



**HAL**  
open science

## Autour de la notion de mesure

Paul Raynaud de Fitte

► **To cite this version:**

Paul Raynaud de Fitte. Autour de la notion de mesure. Probabilités [math.PR]. Université de Rouen, France, 2004. tel-01711326

**HAL Id: tel-01711326**

**<https://hal.science/tel-01711326v1>**

Submitted on 17 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École doctorale Sciences Physiques et Mathématiques pour  
l'Ingénieur

Universités de Rouen et du Havre, INSA de Rouen

---

Autour de la notion de mesure

Mémoire de synthèse présenté pour l'obtention de  
l'Habilitation à Diriger des Recherches  
en Mathématiques

par

Paul RAYNAUD DE FITTE

Travaux exposés le 9 décembre 2004 devant le jury composé de :

Erik J. BALDER	Professeur à l'Université d'Utrecht, Pays-Bas
Vladimir I. BOGACHEV	Professeur à l'Université d'État de Moscou, Russie
Ahmed BOUZIAD	Professeur à l'Université de Rouen
Claude DELLACHERIE	Directeur de Recherches au CNRS, Univ. de Rouen
Roberto FERNÁNDEZ	Professeur à l'Université de Rouen
Henri HEINICH	Professeur à l'INSA de Rouen
Lionel THIBAUT	Professeur à l'Université Montpellier 2
Dalibor VOLNÝ	Professeur à l'Université de Rouen

Rapporteurs :

Erik J. BALDER	Professeur à l'Université d'Utrecht, Pays-Bas
Vladimir I. BOGACHEV	Professeur à l'Université d'État de Moscou, Russie
Jørgen HOFFMANN-JØRGENSEN	Professeur à l'Université d'Aarhus, Danemark
Lionel THIBAUT	Professeur à l'Université Montpellier 2



## Remerciements

Je tiens tout d'abord à remercier les rapporteurs et les membres du jury : Erik Balder, Vladimir I. Bogachev, Jørgen Hoffmann–Jørgensen et Lionel Thibault, qui ont accepté la charge de lire ce mémoire et d'écrire un rapport, Claude Dellacherie qui a accepté de présider le jury, Ahmed Bouziad, Roberto Fernández, Henri Heinich et Dalibor Volný.

Le travail présenté ici a bénéficié de collaborations diverses qui, même lorsqu'elles étaient anciennes, m'ont ouvert de nouveaux horizons. Je remercie tout particulièrement Charles Castaing, qui fut mon directeur de thèse de doctorat, Adam Jakubowski, Mikhail I. Kamenskiĭ et Michel Valadier. Merci aussi à Philippe Andary pour m'avoir initié aux joies de la programmation récursive en Maple, ce qui m'a permis de calculer les 8483 points de la figure 2.2 page 42.



# Table des matières

<b>1 Mesures de Young</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Mesures de Young et convergence stable . . . . .	2
1.3 Critères de compacité . . . . .	10
1.4 Tableaux de résultats sur les convergences stables . . . . .	17
1.5 Convergence en probabilité ou en moyenne d'ordre $p$ . . . . .	19
1.5.1 Convergence en probabilité de mesures de Young . . . . .	19
1.5.2 Convergence en moyenne de mesures de Young . . . . .	23
1.5.3 Produit fibré de mesures de Young . . . . .	28
1.6 Applications en théorie des probabilités . . . . .	29
1.6.1 Solutions faibles d'inclusions d'évolution stochastiques . . . . .	29
1.6.2 Un théorème de la limite centrale . . . . .	34
<b>2 Autour de la loi forte des grands nombres</b>	<b>39</b>
2.1 Moyenne et loi des grands nombres dans un espace métrique . . . . .	39
2.2 Loi des grands nombres pour des fonctions Pettis intégrables . . . . .	51
2.3 La loi des grands nombres comme résultat de compacité . . . . .	55
<b>Références</b>	<b>58</b>
<b>Index des définitions</b>	<b>68</b>
<b>Index des notations</b>	<b>71</b>

Ce travail est essentiellement centré sur la notion de mesure (positive finie) et se décline en deux parties.

La première partie porte sur les mesures de Young, qui sont à la fois des mesures paramétrées et des mesures définies sur un espace produit. On y étudie diverses topologies (topologies stables, topologie de la convergence en mesure) et on en tire quelques applications en théorie des probabilités.

La deuxième partie est consacrée à la notion de moyenne : étude du barycentre au sens de Herer d'une probabilité sur un espace métrique à courbure négative, loi des grands nombres pour cette notion de barycentre, puis loi des grands nombres pour des variables Pettis intégrables d'un espace localement convexe.

Une part importante de ce travail a consisté à mettre en relation les propriétés topologiques ou géométriques de l'espace sous-jacent avec l'étude des propriétés des mesures ou des espaces de mesures sur cet espace.

# Chapitre 1

## Mesures de Young

La présentation qui suit est tirée essentiellement de [CRdFV04, chapitres 1,2,3,4,9] et de [RdF03, CRdF04]. Les applications stochastiques de la section 1.6.1 proviennent de [JKRdF05].

### 1.1 Introduction

Les mesures de Young ont été inventées plusieurs fois en calcul des variations, en théorie du contrôle ou en probabilités, et sous des noms différents : contrôles relaxés, règles, variables aléatoires floues, mesures paramétrées. . . Elles généralisent les fonctions d'une autre manière que les distributions. Un de leurs principaux attraits est qu'on peut les munir d'une topologie pour laquelle on dispose de critères de compacité ne nécessitant que des conditions très faibles et particulièrement simples à vérifier. Ainsi, une suite de fonctions admet "souvent" une limite au sens des mesures de Young.

Commençons par présenter les mesures de Young dans un cas particulier. Soit  $(f_n)$  une suite de fonctions mesurables à valeurs dans  $\mathbb{R}^d$  et définies sur un espace mesuré  $(\Omega, \mathcal{F}, \lambda)$ , où  $\lambda$  est une mesure finie. On dit que  $(f_n)$  converge de manière stable si, pour tout  $A \in \mathcal{F}$  et pour toute fonction continue bornée  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ , la suite  $(\int_A \varphi \circ f_n d\lambda)_n$  converge. La limite de  $(f_n)$  est l'application  $(A, \varphi) \mapsto \lim_n \int_A \varphi \circ f_n d\lambda$ . Il se trouve que cette application peut être identifiée à une mesure  $\mu$  sur  $\Omega \times \mathbb{R}^d$ , définie par

$$\int_{\Omega \times \mathbb{R}^d} \mathbf{1}_A \otimes \varphi d\mu = \lim_n \int_A \varphi \circ f_n d\lambda$$

pour tout  $A$  et tout  $\varphi$  comme ci-dessus (on note  $\mathbf{1}_A$  la fonction indicatrice de  $A$  et  $g \otimes h$  désigne l'application  $(\omega, x) \mapsto g(\omega)h(x)$ ). Il est à remarquer que la marge de  $\mu$  sur  $\Omega$  est exactement  $\lambda$ . Dans cette convergence, chaque  $f_n$



peut également être identifié à une mesure  $\underline{\delta}_{f_n}$  sur  $\Omega \times \mathbb{R}^d$ , plus précisément  $\underline{\delta}_{f_n} = \int_{\Omega} \delta_{\omega} \otimes \delta_{f_n(\omega)} d\lambda(\omega)$ , où  $\delta_x$  désigne la masse de Dirac en  $x$ . Autrement dit, la mesure  $\underline{\delta}_{f_n}$  est portée par le graphe de  $f_n$  et sa marge sur  $\Omega$  est  $\lambda$ . On a alors, pour tout  $A$  et tout  $\varphi$ ,

$$\int_{\Omega \times \mathbb{R}^d} \mathbf{1}_A \otimes \varphi d\underline{\delta}_{f_n} = \int_A \varphi \circ f_n d\lambda.$$

Ainsi, la convergence stable de  $(f_n)_n$  vers  $\mu$  peut être exprimée par

$$\lim_n \int_{\Omega \times \mathbb{R}^d} \mathbf{1}_A \otimes \varphi d\underline{\delta}_{f_n} = \int_{\Omega \times \mathbb{R}^d} \mathbf{1}_A \otimes \varphi d\mu$$

pour tout  $A$  et tout  $\varphi$ . Dans cet exemple, la convergence stable peut être vue comme une convergence dans l'espace des mesures sur  $\Omega \times \mathbb{R}^d$  qui ont pour marge commune  $\lambda$  sur  $\Omega$ . Ce sont bien sûr ces mesures qui sont appelées mesures de Young, ou contrôles relaxés, etc. Ainsi, la convergence stable donne des limites “relaxées” à des suites de fonctions, c'est-à-dire des limites qui ne sont pas nécessairement des fonctions, mais peuvent être des mesures. Par exemple, la convergence stable convergence fournit des solutions généralisées à des problèmes variationnels ou à des problèmes de contrôle [Rou97], ou bien ce que les probabilistes appellent des solutions faibles d'équations différentielles stochastiques [Pel81b, JKRdF05], on en verra un exemple dans la section 1.6.1 page 29. La convergence stable sert aussi dans les théorèmes de limite en probabilité [Jac97, Let98], comme on le verra dans la section 1.6.2.

Pour illustrer le rôle de la compacité, supposons maintenant que  $(f_n)$  soit une suite faiblement convergente dans  $L^1(\Omega, \mathcal{F}, \lambda; \mathbb{R}^d)$ . Alors on montre que  $(f_n)_n$  est séquentiellement relativement compacte pour la topologie stable ; si de plus  $(f_n)$  n'est pas fortement convergente, alors il existe une sous-suite  $(f'_n)$  qui converge de manière stable vers une mesure de Young non associée à une fonction, qui décrit les oscillations de  $(f'_n)_n$  autour de sa limite faible dans  $L^1(\Omega, \mathcal{F}, \lambda; \mathbb{R}^d)$  [Val94, Val99].

## 1.2 Mesures de Young et convergence stable

**Notations générales** Pour tout espace topologique  $\mathbb{T}$ , on note  $\mathfrak{B}(\mathbb{T})$  la tribu borélienne de  $\mathbb{T}$ ,  $\mathfrak{O}(\mathbb{T})$  l'ensemble des parties ouvertes de  $\mathbb{T}$ ,  $\mathfrak{F}(\mathbb{T})$  l'ensemble des parties fermées de  $\mathbb{T}$  et  $\mathfrak{K}(\mathbb{T})$  l'ensemble des parties compactes de  $\mathbb{T}$ . Lorsqu'il n'y a pas d'ambiguïté, on simplifie les notations en écrivant par exemple  $\mathfrak{O}$  au lieu de  $\mathfrak{O}(\mathbb{T})$ . Par *mesure*, on entendra “mesure positive finie”.

Si  $(\mathbb{T}, \mathcal{F})$  est un espace mesuré, on note  $M(\mathbb{T}, \mathcal{F})$  l'ensemble des mesures positives finies sur  $(\mathbb{T}, \mathcal{F})$  et  $P(\mathbb{T}, \mathcal{F})$  l'ensemble des mesures de probabilité sur  $(\mathbb{T}, \mathcal{F})$ . Si  $\mathbb{T}$  est topologique, les notations  $M(\mathbb{T})$  et  $P(\mathbb{T})$  signifieront respectivement  $M(\mathbb{T}, \mathfrak{B}(\mathbb{T}))$  et  $P(\mathbb{T}, \mathfrak{B}(\mathbb{T}))$ . L'espace  $M(\mathbb{T})$  est muni, soit de la *topologie faible*, engendrée par les applications  $\mu \mapsto \mu(\varphi)$ , où  $\varphi$  décrit l'ensemble  $\mathcal{C}_b(\mathbb{T})$  des applications continues bornées de  $\mathbb{T}$  dans  $\mathbb{R}$ , soit de la *topologie étroite*, qui est la topologie la plus faible rendant semicontinues inférieurement les applications  $\mu \mapsto \varphi(\mu)$ , où  $\varphi$  décrit l'ensemble des applications semicontinues inférieurement et bornées de  $\mathbb{T}$  dans  $\mathbb{R}$ . Une mesure  $\mu \in M(\mathbb{T})$  est dite de *Radon* si, pour tout  $B \in \mathfrak{B}(\mathbb{T})$ , on a  $\mu(B) = \sup\{\mu(K); K \in \mathfrak{K}, K \subset B\}$ . L'espace  $\mathbb{T}$  est dit de *Radon* si toute mesure sur  $\mathbb{T}$  est de Radon. Le lecteur pourra consulter [Top70b, GP84, Bog98] pour des compléments sur les mesures boréliennes.

Dans tout ce qui suit,  $(\Omega, \mathcal{F}, P)$  est un espace probabilisé et  $\mathbb{T}$  un espace topologique. On note respectivement  $\pi_\Omega$  et  $\pi_{\mathbb{T}}$  les projections canoniques de  $\Omega \times \mathbb{T}$  sur  $\Omega$  et  $\mathbb{T}$ . Si  $\mu \in M(\Omega \times \mathbb{T})$ , on note  $(\pi_\Omega)_\#(\mu)$  la mesure  $\mu(\cdot \times \mathbb{T})$  image de  $\mu$  par  $\pi_\Omega$ . De même,  $(\pi_{\mathbb{T}})_\#(\mu)$  désigne la mesure  $\mu(\Omega \times \cdot) \in M(\mathbb{T})$ .

**Mesures de Young** On appelle *mesures de Young (de base P)* les mesures de probabilité sur  $\Omega \times \mathbb{T}$  de marge P sur  $\Omega$ , c'est-à-dire les éléments de l'ensemble

$$\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{T}) := \{\mu \in P(\Omega \times \mathbb{T}); \forall A \in \mathcal{F}, \mu(A \times \mathbb{T}) = P(A)\}.$$

Lorsque aucune ambiguïté n'est à craindre, on n'écrira qu'une partie des informations :  $\mathcal{Y}(\mathbb{T})$  ou même simplement  $\mathcal{Y}$ .

Les mesures de Young généralisent les probabilités aléatoires et les variables aléatoires. Examinons quelques sous-ensembles particuliers :

- $\mathcal{Y}_d(\mathbb{T})$  : c'est l'ensemble des mesures de Young *désintégrables*, c'est-à-dire de la forme

$$\mu(A \times B) = \int_A \mu_\omega(B) dP(\omega)$$

pour une application mesurable  $\omega \mapsto \mu_\omega$  de  $\Omega$  dans  $P(\mathbb{T})$  (muni de la topologie étroite), à laquelle on peut identifier  $\mu$  (la probabilité P étant fixée). Si  $\mathbb{T}$  est un espace de Radon et si  $(\Omega, \mathcal{F}, P)$  est complet, alors on a  $\mathcal{Y}_d = \mathcal{Y}$  d'après le théorème de désintégration (voir [HJ71, Val72, Val73]).

- $L^0(\mathbb{T})$  : si  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{T}, \mathfrak{B}(\mathbb{T}))$  est une application mesurable, on appelle *élément aléatoire* de  $\mathbb{T}$  (défini sur  $\Omega$ ) la classe d'équivalence de  $X$  modulo l'égalité P-p.s. Cette classe est encore notée  $X$ . L'ensemble des éléments aléatoires de  $\mathbb{T}$  est noté  $L^0(\mathbb{T})$ . À chaque élément aléatoire

$X$  de  $\mathbb{T}$  défini sur  $\Omega$ , on peut associer la *mesure de Young dégénérée*  $\underline{\delta}_X : \omega \mapsto \delta_{X(\omega)}$ . Ainsi  $L^0(\mathbb{T})$  peut être vu comme un sous-ensemble de  $\mathcal{Y}_d(\mathbb{T})$ . Contrairement aux distributions, qui voient les fonctions de  $\Omega$  dans  $\mathbb{T}$  comme des mesures vectorielles absolument continues par rapport à la mesure de Lebesgue (ce qui suppose  $\Omega = \mathbb{R}^N$  et  $T$  vectoriel), ici les mêmes fonctions sont identifiées à des mesures de probabilité de marge  $P$  particulières sur  $\Omega \times \mathbb{T}$ .

- $P(\mathbb{T})$  : L'ensemble  $P(\mathbb{T})$  peut être vu comme un sous-ensemble de  $\mathcal{Y}_d$  en associant à chaque  $\nu \in P(\mathbb{T})$  la probabilité aléatoire constante  $\omega \mapsto \nu$ . La mesure de Young correspondante est  $\mu = P \otimes \nu$ .

**Topologies stables** On définit des topologies dites *stables*<sup>1</sup> sur  $\mathcal{Y}(\mathbb{T})$  de manière similaire à celle employée pour définir les topologies faible ou étroite sur  $P(\mathbb{T})$ , en remplaçant l'ensemble test de fonctions continues ou semicontinues par un ensemble d'*intégrandes* sur  $\Omega \times \mathbb{T}$ , c'est-à-dire de fonctions mesurables de  $\Omega \times \mathbb{T}$  dans  $[-\infty, +\infty]$ . Précisons d'abord qu'un intégrande  $f$  sur  $\Omega \times \mathbb{T}$  est dit *continu* ou *de Carathéodory* si  $f(\omega, \cdot)$  est continu pour tout  $\omega \in \Omega$ . De même, l'intégrande  $f$  est dit *semicontinu supérieurement* (resp. *inférieurement*) si  $f(\omega, \cdot)$  est semicontinu supérieurement (resp. inférieurement) pour tout  $\omega \in \Omega$ . On obtient naturellement quatre topologies stables.

