

VI.3 Mesure et Intégration

Correction du III.2.1 (page 58)

Si X est dénombrable, toute partie de X est dénombrable, donc réunion dénombrable de singletons. La tribu engendrée par les singletons est donc la tribu discrète $\mathcal{P}(X)$.

Si X n'est pas dénombrable, la tribu engendrée par les singletons contient au moins toutes les parties dénombrables, ainsi que celles de complémentaire dénombrable. Or la famille \mathcal{A} constituées de ces parties est une tribu. En effet, la stabilité par complémentarité est immédiate. Considérons maintenant une famille (A_n) de \mathcal{A} . Si tous les A_n sont dénombrables, alors leur réunion l'est, elle est donc dans \mathcal{A} . Si l'un d'eux, mettons A_1 n'est pas dénombrable, alors A_1^c l'est, donc l'intersection des complémentaires est dénombrable, et donc son complémentaire, qui est l'union des A_n , est dans \mathcal{A} . il s'agit donc bien de la tribu engendrée par les singletons.

Correction du III.2.2 (page 59)

Soit f une application constante : $f(x) = a' \in X'$. On a

$$f^{-1}(A') = \emptyset \text{ si } x \notin A', \quad f^{-1}(A') = X \text{ si } x \in A',$$

d'où $f^{-1}(A') \in \mathcal{A}$, puisque toute tribu contient \emptyset et X , même la plus grossière d'entre elles.

Correction du III.2.3 (page 59)

D'après la définition, l'identité est mesurable si et seulement si tout élément de \mathcal{A}' est dans \mathcal{A} , c'est à dire si $\mathcal{A}' \subset \mathcal{A}$ (\mathcal{A} est plus fine que \mathcal{A}').

Correction du III.2.4 (page 60)

Si \mathcal{A} est la tribu discrète, toute application est mesurable. Si \mathcal{A} est la tribu grossière, alors dès qu'il existe A' tel que $f^{-1}(A')$ est non vide et non identifié à X , l'application f n'est pas mesurable. Les seules applications mesurables sont donc les applications qui sont d'une certaine manière constantes, mais dans un sens un peu particulier : elles sont constantes *autant que* \mathcal{A}' *soit en mesure de distinguer les valeurs*. Plus précisément, il s'agit d'applications telles que, si $f(x) \neq f(y)$, alors nécessairement les éléments de \mathcal{A}' qui contiennent x sont exactement ceux qui contiennent y . Dès que la tribu \mathcal{A}' distingue tous les éléments, c'est-à-dire que pour tous x' et $y' \in X'$ il existe $A' \in \mathcal{A}$ tel que $x \in A'$, $y \notin A'$, alors les seules applications mesurables sont donc les applications constantes $f(x) \equiv a' \in X'$. A l'extrême inverse, si \mathcal{A}' est la tribu grossière, alors toutes les applications sont mesurables.

La propriété énoncée précédemment dans le cas où \mathcal{A} est la tribu grossière est vraie *a fortiori* pour tout \mathcal{A} : si \mathcal{A}' est la tribu grossière, toute application est mesurable (quelle que soit \mathcal{A}). Si \mathcal{A}' est la tribu discrète, une application est mesurable si et seulement si l'image réciproque de toute partie est dans \mathcal{A} . Si \mathcal{A} est discrète, tout est mesurable. S'il existe des ensembles non mesurables sur X , dont on note la collection \mathcal{A}^c (complémentaire de \mathcal{A} dans $\mathcal{P}(X)$), alors les applications non mesurables sont celles pour lesquelles il existe un $B \in \mathcal{A}^c$ image réciproque d'une partie A' de $\mathcal{P}(X')$. De telles parties sont caractérisées par $f(B) \cap f(B^c) = \emptyset$, qui implique alors $f^{-1}(f(B)) = B$, qui est donc non mesurable.

Correction du III.4.1 (page 67)

Les inégalités à vérifier pour que μ^* soit une mesure extérieure sont toutes des tautologies du type $0 \leq 0$, $0 \leq 1$, $1 \leq 1$, $1 \leq +\infty$.

Si X est un singleton, μ^* est une mesure. En revanche si X contient au moins 2 éléments a et b , on a $\mu(\{a\} \cup \{b\}) = 1 < 2 = \mu(\{a\}) + \mu(\{b\})$, il ne s'agit donc pas d'une mesure.

