Faculté des Sciences
Aix-Marseille Université

CTES
Centre de Télé-Enseignement Sciences

MASTER 1 MATHEMATIQUES – INFORMATIQUE
MATHEMATIQUES GENERALES
M1...

Expédition dans la semaine n°	Etape	Code UE	N° d'envoi de l'UE
7	**M1, UE 1.3**	**MIMA-SMAAU11T**	**4**

Nom de l'UE : Mesure, Intégration, Probabilités

Le cours est essentiellement tiré d'un livre disponible sur la page web indiquée ci dessous. Pour chaque semaine, il est proposé d'étudier une partie du cours et de faire des exercices (corrigés).
Deux devoirs sont à rendre afin de bénéficier d'une note de contrôle continu.
note finale=max(note-examen, 1/3(2 note-examen + note-contrôle-continu)).

Contenu de l'envoi : Cours : Chapitres 7, 8 et 9, Exercices
Le chapitre 8 est hors programme, il est intéressant en liaison avec les autres cours

- Guide du travail à effectuer

Semaine 1 (Mesure produit, théorème de Fubini) :
Etudier les paragraphes 7.1, 7.2 (mesure produit) et 7.3 (théorèmes de Fubini-Tonelli et Fubini)
Exercices proposés (avec corrigés) : 7.1, 7.6, 7.7, 7.10

Semaine 2 (Mesure de Lebesgue sur Rn, convolution, changement de variables) :
Etudier les paragraphes 7.4 (mesure de Lebesgue), 7.5 (convolution) et 7.6 (changement de variables)
Exercices proposés (avec corrigés) : 7.11, 7.13, 7.17, 7.18, 7.21, 7.28

Semaine 3 (Vecteurs aléatoires, indépendance) :
Etudier rapidement le chapitre 8 (à la limite du programme de cette UE), sans exercices
Etudier les paragraphes 9.1 (vecteurs gaussiens) et 9.2 (indépendance)
Exercices proposés (avec corrigés) : 9.2, 9.6, 9.7, 9.8, 9.9

Semaine 4 (Vecteurs gaussiens) :
Etudier le paragraphes 9.3 (vecteurs gaussiens)
Exercices proposés (avec corrigés) : 9.12, 9.13, 9.14, 9.16, 9.17

-Coordonnées des enseignants responsables de l'envoi
E. Hillion et T. Gallouet, CMI, 39 rue Joliot Curie, 13453 marseille cedex 13
email : thierry.gallouet@univ-amu.fr, erwan.hillion@univ-amu.fr
Vous pouvez aussi consulter la page web:
http://www.cmi.univ-mrs.fr/~gallouet/tele.d/mip.d
et nous poser des questions par le forum d'ametice ou par email

Chapitre 7

Produits d'espaces mesurés

7.1 Motivation

Au chapitre 2, on a introduit la mesure de Lebesgue sur la tribu des boréliens de $\mathbb{R}$ (notée $\mathcal{B}(\mathbb{R})$), ce qui nous a permis d'exprimer la notion de longueur d'une partie (borélienne) de $\mathbb{R}$. On peut se poser la question de savoir s'il existe une mesure sur une tribu convenable de $\mathbb{R}^2$ qui exprimerait la notion de surface (et une mesure sur une tribu convenable de $\mathbb{R}^3$ qui exprimerait la notion de volume...).

La question est donc : existe-t-il une mesure λ_2 sur une tribu de $\mathbb{R}^2$ contenant $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$, vérifiant :

$$\lambda_2(\mathrm{A} \times \mathrm{B}) = \lambda(\mathrm{A})\lambda(\mathrm{B}),\ \forall \mathrm{A}, \mathrm{B} \in \mathcal{B}(\mathbb{R})\ ?$$

La tribu T_2, sur laquelle on veut définir λ_2, doit donc contenir $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$. On remarque tout d'abord que $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) = \{\mathrm{A} \times \mathrm{B}, \mathrm{A} \in \mathcal{B}(\mathbb{R}), \mathrm{B} \in \mathcal{B}(\mathbb{R})\}$ n'est pas une tribu. En effet, $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ n'est pas stable par passage au complémentaire ni par union (par contre, $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ est stable par intersection dénombrable). On définit alors T_2 comme la tribu engendrée par $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$, qu'on note $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$.

On cherche alors une mesure $\lambda_2 : \mathrm{T}_2 \to \overline{\mathbb{R}}_+$ telle que $\lambda_2(\mathrm{A} \times \mathrm{B}) = \lambda(\mathrm{A})\lambda(\mathrm{B})$ pour tout $\mathrm{A}, \mathrm{B} \in \mathcal{B}(\mathbb{R})$. On peut montrer l'existence et l'unicité de la mesure λ_2 (voir le théorème 7.3). On peut aussi montrer que la tribu T_2 est la tribu borélienne sur $\mathbb{R}^2$, c'est-à-dire la tribu engendrée par les ouverts de $\mathbb{R}^2$ (voir la proposition 7.2).

Une autre question qu'on abordera dans ce chapitre concerne l'intégration des fonctions à plusieurs variables. Considérons par exemple une fonction f définie de $\mathbb{R}^2$

dans $\mathbb{R}$. Sous quelles hypothèses (faciles à vérifier...) peut-on écrire :

$$\int\Big(\int f(x,y)dy\Big)dx = \int\Big(\int f(x,y)dx\Big)dy \ \ ?$$

Une réponse à cette question est apportée par le théorème de Fubini, que nous verrons dans ce chapitre.

On introduira aussi le produit de convolution de deux fonctions, qui sera utile, par exemple, pour démontrer des théorèmes de densité. Il est aussi utilisé pour la résolution d'équations aux dérivées partielles. La convolution est également une notion très utilisée pour de nombreuses applications, par exemple en traitement de signal et d'image.

7.2 Mesure produit

On rappelle ici qu'un espace mesuré (E,T,m) est σ-fini (on dit aussi que m est σ-finie) s'il existe une famille $(A_n)_{n\in\mathbb{N}} \subset T$ telle que $E = \bigcup_{n\in\mathbb{N}} A_n$ et $m(A_n) < +\infty$, pour tout $n \in \mathbb{N}$. L'espace mesuré $(\mathbb{R},\mathcal{B}(\mathbb{R}),\lambda)$ est σ-fini (prendre, par exemple, $A_n = [-n,n]$). Il existe par contre des mesures non σ-finies. L'exemple le plus simple sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ consiste à prendre $m(A) = +\infty$ pour tout $A \in \mathcal{B}(\mathbb{R})$, $A \neq \emptyset$. Un exemple plus intéressant (intervenant pour certains problèmes) consiste à se donner un borélien non vide B de $\mathbb{R}$ (B peut être, par exemple, réduit à un point) et à définir m_B par $m_B(A) = +\infty$ si $A \in \mathcal{B}(\mathbb{R})$ et $A \cap B \neq \emptyset$ et $m_B(A) = 0$ si $A \in \mathcal{B}(\mathbb{R})$ et $A \cap B = \emptyset$.

Définition 7.1 (Tribu produit) *Soient* (E_1,T_1) *et* (E_2,T_2) *des espaces mesurables. On pose* $E = E_1 \times E_2$. *On appelle tribu produit la tribu sur* E *engendrée par* $T_1 \times T_2 = \{A_1 \times A_2,\ A_1 \in T_1,\ A_2 \in T_2\}$. *Cette tribu produit est notée* $T_1 \otimes T_2$.

Un exemple fondamental est $(E_1,T_1) = (E_2,T_2) = (\mathbb{R},\mathcal{B}(\mathbb{R}))$. On va montrer que, dans ce cas, $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

Proposition 7.2 (Tribu $\mathcal{B}(\mathbb{R}^N)$) *Pour tout* $N \geq 2$, *on a* $\mathcal{B}(\mathbb{R}^{N-1}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^N)$.

DÉMONSTRATION – La démonstration est faite pour N = 2 dans l'exercice 2.6). Elle s'adapte facilement pour traiter aussi le cas N > 2 (exercice 7.1). ■

Théorème 7.3 (Mesure produit) *Soient* (E_1,T_1,m_1) *et* (E_2,T_2,m_2) *deux espaces mesurés σ-finis,* $E = E_1 \times E_2$ *et* $T = T_1 \otimes T_2$. *Alors, il existe une et une seule mesure*

m sur T *vérifiant :*

$$m(A_1 \times A_2) = m_1(A_1)m_2(A_2),\ \forall (A_1, A_2) \in T_1 \times T_2; m_i(A_i) < \infty,\ i = 1, 2. \quad (7.1)$$

Cette mesure est notée $m = m_1 \otimes m_2$. De plus, m est σ-finie.

DÉMONSTRATION –

Existence de m. On va construire une mesure m sur T vérifiant (7.1).

Soit A ∈ T. On va montrer, à l'étape 1, que, pour tout $x_1 \in E_1$, on a $1_A(x_1, \cdot) \in \mathcal{M}_+(E_2, T_2)$. On pourra donc poser $f_A(x_1) = \int 1_A(x_1, \cdot) dm_2$, pour tout $x_1 \in E_1$. L'application f_A sera donc une application de E_1 dans $\overline{\mathbb{R}}_+$. On va montrer, à l'étape 2, que $f_A \in \mathcal{M}_+(E_1, T_1)$. On posera alors $m(A) = \int f_A dm_1$. Enfin, il restera à l'étape 3 à montrer que m est bien une mesure vérifiant (7.1) et que m est σ-finie.

Etape 1. Pour $A \in \mathcal{P}(E)$ et $x_1 \in E_1$, on note $S(x_1, A) = \{x_2 \in E_2\,;\,(x_1, x_2) \in A\} \subset E_2$, de sorte que $1_A(x_1, \cdot) = 1_{S(x_1, A)}$.

Soit $x_1 \in E_1$. On pose $\Theta = \{A \in \mathcal{P}(E)\,;\, S(x_1, A) \in T_2\}$.

On remarque tout d'abord que $\Theta \supset T_1 \times T_2$. En effet, si $A = A_1 \times A_2$ avec $A_1 \in T_1$ et $A_2 \in T_2$, on a $S(x_1, A) = A_2 \in T_2$ si $x_1 \in A_1$ et $S(x_1, A) = \emptyset \in T_2$ si $x_1 \notin A_1$.

On remarque ensuite que Θ est une tribu. En effet :

- $\emptyset \in \Theta$ car $S(x_1, \emptyset) = \emptyset \in T_2$,
- Θ est stable par passage au complémentaire. En effet : $S(x_1, A^c) = (S(x_1, A))^c$ (c'est-à-dire $S(x_1, E \setminus A) = E_2 \setminus S(x_1, A)$). On a donc $S(x_1, A^c) \in T_2$ si $A \in \Theta$, ce qui prouve que $A^c \in \Theta$.
- Θ est stable par union dénombrable. Il suffit de remarquer que $S(x_1, \bigcup_{n\in\mathbb{N}} A^{(n)}) = \bigcup_{n\in\mathbb{N}} S(x_1, A^{(n)}) \in T_2$ si $(A^{(n)})_{n\in\mathbb{N}} \subset \Theta$.

L'ensemble Θ est donc une tribu contenant $T_1 \times T_2$, et contient donc $T_1 \otimes T_2 = T$, tribu engendrée par $T_1 \times T_2$. . On a donc $S(x_1, A) \in T_2$ pour tout A ∈ T.

Pour tout A ∈ T, on peut donc définir une application f_A de E_1 dans $\overline{\mathbb{R}}_+$ en posant, pour $x_1 \in E_1$,

$$f_A(x_1) = m_2(S(x_1, A)) = \int 1_{S(x_1,A)} dm_2 = \int 1_A(x_1, \cdot) dm_2 \in \overline{\mathbb{R}}_+. \quad (7.2)$$

Etape 2. Dans cette étape, on démontre que $f_A \in \mathcal{M}_+(E_1, T_1)$ pour tout A ∈ T. Cette étape est plus difficile que la précédente.

On note $\Sigma = \{A \in T\,;\, f_A \in \mathcal{M}_+(E_1, T_1)\}$ et on va montrer que $\Sigma \supset T$ et donc que $\Sigma = T$.

On suppose d'abord que m_2 est finie.

Il est facile de voir que Σ contient $T_1 \times T_2$. En effet, si $A = A_1 \times A_2$ avec $A_1 \in T_1$ et $A_2 \in T_2$, on a alors $f_A = m_2(A_2) 1_{A_1} \in \mathcal{E}_+(E_1, T_1) \subset \mathcal{M}_+(E_1, T_1)$.

On note maintenant $\mathcal{A}$ l'ensemble des réunions finies disjointes d'éléments de $T_1 \times T_2$ ($\mathcal{A}$ s'appelle l'algèbre engendrée par $T_1 \times T_2$, voir l'exercice 7.2). Si $A \in \mathcal{A}$, il existe donc $(A^{(p)})_{p=1,\ldots,n} \subset T_1 \times T_2$ tel que $A^{(p)} \cap A^{(q)} = \emptyset$ si $p \neq q$ et $A = \bigcup_{p=1}^{n} A^{(p)}$. On a alors $f_A(x_1) = m_2(S(x_1, A)) = \sum_{p=1}^{n} m_2(S(x_1, A^{(p)})) = \sum_{p=1}^{n} f_{A^{(p)}} \in \mathcal{M}_+(E_1, T_1)$ car $A^{(p)} \in T_1 \times T_2 \subset \Sigma$. On a donc $\mathcal{A} \subset \Sigma$.

On montre maintenant que Σ est une classe monotone, c'est-à-dire que :

$$(A^{(n)})_{n\in\mathbb{N}} \subset \Sigma,\ A^{(n)} \subset A^{(n+1)}\ \forall n \in \mathbb{N} \Rightarrow \bigcup_{n\in\mathbb{N}} A^{(n)} \in \Sigma \tag{7.3}$$

et

$$(A^{(n)})_{n\in\mathbb{N}} \subset \Sigma,\ A^{(n)} \supset A^{(n+1)}\ \forall n \in \mathbb{N} \Rightarrow \bigcap_{n\in\mathbb{N}} A^{(n)} \in \Sigma. \tag{7.4}$$

Pour démontrer (7.3), on considère une suite $(A^{(n)})_{n\in\mathbb{N}} \subset \Sigma$ telle que $A^{(n)} \subset A^{(n+1)}$ pour tout $n \in \mathbb{N}$. On pose $A = \bigcup_{n\in\mathbb{N}} A^{(n)}$. Soit $x_1 \in E_1$; on a $(S(x_1, A^{(n)}))_{n\in\mathbb{N}} \subset T_2$ (par l'étape 1, car $\Sigma \subset T$), $S(x_1, A^{(n)}) \subset S(x_1, A^{(n+1)})$ pour tout $n \in \mathbb{N}$ et

$$S\left(x_1, \cup_{n\in\mathbb{N}} A^{(n)}\right) = \bigcup_{n\in\mathbb{N}} S(x_1, A^{(n)}).$$

On en déduit, par continuité croissante de m_2, que

$$m_2(S(x_1, A)) = \sup_{n\in\mathbb{N}} m_2(S(x_1, A^{(n)}))$$

et donc que $f_A = \sup_{n\in\mathbb{N}} f_{A^{(n)}}$, ce qui prouve que $f_A \in \mathcal{M}_+(E_1, T_1)$ car $f_{A^{(n)}} \in \mathcal{M}_+(E_1, T_1)$ pour tout $n \in \mathbb{N}$. On a donc $A = \bigcup_{n\in\mathbb{N}} A^{(n)} \in \Sigma$.

La démonstration de (7.4) est similaire, il faut utiliser la continuité décroissante de m_2 au lieu de la continuité croissante. C'est pour utiliser la continuité décroissante de m_2 qu'on a besoin du fait que m_2 est finie.

On a ainsi montré que Σ est une classe monotone contenant l'algèbre $\mathcal{A}$. On peut en déduire (cela fait l'objet de l'exercice 2.13) que Σ contient la tribu engendrée par $\mathcal{A}$ et donc aussi la tribu engendrée par $T_1 \times T_2$ (car $T_1 \times T_2 \subset \mathcal{A}$), c'est-à-dire que Σ contient $T = T_1 \otimes T_2$. On a bien montré, finalement, que $\Sigma = T$.

Il reste maintenant à montrer que $\Sigma = T$ sans l'hypothèse m_2 finie. Comme m_2 est σ-finie, on peut construire une suite $(F_n)_{n\in\mathbb{N}} \subset T_2$ telle que $F_n \subset F_{n+1}$ et $m_2(F_n) < \infty$ pour tout $n \in \mathbb{N}$. Pour $n \in \mathbb{N}$, on définit alors la mesure $m_2^{(n)}$ par $m_2^{(n)}(A_2) = m_2(A_2 \cap F_n)$ pour tout $A_2 \in T_2$. La mesure $m_2^{(n)}$ est finie, l'étape 1 et la première partie de l'étape 2 donne donc que, pour tout $A \in T$, $f_A^{(n)} \in \mathcal{M}_+(E_1, T_1)$ où $f_A^{(n)}$ est définie par 7.2 avec $m_2^{(n)}$ au lieu de m_2 (c'est-à-dire $f_A^{(n)}(x_1) = m_2^{(n)}(S(x_1, A))$ pour tout $x_1 \in E_1$). On conclut alors en remarquant que $f_A^{(n)} \uparrow f_A$ quand $n \to +\infty$, ce qui donne que $f_A \in \mathcal{M}_+(E_1, T_1)$.

On a donc montré que $f_A \in \mathcal{M}_+(E_1, T_1)$ pour tout $A \in T$. Ceci nous permet de définir $m : T \to \overline{\mathbb{R}}_+$ par :

$$m(A) = \int f_A \, dm_1, \text{ pour tout } A \in T. \tag{7.5}$$

Etape 3. Dans cette étape, on montre que m, définie par (7.5), est une mesure sur T et que m vérifie (7.1) et est σ-finie.

On montre d'abord que m est bien une mesure sur T :

- $m(\emptyset) = 0$ car $f_\emptyset(x_1) = m_2(\mathrm{S}(x_1, \emptyset)) = m_2(\emptyset) = 0$.
- (σ-additivité de m) Soit $(\mathrm{A}^{(n)})_{n\in\mathbb{N}} \subset \mathrm{T}$ telle que $\mathrm{A}^{(n)} \cap \mathrm{A}^{(m)} = \emptyset$ si $n \neq m$. On pose $\mathrm{A} = \bigcup_{n\in\mathbb{N}} \mathrm{A}^{(n)}$. Pour $x_1 \in \mathrm{E}_1$, on a :
$$\mathrm{S}(x_1, \mathrm{A}) = \bigcup_{n\in\mathbb{N}} \mathrm{S}(x_1, \mathrm{A}^{(n)}) \text{ et } \mathrm{S}(x_1, \mathrm{A}^{(n)}) \cap \mathrm{S}(x_1, \mathrm{A}^{(m)}) = \emptyset \text{ si } n \neq m.$$
La σ-additivité de m_2 donne alors $m_2(\mathrm{S}(x_1, \mathrm{A})) = \sum_{n\in\mathbb{N}} m_2(\mathrm{S}(x_1, \mathrm{A}^{(n)}))$, c'est-à-dire
$$f_\mathrm{A}(x_1) = \sum_{n\in\mathbb{N}} f_{\mathrm{A}^{(n)}}(x_1).$$
Le premier corollaire du théorème de convergence monotone (corollaire 4.18) donne alors :
$$m(\mathrm{A}) = \int f_\mathrm{A} dm_1 = \sum_{n\in\mathbb{N}} \int f_{\mathrm{A}^{(n)}} dm_1 = \sum_{n\in\mathbb{N}} m(\mathrm{A}^{(n)}),$$
ce qui donne la σ-additivité de m.

On montre maintenant que m vérifie (7.1). Soient $\mathrm{A}_1 \in \mathrm{T}_1$ et $\mathrm{A}_2 \in \mathrm{T}_2$ tels que $m_1(\mathrm{A}_1) < \infty$ et $m(\mathrm{A}_2) < \infty$. On pose $\mathrm{A} = \mathrm{A}_1 \times \mathrm{A}_2$. On a alors $f_\mathrm{A} = m_2(\mathrm{A}_2)1_{\mathrm{A}_1}$ et donc $m(\mathrm{A}) = \int f_\mathrm{A} dm_1 = m_2(\mathrm{A}_2)m_1(\mathrm{A}_1)$.

Il reste à vérifier que m est σ-finie. Comme m_1 et m_2 sont σ-finies, il existe $(\mathrm{B}_1^{(n)})_{n\in\mathbb{N}} \subset \mathrm{T}_1$ et $(\mathrm{B}_2^{(n)})_{n\in\mathbb{N}} \subset \mathrm{T}_2$ tels que $\mathrm{E}_1 = \bigcup_{n\in\mathbb{N}} \mathrm{B}_1^{(n)}$, $\mathrm{E}_2 = \bigcup_{n\in\mathbb{N}} \mathrm{B}_2^{(n)}$ et, pour tout $n \in \mathbb{N}$, $m_1(\mathrm{B}_1^{(n)}) < \infty$ et $m_2(\mathrm{B}_2^{(n)}) < \infty$. Pour $(n, m) \in \mathbb{N}^2$, on pose $\mathrm{C}_{n,m} = \mathrm{B}_1^{(n)} \times \mathrm{B}_2^{(m)}$, de sorte que $\mathrm{E} = \bigcup_{(n,m)\in\mathbb{N}^2} \mathrm{C}_{n,m}$ et $m(\mathrm{C}_{n,m}) = m_1(\mathrm{B}_1^{(n)}) \times m_2(\mathrm{B}_2^{(m)}) < \infty$. Comme $\mathbb{N}^2$ est dénombrable, on en déduit que m est σ-finie.

Unicité de m.

La partie existence de la démonstration donne une mesure m sur T vérifiant (7.1). La partie unicité du théorème peut se montrer avec la proposition 2.31 ; nous développons cette méthode ci-après, ou avec le lemme des classes monotones (exercice 2.13) comme cela est expliqué dans la remarque 7.4.

Soit m et μ deux mesures sur T vérifiant (7.1). Pour montrer que $m = \mu$, on va appliquer la proposition 2.31. On pose :
$$\mathcal{C} = \{\mathrm{A}_1 \times \mathrm{A}_2, \mathrm{A}_1 \in \mathrm{T}_1, \mathrm{A}_2 \in \mathrm{T}_2, m_1(\mathrm{A}_1) < \infty, m_2(\mathrm{A}_2) < \infty\}.$$
Comme m_1 et m_2 sont σ−finies, il est facile de montrer que tout élément de $\mathrm{T}_1 \times \mathrm{T}_2$ est une réunion dénombrable d'éléments de $\mathcal{C}$. On en déduit que $\mathcal{C}$ engendre T. Il est clair que $\mathcal{C}$ est stable par intersection finie et, par (7.1), on a $m = \mu$ sur $\mathcal{C}$. Puis, comme m_1 et m_2 sont σ−finies, il existe deux suites $(\mathrm{E}_{1,n})_{n\in\mathbb{N}} \subset \mathrm{T}_1$ et $(\mathrm{E}_{2,n})_{n\in\mathbb{N}} \subset \mathrm{T}_2$ d'éléments de T_1 et T_2, disjoints deux à deux et t.q. $\mathrm{E}_1 = \bigcup_{n\in\mathbb{N}} \mathrm{E}_{1,n}$, $\mathrm{E}_2 = \bigcup_{n\in\mathbb{N}} \mathrm{E}_{2,n}$ et $m_i(\mathrm{E}_{i,n}) < \infty$ pour tout $i \in \{1, 2\}$ et tout $n \in \mathbb{N}$. Pour $n, m \in \mathbb{N}$, on pose $\mathrm{F}_{n,m} = \mathrm{E}_{1,n} \times \mathrm{E}_{2,m}$. La famille

$(F_{n,m})_{n,m\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathcal{C}$, disjoints deux à deux et t.q. $E = \bigcup_{n,m\in\mathbb{N}} F_{n,m}$ et $m(F_{n,m}) = m_1(E_{1,n})m_2(E_{1,m}) < \infty$. On peut alors utiliser la Proposition 2.31. Elle donne $m = \mu$ sur T et termine la démonstration du théorème.

■

Remarque 7.4 Comme cela a été dit, un autre moyen de montrer la partie unicité du théorème précédent est d'utiliser le lemme des classes monotones (exercice 2.13). Supposons tout d'abord que m_1 et m_2 sont finies. On a alors (par (7.1)) :

$$m(E) = \mu(E) = m_1(E_1)m_2(E_2) < \infty.$$

La condition (7.1) donne également que $m = \mu$ sur $T_1 \times T_2$. On a alors aussi $m = \mu$ sur l'algèbre engendrée par $T_1 \times T_2$, notée $\mathcal{A}$ (cette algèbre a été définie dans la partie existence de la démonstration). En effet, si $A \in \mathcal{A}$, il existe $(A^{(p)})_{p=1,\ldots,n} \subset T_1 \times T_2$ t.q. $A^{(p)} \cap A^{(q)} = \emptyset$ si $p \neq q$ et $A = \bigcup_{p=1}^{n} A^{(p)}$. On a alors, par additivité de m et μ, $m(A) = \sum_{n\in\mathbb{N}} m(A^{(n)}) = \sum_{n\in\mathbb{N}} \mu(A^{(n)}) = \mu(A)$.

On pose maintenant $\Sigma = \{A \in T\,;\, m(A) = \mu(A)\}$. On vient de montrer que $\Sigma \supset \mathcal{A}$. Il est d'autre part facile de voir que Σ est une classe monotone. En effet, les propriétés de continuité croissante et de continuité décroissante appliquées à m et μ permettent facilement de vérifier (7.3) et (7.4) (on utilise ici, pour montrer (7.4), que m et μ sont des mesures finies). Comme dans la partie existence de la démonstration, l'exercice 2.13 donne alors que Σ contient la tribu engendrée par $\mathcal{A}$ et donc que Σ contient $T = T_1 \otimes T_2$, ce qui donne $\Sigma = T$ et donc $m = \mu$.

Dans le cas où m_1 et m_2 ne sont pas finies, mais σ-finies, il existe $(B_1^{(n)})_{n\in\mathbb{N}} \subset T_1$ et $(B_2^{(n)})_{n\in\mathbb{N}} \subset T_2$ t.q. $E_1 = \bigcup_{n\in\mathbb{N}} B_1^{(n)}$, $E_2 = \bigcup_{n\in\mathbb{N}} B_2^{(n)}$ et, pour tout $n \in \mathbb{N}$, $m_1(B_1^{(n)}) < \infty$ et $m_2(B_2^{(n)}) < \infty$. On peut également supposer que $B_1^{(n)} \subset B_1^{(n+1)}$ et $B_2^{(n)} \subset B_2^{(n+1)}$ pour tout $n \in \mathbb{N}$ (il suffit, par exemple, de remplacer $B_i^{(n)}$ par $\bigcup_{p=0}^{n} B_i^{(p)}$). Par un raisonnement analogue à celui fait dans le cas où m_1 et m_2 sont finies, on peut montrer que $m = \mu$ sur $\{A \in T\,;\, A \subset B_1^{(n)} \times B_2^{(n)}\}$. On conclut alors, en utilisant la propriété de continuité croissante, que $m = \mu$ sur T.

Remarque 7.5 Dans le théorème précédente (théorème 7.3), on peut aussi remarquer que :

1. $m(A_1 \times A_2) = m_1(A_1)m_2(A_2) = \infty$ si $A_1 \in T_1$ et $A_2 \in T_2$ avec $m_1(A_1) \neq 0$ et $m(A_2) = \infty$ (ou avec $m_1(A_1) = \infty$ et $m_2(A_2) \neq 0$),
2. $m(A_1 \times A_2) = 0$ si $A_1 \in T_1$ et $A_2 \in T_2$ avec $m_1(A_1) = 0$ et $m(A_2) = \infty$ (ou avec $m_1(A_1) = \infty$ et $m_2(A_2) = 0$).

En effet, on suppose par exemple que $m_1(A_1) = 0$ et $m(A_2) = \infty$. Comme m_2 est σ-finie, on peut construire une suite $(F_n)_{n\in\mathbb{N}} \subset T_2$ t.q. $F_n \subset F_{n+1}$ et $m_2(F_n) < \infty$ pour tout $n \in \mathbb{N}$. On a alors, par continuité croissante de m, $m(A_1 \times A_2) = \lim_{n\to+\infty} m(A_1 \times (A_2 \cap F_n)) = \lim_{n\to+\infty} m_1(A_1)m_2(A_2 \cap F_n) = 0$ (on a d'ailleurs aussi $m_2(A_2 \cap F_n) \uparrow \infty$, ce qui permet de conclure si $0 < m_1(A_1) < \infty$ que $m(A_1 \times A_2) = \infty$). Les autres cas se traitent de manière analogue.

Définition 7.6 (Espace produit)

L'espace (E, T, m)*, construit dans le théorème 7.3, s'appelle l'espace (mesuré) produit des espaces* (E_1, T_1, m_1) *et* (E_2, T_2, m_2).

Un exemple fondamental d'espace produit est l'espace $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ pour $N \geq 2$ que nous verrons dans la section 7.4.

7.3 Théorèmes de Fubini-Tonelli et Fubini

Théorème 7.7 (Fubini-Tonelli) *Soient* (E_1, T_1, m_1) *et* (E_2, T_2, m_2) *des espaces mesurés σ-finis. On note* (E, T, m) *l'espace produit (donc,* $T = T_1 \otimes T_2$ *et* $m = m_1 \otimes m_2$*). Soit* $f : E \to \overline{\mathbb{R}}_+$ *une fonction mesurable positive (i.e.* T*-mesurable positive). Alors :*

1. $f(x_1, \cdot) \in \mathcal{M}_+(E_2, T_2)$ *pour tout* $x_1 \in E_1$,

   on pose

$$\varphi_f(x_1) = \int f(x_1, \cdot)dm_2 = \int f(x_1, x_2)dm_2(x_2) \text{ pour tout } x_1 \in E_1,$$

   de sorte que $\varphi_f : E_1 \to \overline{\mathbb{R}}_+$,

2. $\varphi_f \in \mathcal{M}_+(E_1, T_1)$,
3. $\displaystyle\int f dm = \int \varphi_f dm_1 = \int \Big(\int f(x_1, x_2)dm_2(x_2) \Big) dm_1(x_1)$,
4. *les mêmes résultats sont vrais en inversant les rôles de* m_1 *et* m_2*, de sorte que :*

$$\int \Big(\int f(x_1, x_2)dm_2(x_2) \Big) dm_1(x_1) = \int \Big(\int f(x_1, x_2)dm_1(x_1) \Big) dm_2(x_2).$$

DÉMONSTRATION – la démonstration se fait en plusieurs étapes.

Etape 1. Soit $f = 1_A$, $A \in T$. La partie existence de m de la démonstration du théorème 7.3 donne alors que $\int f dm = m(A) = \int \varphi_f dm_1$.

Plus précisément, on a, pour tout $x_1 \in \mathrm{E}_1$, $f(x_1,\cdot) = 1_{\mathrm{A}}(x_1,\cdot) = 1_{\mathrm{S}(x_1,\mathrm{A})}$, avec

$$\mathrm{S}(x_1,\mathrm{A}) = \{x_2 \in \mathrm{E}_2 \text{ t.q. } (x_1,x_2) \in \mathrm{A}\} \subset \mathrm{E}_2$$

(comme dans la démonstration du théorème 7.3). L'étape 1 de la démonstration (de la partie existence) du théorème 7.3 donne que $\mathrm{S}(x_1,\mathrm{A}) \in \mathrm{T}_2$ pour tout $x_1 \in \mathrm{E}_1$, et donc $f(x_1,\cdot) \in \mathcal{M}_+(\mathrm{E}_2,\mathrm{T}_2)$. Ceci donne le premier item (pour $f = 1_{\mathrm{A}}$) de la conclusion du théorème 7.7.

On pose $\varphi_f(x_1) = \int f(x_1,\cdot)dm_2 = m_2(\mathrm{S}(x_1,\mathrm{A}))$ pour tout $x_1 \in \mathrm{E}_1$. (Cette fonction φ_f est notée f_{A} dans la démonstration du théorème 7.3). L'étape 2 de la démonstration du théorème 7.3 donne que $\varphi_f \in \mathcal{M}_+(\mathrm{E}_1,\mathrm{T}_1)$. Ceci donne le deuxième item (pour $f = 1_{\mathrm{A}}$) de la conclusion du théorème 7.7.

On a alors posé, dans la démonstration du théorème 7.3, $m(\mathrm{A}) = \int \varphi_f dm_1$ et l'étape 3 a montré que m est une mesure sur T vérifiant (7.1) (et la seule mesure sur T vérifiant (7.1), d'après la partie unicité de la démonstration du théorème 7.3). Ceci donne le troisième item (pour $f = 1_{\mathrm{A}}$) de la conclusion du théorème 7.7.

Pour avoir le quatrième item (pour $f = 1_{\mathrm{A}}$) de la conclusion du théorème 7.7, il suffit de remarquer que l'on peut inverser les rôles de m_1 et m_2 dans la démonstration du théorème 7.7. On obtient ainsi que $f(\cdot,x_2) \in \mathcal{M}_+(\mathrm{E}_1,\mathrm{T}_1)$ pour tout $x_2 \in \mathrm{E}_2$. On pose alors $\psi_f(x_2) = \int f(\cdot,x_2)dm_1$. On obtient que $\psi_f \in \mathcal{M}_+(\mathrm{E}_2,\mathrm{T}_2)$. Enfin, on pose $\tilde{m}(\mathrm{A}) = \int \psi_f dm_2$ et on obtient que $\tilde{m}$ est une mesure sur T vérifiant (7.1). La partie unicité de la démonstration du théorème 7.3 donne alors que $m = \tilde{m}$, ce qui est exactement le quatrième item (pour $f = 1_{\mathrm{A}}$) de la conclusion du théorème 7.7.

Etape 2. On prend maintenant $f \in \mathcal{E}_+(\mathrm{E},\mathrm{T})$.
Il existe donc $a_1,\ldots,a_n \in \mathbb{R}_+^*$ et $\mathrm{A}_1,\ldots,\mathrm{A}_n \in \mathrm{T}$ t.q. $f = \sum_{i=1}^n a_i 1_{\mathrm{A}_i}$.

On a alors, pour tout $x_1 \in \mathrm{E}_1$, $f(x_1,\cdot) = \sum_{i=1}^n a_i 1_{\mathrm{A}_i}(x_1,\cdot) \in \mathcal{M}_+(\mathrm{E}_2,\mathrm{T}_2)$ car l'étape 1 donne $1_{\mathrm{A}_i}(x_1,\cdot) \in \mathcal{M}_+(\mathrm{E}_2,\mathrm{T}_2)$ pour tout i. Ce qui donne le premier item de la conclusion du théorème 7.7.

On pose $\varphi_f(x_1) = \int f(x_1,\cdot)dm_2$ pour tout $x_1 \in \mathrm{E}_1$. On a $\varphi_f \in \mathcal{M}_+(\mathrm{E}_1,\mathrm{T}_1)$ car $\varphi_f = \sum_{i=1}^n a_i \varphi_{1_{\mathrm{A}_i}}$ et que $\varphi_{1_{\mathrm{A}_i}} \in \mathcal{M}_+(\mathrm{E}_1,\mathrm{T}_1)$ pour tout i (d'après l'étape 1), ce qui donne le deuxième item de la conclusion du théorème 7.7.

Enfin, on utilise la linéarité de l'intégrale et l'étape 1 pour $f = 1_{\mathrm{A}_i}$, on obtient :

$$\begin{aligned}
\int f dm &= \sum_{i=1}^n a_i m(\mathrm{A}_i) = \sum_{i=1}^n a_i \int \varphi_{1_{\mathrm{A}_i}} dm_1 = \int (\sum_{i=1}^n a_i \varphi_{1_{\mathrm{A}_i}}) dm_1 \\
&= \int \Big(\sum_{i=1}^n a_i \int 1_{\mathrm{A}_i}(x_1,\cdot)dm_2 \Big) dm_1(x_1) = \int \Big(\int f(x_1,\cdot)dm_2 \Big) dm_1(x_1) \\
&= \int \varphi_f dm_1.
\end{aligned}$$

Ce qui donne le troisième item de la conclusion du théorème 7.7.

Pour avoir le quatrième item de la conclusion du théorème 7.7, il suffit de changer les rôles de m_1 et m_2.

Etape 3. On peut enfin prendre $f \in \mathcal{M}_+(E, T)$. Il existe une suite $(f_n)_{n\in\mathbb{N}} \in \mathcal{E}_+(E, T)$ t.q. $f_n \uparrow f$ quand $n \to +\infty$.

On a donc, pour tout $x_1 \in E_1$, $f_n(x_1,\cdot) \uparrow f(x_1,\cdot)$ quand $n \to +\infty$. On en déduit que $f(x_1,\cdot) \in \mathcal{M}_+(E_2, T_2)$ car (d'après l'étape 2) $f_n(x_1,\cdot) \in \mathcal{M}_+(E_2, T_2)$ pour tout $n \in \mathbb{N}$ (ce qui donne le premier item).

Le théorème de convergence monotone (pour m_2) donne que

$$\varphi_{f_n}(x_1) = \int f_n(x_1,\cdot)dm_2 \uparrow \int f(x_1,\cdot)dm_2 = \varphi_f(x_1)$$

pour tout $x_1 \in E_1$. Donc, $\varphi_{f_n} \uparrow \varphi_f$. Comme $\varphi_{f_n} \in \mathcal{M}_+(E_1, T_1)$ (d'après l'étape 2), on en déduit que $\varphi_f \in \mathcal{M}_+(E_1, T_1)$ (ce qui donne le deuxième item).

On applique maintenant le théorème de convergence monotone pour m_1 et pour m, ils donnent :

$$\int \varphi_{f_n} dm_1 \uparrow \int \varphi_f dm_1 \text{ et } \int f_n dm \uparrow \int f dm \text{ quand } n \to +\infty.$$

L'étape 2 donne $\int f_n dm = \int \varphi_{f_n} dm_1$, on en déduit donc que $\int f dm = \int \varphi_f dm_1$, ce qui donne le troisième item de la conclusion du théorème 7.7.

Enfin, ici encore, pour avoir le quatrième item de la conclusion du théorème 7.7, il suffit de changer les rôles de m_1 et m_2. ■

Corollaire 7.8 *Soient* (E_1, T_1, m_1) *et* (E_2, T_2, m_2) *des espaces mesurés σ-finis. On note* (E, T, m) *l'espace produit. Soit* $f : E \to \mathbb{R}$ *une fonction* T*-mesurable. Alors :*

$$f \in \mathcal{L}^1_{\mathbb{R}}(E, T, m) \Leftrightarrow \int \Big(\int |f| dm_2\Big) dm_1 < +\infty \Leftrightarrow \int \Big(\int |f| dm_1\Big) dm_2 < +\infty. \quad (7.6)$$

DÉMONSTRATION – Le corollaire découle immédiatement du théorème 7.7 appliqué à la fonction $|f|$ qui appartient à $\mathcal{M}_+(E, T)$. Dans (7.6), la notation $(\int |f| dm_2) dm_1$ signifie :

$$\Big(\int |f(x_1, x_2)| dm_2(x_2)\Big) dm_1(x_1).$$

La notation est similaire en inversant les rôles de m_1 et m_2. ■

Voici une conséquence immédiate du théorème 7.7 pour la mesurabilité :

Proposition 7.9 *Soient* (E_1, T_1) *et* (E_2, T_2) *deux espaces mesurables. On pose* $E = E_1 \times E_2$ *et* $T = T_1 \otimes T_2$. *Soit* $f \in \mathcal{M}(E, T)$ *(c'est-à-dire* $f : E \to \mathbb{R}$, T*-mesurable). Alors :*

1. $f(x_1,\cdot) \in \mathcal{M}(E_2, T_2)$, *pour tout* $x_1 \in E_1$,

2. $f(\cdot, x_2) \in \mathcal{M}(E_1, T_1)$, *pour tout* $x_2 \in E_2$.

DÉMONSTRATION – La démonstration est facile, il suffit de remarquer que $f = f^+ - f^-$ et que $f^+, f^- \in \mathcal{M}_+(\mathrm{E}, \mathrm{T})$. Le premier item de la conclusion du théorème 7.7 donne alors, pour tout $x_1 \in \mathrm{E}_1$, $f^+(x_1, \cdot) \in \mathcal{M}_+(\mathrm{E}_2, \mathrm{T}_2)$ et $f^-(x_1, \cdot) \in \mathcal{M}_+(\mathrm{E}_2, \mathrm{T}_2)$. Comme $f(x_1, \cdot) = f^+(x_1, \cdot) - f^-(x_1, \cdot)$, on en déduit que $f(x_1, \cdot) \in \mathcal{M}(\mathrm{E}_2, \mathrm{T}_2)$. En changeant les rôles de $(\mathrm{E}_1, \mathrm{T}_1)$ et $(\mathrm{E}_2, \mathrm{T}_2)$, on montre aussi que $f(\cdot, x_2) \in \mathcal{M}(\mathrm{E}_1, \mathrm{T}_1)$, pour tout $x_2 \in \mathrm{E}_2$. ■

Remarque 7.10 La réciproque de la proposition précédente est fausse. Soient $(\mathrm{E}_1, \mathrm{T}_1)$ et $(\mathrm{E}_2, \mathrm{T}_2)$ deux espaces mesurables, $\mathrm{E} = \mathrm{E}_1 \times \mathrm{E}_2$ et $\mathrm{T} = \mathrm{T}_1 \otimes \mathrm{T}_2$. Soit $f : \mathrm{E} \to \mathbb{R}$ t.q.

1. $f(x_1, \cdot) \in \mathcal{M}(\mathrm{E}_2, \mathrm{T}_2)$, pour tout $x_1 \in \mathrm{E}_1$,
2. $f(\cdot, x_2) \in \mathcal{M}(\mathrm{E}_1, \mathrm{T}_1)$, pour tout $x_2 \in \mathrm{E}_2$.

Alors, f n'est pas forcément T-mesurable. Un exemple est donné dans l'exercice 7.4. Un cas particulier intéressant pour laquelle cette réciproque est vraie est donné par la proposition 7.11.

Proposition 7.11 *Soient* $(\mathrm{E}_1, \mathrm{T}_1)$ *et* $(\mathrm{E}_2, \mathrm{T}_2)$ *deux espaces mesurables. On pose* $\mathrm{E} = \mathrm{E}_1 \times \mathrm{E}_2$ *et* $\mathrm{T} = \mathrm{T}_1 \otimes \mathrm{T}_2$. *Soient* $\mathrm{F}_1 \in \mathcal{M}(\mathrm{E}_1, \mathrm{T}_1)$ *et* $\mathrm{F}_2 \in \mathcal{M}(\mathrm{E}_2, \mathrm{T}_2)$. *On définit* $f : \mathrm{E} \to \mathbb{R}$ *par* $f(x_1, x_2) = \mathrm{F}_1(x_1)\mathrm{F}_2(x_2)$ *pour tout* $(x_1, x_2) \in \mathrm{E}$. *Alors* f *est* T*-mesurable (c'est-à-dire* $f \in \mathcal{M}(\mathrm{E}, \mathrm{T})$*).*

DÉMONSTRATION – On procède en trois étapes.

Etape 1. On prend d'abord $\mathrm{F}_1 = 1_{\mathrm{A}_1}$ et $\mathrm{F}_2 = 1_{\mathrm{A}_2}$ avec $\mathrm{A}_1 \in \mathrm{T}_1$ et $\mathrm{A}_2 \in \mathrm{T}_2$. On a alors $f = 1_{\mathrm{A}_1 \times \mathrm{A}_2} \in \mathcal{M}(\mathrm{E}, \mathrm{T})$ car $\mathrm{A}_1 \times \mathrm{A}_2 \in \mathrm{T}_1 \times \mathrm{T}_2 \subset \mathrm{T}_1 \otimes \mathrm{T}_2 = \mathrm{T}$.

Etape 2. On prend maintenant $\mathrm{F}_1 \in \mathcal{E}(\mathrm{E}_1, \mathrm{T}_1)$ et $\mathrm{F}_2 \in \mathcal{E}(\mathrm{E}_2, \mathrm{T}_2)$.

Il existe alors $a_1^{(1)}, \ldots, a_n^{(1)} \in \mathbb{R}$, $\mathrm{A}_1^{(1)}, \ldots, \mathrm{A}_n^{(1)} \in \mathrm{T}_1$, $a_1^{(2)}, \ldots, a_m^{(2)} \in \mathbb{R}$ et $\mathrm{A}_1^{(2)}, \ldots, \mathrm{A}_m^{(2)} \in \mathrm{T}_2$ t.q. :

$$\mathrm{F}_1 = \sum_{i=1}^{n} a_i^{(1)} \mathrm{A}_i^{(1)} \text{ et } \mathrm{A}_i^{(1)} \cap \mathrm{A}_k^{(1)} = \emptyset \text{ si } i \neq k,$$

$$\mathrm{F}_2 = \sum_{j=1}^{m} a_j^{(2)} \mathrm{A}_j^{(2)} \text{ et } \mathrm{A}_j^{(2)} \cap \mathrm{A}_k^{(1)} = \emptyset \text{ si } j \neq k.$$

On a alors $f = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i^{(1)} a_j^{(2)} 1_{\mathrm{A}_i^{(1)} \times \mathrm{A}_j^{(2)}} \in \mathcal{E}(\mathrm{E}, \mathrm{T}) \subset \mathcal{M}(\mathrm{E}, \mathrm{T})$.

