Fonction indicatrice de $\mathbb{Q}$ sur [0;1]

Soit la fonction f *définie sur l'intervalle* $[0;1]$ *par :*

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{si } x \notin \mathbb{Q} \end{cases}$$

On va montrer que cette fonction n'est pas intégrable au sens de Riemann, mais qu'elle est intégrable au sens de Lebesgue.

<u>Preuve</u> : notons

$$\sigma_n = \left\{0 = x_0^{(n)} < x_1^{(n)} < \ldots < x_{n-1}^{(n)} < x_n^{(n)} = 1\right\}$$

une subdivision du segment $[0;1]$. Soit de plus $\xi_n = \{\xi_n(i)\}_{1 \leq i \leq n}$ un ensemble de points tels que

$$x_{i-1}^{(n)} < \xi_n(i) < x_i^{(n)}.$$

On définit alors la somme de Riemann $S(f, \sigma_n, \xi_n)$ associée à la fonction f, à la subdivision σ_n, et à l'ensemble ξ_n par :

$$S(f, \sigma_n, \xi_n) = \sum_{i=1}^{n} f\left(\xi_n(i)\right)\left(x_i^{(n)} - x_{i-1}^{(n)}\right)$$

On sait alors que l'intégrale de Riemann $I_R(f)$ de f, si elle existe, est définie par :

$$I_R(f) = \int_0^1 f(x)dx = \lim_{n \to +\infty} S(f, \sigma_n, \xi_n) \tag{1}$$

De plus, cette définition est indépendante du choix de la suite de subdivisions σ_n et de l'ensemble ξ_n.
Prenons alors la subdivision spécifique suivante :

$$\sigma_n = \left\{0; \frac{1}{n}; \frac{2}{n}; \ldots; \frac{n-1}{n}; 1\right\},$$

et considérons les deux ensembles distincts suivants :

$$\begin{aligned} \xi_n &= \{\xi_n(i)\}_{1 \leq i \leq n} \text{ avec } \xi_n(i) \in \left]x_{i-1}^{(n)}; x_i^{(n)}\right[\cap \mathbb{Q} \\ \xi'_n &= \left\{\xi'_n(i)\right\}_{1 \leq i \leq n} \text{ avec } \xi'_n(i) \in \left]x_{i-1}^{(n)}; x_i^{(n)}\right[- \mathbb{Q} \end{aligned}$$

$\xi_n(i)$ et $\xi'_n(i)$ étant respectivement rationnel et irrationnel, on a :

$$\forall i \in \{1, \ldots, n-1\}, \quad f\left(\xi_n(i)\right) = 1 \text{ et } f\left(\xi'_n(i)\right) = 0. \tag{2}$$

D'autre part, la limite (1) est vraie si et seulement si :

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \geq 0 \text{ tel que } n \geq N_\varepsilon \Rightarrow |I_R(f) - S\left(f, \sigma_n, \xi_n\right)| < \varepsilon.$$

Posons alors $\varepsilon = 1/4$. On a :

$$\exists N \geq 0 \text{ tel que } n \geq N \Rightarrow |I_R(f) - S\left(f, \sigma_n, \xi_n\right)| < 1/4$$

et

$$\exists N' \geq 0 \text{ tel que } n \geq N' \Rightarrow \left|I_R(f) - S\left(f, \sigma_n, \xi'_n\right)\right| < 1/4.$$

Ainsi, pour $n \geq \max(N, N')$:

$$\left|S\left(f, \sigma_n, \xi_n\right) - S\left(f, \sigma_n, \xi'_n\right)\right| \leq |I_R(f) - S\left(f, \sigma_n, \xi_n\right)| + \left|I_R(f) - S\left(f, \sigma_n, \xi'_n\right)\right| < 1/2. \tag{3}$$

