Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-2004**

Annales Corrigées de l'Année 2003-2004

Licence de Mathématiques : Intégration, Analyse de Fourier et Probabilités

Ce polycopié regroupe les devoirs à la maison, D.S. et examens donnés en I.F.P. au cours de l'année universitaire 2003–2004. Tous les énoncés sont accompagnés de solutions entièrement rédigées. Il va de soi que ces corrigés ne pourront être utiles qu'aux lecteurs ayant déjà cherché à résoudre par eux-mêmes les questions posées.

Je remercie tous mes collègues de l'équipe enseignante d'I.F.P. 2003–04, pour leur contribution à ce travail et pour m'avoir donné leur accord pour l'inclusion des énoncés et corrigés correspondants dans ces annales. Le D.M. n° 2 a été pris en charge par Sandra Delaunay et François Recher, le D.M. n° 3 par Raymond Moché et Gijs Tuynman.

L'ensemble du document est disponible sur Internet à l'URL

`http://math.univ-lille1.fr/~suquet/`

Les remarques, critiques et questions des lecteurs seront les bienvenues.

Villeneuve d'Ascq, le 8 septembre 2004

Charles Suquet

Index thématique

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Devoir n° 1

À rendre dans la semaine du 13 octobre 2003

Ex 1. *Probabilité de gagner un jeu sur son service au tennis*
On considère le modèle simplifié suivant du jeu de tennis[1] : le joueur A est au service et affronte B. Un « jeu » est constitué d'un certain nombre d' « échanges » et à la fin de chaque « échange », celui des deux joueurs qui a gagné l'échange marque un point. Le joueur qui gagne le « jeu » est le premier à totaliser au moins 4 points *avec* au moins deux points d'avance sur son adversaire. A peut donc gagner le jeu sur le score de 4 à 0, 4 à 1, 4 à 2, 5 à 3, 6 à 4,... On suppose que tous les échanges sont indépendants[2] et que lors de chaque échange la probabilité pour A de gagner le point reste constante et vaut p. Notre objectif est de calculer en fonction de p la probabilité que A remporte le jeu. On introduit les notations d'événements suivantes.

$$\begin{aligned} G &= \{A \text{ gagne le jeu}\}, \\ A_n &= \{\ A \text{ gagne le } n\text{-ième échange }\}, \\ B_n &= \{\ B \text{ gagne le } n\text{-ième échange }\}, \\ E_{i,j} &= \{\text{Au bout de } i+j \text{ échanges, } A \text{ a } i \text{ points et } B \text{ en a } j\}. \end{aligned}$$

1) Expliquez la décomposition

$$G = E_{4,0} \cup E_{4,1} \cup E_{4,2} \cup \Big(\bigcup_{k=3}^{+\infty} E_{k+2,k} \Big).$$

2) Calculer $P(E_{4,0})$, $P(E_{4,1})$ et $P(E_{4,2})$. *Indication :* on remarquera que $E_{4,1} = A_5 \cap E_{3,1}$ et $E_{4,2} = A_6 \cap E_{3,2}$ et on utilisera l'indépendance des échanges.

3) Expliquer pourquoi pour tout $k \geq 3$, on a :

$$E_{k+2,k} = E_{3,3} \cap \Big(\bigcap_{j=4}^{k} C_j \Big) \cap A_{2k+1} \cap A_{2k+2},$$

où l'on a posé $C_j := (A_{2j-1} \cap B_{2j}) \cup (B_{2j-1} \cap A_{2j})$.

4) Calculer $P(E_{3,3})$, puis $r = P(C_j)$. En déduire $P(E_{k+2,k})$.

1. Il n'est pas nécessaire de connaître les règles classiques du tennis pour faire cet exercice. Elles sont rappelées dans l'énoncé sous une forme permettant de simplifier les écritures.

2. Cette hypothèse est faite pour des raisons de simplification, on pourra en discuter la pertinence après avoir résolu l'exercice.

5) Montrer que les $E_{k+2,k}$ sont deux à deux disjoints et calculer $P\Big(\bigcup\limits_{k=3}^{+\infty} E_{k+2,k}\Big)$.

6) En déduire $P(G)$.

7) L'événement $N = \{\text{le jeu continue indéfiniment sans vainqueur}\}$ s'écrit :

$$N = E_{3,3} \cap \Big(\bigcap_{j=4}^{+\infty} C_j\Big).$$

Montrer que sa probabilité est nulle.

8) Compte tenu du résultat de la question précédente, pouvez vous donner sans calcul la valeur de $P(G)$ lorsque $p = 1/2$? Utilisez cette réponse pour tester la formule trouvée à la question 6.

Ex 2. *Le lemme de Fatou pour les séries*

1) Soient $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites de réels [3]. Montrer que

$$\liminf_{n\to+\infty} a_n + \liminf_{n\to+\infty} b_n \leq \liminf_{n\to+\infty}(a_n + b_n)$$

et que cette inégalité est vérifiée aussi lorsque $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites d'éléments de $\overline{\mathbb{R}}_+$.

2) En déduire le lemme de Fatou pour les séries : si $(u_{n,k})_{n\in\mathbb{N},k\in\mathbb{N}}$ est une suite double d'éléments de $\overline{\mathbb{R}}_+$,

$$\sum_{k=0}^{+\infty} \liminf_{n\to+\infty} u_{n,k} \leq \liminf_{n\to+\infty} \sum_{k=0}^{+\infty} u_{n,k}.$$

3) Donner un exemple où l'inégalité est stricte.

3. *Erratum :* lire « réels positifs », voir corrigé.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Devoir n^o 2 (nouvelle version)
À rendre dans la semaine du 10 novembre 2003

Ex 1. Soit $(\Omega, \mathcal{F}, \mu)$ un ensemble mesuré. Soit $(E_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{F}$ telle que

$$\sum_{n\in\mathbb{N}} \mu(E_n) < +\infty.$$

Notons :

$$A = \{\omega \in \Omega,\ \omega \text{ appartient à une infinité de } E_n\}.$$

1) Exprimer A en fonction des E_n.

2) Montrer que $\mu(A) = 0$.

Ex 2. On note λ la mesure de Lebesgue sur $\mathbb{R}$. Soit A l'ensemble des réels x pour lesquels il existe une infinité de couples d'entiers $(p, q) \in \mathbb{N} \times \mathbb{N}^*$ tels que

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{q^3}$$

Le but de cet exercice est de montrer que $\lambda(A) = 0$.

1) Vérifier que A contient l'ensemble $\mathbb{Q}_+$ des rationnels positifs. A est donc une partie infinie de $\mathbb{R}$.

2) Vérifier que A est un borélien de $\mathbb{R}$ en traduisant sa définition à l'aide d'opérations ensemblistes dénombrables faisant intervenir les intervalles

$$I_{p,q} := \left[\frac{p}{q} - \frac{1}{q^3}, \frac{p}{q} + \frac{1}{q^3}, \right]$$

où $p \in \mathbb{N}$ et $q \in \mathbb{N}^*$.

3) Montrer que $\lambda(A) = 0$ si et seulement si $\lambda(B) = 0$ avec $B = [0, 1[\cap A$.

4) Pour $q \in \mathbb{N}^*$, on définit :

$$B_q := [0, 1[\cap \left(\bigcup_{p\in\mathbb{N}} I_{p,q} \right).$$

Montrer que $x \in [0, 1[$ est dans B si et seulement si x appartient à une infinité d'ensembles B_q.

5) Montrer que

$$\lambda(B_q) \le \frac{2(q+2)}{q^3}.$$

6) En utilisant l'exercice précédent, montrer que $\lambda(B) = 0$.

7) On se propose de montrer que A n'est pas dénombrable. Pour cela, on fixe une suite d'entiers $(j_k)_{k\ge1}$ telle que pour tout $k \ge 1$, $j_{k+1} > 3j_k$ (par exemple la suite définie par $j_k = 4^k$ a cette propriété). On note C l'ensemble des réels x pouvant s'écrire sous la forme

$$x = \sum_{k=1}^{+\infty} \frac{a_k}{2^{j_k}}, \quad (a_k)_{k\ge1} \in \{0,1\}^{\mathbb{N}^*}.$$

Montrer que C est inclus dans A et qu'il est en bijection avec $\{0,1\}^{\mathbb{N}^*}$. Conclure.

Ex 3. Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré tel que $0 < \mu(\Omega) < +\infty$. On dit qu'un événement A de mesure non nulle est un *atome* de $(\Omega, \mathcal{F}, \mu)$ si

$$\forall B \in \mathcal{F}, \quad B \subseteq A \Longrightarrow \mu(B) = 0 \quad \text{ou} \quad \mu(B) = \mu(A).$$

On pourra utiliser le résultat suivant, conséquence directe de la formule de Poincaré : si $A_1, \dots, A_n$ désignent n (≥ 2) événements dont les intersections deux à deux sont μ-négligeables, alors

$$\mu(A_1 \cup \dots \cup A_n) = \mu(A_1) + \dots + \mu(A_n).$$

1) Démontrer que l'on définit une relation d'équivalence $\sim$ dans $\mathcal{F}$ en posant :

$$\forall A, B \in \mathcal{F}, \quad A \sim B \iff \mu(A \Delta B) = 0.$$

Démontrer qu'alors $\mu(A) = \mu(B)$.

2)

2.a) Démontrer que si A est un atome de $(\Omega, \mathcal{F}, \mu)$ et si N est un événement μ-négligeable, $A \Delta N$ est un atome de $(\Omega, \mathcal{F}, \mu)$ équivalent à A.

2.b) Démontrer que si une classe d'équivalence contient un atome, elle ne contient que des atomes qui ont tous la même mesure.

2.c) Démontrer que si A et B sont deux atomes non-équivalents, $\mu(A \cap B) = 0$.

3) On suppose que $(\Omega, \mathcal{F}, \mu)$ possède au moins un atome. On choisit un représentant dans chaque classe d'équivalence formée d'atomes. On note $(A_i,\ i \in I)$ la famille d'atomes ainsi obtenue. Démontrer qu'elle est finie ou infinie dénombrable.

4) On munit $\mathbb{R}$ de la tribu

$$\mathcal{F} = \big\{ B\,;\ B \in \mathrm{Bor}(\mathbb{R})\,,\ B \supseteq [0,1] \quad \text{ou} \quad B^c \supseteq [0,1] \big\},$$

et l'on considère la fonction P sur $\mathcal{F}$ qui ne prend que les valeurs 0 et 1 et vérifie

$$\forall F \in \mathcal{F}, \quad P(F) = 1 \iff F \supseteq [0,1].$$

Démontrer que la fonction P est une probabilité sur $(\mathbb{R}, \mathcal{F})$ et que $[0,1]$ est un atome de $(\mathbb{R}, \mathcal{F}, P)$.

Ex 4. *Version initiale de l'exercice sur les atomes*

Soit $(X, \mathcal{B})$ un espace mesurable. On dit qu'un élément A de $\mathcal{B}$ est un *atome* si A est non vide et si on a :

$$\forall B \in \mathcal{B} \quad B \subset A \Longrightarrow (B = \emptyset \text{ ou } B = A)$$

1) Montrer que la collection des atomes est une famille de parties deux à deux disjointes.

On dit que l'espace mesurable $(X, \mathcal{B})$ est *atomique* si tout élément de $\mathcal{B}$ est réunion d'atomes de $\mathcal{B}$.

2) Montrer que $(X, \mathcal{B})$ est atomique si et seulement si la collection des atomes est une partition de X.

3) Montrer que $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ munie de la mesure de Lebesgue est atomique.

4) On suppose que $\mathcal{B}$ admet un atome A. On définit μ_A sur $\mathcal{B}$ par $\mu_A(B) = 1$ si $A \subset B$ et $\mu_A(B) = 0$ sinon. Montrer que μ_A est une mesure.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Devoir Surveillé, 28 novembre 2003

Conditions de déroulement de l'épreuve :
Durée : 4 heures.
Les calculatrices sont interdites.
Liste exhaustive des documents autorisés :

1. Polycopié de DEUG *Introduction au calcul des probabilités.*
2. Polycopié du *cours* d'IFP 2003-2004 (Chapitres 1 à 4).
3. Une feuille *manuscrite* recto verso au format standard A4, pouvant contenir des extraits du cours à votre choix, à l'exclusion de tout corrigé d'exercice.

Le barème n'est donné qu'à titre indicatif pour vous aider à gérer votre temps et n'a pas valeur contractuelle.

Ex 1. *Continuité presque partout (3 points).*
Soit μ une mesure sur $\big(\mathbb{R}, \mathrm{Bor}(\mathbb{R})\big)$ vérifiant les deux conditions

$$\mu(\{0\}) = 0, \tag{1}$$

$$\forall a, b \in \mathbb{R}, \quad a < b \;\Rightarrow\; \mu(]a, b[) > 0. \tag{2}$$

On note f la fonction définie sur $\mathbb{R}$ par

$$f(x) := \begin{cases} \sin(1/x) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

1) Montrer que f est continue μ-presque partout.

2) Prouver qu'il n'existe aucune fonction g continue *partout* sur $\mathbb{R}$ et égale μ-presque partout à f. *Indication :* on pourra raisonner par l'absurde en montrant que s'il existe une telle g, on peut construire une suite décroissante et convergente vers 0 de réels x_k tels que pour tout k, $f(x_k) = g(x_k)$ avec $f(x_k) > 1/2$ pour k pair et $f(x_k) < -1/2$ pour k impair.

3) *Question facultative hors barême.* On suppose que $h : \mathbb{R} \to \mathbb{R}$ est borélienne et *uniformément* continue sur un borélien E de $\mathbb{R}$ tel que $\mu(\mathbb{R} \setminus E) = 0$. Montrer qu'il existe une fonction g partout continue sur $\mathbb{R}$ telle que $h(x) = g(x)$ pour tout $x \in E$.

Ex 2. *Théorème de Poincaré (7 points).*
Le but de cet exercice est de démontrer le théorème suivant.

Théorème 1. *Soit* $(\Omega, \mathcal{F}, \mathbf{P})$ *un espace probabilisé et* $f : \Omega \to \Omega$*, mesurable* $\mathcal{F}$*-*$\mathcal{F}$*, qui conserve la probabilité* $\mathbf{P}$ *:* $\forall C \in \mathcal{F}$, $\mathbf{P}(f^{-1}(C)) = \mathbf{P}(C)$.

Alors pour tout A *de* $\mathcal{F}$*, il existe* $A' \in \mathcal{F}$ *et inclus dans* A *tel que :*

a) $\mathbf{P}(A') = \mathbf{P}(A)$

b) $\forall \omega \in A'$, $\forall n \in \mathbb{N}^*$, $\exists p \geq n$*, tel que* $f^p(\omega) \in A$*, où* $f^p = \underbrace{f \circ f \circ \cdots \circ f}_{p \text{ fois}}$.

Si on appelle *trajectoire* de ω la suite de ses images $(f^p(\omega))_{p\geq 1}$, on peut exprimer ceci en disant que pour presque tout ω de A, la trajectoire de ω repasse par A une infinité de fois.

1) *Un exemple.* Montrer que les hypothèses du théorème sont vérifiées dans le cas particulier suivant. On prend $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathrm{Bor}([0,1]), \lambda)$, où λ est la restriction de la mesure de Lebesgue sur $\mathbb{R}$ à la tribu borélienne de $[0,1]$. On définit $f : [0,1] \to [0,1]$ par

$$f(x) = 2x\mathbf{1}_{[0,1/2[}(x) + (2x-1)\mathbf{1}_{[1/2,1]}(x).$$

Indication. On admettra que la tribu $\mathrm{Bor}([0,1])$ est engendrée par la famille des intervalles $[a,b]$, $0 \leq a < b \leq 1$ et on commencera par chercher $f^{-1}([a,b])$.

2) On introduit les ensembles

$$A_0 := \{\omega \in A;\ \forall n \in \mathbb{N}^*,\ f^n(\omega) \notin A\}, \quad A_1 := \{\omega \in A;\ \exists n \in \mathbb{N}^*,\ f^n(\omega) \in A\}.$$

Autrement dit, A_0 est l'ensemble des éléments de A dont la trajectoire ne repasse jamais par A, tandis que A_1 est l'ensemble de ceux dont la trajectoire repasse au moins une fois par A. On a clairement $A_1 = A \setminus A_0$. Le but de cette question est d'établir l'égalité $\mathbf{P}(A_1) = \mathbf{P}(A)$. On pose pour tout $n \in \mathbb{N}^*$, $B_n := (f^n)^{-1}(A_0)$.

2. a) Vérifier que A_0 et A_1 et tous les B_n sont dans $\mathcal{F}$.

2. b) Montrer que les B_n sont deux à deux disjoints.

2. c) En déduire que $\mathbf{P}(A_0) = 0$ et conclure.

3) On définit par récurrence $A_k := \{\omega \in A_{k-1};\ \exists n \geq 1,\ f^n(\omega) \in A_{k-1}\}$ pour tout $k \geq 2$. Ainsi A_k est l'ensemble des éléments ω de A dont la trajectoire repasse au moins k fois par A. Montrer que $\mathbf{P}(A_k) = \mathbf{P}(A)$ et achever la démonstration du théorème 1.

Ex 3. *Interlude*[4] *(2 points)*

Soit $(T, \mathcal{T}, \mu)$ un espace mesuré et f une application $T \to \overline{\mathbb{R}}_+$, mesurable $\mathcal{T}$ - $\mathrm{Bor}(\overline{\mathbb{R}}_+)$. On note ν la mesure de densité f par rapport à μ, définie par :

$$\forall B \in \mathcal{T}, \quad \nu(B) = \int_B f \,\mathrm{d}\mu.$$

Montrer que si $\mu(B) = 0$, alors $\nu(B) = 0$. *Indication :* on pourra traiter d'abord le cas où f est bornée, puis généraliser (ce n'est pas la seule méthode possible).

4. Cet exercice est indépendant du problème qui suit, mais son résultat y sera utile.

Problème. *(10 points)*

Dans tout le problème, $(\Omega, \mathcal{F}, \mathbf{P})$ est un espace probabilisé, $p \in]0,1[$ et $(X_k)_{k\geq 1}$ est une suite de variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbf{P})$ et de même loi de Bernoulli de paramètre p. On ne fait pas d'hypothèse générale sur la structure de dépendance de cette suite. Le but du problème est l'étude de la loi de la variable aléatoire Y définie sur $(\Omega, \mathcal{F}, \mathbf{P})$ par :

$$Y := \sum_{k=1}^{+\infty} \frac{X_k}{4^k}. \tag{3}$$

On rappelle que la notation $\{0,1\}^{\mathbb{N}^*}$ désigne l'ensemble de toutes les suites $(a_k)_{k\geq 1}$ dont les termes a_k ne peuvent prendre que les valeurs 0 ou 1. On note λ la mesure de Lebesgue sur $\mathbb{R}$. Les questions sont largement indépendantes, même si certaines utilisent les résultats des questions précédentes que l'on pourra admettre en cas de besoin. La question 8) peut même se traiter en ignorant tout ce qui la précède.

1) Expliquer pourquoi Y est bien définie par (3) et est une variable aléatoire réelle.

2) Calculer $\mathbf{E}Y$ en justifiant votre réponse.

3) Calculer pour $j \in \mathbb{N}^*$ la somme de série $\displaystyle\sum_{k>j} \frac{1}{4^k}$.

4) On désigne par C l'ensemble des réels $x \in [0,1]$ tels que

$$\exists (a_k)_{k\geq 1} \in \{0,1\}^{\mathbb{N}^*}, \quad x = \sum_{k=1}^{+\infty} \frac{a_k}{4^k}. \tag{4}$$

4. a) Montrer que pour $x \in C$, la décomposition (4) est unique. On pourra procéder comme suit. Si $\sum_{k=1}^{+\infty} a_k 4^{-k} = \sum_{k=1}^{+\infty} b_k 4^{-k}$ pour deux suites $(a_k)_{k\geq 1}$ et $(b_k)_{k\geq 1}$ appartenant à $\{0,1\}^{\mathbb{N}^*}$, on pose $j := \sup\{n \in \mathbb{N}^*;\ \forall k \leq n,\ a_k = b_k\}$ et on montre que si j est fini on aboutit à une contradiction.

4. b) En déduire que l'application

$$\varphi : \{0,1\}^{\mathbb{N}^*} \to C \quad (a_k)_{k\geq 1} \longmapsto \sum_{k=1}^{+\infty} \frac{a_k}{4^k}$$

est une bijection. C est-il dénombrable ?

5) On se propose de montrer que $\lambda(C) = 0$. Pour cela on fixe un choix de n premiers chiffres binaires $c_1, \dots, c_n \in \{0,1\}$ et on note $C(c_1, \dots, c_n)$ le sous-ensemble de C formé des x tels que dans le développement (4), $a_i = c_i$ pour $1 \leq i \leq n$. On admet que C et les $C(c_1, \dots, c_n)$ sont des boréliens de $\mathbb{R}$.

5. a) Montrer que $\lambda\big(C(c_1, \dots, c_n)\big) \leq \dfrac{1}{3 \times 4^n}$ et utiliser cette inégalité pour obtenir un majorant de $\lambda(C)$, dépendant de n.

5. b) En déduire que $\lambda(C) = 0$.

6) On note P_Y la loi de Y. Que vaut $P_Y(C)$? La mesure P_Y peut-elle avoir une densité par rapport à λ ?

7) On suppose dans cette question que les variables aléatoires de Bernoulli X_k sont indépendantes (et toujours de même paramètre $p \in]0,1[$). L'étude systématique de l'indépendance n'ayant pas encore été vue en cours, on notera que cette indépendance se traduit par la validité des égalités

$$\mathbf{P}(X_1 = c_1, \dots, X_n = c_n) = \mathbf{P}(X_1 = c_1) \dots \mathbf{P}(X_n = c_n),$$

pour tout $n \geq 2$ et tout choix de chiffres binaires $(c_1, \dots, c_n) \in \{0,1\}^n$. Montrer que dans ce cas $P_Y(\{x\}) = \mathbf{P}(Y = x) = 0$ pour tout $x \in C$. Que peut-on en déduire pour la fonction de répartition F de Y ?

8) Pour les deux cas particuliers suivants où les X_k sont dépendantes, déterminer la loi de Y et représenter graphiquement F.

8. a) $X_k = X_1$ pour tout $k \geq 2$;

8. b) X_1 et X_2 sont indépendantes et pour $k \geq 3$, $X_k = X_1$ si k est impair, $X_k = X_2$ si k est pair.

9) Est-il possible de choisir les X_k, toujours de même loi Bern(p) de sorte que pour tout $x \in C$, $\mathbf{P}(Y = x)$ soit strictement positif ?

10) *Question facultative hors barême. Cette question étant un exercice d'analyse ne sera prise en compte par les correcteurs que pour les étudiants ayant obtenu au moins 10 points sur tout ce qui précède (exercices et problème).* Montrer que C est un fermé de $\mathbb{R}$. On commencera par montrer que si x et x' sont dans C et tels que $|x - x'| < 4^{-j}$, leurs $j-1$ premiers chiffres binaires coïncident. Utiliser ceci pour prouver que si $(x_n)_{n\geq 1}$ est une suite d'éléments de C, convergente dans $\mathbb{R}$, sa limite x appartient encore à C.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Devoir n^o 3
À rendre dans la semaine du 8 décembre 2003

Ex 1.

Démontrer que pour tout entier $n \geq 1$, l'intégrale de Riemann impropre

$$u_n = \int_0^1 \frac{(\log x)^n}{\sqrt{1-x^2}}\, dx$$

est absolument convergente, puis étudier le comportement asymptotique de la suite $(u_n,\ n \geq 1)$.

Ex 2.

Soit $f : [0, +\infty[\longrightarrow [0, +\infty[$ une fonction décroissante λ-intégrable sur $[0, +\infty[$.

1 - Démontrer que

$$x \cdot f(x) \xrightarrow[x \to +\infty]{} 0.$$

Indication. On pourra comparer les intégrales

$$\int_0^{+\infty} f \cdot \mathbf{1}_{[\frac{x}{2}, x]}\, d\lambda = \int_{\frac{x}{2}}^{x} f\, d\lambda \quad \text{et} \quad \int_0^{+\infty} f\, d\lambda.$$

2 - Peut-on trouver $\alpha > 1$ tel que $x^\alpha \cdot f(x) \xrightarrow[x \to +\infty]{} 0$?

Problème

L'intérêt de ce problème est d'étudier des sommes qui pourraient avoir le même rôle d'encadrement et d'approximation que les sommes de Darboux de l'intégrale de Riemann d'une fonction continue sur un intervalle compact dans le cas de l'intégrale de Riemann impropre d'une fonction continue sur $\mathbb{R}^+ = [0, +\infty[$.

Soit f une fonction continue de $\mathbb{R}^+$ dans $\mathbb{R}^+$. Nous lui associons deux suites $(\varphi_n,\ n \geq 1)$ et $(\psi_n,\ n \geq 1)$ de fonctions étagées ainsi définies sur $\mathbb{R}^+$:

$$\forall n \geq 1, \quad \varphi_n = \sum_{k=0}^{n^2-1} a_{n,k} \cdot \mathbf{1}_{[\frac{k}{n}, \frac{k+1}{n}[} \quad \text{et} \quad \psi_n = \sum_{k=0}^{n^2-1} A_{n,k} \cdot \mathbf{1}_{[\frac{k}{n}, \frac{k+1}{n}[} \tag{5}$$

où, pour tous entiers $n \geq 1$ et k, $0 \leq k \leq n^2 - 1$,

$$a_{n,k} = \min\left(f(x)\,,\ \frac{k}{n} \leq x \leq \frac{k+1}{n}\right) \quad \text{et} \quad A_{n,k} = \max\left(f(x)\,,\ \frac{k}{n} \leq x \leq \frac{k+1}{n}\right). \tag{6}$$

Commentaire. Ces suites de fonctions étagées n'ont aucune raison d'être monotones.

1) Démontrer que

$$\forall n \geq 1, \quad \varphi_n \leq f \cdot \mathbf{1}_{[0,n[} \leq \psi_n \tag{7}$$

et que

$$\lim_{n\to+\infty} \varphi_n = \lim_{n\to+\infty} \psi_n = f. \tag{8}$$

Commentaire. Les suites $\left(\int_{\mathbb{R}^+} \varphi_n\, d\lambda\,,\ n \geq 1\right)$ *et* $\left(\int_{\mathbb{R}^+} \psi_n\, d\lambda\,,\ n \geq 1\right)$ *pourraient jouer, pour l'intégrale de Riemann impropre* $\int_{\mathbb{R}^+} f\, d\lambda$*, voir*[5]*, un rôle analogue aux sommes de Darboux. D'après (7),*

$$\int_0^{+\infty} \varphi_n\, d\lambda \leq \int_0^{+\infty} f\, d\lambda \quad \text{et} \quad \int_0^{n} f\, d\lambda \leq \int_0^{+\infty} \psi_n\, d\lambda,$$

mais il n'y a aucune raison que l'on ait $\int_0^{+\infty} f\, d\lambda \leq \int_0^{+\infty} \psi_n\, d\lambda$*, si bien que l'on peut déjà dire que les propriétés d'encadrement des sommes de Darboux ne sont pas vérifiées.*

Nous passons maintenant aux propriétés d'approximation de ces suites d'intégrales.

2) Démontrer que si f est λ-intégrable sur $\mathbb{R}^+$,

$$\int_0^{+\infty} \varphi_n\, d\lambda \xrightarrow[n\to+\infty]{} \int_0^{+\infty} f\, d\lambda. \tag{9}$$

3) Démontrer que la propriété (9) est également vraie dans le cas où f n'est pas λ-intégrable sur $\mathbb{R}^+$.

4) Démontrer que dans le cas où f n'est pas λ-intégrable sur $\mathbb{R}^+$,

$$\int_0^{+\infty} \psi_n\, d\lambda \xrightarrow[n\to+\infty]{} \int_0^{+\infty} f\, d\lambda. \tag{10}$$

5. Il s'agit de $\int_0^{+\infty} f(x)\, dx$, mais on sait que pour une fonction continue ≥ 0, $\int_{\mathbb{R}^+} f\, d\lambda$ et $\int_0^{+\infty} f(x)\, dx$ ont la même signification, à condition d'admettre, ce que nous faisons, qu'une intégrale de Riemann impropre puisse valoir $+\infty$. Dans ce problème (et ceci est vrai aussi pour l'exercice 2, sachant que toute fonction monotone sur un intervalle compact est Riemann-intégrable, Lebesgue-intégrable, les intégrales étant égales) toutes les intégrales, qui sont écrites comme des intégrales de Lebesgue, peuvent être considérées comme des intégrales de Riemann.

5) Nous allons maintenant montrer que la propriété d'approximation (10) peut tomber en défaut dans le cas où f est λ-intégrable sur $\mathbb{R}^+$ en exhibant un contre-exemple.

5.a) On appelle fonction triangulaire de base $[a, b]$ et de hauteur h $(0 \leq a < b,\ h > 0)$ la fonction déffinie sur $\mathbb{R}^+$ par :

$$g(x) = 2h \cdot \frac{x-a}{b-a} \quad \text{si} \quad a \leq x \leq \frac{a+b}{2}\,; \ = 2h \cdot \frac{b-x}{b-a} \quad \text{si} \quad \frac{a+b}{2} \leq x \leq b\,; \ = 0 \quad \text{sinon.}$$

Combien vaut l'intégrale $\displaystyle\int_a^b g\, d\lambda$?

5.b) Pour tout entier $n \geq 1$, on pose

$$s_n = \sum_{k=1}^{n} \frac{1}{k}, \quad m_n = \frac{s_n + s_{n+1}}{2}, \quad a_n = m_n - \frac{1}{4n^3}, \quad b_n = m_n + \frac{1}{4n^3},$$

et on suppose maintenant que pour tout entier $n \geq 1$, f est égale sur $[a_n, b_n]$ à la fonction triangulaire de base $[a_n, b_n]$ et de hauteur $4n$, et qu'elle est nulle ailleurs. Démontrer que

$$\int_0^{+\infty} f\, d\lambda = \sum_{n=1}^{+\infty} \frac{1}{n^2}.$$

5.c) Démontrer que

$$\int_0^{+\infty} \psi_n\, d\lambda \xrightarrow[n \to +\infty]{} +\infty,$$

ce qui montre que la propriété (10) n'est pas vérifiée.

Indication. Si l'intervalle $[\frac{k}{n}, \frac{k+1}{n}]$, $n \geq 1$, $0 \leq k \leq n^2 - 1$, *contient un point de la suite* $(m_n,\ n \geq 1)$, *noté* m_p, *la fonction* ψ_n *vaut au moins* $4p$ *sur l'intervalle* $[\frac{k}{n}, \frac{k+1}{n}[$, *donc*

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} \psi_n\, d\lambda \geq \frac{4p}{n}.$$

L'intégrale $\displaystyle\int_0^{+\infty} \psi_n\, d\lambda$ *est donc minorée par la somme de ces quantités. On démontrera que cette somme tend vers* $+\infty$.

6) La propriété (9) s'étend-elle au cas d'une fonction continue f de $\mathbb{R}^+$ dans $\mathbb{R}$, λ-intégrable sur $\mathbb{R}^+$?

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Devoir nº 4
À rendre dans la semaine du 5 janvier 2004

Ce devoir comporte trois exercices de calcul de lois et deux autres sur la loi des grands nombres. Seul l'exercice 5 nécessite la connaissance de la loi forte des grands nombres. Celle-ci devrait être vue en cours à la rentrée de janvier. Pour vous permettre de traiter cet exercice sans attendre la dernière minute, on donne ci-dessous l'énoncé de la loi forte des grands nombres.

Théorème 2. *On suppose que les variables aléatoires réelles X_k définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ sont indépendantes,* de même loi *et que $\mathbf{E}|X_1| < +\infty$. Alors*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbf{E}X_1. \tag{11}$$

Ex 1. Soient X et Y deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et indépendantes : X est une variable aléatoire discrète à valeurs dans $\mathbb{N}$ dont la loi est donnée par $\mathbf{P}(X = k) = p_k$, $k \in \mathbb{N}$, et Y est de loi uniforme sur $[0, 1]$. Déterminer la fonction de répartition F de la variable aléatoire XY. Que peut on dire de F ?

Ex 2. Soit (X, Y) un vecteur aléatoire dont la loi possède une densité de la forme $f(x, y) = g(x^2 + y^2)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$. Déterminer la loi de la variable aléatoire réelle T définie par

$$T(\omega) := \begin{cases} \frac{Y(\omega)}{X(\omega)} & \text{si } Y(\omega) \neq 0, \\ 0 & \text{sinon.} \end{cases}$$

Ex 3. Un vecteur aléatoire (X, Y) admet sur $\mathbb{R}^2$ une loi de densité

$$f(x, y) = x\mathrm{e}^{-(x+y)}\, \mathbf{1}_{\mathbb{R}_+}(x)\, \mathbf{1}_{\mathbb{R}_+}(y).$$

On pose $U = \min(X, Y)$, $V = \max(X, Y)$.

1) Calculer la densité de la loi du vecteur aléatoire (U, V).
2) En déduire la loi de U et celle de V.
3) Calculer $\mathrm{Cov}(U, V)$.

Ex 4. Soit $(X_n)_{n\geq 3}$ une suite de variables aléatoires discrètes définies sur le même $(\Omega, \mathcal{F}, \mathbf{P})$ et indépendantes. X_n prend les valeurs $-n$, n et 0 avec les probabilités :

$$\mathbf{P}(X_n = n) = \mathbf{P}(X_n = -n) = \frac{1}{2n\ln n} \quad \text{et} \quad \mathbf{P}(X_n = 0) = 1 - \frac{1}{n\ln n}.$$

1) Calculer la variance de X_n. En déduire que la suite $(M_n)_{n\geq 3}$ définie par :

$$M_n := \frac{1}{n}\sum_{i=3}^{n} X_i$$

converge vers 0 dans $L^2(\Omega)$ (et donc aussi en probabilité).

2) En appliquant le lemme de Borel Cantelli aux ensembles $A_n = \{X_n = n\}$ montrer que la suite M_n ne converge pas presque sûrement.
Indication. Si $M_n(\omega)$ converge vers une limite finie ℓ, $X_n(\omega)/n$ converge nécessairement vers 0 (pourquoi ?).

Ex 5.

1) Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. indépendantes de même loi uniforme sur $[0, 1]$. Étudier la limite

$$\lim_{n\to+\infty} \mathbf{E}\left(\frac{X_1^2 + X_2^2 + \cdots + X_n^2}{X_1 + X_2 + \cdots + X_n}\right).$$

On commencera par définir *proprement* la suite de variables aléatoires dont on prend l'espérance.

2) λ_n désignant la mesure de Lebesgue sur $\mathbb{R}^n$, montrer que pour toute fonction f continue sur $[0, 1]$

$$\int_{[0,1]^n} f\Big(\frac{x_1 + x_2 + \cdots + x_n}{n}\Big)\, \mathrm{d}\lambda_n \xrightarrow[n\to+\infty]{} f\Big(\frac{1}{2}\Big).$$

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Examen, 30 janvier 2004

Ce sujet comporte ***6 pages****, dont 3 figures*
Conditions de déroulement de l'épreuve :
Les calculatrices sont interdites.
Liste exhaustive des documents autorisés :

1. Polycopié de DEUG *Introduction au calcul des probabilités.*
2. Polycopié du *cours* d'IFP 2003-2004.
3. Une feuille *manuscrite* recto verso au format standard A4, pouvant contenir des extraits du cours à votre choix, à l'exclusion de tout corrigé d'exercice.

Le barème n'est donné qu'à titre indicatif pour vous aider à gérer votre temps et n'a pas valeur contractuelle.

Ex 1. *(2 points)*
Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles, *de même loi*, telle que $\mathbf{E}|X_1| < +\infty$. On pose $Y_n := \min(X_n, n)$. Montrer que $\mathbf{E}Y_n$ existe et que $\mathbf{E}Y_n$ tend vers $\mathbf{E}X_1$ quand n tend vers l'infini.

Ex 2. *Convergence d'espérances (5 points)*
On note X une variable aléatoire de loi exponentielle de paramètre $a > 0$, ce qui signifie que le loi de X a une densité $f(t) = a\mathrm{e}^{-at}\mathbf{1}_{\mathbb{R}^+}(t)$ par rapport à la mesure de Lebesgue λ. On définit la suite de variables aléatoires $(Y_n)_{n\geq 1}$ par

$$\forall n \in \mathbb{N}^*, \quad Y_n := \left(1 + \frac{X}{n}\right)^n \mathbf{1}_{\{X\geq 0\}}.$$

Ainsi Y_n est une variable aléatoire *positive* et l'expression $\mathbf{E}Y_n$ a toujours un sens comme élément de $\overline{\mathbb{R}}_+$. Le but de l'exercice est d'étudier sa limite quand n tend vers l'infini.

1) Exprimer $\mathbf{E}Y_n$ comme une intégrale sur $[0, +\infty[$ par rapport à λ.

2) Vérifier l'inégalité

$$\forall x \geq 0, \quad \left(1 + \frac{x}{n}\right)^n \leq \mathrm{e}^x.$$

On rappelle que $\ln(1 + u) \leq u$ pour tout $u \in]-1, +\infty[$.

3) On suppose que $a > 1$. En utilisant les deux questions précédentes, montrer que $\mathbf{E}Y_n$ tend vers une limite finie et la calculer en fonction de a.

4) Que peut-on dire de $\lim_{n\to+\infty} \mathbf{E}Y_n$ dans le cas $0 < a \leq 1$?

Ex 3. *Carrés aléatoires (3 points)*

Soit $(X_i)_{i\in\mathbf{N}^*}$ une suite de variables aléatoires, définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$, à valeurs dans $]0,1[$, indépendantes et de même loi uniforme sur $[0,1]$. On pose $S_n = \sum_{i=1}^n X_i$ et on note M_i le *point aléatoire* de coordonnées (S_i, S_i) dans un repère orthonormé du plan. On pose $M_0 = O$ origine du repère. Pour $i \in \mathbf{N}^*$, on note C_i le *carré aléatoire* ouvert de diagonale $M_{i-1}M_i$ (cf. figure 1).

1) On désigne par λ_2 la mesure de Lebesgue de $\mathbb{R}^2$. Soit $E_n = \bigcup_{i=1}^n C_i$. Montrer que quand n tend vers $+\infty$, $n^{-1}\lambda_2(E_n)$ converge presque sûrement vers une limite que l'on déterminera.

2) Soit $a \in]0,1[$ et Δ_a la droite d'équation $y = x + a$. Calculer la probabilité $P(\Delta_a \cap C_i \neq \emptyset)$ (commencer par exprimer l'évènement $\{\Delta_a \cap C_i \neq \emptyset\}$ à l'aide de la variable aléatoire X_i). Trouver la loi de T_n, nombre de C_i $(1 \leq i \leq n)$ ayant une intersection non vide avec Δ_a.

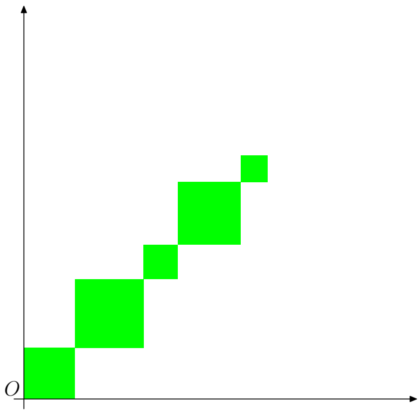

FIG. 1 – Carrés aléatoires

Problème (*13 points*)

Ce problème est consacré à l'algorithme du rejet. La première partie en étudie le principe général. La deuxième met en œuvre cet algorithme pour simuler la loi uniforme sur un « pétale de fleur », puis sur la fleur entière. La méthode du rejet (appelée aussi d'acceptation-rejet) peut être décrite abstraitement comme suit. On suppose que l'on sait générer un vecteur aléatoire M de $\mathbb{R}^d$ suivant une certaine loi μ. On génère alors l'un après l'autre les vecteurs aléatoires $M_1, \ldots, M_n, \ldots$ mutuellement indépendants et de même loi que M, en s'arrêtant au premier d'entre eux qui vérifie une certaine condition $(\mathcal{H}_0)$. Soit T l'indice (aléatoire) correspondant. On a ainsi fabriqué un vecteur (doublement) aléatoire M_T. Comme T est aléatoire, la loi de ce vecteur n'est pas celle de M, c'est une nouvelle loi ν. Si la simulation de M et le test de $(\mathcal{H}_0)$ sont facilement programmables, on dispose ainsi d'une méthode pour générer un vecteur aléatoire de loi ν.