Si X est réduit à un singleton, toutes les parties (il y en a deux au total : X et \emptyset) sont mesurables. Si ça n'est pas un singleton, alors pour toute partie B de X non vide et non égale à X , on a

$$\mu^*(X \cap B) + \mu^*(X \cap B^c) = 1 + 1 > 1 = \mu^*(X),$$

la partie B n'est donc pas mesurable. Les seules parties mesurables pour μ^* sont donc \emptyset et X .

Correction du III.4.2 (page 73)

Toute réunion d'intervalles ouverts dont la somme des longueurs est finie ne pouvant recouvrir \mathbb{R} la mesure extérieure de Lebesgue de \mathbb{R} , qui s'identifie à $\lambda(\mathbb{R})$, est infinie. On peut en revanche écrire \mathbb{R} comme réunion des $] -n, n[$, qui sont tous de mesure finie.

Correction du III.9.1 (page 101)

[N.B. cette correction correspond à une version de l'exercice qui a été modifiée par rapport à celle faite en TD, on ne considère pas des suites extraites mais un ré-ordonnement des termes]

On déduit des hypothèses que x_n tend vers 0, et que x_n n'est pas nulle au-delà d'un certain rang. Par ailleurs, si l'on note I_+ l'ensemble des indices tels que $x_n \geq 0$, et I_- l'ensemble des indices tels que $x_n < 0$, ces deux ensembles sont infinis, et l'on a (du fait de la non convergence absolue)

$$\sum_{I_+} x_n = +\infty, \quad \sum_{I_-} x_n = -\infty.$$

Pour produire une limite infinie, on commence par prendre des indices dans I_+ jusqu'à ce que la série partielle dépasse 1, puis on prend le premier indice de I_- , puis on continue avec I_+ jusqu'à dépasser 2, puis le deuxième indice de I_- etc ... Le nombre d'étape est forcément infinie du fait de la divergence de la série sur I_+ . On parcourt ainsi tous les indices de I_+ et I_- , et la série diverge vers $+\infty$ par construction (noter que les négatifs que l'on ajoute un par un à chaque étape tendent vers 0). On procède symétriquement pour $-\infty$. Pour λ , par exemple positif, on commence par parcourir les indices de I_+ jusqu'à dépasser λ , puis on passe à I_- jusqu'à passer en dessous de λ , puis I_+ , etc... Comme la suite des x_n tend vers 0, on a bien convergence vers λ . Pour avoir un ensemble de valeurs prises dense dans \mathbb{R} , on monte jusqu'à dépasser 1 avec des indices de I_+ , puis on descend sous -2 , puis on remonte au-dessus de 3, etc ... Comme x_n tend vers 0, les progressions de $-n$ à $n+1$ se font avec un pas de plus en plus petit, on finit donc par intersecter n'importe quel intervalle ouvert, aussi petit soit-il.

Cet exercice montre que cela n'a en général aucun sens d'écrire

$$\sum_{n \in I} x_n$$

lorsque I est un ensemble infini dénombrable et que les x_n prennent des valeurs positives et négatives, du fait que l'ordre dans lequel on effectue les additions conditionne fortement le résultat obtenu.

Correction du III.9.2 (page 101)

Pour tout n , on pose

$$I_n = \left\{ i \in I, a_i > \frac{1}{n} \right\}.$$

D'après l'hypothèse, cet ensemble est fini. L'ensemble

$$I_\infty = \{i \in I, a_i > 0\} = \bigcup_{n \in \mathbb{N}} I_n$$

est donc dénombrable comme union dénombrable d'ensembles finis.

Correction du III.9.3 (page 101)

On sait que \mathbb{D} (ou \mathbb{Q}) est dénombrable et dense dans \mathbb{R} , on s'en donne une énumération (q_n) . On considère maintenant la réunion des intervalles ouverts $I_n =]q_n - \varepsilon_n, q_n + \varepsilon_n[$, avec $\varepsilon_n = \varepsilon/2^{n+2}$. Il s'agit d'un ouvert dense, qui est une réunion d'intervalles ouverts (même si certains intervalles de la collection initiale s'intersectent, cela reste une réunion d'intervalles ouverts), dont la longueur totale est majorée par la somme des longueurs des intervalles, qui vaut ε .