Etape 3. On prend enfin $\mathrm{F}_1 \in \mathcal{M}(\mathrm{E}_1, \mathrm{T}_1)$ et $\mathrm{F}_2 \in \mathcal{M}(\mathrm{E}_2, \mathrm{T}_2)$. Il existe $(\mathrm{F}_n^{(1)})_{n \in \mathbb{N}}$ suite de $\mathcal{E}(\mathrm{E}_1, \mathrm{T}_1)$ et $(\mathrm{F}_n^{(2)})_{n \in \mathbb{N}}$ suite de $\mathcal{E}(\mathrm{E}_2, \mathrm{T}_2)$ t.q. $\mathrm{F}_n^{(1)}(x_1) \to \mathrm{F}_1(x_1)$ pour tout $x_1 \in \mathrm{E}_1$ et $\mathrm{F}_n^{(2)}(x_2) \to \mathrm{F}_2(x_2)$ pour tout $x_2 \in \mathrm{E}_2$. On en déduit que $f_n(x_1, x_2) = \mathrm{F}_n^{(1)}(x_1)\mathrm{F}_n^{(2)}(x_2) \to f(x_1, x_2)$ pour tout $(x_1, x_2) \in \mathrm{E}$ et donc que $f \in \mathcal{M}(\mathrm{E}, \mathrm{T})$ car $f_n \in \mathcal{M}(\mathrm{E}, \mathrm{T})$ pour tout $n \in \mathbb{N}$ (étape 2). ■

Théorème 7.12 (Fubini) *Soient* (E_1, T_1, m_1) *et* (E_2, T_2, m_2) *des espaces mesurés* σ*-finis. On note* (E, T, m) *l'espace produit. Soit* f *une fonction* T*-mesurable de* E *dans* $\mathbb{R}$ *(c'est-à-dire* $f \in \mathcal{M}(E, T)$*) et intégrable pour la mesure m, c'est-à-dire* $f \in \mathcal{L}^1_{\mathbb{R}}(E, T, m)$*. Alors :*

1. $f(x_1, \cdot) \in \mathcal{L}^1_{\mathbb{R}}(E_2, T_2, m_2)$ *pour presque tout* $x_1 \in E_1$,

   on pose $\varphi_f(x_1) = \int f(x_1, \cdot) dm_2$ *pour* $x_1 \in E_1$ *t.q.* $f(x_1, \cdot) \in \mathcal{L}^1_{\mathbb{R}}(E_2, T_2, m_2)$*. La fonction* φ_f *est donc définie p.p. sur* E_1 *(et à valeurs dans* $\mathbb{R}$*).*
2. $\varphi_f \in L^1_{\mathbb{R}}(E_1, T_1, m_1)$ *(au sens : il existe* $g \in \mathcal{L}^1_{\mathbb{R}}(E_1, T_1, m_1)$ *t.q.* $f = g$ *p.p.).*
3. $\displaystyle\int f dm = \int \varphi_f dm_1 = \int \Big(\int f(x_1, x_2) dm_2(x_2) \Big) dm_1(x_1),$
4. *les mêmes résultats sont vrais en inversant les rôles de* m_1 *et* m_2*, de sorte que :*

$$\int \Big(\int f(x_1, x_2) dm_2(x_2) \Big) dm_1(x_1) = \int \Big(\int f(x_1, x_2) dm_1(x_1) \Big) dm_2(x_2).$$

DÉMONSTRATION – Comme $f \in \mathcal{M}(E, T)$, on a $f^+, f^- \in \mathcal{M}_+(E, T)$. On peut donc appliquer le théorème de Fubini-Tonelli (théorème 7.7) à f^+ et f^-. Il donne :
1. $f^+(x_1, \cdot), f^-(x_1, \cdot) \in \mathcal{M}_+(E_2, T_2)$, pour tout $x_1 \in E_1$,
2. $\varphi_{f^+}, \varphi_{f^-} \in \mathcal{M}_+(E_1, T_1)$ avec $\varphi_{f^\pm}(x_1) = \int f^\pm(x_1, \cdot) dm_2$ pour tout $x_1 \in E_1$.
3. $\int f^\pm dm = \int \varphi_{f^\pm} dm_1$.

Le premier item donne que $f(x_1, \cdot) = f^+(x_1, \cdot) - f^-(x_1, \cdot) \in \mathcal{M}(E_2, T_2)$ (noter que f, f^+ et f^- sont à valeurs dans $\mathbb{R}$).

Comme $\int f^+ dm < \infty$ et $\int f^- dm < \infty$ (car $f \in \mathcal{L}^1_{\mathbb{R}}(E, T, m)$), le troisième item donne que $\varphi_{f^+} < \infty$ p.p. (sur E_1) et que $\varphi_{f^-} < \infty$ p.p. (sur E_1). On a donc $f^+(x_1, \cdot) \in \mathcal{L}^1_{\mathbb{R}}(E_2, T_2, m_2)$ et $f^-(x_1, \cdot) \in \mathcal{L}^1_{\mathbb{R}}(E_2, T_2, m_2)$ pour presque tout $x_1 \in E_1$. On en déduit donc que $f(x_1, \cdot) = f^+(x_1, \cdot) - f^-(x_1, \cdot) \in \mathcal{L}^1_{\mathbb{R}}(E_2, T_2, m_2)$ pour presque tout $x_1 \in E_1$. Ce qui donne le premier item de la conclusion.

La fonction φ_f est donc définie p.p. sur E_1 et on a $\varphi_f = \varphi_{f^+} - \varphi_{f^-}$ p.p. (on a $\varphi_f(x_1) = \varphi_{f^+}(x_1) - \varphi_{f^-}(x_1)$ en tout point x_1 t.q. $\varphi_{f^+}(x_1) < \infty$ et $\varphi_{f^-}(x_1) < \infty$). Comme $\varphi_{f^+} < \infty$ et $\varphi_{f^-} < \infty$ p.p, on peut trouver $A \in T_1$ t.q. $m_1(A) = 0$ et $\varphi_{f^+} < \infty$ et $\varphi_{f^-} < \infty$ sur $A^c = E_1 \setminus A$. En posant $g = \varphi_{f^+} - \varphi_{f^-}$ sur A^c et $g = 0$ sur A, on a donc $g \in \mathcal{M}(E_1, T_1)$, $g = \varphi_f$ p.p. et $g \in \mathcal{L}^1_{\mathbb{R}}(E_1, T_1, m_1)$ car $\int |g| dm_1 \leq \int \varphi_{f^+} dm_1 + \int \varphi_{f^-} dm_1 < \infty$. Ceci donne le deuxième item de la conclusion (le fait que φ_f appartienne à $L^1_{\mathbb{R}}(E_1, T_1, m_1)$) et donne aussi le troisième item car :

$$\begin{aligned}\int \varphi_f dm_1 = \int g dm_1 &= \int \varphi_{f^+} dm_1 - \int \varphi_{f^-} dm_1 \\ &= \int f^+ dm - \int f^- dm = \int f dm.\end{aligned}$$

Enfin, comme pour le théorème de Fubini-Tonelli, le quatrième item de la conclusion s'obtient en changeant les rôles de m_1 et m_2. ■

Le théorème de Fubini est souvent utilisé sous la forme du corollaire suivant :

Corollaire 7.13 *Soit* (E_1, T_1, m_1) *et* (E_2, T_2, m_2) *des espaces mesurés σ-finis,* (E, T, m) *l'espace produit et* $f : E \to \mathbb{R}$ *une fonction* T*-mesurable t.q. :*

$$\int \Big(\int |f(x_1, x_2)| dm_2(x_2) \Big) dm_1(x_1) < +\infty$$

ou

$$\int \Big(\int |f(x_1, x_2)| dm_1(x_1) \Big) dm_2(x_2) < +\infty.$$

Alors :

$$\int \Big(\int f(x_1, x_2) dm_2(x_2) \Big) dm_1(x_1) = \int \Big(\int f(x_1, x_2) dm_1(x_1) \Big) dm_2(x_2).$$

(Toutes les intégrales ayant bien un sens.)

DÉMONSTRATION – Le corollaire est une conséquence immédiate du théorème 7.12 et de l'équivalence (7.6). ■

Remarque 7.14 (Contre-exemple lié au théorème de Fubini) On cherche ici à construire une fonction pour laquelle la conclusion du théorème de Fubini n'est pas vérifiée. Soit a une fonction (continue) de $\mathbb{R}$ dans $\mathbb{R}$ et $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = a(x)$ si $x \geq 0$ et $x \leq y < 2x$, $f(x, y) = -a(x)$ si $x \geq 0$ et $2x \leq y < 3x$, $f(x, y) = 0$ si $x < 0$ ou $x \geq 0$ et $y \notin [x, 3x]$. On pose $b(x) = x\,a(x)$. On peut montrer que les hypothèses du théorème de Fubini ne sont vérifiées que si $b \in L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. En prenant par exemple $a(x) = 1/(1 + x)^2$, on montre que $\int (\int f(x, y) dy) dx \neq \int (\int f(x, y) dx) dy$ (voir l'exercice 7.5).

7.4 Mesure de Lebesgue sur la tribu des boréliens de $\mathbb{R}^N$

On a déjà vu que $\mathcal{B}(\mathbb{R}^N) = \mathcal{B}(\mathbb{R}^{N-1}) \otimes \mathcal{B}(\mathbb{R})$ pour tout $N \geq 1$ (exercice 2.6 pour $N = 2$ et exercice 7.1). Le paragraphe précédent permet alors de définir la mesure de Lebesgue sur les boréliens de $\mathbb{R}^N$ (c'est-à-dire sur la tribu $\mathcal{B}(\mathbb{R}^N)$, engendrée par les ouverts de $\mathbb{R}^N$) pour tout $N \geq 1$.

Définition 7.15 (Mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$)

1. *La mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^2)$ est la mesure $\lambda \otimes \lambda$, on la note λ_2.*
2. *Par récurrence sur* N, *la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$, $N \geq 3$, est la mesure $\lambda_{N-1} \otimes \lambda$, on la note λ_N.*

On note $L^1(\mathbb{R}^N)$ l'espace $L^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$, et pour $f \in L^1(\mathbb{R}^N)$, on note

$$\int f(x) d\lambda_N(x) = \int f(x) dx.$$

On donne maintenant quelques propriétés de la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$. Il s'agit de propriétés élémentaires ou de généralisations simples de propriétés vues pour la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$. Les démonstrations seront proposées en exercice.

Proposition 7.16 (Propriétés élémentaires de λ_N) *Soit* $N \geq 2$. *On rappelle que λ_N est la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$.*

1. *La mesure λ_N est σ-finie.*
2. *Soit $A_1, \ldots, A_N \in \mathcal{B}(\mathbb{R})$. Alors, $\prod_{i=1}^N A_i \in \mathcal{B}(\mathbb{R}^N)$ et*
$$\lambda_N(\prod_{i=1}^N A_i) = \prod_{i=1}^N \lambda(A_i).$$
3. *Soit $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$ et $\beta_1, \ldots, \beta_N \in \mathbb{R}$ t.q. $\alpha_i < \beta_i$ pour tout $i \in \{1, \ldots, N\}$. Alors :*
$$\lambda_N(\prod_{i=1}^N]\alpha_i, \beta_i[) = \prod_{i=1}^N \lambda(]\alpha_i, \beta_i[) = \prod_{i=1}^N (\beta_i - \alpha_i).$$
4. *Soit* K *un compact de $\mathbb{R}^N$ (noter que $K \in \mathcal{B}(\mathbb{R}^N)$). Alors, $\lambda_N(K) < +\infty$.*
5. *Soit* O *un ouvert non vide de $\mathbb{R}^N$. Alors, $\lambda_N(O) > 0$.*
6. *Soit $f, g \in C(\mathbb{R}^N, \mathbb{R})$. Alors $f = g$ p.p. (c'est-à-dire λ_N-p.p.) implique $f(x) = g(x)$ pour tout $x \in \mathbb{R}^N$.*
7. $C_c(\mathbb{R}^N, \mathbb{R}) \subset \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$. *(En confondant f avec sa classe, on écrira donc souvent $C_c(\mathbb{R}^N, \mathbb{R}) \subset L^1_{\mathbb{R}}(\mathbb{R}^N)$.)*

DÉMONSTRATION – Comme λ_N est une mesure produit, le fait que λ_N est σ-finie est (par récurrence sur N) une conséquence du théorème donnant l'existence (et l'unicité) de la mesure produit (théorème 7.3) car ce théorème donne que le produit de mesures σ-finies est σ-finie.

La démonstration des autres propriétés fait l'objet de l'exercice 7.11. ■

Une propriété très importante de λ_N est sa régularité, c'est-à-dire que pour tout élément A de $\mathcal{B}(\mathbb{R}^N)$ et pour tout $\varepsilon > 0$, il existe O ouvert de $\mathbb{R}^N$ et F fermé de $\mathbb{R}^N$ tels que

$$F \subset A \subset O \text{ et } \lambda_N(O \setminus F) \leq \varepsilon.$$

Cette propriété est une conséquence du fait que toute mesure sur $\mathcal{B}(\mathbb{R}^N)$, finie sur les compacts, est régulière (proposition 7.17).

Proposition 7.17 (Régularité d'une mesure sur $\mathcal{B}(\mathbb{R}^N)$, finie sur les compacts) *Soit m une mesure sur $\mathcal{B}(\mathbb{R}^N)$ t.q. $m(K) < \infty$ pour tout compact K de $\mathbb{R}^N$. (Noter que ceci est vrai pour $m = \lambda_N$.) Alors :*

1. *Pour tout $A \in \mathcal{B}(\mathbb{R}^N)$ et pour tout $\varepsilon > 0$, il existe O ouvert de $\mathbb{R}^N$ et F fermé de $\mathbb{R}^N$ tels que :*

$$F \subset A \subset O \text{ et } m(O \setminus F) \leq \varepsilon.$$

2. *Pour tout $A \in \mathcal{B}(\mathbb{R}^N)$, on a $m(A) = \inf\{m(O), O \text{ ouvert t.q. } A \subset O\}$.*

DÉMONSTRATION – Cette proposition fait l'objet de l'exercice 7.12. ■

On donne maintenant des généralisations au cas de λ_N de propriétés déjà vues pour λ.

Proposition 7.18 (Densité de C_c dans $L^1(\mathbb{R}^N)$) *Pour tout $N \geq 1$, l'espace $C_c(\mathbb{R}^N, \mathbb{R})$ est dense dans $L^1(\mathbb{R}^N)$ (c'est-à-dire que, pour tout $f \in L^1(\mathbb{R}^N)$ et tout $\varepsilon > 0$, il existe $\varphi \in C_c(\mathbb{R}^N, \mathbb{R})$ t.q. $\|f - \varphi\|_1 \leq \varepsilon$).*

DÉMONSTRATION – La démonstration de cette proposition fait l'objet de l'exercice 7.15, elle découle essentiellement de la régularité de λ_N. Cette démonstration est très voisine de celle faite pour le cas $N = 1$, théorème 5.20. ■

Comme cela a déjà été dit après le théorème 5.20, le résultat de densité que nous venons d'énoncer n'est pas limité à la mesure de Lebesgue. Il est vrai pour toute mesure sur $\mathcal{B}(\mathbb{R}^N)$, finie sur les compacts. Il est aussi vrai en remplaçant C_c par C_c^∞. On obtient donc le théorème suivant :

Théorème 7.19 (Densité de C_c^∞ dans $L^1(\mathbb{R}^N)$) *Soit $N \geq 1$ et μ sur une mesure sur $\mathcal{B}(\mathbb{R}^N)$, finie sur les compacts. Alors, l'espace $C_c^\infty(\mathbb{R}^N, \mathbb{R})$ est dense dans $L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mu)$ (c'est-à-dire que, pour tout $f \in L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mu)$ et tout $\varepsilon > 0$, il existe $\varphi \in C_c(\mathbb{R}^N, \mathbb{R})$ t.q. $\|f - \varphi\|_1 \leq \varepsilon$).*

La démonstration de ce théorème fait l'objet de l'exercice 7.16.

Proposition 7.20 (Invariance par translation) *Soient* $N \geq 1$, $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^*$ *et* $\beta_1, \ldots, \beta_N \in \mathbb{R}$.

Pour $x = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$, *on pose* $\varphi(x) = (\alpha_1 x_1 + \beta_1, \ldots, \alpha_N x_N + \beta_N)^t$ *(noter que* φ *est une bijection de* $\mathbb{R}^N$ *dans* $\mathbb{R}^N$*). Pour tout* $A \in \mathcal{B}(\mathbb{R}^N)$, *on a alors* $\varphi(A) = \{\varphi(x), x \in A\} \in \mathcal{B}(\mathbb{R}^N)$ *et* $\lambda_N(\varphi(A)) = \prod_{i=1}^N |\alpha_i| \lambda_N(A)$.

Pour $\alpha_i = 1$ *pour tout i, cette propriété s'appelle invariance par translation de* λ_N.

DÉMONSTRATION – Cette proposition a déjà été vue pour $N = 1$, proposition 2.48. La démonstration de la proposition 2.48 utilisait, par exemple, le fait que tout ouvert est réunion dénombrable d'intervalles ouverts disjoints deux à deux (et la régularité de λ). La démonstration proposée ici pour $N \geq 1$ utilise une récurrence sur N et la partie unicité du théorème 7.3 sur la mesure produit. Elle fait l'objet de l'exercice 7.17.

On peut aussi noter que la partie unicité du théorème 7.3 peut être faite (voir la remarque 7.4) avec le lemme des classes monotones (exercice 2.13). Ce lemme pourrait aussi être utilisé pour démontrer la proposition 2.48 (au lieu du théorème de régularité et du fait que tout ouvert est réunion dénombrable d'intervalles ouverts disjoints deux à deux). ■

Proposition 7.21 (Changement de variables simple)

Soient $N \geq 1$, $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^*$ *et* $\beta_1, \ldots, \beta_N \in \mathbb{R}$.

Pour $x = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$, *on pose* $\varphi(x) = (\alpha_1 x_1 + \beta_1, \ldots, \alpha_N x_N + \beta_N)^t$ *(de sorte que* φ *est une bijection de* $\mathbb{R}^N$ *dans* $\mathbb{R}^N$*). Alors :*

1. *Pour tout* $f \in \mathcal{M}_+ = \mathcal{M}_+(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, *on a* $f \circ \varphi \in \mathcal{M}_+$ *et :*

$$\int f d\lambda_N = \prod_{i=1}^N |\alpha_i| \int (f \circ \varphi) d\lambda_N.$$

2. *Pour tout* $f \in \mathcal{L}^1 = \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$, *on a* $f \circ \varphi \in \mathcal{L}^1$ *et :*

$$\int f d\lambda_N = \prod_{i=1}^N |\alpha_i| \int (f \circ \varphi) d\lambda_N.$$

DÉMONSTRATION – La démonstration est une conséquence simple de la proposition 7.20. Elle fait l'objet de l'exercice 7.18.

Noter aussi que $\prod_{i=1}^N |\alpha_i|$ est la valeur absolue du déterminant de la matrice jacobienne de φ au point x. Cette matrice est notée $D\varphi(x)$, elle ne dépend pas de x pour les applications considérées dans cette proposition. Cette proposition sera généralisée au théorème 7.29. ■

7.5 Convolution

On rappelle que $\mathrm{L}^1(\mathbb{R}^\mathrm{N}) = \mathrm{L}^1_\mathbb{R}(\mathbb{R}^\mathrm{N}, \mathcal{B}(\mathbb{R}^\mathrm{N}), \lambda_\mathrm{N})$ et que, pour $f \in \mathrm{L}^1(\mathbb{R}^\mathrm{N})$, $\int f d\lambda_\mathrm{N} = \int f(x) d\lambda_\mathrm{N}(x) = \int f(x)dx$ (c'est-à-dire que dx signifie toujours $d\lambda_\mathrm{N}(x)$).

On note aussi $\mathcal{L}^1(\mathbb{R}^\mathrm{N}) = \mathcal{L}^1_\mathbb{R}(\mathbb{R}^\mathrm{N}, \mathcal{B}(\mathbb{R}^\mathrm{N}), \lambda_\mathrm{N})$.

Soient $f, g \in \mathrm{L}^1(\mathbb{R}^\mathrm{N})$. On souhaite définir la fonction convoluée de f et g, c'est-à-dire définir $f * g$ par :

$$f * g(x) = \int f(t) g(x-t) dt.$$

La définition de cette fonction nécessite les deux conditions suivantes :

1. Il faut que la définition ne dépende pas des représentants choisis pour f et g.
2. Il faut que, ayant choisi des représentants pour f et g, encore notés f et g, la fonction $g(x-\cdot)f(\cdot)$ appartienne à $\mathrm{L}^1(\mathbb{R}^\mathrm{N})$ (au sens "il existe $h \in \mathcal{L}^1(\mathbb{R}^\mathrm{N})$ t.q. $g(x-\cdot)f(\cdot) = h$ p.p."). Ceci n'est pas immédiat car, en général, le produit deux fonctions intégrables n'est pas une fonction intégrable.

La condition 1 est satisfaite, car, pour $x \in \mathbb{R}^\mathrm{N}$, si f, f_1, g et g_1 sont des fonctions de $\mathbb{R}^\mathrm{N}$ dans $\mathbb{R}$, on a :

$$f = f_1 \text{ p.p.}, \ g = g_1 \text{ p.p. } \Rightarrow f(\cdot)g(x-\cdot) = f_1(\cdot)g_1(x-\cdot) \text{ p.p..} \tag{7.7}$$

(p.p. signifiant ici λ_N-p.p..) En effet, il suffit de remarquer que si $f = f_1$ p.p. et $g = g_1$ p.p., il existe $\mathrm{A}, \mathrm{B} \in \mathcal{B}(\mathbb{R}^\mathrm{N})$ t.q. $\lambda_\mathrm{N}(\mathrm{A}) = \lambda_\mathrm{N}(\mathrm{B}) = 0$, $f = f_1$ sur A^c et $g = g_1$ sur B^c. Pour $x \in \mathbb{R}^\mathrm{N}$, on a alors $f(\cdot)g(x-\cdot) = f_1(\cdot)g_1(x-\cdot)$ sur $\mathrm{A}^c \cap \mathrm{B}^c_x = (\mathrm{A} \cup \mathrm{B}_x)^c$ avec $\mathrm{B}_x = \{x - z,\ z \in \mathrm{B}\}$. On en déduit bien $f(\cdot)g(x-\cdot) = f_1(\cdot)g_1(x-\cdot)$ p.p. car $\lambda_\mathrm{N}(\mathrm{A} \cup \mathrm{B}_x) \le \lambda_\mathrm{N}(\mathrm{A}) + \lambda_\mathrm{N}(\mathrm{B}_x) = \lambda_\mathrm{N}(\mathrm{A}) + \lambda_\mathrm{N}(\mathrm{B}) = 0$ (on utilise ici la propriété d'invariance par translation donnée dans la proposition 7.20).

On en déduit que, si f et f_1 [resp. g et g_1] sont des représentants d'un même élément de $\mathrm{L}^1(\mathbb{R}^\mathrm{N})$, on a $f(.)g(x-.) \in \mathrm{L}^1(\mathbb{R}^\mathrm{N})$ si et seulement si $f_1(.)g_1(x-.) \in \mathrm{L}^1(\mathbb{R}^\mathrm{N})$ et, si $f(.)g(x-.) \in \mathrm{L}^1(\mathbb{R}^\mathrm{N})$, on a $\int f(t)g(x-t)dt = \int f_1(t)g_1(x-t)dt$.

On montre dans la proposition suivante que la condition 2 est satisfaite pour presque tout $x \in \mathbb{R}^\mathrm{N}$.

Proposition 7.22 (Convolution) *Soient $f, g \in \mathrm{L}^1(\mathbb{R}^\mathrm{N})$ (que l'on confond avec l'un de leurs représentants). Alors :*

- *Pour presque tout $x \in \mathbb{R}^N$, la fonction $g(x-\cdot)f(\cdot)$ appartient à $L^1(\mathbb{R}^N)$ (en la confondant avec sa classe). On pose donc : $f * g(x) = \int f(t)g(x-t)dt$. La fonction $f * g$ est donc définie p.p. sur $\mathbb{R}^N$.*
- *$f * g \in L^1(\mathbb{R}^N)$ (au sens "il existe $h \in \mathcal{L}^1(\mathbb{R}^N)$ t.q. $f * g = h$ p.p.").*
- $\|f * g\|_1 \le \|f\|_1 \|g\|_1$.

DÉMONSTRATION – On donne la démonstration pour $N = 1$ (le cas $N > 1$ est similaire, en ayant d'abord montré que $\lambda_{2N} = \lambda_N \otimes \lambda_N$).

On choisit des représentants de f et g, de sorte que $f, g \in \mathcal{L}^1(\mathbb{R}) = \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. On souhaite appliquer le théorème de Fubini (théorème 7.12) à $H : \mathbb{R}^2 \to \mathbb{R}$, définie par $H(x, y) = f(y)g(x-y)$, avec les espaces mesurés $(E_i, T_i, m_i) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ pour $i = 1, 2$.

Comme λ est σ-finie, pour appliquer le théorème de Fubini, il suffit de vérifier que H est $\mathcal{B}(\mathbb{R}^2)$-mesurable et que $\int(\int |H(x, y)|dx)dy < \infty$.

On montre d'abord que H est $\mathcal{B}(\mathbb{R}^2)$-mesurable. On a $H = H_1 \circ \psi$ avec :

$$H_1 \;:\; (x, y) \mapsto f(x)g(y), \;\; \psi \;:\; (x, y) \mapsto (y, x-y).$$

La fonction H_1 est mesurable de $\mathbb{R}^2$ dans $\mathbb{R}$ (car f et g sont mesurables de $\mathbb{R}$ dans $\mathbb{R}$, on applique ici la proposition 7.11) et ψ est mesurable de $\mathbb{R}^2$ dans $\mathbb{R}^2$ car continue ($\mathbb{R}$ et $\mathbb{R}^2$ sont toujours munis de leur tribu borélienne). La fonction H est donc mesurable de $\mathbb{R}^2$ dans $\mathbb{R}$ comme composée de fonctions mesurables. On peut maintenant calculer l'intégrale de $|H|$:

$$\int(\int |H(x, y)|dx)dy = \int(\int |f(y)g(x-y)|dx)dy = \int |f(y)|(\int |g(x-y)|dx)dy.$$

La proposition 7.21 donne $\int |g(x-y)|dx = \int |g(x)|dx = \|g\|_1$, Donc :

$$\int(\int |H(x, y)|dx)dy = \|g\|_1 \int |f(y)|dy = \|g\|_1 \|f\|_1 < \infty.$$

Le théorème de Fubini peut donc s'appliquer. Il donne que $H(x, \cdot) \in \mathcal{L}^1(\mathbb{R})$ pour presque tout $x \in \mathbb{R}$. Donc, $g(x-\cdot)f(\cdot) \in \mathcal{L}^1(\mathbb{R})$ pour presque tout $x \in \mathbb{R}$. Ceci montre bien que $f * g$ est définie p.p.. Le théorème de Fubini donne alors aussi que $f * g \in L^1(\mathbb{R})$ (au sens "il existe $h \in \mathcal{L}^1(\mathbb{R})$ t.q. $f * g = h$ p.p.").

Enfin pour montrer que $\|f * g\|_1 \le \|f\|_1 \|g\|_1$, il suffit de remarquer que :

$$\|f * g\|_1 = \int |\int f(y)g(x-y)dy|dx \le \int(\int |H(x, y)|dx)dy = \|g\|_1 \|f\|_1.$$

■

Remarque 7.23 On a vu précédemment que $L^1(\mathbb{R}^N)$ muni de l'addition (loi de composition interne), de la multiplication par un scalaire (loi de composition externe) et de la norme $\|\cdot\|_1$ est un espace de Banach. L'ajout de la convolution (loi de composition interne) confère à $L^1(\mathbb{R}^N)$ la structure d'algèbre de Banach.

On sait aussi que $C_b(\mathbb{R},\mathbb{R})$ muni de l'addition, de la multiplication interne, de la multiplication par un scalaire et de la norme de la convergence uniforme $\|\cdot\|_u$ est aussi une algèbre de Banach. En fait, nous montrerons par la suite qu'il existe un isomorphisme d'algèbre, appelé transformation de Fourier, entre $L^1(\mathbb{R}^N)$ et son image (par cette transformation) dans $C_b(\mathbb{R}^N,\mathbb{R})$.

Remarque 7.24 On donne ici quelques propriétés supplémentaires de la convolution. Soit $N \geq 1$. Pour $p \in [1,\infty]$, on pose $L^p(\mathbb{R}^N) = L^p_{\mathbb{R}}(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N),\lambda_N)$.

1. Soit $f,g \in L^1(\mathbb{R}^N)$. On a alors $f*g = g*f$ p.p.. Ceci découle de l'invariance par translation de la mesure de Lebesgue (propositions 7.20 et 7.21) et est démontré dans l'exercice 7.21.
2. Soit$1 < p < \infty$. Soient $f \in L^p(\mathbb{R}^N)$ et $g \in L^1(\mathbb{R}^N)$. Alors, $f*g$ est définie p.p. sur $\mathbb{R}^N$, $f*g \in L^p(\mathbb{R}^N)$ et $\|f*g\|_p \leq \|f\|_p\|g\|_1$. Cette propriété fait l'objet de l'exercice 7.23.
3. Soit $p,q \in [1,\infty]$ t.q. $(1/p)+(1/q) = 1$. Soient $f \in L^p(\mathbb{R}^N)$ et $g \in L^q(\mathbb{R}^N)$. Alors, $f*g$ est définie partout sur $\mathbb{R}^N$ et $f*g \in C_b(\mathbb{R}^N,\mathbb{R})$, voir l'exercice 8.7.
4. Soit $p \in [1,\infty]$. Soient $f \in L^p(\mathbb{R}^N)$ et $g \in C^\infty_c(\mathbb{R}^N,\mathbb{R})$. Alors, $f*g$ est définie partout sur $\mathbb{R}^N$ et $f*g \in C^\infty(\mathbb{R}^N,\mathbb{R})$, voir l'exercice 7.22.
5. (Régularisation par convolution) Soit $f : \mathbb{R}^N \to \mathbb{R}$. On dit que $f \in L^1_{\text{loc}}(\mathbb{R}^N)$ si $f1_K \in L^1(\mathbb{R}^N)$ pour tout compact K de R^N. On suppose que $f \in L^1_{\text{loc}}(\mathbb{R}^N)$. Soit $k \in \mathbb{N}$ et $g \in C^k_c(\mathbb{R}^N,\mathbb{R})$. Alors, $f*g$ est définie partout sur $\mathbb{R}^N$ et $f*g \in C^k(\mathbb{R}^N,\mathbb{R})$ (voir l'exercice 7.22, noter que $L^p(\mathbb{R}^N) \subset L^1_{\text{loc}}(\mathbb{R}^N)$).
6. Soit $f,g \in L^1(\mathbb{R}^N)$. On suppose que f et g sont à support compact (f à support compact signifie qu'il existe K, compact de $\mathbb{R}^N$ t.q. $f = 0$ p.p. sur K^c). Alors, la fonction $f*g$ est aussi à support compact. Ceci fait partie de l'exercice 7.21.

La convolution est un outil très utile pour "régulariser" des fonctions. Elle est à la base de résultats de densité fondamentaux que nous verrons dans le chapitre suivant (densité de $C^\infty_c(\mathbb{R}^N,\mathbb{R})$ dans $L^p(\mathbb{R}^N)$ pour $p < \infty$, par exemple).

Il est aussi intéressant de généraliser la convolution de fonctions en convolution de mesures. On commence par remarquer qu'une fonction f dans $L^1_{\text{loc}}(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N),\lambda_N)$ (voir la remarque 7.24) est entièrement déterminée par la mesure qu'elle induit sur les parties boréliennes bornées de $\mathbb{R}^N$, c'est-à-dire par la mesure m définie par $m = f\lambda_N$ (qui est une mesure signée sur Ω si Ω est une partie borélienne bornée de $\mathbb{R}^N$). Ceci est précisé dans le lemme suivant (en remarquant que $\int \varphi dm = \int \varphi f d\lambda_N$ pour tout $\varphi \in C_c(\mathbb{R}^N,\mathbb{R})$).

Lemme 7.25 *Soit* $N \geq 1$ *et* $f,g \in L^1_{loc}(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N),\lambda_N)$ *t.q.* $\int f\varphi d\lambda_N = \int g\varphi d\lambda_N$, *pour tout* $\varphi \in C_c(\mathbb{R}^N,\mathbb{R})$. *Alors,* $f = g$ *p.p..*

DÉMONSTRATION – Soit $M \in \mathbb{N}^*$. On note B_M la boule (fermée) de centre 0 et rayon M dans $\mathbb{R}^N$ et h_M la fonction définie par :

$$h_M(x) = \begin{cases} f(x) - g(x) & \text{si } x \in B_M \text{ et } |f(x) - g(x)| \leq M, \\ \\ 0 & \text{si } x \notin B_M \text{ ou } |f(x) - g(x)| > M, \end{cases}$$

On a $h_M \in L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (car B_M est un compact de $\mathbb{R}^N$). Comme $C_c(\mathbb{R}^N, \mathbb{R})$ est dense dans $L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (théorème 5.20 pour $d = 1$ et théorème 7.18 pour $N \geq 1$), il existe une suite $(\varphi_n)_{n\in\mathbb{N}}$ t.q. $\varphi_n \to h_M$ dans $L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ quand $n \to +\infty$. On peut aussi supposer (quitte à extraire une sous-suite) que $\varphi_n \to h_M$ p.p. (théorème 6.11). Enfin, en remplaçant φ_n par $\max(\min(\varphi_n, M), -M)$ on obtient :

$$\begin{aligned} &\varphi_n \in C_c(\mathbb{R}^N, \mathbb{R}) \text{ pour tout } n \in \mathbb{N}, \\ &\varphi_n \to h_M \text{ p.p.}, \\ &|\varphi_n| \leq M \text{ pour tout } n \in \mathbb{N}. \end{aligned}$$

Comme $\varphi_n \in C_c(\mathbb{R}^N, \mathbb{R})$, on a $\int (f - g)\varphi_n d\lambda_N = 0$. Le théorème de convergence dominée (la domination est par $M1_{B_M}|f - g|$) donne alors $\int h_M(f - g)d\lambda_N = 0$. En faisant maintenant tendre M vers l'infini, le théorème de convergence monotone donne $\int |f - g| d\lambda_N = 0$, et donc $f = g$ p.p. ∎

Pour que la convolution de mesures soit une généralisation de la convolution de fonctions, on souhaite que $m * \mu = (f * g)\lambda_N$, lorsque $m = f\lambda_N$ et $\mu = g\lambda_N$ avec $f, g \in L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (et donc $f * g \in L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$). Noter que m, μ et $m * \mu$ sont des mesures signées. Soient $f, g \in L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$. On pose $m = f\lambda_N$ et $g = g\lambda_N$. Pour toute fonction φ borélienne bornée de $\mathbb{R}^N$ dans $\mathbb{R}$ (par exemple, $\varphi \in C_b(\mathbb{R}^N, \mathbb{R})$),

$$\int (f * g)\varphi d\lambda_N = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x - y)g(y)dy \right) \varphi(x)dx.$$

(On rappelle que dx désigne $d\lambda_N(x)$). En utilisant le théorème de Fubini, qui s'applique bien ici car

$$\int \int |f(x - y)g(y)\varphi(x)|dxdy \leq \|\varphi\|_\infty \|f\|_1 \|g\|_1,$$

et avec le changement de variable $z = x - y$ (pour y fixé), on obtient :

$$\begin{aligned} \int (f * g)\varphi d\lambda_N &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x - y)\varphi(x)dx \right) g(y)dy \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(z)\varphi(z + y)dz \right) g(y)dy. \end{aligned}$$

On a donc :

$$\int (f*g)\varphi d\lambda_{\mathrm{N}} = \int_{\mathbb{R}^{\mathrm{N}}} \left(\int_{\mathbb{R}^{\mathrm{N}}} \varphi(z+y) dm(z) \right) d\mu(y) = \int_{\mathbb{R}^{2\mathrm{N}}} \varphi(y+z) d(m\otimes\mu)(z,y), \tag{7.8}$$

où la dernière égalité découle de la définition de $m\otimes\mu$. Plus précisément, si m et μ sont des mesures finies (c'est-à-dire des applications σ-additives de $\mathcal{B}(\mathbb{R}^{\mathrm{N}})$ dans $\mathbb{R}^+$), la dernière égalité de 7.8 est donnée par le troisième item du théorème de Fubini (théorème 7.12). Si m et μ sont des mesures signées, on se ramène au cas précédent avec la décomposition de Hahn (proposition 2.33) qui donne $m = m^+ - m^-$ et $\mu = \mu^+ - \mu^-$. La mesure $m\otimes\mu$ est alors définie à partir de $m^{\pm}\otimes\mu^{\pm}$.

On est ainsi amené naturellement à définir $m*\mu$ en utilisant le deuxième membre de (7.8) pour définir $\int \varphi d(m*\mu)$.

Définition 7.26 (Convolution de mesures) *Soit* $\mathrm{N}\geq 1$ *et* m, μ *des mesures signées sur* $\mathcal{B}(\mathbb{R}^{\mathrm{N}})$. *On définit la mesure signée* $m*\mu$ *sur* $\mathcal{B}(\mathbb{R}^{\mathrm{N}})$ *par :*

$$m*\mu(\mathrm{A}) = \int_{\mathbb{R}^{2\mathrm{N}}} 1_{\mathrm{A}}(x+y) d(m\otimes\mu)(x,y) \text{ pour tout } \mathrm{A}\in\mathcal{B}(\mathbb{R}^{\mathrm{N}}).$$

où $m\otimes\nu = m^+\otimes\mu^+ + m^-\otimes\mu^- - m^-\otimes\mu^+ - m^+\otimes\mu^-$ *et* $m^{\pm}$ $\mu^{\pm}$ *sont données par la décomposition de Hahn de* m *et* μ *(proposition 2.33).*

Le fait que $m*\mu$ est une mesure signée sur $\mathcal{B}(\mathbb{R}^{\mathrm{N}})$ est facile (la σ-additivité de $m*\mu$ découle, par exemple, du théorème de convergence dominée). On déduit de cette définition la proposition suivante.

Proposition 7.27 *Soit* $\mathrm{N}\geq 1$ *et* m, μ *des mesures signées sur* $\mathcal{B}(\mathbb{R}^{\mathrm{N}})$.

1. *On a alors, pour tout* φ *borélienne bornée de* $\mathbb{R}^{\mathrm{N}}$ *dans* $\mathbb{R}$ *(par exemple,* $\varphi\in C_b(\mathbb{R}^{\mathrm{N}},\mathbb{R})$*) :*

$$\int_{\mathbb{R}^{\mathrm{N}}} \varphi d(m*\mu) = \int_{\mathbb{R}^{2\mathrm{N}}} \varphi(x+y) d(m\otimes\mu)(x,y).$$

2. *Si* m *et* μ *sont des probabilités, la mesure* $m*\mu$ *est aussi une probabilité.*
3. *Si* $f,g\in \mathrm{L}^1(\mathbb{R}^{\mathrm{N}},\mathcal{B}(\mathbb{R}^{\mathrm{N}}),\lambda_{\mathrm{N}})$, $m = f\lambda_{\mathrm{N}}$ *et* $\mu = g\lambda_{\mathrm{N}}$, *on a* $m*\mu = (f*g)\lambda_{\mathrm{N}}$.

DÉMONSTRATION – Le premier item se démontre, comme souvent, en considérant des fonctions étagées, puis en écrivant φ comme limite de fonctions étagées (bornées par la borne supérieure de $|\varphi|$, exercice **??**). Le deuxième item est immédiat en remarquant que $m*\mu(\mathrm{A})\geq 0$ si m et μ sont des mesures (positives) et $m*\mu(\mathbb{R}^{\mathrm{N}}) = m(\mathbb{R}^{\mathrm{N}})\mu(\mathbb{R}^{\mathrm{N}})$. Enfin, le troisième item a été vu avant la proposition 7.27. ■

Remarque 7.28 Il aurait aussi été possible de définir $m*\mu$ grâce au théorème de Riesz dans $C_0(\mathbb{R}^N,\mathbb{R})$ (théorème 5.16 pour $N \geq 1$). Si m, μ sont des mesures signées sur $\mathcal{B}(\mathbb{R}^N)$ (ou, directement, des formes linéaires continues sur $C_0(\mathbb{R}^N,\mathbb{R})$). On définit, l'application L de $C_0(\mathbb{R}^N,\mathbb{R})$ dans $\mathbb{R}$ par :

$$L(\varphi) = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \varphi(x+y)dm(x)d\mu(y), \forall \varphi \in C_0(\mathbb{R}^N,\mathbb{R}).$$

L'application L est une forme linéaire continue sur $C_0(\mathbb{R}^N,\mathbb{R})$. Par le théorème de Riesz, il existe donc une unique mesure de Radon, notée τ (c'est la mesure convoluée de m et μ) t.q. :

$$L(\varphi) = \int \varphi(s)d\tau(s).$$

7.6 Formules de changement de variable

La proposition 7.21 donne un résultat sur les changements de variables "simples". On donne maintenant une généralisation dans le cas où les intégrales portent sur des ouverts bornés de $\mathbb{R}^N$.

Théorème 7.29 (Formules de changement de variable) *Soit* $N \geq 1$, U *et* V *des ouverts bornés de* $\mathbb{R}^N$ *et* φ *un* C^1*-difféomorphisme de* U *dans* V *(i.e.* φ *est une bijection de* U *dans* V, $\varphi \in C^1(U,V)$ *et* $\varphi^{-1} \in C^1(V,U)$*). On note* $D\varphi(y)$ *la matrice jacobienne de* φ *en* y *et* $Det(D\varphi)$ *la fonction* $y \mapsto Det(D\varphi(y))$.

1. Soit $f \in \mathcal{M}_+ = \mathcal{M}_+(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N))$*. Alors,*

$$(f\circ\varphi)|Det(D\varphi)|1_U \in \mathcal{M}_+ \text{ et } \int_V f(x)dx = \int_U f(\varphi(y))|Det(D\varphi(y))|dy.$$

2. Soit $f \in \mathcal{M}(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N))$ *t.q.* $f1_V \in \mathcal{L}^1 = \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N),\lambda_N)$*. Alors,*

$$(f\circ\varphi)|Det(D\varphi)|1_U \in \mathcal{L}^1 \text{ et } \int_V f(x)dx = \int_U f(\varphi(y))|Det(D\varphi(y))|dy.$$

DÉMONSTRATION – Comme φ est de classe C^1, la fonction $(f\circ\varphi)|Det(D\varphi)|1_U$ est mesurable si f est mesurable. Il est plus difficile de montrer l'égalité donnée dans l'item 1 du théorème. Cette démonstration n'est pas faite ici. Elle consiste à se ramener par un procédé de localisation au cas de changements de variable affines.

Le deuxième item du théorème est une conséquence facile du premier. En effet, soit $f \in \mathcal{M}(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N))$ t.q. $f1_V \in \mathcal{L}^1$. En appliquant le premier item à la fonction $|f| \in \mathcal{M}_+$, on obtient que $(f\circ\varphi)|Det(D\varphi)|1_U \in \mathcal{L}^1$. Puis en appliquant le premier item à f^+ et f^- et en faisant la différence, on obtient bien que $\int_V f(x)dx = \int_U f(\varphi(y))|Det(D\varphi(y))|dy$.

■

Un exemple de changement de variable On conclut cette section en donnant l'exemple des coordonnées polaires pour N = 2. Soit $f \in \mathcal{M}_+(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ (ou $f \in \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda_2)$). On veut calculer (par exemple) $\int_{B_1} f(x)dx$, où B_1 est la boule unité (ouverte) de $\mathbb{R}^2$, en passant en coordonnée polaires.

On a donc $B_1 = \{x \in \mathbb{R}^2 ; |x| < 1\}$, où $|\cdot|$ est la norme euclidienne dans $\mathbb{R}^2$, c'est-à-dire $|x|^2 = x_1^2 + x_2^2$ si $x = (x_1, x_2)^t \in \mathbb{R}^2$.