Tous les vecteurs aléatoires et toutes les variables aléatoires considérés dans ce problème seront définis sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

Préliminaire

Soit B un borélien de $\mathbb{R}^d$ tel que $0 < \lambda_d(B) < +\infty$, λ_d désignant la mesure de Lebesgue sur $\mathbb{R}^d$. On dit que le vecteur aléatoire $M : \Omega \to \mathbb{R}^d$ suit la loi uniforme sur B, notation $M \sim \mathrm{Unif}(B)$, si sa loi P_M est donnée par

$$\forall A \in \mathrm{Bor}(\mathbb{R}^d),\ P_M(A) = \mathbf{P}(M \in A) = \frac{\lambda_d(A \cap B)}{\lambda_d(B)}. \tag{12}$$

On rappelle que cette condition équivaut à l'existence pour la mesure P_M d'une densité de la forme $c\mathbf{1}_B$ par rapport à λ_d, c étant une constante.

1) Montrer que si M suit la loi uniforme sur B et si φ est une *bijection linéaire* $\mathbb{R}^d \to \mathbb{R}^d$, $\varphi(M)$ suit la loi uniforme sur $\varphi(B)$.

Partie 1

On suppose désormais que $(M_i)_{i\in\mathbb{N}^*}$ est une suite de vecteurs aléatoires $\Omega \to \mathbb{R}^d$, mutuellement *indépendants* et de *même loi* μ telle que $\mu(B) > 0$. On assimilera pour la commodité du langage le vecteur $M_i(\omega)$ à un point de $\mathbb{R}^d$.

2) Pour tout $\omega \in \Omega$, on pose

$$T(\omega) := \inf\{i \in \mathbb{N}^*;\ M_i(\omega) \in B\},$$

avec la convention $\inf \emptyset = +\infty$. T est donc le numéro du premier point « tombé » dans B. Montrer en vous aidant des indications suivantes que T est une variable aléatoire discrète suivant la loi géométrique de paramètre $p = \mu(B)$. On notera $q := 1 - p$.
Indications. T est une application $\Omega \to \overline{\mathbb{N}}^*$. Sa mesurabilité relativement aux tribus $\mathcal{F}$-$\mathcal{P}(\overline{\mathbb{N}}^*)$ s'établit en vérifiant que pour tout $k \in \overline{\mathbb{N}}^*$, l'ensemble $\{T = k\} = \{\omega \in \Omega;\ T(\omega) = k\}$ appartient à $\mathcal{F}$. Décomposer $\{T = k\}$ à l'aide d'opérations ensemblistes sur des évènements de la forme $M_i^{-1}(A)$, de façon à exploiter la mesurabilité des M_i. Utiliser cette décomposition pour montrer que $\mathbf{P}(T = +\infty) = 0$ et pour calculer $\mathbf{P}(T = k)$ pour $k \in \mathbb{N}^*$.

3) Pour $k \in \mathbb{N}^*$ et $A \in \mathrm{Bor}(\mathbb{R}^d)$, calculer $\mathbf{P}(M_k \in A \cap B \text{ et } T = k)$ en fonction de q et de $\mu(A \cap B)$.

4) On définit $M_T : \Omega \to \mathbb{R}^d$ par

$$M_T(\omega) := \begin{cases} M_{T(\omega)}(\omega) & \text{si } T(\omega) < +\infty, \\ 0 & \text{si } T(\omega) = +\infty. \end{cases}$$

Cette définition équivaut à

$$M_T = \sum_{k\in\overline{\mathbb{N}}^*} M_k \mathbf{1}_{\{T=k\}},$$

ce qui montre que M_T est mesurable $\mathcal{F}$-$\mathrm{Bor}(\mathbb{R}^d)$, c'est donc bien un *vecteur aléatoire*. Prouver que $\mathbf{P}(M_T \in A) = \mathbf{P}(M_1 \in A \mid M_1 \in B)$. *Indication :* partitionner Ω par les évènements $\{T = k\}$.

5) On suppose désormais que B et C sont deux boréliens tels que $B \subset C$ et $0 < \lambda_d(B) < \lambda_d(C) < +\infty$ et que μ est la loi uniforme sur C. Quelle est la loi de M_T ?

6) On dispose seulement d'un générateur de nombres au hasard capable de simuler une suite aussi longue que l'on veut $U_1, \dots, U_\ell$ de variables aléatoires indépendantes de loi uniforme sur $[0,1]$. On s'intéresse au coût de l'algorithme du rejet pour la simulation d'un « échantillon » $M_{T_1}, \dots, M_{T_n}$ de vecteurs aléatoires *indépendants*[6] de même loi que M_T. Ce coût sera mesuré par le nombre S_n de variables U_i utilisées.

On suppose dans cette question que B est *borné*. On peut alors choisir C de la forme $[a_1, b_1] \times \dots \times [a_d, b_d]$.

6 a) Expliquer pourquoi le vecteur aléatoire $M = (X_1, \dots, X_d)$ défini par

$$X_j := a_j + (b_j - a_j)U_j, \quad j = 1, \dots, d,$$

suit la loi uniforme sur C.

6 b) Exprimer S_n en fonction des T_k et expliquer pourquoi la suite de variables aléatoires $n^{-1}S_n$ converge presque sûrement vers une limite constante que l'on exprimera en fonction du rapport $\tau := \lambda_d(C)/\lambda_d(B)$. Dans toute la suite du problème, le coût de l'algorithme du rejet sera la valeur de cette limite presque sûre de S_n/n.

6 c) Comment a-t-on intérêt à choisir C pour minimiser le coût de l'algorithme ? Illustrez votre réponse dans le cas $d = 2$ en prenant pour B le disque unité.

Partie 2

Dans cette partie on s'intéresse à la simulation de vecteurs aléatoires suivant la loi uniforme sur un borélien B de $\mathbb{R}^2$ dont la frontière a pour équation en coordonnées polaires $r = \cos(2t)$. On peut découper B en quatre « pétales » B_0, B_1, B_2, B_3, où

$$B_0 := \big\{(x,y) \in \mathbb{R}^2;\ 0 \leq r \leq \cos(2t),\ -\pi/4 \leq t \leq \pi/4\big\},$$

et les autres pétales se déduisent de B_0 par rotation de centre O et d'angles $\pi/2$, π, $3\pi/2$, cf. figure 2. On ne vous demande pas de justifier la construction de la courbe frontière de B.

7) Vérifier que $\lambda_2(B_0) = \pi/8$. En notant C_0 le rectangle circonscrit à B_0, cf. figure 3, on en déduit (justification non demandée) que

$$\tau = \frac{\lambda_2(C_0)}{\lambda_2(B_0)} = \frac{32}{3\pi\sqrt{6}} \simeq 1{,}386.$$

8) Soit (U,V) un vecteur aléatoire à composantes indépendantes de même loi uniforme sur $[0,1]$. On définit le vecteur aléatoire

$$(X,Y) := \big(U^{1/2}\cos(2\pi V), U^{1/2}\sin(2\pi V)\big).$$

Montrer que (X,Y) suit la loi uniforme sur le disque de centre O et de rayon 1.

6. On admettra qu'il suffit pour cela de lancer n fois l'algorithme du rejet et que les variables aléatoires $T_1, \dots, T_n$ sont indépendantes.

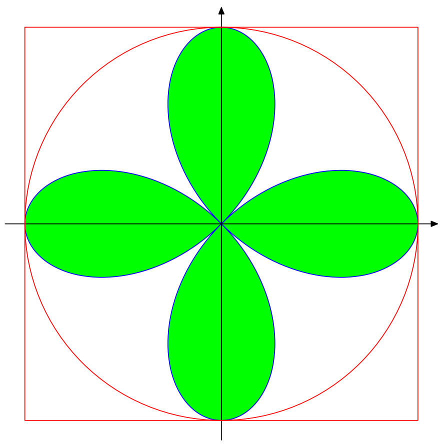

FIG. 2 – Ensemble B, cercle et carré circonscrits

9) Justifiez l'algorithme suivant pour simuler un vecteur aléatoire M'_T de loi uniforme sur B. On commence par générer une variable aléatoire U de loi uniforme sur $[0,1]$ et on définit la variable aléatoire discrète N par

$$N = \sum_{k=0}^{3} k\mathbf{1}_{\{k \leq 4U < k+1\}}.$$

On lance ensuite l'algorithme du rejet avec B_0 et C_0 (cf. question 7) qui nous fournit M_T de loi uniforme sur B_0. On prend alors pour M'_T l'image de M_T par la rotation de centre O et d'angle $N\pi/2$; on notera ρ_k la rotation de centre O et d'angle $k\pi/2$, $k = 0$, 1, 2, 3.

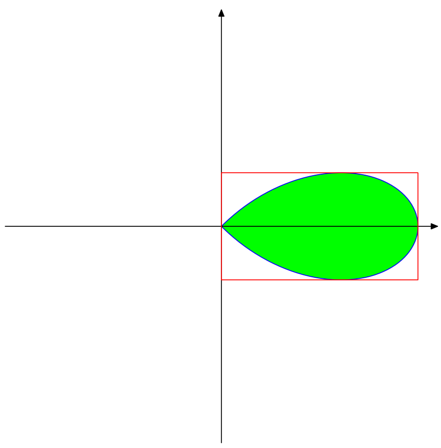

FIG. 3 – Ensemble B_0 et rectangle circonscrit

10) Comparez par leur coût (au sens de la limite p.s. de S_n/n, cf. question 6) les trois algorithmes suivants pour simuler un vecteur aléatoire de loi uniforme sur B :

a) Simulation de la loi uniforme sur le carré $C = [-1,1]^2$ circonscrit à B et algorithme du rejet avec B et C.

b) Simulation de la loi uniforme sur le disque D de centre O et de rayon 1 (pourquoi contient-il B ?) par la méthode de la question 8 et algorithme du rejet avec B et D.

c) La méthode mixte de la question 9.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Examen deuxième session, septembre 2004

Ce sujet comporte ***4 pages****, dont 2 figures.*
Conditions de déroulement de l'épreuve :
Les calculatrices sont interdites.
Liste exhaustive des documents autorisés :

1. Polycopié de DEUG *Introduction au calcul des probabilités.*
2. Polycopié du *cours* d'IFP 2003-2004.
3. Une feuille *manuscrite* recto verso au format standard A4, pouvant contenir des extraits du cours à votre choix, à l'exclusion de tout corrigé d'exercice.

Le barème n'est donné qu'à titre indicatif pour vous aider à gérer votre temps et n'a pas valeur contractuelle.

Ex 1. *Un calcul de série (3 points)*
En fournissant les justifications adéquates, calculer :

$$S := \sum_{k=0}^{+\infty} \int_0^{\pi/2} \left(1 - \sqrt{\sin x}\,\right)^k \cos x \, \mathrm{d}x.$$

Ex 2. *Étude d'une fonctionnelle définie par une espérance (9 points)*
On note $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et $X : \Omega \longrightarrow \mathbb{R}$, une variable aléatoire réelle sur cet espace. On pose

$$\forall t \in \mathbb{R}, \quad \Psi_X(t) := \mathbf{E} \arctan(tX), \tag{13}$$

où arctan désigne la réciproque de la restriction de la fonction tangente à $]-\pi/2, \pi/2[$.

1) Montrer que la fonction Ψ_X est définie sur $\mathbb{R}$ et bornée.

2) Montrer que si X et Y ont même loi, $\Psi_X = \Psi_Y$. Calculer explicitement Ψ_X dans les deux cas particuliers suivants :
a) X suit la loi de Bernoulli de paramètre p ;
b) X suit la loi uniforme sur $[0, 1]$.

3) Que vaut Ψ_X lorsque la loi de X est *symétrique*, *i.e.* X et $-X$ ont même loi ? Peut-on dire que Ψ caractérise la loi de X, c'est-à-dire que $\Psi_X = \Psi_Y$ (égalité de fonctions sur $\mathbb{R}$) implique $P_X = P_Y$ (X et Y ont même loi) ?

4) Montrer que Ψ_X est continue sur $\mathbb{R}$ (quelle que soit la loi de X).

5) Montrer que

$$\lim_{t\to+\infty} \Psi_X(t) = -\frac{\pi}{2}\mathbf{P}(X < 0) + \frac{\pi}{2}\mathbf{P}(X > 0). \tag{14}$$

Indiquer brièvement comment s'adapte votre justification pour fournir la limite en $-\infty$.

6) Démontrer les affirmations suivantes

a) Si $\mathbf{E}|X|$ est finie, Ψ_X est dérivable sur $\mathbb{R}$.

b) Ψ_X reste dérivable sur $\mathbb{R}^*$ même si X n'est pas intégrable.

c) Si X est positive et non intégrable, Ψ_X n'est pas dérivable en 0. *Indication :* on prouvera que si $(t_n)_{n\geq 1}$ est une suite dans $\mathbb{R}^*$, tendant vers 0, $\displaystyle\lim_{n\to+\infty} \frac{\Psi_X(t_n)}{t_n} = +\infty$.

Ex 3. *Simulation par rapport d'uniformes (6 points)*

Le but de cet exercice est de justifier mathématiquement une méthode de simulation d'une variable aléatoire Y de densité f connue. On suppose que l'on sait simuler un vecteur aléatoire (U, V) de loi uniforme sur le borélien

$$A := \left\{(u, v) \in \mathbb{R}^2;\ 0 < u < f\Big(\frac{v}{u}\Big)^{1/2}\right\}. \tag{15}$$

Alors la variable aléatoire $Y := V/U$ suit la loi de densité f. Cette méthode de simulation est appelée méthode du ratio d'uniformes [7].

On note λ_d la mesure de Lebesgue sur $\mathbb{R}^d$. On rappelle que le vecteur aléatoire (U, V) suit la loi uniforme sur A s'il a pour densité $c\mathbf{1}_A$ avec $c = 1/\lambda_2(A)$.

1) Montrer que $\lambda_2(A) = \frac{1}{2}$. *Indication :* on pourra utiliser le changement de variable

$$\varphi :]0, +\infty[\times\mathbb{R} \longrightarrow]0, +\infty[\times\mathbb{R}, \quad (u, v) \longmapsto (x, y) = \varphi(u, v) = \Big(u, \frac{v}{u}\Big).$$

2) Calculer la densité du vecteur aléatoire $(X, Y) := \Big(U, \dfrac{V}{U}\Big)$. En déduire que Y a pour densité f.

3) Cette méthode n'est applicable que si l'on sait effectivement simuler un vecteur de loi uniforme sur A. C'est le cas notamment lorsque A est *borné* [8]. En effet on peut alors inclure A dans un rectangle à côtés parallèles aux axes et appliquer l'algorithme du rejet pour simuler (U, V). Montrer qu'une condition suffisante pour que A soit borné est que la densité f vérifie les deux conditions

$$\sup_{x\in\mathbb{R}} f(x) < +\infty \quad \text{et} \quad \sup_{x\in\mathbb{R}} x^2 f(x) < +\infty. \tag{16}$$

7. Avec un gros abus de langage puisque c'est le vecteur aléatoire (U, V) qui suit une loi uniforme, pas ses composantes (sauf si A est un produit cartésien). Une appelation plus exacte mais de lourdeur dissuasive serait « rapport des composantes d'un vecteur de loi uniforme ».

8. Ceci ne résulte pas automatiquement de la finitude de $\lambda_2(A)$.

Ex 4. *Escargot aléatoire (4 points)*

Soit $(X_i)_{i\in\mathbf{N}^*}$ une suite de variables aléatoires, définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$, indépendantes et de même loi, de carré intégrable ($\mathbf{E}(X_1^2) < +\infty$). On définit par récurrence la suite de variables aléatoires positives $(R_n)_{n\geq 1}$ par

$$R_1 = |X_1|, \quad R_n = \sqrt{R_{n-1}^2 + X_n^2}, \quad n \geq 2.$$

1) Montrer que $(n^{-1/2}R_n)_{n\geq 1}$ converge presque sûrement vers une limite constante τ que l'on exprimera en fonction de $\mathbf{E}(X_1^2)$.

2) Pour illustrer graphiquement cette convergence, on prend les X_i de même loi uniforme sur $[0,1]$ et on construit dans un repère orthonormé du plan la suite $(M_n)_{n\geq 1}$ de points aléatoires par la récurrence suivante. On pose d'abord $M_1 = (X_1, 0)$, puis connaissant M_{n-1} on obtient M_n comme l'unique point tel que $M_{n-1}M_n = X_n$ et que l'angle $(\overrightarrow{M_{n-1}M_n}, \overrightarrow{M_{n-1}O})$ ait pour mesure $+\pi/2$. En d'autre termes, on construit M_n tel que le triangle $OM_{n-1}M_n$ soit rectangle en M_{n-1}, de côtés de l'angle droit $OM_{n-1} = R_{n-1}$ et $M_{n-1}M_n = X_n$ et en tournant toujours dans le sens trigonométrique. On trace ainsi la ligne polygonale $\mathcal{E}_n$ de sommets $M_1, M_2, \ldots, M_n$ (l'escargot aléatoire, cf. figure 4).

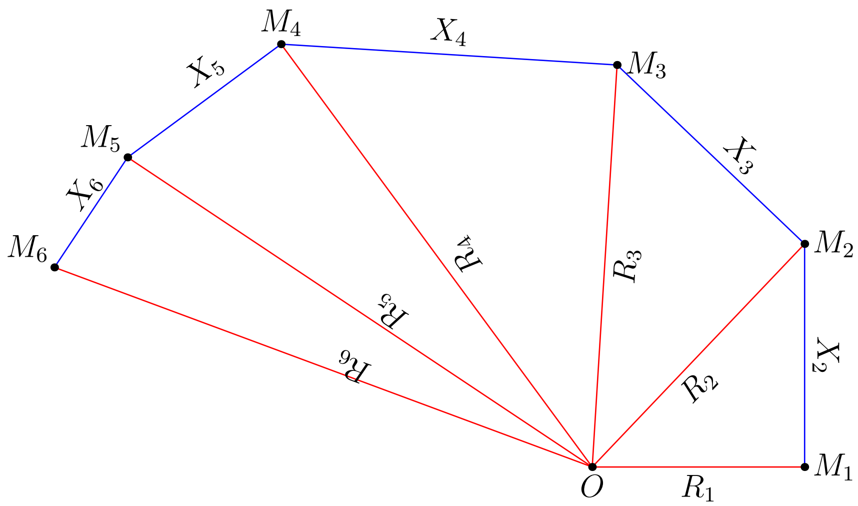

FIG. 4 – Construction de l'escargot aléatoire $\mathcal{E}_6$

On fixe un $\varepsilon > 0$. Montrer que presque sûrement, $\mathcal{E}_n$ est inclus dans le disque de centre O et de rayon $(1+\varepsilon)\sqrt{n/3}$ pour tout n assez grand (cf. figure 5).

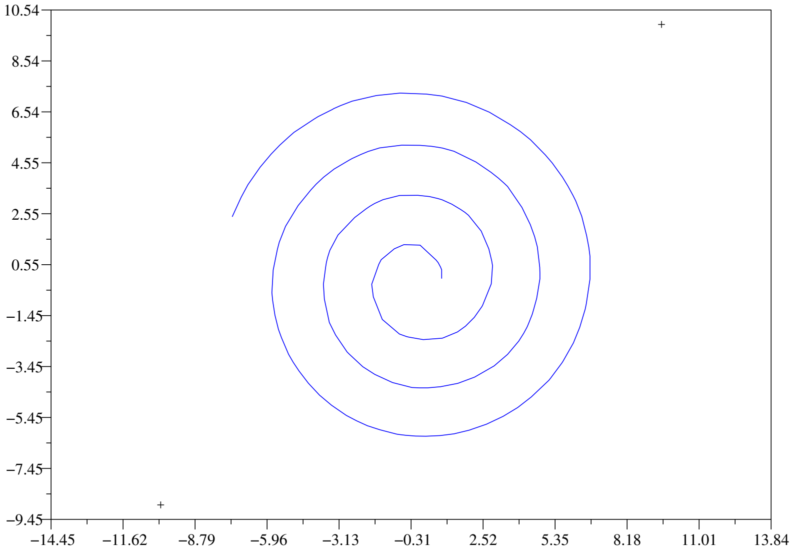

FIG. 5 – Escargot aléatoire $\mathcal{E}_{200}$ (simulation en Scilab)

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé du Devoir n° 1

Ex 1. *Probabilité de gagner un jeu sur son service au tennis*
On considère le modèle simplifié suivant du jeu de tennis : le joueur A est au service et affronte B. Un « jeu » est constitué d'un certain nombre d' « échanges » et à la fin de chaque « échange », celui des deux joueurs qui a gagné l'échange marque un point. Le joueur qui gagne le « jeu » est le premier à totaliser au moins 4 points *avec* au moins deux points d'avance sur son adversaire. On suppose que tous les échanges sont indépendants et que lors de chaque échange la probabilité pour A de gagner le point reste constante et vaut p. Notre objectif est de calculer en fonction de p la probabilité de $G := \{A \text{ gagne le jeu}\}$. On introduit les notations d'événements suivantes.

$$\begin{aligned} A_n &:= \{ A \text{ gagne le } n\text{-ième échange} \}, \quad B_n := \{ B \text{ gagne le } n\text{-ième échange} \}, \\ E_{i,j} &:= \{\text{Au bout de } i+j \text{ échanges, } A \text{ a } i \text{ points et } B \text{ en a } j\}. \end{aligned}$$

1) A peut gagner avec 4 points si le score est 4 à 0 ou 4 à 1 ou 4 à 2, autrement dit si l'un des évènements $E_{4,0}$, $E_{4,1}$ ou $E_{4,2}$ se réalise. On a donc

$$G' := \{ A \text{ gagne le jeu avec 4 points} \} = E_{4,0} \cup E_{4,1} \cup E_{4,2}.$$

L'autre possibilité est que A gagne avec plus de 4 points et cela se produit si et seulement si le jeu se termine sur un score 5 à 3 ou 6 à 4 ou 7 à 5 ou..., bref sur un score de la forme $k+2$ à k pour k décrivant l'ensemble de tous les nombres entiers à partir de 3. Ceci correspond à la réalisation de l'un des évènements de la suite infinie $(E_{k+2,k})_{k\geq 3}$. Ainsi

$$G'' := \{ A \text{ gagne le jeu avec plus de 4 points} \} = \bigcup_{k\geq 3} E_{k+2,k}.$$

Il est clair que $G = G' \cup G''$, d'où la décomposition

$$G = E_{4,0} \cup E_{4,1} \cup E_{4,2} \cup \Big(\bigcup_{k\geq 3} E_{k+2,k} \Big). \tag{17}$$

2) *Calcul de $P(E_{4,0})$, $P(E_{4,1})$ et $P(E_{4,2})$.*
Comme $E_{4,0} = A_1 \cap A_2 \cap A_3 \cap A_4$, l'indépendance des échanges nous donne :

$$P(E_{4,0}) = P(A_1)P(A_2)P(A_3)P(A_4) = p^4.$$

Le gain du jeu par A sur le score de 4 à 1 (évènement $E_{4,1}$) se réalise si et seulement si A gagne le cinquième échange (A_5) et B marque un seul point lors des quatre premiers échanges ($E_{3,1}$). Ainsi $E_{4,1} = A_5 \cap E_{3,1}$. En considérant les 5 premiers échanges comme indépendants [9] on a donc $P(E_{4,1}) = P(E_{3,1})P(A_5) = P(E_{3,1})p$. On remarque que $P(E_{3,1})$ est la probabilité que A ait exactement trois succès dans une suite de 4 épreuves répétées indépendantes, c'est donc $C_4^3 p^3(1-p) = 4p^3(1-p)$. On a donc

$$P(E_{4,1}) = P(E_{3,1})p = 4p^4(1-p).$$

Par un raisonnement analogue, on voit que $E_{4,2} = A_6 \cap E_{3,2}$ et

$$P(E_{4,2}) = P(E_{3,2})P(A_6) = C_5^3 p^3(1-p)^2 p = 10p^4(1-p)^2.$$

3) Pour tout $k \geq 3$, on a :

$$E_{k+2,k} = E_{3,3} \cap \Big(\bigcap_{4\leq j\leq k} C_j \Big) \cap A_{2k+1} \cap A_{2k+2}, \tag{18}$$

où l'on a posé $C_j := (A_{2j-1} \cap B_{2j}) \cup (B_{2j-1} \cap A_{2j})$. Pour justifier cette décomposition, on remarque que $E_{k+2,k}$ est aussi l'évènement « A *gagne* le jeu avec $k+2$ points contre k à B ». Le jeu dure donc dans ce cas $(2k+2)$ échanges, donc au moins 8 échanges puisque $k \geq 3$. Celà implique l'égalité des 2 joueurs au bout des 6 premiers échanges. En effet, compte tenu des règles sur le gain du jeu, les seuls scores possibles au bout de 6 échanges sont 4 à 2 (et alors A remporte le jeu) ou 3 à 3 (les échanges doivent continuer) ou 2 à 4 (et B remporte le jeu). Les scores du type 6 à 0 ou 5 à 1 sont impossibles car alors un des deux joueurs serait arrivé le premier à 4 points avec plus de deux points d'avance et le jeu se serait arrêté avant le sixième échange.

Regardons d'abord le cas particulier de $E_{5,3}$. Cet évènement se réalise si et seulement si les deux joueurs sont à égalité au bout des 6 premiers échanges *et* A remporte le septième et le huitième. Donc

$$P(E_{5,3}) = E_{3,3} \cap A_7 \cap A_8.$$

Ceci est un cas particulier de la formule (18) en convenant qu'une intersection indexée par « $4 \leq j \leq 3$ » est égale à Ω (ce qui est logique puisque qu'aucun j ne vérifie cette condition, on ne met donc aucune restriction d'appartenance à un C_j pour qu'un évènement élémentaire ω soit dans cette « intersection »).

Pour $k \geq 4$, la réalisation de $E_{k+2,k}$ signifie la victoire de A au bout de $2k+2$ échanges avec cette fois $2k+2 \geq 10$. Ceci se produit si et seulement si *chacune* des trois conditions suivantes se réalise (on aura donc l'intersection des trois évènements correspondants) :

- A et B sont à égalité au bout des 6 premiers échanges
- A et B se retrouvent à égalité à l'issue de chaque paire d'échanges depuis la paire $7^{\text{ième}}$, $8^{\text{ième}}$ jusqu'à la paire $(2k-1)^{\text{ième}}$, $(2k)^{\text{ième}}$.
- A gagne les deux derniers échanges : le $(2k+1)^{\text{ième}}$ et le $(2k+2)^{\text{ième}}$.

L'évènement « A et B marquent chacun un point lors de la paire d'échanges $(2j-1)^{\text{ième}}$ et $(2j)^{\text{ième}}$ » s'écrit $C_j = (A_{2j-1} \cap B_{2j}) \cup (B_{2j-1} \cap A_{2j})$ puisque ou bien A marque le premier et B égalise, ou c'est l'inverse. La justification de (18) est donc complète.

9. En toute rigueur, ce n'est pas vrai, mais du point de vue du calcul tout se passe comme s'ils l'étaient, voir la remarque 5 page 33.

4) *Calcul des $P(E_{k+2,k})$.*

L'évènement $E_{3,3}$ se réalise si et seulement si lors des 6 premiers échanges, A marque exactement 3 points (donc B aussi). En considérant ces 6 premiers échanges comme une suite d'épreuves répétées indépendantes, on a immédiatement

$$P(E_{3,3}) = C_6^3 p^3(1-p)^3 = 20p^3(1-p)^3.$$

Les évènements $A_{2j-1} \cap B_{2j}$ et $B_{2j-1} \cap A_{2j}$ sont incompatibles (par exemple parce que le premier implique que A remporte l'échange n° $(2j-1)$ tandis que le deuxième implique que B remporte ce même échange). On a donc

$$P(C_j) = P(A_{2j-1} \cap B_{2j}) + P(B_{2j-1} \cap A_{2j}) = p(1-p) + (1-p)p = 2p(1-p),$$

en utilisant l'indépendance des échanges pour la deuxième égalité. On remarque que la valeur trouvée ne dépend pas de j. On la notera r dans la suite pour alléger les écritures. Finalement, par indépendance des échanges on obtient :

$$\begin{aligned} P(E_{k+2,k}) &= P(E_{3,3}) \times \left(\prod_{j=4}^{k} P(C_j)\right) \times P(A_{2k+1})P(A_{2k+2}) \\ &= 20p^3(1-p)^3 r^{k-3} p^2 \\ &= 20p^5(1-p)^3 r^{k-3}. \end{aligned}$$

Remarquons que cette formule reste valable dans le cas particulier $k = 3$ où l'on a directement

$$P(E_{5,3}) = P(E_{3,3} \cap A_7 \cap A_8) = P(E_{3,3})P(A_7)P(A_8) = 20p^3(1-p)^3p^2.$$

5) Les $E_{k+2,k}$ sont deux à deux disjoints. En effet, soient $k \neq l$, la réalisation de $E_{k+2,k}$ implique que le jeu se termine au $(2k+2)^{\text{ième}}$ échange tandis que celle de $E_{l+2,l}$ implique qu'il se termine au $(2l+2)^{\text{ième}}$ échange. Comme $k \neq l$, on a aussi $2k+2 \neq 2l+2$, donc les évènements $E_{k+2,k}$ et $E_{l+2,l}$ sont incompatibles. On peut ainsi utiliser la σ-additivité :

$$\begin{aligned} P\Big(\bigcup_{k\geq 3} E_{k+2,k}\Big) &= \sum_{k=3}^{+\infty} P(E_{k+2,k}) \\ &= 20p^5(1-p)^3 \sum_{k=3}^{+\infty} r^{k-3} \\ &= 20p^5(1-p)^3 \sum_{j=0}^{+\infty} r^j \\ &= 20p^5(1-p)^3 \frac{1}{1-r}. \end{aligned}$$

La série géométrique de raison r qui intervient dans ce calcul est bien convergente puisque $r = 2p(1-p)$ est dans $[0, 1[$. En effet si $p < 1/2$ alors $2p < 1$ et $2p(1-p) < 1 - p \leq 1$, si $p > 1/2$ alors $1 - p < 1/2$ et on a la même majoration stricte pour r en échangeant les rôles de p et $1 - p$, enfin si $p = 1/2$, $r = 2 \times 1/2 \times 1/2 = 1/2 < 1$. En fait on peut vérifier que le maximum de $2p(1 - p)$ lorsque p parcourt $[0, 1]$ est atteint en $p = 1/2$.

6) Les évènements figurant dans la décomposition (17) correspondent chacun au gain du jeu sur un score différent et sont donc deux à deux incompatibles. On en déduit

$$\begin{aligned} P(G) &= P(E_{4,0}) + P(E_{4,1}) + P(E_{4,2}) + P\Big(\bigcup_{k\geq 3} E_{k+2,k}\Big) \\ &= p^4\Big(1 + 4(1-p) + 10(1-p)^2 + \frac{20p(1-p)^3}{1-2p(1-p)}\Big) =: f(p). \end{aligned} \tag{19}$$

On remarque que pour $p = 0$, $P(G) = 0$ et pour $p = 1$, $P(G) = 1$, ce qui est conforme à l'intuition.

7) L'événement $N = \{$le jeu continue indéfiniment sans vainqueur$\}$ s'écrit :

$$N = E_{3,3} \cap \Big(\bigcap_{j\geq 4} C_j\Big).$$

Il est donc inclus pour tout $n \geq 4$ dans l'évènement

$$N_n := E_{3,3} \cap \Big(\bigcap_{j=4}^{n} C_j\Big),$$

d'où

$$0 \leq P(N) \leq P(N_n) = 20p^3(1-p)^3 r^{n-3}.$$

Ceci étant vrai pour *tout* $n \geq 4$, ces inégalités *larges* se conservent par passage à la limite quand n tend vers l'infini. Comme $0 \leq r < 1$, la limite de $P(N_n)$ est nulle, on en déduit $P(N) = 0$. La probabilité que le jeu continue indéfiniment est nulle (bien que l'évènement N ne soit pas l'ensemble vide).

8) Notons H l'évènement « B gagne le jeu ». Les trois évènements G, H et N forment une partition de Ω, donc $P(G) + P(H) + P(N) = 1$. Comme nous venons de voir que $P(N) = 0$, cette égalité se réduit à $P(G) + P(H) = 1$. Dans le cas particulier où $p = 1/2$, A ne tire aucun avantage de son service, et on doit avoir par symétrie $P(G) = P(H)$, d'où $P(G) = 1/2$. Vérifions le sur la formule générale (19) pour $P(G)$. Elle s'écrit ici

$$P(G) = \frac{1}{16}\Big(1 + 2 + \frac{10}{4} + \frac{20/16}{1 - 1/2}\Big) = \frac{1}{16}\Big(3 + \frac{5}{2} + \frac{5}{2}\Big) = \frac{1}{2}.$$

Remarque 3. On aurait pu dans le cas général calculer $\mathbf{P}(H)$ par la même méthode qui nous a conduit à (19) pour $\mathbf{P}(G)$. Il suffit de remarquer que la probabilité que B gagne l'échange est $q = 1 - p$. Il est clair alors que $P(H) = f(q)$. Comme $P(N) = 0$, on doit avoir $f(q) + f(q) = 1$. Graphiquement ceci se traduit par le fait que la représentation graphique de f admet le point $(1/2, 1/2)$ comme centre de symétrie. La figure 6 (réalisée à l'aide du logiciel Scilab) semble le confirmer. Il faut bien avouer que cette symétrie ne

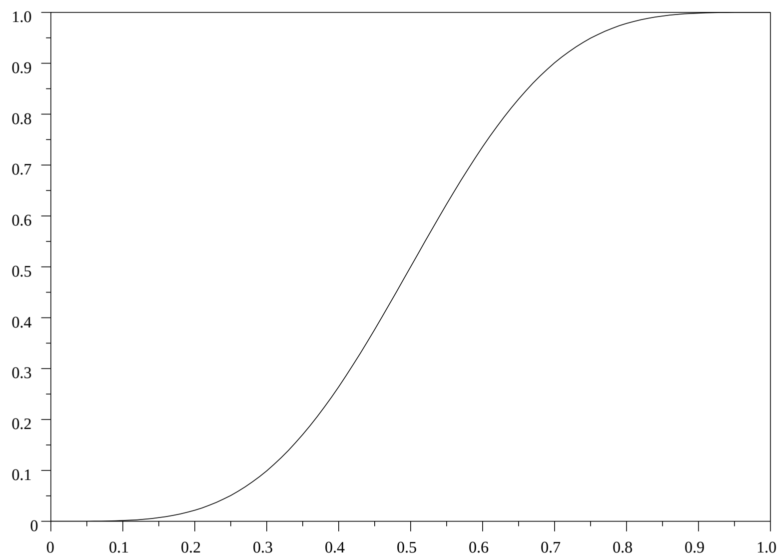

FIG. 6 – Représentation graphique de $P(G) = f(p)$

saute pas aux yeux à la lecture de (19). Pour le vérifier par le calcul, commençons par réécrire (19) sous la forme :

$$f(p) = p^4\Big(1 + 4q + 10q^2 + \frac{20pq^3}{1-2pq}\Big).$$

On a alors

$$f(p) + f(q) = p^4 + q^4 + 4pq(p^3 + q^3) + 10p^2q^2(p^2 + q^2) + \frac{20p^3q^3(p^2+q^2)}{1-2pq}.$$

De là on arrive facilement à $f(p) + f(q) = 1$ grâce aux formules suivantes obtenues en développant $(p+q)^n$ par la formule du binôme en notant que $p + q = 1$:

$$p^2 + q^2 = 1 - 2pq, \quad p^3 + q^3 = 1 - 3pq, \quad p^4 + q^4 = 1 - 4pq + 2p^2q^2.$$

Remarque 4. Un coup d'oeil sur la figure 6 et sur le tableau de valeurs numériques ci-dessous permet de voir l'effet amplificateur des déséquilibres dû aux règles du tennis. Dès que p s'éloigne de 1/2, $f(p)$ se rapproche très vite de 0 ou de 1.

p	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
$f(p)$	0,001	0,022	0,099	0,264	0,500	0,736	0,901	0,978	0,999

Remarque 5. L'hypothèse d'indépendance des échanges n'est en réalité valable que pour les 4 premiers échanges (puisqu'on est sûr qu'ils auront lieu). Pour les autres échanges, en raison de l'arrêt au bout d'un nombre aléatoire d'échanges, on ne peut pas avoir indépendance. Pour s'en convaincre, il suffit de remarquer que les évènements $E_{4,0}$ et A_5 ont des probabilités qui ne valent ni 0 ni 1 et sont *incompatibles* : la réalisation de $E_{4,0}$ impliquant le gain du jeu au bout de 4 échanges, A ne peut gagner le cinquième ! Ils ne sont donc pas indépendants. Pourtant la réalisation de $E_{4,0}$ ne dépend que du résultat des 4 premiers échanges et celle de A_5 que du résultat du cinquième échange. Si ces échanges étaient indépendants, ces deux évènements devraient l'être.

Il y a deux façons de remédier à ce problème. La première est de dire que toutes les probabilités calculées dans ce problème sont les mêmes que si les échanges continuaient indéfiniment après le gain du jeu par l'un des deux joueurs. On aurait ainsi une suite infinie d'épreuves répétées indépendantes. L'autre façon est de faire du conditionnement en chaîne en considérant que ce n'est pas $P(A_n)$ qui est constante égale à p comme le fait abusivement l'énoncé[10] mais les probabilités *conditionnelles* $P(A_n \mid H_{n-1})$ où H_{n-1} représente n'importe quelle suite de résultats détaillés des $n-1$ premiers échanges compatible avec la tenue du $n^{\text{ième}}$ échange. Par exemple pour $n = 5$, on peut mettre à la place de H_4 les évènements $A_1 \cap A_2 \cap B_3 \cap A_4$ ou $B_1 \cap B_2 \cap B_3 \cap A_4$.

10. En fait on doit même clairement avoir $P(A_5) < P(A_4)$ à cause de la possibilité de gain du jeu dès le quatrième échange…

Ex 2. *Le lemme de Fatou pour les séries*

Rappelons d'abord comment on définit la limite inférieure d'une suite $(a_n)_{n\in\mathbb{N}}$ dans $\mathbb{R}$ ou $\overline{\mathbb{R}}$. On pose pour tout $n \in \mathbb{N}$,

$$u_n := \inf\{a_k;\ k \geq n\} = \inf_{k\geq n} a_k.$$

Notons que u_n est un élément de $\overline{\mathbb{R}}$, même si tous les a_k sont finis[11]. En raison de l'inclusion $\{a_k;\ k \geq n+1\} \subset \{a_k;\ k \geq n\}$, on a $u_{n+1} \geq u_n$. Ceci étant vrai pour tout entier n, la suite (u_n) est croissante. Elle converge donc dans $\overline{\mathbb{R}}$ vers une limite u_∞ qui, en raison de la croissance de (u_n), est aussi le supremum de l'ensemble $\{u_n;\ n \in \mathbb{N}\}$. C'est cette limite u_∞ que l'on appelle *limite inférieure* de la suite (a_n). Ainsi

$$\liminf_{n\to+\infty} a_n := u_\infty = \lim_{n\to+\infty} u_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k.$$

1) L'énoncé demandait de montrer que si $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites de réels,

$$\liminf_{n\to+\infty} a_n + \liminf_{n\to+\infty} b_n \leq \liminf_{n\to+\infty}(a_n + b_n) \tag{20}$$

et que cette inégalité reste vérifiée lorsque $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites dans $\overline{\mathbb{R}}_+$.