Correction du III.9.4 (page 101)

a) Soit \mathcal{A} une tribu sur N . Soit $x \in X$. On considère A_x le plus petit élément de \mathcal{A} qui contient x . On considère maintenant y dans A_x . Si $x \notin A_y$, alors $A_x \setminus A_y$ est élément de \mathcal{A} , est strictement inclus dans A_x , et contient x , ce qui est absurde. On a donc nécessairement $x \in A_y$. Alors A_y s'identifie nécessairement à A_x , car sinon l'intersection des deux est strictement incluse dans au moins l'un d'eux, par exemple A_x , ce qui est absurde car alors A_x ne serait plus le plus petit élément de \mathcal{A} contenant x . La partie A_x contient donc les éléments tels que A_x soit le plus petit élément de \mathcal{A} les contenant. On poursuit avec un point hors de A_x (s'il y en a un), et l'on continue jusqu'à avoir recouvert X_N par p parties non vides disjointes, qui constituent donc une partition P . La tribu engendrée par la partition P est contenue dans \mathcal{A} . Pour tout élément A de \mathcal{A} , son intersection avec chacune des parties de la partition est la partie-elle-même, sinon elle ne serait pas minimale. La tribu \mathcal{A} ne contient donc que des unions de parties de P , donc elle est dans la tribu engendrée par P . On établit ainsi une correspondance univoque (pour un ensemble fini, rappelons-le) entre tribus et partitions. Toute tribu sur X_N est engendrée par une partition de X_N , elle est ainsi caractérisée par cette partition.

b) Le cardinal d'une tribu est défini par la granularité de la partition qui l'engendre. Si la partition contient p parties ($1 \leq p \leq N$), alors la tribu contient autant d'éléments que l'ensemble à p éléments contient de parties, c'est à dire 2^p . Une autre manière de voir les choses est de considérer que l'on peut *coder* un élément de la tribu par un mot de p bits, dont chaque bit indique si l'élément en question contient la partie correspondante.

c) Comme on l'a vu au a), compter les tribus revient à compter les partitions. On suppose connu le nombre B_k de tribus sur l'ensemble à k éléments, pour k entre 0 et N . On rajoute maintenant l'élément $x = n + 1$ à X_N pour obtenir X_{N+1} . On se propose d'énumérer les partitions sur X_{N+1} en les classant selon le nombre k d'éléments qui *ne sont pas* dans la partie qui contient x . Pour $k = 0$, on a une seule partition, constituée de l'ensemble X_{N+1} entier. Pour $k = 1$, on a N possibilités pour choisir l'élément qui n'est pas dans la partie contenant x , et chaque choix correspond à une seule partition. Pour $k \geq 2$, on a C_N^k possibilités de choisir les éléments qui ne sont pas dans la partie contenant x . Pour chacun de ces choix, on

a B_k partitions possibles. On a donc

$$B_{N+1} = \sum_{k=0}^N C_N^k B_k,$$

avec $B_0 = 1$ (on considère que l'ensemble vide admet une partition unique, qui est lui-même), et $B_1 = 1$ (tout singleton $\{1\}$ admet une partition unique). On appelle B_N le N -ième nombre de *Bell*, dont on peut montrer qu'il est égal à

$$B_N = \sum_{k=0}^{+\infty} \frac{k^N}{k!}.$$

Correction du III.9.5 (page 101)

On sait que la tribu borélienne est engendrée par les intervalles du type $] - \infty, a]$. Chacun de ces intervalles peut s'écrire comme l'intersection d'intervalles $] - \infty, q_n]$, où q_n est une suite de nombres décimaux qui convergent par valeurs décroissantes vers a . La tribu engendrée par les $] - \infty, q]$, où q est décimal, contient donc les intervalles $] - \infty, a]$, et donc la tribu borélienne, et il s'agit d'un ensemble dénombrable car en bijection avec \mathbb{D} .

Correction du III.9.6 (page 101)

a) On a $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$. Par ailleurs, si $A' \in \mathcal{A}'$, alors

$$f^{-1}((A')^c) = \left(f^{-1}(A')\right)^c \in \mathcal{A},$$

d'où la stabilité par passage au complémentaire. On a enfin, si (A'_n) est une collection d'éléments de \mathcal{A}' ,

$$f^{-1}\left(\bigcup_n A'_n\right) = \bigcup_n f^{-1}(A'_n) \in \mathcal{A}.$$

b) Si \mathcal{A}'' est une tribu sur X' qui rend f mesurable, pour tout A'' dans \mathcal{A}'' , on a $f^{-1}(A'') \in \mathcal{A}$, d'où $A'' \in \mathcal{A}'$. On a donc $\mathcal{A}'' \subset \mathcal{A}'$.

c) On a $\nu(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset)$. Par ailleurs, pour toute collection (A'_n) d'éléments disjoints de \mathcal{A}' , on a

$$\nu\left(\bigcup_n A'_n\right) = \mu\left(f^{-1}\left(\bigcup_n A'_n\right)\right) = \sum_n \mu\left(f^{-1}(A'_n)\right),$$

car les $f^{-1}(A'_n)$ sont disjoints. Il s'agit donc bien d'une mesure qui fait de (X', \mathcal{A}', ν) un espace mesuré.