- $\tau_{\mathcal{Y}}^S$  est la topologie la moins fine qui rende semicontinues supérieurement les applications  $\mu \mapsto \mu(f)$ , où  $f$  décrit l'ensemble des intégrandes bornés semicontinus supérieurement. Autrement dit, une suite généralisée  $(\mu^\alpha)$  dans  $\mathcal{Y}$  converge vers une limite  $\mu \in \mathcal{Y}$  pour  $\tau_{\mathcal{Y}}^S$  si et seulement si  $\limsup_\alpha \mu^\alpha(f) \leq \mu^\infty(f)$  pour tout intégrande  $f$  borné semicontinu supérieurement.
- $\tau_{\mathcal{Y}}^M$  est la topologie la moins fine qui rende continues les applications  $\mu \mapsto \mu(f)$ , où  $f$  décrit l'ensemble des intégrandes de Carathéodory bornés sur  $\Omega \times \mathbb{T}$ .
- $\tau_{\mathcal{Y}}^N$  est la topologie la moins fine qui rende semicontinues supérieurement les applications  $\mu \mapsto \mu(f)$ , où  $f$  décrit l'ensemble des intégrandes de la forme  $f(\omega, x) = \mathbb{1}_A(\omega) \times g(x)$ , où  $A \in \mathcal{F}$  et  $g : \mathbb{T} \rightarrow \mathbb{R}$  est une fonction bornée semicontinue supérieurement.
- $\tau_{\mathcal{Y}}^W$  est la topologie la moins fine qui rende semicontinues supérieurement les applications  $\mu \mapsto \mu(f)$ , où  $f$  décrit l'ensemble des intégrandes de la

---

<sup>1</sup>Cette terminologie provient de Rényi [Rén63, RR58], qui a observé que la convergence dans cette topologie ressemble à la convergence étroite mais possède en plus d'agréables propriétés de stabilité (voir par exemple [CRdFV04, Proposition 9.2.2]). Ce n'est pas une terminologie standard (il n'y en a pas pour l'instant), mais elle est moins ambiguë que les autres utilisées.

forme  $f(\omega, x) = \mathbf{1}_A(\omega) \times g(x)$ , où  $A \in \mathcal{F}$  et  $g : \mathbb{T} \rightarrow \mathbb{R}$  est une fonction bornée continue.

Si une suite généralisée  $(\mu^\alpha)$  dans  $\mathcal{Y}$  converge vers une limite  $\mu \in \mathcal{Y}$  pour  $\tau_{\mathcal{Y}}^\times$ ,  $\times = S, M, N, W$ , on dit que  $(\mu^\alpha)$  converge de manière  $\times$ -stable vers la mesure  $\mu$  et on écrit  $\mu^\alpha \xrightarrow{\times\text{-stable}} \mu^\infty$ .

Les topologies stables sont donc ordonnées comme ci-dessous, où les flèches représentent des implications.

$$\begin{array}{ccc} \tau_{\mathcal{Y}}^S\text{-convergence} & \longrightarrow & \tau_{\mathcal{Y}}^M\text{-convergence} \\ \downarrow & & \downarrow \\ \tau_{\mathcal{Y}}^N\text{-convergence} & \longrightarrow & \tau_{\mathcal{Y}}^W\text{-convergence} \end{array}$$

(pour se souvenir des symboles, le lecteur pourra retenir que M et N représentent des topologies intermédiaires, et que la topologie N-stable est construite comme la topologie étroite, en anglais “narrow”). Le problème qui se pose immédiatement (et dont la solution a des conséquences pratiques importantes) est de savoir dans quels cas les implications réciproques sont vraies. Un premier résultat est donné par le théorème “portmanteau” ci-dessous (c’est une sorte de version paramétrée du célèbre théorème du même nom pour les mesures sur un espace topologique).

Une caractérisation de la topologie  $\tau_{\mathcal{Y}}^S$  fait appel à la notion d’*ensemble aléatoire*, c’est-à-dire d’élément de  $\mathcal{F} \otimes \mathfrak{B}(\mathbb{T})$ . Si  $G$  est un ensemble aléatoire et  $\omega \in \Omega$ , on note  $G(\omega)$  la section de  $G$  en  $\omega$ . On appelle *ensemble aléatoire ouvert* (resp. *fermé*, *compact*) tout ensemble aléatoire  $G$  vérifiant  $G(\omega) \in \mathfrak{G}$  (resp.  $G(\omega) \in \mathfrak{F}$ ,  $G(\omega) \in \mathfrak{K}$ ) pour tout  $\omega \in \Omega$ . On note respectivement  $\underline{\mathfrak{G}}$  (resp.  $\underline{\mathfrak{F}}$ ,  $\underline{\mathfrak{K}}$ ) l’ensemble des ensembles aléatoires ouverts (resp. fermés, compacts).

Rappelons qu’un espace topologique séparé  $\mathbb{T}$  est dit *souslinien* s’il existe un espace polonais  $\mathbb{S}$  et une surjection continue de  $\mathbb{S}$  dans  $\mathbb{T}$ . Si, de plus, cette surjection peut être choisie bijective, on dit que  $\mathbb{T}$  est *lusinien*. Les propriétés des espaces sousliniens que nous utiliserons sont démontrées dans [Sch73]. De nombreux espaces usuels de l’analyse sont lusiniens : outre bien sûr tous les espaces de Banach séparables, il y a aussi par exemple les espaces de distributions  $\mathcal{E}'$ ,  $\mathcal{S}'$ ,  $\mathcal{D}'$ , l’espace  $\mathcal{H}(\mathbb{C})$  des fonctions holomorphes, ou le dual d’un espace de Banach séparable, muni de la topologie faible\* (voir [Sch73, pages 112–117] pour de nombreux autres exemples). L’espace  $\mathbb{T}$  est dit *cosmique* (“Continuous-image Of Separable Metric”, [Mic66]) s’il existe un espace métrique séparable  $\mathbb{S}$  et une application continue surjective de  $\mathbb{S}$  dans  $\mathbb{T}$ . La classe des espaces cosmiques contient donc à la fois celle des espaces sousliniens et celle des espaces métriques séparables. Le lecteur trouvera des caractérisations des espaces cosmiques dans [Mic66, Cal82, Cal84].

Tout espace cosmique régulier est complètement régulier. Signalons que tous les résultats de convergence de mesure de Young que nous donnons dans le cas où  $\mathbb{T}$  est cosmique restent vrais également si  $\mathbb{T}$  est un espace métrique non nécessairement séparable, en supposant tout de même que la mesure de Young limite vérifie  $\mu(\Omega \times \mathbb{T}_0) = 1$  pour une partie séparable  $\mathbb{T}_0$  de  $\mathbb{T}$ .

Le théorème ci-dessous est une extension du théorème “portmanteau” classique pour les mesures boréliennes sur un espace topologique (voir par exemple [Top70b]).

**Théorème 1.2.1 (théorème “portmanteau”)** *Soit  $(\mu^\alpha)$  une suite généralisée dans  $\mathcal{Y}$  et soit  $\mu \in \mathcal{Y}$ .*

A) *Les affirmations A1, A2 et A3 ci-dessous sont équivalentes.*

A1.  $\mu^\alpha \xrightarrow{\text{S-stable}} \mu.$

A2.  $\liminf_\alpha \mu^\alpha(f) \geq \mu(f)$  pour tout intégrande semicontinu inférieurement  $f$  à valeurs dans  $[0, +\infty]$ .

A3.  $\liminf_\alpha \mu^\alpha(G) \geq \mu(G)$  pour tout  $G \in \mathfrak{G}$ .

B) *Soit  $\mathcal{C}$  un ensemble de fonctions positives bornées  $\mathcal{F}$ -mesurables, stable pour la multiplication de deux éléments, contenant la fonction 1 engendrant  $\mathcal{F}$  (e.g.  $\mathcal{C}$  peut consister en l'ensemble des fonctions indicatrices d'un  $\pi$ -système contenant  $\Omega$  qui engendre  $\mathcal{F}$ ). Les affirmations B1 et B2 ci-dessous sont équivalentes. De plus, si  $\mathbb{T}$  est cosmique régulier (donc complètement régulier), et si  $\mathcal{D}$  est une famille filtrante croissante de semidistances sur  $\mathbb{T}$  qui induit la topologie de  $\mathbb{T}$ , alors les affirmations B1, B2 et B3 sont équivalentes.*

B1.  $\mu^\alpha \xrightarrow{\text{W-stable}} \mu.$

B2.  $\liminf_\alpha \mu^\alpha(g \otimes \mathbf{1}_G) \geq \mu(g \otimes \mathbf{1}_G)$  pour tout  $g \in \mathcal{C}$  et tout  $G \in \mathfrak{G}$ .

B3.  $\lim_\alpha \mu^\alpha(g \otimes f) = \mu(g \otimes f)$  pour tout  $g \in \mathcal{C}$ , pour tout  $d \in \mathcal{D}$  et pour toute application  $f : \mathbb{T} \rightarrow [0, 1]$  lipschitzienne par rapport à  $d$ .

(Pour  $g : \Omega \rightarrow \mathbb{R}$  et  $f : \mathbb{T} \rightarrow \mathbb{R}$ , on note  $g \otimes f$  l'application  $(\omega, x) \mapsto g(\omega)f(x)$ .)

C) *Supposons de plus que  $\mathbb{T}$  soit un espace souslinien métrisable. Soit  $d$  une distance compatible avec la topologie de  $\mathbb{T}$ . Alors les affirmations A1 à B3 sont toutes équivalentes à l'affirmation C1 ci-dessous.*

C1.  $\lim_\alpha \mu^\alpha(f \otimes g) = \mu(f \otimes g)$  pour tout  $f \in \mathcal{C}$  et pour toute fonction bornée  $g : \mathbb{T} \rightarrow \mathbb{R}$  qui est lipschitzienne par rapport à  $d$ .

*En particulier, dans ce cas, les topologies stables  $\tau_{\mathcal{Y}}^\times$  ( $\times = \text{S, M, N, W}$ ) coïncident.*

L'implication B1  $\Rightarrow$  A2 est parfois appelée *théorème de semicontinuité*. Sous des hypothèses plus générales que dans la partie C, on peut donner des résultats partiels d'équivalence entre les topologies stables  $\tau_{\mathcal{Y}}^{\times}$ . Par exemple, si  $\mathbb{T}$  est complètement régulier et si la marge de  $\mu$  sur  $\mathbb{T}$  est  $\tau$ -régulière, alors on a  $\mu^{\alpha} \xrightarrow{\text{W-stable}} \mu \Rightarrow \mu^{\alpha} \xrightarrow{\text{N-stable}} \mu$ . Mais le problème de la généralisation du théorème de semicontinuité, c'est-à-dire

*le problème de savoir si  $\tau_{\mathcal{Y}}^{\text{W}}$  coïncide avec  $\tau_{\mathcal{Y}}^{\text{S}}$  dans des espaces plus généraux que les espaces sousliniens métrisables (par exemple les sousliniens réguliers) est toujours ouvert.*

Une autre approche consiste étudier les compacts de  $\tau_{\mathcal{Y}}^{\text{S}}$  (voir la section 1.3), car, sur ces compacts, les topologies  $\tau_{\mathcal{Y}}^{\text{W}}$  et  $\tau_{\mathcal{Y}}^{\text{S}}$  coïncident. On trouvera dans le tableau 1 page 17 une liste de résultats sur le théorème de semicontinuité.

Signalons encore un cas où l'on peut, sans compacité, obtenir sous l'hypothèse de convergence W-stable la convergence d'intégrales de fonctions qui ne sont pas de la forme  $f \otimes g$ . Il nous faut d'abord introduire de nouvelles notations. Soit  $d$  une semidistance continue sur  $\mathbb{T}$ . On note  $\text{BL}_1(\mathbb{T}, d)$  l'ensemble des fonctions  $f : \mathbb{T} \rightarrow [-1, 1]$  lipschitziennes pour  $d$  de module de Lipschitz inférieur ou égal à 1 (c'est-à-dire  $\|f\|_{\text{L}}(d) := \sup_{d(x,y)>0} \frac{|f(x)-f(y)|}{d(x,y)} \leq 1$ ). On note  $\underline{\text{BL}}_1(\Omega, \mathbb{T}, d)$  l'espace des intégrandes  $f$  tels que  $f(\omega, \cdot) \in \text{BL}_1(\mathbb{T}, d)$  pour tout  $\omega \in \Omega$  et  $\underline{\text{BL}}'_1(\Omega, \mathbb{T}, d)$  l'ensemble des éléments  $f$  de  $\underline{\text{BL}}_1(\Omega, \mathbb{T}, d)$  qui sont de la forme  $f = \sum_{i=1}^n \mathbb{1}_{A_i} \otimes g_i$ , où  $(A_i)_{1 \leq i \leq n}$  est une partition mesurable de  $\Omega$  (qui dépend de  $f$ ).

**Proposition 1.2.2 (convergence W-stable et intégrandes lipschitziens)** *Soit  $d$  une semidistance continue sur  $\mathbb{T}$ . Soit  $(\mu^{\alpha})_{\alpha \in \mathbb{A}}$  une suite généralisée dans  $\mathcal{Y}$  qui converge de manière W-stable vers une limite  $\mu \in \mathcal{Y}$ . On suppose qu'il existe une partie séparable  $\mathbb{T}_0$  de  $\mathbb{T}$  telle que  $\mu(\Omega \times \mathbb{T}_0) = 1$ . Alors on a*

$$\lim_{\alpha} \mu^{\alpha}(\varphi \circ f) = \mu^{\infty}(\varphi \circ f)$$

*pour tout intégrande  $f \in \underline{\text{BL}}_1(\Omega, \mathbb{T}, d)$  et pour toute fonction continue  $\varphi : [-1, 1] \rightarrow \mathbb{R}$ .*

**Autres topologies stables, mesures de Young d'ordre  $p$**  On obtient des variantes intéressantes et naturelles des topologies stables en remplaçant la condition de bornitude des intégrandes test par une condition de croissance. Par exemple, si  $p \in [1, +\infty[$  et si  $\mathcal{D}$  est une famille de semidistances sur  $\mathbb{T}$  qui définit la topologie de  $\mathbb{T}$ , on peut considérer les intégrandes  $f$  tels qu'il existe  $d \in \mathcal{D}$  pour lequel on ait  $|f(\omega, \cdot)| \leq 1 + d(x_0, \cdot)^p$ , où  $x_0$  est un élément fixé de  $\mathbb{T}$  (la définition ne dépend pas du choix de  $x_0$ ). On

obtient ainsi, avec des notations évidentes, les topologies stables  $\tau_{\mathcal{Y},\mathcal{D}}^{\times,p}$ , pour  $\times = S, M, N, W$ . Le cadre naturel d'étude de  $\tau_{\mathcal{Y},\mathcal{D}}^{\times,p}$  est le sous-espace

$$\mathcal{Y}_{\mathcal{D}}^p = \{\mu \in \mathcal{Y}; (\forall d \in \mathcal{D}) \int_{\Omega \times \mathbb{T}} d(x_0, \cdot)^p d\mu < \infty\}$$

des *mesures de Young d'ordre  $p$*  (ou *mesures de Young  $p$ -intégrables*) relativement à  $\mathcal{D}$ . Dans le cas où  $\mathbb{T} \subset \mathbb{R}^n$  et  $\mathcal{D}$  a pour unique élément la distance associée à la norme, l'espace  $(\mathcal{Y}_{\mathcal{D}}^p, \tau_{\mathcal{Y},\mathcal{D}}^{\times,p})$  est considéré dans des travaux aussi différents que (notamment) [Bal89b, KP94, PV95, Rou97, DM02, AP03].

Introduisons maintenant une notion d'intégrabilité uniforme qui nous servira pour caractériser la  $\tau_{\mathcal{Y},\mathcal{D}}^{\times,p}$ -convergence. On dit qu'une partie  $\mathfrak{Y} \subset \mathcal{Y}_{\mathcal{D}}^p$  est *uniformément  $p$ -intégrable* si, pour tout  $d \in \mathcal{D}$ ,

$$\lim_{R \rightarrow +\infty} \sup_{\mu \in \mathfrak{Y}} \int_{\Omega \times \{d(a, \cdot) > R\}} d(a, x)^p d\mu(\omega, x) = 0$$

pour un  $a \in \mathbb{T}$  (et donc pour tout  $a \in \mathbb{T}$ ). On dit qu'une suite généralisée  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  dans  $\mathcal{Y}_{\mathcal{D}}^p$  est *asymptotiquement uniformément  $p$ -intégrable* si, pour tout  $d \in \mathcal{D}$ ,

$$\lim_{R \rightarrow +\infty} \limsup_{\alpha} \int_{\Omega \times \{d(a, \cdot) > R\}} d(a, x)^p d\mu^\alpha(\omega, x) = 0$$

pour un  $a \in \mathbb{T}$  (donc pour tout  $a \in \mathbb{T}$ ). Si  $\mathbb{A} = \mathbb{N}$ , la suite  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  est asymptotiquement uniformément  $p$ -intégrable si et seulement si elle est uniformément  $p$ -intégrable.

**Proposition 1.2.3 (Caractérisations de la convergence pour  $\tau_{\mathcal{Y},\mathcal{D}}^{\times,p}$ )** *On suppose que  $\mathbb{T}$  est complètement régulier et que sa topologie est définie par une famille  $\mathcal{D}$  de semidistances. Soient  $p \geq 1$ ,  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  une suite généralisée dans  $\mathcal{Y}_{\mathcal{D}}^p$  et  $\mu \in \mathcal{Y}_{\mathcal{D}}^p$ . On suppose qu'il existe une partie séparable  $\mathbb{T}_0$  de  $\mathbb{T}$  telle que  $\mu(\Omega \times \mathbb{T}_0) = 1$ . Fixons un élément  $a$  de  $\mathbb{T}$ . Alors, pour  $\times = S, M, N, W$ , les assertions suivantes sont équivalentes.*

1.  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  converge vers  $\mu$  pour  $\tau_{\mathcal{Y},\mathcal{D}}^{\times,p}$ .
2.  $\mu^\alpha \xrightarrow{\times\text{-stable}} \mu$  et  $\lim_{\alpha} \int_{\Omega \times \mathbb{T}} d(a, x)^p d\mu^\alpha(x, \omega) = \int_{\Omega \times \mathbb{T}} d(a, x)^p d\mu(x, \omega)$  pour tout  $d \in \mathcal{D}$ .
3.  $\mu^\alpha \xrightarrow{\times\text{-stable}} \mu$  et  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  est asymptotiquement uniformément  $p$ -intégrable.