Un oubli malencontreux rend cette affirmation fausse. Il fallait lire « $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites de réels positifs *». Le contre exemple élémentaire suivant montre que (20) ne peut être vraie pour toutes suites (a_n) et (b_n) de réels. Prenons pour tout $k \in \mathbb{N}$, $a_k := -k$ et $b_k := k$. Alors $\liminf_{n\to+\infty} a_n = -\infty$, $\liminf_{n\to+\infty} b_n = +\infty$ et $\liminf_{n\to+\infty}(a_n+b_n) = 0$. Dans ce cas le membre de gauche de (20) s'écrit « $-\infty+\infty$ », ce qui n'a pas de sens. Pour clarifier la question, nous allons montrer que (20) est vraie pour des suites de réels (pas forcément positifs) (a_n) et (b_n) avec l'hypothèse supplémentaire que les limites inférieures de (a_n) et (b_n) ne sont pas simultanément infinies de signes contraires[12]. Ceci englobe évidemment le cas où (a_n) et (b_n) sont des suites dans $\mathbb{R}_+$ puisqu'alors les limites inférieures sont dans $\overline{\mathbb{R}}_+$.*

Remarquons d'abord que $a_n + b_n$ est bien défini pour tout n (somme de deux réels). Posons pour tout $n \in \mathbb{N}$,

$$u_n := \inf_{k\geq n} a_k, \quad v_n := \inf_{k\geq n} b_k, \quad w_n := \inf_{k\geq n}(a_k + b_k).$$

Les trois suites ainsi définies sont croissantes, donc convergentes dans $\overline{\mathbb{R}}$ vers une limite notée respectivement u_∞, v_∞ et w_∞.

Pour $k \geq n$, on a $u_n \leq a_k$ et $v_n \leq b_k$. Comme a_k et b_k sont réels, ni u_n ni v_n ne peut valoir $+\infty$. S'ils sont tous les deux finis, on a $u_n + v_n \leq a_k + b_k$ (par compatibilité de la relation d'ordre avec l'addition dans $\mathbb{R}$). Si u_n ou v_n vaut $-\infty$, alors $u_n+v_n = -\infty$, d'où $u_n + v_n < a_k + b_k$ puisque $a_k + b_k$ est réel. Ainsi l'inégalité $u_n + v_n \leq a_k + b_k$ est vérifiée

11. On le voit bien sur l'exemple $a_k := -k$.

12. Si vous avez fait l'exercice 3 de la fiche 1 qui proposait de montrer l'inégalité inverse pour les limites supérieures, le bogue de l'énoncé n'a pas du vous échapper.

pour tout $k \geq n$, ce qui implique $u_n + v_n \leq \inf\{(a_k + b_k);\ k \geq n\}$. Le raisonnement fait étant valable pour tout entier n, nous venons d'établir que :

$$\forall n \in \mathbb{N}, \quad u_n + v_n \leq w_n. \tag{21}$$

Comme nous avons supposé que u_∞ et v_∞ ne sont pas simultanément infinis de signes contraires, la suite $(u_n + v_n)$ converge dans $\overline{\mathbb{R}}$ vers $u_\infty + v_\infty$. Le passage à la limite dans $\overline{\mathbb{R}}$ conservant les inégalités *larges*, on déduit alors de (21) que

$$u_\infty + v_\infty \leq w_\infty,$$

ce qui n'est qu'une écriture abrégée de (20).

Pour vérifier (20) dans le cas où (a_n) et (b_n) sont des suites dans $\overline{\mathbb{R}}_+$, on remarque d'abord que la somme de deux éléments de $\overline{\mathbb{R}}_+$ est toujours définie et que u_n, v_n et w_n sont des éléments de $\overline{\mathbb{R}}_+$. On a encore les inégalités $u_n \leq a_k$ et $v_n \leq b_k$ pour tout $k \geq n$. Cette fois u_n et v_n peuvent prendre la valeur $+\infty$, mais cela ne pose pas de problème pour en déduire que $u_n + v_n \leq a_k + b_k$ (puisque $u_n + v_n$ est bien défini comme élément de $\overline{\mathbb{R}}_+$ et par compatibilité de l'ordre avec l'addition dans $\overline{\mathbb{R}}_+$). Comme ci-dessus on en déduit (21) et on peut passer à la limite quand n tend vers $+\infty$: la suite croissante [13] $(u_n + v_n)$ converge toujours dans $\overline{\mathbb{R}}_+$ vers $u_\infty + v_\infty$, d'où le résultat.

2) L'inégalité (20) s'étend à la somme d'un nombre fini de suites dans $\overline{\mathbb{R}}_+$ (récurrence immédiate basée sur l'associativité de l'addition). Ainsi pour tout $j \in \mathbb{N}$,

$$\sum_{i=0}^{j} \liminf_{n\to+\infty} u_{n,i} \leq \liminf_{n\to+\infty} \sum_{i=0}^{j} u_{n,i}. \tag{22}$$

D'autre part les $u_{n,i}$ étant positifs, on a pour tout $j \in \mathbb{N}$ et tout $n \in \mathbb{N}$ l'inégalité suivante dans $\overline{\mathbb{R}}_+$:

$$\sum_{i=0}^{j} u_{n,i} \leq \sum_{i=0}^{+\infty} u_{n,i}.$$

Grâce à la conservation des inégalités larges par passage à la limite inférieure [14], ceci implique :

$$\forall j \in \mathbb{N}, \quad \liminf_{n\to+\infty} \sum_{i=0}^{j} u_{n,i} \leq \liminf_{n\to+\infty} \sum_{i=0}^{+\infty} u_{n,i}. \tag{23}$$

De (22) et (23) on déduit :

$$\forall j \in \mathbb{N}, \quad \sum_{i=0}^{j} \liminf_{n\to+\infty} u_{n,i} \leq \liminf_{n\to+\infty} \sum_{i=0}^{+\infty} u_{n,i}.$$

13. Dans le cas précédent, la suite $(u_n + v_n)$ est aussi une suite croissante de $\overline{\mathbb{R}}$. Elle ne converge vers $u_\infty + v_\infty$ que si cette quantité est bien définie. Regardons à nouveau l'exemple $a_k = -k$ et $b_k = k$. Ici $u_n = -\infty$, $v_n = n$, $u_n + v_n$ est bien défini et vaut $-\infty$ pour tout n, de sorte que la suite croissante $(u_n + v_n)$ est la suite constante égale à $-\infty$, donc convergente vers $u_\infty = -\infty$ et non pas vers $u_\infty + v_\infty$ qui n'est pas défini.

14. Propriété vue en T.D., fiche 1 exercice 3.

Dans cette inégalité, le deuxième membre est une constante (élément de $\overline{\mathbb{R}}_+$), tandis que le premier membre est le terme de rang j d'une suite croissante dans $\overline{\mathbb{R}}_+$, convergente vers $\sum_{i=0}^{+\infty} \liminf_{n\to+\infty} u_{n,i}$. L'inégalité étant valable pour tout entier j, on peut passer à la limite quand j tend vers $+\infty$ en conservant l'inégalité large et on obtient ainsi :

$$\sum_{i=0}^{+\infty} \liminf_{n\to+\infty} u_{n,i} \leq \liminf_{n\to+\infty} \sum_{i=0}^{+\infty} u_{n,i}, \tag{24}$$

ce qui constitue le lemme de Fatou pour les séries à termes dans $\overline{\mathbb{R}}_+$.

3) Voici deux exemples où l'inégalité (24) est stricte. Pour alléger nous noterons G le membre de gauche de (24) et D son membre de droite.
Exemple 1. On prend $u_{n,i} = 1$ si $i = n$ et $u_{n,i} = 0$ si $i \neq n$. Dans ce cas on a immédiatement $\liminf_{n\to+\infty} u_{n,i} = 0$ pour tout i, donc $G = 0$. Pour n fixé, la série $\sum_{i=0}^{+\infty} u_{n,i}$ a un seul terme non nul et sa somme vaut 1 et ne dépend donc pas de n, d'où $D = 1$.
Exemple 2. On prend $u_{n,i} = 1$ si i et n ont même parité et $u_{n,i} = 0$ sinon. On a encore $\liminf_{n\to+\infty} u_{n,i} = 0$ pour tout i, donc $G = 0$. Pour n fixé, la série $\sum_{i=0}^{+\infty} u_{n,i}$ a une infinité de termes valant 1 et a donc pour somme $+\infty$ qui est aussi une constante indépendante de n, d'où $D = +\infty$.

Remarque 6. Le lemme de Fatou qui sera vu ultérieurement en cours s'énonce comme suit. Si $(\Omega, \mathcal{F}, \mu)$ est un espace mesuré et (f_n) une suite de fonctions mesurables $\Omega \to \overline{\mathbb{R}}_+$,

$$\int_\Omega \liminf_{n\to+\infty} f_n \,\mathrm{d}\mu \leq \liminf_{n\to+\infty} \int_\Omega f_n \,\mathrm{d}\mu.$$

L'inégalité (24) n'est qu'un cas particulier de ce lemme général. En effet prenons $\Omega = \mathbb{N}$, $\mathcal{F} = \mathcal{P}(\Omega)$ et définissons $f_n : \mathbb{N} \to \overline{\mathbb{R}}_+$, $k \mapsto f_n(k) := u_{n,k}$. Prenons pour μ la mesure de comptage sur $\mathbb{N}$, définie par $\mu = \sum_{k\in\mathbb{N}} \delta_k$. Alors

$$\int_{\mathbb{N}} f_n \,\mathrm{d}\mu = \sum_{k\in\mathbb{N}} u_{n,k} \quad \text{et} \quad \int_{\mathbb{N}} \liminf_{n\to+\infty} f_n \,\mathrm{d}\mu = \sum_{k\in\mathbb{N}} \liminf_{n\to+\infty} u_{n,k}.$$

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé du Devoir n° 2

Ex 1. *Soit $(\Omega, \mathcal{F}, \mu)$ un ensemble mesuré. Soit $(E_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{F}$ telle que*

$$\sum_{n\in\mathbb{N}} \mu(E_n) < +\infty.$$

Notons :

$$A = \{\omega \in \Omega,\ \omega \text{ appartient à une infinité de } E_n\}.$$

1) *Exprimer A en fonction des E_n.*

$$\begin{aligned} \omega \text{ appartient à une infinité de } E_n &\iff \forall N \geq 0, \exists n \geq N, \omega \in E_n \\ &\iff \omega \in \bigcap_{N\geq 0} \bigcup_{n\geq N} E_n \end{aligned}$$

d'où

$$A = \bigcap_{N\geq 0} \bigcup_{n\geq N} E_n.$$

2) *Montrer que $\mu(A) = 0$.*
Notons pour $N \geq 0$,

$$B_N = \bigcup_{n\geq N} E_n.$$

La suite $(B_N)_{N\in\mathbb{N}}$ est une suite décroissante d'éléments de la tribu $\mathcal{F}$. Elle converge vers

$$A := \bigcap_{N\geq 0} B_N.$$

Pour tout $N \geq 0$, on a

$$\mu(B_N) \leq \sum_{k=N}^{+\infty} \mu(E_k) < +\infty.$$

Donc $\mu(B_0) < +\infty$ et, par continuité monotone séquentielle de la mesure,

$$\mu(A) = \lim_{N\to+\infty} \mu(B_N) \leq \lim_{N\to+\infty} \sum_{k=N}^{+\infty} \mu(E_k) = 0$$

car c'est le reste d'une série convergente. D'où $\mu(A) = 0$.

Ex 2. *On note λ la mesure de Lebesgue sur $\mathbb{R}$. Soit A l'ensemble des réels x pour lesquels il existe une infinité de couples d'entiers $(p,q) \in \mathbb{N} \times \mathbb{N}^*$ tels que*

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{q^3}$$

Le but de cet exercice est de montrer que $\lambda(A) = 0$.

1) *Vérifier que A contient l'ensemble $\mathbb{Q}_+$ des rationnels positifs. A est donc une partie infinie de $\mathbb{R}$.*

Si $x \in \mathbb{Q}_+$, alors il existe deux entiers naturels r et s ($s \neq 0$) tels que $x = \dfrac{r}{s}$. Alors

$$\forall n \in \mathbb{N}^* \qquad \frac{r}{s} = \frac{nr}{ns}$$

d'où

$$\forall n \in \mathbb{N}^* \qquad \left| x - \frac{nr}{ns} \right| = 0 < \frac{1}{(ns)^3}.$$

Il existe donc bien une infinité de couples d'entiers naturels (p,q) ($q \neq 0$) vérifiant la condition demandée.

2) *Vérifier que A est un borélien de $\mathbb{R}$ en traduisant sa définition à l'aide d'opérations ensemblistes dénombrables faisant intervenir les intervalles*

$$I_{p,q} := \left[\frac{p}{q} - \frac{1}{q^3}, \frac{p}{q} + \frac{1}{q^3}, \right]$$

où $p \in \mathbb{N}$ et $q \in \mathbb{N}^*$.

$$\begin{aligned}
& x \in A \\
\Longleftrightarrow \quad & \text{il existe une infinité de } (p,q) \in \mathbb{N} \times \mathbb{N}^* \text{ tels que } \left| x - \frac{p}{q} \right| \leq \frac{1}{q^3} \\
\Longleftrightarrow \quad & \text{il existe une infinité de } (p,q) \in \mathbb{N} \times \mathbb{N}^* \text{ tels que } x \in I_{p,q} \\
\Longleftrightarrow \quad & \text{il existe une infinité de } (p,q) \in \mathbb{N} \times \mathbb{N}^* \text{ tels que } qx - \frac{1}{q^2} \leq p \leq qx + \frac{1}{q^2}
\end{aligned}$$

Or un intervalle de longueur $2/q^2$ contient au plus trois entiers, d'où

$$\begin{aligned}
& x \in A \\
\Longleftrightarrow \quad & \text{il existe une infinité de } q \in \mathbb{N}^* \text{ et il existe } p \text{ vérifiant } qx - \frac{1}{q^2} \leq p \leq qx + \frac{1}{q^2} \\
\Longleftrightarrow \quad & \text{il existe une infinité de } q \in \mathbb{N}^* \text{ tels que } x \in \bigcup_{p \in \mathbb{N}} I_{p,q} \\
\Longleftrightarrow \quad & \forall q \in \mathbb{N}^*, \quad \exists q' \geq q, \quad x \in \bigcup_{p \in \mathbb{N}} I_{p,q'} \\
\Longleftrightarrow \quad & x \in \bigcap_{q \in \mathbb{N}^*} \bigcup_{q' \geq q} \bigcup_{p \in \mathbb{N}} I_{p,q'}
\end{aligned}$$

d'où

$$A = \bigcap_{q \in \mathbb{N}^*} \bigcup_{q' \geq q} \bigcup_{p \in \mathbb{N}} I_{p,q'}$$

est bien un borélien de $\mathbb{R}$.

3) *Montrer que* $\lambda(A) = 0$ *si et seulement si* $\lambda(B) = 0$ *avec* $B = [0, 1[\cap A$.

L'implication $\lambda(A) = 0 \Longrightarrow \lambda(B) = 0$ est évidente, B étant inclus dans A.

Remarquons que pour tout $n \in \mathbb{Z}$ et pour tout $x \in B$:

$$\left| x + n - \frac{p + nq}{q} \right| = \left| x - \frac{p}{q} \right| \leq \frac{1}{q^3}$$

donc $x + n \in A$. De même pour tout x dans A, il existe $n \in \mathbb{Z}$ et $\alpha \in B$ tels que $x = n + \alpha$. On a donc

$$A = \bigcup_{n \in \mathbb{N}} (B + n).$$

Supposons maintenant que $\lambda(B) = 0$. Alors pour tout $n \in \mathbb{Z}$, $\lambda(B + n) = 0$ en raison de l'invariance de λ par translation, d'où

$$\lambda(A) \leq \sum_{n \in \mathbb{Z}} \lambda(B + n) = 0.$$

D'où l'équivalence.

4) *Pour* $q \in \mathbb{N}^*$, *on définit :* $B_q := [0, 1[\cap \left(\bigcup_{p \in \mathbb{N}} I_{p,q} \right)$.

Montrer que $x \in [0, 1[$ *est dans* B *si et seulement si* x *appartient à une infinité d'ensembles* B_q.

D'après la question 2) :

$$\begin{aligned} x \in B &\iff x \in [0, 1[\cap A \\ &\iff \text{il existe une infinité de } q \in \mathbb{N}^* \text{ tels que } x \in \bigcup_{p \in \mathbb{N}} I_{p,q} \bigcap [0, 1[\\ &\iff \text{il existe une infinité de } q \in \mathbb{N}^* \text{ tels que } x \in B_q. \end{aligned}$$

5) *Montrer que*

$$\lambda(B_q) \leq \frac{2(q + 2)}{q^3}.$$

Remarquons que

$$\frac{p}{q} - \frac{1}{q^3} \leq x \leq \frac{p}{q} + \frac{1}{q^3} \iff qx - \frac{1}{q^2} \leq p \leq qx + \frac{1}{q^2} \tag{25}$$

et si $x \in [0, 1[$

$$(25) \Longrightarrow -\frac{1}{q^2} \leq p < q + \frac{1}{q^2} \Longrightarrow 0 \leq p < q + 1.$$

Par conséquent

$$\forall q \geq 1 \qquad \{p \in \mathbb{N},\, I_{p,q} \cap [0,1[\neq \emptyset\} \subset \{p \in \mathbb{N},\, 0 \leq p < q+1\}.$$

D'autre part

$$B_q = [0,1[\bigcap \left(\bigcup_{p \in \mathbb{N}} I_{p,q} \right) = \bigcup_{p \in \mathbb{N}} \left(I_{p,q} \bigcap [0,1[\right).$$

On a donc

$$\lambda(B_q) \leq \sum_{p \in \mathbb{Z}} \lambda \left(I_{p,q} \bigcap [0,1[\right) \leq \sum_{0 \leq p < q+1} \lambda(I_{p,q}) = \frac{2(q+1)}{q^3}.$$

6) *En utilisant l'exercice précédent, montrer que* $\lambda(B) = 0$.

D'après la question précédente :

$$\sum_{q \geq 1} \lambda(B_q) \leq \sum_{q \geq 1} \frac{2(q+2)}{q^3} < +\infty.$$

Il suffit alors d'appliquer l'exercice prédédent pour obtenir $\lambda(B) = 0$ et donc $\lambda(A) = 0$.

7) *On se propose de montrer que A n'est pas dénombrable. Pour cela, on fixe une suite d'entiers $(j_k)_{k \geq 1}$ telle que pour tout $k \geq 1$, $j_{k+1} > 3j_k$ (par exemple la suite définie par $j_k = 4^k$ a cette propriété). On note C l'ensemble des réels x pouvant s'écrire sous la forme*

$$x = \sum_{k=1}^{+\infty} \frac{a_k}{2^{j_k}}, \quad (a_k)_{k \geq 1} \in \{0,1\}^{\mathbb{N}^*}.$$

Montrer que C est inclus dans A et qu'il est en bijection avec $\{0,1\}^{\mathbb{N}^*}$. Conclure.

Soit $x \in C$,

$$x = \sum_{k=1}^{+\infty} \frac{a_k}{2^{j_k}}, \quad (a_k)_{k \geq 1} \in \{0,1\}^{\mathbb{N}^*}.$$

Soit $n \geq 1$, posons

$$x_n = \sum_{k=1}^{n} \frac{a_k}{2^{j_k}},$$

on peut écrire $x_n = p_n/q_n$ avec $q_n = 2^{j_n}$. On a :

$$x - x_n = \sum_{k=n+1}^{+\infty} \frac{a_k}{2^{j_k}} = \frac{1}{2^{j_{n+1}}} \sum_{k=n+1}^{+\infty} \frac{a_k}{2^{j_k - j_{n+1}}} \leq \frac{2}{2^{j_{n+1}}} = \frac{1}{2^{j_{n+1}-1}} \leq \frac{1}{2^{3j_n}},$$

en effet $j_{n+1} > 3j_n$ et comme les j_k sont entiers, $j_{n+1} - 1 \geq 3j_n$. On peut encore écrire

$$\forall n \geq 1, \quad \exists (p_n, q_n) \in \mathbb{N} \times \mathbb{N}^*, \quad 0 \leq x - \frac{p_n}{q_n} \leq \frac{1}{q_n^3}.$$

Ce qui prouve que x appartient à A. Donc $C \subset A$.

Montrons que C n'est pas dénombrable. Pour cela considérons l'application suivante :

$$\varphi : \left\{ \begin{array}{ccc} \{0,1\}^{\mathbb{N}^*} & \longrightarrow & C \\ (a_k)_{k\in\mathbb{N}} & \longmapsto & \displaystyle\sum_{k=1}^{+\infty} \frac{a_k}{2^{j_k}} \end{array} \right.$$

Il est clair par construction que φ est surjective. Montrons que φ est aussi injective. Pour cela considérons deux suites distinctes $(a_k)_{k\in\mathbb{N}}$ et $(b_k)_{k\in\mathbb{N}}$ de $\{0,1\}^{\mathbb{N}^*}$ d'images a et b par φ tels que

$$a = b \quad \Longleftrightarrow \quad \sum_{k=1}^{+\infty} \frac{a_k}{2^{j_k}} = \sum_{k=1}^{+\infty} \frac{b_k}{2^{j_k}}$$

On pose

$$k_0 := \sup\{n \in \mathbb{N} \quad ; \quad \forall k < n \quad a_k = b_k\}.$$

Les suites étant distinctes, k_0 est fini. Alors, en supposant $a_{k_0} = 1$ et $b_{k_0} = 0$ (on sait qu'il sont distincts dans $\{0,1\}$) :

$$a - b = \frac{1}{2^{j_{k_0}}} + \sum_{k=k_0+1}^{+\infty} \frac{a_k - b_k}{2^{j_k}}$$

Mais

$$\begin{aligned} \left| \sum_{k=k_0+1}^{+\infty} \frac{a_k - b_k}{2^{j_k}} \right| &\leq \sum_{k=k_0+1}^{+\infty} \frac{|a_k - b_k|}{2^{j_k}} \leq \sum_{k=k_0+1}^{+\infty} \frac{1}{2^{j_k}} \\ &\leq \frac{1}{2^{j_{k_0+1}}} \sum_{k=k_0+1}^{+\infty} \frac{1}{2^{j_k - j_{k_0+1}}} \\ &\leq \frac{1}{2^{j_{k_0+1}}} \sum_{n=0}^{+\infty} \frac{1}{2^n} \\ &\leq \frac{2}{2^{j_{k_0+1}}} \end{aligned}$$

d'où

$$\frac{1}{2^{j_{k_0}}} - \frac{2}{2^{j_{k_0+1}}} \leq a - b \leq \frac{1}{2^{j_{k_0}}} + \frac{2}{2^{j_{k_0+1}}}$$

et

$$a - b \geq \frac{2^{j_{k_0+1}} - 2^{j_{k_0}+1}}{2^{j_{k_0}+j_{k_0+1}}} > \frac{(2^3-2)2^{j_{k_0}}}{2^{j_{k_0}+j_{k_0+1}}} > 0$$

par croissance des j_k : $j_{k+1} > 3j_k$ $(k \geq 1)$, ce qui contredit l'hypothèse $a = b$. Donc φ est injective.

Par suite φ réalise une bijection entre C et $\{0,1\}^{\mathbb{N}^*}$ qui est non dénombrable. C est donc non dénombrable. Il en est de même pour A.

Ex 3. *Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré tel que $0 < \mu(\Omega) < +\infty$. On dit qu'un événement A de mesure non nulle est un* atome *de $(\Omega, \mathcal{F}, \mu)$ si*

$$\forall B \in \mathcal{F}, \quad B \subseteq A \Longrightarrow \mu(B) = 0 \quad \text{ou} \quad \mu(B) = \mu(A).$$

On pourra utiliser le résultat suivant, conséquence directe de la formule de Poincaré : si $A_1, \ldots, A_n$ désignent n (≥ 2) événements dont les intersections deux à deux sont μ-négligeables, alors

$$\mu(A_1 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n).$$

1) *Démontrer que l'on définit une relation d'équivalence $\sim$ dans $\mathcal{F}$ en posant :*

$$\forall A, B \in \mathcal{F}, \quad A \sim B \iff \mu(A \Delta B) = 0.$$

Démontrer qu'alors $\mu(A) = \mu(B)$.

La relation est clairement réflexive et symétrique. Pour montrer la transitivité, il suffit de constater que :

$$A\Delta C \subset (A\Delta B) \cup (B\Delta C).$$

En effet :

$$\begin{cases} A \cap C^c = (A \cap B \cap C^c) \cup (A \cap B^c \cap C^c) \subset (B \cap C^c) \cup (A \cap B^c) \\ A^c \cap C = (A^c \cap B \cap C) \cup (A^c \cap B^c \cap C) \subset (A^c \cap B) \cup (B^c \cap C) \end{cases}$$

d'où

$$\begin{aligned} A\Delta C &= (A \cap C^c) \cup (A^c \cap C) \\ &\subset (B \cap C^c) \cup (A \cap B^c) \cup (A^c \cap B) \cup (B^c \cap C) = (A\Delta B) \cup (B\Delta C). \end{aligned}$$

Montrons que pour deux éléments A et B de la même classe d'équivalence

$$\mu(A) = \mu(B).$$

On a :

$$A\Delta B = (A \cap B^c) \cup (A^c \cap B).$$

Puisque par définition $\mu(A\Delta B) = 0$, on a $\mu(A \cap B^c) = \mu(A^c \cap B) = 0$. Mais

$$A = (A \cap B) \cup (A \cap B^c) \quad \text{et} \quad B = (B \cap A) \cup (B \cap A^c)$$

d'où

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

2)

2.a) *Démontrer que si A est un atome de $(\Omega, \mathcal{F}, \mu)$ et si N est un événement μ-négligeable, $A\Delta N$ est un atome de $(\Omega, \mathcal{F}, \mu)$ équivalent à A.*

Remarquons tout d'abord que $A\Delta N$ n'est pas de μ-mesure nulle. En effet, s'il l'était, on aurait $A \sim N$ et d'après la question précédente $\mu(A) = \mu(N) = 0$ ce qui est contraire au fait que A soit un atome.
Soit $B \in \mathcal{F}$ tel que $B \subseteq A\Delta N$. Montrons qu'alors $\mu(B) = 0$ ou $\mu(B) = \mu(A\Delta N)$.
On a :

$$B = B \cap (A\Delta N) = (B \cap A \cap N^c) \cup (B \cap A^c \cap N)$$

d'où puisque $B \cap A^c \cap N \subset N$ et que N est de mesure nulle :

$$\mu(B) = \mu(B \cap A \cap N^c).$$

Maintenant $B \cap A \cap N^c \subseteq A$ et comme A est un atome : ou bien

$$\mu(B \cap A \cap N^c) = \mu(B) = 0,$$

ou bien $\mu(B \cap A \cap N^c) = \mu(B) = \mu(A)$, mais

$$\mu(A\Delta N) = \mu(A) + \mu(N) - 2\mu(A \cap N) = \mu(A)$$

d'où $\mu(B) = \mu(A\Delta N)$. Ce qui prouve que $A\Delta N$ est un atome.
Montrons maintenant qu'il est équivalent à A. Pour cela montrons que

$$(A\Delta N)\Delta A = N,$$

on aura alors $\mu((A\Delta N)\Delta A) = 0$.
On a :

$$\begin{aligned}
(A\Delta N)\Delta A &= ((A \cap N^c) \cup (A^c \cap N))\Delta A \\
&= (((A \cap N^c)^c \cap (A^c \cap N)^c) \cap A) \cup (((A \cap N^c) \cup (A^c \cap N)) \cap A^c) \\
&= ((A^c \cup N) \cap (A \cup N^c) \cap A) \cup (A^c \cap N) \\
&= ((A^c \cup N) \cap A) \cup (A^c \cap N) \\
&= (N \cap A) \cup (A^c \cap N) \\
&= N
\end{aligned}$$

2.b) *Démontrer que si une classe d'équivalence contient un atome, elle ne contient que des atomes qui ont tous la même mesure.*

Soit B un élément de la classe d'équivalence contenant l'atome A. Alors $A\Delta B$ est un évènement μ-négligeable. On applique la question précédente : $A\Delta(A\Delta B)$ est un atome à A. Mais $A\Delta(A\Delta B) = B$. On a donc bien que B est un atome. Le fait que tous les atomes équivalents aient la même mesure vient de la question 1).

2.c) *Démontrer que si A et B sont deux atomes non-équivalents, $\mu(A \cap B) = 0$.*

On a $A \cap B \subset A$ qui est un atome, donc ou bien $\mu(A \cap B) = 0$ ou bien $\mu(A \cap B) = \mu(A)$. Supposons $\mu(A \cap B) \neq 0$ donc $\mu(A \cap B) = \mu(A)$. Comme $A \cap B \subset B$ qui est aussi un atome, on a $\mu(A \cap B) = \mu(B)$. Mais $\mu(A\Delta B) = \mu(A) + \mu(B) - 2\mu(A \cap B)$, donc $\mu(A\Delta B) = 0$ ce qui est contraire à l'hypothèse. On a donc $\mu(A \cap B) = 0$.

3) *On suppose que $(\Omega, \mathcal{F}, \mu)$ possède au moins un atome. On choisit un représentant dans chaque classe d'équivalence formée d'atomes. On note $(A_i,\, i \in I)$ la famille d'atomes ainsi obtenue. Démontrer qu'elle est finie ou infinie dénombrable.*

Soit $\tilde{A}$ l'ensemble des $(A_i,\, i \in I)$ et pour tout $n \geq 1$

$$\tilde{A}_n = \left\{ A \in \tilde{A}\,,\, \mu(A) > \frac{M}{n} \right\} \quad \text{où} \quad M = \mu(\Omega) < +\infty.$$

Le cardinal de $\tilde{A}_n$ est strictement inférieur à n. En effet, supposons que $\tilde{A}_n$ contienne n éléments de A disons $A_1, \ldots, A_n$. Alors compte tenu de la définition des A_i et de la question précédente :

$$\mu(A_1 \cup A_2 \cup \ldots \cup A_n) = \mu(A_1) + \mu(A_2) + \cdots + \mu(A_n) > n\frac{M}{n} = M$$

ce qui contredit la définition de M. Or

$$\tilde{A} = \bigcup_{n \geq 1} \tilde{A}_n$$

ce qui prouve que la famille $\tilde{A}$ est au plus dénombrable, comme union dénombrable d'ensemble finis.

4) *On munit $\mathbb{R}$ de la tribu*

$$\mathcal{F} = \big\{ B\,;\, B \in \mathrm{Bor}(\mathbb{R})\,,\, B \supseteq [0,1] \quad \text{ou} \quad B^c \supseteq [0,1] \big\},$$

et l'on considère la fonction P sur $\mathcal{F}$ qui ne prend que les valeurs 0 et 1 et vérifie

$$\forall F \in \mathcal{F}, \quad P(F) = 1 \iff F \supseteq [0,1].$$

Démontrer que la fonction P est une probabilité sur $(\mathbb{R}, \mathcal{F})$ et que $[0,1]$ est un atome de $(\mathbb{R}, \mathcal{F}, P)$.

On a clairement $P(\emptyset) = 0$.

Soit $(F_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{F}$ deux à deux disjoints et $F = \bigcup_{n \in \mathbb{N}} F_n$.

Si pour tout $n \in \mathbb{N}$, $P(F_n) = 0$ alors $[0,1] \subseteq F_n^c$ et $F^c = \bigcap_{n \in \mathbb{N}} F_n^c$ contient $[0,1]$. D'où

$$P(F) = 0 = \sum_{n \in \mathbb{N}} P(F_n).$$

Sinon, les F_n étant disjoints, il existe un unique n_0 tel que $P(F_{n_0}) = 1$. On a alors

$$P(F) = 1 = \sum_{n \in \mathbb{N}} P(F_n).$$

Ce qui prouve que P est une mesure sur $(\mathbb{R}, \mathcal{F})$ et une probabilité puisque $P(\mathcal{F}) = 1$.

On a $P([0,1]) = 1 \neq 0$. Soit $B \subseteq [0,1]$ un élément de $\mathcal{F}$, alors soit $B = [0,1]$ et dans ce cas $P(B) = P([0,1]) = 1$, soit $B \subsetneq [0,1]$ et dans ce cas puisque $B \in \mathcal{F}$, $[0,1] \subseteq B^c$ donc

$$B \subsetneq [0,1] \subseteq B^c,$$

d'où $B = \emptyset$ et $P(B) = 0$. Ceci montre bien que $[0,1]$ est un atome.

Ex 4. Exercice 3 de la version antérieure

Soit $(X, \mathcal{B})$ un espace mesurable. On dit qu'un élément A de $\mathcal{B}$ est un atome si A est non vide et si on a :

$$\forall B \in \mathcal{B} \quad B \subset A \Longrightarrow (B = \emptyset \text{ ou } B = A).$$

1) *Montrer que la collection des atomes est une famille de parties deux à deux disjointes.*

Soit A_1 et A_2 deux atomes distincts. Considérons $A := A_1 \cap A_2$. Si A est non vide, alors $A \subset A_1$ implique $A = A_1$ et de même $A \subset A_2$ implique $A = A_2$. D'où $A_1 = A_2$ ce qui contredit l'hypothèse. Les atomes sont donc deux à deux disjoints.

On dit que l'espace mesurable $(X, \mathcal{B})$ est atomique si tout élément de $\mathcal{B}$ est réunion d'atomes de $\mathcal{B}$.

2) *Montrer que $(X, \mathcal{B})$ est atomique si et seulement si la collection des atomes est une partition de X.*

Supposons l'espace atomique, alors tout élément de $\mathcal{B}$ est réunion d'atomes, donc en particulier X est réunion d'atomes, et comme les atomes sont deux à deux disjoints et qu'ils sont tous inclus dans X, la collection des atomes est une partition de X.

Réciproquement, supposons que $X = \bigcup_{i \in I} A_i$, où $\{A_i, i \in I\}$ est la collection des atomes de $\mathcal{B}$. Soit $B \in \mathcal{B}$, $B \neq \emptyset$, et soit $x \in B$. Comme on a une partition de X, il existe un atome A_x tel que $x \in A_x$. On a alors, $A_x \cap B \neq \emptyset$ et $A_x \cap B \subset A_x$ donc $A_x \cap B = A_x$, on a donc

$$B = \bigcup_{x \in B} \{x\} \subset \bigcup_{x \in B} A_x \subset B,$$

ce qui prouve que B peut s'écrire comme réunion d'atomes de $\mathcal{B}$.

3) *Montrer que $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est atomique.*

Pour cela, nous allons montrer que les singletons sont des boréliens et que ce sont les seuls atomes de $\mathcal{B}(\mathbb{R})$.

Pour tout $x \in \mathbb{R}$,

$$\{x\} = \bigcap_{n \in \mathbb{N}^*} \left] x - \frac{1}{n}, x + \frac{1}{n} \right[\in \mathcal{B}(\mathbb{R}).$$

Soit $A \in \mathcal{B}(\mathbb{R})$. Il est clair que $A \subset \{x\}$ implique que $A = \{x\}$ ou $A = \emptyset$. Ceci montre que les singletons sont des atomes. Réciproquement, si A est un atome de $\mathcal{B}(\mathbb{R})$, il existe $x \in A$. On a $\{x\} \subset A$ d'où $A = \{x\}$. Ce qui prouve que $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ est atomique.

4) *On suppose que $\mathcal{B}$ admet un atome A. On définit μ_A sur $\mathcal{B}$ par $\mu_A(B) = 1$ si $A \subset B$ et $\mu_A(B) = 0$ sinon. Montrons que μ_A est une mesure.*

On a clairement $\mu_A(\emptyset) = 0$.

Soit $(B_i)_{i \in \mathbb{N}}$ est une suite d'éléments de $\mathcal{B}$ deux à deux disjoints. Alors si $A \cap \cup B_i = \emptyset$, alors

$$\mu_A\left(\bigcup_{i \in \mathbb{N}} B_i\right) = 0 = \sum_{i \in \mathbb{N}} \mu_A(B_i).$$

Maintenant s'il existe $x \in A \bigcap \left(\bigcup_{i \in \mathbb{N}} B_i \right)$, alors il existe un unique i_0 tel que $x \in B_{i_0} \cap A$. Mais $B_{i_0} \cap A \subset A$ et A est un atome, donc $B_{i_0} \cap A = A$. D'où $A \subset \cup B_i$ et $\mu_A(\cup B_i) = 1$. D'autre part $A \subset B_{i_0}$ et $\mu_A(B_{i_0}) = 1$. De plus pour tout i différent de i_0, $A \cap B_i = \emptyset$ et $\mu_A(B_i) = 0$. D'où

$$\mu_A \left(\bigcup_{i \in \mathbb{N}} B_i \right) = 1 = \sum_{i \in \mathbb{N}} \mu_A(B_i).$$

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé du Devoir Surveillé du 28 novembre 2003

Ex 1. *Continuité presque partout.*
Soit μ une mesure sur $\big(\mathbb{R}, \mathrm{Bor}(\mathbb{R})\big)$ vérifiant les deux conditions

$$\mu(\{0\}) = 0, \tag{26}$$

$$\forall a, b \in \mathbb{R}, \quad a < b \ \Rightarrow\ \mu(]a, b[) > 0. \tag{27}$$

On note f la fonction définie sur $\mathbb{R}$ par

$$f(x) := \begin{cases} \sin(1/x) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

1) La fonction f est continue sur $\mathbb{R}^*$ comme composée de la fonction $x \mapsto 1/x$, continue sur $\mathbb{R}^*$, avec la fonction sinus qui est continue sur $\mathbb{R}$. Sans qu'il soit besoin d'étudier la continuité de f en zéro, on en déduit que l'ensemble D des points de discontinuité de f est *inclus* dans $\{0\}$. Par hypothèse $\mu(\{0\}) = 0$, donc $\mu(D) = 0$, ce qui signifie que f est continue μ-presque partout.
Remarque. En fait f *n'est pas* continue au point 0 parce que $f(x)$ oscille indéfiniment entre -1 et $+1$ quand x tend vers 0. Plus précisément, notons $u_k = (\pi/2 + k\pi)^{-1}$. Alors $(u_k)_{k\in\mathbb{N}}$ tend vers 0 quand k tend vers $+\infty$ et $f(u_k) = \sin(\pi/2 + k\pi) = (-1)^k$ ne tend pas vers $f(0) = 0$, ce qui suffit pour empêcher f d'être continue au point 0.

2) Supposons qu'il existe une fonction g continue *partout* sur $\mathbb{R}$ et égale μ-presque partout à f. En notant $A := \{x \in \mathbb{R};\ f(x) \neq g(x)\}$, on a donc $\mu(A) = 0$.

Définissons les intervalles I_k par :

$$I_k := \left[\frac{1}{5\pi/6 + k\pi}, \frac{1}{\pi/6 + k\pi}\right], \quad k \in \mathbb{N}.$$

Si $x \in I_k$, $1/x$ appartient à $[\pi/6 + k\pi, 5\pi/6 + k\pi]$. Par conséquent,

a) si k est pair, $\forall x \in I_k$, $f(x) \geq 1/2$;

b) si k est impair, $\forall x \in I_k$, $f(x) \leq -1/2$.

Pour chaque $k \in \mathbb{N}$, I_k a une intersection non vide avec le complémentaire de A. En effet si $A^c \cap I_k = \emptyset$, I_k est inclus dans A d'où $\mu(A) \geq \mu(I_k)$, ce qui est contradictoire avec la nullité de $\mu(A)$ puisque d'après l'hypothèse (27), $\mu(I_k)$ est strictement positif. Ainsi chaque I_k contient au moins un point de A^c, que nous noterons x_k et qui vérifie donc $f(x_k) = g(x_k)$. On voit immédiatement que $\sup I_{k+1} \leq \inf I_k$, ce qui assure la décroissance de (x_k). Comme $0 \leq x_k \leq (\pi/6 + k\pi)^{-1}$, la suite (x_k) converge vers 0. Par continuité de g au point 0, $g(x_k)$ converge vers $g(0)$ et donc $g(x_{k+1}) - g(x_k)$ converge vers 0. On aboutit ainsi à une contradiction puisque :

$$\forall j \in \mathbb{N}^*, \quad g(x_{2j}) - g(x_{2j-1}) \geq \frac{1}{2} - \frac{-1}{2} = 1 > 0.$$

Il n'existe donc aucune fonction g continue *partout* sur $\mathbb{R}$ et égale μ-presque partout à f.