On a bien sûr conservation de la masse totale, du fait que $\nu(X') = \mu(f^{-1}(X')) = \mu(X)$.

Si la mesure μ est la loi de probabilité d'une variable aléatoire X , alors ν est simplement la loi de la variable aléatoire $f(X)$.

c bis) Remarquons en premier lieu, même si ça ne répond pas encore à la question, que s'il existe un singleton $\{x'\}$ dans \mathcal{A}' de mesure nulle, alors l'application constante $f(x) \equiv x'$ affecte une masse nulle à tout $A \in \mathcal{A}$. Il s'agit bien dans ce cas d'une mesure, mais elle est identiquement nulle, la masse n'est donc pas conservée. On se place maintenant pour simplifier dans le cas où $X' = X$, $\mathcal{A}' = \mathcal{A}$, et considérons le cas où il existe un singleton $\{x'\}$ de mesure non nulle. Alors l'application constante considérée précédemment, donne une

même valeur à tous les éléments de \mathcal{A} , ce qui invalide l'additivité (sauf dans des situations très particulières), par exemple dès qu'il existe dans \mathcal{A} deux ensemble disjoints de mesures non nulles.

d) (i) $f(x) = x + c$. Comme vu dans le cours, les translations laissent invariante la mesure de Lebesgue, on a donc $\nu = \lambda$.

(ii) $f(x) = \alpha x$, avec $\alpha \neq 0$. Toute partie A mesurable est dilatée par l'homothétie de rapport α . La mesure d'un objet A dans l'espace d'arrivée est donc $\lambda(A)/\alpha$.

(iii) $f(x) = E(x)$. Cette application balaie en quelque sorte tous les intervalles $[n, n+1[$ et les concentre en $\{n\}$. La mesure image est donc la mesure de comptage sur \mathbb{N} :

$$\nu(A) = \text{Card} \{n \in \mathbb{N}, n \in A\} = \text{Card}(A \cap \mathbb{N}).$$

(iv) $f(x) = 0$. Cette application envoie toute la masse en 0. La mesure image est donc très singulière, il s'agit d'une masse ponctuelle (infinie) en 0 : $\nu(A) = 0$ si $0 \notin A$, $\nu(A) = +\infty$ si $0 \in A$,

(v) En premier lieu, notons que $\nu(A') = 0$ dès que $A' \cap]0, +\infty[= \emptyset$, on a donc

$$\nu(A') = \nu(A' \cap]0, +\infty[).$$

Pour tous a, b , avec $0 < a < b$, la mesure de $]a, b[$ est $\log(b) - \log(a) = \log(b/a)$. On a donc, pour tout $y > 0$, $dy > 0$,

$$\nu(]y, y + dy]) = \log\left(\frac{y + dy}{y}\right) \sim \frac{dy}{y}.$$

La mesure ν a donc sur $]0, +\infty[$ une densité par rapport à la mesure de Lebesgue égale à $1/y$.

e) Soit f et g sont égale μ -presque partout, alors $\mu(f^{-1}(A')) = \mu(g^{-1}(A'))$ pour tout $A' \in \mathcal{B}$.

f) Si μ est une masse ponctuelle, alors $f_{\#}\mu$ est aussi une masse ponctuelle. Si ν ne l'est pas, alors $\Lambda_{\mu, \nu} = \emptyset$. On a un singleton quand les deux sont des masses ponctuelles. Si μ et ν sont toutes deux sommes de n masses ponctuelles de même poids,

$$\mu = \sum_{i=1}^n \delta_{x_i}, \quad \nu = \sum_{j=1}^n \delta_{y_j},$$

où les x_i sont distincts deux à deux, tout comme les y_j , alors $\Lambda_{\mu, \nu}$ contient $n!$ éléments. En effet, les éléments de $\Lambda_{\mu, \nu}$ sont du type $f(x_i) = y_{\varphi(i)}$, où φ est une bijection sur $\llbracket 1, N \rrbracket$ ($\varphi \in S_n$).