À l'aide du théorème 1.2.1, on en déduit le résultat suivant.

**Corollaire 1.2.4** *Si  $\mathbb{T}$  est métrisable souslinien, les topologies  $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p}$  coïncident ( $\times = S, M, N, W$ ).*

Si  $\mathbb{T} = \mathbb{E}$  est un espace vectoriel normé séparable et si  $\mathcal{D}$  ne contient que la distance associée à la norme de  $\mathbb{E}$ , alors  $L_{(\mathbb{T}, \mathcal{D})}^p := \mathcal{Y}_{\mathcal{D}}^p \cap L^0(\Omega, \mathcal{F}, P; \mathbb{T})$  n'est rien d'autre que l'espace  $L^p(\Omega, \mathcal{F}, P; \mathbb{E})$  des fonctions Bochner intégrables de  $(\Omega, \mathcal{F}, P)$  dans  $\mathbb{E}$ . On verra plus loin (corollaire 1.5.7 page 24) que la trace sur  $L^p(\Omega, \mathcal{F}, P; \mathbb{E})$  de la topologie  $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p}$  est la topologie forte de  $L^p(\Omega, \mathcal{F}, P; \mathbb{E})$ .

On retrouve de même les espaces d'Orlicz en remplaçant dans ce qui précède  $d(x_0, \cdot)^p$  par  $\varphi(d(x_0, \cdot))$ , où  $\varphi$  est une fonction de Young [AP04].

**Densité des fonctions dans les mesures de Young** C'est l'une des propriétés les plus importantes des topologies stables, et en fait les mesures de Young ont été historiquement construites comme limites de fonctions. Le résultat ci-dessous a donc une longue histoire, de [You37] à [Bal84b]. Il est légèrement généralisé ici, l'espace considéré n'étant pas nécessairement régulier.

**Théorème 1.2.5 (théorème de densité)** *Supposons que  $\mathbb{T}$  soit un espace de Radon dont les parties compactes  $\mathbb{T}$  sont métrisables (par exemple,  $\mathbb{T}$  est un espace souslinien). Supposons de plus la probabilité  $P$  sans atome. Alors  $L^0$  est dense dans  $\mathcal{Y}$  pour  $\tau_{\mathcal{Y}}^{\times}$ ,  $\times = S, M, N, W$ .*

Après une réduction au cas métrisable compact, la preuve dans [CRdFV04] consiste à approcher toute mesure de Young non dégénérée par une suite de fonctions qui oscille sur des parties mesurables de plus en plus petites.

Une autre approche utilisant le théorème de convergence des martingales est présentée dans [RdFZ02].

On obtient une version analogue pour les mesures de Young  $p$ -intégrables.

**Corollaire 1.2.6** *On suppose que  $\mathbb{T}$  est un espace de Radon à compacts métrisables (par exemple  $\mathbb{T}$  est un espace souslinien régulier), et que  $P$  n'a pas d'atome. Soit  $\mathcal{D}$  une famille de semidistances qui définit la topologie  $\mathbb{T}$ . Alors, pour tout  $p \geq 1$ ,  $L_{(\mathbb{T}, \mathcal{D})}^p$  est dense dans  $\mathcal{Y}_{\mathcal{D}}^p$  pour la topologie  $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p}$ ,  $\times = S, M, N, W$ .*

Comme pour le théorème 1.2.5, on se ramène au cas métrisable compact. Dans ce cas, on a  $\mathcal{Y}_{\mathcal{D}}^p = \mathcal{Y}$ ,  $L_{(\mathbb{T}, \mathcal{D})}^p = L^0$  et  $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p} = \tau_{\mathcal{Y}}^{\times}$ . Le théorème 1.2.5 donne alors le résultat.



**Généralisation aux mesures sur  $\Omega \times \mathbb{T}$**  La définition des topologies stables n'utilise ni le fait que les éléments de  $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{T})$  ont pour masse 1, ni qu'ils ont même marge  $P$ . On peut donc étendre ces topologies à  $M(\Omega \times \mathbb{T})$  en remplaçant  $\mathcal{Y}$  par  $M(\Omega \times \mathbb{T})$  dans les définitions. Les topologies correspondantes sont notées  $\tau_M^\times$ ,  $\times = S, M, N, W$ . C'est le cadre retenu dans les études [Bal01] (qui porte sur  $\tau_M^W$ ) et [RdF03] (qui étudie la compacité pour  $\tau_M^S$ ).

Le théorème “portmanteau” (théorème 1.2.1) admet encore une formulation dans ce cadre général ([RdF03] et [CRdFV04, Remark 2.1.5]) : il suffit d'ajouter, dans les affirmations A3 et B2, que l'on a  $\lim_\alpha \mu_\alpha(\Omega \times \mathbb{T}) = \mu(\Omega \times \mathbb{T})$  et, dans les parties B et C, que l'ensemble  $\mathcal{C}$  contient toutes les fonctions  $\mathbb{1}_A$ ,  $A \in \mathcal{F}$ .<sup>2</sup> En particulier, la topologie  $\tau_M^S$  peut être définie comme la topologie la moins fine qui rende semicontinues inférieurement les applications  $\mu \mapsto \mu(G)$  ( $G \in \mathfrak{G}$ ) et qui rende continue l'application  $\mu \mapsto \mu(\Omega \times \mathbb{T})$ . Ainsi,  $\tau_M^S$  est un cas particulier de la *w-topologie* définie de manière abstraite par F. Topsøe [Top70a].

Remarquons que la trace sur  $M(\mathbb{T})$  de  $\tau_M^\times$  est la topologie étroite si  $\times \in \{S, N\}$  et la topologie faible si  $\times \in \{M, W\}$ . D'autre part, la trace sur  $M(\Omega)$  de  $\tau_M^\times$  est la *s-topologie* [Top70a], c'est-à-dire la topologie la moins fine telle que, pour chaque  $A \in \mathcal{F}$ , l'application  $\mu \mapsto \mu(A)$  soit continue.

Nous reprendrons ce cadre général page 11 dans la présentation des critères de compacité pour les topologies stables.

### 1.3 Critères de compacité

Comme cela a été annoncé dans l'introduction, il existe des conditions suffisantes simples à vérifier pour qu'une partie de  $\mathcal{Y}$  soit compacte pour  $\tau_Y^\times$ . Nous nous contenterons de caractériser les compacts pour  $\tau_Y^S$ , puisqu'ils sont en même temps compacts pour  $\tau_Y^\times$ ,  $\times = M, N, W$ . Par ailleurs, les critères de compacité pour  $\tau_{Y, \mathcal{D}}^{\times, p}$  se déduisent facilement de ceux pour  $\tau_Y^\times$ , d'après la proposition 1.3.1 ci-dessous.

Rappelons qu'un sous-ensemble  $\mathcal{K}$  d'un espace topologique  $\mathbb{T}$  est *compact pour les suites généralisées dans  $\mathbb{T}$*  si toute suite généralisée d'éléments de  $\mathcal{K}$  admet une sous-suite généralisée qui converge dans  $\mathbb{T}$ . De manière équivalente,  $\mathcal{K}$  est compact pour les suites généralisées si toute suite généralisée universelle d'éléments de  $\mathcal{K}$  converge (voir [Kel55] pour les définitions et propriétés des sous-suites généralisées et des suites généralisées universelles).

<sup>2</sup>On peut [RdF03, Theorem 3.2] remplacer cette condition sur  $\mathcal{C}$  par l'hypothèse que l'ensemble  $\{(\pi_\Omega)_\#(\mu^\alpha); \alpha \in \mathbb{A}\}$  est équicontinu (voir la définition page 13)

On dit que  $\mathcal{K}$  est *relativement compact* s'il est contenu dans une partie compacte de  $\mathbb{T}$ . Donc toute partie relativement compacte de  $\mathbb{T}$  est compacte pour les suites généralisées. L'implication réciproque est vraie si  $\mathbb{T}$  est régulier (voir la démonstration dans [PV95] ou dans [OW98]).

Le résultat suivant se déduit facilement de la proposition 1.2.3.

**Proposition 1.3.1 (Parties compactes de  $\mathcal{Y}_{\mathcal{D}}^p$ )** *On suppose que  $\mathbb{T}$  est complètement régulier. Soit  $\mathcal{D}$  une famille de semidistances qui définit la topologie de  $\mathbb{T}$ . Soient  $p \geq 1$  et  $\mathfrak{Y} \subset \mathcal{Y}_{\mathcal{D}}^p$ . Les assertions suivantes sont équivalentes, pour  $\times = S, M, N, W$ .*

1. *L'ensemble  $\mathfrak{Y}$  est compact pour les suites généralisées dans  $(\mathcal{Y}_{\mathcal{D}}^p, \tau_{\mathfrak{Y}; \mathcal{D}}^{\times, p})$ .*
2. *L'ensemble  $\mathfrak{Y}$  est compact pour les suites généralisées dans  $(\mathcal{Y}_{\mathcal{D}}^p, \tau_{\mathfrak{Y}}^{\times})$  et uniformément  $p$ -intégrable par rapport à  $\mathcal{D}$ .*

Le résultat ci-dessous est très utile pour les applications. On dit qu'une partie  $\mathcal{K}$  d'un espace topologique  $\mathbb{T}$  est *séquentiellement relativement compacte* si toute suite d'éléments de  $\mathcal{K}$  admet une sous-suite qui converge dans  $\mathbb{T}$  (cette terminologie n'est pas tout à fait cohérente avec celle que nous avons introduite plus haut, mais elle est conforme à l'usage établi).

**Lemme 1.3.2 (quand la compacité entraîne la compacité séquentielle)** *Soit  $\mathfrak{M}$  une partie relativement compacte de  $\mathcal{Y}$  pour  $\tau_{\mathfrak{Y}}^S$ . Supposons qu'il existe une topologie séparable métrisable moins fine que la topologie de  $\mathbb{T}$  mais qui possède les mêmes boréliens que cette topologie (c'est le cas si, par exemple,  $\mathbb{T}$  est souslinien régulier). Alors  $\mathfrak{M}$  est séquentiellement relativement compact.*

Un autre résultat très utile lié à la compacité séquentielle est la version pour les mesures de Young du théorème de Komlós, due à Balder [Bal89a, Bal90]. Nous ne la présentons pas ici, mais nous l'utiliserons page 33, pour démontrer qu'un certain processus à valeurs mesures, obtenu comme limite stable de processus adaptés, est adapté.

## Critères de Topsøe et de Jacod et Mémin

Nous nous plaçons ici dans le cadre général de l'espace  $(M(\Omega \times \mathbb{T}), \tau_M^S)$ . Nous donnons d'abord une condition nécessaire et suffisante de compacité pour les suites généralisées (critère de Topsøe), dont nous déduisons d'autres critères dans des cas particuliers.

On appelle *pavage* sur  $\Omega \times \mathbb{T}$  tout ensemble non vide de parties de  $\Omega \times \mathbb{T}$ . Énumérons maintenant les propriétés des pavages  $\underline{\mathfrak{G}}(\mathbb{T})$  et  $\underline{\mathfrak{K}}(\mathbb{T})$  dont nous

aurons besoin. La présentation ci-dessous suit celle de [Top70a], pour faciliter la comparaison. On dit que  $\mathbb{T}$  est *sous-métrisable* s'il existe sur  $\mathbb{T}$  une topologie métrisable moins fine. Cela revient à dire qu'il existe une famille dénombrable de fonctions continues à valeurs réelles sur  $\mathbb{T}$  qui sépare les points de  $\mathbb{T}$ .

**Lemme 1.3.3 (Propriétés des pavages  $\underline{\mathcal{G}}$  et  $\underline{\mathcal{K}}$ )** *On suppose que  $\mathbb{T}$  est souslinien sous-métrisable et que  $(\Omega, \mathcal{F})$  est universellement complet.*

I.  $\underline{\mathcal{K}}$  contient  $\emptyset$  et est stable par réunion finie et par intersection dénombrable.

II.  $\underline{\mathcal{G}}$  contient  $\emptyset$  et est stable par union finie et par intersection finie (en fait,  $\underline{\mathcal{G}}$  est également stable par union dénombrable, mais nous n'en aurons pas besoin).

III.  $K \setminus G \in \underline{\mathcal{K}}$  pour tout  $K \in \underline{\mathcal{K}}$  et pour tout  $G \in \underline{\mathcal{G}}$ ,

IV.  $\underline{\mathcal{G}}$  sépare les ensembles dans  $\underline{\mathcal{K}}$ , c'est-à-dire que, pour toute paire  $\{K_1, K_2\}$  d'éléments disjoints de  $\underline{\mathcal{K}}$ , on peut trouver une paire  $\{G_1, G_2\}$  d'éléments disjoints de  $\underline{\mathcal{G}}$  telle que  $K_1 \subset G_1$  et  $K_2 \subset G_2$ .

V'. Soit  $\mathfrak{M}$  une partie de  $M(\Omega \times \mathbb{T})$  telle que  $(\pi_\Omega)_\#(\mathfrak{M})$  soit relativement compact pour la  $s$ -topologie sur  $M(\Omega)$ . Alors  $\mathfrak{M}$  est uniformément  $\sigma$ -régulier sur  $\underline{\mathcal{G}}$  au point  $\emptyset$  par rapport à  $\underline{\mathcal{K}}$ , c'est-à-dire que, pour toute famille dénombrable  $(K_i)_{i \in I}$  d'éléments de  $\underline{\mathcal{K}}$  filtrante décroissante vers  $\emptyset$ , on a

$$\inf_{i \in I} \sup_{\mu \in \mathfrak{M}} \inf_{G \in \underline{\mathcal{G}}, G \supset K_i} \mu(G) = 0$$

(on dit que  $(K_i)_{i \in I}$  est filtrante décroissante vers  $\emptyset$  si  $\bigcap_{i \in I} K_i = \emptyset$  et si, pour tout  $i \in I$  et pour tout  $j \in I$ , il existe  $k \in I$  tel que  $K_k \subset K_i \cap K_j$ ).

**Remarque 1.3.4** Le pavage  $\underline{\mathcal{K}}$  ne vérifie pas la propriété V de [Top70a], c'est-à-dire que  $\underline{\mathcal{K}}$  n'est pas semi-compact (un pavage  $\mathcal{C}$  est dit *semi-compact* si, pour toute famille dénombrable d'éléments de  $\mathcal{C}$  ayant une intersection vide, il existe une sous-famille finie ayant la même intersection). Toutefois la propriété plus faible V' suffit pour obtenir un résultat analogue à [Top70a, Theorem 4].

Avant de déduire du lemme 1.3.3 une adaptation du critère de compacité de Topsøe, il nous faut introduire encore quelques définitions.

On appelle *complétion universelle* de  $\mathcal{F}$  la tribu  $\mathcal{F}^* = \bigcap_{\mu \in M(\Omega)} \mathcal{F}_\mu$ , où  $\mathcal{F}_\mu$  est la tribu  $\mu$ -complétée de  $\mathcal{F}$ , c'est-à-dire la tribu engendrée par  $\mathcal{F}$  et les ensembles  $\mu$ -négligeables. Des notations telles que  $\underline{\mathcal{G}}^*$  ou  $\tau_M^{s*}$  signifient que l'on a remplacé  $\mathcal{F}$  par  $\mathcal{F}^*$  dans les définitions de  $\underline{\mathcal{G}}$  et de  $\tau_M^s$ .

Si  $\mathcal{C}$  et  $\mathcal{E}$  sont deux pavages sur  $\Omega \times \mathbb{T}$ , on dit que  $\mathcal{E}$  *domine*  $\mathcal{C}$  si tout élément de  $\mathcal{C}$  est contenu dans un élément de  $\mathcal{E}$ .

On déduit du lemme 1.3.3 le critère suivant [RdF03], adapté de celui de Topsøe ([Top70a, Corollary 2], voir aussi [Top74, OW98]).

**Théorème 1.3.5 (critère de Topsøe)** *On suppose que  $\mathbb{T}$  est souslinien sous-métrisable. Une partie  $\mathfrak{M}$  de  $M(\Omega \times \mathbb{T})$  est compacte pour les suites généralisées dans  $(M(\Omega \times \mathbb{T}), \tau_M^{S*})$  si et seulement si les conditions (i) et (ii) ci-dessous sont vérifiées :*

- (i) *L'ensemble  $(\pi_\Omega)_\#(\mathfrak{M})$  est compact pour les suites généralisées dans  $M(\Omega)$  muni de la  $s$ -topologie.*
- (ii) *Pour toute sous-famille  $\underline{\mathfrak{G}}'$  de  $\underline{\mathfrak{G}}^*$  qui domine  $\underline{\mathfrak{K}}$  et pour tout  $\epsilon > 0$ , il existe une sous-famille finie  $\underline{\mathfrak{G}}''$  de  $\underline{\mathfrak{G}}'$  telle que, pour tout  $\mu \in \mathfrak{M}$ , on puisse trouver  $G \in \underline{\mathfrak{G}}''$  vérifiant  $\mu(G^c) < \epsilon$ .*

Des critères de compacité pour les suites généralisées dans  $M(\Omega)$  muni de la  $s$ -topologie ont été donnés dans [Top70a, Gän71, AGK76]. En particulier, si une partie  $\mathfrak{M}$  de  $M(\Omega)$  est relativement compacte, elle est *équicontinue*, c'est-à-dire que, pour toute suite décroissante  $(A_n)$  dans  $\mathcal{F}$  telle que  $\bigcap_n A_n = \emptyset$ , on a  $\lim_n \sup_{\mu \in \mathfrak{M}} \mu(A_n) = 0$  [Gän71].