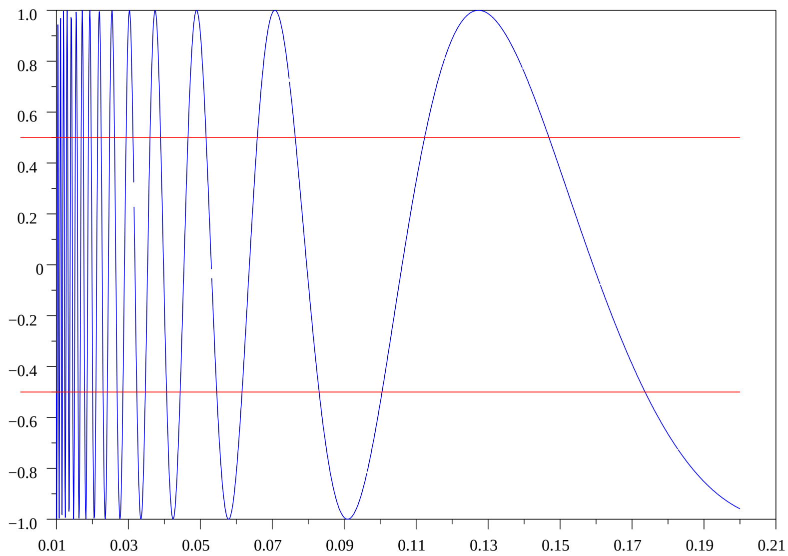

FIG. 7 – Fonction $x \mapsto \sin(1/x)$ au voisinage de 0

3) *Question facultative.* On suppose que $h : \mathbb{R} \to \mathbb{R}$ est borélienne et *uniformément* continue sur un borélien E de $\mathbb{R}$ tel que $\mu(\mathbb{R} \setminus E) = 0$. Montrons qu'il existe une fonction g partout continue sur $\mathbb{R}$ telle que $h(x) = g(x)$ pour tout $x \in E$.

On commence par remarquer que E est dense dans $\mathbb{R}$, par le même argument qu'à la question précédente. En effet pour tout intervalle ouvert $I =]a, b[$ (boule ouverte pour la métrique usuelle de $\mathbb{R}$), $I \cap E \neq \emptyset$ car sinon $I \subset E^c$ et $\mu(I) \leq \mu(E^c) = 0$. La fonction h étant uniformément continue sur une partie dense de l'espace métrique *complet* $\mathbb{R}$ admet un unique prolongement continu[15] $\tilde{h}$ à $\mathbb{R}$. Ainsi la fonction $g := \tilde{h}$ coïncide avec

15. C'est un théorème classique d'analyse : toute fonction h *uniformément* continue sur une partie

h au moins sur E. Les fonctions g et h étant boréliennes, l'ensemble $A := \{x \in \mathbb{R};\ g(x) \neq h(x)\}$ est un borélien de $\mathbb{R}$. Comme $A \subset \mathbb{R} \setminus E$, $\mu(A) = 0$. On a donc bien une fonction g continue partout sur $\mathbb{R}$ et égale à h μ-presque partout sur $\mathbb{R}$.
Remarque. Notons que la fonction $f(x) = \sin(1/x)$ *n'est pas* uniformément continue sur $E = \mathbb{R}^*$.

Ex 2. *Théorème de Poincaré.*
Le but de cet exercice est de démontrer le théorème suivant.

Théorème 7. *Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et $f : \Omega \to \Omega$, mesurable $\mathcal{F}$-$\mathcal{F}$, qui conserve la probabilité $\mathbf{P}$: $\forall C \in \mathcal{F}$, $\mathbf{P}(f^{-1}(C)) = \mathbf{P}(C)$.*
Alors pour tout A de $\mathcal{F}$, il existe $A' \in \mathcal{F}$ et inclus dans A tel que :
a) $\mathbf{P}(A') = \mathbf{P}(A)$
b) $\forall \omega \in A'$, $\forall n \in \mathbb{N}^*$, $\exists p \geq n$, *tel que* $f^p(\omega) \in A$, *où* $f^p = \underbrace{f \circ f \circ \cdots \circ f}_{p \text{ fois}}$.

Si on appelle *trajectoire* de ω la suite de ses images $(f^p(\omega))_{p \geq 1}$, on peut exprimer ceci en disant que pour presque tout ω de A, la trajectoire de ω repasse par A une infinité de fois.

1) Vérifions les hypothèses du théorème sur l'exemple proposé par l'énoncé. On prend $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathrm{Bor}([0,1]), \lambda)$, où λ est la restriction de la mesure de Lebesgue sur $\mathbb{R}$ à la tribu borélienne de $[0,1]$. On définit $f : [0,1] \to [0,1]$ par

$$f(x) = 2x\mathbf{1}_{[0,1/2[}(x) + (2x-1)\mathbf{1}_{[1/2,1]}(x).$$

La fonction f est mesurable comme somme de deux fonctions qui sont chacune produit d'une fonction continue (donc borélienne) par l'indicatrice d'un borélien. La restriction de la mesure de Lebesgue à $\mathrm{Bor}([0,1])$ est bien une probabilité sur $([0,1], \mathrm{Bor}([0,1]))$ puisque $\lambda(\Omega) = \lambda([0,1]) = 1$. Pour vérifier la conservation de la probabilité λ par f, on commence par chercher $f^{-1}([a,b])$ pour $0 \leq a < b \leq 1$.

Chaque $y \in [0,1]$ a exactement deux antécédents par f dans $[0,1]$, l'un est $y/2$ et appartient à $[0,1/2]$, l'autre est $(1+y)/2$ et appartient à $[1/2,1]$. Ceci nous amène à écrire pour $0 \leq a < b \leq 1$,

$$\begin{aligned} f^{-1}([a,b]) &= \big\{x \in [0,1];\ f(x) \in [a,b]\big\} \\ &= \big\{x \in [0,1/2];\ f(x) \in [a,b]\big\} \cup \big\{x \in [1/2,1];\ f(x) \in [a,b]\big\} \\ &= \big\{y/2;\ y \in [a,b]\big\} \cup \big\{(1+y)/2;\ y \in [a,b]\big\} \\ &= \big[a/2, b/2\big] \cup \big[(1+a)/2, (1+b)/2\big]. \end{aligned}$$

Notons au passage que l'on retrouve la mesurabilité de f puisque $f^{-1}([a,b])$ est borélien comme réunion de deux intervalles fermés et que la tribu $\mathrm{Bor}([0,1])$ est engendrée par la famille $\mathcal{C} := \big\{[a,b],\ 0 \leq a, b \leq 1\big\}$.

dense E d'un espace métrique *complet* F admet un unique prolongement $\tilde{h}$ uniformément continu à F. La clé de la preuve est le fait que l'image d'une suite de Cauchy de E par la fonction uniformément continue h est une suite de Cauchy dans $\mathbb{R}$.

De plus les intervalles $\big[a/2, b/2\big]$ et $\big[(1+a)/2, (1+b)/2\big]$ sont disjoints, sauf dans le cas particulier $a = 0$ et $b = 1$ où leur intersection est le singleton $\{1/2\}$ de mesure nulle. On a donc

$$\lambda\big(f^{-1}([a,b])\big) = \lambda\big([a/2, b/2]\big) + \lambda\big([(1+a)/2, (1+b)/2]\big) = \frac{b-a}{2} + \frac{(1+b)-(1+a)}{2} = b-a.$$

Ainsi les probabilités λ et $\lambda \circ f^{-1}$ coïncident sur la π-classe [16] $\mathcal{C}$ qui engendre la tribu $\mathrm{Bor}([0,1])$. Par le théorème d'unicité pour les mesures, elles coïncident donc sur toute la tribu borélienne de $[0,1]$. Ainsi $\lambda \circ f^{-1} = \lambda$, autrement dit f conserve la probabilité λ.

2) On introduit les ensembles

$$A_0 := \{\omega \in A;\ \forall n \in \mathbb{N}^*,\ f^n(\omega) \notin A\}, \quad A_1 := \{\omega \in A;\ \exists n \in \mathbb{N}^*,\ f^n(\omega) \in A\}.$$

A_0 est l'ensemble des éléments de A dont la trajectoire ne repasse jamais par A, tandis que $A_1 = A \setminus A_0$ est l'ensemble de ceux dont la trajectoire repasse au moins une fois par A. Le but de cette question est d'établir que $\mathbf{P}(A_1) = \mathbf{P}(A)$. On pose pour tout $n \in \mathbb{N}^*$, $B_n := (f^n)^{-1}(A_0)$.

2. a) Vérifions que A_0 et A_1 et tous les B_n sont dans $\mathcal{F}$. La mesurabilité étant préservée par la composition des applications, il est clair que f^n est mesurable. L'appartenance des B_n à $\mathcal{F}$ découlera donc immédiatement de l'appartenance de A_0 à $\mathcal{F}$. D'autre part une tribu étant stable par différence ensembliste, l'appartenance de A_1 à $\mathcal{F}$ découlera de celles de A_0 et A à $\mathcal{F}$. Finalement, il ne reste plus qu'à établir que $A_0 \in \mathcal{F}$. Pour ce faire, on réécrit la définition de A_0 à l'aide d'opérations ensemblistes dénombrables sur des éléments de la tribu $\mathcal{F}$:

$$\begin{aligned} A_0 &= \{\omega \in A;\ \forall n \in \mathbb{N}^*,\ f^n(\omega) \notin A\} \\ &= \{\omega \in A;\ \forall n \in \mathbb{N}^*,\ \omega \in (f^n)^{-1}(A^c)\} \\ &= A \cap \Big(\bigcap_{n\in\mathbb{N}^*} (f^n)^{-1}(A^c)\Big). \end{aligned}$$

Comme A appartient à $\mathcal{F}$, A^c aussi et les $(f^n)^{-1}(A^c)$ sont dans $\mathcal{F}$ par mesurabilité des f^n. La tribu $\mathcal{F}$ étant stable par intersections dénombrables, l'appartenance à $\mathcal{F}$ de A_0 est établie.

2. b) Les B_n sont deux à deux disjoints. En effet, prenons deux indices distincts dans $\mathbb{N}^*$ et notons i le plus petit, j le plus grand. On peut écrire $j = i + k$ avec $k \in \mathbb{N}^*$. Soit ω un élément de $B_i \cap B_j$. Alors l'appartenance de ω à B_i signifie que $f^i(\omega) \in A_0$ et implique que $f^k(f^i(\omega))$ n'appartient pas à A et donc *a fortiori* $f^{i+k}(\omega) \notin A_0$. Mais ceci contredit l'appartenance de ω à $B_j = B_{i+k}$. Il n'existe donc aucun élément ω commun aux ensembles B_i et B_j.

2. c) Une récurrence immédiate montre que f^n conserve la probabilité $\mathbf{P}$. Par conséquent on a pour tout $n \in \mathbb{N}^*$, $\mathbf{P}(B_n) = \mathbf{P}\big((f^n)^{-1}(A_0)\big) = \mathbf{P}(A_0)$. Les B_n étant deux à deux

16. Rappelons qu'une π-classe est une famille stable par intersections finies, ce qui est bien le cas de $\mathcal{C}$ en notant que pour $b < a$, $[a,b] = \emptyset$.

disjoints, on a par σ-additivité de $\mathbf{P}$

$$1 \geq \mathbf{P}\Big(\underset{n\in\mathbb{N}^*}{\cup} B_n \Big) = \sum_{n\in\mathbb{N}^*} \mathbf{P}(B_n) \geq 0.$$

La série de terme général *constant* $\mathbf{P}(B_n) = \mathbf{P}(A_0)$ est donc convergente. Ceci ne peut se produire que si $\mathbf{P}(A_0) = 0$.

Comme $A_1 = A \setminus A_0$, on obtient finalement $\mathbf{P}(A_1) = \mathbf{P}(A) - \mathbf{P}(A_0) = \mathbf{P}(A)$.

3) On définit par récurrence $A_k := \{\omega \in A_{k-1};\ \exists n \geq 1,\ f^n(\omega) \in A_{k-1}\}$ pour tout $k \geq 2$. Ainsi A_k est l'ensemble des éléments ω de A dont la trajectoire repasse au moins k fois par A. Pour $k \geq 2$, A_k est défini exactement de la même façon que A_1, à condition de remplacer A par A_{k-1} dans la définition de A_1. Le raisonnement fait à la question précédente et une récurrence immédiate nous fournissent l'appartenance de A_k à $\mathcal{F}$ et l'égalité $\mathbf{P}(A_k) = \mathbf{P}(A_{k-1})$. Ainsi pour tout $k \geq 1$, $\mathbf{P}(A_k) = \mathbf{P}(A_1) = \mathbf{P}(A)$.

Pour achever la démonstration du théorème 7, il suffit de remarquer que la suite $(A_k)_{k\geq 1}$ est décroissante pour l'inclusion. En posant $A' = \underset{k\geq 1}{\cap} A_k$, on a $A' \in \mathcal{F}$ et par continuité séquentielle décroissante de $\mathbf{P}$:

$$\mathbf{P}(A') = \lim_{k\to+\infty} \mathbf{P}(A_k) = \mathbf{P}(A).$$

Enfin si $\omega \in A'$, sa trajectoire repasse au moins k fois dans A et ce pour chaque valeur de $k \in \mathbb{N}^*$. Autrement dit elle repasse une infinité de fois par A.

Ex 3. *Interlude*

Soit $(T, \mathcal{T}, \mu)$ un espace mesuré et f une application $T \to \overline{\mathbb{R}}_+$, mesurable $\mathcal{T}$ - $\mathrm{Bor}(\overline{\mathbb{R}}_+)$. On note ν la mesure de densité f par rapport à μ, définie par :

$$\forall B \in \mathcal{T}, \quad \nu(B) = \int_B f \, \mathrm{d}\mu.$$

Montrons que si $\mu(B) = 0$, alors $\nu(B) = 0$.

Examinons d'abord le cas particulier où f est bornée sur T. Il existe donc une constante $M \in \mathbb{R}_+$ telle que $0 \leq f \leq M$. Par croissance de l'intégrale pour les fonctions mesurables positives on a alors

$$0 \leq \nu(B) = \int_B f \, \mathrm{d}\mu \leq \int_B M \, \mathrm{d}\mu = M\mu(B) = 0.$$

Pour passer au cas général, on peut utiliser la suite (f_n) des troncatures de f définies par $f_n := n\mathbf{1}_{\{f>n\}} + f\mathbf{1}_{\{f\leq n\}}$. Chaque f_n est bornée (par n) et la suite de fonctions mesurables positives (f_n) converge en croissant vers f. Par le théorème de Beppo Levi, $\int_B f_n \, \mathrm{d}\mu \uparrow \int_B f \, \mathrm{d}\mu$ et comme $\int_B f_n \, \mathrm{d}\mu = 0$ pour tout n d'après le cas particulier ci-dessus, on en déduit que $\int_B f \, \mathrm{d}\mu = 0$, d'où $\nu(B) = 0$.

Commentaire. Cet exercice était presque une question de cours et il y a plusieurs variantes à sa solution. On peut par exemple vérifier la propriété pour les indicatrices, puis les fonctions étagées et approximer f mesurable positive par une suite croissante de fonctions étagées en utilisant le théorème de Beppo Levi. Il y a d'ailleurs plus simple quand on a établi le résultat sur les fonctions étagées, c'est d'utiliser la définition de $\int_T f\mathbf{1}_B \,\mathrm{d}\mu$ comme supremum des intégrales des fonctions étagées $u \leq f\mathbf{1}_B$ qui sont donc toutes nulles. La solution la plus rapide consiste à utiliser la croissance et l'homogénéité de l'intégrale des fonctions mesurables positives ($\int_T cg \,\mathrm{d}\mu = c\int_T g\,\mathrm{d}\mu$, y compris pour la constante $c = +\infty$) en écrivant :

$$\int_B f\,\mathrm{d}\mu \leq \int_B (+\infty)\,\mathrm{d}\mu = (+\infty) \times \mu(B) = 0,$$

en raison de la convention « $(+\infty) \times 0 = 0$ ». Il convient néanmoins de remarquer que pour prouver l'homogénéité de l'intégrale dans le cas $c = +\infty$ on utilise d'une façon ou d'une autre le théorème de Beppo Levi.

Problème.

Le but du problème est l'étude de la loi de :

$$Y := \sum_{k=1}^{+\infty} \frac{X_k}{4^k}, \tag{28}$$

où $(X_k)_{k\geq 1}$ est une suite de variables aléatoires de Bernoulli de même paramètre $p \in]0,1[$, définies sur $(\Omega, \mathcal{F}, \mathbf{P})$.

1) Remarquons d'abord que (28) définit bien une application $\Omega \to \mathbb{R}$. En effet pour tout $\omega \in \Omega$, la série numérique de terme général $4^{-k}X_k(\omega)$ converge dans $\mathbb{R}_+$ par comparaison avec la série géométrique de raison $1/4$, en raison de l'encadrement $0 \leq 4^{-k}X_k(\omega) \leq 4^{-k}$ (puisque $X_k(\omega)$ vaut 0 ou 1). De plus cet encadrement nous donne la majoration

$$\forall \omega \in \Omega, \quad 0 \leq Y(\omega) \leq \sum_{k=1}^{+\infty} \frac{1}{4^k} = \frac{1}{4} \times \frac{1}{1 - 1/4} = \frac{1}{3}.$$

L'application Y est donc bornée sur Ω par la constante $1/3$.

Passons à la mesurabilité qui ici, s'entend relativement aux tribus $\mathcal{F}$ et $\mathrm{Bor}(\mathbb{R})$. La mesurabilité de Y découle de celle des X_k (comme variables aléatoires sur $(\Omega, \mathcal{F})$) et de la convergence de la série (28) sur tout Ω. En effet les sommes partielles $S_n := \sum_{k=1}^n 4^{-k}X_k$ de cette série sont mesurables comme combinaisons linéaires de fonctions mesurables (les X_k) et Y est ainsi mesurable comme limite simple (en fait uniforme) sur Ω de la suite de fonctions mesurables $(S_n)_{n\geq 1}$.

La formule (28) définit donc une application mesurable $Y : \Omega \to \mathbb{R}$, vérifiant de plus $Y(\Omega) \subset [0, 1/3]$. Ainsi Y est une variable aléatoire positive bornée.

2) La quantité $\mathbf{E}Y$ existe toujours comme élément de $\overline{\mathbb{R}}_+$ puisque Y est une variable aléatoire positive. On est sûr qu'elle est finie puisque Y est une variable aléatoire bornée. Pour la calculer, le théorème de Beppo Levi sur l'interversion série intégrale pour les

fonctions mesurables positives nous permet d'écrire :

$$\mathbf{E}Y = \mathbf{E}\left\{\sum_{k=1}^{+\infty} \frac{X_k}{4^k}\right\} = \sum_{k=1}^{+\infty} \frac{\mathbf{E}X_k}{4^k} = \sum_{k=1}^{+\infty} \frac{p}{4^k} = \frac{p}{3}.$$

3) Pour $j \in \mathbb{N}^*$ le reste d'ordre j de la série géométrique de raison $1/4$ se calcule comme suit :

$$\sum_{k>j} \frac{1}{4^k} = \frac{1}{4^{j+1}} \sum_{i=0}^{+\infty} \frac{1}{4^i} = \frac{1}{4^{j+1}} \times \frac{1}{1-1/4} = \frac{1}{3 \times 4^j}.$$

4) On désigne par C l'ensemble des réels $x \in [0,1]$ tels que

$$\exists (a_k)_{k\geq 1} \in \{0,1\}^{\mathbb{N}^*}, \quad x = \sum_{k=1}^{+\infty} \frac{a_k}{4^k}. \tag{29}$$

4. a) Étudions l'unicité de la décomposition (29). Pour cela on suppose que pour deux suites $(a_k)_{k\geq 1}$ et $(b_k)_{k\geq 1}$ appartenant à $\{0,1\}^{\mathbb{N}^*}$,

$$\sum_{k=1}^{+\infty} \frac{a_k}{4^k} = \sum_{k=1}^{+\infty} \frac{b_k}{4^k}. \tag{30}$$

Définissons $j := \sup\{n \in \mathbb{N}^*;\ \forall k \leq n,\ a_k = b_k\}$. Si j vaut $+\infty$, cela signifie que pour tout $k \in \mathbb{N}^*$, $a_k = b_k$. Par contre si j est fini, cela signifie que l'ensemble des entiers $k \in \mathbb{N}^*$ tels que $a_k = b_k$ présente au moins un « trou » et que le premier [17] de ces trous est situé sur l'entier $j+1$. Plus formellement :

$$j < +\infty \quad \Leftrightarrow \quad \forall 1 \leq k \leq j,\ a_k = b_k \text{ et } a_{j+1} \neq b_{j+1}. \tag{31}$$

On peut supposer sans perte de généralité que $a_{j+1} = 1$ et $b_{j+1} = 0$ (quitte à permuter les notations a et b). De (30) et (31) on déduit alors

$$\frac{1}{4^{j+1}} + \sum_{k=j+2}^{+\infty} \frac{a_k}{4^k} = \sum_{k=j+2}^{+\infty} \frac{b_k}{4^k},$$

ce qui implique, en notant que $|b_k - a_k|$ ne peut valoir que 0 ou 1,

$$\frac{1}{4^{j+1}} = \sum_{k=j+2}^{+\infty} \frac{b_k - a_k}{4^k} \leq \sum_{k=j+2}^{+\infty} \frac{|b_k - a_k|}{4^k} \leq \sum_{k=j+2}^{+\infty} \frac{1}{4^k} = \frac{1}{3 \times 4^{j+1}}.$$

On aboutit ainsi à une absurdité, ce qui montre que j ne peut être fini et établit l'unicité de la décomposition (29).

17. Rien ne permet d'affirmer que $a_k \neq b_k$ pour tout $k > j$, contrairement à ce que les correcteurs ont pu lire dans trop de copies !

4. b) L'application

$$\varphi : \{0,1\}^{\mathbb{N}^*} \to C \quad (a_k)_{k\geq 1} \longmapsto \sum_{k=1}^{+\infty} \frac{a_k}{4^k}$$

est clairement surjective par la définition même de C. L'unicité de la décomposition (29) montre qu'elle est injective. C'est donc une bijection de $\{0,1\}^{\mathbb{N}^*}$ sur C. Comme $\{0,1\}^{\mathbb{N}^*}$ n'est pas dénombrable, on en déduit que C ne l'est pas.

5) On se propose de montrer que $\lambda(C) = 0$. Pour cela on fixe un choix de n premiers chiffres binaires $c_1, \dots, c_n \in \{0,1\}$ et on note $C(c_1, \dots, c_n)$ le sous-ensemble de C formé des x tels que dans le développement (29), $a_i = c_i$ pour $1 \leq i \leq n$. On admet que C et les $C(c_1, \dots, c_n)$ sont des boréliens de $\mathbb{R}$.

5. a) En notant $s_n := \sum_{k=1}^{n} c_k 4^{-k}$, on voit que tout $x \in C(c_1, \dots, c_n)$ a un développement (29) de la forme

$$x = s_n + \sum_{k=n+1} \frac{a_k}{4^k},$$

où les $(a_k)_{k>n}$ sont dans $\{0,1\}$. En utilisant la question 3, on en déduit que

$$\forall x \in C(c_1, \dots, c_n), \qquad s_n \leq x \leq s_n + \frac{1}{3 \times 4^n}.$$

Par conséquent on a $C(c_1, \dots, c_n)$ est inclus dans $\left[s_n, s_n + \frac{1}{3\times 4^n}\right]$, d'où

$$\lambda\big(C(c_1, \dots, c_n)\big) \leq \lambda\left(\left[s_n, s_n + \frac{1}{3 \times 4^n}\right]\right) = \frac{1}{3 \times 4^n}.$$

Remarquons que si les bornes de l'intervalle dépendent, *via* la valeur de s_n, du choix des chiffres $c_1, \dots, c_n$, le majorant obtenu ne dépend que de n et pas du choix des n premiers chiffres. En considérant tous les choix possibles $\mathbf{c} = (c_1, \dots, c_n)$ des n premiers chiffres, on obtient pour C la décomposition

$$C = \underset{\mathbf{c}\in\{0,1\}^n}{\cup} C(\mathbf{c}).$$

Par sous-additivité finie de λ, on en déduit que

$$\lambda(C) \leq \sum_{\mathbf{c}\in\{0,1\}^n} \lambda\big(C(\mathbf{c})\big) \leq \operatorname{card}(\{0,1\}^n) \times \frac{1}{3 \times 4^n} = \frac{2^n}{3 \times 4^n} = \frac{1}{3 \times 2^n}$$

5. b) L'inégalité $\lambda(C) \leq 2^{-n}$ étant vérifiée pour tout n, on obtient en faisant tendre n vers l'infini, $\lambda(C) = 0$.

6) On note P_Y la loi de Y. Puisque C est un borélien de $\mathbb{R}$, $P_Y(C)$ est bien défini et s'écrit $P_Y(C) = P\big(Y^{-1}(C)\big)$. Il est clair d'autre part que pour tout $\omega \in \Omega$, $Y(\omega)$ appartient à C. On a donc[18] $Y^{-1}(C) = \Omega$ et $P_Y(C) = 1$. D'après l'exercice 3, la mesure P_Y ne peut avoir de densité rapport à λ car s'il en était ainsi on devrait avoir $P_Y(C) = 0$ puisque $\lambda(C) = 0$.

18. Attention, on a seulement $Y(\Omega) \subset C$ et il se peut que l'inclusion soit stricte. Par exemple la variable Y construite à la question 8.a) ci-dessous ne prend que deux valeurs distinctes, 0 et 1/3, donc $Y(\Omega) = \{0, 1/3\} \subsetneq C$.

7) On suppose dans cette question que les variables aléatoires de Bernoulli X_k sont indépendantes (et toujours de même paramètre $p \in]0,1[$). Cette indépendance se traduit par la validité des égalités

$$\mathbf{P}(X_1 = c_1, \dots, X_n = c_n) = \mathbf{P}(X_1 = c_1) \dots \mathbf{P}(X_n = c_n),$$

pour tout $n \geq 2$ et tout choix de chiffres binaires $(c_1, \dots, c_n) \in \{0,1\}^n$.

Fixons $x \in C$ et notons $x = \sum_{k\geq 1} a_k 4^{-k}$ son unique développement (29), autrement dit, $(a_k)_{k\geq 1} = \varphi^{-1}(x)$. En raison de l'unicité du développement (29), on a pour tout $n \geq 1$ l'implication

$$Y(\omega) = x \quad \Rightarrow \quad \forall 1 \leq k \leq n, \quad X_k(\omega) = a_k.$$

Cette implication se traduit par l'inclusion d'évènements

$$\{Y = x\} \subset \bigcap_{k=1}^{n} \{X_k = a_k\}.$$

Cette inclusion est vérifiée pour tout $n \geq 1$. Grâce à l'indépendance des X_k, on en déduit

$$\forall n \geq 1, \quad \mathbf{P}(Y = x) = \prod_{k=1}^{n} \mathbf{P}(X_k = a_k) = p^{s(n)}(1-p)^{n-s(n)}, \tag{32}$$

où l'on a noté $s(n) := \sum_{k=1}^{n} a_k$ (*i.e.* $s(n)$ est le nombre de chiffres 1 dans la suite finie $a_1, \dots, a_n$). Posons $r := \max(p, 1-p)$; comme p est compris strictement entre 0 et 1, on a aussi $1 - p < 1$ et $0 < r < 1$. Revenant à (32), on obtient

$$\forall n \geq 1, \quad \mathbf{P}(Y = x) \leq r^n,$$

d'où en faisant tendre n vers l'infini

$$\mathbf{P}(Y = x) = 0.$$

Comme x était quelconque dans C, on vient d'établir l'égalité $P_Y(\{x\}) = \mathbf{P}(Y = x) = 0$ pour tout $x \in C$. Cette égalité est vérifiée aussi pour tout réel $x \notin C$ puisqu'alors $\{x\} \subset \mathbb{R} \setminus C$ et que d'après la question 6, $P_Y(\mathbb{R} \setminus C) = 1 - P_Y(C) = 1 - 1 = 0$. Finalement $\mathbf{P}(Y = x) = 0$ pour tout $x \in \mathbb{R}$.

On en déduit que la fonction de répartition F de Y est continue sur $\mathbb{R}$. On vient ainsi de construire à (relativement) peu de frais une loi de variable aléatoire P_Y ayant une f.d.r. continue et pas de densité par rapport à la mesure de Lebesgue, ce que l'on appelle une *loi singulière* (en fait toute une famille de telles lois, paramétrée par $p \in]0,1[$).

8) Détermination de la loi de Y dans deux cas simples où les X_k sont dépendantes.

8. a) $X_k = X_1$ pour tout $k \geq 2$. Dans ce cas on voit immédiatement que :

$$Y = \sum_{k=1}^{+\infty} \frac{X_1}{4^k} = X_1 \sum_{k=1}^{+\infty} \frac{1}{4^k} = \frac{1}{3} X_1.$$

Ainsi Y est une variable aléatoire discrète ne prenant que les deux valeurs 0 et $1/3$, avec probabilités respectives $1-p$ et p.

8. b) X_1 et X_2 sont indépendantes et pour $k \geq 3$, $X_k = X_1$ si k est impair, $X_k = X_2$ si k est pair. Là encore Y est une variable aléatoire discrète ne pouvant prendre que quatre valeurs distinctes, correspondant aux 4 valeurs possibles prises par le couple (X_1, X_2). Plus précisément on a

$$Y = X_1 \sum_{j=0}^{+\infty} \frac{1}{4^{2j+1}} + X_2 \sum_{j=1}^{+\infty} \frac{1}{4^{2j}} = \frac{1}{4}\frac{1}{1-1/16}X_1 + \frac{1}{16}\frac{1}{1-1/16}X_2 = \frac{4}{15}X_1 + \frac{1}{15}X_2.$$

En utilisant l'indépendance de X_1 et X_2, on peut décrire complètement la loi de Y par le tableau suivant.

Valeur de (X_1, X_2)	Valeur y_k de Y	$\mathbf{P}(Y = y_k)$
$(0,0)$	0	$(1-p)^2$
$(0,1)$	$1/15$	$(1-p)p$
$(1,0)$	$4/15$	$p(1-p)$
$(1,1)$	$1/3$	p^2

9) Il est *impossible* de choisir les X_k, toujours de même loi Bern(p) de sorte que pour tout $x \in C$, $\mathbf{P}(Y = x)$ soit strictement positif. En effet si $\mathbf{P}(Y = x) > 0$, la fonction de répartition de Y présente une discontinuité au point x. Or on sait que l'ensemble des points de discontinuité d'une fonction de répartition est au plus dénombrable et d'après la question 4.b), l'ensemble C est infini non dénombrable.

10) Nous allons montrer que C est un *fermé* de $\mathbb{R}$ en prouvant que si (x_n) est une suite d'éléments de C convergente vers un réel x, alors x admet un développement en base 4 dont tous les chiffres soient des 0 ou des 1, d'où $x \in C$.

Rappel sur les développements en base 4. Le calcul déjà fait à la question 3) nous donne l'égalité $\sum_{k>j} 3.4^{-k} = 4^{-j}$. Cette égalité est la cause de l'existence de deux développements en base 4 pour les réels de $[0,1]$ de la forme $l4^{-n}$ avec l et n entiers. Le développement en base 4 est unique pour les réels qui ne sont pas de cette forme. Plus précisément, si $x = l4^{-n}$ avec l non divisible par 4, il admet :

– Un développement *propre*

$$x = \sum_{k=1}^{n} \frac{c_k}{4^k},$$

où les $c_k \in \{0,1,2,3\}$ sont donnés par le développement en base 4 de l'entier l et $c_n \neq 0$;

– Un développement *impropre*

$$\sum_{k=1}^{n} \frac{c'_k}{4^k} + \sum_{k=n+1}^{+\infty} \frac{3}{4^k}$$

avec $c'_k = c_k$ pour tout $k < n$ et $c'_n = c_n - 1$.

Par conséquent si un réel $x \in [0,1]$ admet un développement en base 4 de la forme $x = \sum_{k=1}^{+\infty} b_k 4^{-k}$ et si pour un certain j, $b_j < 3$, alors il y a *unicité* des $j-1$ premiers chiffres de x en base 4. En effet, ou bien x a un seul développement et c'est trivial, ou bien x en a deux et dans son développement impropre, la suite ininterrompue de chiffres 3 ne peut commencer qu'après le rang j.

L'idée de la preuve que nous allons donner de la fermeture de C est que si la suite (x_n) d'éléments de C converge vers un certain réel x, la suite des $j-1$ premiers chiffres de x_n devient constante à partir d'un certain rang et que ces $j-1$ premiers chiffres sont hérités par x.

Considérons d'abord pour $j \geq 1$ fixé, deux éléments y et y' de C tels que $|y-y'| < 4^{-j}$. Ils admettent des développements respectifs $\sum_{k=1}^{+\infty} a_k 4^{-k}$ et $\sum_{k=1}^{+\infty} a'_k 4^{-k}$ avec pour tout $k \geq 1$, $a_k, a'_k \in \{0,1\}$. Supposons qu'il existe un $k < j$ tel que $a_k \neq a'_k$ et posons $m := \min\{k < j;\ a_k \neq a'_k\}$. Quitte à échanger les rôles de y et y', on peut toujours supposer que $a_m = 1$ et $a'_m = 0$. On a alors

$$y = \sum_{k=1}^{m-1} \frac{a_k}{4^k} + \frac{1}{4^m} + \sum_{k>m} \frac{a_k}{4^k} \quad \text{et} \quad y' = \sum_{k=1}^{m-1} \frac{a_k}{4^k} + \frac{0}{4^m} + \sum_{k>m} \frac{a'_k}{4^k},$$

d'où l'égalité

$$\frac{1}{4^m} = \sum_{k>m} \frac{a_k - a'k}{4^k} - (y-y').$$

En notant que $a_k - a'_k$ ne peut prendre que les valeurs -1, 0 ou 1, que $|y-y'| < 4^{-j}$ et que $m+1 \leq j$, on en déduit la majoration

$$\frac{1}{4^m} < \sum_{k>m} \frac{a'_k}{4^k} + \frac{1}{4^j} = \frac{1}{3 \times 4^m} + \frac{1}{4^j} \leq \left(\frac{1}{3} + \frac{1}{4}\right) \frac{1}{4^m} < \frac{1}{4^m},$$

ce qui est absurde. Donc il n'existe pas de $k < j$ tel que $a_k \neq a'_k$, autrement dit les $j-1$ premiers chiffres de y et y' sont identiques.

Comme la suite (x_n) d'éléments de C converge vers x, on peut trouver pour tout entier $j \geq 1$, un entier n_j tel que

$$\forall n \geq n_j, \quad |x - x_n| < \frac{1}{2 \times 4^j},$$

d'où par inégalité triangulaire $|x_{n_j} - x_n| < 4^{-j}$. D'après le paragraphe précédent, on en déduit que x_{n_j} et x_n ont les mêmes $j-1$ premiers chiffres que nous noterons $a_{j,k}$, $1 \leq k < j$. Ainsi

$$\forall n \geq n_j, \quad x_n = \sum_{k=1}^{j-1} \frac{a_{j,k}}{4^k} + \varepsilon_{j,n},$$

avec les aj,k dans $\{0,1\}$ et $\varepsilon_{j,n} \leq \frac{1}{3} 4^{-j+1}$. En faisant tendre n vers l'infini et en notant que les $j-1$ premiers termes ne dépendent pas de n, la convergence de x_n vers x entraîne celle de $\varepsilon_{j,n}$ vers un certain ε_j et on obtient la décomposition

$$x = \sum_{k=1}^{j-1} \frac{a_{j,k}}{4^k} + \varepsilon_j, \quad \text{avec } \varepsilon_j \leq \frac{1}{3}\frac{1}{4^{j-1}} < \frac{1}{4^{j-1}}. \tag{33}$$

Le développement de ε_j en base 4 (qu'il soit propre ou impropre) a forcément ses $j-1$ premiers chiffres nuls car sinon $\varepsilon_j \geq 4^{-j+1}$, ce qui contredit l'inégalité ci-dessus. On peut donc écrire

$$\varepsilon_j = \sum_{k=j}^{+\infty} \frac{c_k}{4^k}.$$

Si $c_j = 3$, alors $\varepsilon_j \geq 3 \times 4^{-j}$, ce qui contredit l'inégalité $\varepsilon_j \leq \frac{4}{3}4^{-j}$, donc $c_j \leq 2$. On en déduit que dans le développement de x en base 4 donné par (33), les $j-1$ premiers chiffres sont uniques, que ce développement soit propre ou impropre.

Les $j-1$ premiers chiffres de tout développement de x en base 4 sont donc des 0 ou des 1. Comme ce raisonnement a été fait pour j quelconque, cela signifie que *tous* les chiffres de ce développment sont des 0 ou des 1, donc que $x \in C$. Au passage, on voit *a posteriori* que x ne peut avoir de développement impropre.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé du Devoir n° 3

Ex 1. Notons d'abord que pour $x \in]0,1[$ on a $|\frac{(\ln(x))^n}{\sqrt{1-x^2}}| = (-1)^n \frac{(\ln(x))^n}{\sqrt{1-x^2}}$ et donc si cette fonction est intégrable (convergente) sur $]0,1[$, elle est absolument convergente. Il y a deux problèmes : en 0 et en 1. On coupe donc l'intégrale en deux partie : $\int_0^{1/2}$ et $\int_{1/2}^1$.

Vu que $\lim_{x\downarrow 0}\sqrt{1-x^2} = 1$, les intégrales $\int_0^{1/2} \frac{(\ln(x))^n}{\sqrt{1-x^2}}\, dx$ et $\int_0^{1/2}(\ln(x))^n\, dx$ sont de même nature. En notant $I_n(x) = \int_1^x (\ln(y))^n\, dy$ une primitive de $(\ln(x))^n$, on trouve par une intégration par partie la relation $I_n(x) = x(\ln(x))^{n+1} - nI_{n-1}$. Vu que $I_0 = x - 1$, on démontre par récurrence la formule

$$I_n(x) = (-1)^{n+1}n! + x \sum_{k=0}^{n}(-1)^{n-k}\frac{n!}{k!}(\ln(x))^k.$$

En utilisant que $\lim_{x\downarrow 0} x(\ln(x))^k = 0$, on en déduit que $\lim_{\epsilon\downarrow 0}\int_\epsilon^{1/2}(\ln(x))^n\, dx$ existe.

Vu que $\lim_{x\uparrow 1}\frac{-\ln(x)}{1-x} = 1$, les intégrales $\int_{1/2} \frac{(\ln(x)^n}{\sqrt{1-x^2}}\, dx$ et $\int_{1/2}^1 (1-x)^{n-1/2}\, dx$ sont de même nature. Vu que $-(1-x)^{n+1/2}/(n+1/2)$ est une primitive de $(1-x)^{n-1/2}$ et que $n+1/2 > 0$, l'intégrale $\int_{1/2}^1(1-x)^{n-1/2}\, dx$ est convergente.

En rassemblant ces deux résultats, on conclut que $\int_0^1 \frac{(\ln(x))^n}{\sqrt{1-x^2}}\, dx$ est convergente et donc, par l'argument donné au début, elle est absolument convergente.

Pour $x \in]0,1/e[$ la suite $|\frac{(\ln(x))^n}{\sqrt{1-x^2}}|$ est croissante avec limite $+\infty$. En appliquant le théorème de Beppo Levi, on conclut que $\lim_{n\to\infty}\int_0^{1/e} |\frac{(\ln(x))^n}{\sqrt{1-x^2}}|\, dx = +\infty$.