Correction du III.9.7 (page 102)

a) On a bien $\mu_0(\emptyset) = 0$ et pour toute collection disjointe de A_i mesurables, on a

$$\mu_0(\cup A_i) = \text{Card}((\cup A_i) \cap \mathbb{Z}) = \sum_i \text{Card}(A_i \cap \mathbb{Z}),$$

il s'agit donc bien d'une mesure, infinie car $\mathbb{Z} \subset \mathbb{R}$ est infini, mais σ -finie car \mathbb{R} s'écrit comme réunion des intervalles $[-n, n]$.

Les ensembles de mesure pleines pour μ_0 sont simplement les ensembles qui contiennent \mathbb{Z} .

Deux fonctions sont égales μ_0 -presque partout si et seulement si elles prennent les mêmes valeurs sur les entiers, ce qui laisse évidemment de la marge pour les choisir très différentes hors des entiers. Par exemple la fonction $\sin(\pi x)$ est nulle μ_0 -presque partout.

On peut construire une version finie de cette mesure en introduisant des poids, par exemple

$$\tilde{\mu}_0 = \sum_{k \in \mathbb{Z}} \frac{1}{2^{|k|}} \delta_k,$$

qui donne à \mathbb{R} (ou à n'importe quelle partie qui contient les entiers), une masse égale à 3.

b) la mesure μ_n est du même type que μ_0 , elle en a les mêmes propriétés. On n'a pas absolue continuité relativement à λ car les singletons $\{k/2^n\}$ sont de masse nulle pour λ , non nulle pour μ_n . Inversement les intervalles $]k/2^n, (k+1)/2^n[$ ont une masse non nulle pour λ , nulle pour μ_n .

Pour tout intervalle $]a, b[$ le nombre d'entiers de la forme $k/2^n$ qu'il contient est équivalent à $2^n(b-a)$, ce qui assure la convergence demandée.

c) Considérons par exemple l'ensemble A des nombres de l'intervalle $]0, 1[$ qui ne sont pas de la forme $k/2^n$. C'est ensemble est de mesure pleine sur $]0, 1[$, car on a enlevé un ensemble dénombrable, donc de mesure nulle. On a ainsi $\lambda(A) = 1$, et pourtant $\mu_n(A) = 0$ pour tout n .

Correction du III.9.8 (page 103)

a) On a

$$A \cup B = (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B) \text{ (union disjointe),}$$

d'où

$$\begin{aligned} \mu(A \cup B) &= \mu(A) - \mu(A \cap B) + \mu(B) - \mu(A \cap B) + \mu(A \cap B) \\ &= \mu(A) + \mu(B) - \mu(A \cap B). \end{aligned}$$

b) On écrit

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

On développe $\mu(A \cup B)$ d'après ce qui précède, et l'on écrit

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

c) On peut montrer par récurrence la formule générale

$$\begin{aligned} \mu\left(\bigcup_{n=1}^N A_n\right) &= \sum_{1 \leq n \leq N} \mu(A_n) - \sum_{1 \leq i_1 < i_2 \leq N} \mu(A_{i_1} \cap A_{i_2}) + \sum_{1 \leq i_1 < i_2 < i_3 \leq N} \mu(A_{i_1} \cap A_{i_2} \cap A_{i_3}) \\ &\quad + (-1)^N \mu(A_1 \cap A_2 \cap \cdots \cap A_N). \end{aligned}$$

Correction du III.9.9 (page 103)

a) Il s'agit d'une mesure de comptage, dont on vérifie immédiatement que c'est bien une

mesure. Tout ouvert non vide contient un intervalle ouvert, donc une infinité de rationnels, sa mesure est donc infinie. Et la mesure de l'ouvert \emptyset est 0.

On note A_n l'ensemble des rationnels qui s'écrivent $\pm a/b$, avec a et b des entiers naturels entre 0 et n . on a $\mu(A_n \cap [-n, n]) < +\infty$, et l'union des $A_n \cap [-n, n]$ recouvre \mathbb{R} .

b) On a $\mu(]0, 1[\cap \mathbb{Q}) = +\infty$ et $\lambda(]0, 1[\cap \mathbb{Q}) = 0$ ce qui invalide $\mu \ll \lambda$, et dans l'autre sens $\mu(]0, 1[\cap \mathbb{Q}^c) = 0$ et $\lambda(]0, 1[\cap \mathbb{Q}^c) = 1$, ce qui invalide $\lambda \ll \mu$.