**Idée de la preuve du théorème 1.3.5.** La partie directe est exactement comme dans [Top74]. Supposons  $\mathfrak{M}$  compact pour les suites généralisées dans  $(M(\Omega \times \mathbb{T}), \tau_M^{S*})$ . La propriété (i) est trivialement vérifiée. Supposons que (ii) ne le soit pas : il existe une famille  $\underline{\mathfrak{G}}' = (G_K)_{K \in \underline{\mathfrak{K}}}$  d'éléments de  $\underline{\mathfrak{G}}$  telle que  $G_K \supset K$  pour chaque  $K \in \underline{\mathfrak{K}}$  et que, pour toute sous-famille finie  $\underline{\mathfrak{G}}''$  de  $\underline{\mathfrak{G}}'$ , il existe  $\mu \in \mathfrak{M}$  vérifiant  $\min\{\mu(G^c); G \in \underline{\mathfrak{G}}''\} > \epsilon$ . Mais alors la famille

$$(\mathcal{O}_{G_K})_{K \in \underline{\mathfrak{K}}} := (\{\mu \in M(\Omega \times \mathbb{T}); \mu(G_K^c) < \epsilon\})_{K \in \underline{\mathfrak{K}}}$$

est un recouvrement ouvert de  $M(\Omega \times \mathbb{T})$ . D'après [OW98, Proposition 1], on peut en extraire un sous-recouvrement fini de  $\mathfrak{M}$ , ce qui conduit à une contradiction.

Réciproquement, supposons que  $\mathfrak{M}$  vérifie (i) et (ii). Soit  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  une suite généralisée universelle dans  $\mathfrak{M}$ . D'après (i), on a  $\sup_{\mu \in \mathfrak{M}} \mu(\Omega \times \mathbb{T}) < +\infty$ . L'application  $\nu : \underline{\mathfrak{G}} \rightarrow [0, +\infty[$  définie par

$$(\forall G \in \underline{\mathfrak{G}}) \quad \nu(G) = \lim_{\alpha} \mu^\alpha(G)$$

est bornée d'après (i). La propriété  $V'$  du lemme 1.3.3 et l'hypothèse (i) entraînent que, pour toute famille dénombrable  $(K_i)_{i \in I}$  d'éléments de  $\underline{\mathfrak{K}}$  filtrante décroissante vers  $\emptyset$ , on a

$$\inf_{i \in I} \inf_{G \in \underline{\mathfrak{G}}, G \supset K_i} \nu(G) \leq \inf_{i \in I} \sup_{\mu \in \mathfrak{M}} \inf_{G \in \underline{\mathfrak{G}}, G \supset K_i} \mu(G) = 0 = \nu(\emptyset).$$

c'est-à-dire que l'application  $\nu$  est  $\sigma$ -régulière en  $\emptyset$  par rapport à  $\underline{\mathfrak{K}}$ . Il découle alors de [Top70a, Theorem 2] et des propriétés I à IV du lemme 1.3.3

que l'application  $\mu : \mathcal{F} \otimes \mathfrak{B}(\mathbb{T}) \rightarrow [0, +\infty[$  définie par

$$(\forall B \in \mathcal{F} \otimes \mathfrak{B}(\mathbb{T})) \quad \mu(B) = \sup_{K \subset B, K \in \mathfrak{K}} \inf_{G \in \underline{\mathfrak{G}}, G \supset K} \nu(G)$$

est une mesure sur  $\mathcal{F} \otimes \mathfrak{B}(\mathbb{T})$ . De plus, on a évidemment

$$(\forall G \in \underline{\mathfrak{G}}) \quad \mu(G) \leq \liminf_{\alpha} \mu^{\alpha}(G).$$

Pour montrer que  $(\mu^{\alpha})_{\alpha}$  converge vers  $\mu$ , il ne reste plus qu'à vérifier que

$$\mu(\Omega \times \mathbb{T}) = \lim_{\alpha} \mu^{\alpha}(\Omega \times \mathbb{T}).$$

Supposons le contraire : il existe  $\epsilon > 0$  et, pour chaque  $K \in \mathfrak{K}$ , un élément  $G_K$  de  $\underline{\mathfrak{G}}$  tel que  $G_K \supset K$  et  $\lim_{\alpha} \mu^{\alpha}(G_K) + 2\epsilon \leq \lim_{\alpha} \mu^{\alpha}(\Omega \times \mathbb{T})$ . On obtient alors une contradiction avec (ii) en prenant  $\underline{\mathfrak{G}}' = \{G_K; K \in \mathfrak{K}\}$ .  $\square$

On déduit du théorème 1.3.5 des généralisations du critère de compacité de Prokhorov pour les mesures boréliennes. Définissons d'abord des notions de tension pour les mesures de Young.

On dit qu'une partie  $\mathfrak{M}$  de  $M(\Omega)$  est *souplement tendue* si, pour tout  $\epsilon > 0$ , on peut trouver un compact aléatoire  $K \in \mathfrak{K}$  tel que  $\sup_{\mu \in \mathfrak{M}} \mu(K^c) < \epsilon$ . On dit que  $\mathfrak{M}$  est *strictement tendue* si, pour tout  $\epsilon > 0$ , on peut trouver un compact  $K \in \mathfrak{K}$  tel que  $\sup_{\mu \in \mathfrak{M}} \mu(\Omega \times K^c) < \epsilon$ . Donc  $\mathfrak{M}$  est strictement tendu si et seulement si  $(\pi_{\mathbb{T}})_{\#}(\mathfrak{M})$  est tendu au sens usuel. Une autre formulation de ces notions de tension a été introduite par Balder [Bal84a] et fait appel aux intégrandes inf-compacts.

Le théorème 1.3.5 nous donne immédiatement le résultat suivant, qui généralise [Bal01, Theorem 5.2] dans la mesure où  $\mathbb{T}$  n'est pas supposé régulier et où la compacité obtenue l'est pour une topologie a priori plus fine que la "ws-topology"  $\tau_{\mathbb{M}}^w$ .

**Proposition 1.3.6 (critère de Prokhorov)**

- A) *Supposons que  $\mathbb{T}$  soit souslinien sous-métrisable. Soit  $\mathfrak{M}$  une partie souplement tendue de  $M(\Omega \times \mathbb{T})$  telle que  $(\pi_{\Omega})_{\#}(\mathfrak{M})$  soit compact pour les suites généralisées dans  $M(\Omega)$  muni de la  $s$ -topologie. Alors  $\mathfrak{M}$  est compact pour les suites généralisées dans  $(M(\Omega \times \mathbb{T}), \tau_{\mathbb{M}}^{s*})$  (et donc aussi dans  $(M(\Omega \times \mathbb{T}), \tau_{\mathbb{M}}^s)$ ).*
- B) *Supposons maintenant seulement que les compacts de  $\mathbb{T}$  soient métrisables. Soit  $\mathfrak{M}$  une partie strictement tendue de  $M(\Omega \times \mathbb{T})$  telle que  $(\pi_{\Omega})_{\#}(\mathfrak{M})$  soit compact pour les suites généralisées dans  $M(\Omega)$  muni de la  $s$ -topologie. Alors  $\mathfrak{M}$  est compact pour les suites généralisées dans  $(M(\Omega \times \mathbb{T}), \tau_{\mathbb{M}}^{s*})$  (et donc aussi dans  $(M(\Omega \times \mathbb{T}), \tau_{\mathbb{M}}^s)$ ).*

Rappelons qu'une mesure  $\mu \in M(\mathbb{T})$  est dite  $\tau$ -régulière si, pour toute suite filtrante décroissante  $(F_\alpha)$  de fermés de  $\mathbb{T}$ , on a  $\mu(\cap_\alpha F_\alpha) = \inf_\alpha \mu(F_\alpha)$ . Toute mesure de Radon est  $\tau$ -régulière. L'espace  $\mathbb{T}$  est dit de *Prokhorov* si toute ensemble étroitement compact de mesures  $\tau$ -régulières est tendu (voir par exemple [Bog98]). Les espaces polonais sont des espaces de Prokhorov, d'autres exemples sont donnés dans [Top74, Whe83]. La caractérisation la plus générale des espaces de Prokhorov est donnée dans [Bou02].

Le critère de Prokhorov prend une forme particulièrement agréable dans le cas où  $\mathbb{T}$  est un espace de Prokhorov. Le résultat ci-dessous généralise un critère de Jacod et Mémin ([JM81a, Théorème 2.8], voir aussi [Sch75, Theorem 3.10] et [Bal01, Theorem 5.2]), qui était connu dans le cas où  $\mathbb{T}$  est polonais.

**Corollaire 1.3.7 (critère de Jacod et Mémin)** *On suppose  $\mathbb{T}$  souslinien sous-métrisable et de Prokhorov. Une partie  $\mathfrak{M}$  de  $M(\Omega \times \mathbb{T})$  est compacte pour les suites généralisées dans  $(M(\Omega \times \mathbb{T}), \tau_M^s)$  si et seulement si les conditions (a) et (b) ci-dessous sont vérifiées.*

- (a)  $(\pi_\Omega)_\#(\mathfrak{M})$  est compact pour les suites généralisées dans  $M(\Omega)$  muni de la  $s$ -topologie.
- (b)  $(\pi_{\mathbb{T}})_\#(\mathfrak{M})$  est compact pour les suites généralisées dans  $M(\mathbb{T})$  muni de la topologie étroite.

**Application aux mesures de Young** Les critères précédents admettent des versions triviales dans  $(\mathcal{Y}, \tau_{\mathcal{Y}}^s)$ , puisque la projection sur  $M(\Omega)$  de toute partie de  $\mathcal{Y}$  est  $\{P\}$ , donc toujours compacte. On a en particulier le résultat suivant.

**Théorème 1.3.8 (critères de compacité pour les mesures de Young)** *On suppose que  $\mathbb{T}$  est souslinien sous-métrisable.*

1. *Toute partie simplement tendue de  $\mathcal{Y}$  est compacte pour les suites généralisées dans la topologie  $\tau_{\mathcal{Y}}^s$ .*
2. *Si  $\mathbb{T}$  est un espace de Prokhorov, une partie  $\mathfrak{M}$  de  $\mathcal{Y}$  est compacte pour les suites généralisées dans la topologie  $\tau_{\mathcal{Y}}^s$  si et seulement si  $(\pi_{\mathbb{T}})_\#(\mathfrak{M})$  est compact pour les suites généralisées dans la topologie étroite.*
3. *Toute partie relativement compacte de  $(\mathcal{Y}, \tau_{\mathcal{Y}}^s)$  est séquentiellement relativement compacte.*

On en déduit immédiatement un résultat bien connu dans le cas polonais ([Bal84a, page 573], [Jaw84, théorème 2.4]) :

**Corollaire 1.3.9 (équivalence des notions de tension)** *Supposons que  $\mathbb{T}$  soit un espace de Prokhorov souslinien sous-métrisable. Alors toute partie simplement tendue de  $\mathcal{Y}$  est strictement tendue.*

On obtient une variante utile du corollaire 1.3.9, par une méthode différente [CRdFV04, Theorem 4.3.14], inspirée d'une idée de Valadier [Jaw84, théorème 2.4]. On dit que  $\mathbb{T}$  admet une *suite cofinale de compacts métrisables* s'il existe une suite  $(K_n)$  de compacts métrisables de  $\mathbb{T}$  telle que tout compact de  $\mathbb{T}$  soit contenu dans l'un des  $K_n$ . C'est le cas si, par exemple,  $\mathbb{T}$  est le dual  $*$ -faible d'un espace de Banach séparable ; or un tel espace n'est de Prokhorov que s'il est de dimension finie [Fer94].

**Proposition 1.3.10** *Supposons que  $\mathbb{T}$  admette une suite cofinale de compacts métrisables. Alors toute partie simplement tendue de  $\mathcal{Y}$  est strictement tendue.*

**Idée de la preuve.** On munit l'ensemble  $\mathfrak{K}$  de la topologie de Vietoris, c'est-à-dire de la topologie engendrée par les ensembles

$$U^- = \{F \in \mathfrak{F}; F \cap U \neq \emptyset\}$$

et

$$U^+ = \{F \in \mathfrak{F}; F \subset U\},$$

où  $U$  décrit  $\mathfrak{G}$ . D'après une caractérisation due à Michael [Mic51, Theorem 2.5.2] (pour la condition suffisante) et Christensen [Chr74, Theorem 3.1] (pour la condition nécessaire), une partie  $\mathfrak{C}$  de  $\mathfrak{K}$  est relativement compacte si et seulement s'il existe  $K \in \mathfrak{K}$  tel que l'on ait  $C \subset K$  pour tout  $C \in \mathfrak{C}$ . Sous les hypothèses de la proposition, on montre que  $\mathfrak{K}$  est un espace de Radon [CRdFV04, Theorem 4.3.14].

Soit  $\mathfrak{M}$  une partie simplement tendue de  $\mathcal{Y}$  et soit  $\epsilon > 0$ . Il existe  $\underline{K} \in \mathfrak{K}$  tel que  $\sup_{\mu \in \mathfrak{M}} \mu(\underline{K}) \geq 1 - \epsilon/2$ . Or  $\underline{K}$  peut être vu comme un élément aléatoire de  $\mathfrak{K}$ , dont la loi est donc tendue puisque  $\mathfrak{K}$  est un espace de Radon. D'après le résultat de Michael, il existe donc un compact  $K \in \mathfrak{K}$  tel que l'on ait  $P\{\underline{K}(\omega) \subset K\} \geq 1 - \epsilon$ . On en déduit immédiatement  $\sup_{\mu \in \mathfrak{M}} \mu(\Omega \times K) \geq 1 - \epsilon$ , donc  $\mathfrak{M}$  est strictement tendu.  $\square$

## 1.4 Tableaux de résultats sur les convergences stables

On rassemble dans cette partie des résultats qui précisent ceux des sections précédentes.

TABLEAU 1: THEOREME DE SEMICONTINUITÉ

$\mu^\alpha \xrightarrow{\text{W-stable}} \mu^\infty \Rightarrow \mu^\alpha \xrightarrow{\text{S-stable}} \mu^\infty$	
Condition suffisante	Preuve dans [CRdFV04]
$\mathbb{T}$ est souslinien métrisable	Portmanteau Theorem (Theorem 2.1.3-G page 25)
$(\mu^\alpha)_\alpha$ est strictement tendu et $P_\tau(\mathbb{T}) = P(\mathbb{T})$ (e.g. $\mathbb{T}$ est de Radon ou héréditairement Lindelöf) et $\mathbb{T}$ est complètement régulier et les parties compactes de $\mathbb{T}$ sont métrisables  ou  $(\mu^\alpha)_\alpha$ est souplement tendu et $\mathbb{T}$ est souslinien régulier	Theorem 4.3.8 page 94
$(\mu^\alpha)_\alpha$ est une suite et $\mathbb{T}$ est un espace séquentiellement de Prohorov tel que $P_t(\mathbb{T}) = P_\tau(\mathbb{T}) = P(\mathbb{T})$ (e.g. $\mathbb{T}$ est un espace métrique ou un espace $k_\omega$ sous-métrisable)	Theorem 4.5.1 page 104



TABLEAU 2: CRITÈRES DE RELATIVE COMPACTITÉ SÉQUENTIELLE

$\mathcal{K} \subset \mathcal{Y}$ est $\tau_{\mathcal{Y}}^s$ -séquentiellement relativement compact	
Condition suffisante	Preuve dans [CRdFV04]
$\mathcal{K}$ est strictement tendu et les parties compactes de $\mathbb{T}$ sont métrisables ou $\mathcal{K}$ est souplement tendu et $\mathbb{T}$ est souslinien sous-métrisable	Theorem 4.3.5 (Prohorov Criterion) page 92
$\mathcal{K}$ est souplement tendu et $\mathbb{T}$ est souslinien et $\mathfrak{K}(\mathbb{T})$ est de Radon (e.g. $\mathbb{T}$ a une suite cofinale de parties compactes métrisables)	Theorem 4.3.5 (Prohorov Criterion) page 92 et Theorem 4.3.13 page 97 (voir aussi Theorem 4.3.14)

$\mathcal{K}$ est $\tau_{\mathcal{Y}}^s$ -séquentiellement relativement compact $\Rightarrow$ $\mathcal{K}$ est $\tau_{\mathcal{Y}}^s$ -relativement compact	
Condition suffisante	Preuve dans [CRdFV04]
$\mathbb{T}$ est souslinien métrisable et $\mathcal{F}$ est essentiellement dénombrablement engendré (donc $\tau_{\mathcal{Y}}^s$ est métrisable)	Remark 4.5.3 page 105
$\mathbb{T}$ est un espace polonais ou bien un espace séquentiellement de Prohorov possédant une suite cofinale de compacts métrisables (e.g. $\mathbb{T}$ est un espace $k_{\omega}$ sous-métrisable)	Proposition 4.5.2 page 105

$\mathcal{K}$ est $\tau_{\mathcal{Y}}^s$ -relativement compact $\Rightarrow$ $\mathcal{K}$ est $\tau_{\mathcal{Y}}^s$ -séquentiellement relativement compact	
Condition suffisante	Preuve dans [CRdFV04]
$\mathbb{T}$ est souslinien sous-métrisable	Lemma 4.1.1 page 84

## 1.5 Convergence en probabilité ou en moyenne d'ordre $p$

Dans toute cette partie, on suppose pour simplifier que  $\mathbb{T}$  est un espace cosmique régulier (donc complètement régulier) dont la topologie est définie par un ensemble  $\mathcal{D}$  de semidistances.