Comme $\int_0^1 |\frac{(\ln(x))^n}{\sqrt{1-x^2}}|\, dx \geq \int_0^{1/e} |\frac{(\ln(x))^n}{\sqrt{1-x^2}}|\, dx$, on en déduit que $\lim_{n\to\infty}\int_0^1 |\frac{(\ln(x))^n}{\sqrt{1-x^2}}|\, dx = +\infty$. En utilisant de nouveau l'argument donné au début, on obtient que $\lim_{n\to\infty} u_{2n} = +\infty$ et $\lim_{n\to\infty} u_{2n+1} = -\infty$.

Ex 2.

1) $\forall x \geq 0 : 0 \leq f \cdot \mathbf{1}_{[\frac{x}{2},x]} \leq f$ et $\lim\limits_{x\to\infty} f \cdot \mathbf{1}_{[\frac{x}{2},x]} = 0$. Il en résulte, d'après la version continue du théorème de convergence dominée, que $\lim\limits_{x\to\infty} \int_{x/2}^{x} f\, d\lambda = 0$. De plus, comme f est décroissante, pour tout $x \geq 0$ on a $0 \leq \frac{x}{2} \cdot f(x) \leq \int_{x/2}^{x} f\, d\lambda$, d'où le résultat.

2) En général non. Considérons par example la fonction $f(x) = \frac{\mathbf{1}_{[2,\infty[}(x)}{x(\ln(x))^2}$. Alors, sachant que $\frac{-1}{\ln(x)}$ est une primitive de $\frac{1}{x(\ln(x))^2}$, on trouve $\int_0^A f(x)\, dx = \frac{1}{\ln(2)} - \frac{1}{\ln(A)}$. On en déduit que f est décroissante et λ-intégrable sur $[0,\infty[$, mais pour aucun $\alpha > 1$ a-t-on $\lim\limits_{x\to\infty} x^\alpha \cdot f(x) = 0$.

Problème

1) Pour $x \geq n, \forall k = 0, \ldots, n^2-1 : \mathbf{1}_{[\frac{k}{n},\frac{k+1}{n}[}(x) = 0$ et donc $\varphi_n(x) = \psi_n(x) = 0$. Pour $0 \leq x < n$ il existe un unique $k \in \{0, \ldots, n^2-1\}$ tel que $x \in [\frac{k}{n}, \frac{k+1}{n}[$ et donc $\varphi_n(x) = a_{n,k}$ et $\psi_n(x) = A_{n,k}$. Par définition de $a_{n,k}$ et $A_{n,k}$ on a les inégalités $\varphi_n(x) \leq f(x) \leq \psi_n(x)$, ce qui démontre la formule (7).

Soit $x \geq 0$ arbitraire. La continuité de f en x nous dit que $\forall \epsilon > 0 \exists \delta > 0 \forall y$: $|x-y| < \delta \implies |f(x) - f(y)| < \epsilon$. En choisissant, pour un ϵ donné, un $N \in \mathbb{N}$ tel que $N > x$ et $\frac{1}{N} < \delta$ (pour le δ donné par la continuité de f), et en choisissant pour chaque $n \geq N$ l'unique $k \in \{0, \ldots, n^2-1\}$ tel que $x \in [\frac{k}{n}, \frac{k+1}{n}[$, on obtient le résultat que $\forall \epsilon > 0 \exists N \forall n \geq N \forall y \in [\frac{k}{n}, \frac{k+1}{n}] : |f(x) - f(y)| < \epsilon$, simplement parce que $y \in [\frac{k}{n}, \frac{k+1}{n}[$ implique que $|x-y| \leq \frac{1}{n} \leq \frac{1}{N} < \delta$ (N.B. y appartient à l'intervalle fermé !). On en déduit que $|A_{n,k} - f(x)| < \epsilon$ et $|f(x) - a_{n,k}| < \epsilon$, c'est-à-dire $\forall \epsilon > 0 \exists N \forall n \geq N : |\psi_n(x) - f(x)| < \epsilon$ et $|f(x) - \varphi_n(x)| < \epsilon$. D'où le résultat (8).

2) La formule (9) est une conséquence immédiate du théorème de convergence dominée car les fonctions positives φ_n sont dominées par la fonction intégrable f.

3) En appliquant le lemme de Fatou on obtient $\infty = \int_0^\infty f\, d\lambda \leq \liminf\limits_{n\to\infty} \int_0^\infty \varphi_n\, d\lambda$. Comme on a toujours $\liminf\limits_{n\to\infty} \int_0^\infty \varphi_n\, d\lambda \leq \limsup\limits_{n\to\infty} \int_0^\infty \varphi_n\, d\lambda$, on obtient l'égalité

$$\liminf_{n\to\infty} \int_0^\infty \varphi_n\, d\lambda = \limsup_{n\to\infty} \int_0^\infty \varphi_n\, d\lambda = \infty,$$

et donc $\lim\limits_{n\to\infty} \int_0^\infty \varphi_n\, d\lambda = \infty = \int_0^\infty f\, d\lambda$.

4) On a par (7) l'inégalité $\int_0^\infty f \cdot \mathbf{1}_{[0,n[}\, d\lambda \leq \int_0^\infty \psi_n\, d\lambda$. Par hypothèse on a $\lim\limits_{n\to\infty} \int_0^\infty f \cdot \mathbf{1}_{[0,n[}\, d\lambda = \infty$, et donc $\lim\limits_{n\to\infty} \int_0^\infty \psi_n\, d\lambda = \infty = \int_0^\infty f\, d\lambda$.

5)

5a) La surface d'un triangle est la moitié du produit de la base et de la hauteur, donc $\int_0^\infty g\, d\lambda = \int_a^b g\, d\lambda = \frac{1}{2}h(b-a)$, ce qui se démontre aussi par un calcul direct de l'intégrale.

5b) Montrons d'abord que les triangles ont des bases disjointes, c'est-à-dire que $b_n \leq a_{n+1}$ pour tout $n \geq 1$. Un calcul direct montre $a_{n+1} - b_n = (\frac{1}{2(n+1)} - \frac{1}{4n^3}) + (\frac{1}{2(n+2)} - \frac{1}{4(n+1)^3})$. On constate que $\frac{1}{2(n+1)} - \frac{1}{4n^3} = \frac{2n^3-(n+1)}{4(n+1)n^3} \geq 0$ pour tout $n \geq 1$ et donc $b_n \leq a_{n+1}$ pour tout $n \geq 1$. En notant g_n la fonction triangulaire de base $[a_n, b_n]$ et de hauteur $4n$, on constate que $f = \lim_{N\to\infty} \sum_{n=1}^{N} g_n$. Le théorème de Beppo Levi s'applique et on obtient $\int_0^\infty f\,d\lambda = \sum_{n=1}^\infty \frac{1}{2} 4n(b_n - a_n) = \sum_{n=1}^\infty \frac{1}{n^2}$.

5c) Notons d'abord que $m_1 = \frac{5}{4}$, $m_2 = \frac{3}{5}$ et $m_{\ell+1} - m_\ell = \frac{1}{2}(\frac{1}{\ell+1} + \frac{1}{\ell+2}) < \frac{1}{\ell+1}$. Par récurrence on en déduit que $m_\ell < \ell$ pour tout $\ell \geq 2$. Soit maintenant $n \geq 1$ fixé. Pour $1 \leq \ell \leq n$ soit $k_\ell \in \{0, \ldots, n^2 - 1\}$ l'unique entier tel que $m_\ell \in [\frac{k_\ell}{n}, \frac{k_\ell + 1}{n}[$. On en déduit que $A_{n,k_\ell} \geq 4\ell$ (N.B. on n'est pas sûr qu'une partie (importante) du triangle suivant ne se trouve pas dans l'intervalle $[\frac{k_\ell}{n}, \frac{k_\ell+1}{n}[$, d'où seulement l'inégalité). L'inégalité $m_{\ell+1} - m_\ell = \frac{1}{2}(\frac{1}{\ell+1} + \frac{1}{\ell+2}) > \frac{1}{\ell+2}$ montre que pour $1 \leq \ell \leq n-2$ tous les k_ℓ sont distincts (car deux m_ℓ ne peuvent pas être dans un même intervalle de longueur $\frac{1}{n}$). Ceci démontre que la suite $(k_\ell)_{1\leq\ell\leq n-2}$ est une suite extraite de la suite $0, \ldots, n^2 - 1$. Vu que $\int_0^\infty \psi_n\,d\lambda = \sum_{k=0}^{n^2-1} A_{n,k}/n$, on obtient les inégalités suivantes : $\int_0^\infty \psi_n\,d\lambda \geq \sum_{\ell=1}^{n-2} A_{n,k_\ell}/n \geq \sum_{\ell=1}^{n-2} 4\ell/n = 2(n-2)(n-1)/n$, ce qui tend vers ∞ quand n tend vers ∞. Et donc $\lim_{n\to\infty} \int_0^\infty \psi_n\,d\lambda = \infty$.

6) Non. Si on prend $g = -f$ avec f la fonction définie en 5b), on a évidemment $\varphi_n^g = -\psi_n^f$ (où l'exposant indique par rapport à quelle fonction on définie les fonctions φ_n et ψ_n), et donc $\lim_{n\to\infty} \int_0^\infty \varphi_n^g\,d\lambda = -\infty \neq \int_0^\infty g\,d\lambda$.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé du Devoir n° 4

Ex 1. Soient X et Y deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et indépendantes : X est une variable aléatoire discrète à valeurs dans $\mathbb{N}$ dont la loi est donnée par $\mathbf{P}(X = k) = p_k$, $k \in \mathbb{N}$, et Y est de loi uniforme sur $[0, 1]$. On étudie la fonction de répartition F de XY.

Pour tout $x \in \mathbb{R}$, on a $F(x) = \mathbf{P}(XY \leq x)$. Comme XY est une variables aléatoire positive, il est clair que $F(x) = 0$ pour tout $x < 0$. Pour calculer $\mathbf{P}(XY \leq x)$ lorsque x est positif, l'idée est de conditionner par les valeurs possibles de X, de façon à pouvoir exploiter l'indépendance de X et Y.

Supposons dans un premier temps que $\mathbf{P}(X = k) \neq 0$ pour tout $k \in \mathbb{N}$. On dispose alors d'une partition de Ω en la suite $(\{X = k\})_{k\in\mathbb{N}}$ d'évènements de probabilité non nulle. La formule de conditionnement par les cas possibles (cf. [ICP], Chap. 2) s'écrit ici :

$$\mathbf{P}(XY \leq x) = \sum_{k\in\mathbb{N}} \mathbf{P}(XY \leq x \mid X = k)\mathbf{P}(X = k). \tag{34}$$

Le terme indexé par $k = 0$ se calcule directement :

$$\mathbf{P}(XY \leq x \mid X = 0)p_0 = \mathbf{P}(0 \times Y \leq x \mid X = 0)p_0 = \mathbf{P}(\Omega \mid X = 0)p_0 = p_0,$$

puisque $x \geq 0$. Pour $k > 0$, on a

$$\mathbf{P}(XY \leq x \mid X = k) = \mathbf{P}\Big(Y \leq \frac{x}{k} \mid X = k\Big) = \mathbf{P}\Big(Y \leq \frac{x}{k}\Big),$$

la deuxième égalité étant due à *l'indépendance* des variables aléatoires X et Y qui entraîne celle des évènements $\{Y \leq x/k\}$ et $\{X = k\}$. En reportant dans (34), on obtient

$$\forall x \geq 0, \quad \mathbf{P}(XY \leq x) = p_0 + \sum_{k\in\mathbb{N}^*} p_k \mathbf{P}\Big(Y \leq \frac{x}{k}\Big). \tag{35}$$

À ce stade nous n'avons pas encore utilisé notre connaissance de la loi de Y (sauf pour en déduire la positivité de Y). La formule (35) est donc valable pour n'importe quelle variable aléatoire positive Y indépendante de X.

Puisque Y suit la loi uniforme sur $[0, 1]$, on a pour tout $x \geq 0$,

$$\mathbf{P}\Big(Y \leq \frac{x}{k}\Big) = P_Y\big(]-\infty, x/k]\big) = \frac{\lambda\big(]-\infty, x/k] \cap [0, 1]\big)}{\lambda\big([0, 1]\big)} = \min\Big(1, \frac{x}{k}\Big).$$

On obtient donc l'expression suivante pour F :

$$\forall x \geq 0, \quad F(x) = p_0 + \sum_{k\in\mathbb{N}^*} p_k \min\Big(1, \frac{x}{k}\Big). \tag{36}$$

Pour établir (36), nous nous sommes restreints au cas où $p_k > 0$ pour tout $k \in \mathbb{N}$, afin d'éviter le conditionnement par un évènement de probabilité nulle. Il n'est pas difficile de lever cette restriction. Notons

$$I := \{k \in \mathbb{N};\ p_k > 0\}, \quad I^c := \mathbb{N} \setminus I, \quad \Omega' := \{\omega \in \Omega;\ X(\omega) \in I\}.$$

Il est clair que $\mathbf{P}(\Omega') = 1$, d'où $\mathbf{P}(XY \leq x) = \mathbf{P}(\{XY \leq x\} \cap \Omega')$. De plus la suite $(\{X = k\})_{k\in I}$ forme une partition de Ω' en évènements de probabilité non nulle. On en déduit

$$\begin{aligned}
\mathbf{P}(XY \leq x) = \mathbf{P}(\{XY \leq x\} \cap \Omega') &= \sum_{k\in I} \mathbf{P}(\{XY \leq x\} \cap \Omega' \mid X = k)\mathbf{P}(X = k) \\
&= \sum_{k\in I} \mathbf{P}(XY \mid X = k)\mathbf{P}(X = k),
\end{aligned}$$

en notant que si $k \in I$, $\{X = k\}$ est inclus dans Ω', ce qui explique la disparition de Ω' dans la dernière égalité. À partir de là, il est clair que l'on aboutit à l'analogue de (36), avec une indexation par I au lieu de $\mathbb{N}$. En fait la formule (36) elle même reste valable puisqu'elle ne contient plus de probabilités conditionnelles et que les termes indexés par I^c sont nuls.

La formule (36) permet de voir que F est continue sur $[0, +\infty[$. En effet, les fonctions $g_k : x \mapsto \min(1, x/k)$ sont continues sur $[0, +\infty[$ et uniformément bornées $0 \leq g_k \leq 1$. La série au second membre de (36) est convergente normalement (donc aussi uniformément) sur $[0, +\infty[$, puisque

$$\sum_{k=1}^{+\infty} \sup_{x\geq 0} |p_k g_k(x)| \leq \sum_{k=1}^{+\infty} p_k \leq 1 < +\infty.$$

Sa somme est donc bien une fonction continue sur $[0, +\infty[$. Finalement, F est continue sur $\mathbb{R}$ si $p_0 = 0$. Sinon elle a une seule discontinuité, de première espèce, au point 0 avec saut d'amplitude p_0.

Une autre information contenue dans (36) est que F est *affine* sur chaque intervalle $[j, j+1[$, $j \in \mathbb{N}$. En effet en remarquant que $g_k(x) = 1$ si $x \geq k$ et $g_k(x) = x/k$ si $x < k$, on obtient :

$$\forall x \in [j, j+1[, \quad F(x) = \sum_{k=0}^{j} p_k + x \sum_{k=j+1}^{+\infty} \frac{p_k}{k} = a_j + b_j x. \tag{37}$$

Le graphe de F sur $[0, +\infty[$ est donc une *ligne polygonale* avec sommets aux points d'abscisses entières. Ce graphe admet bien sûr comme asymptote la droite horizontale d'équation $y = 1$ (comme le graphe de n'importe quelle fonction de répartition). La

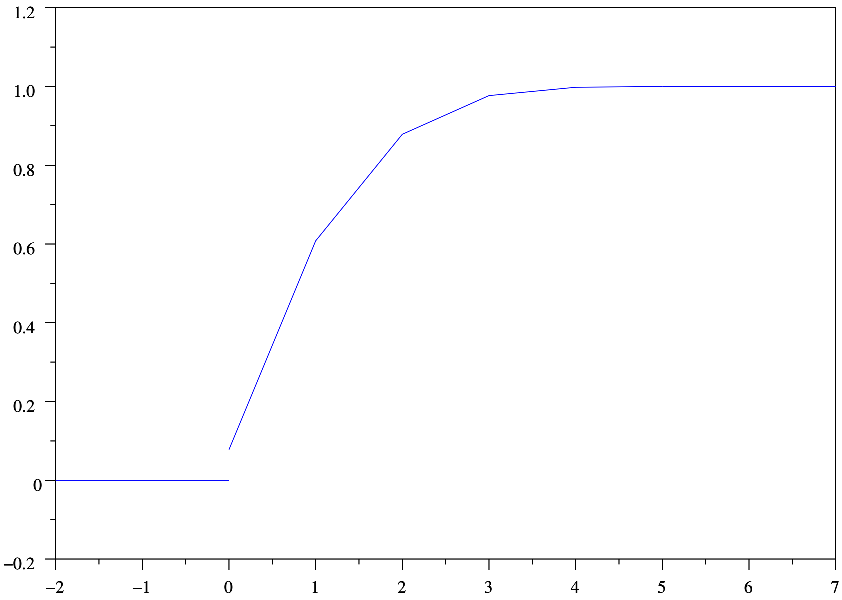

FIG. 8 – F pour $X \sim \mathrm{Bin}(5, 0.4)$

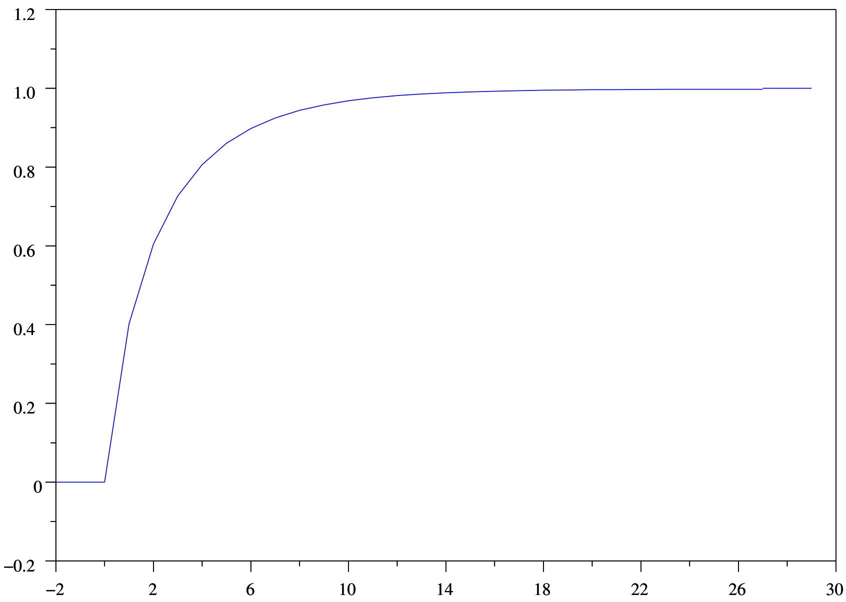

FIG. 9 – F pour $X \sim \mathrm{Geom}(0.2)$

pente du graphe entre j et $j+1$ est constante et vaut b_j. Comme la suite $(b_j)_{j\in\mathbb{N}}$ est décroissante, la restriction de F à $[0,+\infty[$ est une fonction *concave*[19].

Dans le cas où F est continue (*i.e.* si $p_0 = 0$), on peut se demander si la loi de XY admet une densité par rapport à la mesure de Lebesgue. La réponse est affirmative puisque F est alors C^1 par morceaux (Prop. 4.51 du cours). Dans le résultat démontré en cours, le nombre de « morceaux » était fini, ici il peut être infini dénombrable (les morceaux étant les intervalles $]j,j+1[$, $j\in\mathbb{N}$). Il facile d'étendre la preuve vue en cours à ce nouveau cas (vérification laissée en exercice). En écrivant F sous la forme

$$F(x) = \sum_{j=0}^{+\infty}(a_j + b_j x)\mathbf{1}_{[j,j+1[}(x), \quad (x\geq 0), \tag{38}$$

on obtient la densité $f = F'$ λ-p.p. en remarquant que F est dérivable sur $\mathbb{R}\setminus\mathbb{N}$ et que

$$\forall x\in\mathbb{R}_+\setminus\mathbb{N}, \quad F'(x) = \sum_{j=0}^{+\infty} b_j\mathbf{1}_{]j,j+1[}(x).$$

Il n'y a aucun problème pour dériver la série terme à terme, car pour x_0 fixé dans $\mathbb{R}_+\setminus\mathbb{N}$, il existe un unique $k\in\mathbb{N}$ tel que $k < x_0 < k+1$ et pour tous les x du voisinage $]k,k+1[$ de x_0, la série (38) a tous ses termes nuls sauf celui d'indice k, autrement dit, $F(x) = a_k + b_k x$ sur $]k,k+1[$.

Ex 2. Soit (X,Y) un vecteur aléatoire dont la loi possède une densité de la forme $f(x,y) = g(x^2+y^2)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$. Déterminer la loi de la variable aléatoire réelle T définie par

$$T(\omega) := \begin{cases} \frac{Y(\omega)}{X(\omega)} & \text{si } X(\omega)\neq 0,\\ 0 & \text{sinon.}\end{cases}$$

Complétons T en un vecteur aléatoire (S,T) de même dimension que (X,Y) par adjonction de la variable aléatoire $S := X$. On va chercher la loi de (S,T) par la méthode du changement de variable. La loi de T, deuxième loi marginale, s'obtiendra alors par intégration partielle de la densité du vecteur. On a ainsi $(S,T) = \varphi(X,Y)$, où φ est définie sur $\mathbb{R}^2$ par

$$(s,t) = \varphi(x,y) = \begin{cases}(x, y/x) & \text{si } x\neq 0,\\ (0,0) & \text{si } x = 0.\end{cases}$$

On vérifie facilement que φ restreinte à l'ouvert $W := \{(x,y)\in\mathbb{R}^2;\ x\neq 0\}$ est une bijection de cet ouvert sur lui même. Son inverse s'écrit

$$(x,y) = \varphi^{-1}(s,t) = (s,st), \quad (s,t)\in\varphi(W) = W.$$

19. Cette concavité peut d'ailleurs se lire directement sur la formule (36). En effet chaque g_k est concave sur $[0,+\infty[$, ce qui s'écrit $g_k(ta+(1-t)b)\geq tg_k(a)+(1-t)g_k(b)$ pour $0\leq a<b<+\infty$ et tout $t\in[0,1]$. Cette inégalité est héritée par la série $\sum_{k\geq 1} p_k g_k$ en raison de la positivité des p_k.

Le déterminant jacobien de φ^{-1} est donc

$$\mathrm{Jac}(\varphi^{-1})(s,t) = \begin{vmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ t & s \end{vmatrix} = s.$$

Les 4 dérivées partielles de φ^{-1} sont des fonctions continues de (s,t) et le jacobien ne s'annule en aucun point de W, on a donc bien un C^1 difféomorphisme de W sur W. Remarquons aussi que $\mathbf{P}\big((X,Y)\notin W\big) = \mathbf{P}(X=0) = 0$ parce que X est une variable aléatoire à densité. On aura donc pour toute fonction H mesurable positive,

$$\int_{\mathbb{R}^2} H \,\mathrm{d}P_{(X,Y)} = \int_W H\,\mathrm{d}P_{(X,Y)}. \tag{39}$$

Soit h borélienne $\mathbb{R}^2 \to \mathbb{R}_+$, calculons de deux façons $\mathbf{E}h(S,T)$. D'abord par transfert $\Omega \xrightarrow{(S,T)} \mathbb{R}^2$,

$$\mathbf{E}h(S,T) = \int_{\mathbb{R}^2} h(s,t)\,\mathrm{d}P_{(S,T)}(s,t). \tag{40}$$

D'autre part comme $\mathbf{E}h(S,T) = \mathbf{E}h\big(\varphi(X,Y)\big)$, le transfert $\Omega \xrightarrow{(X,Y)} \mathbb{R}^2$, (39) et l'expression de la densité f de (X,Y) nous donnent

$$\begin{aligned} \mathbf{E}h(S,T) &= \int_{\mathbb{R}^2} h\big(\varphi(x,y)\big)\,\mathrm{d}P_{(X,Y)}(x,y) \\ &= \int_W h\big(\varphi(x,y)\big)\,\mathrm{d}P_{(X,Y)}(x,y) \\ &= \int_W h\big(\varphi(x,y)\big) g(x^2+y^2)\,\mathrm{d}\lambda_2(x,y). \end{aligned} \tag{41}$$

Le changement de variable φ est un C^1 difféomorphisme de l'ouvert W sur lui même. Appliqué à l'intégrale (41), il conduit à

$$\mathbf{E}h(S,T) = \int_W h(s,t) g\big(s^2(1+t^2)\big)|s|\,\mathrm{d}\lambda_2(s,t). \tag{42}$$

Les égalités (40) et (42) étant valables pour toute h mesurable positive, leur comparaison montre que la loi $P_{(S,T)}$ du vecteur aléatoire (S,T) est la mesure à densité par rapport à λ_2 :

$$f_{S,T}:\ (s,t) \longmapsto g\big(s^2(1+t^2)\big)|s|\mathbf{1}_W(s,t).$$

La densité f_T de la loi marginale P_T s'en déduit par intégration partielle de $f_{S,T}$:

$$f_T(t) = \int_{\mathbb{R}} g\big(s^2(1+t^2)\big)|s|\mathbf{1}_W(s,t)\,\mathrm{d}\lambda_1(s) \tag{43}$$

En se souvenant que $W = \{(s,t)\in\mathbb{R}^2;\ s\neq 0\}$, on voit que $\mathbf{1}_W(s,t) = \mathbf{1}_{\mathbb{R}^*}(s)$ (quel que soit $t\in\mathbb{R}$), d'où

$$f_T(t) = \int_{]-\infty,0[} g\big(s^2(1+t^2)\big)|s|\,\mathrm{d}\lambda_1(s) + \int_{]0,+\infty[} g\big(s^2(1+t^2)\big)|s|\,\mathrm{d}\lambda_1(s). \tag{44}$$

Pour calculer ces deux intégrales, rappelons la formule de changement de variable pour une intégrale relative à λ_1. Si I est un intervalle ouvert de $\mathbb{R}$ et $\psi : I \to \mathbb{R}$ une application C^1 et strictement monotone sur I, alors pour toute g mesurable positive ou λ_1-intégrable sur $\psi(I)$, on a

$$\int_I g\big(\psi(s)\big)|\psi'(s)|\,\mathrm{d}\lambda_1(s) = \int_{\psi(I)} g(u)\,\mathrm{d}\lambda_1(u). \tag{45}$$

On peut appliquer cette formule avec $I =]-\infty, 0[$ puis $I =]0, +\infty[$ et $\psi(s) = cs^2$ où $c > 0$ est une constante. On obtient ainsi $\int_I g(cs^2)|s|\,\mathrm{d}\lambda_1(s) = (2c)^{-1}\int_{]0,+\infty[} g\,\mathrm{d}\lambda_1$. En revenant à (44), on voit qu'en prenant $c = 1+t^2$ (qui est bien une constante par rapport à la variable d'intégration s), on obtient :

$$f_T(t) = \frac{1}{1+t^2}\int_{]0,+\infty[} g\,\mathrm{d}\lambda_1. \tag{46}$$

Il nous reste à déterminer la constante $a := \int_{]0,+\infty[} g\,\mathrm{d}\lambda_1$. Une première façon serait d'écrire que f_T étant une densité de probabilité, $\int_{\mathbb{R}} f_T\,\mathrm{d}\lambda_1 = 1$, ce qui donne $a = \frac{1}{\pi}$ après un calcul élémentaire, utilisant l'intégrale de Riemann généralisée $\int_{-\infty}^{+\infty}(1+t^2)^{-1}\,\mathrm{d}t$. Il n'est pas superflu de chercher la valeur de a par la deuxième méthode, ne serait ce que pour détecter une éventuelle erreur de calcul. On écrit cette fois que $f(x,y) = g(x^2+y^2)$ étant une densité de probabilité par rapport à λ_2, on doit avoir

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}\lambda_2(x,y) = \int_{\mathbb{R}^2} g(x^2+y^2)\,\mathrm{d}\lambda_2(x,y) = 1.$$

Par passage en coordonnées polaires dans cette dernière intégrale, on obtient :

$$\begin{aligned}
1 = \int_{]0,+\infty[\times]0,2\pi[} g(r^2)r\,\mathrm{d}\lambda_2(r,\theta) &= \int_{]0,+\infty[} g(r^2)r\left\{\int_{]0,2\pi[}\mathrm{d}\lambda_1(\theta)\right\}\mathrm{d}\lambda_1(r)\\
&= 2\pi\int_{]0,+\infty[} g(r^2)r\,\mathrm{d}\lambda_1(r)\\
&= 2\pi\times\frac{1}{2}\int_{]0,+\infty[} g\,\mathrm{d}\lambda_1,
\end{aligned}$$

en utilisant à nouveau la formule de changement de variable (45) avec $\psi(r) = r^2$. On retrouve bien (avec soulagement) $\int_{]0,+\infty[} g\,\mathrm{d}\lambda_1 = 1/\pi$.

En conclusion la loi de la variable aléatoire T admet pour densité par rapport à λ_1 la fonction

$$f_T : \mathbb{R} \to \mathbb{R}_+, \quad t \mapsto \frac{1}{\pi(1+t^2)}.$$

T suit donc la loi de Cauchy.

Ex 3. Un vecteur aléatoire (X,Y) admet sur $\mathbb{R}^2$ une loi de densité

$$f(x,y) = x\mathrm{e}^{-(x+y)}\,\mathbf{1}_{\mathbb{R}_+}(x)\,\mathbf{1}_{\mathbb{R}_+}(y).$$

On pose $U = \min(X, Y)$, $V = \max(X, Y)$ et on cherche les densités des lois de (U, V), U, V ainsi que la valeur de $\mathrm{Cov}(U, V)$.

En préliminaire, notons que f est bien une densité de probabilité relativement à λ_2. En effet f est mesurable positive et en appliquant le théorème de Fubini pour les fonctions à variables séparées et la conversion de l'intégrale de Lebesgue en intégrale de Riemann pour une fonction continue d'intégrale généralisée absolument convergente, on obtient :

$$\begin{aligned}
\int_{\mathbb{R}^2} f \,\mathrm{d}\lambda_2 &= \left\{\int_{[0,+\infty[} x\mathrm{e}^{-x}\,\mathrm{d}\lambda_1(x)\right\}\left\{\int_{[0,+\infty[} \mathrm{e}^{-y}\,\mathrm{d}\lambda_1(y)\right\} \\
&= \left\{\int_0^{+\infty} x\mathrm{e}^{-x}\,\mathrm{d}x\right\}\left\{\int_0^{+\infty} \mathrm{e}^{-y}\,\mathrm{d}y\right\} \\
&= \left\{\left[-x\mathrm{e}^{-x}\right]_0^{+\infty} + \int_0^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x\right\}\left[-\mathrm{e}^{-y}\right]_0^{+\infty} = 1.
\end{aligned}$$

Le fait que f soit de la forme $f_1 \otimes f_2$ nous dit que X et Y sont des variables aléatoires indépendantes de densité respectives $x \mapsto c_1 x\mathrm{e}^{-x}\mathbf{1}_{\mathbb{R}_+}(x)$ et $y \mapsto c_2\mathrm{e}^{-y}\mathbf{1}_{\mathbb{R}_+}(y)$, les constantes c_1 et c_2 étant liées par $c_1c_2 = 1$. On voit immédiatement que $c_2 = 1$ en écrivant que $\int_{\mathbb{R}} c_2\mathrm{e}^{-y}\mathbf{1}_{\mathbb{R}_+}(y)\,\mathrm{d}\lambda_2(y) = 1$, donc c_1 vaut aussi 1.

1) *Calcul de la densité de la loi $P_{(U,V)}$ de (U, V).* Notons

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (u, v) = \varphi(x, y) = \big(\min(x, y), \max(x, y)\big).$$

Soit $h : \mathbb{R}^2 \to \mathbb{R}_+$ borélienne quelconque. On calcule de deux façons $\mathbf{E}h(U, V)$.

Première façon : par transfert $\Omega \xrightarrow{(U,V)} \mathbb{R}^2$.

$$\mathbf{E}h(U, V) = \int_\Omega h\big(U(\omega), V(\omega)\big)\,\mathrm{d}\mathbf{P}(\omega) = \int_{\mathbb{R}^2} h(u, v)\,\mathrm{d}P_{(U,V)}(u, v). \tag{47}$$

Deuxième façon : en notant que $\mathbf{E}h(U, V) = \mathbf{E}\varphi(X, Y)$, on a par le transfert $\Omega \xrightarrow{(X,Y)} \mathbb{R}^2$,

$$\mathbf{E}h(U, V) = \int_\Omega (h \circ \varphi)\big(X(\omega), Y(\omega)\big)\,\mathrm{d}\mathbf{P}(\omega) = \int_{\mathbb{R}^2} h\big(\varphi(x, y)\big)\,\mathrm{d}P_{(X,Y)}(x, y). \tag{48}$$

La mesure $P_{(X,Y)}$ ayant pour densité f par rapport à λ_2, on déduit de (48) l'égalité

$$\mathbf{E}h(U, V) = \int_{\mathbb{R}_+^2} h\big(\varphi(x, y)\big)x\mathrm{e}^{-(x+y)}\,\mathrm{d}\lambda_2(x, y). \tag{49}$$

On ne peut pas utiliser directement φ pour faire le changement de variable car elle n'est pas injective sur $\mathbb{R}_+^2$: tout point $(x, y) \in \mathbb{R}_+^2$ a même image par φ que son symétrique $(y, x) \in \mathbb{R}_+^2$. On introduit alors la découpage suivant

$$\mathbb{R}_+^2 = \Delta \cup \Delta' \cup D,$$

où

$$\Delta := \{(x,y)\in\mathbb{R}^2;\ 0<x<y\},\quad \Delta' := \{(x,y)\in\mathbb{R}^2;\ 0<y<x\}$$

et D est l'union des trois demi-droites $\{y=0, x\geq 0\}$, $\{x=y,\ x\geq 0\}$ et $\{x=0, y\geq 0$. Comme $\lambda_2(D)=0$, on obtient en effectuant ce découpage de l'ensemble d'intégration dans (49) :

$$\mathbf{E}h(U,V) = \int_\Delta h\big(\varphi(x,y)\big)x\mathrm{e}^{-(x+y)}\,\mathrm{d}\lambda_2(x,y) + \int_{\Delta'} h\big(\varphi(x,y)\big)x\mathrm{e}^{-(x+y)}\,\mathrm{d}\lambda_2(x,y). \tag{50}$$

L'intérêt de ce découpage est que la restriction de φ à chacun des ouverts Δ et Δ' est injective et d'expression simple. En effet,

$$\forall (x,y)\in\Delta,\quad \varphi(x,y)=(x,y),\quad \forall(x,y)\in\Delta',\quad \varphi(x,y)=(y,x).$$

Le changement de variable $(x,y)\mapsto(u,v)=\varphi(x,y)$ pour la première intégrale dans (50) est immédiat puisque la restriction de φ à Δ est l'identité. Il s'écrit :

$$\int_\Delta h\big(\varphi(x,y)\big)x\mathrm{e}^{-(x+y)}\,\mathrm{d}\lambda_2(x,y) = \int_\Delta h(u,v)u\mathrm{e}^{-(u+v)}\,\mathrm{d}\lambda_2(u,v). \tag{51}$$

Pour la deuxième intégrale dans (50), on considère la *bijection linéaire* $\psi : \mathbb{R}^2\to\mathbb{R}^2$, $(x,y)\mapsto(y,x)$ dont la restriction à Δ' coïncide avec φ. La matrice de ψ dans la base canonique est :

$$m(\psi)=\begin{pmatrix}0&1\\1&0\end{pmatrix}\qquad\text{donc}\qquad |\det\psi|=1.$$

Par la formule de changement de variable linéaire et bijectif on a donc en notant que $\psi(\Delta')=\Delta$ et que cette fois $x=v$ et $y=u$:

$$\int_{\Delta'} h\big(\varphi(x,y)\big)x\mathrm{e}^{-(x+y)}\,\mathrm{d}\lambda_2(x,y) = \int_\Delta h(u,v)v\mathrm{e}^{-(u+v)}\,\mathrm{d}\lambda_2(u,v). \tag{52}$$

En rassemblant (50), (51) et (52), on aboutit à

$$\begin{aligned}\mathbf{E}h(U,V) &= \int_\Delta h(u,v)(u+v)\mathrm{e}^{-(u+v)}\,\mathrm{d}\lambda_2(u,v)\\ &= \int_{\mathbb{R}^2} h(u,v)(u+v)\mathrm{e}^{-(u+v)}\mathbf{1}_\Delta(u,v)\,\mathrm{d}\lambda_2(u,v). \qquad (53)\end{aligned}$$

Les égalités (47) et (53) étant valables pour toute fonction borélienne positive h leur comparaison nous permet de conclure que la loi $P_{(U,V)}$ est la mesure admettant par rapport à λ_2 la densité

$$f_{(U,V)}(u,v)=(u+v)\mathrm{e}^{-(u+v)}\mathbf{1}_\Delta(u,v),$$

où Δ est l'ouvert $\{(u,v)\in\mathbb{R}^2;\ 0<u<v\}$.

2) *Calcul des densités f_U de P_U et f_V de P_V.* Les densités par rapport à λ_1 des lois marginales de (U, V) s'obtiennent par intégration partielle de $f_{(U,V)}$. Dans ces intégrations, nous utiliserons les identités

$$\forall (u, v) \in \mathbb{R}^2, \quad \mathbf{1}_\Delta(u, v) = \mathbf{1}_{\Delta_u}(v) = \mathbf{1}_{\Delta_v}(u).$$

Il convient donc de commencer par préciser les sections Δ_u et Δ_v. On voit immédiatement que :

$$\text{a)}\ \ \Delta_u = \begin{cases} \emptyset & \text{si } u \leq 0 \\ ]u, +\infty[& \text{si } u > 0 \end{cases} \qquad \text{b)}\ \ \Delta_v = \begin{cases} \emptyset & \text{si } v \leq 0 \\ ]0, v[& \text{si } v > 0. \end{cases} \tag{54}$$

Calcul de f_U. Pour tout $u \in \mathbb{R}$,

$$f_U(u) = \int_{\mathbb{R}} (u + v)\mathrm{e}^{-(u+v)} \mathbf{1}_\Delta(u, v) \,\mathrm{d}\lambda_1(v) = \int_{\mathbb{R}} (u + v)\mathrm{e}^{-(u+v)} \mathbf{1}_{\Delta_u}(v) \,\mathrm{d}\lambda_1(v).$$

D'après (54.a), cette intégrale vaut 0 si $u \leq 0$ et peut s'écrire pour $u > 0$ sous la forme

$$f_U(u) = \int_{]u,+\infty[} (u + v)\mathrm{e}^{-(u+v)} \,\mathrm{d}\lambda_1(v) = \int_u^{+\infty} (u + v)\mathrm{e}^{-(u+v)} \,\mathrm{d}v,$$

puisque l'intégrale de Riemann généralisée ainsi introduite est celle d'une fonction continue sur $]u, +\infty[$ et converge absolument. On achève le calcul en intégrant par parties :

$$f_U(u) = \left[-(u + v)\mathrm{e}^{-(u+v)}\right]_u^{+\infty} + \int_u^{+\infty} \mathrm{e}^{-(u+v)} \,\mathrm{d}v = 2u\mathrm{e}^{-2u} + \left[-\mathrm{e}^{-(u+v)}\right]_u^{+\infty} = (2u+1)\mathrm{e}^{-2u}.$$

Finalement, la densité f_U est donnée par

$$\forall u \in \mathbb{R}, \quad f_U(u) = (2u + 1)\mathrm{e}^{-2u} \mathbf{1}_{]0,+\infty[}(u). \tag{55}$$

Calcul de f_V. Pour tout $v \in \mathbb{R}$,

$$f_V(v) = \int_{\mathbb{R}} (u + v)\mathrm{e}^{-(u+v)} \mathbf{1}_\Delta(u, v) \,\mathrm{d}\lambda_1(u) = \int_{\mathbb{R}} (u + v)\mathrm{e}^{-(u+v)} \mathbf{1}_{\Delta_v}(u) \,\mathrm{d}\lambda_1(u).$$

D'après (54.b), cette intégrale vaut 0 si $v \leq 0$ et peut s'écrire pour $v > 0$ sous la forme

$$f_V(v) = \int_{]0,v[} (u + v)\mathrm{e}^{-(u+v)} \,\mathrm{d}\lambda_1(u) = \int_0^v (u + v)\mathrm{e}^{-(u+v)} \,\mathrm{d}u,$$

la conversion en intégrale de Riemann ordinaire étant justifiée par la continuité de $u \mapsto (u + v)\mathrm{e}^{-(u+v)}$ sur $[0, v]$. On achève le calcul en intégrant par parties :

$$\begin{aligned} f_V(v) = \left[-(u + v)\mathrm{e}^{-(u+v)}\right]_0^v + \int_0^v \mathrm{e}^{-(u+v)} \,\mathrm{d}u &= -2v\mathrm{e}^{-2v} + v\mathrm{e}^{-v} \left[-\mathrm{e}^{-(u+v)}\right]_0^v \\ &= (v + 1)\mathrm{e}^{-v} - (2v + 1)\mathrm{e}^{-2v}. \end{aligned}$$

Finalement, la densité f_V est donnée par

$$\forall v \in \mathbb{R}, \quad f_V(v) = \left((v + 1)\mathrm{e}^{-v} - (2v + 1)\mathrm{e}^{-2v}\right) \mathbf{1}_{]0,+\infty[}(v). \tag{56}$$

Remarque : L'étudiant consciencieux ne manquera pas ici de vérifier la positivité de f_V, par exemple en étudiant le signe de $g(v) := 1 + v - (2v + 1)\mathrm{e}^{-v}$ sur $[0, +\infty[$.