On pose $L = \{(x_1, 0)^t, x_1 \in [0,1[\}$ et on remarque que $\lambda_2(L) = \lambda([0,1[) \times \lambda(\{0\}) = 0$. Donc, en posant $V = B_1 \setminus L$, on a :

$$\int_{B_1} f(x)dx = \int_{B_1\setminus L} f(x)dx = \int_V f(x)dx (= \int_V f d\lambda_2).$$

On pose aussi $U =]0,1[\times]0,2\pi[$, de sorte que U et V sont de ouverts bornés de $\mathbb{R}^2$. L'application $\varphi : (r,\theta)^t \mapsto (r\cos\theta, r\sin\theta)^t$ est alors une bijection de U dans V. Elle est de classe C^1 et son inverse est de classe C^1 (φ et φ^{-1} sont même de classe C^∞). On peut calculer la matrice jacobienne de φ et son déterminant :

$$D\varphi(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}, \ |Det(D\varphi(r,\theta))| = r.$$

On peut donc appliquer le théorème 7.29, il donne :

$$\int_{B_1} f(x)dx = \int_V f(x)dx = \int_U f(r\cos\theta, r\sin\theta) r d\lambda_2(r,\theta)$$
$$= \int_{]0,1[\times]0,2\pi[} f(r\cos\theta, r\sin\theta) r d\lambda_2(r,\theta)$$

En appliquant maintenant le théorème de Fubini-Tonelli pour évaluer la dernière intégrale (si $f \in \mathcal{L}^1$ au lieu de $f \in \mathcal{M}_+$, on raisonne d'abord sur $|f|$ puis on utilise le théorème de Fubini), on obtient :

$$\int_{B_1} f(x)dx = \int_0^{2\pi} (\int_0^1 f(r\cos\theta, r\sin\theta) r dr) d\theta.$$

Si $f(x)$ ne dépend que de $|x|$, c'est-à-dire s'il existe ψ t.q. $f(x) = \psi(|x|)$, on obtient alors :

$$\int_{B_1} f(x)dx = 2\pi \int_0^1 \psi(r) r dr.$$

En particulier, on voit que $f1_{B_1} \in \mathcal{L}^1(\mathbb{R}^2)$ si et seulement si $g \in \mathcal{L}^1_{\mathbb{R}}(]0,1[,\mathcal{B}(\mathbb{R}),\lambda)$, avec g définie par $g(r) = r\psi(r)$ pour $r \in]0,1[$.

Prenons toujours $f \in \mathcal{M}_+(\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2))$ (ou bien $f \in \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2),\lambda_2)$). Le raisonnement que nous venons de faire pour B_1 peut être fait pour $B_a = \{x \in \mathbb{R}^2 ; |x| < a\}$ avec $a > 0$. On obtient alors, pour tout $a > 0$:

$$\int_{B_a} f(x)dx = \int_0^{2\pi} (\int_0^a f(r\cos\theta, r\sin\theta)rdr)d\theta. \tag{7.9}$$

En prenant $a = n$, $n \in \mathbb{N}^*$, dans (7.9), on obtient aussi, quand $n \to +\infty$ (avec le théorème de convergence monotone si $f \in \mathcal{M}_+$ et en raisonnement avec $f^{\pm}$ si $f \in \mathcal{L}^1$) :

$$\int f(x)dx = \int_0^{2\pi} (\int_0^{\infty} f(r\cos\theta, r\sin\theta)rdr)d\theta. \tag{7.10}$$

7.7 Exercices

7.7.1 Mesure produit

Exercice 7.1 (Tribu borélienne sur $\mathbb{R}^n$) Montrer que, pour tout $n \geq 2$, on a $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}^{n-1}) \otimes \mathcal{B}(\mathbb{R})$. [S'inspirer de la démonstration faite pour $n = 2$ dans l'exercice 2.6.]

Corrigé – *On note* $T = \mathcal{B}(\mathbb{R}^{n-1}) \otimes \mathcal{B}(\mathbb{R})$*, c'est-à-dire la tribu (sur* $\mathbb{R}^n$*) engendrée par* $\{A \times B ; A \in \mathcal{B}(\mathbb{R}^{n-1}), B \in \mathcal{B}(\mathbb{R})\}$. *On veut montrer que* $T = \mathcal{B}(\mathbb{R}^n)$.

Etape 1, $\mathcal{B}(\mathbb{R}^n) \subset T$**.**

Soit O *un ouvert de* $\mathbb{R}^n$*. On va montrer que* $O \in T$*. On suppose* $O \neq \emptyset$ *(on sait déjà que* $\emptyset \in T$*). Pour tout* $x = (x_1,\ldots,x_n)^t \in O$*, il existe* $r > 0$ *t.q.* $\prod_{i=1}^n]x_i - r, x_i + r[\subset O$*. Comme les rationnels sont denses dans* $\mathbb{R}$*, on peut trouver, pour tout* $i \in \{1,\ldots,n\}$*,* $y_i \in \mathbb{Q} \cap]x_i - r, x_i[$ *et* $z_i \in \mathbb{Q} \cap]x_i, x_i + r[$*. On a donc* $x \in \prod_{i=1}^n]y_i, z_i[\subset O$.

On note alors $I = \{(y,z) \in \mathbb{Q}^{2n} ; \prod_{i=1}^n]y_i, z_i[\subset O\}$ *(avec* $y = (y_1,\ldots,y_n)^t$ *et* $z = (z_1,\ldots,z_n)^t$*). Pour tout* $x \in O$*, il existe donc* $(y,z) \in I$ *t.q.* $x \in \prod_{i=1}^n]y_i, z_i[$*. On en déduit que* $O = \bigcup_{(y,z)\in I} \prod_{i=1}^n]y_i, z_i[$.

L'ensemble $\prod_{i=1}^{n-1}]y_i, z_i[$ *est un ouvert de* $\mathbb{R}^{n-1}$*, il appartient donc à* $\mathcal{B}(\mathbb{R}^{n-1})$ *(qui est la tribu engendrée par les ouverts de* $\mathbb{R}^{n-1}$*). L'ensemble* $]y_n, z_n[$ *appartient à* $\mathcal{B}(\mathbb{R})$*. Donc,* $\prod_{i=1}^n]y_i, z_i[\in \mathcal{B}(\mathbb{R}^{n-1}) \times \mathcal{B}(\mathbb{R})$*. Comme* I *est au plus dénombrable (car* $\mathbb{Q}^{2n}$ *est dénombrable), on en déduit que* $O \in T$*. On a ainsi montré que* T *est une tribu contenant tous les ouverts de* $\mathbb{R}^n$*, et donc contenant la tribu engendrée par les ouverts de* $\mathbb{R}^n$ *(c'est-à-dire* $\mathcal{B}(\mathbb{R}^n)$*). Donc,* $\mathcal{B}(\mathbb{R}^n) \subset T$.

Etape 2, $\mathcal{B}(\mathbb{R}^n) \supset T$**.**

On reprend ici aussi le même démarche que dans l'exercice 2.6.

1. *Soit* A *un ouvert de* $\mathbb{R}^{n-1}$ *et* $T_1 = \{B \in \mathcal{B}(\mathbb{R}); A \times B \in \mathcal{B}(\mathbb{R}^n)\}$. *On montre d'abord que* T_1 *est une tribu (sur* $\mathbb{R}$*)*

   — $\emptyset \in T_1$ *car* $A \times \emptyset = \emptyset \in \mathcal{B}(\mathbb{R}^n)$.

   — *On montre ici que* T_1 *est stable par passage au complémentaire. Soit* $B \in T_1$*, on a donc* $B^c \in \mathcal{B}(\mathbb{R})$ *et* $A \times B^c = A \times (\mathbb{R} \setminus B) = (A \times \mathbb{R}) \setminus (A \times B)$. *Or,* $(A \times \mathbb{R})$ *est un ouvert de* $\mathbb{R}^n$ *(car* A *est un ouvert de* $\mathbb{R}^{n-1}$ *et* $\mathbb{R}$ *est un ouvert de* $\mathbb{R}$*), on a donc* $(A \times \mathbb{R}) \in \mathcal{B}(\mathbb{R}^n)$. *D'autre part,* $(A \times B) \in \mathcal{B}(\mathbb{R}^n)$ *(car* $B \in T_1$*). Donc,* $A \times B^c = (A \times \mathbb{R}) \setminus (A \times B) \in \mathcal{B}(\mathbb{R}^n)$. *Ce qui prouve que* $B^c \in T_1$ *et donc que* T_1 *est stable par passage au complémentaire.*

   — *Enfin,* T_1 *est stable par union dénombrable. En effet, si* $(B_p)_{p\in\mathbb{N}} \subset T_1$*, on a* $A \times (\bigcup_{p\in\mathbb{N}} B_p) = \bigcup_{p\in\mathbb{N}} A \times B_p \in \mathcal{B}(\mathbb{R}^n)$ *(car* $A \times B_p \in \mathcal{B}(\mathbb{R}^n)$ *pour tout* $p \in \mathbb{N}$*). Donc,* $\bigcup_{p\in\mathbb{N}} B_p \in T_1$.

   On a donc montré que T_1 *est une tribu.*

   On montre maintenant que T_1 *contient les ouverts de* $\mathbb{R}$.

   Soit B *un ouvert de* $\mathbb{R}$. *On a donc* $B \in \mathcal{B}(\mathbb{R})$ *et, comme* $A \times B$ *est un ouvert de* $\mathbb{R}^n$*, on a* $A \times B \in \mathcal{B}(\mathbb{R}^n)$. *On a donc* $B \in T_1$.

   T_1 *est donc une tribu contenant les ouverts de* $\mathbb{R}$*, donc contenant* $\mathcal{B}(\mathbb{R})$. *Donc,* $T_1 = \mathcal{B}(\mathbb{R})$.

   La conséquence de ce résultat est :

   $$A \text{ ouvert de } \mathbb{R}^{n-1} \text{ et } B \in \mathcal{B}(\mathbb{R}) \Rightarrow A \times B \in \mathcal{B}(\mathbb{R}^n). \tag{7.11}$$

2. *Soit* $B \in \mathcal{B}(\mathbb{R})$ *et* $T_2 = \{A \in \mathcal{B}(\mathbb{R}^{n-1}); A \times B \in \mathcal{B}(\mathbb{R}^n)\}$. *On va montrer que* $T_2 = \mathcal{B}(\mathbb{R}^{n-1})$.

   On commence par remarquer que (7.11) donne que T_2 *contient les ouverts de* $\mathbb{R}^{n-1}$. *En effet, soit* A *un ouvert de* $\mathbb{R}^{n-1}$*, la propriété (7.11) donne que* $A \times B \in \mathcal{B}(\mathbb{R}^n)$*, et donc* $A \in T_2$.

   On montre maintenant que T_2 *est une tribu (on en déduira que* $T_2 = \mathcal{B}(\mathbb{R}^{n-1})$*).*

   — $\emptyset \in T_2$ *car* $\emptyset \times B = \emptyset \in \mathcal{B}(\mathbb{R}^n)$.

   — *On montre ici que* T_1 *est stable par passage au complémentaire. Soit* $A \in T_2$*, on a* $A^c \in \mathcal{B}(\mathbb{R}^{n-1})$ *et* $A^c \times B = (\mathbb{R}^{n-1} \times B) \setminus (A \times B)$. *La propriété (7.11) donne* $(\mathbb{R}^{n-1} \times B) \in \mathcal{B}(\mathbb{R}^n)$ *car* $\mathbb{R}^{n-1}$ *est un ouvert de* $\mathbb{R}^{n-1}$. *D'autre part,* $(A \times B) \in \mathcal{B}(\mathbb{R}^n)$ *(car* $A \in T_2$*). Donc,* $A^c \times B \in \mathcal{B}(\mathbb{R}^n)$. *Ce qui prouve que* $A^c \in T_2$ *et donc que* T_2 *est stable par passage au complémentaire.*

   — *Enfin,* T_2 *est stable par union dénombrable. En effet, si* $(A_p)_{p\in\mathbb{N}} \subset T_2$*, on a* $(\bigcup_{p\in\mathbb{N}} A_p) \times B = \bigcup_{p\in\mathbb{N}} (A_p \times B) \in \mathcal{B}(\mathbb{R}^n)$ *(car* $A_p \times B \in \mathcal{B}(\mathbb{R}^n)$ *pour tout* $p \in \mathbb{N}$*). Donc,* $\bigcup_{p\in\mathbb{N}} A_p \in T_2$.

   T_2 *est donc une tribu (sur* $\mathbb{R}^{n-1}$*) contenant les ouverts de* $\mathbb{R}^{n-1}$*, ce qui prouve que* $T_2 \supset \mathcal{B}(\mathbb{R}^{n-1})$ *et donc, finalement,* $T_2 = \mathcal{B}(\mathbb{R}^{n-1})$.

   On a donc obtenu le résultat suivant :

   $$A \in \mathcal{B}(\mathbb{R}^{n-1}), B \in \mathcal{B}(\mathbb{R}) \Rightarrow A \times B \in \mathcal{B}(\mathbb{R}^n). \tag{7.12}$$

3. *On montre maintenant que* $T \subset \mathcal{B}(\mathbb{R}^n)$ *(et donc que* $T = \mathcal{B}(\mathbb{R}^n)$*).*

   Grâce à (7.12), on a $\{A \times B;\ A \in \mathcal{B}(\mathbb{R}^{n-1}),\ B \in \mathcal{B}(\mathbb{R})\} \subset \mathcal{B}(\mathbb{R}^n)$. *On en déduit que* $T \subset \mathcal{B}(\mathbb{R}^n)$. *On a donc bien, avec l'étape 1,* $T = \mathcal{B}(\mathbb{R}^n)$.

Exercice 7.6 (Intégrale d'une fonction positive) Soit (E, T, m) un espace mesuré σ-fini, et $f : E \to \mathbb{R}_+$ une application mesurable. On pose $F = 1_A$ avec $A = \{(t, x) \in \mathbb{R} \times E;\ 0 < t < f(x)\}$.

1. Montrer que F est $\mathcal{B}(\mathbb{R}) \otimes T$- mesurable

Corrigé – *Comme $f \in \mathcal{M}_+$, il existe $(f_n) \in \mathcal{E}_+$ t.q. $f_n \uparrow f$ quand $n \to +\infty$. On a alors $A = \bigcup_{n \in \mathbb{N}} A_n$ avec $A_n = \{(t, x) \in \mathbb{R}_+ \times \mathbb{R};\ 0 < t < f_n(x)\}$. Pour montrer que F est mesurable (ce qui équivalent à montrer que $A \in \mathcal{B}(\mathbb{R}) \otimes T$), il suffit donc de montrer que $A_n \in \mathcal{B}(\mathbb{R}) \otimes T$ pour tout $n \in \mathbb{N}$.*

Soit donc $n \in \mathbb{N}$. Il existe $a_1, \ldots, a_p \in \mathbb{R}_+^*$ et $B_1, \ldots, B_p \in T$ t.q. $B_i \cap B_j = \emptyset$ si $i \neq j$ et $f_n = \sum_{i=1}^{p} a_i 1_{B_i}$. On a donc $A_n = \bigcup_{i=1}^{p}]0, a_i[\times B_i \in \mathcal{B}(\mathbb{R}) \otimes T$ car $]0, a_i[\times B_i \in \mathcal{B}(\mathbb{R}) \times T \subset \mathcal{B}(\mathbb{R}) \otimes T$ pour tout i.

2. Montrer que $\int f dm = \int_0^{+\infty} m(\{x \in E; f(x) > t\}) dt$ et que $\lambda \otimes m(A) = \int f dm$. [Utiliser le théorème de Fubini-Tonelli.]

Corrigé – *On peut appliquer le théorème de Fubini-Tonelli (théorème 7.7) à la fonction F, il donne :*

$$\int F d(\lambda \otimes A) = \int \left(\int F(t, x) d\lambda(t) \right) dm(x) = \int \left(\int F(t, x) dm(x) \right) d\lambda(t). \tag{7.13}$$

Pour $x \in E$, $\int F(t, x) d\lambda(t) = \lambda(]0, f(x)[) = f(x)$.

Pour $t \in \mathbb{R}$. Si $t \leq 0$, on a $F(t, \cdot) = 0$. Donc, $\int F(t, x) dm(x) = 0$. Si $t > 0$, on a $F(t, \cdot) = 1_{\{f > t\}}$. Donc, $\int F(t, x) dm(x) = m(\{f > t\})$.

On déduit donc de (7.13) :

$$\int F d(\lambda \otimes A) = \int f(x) dm(x) = \int_{\mathbb{R}_+} m(\{f > t\}) d\lambda(t) = \int_0^{\infty} m(\{x \in E;\ f(x) > t\}) dt.$$

Comme $F = 1_A$, on a aussi $\int F d(\lambda \otimes m) = \lambda \otimes m(A)$, ce qui termine cette question.

3. Montrer que $\int f dm = \int_0^{+\infty} m(\{x \in E; f(x) \geq t\}) dt$.

Corrigé – *On reprend le raisonnement de la question précédente en remplaçant F par $G = 1_B$ avec $B = \{(t, x) \in \mathbb{R} \times E;\ 0 < t \leq f(x)\}$. On remarque d'abord que B est $\mathcal{B}(\mathbb{R}) \otimes T$-mesurable. En effet, on a $B = \bigcap_{n \in \mathbb{N}^*} B_n$ avec $B_n = \{(t, x) \in \mathbb{R} \times E;\ 0 < t < f_n(x)\}$ et $f_n = f + \frac{1}{n}$. Comme $f_n \in \mathcal{M}_+$, la première question donne $B_n \in \mathcal{B}(\mathbb{R}) \otimes T$. On en déduit que $B \in B(\mathbb{R}) \otimes T$. On peut maintenant appliquer le théorème de Fubini-Tonelli à la fonction G, il donne :*

$$\int \left(\int G(t, x) d\lambda(t) \right) dm(x) = \int \left(\int G(t, x) dm(x) \right) d\lambda(t). \tag{7.14}$$

Pour $x \in E$, $\int G(t, x) d\lambda(t) = \lambda(]0, f(x)]) = f(x)$.

Pour $t \in \mathbb{R}$. Si $t \leq 0$, on a $G(t, \cdot) = 0$. Donc, $\int G(t, x) dm(x) = 0$. Si $t > 0$, on a $G(t, \cdot) = 1_{\{f \geq t\}}$. Donc, $\int G(t, x) dm(x) = m(\{f \geq t\})$.

On déduit donc de (7.14) :

$$\int f(x)dm(x) = \int_0^\infty m(\{x \in E;\ f(x) \geq t\})dt.$$

Exercice 7.7 (Espace L^p faible) On note $\mathcal{L}^1$ l'espace des fonctions intégrables de $\mathbb{R}$ dans $\mathbb{R}$ pour la mesure de Lebesgue et $\mathcal{L}^2$ l'espace des fonctions de carré intégrable de $\mathbb{R}$ dans $\mathbb{R}$ pour la mesure de Lebesgue.

Définition 7.30 *Soit $p \in \{1, 2\}$. Une fonction borélienne f de $\mathbb{R}$ dans $\mathbb{R}$ appartient à l'espace M^p si il existe $C \in \mathbb{R}$ tel que, pour tout $a > 0$,*

$$\lambda(\{|f| \geq a\}) \leq \frac{C}{a^p}. \tag{7.15}$$

Si $f \in M^p$, on note $C_f = \sup_{a>0} a^p \lambda(\{|f| \geq a\})$.

1. Montrer que M^1 et M^2 sont des espaces vectoriels (sur $\mathbb{R}$).

Corrigé – *Soit $p \in \{1, 2\}$.*
Soit $f \in M^p$ et $\alpha \in \mathbb{R}^*$. On remarque que, pour tout $a > 0$,

$$\{|\alpha f| \geq a\} = \{|f| \geq \frac{a}{|\alpha|}\} \leq \frac{C_f |\alpha|^p}{a^p},$$

et donc $\alpha f \in M^p$. Pour $\alpha = 0$, On a aussi $\alpha f \in M^p$. On a donc $\alpha f \in M^p$ pour tout $f \in M^p$ et tout $\alpha \in \mathbb{R}$.
Soit maintenant $f, g \in M^p$. Pour tout $a > 0$, on a $\{|f+g| \geq a\} \subset \{|f| \geq a/2\} \cup \{|g| \geq a/2\}$ et donc

$$\lambda(\{|f+g| \geq a\}) \leq \lambda(\{|f| \geq a/2\}) + \lambda(\{|g| \geq a/2\}) \leq 2^p(C_f + C_g)a^p,$$

ce qui prouve que $f + g \in M^p$ et finalement que M^p est bien un espace vectoriel sur $\mathbb{R}$.

2. Montrer que $\mathcal{L}^p \subset M^p$, pour $p = 1$ et $p = 2$.

Corrigé – *Soient $p \in \{1, 2\}$ et $f \in \mathcal{L}^p$. Pour tout $a > 0$, on a $a 1_{\{|f| \geq a\}} \leq |f|$ p.p. et donc $a^p \lambda(\{|f| \geq a\} \leq \|f\|_p^p$, ce qui prouve que $f \in M^p$.*

3. Donner un exemple de fonction qui appartient à M^1 mais n'appartient pas à $\mathcal{L}^1$.

Corrigé – *On définit f par $f(x) = 1/x$ si $x > 0$ et $f(x) = 0$ si $x \leq 0$. On a $f \notin \mathcal{L}^1$ et, pour tout $a > 0$, $\{|f| > a\} = \{x;\ 0 < x < 1/a\}$ et donc $\lambda(\{|f| > a\}) = 1/a$. Ce qui prouve que $f \in M^1$.*

4. Soit $f \in M^2$ et K un compact de $\mathbb{R}$. Montrer que $f 1_K \in \mathcal{L}^1$. [On pourra utiliser le fait (démontré dans l'exercice 7.6) que pour toute fonction borélienne g de $\mathbb{R}$ dans $\mathbb{R}$, on a $\int |g(x)|dx = \int_0^\infty \lambda(\{|g| \geq a\})da$.]

Corrigé – *On pose $g = f1_K$ et on remarque que, pour tout $a > 0$, $\lambda(\{|g| \geq a\}) \leq \min(\lambda(K), C_f/a^2)$. On en déduit*

$$\int_K |f| d\lambda = \int |g| d\lambda = \int_0^\infty \lambda(\{|g| \geq a\}) da \leq \lambda(K) + \int_1^\infty \frac{C_f}{a^2} da < +\infty,$$

et donc que $f1_K \in \mathcal{L}^1$.

5. Donner un exemple pour lequel $f \in M^2$ et $f \notin \mathcal{L}^2$.

Corrigé – *On définit f par $f(x) = 1/\sqrt{x}$ si $x > 0$ et $f(x) = 0$ si $x \leq 0$. On a $f \notin \mathcal{L}^2$ et, pour tout $a > 0$, $\{|f| > a\} = \{x; 0 < x < 1/a^2\}$ et donc $\lambda(\{|f| > a\}) = 1/a^2$. Ce qui prouve que $f \in M^2$.*

6. Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de M^2 et $f \in M^2$. on suppose que, pour tout $n \in \mathbb{N}$, il existe δ_n tel que $\lim_{n\to+\infty} \delta_n = 0$ et

$$\lambda(\{|f_n - f| \geq a\}) \leq \frac{\delta_n}{a^2} \text{ pour tout } a > 0 \text{ et tout } n \in \mathbb{N}.$$

Montrer que $\lim_{n\to+\infty} \int_K |f_n(x) - f(x)| dx = 0$ pour tout compact K de $\mathbb{R}$.

Corrigé – *Soit K un compact de $\mathbb{R}$. On pose $g_n = (f_n - f)1_K$ et on remarque encore que, pour tout $a > 0$, $\lambda(\{|g_n| \geq a\}) \leq \min(\lambda(K), \delta_n/a^2)$. On en déduit que pour $\eta > 0$ et tout $n \in \mathbb{N}$,*

$$\int_K |f_n - f| d\lambda = \int |g_n| d\lambda = \int_0^\infty \lambda(\{|g_n| \geq a\}) da \leq \eta\lambda(K) + \int_\eta^\infty \frac{\delta_n}{a^2} da.$$

Soit $\varepsilon > 0$. On choisir d'abord $\eta > 0$ pour avoir $\eta\lambda(K) \leq \varepsilon$.
Puis, le fait que $\lim_{n\to+\infty} \delta_n = 0$ et que $\int_\eta^\infty 1/a^2 da < +\infty$ donnent l'existence de n_0 tel que

$$n \geq n_0 \Rightarrow \delta_n \int_\eta^{+\infty} \frac{1}{a_2} da \leq \varepsilon.$$

On a donc

$$n \geq n_0 \Rightarrow \int_K |f_n - f| d\lambda \leq 2\varepsilon.$$

La convergence vers 0 de $\int_K |f_n(x) - f(x)| dx$ est bien démontrée.

Exercice 7.10 (Intégrale de Dirichlet)

1. Vérifier que si $n \geq 1$,, $\displaystyle\int_0^n \frac{\sin x}{x} dx = \int_0^n (\int_0^\infty e^{-xt} dt) \sin x dx.$

Corrigé – *Soit $n \in \mathbb{N}$. En posant $\frac{\sin x}{x} = 1$ pour $x = 0$, la fonction $x \mapsto \frac{\sin x}{x}$ est continue sur $[0, n]$, elle est donc bien intégrable pour l'espace mesuré $([0, n], \mathcal{B}([0, n]), \lambda)$.*

Pour tout $x > 0$, on a $e^{x\cdot} \in \mathcal{M}_+(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. Pour calculer $\int_0^\infty e^{-xt} dt$, on remarque que, pour tout $p \in \mathbb{N}$, on a $\int_0^p e^{-xt} dt = [\frac{-e^{-xt}}{x}]_0^p = \frac{1}{x} - \frac{e^{-xp}}{x}$. Quand $p \to \infty$, on en déduit, avec le théorème de convergence monotone, que $\int_0^\infty e^{-xt} dt = \frac{1}{x}$. On obtient bien :

$$\int_0^n \frac{\sin x}{x} dx = \int_0^n (\int_0^\infty e^{-xt} dt) \sin x dx.$$

2. Calculer $\displaystyle \mathrm{F}_n(t) = \int_0^n e^{-xt} \sin x dx$ $(t \geq 0)$.

Corrigé – *Pour $t = 0$, on a $\int_0^n e^{-xt} \sin x dx = \int_0^n \sin x dx = 1 - \cos n$. On a donc $\mathrm{F}_n(0) = 1 - \cos n$.*

Soit maintenant $t > 0$. Comme les fonctions $x \mapsto e^{-xt}$ et $x \mapsto \sin x$ sont indéfiniment dérivables sur $\mathbb{R}$, on peut calculer $\int_0^n e^{-xt} \sin x dx$ en intégrant deux fois par parties :

$$\begin{aligned}\int_0^n e^{-xt} \sin x dx &= -\int_0^n t e^{-xt} \cos x dx - [e^{-xt} \cos x]_0^n \\ &= -\int_0^n t^2 e^{-xt} \sin x dx - [t e^{-xt} \sin x]_0^n - [e^{-xt} \cos x]_0^n.\end{aligned}$$

Ce qui donne :

$$(t^2 + 1) \int_0^n e^{-xt} \sin x dx = 1 - t e^{-nt} \sin n - e^{-nt} \cos n.$$

et donc :

$$\mathrm{F}_n(t) = \int_0^n e^{-xt} \sin x dx = \frac{1 - t e^{-nt} \sin n - e^{-nt} \cos n}{t^2 + 1}.$$

3. Calculer $\displaystyle\lim_{n \to +\infty} \int_0^\infty \mathrm{F}_n(t) dt$. ($\mathrm{F}_n$ est définie à la question précédente.)

Corrigé – *Pour tout $n \in \mathbb{N}$, F_n est continue de $\mathbb{R}_+$ dans $\mathbb{R}$, elle est donc mesurable de $\mathbb{R}_+$ dans $\mathbb{R}$.*

On a, pour tout $n \in \mathbb{N}$ et tout $t \geq 0$, $t e^{-nt} \leq t e^{-t} \leq 1/e$. On en déduit :

$$|\mathrm{F}_n(t)| \leq (2 + \frac{1}{e}) \frac{1}{t^2 + 1} \text{ pour tout } t \in \mathbb{R}+ \text{ et pour tout } n \in \mathbb{N}.$$

(en fait, on a même $0 \leq \mathrm{F}_n(t) \leq \frac{1}{t^2+1}$.) Comme $t \mapsto \frac{1}{t^2+1}$ est intégrable pour la mesure de Lebesgue sur les boréliens de $\mathbb{R}_+$, ceci donne $\mathrm{F}_n \in \mathcal{L}^1(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda)$ pour tout $n \in \mathbb{N}$.

Enfin comme $F_n(t) \to \frac{1}{t^2+1}$ *quand* $n \to +\infty$, *pour tout* $t > 0$, *on peut appliquer le théorème de convergence dominée à la suite* $(F_n)_{n\in\mathbb{N}}$, *il donne :*

$$\int_0^\infty F_n(t)dt \to \int_0^\infty \frac{1}{t^2+1}dt = \frac{\pi}{2}, \text{ quand } n \to +\infty.$$

4. En déduire que $\lim_{n\to+\infty} \int_0^n \frac{\sin x}{x}dx = \frac{\pi}{2}$.

Corrigé – *Soit* $n \in \mathbb{N}$. *On définit* H *de* $\mathbb{R}^2$ *dans* $\mathbb{R}$ *par*

$$H(x,t) = e^{-xt}\sin x 1_{[0,n]}(x)1_{[0,\infty]}(t).$$

La fonction H *est* $\mathcal{B}(\mathbb{R}^2)$*-mesurable et elle appartient à* $\mathcal{L}^1(\mathbb{R}^2)$ *car, par le théorème de Fubini-Tonelli :*

$$\int |H(x,t)|d\lambda_2(x,t) \le \int_0^n |\sin x|(\int_0^\infty e^{-xt}dt)dx = \int_0^n \frac{|\sin x|}{x}dx < \infty,$$

car la fonction $x \mapsto \frac{|\sin x|}{x}$ *est continue que* $[0,n]$ *en posant* $\frac{|\sin x|}{x} = 1$ *pour* $x = 0$, *elle est donc bien intégrable pour l'espace mesuré* $([0,n],\mathcal{B}([0,n]),\lambda)$.

On peut donc appliquer le théorème de Fubini à la fonction H, *il donne, avec la première question :*

$$\int_0^n \frac{\sin x}{x}dx = \int_0^n(\int_0^\infty e^{-xt}dt)\sin x dx = \int_0^\infty(\int_0^n e^{-xt}\sin x dx)dt$$
$$= \int_0^\infty F_n(t)dt.$$

La question 3 donne alors $\lim_{n\to+\infty}\int_0^n \frac{\sin x}{x}dx = \lim_{n\to+\infty}\int_0^\infty F_n(t)dt = \frac{\pi}{2}$.

7.7.3 Mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$

Exercice 7.11 (Propriétés élémentaires de λ_N) Soit $N \ge 2$. On rappelle que λ_N est la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$.

1. Soit $A_1,\ldots,A_N \in \mathcal{B}(\mathbb{R})$.
Montrer que $\prod_{i=1}^N A_i \in \mathcal{B}(\mathbb{R}^N)$ et $\lambda_N(\prod_{i=1}^N A_i) = \prod_{i=1}^N \lambda(A_i)$.

Corrigé – *On va démontrer cette question en supposant tout d'abord que* $\lambda(A_i) < \infty$ *pour tout* i. *Le cas général s'obtient alors en utilisant* $A_i \cap [-p,p]$ *au lieu de* A_i *et en faisant ensuite tendre* p *vers l'infini. On obtient bien la propriété voulue (en convenant que* $0 \times \infty = 0$*). Cette méthode est décrite dans la remarque 7.5.*

On démontre donc, par récurrence sur N, *la propriété suivante :*

$$\begin{aligned}&A_1,\ldots,A_N \in \mathcal{B}(\mathbb{R}),\ \lambda(A_i) < \infty \text{ pour tout } i \Rightarrow\\ &\prod_{i=1}^N A_i \in \mathcal{B}(\mathbb{R}^N) \text{ et } \lambda_N(\prod_{i=1}^N A_i) = \prod_{i=1}^N \lambda(A_i).\end{aligned} \tag{7.18}$$

La propriété (7.18) est vraie pour $N = 2$. *En effet, on sait que* $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$ *(proposition 7.2) et que* $\lambda_2 = \lambda\otimes\lambda$ *(définition 7.15). On a donc bien (avec la définition* *d'une mesure produit, théorème 7.3) :*

$$A_1, A_2 \in \mathcal{B}(\mathbb{R}),\ \lambda(A_1) < \infty,\ \lambda(A_2) < \infty$$
$$\Rightarrow A_1 \times A_2 \in \mathcal{B}(\mathbb{R}^2) \text{ et } \lambda_2(A_1 \times A_2) = \lambda(A_1)\lambda(A_2).$$

On suppose maintenant que la propriété (7.18) est vraie pour un certain $N \geq 2$, *et on la démontre pour* $N+1$.

Soit donc $A_1, \ldots, A_{N+1} \in \mathcal{B}(\mathbb{R})$ *t.q.* $\lambda(A_i) < \infty$ *pour tout* i.
Par (7.18), on a $\prod_{i=1}^N A_i \in \mathcal{B}(\mathbb{R}^N)$ *et* $\lambda_N(\prod_{i=1}^N A_i) = \prod_{i=1}^N \lambda(A_i)$. *On rappelle que* $\mathcal{B}(\mathbb{R}^{N+1}) = \mathcal{B}(\mathbb{R}^N) \otimes \mathcal{B}(\mathbb{R})$ *(proposition 7.2) et que* $\lambda_{N+1} = \lambda_N \otimes \lambda$ *(définition 7.15).*
On en déduit que $\prod_{i=1}^{N+1} A_i = (\prod_{i=1}^N A_i) \times A_{N+1} \in \mathcal{B}(\mathbb{R}^N) \times \mathcal{B}(\mathbb{R}) \subset \mathcal{B}(\mathbb{R}^N) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^{N+1})$ *et*

$$\lambda_{N+1}(\prod_{i=1}^{N+1} A_i) = \lambda_N(\prod_{i=1}^{N} A_i)\lambda(A_{N+1}) = \prod_{i=1}^{N+1} \lambda(A_i).$$

ce qui donne bien (7.18) avec $N+1$ *au lieu de* N *et termine donc la démonstration par récurrence.*

2. Soit $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$ et $\beta_1, \ldots, \beta_N \in \mathbb{R}$ t.q. $\alpha_i < \beta_i$ pour tout $i \in \{1, \ldots, N\}$. Montrer que
$$\lambda_N(\prod_{i=1}^N]\alpha_i, \beta_i[) = \prod_{i=1}^N \lambda(]\alpha_i, \beta_i[).$$

   Corrigé – *Cette question est une conséquence immédiate de la précédente en prenant* $A_i =]\alpha_i, \beta_i[$.

3. Soit K est un compact de $\mathbb{R}^N$ (noter que $K \in \mathcal{B}(\mathbb{R}^N)$). Montrer que $\lambda_N(K) < +\infty$.

   Corrigé – *Comme* K *est compact, il est borné. Il existe donc* $a \in \mathbb{R}_+^*$ *t.q.* $K \subset \prod_{i=1}^N]-a, a[$. *On en déduit que* $\lambda_N(K) \leq \lambda_N(\prod_{i=1}^N]-a, a[) = \prod_{i=1}^N \lambda(]-a, a[) = (2a)^N < \infty$.

4. Soit O un ouvert non vide de $\mathbb{R}^N$. Montrer que $\lambda_N(O) > 0$.

   Corrigé – *Soit* $x = (x_1, \ldots, x_N)^t \in O$. *Comme* O *est ouvert, il existe* $\varepsilon > 0$ *t.q.* $\prod_{i=1}^N]x_i - \varepsilon, x_i + \varepsilon[\subset O$. *On a donc* $\lambda_N(O) \geq \lambda_N(\prod_{i=1}^N]x_i - \varepsilon, x_i + \varepsilon[) = \prod_{i=1}^N \lambda(]x_i - \varepsilon, x_i + \varepsilon[) = (2\varepsilon)^N > 0$.

5. Soit $f, g \in C(\mathbb{R}^N, \mathbb{R})$. Montrer que $f = g$ p.p. (c'est-à-dire λ_N-p.p.) implique $f(x) = g(x)$ pour tout $x \in \mathbb{R}^N$.

   Corrigé – *Soit* $O = \{f \neq g\} = \{x \in \mathbb{R}^N; f(x) \neq g(x)\}$. *Comme* f *et* g *sont continues,* O *est ouvert. Comme* $f = g$ *p.p., on a nécessairement* $\lambda_N(O) = 0$. *Enfin, la question précédente donne alors que* $O = \emptyset$ *et donc que* $f(x) = g(x)$ *pour tout* $x \in \mathbb{R}^N$.

6. Montrer que $C_c(\mathbb{R}^N, \mathbb{R}) \subset \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$.

Corrigé – *Soit $f \in C_c(\mathbb{R}^N, \mathbb{R})$. Comme f est continue, on a f mesurable de $\mathbb{R}^N$ dans $\mathbb{R}$ ($\mathbb{R}^N$ et $\mathbb{R}$ étant munis de leur tribu borélienne, on dit aussi que f est borélienne). Comme f est à support compact, il existe $a \in \mathbb{R}_+^*$ t.q. $f = 0$ sur K^c avec $K = \prod_{i=1}^N [-a, a]$. Enfin, f est bornée, il existe donc M t.q. $|f(x)| \le M$ pour tout $x \in \mathbb{R}^N$. On en déduit que $\int |f| d\lambda_N \le M(2a)^N < \infty$ et donc que $f \in \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$.*

Exercice 7.12 (Régularité de λ_N) Soit m une mesure sur $\mathcal{B}(\mathbb{R}^N)$ t.q. $m(K) < \infty$ pour tout compact K de $\mathbb{R}^N$. (noter que ceci est vrai pour $m = \lambda_N$.)

1. Soient $A \in \mathcal{B}(\mathbb{R}^N)$ et $\varepsilon > 0$. Montrer qu'il existe O ouvert de $\mathbb{R}^N$ et F fermé de $\mathbb{R}^N$ tels que :
$$F \subset A \subset O \text{ et } m(O \setminus F) \le \varepsilon.$$

Corrigé – *On reprend ici la démonstration de la régularité d'une mesure sur $\mathcal{B}(\mathbb{R})$, finie sur les compacts (théorème 2.43).*

On appelle T l'ensemble des $A \in \mathcal{B}(\mathbb{R}^N)$ t.q. pour tout $\varepsilon > 0$, il existe O ouvert et F fermé vérifiant $F \subset A \subset O$ et $m(O \setminus F) \le \varepsilon$. On va montrer que T est une tribu contenant $\mathcal{C} = \{\prod_{i=1}^N]a_i, b_i[, -\infty < a_i < b_i < \infty$ pour tout $i \in \{1, \ldots, N\}\}$. Comme $\mathcal{C}$ engendre $\mathcal{B}(\mathbb{R}^N)$ (voir l'exercice 2.7, il est même démontré qu'on peut, dans la définition de $\mathcal{C}$, se limiter au cas où les a_i et b_i sont dans $\mathbb{Q}$), ceci donnera $T = \mathcal{B}(\mathbb{R}^N)$.

On démontre tout d'abord que $\mathcal{C} \subset T$. Soit $-\infty < a_i < b_i < \infty$ pour tout $i \in \{1, \ldots, N\}$ et $A = \prod_{i=1}^N]a_i, b_i[$. Soit $\varepsilon > 0$. On veut montrer qu'il existe O ouvert et F fermé t.q. $F \subset A \subset O$ et $m(O \setminus F) \le \varepsilon$.

Soit $n_0 \in \mathbb{N}$ t.q. $(2/n_0) < b_i - a_i$ pour tout $i \in \{1, \ldots, N\}$. Pour $n \ge n_0$, on pose $F_n = \prod_{i=1}^N [a_i + (1/n), b_i - (1/n)]$ et $O = A$. On a bien F_n fermé, O ouvert et $F_n \subset A \subset O$. On remarque ensuite que $O \setminus F_n \subset C_n$ avec :
$$C_n = \bigcup_{q=1}^N C_{n,q}, \;\; C_{n,q} = \prod_{i=1}^N I_{i,q}^{(n)},$$
$$I_{i,q}^{(n)} =]a_i, b_i[\text{ si } i \ne q, \; I_{q,q}^{(n)} =]a_q, a_q + \frac{1}{n}[\cup]b_q - \frac{1}{n}, b_q[.$$
Soit $q \in \{1, \ldots, N\}$. On a $C_{n+1,q} \subset C_{n,q}$ (pour tout $n \ge n_0$), $\bigcap_{n \ge n_0} C_{n,q} = \emptyset$ et, comme m est finie sur les compacts,
$$m(C_{n,q}) \le m(\prod_{i=1}^N [a_i, b_i]) < \infty.$$
On peut donc utiliser la propriété de continuité décroissante d'une mesure. On obtient $m(C_{n,q}) \to 0$ quand $n \to +\infty$. On a alors aussi $m(C_n) \le \sum_{q=1}^N m(C_{n,q}) \to 0$ quand $n \to +\infty$. Il existe donc n t.q. $m(C_n) < \varepsilon$. On prenant $F = F_n$, on a bien alors O ouvert, F fermé et $m(O \setminus F) \le \varepsilon$, ce qui montre que $A \in T$ et donc que $\mathcal{C} \subset T$.

On montre maintenant que T est une tribu. On remarque tout d'abord que $\emptyset \in T$ (il suffit de prendre $F = O = \emptyset$) et que T est stable par passage au complémentaire (car, si $F \subset A \subset O$, on a $O^c \subset A^c \subset F^c$ et $F^c \setminus O^c = O \setminus F$). Il reste à montrer que T est stable par union dénombrable.

Soit $(A_n)_{n\in\mathbb{N}} \subset T$ *et* $A = \bigcup_{n\in\mathbb{N}} A_n$*. On veut montrer que* $A \in T$*. On va commencer par traiter le cas (simple) où* $m(A) < \infty$ *puis le cas (plus difficile) où* $m(A) = \infty$*.*

Premier cas. *On suppose que* $m(A) < \infty$*. La démonstration est ici identique à celle faite pour* $N = 1$*. Soit* $\varepsilon > 0$*. Pour tout* $n \in \mathbb{N}$*, il existe* O_n *ouvert et* F_n *fermé t.q.* $F_n \subset A_n \subset O_n$ *et* $m(O_n \setminus F_n) \le (\varepsilon/2^n)$*. On pose* $O = \bigcup_{n\in\mathbb{N}} O_n$ *et* $\tilde{F} = \bigcup_{n\in\mathbb{N}} F_n$*. On a* $\tilde{F} \subset A \subset O$*,* $m(O \setminus \tilde{F}) \le 2\varepsilon$*, car* $(O \setminus \tilde{F}) \subset \bigcup_{n\in\mathbb{N}}(O_n \setminus F_n)$*, et* O *ouvert mais* $\tilde{F}$ *n'est pas nécessairement fermé…*

Cependant, puisque $m(A) < \infty$*, on a aussi* $m(\tilde{F}) < \infty$*. Par continuité croissante de m on a* $m(\bigcup_{p=0}^n F_n) \to m(\tilde{F})$*, quand* $n \to \infty$*, d'où (puisque* $m(\tilde{F}) < \infty$*)* $m(\tilde{F}) - m(\bigcup_{p=0}^n F_n) \to 0$*. On prend alors* $F = \bigcup_{p=0}^n F_n$ *avec n assez grand pour que* $m(\tilde{F}) - m(F) \le \varepsilon$*. On a bien* $F \subset A \subset O$*,* O *ouvert,* F *fermé et* $m(O \setminus F) \le 3\varepsilon$*. Ceci prouve que* $A \in T$*.*

Deuxième cas. *On suppose maintenant que* $m(A) = \infty$ *(et le raisonnement précédent n'est plus correct si* $m(\tilde{F}) = \infty$*). On raisonne en trois étapes, en adaptant la démonstration faite pour* $N = 1$ *:*

(a) Soit $p = (p_1, \dots, p_N) \in \mathbb{Z}^N$*. On remarque d'abord que* $A_n \cap \prod_{i=1}^N [p_i, p_i + 1[\in T$ *pour tout* $n \in \mathbb{N}$*. En effet, soit* $n \in \mathbb{N}$ *et* $\varepsilon > 0$*. Il existe* O *ouvert et* F *fermé t.q.* $F \subset A_n \subset O$ *et* $m(O \setminus F) \le \varepsilon$*. Pour* $k \in \mathbb{N}^*$*, on a donc*

$$F_k = F \cap \prod_{i=1}^N [p_i, p_i + 1 - \frac{1}{k}] \subset A_n \cap \prod_{i=1}^N [p_i, p_i + 1[\subset O_k = O \cap \prod_{i=1}^N]p_i - \frac{1}{k}, p_i + 1[.$$

On a F_k *fermé,* O_k *ouvert et* $(O_k \setminus F_k) \subset (O \setminus F) \cup D_k$*, avec :*

$$D_k = \bigcup_{q=1}^N D_{k,q},\quad D_{k,q} = \prod_{i=1}^N J_{i,q}^{(k)},$$
$$J_{i,q}^{(k)} =]p_i - \frac{1}{k}, p_i + 1[\text{ si } i \ne p,$$
$$J_{q,q}^{(k)} =]p_q - \frac{1}{k}, p_q[\cup]p_q + 1 - \frac{1}{k}, p_q + 1[.$$

En utilisant la continuité décroissante de m et le fait que m est finie sur les compacts (ce qui donne $m(D_k) \le m(\prod_{i=1}^N [p_i - 1, p_i + 1]) < \infty$*), on démontre (comme pour les* C_n *précédemment) que* $m(D_k) \to 0$ *quand* $k \to \infty$*. Il existe donc* $k \in \mathbb{N}^*$ *t.q.* $m(D_k) \le \varepsilon$ *et donc* $m(O_k \setminus F_k) \le m(O \setminus F) + m(D_k) \le 2\varepsilon$*. Ce qui donne bien que* $A_n \cap \prod_{i=1}^N [p_i, p_i + 1[\in T$*.*

(b) Soit $p = (p_1, \dots, p_N) \in \mathbb{Z}^N$*. Comme* $m(A \cap \prod_{i=1}^N [p_i, p_i + 1[) < \infty$*, on peut maintenant utiliser le premier cas avec* $A \cap \prod_{i=1}^N [p_i, p_i + 1[= \bigcup_{n\in\mathbb{N}} (A_n \cap \prod_{i=1}^N [p_i, p_i + 1[)$*. Il donne que* $A \cap \prod_{i=1}^N [p_i, p_i + 1[\in T$ *pour tout* $p = (p_1, \dots, p_N) \in \mathbb{Z}^N$*.*

(c) On montre enfin que $A \in T$*. Soit* $\varepsilon > 0$*. Pour tout* $p = (p_1, \dots, p_N) \in \mathbb{Z}^N$*, il existe un ouvert* O_p *et un fermé* F_p *t.q.* $F_p \subset A \cap \prod_{i=1}^N [p_i, p_i + 1[\subset O_p$ *et* $m(O_p \setminus F_p) \le \varepsilon/(2^{|p|})$*, en posant* $|p| = \sum_{i=1}^N |p_i|$*. On prend* $O = \bigcup_{p\in\mathbb{Z}^N} O_p$ *et* $F = \bigcup_{p\in\mathbb{Z}^N} F_p$*. On obtient* $F \subset A \subset O$*,* $m(O \setminus F) \le 3^N \varepsilon$ *et* O *est ouvert. Il reste à montrer que* F *est fermé.*

Soit $(x_n)_{n\in\mathbb{N}} \subset \mathrm{F}$ *t.q.* $x_n \to x$ *(dans* $\mathbb{R}^{\mathrm{N}}$*) quand* $n \to +\infty$*. On veut montrer que* $x \in \mathrm{F}$*. Il existe* $p = (p_1, \ldots, p_{\mathrm{N}}) \in \mathbb{Z}^{\mathrm{N}}$ *t.q.* $x \in \prod_{i=1}^{\mathrm{N}}]p_i - 1, p_i + 1[$*. Il existe donc* $n_0 \in \mathbb{N}$ *t.q.* $x_n \in \prod_{i=1}^{\mathrm{N}}]p_i - 1, p_i + 1[$ *pour tout* $n \geq n_0$*. Comme* $x_n \in \bigcup_{q\in\mathbb{Z}^{\mathrm{N}}} \mathrm{F}_q$ *et que* $\mathrm{F}_q \subset \prod_{i=1}^{\mathrm{N}} [q_i, q_i + 1[$ *pour tout* $q = (q_1, \ldots, q_{\mathrm{N}})^t \in \mathrm{Z}^{\mathrm{N}}$*, on a donc* $x_n \in \bigcup_{q\in\mathrm{E}_p} \mathrm{F}_q$*, pour tout* $n \geq n_0$*, où* $\mathrm{E}_q = \{q = (q_1, \ldots, q_{\mathrm{N}})^t \in \mathbb{Z}^{\mathrm{N}}\,;\, q_i \in \{p_i, p_i - 1\}$ *pour tout* $i \in \{1, \ldots, \mathrm{N}\}\}$*. Comme* E_q *est de cardinal fini et que* F_q *est fermé pour tout* q*, l'ensemble* $\bigcup_{q\in\mathrm{E}_p} \mathrm{F}_q$ *est donc aussi fermé, on en déduit que* $x \in \bigcup_{q\in\mathrm{E}_p} \mathrm{F}_q \subset \mathrm{F}$ *et donc que* F *est fermé.*

Ceci montre bien que $\mathrm{A} \in \mathrm{T}$ *et termine la démonstration du fait que* T *est une tribu. Comme cela a déjà été dit, on en déduit que* $\mathrm{T} = \mathcal{B}(\mathbb{R}^{\mathrm{N}})$.