### 1.5.1 Convergence en probabilité de mesures de Young

#### Convergence en probabilité d'éléments aléatoires

Soit  $(X_\alpha)$  une suite généralisée dans  $L^0(\mathbb{T})$  et soit  $X \in L^0(\mathbb{T})$ . On dit que  $(X_\alpha)$  converge en probabilité vers  $X$ , et on note

$$X_\alpha \xrightarrow{\text{probabilité}} X$$

si on a

$$(1.5.1) \quad (\forall d \in \mathcal{D}) \quad (\forall \epsilon > 0) \quad \lim_{\alpha} \mathbb{P} \{d(X_\alpha, X) > \epsilon\} = 0.$$

Un calcul classique et facile montre que (1.5.1) équivaut à

$$(1.5.2) \quad (\forall d \in \mathcal{D}) \quad \lim_{\alpha} \int_{\Omega} \min(d(X_\alpha, X), 1) d\mathbb{P} = 0.$$

J. Hoffmann-Jørgensen a donné une définition de la convergence en probabilité qui ne dépend que de la topologie de  $\mathbb{T}$  et non de la structure uniforme associée à  $\mathcal{D}$  : plus précisément, soit  $C(\mathbb{T}, [0, 1])$  l'espace des fonctions continues de  $\mathbb{T}$  dans  $[0, 1]$ . Alors (1.5.1) et (1.5.2) équivalent à

$$(1.5.3) \quad (\forall \varphi \in C(\mathbb{T}, [0, 1])) \quad \lim_{\alpha} \int_{\Omega} |\varphi(X_\alpha) - \varphi(X)| d\mathbb{P} = 0.$$

([HJ91, Theorem 7.4] et [HJ98, Corollary 4.7]). (Cette définition a aussi l'avantage de conserver un sens même si  $\mathbb{T}$  n'est pas complètement régulier.) D'autre part, pour toute variable aléatoire réelle définie sur  $\Omega$ , on a

$$(1.5.4) \quad \int_{\Omega} |Y| d\mathbb{P} \geq \sup_{A \in \mathcal{F}} \left| \int_A Y d\mathbb{P} \right| \geq \frac{1}{2} \int_{\Omega} |Y| d\mathbb{P},$$

donc (1.5.3) revient à

$$(1.5.5) \quad (\forall \varphi \in C(\mathbb{T}, [0, 1])) \quad \lim_{\alpha} \sup_{A \in \mathcal{F}} \int_A (\varphi(X_\alpha) - \varphi(X_\infty)) d\mathbb{P} = 0.$$

On voit donc que l'on a en particulier, en enlevant le supremum et en traduisant (1.5.5) dans le langage des mesures de Young,

$$(1.5.6) \quad (\forall A \in \mathcal{F}) \quad (\forall \varphi \in C(\mathbb{T}, [0, 1])) \quad \lim_{\alpha} \underline{\delta}_{X_{\alpha}}(\mathbf{1}_A \otimes \varphi) = \underline{\delta}_X(\mathbf{1}_A \otimes \varphi).$$

Donc, si  $(X_{\alpha})$  converge en probabilité vers  $X$ , (1.5.6) montre que  $\underline{\delta}_{X_{\alpha}}$  converge de manière W-stable vers  $\underline{\delta}_X$ . Réciproquement, supposons que  $\underline{\delta}_{X_{\alpha}}$  converge de manière W-stable vers  $\underline{\delta}_X$ . Soit  $d \in \mathcal{D}$  et soit  $g \in \underline{\text{BL}}_1(\Omega, \mathbb{T}, d)$  l'intégrande défini par

$$g(\omega, x) = d(x, X(\omega)).$$

D'après la proposition 1.2.2, on a

$$\lim_{\alpha} \int d(X_{\alpha}, X) dP = \lim_{\alpha} \int g d\underline{\delta}_{X_{\alpha}} = \int g d\underline{\delta}_X = 0,$$

donc  $(X_{\alpha})_{\alpha \in \mathbb{A}}$  converge en probabilité vers  $X$ . Ainsi, la topologie de la convergence en probabilité sur  $L^0$  coïncide avec la trace sur  $L^0$  de la convergence W-stable. Cela signifie que l'on peut supprimer le supremum dans (1.5.5) et exiger seulement que la convergence ait lieu pour chaque  $A \in \mathcal{F}$ . On verra un résultat plus fort dans le théorème 1.5.4.

Si  $d$  est une semidistance sur  $\mathbb{T}$ , on note  $\Delta_{\text{prob}}^{(d)}$  la semidistance sur  $L^0$  définie par

$$\Delta_{\text{prob}}^{(d)}(X, Y) = \mathbf{E}(\min(d(X, Y), 1)).$$

La convergence en probabilité est donc la convergence pour la topologie  $\tau_{L^0}^{\text{prob}}$  induite par les semidistances  $\Delta_{\text{prob}}^{(d)}$  ( $d \in \mathcal{D}$ ). On remarque aussi que, si  $d$  est une distance sur  $\mathbb{T}$ , alors  $\Delta_{\text{prob}}^{(d)}$  est une distance sur  $L^0(\mathbb{T})$ , et si  $d$  induit la topologie de  $\mathbb{T}$ , alors  $\Delta_{\text{prob}}^{(d)}$  induit  $\tau_{L^0}^{\text{prob}}$ .

## Convergence en probabilité de mesures de Young

On définit naturellement la topologie  $\tau_{\mathcal{Y}_d}^{\text{prob}}$  de la convergence en probabilité sur le sous-espace  $\mathcal{Y}_d$  des mesures de Young désintégrables, vues comme des éléments aléatoires de l'espace  $P(\mathbb{T})$  muni de la topologie faible. L'espace  $\mathbb{T}$  étant supposé cosmique régulier,  $P(\mathbb{T})$  est également cosmique régulier pour la topologie faible [CRdFV04, Proposition 1.3.2], donc le cadre décrit plus haut s'applique parfaitement. On verra plus loin que la topologie  $\tau_{\mathcal{Y}_d}^{\text{prob}}$  a un prolongement naturel à  $\mathcal{Y}$ .

Parmi les ensembles de semidistances qui induisent la topologie faible de  $P(\mathbb{T})$ , il en est deux particulièrement intéressants. Le premier est donné par la définition : c'est celui des semidistances  $d^{(f)} : (\mu, \nu) \mapsto |\mu(f) - \nu(f)|$ , où  $f$

décrit  $C(\mathbb{T}, [0, 1])$ . Ainsi, la topologie  $\tau_{\mathcal{Y}_d}^{\text{prob}}$  est induite par les semidistances  $\Delta_{\text{prob}}^{(d^{(f)})}$  ( $f \in C(\mathbb{T}, [0, 1])$ ) définies par

$$\Delta_{\text{prob}}^{(d^{(f)})}(\mu, \nu) = \int_{\Omega} |\mu_{\omega}(f) - \nu_{\omega}(f)| dP(\omega).$$

D'après (1.5.4), on voit que les semidistances  $\Delta_{\text{prob}}^{(d^{(f)})}$  sont équivalentes aux semidistances  $\tilde{\Delta}_{\text{prob}}^{(f)}$  définies par

$$(1.5.7) \quad \tilde{\Delta}_{\text{prob}}^{(f)}(\mu, \nu) = \sup_{A \in \mathcal{F}} |\mu(\mathbf{1}_A \otimes f) - \nu(\mathbf{1}_A \otimes f)|.$$

L'intérêt de l'égalité (1.5.7) est qu'elle conserve un sens même si  $\mu$  et  $\nu$  ne sont pas désintégrables. On obtient donc une extension naturelle de  $\tau_{\mathcal{Y}_d}^{\text{prob}}$  à  $\mathcal{Y}$  : la *topologie de la convergence en probabilité*  $\tau_{\mathcal{Y}}^{\text{prob}}$  sur  $\mathcal{Y}$  est la topologie induite par les semidistances  $\tilde{\Delta}_{\text{prob}}^{(d^{(f)})}$  ( $f \in C(\mathbb{T}, [0, 1])$ ) étendues à  $\mathcal{Y}$ .

Le second ensemble de semidistances qui induit la topologie faible de  $P(\mathbb{T})$  est celui des semidistances de Dudley  $\Delta_{\text{BL}}^{(d)}$  ( $d \in \mathcal{D}$ )

$$\Delta_{\text{BL}}^{(d)}(\mu, \nu) = \sup_{f \in \text{BL}_1(\mathbb{T}, d)} |\mu(f) - \nu(f)|$$

(voir [Dud66]). La topologie  $\tau_{\mathcal{Y}}^{\text{prob}}$  est donc induite par les semidistances  $\underline{\Delta}_{\text{BL}}^{(d)}$  ( $d \in \mathcal{D}$ ) définies par

$$\underline{\Delta}_{\text{BL}}^{(d)}(\mu, \nu) = \int_{\Omega} \Delta_{\text{BL}}^{(d)}(\mu_{\omega}, \nu_{\omega}) dP(\omega).$$

On appelle  $\underline{\Delta}_{\text{BL}}^{(d)}$  la *semidistance de Dudley paramétrée* associée à  $d$ .

**Théorème 1.5.1** *On suppose que  $\mathcal{D}$  est filtrant croissant. La topologie  $\tau_{\mathcal{Y}_d}^{\text{prob}}$  est induite par la famille  $(\underline{\Delta}_{\text{BL}}^{(d)})_{d \in \mathcal{D}}$  et on a, pour tout  $d \in \mathcal{D}$  et tout  $\mu, \nu \in \mathcal{Y}$ ,*

$$(1.5.8) \quad \underline{\Delta}_{\text{BL}}^{(d)}(\mu, \nu) = \sup_{f \in \underline{\text{BL}}_1(\Omega, \mathbb{T}, d)} (\mu(f) - \nu(f)) = \sup_{f \in \underline{\text{BL}}'_1(\Omega, \mathbb{T}, d)} (\mu(f) - \nu(f)).$$

*Si  $d$  est une distance, alors  $\underline{\Delta}_{\text{BL}}^{(d)}$  est une distance sur  $\mathcal{Y}$ .*

On obtient donc deux définitions de  $\underline{\Delta}_{\text{BL}}^{(d)}(\mu, \nu)$  qui conservent un sens même si  $\mu$  et  $\nu$  ne sont pas désintégrables.

**Théorème 1.5.2 (semidistances de Dudley paramétrées sur  $\mathcal{Y}$ )** *On suppose que  $\mathcal{D}$  est filtrant croissant. Pour chaque  $d \in \mathcal{D}$  et pour tout  $\mu, \nu \in \mathcal{Y}$ , la deuxième égalité dans (1.5.8) reste vraie si  $\nu$  ne sont pas désintégrables, c'est-à-dire qu'on a*

$$\sup_{f \in \underline{\text{BL}}_1(\Omega, \mathbb{T}, d)} (\mu(f) - \nu(f)) = \sup_{f \in \underline{\text{BL}}'_1(\Omega, \mathbb{T}, d)} (\mu(f) - \nu(f)).$$

*On peut donc prendre (1.5.8) comme définition des semidistances  $\underline{\Delta}_{\text{BL}}^{(d)}$  sur l'espace  $\mathcal{Y}$  tout entier. De plus, la topologie  $\tau_{\mathcal{Y}}^{\text{prob}}$  de la convergence en probabilité sur  $\mathcal{Y}$  est induite par les semidistances  $\underline{\Delta}_{\text{BL}}^{(d)}$  ( $d \in \mathcal{D}$ ).*

**Remarque 1.5.3 (comparaison des caractérisations de  $\tau_{\mathcal{Y}}^{\text{w}}$  et  $\tau_{\mathcal{Y}}^{\text{prob}}$ )**

La première définition de  $\tau_{\mathcal{Y}}^{\text{prob}}$  a montré que la différence entre les topologies  $\tau_{\mathcal{Y}}^{\text{prob}}$  et  $\tau_{\mathcal{Y}}^{\text{w}}$  tient au terme  $\sup_{A \in \mathcal{F}}$  dans (1.5.7). De même, d'après le théorème 1.5.2, si  $\mathcal{D}$  est filtrant croissant, alors  $\tau_{\mathcal{Y}}^{\text{prob}}$  est la topologie de la convergence uniforme sur les ensembles  $\underline{\text{BL}}_1(\Omega, \mathbb{T}, d)$  ( $d \in \mathcal{D}$ ), tandis que, d'après le théorème 1.2.1 (condition B3)  $\tau_{\mathcal{Y}}^{\text{w}}$  est la topologie de la convergence simple sur  $\cup_{d \in \mathcal{D}} \underline{\text{BL}}_1(\Omega, \mathbb{T}, d)$ .

**Théorème 1.5.4 (comparaison de  $\tau_{\mathcal{Y}}^{\text{prob}}$  avec  $\tau_{\mathcal{Y}}^{\text{w}}$ )**

1. *La topologie  $\tau_{\mathcal{Y}}^{\text{prob}}$  est plus fine que  $\tau_{\mathcal{Y}}^{\text{w}}$ .*
2.  *$\tau_{\mathcal{Y}}^{\text{prob}}$  est strictement plus fine que  $\tau_{\mathcal{Y}}^{\text{w}}$  si et seulement si  $\mathbb{T}$  a au moins deux éléments et  $(\Omega, \mathcal{F}, \mathbb{P})$  a une partie non atomique.*
3. *Ces deux topologies coïncident sur  $L^0(\mathbb{T})$ . De plus, si  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  est une suite généralisée dans  $\mathcal{Y}_d$  et si  $X \in L^0$ , alors  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  converge de manière  $W$ -stable vers  $\mu^\infty = \underline{\delta}_X$  si et seulement si  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  converge en probabilité vers  $\underline{\delta}_X$ .*
4.  *$L^0(\mathbb{T})$  est fermé pour  $\tau_{\mathcal{Y}}^{\text{prob}}$  dans  $L^0(\mathbb{P}(\mathbb{T})) = \mathcal{Y}_d$ .*
5. *Si  $\mathbb{T}$  est polonais ou si  $\mathbb{T}$  admet une suite cofinale de compacts métrisables, alors  $(\mathcal{Y}, \tau_{\mathcal{Y}}^{\text{prob}})$  et  $(\mathcal{Y}, \tau_{\mathcal{Y}}^{\text{w}})$  ont les mêmes parties convexes séquentiellement fermées.*

**Remarque 1.5.5** Supposons que  $\mathbb{T}$  soit souslinien (régulier) et que  $\mathbb{P}$  soit sans atome. D'après le théorème 1.2.5,  $L^0$  est dense dans  $\mathcal{Y}_d = \mathcal{Y}$  muni de  $\tau_{\mathcal{Y}}^{\text{w}}$ , alors qu'il est fermé dans  $\mathcal{Y}$  pour la topologie  $\tau_{\mathcal{Y}}^{\text{prob}}$  et que ces deux topologies coïncident sur  $L^0$ .

On retrouve la même situation avec le sous espace  $\mathbf{C}$  de l'espace de Skorokhod  $\mathbf{D}$  (voir page 36) :  $\mathbf{C}$  est dense dans  $\mathbf{D}$  pour la topologie  $J_1$ , alors qu'il est fermé dans  $\mathbf{D}$  pour la convergence uniforme sur les compacts et que ces deux topologies coïncident sur  $\mathbf{C}$ .

## 1.5.2 Convergence en moyenne de mesures de Young

Soit  $p \geq 1$  fixé. Nous allons définir sur  $\mathcal{Y}_{\mathcal{D}}^p$  une topologie  $\tau_{\mathcal{Y},\mathcal{D}}^p$  plus fine que  $\tau_{\mathcal{Y}}^{\text{prob}}$  et qui induit sur  $L_{(\mathbb{T},\mathcal{D})}^p := \mathcal{Y}_{\mathcal{D}}^p \cap L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{T})$  la même topologie que  $\tau_{\mathcal{Y},\mathcal{D}}^{\times,p}$ . On verra ensuite dans le corollaire 1.5.7 que, lorsque  $\mathbb{T} = \mathbb{E}$  est un espace normé la trace sur  $L_{(\mathbb{T},\mathcal{D})}^p$  de  $\tau_{\mathcal{Y},\mathcal{D}}^p$  est la topologie forte usuelle de  $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{E})$ .

Pour tout  $\mu$  et tout  $\nu$  dans  $\mathbb{P}(\mathbb{T})$ , on note

$$D(\mu, \nu) = \{ \pi \in \mathbb{P}(\mathbb{T} \times \mathbb{T}); \pi(\cdot \times \mathbb{T}) = \mu \text{ et } \pi(\mathbb{T} \times \cdot) = \nu \}.$$

Si  $c : \mathbb{T} \times \mathbb{T} \rightarrow [0, +\infty]$  est une fonction mesurable, on définit la *fonctionnelle de Kantorovich*  $\Delta_{\text{KR}}^{(c)} : \mathbb{P}(\mathbb{T}) \times \mathbb{P}(\mathbb{T}) \rightarrow [0, +\infty]$  associée à  $c$  par

$$\Delta_{\text{KR}}^{(c)}(\mu, \nu) := \inf_{\pi \in D(\mu, \nu)} \int_{\mathbb{T} \times \mathbb{T}} c(x, y) d\pi(x, y).$$

Pour tout  $d \in \mathcal{D}$ , on définit une semidistance  $\Delta_{\text{KR}}^{(p,d)} : \mathbb{P}(\mathbb{T}) \times \mathbb{P}(\mathbb{T}) \rightarrow [0, +\infty]$  en posant

$$\Delta_{\text{KR}}^{(p,d)}(\mu, \nu) := \left( \Delta_{\text{KR}}^{(dp)} \right)^{1/p} = \inf_{\pi \in D(\mu, \nu)} \left( \int_{\mathbb{T} \times \mathbb{T}} d(x, y)^p d\pi(x, y) \right)^{1/p}.$$

On appelle cette semidistance la *semidistance de Wasserstein* associée à  $d$ . Cette semidistance est bien sûr à valeurs finies sur

$$\mathbb{P}_d^p(\mathbb{T}) := \left\{ \mu \in \mathbb{P}(\mathbb{T}); \int_{\mathbb{T}} d(x_0, \cdot)^p d\mu < \infty \right\}.$$

C'est une distance sur  $\mathbb{P}_d^p(\mathbb{T})$  si et seulement si  $d$  est une distance. On montre que la topologie induite par la famille  $(\Delta_{\text{KR}}^{(p,d)}(\mu, \nu))_{d \in \mathcal{D}}$  induit une topologie plus fine que la convergence faible, et qui coïncide avec celle-ci si les éléments de  $\mathcal{D}$  sont bornés. Le lecteur pourra consulter [Rac91] sur le sujet des distances de Wasserstein.