Or, on a d'autre part, d'après (2) :

$$S\left(f,\sigma_n,\xi_n\right)=\sum_{i=1}^{n} f\left(\xi_n(i)\right)\frac{1}{n}=\frac{1}{n}\sum_{i=1}^{n}1=1$$

et

$$S\left(f,\sigma_n,\xi_n'\right)=\sum_{i=1}^{n} f\left(\xi_n'(i)\right)\frac{1}{n}=\frac{1}{n}\sum_{i=1}^{n}0=0\ ,$$

et donc

$$\left|S\left(f,\sigma_n,\xi_n\right)-S\left(f,\sigma_n,\xi_n'\right)\right|=1\ ,$$

ce qui contredit l'inégalité (3). Ainsi la limite $I_R(f)$ dans (1) diffère selon l'ensemble ξ_n considéré, ce qui n'est pas possible.
La fonction f n'est donc pas Riemann-intégrable.
Par contre, l'intégrale de Lebesgue $I_L(f)$ de f s'écrit [1]:

$$\begin{aligned} I_L(f) &= \int_{[0;1]} f d\lambda \\ &= \int_{[0;1]\cap\mathbb{Q}} f d\lambda + \int_{[0;1]-\mathbb{Q}} f d\lambda \\ &= \int_{[0;1]\cap\mathbb{Q}} 1 d\lambda + \int_{[0;1]-\mathbb{Q}} 0 d\lambda \\ &= 1\times\lambda\left([0;1]\cap\mathbb{Q}\right)+0\times\lambda\left([0;1]-\mathbb{Q}\right) \\ &= 1\times 0+0\times 1 \\ &= 0 \end{aligned} \tag{4}$$

Ainsi, la fonction f est bien Lebesgue-intégrable, et son intégrale est nulle. En réalité, cela provient simplement du fait que f est une fonction *nulle presque partout* : on aboutirait au même résultat avec la fonction g définie sur $[0;1]$ par

$$g(x)=\begin{cases} 1 & \text{si } x\in\mathcal{C} \\ 0 & \text{si } x\notin\mathcal{C}\end{cases}$$

où $\mathcal{C}$ est *l'ensemble de Cantor*, dont on montre ci-dessous qu'il est de mesure de Lebesgue nulle bien qu'il soit non-dénombrable.

Ensemble de Cantor

Notons C_0 *l'intervalle* $[0;1]$. *On construit* C_1 *à partir de* C_0 *en enlevant l'intervalle central* $]1/3;2/3[$. *On a donc* $C_1=[0;1/3]\cup[2/3;1]$.
De la même façon, on définit C_2 *en enlevant de* C_1 *les intervalles* $]1/9;2/9[$ *et* $]7/9;8/9[$, *soit* $C_2=[0;1/9]\cup[2/9;3/9]\cup[6/9;7/9]\cup[8/9;1]$.
On construit ainsi de proche en proche C_n *à partir de* C_{n-1} *en enlevant dans chaque intervalle de* C_{n-1} *(de longueur* $1/3^{n-1}$*) l'intervalle central (de longueur* $1/3^n$*).*
On pose alors :

$$\mathcal{C}=\bigcap_{n=1}^{+\infty}C_n$$

[1] Noter que l'expression de $I_L(f)$ donnée dans (4) provient de la construction même de l'intégrale de Lebesgue à partir de fonctions dites *étagées*, ou *simples*, c'est-à-dire prenant un ensemble fini de valeurs, sur des ensembles mesurables disjoints - mais pas nécessairement adjacents.

L'ensemble $\mathcal{C}$ est appelé ***ensemble de Cantor****. Montrons maintenant qu'il possède les deux propriétés suivantes :*

