dénombrable.

Une mesure $\mu \geq 0$ σ-additive sur $(\mathrm{X}, \mathcal{F})$ est une application $\mathcal{F} \to [0, +\infty]$ telle que $\mu(\emptyset) = 0$, et telle que pour toute famille finie ou dénombrable $(\mathrm{A}_i)_{i \in \mathrm{I}} \subset \mathcal{F}$ d'ensembles *deux à deux disjoints*, on ait

$$\mu\Bigl(\bigcup_{i \in \mathrm{I}} \mathrm{A}_i \Bigr) = \sum_{i \in \mathrm{I}} \mu(\mathrm{A}_i).$$

Exemples.

— Exemples sur un ensemble dénombrable I, égal à $\mathbb{Z}$ ou à $\mathbb{N}$ par exemple : la tribu $\mathcal{F}$ est la famille $\mathcal{P}(\mathrm{I})$ de toutes les parties de I ; il suffit de donner la mesure des singletons, on pourra ensuite trouver la mesure d'un sous-ensemble quelconque $\mathrm{A} \subset \mathrm{I}$ par la formule

$$\mu(\mathrm{A}) = \sum_{n \in \mathrm{A}} \mu\bigl(\{n\}\bigr).$$

Remarque : c'est une mesure σ-additive à cause du résultat sur le découpage arbitraire des sommations infinies.

La mesure de comptage sur $\mathbb{Z}$ est définie par $\mu\bigl(\{n\}\bigr) = 1$ pour tout $n \in \mathbb{Z}$. La mesure d'un ensemble $\mathrm{A} \subset \mathbb{Z}$ est le nombre d'éléments de cet ensemble, fini ou égal à $+\infty$.

La *mesure de Poisson* de paramètre $\theta > 0$ sur $\mathbb{N}$ est définie par

$$\forall n \in \mathbb{N}, \quad \mu\bigl(\{n\}\bigr) = \mathrm{e}^{-\theta} \,\frac{\theta^n}{n!}.$$

C'est une *mesure de probabilité* sur l'ensemble $\mathbb{N}$, c'est-à-dire que la mesure de l'espace entier est égale à 1.

— L'exemple qui sera le plus important pour nous est celui de la *mesure de Lebesgue*, notée λ ; son existence n'est pas évidente, et sera admise, au moins pour l'instant : il existe une tribu $\mathcal{F}$ de parties de $\mathbb{R}$, contenant les intervalles, et une mesure λ sur $\mathcal{F}$ telle que

$$\lambda([a, b]) = b - a$$

pour tout $a < b$ réels.

Remarque sur la complexité de cette tribu : elle contient les singletons $[a, a]$, qui sont de mesure nulle $a - a$, donc elle contient tous les ensembles dénombrables, tels que Q par exemple ; le problème de leur mesure est décidé par la σ-additivité imposée par la définition,

$$\lambda(\mathrm{Q}) = \sum_{n \geqslant 0} \lambda(\{r_n\}) = 0.$$

On voit que tout ouvert U de $\mathbb{R}$ est réunion *dénombrable* d'intervalles $[a, b]$, donc la tribu $\mathcal{F}$ contient tous les ouverts, et tous les fermés par passage au complémentaire. Mais l'histoire ne s'arrête pas là : on aura aussi les intersections $\mathrm{U} \cap \mathrm{F}$ d'un ouvert et d'un fermé, qu'on peut inclure dans la classe plus large des réunions dénombrables de fermés, les complémentaires de cette nouvelle classe, à nouveau les réunions dénombrables de la dernière classe obtenue, etc.

Il y a du flou sur la tribu $\mathcal{F}$ mentionnée ci-dessus. On introduira une tribu bien précise, la *tribu borélienne*, qui est la *plus petite* tribu contenant les intervalles de $\mathbb{R}$.