Définissons maintenant des distances de Wasserstein paramétrées. Pour tout  $\mu$  et tout  $\nu$  dans  $\mathcal{Y}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{T})$ , on note

$$\underline{D}(\mu, \nu) = \{ \pi \in \mathcal{Y}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{T} \times \mathbb{T}); \pi(\cdot \times \cdot \times \mathbb{T}) = \mu \text{ et } \pi(\cdot \times \mathbb{T} \times \cdot) = \nu \}.$$

Pour  $c : \mathbb{T} \times \mathbb{T} \rightarrow [0, +\infty]$  mesurable et  $\mu, \nu \in \mathcal{Y}$ , on pose alors

$$\underline{\Delta}_{\text{KR}}^{(c)}(\mu, \nu) = \inf_{\pi \in \underline{D}(\mu, \nu)} \int_{\Omega \times \mathbb{T} \times \mathbb{T}} c(x, y) d\pi(\omega, x, y).$$

Pour chaque  $d \in \mathcal{D}$ , on appelle *semidistance de Wasserstein paramétrée* associée à  $d$  la semidistance sur  $\mathcal{Y}$  (à valeurs finies sur  $\mathcal{Y}_\mathcal{D}^p$ ) définie par

$$\underline{\Delta}_{\text{KR}}^{(p,d)}(\mu, \nu) := \left( \underline{\Delta}_{\text{KR}}^{(dp)} \right)^{1/p} = \inf_{\pi \in D(\mu, \nu)} \left( \int_{\mathbb{T} \times \mathbb{T}} d(x, y)^p d\pi(x, y) \right)^{1/p}.$$

On définit alors la *topologie*  $\tau_{\mathcal{Y}, \mathcal{D}}^p$  de la convergence en moyenne d'ordre  $p$  sur  $\mathcal{Y}_\mathcal{D}^p$  comme étant la topologie induite par les semidistances  $\underline{\Delta}_{\text{KR}}^{(p,d)}$  ( $d \in \mathcal{D}$ ).

On peut donner diverses caractérisations de la convergence en moyenne d'ordre  $p$ , dans le même esprit que la proposition 1.2.3 page 8.

**Théorème 1.5.6** *Soient  $p \geq 1$ ,  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  une suite généralisée dans  $\mathcal{Y}_\mathcal{D}^p$  et  $\mu \in \mathcal{Y}_\mathcal{D}^p$ . Soit  $d \in \mathcal{D}$  et  $x_0 \in \mathbb{T}$  fixés. Les affirmations suivantes sont équivalentes.*

1.  $\mu^\alpha \xrightarrow{\tau_{\mathcal{Y}, \mathcal{D}}^p} \mu$ .
2.  $\mu^\alpha \xrightarrow{\text{probabilité}} \mu$  et  $\lim_\alpha \int_{\Omega \times \mathbb{T}} d(a, x)^p d\mu^\alpha(\omega, x) = \int_{\Omega \times \mathbb{T}} d(a, x)^p d\mu(\omega, x)$  pour tout  $d \in \mathcal{D}$ .
3.  $\mu^\alpha \xrightarrow{\text{probabilité}} \mu$  et  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  est asymptotiquement uniformément  $p$ -intégrable.
4.  $\lim_\alpha \sup_{A \in \mathcal{F}} |\mu^\alpha(\mathbf{1}_A \otimes f) - \mu(\mathbf{1}_A \otimes f)| = 0$  pour tout  $f : \mathbb{T} \rightarrow [0, +\infty[$  tel que  $f \leq 1 + d(x_0, \cdot)^p$ .

En comparant la condition 3 du théorème 1.5.6 avec l'affirmation 3 du théorème 1.5.4, on obtient immédiatement le résultat ci-dessous.

**Corollaire 1.5.7** *Les topologies  $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p}$  et  $\tau_{\mathcal{Y}, \mathcal{D}}^p$  coïncident sur  $L_{(\mathbb{T}, \mathcal{D})}^p$ .*

*En particulier, si  $\mathbb{T}$  est un espace vectoriel normé séparable et si  $\mathcal{D}$  a pour unique élément la distance associée à la norme de  $\mathbb{E}$ , la trace sur  $L_{(\mathbb{E}, \|\cdot\|)}^p = L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{E})$  de la topologie  $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p}$  est la topologie forte usuelle de  $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{E})$ .*

La preuve du théorème 1.5.6 fait appel au lemme ci-dessous, dont la démonstration se trouve dans [DPRdF].

**Lemme 1.5.8** *On suppose que  $\mathbb{T}$  est souslinien. Soit  $c : \mathbb{T} \times \mathbb{T} \rightarrow [0, +\infty[$  une fonction semicontinue inférieurement. Alors, pour tout  $\mu$  et tout  $\nu$  dans  $\mathcal{Y}$ , l'application  $\omega \mapsto \Delta_{\text{KR}}^{(c)}(\mu_\omega, \nu_\omega)$  est  $\mathcal{F}^*$ -mesurable et il existe une mesure de Young  $\lambda \in \mathcal{Y}(\Omega, \mathcal{F}^*, \mathbb{P}; \mathbb{T} \times \mathbb{T})$  telle que*

$$\Delta_{\text{KR}}^{(c)}(\mu_\omega, \nu_\omega) = \int_{\mathbb{T} \times \mathbb{T}} c(x, y) d\lambda_\omega(x, y) \quad \text{P-}p.s.$$

De plus, on a

$$\underline{\Delta}_{\text{KR}}^{(c)}(\mu, \nu) = \int_{\Omega \times \mathbb{T} \times \mathbb{T}} c(x, y) d\lambda(\omega, x, y) = \int_{\Omega} \Delta_{\text{KR}}^{(c)}(\mu_{\omega}, \nu_{\omega}) dP(\omega).$$

### Idée de la preuve du théorème 1.5.6.

1  $\Rightarrow$  2. On commence par un passage au quotient  $\mathbb{T}/d$  comme dans [CRdF00, pages 125–126] ou dans [CRdFV04, Theorem 4.1]. On se ramène ainsi au cas où  $(\mathbb{T}, d)$  est un espace métrique séparable. On peut de plus, comme dans [CRdF00, page 126], raisonner dans le complété de  $(\mathbb{T}, d)$  (noté encore  $(\mathbb{T}, d)$ ). On suppose donc sans perdre en généralité que  $(\mathbb{T}, d)$  est un espace métrique séparable complet. En particulier,  $\mathbb{T}$  est souslinien et on peut appliquer le lemme 1.5.8 avec  $c = d^p$  : on a donc  $\lim_{\alpha} \int_{\Omega} \Delta_{\text{KR}}^{(p,d)}(\mu_{\omega}^{\alpha}, \mu_{\omega}) dP = 0$ . On en déduit

$$\lim_{\alpha} \int_{\Omega} \Delta_{\text{KR}}^{(1, \min(d,1))}(\mu_{\omega}^{\alpha}, \mu_{\omega}) dP = 0.$$

Comme  $\Delta_{\text{KR}}^{(1, \min(d,1))}$  est une distance bornée sur  $P(\mathbb{T})$  qui induit la topologie faible, la suite  $(\mu^{\alpha})$  converge donc en probabilité vers  $\mu$ . La preuve de la  $p$ -intégrabilité uniforme est semblable à [CRdF00, page 125].

2  $\Leftrightarrow$  3. Très facile et peut également se déduire de l'équivalence analogue dans la proposition 1.2.3.

3  $\Rightarrow$  1. Comme dans la preuve de 1  $\Rightarrow$  2, on suppose sans perdre en généralité que  $(\mathbb{T}, d)$  est un espace métrique séparable complet. Pour tout  $R > 0$ , notons  $d_R$  la distance  $\min(d, R)$ . Comme  $(\mu^{\alpha})$  converge en probabilité vers  $\mu$ , on a

$$(1.5.9) \quad \lim_{\alpha} \underline{\Delta}_{\text{KR}}^{(1, d_R)}(\mu^{\alpha}, \mu) = \lim_{\alpha} \int_{\Omega} \Delta_{\text{KR}}^{(1, d_R)}(\mu_{\omega}^{\alpha}, \mu_{\omega}) dP(\omega)$$

puisque la distance  $\Delta_{\text{KR}}^{(1, d_R)}$  est bornée et induit la convergence faible sur  $P(\mathbb{T}, d)$ . Soit  $\epsilon > 0$  et choisissons  $R > 0$  tel que

$$\begin{aligned} \limsup_{\alpha} \int_{\Omega \times \{d(x_0, \cdot) > R\}} d(x_0, x)^p d\mu^{\alpha}(\omega, x) &\leq \epsilon \\ \text{et} \int_{\Omega \times \{d(x_0, \cdot) > R\}} d(x_0, x)^p d\mu(\omega, x) &\leq \epsilon. \end{aligned}$$

Pour chaque indice  $\alpha$ , soit  $\lambda^{\alpha}$  la mesure de Young donnée par le lemme 1.5.8 avec  $c = d^p$  et  $\lambda^{R, \alpha}$  celle obtenue pour  $c = d_R^p$ . On a alors, en notant  $B$  la



boule fermée de centre  $x_0$  et de rayon  $R$ , on a

$$\begin{aligned} & \left( \underline{\Delta}_{\text{KR}}^{(p,d)}(\mu^\alpha, \mu) \right)^p \\ &= \int_{\Omega \times \mathbb{T} \times \mathbb{T}} d(x, y)^p d\lambda^\alpha(\omega, x, y) \leq \int_{\Omega \times \mathbb{T} \times \mathbb{T}} d(x, y)^p d\lambda^{R,\alpha}(\omega, x, y) \end{aligned}$$

d'où

$$\begin{aligned} & \left( \underline{\Delta}_{\text{KR}}^{(p,d)}(\mu^\alpha, \mu) \right)^p \\ & \leq \int_{\Omega \times \mathbb{T} \times \mathbb{T}} d_R(x, y)^p d\lambda^{R,\alpha}(\omega, x, y) + \int_{\Omega \times B \times B^c} d(x, y)^p d\lambda^{R,\alpha}(\omega, x, y) \\ & \quad + \int_{\Omega \times B^c \times B} d(x, y)^p d\lambda^{R,\alpha}(\omega, x, y) + \int_{\Omega \times B^c \times B^c} d(x, y)^p d\lambda^{R,\alpha}(\omega, x, y) \\ & \leq R^{p-1} \int_{\Omega \times \mathbb{T} \times \mathbb{T}} d_R(x, y) d\lambda^{R,\alpha}(\omega, x, y) \\ & \quad + 2^p \int_{\Omega \times B \times B^c} d(x_0, y)^p d\lambda^{R,\alpha}(\omega, x, y) + 2^p \int_{\Omega \times B^c \times B} d(x, x_0)^p d\lambda^{R,\alpha}(\omega, x, y) \\ & \quad + 2^{p-1} \int_{\Omega \times B^c \times B^c} (d(x_0, y)^p + d(x, x_0)^p) d\lambda^{R,\alpha}(\omega, x, y) \\ & \leq \underline{\Delta}_{\text{KR}}^{(1,d_R)}(\mu^\alpha, \mu) \\ & \quad + 2^{p+1} \int_{\Omega \times B^c} d(x_0, y)^p d\mu(\omega, y) + 2^{p+1} \int_{\Omega \times B^c} d(x_0, x)^p d\mu^\alpha(\omega, x) \\ & \leq \underline{\Delta}_{\text{KR}}^{(1,d_R)}(\mu^\alpha, \mu) + 2^{p+2}\epsilon. \end{aligned}$$

On en déduit l'implication voulue d'après (1.5.9).

3  $\Rightarrow$  4. Immédiat par troncation.

4  $\Rightarrow$  2. Il suffit d'appliquer 4 avec  $f(x) = d(x_0, \cdot)^p$ .  $\square$

## Mesures de Young d'ordre 1 et distance de Wasserstein paramétrée

Dans le cas  $p = 1$ , le théorème de dualité de Kantorovich–Rubinšteïn permet une autre expression des semidistances  $\Delta_{\text{KR}}^{(p,d)}(\mu, \nu)$  ( $d \in \mathcal{D}$ ). Rappelons d'abord un résultat célèbre, sous une forme légèrement moins générale que celle donnée par Levin [Lev84] (voir aussi [RR98, Theorem 4.6.6]).

**Théorème (théorème de dualité de Kantorovich–Rubinšteïn)** *On suppose que  $\mathbb{T}$  est souslinien régulier. Soit  $c : \mathbb{T} \times \mathbb{T} \rightarrow [0, +\infty]$  une application mesurable. Pour  $(\mu, \nu) \in \mathcal{P}(\mathbb{T}) \times \mathcal{P}(\mathbb{T})$ , on note*

$$\Delta_{\text{L}}^{(c)}(\mu, \nu) := \sup_{f \in \text{Lip}_{\mathbb{T}}^{(c)}} (\mu(f) - \nu(f))$$

où

$$\text{Lip}_{\mathbb{T}}^{(c)} = \{u \in \mathcal{C}_b(\mathbb{T}); (\forall x, y \in \mathbb{T}) |u(x) - u(y)| \leq c(x, y)\}.$$

Alors  $\Delta_{\text{KR}}^{(c)}(\mu, \nu) = \Delta_{\text{L}}^{(c)}(\mu, \nu)$  pour tout  $\mu, \nu \in \mathcal{P}(\mathbb{T})$  si et seulement si

$$(1.5.10) \quad (\forall x, y \in \mathbb{T}) \quad c(x, y) = \sup_{u \in \text{Lip}_{\mathbb{T}}^{(c)}} |u(x) - u(y)|.$$

Toute fonction  $c : \mathbb{T} \times \mathbb{T} \rightarrow [0, +\infty]$  vérifiant (1.5.10) est semicontinue inférieurement, puisque elle est le supremum d'un ensemble de fonctions continues. Toute semidistance continue  $c$  sur  $\mathbb{T}$  vérifie (1.5.10) [RR98, Corollary 4.5.7]. Si  $\mathbb{T}$  est compact, toute semidistance semicontinue inférieurement  $c$  sur  $\mathbb{T}$  vérifie (1.5.10) [RR98, Remark 4.5.6].

Le résultat ci-dessous provient de [DPRdF]. On note

$$\mathcal{Y}_c^1(\Omega, \mathcal{F}, \mathcal{P}; \mathbb{T}) = \{\mu \in \mathcal{Y}; \int_{\Omega \times \mathbb{T}} c(x, x_0) d\mu(\omega, x) < +\infty\}$$

où  $x_0$  est un élément fixé de  $\mathbb{T}$  (d'après (1.5.10) cette définition ne dépend pas du choix de  $x_0$ ). On note également  $\underline{\text{Lip}}^{(c)}$  l'ensemble des intégrandes mesurables  $f : \Omega \times \mathbb{T} \rightarrow \mathbb{R}$  vérifiant  $f(\omega, \cdot) \in \text{Lip}_{\mathbb{T}}^{(c)}$  pour tout  $\omega \in \Omega$ . On pose enfin, pour  $\mu, \nu \in \mathcal{Y}_c^1$ ,

$$\underline{\Delta}_{\text{L}}^{(c)}(\mu, \nu) = \sup_{f \in \underline{\text{Lip}}^{(c)}} (\mu(f) - \nu(f)).$$

**Théorème 1.5.9 (théorème de Kantorovich–Rubinšteïn paramétré)**

On suppose que  $\mathbb{T}$  est souslinien régulier. Soit  $c : \mathbb{T} \times \mathbb{T} \rightarrow [0, +\infty[$  vérifiant (1.5.10). Soient  $\mu, \nu \in \mathcal{Y}_c^1$ . Soit  $\lambda \in \mathcal{Y}(\Omega, \mathcal{F}^*, \mathcal{P}; \mathbb{T} \times \mathbb{T})$  comme dans le lemme 1.5.8. On a

$$\underline{\Delta}_{\text{KR}}^{(c)}(\mu, \nu) = \int_{\Omega \times \mathbb{T} \times \mathbb{T}} c(x, y) d\lambda(\omega, x, y) = \underline{\Delta}_{\text{L}}^{(c)}(\mu, \nu).$$

En particulier, si  $\mathcal{F} = \mathcal{F}^*$ , on a  $\lambda \in \underline{D}(\mu, \nu)$ , donc l'infimum dans la définition de  $\underline{\Delta}_{\text{KR}}^{(c)}(\mu, \nu)$  est atteint.

Ce théorème a des applications en dehors du cas où  $c$  est une semidistance continue : on trouvera par exemple dans [DPRdF] une application au couplage de variables aléatoires, dans laquelle  $c$  est la distance discrète.