1) l'ensemble de Cantor est de mesure de Lebesgue nulle :
En effet, chaque ensemble C_n est constitué de 2^n intervalles fermés notés $(I_{n,k})_{1\leq k\leq 2^n}$, tous de longueur (donc de mesure de Lebesgue) $1/3^n$.
Chaque ensemble C_n est donc fermé comme réunion finie d'intervalles fermés. Ainsi, $\mathcal{C}$ est également fermé (intersection dénombrable de fermés).
De plus, la suite $(C_n)_{n\in\mathbb{N}}$ est une suite décroissante, c'est-à-dire $C_{n+1}\subset C_n$. On a donc :

$$\lambda(C_{n+1})\leq\lambda(C_n)$$

où λ désigne la mesure de Lebesgue. D'après le *théorème de continuité décroissante*, on a alors :

$$\lambda(C)=\lambda(\bigcap_{n=1}^{+\infty}C_n)=\lim_{n\to+\infty}\lambda(C_n)=\lim_{n\to+\infty}\sum_{k=1}^{2^n}1/3^n=\lim_{n\to+\infty}\left(\frac{2}{3}\right)^n=0$$

2) l'ensemble de Cantor n'est pas dénombrable :
On commence par montrer le résultat suivant :
lemme : *tout élément x de $\mathcal{C}$ peut s'écrire de façon unique sous la forme suivante (appelée "décomposition dyadique") :*

$$x=\sum_{n=1}^{+\infty}\frac{x_n}{3^n}\text{ avec }x_n\in\{0,2\}\tag{5}$$

Réciproquement, tout réel x de la forme (5) *est élément de $\mathcal{C}$.*
Preuve : $\mathcal{C}$ est constitué des bornes des intervalles $(I_{n,k})_{1\leq k\leq 2^n,n\geq 1}$. Tout élément x de $\mathcal{C}$ peut donc s'écrire sous la forme

$$x=\sum_{n=1}^{+\infty}\frac{x_n}{3^n}\text{ avec }x_n\in\{0,1,2\}$$

On fait ainsi correspondre à x la suite $(x_1,x_2,\ldots,x_n,\ldots)\in\{0,1,2\}^{\mathbb{N}}$.
On va alors montrer que tout élément de $\mathcal{C}$ associé à une suite du type $(x_1,x_2,\ldots,x_{n_0-1},1,x_{n_0+1},\ldots)$ peut en fait être représenté par la suite $(x_1,x_2,\ldots,\ x_{n_0-1},0,2,2,\ldots,2,\ldots)$ ou par la suite $(x_1,x_2,\ldots,x_{n_0-1},2,0,0,\ldots,0,\ldots)$.
Désignons donc par n_0 le premier entier (s'il existe) tel que $x_{n_0}=1$. Alors, x est une borne supérieure d'un intervalle I_{n_0,k_0}. De plus, on a :

$$x-\sum_{n=1}^{n_0}\frac{x_n}{3^n}\leq\sum_{n=n_0+1}^{+\infty}\frac{2}{3^n}=\frac{2}{3^{n_0+1}}\sum_{n=0}^{+\infty}\frac{1}{3^n}=\frac{1}{3^{n_0}}=\lambda(I_{n_0,k_0})$$

Supposons alors que les $(x_n)_{n\geq n_0+1}$ ne sont pas tous nuls, et qu'il existe $n\geq n_0+1$ tel que $x_n\neq 2$. On a alors

$$0<x-\sum_{n=1}^{n_0}\frac{x_n}{3^n}<\frac{1}{3^{n_0}}\ .\tag{6}$$

Or, $\sum\limits_{n=1}^{n_0}\frac{x_n}{3^n}$ représente la borne supérieure de l'intervalle I_{n_0,k_0}, et, par construction de $\mathcal{C}$, l'intervalle ouvert de longueur $\frac{1}{3^{n_0}}$ se trouvant à droite de ce point n'appartient pas à $\mathcal{C}$ (puisque cet intervalle a été enlevé lors de la construction de C_{n_0+1} à partir de C_{n_0}). Mais d'après (6), x "tombe" dans cet intervalle. Cela implique que $x\notin\mathcal{C}$.
Ainsi, x ne peut être représenté que par l'une des deux suites suivantes

$$(x_n^1)_{n\in\mathbb{N}}=(x_1,x_2,\ldots,x_{n_0-1},1,0,0,\ldots,0,\ldots)$$

ou

$$\left(x_n^2\right)_{n\in\mathbb{N}} = (x_1, x_2, \ldots, x_{n_0-1}, 1, 2, 2, \ldots, 2, \ldots)$$

