

Théorème de Lebesgue

Source : Kolmogorov, Fomine, *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*, page 319.

Théorème de Lebesgue : Toute fonction continue monotone f définie sur un segment $[a, b]$ est presque partout dérivable.

Le théorème que nous allons démontrer nécessite la définition et les deux lemmes suivants :

Définition : Un point $x \in [a, b]$ est dit invisible à droite pour la fonction g continue s'il existe $\xi \in]x, b]$ tel que $g(x) < g(\xi)$, et invisible à gauche s'il existe $\xi \in [a, x[$ tel que $g(x) < g(\xi)$.

Lemme de Riesz : Pour toute fonction continue g sur un intervalle $[a, b]$, l'ensemble des points invisibles à droite est ouvert dans $[a, b]$, et par conséquent, peut être représenté comme une union finie ou dénombrable d'ouverts deux à deux disjoints $]a_k, b_k[$ (et peut-être d'un intervalle semi-ouvert contenant l'extrémité a). Pour chacun de ces intervalles, on a $g(a_k) \leq g(b_k)$.

Démonstration : Si x_0 est un point invisible à droite pour g , en vertu de la continuité de g , tous les points assez voisins de x_0 jouissent de la même propriété. Ces points forment donc un ouvert de $[a, b]$ (voir notamment la *Topologie* de Queffélec pour une description des ouverts de $\mathbb{R}$).

Soit $]a_k, b_k[$ l'un des intervalles composant cet ensemble. Supposons par l'absurde que $g(a_k) > g(b_k)$. Il existe alors à l'intérieur de cet intervalle un point x_0 tel que $g(x_0) > g(b_k)$. Soit x^* le plus à droite des points de cet $]a_k, b_k[$ pour lequel $g(x) = g(x_0)$.

Comme $x^* \in]a_k, b_k[$, il existe un point $\xi > x^*$ ($\xi \in [x_0, b]$) tel que $g(\xi) > g(x^*)$. Par définition de x^* , $\xi \notin]a_k, b_k[$, car $g(x^*) = g(x_0) > g(b_k)$, et le théorème des valeurs intermédiaires nous donnerait $x^{**} \in]\xi, b_k[$ tel que $g(x^{**}) = g(x_0)$. Donc $\xi > b_k$.

Mais alors $g(b_k) < g(x_0) < g(\xi)$, ce qui est absurde, car b_k n'est pas invisible à droite. Donc $g(a_k) \leq g(b_k)$.

CQFD.

En réalité, $g(a_k) = g(b_k)$, sauf lorsque $a_k = a$.

Remarque : Les points invisibles à gauche vérifient le même théorème, sauf que cette fois, $g(b_k) \leq g(a_k)$. En effet, x_0 est invisible à gauche pour $g \iff a + b - x_0$ est invisible à droite pour $x \mapsto g(a + b - x)$, et le résultat découle de ce qui précède.

Lemme : Soit un ensemble mesurable A sur un segment $[a, b]$ tel qu'il existe $\rho \in]0, 1[$ vérifiant que pour tout intervalle $]\alpha, \beta[\subset [a, b]$, on ait l'inégalité : $\mu(A \cap]\alpha, \beta[) \leq \rho(\beta - \alpha)$. Alors $\mu(A) = 0$.

Démonstration : Posons $t \stackrel{\text{déf}}{=} \mu(A)$. Pour tout $\varepsilon > 0$, il existe un ensemble ouvert Ω , réunion dénombrable d'intervalles $]\alpha_m, \beta_m[$ tel que $A \subset \Omega$ et $\sum_m (\beta_m - \alpha_m) < t + \varepsilon$ (voir la définition de l'intégrale de Lebesgue rappelée par le Kolmogorov-Fomine). Posons $t_m \stackrel{\text{déf}}{=} \mu(A \cap]\alpha_m, \beta_m[)$.

Il est clair que $t = \sum_m t_m$. D'après l'hypothèse, $t_m \leq \rho(\beta_m - \alpha_m)$. Par conséquent,

$$t \leq \rho \sum_m (\beta_m - \alpha_m) < \rho(t + \varepsilon),$$

et comme ε est arbitraire, $t \leq \rho t$. Comme $\rho \in]0, 1[$, on obtient bien $\mu(A) = t = 0$.

CQFD.