2. Soit $\mathrm{A} \in \mathcal{B}(\mathbb{R}^{\mathrm{N}})$. Déduire de la question précédente que $m(\mathrm{A}) = \inf\{m(\mathrm{O}), \mathrm{O}$ ouvert t.q. $\mathrm{A} \subset \mathrm{O}\}$.

Corrigé – *Par monotonie de* m *on a* $m(\mathrm{A}) \leq m(\mathrm{O})$ *si* $\mathrm{A} \subset \mathrm{O}$*, donc* $m(\mathrm{A}) \leq \inf\{m(\mathrm{O}), \mathrm{O}$ *ouvert t.q.* $\mathrm{A} \subset \mathrm{O}\}$*. Il reste donc à montrer que* $m(\mathrm{A}) \geq \inf\{m(\mathrm{O}), \mathrm{O}$ *ouvert t.q.* $\mathrm{A} \subset \mathrm{O}\}$.

Soit $\varepsilon > 0$*, d'après la question précédente, il existe* O *ouvert et* F *fermé t.q* $\mathrm{F} \subset \mathrm{A} \subset \mathrm{O}$ *et* $m(\mathrm{O} \setminus \mathrm{F}) \leq \varepsilon$*. On a donc aussi* $m(\mathrm{O} \setminus \mathrm{A}) \leq \varepsilon$ *et donc* $m(\mathrm{O}) = m(\mathrm{A}) + m(\mathrm{O} \setminus \mathrm{A}) \leq m(\mathrm{A}) + \varepsilon$*. Ceci montre bien que* $\inf\{m(\mathrm{O}), \mathrm{O}$ *ouvert t.q.* $\mathrm{A} \subset \mathrm{O}\} \leq m(\mathrm{A})$.

Exercice 7.13 (Fonction de Carathéodory)

Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$. On suppose que pour tout $s \in \mathbb{R}$ l'application $x \mapsto f(x,s)$ est borélienne (de $\mathbb{R}$ dans $\mathbb{R}$) et que pour tout $x \in \mathbb{R}$ l'application $s \mapsto f(x,s)$ est continue (de $\mathbb{R}$ dans $\mathbb{R}$). Pour tout $n \in \mathbb{N}^*$ on définit f_n de $\mathbb{R}^2$ dans $\mathbb{R}$ par $f_n(x,s) = f(x, \frac{i}{n})$ si $\frac{i}{n} \leq s < \frac{i+1}{n}$, $i \in \mathbb{Z}$.

1. Soit $n \in \mathbb{N}^*$. Soit $\mathrm{B} \in \mathcal{B}(\mathbb{R})$. Montrer que $f_n^{-1}(\mathrm{B}) \in \mathcal{B}(\mathbb{R}^2)$.

Corrigé – *Pour* $i \in \mathbb{Z}$*, on note* φ_i *la fonction (de* $\mathbb{R}$ *dans* $\mathbb{R}$*)* $x \mapsto f(x, i/n)$*. (Noter que* n *est fixé pour cette question.) Cette fonction* φ_i *est borélienne par hypothèse.*
On remarque maintenant que $f_n^{-1}(\mathrm{B}) = \cup_{i\in\mathbb{Z}} \varphi_i^{-1}(\mathrm{B}) \times [\frac{i}{n}, \frac{i+1}{n}[$.
En effet, soit $(x,s) \in f_n^{-1}(\mathrm{B})$*. Il existe* $i \in \mathbb{Z}$ *tel que* $s \in [i/n, (i+1)/n[$ *et donc* $f_n(x,s) = f(x, \frac{i}{n}) = \varphi_i(x)$*. On a donc* $x \in \varphi_i^{-1}(\mathrm{B})$ *et*

$$(x,s) \in \varphi_i^{-1}(\mathrm{B}) \times [\frac{i}{n}, \frac{i+1}{n}[\subset \bigcup_{j\in\mathbb{Z}} \varphi_j^{-1}(\mathrm{B}) \times [\frac{j}{n}, \frac{j+1}{n}[.$$

Réciroquement, si il existe $i \in \mathbb{Z}$ *tel que* $(x,s) \in \varphi_i^{-1}(\mathrm{B}) \times [\frac{i}{n}, \frac{i+1}{n}[$*, on a alors* $f_n(x,s) = f(x, i/n) = \varphi_i(x) \in \mathrm{B}$ *et donc* $(x,s) \in f_n^{-1}(\mathrm{B})$.
On a donc bien montré que $f_n^{-1}(\mathrm{B}) = \cup_{i\in\mathbb{Z}} \varphi_i^{-1}(\mathrm{B}) \times [\frac{i}{n}, \frac{i+1}{n}[$.
Comme, pour tout $i \in \mathbb{Z}$*,* φ_i *est borélienne, on a* $\varphi_i^{-1}(\mathrm{B}) \in \mathcal{B}(\mathbb{R})$ *et donc* $\varphi_i^{-1}(\mathrm{B}) \times [\frac{i}{n}, \frac{i+1}{n}[\in \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \subset \mathcal{B}(\mathbb{R}^2)$*. La stabilité de* $\mathcal{B}(\mathbb{R}^2)$ *par union dénombrable donne alors* $f_n^{-1}(\mathrm{B}) \in \mathcal{B}(\mathbb{R}^2)$.

2. Montrer que, pour tout $x, s \in \mathbb{R}$, $f_n(x,s) \to f(x,s)$ quand $n \to +\infty$. En déduire que f est une fonction borélienne de $\mathbb{R}^2$ dans $\mathbb{R}$.

Corrigé – *Soient $x, s \in \mathbb{R}$. Pour tout $n \in \mathbb{N}^*$, il existe i_n tel que $i_n/n \le s < (i_n+1)/n$ et on a $f_n(x) = f(x, i_n/n)$ et donc $f_n(x,s) - f(x,s) = f(s, i_n/n) - f(x)$.*

Comme $|i_n/n - s| \le 1/n$, on a $\lim_{n\to+\infty} i_n/n = s$ et donc, gràce à la continuité de $f(x,\cdot)$, $\lim_{n\to+\infty} f_n(x,s) = f(x,s)$.

La suite $(f_n)_{n\in\mathbb{N}^*}$ converge donc simplement vers f. La 1ere question montre que f_n est borélienne pour tout $n \in \mathbb{N}^*$, on en déduit que f est aussi borélienne (voir la proposition 3.19 sur la stabilité de l'ensemble des fonctions mesurables).

Exercice 7.15 (Densité de C_c dans $L^1(\mathbb{R}^N)$ pour la mesure de Lebesgue) Montrer que l'espace $C_c(\mathbb{R}^N, \mathbb{R})$ ($N \geq 1$) est dense dans $L^1(\mathbb{R}^N)$ (c'est-à-dire que, pour tout $f \in L^1(\mathbb{R}^N)$ et tout $\varepsilon > 0$, il existe $\varphi \in C_c(\mathbb{R}^N, \mathbb{R})$ t.q. $\|f - \varphi\|_1 \leq \varepsilon$). [S'inspirer de la démonstration faite pour le cas $N = 1$, théorème 5.20.]

Corrigé – *On reprend la démonstration du théorème 5.20, les modifications à apporter sont mineures. Le point essentiel est la régularité de λ_N (proposition 7.17).*

Etape 1. *On suppose ici que $f = 1_A$ avec $A \in \mathcal{B}(\mathbb{R}^N)$ et $\lambda_N(A) < +\infty$.*

Soit $\varepsilon > 0$. Comme λ_N est une mesure régulière (proposition 7.17), il existe un ouvert O et un fermé F tels que $F \subset A \subset O$ et $\lambda_N(O \setminus F) \leq \varepsilon$. Pour $n \in \mathbb{N}^*$, on pose $F_n = F \cap B_n$, ou B_n est la boule fermée de centre 0 et de rayon n. Comme dans le cas $N = 1$, Il existe n_0 tel que $\lambda_N(F \setminus F_{n_0}) \leq \varepsilon$. On pose $K = F_{n_0}$ et on obtient donc $K \subset F \subset A \subset O$, ce qui donne

$$\lambda_N(O \setminus K) \leq \lambda_N(O \setminus F) + \lambda_N(F \setminus K) \leq 2\varepsilon.$$

On a donc trouvé un compact K et un ouvert O tels que $K \subset A \subset O$ et $\lambda_N(O \setminus K) \leq 2\varepsilon$. Ceci va nous permettre de construire $\varphi \in C_c(\mathbb{R}^N, \mathbb{R})$ telle que $\|f - \varphi\|_1 \leq 2\varepsilon$.

On pose

$$d = d(K, O^c) = \inf\{d(x, y), x \in K, y \in O^c\}.$$

On montre, comme dans le cas $N = 1$ que $d > 0$ et on définit alors la fonction φ par

$$\forall x \in \mathbb{R}^N, \varphi(x) = \frac{1}{d}(d - d(x, K))^+ \text{ avec } d(x, K) = \inf\{d(x, y), y \in K\}.$$

La fonction φ est continue car $x \mapsto d(x, K)$ est continue (et même lipschitzienne car $|d(x, K) - d(y, K)| \leq |x - y|$). Elle est à support compact car il existe $A > 0$ tel que $K \subset B_A$ et on remarque alors que $\varphi = 0$ sur B^c_{A+d}. On a donc $\varphi \in C_c(\mathbb{R}^N, \mathbb{R})$. Enfin, on remarque que $\varphi = 1$ sur K, $\varphi = 0$ sur O^c et $0 \leq \varphi \leq 1$ (partout). On en déduit que $f - \varphi = 0$ sur $K \cup O^c$ et $0 \leq |f - \varphi| \leq 1$, ce qui donne

$$\|f - \varphi\|_1 \leq \lambda_N(O \setminus K) \leq 2\varepsilon,$$

et termine donc la première (et principale) étape.

Les étapes suivantes (étapes 2, 3 et 4) sont identiques à celles du cas $N = 1$ en remplaçant $C_c(\mathbb{R}, \mathbb{R})$ par $C_c(\mathbb{R}^N, \mathbb{R})$ et λ par λ_N.

Exercice 7.17 (Invariance par translation de λ_N) Soient $N \geq 1$, $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^*$ et $\beta_1, \ldots, \beta_N \in \mathbb{R}$. Pour $x = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$, on pose $\varphi(x) = (\alpha_1 x_1 + \beta_1, \ldots, \alpha_N x_N + \beta_N)^t$, de sorte que φ est une bijection de $\mathbb{R}^N$ dans $\mathbb{R}^N$.

1. Soit $A \in \mathcal{B}(\mathbb{R}^N)$, montrer que $\varphi(A) = \{\varphi(x), x \in A\} \in \mathcal{B}(\mathbb{R}^N)$.

   Corrigé – *On pose* $T = \{A \in \mathcal{B}(\mathbb{R}^N)$ *t.q.* $\varphi(A) \in \mathcal{B}(\mathbb{R}^N)\}$.

   Comme φ est bijective, il est facile de montrer que T est une tribu. En effet, il suffit de remarquer que $\varphi(\mathbb{R}^N) = \mathbb{R}^N$, $\varphi(A^c) = (\varphi(A))^c$ et $\varphi(\bigcup_{n\in\mathbb{N}} A_n) = \bigcup_{n\in\mathbb{N}} \varphi(A_n)$.

   Comme φ est continue, φ transforme les compacts en compacts. On note $\mathcal{C}$ l'ensemble des compacts de $\mathbb{R}^N$, on a donc $\mathcal{C} \subset T$ (on rappelle que les compacts sont des boréliens). Comme l'ensemble des compacts de $\mathbb{R}^N$ engendre $\mathcal{B}(\mathbb{R}^N)$ (noter que tout ouvert peut s'écrire comme une réunion dénombrable de compacts), on a donc $T = \mathcal{B}(\mathbb{R}^N)$, ce qui donne bien $\varphi(A) \in \mathcal{B}(\mathbb{R}^N)$ pour tout $A \in \mathcal{B}(\mathbb{R}^N)$.

2. Montrer que $\lambda_N(\varphi(A)) = \prod_{i=1}^N |\alpha_i| \lambda_N(A)$ pour tout $A \in \mathcal{B}(\mathbb{R}^N)$. [On pourra faire une récurrence sur N : la proposition 2.48 donne le résultat pour la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$, notée λ. On suppose que le résultat est vrai pour λ_{N-1} (et pour toute famille $\alpha_1, \ldots, \alpha_{N-1} \in \mathbb{R}^*$, $\beta_1, \ldots, \beta_{N-1} \in \mathbb{R}$). On le démontre alors pour λ_N en posant $m(A) = (\prod_{i=1}^N |\alpha_i|)^{-1} \lambda_N(\varphi(A))$ et en montrant que m est une mesure sur $\mathcal{B}(\mathbb{R}^N)$ égale à λ_N sur $\mathcal{B}(\mathbb{R}^{N-1}) \otimes \mathcal{B}(\mathbb{R})$. On utilise pour conclure la partie unicité du théorème 7.3 sur la mesure produit.]

   Corrigé – *On procède par récurrence sur N. La proposition 2.48 donne le résultat pour la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$, c'est-à-dire pour $N = 1$ en posant $\lambda_1 = \lambda$. On suppose maintenant que le résultat est vrai pour $N-1$ avec un certain $N \geq 2$, c'est-à-dire que $\lambda_{N-1}(\psi(B)) = \prod_{i=1}^{N-1} |\alpha_i| \lambda_{N-1}(B)$ pour tout $B \in \mathcal{B}(\mathbb{R}^{N-1})$, $\alpha_1, \ldots, \alpha_{N-1} \in \mathbb{R}^*$, $\beta_1, \ldots, \beta_{N-1} \in \mathbb{R}$ et ψ définie par $\psi(y) = (\alpha_1 y_1 + \beta_1, \ldots, \alpha_{N-1} y_{N-1} + \beta_{N-1})^t$ pour tout $y = (y_1, \ldots, y_{N-1})^t \in \mathbb{R}^{N-1}$, et on démontre le résultat pour N.*

   Soit donc $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^*$, $\beta_1, \ldots, \beta_N \in \mathbb{R}$ et φ définie par

   $$\varphi(x) = (\alpha_1 x_1 + \beta_1, \ldots, \alpha_N x_N + \beta_N)^t \text{ pour tout } x = (x_1, \ldots, x_N)^t \in \mathbb{R}^N.$$

   Pour $A \in \mathcal{B}(\mathbb{R}^N)$, on pose $m(A) = (\prod_{i=1}^N |\alpha_i|)^{-1} \lambda_N(\varphi(A))$. On montre tout d'abord que m est une mesure. On a bien $m(\emptyset) = 0$ car $\varphi(\emptyset) = \emptyset$. Puis, soit $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^N)$ avec $A_n \cap A_m = \emptyset$ si $n \neq m$. On a

   $$\varphi(\bigcup_{n\in\mathbb{N}} A_n) = \bigcup_{n\in\mathbb{N}} \varphi(A_n) \text{ avec } \varphi(A_n) \cap \varphi(A_m) = \emptyset \text{ si } n \neq m \text{ (car } \varphi \text{ est bijective).}$$

   Donc, $\lambda_N(\varphi(\bigcup_{n\in\mathbb{N}} A_n)) = \sum_{n\in\mathbb{N}} \lambda_N(\varphi(A_n))$ et donc $m(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} m(A_n)$. Ce qui prouve la σ-additivité de m et donc le fait que m est une mesure sur $\mathcal{B}(\mathbb{R}^N)$.

   On montre maintenant que $m(A_1 \times A_2) = \lambda_{N-1}(A_1)\lambda(A_2)$ pour tout $A_1 \in \mathcal{B}(\mathbb{R}^{N-1})$ et pour tout $A_2 \in \mathcal{B}(\mathbb{R})$ t.q. $\lambda_{N-1}(A_1) < \infty$ et $\lambda(A_2) < \infty$. Soit donc $A_1 \in \mathcal{B}(\mathbb{R}^{N-1})$ et $A_2 \in \mathcal{B}(\mathbb{R})$ t.q. $\lambda_{N-1}(A_1) < \infty$ et $\lambda(A_2) < \infty$. On a

   $$m(A_1 \times A_2) = (\prod_{i=1}^N |\alpha_i|)^{-1} \lambda_N(\varphi(A_1 \times A_2)).$$

On pose $\psi(y) = (\alpha_1 y_1 + \beta_1, \ldots, \alpha_{N-1} y_{N-1} + \beta_{N-1})^t$*, pour tout* $y = (y_1, \ldots, y_{N-1})^t \in \mathbb{R}^{N-1}$*, et* $\tau(z) = \alpha_N z + \beta_N$*, pour tout* $z \in \mathbb{R}$*. On a donc* $\varphi(A_1 \times A_2) = \psi(A_1) \times \tau(A_2)$*. L'hypothèse de récurrence et la proposition 2.48 donne que* $\lambda_{N-1}(\psi(A_1)) = \prod_{i=1}^{N-1} |\alpha_i| \lambda_{N-1}(A_1) < \infty$ *et que* $\lambda(\tau(A_2)) = \alpha_N \lambda(A_2) < \infty$*. Comme* $\lambda_N = \lambda_{N-1} \otimes \lambda$ *(car c'est la définition de* λ_N*) on en déduit :*

$$\lambda_N(\varphi(A_1 \times A_2)) = \lambda_{N-1}(\psi(A_1))\lambda(\tau(A_2)) = \prod_{i=1}^{N-1} |\alpha_i| \lambda_{N-1}(A_1) \alpha_N \lambda(A_2),$$

et donc :

$$m(A_1 \times A_2) = (\prod_{i=1}^{N} |\alpha_i|)^{-1} \lambda_N(\varphi(A_1 \times A_2)) = \lambda_{N-1}(A_1)\lambda(A_2).$$

On peut maintenant conclure. Comme m est une mesure sur $\mathcal{B}(\mathbb{R}^N) = \mathcal{B}(\mathbb{R}^{N-1}) \otimes \mathcal{B}(\mathbb{R})$*vérifiant* $m(A_1 \times A_2) = \lambda_{N-1}(A_1)\lambda(A_2)$ *pour tout* $A_1 \in \mathcal{B}(\mathbb{R}^{N-1})$ *et pour tout* $A_2 \in \mathcal{B}(\mathbb{R})$ *t.q.* $\lambda_{N-1}(A_1) < \infty$ *et* $\lambda(A_2) < \infty$*, la partie unicité du théorème 7.3 donne que* $m = \lambda_{N-1} \otimes \lambda$*, c'est-à-dire* $m = \lambda_N$*. On a donc bien* $\lambda_N(\varphi(A)) = \prod_{i=1}^{N} |\alpha_i| \lambda_N(A)$ *pour tout* $A \in \mathcal{B}(\mathbb{R}^N)$*.*

Exercice 7.18 (Changement de variable simple) Soient $N \geq 1$, $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^*$ et $\beta_1, \ldots, \beta_N \in \mathbb{R}$. Pour $x = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$. On pose $\varphi(x) = (\alpha_1 x_1 + \beta_1, \ldots, \alpha_N x_N + \beta_N)^t$ (de sorte que φ est une bijection de $\mathbb{R}^N$ dans $\mathbb{R}^N$).

1. Soit $f \in \mathcal{E}_+ = \mathcal{E}_+(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, montrer que $f \circ \varphi \in \mathcal{E}_+$ et que

$$\int f d\lambda_N = \prod_{i=1}^{N} |\alpha_i| \int (f \circ \varphi) d\lambda_N.$$

[Utiliser l'exercice 7.17.]

Corrigé – *Soit* $A \in \mathcal{B}(\mathbb{R}^N)$ *et* $f = 1_A$*. On a alors* $f \circ \varphi = 1_B$ *avec* $B = \varphi^{-1}(A)$*. En appliquant l'exercice 7.17à l'inverse de* φ*, noté* ψ*, on a donc* $\lambda_N(B) = \lambda_N(\psi(A)) = (\prod_{i=1}^{N} |\alpha_i|)^{-1} \lambda_N(A)$*, c'est-à-dire :*

$$\int f d\lambda_N = \lambda_N(A) = (\prod_{i=1}^{N} |\alpha_i|)\lambda_N(B) = (\prod_{i=1}^{N} |\alpha_i|) \int f \circ \varphi d\lambda_N.$$

Soit maintenant $f \in \mathcal{E}_+ \setminus \{0\}$*. Il existe donc* $a_1, \ldots, a_n \in \mathbb{R}_+^*$ *et* $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}^N)$ *t.q.* $f = \sum_{i=1}^{n} a_i f_i$*, avec* $f_i = 1_{A_i}$*. On a alors, par linéarité de l'intégrale :*

$$\int f d\lambda_N = \sum_{i=1}^{n} a_i \int f_i d\lambda_N = (\prod_{i=1}^{N} |\alpha_i|) \sum_{i=1}^{n} a_i \int f_i \circ \varphi d\lambda_N$$
$$= (\prod_{i=1}^{N} |\alpha_i|) \int \sum_{i=1}^{n} a_i f_i \circ \varphi d\lambda_N = (\prod_{i=1}^{N} |\alpha_i|) \int f \circ \varphi d\lambda_N.$$

(Ce qui est, bien sûr, aussi vrai si $f = 0$*.)*

2. Soit $f \in \mathcal{M}_+ = \mathcal{M}_+(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, montrer que $f \circ \varphi \in \mathcal{M}_+$ et que $\int f d\lambda_N = \prod_{i=1}^N |\alpha_i| \int (f \circ \varphi) d\lambda_N$.

   Corrigé – *Il existe $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}_+$ t.q. $f_n \uparrow f$ quand $n \to +\infty$. On a donc aussi $(f_n \circ \varphi)_{n\in\mathbb{N}} \subset \mathcal{E}_+$ et $f_n \circ \varphi \uparrow f \circ \varphi$ quand $n \to +\infty$ (ce qui donne, en particulier que $f \circ \varphi \in \mathcal{M}_+$). La question précédente donne :*

   $$\int f_n d\lambda_N = (\prod_{i=1}^N |\alpha_i|) \int f_n \circ \varphi d\lambda_N.$$

   La définition de l'intégrale sur $\mathcal{M}_+$ donne alors, quand $n \to +\infty$:

   $$\int f d\lambda_N = (\prod_{i=1}^N |\alpha_i|) \int f \circ \varphi d\lambda_N.$$

3. Soit $f \in \mathcal{L}^1 = \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$, montrer que $f \circ \varphi \in \mathcal{L}^1$ et que $\int f d\lambda_N = \prod_{i=1}^N |\alpha_i| \int (f \circ \varphi) d\lambda_N$.

   Corrigé – *Comme f est mesurable de $\mathbb{R}^N$ dans $\mathbb{R}$ et φ est mesurable de $\mathbb{R}^N$ dans $\mathbb{R}^N$ ($\mathbb{R}^N$ et $\mathbb{R}$ étant munis de leur tribu borélienne), on a bien $f \circ \varphi$ mesurable de $\mathbb{R}^N$ dans $\mathbb{R}$.*

   En appliquant la question précédente à la fonction $|f|$ on obtient que $\int |f| \circ \varphi d\lambda_N = \int |f \circ \varphi| d\lambda_N < \infty$ car $\int |f| d\lambda_N < \infty$. On a donc $f \circ \varphi \in \mathcal{L}^1$.

   Enfin, en remarquant que $(f \circ \varphi)^+ = f^+ \circ \varphi$ et $(f \circ \varphi)^- = f^- \circ \varphi$ et en utilisant la question précédente avec f^+ et f^-, on obtient :

   $$\int f^+ d\lambda_N = (\prod_{i=1}^N |\alpha_i|) \int f^+ \circ \varphi d\lambda_N,$$
   $$\int f^- d\lambda_N = (\prod_{i=1}^N |\alpha_i|) \int f^- \circ \varphi d\lambda_N.$$

   En faisant la différence, on en déduit :

   $$\int f d\lambda_N = (\prod_{i=1}^N |\alpha_i|) \int f \circ \varphi d\lambda_N.$$

Exercice 7.21 (Propriétés élémentaires de la convolution) Soit $f, g \in \mathrm{L}^1(\mathbb{R}^\mathrm{N}) = \mathrm{L}^1_\mathbb{R}(\mathbb{R}^\mathrm{N}, \mathcal{B}(\mathbb{R}^\mathrm{N}), \lambda_\mathrm{N})$.

1. Montrer que $f * g = g * f$ p.p.. [Utiliser l'invariance par translation de la mesure de Lebesgue et sa conséquence pour les changements de variable simples (propositions 7.20 et 7.21).]

   Corrigé – *On confond, comme d'habitude f (et g) avec l'un de ses représentants, et on choisit comme représentant de $f * g$ (qui est définie comme un élément de $\mathrm{L}^1(\mathbb{R}^\mathrm{N})$) la fonction définie par*
   $$f * g(x) = \int f(y) g(x-y) d\lambda_\mathrm{N}(y) \text{ si } f(\cdot)g(x-\cdot) \in \mathcal{L}^1(\mathbb{R}^\mathrm{N}).$$
   On sait que cette fonction est définie p.p. car $f(\cdot)g(x-\cdot) \in \mathcal{L}^1(\mathbb{R}^\mathrm{N})$ pour presque tout $x \in \mathbb{R}^\mathrm{N}$. On choisit de manière analogue comme représentant de $g * f$ la fonction définie par
   $$g * f(x) = \int g(y) f(x-y) d\lambda_\mathrm{N}(y) \text{ si } g(\cdot)f(x-\cdot) \in \mathcal{L}^1(\mathbb{R}^\mathrm{N}).$$
   Soit $x \in \mathbb{R}^\mathrm{N}$ un point pour lequel $f * g$ est définie. On a donc $f(\cdot)g(x-\cdot) \in \mathcal{L}^1(\mathbb{R}^\mathrm{N})$. On pose $h(\cdot) = f(\cdot)g(x-\cdot)$. On utilise alors la proposition 7.21 (changement de variable simple) avec φ définie par $\varphi(y) = -y + x$ pour tout $y \in \mathbb{R}^\mathrm{N}$. Elle donne $h \circ \varphi \in \mathcal{L}^1(\mathbb{R}^\mathrm{N})$ et :
   $$\int h(\varphi(y)) d\lambda_\mathrm{N}(y) = \int h(y) d\lambda_\mathrm{N}(y).$$
   Comme $h(\varphi(y)) = f(\varphi(y))g(x - \varphi(y)) = f(x-y)g(y)$, on en déduit que $g * f$ est définie au point x et que $g * f(x) = f * g(x)$. Ceci montre bien que $f * g = g * f$ p.p..

2. On suppose que f et g sont à support compact (f à support compact signifie qu'il existe K, compact de $\mathbb{R}^\mathrm{N}$, t.q. $f = 0$ p.p. sur K^c). Montrer que la fonction $f * g$ est alors aussi à support compact. [On désigne par $\mathrm{B}(0, \alpha)$ la boule ouverte de centre 0 et de rayon α. Comme f et g sont à support compact, il existe a et $b \in \mathbb{R}_+$ tels que $f = 0$ p.p. sur $\mathrm{B}(0, a)^c$ et $g = 0$ p.p. sur $\mathrm{B}(0, b)^c$. Montrer que $f * g = 0$ p.p. sur $\mathrm{B}(0, a+b)^c$.]

   Corrigé – *Comme pour la question précédente, on confond f (et g) avec l'un de ses représentants, et on choisit comme représentant de $f * g$ la fonction définie par $f * g(x) = \int f(y)g(x-y) d\lambda_\mathrm{N}(y)$ si $f(\cdot)g(x-\cdot) \in \mathcal{L}^1(\mathbb{R}^\mathrm{N})$.*

   Soit $a, b \in \mathbb{R}_+$ t.q. $f = 0$ p.p. sur $\mathrm{B}(0, a)^c$ et $g = 0$ p.p. sur $\mathrm{B}(0, b)^c$. ($\mathbb{R}^\mathrm{N}$ est muni d'une norme, notée $\|\cdot\|$.) Soit $x \in \mathrm{B}(0, a+b)^c$. On va montrer que $f(\cdot)g(x-\cdot) = 0$ p.p. (et donc que $f * g(x) = 0$, noter aussi que $f(\cdot)g(x-\cdot)$ est mesurable car f et g le sont).

   Comme $f = 0$ p.p. sur $\mathrm{B}(0, a)^c$, on a aussi $f(\cdot)g(x-\cdot) = 0$ p.p. sur $\mathrm{B}(0, a)^c$. Comme $g = 0$ p.p. sur $\mathrm{B}(0, b)^c$, on a $f(\cdot)g(x-\cdot) = 0$ p.p. sur $\mathrm{B}(x, b)^c$ (car $y \in \mathrm{B}(x, b)^c \iff (x-y) \in \mathrm{B}(0, b)^c$). On a donc :
   $$f(\cdot)g(x-\cdot) = 0 \text{ p.p. sur } \mathrm{B}(0, a)^c \cup \mathrm{B}(x, b)^c. \tag{7.21}$$

Or $B(0,a) \cap B(x,b) = \emptyset$ *car* $y \in B(0,a) \cap B(x,b)$ *implique* $\|y\| < a$ *et* $\|x-y\| < b$ *et donc* $\|x\| = \|x-y+y\| < a+b$, *ce qui contredit* $x \in B(0,a+b)^c$. *On a donc* $B(0,a)^c \cup B(x,b)^c = (B(0,a) \cap B(x,b))^c = \mathbb{R}^N$ *et (7.21) donne alors* $f(\cdot)g(x-\cdot) = 0$ *p.p. et donc* $f*g$ *est définie au point* x *et* $f*g(x) = 0$.

On a bien montré que $f*g = 0$ *p.p. sur* $B(0,a+b)^c$ *et donc que* $f*g$ *est à support compact.*

Exercice 7.22 (Convolution $L^p - C_c^\infty$)

Soit $1 \le p \le \infty$. Soit $f \in \mathcal{L}^p(\mathbb{R}^N) = \mathcal{L}^p_{\mathbb{R}}(\mathbb{R}^N, B(\mathbb{R}^N), \lambda_N)$ (ou $f \in \mathcal{L}^1_{\text{loc}}(\mathbb{R}^N)$) et $\rho \in C_c^\infty(\mathbb{R}^N, \mathbb{R})$. On pourra se limiter au cas $N = 1$.

1. Montrer que pour tout $x \in \mathbb{R}^N$, la fonction $f(\cdot)\rho(x-\cdot)$ appartient à $\mathcal{L}^1(\mathbb{R}^N)$. On pose alors

$$f * \rho(x) = \int f(\cdot)\rho(x-\cdot)d\lambda_N.$$

Corrigé – *On rappelle que* $f \in \mathcal{L}^1_{loc}(\mathbb{R}^N)$ *si* $f \in \mathcal{M} = \mathcal{M}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ *et* $f1_K \in \mathcal{L}^1(\mathbb{R}^N)$ *pour tout compact* K *de* $\mathbb{R}^N$. *On a donc* $\mathcal{L}^p(\mathbb{R}^N) \subset \mathcal{L}^1_{loc}(\mathbb{R}^N)$ *(pour tout* $1 \le p \le \infty$*) car si* $f \in \mathcal{L}^p(\mathbb{R}^N)$, *on a bien* $f \in \mathcal{M}$ *et, grâce à l'inégalité de Hölder,* $f1_K \in \mathcal{L}^1(\mathbb{R}^N)$ *(et* $\|f1_K\|_1 \le \|f\|_p \|1_K\|_q < \infty$,*avec* $q = p/(p-1)$*).*

On suppose donc dans la suite de ce corrigé que $f \in \mathcal{L}^1_{loc}(\mathbb{R}^N)$ *car cette hypothèse est plus générale que* $f \in \mathcal{L}^p(\mathbb{R}^N)$. *Pour simplifier la rédaction, on se limite au cas* $N = 1$.

Soit $x \in \mathbb{R}$. *La fonction* $f(\cdot)\rho(x-\cdot)$ *est mesurable (c'est-à-dire ici borélienne car* $\mathbb{R}$ *est muni de sa tribu de Borel) car* f *est mesurable et* $\rho(x-\cdot)$ *est mesurable (car continue). Comme* ρ *est à support compact, il existe* $a \in \mathbb{R}_+$ *t.q.* $\rho = 0$ *sur* $[-a,a]^c$. *On a donc* $\rho(x-\cdot) = 0$ *sur* K_x^c *avec* $K_x = [x-a, x+a]$, *ce qui permet de montrer que* $f(\cdot)\rho(x-\cdot) \in \mathcal{L}^1(\mathbb{R})$ *car* $|f(\cdot)\rho(x-\cdot)| \le |f1_{K_x}|\|\rho\|_u$, *avec* $\|\rho\|_u = \max\{|\rho(z)|, z \in \mathbb{R}\} < \infty$, *et* $f1_{K_x} \in \mathcal{L}^1(\mathbb{R})$.

La fonction $f*\rho$ *est donc définie sur tout* $\mathbb{R}$.

2. Montrer que $f * \rho \in C^\infty(\mathbb{R}^N, \mathbb{R})$.

Corrigé – *Comme dans la question précédente, on va utiliser* $a \in \mathbb{R}_+$ *t.q.* $\rho = 0$ *sur* $[-a,a]^c$ *et* $\|\rho\|_u = \max\{|\rho(z)|, z \in \mathbb{R}\}$ *(on va aussi utiliser les normes des dérivées de* ρ, $\|\rho^{(k)}\|_u$*). On raisonne maintenant en trois étapes.*

Etape 1. *On commence par montrer que* $f*\rho$ *est continue en tout point de* $\mathbb{R}$. *Soit* $x_0 \in \mathbb{R}$. *La continuité de* $f*\rho$ *en* x_0 *découle du théorème de continuité sous le signe* $\int$, *théorème 4.52. En effet, on pose, pour* $x, y \in \mathbb{R}$ *:*

$$F(x,y) = f(y)\rho(x-y).$$

On a $F(x,\cdot) \in \mathcal{L}^1(\mathbb{R})$ *pour tout* $x \in \mathbb{R}$ *et* $f*\rho(x) = \int F(x,\cdot)d\lambda$. *La fonction* F *vérifie alors les 2 hypothèses suivantes :*

(a) $x \mapsto F(x,y)$ *est continue en* x_0*, pour tout* $y \in \mathbb{R}$*,*

(b) $|F(x,y)| \le G(y)$ *pour tout* $x \in]x_0-1, x_0+1[$ *et pour tout* $y \in \mathbb{R}$*, en prenant* $G = |f1_K| \|\rho\|_u$ *avec* $K = [x_0-1-a, x_0+1+a]$*.*

Comme K *est compact, on a* $G \in \mathcal{L}^1(\mathbb{R})$ *et le théorème 4.52 donne bien la continuité de* $f * \rho$ *en* x_0*.*

Etape 2. *On montre maintenant que* $f * \rho$ *est dérivable en tout point et que* $(f*\rho)' = f*\rho'$*. Soit* $x_0 \in \mathbb{R}$*. Pour montrer la dérivabilité de* $f*\rho$ *en* x_0*, on utilise le théorème de dérivabilité sous le signe* $\int$*, théorème 4.53. On reprend la même fonction* F *que dans l'étape 1, elle vérifie les 2 hypothèses suivantes :*

(a) $x \mapsto F(x,y)$ *est dérivable pour tout* $x \in]x_0-1, x_0+1[$ *et pour tout* $y \in \mathbb{R}$*,*

(b) $|\frac{\partial F}{\partial x}(x,y)| = |f(y)\rho'(x-y)| \le H(y)$ *pour tout* $x \in]x_0-1, x_0+1[$ *et pour tout* $y \in \mathbb{R}$*, en prenant* $H = |f1_K| \|\rho'\|_u$ *avec* $K = [x_0-1-a, x_0+1+a]$*.*

Comme K *est compact, on a* $H \in \mathcal{L}^1(\mathbb{R})$ *et le théorème 4.53 donne bien la dérivabilité de* $f*\rho$ *en* x_0 *et le fait que* $(f*\rho)'(x_0) = f*\rho'(x_0)$*.*

Etape 3. *On montre enfin que* $f*\rho \in C^\infty(\mathbb{R},\mathbb{R})$*. Pour cela, on va montrer, par récurrence sur k, que* $f*\rho \in C^k(\mathbb{R},\mathbb{R})$ *et* $(f*\rho)^{(k)} = f*\rho^{(k)}$ *pour tout* $k \in \mathbb{N}$ *(ce qui donne bien* $f*\rho \in C^\infty(\mathbb{R},\mathbb{R})$*).*

L'étape 1 montre que $f*\rho \in C^0(\mathbb{R},\mathbb{R})$ *(et on a bien* $(f*\rho)^{(0)} = f*\rho = f*\rho^{(0)}$*). On suppose maintenant que* $f*\rho \in C^k(\mathbb{R},\mathbb{R})$ *et* $(f*\rho)^{(k)} = f*\rho^{(k)}$ *pour un certain* $k \in \mathbb{N}$*. L'étape 2 appliquée à* $\rho^{(k)}$ *(qui appartient aussi à* $C_c^\infty(\mathbb{R},\mathbb{R})$*) au lieu de* ρ *donne alors que* $f*\rho^{(k)}$ *est dérivable et que sa dérivée est* $f*\rho^{(k+1)}$*. L'étape 1 appliquée à* $\rho^{(k+1)}$ *donne que* $f*\rho^{(k+1)} \in C(\mathbb{R},\mathbb{R})$*. On a donc bien finalement montré que* $f*\rho \in C^{k+1}(\mathbb{R},\mathbb{R})$ *et* $(f*\rho)^{(k+1)} = f*\rho^{(k+1)}$*, ce qui termine la récurrence.*

3. On suppose maintenant que f est à support compact, c'est-à-dire qu'il existe un compact de $\mathbb{R}$, noté K, t.q. $f = 0$ p.p. sur K^c, montrer que $f*\rho$ est aussi à support compact.

Corrigé – *Comme* f *et* ρ *sont à support compact, on démontre que* $f*\rho$ *est aussi à support compact, comme cela a été fait dans l'exercice 7.21. Plus précisément, si* $f = 0$ *p.p. sur* $B(0,a)^c$ *et* $\rho = 0$ *sur* $B(o,b)^c$*, on a* $f*\rho = 0$ *sur* $B(0,a+b)^c$ *(voir l'exercice 7.21).*

Exercice 7.23 (Inégalité de Young) Soient $1 < p < +\infty, f \in L^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $g \in L^p_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$. Montrer que $f*g$ est définie p.p., que $f*g \in L^p_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et que $\|f*g\|_p \le \|f\|_1 \|g\|_p$. [Écrire

$$\int \Big(\int |f(x-y)g(y)|dy\Big)^p dx = \int \Big(\int |f(x-y)|^{\frac{1}{q}} |f(x-y)|^{\frac{1}{p}} |g(y)| dy\Big)^p dx,$$

avec $q = p/(p-1)$. Appliquer l'inégalité de Hölder puis le théorème de Fubini-Tonelli.]

Corrigé – *Pour simplifier les notations, on ne traite ici que le cas* $\mathrm{N} = 1$. *On suppose aussi que* f *et* g *ne sont pas nulles p.p. (sinon, il est immédiat que* $f * g$ *est définie partout et* $f * g(x) = 0$ *pour tout* $x \in \mathbb{R}$*).*

On confond f *et* g *avec l'un de leurs représentants.*

Pour $x, y \in \mathbb{R}$*, on pose* $\mathrm{H}(x, y) = g(y) f(x - y)$*. La première partie de la démonstration de la proposition sur la convolution (proposition 7.22) montre que* H*, et donc* $|\mathrm{H}|$*, est* $\mathcal{B}(\mathbb{R}^2)$*-mesurable. On peut alors utiliser les deux premières conclusions du théorème de Fubini-Tonelli (théorème 7.7) pour affirmer que* $|g(\cdot) f(x - \cdot)| \in \mathcal{M}_+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *pour tout* $x \in \mathbb{R}$ *et que* $\varphi \in \mathcal{M}_+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *avec* φ *définie par* $\varphi(x) = \int |g(\cdot) f(x - \cdot)| d\lambda$ *pour tout* $x \in \mathbb{R}$*. Par composition de fonctions mesurables, on a donc aussi* $\varphi^p \in \mathcal{M}_+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*.*

Soit $x \in \mathbb{R}$*. Les fonctions* $|f(x - \cdot)|^{\frac{1}{q}}$ *et* $|f(x - \cdot)|^{\frac{1}{p}} |g(\cdot)|$ *sont aussi mesurables. On peut alors utiliser l'inégalité de Hölder avec* $q = p/(p - 1)$*. elle donne :*

$$
\begin{aligned}
(\varphi(x))^p &= \Big(\int |f(x - \cdot)|^{\frac{1}{q}} |f(x - \cdot)|^{\frac{1}{p}} |g(\cdot)| d\lambda\Big)^p \\
&\le \Big(\int |f(x - \cdot)| d\lambda\Big)^{\frac{p}{q}} \Big(\int |f(x - \cdot)| |g(\cdot)|^p d\lambda\Big) \qquad (7.22) \\
&\le \|f\|_1^{\frac{p}{q}} \Big(\int |f(x - \cdot)| |g(\cdot)|^p d\lambda\Big).
\end{aligned}
$$

Noter que (7.22) est vraie si $|f(x - \cdot)| |g(\cdot)|^p \in \mathcal{L}^1(\mathbb{R})$ *et si* $|f(x - \cdot)| |g(\cdot)|^p \notin \mathcal{L}^1(\mathbb{R})$*. Dans ce dernier cas on obtient seulement* $\varphi(x)^p \le \infty$*. On a aussi utilisé la proposition 7.21 pour dire que* $\int |f(x - \cdot)| d\lambda = \int |f| d\lambda = \|f\|_1$*.*

On peut maintenant utiliser le théorème de Fubini-Tonelli (théorème 7.7) avec les fonctions $|f|$ *et* $|g|^p$*. Il donne :*

$$
\begin{aligned}
\int \varphi(x)^p d\lambda(x) &\le \|f\|_1^{\frac{p}{q}} \int \Big(\int |f(x - \cdot)| |g(\cdot)|^p d\lambda\Big) d\lambda(x) \\
&\le \|f\|_1^{\frac{p}{q}} \int \Big(\int |f(x - y)| |g(y)|^p d\lambda(x)\Big) d\lambda(y) \qquad (7.23) \\
&\le \|f\|_1^{\frac{p}{q}} \|f\|_1 \|g\|_p^p = \|f\|_1^p \|g\|_p^p.
\end{aligned}
$$

Ceci donne que $\varphi \in \mathcal{L}^p(\mathbb{R})$ *et, en particulier que* $\varphi(x) < \infty$ *p.p., c'est-à-dire* $\lambda(\mathrm{A}) = 0$ *avec* $\mathrm{A} = \{\varphi = \infty\}$ *(noter que* $\mathrm{A} \in \mathcal{B}(\mathbb{R})$ *puisque* $\varphi \in \mathcal{M}_+$*). Pour tout* $x \in \mathrm{A}^c$*, on a donc* $g(\cdot) f(x - \cdot) \in \mathcal{L}^1(\mathbb{R})$ *et on peut définir* $g * f(x)$ *par* $g * f(x) = \int g(y) f(x - y) d\lambda(y)$*.*

La fonction $g * f$ *est donc définie p.p.. On remarque maintenant qu'elle est égale p.p. à une fonction mesurable. En effet, les fonctions* H^+ *et* H^- *sont* $\mathcal{B}(\mathbb{R}^2)$*-mesurables (d'après la première partie de la démonstration de la proposition 7.22, on rappelle que* $\mathrm{H}(x, y) = g(y) f(x - y)$*). Les deux premières conclusions du théorème 7.7) donnent alors* $(g(\cdot) f(x - \cdot)^{\pm} \in \mathcal{M}_+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *pour tout* $x \in \mathbb{R}$ *et que* $\varphi_1, \varphi_2 \in \mathcal{M}_+(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *avec* φ_1 *et* φ_2 *définies par, pour* $x \in \mathbb{R}$*,*

$$
\varphi_1(x) = \int (g(\cdot) f(x - \cdot))^+ d\lambda, \ \ \varphi_2(x) = \int (g(\cdot) f(x - \cdot))^- d\lambda.
$$

On pose alors $h = \varphi_1 - \varphi_2$ *sur* A^c *et* $h = 0$ *sur* A*. On a bien que* h *est mesurable (de* $\mathbb{R}$ *dans* $\mathbb{R}$*) et* $h = g * f$ *p.p..*

On remarque enfin que $h \in \mathcal{L}^p(\mathbb{R})$ car $|h| \le \varphi$ p.p. et $\varphi \in \mathcal{L}^p(\mathbb{R})$. On en déduit bien que $g * f \in \mathrm{L}^p(\mathbb{R})$ (avec la confusion habituelle entre $g * f$ et la classe de h dans $\mathrm{L}^p(\mathbb{R})$).