3) *Calcul de* $\mathrm{Cov}(U,V)$. Au vu des formules (55) et (56), il est clair que U et V sont de carré intégrable. La covariance de (U,V) est donc bien définie. Calculons la par la formule $\mathrm{Cov}(U,V)=\mathbf{E}(UV)-\mathbf{E}U\mathbf{E}V$. Pour éviter la répétition fastidieuse d'intégrations par parties, rappelons en préambule que

$$\forall n\in\mathbb{N},\qquad \int_0^{+\infty} t^n\mathrm{e}^{-t}\,\mathrm{d}t=\Gamma(n+1)=n!$$

Les intégrales par rapport à λ_1 apparaissant dans les calculs ci-dessous concernent à chaque fois des fonctions continues sur $[0,+\infty[$, d'intégrales généralisées de Riemann absolument convergentes sur $[0,+\infty[$. Ceci légitime la conversion de ces intégrales de Lebesgue en intégrales généralisées de Riemann.

Calcul de $\mathbf{E}U$ *et de* $\mathbf{E}V$.

$$\begin{aligned}\mathbf{E}U=\int_{\mathbb{R}} uf_U(u)\,\mathrm{d}\lambda_1(u) &= \int_{]0,+\infty[} u(2u+1)\mathrm{e}^{-2u}\,\mathrm{d}\lambda_1(u)\\ &= \int_0^{+\infty}(2u^2+u)\mathrm{e}^{-2u}\,\mathrm{d}u \quad (\text{on pose } t=2u)\\ &= \frac{1}{4}\int_0^{+\infty}(t^2+t)\mathrm{e}^{-t}\,\mathrm{d}t=\frac{1}{4}(2!+1!)=\frac{3}{4}.\end{aligned}$$

De même,

$$\begin{aligned}\mathbf{E}V=\int_{\mathbb{R}} vf_V(v)\,\mathrm{d}\lambda_1(v) &= \int_0^{+\infty} v(v+1)\mathrm{e}^{-v}\,\mathrm{d}v-\int_0^{+\infty} v(2v+1)\mathrm{e}^{-2v}\,\mathrm{d}v\\ &= 2!+1!-\frac{3}{4}=\frac{9}{4}.\end{aligned}$$

Calcul de $\mathbf{E}(UV)$. Puisque l'on connaît la densité de $P_{(U,V)}$, il est assez naturel de calculer $\mathbf{E}(UV)$ en utilisant le transfert $\Omega\xrightarrow{(U,V)}\mathbb{R}^2$, ce qui s'écrit ici :

$$\mathbf{E}(UV)=\int_\Omega U(\omega)V(\omega)\,\mathrm{d}\mathbf{P}(\omega)=\int_{\mathbb{R}^2} uvf_{(U,V)}(u,v)\,\mathrm{d}\lambda_2(u,v),$$

d'où

$$\mathbf{E}(UV)=\int_\Delta uv(u+v)\mathrm{e}^{-(u+v)}\,\mathrm{d}\lambda_2(u,v).$$

Cette intégrale se calcule ensuite grâce au théorème de Fubini (calcul laissé en exercice).

Il y a néanmoins ici une méthode plus astucieuse basée sur l'expression de UV à l'aide de X et Y :

$$UV=\min(X,Y)\max(X,Y)=XY.$$

Ceci nous permet d'exploiter *l'indépendance* de X et Y (cf. le paragraphe préliminaire avant la solution de la question 1)) pour écrire

$$\mathbf{E}(UV)=\mathbf{E}(XY)=\mathbf{E}X\mathbf{E}Y=\left\{\int_0^{+\infty}x^2\mathrm{e}^{-x}\,\mathrm{d}x\right\}\left\{\int_0^{+\infty}y\mathrm{e}^{-y}\,\mathrm{d}y\right\}=2!\,1!=2.$$

Finalement,

$$\mathrm{Cov}(U,V)=2-\frac{3}{4}\times\frac{9}{4}=\frac{5}{16}.$$

Ex 4. Soit $(X_n)_{n\geq 3}$ une suite de variables aléatoires discrètes définies sur le même $(\Omega, \mathcal{F}, \mathbf{P})$ et indépendantes. X_n prend les valeurs $-n$, n et 0 avec les probabilités :

$$\mathbf{P}(X_n = n) = \mathbf{P}(X_n = -n) = \frac{1}{2n\ln n} \quad \text{et} \quad \mathbf{P}(X_n = 0) = 1 - \frac{1}{n\ln n}.$$

1) Dans cette question, on montre que la suite $(M_n)_{n\geq 3}$ définie par :

$$M_n := \frac{1}{n}\sum_{i=3}^{n} X_i$$

converge vers 0 dans $L^2(\Omega)$ (et donc aussi en probabilité).

Commençons par calculer la variance de X_n, dont l'existence est évidente puisque X_n est une variable aléatoire bornée. On note d'abord que par symétrie de la loi de X_n,

$$\mathbf{E}X_n = -n\mathbf{P}(X_n = -n) + 0 \times \mathbf{P}(X_n = 0) + n\mathbf{P}(X_n = n) = 0.$$

Par conséquent $\operatorname{Var} X_n = \mathbf{E}(X_n^2)$, d'où

$$\operatorname{Var} X_n = (-n)^2\mathbf{P}(X_n = -n) + 0^2 \times \mathbf{P}(X_n = 0) + n^2\mathbf{P}(X_n = n) = \frac{n}{\ln n}.$$

Par indépendance des X_i,

$$\operatorname{Var} M_n = \frac{1}{n^2}\operatorname{Var}\left(\sum_{i=3}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=3}^{n}\operatorname{Var} X_i = \frac{1}{n^2}\sum_{i=3}^{n}\frac{i}{\ln i}.$$

Comme les X_n sont d'espérance nulle, il en est de même pour M_n, par linéarité de l'espérance. Par conséquent, $\operatorname{Var} M_n = \mathbf{E}(M_n^2) = \|M_n - 0\|^2_{\mathrm{L}^2(\Omega)}$ et pour prouver que M_n converge vers 0 dans $\mathrm{L}^2(\Omega)$, il suffit d'établir que

$$v_n := \frac{1}{n^2}\sum_{i=3}^{n}\frac{i}{\ln i} \xrightarrow[n\to+\infty]{} 0. \tag{57}$$

Remarquons d'abord que s'il n'y avait pas de logarithmes au dénominateur, la limite serait $1/2$ puisque pour tout entier $j \geq 1$,

$$\sum_{i=1}^{j} i = \frac{j(j+1)}{2}. \tag{58}$$

Si le terme général dans (57) s'écrivait $\frac{i}{\ln n}$ au lieu de $\frac{i}{\ln i}$, on déduirait de (58) que $v_n = O(1/\ln n)$ et la convergence (57) serait établie. Or on peut se ramener à cette situation, au moins pour le bloc de termes indexé par $n^{1/2} < i \leq n$ puisque $\ln i$ est alors minoré par $\ln(n^{1/2}) = \frac{1}{2}\ln n$. Ceci nous amène à proposer le découpage et les majorations suivants (noter que $\ln 3 > 1$).

$$\begin{aligned} v_n = \frac{1}{n^2}\sum_{3\leq i\leq n^{1/2}}\frac{i}{\ln i} + \frac{1}{n^2}\sum_{n^{1/2}<i\leq n}\frac{i}{\ln i} &\leq \frac{1}{n^2}\sum_{3\leq i\leq n^{1/2}} i + \frac{2}{n^2\ln n}\sum_{n^{1/2}<i\leq n} i \\ &\leq \frac{1}{n^2}\frac{n^{1/2}(n^{1/2}+1)}{2} + \frac{2}{n^2\ln n}\frac{n(n+1)}{2}. \end{aligned}$$

Après simplifications, on voit ainsi que $v_n = O(1/\ln n)$, ce qui implique la convergence (57).

La suite de variables aléatoires (M_n) converge donc dans $\mathrm{L}^2(\Omega)$ vers 0.

2) On montre en raisonnant par l'absurde que (M_n) ne peut converger presque sûrement vers une variable aléatoire réelle Y, c'est à dire vers une application mesurable $Y : \Omega \to \mathbb{R}$. Supposons qu'il existe une telle variable aléatoire Y. Alors en posant

$$\Omega' := \big\{\omega \in \Omega;\ M_n(\omega) \xrightarrow[n\to+\infty]{} Y(\omega)\big\},$$

on a $\mathbf{P}(\Omega') = 1$. Soit ω quelconque dans Ω'. En écrivant pour $n \geq 4$,

$$\frac{X_n(\omega)}{n} = \frac{1}{n}\sum_{i=3}^{n} X_i(\omega) - \frac{1}{n}\sum_{i=3}^{n-1} X_i(\omega) = M_n(\omega) - \frac{n-1}{n}M_{n-1}(\omega),$$

on voit que $n^{-1}X_n(\omega)$ converge vers $Y(\omega) - Y(\omega) = 0$ (noter que Y étant à valeurs dans $\mathbb{R}$, $Y(\omega)$ est *fini*). Comme $\mathbf{P}(\Omega') = 1$ et ω était quelconque dans Ω', on en déduit que $n^{-1}X_n$ converge presque sûrement vers 0. Pour montrer que cette dernière convergence est impossible, on invoque le deuxième lemme de Borel-Cantelli en remarquant que les évènements $A_n := \{X_n = n\}$ forment une suite indépendante et que

$$\sum_{n=3}^{+\infty} \mathbf{P}(A_n) = \sum_{n=3}^{+\infty} \frac{1}{2n\ln n} = +\infty.$$

On en déduit que

$$\mathbf{P}\big(\limsup_n A_n\big) = \mathbf{P}\big(\limsup_n \{X_n = n\}\big) = 1,$$

autrement dit que presque sûrement la suite $(n^{-1}X_n)_{n\geq3}$ prend une infinité de fois la valeur 1. Ceci interdit sa convergence presque sûre vers 0. Ainsi l'hypothèse de convergence presque sûre de M_n vers une variable aléatoire réelle Y conduit à une contradiction.

Remarque 1. Dans le raisonnement ci-dessus, la finitude de $Y(\omega)$ jouait un rôle clé. On peut donc légitimement se demander si M_n ne pourrait pas converger p.s. vers une variable aléatoire Y à valeurs dans $\overline{\mathbb{R}}$. Dans ce cas en appliquant le raisonnement ci-dessus pour les ω tels que $|Y(\omega)| < +\infty$, on voit que nécessairement $\mathbf{P}(|Y| < +\infty) = 0$, d'où la convergence presque sûre de $|M_n|$ vers $+\infty$. Ceci est encore impossible car d'après la question 1), $(|M_n|)_{n\geq3}$ possède une sous-suite $(|M_{n_k}|)_{k\geq1}$ qui converge presque sûrement vers 0. On devrait donc avoir sur un évènement Ω'' de probabilité 1 (donc non vide !),

$$\forall\omega \in \Omega'',\quad |M_{n_k}(\omega)| \xrightarrow[k\to+\infty]{} 0 \quad \text{et} \quad |M_{n_k}(\omega)| \xrightarrow[k\to+\infty]{} +\infty,$$

ce qui est absurde.

Remarque 2. Cet exemple de suite (M_n) vérifiant une loi faible des grands nombres (convergence en probabilité) mais pas la loi forte des grands nombres (convergence p.s.) permet de voir la finesse de la loi forte des grands nombres de Kolmogorov. D'après ce théorème, une condition suffisante pour que M_n converge p.s. vers 0 est que la série de terme général $n^{-2}\operatorname{Var} X_n$ converge. Or ici $n^{-2}\operatorname{Var} X_n = (n \ln n)^{-1}$. On voit qu'on manque la convergence de justesse. Si on considère pour un $\varepsilon > 0$ la suite de variables aléatoires discrètes indépendantes X'_n dont la loi est donnée par

$$\mathbf{P}(X'_n = n) = \mathbf{P}(X'_n = -n) = \frac{1}{2n(\ln n)^{1+\varepsilon}} \quad \text{et} \quad \mathbf{P}(X'_n = 0) = 1 - \frac{1}{n(\ln n)^{1+\varepsilon}},$$

la loi forte des grands nombres de Kolmogorov nous donne immédiatement la convergence p.s. vers 0 de la suite correspondante des moyennes arithmétiques M'_n.

Ex 5. *Deux applications de la l.f.g.n. pour des v.a. uniformes*

1) Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. définies sur les même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$, indépendantes, de même loi uniforme sur $[0,1]$. Rappelons que la loi uniforme sur $[0,1]$ a une fonction de répartition F donnée par

$$F(x) = \mathbf{P}(X_1 \in]-\infty, x)) = \frac{\lambda(]-\infty, x) \cap [0,1])}{\lambda([0,1]} = \begin{cases} 0 & \text{si } x \leq 0 \\ \min(x,1) & \text{si } x > 0. \end{cases} \tag{59}$$

et qu'elle admet pour densité par rapport à la mesure de Lebesgue λ la fonction $\mathbf{1}_[0,1]$. On pose $S_n := X_1 + \cdots + X_n$ et $V_n = X_1^2 + \cdots + X_n^2$. On définit la suite de variables aléatoires $(Z_n)_{n\geq 1}$ sur $(\Omega, \mathcal{F}, \mathbf{P})$ en posant :

$$Z_n(\omega) = \begin{cases} \frac{V_n(\omega)}{S_n(\omega)} & \text{si } S_n(\omega) \neq 0, \\ 0 & \text{si } S_n(\omega) = 0. \end{cases}$$

Remarquons tout de suite que $\mathbf{P}(S_n = 0) = 0$. En effet pour que S_n soit nul, il est nécessaire que l'un au moins des X_i $(1 \leq i \leq n)$ soit négatif ou nul, d'où $\{S_n = 0\} \subset \bigcup_{i=1}^{n} \{X_i \leq 0\}$. Par sous-additivité finie de $\mathbf{P}$, on en déduit

$$\mathbf{P}(S_n = 0) \leq \sum_{i=1}^{n} \mathbf{P}(X_i \leq 0) = 0,$$

puisque $\mathbf{P}(X_i \leq 0) = 0$ d'après (59). On en déduit que $\Omega_0 := \cap_{n\geq 1}\{S_n \neq 0\}$ est de probabilité 1 comme intersection dénombrable d'évènements de probabilité 1.

Pour étudier la limite de $\mathbf{E}Z_n$ quand n tend vers $+\infty$, on montre d'abord que Z_n converge presque sûrement vers une constante en écrivant $Z_n = (n^{-1}V_n)/(n^{-1}S_n)$ sur Ω_0 et en appliquant la loi forte des grands nombre au numérateur et au dénominateur. En suite on justifie l'interversion limite intégrale par convergence dominée. Voici les détails de ce programme.

Les X_i sont indépendantes, de même loi et $\mathbf{E}|X_1| < +\infty$ (car X_1 est bornée). Par la loi forte des grands nombres (cas i.i.d.) on a donc

$$\frac{1}{n} S_n \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbf{E}X_1 = \int_{\mathbb{R}} x \mathbf{1}_{[0,1]}(x)\, \mathrm{d}\lambda(x) = \int_0^1 x\, \mathrm{d}x = \frac{1}{2}. \tag{60}$$

Les X_i^2 héritent de l'indépendance des X_i, sont aussi de même loi et $\mathbf{E}(X_1^2) < +\infty$. La loi forte des grands nombres nous donne ici

$$\frac{1}{n}V_n \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbf{E}(X_1^2) = \int_{\mathbb{R}} x^2 \mathbf{1}_{[0,1]}(x)\,\mathrm{d}\lambda(x) = \int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}. \tag{61}$$

Définissons maintenant

$$\Omega_1 := \Big\{\omega \in \Omega;\ n^{-1}S_n(\omega) \xrightarrow[n\to+\infty]{} \frac{1}{2}\Big\}, \quad \Omega_2 := \Big\{\omega \in \Omega;\ n^{-1}V_n(\omega) \xrightarrow[n\to+\infty]{} \frac{1}{3}\Big\}.$$

D'après (60) et (61), Ω_1 et Ω_2 sont de probabilité 1. Par conséquent l'évènement $\Omega' := \Omega_0\cap\Omega_1\cap\Omega_2$ est de probabilité 1 comme intersection de trois évènements de probabilité 1. Par construction de Ω_0, Ω_1 et Ω_2, on a

$$\forall \omega \in \Omega', \quad Z_n(\omega) = \frac{V_n(\omega)}{S_n(\omega)} = \frac{n^{-1}V_n(\omega)}{n^{-1}S_n(\omega)} \xrightarrow[n\to+\infty]{} \frac{1/3}{1/2} = \frac{2}{3}.$$

Puisque $\mathbf{P}(\Omega') = 1$, nous venons ainsi d'établir la convergence presque sûre de Z_n vers la constante $2/3$.

Considérons maintenant $\Omega_3 := \bigcap_{i\geq 1}\{0 \leq X_i \leq 1\}$. Il est de probabilité 1 comme intersection *dénombrable* d'évènements de probabilité 1. Cet Ω_3 nous sert à vérifier la domination presque sûre de Z_n par la constante 1 en notant que

$$\forall \omega \in \Omega_3,\ \forall i \geq 1, \quad 0 \leq X_i(\omega) \leq 1, \text{ d'où } 0 \leq X_i(\omega)^2 \leq X_i(\omega).$$

Donc pour tout $\omega \in \Omega_3$ et tout $n \geq 1$, $0 \leq V_n(\omega) \leq S_n(\omega)$ d'où $0 \leq Z_n \leq 1$ (noter que sur $\Omega_3 \cap \Omega_0^c$, $Z_n = 0$ pour les n tels que $S_n = 0$). Nous avons ainsi montré la domination presque sûre de la suite (Z_n) par la constante (donc $\mathbf{P}$ intégrable) 1. Comme (Z_n) converge presque sûrement vers $2/3$, on peut conclure par le théorème de convergence dominée que :

$$\lim_{n\to+\infty} \mathbf{E}Z_n = \mathbf{E}\Big(\frac{2}{3}\Big) = \frac{2}{3}.$$

2) En gardant les notations de la question précédente, la loi de $(X_1, \dots, X_n)$ a pour densité $f_{X_1}^{\otimes n} = \mathbf{1}_{[0,1]^n}$ par rapport à λ_n. Comme f est continue sur le compact $[0,1]$, elle y est bornée. Pour $\omega \in \Omega_3$, $n^{-1}S_n(\omega) \in [0,1]$. Par continuité de f, $f\big(n^{-1}S_n(\omega)\big)$ converge vers $f(1/2)$ pour tout $\omega \in \Omega_1 \cap \Omega_3$ et cette convergence est dominée par la constante $\|f\|_\infty$ qui est $\mathbf{P}$-intégrable. En définissant $W_n := f(S_n/n)$ sur $\Omega_1\cap\Omega_3$ et $W_n := 0$ sur son complémentaire, on a par convergence dominée $\mathbf{E}W_n \xrightarrow[n\to+\infty]{} \mathbf{E}f(1/2) = f(1/2)$. Enfin en appliquant le théorème de transfert, on remarque que

$$\mathbf{E}W_n = \int_\Omega W_n\,\mathrm{d}\mathbf{P} = \int_{\Omega_3} f\Big(\frac{S_n}{n}\Big)\,\mathrm{d}\mathbf{P} = \int_{[0,1]^n} f\Big(\frac{x_1 + \cdots + x_n}{n}\Big)\,\mathrm{d}\lambda_n(x_1, \dots, x_n),$$

ce qui permet de traduire la convergence de $\mathbf{E}W_n$ vers $f(1/2)$ en le résultat d'analyse proposé par l'énoncé.

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé de l'examen du 30 janvier 2004

Ex 1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles, *de même loi*, telle que $\mathbf{E}|X_1| < +\infty$. On pose $Y_n := \min(X_n, n)$.

L'existence de $\mathbf{E}Y_n$ résulte de la *mesurabilité* de Y_n (comme min de deux applications mesurables) et de *l'intégrabilité* de X_1 via l'inégalité $|Y_n| \leq |X_n|$. Cette inégalité se vérifie comme suit. Si $X_n(\omega) < n$, on a $Y_n(\omega) = X_n(\omega)$ et $|Y_n(\omega)| = |X_n(\omega)|$. Si $X_n(\omega) \geq n$, on a $0 < Y_n(\omega) = n \leq X_n(\omega)$, d'où $|Y_n(\omega)| \leq |X_n(\omega)|$. En intégrant l'inégalité $|Y_n| \leq |X_n|$ sur Ω, on obtient $\mathbf{E}|Y_n| \leq \mathbf{E}|X_n| = \mathbf{E}|X_1| < +\infty$. L'égalité $\mathbf{E}|X_n| = \mathbf{E}|X_1|$ résulte du fait que X_n a même loi que X_1.

Maintenant que l'existence de $\mathbf{E}Y_n$ est assurée, l'égalité des lois de X_n et X_1 nous permet d'écrire (cf. remarque 1 ci-dessous) :

$$\forall n \in \mathbb{N}^*, \quad \mathbf{E}Y_n = \mathbf{E}\min(X_n, n) = \mathbf{E}\min(X_1, n). \tag{62}$$

Posons $Z_n := \min(X_1, n)$. Avec la même justification que ci-dessus, on a $|Z_n| \leq |X_1|$, ce qui montre que la *suite* $(Z_n)_{n\in\mathbb{N}^*}$ est *dominée* par la variable aléatoire intégrable $|X_1|$. D'autre part Z_n converge vers X_1 partout sur Ω quand n tend vers l'infini. En effet puisque X_1 est une variable aléatoire à valeurs dans $\mathbb{R}$, pour tout $\omega \in \Omega$, $X_1(\omega)$ est *fini* et on peut trouver un entier $n_0 = n_0(\omega) > X_1(\omega)$. Pour tout $n \geq n_0$, $Z_n(\omega) = X_1(\omega)$, la suite $\big(Z_n(\omega)\big)_{n\in\mathbb{N}^*}$ converge ainsi vers $X_1(\omega)$. Le théorème de convergence dominée appliqué à $(Z_n)_{n\in\mathbb{N}^*}$ et les égalités (62) nous permettent de conclure à la convergence de $\mathbf{E}Y_n$ vers $\mathbf{E}X_1$.

Remarque 1. Si X_n a même loi que X_1, $\mathbf{E}h(X_n) = \mathbf{E}h(X_1)$ pour toute application $h : \mathbb{R} \to \mathbb{R}$ qui est borélienne positive ou P_{X_1} intégrable sur $\mathbb{R}$. En effet grâce à l'égalité de mesures $P_{X_n} = P_{X_1}$, on peut écrire en utilisant deux fois le théorème de transfert :

$$\mathbf{E}h(X_1) = \int_\Omega h(X_1(\omega))\,\mathrm{d}\mathbf{P}(\omega) = \int_\mathbb{R} h(x)\,\mathrm{d}P_{X_1}(x) = \int_\mathbb{R} h(x)\,\mathrm{d}P_{X_n}(x) = \mathbf{E}h(X_n).$$

Cette propriété a été utilisée d'abord avec $h : x \mapsto |x|$, puis avec $h : x \mapsto \min(x, n)$.

Remarque 2. Le résultat démontré reste vrai avec X_1 variable aléatoire à valeurs dans $\overline{\mathbb{R}}$ au lieu de $\mathbb{R}$. En effet on sait qu'une telle variable aléatoire intégrable est finie presque partout sur Ω, donc que $Z_n(\omega)$ converge vers $X_1(\omega)$ presque partout sur Ω, ce qui suffit pour satisfaire la condition de convergence dans le théorème de convergence dominée. En fait on a encore convergence partout de Z_n vers X_1, car si $X_1(\omega) = +\infty$, alors $Z_n(\omega) = n$ pour tout n et la limite de $Z_n(\omega)$ est $+\infty$, ce qui vaut encore $X_1(\omega)$.

Ex 2. *Convergence d'espérances*

On note X une variable aléatoire de loi exponentielle de paramètre $a > 0$, ce qui signifie que le loi de X a une densité $f(t) = ae^{-at}\mathbf{1}_{\mathbb{R}^+}(t)$ par rapport à la mesure de Lebesgue λ. On définit la suite de variables aléatoires $(Y_n)_{n\geq 1}$ par

$$\forall n \in \mathbb{N}^*, \quad Y_n := \left(1 + \frac{X}{n}\right)^n \mathbf{1}_{\{X\geq 0\}}.$$

Ainsi Y_n est une variable aléatoire *positive* et l'expression $\mathbf{E}Y_n$ a toujours un sens comme élément de $\overline{\mathbb{R}}_+$. Le but de l'exercice est d'étudier sa limite quand n tend vers l'infini.

1) Par le transfert $\Omega \xrightarrow{X} \mathbb{R}$, on obtient en notant $P_X = \mathbf{P} \circ X^{-1}$ la loi de X :

$$\begin{aligned} \mathbf{E}Y_n = \int_\Omega Y_n(\omega)\,\mathrm{d}\mathbf{P}(\omega) &= \int_\Omega \left(1 + \frac{X(\omega)}{n}\right)^n \mathbf{1}_{\mathbb{R}_+}(X(\omega))\,\mathrm{d}\mathbf{P}(\omega) \\ &= \int_{\mathbb{R}} \left(1 + \frac{x}{n}\right)^n \mathbf{1}_{\mathbb{R}_+}(x)\,\mathrm{d}P_X(x) \\ &= \int_{[0,+\infty[} \left(1 + \frac{x}{n}\right)^n ae^{-ax}\,\mathrm{d}\lambda(x). \end{aligned} \tag{63}$$

2) La fonction $h :]-1, +\infty[\to \mathbb{R}$, $u \mapsto v = \ln(1+u)$ est *concave* car pour tout $u > -1$, $h''(u) = -(1+u)^{-2} < 0$. Il en résulte que la courbe d'équation $v = h(u)$ est en tout point au dessous de sa tangente.

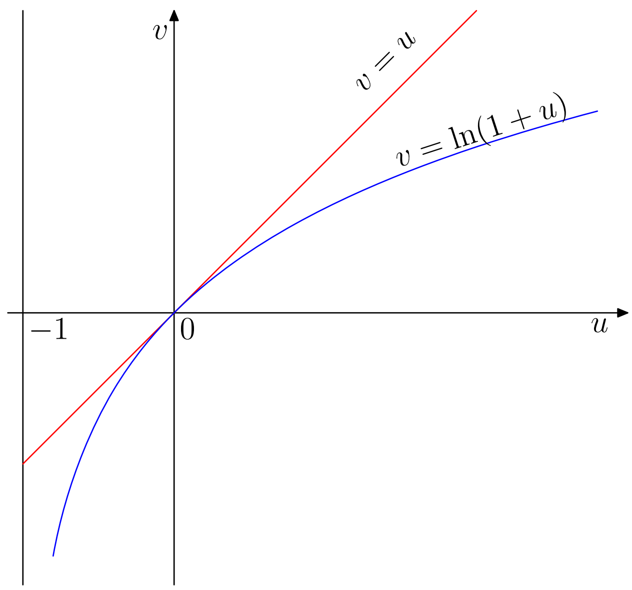

FIG. 10 – Courbe $v = \ln(1+u)$ et sa tangente en 0

Ceci est vrai en particulier au point d'abscisse $u_0 = 0$ et se traduit alors par l'inégalité

$$\forall u \in]-1, +\infty[, \quad \ln(1+u) \leq u. \tag{64}$$

En prenant $u = x/n$, avec $x \geq 0$ quelconque dans (64), on en déduit $n\ln(1 + x/n) \leq x$, d'où par croissance de la fonction exponentielle $\exp\bigl(n\ln(1 + x/n)\bigr) \leq \exp(x)$, ce qui s'écrit encore :

$$\forall x \geq 0, \quad \Bigl(1 + \frac{x}{n}\Bigr)^n \leq \mathrm{e}^x. \tag{65}$$

3) On suppose que $a > 1$. Considérons la suite des fonctions mesurables positives f_n définies sur $[0, +\infty[$ par $f_n(x) := (1 + x/n)^n a\mathrm{e}^{-ax}$. D'après (63), $\mathbf{E}Y_n = \int_{\mathbb{R}_+} f_n \,\mathrm{d}\lambda$. Nous allons établir la convergence de $\mathbf{E}Y_n$ en montrant que $(f_n)_{n\geq 1}$ vérifie les hypothèses du théorème de convergence dominée relativement à l'espace mesuré $(\mathbb{R}_+, \mathrm{Bor}(\mathbb{R}_+), \lambda)$.

a) *Convergence de $(f_n)_{n\geq 1}$ λ-p.p. sur $\mathbb{R}_+$.* Il est bien connu[20] que $(1+x/n)^n$ converge vers e^x quand n tend vers $+\infty$. Par conséquent la suite de fonctions mesurables $(f_n)_{n\geq 1}$ converge sur tout $\mathbb{R}_+$ vers la fonction $f_\infty : x \mapsto a\mathrm{e}^{(1-a)x}$.

b) *Domination de $(f_n)_{n\geq 1}$ par g λ-intégrable.* D'après (65), on a pour tout $x \in \mathbb{R}_+$, $0 \leq f_n(x) \leq a\mathrm{e}^{(1-a)x}$. La suite $(f_n)_{n\geq 1}$ est donc *dominée* par la fonction $g = f_\infty$. Cette fonction est *continue et positive* sur $\mathbb{R}_+$, on a donc $\int_{\mathbb{R}_+} g \,\mathrm{d}\lambda = \int_0^{+\infty} g(x)\,\mathrm{d}x$ (égalité dans $\overline{\mathbb{R}}_+$). Or l'intégrale de Riemann généralisée $\int_0^{+\infty} g(x)\,\mathrm{d}x$ est convergente, comme le montre le calcul élémentaire suivant :

$$\begin{aligned}
\int_0^{+\infty} a\mathrm{e}^{(1-a)x}\,\mathrm{d}x = \lim_{b\to+\infty}\int_0^b a\mathrm{e}^{(1-a)x}\,\mathrm{d}x &= \lim_{b\to+\infty}\Bigl[\frac{a}{1-a}\mathrm{e}^{(1-a)x}\Bigr]_0^b \\
&= \frac{a}{1-a}\bigl(\lim_{b\to+\infty}\mathrm{e}^{(1-a)b} - 1\bigr) \\
&= \frac{a}{a-1} < +\infty,
\end{aligned}$$

en notant que $\mathrm{e}^{(1-a)b}$ tend vers 0 quand b tend vers $+\infty$ parce que $1 - a < 0$. On en déduit que g est λ-intégrable sur $\mathbb{R}_+$.

Ainsi les deux hypothèses du théorème de convergence dominée sont satisfaites par $(f_n)_{n\geq 1}$ et on peut intervertir limite et intégrale :

$$\lim_{n\to+\infty}\int_{\mathbb{R}_+} f_n\,\mathrm{d}\lambda = \int_{\mathbb{R}_+} f_\infty\,\mathrm{d}\lambda = \frac{a}{a-1}.$$

On en conclut que

$$\forall a > 1, \quad \lim_{n\to+\infty}\mathbf{E}Y_n = \frac{a}{a-1}.$$

4) Dans le cas $0 < a \leq 1$, on remarque que $\int_0^{+\infty} f_\infty(x)\,\mathrm{d}x = +\infty$. C'est évident directement si $a = 1$ puisqu'alors f_∞ est la fonction constante égale à 1 sur $\mathbb{R}_+$. Dans le cas $0 < a < 1$, cela résulte du calcul ci-dessus puisqu'alors $1 - a > 0$ et $\mathrm{e}^{(1-a)b}$ tend vers $+\infty$ quand b tend vers $+\infty$. On a donc $\int_{\mathbb{R}_+} f_\infty\,\mathrm{d}\lambda = +\infty$. En combinant cette remarque avec le lemme de Fatou, il vient :

$$+\infty = \int_{\mathbb{R}_+} f_\infty\,\mathrm{d}\lambda = \int_{\mathbb{R}_+}\liminf_{n\to+\infty} f_n\,\mathrm{d}\lambda \leq \liminf_{n\to+\infty}\int_{\mathbb{R}_+} f_n\,\mathrm{d}\lambda = \liminf_{n\to+\infty}\mathbf{E}Y_n.$$

20. Sinon prendre le logarithme et utiliser $\ln(1+x/n) \sim x/n$ pour voir que si $x > 0$, $n\ln(1+x/n) \sim x$, ce qui lève la forme indéterminée, le cas $x = 0$ étant évident directement.

Ainsi $\liminf_{n\to+\infty} \mathbf{E}Y_n = +\infty$, ce qui nous permet de conclure que

$$\forall a \in]0,1], \quad \lim_{n\to+\infty} \mathbf{E}Y_n = +\infty.$$

Remarque. Il y a moyen d'unifier les solutions proposées ci-dessus dans les cas $a > 1$ et $0 < a \leq 1$ en utilisant le théorème de Beppo Levi au lieu du théorème de convergence dominée et du lemme de Fatou. Pour cela nous allons vérifier la *croissance de la suite* $(f_n)_{n\geq 1}$ de fonctions mesurables positives, ce qui se réduit à celle de la suite $(g_n)_{n\geq 1}$, en posant $g_n(x) := (1 + x/n)^n$. La fonction logarithme néperien étant croissante, la croissance de la suite $(g_n)_{n\geq 1}$ équivaut à celle de $(\ln g_n)_{n\geq 1}$. Une façon commode d'établir cette dernière croissance est d'utiliser la propriété suivante des fonctions *concaves*. Soit h concave sur un intervalle I de $\mathbb{R}$, $\mathcal{C}$ la représentation graphique de h et $a < b$ deux réels de I. Alors la corde de $\mathcal{C}$ entre les points d'abscisses a et b a une *pente* qui, a étant fixé, est une fonction *décroissante* de b. Cette pente varie ainsi en sens contraire de b, donc croît lorsque b se rapproche de a. En appliquant ceci à la fonction concave $h : u \mapsto \ln(1+u)$ avec $a = 0$ et $b = x/n$ pour $x > 0$ fixé, la pente $p_n(x)$ de la corde entre a et b s'écrit :

$$p_n(x) = \frac{\ln(1 + x/n)}{x/n} = \frac{1}{x} \ln g_n(x).$$

La croissance de la suite de pentes $\big(p_n(x)\big)_{n\geq 1}$ implique ainsi celle de la suite $\big(\ln g_n(x)\big)_{n\geq 1}$

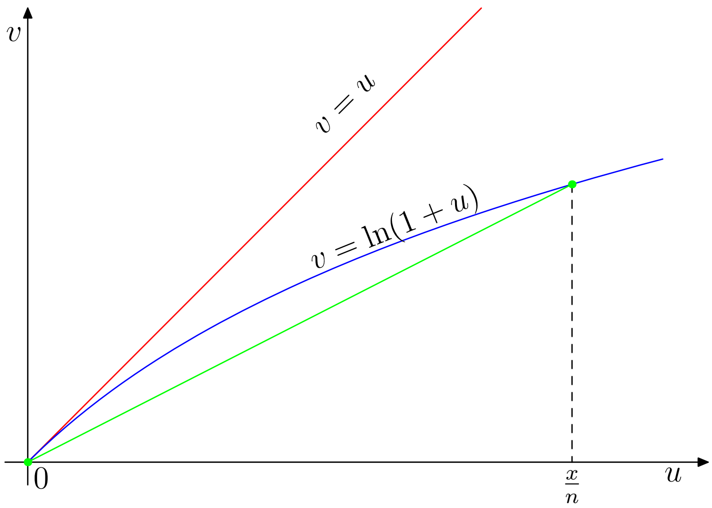

FIG. 11 – Courbe $v = \ln(1 + u)$ et sa corde entre 0 et x/n

puis celle de $\big(g_n(x)\big)_{n\geq 1}$ et ce résultat est valable pour tout $x > 0$. Le cas $x = 0$ est évident directement puisque $\big(g_n(0)\big)_{n\geq 1}$ est la suite constante de valeur 1. Nous avons donc établi la croissance de la suite de fonctions mesurables positives $(f_n)_{n\geq 1}$. La convergence de cette suite en tout point de $\mathbb{R}_+$ a été établie à la question 3, on peut appliquer le théorème de Beppo Levi et obtenir :

$$\mathbf{E}Y_n = \int_{\mathbb{R}_+} f_n \,\mathrm{d}\lambda \uparrow \int_{\mathbb{R}_+} f_\infty \,\mathrm{d}\lambda = \int_0^{+\infty} a\mathrm{e}^{(1-a)x} \,\mathrm{d}x = \begin{cases} a/(a-1) & \text{si } a > 1, \\ +\infty & \text{si } 0 < a \leq 1. \end{cases}$$

Ex 3. *Carrés aléatoires*

Soit $(X_i)_{i\in\mathbf{N}^*}$ une suite de variables aléatoires, définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$, à valeurs dans $]0,1[$, indépendantes et de même loi uniforme sur $]0,1[$. On pose $S_n = \sum_{i=1}^n X_i$ et on note M_i le *point aléatoire* de coordonnées (S_i, S_i) dans un repère orthonormé du plan. On pose $M_0 = O$ origine du repère. Pour $i \in \mathbf{N}^*$, on note C_i le *carré aléatoire* ouvert de diagonale $M_{i-1}M_i$.

1) Le carré aléatoire C_i ayant pour côté $S_i - S_{i-1} = X_i$ a pour aire $\lambda_2(C_i) = X_i^2$.