Effectuons à présent un rappel sur les limites inférieures (que l'on utilise en général plutôt avec des suites), dans le cadre des fonction :

$$\liminf_{x \rightarrow x_0^+} \left(\frac{f(x) - f(x_0)}{x - x_0} \right) \stackrel{\text{déf}}{=} \sup_{\varepsilon > 0} \left(\inf_{x \in]x_0, x_0 + \varepsilon[} \left(\frac{f(x) - f(x_0)}{x - x_0} \right) \right).$$

On définit de même les limites inférieures à gauche, et supérieures à gauche et à droite.

Nous pouvons à présent démontrer le théorème de Lebesgue :

Démonstration du théorème de Lebesgue : Comme la dérivée de f en $x_0 \in]a, b[$ est la limite du taux d'accroissement $\frac{f(x) - f(x_0)}{x - x_0}$ quand $x \rightarrow x_0$, on introduit :

$$\begin{aligned} \Lambda_d &= \limsup_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0}, & \lambda_d &= \liminf_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0}, \\ \Lambda_g &= \limsup_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}, & \lambda_g &= \liminf_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}. \end{aligned}$$

Il est clair que $\lambda_d \leq \Lambda_d$ et $\lambda_g \leq \Lambda_d$ dans $\overline{\mathbb{R}}$. De plus, $f'(x_0)$ existe si, et seulement si,

$$-\infty < \lambda_g = \lambda_d = \Lambda_g = \Lambda_d < +\infty.$$

Quitte à prendre $-f$ au lieu de f , on peut supposer f croissante. On pose alors $f^* : x \mapsto -f(-x)$, qui est continue et croissante sur $[-b, -a]$. on montre alors facilement que :

$$\Lambda_g(x) = \Lambda_d^*(-x), \text{ et } \lambda_d(x) = \lambda_g^*(-x), \quad \forall x \in]a, b[.$$

Si l'on a démontré que μ -presque partout, $\Lambda_d < \infty$ et $\lambda_g \geq \Lambda_d$, pour toute fonction continue croissante, on obtient l'inégalité $\lambda_d = \lambda_g^*(-x) \geq \Lambda_d^*(-x) = \Lambda_g(x)$, puis par croissance de f :

$$-\infty < 0 \leq \Lambda_d \leq \lambda_g \leq \Lambda_g \leq \lambda_d \leq \Lambda_d < \infty \quad \mu\text{-pp},$$

et donc f est presque partout dérivable.

(i) Tout d'abord, si $\Lambda_d(x_0) = \infty$, c'est que $\forall c > 0, \exists \xi > x_0$ tel que :

$$\begin{aligned} \frac{f(\xi) - f(x_0)}{\xi - x_0} &> c, \quad \text{càd} \quad f(\xi) - f(x_0) > c(\xi - x_0) \\ &\quad \text{càd} \quad f(\xi) - c\xi > f(x_0) - cx_0. \end{aligned}$$

Donc x_0 est inversible à droite pour la fonction $g : x \mapsto f(x) - cx$.

D'après le lemme de Riesz, l'ensemble de ces points s'écrit $\bigsqcup_{k \in \mathbb{N}}]a_k, b_k[$, et pour chaque intervalle $]a_k, b_k[$, on a l'inégalité $f(a_k) - ca_k \leq f(b_k) - cb_k$, c'est-à-dire $f(b_k) - f(a_k) \geq c(b_k - a_k)$.

Comme $\{\Lambda_d = \infty\} \subset \bigsqcup_{k \in \mathbb{N}}]a_k, b_k[$, et divisant par c et en sommant, par croissance de f ,

$$\mu(\Lambda_d = \infty) \leq \sum_{k \in \mathbb{N}} (b_k - a_k) \leq \sum_{k \in \mathbb{N}} \frac{f(b_k) - f(a_k)}{c} \leq \frac{f(b) - f(a)}{c}.$$

Ceci étant vrai pour tout $c > 0$, en faisant tendre c vers $+\infty$, on trouve que $\mu(\Lambda_d = \infty) = 0$.