Le fait que $\|g * f\|_p \le \|f\|_1 \|g\|_p$ est conséquence immédiate de (7.23) puisque $g * f \le \varphi$ p.p..

Enfin, le raisonnement fait dans la première question de l'exercice (7.21) montre que, pour tout $x \in \mathbb{R}$, $f(\cdot)g(x-\cdot) \in \mathcal{L}^1(\mathbb{R}^N)$ si et seulement $f(x-\cdot)g(\cdot) \in \mathcal{L}^1(\mathbb{R}^N)$ et que ces deux fonctions ont alors même intégrale. Ceci permet de montrer que $f * g$ est aussi définie p.p. et que $f * g = g * f$ p.p.

Exercice 7.24 (Itérations de convolution) Soient $\mathrm{L}^1 = \mathrm{L}^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $f \in \mathrm{L}^1$ t.q. $f = 0$ p.p. sur $\mathbb{R}_-$. On pose $f^{*1} = f$ et pour $n > 1$, $f^{*n} = f^{*(n-1)} * f$.

Pour $\alpha \ge 0$, on pose $g(\alpha) = \displaystyle\int_0^{+\infty} e^{-\alpha t}|f(t)|dt$.

1(a) Montrer que f^{*n} est bien définie pour tout $n \in \mathbb{N}^*$, et que $f^{*n} = 0$ p.p. sur $\mathbb{R}_-$.

Corrigé – *On montre par récurrence que f^{*n} est bien définie pour tout $n \in \mathbb{N}^*$, appartient à L^1 et que $f^{*n} = 0$ p.p. sur $\mathbb{R}_-$.*
En effet, on a bien $f^{*1} = f \in \mathrm{L}^1$ et $f^{*1} = f = 0$ p.p. sur $\mathbb{R}_-$.
Puis on suppose (hypothèse de récurrence) que $f^{*n} \in \mathrm{L}^1$ et $f^{*n} = 0$ p.p. sur $\mathbb{R}_-$. On remarque alors que f^{*n+1} est bien définie comme étant le produit de convolution de deux éléments de L^1 et appartient à L^1 (proposition 7.22). Puis pour $x < 0$, on a $f^{*n}(x-\cdot)f(\cdot) = 0$ p.p. sur $\mathbb{R}_-$ (en remarquant en $x - y < 0$ pour $y > 0$) et donc $f^{*n+1}(x) = 0$. Ce qui termine la récurrence.

(b) Montrer, par récurrence sur n, que $\int_0^{+\infty} e^{-\alpha t}|f^{*n}(t)|dt \le (g(\alpha))^n$, pour tout $\alpha \ge 0$ et tout $n \ge 1$.

Corrigé – *Soit $\alpha \ge 0$. On montre par récurrence sur n que $\int_{\mathbb{R}} e^{-\alpha t}|f^{*n}(t)|dt \le (g(\alpha))^n$. Cette inégalité est vraie pour $n = 1$ puisque c'est la définition de g et que $f = 0$ p.p. sur $\mathbb{R}_-$.*
On suppose maintenant que $\int_{\mathbb{R}} e^{-\alpha t}|f^{*n}(t)|dt \le (g(\alpha))^n$. On a alors, en utilisant le thèorème de Fubini-Tonelli (théorème 7.7) puis le changement de variable $t - s = \tau$

$$\begin{aligned}
\int_{\mathbb{R}} e^{-\alpha t}|f^{*(n+1)}(t)|dt &= \int_{\mathbb{R}} e^{-\alpha t}\Big|\int_{\mathbb{R}} f^{*n}(t-s)f(s)ds\Big|dt \\
&\le \int_{\mathbb{R}} e^{-\alpha t}\Big(\int_{\mathbb{R}} |f^{*n}(t-s)||f(s)|ds\Big)dt \\
&= \int_{\mathbb{R}} e^{-\alpha s}|f(s)|\Big(\int_{\mathbb{R}} e^{-\alpha(t-s)}|f^{*n}(t-s)|dt\Big)ds
\end{aligned}$$
$$= \int_{\mathbb{R}} e^{-\alpha s}|f(s)|\Big(\int_{\mathbb{R}} e^{-\alpha\tau}|f^{*n}(\tau)|d\tau\Big)ds \le \int_{\mathbb{R}} e^{-\alpha s}|f(s)|(g(\alpha))^n ds \le (g(\alpha))^{n+1}.$$

(c) Montrer que $\int_0^x |f^{*n}(t)|dt \le e^{\alpha x}(g(\alpha))^n$, pour tout $\alpha \ge 0$ tout $n \ge 1$ et tout $x \ge 0$.

Corrigé – *If suffit ici de remarquer que $e^{-\alpha x} \le e^{-\alpha t}$ pour $0 < t < x$, on en déduit, pour tout n, avec la question précédente,*

$$e^{-\alpha x} \int_0^x |f^{*n}(t)| dt \le \int_0^x e^{-\alpha t} |f^{*n}(t)| dt \le \int_0^{+\infty} e^{-\alpha t} |f^{*n}(t)| dt \le g(\alpha)^n.$$

Ce qui donne bien l'inégalité recherchée.

2. Soit $h \in C_b(\mathbb{R}, \mathbb{R})$. Montrer que $h * f(x)$ est définie pour tout $x \in \mathbb{R}$ et que $h * f \in C_b(\mathbb{R}, \mathbb{R})$. Remarquer de même que $h * f^{*n} \in C_b(\mathbb{R}, \mathbb{R})$. On suppose maintenant que, $h = 0$ sur $\mathbb{R}_-$. Montrer que, pour tout $x \in \mathbb{R}$, $h * f^{*n}(x) \to 0$ quand $n \to +\infty$.

Corrigé – *On pose* $\mathrm{M} = \sup_{x \in \mathbb{R}} |h(x)|$. *Soit $n \in \mathbb{N}^*$. Pour tout $x \in \mathbb{R}$, la fonction $y \mapsto h(x-y) f^{*n}(y)$ est intégrable (car elle est mesurable de dominée par* $\mathrm{M}|f^{*n}|$ *qui est intégrable). La fonction $h * f^{*n}$ est donc définie pour tout $x \in \mathbb{R}$. Cette fonction est bornée car $|h * f^{*n}(x)| \le \mathrm{M} \|f^{*n}\|_1 < +\infty$. Enfin, elle est continue d'après le théorème de continuité sous le signe intégrale, théorème 4.52.*

On suppose maintenant que, $h = 0$ sur $\mathbb{R}_-$. Comme $f^{*n} = 0$ p.p. sur $\mathbb{R}_-$, $h * f^{*n}(x) = 0$ pour tout $x \le 0$ et tout $n \in \mathbb{N}^*$.

Soit maintenant $x > 0$. On a, avec la question précédente, pour tout $n \in \mathbb{N}^*$ et $\alpha \ge 0$,

$$|h * f^{*n}(x)| = |\int_0^x h(x-t) f^{*n}(t) dt| \le \mathrm{M} \int_0^x |f^{*n}(t)| dt \le \mathrm{M} e^{\alpha x} g(\alpha)^n. \qquad (7.24)$$

Il suffit maintenant de remarquer que $\lim_{\alpha \to +\infty} g(\alpha) = 0$ *(c'est une conséquence du théorème de convergence dominée car* $\lim_{\alpha \to +\infty} e^{-\alpha t} = 0$ *et $|e^{-\alpha t} f(t)| \le |f((t)|$ pour tout $t > 0$). On peut donc choisir $\alpha \ge 0$ tel que $g(\alpha) < 1$. On déduit alors de l'inégalité* (7.24) $\lim_{n \to +\infty} h * f^{*n}(x) = 0$.

Exercice 7.25 Soient μ et ν deux mesures finies sur $\mathcal{B}(\mathbb{R})$. Pour $\mathrm{A} \in \mathcal{B}(\mathbb{R})$, on pose

$$\mu * \nu(\mathrm{A}) = \int_{\mathbb{R}^2} 1_{\mathrm{A}}(x+y) \, d(\mu \otimes \nu)(x, y).$$

1. Montrer que $\mu * \nu$ est une mesure finie sur $\mathcal{B}(\mathbb{R})$.

Corrigé – *Selon le théorème 7.3, $\mu \otimes \nu$ est mesure finie sur les boréliens de $\mathbb{R}^2$ et* $\mu \otimes \nu(\mathbb{R}^2) = \mu(\mathbb{R}) \nu(\mathbb{R})$ *(cf. égalité* (7.1) *avec* $\mathrm{A}_1 = \mathrm{A}_2 = \mathbb{R}$*)*

On remarque maintenant que si φ est une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$ (c'est-à-dire mesurable quand $\mathbb{R}$ est muni de la tribu de Borel) la fonction $(x, y) \mapsto \varphi(x, y)$ est borélienne de $\mathbb{R}^2$ dans $\mathbb{R}$ car c'est la composée de l'application continue (donc borélienne) $(x, y) \mapsto x + y$ (de $\mathbb{R}^2$ dans $\mathbb{R}$) avec φ. Si φ est de plus positive ou bornée $\int_{\mathbb{R}^2} \varphi(x+y) d(\mu \otimes \nu)(x, y)$ est alors bien définie (et appartient à $\mathbb{R}$ ou éventuellement $\overline{\mathbb{R}}_+$ dans le cas φ positive non bornée).

Comme $\mu * \nu(\mathrm{A}) = \mu \otimes \nu(\mathbb{R}^2) < +\infty$, *il suffit de vérifier la σ-additivité de $\mu * \nu$ pour montrer que $\mu * \nu$ est une mesure. Soit $(\mathrm{A}_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{B}(\mathbb{R})$ disjoints deux à deux, et* $\mathrm{A} = \cup_{n \in \mathbb{N}} \mathrm{A}_n$. *Comme* $1_{\mathrm{A}} = \sum_{n \in \mathbb{N}} 1_{\mathrm{A}_n}$, *le téorème de convergence* *monotone (ou sa conséquence, corollaire 4.18) donne*

$$\mu * \nu(\mathrm{A}) = \int_{\mathbb{R}^2} \sum_{n\in\mathbb{N}} 1_{\mathrm{A}_n}(x+y)\, d(\mu\otimes\nu)(x,y) = \sum_{n\in\mathbb{N}} \int_{\mathbb{R}^2} 1_{\mathrm{A}_n}(x+y)\, d(\mu\otimes\nu)(x,y)$$
$$= \sum_{n\in\mathbb{N}} \mu * \nu(\mathrm{A}_n).$$

2. Montrer que

$$\int_{\mathbb{R}} \varphi(z)\, d(\mu*\nu)(z) = \int_{\mathbb{R}^2} \varphi(x+y)\, d(\mu\otimes\nu)(x,y), \tag{7.25}$$

pour toute fonction φ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne positive ou borélienne bornée.

Corrigé – *L'égalité* (7.25) *est vrai si* $\varphi = 1_{\mathrm{A}}$, $\mathrm{A} \in \mathcal{B}(\mathrm{R})$ *(par la définition de* $\mu*\nu$*). Par "linéarité positive" (proposition 4.10) pour les deux mesures* $\mu*\nu$ *et* $\mu\otimes\nu$ *elle est aussi vraie si* φ *est étagée positive, c'est-à-dire* $\varphi \in \mathcal{E}_+(\mathbb{R},\mathbb{R})$ *(en fait, on peut aussi remarquer sur* $(x,y) \mapsto \varphi(x+y)$ *est étagée positive de* $\mathbb{R}^2$ *dans* $\mathbb{R}$ *et utiliser la définition de l'intégrale sur* $\mathcal{E}_+$, $\mathbb{R}$ *et* $\mathbb{R}^2$ *sont toujours munis de leur tribu boréleinne).*
Soit maintenant φ *borélienne positive de* $\mathbb{R}$ *dans* $\mathbb{R}$*. La fonction* φ *est alors limite croissante d'une suite de fonctions étagées positives et on obtient* (7.25) *avec le théorème de convergence monotone pour les deux mesures* $\mu*\nu$ *et* $\mu\otimes\nu$*. (En fait, on peut aussi utiliser directement la définition de l'intégrale sur* $\mathcal{M}_+$*.)*
Enfin, si φ *est une fonction de* $\mathbb{R}$ *dans* $\mathbb{R}$ *borélienne bornée, elle est donc intégrable pour la mesure* $\mu*\nu$ *(car c'est une mesure finie). En utilisant* (7.25) *avec* φ^+ *et* φ^-*, on en déduit que la fonction* $(x,y) \mapsto \varphi(x,y)$ *est intégrable pour la mesure* $\mu\otimes\nu$ *et que* (7.25) *est vraie.*

3. Montrer que si μ et ν sont des probabilités, alors $\mu*\nu$ est une probabilité.

Corrigé – *Il suffit ici de remarquer que* $\mu*\nu(\mathbb{R}) = \mu\otimes\nu(\mathbb{R}^2) = \mu(\mathbb{R})\nu(\mathbb{R}) = 1$ *si* μ *et* ν *sont des probabilités.*

4. Montrer que si μ et ν sont des probabiltés de densités respectives f et g (par rapport à la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$), alors $\mu*\nu$ est la probabilité de densité $f*g$ par rapport à la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$.

Corrigé – *On reprend ici le preuve faite dans le paragraphe 7.5. Soit* φ *une fonction de* $\mathbb{R}$ *dans* $\mathbb{R}$ *borélienne bornée. En utilisant* $\mu = f\lambda$, $\nu = g\lambda$ *et le théorème de Fubini, on obtient*

$$\int_{\mathbb{R}} \varphi d(\mu*\nu) = \int_{\mathbb{R}^2} \varphi(x+y) d(\mu\otimes\nu) = \int_{\mathbb{R}} (\int_{\mathbb{R}} \varphi(x+y) f(x)dx) g(y)dy.$$

Le changement de variable $x+y = z$ *dans l'intégrale sur* x *puis le théorème de Fubini donnent*

$$\int_{\mathbb{R}} \varphi d(\mu*\nu) = \int_{\mathbb{R}} (\int_{\mathbb{R}} \varphi(z) f(z-y)dz) g(y)dy = \int_{\mathbb{R}} (\int_{\mathbb{R}} f(z-y)g(y)dy)\varphi(z)dz$$
$$= \int_{\mathbb{R}} f*g(z)\varphi(z)dz.$$

Ceci prouve bien que $\mu*\nu = (f*g)\lambda$.

Exercice 7.26 (Fonction de Carathéodory et composition)

Soit $N, p, q \in \mathbb{N}^\star$ et Ω un ouvert de $\mathbb{R}^N$. Soit a une application de $\Omega \times \mathbb{R}^p$ dans $\mathbb{R}^q$. On suppose que a est une fonction de Carathéodory, c'est-à-dire que $a(\cdot, s)$ est borélienne pour tout $s \in \mathbb{R}^p$ et $a(x, \cdot)$ est continue pour tout $x \in \Omega$.

1. Montrer que la fonction a est borélienne de $\Omega \times \mathbb{R}^p$ dans $\mathbb{R}^q$. (Noter que ceci peut être faux si a était seulement borélienne par rapport à chacun de ses arguments, un exemple est donné dans l'exercice 7.4.)

   Corrigé – *Pour $n \in \mathbb{N}^*$, on considère une partition dénombrable de $\mathbb{R}^p$ formée de boréliens de diamètre inférieur à $1/n$ (une telle partition est possible en utilisant, par exemple, un quadrillage de $\mathbb{R}^p$.*

   Pour $n \in \mathbb{N}^*$, on note $A_{k,n}$, $k \in \mathbb{N}$, les éléments de cette partition (avec $A_{k,n}$ non vide pour tout k). Pour tout $k \in \mathbb{N}$ on choisit un point $s_{k,n} \in A_{k,n}$ et on définit la fonction a_n de $\Omega \times \mathbb{R}^p$ dans $\mathbb{R}^q$ par

   $$a_n(x, s) = a(x, s_{k,n}) \text{ si } s \in A_{k,n}.$$

   Si $s \in A_{k,n}$, on a $|s - s_{k,n}| \le 1/n$ (car le diamètre de $A_{k,n}$ est inférieur à $1/n$). Comme a est continue par rapport à son deuxième argument, on a pour tout $(x, s) \in \Omega \times \mathbb{R}^p$, $lim_{n\to+\infty} a_n(x, s) = a(x, s)$.

   On remarque maintenant que a_n est (pour tout n) une fonction borélienne. En effet $a_n(x, s) = \sum_{k\in\mathbb{N}} a(x, s_{k,n}) 1_{A_{k,n}}(s)$ qui est bien une fonction borélienne comme limite d'une somme et de produits de fonctions boréliennes (voir la proposition 3.19). On a utilisé ici le fait que $a(\cdot, s)$ est borélienne, de Ω dans $\mathbb{R}^q$, pour tout $s \in \mathbb{R}^p$.

   Finalement, a est borélienne comme limite de fonctions boréliennes.

   Remarque : En pratique, on peut remplacer l'hypothèse "$a(x, \cdot)$ est continue pour tout $x \in \Omega$" par l'hypothèse "$a(x, \cdot)$ est continue pour presque tout $x \in \Omega$". Le raisonnement précédent permet alors de montrer qu'il existe $\bar{a}$ borélienne telle que $a = \bar{a}$ p.p., ce qui est suffisant du point de vue de la théorie de l'intégration.

2. Soit v est une fonction borélienne de Ω dans $\mathbb{R}^p$. Montrer que la fonction $x \mapsto a(x, v(x))$ est alors borélienne de Ω dans $\mathbb{R}^q$.

   Corrigé – *On note b la fonction de Ω dans $\Omega \times \mathbb{R}^p$ définie par $b(x) = (x, v(x))$. La fonction b est borélienne car ses deux composantes sont boréliennes (voir la proposition 4.63). On en déduit que la fonction $x \mapsto a(x, v(x))$ est borélienne car cette fonction est la composition de b avec a (c'est-à-dire la fonction $x \mapsto a(b(x))$) qui sont boréliennes.*

3. Soient v_1, v_2 deux fonctions de Ω dans $\mathbb{R}^p$. On suppose que $v_1 = v_2$ p.p.. Montrer que les fonctions $x \mapsto a(x, v_1(x))$ et $x \mapsto a(x, v_2(x))$ sont égales p.p. sur Ω.

   Corrigé – *Soit $A \in B(\mathbb{R})$ $(A \subset \Omega)$ tel que $\lambda(A) = 0$ et $v_1 = v_2$ sur $\Omega \setminus A$. On a alors $a(x, v_1(x)) = a(x, v_2(x))$ pour tout $x \in \Omega \setminus A$ et donc $a(x, v_1(x)) = a(x, v_2(x))$ p.p. sur Ω.*

Exercice 7.28 (Coordonnées polaires)

1. Calculer $\int_{(\mathbb{R}^+)^2} e^{-(x^2+y^2)}\,dx\,dy$ (on rappelle que $dx\,dy$ désigne $d\lambda_2(x,y)$).
   [On pourra utiliser le passage en coordonnées polaires.]

   Corrigé – *Pour $x, y \in \mathbb{R}$, on pose $f(x,y) = e^{-(x^2+y^2)}1_{\mathbb{R}_+}(x)1_{\mathbb{R}_+}(y)$. On a $f \in \mathcal{M}_+(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ (noter que f est le produit de fonctions mesurables). On peut donc lui appliquer la formule (7.10) :*

   $$\int f(x,y)d\lambda_2(x,y) = \int_0^{\frac{\pi}{2}}\left(\int_0^\infty f(r\cos\theta, r\sin\theta)r\,dr\right)d\theta.$$

   On a donc :

   $$\int_{(\mathbb{R}^+)^2} e^{-(x^2+y^2)}d\lambda_2(x,y) = \frac{\pi}{2}\int_0^\infty e^{-r^2}r\,dr.$$

   Puis, comme $\int_0^n e^{-r^2}r\,dr = \frac{1}{2}(1-e^{-n^2})$ et que, par le théorème de convergence monotone,

   $$\lim_{n\to+\infty}\int_0^n e^{-r^2}r\,dr = \int_0^\infty e^{-r^2}r\,dr,$$

   on en déduit $\int_0^\infty e^{-r^2}r\,dr = \frac{1}{2}$ et donc

   $$\int_{(\mathbb{R}^+)^2} e^{-(x^2+y^2)}d\lambda_2(x,y) = \frac{\pi}{4}.$$

2. Calculer $\displaystyle\int_{\mathbb{R}^+} e^{-x^2}dx$.

   Corrigé – *On applique le théorème de Fubini–Tonelli (théorème 7.7) à la fonction f définie à la question précédente. Il donne :*

   $$\int f(x,y)d\lambda_2(x,y) = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x,y)dy\right)dx,$$

   et donc

   $$\frac{\pi}{4} = \int_{(\mathbb{R}^+)^2} e^{-(x^2+y^2)}d\lambda_2(x,y) = \int_0^\infty\left(\int_0^\infty e^{-y^2}dy\right)e^{-x^2}dx = \left(\int_0^\infty e^{-x^2}dx\right)^2.$$

   On en déduit que $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$.

Chapitre 8

Densité, séparabilité et compacité

Ce chapitre est consacré en majeure partie aux espaces $L^p_{\mathbb{R}}(\Omega, \mathcal{B}(\Omega), \lambda_N)$ où Ω un ouvert de $\mathbb{R}^N$, $N \geq 1$, $\mathcal{B}(\Omega)$ est la tribu borélienne de Ω, λ_N désigne la restriction à $\mathcal{B}(\Omega)$ de la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^N)$ (aussi notée λ_N) et $1 \leq p \leq \infty$.

On notera toujours $L^p(\Omega) = L^p_{\mathbb{R}}(\Omega, \mathcal{B}(\Omega), \lambda_N)$.

8.1 Théorèmes de densité pour les espaces $L^p(\Omega)$

Nous avons vu au chapitre précédent que les espaces L^p sont très intéressants car ils sont en particulier complets. Cependant, les éléments de ces espaces sont des objets avec lesquels il est malaisé de travailler. Pour démontrer des propriétés sur ces objets, on travaille très souvent par densité : on travaille sur des fonctions régulières, qui sont faciles à manipuler, et on passe ensuite à la limite, à condition d'avoir établi au préalable un résultat de densité, qui nous permet justement ce passage à la limite.

8.1.1 Densité des fonctions $C_c(\Omega, \mathbb{R})$ dans $L^p(\Omega)$

Définition 8.1 (Fonction à support compact) *Soient* $N \geq 1$, Ω *un ouvert de* $\mathbb{R}^N$ *et* f *une fonction définie de* Ω *dans* $\mathbb{R}$. *On dit que* f *est à support compact (dans* Ω*) s'il existe un compact* $K \subset \Omega$ *tel que* $f = 0$ *sur* $\Omega \setminus K$.

On note souvent $\mathrm{supp}(f)$ l'adhérence dans Ω de l'ensemble des $x \in \Omega$ t.q. $f(x) \neq 0$. On peut montrer que f est à support compact si et seulement si $\mathrm{supp}(f)$ est compact.

Définition 8.2 ($C_c^\infty(\Omega,\mathbb{R})$) *Soient* $N \geq 1$, Ω *un ouvert de* $\mathbb{R}^N$ *et* f *une fonction définie de* Ω *dans* $\mathbb{R}$. *On dit que* $f \in C_c^\infty(\Omega,\mathbb{R})$ *si*

— f *est de classe* C^∞ *(de* Ω *dans* $\mathbb{R}$*).*

— f *est à support compact (dans* Ω*).*

On note aussi $\mathcal{D}(\Omega) = C_c^\infty(\Omega,\mathbb{R})$.

Remarque 8.3 Si $N = 1$ et $\Omega =]0,1[$, la fonction f définie par $f(x) = x(x-1)$ est de classe C^∞ sur Ω, mais elle n'est pas à support compact. En effet, il n'existe pas de compact inclus dans $]0,1[$ t.q. f soit nulle en dehors de ce compact.

Par contre, si f est de classe C^∞ sur $]0,1[$ et s'il existe $\varepsilon > 0$ tel que $f(x) = 0$ pour $x \in]0,\varepsilon[$ et pour $x \in]1-\varepsilon,1[$, alors $f \in C_c^\infty(\Omega,\mathbb{R})$.

Théorème 8.4 (Densité de $C_c(\Omega,\mathbb{R})$ dans $L^p(\Omega)$) *Soient* $N \geq 1$, $p \in [1,+\infty[$ *et* Ω *un ouvert de* $\mathbb{R}^N$ *(par exemple,* $\Omega = \mathbb{R}^N$*). Alors :*

$C_c(\Omega,\mathbb{R})$ *est dense dans* $L^p(\Omega)$ *c'est-à-dire :*

$$\forall f \in L^p(\Omega),\ \forall \varepsilon > 0,\ \exists \varphi \in C_c(\Omega,\mathbb{R}) \text{ t.q. } \|f-\varphi\|_p \leq \varepsilon.$$

DÉMONSTRATION – La démonstration de ce résultat est faite dans l'exercice 6.5 pour le cas $\Omega = \mathbb{R}$. La généralisation donnée ici se démontre de manière très voisine (grâce au résultat de régularité, proposition 7.17), voir l'exercice 8.1. ■

Une conséquence importante du théorème 8.4 est la continuité en moyenne que l'on donne maintenant.

Théorème 8.5 (Continuité en moyenne) *Soient* $N \geq 1$, $p \in [1,+\infty[$, *et* $f \in L^p(\mathbb{R}^N)$. *Alors,* $\|f(\cdot+h) - f\|_p \to 0$ *quand* $h \to 0$, *c'est-à-dire :*

$$\int |f(x+h) - f(x)|^p dx \to 0, \text{ quand } h \to 0.$$

La démonstration est ici encore très similaire à la démonstration vue pour $N = 1$ dans l'exercice 6.5, elle est proposée dans l'exercice 8.2.

Remarque 8.6 (Attention à L^∞ !) Les deux résultats précédents sont faux dans L^∞. Considérer par exemple le cas $N = 1$ et la fonction $f = 1_{\mathbb{R}_+}$ (qui appartient à $L^\infty(\mathbb{R})$, en confondant f avec sa classe). On peut montrer que (voir l'exercice 8.4) :

$$\forall \varphi \in C(\mathbb{R},\mathbb{R}),\ \|f-\varphi\|_\infty \geq \frac{1}{2},$$
$$\forall h>0,\ \|f(\cdot+h)-f\|_\infty = 1.$$

8.1.2 Régularisation par convolution

Si $a \in \mathbb{R}_+$, on note B_a la boule fermée de centre 0 et de rayon a de $\mathbb{R}^N$.

Définition 8.7 (Espace $L^1_{\mathbf{loc}}$ des fonctions localement intégrales) *Soient* $N \geq 1$ *et* Ω *un ouvert de* $\mathbb{R}^N$*. Soit* $f : \Omega \to \mathbb{R}$*. On dit que* f *est localement intégrable sur* Ω *si* $f 1_K \in L^1(\Omega)$ *(au sens "il existe* $g \in \mathcal{L}^1(\Omega)$ *t.q.* $f = g$ *p.p. sur* K*") pour tout compact* $K \in \Omega$.

On note $L^1_{loc}(\Omega)(= L^1_{loc}(\Omega,\mathcal{B}(\Omega),\lambda_N))$ *l'ensemble des fonctions localement intégrables sur* Ω.

Remarque 8.8 Soient $N \geq 1$ et Ω un ouvert de $\mathbb{R}^N$. Pour tout p tel que $1 \leq p \leq +\infty$, on a $L^p(\Omega) \subset L^1_{\mathrm{loc}}(\Omega)$ (ceci est une conséquence immédiate du résultat d'inclusion entre les espaces L^p, proposition 6.25).

Définition 8.9 (Famille régularisante) *Soit* $N \geq 1$ *et soit* $\rho \in C_c^\infty(\mathbb{R}^N,\mathbb{R})$ *t.q.* $\rho \geq 0$, $\{x \in \mathbb{R}^N\,;\, \rho(x) \neq 0\} \subset B_1$ *et* $\int \rho(x)dx = 1$*. On appelle famille régularisante (ou famille de noyaux régularisants) la famille de fonctions* $(\rho_n)_{n\in\mathbb{N}^*} \subset C_c^\infty(\mathbb{R}^N,\mathbb{R})$ *définie par :* $\rho_n(x) = n^N\rho(nx), x \in \mathbb{R}^N$, $n \in \mathbb{N}^*$.

Remarque 8.10 Dans la définition précédente, il est facile de vérifier que

$$\{x \in \mathbb{R}^N\ ;\ \rho_n(x) \neq 0\} \subset B_{\frac{1}{n}} \text{ et } \int \rho_n(x)dx = 1.$$

Notons qu'il existe bien des fonctions vérifiant les propriétés demandées pour ρ dans la définition 8.9. Pour $N = 1$, par exemple, il suffit de prendre

$$\rho(x) = \begin{cases} \alpha \exp(\frac{1}{x^2-1}) \text{ si } x \in]-1,1[, \\ 0 \text{ si } x \notin]-1,1[, \end{cases}$$

avec $\alpha > 0$ choisi pour avoir $\int \rho(x)dx = 1$.

Lemme 8.11 (Régularisation par convolution) *Soient* $N \geq 1$, $(\rho_n)_{n\in\mathbb{N}^*}$ *une famille régularisante et* $f \in L^1_{loc}(\mathbb{R}^N)$*. Alors,* $f * \rho_n$ *est définie partout sur* $\mathbb{R}^N$ *et* $f * \rho_n \in C^\infty(\mathbb{R}^N,\mathbb{R})$*. De plus, s'il existe* $a > 0$ *t.q.* $f = 0$ *p.p. sur* B_a^c*, on a alors* $f * \rho_n = 0$ *sur* $B^c_{a+\frac{1}{n}}$ *(*$f * \rho_n$ *est donc à support compact).*

DÉMONSTRATION – La démonstration du fait que $f * \rho_n \in C^\infty(\mathbb{R}^N,\mathbb{R})$ est donnée dans l'exercice 7.22. La seconde partie est donnée dans les exercices 7.22 et 7.21. L'indication de la seconde question de l'exercice 7.21 donne le support indiqué ici pour $f * \rho_n$. ■

Proposition 8.12 (Densité de $C^\infty(\mathbb{R}^N,\mathbb{R})$ dans $L^p(\mathbb{R}^N)$) *Soit $N \geq 1$, $(\rho_n)_{n\in\mathbb{N}^*}$ une famille régularisante. Soient $p \in [1,+\infty[$ et $f \in L^p(\mathbb{R}^N)$. Alors,*

$$f * \rho_n \to f \text{ dans } L^p(\mathbb{R}^N) \text{ lorsque } n \to +\infty.$$

DÉMONSTRATION – La démonstration est une conséquence du théorème de continuité en moyenne (théorème 8.5). On choisit un représentant de f, encore noté f. Pour $n \in \mathbb{N}^*$, on pose $f_n = f * \rho_n$. Pour $x \in \mathbb{R}^N$, on a :

$$\begin{aligned} f_n(x) - f(x) &= \int f(x-y)\rho_n(y)dy - f(x)\int \rho_n(y))dy \\ &= \int (f(x-y) - f(x))\rho_n(y)dy, \end{aligned}$$

et donc :

$$|f_n(x) - f(x)|^p \leq \left(\int |f(x-y) - f(x)|\rho_n(y)dy\right)^p.$$

Pour $p > 1$, on utilise l'inégalité de Höder en écrivant $\rho_n = \rho_n^{\frac{1}{p}}\rho_n^{\frac{1}{q}}$ (avec $q = p/(p-1)$) et on obtient (ce qui est aussi immédiatement vrai pour $p = 1$) :

$$\begin{aligned} |f_n(x) - f(x)|^p &\leq \int |f(x-y) - f(x)|^p \rho_n(y)dy\left(\int \rho_n(y)dy\right)^{\frac{p}{q}} && (8.1) \\ &\leq \int |f(x-y) - f(x)|^p \rho_n(y)dy. && (8.2) \end{aligned}$$

On remarque maintenant que l'application $(x,y)^t \mapsto (f(y) - f(x))$ est mesurable de $\mathbb{R}^2$ dans $\mathbb{R}$ (munis de leurs tribus boréliennes respectives) ; ceci est une conséquence (par exemple) de la proposition 7.11 et de la mesurabilité de la somme d'applications mesurables. L'application $(x,y)^t \mapsto (x, x-y)$ est mesurable de $\mathbb{R}^2$ dans $\mathbb{R}^2$ (car continue). Par composition, l'application $(x,y)^t \mapsto (f(x-y) - f(x))$ est donc mesurable de $\mathbb{R}^2$ dans $\mathbb{R}$. On en déduit, en utilisant une nouvelle fois la mesurabilité de la composée d'applications mesurables (et du produit d'applications mesurables), que $(x,y)^t \mapsto |f(x-y) - f(x)|^p\rho_n(y)$ est mesurable (positive) de $\mathbb{R}^2$ dans $\mathbb{R}$. On peut donc appliquer le théorème de Fubini-Tonelli pour déduire de (8.2) que :

$$\begin{aligned} \int |f_n(x) - f(x)|^p dx &\leq \int \left(\int |f(x-y) - f(x)|^p \rho_n(y)dy\right)dx \\ &= \int \left(\int |f(x-y) - f(x)|^p \rho_n(y)dx\right)dy && (8.3) \\ &= \int_{B_{\frac{1}{n}}} \|f(\cdot - y) - f\|_p^p \rho_n(y)dy. \end{aligned}$$

On utilise maintenant le théorème 8.5. Soit $\varepsilon > 0$. Il existe $\eta > 0$ t.q. :

$$h \in \mathbb{R}^N,\ |h| \le \eta\ \Rightarrow \|f(\cdot - h) - f\|_p \le \varepsilon.$$

On déduit donc de (8.3) que :

$$\frac{1}{n} \le \eta \Rightarrow \|f_n - f\|_p \le \varepsilon.$$

Ce qui termine la démonstration. ■

La proposition 8.12 nous dit que l'espace $C^\infty(\mathbb{R}^N, \mathbb{R})$ est dense dans l'espace $L^p(\mathbb{R}^N)$. Il est très souvent plus facile de travailler avec des fonctions à support compact. De fait, les deux résultats précédents permettent de démontrer le théorème de densité de $C_c^\infty(\mathbb{R}^N, \mathbb{R})$ dans $L^p(\mathbb{R}^N)$:

Théorème 8.13 (Densité de $C_c^\infty(\mathbb{R}^N, \mathbb{R})$ dans $L^p(\mathbb{R}^N)$) *Soient* $N \ge 1$ *et* $p \in [1, +\infty[$, $C_c^\infty(\mathbb{R}^N, \mathbb{R})$ *est dense dans* $L^p(\mathbb{R}^N)$.

DÉMONSTRATION – La démonstration proposée ici utilise une méthode dite de troncature et régularisation.

Pour $f \in L^p(\mathbb{R}^N)$, on dit que f est à support compact s'il existe K compact de $\mathbb{R}^N$ t.q. $f = 0$ p.p. sur K^c. On note A l'ensemble des fonctions $f \in L^p(\mathbb{R}^N)$ à support compact.

Etape 1. On montre dans cette étape que A est dense dans $L^p(\mathbb{R}^N)$. Soit $f \in L^p(\mathbb{R}^N)$. Pour $n \in \mathbb{N}$, on pose $f_n = f 1_{B_n}$. Comme $f_n \to f$ p.p. quand $n \to +\infty$ et $|f_n| \le |f|$ p.p. (pour tout $n \in \mathbb{N}$), on peut appliquer le théorème de convergence dominée dans L^p (on utilise ici le fait que $p < \infty$). Il donne que $f_n \to f$ dans $L^p(\mathbb{R}^N)$ quand $n \to +\infty$. Comme $(f_n)_{n\in\mathbb{N}} \subset A$, on a bien montré la densité de A dans $L^p(\mathbb{R}^N)$.

Etape 2. Soit maintenant $f \in A$. Pour conclure la démonstration, il suffit de montrer qu'il existe une suite $(f_n)_{n\in\mathbb{N}^\star} \subset C_c^\infty(\mathbb{R}^N, \mathbb{R})$ t.q. $f_n \to f$ dans $L^p(\mathbb{R}^N)$ quand $n \to +\infty$. Or cette suite est donné avec $f_n = f * \rho_n$ où $(\rho_n)_{n\in\mathbb{N}^\star}$ est une famille régularisante. En effet, le lemme 8.11 donne que $(f_n)_{n\in\mathbb{N}^\star} \subset C_c^\infty(\mathbb{R}^N, \mathbb{R})$ et la proposition 8.12 donne que $f_n \to f$ dans $L^p(\mathbb{R}^N)$ quand $n \to +\infty$. ■

On rappelle (remarque 8.6) que $C_c(\mathbb{R}^N, \mathbb{R})$ (et donc aussi $C_c^\infty(\mathbb{R}^N, \mathbb{R})$) n'est pas dense dans $L^\infty(\mathbb{R}^N)$. Il est intéressant aussi de remarquer que le théorème précédent (théorème 8.13) est encore vrai si on remplace la mesure de Lebesgue par une mesure sur les boréliens de $\mathbb{R}^N$ ($N \ge 1$), finie sur les compacts. Toutefois la démonstration donnée ici doit alors être modifiée. Ceci est fait dans l'exercice 7.16 pour $p = 1$ (et la généralisation pour traiter tous les cas $p \in [1, \infty[$ est assez simple).

8.1.3 Densité de $C_c^\infty(\Omega, \mathbb{R})$ dans $L^p(\Omega)$

On a aussi un résultat de densité pour les fonctions définies sur un ouvert de $\mathbb{R}^N$.

Théorème 8.14 (Densité de $\mathcal{D}(\Omega)$ dans $L^p(\Omega)$) *Soient* $N \geq 1$, $p \in [1,+\infty[$ *et* Ω *un ouvert de* $\mathbb{R}^N$*. Alors,* $\mathcal{D}(\Omega) = C_c^\infty(\Omega,\mathbb{R})$ *est dense dans* $L^p(\Omega)$.

DÉMONSTRATION – Pour $\Omega \neq \mathbb{R}^N$, on pose $K_n = \{x \in \mathbb{R}^N \,;\, d(x,\Omega^c) \geq \frac{1}{n}\} \cap B_n$ si $n \in \mathbb{N}^*$.

Soient $f \in L^p(\Omega)$ et $\varepsilon > 0$. On remarque d'abord (en utilisant, comme pour le théorème précédent, le théorème de convergence dominée dans L^p) que $f 1_{K_n} \to f$ dans $L^p(\Omega)$ quand $n \to +\infty$. On peut donc choisir $n_0 \in \mathbb{N}^*$ t.q. $\|f - f_{n_0}\|_p \leq \varepsilon$.

On pose maintenant $g = f_{n_0}$. On peut considérer que $g \in L^p(\mathbb{R}^N)$ et on a $g = 0$ p.p. sur K^c avec $K = K_{n_0}$. En prenant une famille régularisante, $(\rho_n)_{n\in\mathbb{N}^*}$, le lemme 8.11 et la proposition 8.12 donnent que $g * \rho_n \to g$ dans $L^p(\mathbb{R}^N)$ quand $n \to +\infty$ et $(g * \rho_n)_{n\in\mathbb{N}^*} \subset C_c^\infty(\mathbb{R}^N,\mathbb{R})$. Il suffit alors de remarquer que la restriction de $g * \rho_n$ à Ω est à support compact dans Ω dès que $n > n_0$ pour conclure la démonstration.

■

Ici aussi, le théorème 8.14 est encore vrai si on remplace la mesure de Lebesgue par une mesure sur les boréliens de Ω, finie sur les compacts. La démonstration donnée ici doit alors être modifiée (voir l'exercice 7.16).

8.2 Séparabilité de $L^p(\Omega)$

Proposition 8.15 *Soient* $N \geq 1$, Ω *un ouvert de* $\mathbb{R}^N$ *et* p *tel que* $1 \leq p < +\infty$*. Alors, l'espace* $L^p(\Omega)$ *est séparable.*

La démonstration fait l'objet de l'exercice 8.5.

Les espaces du type L^∞ ne sont, en général, pas séparables. L'exercice 8.6 montre que, par exemple, $L^\infty_\mathbb{R}(\mathbb{R}, T, \lambda)$ n'est pas séparable. Une adaptation simple de la démonstration de l'exercice 8.6 montre aussi que l'espace $C_b(\mathbb{R}^d,\mathbb{R})$ (ensemble des fonctions continues bornées de $\mathbb{R}^d$ dans $\mathbb{R}$), muni de la norme de la convergence uniforme, n'est pas séparable. Par contre l'espace $C_0(\mathbb{R}^d,\mathbb{R})$ (ensemble des fonctions continues de $\mathbb{R}^d$ dans $\mathbb{R}$ tendant vers 0 quand le norme de la variable tend vers l'infini), muni de la même norme, est séparable. Ceci est une conséquence des deux premières questions de l'exercice 8.5 (car C_c est dense dans C_0).

8.3 Compacité dans les espaces $L^p(\Omega)$

Soit E un espace vectoriel normé et A une partie de E ; on rappelle que A est (séquentiellement) compact si de toute suite d'éléments de A on peut extraire une sous-suite qui converge. Notons que cette notion de compacité séquentielle est équivalente à la notion de compacité de Borel -Lebesgue (*i.e.* de tout recouvrement de A par des ouverts on peut extraire un sous-recouvrement fini) dès que E est un espace métrique (ce qui est notre cas, car E est un espace vectoriel normé).

Une partie A de E est dite relativement compacte si son adhérence est compacte (ou encore s'il existe un compact K de E tel que $A \subset K$).

Dans le cas où E est un espace de dimension finie, A est compacte si et seulement si A est fermée bornée, et A est relativement compacte si et seulement si A est bornée.

Ces deux caractérisations sont fausses si $\dim(E) = +\infty$. On sait par un théorème de Riesz (différent des deux théorèmes de Riesz donnés dans ce livre !) que la boule unité fermée d'un e.v.n. E est compacte si et seulement si la dimension de E est finie.