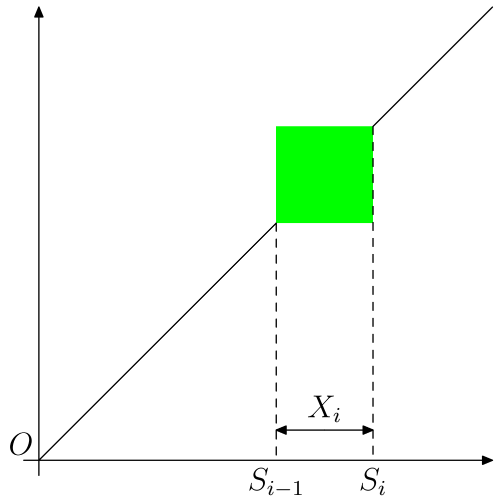

FIG. 12 – Carré aléatoire C_i d'aire $\lambda_2(C_i) = X_i^2$

Les C_i étant deux à deux disjoints (on a pris les carrés ouverts), on a par additivité de λ_2,

$$\frac{1}{n}\lambda_2(E_n) = \frac{1}{n}\sum_{i=1}^n X_i^2. \tag{66}$$

Les X_i sont indépendantes et de même loi. Les X_i^2 héritent de ces deux propriétés puisque $X_i^2 = h(X_i)$ où h est la fonction [21] borélienne $x \mapsto x^2$. De plus la variable X_1 étant bornée ($0 \le X_1 \le 1$) a des moments de tout ordre d'où en particulier $\mathbf{E}X_1^2 < +\infty$. La suite $(X_i^2)_{i\ge1}$ vérifie donc toutes les hypothèses de la loi forte des grands nombres de Khintchine, d'où

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbf{E}X_1^2. \tag{67}$$

Pour calculer cette limite, on rappelle que la loi uniforme sur $]0,1[$ a pour densité par rapport à λ_1 la fonction $\mathbf{1}_{]0,1[}$. Par transfert $\Omega \xrightarrow{X_1} \mathbb{R}$, on en déduit

$$\mathbf{E}X_1^2 = \int_\Omega X_1(\omega)^2 \,\mathrm{d}\mathbf{P}(\omega) = \int_{\mathbb{R}} x^2 \,\mathrm{d}P_X(x) = \int_{\mathbb{R}} x^2 \mathbf{1}_{]0,1[}(x) \,\mathrm{d}\lambda_1(x) = \int_{]0,1[} x^2 \,\mathrm{d}\lambda_1(x). \tag{68}$$

21. L'indépendance serait conservée même pour des $h_i(X_i)$ avec h_i borélienne. Par contre avec des h_i dépendant de i, on ne pourrait plus affirmer que les $h_i(X_i)$ ont même loi.

La fonction $h : x \mapsto x^2$ étant continue sur l'intervalle *fermé* $[0,1]$, ses intégrales au sens de Lebesgue (relativement à λ_1) et de Riemann coïncident, d'où

$$\int_{]0,1[} x^2 \,\mathrm{d}\lambda_1(x) = \int_0^1 x^2 \,\mathrm{d}x = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}. \tag{69}$$

En rassemblant (66)–(69), on conclut que

$$\frac{1}{n}\lambda_2(E_n) \xrightarrow[n\to+\infty]{\text{p.s.}} \frac{1}{3}.$$

2) Soit $a \in]0,1[$ et Δ_a la droite d'équation $y = x + a$. On se propose de calculer la probabilité $\mathbf{P}(\Delta_a \cap C_i \neq \emptyset)$ et d'en déduire la loi de T_n, nombre de C_i $(1 \leq i \leq n)$ ayant une intersection non vide avec Δ_a.

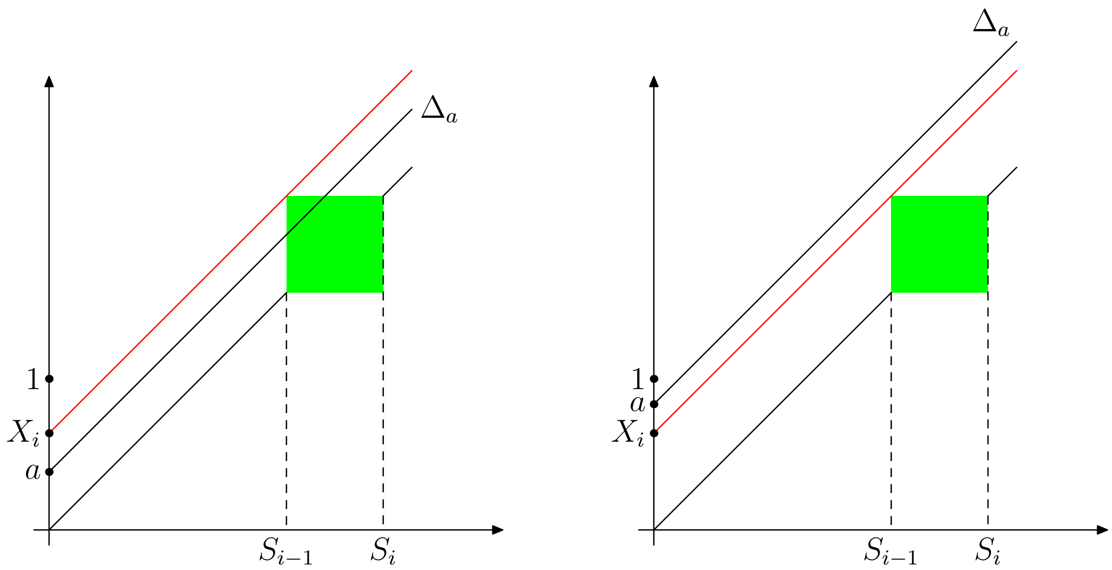

FIG. 13 – $C_i \cap \Delta_a \neq \emptyset$ si et seulement si $X_i > a$

Pour cela on commence par remarquer que Δ_a intersecte le carré C_i si et seulement si $a < X_i$, voir figure 13. En rappelant que X_i suit la loi uniforme sur $]0,1[$, on en déduit :

$$\mathbf{P}(\Delta_a \cap C_i \neq \emptyset) = \mathbf{P}(X_i > a) = \frac{\lambda_1(]a,+\infty[\cap]0,1[)}{\lambda_1(]0,1[)} = \lambda_1(]a,1[) = 1 - a.$$

L'égalité d'évènements $\{\Delta_a \cap C_i \neq \emptyset\} = \{X_i > a\}$ permet d'écrire

$$T_n = \sum_{i=1}^n Y_i \quad \text{avec} \quad Y_i := \mathbf{1}_{\{X_i > a\}}.$$

Les Y_i s'écrivent $Y_i = h(X_i)$ où h est la fonction mesurable $\mathbf{1}_{]a,+\infty[}$. Elles héritent donc de l'indépendance des X_i. De plus les Y_i sont des variables aléatoires de Bernoulli de même paramètre $p = \mathbf{E}Y_i = \mathbf{P}(X_i > a) = 1 - a$. Leur somme T_n suit donc la loi binomiale de paramètres n et $p = 1 - a$. Autrement dit,

$$T_n(\Omega) = \{0, 1, \ldots, n\} \quad \text{et} \quad \forall k \in T_n(\Omega),\ \mathbf{P}(T_n = k\} = C_n^k (1-a)^k a^{n-k}.$$

Problème

Ce problème est consacré à l'algorithme du rejet. La méthode du rejet (appelée aussi d'acceptation-rejet) peut être décrite abstraitement comme suit. On suppose que l'on sait générer un vecteur aléatoire M de $\mathbb{R}^d$ suivant une certaine loi μ. On génère alors l'un après l'autre les vecteurs aléatoires $M_1, \ldots, M_n, \ldots$ mutuellement indépendants et de même loi que M, en s'arrêtant au premier d'entre eux qui vérifie une certaine condition $(\mathcal{H}_0)$. Soit T l'indice (aléatoire) correspondant. On a ainsi fabriqué un vecteur (doublement) aléatoire M_T. Comme T est aléatoire, la loi de ce vecteur n'est pas celle de M, c'est une nouvelle loi ν. Si la simulation de M et le test de $(\mathcal{H}_0)$ sont facilement programmables, on dispose ainsi d'une méthode pour générer un vecteur aléatoire de loi ν.

Préliminaire

Soit B un borélien de $\mathbb{R}^d$ tel que $0 < \lambda_d(B) < +\infty$, λ_d désignant la mesure de Lebesgue sur $\mathbb{R}^d$. On dit que le vecteur aléatoire $M : \Omega \to \mathbb{R}^d$ suit la loi uniforme sur B, notation $M \sim \mathrm{Unif}(B)$, si sa loi P_M est donnée par

$$\forall A \in \mathrm{Bor}(\mathbb{R}^d),\ P_M(A) = \mathbf{P}(M \in A) = \frac{\lambda_d(A \cap B)}{\lambda_d(B)}. \tag{70}$$

On rappelle que cette condition équivaut à l'existence pour la mesure P_M d'une densité de la forme $c\mathbf{1}_B$ par rapport à λ_d, avec une constante $c > 0$.

1) Soit M un vecteur aléatoire de loi uniforme sur B et φ est une *bijection linéaire* $\mathbb{R}^d \to \mathbb{R}^d$. Vérifions que $\varphi(M)$ suit la loi uniforme sur $\varphi(B)$. On sait (cf. cours, chapitre 5) que $\lambda_d \circ \varphi^{-1} = c\lambda_d$, où la constante $c = |\det(\varphi^{-1})|$ est strictement positive. Soit A' un borélien quelconque de $\mathbb{R}^d$. Puisque M suit la loi uniforme sur B, on a :

$$\mathbf{P}\big(\varphi(M) \in A'\big) = \mathbf{P}\big(M \in \varphi^{-1}(A')\big) = \frac{\lambda_d\big(\varphi^{-1}(A') \cap B\big)}{\lambda_d(B)} \tag{71}$$

Posons $B' := \varphi(B)$. Comme φ est *injective*, on a $B = \varphi^{-1}(B')$, voir la justification ci-dessous. On peut alors écrire $\varphi^{-1}(A') \cap B = \varphi^{-1}(A') \cap \varphi^{-1}(B') = \varphi^{-1}(A' \cap B')$. En reportant dans (71), on obtient

$$\mathbf{P}\big(\varphi(M) \in A'\big) = \frac{(\lambda_d \circ \varphi^{-1})(A' \cap B')}{(\lambda_d \circ \varphi^{-1})(B')} = \frac{c\lambda_d(A' \cap B')}{c\lambda_d(B')} = \frac{\lambda_d(A' \cap B')}{\lambda_d(B')},$$

ce qui montre que $\varphi(M)$ suit la loi uniforme sur B', puisque A' était quelconque.

Justification de $\varphi^{-1}(B') = B$. Par définition de l'inverse *ensembliste* φ^{-1} et de l'image ensembliste $B' = \varphi(B)$, on a

$$\varphi^{-1}(B') = \{x \in \mathbb{R}^d;\ \varphi(x) \in B'\} = \{x \in \mathbb{R}^d;\ \exists y \in B,\ \varphi(x) = \varphi(y)\}. \tag{72}$$

Comme φ est injective, l'égalité $\varphi(x) = \varphi(y)$ équivaut à $x = y$. En reportant dans (72), on obtient

$$\varphi^{-1}(B') = \{x \in \mathbb{R}^d;\ \exists y \in B,\ x = y\} = B.$$

Remarque. Les seules propriétés de φ utilisées dans cette démonstration sont $\lambda_d \circ \varphi^{-1} = c\lambda_d$ et l'injectivité de φ. On en déduit immédiatement que si h est une injection mesurable $\mathbb{R}^d \to \mathbb{R}^d$ telle que $\lambda_d \circ h^{-1} = c\lambda_d$ pour une certaine constante $c > 0$, alors $h(M)$ suit la loi uniforme sur $h(B)$. Ceci s'applique notamment au cas où h est une bijection affine $\mathbb{R}^d \to \mathbb{R}^d$, c'est-à-dire la composée d'une bijection linéaire φ avec une translation τ.

Partie 1

On suppose désormais que $(M_i)_{i\in\mathbb{N}^*}$ est une suite de vecteurs aléatoires $\Omega \to \mathbb{R}^d$, mutuellement *indépendants* et de *même loi* μ telle que $\mu(B) > 0$. On assimilera pour la commodité du langage le vecteur $M_i(\omega)$ à un point de $\mathbb{R}^d$.

2) Pour tout $\omega \in \Omega$, on pose

$$T(\omega) := \inf\{i \in \mathbb{N}^*;\ M_i(\omega) \in B\},$$

avec la convention $\inf \emptyset = +\infty$. T est donc le numéro du premier point « tombé » dans B. Nous allons montrer que T est une variable aléatoire discrète suivant la loi géométrique de paramètre $p = \mu(B)$. On note $q := 1 - p$.

Commençons par justifier la mesurabilité de T, application de Ω dans $\overline{\mathbb{N}^*} = \mathbb{N}^* \cup \{+\infty\}$. Les tribus concernées par cette mesurabilité sont $\mathcal{F}$ et $\mathcal{P}(\overline{\mathbb{N}^*})$. En raison de la dénombrabilité de $\overline{\mathbb{N}^*}$, il suffit de vérifier que

$$\forall k \in \overline{\mathbb{N}^*}, \quad T^{-1}(\{k\}) \in \mathcal{F}. \tag{73}$$

Si $k \in \mathbb{N}^*$, on a

$$\{T = k\} = \{\forall i < k,\ M_i \notin B \text{ et } M_k \in B\} = \Bigg(\bigcap_{1\le i<k} M_i^{-1}(B^c)\Bigg) \cap M_k^{-1}(B), \tag{74}$$

tandis que dans le cas particulier $k = +\infty$,

$$\{T = +\infty\} = \{\forall i \in \mathbb{N}^*,\ M_i \notin B\} = \bigcap_{i\in\mathbb{N}^*} M_i^{-1}(B^c). \tag{75}$$

Les M_i étant des vecteurs aléatoires dans $\mathbb{R}^d$, donc mesurables $\mathcal{F}$-Bor$(\mathbb{R}^d)$, l'image réciproque par M_i d'un borélien quelconque de $\mathbb{R}^d$ est un élément de $\mathcal{F}$. Il résulte alors de (74) et de (75) que $T^{-1}(\{k\})$ est intersection finie ou dénombrable d'éléments de $\mathcal{F}$, ce qui établit (73). Ainsi T est bien une variable aléatoire discrète à valeurs dans $\overline{\mathbb{N}^*}$.

Sa loi est caractérisée par les $\mathbf{P}(T=k)$. Posons $p=\mu(B)=\mathbf{P}(M_i\in B)$. Si $k\in\mathbb{N}^*$, (74) et l'indépendance des M_i nous donnent

$$\forall k\in\mathbb{N}^*,\quad \mathbf{P}(T=k)=(1-p)^{k-1}p. \tag{76}$$

Calculons $\mathbf{P}(T=+\infty)$. Par hypothèse $\mu(B)>0$, donc $1-p=\mu(B^c)=\mathbf{P}\big(M_i^{-1}(B^c)\big)$ est strictement inférieur à 1. En remarquant que pour tout $n\in\mathbb{N}^*$ on a l'inclusion $\cap_{i\in\mathbb{N}^*}M_i^{-1}(B^c)\subset\cap_{i\leq n}M_i^{-1}(B^c)$, on a par indépendance $\mathbf{P}(T=+\infty)\leq(1-p)^n$, d'où en faisant tendre n vers l'infini,

$$\mathbf{P}(T=+\infty)=0. \tag{77}$$

Compte-tenu de (76) et (77), la loi de T peut s'écrire

$$P_T=\sum_{k\in\mathbb{N}^*}pq^{k-1}\delta_k,$$

c'est donc bien la loi géométrique de paramètre $p=\mu(B)$.

3) Calculons maintenant $\mathbf{P}(M_k\in A\cap B \text{ et } T=k)$ pour $k\in\mathbb{N}^*$ et $A\in\operatorname{Bor}(\mathbb{R}^d)$. Comme $A\cap B$ est inclus dans B, on a clairement

$$\begin{aligned}\{M_k\in A\cap B \text{ et } T=k\} &= \{\forall i<k,\ M_i\notin B \text{ et } M_k\in A\cap B\}\\ &= \Big(\bigcap_{1\leq i<k}M_i^{-1}(B^c)\Big)\cap M_k^{-1}(A\cap B),\end{aligned}$$

d'où par indépendance des M_i,

$$\mathbf{P}(M_k\in A\cap B \text{ et } T=k)=q^{k-1}\mathbf{P}\big(M_k^{-1}(A\cap B)\big)=q^{k-1}\mu(A\cap B). \tag{78}$$

4) On définit $M_T:\Omega\to\mathbb{R}^d$ par

$$M_T(\omega):=\begin{cases}M_{T(\omega)}(\omega) & \text{si } T(\omega)<+\infty,\\ 0 & \text{si } T(\omega)=+\infty.\end{cases}$$

Cette définition équivaut à

$$M_T=\sum_{k\in\overline{\mathbb{N}}^*}M_k\mathbf{1}_{\{T=k\}},$$

ce qui montre que M_T est mesurable $\mathcal{F}$-$\operatorname{Bor}(\mathbb{R}^d)$, c'est donc bien un *vecteur aléatoire*. Pour trouver sa loi, calculons $\mathbf{P}(M_T\in A)$ pour A borélien quelconque de $\mathbb{R}^d$. En partitionnant Ω par les évènements $\{T=k\}$, $k\in\overline{\mathbb{N}}^*$, on a

$$\mathbf{P}(M_T\in A)=\sum_{k\in\overline{\mathbb{N}}^*}\mathbf{P}(M_T\in A \text{ et } T=k). \tag{79}$$

En raison de l'inclusion $\{M_T\in A \text{ et } T=+\infty\}\subset\{T=+\infty\}$ et de (77), le terme indexé par $k=+\infty$ dans cette « série » est nul. Pour calculer le terme général indexé par $k \in \mathbb{N}^*$, exprimons l'évènement $E_{A,k} := \{M_T \in A \text{ et } T = k\}$ à l'aide de M_k et de T. Il est commode de raisonner sur un élément quelconque ω de $E_{A,k}$:

$$\begin{aligned} \omega \in E_{A,k} \quad &\Leftrightarrow \quad M_{T(\omega)}(\omega) \in A \text{ et } T(\omega) = k \\ &\Leftrightarrow \quad M_k(\omega) \in A \text{ et } T(\omega) = k \text{ et } M_k(\omega) \in B \\ &\Leftrightarrow \quad M_k(\omega) \in A \cap B \text{ et } T(\omega) = k. \end{aligned}$$

On voit ainsi que

$$E_{A,k} = \{M_k \in A \cap B \text{ et } T = k\}.$$

D'après (78), la probabilité de cet évènement est $q^{k-1}\mu(A \cap B)$. En reportant ce résultat dans (79), on obtient (en notant [22] que $0 \leq q < 1$)

$$\mathbf{P}(M_T \in A) = \sum_{k \in \mathbb{N}^*} q^{k-1}\mu(A \cap B) = \mu(A \cap B) \sum_{j \in \mathbb{N}} q^j = \frac{\mu(A \cap B)}{1-q}.$$

En se souvenant que $1 - q = p = \mu(B)$, on aboutit à

$$\mathbf{P}(M_T \in A) = \frac{\mu(A \cap B)}{\mu(B)} = \mu(A \mid B). \tag{80}$$

Ceci étant valable pour tout borélien A de $\mathbb{R}^d$, on en conclut que la loi du vecteur aléatoire M_T est la mesure de probabilité conditionnelle $\nu := \mu(\,.\, \mid B)$.

5) Soient B et C deux boréliens tels que $B \subset C$ et $0 < \lambda_d(B) < \lambda_d(C) < +\infty$ et soit μ la loi uniforme sur C. Alors pour tout $A \in \mathrm{Bor}(\mathbb{R}^d)$, (80) s'écrit :

$$\mathbf{P}(M_T \in A) = \frac{\mu(A \cap B)}{\mu(B)} = \frac{\frac{\lambda_d(A \cap B \cap C)}{\lambda_d(C)}}{\frac{\lambda_d(B \cap C)}{\lambda_d(C)}} = \frac{\lambda_d(A \cap B \cap C)}{\lambda_d(B \cap C)} = \frac{\lambda_d(A \cap B)}{\lambda_d(B)},$$

ce qui montre que M_T suit la loi uniforme sur B.

6) On dispose seulement d'un générateur de nombres au hasard capable de simuler une suite aussi longue que l'on veut $U_1, \ldots, U_\ell$ de variables aléatoires indépendantes de loi uniforme sur $[0, 1]$. On s'intéresse au coût de l'algorithme du rejet pour la simulation d'un « échantillon » $M_{T_1}, \ldots, M_{T_n}$ de vecteurs aléatoires *indépendants* [23] de même loi que M_T. Ce coût sera mesuré par le nombre S_n de variables U_i utilisées.

On suppose dans cette question que B est *borné*. On peut alors choisir C de la forme $[a_1, b_1] \times \cdots \times [a_d, b_d]$.

6 a) Le vecteur aléatoire $M = (X_1, \ldots, X_d)$ défini par

$$X_j := a_j + (b_j - a_j)U_j, \quad j = 1, \ldots, d,$$

22. Comme $\mu(B) > 0$, $q = 1 - \mu(B) < 1$. Soit dit en passant, on peut supposer aussi que $\mu(B) < 1$ et donc que $q > 0$ sans perdre de généralité. En effet si $\mu(B) = 1$, on est sûr que M_i va tomber dans B dès la première tentative, autrement dit, $\mathbf{P}(M_T = M_1) = 1$. Dans ce cas l'algorithme du rejet est sans intérêt puisque M_T a même loi que M_1 que l'on sait simuler *a priori*.

23. On admettra qu'il suffit pour cela de lancer n fois l'algorithme du rejet et que les variables aléatoires $T_1, \ldots, T_n$ sont indépendantes.

suit la loi uniforme sur C. Pour le voir on remarque que $M = h(U)$, avec $U := (U_1, \dots, U_d)$ et $h = \tau \circ \varphi$, où l'on a noté

$$\varphi : \mathbb{R}^d \to \mathbb{R}^d, \quad (u_1, \dots, u_d) \longmapsto (x_1, \dots, x_d) = \big((b_1 - a_1)u_1, \dots, (b_d - a_d)u_d\big)$$

et

$$\tau : \mathbb{R}^d \to \mathbb{R}^d, \quad (x_1, \dots, x_d) \longmapsto (a_1, \dots, a_d) + (x_1, \dots, x_d).$$

La translation τ est une bijection mesurable vérifiant $\lambda_d \circ \tau^{-1} = \lambda_d$. De son côté, l'application linéaire φ est *bijective* car l'hypothèse $\lambda_d(B) > 0$ et l'inclusion $B \subset [a_1, b_1] \times \dots \times [a_d, b_d]$ impliquent qu'aucun $(b_j - a_j)$ ne peut être nul. Ainsi l'application $h = \tau \circ \varphi$ est une bijection mesurable telle que $\lambda_d \circ h^{-1} = c\lambda_d$ avec $c > 0$. Elle transforme l'hypercube $[0, 1]^d$ en le pavé $C = [a_1, b_1] \times \dots \times [a_d, b_d]$. D'après la remarque faite à la question 1, il nous suffit alors de vérifier que U suit la loi uniforme sur $[0, 1]^d$ pour en déduire que $M = h(U)$ suit la loi uniforme sur $C = h(U)$. Les U_i de même loi uniforme sur $[0, 1]$ ont chacune pour densité $f_i := \mathbf{1}_{[0,1]}$ par rapport à λ_1. Par indépendance des U_i, le vecteur $U = (U_1, \dots, U_d)$ a pour densité par rapport à λ_d :

$$f_U = f_1 \otimes \dots \otimes f_d = \mathbf{1}_{[0,1]} \otimes \dots \otimes \mathbf{1}_{[0,1]} = \mathbf{1}_{[0,1]^d},$$

ce qui est bien la densité de la loi uniforme sur $[0, 1]^d$.

6 b) Pour générer M_{T_k}, il faut générer T_k vecteurs M_i et la fabrication de chaque M_i consomme d variables aléatoires U_j uniformes sur $[0, 1]$. La simulation de l'échantillon $M_{T_1}, \dots, M_{T_n}$ consomme donc en tout

$$S_n = (T_1 + \dots + T_n)d$$

variables aléatoires de loi uniforme sur $[0, 1]$.

Les variables aléatoires T_k sont indépendantes et de même loi géométrique de paramètre

$$p = \mu(B) = \frac{\lambda_d(B \cap C)}{\lambda_d(C)} = \frac{\lambda_d(B)}{\lambda_d(C)} = \frac{1}{\tau}.$$

Elles ont une espérance finie $\mathbf{E}T_1 = 1/p = \tau$. Par la loi forte des grands nombres de Khintchine, on a donc

$$\frac{S_n}{n} \xrightarrow[n \to +\infty]{\text{p.s.}} \tau d.$$

6 c) Si l'on mesure le coût de l'algorithme par cette limite τd, il est clair que la minimisation de ce coût revient à celle de τ.

Puisque $B \subset C$, on a $\tau \geq 1$. Il est clair que le choix $\tau = 1$ n'est pas réaliste (cf. note n$^\text{o}$ 22 au bas de la page 85) puisqu'alors l'algorithme du rejet serait inutile. En fait il s'agit ici de minimiser τ sur la classe des pavés $C = [a_1, b_1] \times \dots \times [a_d, b_d]$ contenant B. On peut voir que le minimum de τ est atteint pour le pavé $C_0 = [\alpha_1, \beta_1] \times \dots \times [\alpha_d, \beta_d]$ où les α_j, β_j sont définis comme suit. On note π_j la projection sur la j^e coordonnée :

$$\pi_j : (x_1, \dots, x_j, \dots, x_d) \longmapsto x_j, \quad 1 \leq j \leq d$$

et on pose

$$\alpha_j := \inf\{\pi_j(x);\ x \in B\}, \quad \beta_j := \sup\{\pi_j(x);\ x \in B\}, \quad 1 \leq j \leq d.$$

Pour tout $j = 1, \dots, d$, l'inclusion $B \subset C = [a_1, b_1] \times \cdots \times [a_d, b_d]$ implique $\pi_j(B) \subset \pi_j(C) = [a_j, b_j]$. De cette dernière inclusion entre parties de $\mathbb{R}$, on déduit les inégalités $\inf \pi_j(B) \geq a_j$ et $\sup \pi_j(B) \leq b_j$, d'où $[\alpha_j, \beta_j] \subset [a_j, b_j]$ et finalement $C_0 \subset C$. D'autre part il est clair que C_0 contient B puisque si x est un élément quelconque de B, en l'écrivant $x = (\pi_1(x), \dots, \pi_d(x))$, on voit que pour tout $j = 1, \dots, d$, on a $\alpha_j \leq \pi_j(x) \leq \beta_j$ et donc que x appartient à C_0.

Dans le cas particulier où $d = 2$ et B est le disque unité, on voit immédiatement que C_0 est le carré $[-1, 1]^2$ circonscrit au disque unité. Sur cet exemple la valeur minimale de τ est donc $\lambda_2(C_0)/\lambda_2(B) = 4/\pi \simeq 1{,}273$.

Partie 2

On s'intéresse maintenant à la simulation de vecteurs aléatoires de loi uniforme sur un borélien B de $\mathbb{R}^2$ dont la frontière a pour équation en coordonnées polaires $r = \cos(2t)$. On peut découper B en quatre « pétales » B_0, B_1, B_2, B_3, où

$$B_0 := \left\{(x, y) \in \mathbb{R}^2;\ 0 \leq r \leq \cos(2t),\ -\pi/4 \leq t \leq \pi/4\right\},$$

et les autres pétales se déduisent de B_0 par rotation de centre O et d'angles $\pi/2$, π, $3\pi/2$, cf. figure 14.

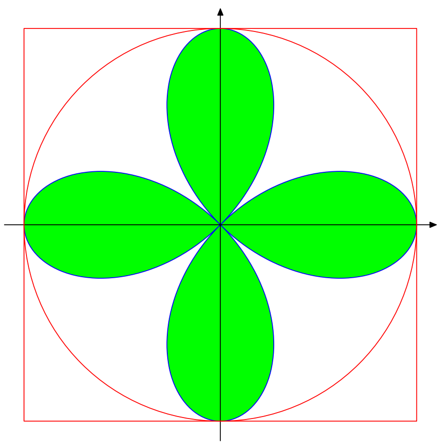

FIG. 14 – Ensemble B, cercle et carré circonscrits

7) *Calcul de l'aire $\lambda_2(B_0)$.* Commençons par préciser la définition de B_0 en notant φ le passage des coordonnées cartésiennes aux coordonnées polaires qu'il est commode ici de définir comme suit. Notons W et W' les ouverts

$$W := \mathbb{R}^2 \setminus]-\infty, 0] \times \{0\}, \quad W' :=]0, +\infty[\times]-\pi, \pi[.$$

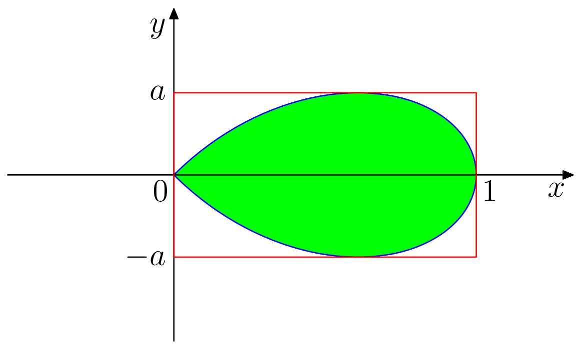

FIG. 15 – Ensemble B_0 et rectangle circonscrit $C_0 = [0,1] \times [-a, a]$

Posons en outre
$$W_0 := W \cup \{(0,0)\}, \quad W_0' := W' \cup \{(0,0)\}.$$
On définit $\varphi : W_0 \to W_0'$ en posant $\varphi(0,0) := 0$ et pour $(x,y) \in W$, $\varphi(x,y) = (r,t)$ où (r,t) est l'unique élément de W' tel que $x = r\cos t$ et $y = r \sin t$. Ainsi définie, φ est une *bijection* de W_0 sur W_0' et un C^1 difféomorphisme de l'*ouvert* W sur l'*ouvert* W'. Définissons maintenant

$$B_0' := \big\{(r,t) \in W_0';\ 0 \le r \le \cos(2t),\ -\pi/4 \le t \le \pi/4\big\}.$$

Avec ces notations, on peut réécrire la définition de B_0 sous la forme

$$B_0 = \big\{(x,y) \in W_0;\ \varphi(x,y) \in B_0'\big\}.$$

Pour calculer $\lambda_2(B_0)$, on utilise le passage en coordonnées polaires :

$$\lambda_2(B_0) = \int_{\mathbb{R}^2} \mathbf{1}_{B_0}(x,y)\,\mathrm{d}\lambda_2(x,y) = \int_W \mathbf{1}_{B_0}(x,y)\,\mathrm{d}\lambda_2(x,y) = \int_{W'} \mathbf{1}_{B_0}\big(\varphi^{-1}(r,t)\big) r\,\mathrm{d}\lambda_2(r,t).$$

En raison de l'équivalence entre « $\varphi^{-1}(r,t) \in B_0$ » et « $(r,t) \in \varphi(B_0) = B_0'$ », on a $\mathbf{1}_{B_0}\big(\varphi^{-1}(r,t)\big) = \mathbf{1}_{B_0'}(r,t)$, d'où

$$\lambda_2(B_0) = \int_{]0,+\infty[\times]-\pi,\pi[} \mathbf{1}_{B_0'}(r,t) r\,\mathrm{d}(\lambda_1 \otimes \lambda_1)(r,t).$$

Cette dernière intégrale se calcule grâce au théorème de Fubini-Tonelli. La section à t fixé de B_0' étant $[0, \cos(2t)]$, on obtient en utilisant aussi la coïncidence entre intégrales de Lebesgue et de Riemann pour une fonction continue sur un intervalle fermé borné :

$$\lambda_2(B_0) = \int_{[-\pi/4,\pi/4]} \int_{]0,\cos(2t)]} r\,\mathrm{d}\lambda_1(r)\,\mathrm{d}\lambda_1(t) = \int_{-\pi/4}^{\pi/4} \left[\frac{r^2}{2}\right]_0^{\cos(2t)} \mathrm{d}t = \int_{-\pi/4}^{\pi/4} \frac{\cos^2(2t)}{2}\,\mathrm{d}t.$$

Cette dernière intégrale se calcule par linéarisation grâce à la formule $1 + \cos u = 2\cos^2(u/2)$:
$$\lambda_2(B_0) = \frac{1}{4} \int_{-\pi/4}^{\pi/4} \big(1 + \cos(4t)\big)\,\mathrm{d}t = \frac{\pi}{8}.$$

Déterminons maintenant le rectangle C_0 minimal contenant B_0. Il s'écrit $C_0 = [0,1] \times [-a,a]$, où $a := \sup\{y;\ (x,y) \in B_0\}$. Pour cela on cherche les points à tangente horizontale de la frontière de B_0 d'équation polaire $r = \cos(2t)$, $-\pi/4, \leq t \leq \pi/4$. Une représentation paramétrique de cette frontière étant donnée par

$$\begin{cases} x(t) = \cos(2t)\cos t, \\ y(t) = \cos(2t)\sin t, \end{cases}$$

on est ramené à la résolution de l'équation $y'(t) = 0$. Par symétrie, on peut se limiter à $t \in [0, \pi/4]$. Le calcul élémentaire

$$y'(t) = -2\sin(2t)\sin t + \cos(2t)\cos t = -4\sin^2 t\cos t + (1-2\sin^2 t)\cos t = (1-6\sin^2 t)\cos t,$$

montre que l'équation $y'(t) = 0$ a pour unique solution sur $[0, \pi/4]$, l'unique réel t_0 tel que

$$\sin t_0 = \frac{1}{\sqrt{6}}.$$

L'étude du signe de $y'(t)$ pour $0 \leq t \leq t_0$ et $t_0 \leq t \leq \pi/4$ montre que $y(t_0)$ réalise le maximum pour les valeurs de $y(t)$. On a donc $a = y(t_0)$. Pour achever le calcul de a, on remarque que $1 - \cos(2t_0) = 2\sin^2(t_0) = 2/6 = 1/3$, d'où $\cos(2t_0) = 2/3$ et

$$a = y(t_0) = \cos(2t_0)\sin t_0 = \frac{2}{3\sqrt{6}}.$$

Finalement, $\lambda_2(C_0) = 2a$ et

$$\tau = \frac{\lambda_2(C_0)}{\lambda_2(B_0)} = \frac{32}{3\pi\sqrt{6}} \simeq 1{,}386.$$

8) Soit (U,V) un vecteur aléatoire à composantes indépendantes de même loi uniforme sur $[0,1]$. On définit le vecteur aléatoire

$$(X,Y) := \big(U^{1/2}\cos(2\pi V), U^{1/2}\sin(2\pi V)\big).$$

Montrons que (X,Y) suit la loi uniforme sur le disque de centre O et de rayon 1. Pour cela on calcule de deux façons $\mathbf{E}h(X,Y)$, h désignant une fonction borélienne positive quelconque.

1re façon, par transfert $\Omega \xrightarrow{(X,Y)} \mathbb{R}^2$:

$$\mathbf{E}h(X,Y) = \int_\Omega h\big(X(\omega), Y(\omega)\big)\,\mathrm{d}\mathbf{P}(\omega) = \int_{\mathbb{R}^2} h(x,y)\,\mathrm{d}P_{(X,Y)}(x,y). \tag{81}$$

2e façon, par transfert $\Omega \xrightarrow{(U,V)} \mathbb{R}^2$:

$$\begin{aligned} \mathbf{E}h(X,Y) &= \int_\Omega h\big(U(\omega)^{1/2}\cos(2\pi V(\omega)), U(\omega)^{1/2}\sin(2\pi V(\omega))\big)\,\mathrm{d}\mathbf{P}(\omega) \\ &= \int_{\mathbb{R}^2} h\big(u^{1/2}\cos(2\pi v), u^{1/2}\sin(2\pi v)\big)\,\mathrm{d}P_{(U,V)}(u,v) \qquad (82) \\ &= \int_{]0,1[^2} h\big(u^{1/2}\cos(2\pi v), u^{1/2}\sin(2\pi v)\big)\,\mathrm{d}\lambda_2(u,v). \qquad (83) \end{aligned}$$

Le passage de (82) à (83) exploite le fait que U et V étant indépendantes et de même loi uniforme sur $[0,1]$, le vecteur aléatoire (U,V) a pour densité $\mathbf{1}_{[0,1]^2}$ par rapport à λ_2. De plus, on peut dans (83) intégrer sur le carré ouvert $]0,1[^2$ au lieu de $[0,1]^2$ car la frontière de ce dernier est la réunion de quatre segments de λ_2-mesure nulle.

Il est alors naturel de poser dans (83),

$$(x,y) = \big(u^{1/2}\cos(2\pi v), u^{1/2}\sin(2\pi v)\big). \tag{84}$$

Pour préciser ce changement de variables, les notations suivantes nous seront utiles :

$$D := \big\{(x,y)\in\mathbb{R}^2;\ 0\le x^2+y^2<1\big\},\quad \Delta := D\setminus [0,1[\times\{0\}.$$

L'application $\varphi : (u,v)\mapsto (x,y)$ où (x,y) est donné par (84) réalise clairement une bijection de l'ouvert $]0,1[^2$ sur l'ouvert Δ. Ses quatre dérivées partielles apparaissant dans le calcul de son jacobien ci-dessous sont visiblement continues sur $]0,1[^2$, donc φ est de classe C^1.

$$\operatorname{Jac}(\varphi)(u,v) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix} = \begin{vmatrix} \frac{1}{2}u^{-1/2}\cos(2\pi v) & -2\pi u^{1/2}\sin(2\pi v) \\ \frac{1}{2}u^{-1/2}\sin(2\pi v) & 2\pi u^{1/2}\cos(2\pi v)\end{vmatrix} = \pi.$$

On voit ainsi que $\operatorname{Jac}(\varphi)(u,v)$ ne s'annule en aucun point de $]0,1[^2$. Le théorème d'inversion globale nous dit alors que φ est un C^1 difféomorphisme de $]0,1[^2$ sur Δ. De plus on a la relation

$$\operatorname{Jac}(\varphi^{-1})(x,y) = \frac{1}{\operatorname{Jac}(\varphi)(\varphi^{-1}(x,y))} = \frac{1}{\pi}.$$

La formule de changement de variable par le C^1-difféomorphisme φ appliquée à l'intégrale (83) nous donne finalement :

$$\mathbf{E}h(X,Y) = \int_\Delta h(x,y)\frac{1}{\pi}\,\mathrm{d}\lambda_2(x,y) = \int_D h(x,y)\frac{1}{\pi}\,\mathrm{d}\lambda_2(x,y) \tag{85}$$

$$= \int_{\mathbb{R}^2} h(x,y)\frac{1}{\pi}\mathbf{1}_D(x,y)\,\mathrm{d}\lambda_2(x,y). \tag{86}$$

La deuxième égalité dans (85) est légitime parce que D s'obtient en réunissant à Δ le segment $[0,1[\times\{0\}$ qui est de λ_2 mesure nulle.

La comparaison de (81) et (86) nous permet de conclure que la loi $P_{(X,Y)}$ est la mesure de densité $\frac{1}{\pi}\mathbf{1}_D$ par rapport à λ_2, c'est donc bien la loi uniforme sur D (noter au passage que $\lambda_2(D)=\pi$).

Remarque. $P_{(X,Y)}$ est aussi la loi uniforme sur le disque *fermé* $\overline{D}$, c'est d'ailleurs pour cela que l'énoncé ne précisait pas si le disque considéré était ouvert ou fermé. Il s'agit d'une propriété générale. Si μ est la loi uniforme sur un borélien B de $\mathbb{R}^d$, c'est aussi la loi uniforme sur tout borélien B' qui coïncide avec B, à un ensemble de λ_d mesure nulle près. Il suffit d'examiner la définition (70) pour s'en convaincre.

9) L'algorithme suivant simule un vecteur aléatoire M'_T de loi uniforme sur B en combinant l'algorithme du rejet relatif à B_0 et C_0 avec une rotation aléatoire. On génère d'abord U de loi uniforme sur $[0,1]$ et on définit

$$N := \sum_{k=0}^{3} k\mathbf{1}_{\{k\leq 4U<k+1\}}.$$

L'algorithme du rejet avec B_0 et C_0 nous fournit M_T de loi uniforme sur B_0 et on prend pour M'_T l'image de M_T par la rotation de centre O et d'angle $N\pi/2$.

Pour $k \in \mathbb{Z}$, nous noterons ρ_k la rotation de centre O et d'angle $k\pi/2$. Les ρ_k sont des bijections mesurables et vérifient $\rho_k^{-1} = \rho_{-k}$. De plus λ_2 est invariante par toute rotation.