Correction du III.9.10 (page 103)

a) L'ensemble A est la collection des intervalles $[k \times 10^{-n+1}, (k+1/2) \times 10^{-n+1}[$. L'intersection d'un intervalle de longueur 10^{-n+1} avec A est donc de longueur totale $10^{-n+1}/2$, d'où le résultat pour les intervalles dont la longueur est un multiple entier de cette quantité.

Soit maintenant $I =]a, b[$ un intervalle. Pour tout N , on a

$$b - a = k_N \times 10^{-N} + \varepsilon_N$$

avec $(E(x))$ est la partie entière de x

$$k_N = E((b - a) \times 10^N), \quad \varepsilon_N = b - a - k_N \times 10^{-N} \in [0, 10^{-N}[.$$

l'intervalle s'écrit donc comme la réunion d'un intervalle dont la longueur est multiple de 10^{-N} , avec un intervalle de longueur $\varepsilon_N < 10^{-N}$. On a donc

$$\lambda(]a, b[\cap A_{N+1}) \geq \frac{b - a - \varepsilon_N}{2} \geq \frac{b - a}{2} - \frac{1}{2} 10^{-N},$$

et

$$\lambda(]a, b[\cap A_{N+1}) \leq \frac{b - a - \varepsilon_N}{2} + \varepsilon_N \leq \frac{b - a}{2} + 10^{-N},$$

d'où la convergence de $\lambda(]a, b[\cap A_{N+1})$ vers $(b - a)/2$.

b) On se restreint à l'intervalle $]0, 1[$, en considérant l'ensemble $A' = A \cap]0, 1[$. D'après les hypothèses on a $\lambda(A') = 1/2$. L'ensemble A' étant mesurable, sa mesure s'identifie à sa mesure extérieure

$$\lambda(A') = \inf_{C_{A'}} \left(\sum_i (b_i - a_i) \right),$$

avec

$$C_{A'} = \left\{ (]a_i, b_i[)_{i \in \mathbb{N}}, A' \subset \bigcup_{\mathbb{N}}]a_i, b_i[\right\}.$$

Il existe donc une collection d'intervalles ouverts dont l'union contient A' telle que

$$\sum_i (b_i - a_i) \leq 3/4.$$

On note U la réunion ci-dessus. Comme $A' \subset U$, on a

$$\frac{1}{2} = \lambda(A') = \lambda(A' \cap U) \leq \sum_i \lambda(A' \cap]a_i, b_i[) = \frac{1}{2} \sum_n (b_i - a_i) = 3/8,$$

ce qui est absurde.

Correction du III.9.12 (page 104)

La fonction f étant dérivable, elle est continue, et donc mesurable (proposition III.5.4, page 79). Pour tout $n \geq 1$, la fonction g_n définie par

$$g_n(x) = \frac{f(x + 1/n) - f(x)}{1/n}$$

est elle-même continue, donc mesurable. La fonction f' s'exprime

$$f'(x) = \lim_{n \rightarrow +\infty} g_n(x) = \limsup_{n \rightarrow +\infty} g_n(x).$$

L'application f' est donc mesurable d'après la proposition III.5.2, page 77.

Correction du III.9.11 (page 104)

a) L'ensemble K est une intersection de fermés (comme unions finies d'intervalles fermés), il s'agit donc d'un fermé, qui est borné par construction, donc d'un compact. Si K contient un intervalle ouvert, cet intervalle est dans chacune des K_n , réunion d'intervalles de longueur $1/3^n$, sa longueur est donc inférieure à tout $1/3^n$, donc nécessairement nulle.

b) À toute suite $a = (a_n)_{n \geq 1}$ dans $\{0, 1\}$, on peut associer

$$x_a = \sum_{n=0}^{+\infty} 2 a_n \frac{1}{3^n}.$$

Le réel x_a appartient à K . En effet, la série partielle définit x_a^n qui est l'extrémité gauche de l'un des intervalles de K_n . La suite x_a^n est donc dans K , et elle est de Cauchy par construction, elle converge donc dans \mathbb{R} , donc dans K car K est fermé. Cette application $\{0, 1\}^{\mathbb{N}^*} \rightarrow K$ est injective. Or l'ensemble de départ s'identifie à l'ensemble des parties de \mathbb{N}^* , qui est non dénombrable.

c) On a $\lambda(K_n) = 2^n/3^n$, et $K \subset K_n$ pour tout n , d'où $\lambda(K) \leq 2^n/3^n \rightarrow 0$.