Dans le cas de $\left(x_n^1\right)_{n\in\mathbb{N}}$, on a :

$$x = \sum_{n=1}^{n_0-1} \frac{x_n}{3^n} + \frac{1}{3^{n_0}} \ ,$$

que l'on peut écrire sous la forme

$$x = \sum_{n=1}^{n_0-1} \frac{x_n}{3^n} + \frac{0}{3^{n_0}} + \sum_{n=n_0+1}^{+\infty} \frac{2}{3^n}$$

Ainsi, la suite $\left(x_n^1\right)_{n\in\mathbb{N}}$ peut également s'écrire

$$\left(x_n^1\right)_{n\in\mathbb{N}} = (x_1, x_2, \ldots, x_{n_0-1}, 0, 2, 2, \ldots, 2, \ldots)$$

Dans le cas de $\left(x_n^2\right)_{n\in\mathbb{N}}$, on a :

$$\begin{aligned} x &= \sum_{n=1}^{n_0-1} \frac{x_n}{3^n} + \frac{1}{3^{n_0}} + \sum_{n=n_0+1}^{+\infty} \frac{2}{3^n} \\ &= \sum_{n=1}^{n_0-1} \frac{x_n}{3^n} + \frac{2}{3^{n_0}} \end{aligned}$$

Ainsi, la suite $\left(x_n^2\right)_{n\in\mathbb{N}}$ s'écrit aussi sous la forme

$$\left(x_n^2\right)_{n\in\mathbb{N}} = (x_1, x_2, \ldots, x_{n_0-1}, 2, 0, 0, \ldots, 0, \ldots)$$

On a donc bien montré que $(x_1, x_2, \ldots, x_{n_0-1}, 1, x_{n_0+1}, \ldots)$ pouvait être représenté par la suite $(x_1, x_2, \ldots, x_{n_0-1}, 0, 2, 2, \ldots, 2, \ldots)$ ou par la suite $(x_1, x_2, \ldots, x_{n_0-1}, 2, 0, 0, \ldots, 0, \ldots)$.
Ainsi, tout élément de $\mathcal{C}$ s'écrit sous la forme (5). De plus, cette décomposition est unique.
En effet, supposons que x puisse être représenté par deux suites différentes $\left(x_n^1\right)_{n\in\mathbb{N}}$ et $\left(x_n^2\right)_{n\in\mathbb{N}}$. Soit $m = \min\left\{n \in \mathbb{N} \,\middle|\, x_n^1 \neq x_n^2\right\}$. On a alors $\left(x_m^1, x_m^2\right) = (2, 0)$ ou $\left(x_m^1, x_m^2\right) = (0, 2)$. On peut supposer que $\left(x_m^1, x_m^2\right) = (2, 0)$. On obtient :

$$x = \sum_{n=1}^{m-1} \frac{x_n^1}{3^n} + \frac{2}{3^m} + \sum_{n=m+1}^{+\infty} \frac{x_n^1}{3^n} = \sum_{n=1}^{m-1} \frac{x_n^2}{3^n} + \frac{0}{3^m} + \sum_{n=m+1}^{+\infty} \frac{x_n^2}{3^n}$$

d'où :

$$\frac{2}{3^m} = \sum_{n=m+1}^{+\infty} \frac{x_n^2 - x_n^1}{3^n} \tag{7}$$

Or, s'il existe $n \geq m+1$ tel que $x_n^2 - x_n^1 \neq 2$, on a :

$$\sum_{n=m+1}^{+\infty} \frac{x_n^2 - x_n^1}{3^n} < \sum_{n=m+1}^{+\infty} \frac{2}{3^n} = \frac{2}{3^m} \ ,$$

et l'égalité (7) ne peut avoir lieu. On doit donc avoir

$$\forall n \geq m+1, \quad x_n^1 = 0 \text{ et } x_n^2 = 2.$$

Ainsi, les suites $\left(x_n^1\right)_{n\in\mathbb{N}}$ et $\left(x_n^2\right)_{n\in\mathbb{N}}$ doivent être de la forme :