On s'intéresse ici au cas $E = L^p(\Omega)$ (Ω ouvert non vide de $\mathbb{R}^N$), espace vectoriel normé de dimension infinie, et on voudrait caractériser les parties relativement compactes ; en particulier, étant donnée une suite de fonctions de $L^p(\Omega)$, sous quelles hypothèses peut-on en extraire une sous-suite qui converge ? Une condition nécessaire évidente est que la partie considérée soit bornée (une partie relativement compacte est toujours bornée). La deuxième condition est, pour $1 \le p < +\infty$ et Ω bornée, que la partie soit équicontinue en moyenne, au sens précisé dans le théorème suivant :

Théorème 8.16 (Kolmogorov) *Soit* $\Omega \in \mathcal{B}(\mathbb{R}^N)$ *et* $1 \le p < +\infty$ *; on considère ici l'espace mesuré* $(E, T, m) = (\Omega, \mathcal{B}(\mathbb{R}^N), \lambda_N)$. *Soit* $A \subset L^p(\Omega)$ *;* A *est relativement compacte si et seulement si :*

1. *Il existe* $C \in \mathbb{R}$ *t.q.* $\|f\|_p \le C$ *pour tout* $f \in A$,
2. *la partie* A *est équicontinue en moyenne, c'est-à-dire que pour tout* $\varepsilon > 0$, *il existe* $\delta > 0$ *t.q. :*
$$\text{pour tout } f \in A,\ |h| \le \delta \Rightarrow \|\tilde{f}(\cdot + h) - \tilde{f}\|_p \le \varepsilon,$$
3. *la partie* A *est équi-petite à l'infini, c'est-à-dire que pour tout* $\varepsilon > 0$, *il existe* $a \in \mathbb{R}_+$, *t.q.*
$$\int_{B_a^c} |\tilde{f}|^p \, dm \le \varepsilon \text{ pour tout } f \in A,$$
où $\tilde{f}$ *est la prolongée de* f *par* 0 *en dehors de* Ω.

DÉMONSTRATION – La démonstration de ce théorème se fait en utilisant la densité de l'espace de fonctions $C_c(\Omega, \mathbb{R})$ dans $L^p(\Omega)$ et le théorème d'Ascoli que nous rappelons ci après (théorème 8.17). Elle est laissée à titre d'exercice, le cas $p = 1$ et $\Omega =]0, 1[$ est donné dans l'exercice 8.9. ■

Dans le cas où Ω est un intervalle borné de $\mathbb{R}$, le théorème 8.16 peut s'énoncer plus simplement sans introduire $\tilde{f}$. Ceci est développé dans l'exercice 8.10.

Théorème 8.17 (Ascoli) *Soient* K *une partie compacte de* $\mathbb{R}$ *et* A *une partie de l'espace vectoriel* $C(K,\mathbb{R})$ *muni de la norme uniforme ;* A *est relativement compacte si et seulement si* A *est bornée et uniformément équicontinue, i.e.*

$\forall\ \varepsilon > 0, \exists \delta > 0\,;\ \forall\ x, y \in K, |x - y| \le \delta \Rightarrow |f(x) - f(y)| \le \varepsilon, \forall\ f \in A.$

Corollaire 8.18 *Soient* $\Omega \in \mathcal{B}(\mathbb{R}^N)$, $1 \le p < +\infty$ *et* $(f_n)_{n\in\mathbb{N}} \subset L^p(\Omega)$. *On suppose que la famille* $A = \{f_n,\ n \in \mathbb{N}\}$ *vérifie les conditions* 1, 2 *et* 3 *du théorème de Kolmogorov. On peut alors extraire de* $(f_n)_{n\in\mathbb{N}}$ *une sous-suite, notée* $(f_{\varphi(n)})_{n\in\mathbb{N}}$, *et il existe* $f \in L^p(\Omega)$ *t.q.* $f_{\varphi(n)} \to f$ *dans* $L^p(\Omega)$ *quand* $n \to +\infty$.

8.4 Compacité faible-$*$

On a introduit au chapitre 6 la convergence faible-$*$ dans le dual d'un espace de Banach. On a une propriété de compacité très utile lorsque l'espace de Banach considéré est séparable :

Proposition 8.19 (Compacité faible-$*$ séq. des bornés de E' si E est séparable) *Soit* F *un espace de Banach séparable et* F' *son dual topologique. Soit* $(T_n)_{n\in\mathbb{N}}$ *une suite bornée de* F'. *Alors, il existe une sous-suite* $(T_{n_k})_{k\in\mathbb{N}}$ *et* $T \in F'$ *t.q. la sous-suite* $(T_{n_k})_{k\in\mathbb{N}}$ *converge vers* T $*$*-faiblement dans* F' *(i.e.* $T_{n_k}(u) \to T(u)$ *(dans* $\mathbb{R}$*) pour tout élément* u *de* F.*)*

DÉMONSTRATION – La démonstration va se faire en trois étapes. L'étape principale est probablement la deuxième étape (qui décrit le procédé diagonal).

On commence par remarquer que, comme F est séparable, il existe une partie A de F, dénombrable et dense. Comme A est dénombrable, il existe uns suite $(f_p)_{p\in\mathbb{N}^*}$ t.q. $A = \{f_p,\ p \in \mathbb{N}^*\}$. On note aussi $C = \sup_{n\in\mathbb{N}} \|T_n\|_{F'}$ (on a $C < +\infty$ car la suite $(T_n)_{n\in\mathbb{N}}$ est bornée dans F').

Etape 1. Dans cette première étape, on montre, par récurrence, l'existence d'une suite d'applications strictement croissantes $(\psi_p)_{p\in\mathbb{N}^*}$ de $\mathbb{N}$ dans $\mathbb{N}$ t.q., pour tout $p \in \mathbb{N}^*$, la suite $(T_{\psi_1\circ\ldots\circ\psi_p(n)}(f_p))_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$.

L'existence de ψ_1 découle du fait que la suite $(T_n(f_1))_{n\in\mathbb{N}}$ est bornée dans $\mathbb{R}$ (en effet, on a $|T_n(f_1)| \le C\|f_1\|_F$). Puis, pour $p \ge 1$, en supposant $\psi_1, \ldots, \psi_p$ construits, on utilise le fait que la suite $(T_{\psi_1\circ\ldots\circ\psi_p(n)}(f_{p+1}))_{n\in\mathbb{N}}$ est bornée dans $\mathbb{R}$. On obtient l'existence d'une application strictement croissante ψ_{p+1} de $\mathbb{N}$ dans $\mathbb{N}$ t.q. la suite $(T_{\psi_1\circ\ldots\circ\psi_{p+1}(n)}(f_{p+1}))_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$.

Etape 2 (procédé diagonal). On note φ_n l'application $\psi_1 \circ \ldots \circ \psi_n$. La deuxième étape consiste à définir ψ de $\mathbb{N}$ dans $\mathbb{N}$ en posant

$$\psi(n) = \varphi_n(n), \text{ pour } n \in \mathbb{N}.$$

La fonction ψ est strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. On va montrer que, pour tout $p \in \mathbb{N}^*$, la suite $(\mathrm{T}_{\psi(n)}(f_p))_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$. En effet, soit $p \in \mathbb{N}^*$. Pour $n > p$, on a :

$$\psi(n) = \psi_1 \circ \ldots \circ \psi_p(\psi_{p+1} \circ \ldots \circ \psi_n(n)) = \varphi_p(\psi_{p+1} \circ \ldots \circ \psi_n(n)).$$

La suite $(\mathrm{T}_{\psi(n)}(f_p))_{n\in\mathbb{N}}$ est donc extraite de la suite $(\mathrm{T}_{\varphi_p(n)}(f_p))_{n\in\mathbb{N}}$ à partir de $n = p$. Ceci prouve que $(\mathrm{T}_{\psi(n)}(f_p))_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$.

Etape 3. On utilise maintenant la densité de A dans F. L'étape 2 nous donne, pour tout $g \in \mathrm{A}$, la convergence dans $\mathbb{R}$ de la suite $(\mathrm{T}_{\psi(n)}(g))_{n\in\mathbb{N}}$. La densité de A dans F et le fait que la suite $(\mathrm{T}_n)_{n\in\mathbb{N}}$ est bornée dans F′ nous permet d'en déduire que la suite $(\mathrm{T}_{\psi(n)}(f))_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$ pour tout $f \in \mathrm{F}$. En effet, soit $f \in \mathrm{F}$. Pour tout $g \in \mathrm{A}$ et tout $m, n \in \mathbb{N}$, on a

$$|\mathrm{T}_{\psi(n)}(f) - \mathrm{T}_{\psi(m)}(f)| \le 2\mathrm{C}\|f - g\|_{\mathrm{F}} + |\mathrm{T}_{\psi(n)}(g) - \mathrm{T}_{\psi(m)}(g)|.$$

On en déduit facilement que la suite $(\mathrm{T}_{\psi(n)}(f))_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{R}$ et donc est convergente.

Pour tout $f \in \mathrm{F}$, on note $\mathrm{T}(f)$ la limite (dans $\mathbb{R}$) de la suite $(\mathrm{T}_{\psi(n)}(f))_{n\in\mathbb{N}}$. Comme les applications T_n sont linéaires, l'application T est aussi linéaire (de F dans $\mathbb{R}$). Puis, comme $\|\mathrm{T}_n\|_{\mathrm{F}'} \le \mathrm{C}$ pour tout $n \in \mathbb{N}$, on a aussi $|\mathrm{T}(f)| \le \mathrm{C}\|f\|_{\mathrm{F}}$ pour tout $f \in \mathrm{F}$. On a donc $\mathrm{T} \in \mathrm{F}'$ et, finalement, $\mathrm{T}_{\psi(n)} \to \mathrm{T}$ ∗-faiblement dans F′, quand $n \to +\infty$. Ce qui termine la démonstration. ■

La proposition 8.19 s'applique aux suites bornées de $\mathrm{L}^\infty(\Omega)$. En effet, grâce au théorème 6.70 et à la proposition 8.15, l'espace $\mathrm{L}^\infty(\Omega)$ peut être identifié au dual de l'espace $\mathrm{L}^1(\Omega)$, qui est un espace séparable (Ω est ici un borélien de $\mathbb{R}^{\mathrm{N}}$). On a donc, par exemple, le résultat suivant :

Proposition 8.20 (Compacité faible-∗ séquentielle des bornés de L^∞) *Soit $d \ge 1$. On note L^∞ l'espace $\mathrm{L}^\infty_{\mathbb{R}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Soit $(u_n)_{n\in\mathbb{N}}$ une suite bornée de L^∞, i.e. telle qu'il existe $\mathrm{M} \in \mathbb{R}_+$ t.q. $\|u_n\|_\infty \le \mathrm{M}$ pour tout $n \in \mathbb{N}$. Alors, il existe une sous-suite, encore notée $(u_n)_{n\in\mathbb{N}}$, et il existe $u \in \mathrm{L}^\infty$ t.q. $u_n \to u$ ∗-faiblement dans L^∞.*

Dans le cas $p \in]1, +\infty[$, on peut écrire une propriété de compacité faible :

Proposition 8.21 (Compacité faible séquentielle des bornés de L^p) *Soit $d \ge 1$ et soit $p \in]1, +\infty[$. On note L^p l'espace $\mathrm{L}^p_{\mathbb{R}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Soit $(u_n)_{n\in\mathbb{N}}$ une suite bornée de L^p, i.e. telle qu'il existe $\mathrm{M} \in \mathbb{R}_+$ t.q. $\|u_n\|_p \le \mathrm{M}$ pour tout $n \in \mathbb{N}$. Alors, il existe une sous-suite, encore notée $(u_n)_{n\in\mathbb{N}}$, et il existe $u \in \mathrm{L}^p$ t.q. $u_n \to u$ faiblement dans L^p, c'est-à-dire t.q. $\int (u_n v - uv) dm \to 0$ pour tout $v \in \mathrm{L}^q$, où $q = \frac{p}{p-1}$.*

On donne maintenant une conséquence, très utile pour les probabilités, de la proposition 8.19. Cette conséquence a déjà été évoquée dans la remarque 6.95.

Proposition 8.22 (Convergence étroite d'une suite tendue) *Soit $d \geq 1$ et $(m_n)_{n\in\mathbb{N}}$ une suite de mesures finies sur $\mathcal{B}(\mathbb{R}^d)$. On suppose que $\sup_{n\in\mathbb{N}} m_n(\mathbb{R}^d) < +\infty$. Il existe alors une sous-suite, encore notée $(m_n)_{n\in\mathbb{N}}$, et une mesure finie sur $\mathcal{B}(\mathbb{R}^d)$ notée m t.q.*

$$\int \varphi dm_n \to \int \varphi dm \text{ pour tout } \varphi \in C_0(\mathbb{R}^d, \mathbb{R}).$$

De plus, si la suite $(m_n)_{n\in\mathbb{N}}$ est tendue, on a alors $m_n \to m$ étroitement. (Ce dernier résultat est aussi vrai si on remplace l'hypothèse "$(m_n)_{n\in\mathbb{N}}$ est tendue" par "$m_n(\mathbb{R}^d) \to m(\mathbb{R}^d)$").

DÉMONSTRATION – On rappelle que $C_0(\mathbb{R}^d, \mathbb{R}) = \{\varphi \in C(\mathbb{R}^d, \mathbb{R}), \varphi(x) \to 0$ quand $|x| \to \infty\}$. On munit $C_0(\mathbb{R}^d, \mathbb{R})$ de la norme de la convergence uniforme, c'est-à-dire $\|\varphi\|_u = \sup\{|\varphi(x)|, x \in \mathbb{R}^d\}$. L'espace $C_0(\mathbb{R}^d, \mathbb{R})$, muni de cette norme, est un espace de Banach séparable (alors que l'espace $C_b(\mathbb{R}^d, \mathbb{R})$ muni de cette même norme n'est pas séparable).

On note E l'espace $C_0(\mathbb{R}^d, \mathbb{R})$ (muni de la norme de la convergence uniforme). Pour $n \in \mathbb{N}$ et $\varphi \in E$, on pose $L_n(\varphi) = \int \varphi dm_n$. La suite $(L_n)_{n\in\mathbb{N}}$ est donc bornée dans E', car on a

$$\|L_n\|_{E'} \leq m_n(\mathbb{R}^d).$$

Il existe donc une sous-suite de la suite $(m_n)_{n\in\mathbb{N}}$, encore notée $(m_n)_{n\in\mathbb{N}}$, et il existe $L \in E'$ t.q.

$$\int \varphi dm_n \to L(\varphi) \text{ pour tout } \varphi \in E.$$

L'application L est donc une application linéaire positive (c'est-à-dire $\varphi \geq 0 \Rightarrow L(\varphi) \geq 0$) de E dans $\mathbb{R}$. D'après le chapitre 5, il existe alors une mesure finie m sur les boréliens de $\mathbb{R}^d$ t.q. $L(\varphi) = \int \varphi dm$ pour tout $\varphi \in E$ (en fait le théorème 5.12 donne ce résultat pour $d = 1$, la démonstration est semblable pour $d > 1$). Ceci donne bien le résultat annoncé, c'est-à-dire

$$\int \varphi dm_n \to \int \varphi dm \text{ pour tout } \varphi \in C_0(\mathbb{R}^d, \mathbb{R}).$$

Il est intéressant de noter que si m_n est pour tout $n \in \mathbb{N}$ une probabilité, la mesure m n'est pas forcément une probabilité et la suite $(m_n)_{n\in\mathbb{N}}$ peut ne pas converger étroitement vers m. (Un exemple pour $d = 1$ consiste à prendre $m_n = \delta_n$ et $m = 0$.)

Toutefois, si on suppose que la suite $(m_n)_{n\in\mathbb{N}}$ est tendue (ou si on suppose que $m_n(\mathbb{R}^d) \to m(\mathbb{R})$), la proposition 6.94 (ou la proposition 6.87 pour le cas $m_n(\mathbb{R}^d) \to m(\mathbb{R})$) donne la convergence étroite de m_n vers m, c'est-à-dire

$$\int \varphi dm_n \to \int \varphi dm \text{ pour tout } \varphi \in C_b(\mathbb{R}^d, \mathbb{R}).$$

■

Du point de vue des probabilités, la conséquence de la proposition 8.22 est que si $(X_n)_{n\in\mathbb{N}}$ est une suite de v.a.r. dont la suite des lois est tendue, il existe alors une v.a.r.

X et une sous-suite de la suite X_n telle que cette sous-suite converge en loi vers X (cf. proposition 9.21 pour le cas de vecteurs aléatoires).

Chapitre 9

Vecteurs aléatoires

9.1 Définition, propriétés élémentaires

Définition 9.1 (Vecteur aléatoire) *Soit $d \geq 1$ et $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle vecteur aléatoire (souvent noté v.a.) de dimension d une application mesurable de Ω dans $\mathbb{R}^d$ (où $\mathbb{R}^d$ est muni de la tribu borélienne).*

Noter que la notation "v.a." signifie indifféremment "variable(s) aléatoire(s)" ou "vecteur(s) aléatoire(s)".

Proposition 9.2 *Soit $d > 1$ et $(\Omega, \mathcal{A})$ un espace probabilisable. Soit X une application de Ω dans $\mathbb{R}^d$. On note $\mathrm{X}_1, \ldots, \mathrm{X}_d$ les composantes de X (de sorte de $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$). On note $\sigma(\mathrm{X})$ la tribu engendrée par X (et $\sigma(\mathrm{X}_i)$, pour $i = 1, \ldots, d$ la tribu engendrée par X_i). On a alors :*

1. *$\sigma(\mathrm{X})$ est la plus petite tribu contenant les tribus $\sigma(\mathrm{X}_1), \ldots, \sigma(\mathrm{X}_d)$.*
2. *X est un v.a. si et seulement si X_i est, pour tout $i \in \{1, \ldots, d\}$, une v.a.r..*

DÉMONSTRATION – On rappelle que $\sigma(\mathrm{X}) = \{\mathrm{X}^{-1}(\mathrm{B}), \mathrm{B} \in \mathcal{B}(\mathbb{R}^d)\}$ et que, pour tout i, $\sigma(\mathrm{X}_i) = \{\mathrm{X}_i^{-1}(\mathrm{A}), \mathrm{A} \in \mathcal{B}(\mathbb{R})\}$. On note $\mathcal{C} = \{\prod_{i=1}^d \mathrm{A}_i, \mathrm{A}_i \in \mathcal{B}(\mathbb{R})$ pour tout $i\}$. On rappelle que $\mathcal{C} \subset \mathcal{B}(\mathbb{R}^d)$ (voir l'exercice 7.11) et que $\mathcal{C}$ engendre $\mathcal{B}(\mathbb{R}^d)$ (c'est même encore vrai si on se limite à prendre pour A_i des intervalles, ouverts, voir l'exercice 2.7).

On note T la plus petite tribu contenant les tribus $\sigma(\mathrm{X}_1), \ldots, \sigma(\mathrm{X}_d)$. Soit $i \in \{1, \ldots, d\}$ et $\mathrm{A} \in \mathcal{B}(\mathbb{R})$. En prenant $\mathrm{B} = \prod_{j=1}^d \mathrm{A}_j$ avec $\mathrm{A}_j = \mathbb{R}$ si $j \neq i$ et $\mathrm{A}_i = \mathrm{A}$, on a $\mathrm{X}^{-1}(\mathrm{B}) \in \sigma(\mathrm{X})$ (car $\mathrm{B} \in \mathcal{C} \subset \mathcal{B}(\mathbb{R}^d)$) et $\mathrm{X}^{-1}(\mathrm{B}) = \mathrm{X}_i^{-1}(\mathrm{A})$. On en déduit $\mathrm{X}_i^{-1}(\mathrm{A}) \in \sigma(\mathrm{X})$ et donc $\sigma(\mathrm{X}_i) \subset \sigma(\mathrm{X})$ pour tout $i \in \{1, \ldots, d\}$. Comme $\sigma(\mathrm{X})$ est une tribu, on a donc $\mathrm{T} \subset \sigma(\mathrm{X})$.

On montre maintenant l'inclusion inverse (c'est-à-dire $\mathrm{T} \supset \sigma(\mathrm{X})$). On remarque que, si $\mathrm{B} \in \mathcal{C}$ on a $\mathrm{B} = \prod_{i=1}^{d} \mathrm{A}_i$ avec des A_i appartenant à $\mathcal{B}(\mathbb{R})$. On a donc $\mathrm{X}^{-1}(\mathrm{B}) = \bigcap_{i=1}^{d} \mathrm{X}_i^{-1}(\mathrm{A}_i) \in \mathrm{T}$ (car $\mathrm{X}_i^{-1}(\mathrm{A}_i) \in \sigma(\mathrm{X}_i) \subset \mathrm{T}$). Or, il est facile de voir que $\{\mathrm{B} \in \mathcal{B}(\mathbb{R}^d)$, t.q. $\mathrm{X}^{-1}(\mathrm{B}) \in \mathrm{T}\}$ est une tribu. Cette tribu contient $\mathcal{C}$, elle contient donc la tribu engendrée par $\mathcal{C}$, c'est-à-dire $\mathcal{B}(\mathbb{R}^d)$. On vient donc de montrer que $\{\mathrm{B} \in \mathcal{B}(\mathbb{R}^d)$, t.q. $\mathrm{X}^{-1}(\mathrm{B}) \in \mathrm{T}\} = \mathcal{B}(\mathbb{R}^d)$, c'est-à-dire que $\mathrm{X}^{-1}(\mathrm{B}) \in \mathrm{T}$ pour tout $\mathrm{B} \in \mathcal{B}(\mathbb{R}^d)$, ou encore que $\sigma(\mathrm{X}) \subset \mathrm{T}$. On a bien, finalement, $\sigma(\mathrm{X}) = \mathrm{T}$, ce qui donne le premier item de la proposition.

Le deuxième item est un conséquence facile du premier. En effet, si X est un v.a., on a pour tout i, $\sigma(\mathrm{X}_i) \subset \sigma(\mathrm{X}) \subset \mathcal{A}$ et donc X_i est une v.a.r.. Réciproquement ; si X_i est, pour tout i, une v.a.r., on a $\sigma(\mathrm{X}_i) \subset \mathcal{A}$ pour tout i. Comme $\mathcal{A}$ est une tribu, on a donc $\mathcal{A} \supset \mathrm{T}$ et comme $\mathrm{T} = \sigma(\mathrm{X})$ on en déduit que X est un v.a.. ■

La démonstration de la proposition 9.2 n'utilise pas vraiment le fait que les X_i soient des applications à valeurs dans $\mathbb{R}$. Elle se généralise donc facilement au cas où les X_i sont des applications à valeurs dans $\mathbb{R}^{d_i}$.

Proposition 9.3 *Soit $p \in \mathbb{N}$, $p \geq 2$, $d_1, \ldots, d_p \in \mathbb{N}^*$. soit $(\Omega, \mathcal{A})$ un espace probabilisable et, pour tout $i \in \{1, \ldots, p\}$, X_i une application de Ω dans $\mathbb{R}^{d_i}$. On note $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$, de sorte que X est une application de Ω dans $\mathbb{R}^d$ avec $d = d_1 + \ldots + d_p$. Soit $d \geq 1$ et $(\Omega, \mathcal{A})$ un espace probabilisable. Soit X une application de Ω dans $\mathbb{R}^d$. On note $\sigma(\mathrm{X})$ la tribu engendrée par X (et $\sigma(\mathrm{X}_i)$, pour $i = 1, \ldots, p$ la tribu engendrée par X_i). On a alors :*

1. *$\sigma(\mathrm{X})$ est la plus petite tribu contenant les tribus $\sigma(\mathrm{X}_1), \ldots, \sigma(\mathrm{X}_p)$.*
2. *X est un v.a. (de dimension d) si et seulement si X_i est, pour tout $i \in \{1, \ldots, p\}$, un v.a. (de dimension d_i).*

Définition 9.4 (Loi d'un v.a.) *Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d \geq 1$ et X un v.a. de dimension d. On appelle loi de probabilité de X, et on note cette loi P_X, la probabilité sur $\mathcal{B}(\mathbb{R}^d)$ image par X de P (c'est-à-dire $\mathrm{P}_\mathrm{X}(\mathrm{A}) = \mathrm{P}(\mathrm{X}^{-1}(\mathrm{A})) = \mathrm{P}(\{\mathrm{X} \in \mathrm{A}\})$ pour tout $\mathrm{A} \in \mathcal{B}(\mathbb{R}^d)$). Si $(\mathrm{X}_1, \ldots, \mathrm{X}_d)$ sont les composantes de X, La probabilité P_X est aussi appelée loi conjointe de $(\mathrm{X}_1, \ldots, \mathrm{X}_d)$.*

Un exemple important est la loi normale multidimensionnelle.

Définition 9.5 (Loi normale multidimensionnelle) *Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d \geq 1$ et X un v.a. de dimension d. Soit $m \in \mathbb{R}^d$ et D une matrice (à coefficients réels, de taille $d \times d$) s.d.p. (c'est-à-dire symétrique définie positive). Le v.a. X a pour* *loi* $\mathcal{N}(m, \mathrm{D})$ *(loi normale de paramètre* m *et* D*) si* $\mathrm{P_X} = f\lambda_d$ *avec :*

$$f(x) = \frac{1}{(2\pi)^{d/2}\sqrt{det(\mathrm{D})}} \exp(-\frac{1}{2}(x-m)^t \mathrm{D}^{-1}(x-m)) \text{ pour } x \in \mathbb{R}^d.$$

Si D *est seulement semi–définie positive (c'est-à-dire* $\mathrm{D}u \cdot u \geq 0$ *pour tout* $u \in \mathbb{R}^d$*) au lieu d'être définie positive (c'est-à-dire* $\mathrm{D}u \cdot u > 0$ *pour tout* $u \in \mathbb{R}^d$, $u \neq 0$*), la loi normale* $\mathcal{N}(m, \mathrm{D})$ *est aussi définie, voir la proposition 9.33, mais ce n'est pas une loi de densité (par rapport à la mesure de Lebesgue), voir l'exercice 10.14 (et l'exercice 9.16 qui donne un exemple de loi normale bidimensionnelle qui n'est pas de densité).*

La fonction f donnée dans la défintion 9.5 est bien positive et d'intégrale (de Lebesgue) égale à 1 (ceci sera d'aileurs démontré dans la proposition 9.33). Nous verrons aussi dans la proposition 9.33 que si X est un v.a. suivant une loi normale multidimensionnelle, X est un v.a. gaussien. L'extension de la définition de la loi normale multidimensionnelle au cas D non inversible permettra alors de dire que les v.a. gaussiens sont exactement ceux qui suivent une loi normale multidimensionnelle (proposition 9.33).

Le fait que les composantes du v.a. X suivent une loi normale ne donne pas que X suit une loi normale multidimensionnelle (voir, par exemple, l'exercice 10.15). Mais, si les composantes du v.a. X suivent une loi normale et sont indépendantes, le v.a. X suit alors une loi normale multidimensionnelle (cf. l'exercice 9.10 ou l'exercice 10.15).

Définition 9.6 (I-ème projecteur) *Soit* $d \geq 1$. *On appelle i-ème projecteur de* $\mathbb{R}^d$ *l'application* π_i*, de* $\mathbb{R}^d$ *dans* $\mathbb{R}$*, qui a un vecteur de* $\mathbb{R}^d$ *fait correspondre sa i-ème composante dans la base canonique de* $\mathbb{R}^d$.

Définition 9.7 (Probabilité marginale) *Soit* $d > 1$.

1. *Soit* p *une probabilité sur les boréliens de* $\mathbb{R}^d$ *et* $i \in \{1, \ldots, d\}$. *On appelle i-ème probabilité marginale de* p*, la probabilité sur les boréliens de* $\mathbb{R}$ *défini par* $p_i(\mathrm{B}) = p(\pi_i^{-1}(\mathrm{B}))$ *pour tout* $\mathrm{B} \in \mathcal{B}(\mathbb{R})$. *(On a donc, par exemple,* $p_1(\mathrm{B}) = p(\mathrm{B} \times \mathbb{R}^{d-1})$.*)*
2. *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* X *un vecteur aléatoire de dimension* d. *On appelle i-ème loi de probabilité marginale du vecteur aléatoire* X *(ou i-ème probabilité marginale du vecteur aléatoire* X*) la probabilité marginale de* $\mathrm{P_X}$ *(c'est donc aussi la mesure image de* $\mathrm{P_X}$ *par le i-ème projecteur* π_i*). La remarque suivante montre que cette probabilité marginale est en fait la loi de* X_i *notée* $\mathrm{P}_{\mathrm{X}_i}$.

Remarque 9.8 Soit $(\Omega, \mathcal{A}, \mathrm{P})$ espace probabilisé, $d \geq 1$ et $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$ un vecteur aléatoire de dimension d. On note q_i la i-ème probabilité marginale du vecteur aléatoire X. Soit $\mathrm{B} \in \mathrm{T}$, par définition de la loi marginale, on a :

$$q_i(\mathrm{B}) = \mathrm{P_X}(\pi_i^{-1}(\mathrm{B})) = \mathrm{P}(\mathrm{X}^{-1}(\pi_i^{-1}(\mathrm{B}))).$$

Comme $\mathrm{X}_i = \pi_i \circ \mathrm{X}$, on a donc :

$$q_i(\mathrm{B}) = \mathrm{P}(\mathrm{X}_i^{-1}(\mathrm{B})).$$

La probabilité q_i est donc aussi la loi de la variable aléatoire X_i.

Remarque 9.9 Soit $(\Omega, \mathcal{A}, \mathrm{P})$ espace probabilisé, $d > 1$ et $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$ un vecteur aléatoire de dimension d. La connaissance de $\mathrm{P_X}$ entraîne la connaissance des $\mathrm{P}_{\mathrm{X}_i}$. La réciproque est en général fausse.

On définit la densité d'une loi de manière analogue au cas scalaire.

Définition 9.10 (Loi de densité) *Soit p une probabilité sur $\mathcal{B}(\mathbb{R}^d)$ ($d \geq 1$), on dit que p est une probabilité de densité (par rapport à la mesure de Lebesgue) s'il existe une application borélienne f de $\mathbb{R}^d$ dans $\mathbb{R}_+$ t.q. $p = f\lambda_d$.*

De même que dans le cas scalaire, on a un théorème qui donne la correspondance entre l'intégrale par rapport à la probabilité P d'une fonction de la v.a.r. X et l'intégrale de cette fonction par rapport à la probabilité $\mathrm{P_X}$:

Théorème 9.11 (Loi image) *Soit $d \geq 1$, $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, X un v.a. de dimension d et $\mathrm{P_X}$ la loi de X. Soit φ une application borélienne de $\mathbb{R}^d$ dans $\mathbb{R}$. On a alors (On rappelle que $\varphi(\mathrm{X})$ désigne $\varphi \circ \mathrm{X}$) :*

1. *$\varphi(\mathrm{X}) \in \mathrm{L}^1_{\mathbb{R}}(\Omega, \mathcal{A}, \mathrm{P})$ si et seulement si $\varphi \in \mathrm{L}^1_{\mathbb{R}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathrm{P_X})$,*
2. *L'égalité*
$$\int_\Omega \varphi(\mathrm{X}) d\mathrm{P} = \int_{\mathbb{R}^d} \varphi d\mathrm{P_X},$$
est vraie si φ prend ses valeurs dans $\mathbb{R}_+$, ou si φ est bornée ou encore si $\varphi(\mathrm{X}) \in \mathrm{L}^1_{\mathbb{R}}(\Omega, \mathcal{A}, \mathrm{P})$.

Définition 9.12 (Fonction de répartition)

Soit $d \geq 1$, $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$ un v.a. de dimension d et $\mathrm{P_X}$ la loi de X. On appelle fonction de répartition du vecteur aléatoire X la fonction définie de $\mathbb{R}^d$ dans $[0, 1]$ par : $\mathrm{F_X}(t_1, \ldots, t_d) = \mathrm{P}(\{\mathrm{X}_1 \leq t_1, \ldots, \mathrm{X}_d \leq t_d\})$.

Proposition 9.13 *Soit* $d \geq 1$, $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* X *un v.a. de dimension* d *de fonction de répartition* F_X. *Alors :*

1. $0 \leq \mathrm{F}_\mathrm{X}(t) \leq 1$ *pour tout* $t \in \mathbb{R}^d$ *;*
2. *Si* $t, t' \in \mathbb{R}^d$, $t \leq t'$ *(i.e.* $t_i \leq t'_i$, *pour tout* $i = 1, \ldots, d$*),* $\mathrm{F}_\mathrm{X}(t) \leq \mathrm{F}_\mathrm{X}(t')$ *;*
3. F_X *est continue à droite en tout point ;*
4. $\mathrm{F}_\mathrm{X}(t_1, \ldots, t_d) \to 1$ *lorsque* $(t_1, \ldots, t_d) \to (+\infty, \ldots, +\infty)$ *;*
5. $\mathrm{F}_\mathrm{X}(t_1, \ldots, t_d) \to 0$ *lorsque* $t_i \to -\infty$ *(à* i *fixé).*

La démonstration découle facilement des propriétés d'une mesure (monotonie, continuité croissante et continuité décroissante).

Proposition 9.14 *Soit* $d \geq 1$, $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* X *un v.a. de dimension* d *de fonction de répartition* F_X. *Si* $\mathrm{F}_\mathrm{X} \in \mathrm{C}^d(\mathbb{R}^d, [0,1])$, *alors* P_X *est une probabilité de densité (par rapport à Lebesgue) et cette densité, notée* f_X, *vérifie*

$$f_\mathrm{X} = \frac{\partial^d \mathrm{F}_\mathrm{X}}{\partial x_1 \ldots \partial x_d} \quad \lambda_d\text{-p.p.}. \tag{9.1}$$

DÉMONSTRATION – On suppose que $d = 1$. Pour tout $a, b \in \mathbb{R}$, $a < b$, la définition de F_X donne $\mathrm{P}_\mathrm{X}(]a,b]) = \mathrm{F}_\mathrm{X}(b) - \mathrm{F}_\mathrm{X}(a)$. Mais, comme F_X est de classe C^1, on a $\mathrm{F}_\mathrm{X}(b) - \mathrm{F}_\mathrm{X}(a) = \int_a^b \mathrm{F}'_\mathrm{X}(t)dt$. En notant m la mesure de densité F'_X par rapport à λ, on a donc $\mathrm{P}_\mathrm{X}(]a,b]) = m(]a,b])$ pour tout $a, b \in \mathbb{R}$, $a < b$, ce qui est suffisant pour dire que $\mathrm{P}_\mathrm{X} = m$ (en utilisant, par exemple, la proposition 2.31).

Pour $d > 1$, on note m la mesure de densité $\partial^d \mathrm{F}_\mathrm{X} / \partial x_1 \ldots \partial x_d$ par rapport à λ_d. Un raisonnement voisin du précédent donne $m(\prod_{i=1}^d]a_i, b_i]) = \mathrm{P}_\mathrm{X}(\prod_{i=1}^d]a_i, b_i])$ pour tout $a_i, b_i \in \mathbb{R}$, $a_i < b_i$, $i = 1, \ldots, d$. On en déduit $m = \mathrm{P}_\mathrm{X}$ avec la proposition 2.31.

■

Définition 9.15 (Espérance) *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $d > 1$ *et* $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$ *un vecteur aléatoire de dimension* d. *On suppose que* $\mathrm{E}(|\mathrm{X}_i|) < \infty$ *pour tout* $i \in \{1, \ldots, d\}$. *L'espérance de* X, *notée* E(X), *est alors le vecteur de* $\mathbb{R}^d$ *dont les composantes sont les espérances des* X_i, *c'est-à-dire* $\mathrm{E}(\mathrm{X}) = (\mathrm{E}(\mathrm{X}_1), \ldots, \mathrm{E}(\mathrm{X}_d))^t$.

Remarque 9.16 Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d > 1$ et $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)^t$ un vecteur aléatoire de dimension d t.q. $\mathrm{E}(|\mathrm{X}_i|) < \infty$ pour tout $i \in \{1, \ldots, d\}$. Soit $u \in \mathbb{R}^d$. L'application $u \cdot \mathrm{X}$ (qui à $\omega \in \Omega$ associe $u \cdot \mathrm{X}(\omega)$, on rappelle que $\xi \cdot \eta$ est le produit scalaire canonique de ξ et η dans $\mathbb{R}^d$) est une v.a.r. intégrable et il est clair que $\mathrm{E}(u \cdot \mathrm{X}) = u \cdot \mathrm{E}(\mathrm{X})$. L'application $u \mapsto \mathrm{E}(u \cdot \mathrm{X})$ est donc l'application linéaire sur $\mathbb{R}^d$ représentée par E(X).

Définition 9.17 (Variance, covariance) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d > 1$ *et* $X = (X_1, \ldots, X_d)^t$ *un vecteur aléatoire de dimension* d*. On suppose que* $E(X_i^2) < \infty$ *pour tout* $i \in \{1, \ldots, d\}$*. On définit alors la matrice de covariance de* X*, notée* $Cov(X)$*, comme la matrice dont le coefficient* i, j *est donné par :*

$$Cov(X)_{i,j} = E((X_i - E(X_i))(X_j - E(X_j))) \text{ pour tout } i, j \in \{1, \ldots, d\}.$$

On a donc $C_{i,i} = Var(X_i)$ *et* $C_{i,j} = Cov(X_i, X_j)$ *(noter d'ailleurs que* $Cov(X_i, X_i) = Var(X_i)$*). Enfin, on peut aussi noter que* $Cov(X)$ *est l'espérance de la matrice* $(X - E(X))(X - E(X))^t$.

Remarque 9.18 Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $d > 1$ et $X = (X_1, \ldots, X_d)^t$ un vecteur aléatoire de dimension d t.q. $E(X_i^2) < \infty$ pour tout $i \in \{1, \ldots, d\}$. Pour tout $u \in \mathbb{R}^d$, on a donc $E((u \cdot X)^2) < \infty$ et il est facile de voir (voir l'exercice 9.6) que $\mathrm{Var}(u \cdot X) = u^t \mathrm{Cov}(X) u$. La matrice Cov(X) est donc la matrice de la forme quadratique $u \mapsto \mathrm{Var}(u \cdot X)$, définie sur $\mathbb{R}^d$. Cette matrice est donc symétrique et semi–définie positive. On peut aussi noter que, pour tout $u, v \in \mathbb{R}^d$, on a $u^t \mathrm{Cov}(X) v = E((u \cdot X - u \cdot E(X))(v \cdot X - v \cdot E(X)))$.

Comme dans le cas des v.a.r., on définit la convergence en loi de v.a. à partir de la convergence étroite (ou vague, puisque c'est équivalent pour des probabilités) des lois des v.a.. La convergence étroite est définie dans la définition 6.86.

Définition 9.19 (Convergence en loi de v.a.) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$, $(X_n)_{n \in \mathbb{N}}$ *une suite de v.a. de dimension* d *et* X *un v.a. de dimension* d*. On dit que* $X_n \to X$ *en loi, quand* $n \to +\infty$*, si :*

$$\int_\Omega \varphi(X_n) dP \to \int_\Omega \varphi(X) dP \text{ pour tout } \varphi \in C_b(\mathbb{R}^d, \mathbb{R}).$$

(Ce que est équivalent à dire que $P_{X_n} \to P_X$ *étroitement.)*

Comme pour les v.a.r., il est possible de définir la convergence en loi pour une suite de v.a. définies sur des espaces probabilisés différents (c'est-à-dire que X_n *est définie sur l'espace probabilisé* $(\Omega_n, \mathcal{A}_n, P_n)$ *dépendant de n, et* X *est définie sur* $(\Omega, \mathcal{A}, P)$*). Il suffit que tous les v.a. aient même dimension. Nous utiliserons parfois implicitement cette définition plus générale.*

Ici, comme dans le cas des v.a.r., on remarque que la convergence en loi donne la tension de la suite des lois. Ceci est donné dans la proposition 9.20.

Proposition 9.20 *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $d \geq 1$, $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *une suite de v.a. de dimension* d *et* X *un v.a. de dimension* d. *On suppose que* $\mathrm{X}_n \to \mathrm{X}$ *en loi. La suite* $(\mathrm{P}_{\mathrm{X}_n})_{n\in\mathbb{N}}$ *est alors tendue, c'est-à-dire*

$$\lim_{a\to+\infty} \mathrm{P}(\{|\mathrm{X}_n| > a\}) = 0 \text{ uniformément par rapport à } n.$$

DÉMONSTRATION – Le résultat est donnée par proposition 6.93 en prenant $m_n = \mathrm{P}_{\mathrm{X}_n}$. ■

La proposition 9.20 admet une réciproque partielle qui est due à la proposition 8.22. Nous donnons maintenant cette réciproque (partielle).

Proposition 9.21 *Soit* $d \geq 1$, $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *une suite de v.a. de dimension* d. *On suppose que la suite des lois des* X_n *est tendue. Il existe alors un v.a.* X *(de dimension* d*) et une sous-suite de la suite* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *telle que cette sous-suite converge en loi vers* X.

DÉMONSTRATION – Pour $n \in \mathbb{N}$, on note m_n la loi de X_n. Comme $(m_n)_{n\in\mathbb{N}}$ est une suite tendue de probabilités sur $\mathcal{B}(\mathbb{R}^d)$, la proposition 8.22 donne l'existence d'une probabilité m sur $\mathcal{B}(\mathbb{R}^d)$ t.q. $m_n \to m$ étroitement. En prenant un v.a. X t.q. $\mathrm{P}_{\mathrm{X}} = m$ (X n'est pas nécessairement défini sur le même espace probabilisé), on a donc la convergence en loi de X_n vers X. ■

9.2 Indépendance

Définition 9.22 *Soit* $p \in \mathbb{N}$, $p \geq 2$, $d_1, \ldots, d_p \in \mathbb{N}^*$. *soit* $(\Omega, \mathcal{A}, p)$ *un espace probabilisé et, pour tout* $i \in \{1, \ldots, p\}$, X_i *un v.a. de dimension* d_i. *Les v.a.* $\mathrm{X}_1, \ldots, \mathrm{X}_p$ *sont dits indépendants si les tribus* $\sigma(\mathrm{X}_1), \ldots, \sigma(\mathrm{X}_p)$ *(engendrées par* $\mathrm{X}_1, \ldots, \mathrm{X}_p$*) sont indépendantes (cf définition 2.57, p. 69)*

La proposition suivante donne des opérations possibles sur l'indépendance.

Proposition 9.23 *Soit* $p \in \mathbb{N}$, $p \geq 2$, $d_1, \ldots, d_p \in \mathbb{N}^*$. *soit* $(\Omega, \mathcal{A})$ *un espace probabilisable et, pour tout* $i \in \{1, \ldots, p\}$, X_i *un v.a. de dimension* d_i. *On suppose que les v.a.* $\mathrm{X}_1, \ldots, \mathrm{X}_p$ *sont indépendants. On se donne maintenant une suite strictement croissante* $p_0, \ldots, p_q$ *(*$q \geq 2$*) t.q.* $1 = p_0 < \ldots < p_q = p + 1$ *et, pour* $i \in \{1, \ldots, q\}$, *on note* Y_i *le v.a.* $(\mathrm{X}_{p_{i-1}}, \ldots, \mathrm{X}_{p_i - 1})^t$, *qui est donc un v.a. de dimension* $r_i = d_{p_{i-1}} + \ldots + d_{p_i - 1}$. *On a alors*

1. Les v.a. $\mathrm{Y}_1, \ldots, \mathrm{Y}_q$ *sont indépendantes,*

2. *Si φ_i est, pour tout $i \in \{1,\dots,q\}$, une application borélienne de $\mathbb{R}^{r_i}$ dans $\mathbb{R}^{s_i}$ (avec $s_i \in \mathbb{N}^\star$), les v.a. $\varphi_1(Y_1)$, …, $\varphi_q(Y_q)$ sont indépendants.*

DÉMONSTRATION – le premier item est une conséquence immédiate de la proposition 2.59 (sur l'indépendance de tribus) et de la proposition 9.3. Le second item est une conséquence immédiate du premier car $\sigma(\varphi_i(Y_i)) \subset \sigma(Y_i)$ pour tout i. ■

Remarque 9.24 Soit $(\Omega, \mathcal{A})$ un espace probabilisable et X, Y deux v.a. (pas nécessairement de même dimension). La proposition 9.23 montre que si X et Y sont indépendants, toute composante de X est indépendante de toute composante de Y. Réciproquement, si toute composante de X est indépendante de toute composante de Y, on en déduit que X et Y sont indépendants. Ceci est encore une conséquence simple de la proposition 2.59.

On donne maintenant une généralisation de la proposition 4.59.

Proposition 9.25 *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $n \geq 2$ et $X_1,\dots,X_n$ des v.a. indépendants, de dimension $d_1,\dots,d_n \in \mathbb{N}^\star$.*

1. *Soit, pour tout $i \in \{1,\dots,n\}$, une fonction borélienne, notée φ_i, de $\mathbb{R}^{d_i}$ dans $\overline{\mathbb{R}}_+$. On a alors :*

$$E\Big(\prod_{i=1}^{n} \varphi_i(X_i)\Big) = \prod_{i=1}^{n} E(\varphi_i(X_i)). \tag{9.2}$$

(En convenant qu'un produit de termes est nul si l'un des termes est nul.)

2. *Soit, pour tout $i \in \{1,\dots,n\}$, φ_i une fonction borélienne de $\mathbb{R}^{d_i}$ dans $\mathbb{R}$. On suppose que $\varphi(X_i)$ est intégrable pour tout $i = 1,\dots,n$. La v.a.r. $\prod_{i=1}^{n} \varphi_i(X_i)$ est alors intégrable et l'égalité (9.2) est vraie.*

3. *Soit, pour tout $i \in \{1,\dots,n\}$, φ_i une fonction borélienne bornée de $\mathbb{R}^{d_i}$ dans $\mathbb{R}$. Alors, l'égalité (9.2) est vraie.*

N.B. Comme dans la proposition 4.59, l'item 3 est donc une condition nécessaire suffisante pour que les v.a. $X_1,\dots,X_n$ soient indépendants.