(ii) Soient à présent $0 < c < D < \infty$, et $\rho = c/D$. Nous introduisons également l'ensemble $E_{c,D} = \{\Lambda_d > D \text{ et } \lambda_g < c\}$. Soit également $]\alpha, \beta[\subset [a, b]$. Un point $x_0 \in]\alpha, \beta[$ tel que $\lambda_g(x_0) < c$ vérifie $\frac{f(\xi) - f(x_0)}{\xi - x_0} < c$ pour un certain $\xi < x_0$, càd $f(\xi) - c\xi > f(x_0) - cx_0$.

Donc x_0 est invisible à gauche la fonction $x \mapsto f(x) - cx$, donc par le lemme de Riesz, l'ensemble de ces points est $\bigsqcup_{k \in \mathbb{N}}]\alpha_k, \beta_k[$, avec

$$\forall k \in \mathbb{N}, f(\alpha_k) - c\alpha_k \geq f(\beta_k) - c\beta_k, \quad \text{càd} \quad f(\beta_k) - f(\alpha_k) \leq c(\beta_k - \alpha_k).$$

Pour $k \in \mathbb{N}$, considérons $\Omega_k = \{\Lambda_d > D\} \cap]\alpha_k, \beta_k[$. En raisonnant comme précédemment, toujours d'après le lemme de Riesz, $\Omega_k \subset \bigsqcup_{j \in \mathbb{N}}]\alpha_{k,j}, \beta_{k,j}[$: ensemble des points invisibles à droite pour $x \mapsto f(x) - Dx$. On a donc, $\forall j \in \mathbb{N}$,

$$\beta_{k,j} - \alpha_{k,j} \leq \frac{1}{D}(f(\beta_{k,j}) - f(\alpha_{k,j})).$$

Donc $E_{c,D} \cap]\alpha, \beta[\subset \bigsqcup_{k,j \in \mathbb{N}}]\alpha_{k,j}, \beta_{k,j}[$, et on a

$$\begin{aligned} \mu(E_{c,D} \cap]\alpha, \beta[) &\leq \sum_{k,j \in \mathbb{N}} (\beta_{k,j} - \alpha_{k,j}) \\ &\leq \frac{1}{D} \sum_{k,j \in \mathbb{N}} (f(\beta_{k,j}) - f(\alpha_{k,j})) \\ &\leq \frac{1}{D} \sum_{k \in \mathbb{N}} \sum_{j \in \mathbb{N}} (f(\beta_{k,j}) - f(\alpha_{k,j})) \\ &\leq \frac{1}{D} \sum_{k \in \mathbb{N}} (f(\beta_k) - f(\alpha_k)) \\ &\leq \frac{c}{D} \sum_{k \in \mathbb{N}} (\beta_k - \alpha_k) \\ &\leq \rho(\beta - \alpha). \end{aligned}$$

Or, comme $\rho \in]0,1[$, par le lemme, $\mu(E_{c,D}) = 0$. Enfin,

$$\{\Lambda_d > \lambda_g\} \subset \bigcup_{\substack{c,D \in \mathbb{Q} \\ c < D}} \{\Lambda_d > D > c > \lambda_g\} = \bigcup_{\substack{c,D \in \mathbb{Q} \\ c < D}} E_{c,D}.$$

Donc par continuité croissante de la mesure de Lebesgue μ , $\mu(\Lambda_d > \lambda_g) = 0$.

Moralité : $\Lambda_d < \infty$ et $\lambda \geq \Lambda_d$ μ -presque partout, d'où f est presque partout dérivable.

CQFD.

Remarque culturelle : Le "vrai" théorème de Lebesgue est plus général : l'hypothèse de continuité n'est pas nécessaire. pour le démontrer, on suit le même raisonnement, en définissant cette fois les points invisibles à droite comme étant :

$$\{x \in [a,b] \mid \exists \xi > x_0 ; \max(g(x_0^+), g(x_0), g(x_0^-)) < g(\xi)\}.$$

Le lemme de Riesz se généralise à cette définition, et avec lui le théorème de Lebesgue.

Leçons concernées :

- 228 - *Continuité, dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications*
- 229 - *Fonctions monotones. Fonctions convexes. Exemples et applications.*

Remarque : Ce développement est assez long, avec de nombreuses notations à introduire (mais peut-être peut-on en introduire certaines dans le plan ?). La définition et les deux lemmes qui suivent sont à mettre dans le plan et à admettre bien évidemment (mais il est bon d'en connaître les preuves, au moins dans les grandes lignes).