**Corollaire 1.5.10** Si  $\mathbb{T}$  est souslinien régulier, la topologie  $\tau_{\mathcal{Y}, \mathcal{D}}^p$  de la convergence en moyenne est induite par les semidistances  $\underline{\Delta}_{\text{L}}^{(d)}$  ( $d \in \mathcal{D}$ ).

### 1.5.3 Produit fibré de mesures de Young

Soient  $\mathbb{S}$  et  $\mathbb{T}$  des espaces cosmiques réguliers et soient  $\mu \in \mathcal{Y}_d(\Omega, \mathbb{S})$  et  $\nu \in \mathcal{Y}_d(\Omega, \mathbb{T})$ . On appelle *produit fibré* de  $\mu$  et  $\nu$  la mesure de Young  $\mu \underline{\otimes} \nu \in \mathcal{Y}_d(\Omega, \mathcal{F}, P; \mathbb{S} \times \mathbb{T})$  définie par

$$(\mu \underline{\otimes} \nu)_\omega = \mu_\omega \otimes \nu_\omega$$

pour tout  $\omega \in \Omega$ . Cette définition est cohérente car, d'après l'hypothèse sur  $\mathbb{S}$  et  $\mathbb{T}$ , on a  $\mathfrak{B}(\mathbb{S}) \otimes \mathfrak{B}(\mathbb{T}) = \mathfrak{B}(\mathbb{S} \times \mathbb{T})$ .

Le théorème suivant a une longue histoire et de nombreuses applications [Fis70, Bal88, Val90, Val94, Tat02]. On en verra une application dans la section 1.6.2. Dans tous les articles cités ci-dessus (sauf [Tat02]),  $(\nu^\alpha)_\alpha$  converge de manière  $W$ -stable vers une mesure de Young dégénérée. Mais, d'après la partie 3 du théorème 1.5.4, cette hypothèse est contenue dans l'hypothèse 1b ci-dessous.

#### Théorème 1.5.11 (théorème du produit fibré)

1. *Soient  $\mathbb{S}$  et  $\mathbb{T}$  des espaces cosmiques. Soient  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  et  $(\nu^\alpha)_{\alpha \in \mathbb{A}}$  des suites généralisées dans  $\mathcal{Y}_d(\mathbb{S})$  et  $\mathcal{Y}_d(\mathbb{T})$  respectivement (avec les mêmes indices). Supposons que*
  - 1a.  $(\mu^\alpha)_{\alpha \in \mathbb{A}}$  converge de manière  $W$ -stable vers  $\mu \in \mathcal{Y}_d(\mathbb{S})$ ,
  - 1b.  $(\nu^\alpha)_{\alpha \in \mathbb{A}}$  converge en probabilité vers  $\nu \in \mathcal{Y}_d(\mathbb{T})$ .*Alors  $(\mu^\alpha \underline{\otimes} \nu^\alpha)_{\alpha \in \mathbb{A}}$  converge de manière  $W$ -stable vers  $\mu \underline{\otimes} \nu$ .*
2. *Si de plus  $(\mu^\alpha)_\alpha$  converge en probabilité, alors  $(\mu^\alpha \underline{\otimes} \nu^\alpha)_\alpha$  converge en probabilité vers  $\mu \underline{\otimes} \nu$ .*
3. *Soit  $p \geq 1$ . On suppose  $\mathcal{D}$  filtrant croissant. La topologie de  $\mathbb{T} \times \mathbb{T}$  est donc définie par la famille  $\tilde{\mathcal{D}}$  de semidistances  $\tilde{d}((x, x'), (y, y')) = d(x, y) + d(x', y')$  ( $d \in \mathcal{D}$ ). Si  $(\mu^\alpha)$  converge pour  $\tau_{\mathcal{Y}, \mathcal{D}}^{w,p}$  (resp. pour  $\tau_{\mathcal{Y}, \mathcal{D}}^p$ ) vers  $\mu$  et si  $(\nu^\alpha)$  converge pour  $\tau_{\mathcal{Y}, \mathcal{D}}^p$  vers  $\nu$ , alors  $(\mu^\alpha \underline{\otimes} \nu^\alpha)$  converge pour  $\tau_{\mathcal{Y}, \mathcal{D}}^{w,p}$  (resp. pour  $\tau_{\mathcal{Y}, \mathcal{D}}^p$ ) vers  $\mu \underline{\otimes} \nu$ .*

La démonstration des parties 1 et 2 est donnée dans [CRdFV04] et repose sur les caractérisations des convergences pour  $\tau_{\mathcal{Y}}^w$  et  $\tau_{\mathcal{Y}}^{\text{prob}}$  à l'aide des intégrandes lipschitziens : théorème 1.2.1, proposition 1.2.2 et théorème 1.5.1. La partie 3 s'en déduit immédiatement, en vérifiant que  $(\mu^\alpha \underline{\otimes} \nu^\alpha)_\alpha$  est asymptotiquement uniformément  $p$ -intégrable.

## 1.6 Applications en théorie des probabilités

Deux applications des mesures de Young sont présentées ici. La première utilise le critère de compacité de Prokhorov pour définir une solution faible pour une inclusion différentielle stochastique. La deuxième fait appel au théorème du produit fibré pour démontrer un théorème de la limite centrale pour un nombre aléatoire de vecteurs aléatoires.

### 1.6.1 Solutions faibles d'inclusions d'évolution stochastiques

Les mesures de Young fournissent un moyen assez naturel de construire des solutions “faibles” (au sens probabiliste) d'équations différentielles stochastiques. C'est en gros l'idée qui est à l'œuvre dans la notion de *solution-mesure* chez Jacod et Mémmin [JM81b], et les mesures de Young ont été réinventées dans ce contexte par Pellaumail [Pel81b, Pel81a] sous le nom de règles. Voici un exemple, tiré de [JKRdF05].

Dans ce qui suit,  $T \in ]0, +\infty[$  est un horizon fixé et  $\mathfrak{F} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$  est une base stochastique vérifiant les conditions habituelles. On se donne deux espaces de Hilbert séparables  $\mathbb{H}$  et  $\mathbb{U}$  et un  $(\mathcal{F}_t)$ -mouvement brownien  $W$  (éventuellement cylindrique)  $\mathbb{U}$ . Soit  $\mathbb{L}$  l'espace des opérateurs de Hilbert-Schmidt de  $\mathbb{U}$  dans  $\mathbb{H}$ . On considère le problème de Cauchy

$$(1.6.1) \quad \begin{cases} dX_t \in AX_t + F(t, X_t) dt + G(t, X_t) dW(t) \\ X(0) = \xi \end{cases}$$

où  $X$  est à valeurs dans  $\mathbb{H}$ ,  $A$  est un opérateur linéaire sur  $\mathbb{H}$  et  $F$  et  $G$  sont des applications mesurables à valeurs convexes compactes, respectivement dans  $\mathbb{H}$  et dans  $\mathbb{L}$ . On se donne une fois pour toutes un nombre  $p > 2$ . Pour tout  $t \in [0, T]$ ,  $\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})$  désigne l'espace des processus continus  $(\mathcal{F}_t)$ -adaptés  $X$  à valeurs dans  $\mathbb{H}$  tels que

$$\|X\|_{\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})}^p := \mathbf{E} \left( \sup_{0 \leq s \leq t} \|X(s)\|_{\mathbb{H}}^p \right) < +\infty.$$

Si  $\mathbb{E}$  est un espace de Banach, on note  $\mathfrak{K}_c^*(\mathbb{E})$  l'ensemble des compacts convexes non vides de  $\mathbb{E}$ , et  $\text{Hausd}_{\mathbb{E}}$  la distance de Hausdorff sur  $\mathfrak{K}_c^*(\mathbb{E})$  (la définition de cette distance est rappelée page 40). On se donne également les hypothèses suivantes.

(HS)  $A$  engendre un semigroupe  $(S(t))_{t \geq 0}$  de classe  $\mathcal{C}_0$ . En particulier, il existe  $M > 0$  et  $\beta \in ]-\infty, +\infty[$  tels que, pour tout  $t \geq 0$ ,

$$(1.6.2) \quad \|S(t)\| \leq M e^{\beta t}.$$

Pour  $t \in [0, T]$ , on note  $M_t = \sup_{0 \leq s \leq t} M e^{\beta s}$ .

(HFG)  $F : [0, T] \times \mathbb{H} \rightarrow \mathfrak{K}_c^*(\mathbb{H})$  et  $G : [0, T] \times \mathbb{H} \rightarrow \mathfrak{K}_c^*(\mathbb{L})$  sont des applications mesurables vérifiant les conditions (i) et (ii) ci-dessous.

(i) Il existe une constante  $C_{\text{croiss}} > 0$  telle que, pour tout  $(t, x) \in [0, T] \times \mathbb{H}$ ,

$$\begin{aligned} \text{Hausd}_{\mathbb{H}}(0, F(t, x)) &\leq C_{\text{croiss}}(1 + \|x\|), \\ \text{Hausd}_{\mathbb{L}}(0, G(t, x)) &\leq C_{\text{croiss}}(1 + \|x\|). \end{aligned}$$

(ii) Pour tout  $(t, x, y) \in [0, T] \times \mathbb{H} \times \mathbb{H}$ ,

$$\begin{aligned} (\text{Hausd}_{\mathbb{H}}(F(t, x), F(t, y)))^p &\leq L(t, \|x - y\|^p), \\ (\text{Hausd}_{\mathbb{L}}(G(t, x), G(t, y)))^p &\leq L(t, \|x - y\|^p), \end{aligned}$$

où  $L : [0, T] \times [0, +\infty] \rightarrow [0, +\infty]$  est une fonction continue donnée telle que

(a) pour tout  $t \in [0, T]$ , l'application  $L(t, \cdot)$  est croissante et concave,

(b) pour toute application mesurable  $z : [0, T] \rightarrow [0, +\infty]$  et pour toute constante  $K > 0$ , l'implication suivante est vraie :

$$(1.6.3) \quad \left[ (\forall t \in [0, T]) \ z(t) \leq K \int_0^t L(s, z(s)) \, ds \right] \Rightarrow z = 0.$$

En particulier, on a  $L(t, 0) = 0$  pour tout  $t \in [0, T]$ , donc (HFG)-(ii) entraîne que, pour tout  $t \in [0, T]$ , les applications  $F(t, \cdot)$  et  $G(t, \cdot)$  sont continues. Ce type de fonction  $L$  est considéré dans, par exemple, [Yam81, Rod84, Man90, AKP<sup>+</sup>92, Tan92, Bar98, BB02]. On peut en trouver des exemple concrets dans [Har73, Section 6 du Chapitre 3].

(HI)  $\xi \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}|_{\mathcal{F}_0})$ .

Rappelons (voir [DPZ92a, DPZ92b]) que, sous l'hypothèse (HS), il existe une constante  $C_{\text{Conv}}$  telle que, pour tout processus prévisible  $Z \in L^p(\Omega \times [0, T]; \mathbb{L})$ , on ait

$$(1.6.4) \quad \mathbf{E} \left[ \sup_{s \leq t} \left\| \int_0^s S(s-r) Z(r) \, dW(r) \right\|^p \right] \leq C_{\text{Conv}} t^{(p/2)-1} \mathbf{E} \int_0^t \|Z(s)\|_{\mathbb{L}}^p \, ds.$$

**Définition 1.6.1** On dit qu'un processus  $X \in \mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$  est une *solution d'évolution (forte)* de (1.6.1) s'il existe des processus prévisibles  $f$  et  $g$  définis sur  $\mathfrak{F}$  tels que l'on ait

$$(1.6.5) \quad \begin{cases} X(t) \in S(t)\xi + \int_0^t S(t-s)f(s) ds + \int_0^t S(t-s)g(s) dW(s) \\ f(s) \in F(s, X(s)) \text{ P-p.p.} \\ g(s) \in G(s, X(s)) \text{ P-p.p.} \end{cases}$$

On dit qu'un processus  $X$  est une *solution d'évolution faible* ou *solution-mesure d'évolution* de (1.6.1) s'il existe une base stochastique  $\underline{\mathfrak{F}} = (\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_{t, \mu})$  telle que

1.  $\underline{\Omega}$  est de la forme  $\underline{\Omega} = \Omega \times \Omega'$ ,  $\underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$  pour une tribu  $\mathcal{F}'$  sur  $\Omega'$ ,  $\underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'_t$  pour une filtration continue à droite  $(\mathcal{F}'_t)$  sur  $(\Omega', \mathcal{F}')$  et la probabilité  $\mu$  vérifie  $\mu(A \times \Omega') = P(A)$  pour tout  $A \in \mathcal{F}$ ,
2. le processus  $W$  est un mouvement brownien sur  $\underline{\mathfrak{F}}$  (on identifie ici toute variable aléatoire  $X$  définie sur  $\Omega$  avec la variable aléatoire  $(\omega, \omega') \mapsto X(\omega)$  définie sur  $\underline{\Omega}$ ),
3.  $X \in \mathcal{N}_c^p(\underline{\mathfrak{F}}, [0, T]; \mathbb{T})$  et il existe des processus prévisibles  $f$  et  $g$  définis sur  $\underline{\mathfrak{F}}$  vérifiant (1.6.5).

Les qualificatifs “fort” et “faible” sont donc à prendre au sens probabiliste. La terminologie *solution-mesure* est celle de Jacod et Mémmin [JM81b]. Si  $\mathcal{F}'$  est la tribu borélienne d'une topologie sur  $\Omega'$ , on peut voir une solution-mesure comme une mesure de Young : c'est le point de vue adopté par Pellaumail [Pel81b, Pel81a].

**Théorème 1.6.2** *Sous les hypothèses (HS), (HFG) et (HI), l'inclusion (1.6.1) admet une solution d'évolution faible.*

Dans le cas où  $\mathbb{H}$  et  $\mathbb{U}$  sont de dimension finie, une adaptation facile de la preuve de ce théorème et l'utilisation du point de Steiner montrent que (1.6.1) admet une solution d'évolution forte.

**Idée de la preuve du théorème 1.6.2.** On note  $\Phi$  l'application qui, à tout processus adapté  $X$  à trajectoires continues dans  $\mathbb{H}$  tel que  $\mathbf{E} \int_0^T \|X(s)\|^p ds < +\infty$ , associe l'ensemble des processus de la forme

$$S(t)\xi + \int_0^t S(t-s)f(s) ds + \int_0^t S(t-s)g(s) dW(s),$$

où  $f$  et  $g$  sont des sélections prévisibles de  $(\omega, t) \mapsto F(t, X(\omega, t))$  et  $(\omega, t) \mapsto G(t, X(\omega, t))$  respectivement. Soit  $\mathbf{C}([0, T]; \mathbb{H})$  l'espace des applications continues de  $[0, T]$  dans  $\mathbb{H}$ . La preuve du théorème 1.6.2 repose sur les deux lemmes qui suivent.

**Lemme 1.6.3** *Soit  $\Lambda$  un ensemble de processus continus  $(\mathcal{F}_t)$ -adaptés sur  $\mathbb{H}$ . On suppose que chaque élément de  $\Lambda$  appartient à  $L^p(\Omega \times [0, T]; \mathbb{H})$  et que  $\Lambda$ , en tant qu'ensemble de variables aléatoires à valeurs dans  $\mathbf{C}([0, T]; \mathbb{H})$ , est tendu. On suppose de plus que  $G$  vérifie (HFG)-(i). Alors  $\Phi(\Lambda)$  est un ensemble tendu d'éléments aléatoires de  $\mathbf{C}([0, T]; \mathbb{H})$ .*

Si  $(\mathbb{E}, d)$  est un espace métrique et  $\Lambda \subset \mathbb{E}$ , on dit qu'une partie  $\Lambda'$  de  $\mathbb{E}$  est un  $\epsilon$ -réseau de  $\Lambda$  si on a  $\inf_{X \in \Lambda} d(X, \Lambda') \leq \epsilon$  (donc  $\Lambda'$  n'est pas nécessairement contenu dans  $\Lambda$ ). Soit  $\Lambda$  une partie de  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . Pour tout  $s \in [0, T]$ , soit  $\Lambda_s \subset \mathcal{N}_c^p(\mathfrak{F}, [0, s]; \mathbb{H})$  l'ensemble des restrictions à  $[0, s]$  d'éléments de  $\Lambda$ . On note

$$\Psi(\Lambda)(s) := \inf \{ \epsilon > 0; \Lambda_s \text{ has a tight } \epsilon\text{-net in } \mathcal{N}_c^p(\mathfrak{F}, [0, s]; \mathbb{H}) \}.$$

L'ensemble  $\Lambda$  est donc tendu si et seulement si  $\Psi(\Lambda)(T) = 0$ . On note

$$\Psi(\Lambda) := (\Psi(X)(s))_{0 \leq s \leq T}.$$

La famille  $(\Psi(X)(s))_{0 \leq s \leq T}$  est appelée *mesure de non compacité* de  $\Lambda$  (sur le sujet des mesures de non compacité, on peut consulter [AKP<sup>+</sup>92, KOZ01]).

La preuve du lemme ci-dessous s'inspire de [AKP<sup>+</sup>92, Lemma 4.2.6] et fait appel au lemme 1.6.3.

**Lemme 1.6.4** *Soit  $\Lambda$  une partie bornée de  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . On suppose que l'hypothèse (HFG) est vérifiée. Alors on a*

$$\Psi^p(\Phi \circ \Lambda)(t) \leq k \int_0^t L(s, \Psi^p(\Lambda))(s) ds$$

pour une constante  $k$  qui ne dépend que de  $T, p, M_T$  et  $C_{\text{Conv}}$ .