La démonstration suit pas à pas celle de la proposition 4.59, sans difficulté supplémentaire.

Remarque 9.26 Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. La proposition 9.25 donne aussi des résultats quand les fonctions φ_i sont à la valeurs dans $\mathbb{R}^{r_i}$ avec $r_i \neq 1$. Par exemple, soit X, Y deux v.a. indépendants de dimensions d. On suppose X et Y intégrables (c'est-à-dire $E(|X|) < \infty$ et $E(|Y|) < \infty$). La v.a.r. $X \cdot Y$ est alors intégrable et $E(X \cdot Y) = E(X) \cdot E(Y)$.

Plus généralement, soit X est un v.a. de dimension d, Y est un v.a. de dimension $\bar{d}$, X, Y indépendants. On suppose que φ est une fonction borélienne de $\mathbb{R}^d$ dans $\mathbb{R}^r$ et ψ une fonction borélienne de $\mathbb{R}^{\bar{d}}$ dans $\mathbb{R}^r$ t.q. $\varphi \circ X$ et $\psi \circ Y$ soient intégrables. On a alors $\varphi \circ X \cdot \psi \circ Y$ intégrable et $E(\varphi \circ X \cdot \psi \circ Y) = E(\varphi \circ X) \cdot E(\psi \circ Y)$.

Enfin, si $X_1, \ldots, X_n$ sont des v.a. indépendants de dimension d ($d \geq 1$), la formule donnée dans la proposition 6.97 se généralise et donne :

$$\mathrm{Cov}(X_1 + \ldots + X_n) = \sum_{i=1}^{n} \mathrm{Cov}(X_i).$$

Pour le montrer, il suffit de traiter le cas $n = 2$ (qui est traité dans l'exercice 9.8) et de faire une récurrence sur n.

On donne maintenant la généralisation de la proposition 4.61.

Proposition 9.27 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $n \geq 2$ *et* $X_1, \ldots, X_n$ *des v.a. de dimension* $d_1, \ldots, d_n \in \mathbb{N}^*$. *Ces v.a. sont indépendants si et seulement si on a, pour tout famille* $\{\varphi_i,\ i = 1, \ldots n\}$ *t.q.* $\varphi_i \in C_c(\mathbb{R}^{d_i}, \mathbb{R})$ *pour tout* i,

$$E\Big(\prod_{i=1}^{n} \varphi_i(X_i)\Big) = \prod_{i=1}^{n} E(\varphi_i(X_i)), \tag{9.3}$$

(En convenant qu'un produit de termes est nul si l'un des termes est nul.)

Ici encore, la démonstration suit pas à pas celle de la proposition 4.61, sans difficulté supplémentaire.

Théorème 9.28 (Loi d'un couple de v.a. indépendantes) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé.*

1. *Soit* X, Y *deux v.a.r.. Les v.a.r.* X *et* Y *sont indépendantes si et seulement si la loi du v.a.* $(X, Y)^t$ *(ou du couple* (X, Y)*) est* $P_{(X,Y)^t} = P_X \otimes P_Y$.
2. *Soit* $X_1, \ldots, X_n$ *(*$n \geq 2$*)* n *v.a. de dimension* $d_1, \ldots, d_n \in \mathbb{N}^*$.
   On pose $Z = (X_1, \ldots, X_n)^t$ *(Z est donc un v.a. de dimension* $d_1 + \ldots + d_n$*). Les v.a.* $X_1, \ldots, X_n$ *sont indépendantes si et seulement si la loi du v.a. Z est* $P_Z = P_{X_1} \otimes \ldots \otimes P_{X_n}$.

DÉMONSTRATION – La démonstration du premier item fait partie de l'exercice 9.9. le second item est une généralisation assez simple (en faisant, par exemple, une récurrence sur n pour se ramener au cas $n = 2$). ■

On utilise souvent en théorie des probabilités une suite (finie ou infinie) de v.a.r. indépendantes ayant des lois prescrites. Le théorème suivant montre qu'il existe effectivement un espace probabilisé et une suite de variables aléatoires réelles indépendantes sur cet espace ayant des lois prescrites.

Théorème 9.29 (Existence de v.a.r. indépendantes de lois prescrites)

Soit $(p_n)_{n\in\mathbb{N}^*}$ une suite de probabilités sur $\mathcal{B}(\mathbb{R})$. On a alors :

1. *Pour tout $n \in \mathbb{N}^*$, il existe un espace probabilisé, noté $(\Omega, \mathcal{A}, \mathrm{P})$, et une suite finie de v.a.r. indépendantes, notées $\mathrm{X}_1, \ldots, \mathrm{X}_n$, t.q. $\mathrm{P}_{\mathrm{X}_k} = p_k$ pour tout $k \in \{1, \ldots, n\}$.*
2. *Il existe un espace probabilisé, noté $(\Omega, \mathcal{A}, \mathrm{P})$, et une suite de v.a.r. indépendantes, notée $(\mathrm{X}_n)_{n\in\mathbb{N}^*}$ t.q. $\mathrm{P}_{\mathrm{X}_n} = p_n$ pour tout $n \in \mathbb{N}^*$.*

DÉMONSTRATION – Nous démontrons ici le premier item. Le second (plus difficile) sera admis. Soit $n \in \mathbb{N}^*$. Un raisonnement par récurrence utilisant le théorème d'existence (et unicité) de la mesure produit (théorème 7.3) permet de construire une mesure p sur $\mathcal{B}(\mathbb{R}^n)$ vérifiant, pour toute famille $\mathrm{A}_1, \ldots, \mathrm{A}_n$ de boréliens de $\mathbb{R}$:

$$p(\prod_{i=1}^{n} \mathrm{A}_i) = \prod_{i=1}^{n} p_i(\mathrm{A}_i).$$

On a donc $p = p_1 \otimes \ldots \otimes p_n$.

On prend alors $(\Omega, \mathcal{A}, \mathrm{P}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), p)$ et, pour tout $i = 1, \ldots, n$, X_i est l'application qui a $\omega \in \mathbb{R}^n$ associe sa i-ième composante. Enfin, on note $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)^t$, de sorte que X est un v.a. de dimension n.

Pour tout $\omega \in \mathbb{R}^n$, on a $\mathrm{X}(\omega) = \omega$, ceci prouve que $\mathrm{P}_\mathrm{X} = \mathrm{P}$. Puis, pour tout $i \in \{1, \ldots, n\}$, on a $\mathrm{X}(\omega) = \omega_i$ (où ω_i désigne la i-ième composante de ω). On en déduit que $\mathrm{P}_{\mathrm{X}_i} = p_i$. Enfin, comme $p = p_1 \otimes \ldots \otimes p_n$, on a aussi $\mathrm{P}_\mathrm{X} = p_{\mathrm{X}_1} \otimes \ldots \otimes p_{\mathrm{X}_n}$. Le théorème 9.28 donne donc que les v.a.r. $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes. ■

On s'intéresse maintenant à la somme de variables aléatoires indépendantes.

Proposition 9.30 (Loi de la somme de v.a. indépendantes) *Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d \geq 1$ et X, Y deux v.a. indépendants de dimension d. Alors, $\mathrm{P}_{\mathrm{X}+\mathrm{Y}} = \mathrm{P}_\mathrm{X} * \mathrm{P}_\mathrm{Y}$.*

DÉMONSTRATION – Soit φ une fonction borélienne bornée de $\mathbb{R}^d$ dans $\mathbb{R}$. Comme X et Y sont indépendants, on a $\mathrm{P}_{(\mathrm{X},\mathrm{Y})} = \mathrm{P}_\mathrm{X} \otimes \mathrm{P}_\mathrm{Y}$ (par le théorème 9.28) et donc :

$$\int_\Omega \varphi(\mathrm{X}+\mathrm{Y}) d\mathrm{P} = \int_{\mathbb{R}^{2d}} \varphi(x+y) d\mathrm{P}_{(\mathrm{X},\mathrm{Y})}(x,y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x+y) d\mathrm{P}_\mathrm{X}(x) d\mathrm{P}_\mathrm{Y}(y).$$

La définition de la convolution de mesure donne (voir la définition 7.26 et la proposition 7.27 :

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x+y) d\mathrm{P}_\mathrm{X}(x) d\mathrm{P}_\mathrm{Y}(y) = \int_{\mathbb{R}^d} \varphi(z) d(\mathrm{P}_\mathrm{X} * \mathrm{P}_\mathrm{Y})(z).$$

On en déduit que $\int_\Omega \varphi(X+Y)dP = \int_{\mathbb{R}^d} \varphi(z)d(P_X * P_Y)(z)$, c'est-à-dire $P_{X+Y} = P_X * P_Y$. ■

9.3 Vecteurs gaussiens

Soit $m \in \mathbb{R}$ et $\sigma > 0$. On rappelle que la loi normale $\mathcal{N}(m,\sigma^2)$ est la probabilité (sur $\mathcal{B}(\mathbb{R})$) de densité f avec, pour $x \in \mathbb{R}$,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Pour introduire les v.a. gaussiens, il est utile de définir aussi la loi normale $\mathcal{N}(m,0)$. Cette loi normale est la probabilité (sur $\mathcal{B}(\mathbb{R})$) δ_m (appelée mesure de Dirac au point m). Comme elle vérifie, pour tout fonction φ borélienne de $\mathbb{R}$ dans $\mathbb{R}$, $\int \varphi d\delta_m = \varphi(m)$, il est facile de vérifier que $\mathcal{N}(m,0)$ est la limite étroite, quand $\sigma \to 0$, $\sigma > 0$, de $\mathcal{N}(m,\sigma^2)$.

Définition 9.31 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$ *et* X *un v.a. de dimension* d*. le v.a.* X *est un v.a. gaussien si* $u \cdot X$ *est, pour tout* $u \in \mathbb{R}^d$*, une v.a.r. gaussienne.*

Remarque 9.32 Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $d > 1$ et X un v.a. de dimension d. On suppose que chaque composante de X est une v.a.r. gaussienne. Alors, le vecteur X n'est pas nécessairement gaussien. Mais, si les composantes de X sont indépendantes, le vecteur X est alors un v.a. gaussien. On peut aussi montrer que si X est un v.a. gaussien et que les composantes de X sont indépendantes deux à deux (ou si on a seulement $\mathrm{Cov}(X_i, X_j) = 0$ pour tout $i, j \in \{1, \ldots, d\}$ tels que $i \neq j$), alors les composantes de X sont indépendantes (voir l'exercice 10.15 pour tous ces résultats).

Les v.a. gaussiens sont les vecteurs qui suivent un loi normale multidimensionnelle, dont la définition est donnée dans la définition 9.5, à condition d'étendre convenablement la définition 9.5 au cas D non inversible. C'est l'objectif de la proposition suivante.

Proposition 9.33 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$ *et* X *un v.a. de dimension* d*.*

1. *Soit* $m \in \mathbb{R}^d$ *et* D *une matrice s.d.p. (de taille* $d \times d$*). Si* $X \sim \mathcal{N}(m, D)$*, où* $\mathcal{N}(m, D)$ *est définie par la définition 9.5, alors* X *est un vecteur gaussien,* $E(X) = m$ *et* $Cov(X) = D$.
2. *Si* X *est un vecteur gaussien, on pose* $m = E(X)$ *et* $D = Cov(X)$*. Alors, la loi de* X *ne dépend que de* m *et* D*. Si* D *est inversible on a* $X \sim \mathcal{N}(m, D)$ *(où* $\mathcal{N}(m, D)$ *est définie par la définition 9.5).*

Ceci permet de définir $\mathcal{N}(m,\mathrm{D})$ dans le cas où D est seulement symétrique semi–définie positive (et $m \in \mathbb{R}^d$). On définit $\mathcal{N}(m,\mathrm{D})$ comme étant la loi d'un vecteur gaussien de dimension d t.q. $m = \mathrm{E}(\mathrm{X})$ et $\mathrm{D} = Cov(\mathrm{X})$ (on peut montrer qu'un tel vecteur existe).

DÉMONSTRATION –

1. Démonstration du premier item.

   On suppose que $\mathrm{X} \sim \mathcal{N}(m,\mathrm{D})$ avec $m \in \mathbb{R}^d$ et D une matrice s.d.p.. Comme D est une matrice s.d.p., il existe une matrice diagonale M dont les termes diagonaux sont strictement positifs (ce sont les valeurs propres de M) et une matrice orthogonale P (c'est-à-dire que $\mathrm{P}^{-1} = \mathrm{P}^t$) telles que $\mathrm{D} = \mathrm{P}^t\mathrm{MP}$. On note N la matrice diagonale dont les termes diagonaux sont les racines des termes diagonaux de M, de sorte que $\mathrm{N}^2 = \mathrm{M}$, $\mathrm{D} = \mathrm{P}^t\mathrm{NNP}$ et donc, en notant $\bar{\mathrm{N}} = \mathrm{N}^{-1}$, $\mathrm{D}^{-1} = \mathrm{P}^t\bar{\mathrm{N}}\bar{\mathrm{N}}\mathrm{P} = (\bar{\mathrm{N}}\mathrm{P})^t\bar{\mathrm{N}}\mathrm{P}$. On remarque aussi que $\det(\mathrm{D}) = \det(\mathrm{M})$ et donc $\det(\mathrm{N}) = \sqrt{\det(\mathrm{D})}$.

   On montre tout d'abord que $\mathrm{E}(\mathrm{X}) = m$. On a

$$\mathrm{E}(\mathrm{X}) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\mathrm{D})}} \int_{\mathbb{R}^d} x e^{-\frac{1}{2}(x-m)^t\mathrm{D}^{-1}(x-m)} dx,$$

   ce qui donne, avec le changement de variable $x - m = \mathrm{P}^t\mathrm{N}y$, dont le jacobien est $\det(\mathrm{N})$ (c'est-à-dire $\sqrt{\det(\mathrm{D})}$),

$$\mathrm{E}(\mathrm{X}) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{P}^t\mathrm{N}y e^{-\frac{1}{2}y^t y} dy + \frac{m}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{1}{2}y^t y} dy.$$

   En notant $y_1,\ldots,y_d$ les composantes de y, on a $y^t y = \sum_{i=1}^d y_i^2$ et donc

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}y^t y} dy = \prod_{i=1}^d \int_{\mathbb{R}} e^{-\frac{1}{2}y_i^2} dy_i = (2\pi)^{d/2}$$

   et, pour tout $i \in \{1,\ldots,d\}$, $\int_{\mathbb{R}} y_i e^{-\frac{1}{2}y^t y} dy_i = 0$, ce qui donne, par linéarité de l'intégrale,

$$\int_{\mathbb{R}^d} \mathrm{P}^t\mathrm{N}y e^{-\frac{1}{2}y^t y} dy = \mathrm{P}^t\mathrm{N} \int_{\mathbb{R}^d} y e^{-\frac{1}{2}y^t y} dy = 0.$$

   On obtient bien finalement $\mathrm{E}(\mathrm{X}) = m$.

   On montre maintenant que $\mathrm{D} = \mathrm{Cov}(\mathrm{X})$. Le simple fait que X soit un vecteur aléatoire de dimension d tel que $\mathrm{E}(|\mathrm{X}|^2) < +\infty$ nous donne que, pour tout $u \in \mathbb{R}^d$, $\mathrm{Var}(u\cdot\mathrm{X}) = u^t\mathrm{Cov}(\mathrm{X})u$ (voir l'exercice 9.6). Comme D et Cov(X) sont des matrices symétriques, il suffit donc de montrer que $\mathrm{Var}(u\cdot\mathrm{X}) = u^t\mathrm{D}u$ pour tout $u \in \mathbb{R}^d$ pour conclure que $\mathrm{D} = \mathrm{Cov}(\mathrm{X})$.

   Soit $u \in \mathbb{R}^d$. On a

$$\mathrm{Var}(u\cdot\mathrm{X}) = \mathrm{Var}((\mathrm{X}-m)\cdot u) = \mathrm{E}(u^t(x-m)(x-m)^t u) =$$
$$u^t\Big(\int_{\mathbb{R}^d} \frac{1}{(2\pi)^{d/2}\sqrt{\det(\mathrm{D})}}(x-m)(x-m)^t e^{-\frac{1}{2}(x-m)^t\mathrm{D}^{-1}(x-m)} dx\Big)u. \quad (9.4)$$

Avec le changement de variable $x - m = P^t N y$ on obtient

$$\int_{\mathbb{R}^d} \frac{1}{\sqrt{\det(D)}}(x-m)(x-m)^t e^{-\frac{1}{2}(x-m)^t D^{-1}(x-m)} dx = P^t N\Big(\int_{\mathbb{R}^d} y y^t e^{-\frac{1}{2}y^t y} dy\Big) NP.$$

Pour $i \neq j$, on a $\int_{\mathbb{R}^d} y_i y_j e^{-\frac{1}{2}y^t y} dy = 0$ et, pour $i = j$, $\int_{\mathbb{R}^d} y_i^2 e^{-\frac{1}{2}y^t y} dy = (2\pi)^{d/2}$. On en déduit, en notant I_d la matrice identité de taille $d \times d$,

$$\int_{\mathbb{R}^d} y y^t e^{-\frac{1}{2}y^t y} dy = (2\pi)^{d/2} I_d,$$

et donc, en revenant à 9.4, $\mathrm{Var}(u \cdot X) = u^t P^t NNP = u^t D u$. Ce qui prouve bien que $D = \mathrm{Cov}(X)$.

Il reste à montrer que X est un vecteur gaussien, c'est-à-dire que, pour tout $u \in \mathbb{R}^d$, $u \cdot X$ est une v.a.r. gaussienne, ce qui est équivalent à montrer que $u \cdot (X - m)$ est une v.a.r. gaussienne centrée. Le cas $u = 0$ est immédiat (on obtient la loi $\mathcal{N}(0,0)$). On peut donc se limiter à $u \neq 0$ et même u tel que $|u| = 1$ (car si Z est un v.a.r. gaussienne centrée, αZ est encore une v.a.r. gaussienne centrée pour tout $\alpha \in \mathbb{R}$).

Soit donc $u \in \mathbb{R}^d$ tel que $|u| = 1$. On va montrer que $u \cdot (X - m)$ est une v.a.r. gaussienne centrée.

Soit φ une fonction borélienne bornée positive de $\mathbb{R}$ dans $\mathbb{R}$. On a, en utilisant le changement de variables $(x - m) = y$,

$$E(\varphi(u \cdot (X - m)) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(D)}} \int_{\mathbb{R}^d} \varphi(u \cdot y) e^{-\frac{1}{2}y^t D^{-1} y} dy.$$

Comme $|u| = 1$, on peut choisir une base orthonormée de $\mathbb{R}^d$ pour laquelle u est le premier vecteur. Il existe alors une matrice orthogonale Q (appelée matrice de passage, elle est formée par les vecteurs de cette nouvelle base) telle que, pour tout $y \in R^d$, $y = Qz$ avec $z_1 = u \cdot y$. Le changement de variables $y = Qz$ donne alors dans la formule précédente, en notant que $|\det(Q)| = 1$ et en posant $\bar{D} = Q^t D^{-1} Q$,

$$E(\varphi(u \cdot (X - m)) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(D)}} \int_{\mathbb{R}^d} \varphi(z_1) e^{-\frac{1}{2}z^t \bar{D} z} dz.$$

La matrice $\bar{D}$ est encore une matrice s.d.p.. On peut utiliser sa décomposition de Cholesky (plus exactement la décomposition de Cholesky obtenue en changeant l'ordre des inconnues). Il existe une matrice inversible triangulaire inférieure, notée L, telle que $\bar{D} = L^t L$. Les coefficients diagonaux de L sont strictement positifs et $\det(L) = \sqrt{\det(\bar{D})} = 1/\sqrt{\det(D)}$. Le changement de variables $y = Lz$ donne alors, en notant α le terme de L en première ligne et première colonne,

$$E(\varphi(u \cdot (X - m)) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(\frac{y_1}{\alpha}) e^{-\frac{1}{2}y^t y} dy.$$

Enfin, en intégrant (grâce au théorème de Fubini-Tonelli) d'abord par rapport à $y_2, \ldots, y_n$ et en utilisant le changement de variable $y_1 = \alpha x_1$, on obtient

$$E(\varphi(u \cdot (X - m)) = \frac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} \alpha\varphi(x_1) e^{-\frac{1}{2}\alpha^2 x_1^2} dx_1.$$

Ceci qui prouve que $u \cdot (X - m) \sim \mathcal{N}(0, \alpha^{-2})$ et termine la démonstration du premier item.

2. Démonstration du deuxième item. On suppose que X est un vecteur gaussien et on pose $m = E(X)$ et $D = \mathrm{Cov}(X)$. On montrera dans la proposition 10.23 que la loi de X ne dépend que de m et D. (On admet ce résultat ici.)

   Si D est inversible (et donc D est s.d.p.), soit Y un v.a. de dimension d tel que $Y \sim \mathcal{N}(m, D)$. D'après le premier item de cette proposition, Y est un vecteur gaussien, $E(Y) = m$ et $\mathrm{Cov}(Y) = D$. comme la loi d'un vecteur gaussien ne dépend que de son espérance et de sa covariance, les vecteurs gaussiens X et Y ont même loi. On a donc $X \sim \mathcal{N}(m, D)$

■

Remarque 9.34 La notion de vecteur aléatoire gaussien (et de loi normale multidimensionnelle) généralise la notion de variable aléatoire gaussienne (et de loi normale). Le théorème central limite que nous avons énoncé dans la remarque 6.101 se généralise au cas vectoriel. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $d \geq 1$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. de dimension d i.i.d.. On suppose que $E(|X_1|^2) < \infty$. On pose $E(X_1) = m$, $D = \mathrm{Cov}(X_1)$ et, pour $n \in \mathbb{N}^*$,

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - m).$$

La suite $(Y_n)_{n\in\mathbb{N}^*}$ converge alors en loi vers tout v.a. Y dont la loi est la normale multidimensionnelle $\mathcal{N}(0, D)$. (Voir la définition 9.5 et la proposition 9.33 pour la définition de cette loi normale multidimensionnelle.) Ceci sera démontré au chapitre 10 (théorème 10.24) en utilisant la transformation de Fourier.

9.4 Exercices

9.4.1 Définition, propriétés élémentaires

Exercice 9.1 (Application mesurable)

Soient (E, T) un espace mesurable et $(f_k)_{k=1,\ldots,N}$ une famille de fonctions mesurables de E dans $\mathbb{R}$ muni de la tribu borélienne. Montrer que l'application f définie de E dans $\mathbb{R}^N$ par : $(x, \ldots, x) \mapsto f(x) = (f_1(x), \ldots f_N(x))$ est mesurable.

Exercice 9.2 (Probabilités marginales) Cet exercice montre en particulier (avec les questions 2 et 3) que la connaissance des lois de probabilité marginales ne détermine pas forcément la loi de probabilité.

On note λ_2 la mesure de Lebesgue sur les boréliens de $\mathbb{R}^2$ et $\delta_{(a,b)}$ la mesure de Dirac en (a, b). Soit p une probabilité sur les boréliens de $\mathbb{R}^2$, on note p_1 et p_2 ses probabilités marginales.

1. Soit $(a,b) \in \mathbb{R}^2$. Montrer que $p = \delta_{(a,b)}$ si et seulement si $p_1 = \delta_a$ et $p_2 = \delta_b$.

   Corrigé – *On suppose que* $p = \delta_{(a,b)}$. *On a alors, pour tout* $A \in \mathcal{B}(\mathbb{R})$, $p_1(A) = p(A \times \mathbb{R}) = 1$ *si* $a \in A$ *et* 0 *si* $a \notin A$. *On en déduit que* $p_1 = \delta_a$. *Un raisonnement analogue donne* $p_2 = \delta_b$.

   On suppose maintenant que $p_1 = \delta_a$ *et* $p_2 = \delta_b$. *On remarque alors que*
   $$\{a,b\}^c \subset (\{a\}^c \times \mathbb{R}) \cup (\mathbb{R} \times \{b\}^c)$$
   et donc
   $$p(\{a,b\}^c) \le p(\{a\}^c \times \mathbb{R}) + p(\mathbb{R} \times \{b\}^c) = p_1(\{a\}^c) + p_2(\{b\}^c) = 0.$$
   On en déduit que $p = \delta_{(a,b)}$.

2. On suppose que $p = \mathcal{U}([0,1]^2)$ (c'est-à-dire $p = 1_{]0,1[^2}\lambda_2$) Montrer que $p_1 = p_2 = \mathcal{U}([0,1])$ (c'est-à-dire $p_1 = p_2 = 1_{]0,1[}\lambda$).

   Corrigé – *Soit* $B \in \mathcal{B}(\mathbb{R})$, *on a*
   $$p_1(B) = p(B \times \mathbb{R}) = \int_{]0,1[^2} 1_B(x) d\lambda_2(x,y) = \lambda(B \cap]0,1[)$$
   et donc $p_1 = 1_{]0,1[}\lambda$. *Un raisonnement analogue donne* $p_2 = 1_{]0,1[}\lambda$.

3. Pour $A \in \mathcal{B}(\mathbb{R}^2)$, on pose $p(A) = \lambda(\{t \in [0,1],\ (t,t) \in A\})$. Montrer que p est une probabilité sur les boréliens de $\mathbb{R}^2$ et que $p_1 = p_2 = \mathcal{U}([0,1])$.

   Corrigé – *On remarque tout d'abord que si* $A \in \mathcal{B}(\mathbb{R}^2)$ *l'ensemble* $\{t \in [0,1],\ (t,t) \in A\}$ *est un borélien de* $\mathbb{R}$ *car c'est la projection sur la première coordonnée d'un borélien de* $\mathbb{R}^2$. *On remarque ensuite que* $p(\mathbb{R}^2) = 1$ *et que p est σ-additive. Ceci montre que p est une probabilité.*

   On calcule maintenant p_1 *et* p_2. *Soit* $B \in \mathcal{B}(\mathbb{R})$, *on a*
   $$p_1(B) = p(B \times \mathbb{R}) = \lambda(\{t \in B \cap [0,1]\})$$
   et donc $p_1 = \mathcal{U}([0,1])$. *Un raisonnement analogue donne* $p_2 = 1_{]0,1[}\lambda$.

   N.B. On peut remarquer que $p \neq \mathcal{U}([0,1]^2)$...

Exercice 9.3 (Somme de v.a.r. avec lois de Poisson) Soient X_1 et X_2 deux v.a.r. sur l'espace probabilisé (E, T, p) suivant des lois de Poisson de paramètre λ_1 et λ_2 respectivement, montrer que $X_1 + X_2$ suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$. (On rappelle que la loi de Poisson de paramètre λ est donnée par $\sum_{k \in \mathbb{N}} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$, où δ_k désigne la mesure de Dirac en k).

Exercice 9.4 (Exemple de loi pour le max de deux v.a.r.) On considère deux variables aléatoires réelles X et Y qui ont pour loi de probabilité conjointe la loi dans le plan $\mathbb{R}^2$ de densité $f(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$ par rapport à la mesure de Lebesgue de $\mathbb{R}^2$, σ étant un nombre réel donné non nul. Quelle est la loi de probabilité de la variable aléatoire $Z = \max(|X|, |Y|)$?

Corrigé – *On peut supposer* $\sigma > 0$. *Soit* φ *une fonction de* $\mathbb{R}$ *dans* $\mathbb{R}$, *borélienne bornée. On va chercher la loi de* Z *en exprimant convenablement* $E(\varphi(Z))$ *grâce à la connaissance de la loi du couple* (X, Y).

$$E(\varphi(Z)) = \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(\max(|x|,|y|))\frac{1}{2\pi\sigma^2}\exp\Big(-\frac{x^2+y^2}{2\sigma^2}\Big)dxdy.$$

On en déduit facilement que

$$E(\varphi(Z)) = \frac{8}{2\pi\sigma^2}\int_0^\infty \varphi(x)\exp(-\frac{x^2}{2\sigma^2})\Big(\int_0^x \exp\Big(-\frac{y^2}{2\sigma^2}\Big)dy\Big)dx.$$

Pour $u \ge 0$, *on pose* $e(u) = \int_0^u \exp(-\frac{\xi^2}{2})d\xi$. *On a alors*

$$E(\varphi(Z)) = \frac{4}{\pi\sigma}\int_0^\infty \varphi(x)\exp(-\frac{x^2}{2\sigma^2})e(\frac{x}{\sigma})dx.$$

La loi de Z *a donc une densité par rapport à la mesure de Lebesgue de* $\mathbb{R}$ *et cette densité* g *est donnée par*

$$g(x) = \frac{4}{\pi\sigma}1_{\mathbb{R}_+}(x)\exp(-\frac{x^2}{2\sigma^2})e(\frac{x}{\sigma}).$$

Exercice 9.5 (Calcul de lois) Soit (Ω, T, P) un espace probabilisé et (X, Y) un vecteur aléatoire (de dimension 2) de loi de densité

$$f(x,y) = \frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}.$$

1. Déterminer la densité de la loi de la v.a.r. $X^2 + Y^2$.

Corrigé – *On pose* $Z = X^2 + Y^2$. *Soit* φ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$. *On a*

$$E(\varphi(Z)) = \int_{\mathbb{R}^2}\varphi(x^2+y^2)\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}d(x,y).$$

On peut passer dans cette intégrale en coordonnées polaires. On obtient

$$E(\varphi(Z)) = \frac{1}{2\pi}\int_0^{+\infty}\int_0^{2\pi}\varphi(r^2)e^{-\frac{1}{2}r^2}rd\theta dr = \int_0^{+\infty}\varphi(r^2)e^{-\frac{1}{2}r^2}rdr.$$

Le changement de variable $r^2 = s$ *donne alors*

$$E(\varphi(Z)) = \int_0^{+\infty}\varphi(s)e^{-\frac{1}{2}s}\frac{1}{2}ds.$$

La v.a.r. Z *a donc une densité. Cette densité est donnée par la fonction* g *définie par :* $g(y) = \frac{1}{2}e^{-\frac{1}{2}s}$ *si* $y > 0$ *et* $g(y) = 0$ *si* $y \le 0$.

2. Déterminer la densité de la loi de la v.a.r. $|X| + |Y|$.

Corrigé – *On pose* $T = |X| + |Y|$. *Soit* φ *une fonction borélienne bornée positive de* $\mathbb{R}$ *dans* $\mathbb{R}$. *On a*

$$E(\varphi(T)) = \int_{\mathbb{R}^2}\varphi(|x|+|y|)\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}d(x,y).$$

La parité en x et en y de la fonction que l'on intègre sur $\mathbb{R}^2$ nous permet d'écrire

$$E(\varphi(T)) = 4\int_{(\mathbb{R}_+)^2} \varphi(x+y)\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}d(x,y).$$

Le théorème de Fubini-Tonelli nous donne alors

$$E(\varphi(T)) = 4\int_0^{+\infty}\Big(\int_0^{+\infty}\varphi(x+y)\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}dy\Big)dx.$$

Pour tout $x \in \mathbb{R}_+$ fixé, on utilise le changement de variable $x+y = z$ dans l'intégrale par rapport à y. On obtient

$$E(\varphi(T)) = 4\int_0^{+\infty}\Big(\int_0^{+\infty}\varphi(z)\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+(z-x)^2)}1_{]x,+\infty[}(z)dz\Big)dx.$$

On remarque maintenant que $\frac{1}{2}(x^2+(z-x)^2) = x^2-zx+\frac{1}{2}z^2 = (x-\frac{1}{2}z)^2+\frac{1}{4}z^2$. Ceci donne en réutilisant le théorème de Fubini-Tonelli et le fait que $1_{]x,+\infty[}(z) = 1_{]0,z[}(x)$

$$E(\varphi(T)) = 4\int_0^{+\infty}\varphi(z)e^{-\frac{1}{4}z^2}\Big(\int_0^z \frac{1}{2\pi}e^{-(x-\frac{1}{2}z)^2)}dx\Big)dz.$$

Le changement de variable $x-\frac{1}{2}z = \frac{t}{\sqrt{2}}$ nous donne

$$\int_0^z \frac{1}{2\pi}e^{-(x-\frac{1}{2}z)^2)}dx = \int_{-\frac{\sqrt{2}}{2}z}^{\frac{\sqrt{2}}{2}z}\frac{1}{2\sqrt{2}\pi}e^{-\frac{t^2}{2}}dt = \int_0^{\frac{\sqrt{2}}{2}z}\frac{1}{\sqrt{2}\pi}e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2}\pi}e(\frac{\sqrt{2}}{2}z),$$

en posant pour $s \geq 0$, $e(s) = \int_0^s e^{-\frac{t^2}{2}}dt$. On obtient donc

$$E(\varphi(T)) = \int_0^{+\infty}\varphi(z)\frac{2\sqrt{2}}{\pi}e(\frac{\sqrt{2}}{2}z)e^{-\frac{1}{4}z^2}dz.$$

La v.a.r. T a donc une densité. Cette densité est donnée par la fonction g définie par : $g(y) = \frac{2\sqrt{2}}{\pi}e(\frac{\sqrt{2}}{2}z)e^{-\frac{1}{4}z^2}$ si $y > 0$ et $g(y) = 0$ si $y \leq 0$.

Exercice 9.6 (Matrice des moments, matrice de covariance) Soit $(\Omega, \mathcal{A}, P)$ est un espace probabilisé. Soit X un vecteur aléatoire de dimension d ($d \geq 1$), dont toutes les coordonnées (notées X_i, $i = 1,\ldots,d$) sont supposées être de carré intégrable. La matrice des moments d'ordre 2 du v.a. X est la matrice $E(XX^t)$, dont le terme (i,j) est le réel $E(X_iX_j)$, et on rappelle que la matrice de covariance de X, notée Cov(X), est la matrice $E((X-E(X))(X-E(X))^t) = E(XX^t)-E(X)E(X)^t$, dont le terme (i,j) est la covariance des v.a.r. X_i et X_j. (La notation X^t désigne le transposé du vecteur X.)

1. Soit $u \in \mathbb{R}^d$, montrer que $u^tE(XX^t)u = E((u\cdot X)^2)$ et que $u^t\mathrm{Cov}(X)u = \mathrm{Var}(u\cdot X)$ (On rappelle que $u\cdot X = \sum_{i=1}^d u_iX_i = u^tX$).

Corrigé – *L'application $Z \mapsto E(Z)$ est linéaire, on en déduit que*

$$u^tE(XX^t)u = E(u^tXX^tu) = E((u\cdot X)^2).$$

Cette égalité appliquée au vu v.a. $X - E(X)$ donne $u^tCov(X)u = Var(u\cdot X)$ car $u\cdot X - E(u\cdot X) = u\cdot(X-E(X))$.

2. Montrer que les matrices $\mathrm{E}(\mathrm{XX}^t)$ et $\mathrm{Cov}(\mathrm{X})$ sont symétriques et semi–définies positives.

Corrigé – *Il est facile de voir que* $\mathrm{E}(\mathrm{XX}^t)$ *et* $Cov(\mathrm{X})$ *sont des matrices symétriques (car pour toutes v.a.r.* Y,Z *on a* $\mathrm{E}(\mathrm{YZ}) = \mathrm{E}(\mathrm{ZY})$*). La question précédente donne* $u^t\mathrm{E}(\mathrm{XX}^t)u \geq 0$ *et* $u^t Cov(\mathrm{X})u \geq 0$ *pour tout* $u \in \mathbb{R}^d$*, ce qui signifie que les matrices* $\mathrm{E}(\mathrm{XX}^t)$ *et* $Cov(\mathrm{X})$ *sont semi–définies positives.*

3. Montrer que si A est une matrice $k \times d$ et b un vecteur de $\mathbb{R}^k$, $\mathrm{Y} = \mathrm{AX} + b$, alors $\mathrm{E}(\mathrm{Y}) = \mathrm{AE}(\mathrm{X}) + b$ et $\mathrm{Cov}(\mathrm{Y}) = \mathrm{ACov}(\mathrm{X})\mathrm{A}^t$.

Corrigé – *Comme* X *est un v.a. de carré intégrable, la fonction* Y *est aussi un v.a. de carré intégrable. En utilisant la linéarité de l'application* $\mathrm{Z} \mapsto \mathrm{E}(\mathrm{Z})$*, on calcule son espérance et sa variance :*

$$\begin{aligned}
\mathrm{E}(\mathrm{Y}) &= \mathrm{E}(\mathrm{AX}+b) = \mathrm{AE}(\mathrm{X}) + \mathrm{E}(b) = \mathrm{AE}(\mathrm{X}) + b,\\
Cov(\mathrm{Y}) &= \mathrm{E}([\mathrm{A}(\mathrm{X}-\mathrm{E}(\mathrm{X}))][\mathrm{A}(\mathrm{X}-\mathrm{E}(\mathrm{X}))]^t)\\
&= \mathrm{E}(\mathrm{A}[\mathrm{X}-\mathrm{E}(\mathrm{X})][\mathrm{X}-\mathrm{E}(\mathrm{X})]^t\mathrm{A}^t)\\
&= \mathrm{AE}([\mathrm{X}-\mathrm{E}(\mathrm{X})][\mathrm{X}-\mathrm{E}(\mathrm{X})]^t)\mathrm{A}^t = \mathrm{A}Cov(\mathrm{X})\mathrm{A}^t.
\end{aligned}$$

4. Montrer que si $\mathrm{Cov}(\mathrm{X})$ n'est pas inversible, alors le v.a. X prend p.s. ses valeurs dans un sous–espace affine de $\mathbb{R}^d$ de dimension (inférieure ou) égale à $d-1$, et que la loi de X n'est pas absolument continue par rapport à la mesure de Lebesgue de $\mathbb{R}^d$, notée λ_d (*i.e.* cette loi n'est pas à densité dans $\mathbb{R}^d$).

Corrigé – *Si* $Cov(\mathrm{X})$ *n'est pas inversible, il existe* $u \in \mathbb{R}^d$*,* $u \neq 0$*, t.q.* $Cov(\mathrm{X})u = 0$*. On a donc* $Var(u \cdot \mathrm{X}) = u^t Cov(\mathrm{X})u = 0$*, ce qui donne* $u \cdot \mathrm{X} = \mathrm{E}(u \cdot \mathrm{X})$ *p.s.. On pose* $\beta = \mathrm{E}(u \cdot \mathrm{X})$ *et on a donc :*

$$\mathrm{X} \in \mathrm{H} \textit{ p.s. avec } \mathrm{H} = \{v \in \mathbb{R}^d \textit{ t.q. } u \cdot v = \beta\}.$$

Comme $u \neq 0$*, l'ensemble* H *est un sous–espace affine de* $\mathbb{R}^d$ *de dimension égale à* $d-1$.

L'ensemble H *étant de dimension* $d-1$*, on a* $\lambda_d(\mathrm{H}) = 0$*. Or on a* $\mathrm{P_X}(\mathrm{H}) = \mathrm{P}(\mathrm{X} \in \mathrm{H}) = 1 \neq 0$*. On en déduit que* $\mathrm{P_X}$ *n'est pas absolument continue par rapport à* λ_d *et donc que* $\mathrm{P_X}$ *n'est pas une loi de densité par rapport à* λ_d *(voir le théorème 6.78).*

5. Montrer que si trois points x, y et z de $\mathbb{R}^2$ ne sont pas alignés, tout v. a. X de dimension 2 tel que $\mathrm{P}(\mathrm{X} = x) > 0$, $\mathrm{P}(\mathrm{X} = y) > 0$ et $\mathrm{P}(\mathrm{X} = z) > 0$ a une matrice de covariance non dégénérée.

Corrigé – *On raisonne par l'absurde. Soit x, y et z trois points de* $\mathbb{R}^2$ *non alignés et* X *un v.a. de dimension 2 tel que* $\mathrm{P}(\mathrm{X} = x) > 0$*,* $\mathrm{P}(\mathrm{X} = y) > 0$ *et* $\mathrm{P}(\mathrm{X} = z) > 0$*. On suppose que la matrice de covariance de* X *est dégénérée. La question précédente donne alors qu'il existe une droite de* $\mathbb{R}^2$ *(c'est-à-dire un sous–espace affine de* $\mathbb{R}^2$ *de dimension égale à* 1*), notée* D*, t.q.* $\mathrm{X} \in \mathrm{D}$ *p.s.. Comme* $\mathrm{P}(\mathrm{X} = x) > 0$*, on a donc* $\{\mathrm{X} = x\} \cap \mathrm{D} \neq \emptyset$*. En prenant* $\omega \in \{\mathrm{X} = x\} \cap \mathrm{D}$*, on déduit* $x = \mathrm{X}(\omega) \in \mathrm{D}$*. On montre de même que* $y, z \in \mathrm{D}$*, ce qui est impossible car les trois points x, y et z ne sont pas alignés.*

Exercice 9.7 (Famille de mesures tendue) Soit μ et ν deux probabilités sur les boréliens de $\mathbb{R}$.

1. Montrer qu'il existe p, probabilité sur les boréliens de $\mathbb{R}^2$, t.q.

$$\begin{aligned} &p(\mathrm{A}\times\mathbb{R}) = \mu(\mathrm{A}) \text{ pour tout } \mathrm{A}\in\mathcal{B}(\mathbb{R}),\\ &p(\mathbb{R}\times\mathrm{A}) = \nu(\mathrm{A}) \text{ pour tout } \mathrm{A}\in\mathcal{B}(\mathbb{R}). \end{aligned} \tag{9.5}$$

(On dit alors que μ et ν sont les probabilités marginales de p.)

Corrigé – *On prend $p=\mu\otimes\nu$. La mesure p est bien une probabilité sur les boréliens de $\mathbb{R}^2$ vérifiant (9.5).*

Dans la suite, on note E l'ensemble des probabilités sur $\mathcal{B}(\mathbb{R}^2)$ vérifiant (9.5).

2. Pour $a>0$, on pose $\mathrm{B}_a=\{x\in\mathbb{R}^2 \text{ t.q. } |x|\le a\}$ (où $|\cdot|$ désigne la norme euclidienne dans $\mathbb{R}^2$). Soit $\varepsilon>0$. Montrer qu'il existe $a>0$ (ne dépendant que ε, μ et ν) t.q.,

$$p(\mathrm{B}_a^c)\le\varepsilon \text{ pour tout } p\in\mathrm{E}.$$

(On dit que l'ensemble E est tendu.)

Corrigé – *Grâce à la continuité décroissante d'une mesure, il existe $n\in\mathbb{N}$ (ne dépendant que de ε et μ) et $m\in\mathbb{N}$ (ne dépendant que de ε et ν) t.q.*

$$\mu(\{x_1\in\mathbb{R} \text{ t.q. } |x_1|\ge n\})\le\varepsilon \text{ et } \nu(\{x_2\in\mathbb{R} \text{ t.q. } |x_2|\ge m\})\le\varepsilon.$$

On choisit alors $a>0$ t.q. $a\ge\sqrt{2}n$ et $a\ge\sqrt{2}m$ (le nombre a ne dépend donc que ε, μ et ν) de sorte que

$$\begin{aligned} &\mathrm{B}_a^c=\{x=(x_1,x_2) \text{ t.q. } x_1^2+x_2^2>a^2\}\subset\\ &\{x=(x_1,x_2)\in\mathbb{R}^2 \text{ t.q. } |x_1|\ge n\}\cup\{x=(x_1,x_2)\in\mathbb{R}^2 \text{ t.q. } |x_2|\ge m\}, \end{aligned}$$

c'est-à-dire

$$\mathrm{B}_a^c\subset\Big(\{x_1\in\mathbb{R} \text{ t.q. } |x_1|\ge n\}\times\mathbb{R}\Big)\cup\Big(\mathbb{R}\times\{x_2\in\mathbb{R} \text{ t.q. } |x_2|\ge m\}\Big).$$

Soit $p\in\mathrm{E}$, par σ-sous additivité de p et en utilisant (9.5), on a donc

$$\begin{aligned} p(\mathrm{B}_a^c)&\le p(\{x_1\in\mathbb{R} \text{ t.q. } |x_1|\ge n\}\times\mathbb{R})+p(\mathbb{R}\times\{x_2\in\mathbb{R} \text{ t.q. } |x_2|\ge m\})\\ &=\mu(\{x_1\in\mathbb{R} \text{ t.q. } |x_1|\ge n\})+\nu(\{x_2\in\mathbb{R} \text{ t.q. } |x_2|\ge m\})\le 2\varepsilon, \end{aligned}$$

ce qui termine cette question.

3. Soit $(p_n)_{n\in\mathbb{N}}$ une suite d'éléments de E. Monter qu'il existe p dans E et une sous-suite de la suite $(p_n)_{n\in\mathbb{N}}$, encore notée $(p_n)_{n\in\mathbb{N}}$, t.q.

$$p_n\to p \text{ étroitement, quand } n\to+\infty.$$

[On pourra utiliser les propositions du cours sur les suites de mesures tendues.]