Correction du III.9.13 (page 104)

a) Le fait que $f(\cdot, t)$ soit mesurable, et la condition (ii) assurent que l'intégrale est bien définie. Il s'agit maintenant de montrer la continuité. Soit t_n une suite de réels qui tend vers t , on pose $f_n(x) = f(x, t_n)$. La continuité de f par rapport à t assure la convergence simple de $f(\cdot, t_n)$ vers $f(\cdot, t)$. La condition (iii) permet d'appliquer le théorème de convergence dominée, qui assure la convergence des intégrales :

$$\int_X f(x, t_n) d\mu(x) \rightarrow \int_X f(x, t) d\mu(x).$$

b) Soit $t \in I$ et $\varepsilon > 0$ tel que $t \pm \varepsilon \in I$. On définit (la fonction h introduite ci-dessous est définie pour ce t particulier)

$$h : (x, s) \in X \times]-\varepsilon, \varepsilon[\mapsto h(x, s) = \begin{cases} \frac{f(x, t+s) - f(x, t)}{s} & \text{si } s \neq 0 \\ \frac{\partial f}{\partial t}(x, t) & \text{si } s = 0 \end{cases}$$

La fonction $h(\cdot, s)$ est mesurable pour tout $s \neq 0$, comme $h(x, 0)$ est la limite de $h(x, s)$ quand s tend vers 0, la fonction $h(\cdot, 0)$ est également mesurable. La fonction $s \mapsto h(x, s)$

est continue pour tout x d'après l'hypothèse de différentiabilité de f . On a par ailleurs $|h(x, s)| \leq g(x)$ d'après le théorème des accroissements finis. On peut donc appliquer la première question, qui assure

$$\frac{F(t+s) - F(t)}{s} = \int_X h(x, s) d\mu(x) \longrightarrow \int_X h(x, 0) d\mu(x) = \int_X \frac{\partial f}{\partial t}(x, t) d\mu(x).$$

qui exprime la dérivabilité de F .

Correction du III.9.14 (page 105)

a) On a

$$\mu_1(A_1) = \gamma(\pi_1^{-1}(A_1)) = \gamma(A_1 \times X_2),$$

et de même $\mu_2(A_2) = \gamma(X_1 \times A_2)$. On a en particulier $\mu_1(X_1) = \gamma(X_1 \times X_2) = \mu_2(X_2)$.

On n'a pas en général $\gamma = \mu_1 \otimes \mu_2$. Si l'on considère par exemple $X_1 = X_2 = \{0, 1\}$, et

$$\gamma(\{0, 0\}) = \gamma(\{1, 1\}) = 1/2, \quad \gamma(\{1, 0\}) = \gamma(\{0, 1\}) = 1/2.$$

On a $\mu_1(\{0\}) = \mu_1(\{1\}) = \mu_2(\{0\}) = \mu_2(\{1\}) = 1/2$, de telle sorte que $\gamma \neq \mu_1 \otimes \mu_2$.

Si γ est la loi d'une variable aléatoire $(Y_1, Y_2) \in X_1 \times X_2$, μ_i est simplement la loi de Y_i . La situation $\gamma = \mu_1 \otimes \mu_2$ correspond au cas de variables aléatoires indépendantes.

b) Dans le cas où $X_1 = X_2 = \llbracket 1, N \rrbracket$, les singletons sont des rectangles, et la mesure γ est entièrement déterminée par sa valeurs sur les singletons. La mesure s'identifie donc à une matrice $\gamma = (\gamma_{ij})$, avec $\gamma_{ij} = \gamma(\{(i, j)\})$. La première marginale μ_1 correspond à la somme des éléments des lignes :

$$(\mu_1)_i = \sum_{j=1}^N \gamma_{ij},$$

et μ_2 la somme des éléments des colonnes successives.

c) On peut toujours prendre pour γ la mesure produit $\mu_1 \otimes \mu_2$, qui correspondrait à des variables aléatoires indépendantes dans le cas de mesures de probabilité).

d) On peut identifier les plans de transport à des matrices $\gamma = (\gamma_{ij})$, et on a

$$\Pi_{\mu_1, \mu_2} = \left\{ \gamma = (\gamma_{ij}) \in \mathbb{R}_+^{n \times m}, \sum_{i=1}^n \gamma_{ij} = \mu_2^j, \sum_{j=1}^m \gamma_{ij} = \mu_1^i \right\}.$$

Il s'agit donc d'un ensemble de matrices dont la somme des éléments de chaque ligne, et de chaque colonne, est fixée. C'est un convexe (intersection d'un sous-espace vectoriel avec le convexe $\mathbb{R}_+^{n \times m}$). C'est un singleton dès que l'une des mesures (arrivée ou départ) est concentrée en un point unique. Dès qu'on a deux points chargés de part et d'autre, il existe plusieurs plans de transport.