La variable aléatoire discrète N suit la loi uniforme sur $N(\Omega) = \{0,1,2,3\}$. En effet

$$\forall j \in \{1,2,3,4\}, \quad \mathbf{P}(N=j) = \mathbf{P}(j \leq 4U < j+1) = \mathbf{P}\Big(U \in \Big[\frac{j}{4}, \frac{j+1}{4}\Big[\Big) = \frac{1}{4}. \quad (87)$$

Pour déterminer la loi de M'_T, on calcule $\mathbf{P}(M'_T \in A)$ pour A borélien quelconque de $\mathbb{R}^2$. Les justifications sont exposées après le calcul.

$$\begin{aligned}
\mathbf{P}(M'_T \in A) &= \sum_{k=0}^{3} \mathbf{P}(M'_T \in A \text{ et } N=k) && (88)\\
&= \sum_{k=0}^{3} \mathbf{P}\big(\rho_k(M_T) \in A \text{ et } N=k\big) && (89)\\
&= \sum_{k=0}^{3} \mathbf{P}\big(M_T \in \rho_{-k}(A)\big)\mathbf{P}\big(N=k\big) && (90)\\
&= \sum_{k=0}^{3} \frac{\lambda_2\big(\rho_{-k}(A)\cap B_0\big)}{4\lambda_2(B_0)} && (91)\\
&= \sum_{k=0}^{3} \frac{\lambda_2\big(A\cap B_k\big)}{\lambda_2(B)} = \frac{\lambda_2\big(A\cap B\big)}{\lambda_2(B)}. && (92)
\end{aligned}$$

Ceci étant vrai pour tout $A \in \mathrm{Bor}(\mathbb{R}^2)$, M'_T suit la loi uniforme sur B.
Justifications.

(88) : on a partitionné Ω suivant les 4 valeurs possibles de N.

(89) : pour tout $\omega \in \Omega$, on a les équivalences logiques

$$\begin{aligned}
M'_T(\omega) \in A \text{ et } N(\omega)=k &\Leftrightarrow M'_T(\omega) \in A \text{ et } N(\omega)=k \text{ et } M'_T(\omega) = \rho_k\big(M_T(\omega)\big)\\
&\Leftrightarrow \rho_k\big(M_T(\omega)\big) \in A \text{ et } N(\omega)=k.
\end{aligned}$$

L'égalité d'évènements $\{M'_T \in A \text{ et } N=k\} = \{\rho_k(M_T) \in A \text{ et } N=k\}$ en découle.

(90) : indépendance de M_T et N et équivalence entre $\rho_k(M) \in A$ et $M \in \rho_{-k}(A)$, car $\rho_{-k}(\rho_k(M)) = M$.

(91) : par la question 5, M_T suit la loi uniforme sur B_0, on peut donc appliquer (70). N suit la loi uniforme sur $\{0, 1, 2, 3\}$, cf. (87).

(92) : pour la première égalité, au numérateur on utilise l'invariance de λ_2 par rotation qui légitime l'égalité $\lambda_2(E) = \lambda_2(\rho_k(E))$ pour le borélien $E = A_{-k} \cap B_0$, où $A_{-k} := \rho_{-k}(A)$. De plus comme ρ_k est injective, $\rho_k(A_{-k} \cap B_0) = \rho_k(A_{-k}) \cap \rho_k(B_0)$. Enfin $\rho_k(B_0) = B_k$ et ρ_k étant la rotation inverse de ρ_{-k}, $\rho_k(A_{-k}) = A$. Pour le dénominateur, on remarque que $B \setminus \{O\}$ est l'union disjointe des $B_k \setminus \{O\}$ et que $\lambda_2(B_k) = \lambda_2\big(\rho_k(B_0)\big) = \lambda_2(B_0)$ en raison de l'invariance de λ_2 par rotation. On en déduit que $\lambda_2(B) = 4\lambda_2(B_0)$. La deuxième égalité se justifie de manière analogue en remarquant que $A \cap B$, privé éventuellement de O est l'union disjointe des $A \cap B_k$ privés éventuellement de O.

10) Comparons par leur coût (au sens de la limite p.s. de S_n/n, cf. question 6) les trois algorithmes suivants pour simuler un vecteur aléatoire de loi uniforme sur B :

a) Simulation de la loi uniforme sur le carré $C = [-1, 1]^2$ circonscrit à B et algorithme du rejet avec B et C.

b) Simulation de la loi uniforme sur le disque D de centre O et de rayon 1 par la méthode de la question 8 et algorithme du rejet avec B et $\overline{D}$. Il est clair que le disque fermé $\overline{D}$ contient B puisque les coordonnées polaires de ses points sont caractérisées par $0 \leq r \leq 1$ tandis que celles des points de B sont caractériséees par $0 \leq r \leq |\cos(2t)|$.

c) La méthode mixte de la question 9.

D'après la question 6) le coût des algorithmes a) et b) est τd avec $d = 2$ et τ valant respectivement

$$\tau_a = \frac{\lambda_2([-1, 1]^2)}{\lambda_2(B)} = \frac{4}{\pi/2} = \frac{8}{\pi}, \quad \tau_b = \frac{\lambda_2(D)}{\lambda_2(B)} = \frac{\pi}{\pi/2} = 2.$$

Pour la méthode de la question 9, la génération du vecteur aléatoire M'_{T_i} consomme $1 + 2T_i$ v.a. uniformes sur $[0, 1]$: une pour N_i et $2T_i$ pour M_{T_i}. On a donc

$$S_n = n + 2\sum_{i=1}^{n} T_i$$

et par la loi forte des grands nombres

$$\frac{S_n}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} 1 + 2\frac{\lambda_2(C_0)}{\lambda_2(B_0)} = 1 + \frac{64}{3\pi\sqrt{6}}$$

Les algorithmes a), b), c) sont rangés par ordre décroissant de coût puisque

$$\text{coût}(a) = \frac{16}{\pi} \simeq 5{,}093; \quad \text{coût}(b) = 4; \quad \text{coût}(c) = 1 + \frac{64}{3\pi\sqrt{6}} \simeq 3{,}772.$$

Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2003-04**

Corrigé de l'examen de deuxième session, septembre 2004

Ex 1. *Calcul de la série* $S := \sum_{k=0}^{+\infty} \int_0^{\pi/2} \left(1 - \sqrt{\sin x}\,\right)^k \cos x \, \mathrm{d}x.$

On remarque d'abord que le terme général I_k de cette série est l'intégrale de Riemann d'une fonction *continue* sur *l'intervalle fermé borné* $[0, \pi/2]$, on peut donc la convertir en intégrale de Lebesgue :

$$I_k = \int_0^{\pi/2} \left(1 - \sqrt{\sin x}\,\right)^k \cos x \, \mathrm{d}x = \int_{[0,\pi/2]} \left(1 - \sqrt{\sin x}\,\right)^k \cos x \, \mathrm{d}\lambda(x).$$

Pour $x \in [0, \pi/2]$, $\sin x$ et $\cos x$ sont dans $[0, 1]$ et ceci entraîne la positivité de l'intégrande dans I_k. Par le théorème d'interversion série intégrale pour les fonctions mesurables positives, on en déduit

$$S = \int_{[0,\pi/2]} \left\{ \sum_{k=0}^{+\infty} \left(1 - \sqrt{\sin x}\,\right)^k \cos x \right\} \mathrm{d}\lambda(x),$$

égalité vraie *a priori* dans $\overline{\mathbb{R}}_+$.

Pour $x \in]0, \pi/2]$, la série entre accolades est une série géométrique de raison $q(x) = (1 - \sqrt{\sin x}) \in [0, 1[$ et de premier terme $\cos x$. Elle est donc convergente et

$$\forall x \in]0, \pi/2], \quad \sum_{k=0}^{+\infty} \left(1 - \sqrt{\sin x}\,\right)^k \cos x = \frac{\cos x}{1 - q(x)} = \frac{\cos x}{\sqrt{\sin x}}.$$

Comme le singleton $\{0\}$ est λ-négligeable, on a $\int_{[0,\pi/2]} \{\dots\} \, \mathrm{d}\lambda = \int_{]0,\pi/2]} \{\dots\} \, \mathrm{d}\lambda$, d'où

$$S = \int_{]0,\pi/2]} \frac{\cos x}{\sqrt{\sin x}} \, \mathrm{d}\lambda(x).$$

La fonction $f : x \mapsto \cos x (\sin x)^{-1/2}$ est continue et positive sur $]0, \pi/2]$ et tend vers $+\infty$ à droite en 0, en étant équivalente à $x^{-1/2}$. L'intégrale de Riemann généralisée $\int_0^{\pi/2} f(x) \, \mathrm{d}x$ est donc absolument convergente et

$$S = \int_{]0,\pi/2]} \frac{\cos x}{\sqrt{\sin x}} \, \mathrm{d}\lambda(x) = \int_0^{\pi/2} \frac{\cos x}{\sqrt{\sin x}} \, \mathrm{d}x.$$

La fonction f étant de la forme $u^{-1/2}u'$ admet pour primitive $2u^{1/2}$, d'où finalement :

$$S = \lim_{\varepsilon\to 0^+} \int_\varepsilon^{\pi/2} \frac{\cos x}{\sqrt{\sin x}}\,\mathrm{d}x = \lim_{\varepsilon\to 0^+} \Big[2\sqrt{\sin x}\Big]_\varepsilon^{\pi/2} = 2.$$

Ex 2. *Étude de* $t \mapsto \mathbf{E}\arctan(tX)$

On note X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et on pose

$$\forall t \in \mathbb{R}, \quad \Psi_X(t) := \mathbf{E}\arctan(tX), \tag{93}$$

où arctan désigne la réciproque de la restriction de la fonction tangente à $]-\pi/2, \pi/2[$.

1) La variable aléatoire réelle X est par définition une application mesurable $\mathcal{F}$-Bor($\mathbb{R}$). Pour tout $t \in \mathbb{R}$, l'application $a_t : \mathbb{R} \to \mathbb{R}$, $x \mapsto \arctan(tx)$ est continue donc borélienne. L'application composée $a_t \circ X : \Omega \to \mathbb{R}$, $\omega \mapsto \arctan(tX(\omega))$ est donc mesurable $\mathcal{F}$-Bor($\mathbb{R}$). C'est donc une *variable aléatoire*, bornée puisque prenant ses valeurs dans $]-\pi/2, \pi/2[$. Par conséquent son espérance existe et ceci étant vrai pour t réel quelconque, la fonction Ψ_X est définie sur $\mathbb{R}$ et bornée. Remarquons au passage que par linéarité de l'espérance et imparité de la fonction arctan, Ψ_X est toujours impaire.

2) Par le transfert $\Omega \xrightarrow{X} \mathbb{R}$, on a en notant P_X la loi de X (mesure image $\mathbf{P}\circ X^{-1}$) :

$$\mathbf{E}\arctan(tX) = \int_\Omega \arctan(tX(\omega))\,\mathrm{d}\mathbf{P}(\omega) = \int_\mathbb{R} \arctan(tx)\,\mathrm{d}P_X(x).$$

Dans l'intégrale $I = \int_\mathbb{R} \arctan(tx)\,\mathrm{d}P_X(x)$, le symbole x est une variable muette que l'on peut remplacer par n'importe quelle lettre (par exemple y) et on peut tout aussi bien écrire $I = \int_\mathbb{R} \arctan(t\,.)\,\mathrm{d}P_X$. Si X et Y ont même loi, $P_X = P_Y$ et

$$\Psi_X(t) = \int_\mathbb{R} \arctan(t\,.)\,\mathrm{d}P_X = \int_\mathbb{R} \arctan(t\,.)\,\mathrm{d}P_Y = \Psi_Y(t).$$

Cette égalité étant vérifiée pour tout t réel, on a bien $\Psi_X = \Psi_Y$.

Ainsi la fonctionnelle Ψ_X ne dépend de la variable aléatoire X que par sa loi P_X. Calculons la dans les deux cas particuliers suivants :

a) $P_X = \mathrm{Bern}(p)$. Alors $P_X = (1-p)\delta_0 + p\delta_1$, d'où

$$\Psi_X(t) = \int_\mathbb{R} \arctan(t\,.)\,\mathrm{d}\big((1-p)\delta_0 + p\delta_1\big) = (1-p)\arctan(0) + p\arctan(t) = p\arctan(t).$$

b) $P_X = \mathrm{Unif}([0,1])$. Alors P_X est la restriction [24] de la mesure de Lebesgue à $[0,1]$, d'où

$$\Psi_X(t) = \int_{[0,1]} \arctan(tx)\,\mathrm{d}\lambda(x) = \int_0^1 \arctan(tx)\,\mathrm{d}x,$$

la conversion de l'intégrale de Lebesgue en intégrale de Riemann se justifiant ici par la continuité de l'intégrande sur l'intervalle fermé borné $[0,1]$. Cette intégrale de Riemann

24. Plus précisément la mesure de densité $\mathbf{1}_{[0,1]}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$.

est nulle pour $t = 0$. Pour $t \neq 0$, on la calcule en enchaînant le changement de variable $u = tx$ et une intégration par parties.

$$\begin{aligned}\int_0^1 \arctan(tx)\,\mathrm{d}x = \frac{1}{t}\int_0^t \arctan u\,\mathrm{d}u &= \frac{1}{t}\left[u\arctan u\right]_0^t - \frac{1}{t}\int_0^t \frac{u}{1+u^2}\,\mathrm{d}u \\ &= \arctan t - \frac{1}{2t}\left[\ln(1+u^2)\right]_0^t.\end{aligned}$$

Finalement $\Psi_X(0) = 0$ et

$$\forall t \in \mathbb{R}^*, \quad \Psi_X(t) = \arctan t - \frac{\ln(1+t^2)}{2t}.$$

Remarquons au passage que Ψ_X est continue en 0, puisque $\lim_{x\to 0}\Psi_X(t) = 0$.

3) Lorsque la loi de X est *symétrique*, *i.e.* X et $-X$ ont même loi, on a clairement (cf. question 2) :

$$\forall t \in \mathbb{R}, \quad \Psi_X(t) = \Psi_{-X}(t) = \mathbf{E}\arctan(-tX) = -\Psi_X(t),$$

d'où $\Psi_X(t) = 0$. Ainsi Ψ_X est la fonction nulle chaque fois que X a une loi symétrique. Il en résulte que la fonctionnelle Ψ ne caractérise pas la loi puisque si X et Y ont des lois symétriques différentes, par exemple $P_X = \mathrm{Unif}[-1,1]$ et $P_Y = \mathfrak{N}(0,1)$, $\Psi_X = \Psi_Y = 0$.

4) La continuité de Ψ_X sur $\mathbb{R}$ est une application immédiate du théorème de continuité sous le signe somme en notant que

a) pour tout $\omega \in \Omega$, l'application $t \mapsto \arctan(tX(\omega))$ est continue sur $\mathbb{R}$;

b) l'application $\omega \mapsto \arctan(tX(\omega))$ est dominée sur Ω par la fonction constante (donc $\mathbf{P}$-intégrable) $\pi/2$, et ce *uniformément par rapport à* $t \in \mathbb{R}$.

Il est clair que cette justification n'utilise aucune hypothèse sur la loi de X.

5) Pour montrer que

$$\lim_{t\to+\infty}\Psi_X(t) = -\frac{\pi}{2}\mathbf{P}(X<0) + \frac{\pi}{2}\mathbf{P}(X>0), \tag{94}$$

notons $(t_n)_{n\geq 1}$ une suite tendant vers $+\infty$. L'étude de la convergence de $\big(\arctan(t_nX)\big)_{n\geq 1}$ nous amène naturellement à distinguer les trois cas suivants correspondant à la partition de Ω en les évènements $\{X<0\}$, $\{X=0\}$ et $\{X>0\}$. On voit immédiatement que

$$\arctan(t_nX(\omega)) \xrightarrow[n\to+\infty]{} \begin{cases} -\frac{\pi}{2} & \text{si } X(\omega) < 0, \\ 0 & \text{si } X(\omega) = 0, \\ \frac{\pi}{2} & \text{si } X(\omega) > 0. \end{cases}$$

Autrement dit la suite de fonctions mesurables $\big(\arctan(t_nX)\big)_{n\geq 1}$ converge sur tout Ω vers $-\frac{\pi}{2}\mathbf{1}_{\{X<0\}} + \frac{\pi}{2}\mathbf{1}_{\{X>0\}}$. Cette convergence est *dominée* par la fonction constante (donc $\mathbf{P}$-intégrable) $\omega \mapsto \frac{\pi}{2}$. Le théorème de convergence dominée nous donne alors

$$\lim_{n\to+\infty}\Psi_X(t_n) = \mathbf{E}\Big(-\frac{\pi}{2}\mathbf{1}_{\{X<0\}} + \frac{\pi}{2}\mathbf{1}_{\{X>0\}}\Big) = -\frac{\pi}{2}\mathbf{P}(X<0) + \frac{\pi}{2}\mathbf{P}(X>0).$$

Le choix de la suite $(t_n)_{n\geq 1}$ tendant vers $+\infty$ étant arbitraire, on en déduit (94).

Pour trouver la limite de Ψ_X en $-\infty$, on peut soit reprendre l'argumentation précédente en changeant seulement le signe de la limite dans les cas 1 et 3, soit utiliser l'imparité de Ψ_X. On obtient ainsi :

$$\lim_{t\to-\infty} \Psi_X(t) = \frac{\pi}{2}\mathbf{P}(X<0) - \frac{\pi}{2}\mathbf{P}(X>0). \tag{95}$$

6) *Étude de la dérivabilité de* Ψ_X

Montrer que Ψ_X est dérivable sur $\mathbb{R}^*$ quelle que soit la loi de X et qu'elle est dérivable en 0 si $\mathbf{E}|X|$ est finie. Que peut-on dire de la continuité de la dérivée de Ψ_X ?

Posons

$$f : \mathbb{R}\times\Omega \longrightarrow \mathbb{R}, \quad (t,\omega) \longmapsto \arctan\big(tX(\omega)\big).$$

L'étude de la dérivabilité de Ψ_X *via* le théorème de dérivation sous le signe somme repose sur les constatations suivantes.

1. Pour tout $t\in\mathbb{R}$, $f(t,\,.\,)$ est $\mathbf{P}$-intégrable sur Ω comme fonction bornée.
2. Pour tout $\omega\in\Omega$, $f(\,.\,,\omega)$ est dérivable sur $\mathbb{R}$.
3. Pour tout $t\in\mathbb{R}$,
$$\frac{\partial}{\partial t} f(t,\omega) = \frac{X(\omega)}{1+t^2X(\omega)^2}.$$

Pour appliquer le théorème de dérivation sous le signe somme, nous cherchons une majoration uniforme en t de $\left|\frac{\partial f}{\partial t}\right|$ par une fonction $g : \Omega \to \overline{\mathbb{R}}_+$ intégrable sur Ω. Si X est elle-même intégrable, il suffit d'utiliser la majoration évidente :

$$\forall t\in\mathbb{R},\ \forall\omega\in\Omega, \quad \left|\frac{X(\omega)}{1+t^2X(\omega)^2}\right| = \frac{|X(\omega)|}{1+t^2X(\omega)^2} \leq |X(\omega)|. \tag{96}$$

Le théorème de dérivation sous le signe somme (appliqué en version « globale ») nous dit alors que si X est intégrable, Ψ_X est dérivable sur $\mathbb{R}$ et

$$\forall t\in\mathbb{R}, \quad (\Psi_X)'(t) = \mathbf{E}\left(\frac{X}{1+t^2X^2}\right). \tag{97}$$

Que se passe-t-il si X n'est plus intégrable ? On remarque que dans (96), le majorant $|X(\omega)|$ est atteint pour $t=0$. On ne peut donc pas trouver de majorant plus petit uniforme en t sur un voisinage de 0. Par conséquent la dérivabilité en 0 ne peut s'obtenir par le théorème de dérivabilité sous le signe somme.

Essayons néanmoins de prouver la dérivabilité sur $\mathbb{R}^*$ en fixant un $t_0\neq 0$ et en cherchant un voisinage V_0 de t_0 sur lequel $\left|\frac{\partial f}{\partial t}\right|$ soit majoré uniformément en t par une fonction intégrable sur Ω. Puisque Ψ_X est impaire, on ne perd pas de généralité en supposant $t_0>0$. On utilise alors les majorations suivantes.

$$\text{Si } |X(\omega)|\leq 1, \qquad \frac{|X(\omega)|}{1+t^2X(\omega)^2} \leq |X(\omega)| \leq 1.$$

$$\text{Si } |X(\omega)|> 1, \qquad \frac{|X(\omega)|}{1+t^2X(\omega)^2} \leq \frac{|X(\omega)|}{t^2X(\omega)^2} = \frac{1}{t^2|X(\omega)|} \leq \frac{1}{t^2}.$$

On peut résumer ces deux cas en notant que

$$\forall t \in \mathbb{R}^*,\ \forall \omega \in \Omega, \quad \Big|\frac{\partial}{\partial t} f(t,\omega)\Big| \leq \max\Big(1, \frac{1}{t^2}\Big).$$

Finalement, en prenant comme voisinage de t_0, $V_0 :=]t_0/2, +\infty[$, on a majoration uniforme sur V_0 de $\big|\frac{\partial f}{\partial t}\big|$ par la fonction constante (donc $\mathbf{P}$-intégrable) $\omega \mapsto \max(1, 4t_0^{-2})$. Le théorème de dérivation sous le signe somme (version « locale ») nous donne alors la dérivabilité de Ψ_X sur V_0, la dérivée se calculant par (97). Comme $t_0 \neq 0$ était arbitraire, cette formule reste vraie pour tout $t \in \mathbb{R}^*$.

Le fait que pour X non intégrable, le théorème de dérivation sous le signe somme ne s'applique pas sur un voisinage de 0 n'interdit pas *a priori* à Ψ_X d'être dérivable en 0. D'ailleurs il est facile d'exhiber des exemples de cette situation en prenant pour X n'importe quelle variable aléatoire non intégrable et de loi symétrique (puisqu'alors Ψ_X est identiquement nulle sur $\mathbb{R}$, donc trivialement C^∞ sur $\mathbb{R}$). À partir de cette famille de variables aléatoires symétriques non intégrables, on peut construire une famille encore plus grande de variables aléatoires Z non intégrables telles que Ψ_Z soit dérivable en 0. Il suffit de prendre X symétrique non intégrable, Y intégrable et Z telle que $P_Z = pP_X + (1-p)P_Y$ pour un $p \in]0,1[$ (exercice : vérifiez cette affirmation et proposez une méthode pour construire explicitement une telle Z à partir de X et Y).

Nous allons maintenant prouver que si X est *positive non intégrable*, Ψ_X est non dérivable en 0. Ceci établira l'optimalité de la dérivabilité de Ψ_X sur $\mathbb{R}^*$ dans le cas où l'on ne fait aucune hypothèse sur la loi de X. Dans ce but, on montre (en notant que $\Psi_X(0) = 0$) que pour une suite quelconque (t_n) dans $\mathbb{R}^*$ et tendant vers 0,

$$\lim_{n\to+\infty} \frac{\Psi_X(t_n)}{t_n} = +\infty. \tag{98}$$

Pour une telle suite (t_n) on a

$$\forall \omega \in \Omega, \quad \frac{\arctan\big(t_n X(\omega)\big)}{t_n} \xrightarrow[n\to+\infty]{} X(\omega).$$

Les fonctions $\arctan(t_n X)$ étant mesurables positives, cette convergence combinée avec le lemme de Fatou nous donne alors[25] :

$$\liminf_{n\to+\infty} \int_\Omega \frac{\arctan\big(t_n X(\omega)\big)}{t_n} \,\mathrm{d}P(\omega) \geq \int_\Omega X(\omega)\,\mathrm{d}P(\omega) = +\infty.$$

Ceci prouve (98), puisque si la limite inférieure d'une suite dans $\overline{\mathbb{R}}$ vaut $+\infty$, la suite a pour limite $+\infty$.

25. Rappelons que si X est mesurable positive, l'écriture $\int_\Omega X(\omega)\,\mathrm{d}P(\omega)$ a toujours un sens comme élément de $\overline{\mathbb{R}}_+$.

Ex 3. *Simulation par rapport d'uniformes (6 points)*

On se propose de justifier la méthode suivante de simulation d'une variable aléatoire Y de densité f connue. On suppose que l'on sait simuler un vecteur aléatoire (U,V) de loi uniforme sur le borélien

$$A := \Big\{(u,v)\in\mathbb{R}^2;\ 0<u<f\Big(\frac{v}{u}\Big)^{1/2}\Big\}. \tag{99}$$

Alors la variable aléatoire $Y := V/U$ suit la loi de densité f.

1) *Calcul de* $\lambda_2(A)$. En notant l'inclusion $A\subset]0,+\infty[\times\mathbb{R}$, on a

$$\lambda_2(A) = \int_{\mathbb{R}^2} \mathbf{1}_A(u,v)\,\mathrm{d}\lambda_2(u,v) = \int_{]0,+\infty[\times\mathbb{R}} \mathbf{1}_A(u,v)\,\mathrm{d}\lambda_2(u,v). \tag{100}$$

Considérons le changement de variable

$$\varphi :]0,+\infty[\times\mathbb{R} \longrightarrow]0,+\infty[\times\mathbb{R},\quad (u,v)\longmapsto (x,y)=\varphi(u,v)=\Big(u,\frac{v}{u}\Big).$$

C'est une bijection de l'ouvert $D=]0,+\infty[\times\mathbb{R}$ sur lui même. En effet pour tout $(u,v)\in D$, u est strictement positif donc $(u,v/u)$ est dans D. Ainsi $\varphi(D)$ est inclus dans D. D'autre part pour tout couple $(x,y)\in D$, il existe un unique couple $(u,v)\in D$ tel que $\varphi(u,v)=(x,y)$, ce couple s'obtient immédiatement en résolvant le système $x=u$ et $y=v/u$, d'inconnues u et v, ce qui donne $u=x$ et $v=xy$.

L'application réciproque φ^{-1} est donnée par

$$\varphi^{-1} :]0,+\infty[\times\mathbb{R} \longrightarrow]0,+\infty[\times\mathbb{R},\quad (x,y)\longmapsto (u,v)=\varphi^{-1}(x,y)=(x,xy).$$

Les quatre dérivées partielles de φ^{-1} apparaissant dans le calcul de son jacobien ci-dessous sont visiblement continues sur D, donc φ^{-1} est de classe C^1.

$$\operatorname{Jac}(\varphi^{-1})(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ y & x \end{vmatrix} = x.$$

On voit ainsi que $\operatorname{Jac}(\varphi^{-1})(x,y)$ ne s'annule en aucun point de D. Le théorème d'inversion globale nous dit alors que φ est un C^1 difféomorphisme de D sur D.

Le changement de variable φ dans (100) s'écrit :

$$\lambda_2(A) = \int_{]0,+\infty[\times\mathbb{R}} \mathbf{1}_A\big(\varphi^{-1}(x,y)\big)x\,\mathrm{d}\lambda_2(x,y) = \int_{]0,+\infty[\times\mathbb{R}} \mathbf{1}_{\varphi(A)}(x,y)x\,\mathrm{d}\lambda_2(x,y). \tag{101}$$

On détermine $B:=\varphi(A)$ en notant que pour $(x,y)\in D$,

$$\varphi^{-1}(x,y)\in A \Leftrightarrow (x,xy)\in A \Leftrightarrow 0<x<f(y)^{1/2} \Leftrightarrow 0<x^2<f(y).$$

La dernière équivalence utilise la positivité stricte de x due à l'appartenance de (x,y) à D. Ainsi $B=\{(x,y)\in D;\ 0<x^2<f(y)\}$. En reportant ceci dans (101), on obtient en

utilisant successivement le théorème de Fubini-Tonelli, le changement de variable $t = x^2$ et le fait que f est une densité de probabilité ($\int_{\mathbb{R}} f \,\mathrm{d}\lambda_1 = 1$) :

$$\begin{aligned}
\lambda_2(A) &= \int_{\mathbb{R}} \left\{ \int_{]0,+\infty[} \mathbf{1}_{]0,f(y)[}(x^2) x \,\mathrm{d}\lambda_1(x) \right\} \mathrm{d}\lambda_1(y) \\
&= \int_{\mathbb{R}} \left\{ \int_{]0,+\infty[} \mathbf{1}_{]0,f(y)[}(t) \frac{1}{2} \,\mathrm{d}\lambda_1(t) \right\} \mathrm{d}\lambda_1(y) \\
&= \frac{1}{2} \int_{\mathbb{R}} f(y) \,\mathrm{d}\lambda_1(y) = \frac{1}{2}.
\end{aligned}$$

Le borélien A étant de mesure de Lebesgue *finie*, il est légitime de parler de la loi uniforme sur A.

2) Pour obtenir la loi du couple $(X,Y) := \Big(U, \dfrac{V}{U}\Big)$, on calcule de deux façons $\mathbf{E}h(X,Y)$ pour $h : \mathbb{R}^2 \to \mathbb{R}_+$ borélienne positive arbitraire.

1re façon, par transfert $\Omega \xrightarrow{(X,Y)} \mathbb{R}^2$.

$$\mathbf{E}h(X,Y) = \int_{\Omega} h\big(X(\omega), Y(\omega)\big) \,\mathrm{d}\mathbf{P}(\omega) = \int_{\mathbb{R}^2} h(x,y) \,\mathrm{d}P_{(X,Y)}(x,y). \tag{102}$$

2e façon, par transfert $\Omega \xrightarrow{(U,V)} \mathbb{R}^2$. En utilisant les mêmes notations qu'à la question précédente, on obtient :

$$\begin{aligned}
\mathbf{E}h(X,Y) &= \int_{\Omega} h\left(U(\omega), \frac{V(\omega)}{U(\omega)}\right) \mathrm{d}\mathbf{P}(\omega) \\
&= \int_{\mathbb{R}^2} h\Big(u, \frac{v}{u}\Big) \,\mathrm{d}P_{(U,V)}(u,v) && (103) \\
&= \int_{D} h\Big(u, \frac{v}{u}\Big) 2\mathbf{1}_A(u,v) \,\mathrm{d}\lambda_2(u,v) && (104) \\
&= \int_{D} h(x,y) 2x\mathbf{1}_B(x,y) \,\mathrm{d}\lambda_2(x,y) && (105) \\
&= \int_{\mathbb{R}^2} h(x,y) 2x\mathbf{1}_B(x,y) \,\mathrm{d}\lambda_2(x,y). && (106)
\end{aligned}$$

Justifications.

(103) : Transfert $\Omega \xrightarrow{(U,V)} \mathbb{R}^2$.

(104) : $P_{(U,V)}$ est la loi uniforme sur A, de densité $\lambda_2(A)^{-1}\mathbf{1}_A$ et A est inclus dans l'ouvert D.

(105) : On applique le changement de variable φ étudié à la question précédente, $\varphi(D) = D$, $\varphi(A) = B$ et $|\operatorname{Jac}(\varphi^{-1})(x,y)| = x$.

La comparaison de (102) et (106) nous permet de conclure que la loi $P_{(X,Y)}$ est la mesure de densité par rapport à λ_2 :

$$g : (x,y) \longmapsto 2x\mathbf{1}_B(x,y), \quad \text{où} \quad B = \{(x,y) \in D;\ 0 < x^2 < f(y)\}.$$

Il en résulte que $Y = V/U$ est à densité par rapport à λ_1 et que cette densité g_Y s'obtient par intégration partielle de g :

$$g_Y(y) = \int_{\mathbb{R}} g(x,y)\,\mathrm{d}\lambda_1(x) = \int_{\mathbb{R}} \mathbf{1}_{]0,f(y)[}(x^2)2x\,\mathrm{d}\lambda_1(x) = \int_{\mathbb{R}} \mathbf{1}_{]0,f(y)[}(t)\,\mathrm{d}\lambda_1(t) = f(y).$$

Ainsi la variable aléatoire $Y = V/U$ a bien pour densité f, ce qui achève la justification de la méthode du rapport d'uniformes pour simuler Y.

3) Cette méthode n'est applicable que si l'on sait effectivement simuler un vecteur de loi uniforme sur A. C'est le cas notamment *via* l'algorithme du rejet lorsque A est *borné*. Montrons qu'une condition suffisante pour que A soit borné est que la densité f vérifie

$$\sup_{x\in\mathbb{R}} f(x) < +\infty \quad \text{et} \quad \sup_{x\in\mathbb{R}} x^2 f(x) < +\infty. \tag{107}$$

Supposons donc qu'il existe deux constantes réelles positives a et b telles que

$$\forall t \in \mathbb{R}, \quad f(t) \leq a, \quad t^2 f(t) \leq b.$$

Soit (u,v) un point de A. Alors l'inégalité $0 < u < f(v/u)^{1/2}$ implique immédiatement que $0 < u < a^{1/2}$. D'autre part on a $0 < u^2 < f(v/u)$, ce qui en posant $t = v/u$ s'écrit encore $0 < (v/t)^2 < f(t)$, d'où $0 < v^2 < t^2 f(t) \leq b$. On en déduit $-b^{1/2} < v < b^{1/2}$. Comme (u,v) était quelconque dans A, nous avons ainsi établi l'inclusion

$$A \subset [0, a^{1/2}] \times [-b^{1/2}, b^{1/2}].$$

Ainsi les conditions (107) impliquent la bornitude de A.

Ex 4. *Escargot aléatoire (4 points)*

Soit $(X_i)_{i\in\mathbf{N}^*}$ une suite de variables aléatoires, définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$, indépendantes et de même loi, de carré intégrable ($\mathbf{E}(X_1^2) < +\infty$). On définit par récurrence la suite de variables aléatoires positives $(R_n)_{n\geq 1}$ par

$$R_1 = |X_1|, \quad R_n = \sqrt{R_{n-1}^2 + X_n^2}, \quad n \geq 2.$$

1) Pour étudier le comportement asymptotique de $(n^{-1/2}R_n)$, on commence par remarquer que :

$$R_n^2 = R_{n-1}^2 + X_n^2 \quad \text{d'où} \quad R_n^2 = \sum_{i=1}^n X_i^2.$$

Ainsi $\left(n^{-1}R_n^2\right)_{n\geq 1}$ apparaît comme la suite des moyennes arithmétiques des termes de la suite $(X_i^2)_{i\in\mathbf{N}^*}$. Comme ces variables s'écrivent $X_i^2 = h(X_i)$ où h est la fonction mesurable $h : x \mapsto x^2$, la suite $(X_i^2)_{i\in\mathbf{N}^*}$ hérite de l'indépendance des X_i. Comme h ne dépend pas de i et les X_i ont même loi, il en va de même pour les X_i^2. Enfin, $\mathbf{E}(X_i^2)$ est finie par hypothèse. Les X_i^2 vérifient donc les hypothèse de la loi forte des grands nombres (cas i.i.d.), d'où

$$\frac{R_n^2}{n} = \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbf{E}(X_1^2). \tag{108}$$

Par continuité de la fonction $g : \mathbb{R}_+ \to \mathbb{R}_+$, $x \mapsto \sqrt{x}$, on en déduit immédiatement que

$$\frac{R_n}{n^{1/2}} \xrightarrow[n\to+\infty]{\text{p.s.}} \sqrt{\mathbf{E}(X_1^2)} =: \tau. \tag{109}$$

2) Les X_i étant maintenant de même loi uniforme sur $[0, 1]$, on construit dans un repère orthonormé du plan la suite $(M_n)_{n\geq 1}$ de points aléatoires par la récurrence suivante. On pose d'abord $M_1 = (X_1, 0)$, puis connaissant M_{n-1} on obtient M_n comme l'unique point tel que $M_{n-1}M_n = X_n$ et que l'angle $(\overrightarrow{M_{n-1}M_n}, \overrightarrow{M_{n-1}O})$ ait pour mesure $+\pi/2$. On trace ainsi la ligne polygonale $\mathcal{E}_n$ de sommets $M_1, M_2, \dots, M_n$ (l'escargot aléatoire). On fixe un $\varepsilon > 0$ et on se propose de montrer que presque sûrement, $\mathcal{E}_n$ est inclus dans le disque de centre O et de rayon $(1 + \varepsilon)\sqrt{n/3}$ pour tout n assez grand.

Commençons par calculer τ lorsque X_1 suit la loi uniforme sur $[0, 1]$.

$$\mathbf{E}X_1^2 = \int_0^1 x^2 \,\mathrm{d}x = \frac{1}{3}, \quad \text{d'où} \quad \tau = \frac{1}{\sqrt{3}}.$$

Remarquons ensuite que la suite (R_n) étant croissante, $\mathcal{E}_n$ est inclus dans le disque fermé D_n de centre O et de rayon R_n. En effet puisque $OM_i = R_i$, tous les M_i pour $i \leq n$ sont dans D_n. De plus si $i < n$, M_i et M_{i+1} sont tous deux dans D_n qui par *convexité* contient tout le segment d'extrémités M_i et M_{i+1}. Ainsi D_n contient toute la ligne polygonale $\mathcal{E}_n$.

Notons $\Omega' := \{\omega \in \Omega;\ \lim_{n\to+\infty} n^{-1/2}R_n(\omega) = \tau\}$. On a alors

$$\forall \omega \in \Omega',\ \forall \varepsilon > 0,\ \exists n_0 = n_0(\omega, \varepsilon),\ \forall n \geq n_0, \quad n^{-1/2}R_n(\omega) \leq \tau(1 + \varepsilon).$$

Pour ε_0 fixé, on a clairement l'inclusion d'évènements

$$\Omega' \subset \big\{\omega \in \Omega,\ \exists n_0 = n_0(\omega, \varepsilon_0),\ \forall n \geq n_0,\ n^{-1/2}R_n(\omega) \leq \tau(1 + \varepsilon)\big\} =: A_n.$$

Pour tout $\omega \in A_n$, on a

$$\forall n \geq n_0(\omega, \varepsilon_0) \quad R_n(\omega) \leq n^{1/2}\tau(1 + \varepsilon), \quad \text{d'où} \quad \mathcal{E}_n(\omega) \subset D_n \subset D\Big(O, (1 + \varepsilon)\sqrt{\frac{n}{3}}\Big).$$

Par la loi forte des grands nombres (question précédente), $\mathbf{P}(\Omega') = 1$, d'où $\mathbf{P}(A_n) = 1$. On a ainsi établi que presque sûrement, $\mathcal{E}_n$ est inclus dans le disque de centre O et de rayon $(1 + \varepsilon)\sqrt{n/3}$ pour tout n assez grand.

La figure 16 illustre ceci. On a choisi pour cette simulation $n = 200$, $\varepsilon = 0{,}05$. Les deux cercles tracés ont respectivement pour rayon R_n et $(1 + \varepsilon)\sqrt{n/3}$.

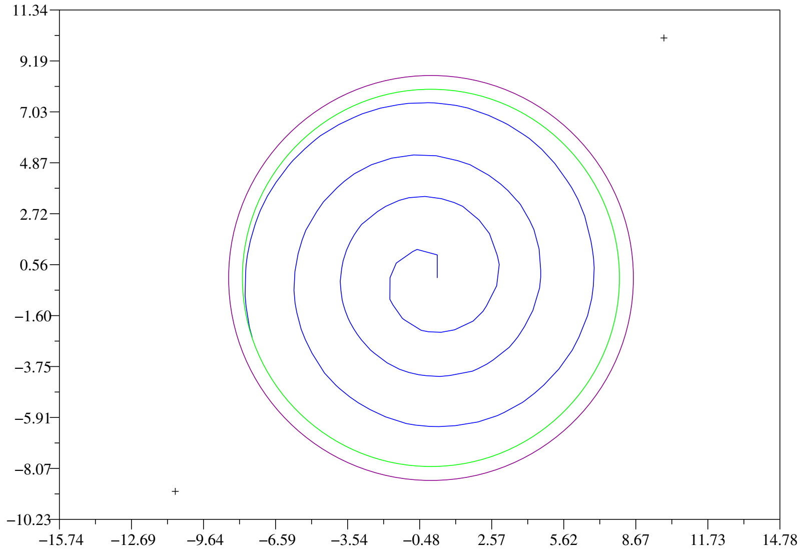

FIG. 16 – Escargot aléatoire $\mathcal{E}_{200}$ (simulation en Scilab)