$$\left(x_n^1\right)_{n\in\mathbb{N}} = (x_1, x_2, \ldots, x_{m-1}, 2, 0, 0, \ldots, 0, \ldots) \tag{8}$$

et

$$\left(x_n^2\right)_{n\in\mathbb{N}} = (x_1, x_2, \ldots, x_{m-1}, 0, 2, 2, \ldots, 2, \ldots) \ ,$$

et ces deux écritures sont équivalentes d'après (7). On choisira alors dans ce cas d'associer à x la représentation (8). La forme (5) est alors unique.
Réciproquement, soit x vérifiant la relation (5). On a alors :

$$x = \lim_{N\to+\infty} \widehat{x}_N \text{ avec } \widehat{x}_N = \sum_{n=1}^{N} \frac{x_n}{3^n}$$

Or, d'après la construction des ensembles $(C_n)_{n\in\mathbb{N}}$, on a $\widehat{x}_N \in C_N$, $\forall N \in \mathbb{N}$. Puisque $\mathcal{C} = \lim_{N\to+\infty} C_N$, on a bien $x \in \mathcal{C}$. Le lemme est ainsi démontré.
Etant donné le lemme, $\mathcal{C}$ est dénombrable si et seulement si il existe une fonction ϕ bijective de $\mathbb{N}$ dans $\mathcal{C}$ telle que :

$$\begin{aligned} \phi : \quad & \mathbb{N} \longrightarrow \mathcal{C} \\ & n \longmapsto x = \sum_{k=1}^{+\infty} \frac{x_k^{(n)}}{3^k} \end{aligned}$$

On utilise maintenant le *procédé diagonal de Cantor*. A chaque entier $n = 0, 1, 2, \ldots$ est associé une suite $\left(x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots\right) \in \{0, 2\}^{\mathbb{N}}$ qui représente ses "coordonnées" dans $\mathcal{C}$:

$$\begin{array}{ccc} n = 1 & \longrightarrow & x_1^{(1)}, x_2^{(1)}, \ldots, x_k^{(1)}, \ldots \\ n = 2 & \longrightarrow & x_1^{(2)}, x_2^{(2)}, \ldots, x_k^{(2)}, \ldots \\ \vdots & \vdots & \vdots \\ n & \longrightarrow & x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots \end{array}$$

Considérons alors les termes diagonaux $\left(x_n^{(n)}\right)_{n\in\mathbb{N}}$. On définit :

$$x_n = \begin{cases} 0 & \text{si } x_n^{(n)} = 2 \\ 2 & \text{si } x_n^{(n)} = 0 \end{cases} \tag{9}$$

On pose alors

$$x = \sum_{n=1}^{+\infty} \frac{x_n}{3^n}$$

Ainsi, d'après le lemme, on a $x \in \mathcal{C}$ (car on a bien $x_n \in \{0, 2\}$).

Par contre, on a $x \notin \phi(\mathbb{N})$. En effet, si on avait $x \in \phi(\mathbb{N})$, il existerait $n_0 \in \mathbb{N}$ (unique) tel que $\phi(n_0) = x$.
Or, à $\phi(n_0)$ est associé la suite $\left(x_1^{(n_0)}, x_2^{(n_0)}, \ldots, x_{n_0}^{(n_0)}, \ldots\right)$, et à x la suite $(x_1, x_2, \ldots, x_{n_0}, \ldots)$. Mais d'après (9), on a

$$x_{n_0} \neq x_{n_0}^{(n_0)} \ ,$$

ce qui contredit le fait que $x = \phi(n_0)$ (puisque la décomposition dyadique est unique). Donc $x \notin \phi(\mathbb{N})$.
Ainsi, il n'existe pas de bijection entre $\mathbb{N}$ et $\mathcal{C}$, ce qui montre que $\mathcal{C}$ n'est pas dénombrable.