Corrigé – *On utilise ici la proposition 8.22 Comme la suite de probabilités* $(p_n)_{n\in\mathbb{N}}$ *est tendue, on obtient l'existence d'une mesure* p *sur* $\mathcal{B}(\mathbb{R}^2)$ *et d'une sous-suite, encore notée* $(p_n)_{n\in\mathbb{N}}$, *t.q.* $p_n \to p$ *étroitement, quand* $n \to +\infty$. *Comme* $p_n(\mathbb{R}^2) = 1$ *pour tout* $n \in \mathbb{N}$, *on a aussi* $p(\mathbb{R}^2) = 1$. *La mesure* p *est donc une probabilité. Il reste à montrer que* p *vérifie (9.5).*

Pour $\mathrm{A} \in \mathcal{B}(\mathbb{R})$, *on note* $m(\mathrm{A}) = p(\mathrm{A}\times\mathbb{R})$. *Il est facile de voir que* m *est une probabilité sur* $\mathcal{B}(\mathbb{R})$. *Pour* $\varphi = 1_{\mathrm{A}}$, $\mathrm{A} \in \mathcal{B}(\mathbb{R})$, *on a, par définition de* m,

$$\int_{\mathbb{R}} \varphi dm = \int_{\mathbb{R}^2} \varphi(x) dp(x,y). \tag{9.6}$$

Par linéarité de l'intégrale, (9.6) est encore vraie pour φ *étagée positive (de* $\mathbb{R}$ *dans* $\mathbb{R}$*). Puis, par convergence monotone, (9.6) est vraie pour* φ *borélienne positive. Enfin, en décomposant* φ *en* $\varphi^+ - \varphi^-$, *on obtient que (9.6) est vraie pour* φ *borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$ *(et donc, en particulier, pour* $\varphi \in \mathrm{C}_b(\mathbb{R},\mathbb{R})$*).*

Soit $\varphi \in \mathrm{C}_b(\mathbb{R},\mathbb{R})$. *Le raisonnement précédent appliqué à* p_n *au lieu de* p *donne*

$$\int_{\mathbb{R}^2} \varphi(x) dp_n(x,y) = \int_{\mathbb{R}} \varphi(x) d\mu(x).$$

Quand $n \to +\infty$, *comme* $p_n \to p$ *étroitement, on en déduit que*

$$\int_{\mathbb{R}^2} \varphi(x) dp(x,y) = \int_{\mathbb{R}} \varphi(x) d\mu(x).$$

c'est-à-dire , avec (9.6),

$$\int_{\mathbb{R}} \varphi dm = \int_{\mathbb{R}} \varphi d\mu.$$

On en déduit que $m = \mu$ *(en utilisant, par exemple, la proposition 5.8). Ceci prouve que* p *vérifie la première condition de (9.5). Un raisonnement analogue donne la deuxième condition de (9.5).*

4. (Question indépendante des précédentes) Soit $p \in \mathrm{E}$ et (U,V) un vecteur aléatoire de dimension 2 dont la loi est p. Montrer que la loi de U est μ et que la loi de V est ν.

Corrigé – *Soit* φ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$ *(en fait, on peut aussi se limiter ici à prendre* φ *dans* $\mathrm{C}_b(\mathbb{R},\mathbb{R})$*). Comme la loi de* (U,V) *est* p, *on a*

$$\mathrm{E}(\varphi(\mathrm{U})) = \int_{\mathbb{R}^2} \varphi(u) dp(u,v).$$

Puis, comme $p(\mathrm{A}\times\mathbb{R}) = \mu(\mathrm{A})$ *pour tout* $\mathrm{A} \in \mathcal{B}(\mathbb{R})$, *on en déduit grâce à (9.6) (qui est vraie pour* φ *borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*)*

$$\mathrm{E}(\varphi(\mathrm{U})) = \int_{\mathbb{R}^2} \varphi(u) d\mu(u).$$

Ceci prouve que la loi de U *est* μ. *Un raisonnement analogue donne que la loi de* V *est* ν.

9.4.2 Indépendance

Exercice 9.8 (Covariance d'une somme de v.a.i.) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ est un espace probabilisé et X, Y deux vecteurs aléatoires indépendants de dimension d ($d \geq 1$). Montrer que $\mathrm{Cov}(\mathrm{X}+\mathrm{Y}) = \mathrm{Cov}(\mathrm{X})+\mathrm{Cov}(\mathrm{Y})$.

Corrigé – *Comme* X *et* Y *sont des v.a. indépendants, toute composante de* X *est indépendante de toute composante de* Y *(voir, par exemple, la remarque 9.24). On en déduit :*

$$E([X - E(X)][Y - E(Y)]^t) = E([Y - E(Y)][X - E(X)]^t) = 0$$

et donc :

$$\begin{aligned} Cov(X+Y) &= E([X - E(X) + Y - E(Y)][X - E(X) + Y - E(Y)]^t) \\ &= E([X - E(X)][X - E(X)]^t) + E([X - E(X)][Y - E(Y)]^t) \\ &\qquad + E([Y - E(Y)][X - E(X)]^t) + E([Y - E(Y)][Y - E(Y)]^t) \\ &= E([X - E(X)][X - E(X)]^t) + E([Y - E(Y)][Y - E(Y)]^t) \\ &= Cov(X) + Cov(Y). \end{aligned}$$

Exercice 9.9 (Loi d'un couple de v.a. indépendantes) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé et X, Y deux v.a.r..

1. Montrer que X et Y sont indépendantes si et seulement si la loi du couple (X, Y) est $\mathrm{P}_{(\mathrm{X},\mathrm{Y})} = \mathrm{P}_\mathrm{X} \otimes \mathrm{P}_\mathrm{Y}$.

   Corrigé – *On suppose tout d'abord que* X *et* Y *sont indépendantes. Soit* $A, B \in \mathcal{B}(\mathbb{R})$. *On a, par définition de* $P_{(X,Y)}$, $P_{(X,Y)}(A \times B) = P(X \in A, Y \in B)$. *Comme* X *et* Y *sont indépendantes, on a* $P(X \in A, Y \in B) = P(X \in A)P(X \in B)$, *ce qui donne :*

   $$P_{(X,Y)}(A \times B) = P_X(A)P_Y(B).$$

   Or, $P_X \otimes P_Y$ *vérifie aussi* $P_X \otimes P_Y(A \times B) = P_X(A)P_Y(B)$. *La partie unicité du théorème 7.3 donne alors* $P_{(X,Y)} = P_X \otimes P_Y$.

   Réciproquement, on suppose maintenant que $P_{(X,Y)} = P_X \otimes P_Y$, *on a donc, pour tout* $A, B \in \mathcal{B}(\mathbb{R})$, $P(X \in A, Y \in B) = P_{(X,Y)}(A \times B) = P_X \otimes P_Y(A \times B) = P_X(A)P_Y(B)$, *ce qui prouve que les tribus engendrées par* X *et* Y *sont indépendantes, c'est-à-dire que* X *et* Y *sont indépendantes.*

2. On suppose que X et Y ont des densités par rapport à λ : $\mathrm{P}_\mathrm{X} = f\lambda$ et $\mathrm{P}_\mathrm{Y} = g\lambda$, avec $f, g \in \mathrm{L}^1_\mathbb{R}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ (et positives p.p.). Montrer que X et Y sont indépendantes si et seulement si la loi du couple (X, Y) a pour densité la fonction $(x, y) \mapsto f(x)g(y)$ par rapport à λ_2.

   Corrigé – *Comme* $P_X = f\lambda$ *et* $P_Y = g\lambda$, *on a, pour tout* $A, B \in \mathcal{B}(\mathbb{R})$,

   $$P_X \otimes P_Y(A \times B) = P_X(A)P_Y(B) = \int_A f d\lambda \int_B g d\lambda.$$

Le théorème 7.7 donne alors :

$$P_X \otimes P_Y(A \times B) = \int_{A\times B} f(x)g(y)d\lambda_2(x,y).$$

Ce qui donne $P_X \otimes P_Y = F\lambda_2$, *avec* $F(x,y) = f(x)g(y)$ *pour* $x,y \in \mathbb{R}$. *la mesure* $P_X \otimes P_Y$ *est donc la mesure de densité* F *par rapport à* λ_2. *La question 2 est alors une conséquence immédiate de la première question.*

Exercice 9.10 (V.a. suivant une loi normale multi–dimensionnelle) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $X = (X_1, \ldots, X_d)$ $(d > 1)$ un v.a. de dimension d. On suppose que les X_i sont des v.a.r. indépendantes et que $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$, avec $m_i \in \mathbb{R}$ et $\sigma_i > 0$ pour tout i. Montrer que X suit une loi normale multidimensionnelle et donner m et D t.q. $X \sim \mathcal{N}(m, D)$.

Corrigé – *Comme les* X_i $(i = 1, \ldots, d)$ *sont des v.a.r. indépendantes, la loi de* X *est le produit des lois* P_{X_i}, $i = 1, \ldots, d$, *c'est-à-dire* $P_X = P_{X_1} \otimes \ldots \otimes P_{X_d}$. *Cette propriété est donnée dans le théorème 9.28 (elle découle du fait que* $P(X_1 \in A_1, \ldots, X_d \in A_d) = P(X_1 \in A_1)\ldots P(X_d \in A_d)$ *si* $A_1, \ldots A_d$ *sont* d *boréliens de* $\mathbb{R}$*). Ceci donne que pour toute fonction* φ *borélienne bornée de* $\mathbb{R}^d$ *dans* $\mathbb{R}$, *on a :*

$$\int_\Omega \varphi(X)dP = \int_{\mathbb{R}^d} \varphi(x)dP_X(x) = \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} \varphi(x_1, \ldots, x_d)dP_{X_1}(x_1)\ldots dP_{X_d}(x_d).$$

Comme, pour $i = 1, \ldots, d$, $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$, *on a* $P_{X_i} = f_i\lambda$ *avec, pour* $x \in \mathbb{R}$,

$$f_i(x) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-m_i)^2}{2\sigma_i^2}}.$$

On en déduit

$$\int_\Omega \varphi(X)dP = \frac{(2\pi)^{-\frac{d}{2}}}{\prod_{i=1}^d \sigma_i} \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} \varphi(x_1, \ldots, x_d) e^{-\sum_{i=1}^d \frac{(x_i-m_i)^2}{2\sigma_i^2}} dx_1 \ldots dx_d.$$

Ce qui donne :

$$\int_\Omega \varphi(X)dP = (2\pi)^{-\frac{d}{2}} \frac{1}{\sqrt{det D}} \int_{\mathbb{R}^d} \varphi(x) e^{-\frac{1}{2}(x-m)^t D^{-1}(x-m)} dx,$$

où dx *désigne l'intégration par rapport à la mesure de Lebesgue sur* $\mathbb{R}^d$ *(c'est-à-dire* $dx = d\lambda_d(x)$*),* D *désigne la matrice diagonale dont les termes diagonaux sont* $\sigma_1^2, \ldots, \sigma_d^2$ *et* $m = (m_1, \ldots, m_d)^t$. *Ceci montre bien que* X *suit une loi normale multi–dimensionnelle avec* $X \sim \mathcal{N}(m, D)$ *(voir la définition 9.5).*

Exercice 9.11 (Somme de v.a. indépendantes et convolution) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a.r. indépendantes.

1. On suppose que la loi de X a une densité, notée f, par rapport à la mesure de Lebesgue. Montrer que la loi de X + Y a aussi une densité (par rapport à la mesure de Lebesgue) et l'exprimer en fonction de f et de la loi de Y.

Corrigé – *Comme les* X *et* Y *sont des v.a.r. indépendantes, la loi du couple* (X,Y) *est* $P_X \otimes P_Y$ *(théorème 9.28). On a donc, pour toute fonction* φ *borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$ *:*

$$\int_\Omega \varphi(X+Y)dP = \int_{\mathbb{R}^2} \varphi(x+y)dP_{(X,Y)}(x,y)$$
$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x+y)dP_X(x)dP_Y(y).$$

Comme $P_X = f\lambda$*, on a alors :*

$$\int_\Omega \varphi(X+Y)dP = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} \varphi(x+y)f(x)dx\right)dP_Y(y).$$

A y *fixé, on utilise le changement de variable* $x+y=z$*, on obtient :*

$$\int_\Omega \varphi(X+Y)dP = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} \varphi(z)f(z-y)dz\right)dP_Y(y),$$

et donc, en utilisant le théorème de Fubini (théorème 7.12) ou le théorème de Fubini-Tonelli (théorème 7.7) si on se limite à des fonctions φ *positives (ce que l'on peut faire) :*

$$\int_\Omega \varphi(X+Y)dP = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(z-y)dP_Y(y)\right)\varphi(z)dz = \int_{\mathbb{R}} g(z)\varphi(z)dz,$$

avec $g(z) = \int_{\mathbb{R}} f(z-y)dP_Y(y)$ *pour tout* $z \in \mathbb{R}$*. Ceci montre que* $(X+Y)$ *est une v.a.r. dont la loi a pour densité la fonction* g *(par rapport à la mesure de Lebesgue).*

2. On suppose que $X \sim \mathcal{N}(\mu,\sigma^2)$ et $Y \sim \mathcal{N}(m,s^2)$ $(m,\mu,s,\sigma \in \mathbb{R})$. Montrer que $X+Y \sim \mathcal{N}(\mu+m,\sigma^2+s^2)$ (le signe "$\sim$" signifie "a pour loi").

Corrigé – *On peut supposer, bien sûr,* $s \geq 0$ *et* $\sigma \geq 0$ *(car* σ *et* σ *n'interviennent que par leur carré). Nous allons commencer par deux cas particuliers faciles.*

Cas 1 *Si* $s=\sigma=0$*, on a* $X=\mu$ *p.s. et* $Y=m$ *p.s. et donc* $X+Y=m+\mu$ *p.s., ce qui donne bien* $X+Y \sim \mathcal{N}(m+\mu,0)$*.*

Cas 2 *Si* $s=0$ *et* $\sigma>0$*, on a* $Y=m$ *p.s. et donc* $X+Y=X+m$ *p.s.. La densité de* $X+Y$ *(par rapport à Lebesgue) est alors la densité de* X *translatée de* m*, ce qui donne bien* $X+Y \sim \mathcal{N}(m+\mu,\sigma^2)$*. De même, si* $s>0$ *et* $\sigma=0$*, on a* $X+Y \sim \mathcal{N}(m+\mu,s^2)$*.*

Cas 3 *On suppose maintenant que* $s,\sigma>0$*. L'énoncé de cet exercice suggère (implicitement) de faire cette question en utilisant la question précédente. Nous allons procéder ainsi dans ce corrigé (mais d'autres méthodes sont possibles, voir la N.B. ci après). La question précédente donne que* $(X+Y)$ *est une v.a.r. dont la loi a pour densité la fonction* g *(par rapport à la mesure de Lebesgue) vérifiant, pour* $x \in \mathbb{R}$ *:*

$$g(x) = \frac{1}{\sqrt{2\pi}s}\frac{1}{\sqrt{2\pi}\sigma}\int_{\mathbb{R}} e^{-\frac{(x-y-\mu)^2}{2\sigma^2}}e^{-\frac{(y-m)^2}{2s^2}}dy.$$

Soit $x \in \mathbb{R}$*. On pose* $\bar{x} = x-m-\mu$ *et on utilise le changement de variable* $z = \frac{y-m}{s}$*. On obtient :*

$$g(x) = \frac{1}{2\pi s\sigma}\int_{\mathbb{R}} e^{-\frac{(\bar{x}-sz)^2}{2\sigma^2}}e^{-\frac{z^2}{2}}dz = \frac{1}{2\pi\sigma}\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(\bar{x}^2-2sz\bar{x}+(s^2+\sigma^2)z^2)}dz.$$

En remarquant que $(\bar{x}^2 - 2sz\bar{x} + (s^2+\sigma^2)z^2) = (\sqrt{s^2+\sigma^2}z - \frac{1}{\sqrt{s^2+\sigma^2}}s\bar{x})^2 + (\frac{\sigma^2}{s^2+\sigma^2}\bar{x}^2)$, *on a aussi, avec le changement de variable* $\frac{\sqrt{s^2+\sigma^2}}{\sigma}z - \frac{1}{\sigma\sqrt{s^2+\sigma^2}}s\bar{x} = \bar{z}$:

$$g(x) = \frac{1}{2\pi\sigma}e^{-\frac{1}{2(s^2+\sigma^2)}\bar{x}^2}\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(\sqrt{s^2+\sigma^2}z - \frac{1}{\sqrt{s^2+\sigma^2}}s\bar{x})^2}dz$$
$$= \frac{1}{2\pi\sqrt{s^2+\sigma^2}}e^{-\frac{1}{2(s^2+\sigma^2)}\bar{x}^2}\int_{\mathbb{R}} e^{-\frac{1}{2}\bar{z}^2}d\bar{z}.$$

Comme $\int_{\mathbb{R}} e^{-\frac{1}{2}\bar{z}^2}d\bar{z} = \sqrt{2\pi}$ *et* $\bar{x} = x - m - \mu$, *on a donc*

$$g(z) = \frac{1}{\sqrt{2\pi}\sqrt{s^2+\sigma^2}}e^{-\frac{1}{2(s^2+\sigma^2)}(x-m-\mu)^2}.$$

Ce qui donne bien $X + Y \sim \mathcal{N}(\mu + m, \sigma^2 + s^2)$.

N.B. Cette question peut aussi être résolue en utilisant la transformée de Fourier que nous verrons au chapitre 10. Elle est résolue ainsi (avec, plus généralement, $aX + bY$, $a, b \in \mathbb{R}$, *au lieu de* $X + Y$*) dans l'exercice 10.15, question 1(a)).*

Exercice 9.12 (Calcul de π)

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $(U_n)_{n\in\mathbb{N}^*}$ et $(V_n)_{n\in\mathbb{N}^*}$ deux suites de variables aléatoires réelles de loi uniforme sur l'intervalle $[0,1]$. On suppose que toutes ces v.a. sont indépendantes (dans leur ensemble). On pose pour tout $n \geq 1$:

$$X_n = 1 \text{ si } U_n^2 + V_n^2 \leq 1 \text{ et } X_n = 0 \text{ sinon,}$$

et

$$Z_n = 4\frac{X_1 + \cdots + X_n}{n}.$$

1. Déterminer, pour tout $n \in \mathbb{N}^*$, la loi de X_n.

Corrigé – *On note* $D = \{(x,y) \in \mathbb{R}^2, \text{ t.q. } x^2 + y^2 \leq 1\}$ *et* $\psi = 1_D$, *de sorte que* ψ *est une fonction borélienne de* $\mathbb{R}^2$ *dans* $\mathbb{R}$ *(car* $D \in \mathcal{B}(\mathbb{R}^2)$*) et que* $X_n = 1_D(U_n, V_n) = \psi(U_n, V_n)$ *pour* $n \in \mathbb{N}^*$.

Soit $n \in \mathbb{N}^*$. *Comme* U_n *et* V_n *sont des v.a.r. et que* ψ *est borélienne de* $\mathbb{R}^2$ *dans* $\mathbb{R}$, *la fonction* X_n *est donc une v.a.r.. Comme* X_n *ne prend que les valeurs* 1 *et* 0, *la loi de* X_n *est* $P_{X_n} = p\delta_1 + (1-p)\delta_0$ *où* $p = P(U_n^2 + V_n^2 \leq 1)$. *Pour calculer p, on utilise le fait que* U_n *et* V_n *sont indépendantes, on obtient (avec le théorème 9.28 sur la loi d'un couple de v.a.) :*

$$P(U_n^2 + V_n^2 \leq 1) = \int_\Omega 1_D(U_n, V_n)dP = \int_{\mathbb{R}}\int_{\mathbb{R}} 1_D(x,y)dP_{U_n}(x)dP_{V_n}(y).$$

Comme $U_n \sim \mathcal{U}(0,1)$ *et* $V_n \sim \mathcal{U}(0,1)$, *on en déduit (en utilisant les coordonnées polaires) :*

$$p = \int_0^1\int_0^1 1_D(x,y)dxdy = \frac{\pi}{2}\int_0^1 rdr = \frac{\pi}{4}.$$

Donc $p_{X_n} = \frac{\pi}{4}\delta_1 + (1 - \frac{\pi}{4})\delta_0$.

2. Montrer que la suite $(Z_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers π.

Corrigé – *On utilise les notations introduites dans la première question. Pour $n\in\mathbb{N}^*$, on pose $Y_n = 4X_n$, c'est-à-dire $Y_n = 4\psi(U_n, V_n)$. Comme les v.a.r. $(U_n)_{n\in\mathbb{N}^*}$, $(V_n)_{n\in\mathbb{N}^*}$ sont indépendantes (dans leur ensemble), la proposition 9.23 donne la suite $(Y_n)_{n\in\mathbb{N}^*}$ est une suite de v.a.r. indépendantes. De plus, comme $Y_n = 4X_n$, on a $P_{Y_n} = \frac{\pi}{4}\delta_4 + (1-\frac{\pi}{4})\delta_0$. La suite Y_n est donc une suite de v.a.r.i.i.d. de carrés intégrables (on a $E(Y_n) = \pi$ et $E(Y_n^2) = 4\pi$). On peut appliquer la proposition 6.98 (loi faible des grands nombres), il donne que Z_n tend en probabilité vers $E(Y_1)$, c'est-à-dire vers π.*

3. Soit $\alpha \in]0,1[$ et $\varepsilon > 0$. A l'aide de l'inégalité de Tchebychev, donner, en fonction de α et ε, une valeur de $n_0 \in \mathbb{N}^*$ t.q.

$$n \geq n_0 \Rightarrow P[|Z_n - \pi| \geq \varepsilon] \leq \alpha.$$

Corrigé – *On reprend ici la démonstration de la proposition 6.98.*

En utilisant l'inégalité de Bienaymé Tchebychev (lemme 4.57), on a :

$$p(|Z_n - \pi| \geq \varepsilon) = p((Z_n - \pi)^2 \geq \varepsilon^2) \leq \frac{1}{\varepsilon^2} E((Z_n - \pi)^2).$$

Puis, en posant $S_n = \sum_{i=1}^n Y_i$, on a $Z_n = \frac{S_n}{n}$ et donc $E((Z_n-\pi)^2) = \frac{1}{n^2}E((S_n - n\pi)^2) = \frac{Var(S_n)}{n^2}$. La proposition 6.97 donne $Var(S_n) = nVar(Y_1) = n\pi(4-\pi)$. On en déduit finalement

$$p(|Z_n - \pi| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\frac{n\pi(4-\pi)}{n^2} = \frac{\pi(4-\pi)}{n\varepsilon^2}.$$

Il suffit donc de prendre n_0 t.q. $\frac{\pi(4-\pi)}{n_0\varepsilon^2} \leq \alpha$ pour avoir $P(|Z_n - \pi| \geq \varepsilon) \leq \alpha$ si $n \geq n_0$.

Exercice 9.13 (Indépendance des composantes d'un v.a.) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a. de dimension 2. On note X_1, X_2 les composantes de X et Y_1, Y_2 les composantes de Y. On suppose que X et Y ont même loi.

1. Montrer que X_1 et Y_1 ont même loi.

Corrigé – *Soit $\varphi \in C_b(\mathbb{R},\mathbb{R})$. En utilisant la définition de P_X et P_Y et le fait que $P_X = P_Y$, on obtient*

$$\int_\Omega \varphi(X_1)dP = \int_{\mathbb{R}^2} \varphi(x_1)dP_X(x_1,x_2) = \int_{\mathbb{R}^2} \varphi(x_1)dP_Y(x_1,x_2) = \int_\Omega \varphi(Y_1)dP.$$

Comme

$$\int_{\mathbb{R}} \varphi dP_{X_1} = \int_\Omega \varphi(X_1)dP \text{ et } \int_{\mathbb{R}} \varphi dP_{Y_1} = \int_\Omega \varphi(Y_1)dP,$$

on en déduit que $P_{X_1} = P_{Y_1}$.

2. On suppose que X_1 et X_2 sont des v.a.r. indépendantes. Montrer que Y_1 et Y_2 sont aussi des v.a.r. indépendantes.

Corrigé – *Comme* X_1 *et* X_2 *sont indépendantes, on a* $P_X = P_{X_1} \otimes P_{X_2}$ *(théorème 9.28). Comme* $P_X = P_Y$*, on a aussi, d'après la première question,* $P_{X_1} = P_{Y_1}$ *et de même* $P_{X_2} = P_{Y_2}$*. On en déduit que* $P_Y = P_{Y_1} \otimes P_{Y_2}$*, ce qui prouve que* Y_1 *et* Y_2 *sont des v.a.r. indépendantes (théorème 9.28).*

Exercice 9.14 (Mesure produit et différence de deux v.a.r.i.i.d.) 1. Soit m une mesure sur les boréliens de $\mathbb{R}$. On suppose que $m(\mathbb{R}) > 0$. Soit $\eta > 0$. Montrer qu'il existe $a \in \mathbb{R}$ t.q. $m(]a-\eta, a+\eta[) > 0$. En déduire que

$$m \otimes m(A) = \int m(]y-2\eta, y+2\eta[)dm(y) \geq m(]a-\eta, a+\eta[)^2 > 0,$$

avec $A = \{(x,y) \in \mathbb{R}^2 \text{ t.q. } |x-y| < 2\eta\}$. (On rappelle que $m \otimes m$ est une mesure sur $\mathcal{B}(\mathbb{R}^2)$.)

Corrigé – *Pour* $i \in \mathbb{Z}$*, on pose* $a_i = i\eta$*, de sorte que* $\mathbb{R} = \bigcup_{i\in\mathbb{Z}}]a_i - \eta, a_i + \eta[$*. Par σ-sous additivité de m, on a donc* $0 < m(\mathbb{R}) \leq \sum_{i\in\mathbb{Z}} m(]a_i-\eta, a_i+\eta[)$*. Il existe donc* $i \in \mathbb{Z}$ *t.q.* $m(]a_i - \eta, a_i + \eta[) > 0$.

On remarque d'abord que A *est un borélien de* $\mathbb{R}^2$ *car* A *est un ouvert de* $\mathbb{R}^2$*. Par définition de* $m \otimes m$*, on a*

$$m \otimes m(A) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} 1_A(x,y)dm(x) \right) dm(y) = \int m(]y-2\eta, y+2\eta[)dm(y).$$

pour $y \in]a-\eta, a+\eta[$*, on a* $]a-\eta, a+\eta[\subset [y-2\eta, y+2\eta[$ *et donc* $m(]y-2\eta, y+2\eta[) \geq m(]a-\eta, a+\eta[)$*. On a donc*

$$\int m(]y-2\eta, y+2\eta[)dm(y) \geq \int_{]a-\eta,a+\eta[} m(]a-\eta, a+\eta[)dm$$
$$= m(]a-\eta, a+\eta[)^2 > 0.$$

2. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a.r. indépendantes et de même loi. Soit $\varepsilon > 0$. Montrer que

$$P(|X-Y| < \varepsilon) = P(\{\omega \in \Omega, |X(\omega) - Y(\omega)| < \varepsilon\}) > 0.$$

[On pourra utiliser la première question avec $m = P_X = P_Y$.]

Corrigé – *On pose* $B = \{\omega \in \Omega \text{ t.q. } |X(\omega) - Y(\omega)| < \varepsilon\}$ *et* $\bar{B} = \{(x,y) \in \mathbb{R}^2 \text{ t.q. } |x-y| < \varepsilon\}$*. On note* $P_{(X,Y)}$ *la loi du v.a. (de dimension 2)* (X,Y)*. On a*

$$P(|X-Y| < \varepsilon) = \int_\Omega 1_B dP = \int_{\mathbb{R}^2} 1_{\bar{B}} dP_{(X,Y)}.$$

Comme X *et* Y *sont indépendantes, on a, en notant m la loi commune a* X *et* Y*,* $P_{(X,Y)} = m \otimes m$*. On a donc (avec la première question avec* $2\eta = \varepsilon$ *et en remarquant que* $m(\mathbb{R}) = 1 > 0$*),*

$$P(|X-Y| < \varepsilon) = m \otimes m(\bar{B}) > 0.$$

Exercice 9.15 (Les pièges de l'indépendance) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a.r. indépendantes. On suppose que X a pour loi la loi uniforme sur $[0,1]$.

Pour $x \in \mathbb{R}$, on pose $e(x) = \max\{n \in \mathbb{Z},\ n \le x\}$ et $r(x) = x - e(x)$.

Soit f une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$. On définit la v.a.r. Z par la formule $Z = r(X + f(Y))$ (c'est-à-dire $Z(\omega) = r(X(\omega) + f(Y(\omega))$ pour tout $\omega \in \Omega$).

1. Montrer que Z a pour loi la loi uniforme sur $[0,1]$. [Soit φ une fonction borélienne bornée de $[0,1[$ dans $\mathbb{R}$. On note $\bar\varphi$ la fonction de $\mathbb{R}$ dans $\mathbb{R}$ obtenue en prolongeant φ par périodicité (de période 1), c'est-à-dire $\bar\varphi(x+n) = \varphi(x)$ si $x \in [0,1[$ et $n \in \mathbb{Z}$. On pourra commencer par montrer que $E(\varphi(Z)) = \int_{\mathbb{R}} \int_0^1 \bar\varphi(x + f(y))\, dx\, dm(y)$, où m est la loi de Y, c'est-à-dire $m = P_Y$.]

   Corrigé – *La fonction r est borélienne de $\mathbb{R}$ dans $\mathbb{R}$. Par composition de fonctions mesurables on en déduit donc que Z est bien une v.a.r.. On remarque aussi que Z ne prend ses valeurs que dans l'intervalle $[0,1[$.*

   Soit φ une fonction borélienne bornée de $[0,1[$ dans $\mathbb{R}$ et $\bar\varphi$ la fonction de $\mathbb{R}$ dans $\mathbb{R}$ obtenue en prolongeant φ par périodicité de période 1. Comme les v.a.r. X et Y sont indépendantes, la loi du couple (X, Y) est le produit tensoriel des lois de X et Y, on a donc
   $$E(\varphi(Z)) = \int_{\mathbb{R}} \int_0^1 \varphi(r(x + f(y)))dxdm(y).$$
   Comme $z = r(z) + n$ avec $n \in \mathbb{Z}$, on a $\bar\varphi(z) = \varphi(r(z))$ pour tout $z \in \mathbb{R}$ et donc
   $$E(\varphi(Z)) = \int_{\mathbb{R}} \int_0^1 \bar\varphi(x + f(y))dxdm(y).$$
   Pour tout $y \in \mathbb{R}$, on utilise maintenant le changement de variable $\xi = x + f(y)$ dans l'intégrale par rapport à x. Puis, on utilise la périodicité de $\bar\varphi$. On obtient ainsi
   $$\begin{aligned} E(\varphi(Z)) = \int_{\mathbb{R}} \Big(\int_{f(y)}^{f(y)+1} \bar\varphi(\xi)d\xi \Big) dm(y) &= \int_{\mathbb{R}} \Big(\int_0^1 \bar\varphi(\xi)d\xi \Big) dm(y) \\ &= \int_{\mathbb{R}} \Big(\int_0^1 \varphi(\xi)d\xi \Big) dm(y). \end{aligned}$$
   Comme m est une probabilité, on a donc finalement
   $$E(\varphi(Z)) = \int_0^1 \varphi(\xi)d\xi.$$
   Cette dernière égalité est valable pour toute fonction borélienne bornée de $\mathbb{R}$ dans $\mathbb{R}$ (on rappelle que Z prend ses valeurs dans $[0,1[$), elle donne bien que Z a pour loi la loi uniforme sur $[0,1]$.

2. Montrer que Z et Y sont indépendantes. [Soit φ une fonction borélienne bornée de $[0,1[\times\mathbb{R}$ dans $\mathbb{R}$. On pourra s'inspirer de la question précédente pour exprimer $E(\varphi(Z, Y))$.]

Corrigé – *Pour montrer l'indépendance de* Z *et* Y, *il suffit de montrer que la loi du couple* (Z, Y) *est le produit tensoriel des lois de* Z *et* Y.

Soit φ *une fonction borélienne bornée de* $[0,1[\times\mathbb{R}$ *dans* $\mathbb{R}$. *On a, grâce à l'indépendance de* X *et* Y,

$$E(\varphi(Z,Y)) = \int_{\mathbb{R}}\int_0^1 \varphi(r(x+f(y)),y)dxdm(y).$$

Ici aussi, on prolonge φ *par périodicité (en* x*) de période* 1, *c'est-à-dire que l'on pose* $\bar\varphi(x+n,y) = \varphi(x,y)$ *si* $x \in [0,1[$, $n \in \mathbb{Z}$ *et* $y \in \mathbb{R}$. *On obtient*

$$E(\varphi(Z,Y)) = \int_{\mathbb{R}}\int_0^1 \bar\varphi(x+f(y),y)dxdm(y).$$

A y *fixé, on utilise le changement de variable* $\xi = x+f(y)$. *Puis on utilise la périodicité de* $\bar\varphi$. *On a alors*

$$\begin{aligned}E(\varphi(Z,Y)) = \int_{\mathbb{R}}\Big(\int_{f(y)}^{f(y)+1}\bar\varphi(\xi,y)d\xi\Big)dm(y) &= \int_{\mathbb{R}}\Big(\int_0^1\bar\varphi(\xi,y)d\xi\Big)dm(y)\\ &= \int_{\mathbb{R}}\int_0^1 \varphi(\xi,y)d\xi dm(y).\end{aligned}$$

Finalement, pour tout fonction borélienne bornée de $\mathbb{R}^2$ *dans* $\mathbb{R}$ *on a* $E(\varphi(Z,Y)) = \int_{\mathbb{R}}\int_0^1 \varphi(\xi,y)d\xi dm(y)$. *Ceci montre bien que la loi du couple* (Z,Y) *est le produit tensoriel de la loi uniforme avec la loi de* Y *et donc de la loi de* Z *avec la loi de* Y. *les v.a.r.* Z *et* Y *sont donc indépendantes.*

3. Montrer, en donnant un exemple, que Z et Y peuvent ne pas être indépendantes si on retire l'hypothèse que X a pour loi la loi uniforme sur $[0,1]$.

Corrigé – *On suppose, par exemple, que* Y *a pour loi la loi uniforme sur* $[0,1]$, *que* $f(s) = s$ *pour tout* $s \in \mathbb{R}$ *et que* X *ne prend que des valeurs entières. On a alors* $Z = r(X+f(Y)) = r(Y) = Y$ *p.s.. On en déduit que* Z *et* Y *ne sont pas indépendantes. En effet, on a, par exemple,*

$$P(\{Z\in[0,1/2[\text{ et } Y\in[0,1/2]\}) = 1/2 \neq 1/4 = P(\{Z\in[0,1/2]\})P(\{Y\in[0,1/2]\}).$$

9.4.3 Vecteurs gaussiens

Exercice 9.16 (Loi du couple (X,X) **si** $X \sim \mathcal{N}(0,1)$**)**

Pour tout A borélien de $\mathbb{R}^2$, on pose $T(A) = \{x \in \mathbb{R} \text{ t.q. } (x,x)^t \in A\}$.

1. Montrer que, pour tout A un borélien de $\mathbb{R}^2$, $T(A)$ est un borélien de $\mathbb{R}$.

Corrigé – *L'application* $x \mapsto (x,x)^t$ *est continue de* $\mathbb{R}$ *dans* $\mathbb{R}^2$, *elle donc borélienne. Comme* $T(A)$ *est l'image réciproque de* A *par cette application, on a bien* $T(A) \in \mathcal{B}(\mathbb{R})$.

Pour tout A borélien de $\mathbb{R}^2$, on pose $m(A) = \lambda(T(A))$ (où λ est mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$).

On a ainsi défini une application m de $\mathcal{B}(\mathbb{R}^2)$ dans $\overline{\mathbb{R}}_+$.

2. Montrer que l'application m (définie ci-dessus) est une mesure sur $\mathcal{B}(\mathbb{R}^2)$ et que pour toute application φ borélienne positive de $\mathbb{R}^2$ dans $\mathbb{R}$ on a :

$$\int_{\mathbb{R}^2} \varphi(x,y)dm(x,y) = \int_{\mathbb{R}} \varphi(x,x)dx.$$

Corrigé – *Pour montrer que m est une mesure, on remarque que $m(\emptyset) = 0$ (car $\mathrm{T}(\emptyset) = \emptyset$) et m est σ-additive (ce qui découle simplement du fait que $\mathrm{A} \cap \mathrm{B} = \emptyset \Rightarrow \mathrm{T}(\mathrm{A}) \cap \mathrm{T}(\mathrm{B}) = \emptyset$).*

On montre ensuite que $\int_{\mathbb{R}^2} \varphi(x,y)dm(x,y) = \int_{\mathbb{R}} \varphi(x,x)dx$ pour $\varphi = 1_{\mathrm{A}}$, avec $\mathrm{A} \in \mathcal{B}(\mathbb{R}^2)$ (c'est une conséquence immédiate de la définition de m), puis pour φ étagée positive (par linéarité), puis pour φ borélienne positive (car une telle fonction est limite croissante de fonctions étagées positives).

3. Soit $(\Omega, \mathcal{A}, \mathrm{P})$ est un espace probabilisé. et X une v.a.r. t.q. $\mathrm{X} \sim \mathcal{N}(0,1)$.

(a) On pose $\mathrm{Z} = (\mathrm{X},\mathrm{X})^t$. Montrer que le v.a. Z est un vecteur gaussien et donner, pour $a \in \mathbb{R}^2$, la loi de $a \cdot \mathrm{Z}$ (en fonction de a).

Corrigé – *Soit $a = (a_1, a_2)^t \in \mathbb{R}^2$. On pose $\alpha = a_1 + a_2$, de sorte que $a \cdot \mathrm{Z} = \alpha \mathrm{X}$.*

Si $\alpha = 0$, on a $\mathrm{P}_{a \cdot \mathrm{Z}} = \mathrm{P}_0 = \mathcal{N}(0,0)$ (qui est bien une loi gaussienne).

Si $\alpha \neq 0$. Soit φ une application borélienne bornée de $\mathbb{R}$ ans $\mathbb{R}$ on a :

$$\int_\Omega \varphi(a \cdot \mathrm{Z})d\mathrm{P} = \int_{\mathbb{R}} \varphi(\alpha x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \int_{\mathbb{R}} \varphi(y)\frac{1}{\sqrt{2\pi}|\alpha|}e^{-\frac{y^2}{2\alpha^2}}dy.$$

Ce qui prouve que $a \cdot \mathrm{Z} \sim \mathcal{N}(0,\alpha^2)$. Le v.a. Z est donc bien un vecteur gaussien.

(b) Montrer la loi du v.a. $(\mathrm{X},\mathrm{X})^t$ a une densité par rapport à la mesure m définie dans les questions précédentes et donner cette densité. En déduire que la loi du v.a. $(\mathrm{X},\mathrm{X})^t$ n'a pas de densité par rapport à la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^2)$.

Corrigé – *Soit φ borélienne bornée de $\mathbb{R}^2$ dans $\mathbb{R}$. On a :*

$$\int_\Omega \varphi(\mathrm{X},\mathrm{X})d\mathrm{P} = \int_{\mathbb{R}} \varphi(x,x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.$$

On en déduit que la loi du v.a. $(\mathrm{X},\mathrm{X})^t$ est fm avec f borélienne de $\mathbb{R}^2$ dans $\mathbb{R}$ et t.q. $f(x,x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ pour $x \in \mathbb{R}$. Comme m est étrangère à λ_2, loi du v.a. $(\mathrm{X},\mathrm{X})^t$ n'a pas de densité par rapport à λ_2 (qui est la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^2)$).

N.B. La loi de $(\mathrm{X},\mathrm{X})^t$ est une loi normale bidimensionnelle car le vecteur $(\mathrm{X},\mathrm{X})^t$ est gaussien (voir la proposition 9.33) mais ce n'est pas une loi de densité par rapport à la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^2)$.

Exercice 9.17 (Sur le minimum de deux v.a.r. de loi exponentielle) Soit (Ω, T, P) un espace probabilisé, X_1 une v.a.r. dont la loi est la loi exponentielle de paramètre 1 et X_2 une v.a.r. dont la loi est la loi exponentielle de paramètre 2 (soit $\lambda > 0$, on rappelle que la loi exponentielle de paramètre λ est la loi de densité f_λ avec $f_\lambda(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$ et $f_\lambda(x) = 0$ pour $x < 0$). On suppose que X_1 et X_2 sont indépendantes.

Pour $\omega \in \Omega$, on pose

$$Z(\omega) = X_1(\omega) \text{ et } N(\omega) = 1, \text{ si } X_1(\omega) \leq X_2(\omega),$$
$$Z(\omega) = X_2(\omega) \text{ et } N(\omega) = 2, \text{ si } X_2(\omega) < X_1(\omega).$$

1. Montrer que, pour tout $x \in \mathbb{R}$, $\{Z \geq x\} = \{X_1 \geq x\} \cap \{X_2 \geq x\}$. Calculer la fonction de répartition de la loi de Z et en déduire la loi de Z.

   Corrigé – *Soit $x \in \mathbb{R}$.*
   Soit $\omega \in \Omega$. On $Z(\omega) = \min\{X_1(\omega), X_2(\omega)\}$ et donc $Z(\omega) \geq x$ si et seulement si $X_1(\omega) \geq x$ et $X_2(\omega) \geq x$. Ceci donne bien
   $$\{Z \geq x\} = \{X_1 \geq x\} \cap \{X_2 \geq x\}.$$
   Comme X_1 et X_2 sont indépendantes, on a donc
   $$P(\{Z \geq x\}) = P(\{X_1 \geq x\})P(\{X_2 \geq x\}). \tag{9.7}$$
   Si Y est une v.a.r. de loi exponentielle de paramètre λ ($\lambda > 0$). On a alors $P(\{Y \geq x\}) = \int_{x^+}^{\infty} \lambda e^{-\lambda t} dt = e^{-\lambda x^+}$.
   La formule (9.7) donne donc $P(\{Z \geq x\}) = e^{-3x^+}$. La fonction de répartition de Z est donc la fonction $x \mapsto 1 - e^{-3x^+}$. Ceci prouve que Z a pour loi la loi exponentielle de paramètre 3. (On rappelle que la fonction de répartition d'une v.a.r. détermine entièrement la loi de cette v.a.r..)

2. Montrer que $P(\{X_1 = X_2\}) = 0$. En déduire que presque sûrement il existe un unique i t.q. $X_i = Z$.

   Corrigé – *Comme X_1 et X_2 sont indépendantes, on $P_{(X_1,X_2)} = P_{X_1} \otimes P_{X_2}$. On a donc, avec le théorème de Fubini,*
   $$P(\{X_1 = X_2\}) = \int_\Omega 1_{\{X_1 = X_2\}} dP = \int_{\mathbb{R}^2} 1_{\{(x,y) \in \mathbb{R}^2, x=y\}} f_1(x) f_2(y) d(x,y)$$
   $$= \int_0^\infty e^{-x} \Big(\int_x^x 2e^{-2y} dy \Big) dx = 0.$$
   En dehors de l'ensemble $\{X_1 = X_2\}$ (qui de probabilité nulle) il existe un seul i pour lequel $Z = \min\{X_1, X_2\}$. On a donc bien presque sûrement un unique i t.q. $X_i = Z$.

3. Donner la loi de N.

Corrigé – *la v.a.r.* N *ne prend que deux valeurs,* 0 *et* 1. *On calcule* $P(\{N=1\})$ *et* $P(\{N=2\})$. *On a*

$$P(\{N=1\}) = P(\{X_1 \le X_2\}) = \int_\Omega 1_{\{X_1 \le X_2\}} dP = \int_{\mathbb{R}^2} 1_{\{(x,y)\in\mathbb{R}^2, x\le y\}} f_1(x) f_2(y) d(x,y)$$

$$= \int_0^\infty e^{-x} (\int_x^{+\infty} 2e^{-2y} dy) dx = \int_0^\infty e^{-3x} dx = \frac{1}{3}.$$

On a donc $P(\{N=2\}) = \frac{2}{3}$ *et la loi de* N *est* $P_N = \frac{1}{3}\delta_1 + \frac{2}{3}\delta_2$.

4. Les v.a.r. Z et N sont elles indépendantes (justifier la réponse) ?

Corrigé – *Soit* φ *et* ψ *deux fonctions boréliennes bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$. *On compare* $E(\varphi(N)\psi(Z))$ *et* $E(\varphi(N))E(\psi(Z))$.
On a

$$E(\varphi(N)\psi(Z)) = \int_\Omega \varphi(1)\psi(X_1) 1_{\{X_1 \le X_2\}} dP + \int_\Omega \varphi(2)\psi(X_2) 1_{\{X_1 > X_2\}} dP$$

$$= \varphi(1) \int_{\mathbb{R}^2} 1_{\{(x,y)\in\mathbb{R}^2, x\le y\}} \psi(x) f_1(x) f_2(y) d(x,y)$$

$$+\varphi(2) \int_{\mathbb{R}^2} 1_{\{(x,y)\in\mathbb{R}^2, x> y\}} \psi(y) f_1(x) f_2(y) d(x,y)$$

$$= \varphi(1) \int_0^\infty e^{-x} \psi(x) (\int_x^{+\infty} 2e^{-2y} dy) dx$$

$$+\varphi(2) \int_0^\infty 2e^{-2y} \psi(y) (\int_y^{+\infty} e^{-x} dy) dx$$

$$= (\varphi(1) + 2\varphi(2)) \int_0^\infty e^{-3x} \psi(x) dx.$$

En prenant $\varphi = 1$ *dans la formule précédente on a* $E(\psi(Z)) = 3\int_0^\infty e^{-3x}\psi(x)dx$.
Comme $E(\varphi(N)) = \varphi(1)P(\{N=1\}) + \varphi(2)P(\{N=2\}) = \frac{1}{3}(\varphi(1)+2\varphi(2))$ *on en déduit que* $E(\varphi(N)\psi(Z)) = E(\varphi(N))E(\psi(Z))$.
Les v.a.r. Z *et* N *sont donc indépendantes.*