e) L'application considérée est linéaire donc continue, elle est donc minorée sur le compact Π_{μ_1, μ_2} , et atteint sa borne inférieure en au moins un point. L'application n'est pas strictement convexe, il n'y a pas de raison que ce point soit unique. De fait il ne l'est pas en général, c'est clair dans des cas dégénérés (par exemple si tous les coûts sont égaux à une même valeur),

mais aussi par exemple dans le cas de points (au moins 2 points à l'arrivée et au départ) appartenant à une même droite, avec $c_{ij} = \|y_j - x_i\|$. On pourra considérer par exemple (on se place sur la droite réelle) $x_1 = 0$, $x_2 = 1$, $y_1 = 1$, $y_2 = 2$, avec des masses égales. On peut laisser la masse en x_2 sur place (en y_1), et envoyer la masse en x_1 en y_2 , on translate de 1 la mesure. Noter que si le coût est quadratique en la distance, alors on a une solution unique (qui correspond à la translation).

f) Dans ce cas de figure, le problème consistant à trouver une solution du problème de minimisation revient à trouver un protocole d'envoi de la masse portée par les points de production vers les points de consommation de façon à minimiser le coût total de transport. Par exemple si c_{ij} correspond à la longueur du parcours routier entre x_i et y_j , cela revient à minimiser la longueur totale du parcours, et donc en première approximation l'essence consommée⁴.

g) On peut simplement définir l'opposé de la valeur ajoutée

$$A = - \sum_{ij} u_{ij} \gamma_{ij}.$$

Minimiser A revient à maximiser la valeur ajoutée, et le problème obtenu est du type des précédents, avec des coûts c_{ij} égaux aux opposés des productivités $i - j$.

Correction du III.8.1 (page 99)

Correction du III.8.2 (page 100)

Correction du III.8.3 (page 100)

On considère la suite $f_n(x) = \mathbb{1}_{]0,n[}(x)$, qui est dans la boule unité fermée de L^∞ . Deux termes différents de cette suites sont toujours à distance 1, il est donc exclu que l'on puisse en extraire une sous-suite convergente. La non compacité ne vient pas de la non compacité de \mathbb{R} , on peut construire une suite analogue de fonction à support compact, par exemple $f_n(x) = \mathbb{1}_{]0,1-1/n[}(x)$.

Correction du III.8.4 (page 100)

a) Soit $f \in L^p(I)$. On a, pour tout x dans I tel que $|f(x)| \geq 1$,

$$|f(x)| \leq |f(x)| |f(x)|^{p-1} = |f(x)|^p,$$

d'où l'on déduit que

$$|f(x)| \leq \max(1, |f(x)|^p),$$

et donc $|f(x)|$ est intégrable.

b) Si l'intervalle n'est pas borné, par exemple si $I =]1, +\infty[$, la fonction $f(x) = 1/x$ est dans L^p pour tout $p > 1$, mais pas dans $L^1(I)$.

4. Le problème réel est en général plus complexe, dans la mesure où le coût s'écrit, pour de petites masses à transporter, comme un coût fixe, dû au fait que l'on fait partir un camion. L'approche proposée correspond au cas où l'on a des grandes masses de produit à transporter, et donc un nombre de camions important à mobiliser (que l'on peut alors considérer comme une variable réelle).

Correction du III.8.5 (page 100)

a) On considère la suite de fonctions définie par

$$g_n(x) = |f(x) - T_n \circ f(x)|^p.$$

On a

$$|g_n(x)| \leq |f(x)|^p$$

qui est intégrable. Et $g_n(x)$ converge vers 0 pour presque tout x . On a donc convergence (dominée) de l'intégrale de g vers 0, d'où la convergence en norme L^p de $T_n \circ f$ vers f .

b) On peut considérer de la même manière

$$h_n(x) = |f(x) - \chi_n(x)f(x)|^p.$$

Le même raisonnement assure la convergence vers f de $\chi_n f$ vers f .

c) Le même raisonnement s'applique à $\chi_n T_n \circ f$.

d) Pour $p = +\infty$, $T_n \circ f$ est presque partout égal à f pour n assez grand, on a donc bien convergence en norme L^∞ . Pour $\chi_n f$ en revanche on n'a pas convergence. Pour $f \equiv 1$ par exemple, on a $\|f - \chi_n f\|$ identiquement égal à 1.