*Première étape de la preuve du théorème 1.6.2* On commence par construire suivant le schéma de Tonelli une suite tendue  $(X_n)$  de solutions approchées : pour chaque entier  $n \geq 1$ , on définit un processus  $\tilde{X}_n$  sur  $[-1, T]$  par en posant  $\tilde{X}_n(t) = 0$  si  $t \leq 0$  et, pour tout  $t \geq 0$ ,

$$\tilde{X}_n(t) = S(t)\xi + \int_0^t S(t - (s - \frac{1}{n}))f_n(s) ds + \int_0^t S(t - (s - \frac{1}{n}))g_n(s) dW(s),$$

où  $f_n : \Omega \times [0, T] \rightarrow \mathbb{H}$  et  $g_n : \Omega \times [0, T] \rightarrow \mathbb{L}$  sont des processus prévisibles et

$$f_n(s) \in F(s, \tilde{X}_n(s - 1/n)) \text{ P-p.s. et } g_n(s) \in G(s, \tilde{X}_n(s - 1/n)) \text{ P-p.s.}$$

On pose alors  $X_n(t) = \xi$  pour  $t \leq 1/n$  et, pour  $t \in [1/n, T]$ ,

$$\begin{aligned} X_n(t) &= \tilde{X}_n(t - 1/n) \\ &= S(t - 1/n)\xi + \int_0^{t-1/n} S(t-s)f_n(s) ds + \int_0^{t-1/n} S(t-s)g_n(s) dW(s). \end{aligned}$$

Un calcul faisant appel à (1.6.2), (1.6.4) et au lemme de Gronwall montre que  $(X_n)$  est borné dans  $L^p(\Omega \times [0, T]; \mathbb{H})$ . Ensuite, à l'aide du lemme 1.6.4, on obtient

$$(\Psi(\cup_n \{X_n\})(t))^p \leq M_T^p \Psi(\cup_n \Phi(X_n)(t))^p \leq M_T^p k \int_0^t L(s, \Psi^p(\cup_n \{X_n\})(s)) ds.$$

On conclut en utilisant (1.6.3) que  $(X_n)$  est un ensemble tendu d'éléments aléatoires de  $\mathbf{C}([0, T]; \mathbb{H})$ .

*Deuxième étape de la preuve du théorème 1.6.2* On construit maintenant une solution faible d'évolution de (1.6.1).

D'après le critère de compacité Prokhorov pour les mesures de Young (proposition 1.3.6 et lemme 1.3.2), on peut extraire une sous-suite de  $(X_n)$  qui converge de manière S-stable vers une mesure de Young  $\mu \in \mathcal{Y}(\Omega, \mathcal{F}, P; \mathbf{C}([0, T]; \mathbb{H}))$ . Pour simplifier l'écriture, on note encore  $(X_n)$  cette sous-suite.

Il est commode de représenter la mesure de Young limite  $\mu$  par une variable aléatoire définie sur un espace probabilisé étendu. Soit  $\mathcal{C}$  la tribu borélienne de  $\mathbf{C}([0, T]; \mathbb{H})$ . Pour chaque  $t \in [0, T]$ , soit  $\mathcal{C}_t$  la sous-tribu de  $\mathcal{C}$  engendrée par  $\mathbf{C}([0, t]; \mathbb{H})$ . On définit une base stochastique  $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$  par

$$\underline{\Omega} = \Omega \times \mathbf{C}([0, T]; \mathbb{H}), \quad \underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{C} \quad \underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{C}_t$$

et on définit  $X_\infty$  sur  $\underline{\Omega}$  par  $X_\infty(\omega, u) = u$ . Il est immédiat que  $\mu$  a pour marge  $L(X_\infty)$  sur  $\mathbb{H}$  et que  $X_\infty$  est  $(\underline{\mathcal{F}}_t)$ -adapté. Les variables aléatoires  $X_n$  peuvent être considérées comme définies sur  $\underline{\Omega}$ , en notant  $X_n(\omega, u) := X_n(\omega)$  ( $n \in \mathbb{N}$ ). De plus,  $X_n$  est  $(\underline{\mathcal{F}}_t)$ -adapté pour chaque  $n$ , et  $W$  est  $(\underline{\mathcal{F}}_t)$ -adapté. D'après une version du théorème de Komlós pour les mesures de Young due à Balder [Bal89a, Bal90], on peut extraire de toute sous-suite de  $(X_n)$  une sous-suite  $(X'_n)$  qui K-converge vers  $\mu$ , c'est-à-dire que l'on a, pour toute sous-suite  $(X''_n)$  de  $(X'_n)$ ,

$$\frac{1}{n} \sum_{i=1}^n \delta_{X''_i(\omega)} \xrightarrow{\text{étroitement}} \mu_\omega \text{ p.S.}$$

(où  $\omega \mapsto \mu_\omega$  est la désintégration de  $\mu$ ). Ceci entraîne que, pour tout  $A \in \mathcal{C}_t$ , l'application  $\omega \mapsto \mu_\omega(A)$  est  $\mathcal{F}_t$ -mesurable. On en déduit facilement que, sous



la probabilité  $\mu$ ,  $W$  est à accroissements indépendants et donc que  $W$  est un processus  $(\mathcal{F}_t)$ -brownien.

Pour obtenir (1.6.5), on choisit des solutions particulières  $f_n$  et  $g_n$ . D'après un résultat de Kucia [Kuc98], on peut trouver des sélections mesurables  $\mathfrak{f}$  et  $\mathfrak{g}$  de  $F$  et  $G$  respectivement telles que  $\mathfrak{f}(t, \cdot)$  et  $\mathfrak{g}(t, \cdot)$  soient continues pour chaque  $t \in [0, T]$ . On note  $\widehat{N} = \mathbb{N} \cup \{\infty\}$  et on pose, pour chaque  $n \in \widehat{N}$  et chaque  $t \in [0, T]$ ,

$$f_n(t) = \mathfrak{f}(t, X_n(t)), \quad g_n(t) = \mathfrak{g}(t, X_n(t)),$$

$$Z_n(t) = -X_n(t) + S(t)\xi + \int_0^{t-1/n} S(t-s)f_n(s) ds + \int_0^{t-1/n} S(t-s)g_n(s) dW(s)$$

(avec  $1/\infty := 0$ ). La suite  $(X_n, W, f_n, g_n)$  converge en loi vers  $(X_\infty, W, f_\infty, g_\infty)$  dans  $\mathbf{C}([0, T]; \mathbb{H}) \times \mathbf{C}([0, T]; \mathbb{U}) \times L^p([0, T]; \mathbb{H}) \times L^p([0, T]; \mathbb{L})$ . En utilisant l'hypothèse  $p > 2$ , on montre que, pour chaque  $t \in [0, T]$ , la suite  $(Z_n(t))$  converge en loi vers  $Z_\infty(t)$ . Mais d'autre part il découle de (HFG)-(i) et du fait que la suite  $(X_n)$  est bornée dans  $L^p(\Omega \times [0, T]; \mathbb{H})$  que  $Z_n(t)$  converge vers 0 en probabilité. Pour tout  $t \in [0, T]$ , on a donc  $Z_\infty(t) = 0$  p.s. Comme  $Z_\infty$  est à trajectoires continues, cela entraîne que  $Z_\infty = 0$  p.s. Donc  $X$  est une solution faible d'évolution de (1.6.1).  $\square$

Une solution d'évolution forte de (1.6.1) a été obtenue par Da Prato et Frankowska [DPF94] dans le cas où  $F$  et  $G$  vérifient une condition de Lipschitz par rapport à la deuxième variable. Cette condition est nettement affaiblie dans le théorème 1.6.2 présenté ici. En revanche, nous avons supposé les applications  $F$  et  $G$  déterministes et à valeurs convexes compactes, alors que dans [DPF94] elles sont aléatoires et à valeurs fermées non nécessairement bornées.

## 1.6.2 Un théorème de la limite centrale

Dans cette partie,  $\mathbb{E}$  est un espace vectoriel topologique localement convexe. On suppose de plus que  $\mathbb{E}$  est lusinien, c'est-à-dire qu'il existe une topologie polonaise sur  $\mathbb{E}$  plus fine que la topologie de  $\mathbb{E}$ . On se donne une suite  $(X_n)_n$  de vecteurs aléatoires de  $\mathbb{E}$ . Pour tout  $n \geq 1$ , on note  $S_n = \sum_{i=1}^n X_i$ . On se donne également une suite  $(C_n)_n$  dans  $]0, +\infty[$  vérifiant

$$(1.6.6) \quad \lim_n C_n = +\infty \quad \text{et} \quad \lim_n \left(1 - \frac{C_{n-m}}{C_n}\right) = 0 \quad \text{pour tout } m \text{ fixé.}$$

Par exemple,  $C_n = \frac{1}{n^\alpha}$  convient pour  $\alpha > 0$ .

Pour tout  $n \geq 1$ , on pose  $Z_n = \frac{S_n}{C_n}$ . On dit que  $(X_n)$  vérifie le *théorème de la limite centrale (TLC)* (pour la suite  $C_n$ ) s'il existe une probabilité  $\nu$  sur  $\mathbb{E}$  telle que  $(L(Z_n))$  converge faiblement vers  $\nu$ . On dit que  $(X_n)$  vérifie le *théorème de la limite centrale stable (TLC stable)* (pour la suite  $C_n$ ) s'il existe une mesure de Young  $\mu \in \mathcal{Y}$  telle que  $(\delta_{Z_n})$  converge vers  $\mu$  pour  $\tau_y^w$ . Le TLC stable entraîne donc le TLC, mais par ailleurs on peut souvent déduire un TLC stable d'un TLC : supposons par exemple que  $(X_n)$  soit une suite  $\alpha$ -mélangeante et que, pour tous  $n$  et  $p$ ,  $S_n$  ait la même loi  $S_{p+n} - S_p$ . Alors, si  $(L(Z_n))$  converge faiblement vers une probabilité  $\nu$ , on montre facilement que  $(\delta_{Z_n})$  converge vers  $P \otimes \nu$  pour  $\tau_y^w$ .

Nous nous intéressons ici au TCL pour un nombre aléatoire de vecteurs aléatoires, c'est-à-dire à la convergence stable ou en loi de  $(Z_{\eta_n})$ , où  $(\eta_n)$  est une suite d'entiers qui converge en probabilité vers  $+\infty$ . Nous allons prouver un résultat de ce type pour une version plus forte du TCL, appelée *TCL fonctionnel*. Dans ce cadre, le théorème 1.5.11 sur le produit fibré permet de réinterpréter dans un cadre topologique général les idées de Billingsley [Bil68, Section 17], développées par Fischler [Fis76] puis Aldous [Ald78].

Auparavant, quelques nouvelles définitions sont nécessaires. On appelle *espace de Skorokhod* et on note  $\mathbf{D}(\mathbb{R}^+; \mathbb{E})$  (ou simplement  $\mathbf{D}$ ) l'ensemble des applications de  $\mathbb{R}^+$  dans  $\mathbb{E}$  qui sont continues à droite et ont en chaque point une limite à gauche. On munit  $\mathbf{D}(\mathbb{R}^+; \mathbb{E})$  d'une topologie séparable complètement régulière appelée  $J_1$  que l'on peut définir de la manière suivante. Soit  $\mathcal{D}$  une famille de semidistances qui définissent la topologie de  $\mathbb{E}$ . On associe à chaque  $d \in \mathcal{D}$  une semidistance  $\tilde{d}$  sur  $\mathbf{D}(\mathbb{R}^+; \mathbb{E})$  de sorte que

1. la topologie  $J_1$  sur  $\mathbf{D}$  est définie par l'ensemble de semidistances  $\tilde{\mathcal{D}} = \{\tilde{d}; d \in \mathcal{D}\}$ ,
2. pour toute suite  $(F_n)_{n \in \mathbb{N}}$  dans  $\mathbf{D}$  et pour tout  $d \in \mathcal{D}$  la suite  $(F_n)_{n \in \mathbb{N}}$  converge dans  $(\mathbf{D}, \tilde{d})$  vers une limite  $F \in \mathbf{D}$  si et seulement s'il existe une suite  $(u_n)_n$  de bijections continues de  $\mathbb{R}^+$  dans  $\mathbb{R}^+$  qui converge uniformément sur les parties bornées de  $\mathbb{R}^+$  vers l'application identique  $\text{Id}_{\mathbb{R}^+}$  et telle que  $(F_n \circ u_n)$  converge vers  $F$   $d$ -uniformément sur les parties bornées de  $\mathbb{R}^+$ .

Cette topologie est une variante, due à Lindvall dans le cas  $\mathbb{E} = \mathbb{R}$ , de la *topologie  $J_1$  de Skorokhod* ([Lin73], voir aussi [Jac85]). La généralisation à un espace complètement régulier  $\mathbb{E}$  est due à Mitoma [Mit83] et Jakubowski [Jak86]. On montre [CRdFV04, Lemma 9.4.4] que, comme  $\mathbb{E}$  est lusilien, l'espace  $\mathbf{D}$  est souslinien.

Un sous-espace intéressant de  $\mathbf{D}(\mathbb{R}^+; \mathbb{E})$  est l'espace  $\mathbf{C}(\mathbb{R}^+; \mathbb{E})$  (ou simplement  $\mathbf{C}$  s'il n'y a pas d'ambiguïté) des applications continues de  $\mathbb{R}^+$  dans

$\mathbb{E}$ . La trace de  $J_1$  sur  $\mathbf{C}$  est la topologie de la convergence uniforme sur les parties bornées de  $\mathbb{R}^+$ . Toutefois,  $\mathbf{C}$  est dense dans  $\mathbf{D}$  pour  $J_1$ , alors que c'est un sous-espace fermé de  $\mathbf{D}$  muni de la topologie de la convergence uniforme sur les parties bornées de  $\mathbb{R}^+$ . On retrouve ainsi une situation semblable à celle du sous espace  $L^0$  de  $\mathcal{Y}_d$ , qui est dense dans  $\mathcal{Y}_d$  pour la topologie stable mais fermé pour la topologie de la convergence en mesure, alors que ces deux topologies ont la même trace sur  $L^0$  (voir la remarque 1.5.5 page 22).

Avec les notations  $(X_n)$ ,  $(S_n)$ ,  $(C_n)$  et  $(Z_n)$  précédentes, on définit une suite  $(G_n)$  d'éléments aléatoires de  $\mathbf{D}$  en posant, pour tout  $n \geq 1$ , tout  $\omega \in \Omega$  et tout  $t \in \mathbb{R}^+$ ,

$$G_n(\omega, t) = G_n(\omega)(t) = \frac{S_{[nt]}(\omega)}{C_n},$$

où  $[nt]$  désigne la partie entière de  $nt$ . On dit alors que  $(X_n)$  vérifie le *principe d'invariance* ou *théorème de la limite centrale fonctionnel* (pour la suite  $C_n$ ) si  $(G_n)_n$  converge en loi dans  $\mathbf{D}(\mathbb{R}^+; \mathbb{E})$  vers une probabilité  $\nu \in \mathbf{P}(\mathbf{D}(\mathbb{R}^+; \mathbb{E}))$ . On a donc alors  $L(Z_n) = L(G_n(\cdot, 1)) \rightarrow \nu_1$ , où  $\nu_1$  est l'image de  $\nu$  par la projection  $F \mapsto F(1)$  de  $\mathbf{D}$  sur  $\mathbb{E}$ .

Si de plus la convergence de  $(G_n)$  vers  $\nu$  est stable, c'est-à-dire s'il existe  $\mu \in \mathcal{Y}(\Omega, \mathcal{F}, \mathbf{P}; \mathbf{D}(\mathbb{R}^+; \mathbb{E}))$  tel que

$$\underline{\delta}_{G_n} \xrightarrow{\text{W-stable}} \mu,$$

on dit alors que  $(X_n)$  vérifie le *principe d'invariance stable* ou *théorème de la limite centrale fonctionnel stable* (pour la suite  $C_n$ ).

**Théorème 1.6.5 (TCL fonctionnel stable pour un nombre aléatoire de vecteurs aléatoires)** *Avec les notations  $(X_n)$ ,  $(S_n)$ ,  $(C_n)$ ,  $(Z_n)$  et  $(G_n)$  précédentes, et en supposant toujours que  $\mathbb{E}$  est lusinien, soit  $\alpha > 0$  et supposons que  $C_n = n^\alpha$  pour chaque  $n \geq 1$ . Soient  $\theta$  une variable aléatoire réelle telle  $\theta > 0$  p.s.,  $(a_n)_n$  une suite de nombres dans  $]0, +\infty[$  convergeant vers  $+\infty$  et  $(\eta_n)$  une suite de variables aléatoires à valeurs dans  $]0, +\infty[$ . On suppose de plus que*

(a)  $(X_n)_{n \geq 1}$  vérifie un TLC fonctionnel stable : il existe  $\gamma \in \mathcal{Y}(\mathbf{D}(\mathbb{R}^+; \mathbb{E}))$  tel que l'on ait

$$G_n \xrightarrow{\text{W-stable}} \gamma,$$

(b)

$$\frac{\eta_n}{a_n} \xrightarrow{\text{probabilité}} \theta.$$

On a alors

$$G_{\eta_n} \xrightarrow{\text{W-stable}} \gamma.$$

La preuve est très simple dans le cas où  $(\eta_n)$  est indépendant de  $(X_n)$ , et dans ce cas l'hypothèse (b) peut être remplacée par  $\eta_n \xrightarrow{\text{probabilité}} +\infty$ .

**Idée de la preuve.** Posons, pour chaque  $(\omega, t) \in \Omega \times \mathbb{R}^+$  et pour chaque  $n \geq 1$ ,

$$\theta_n(\omega) = \frac{\eta_n(\omega)}{a_n} \text{ et } F_n(\omega, t) = F_n(\omega)(t) = \theta_n(\omega) t.$$

On remarque que chaque  $F_n$  prend ses valeurs dans la partie  $\mathbf{C}_0^\infty$  de  $\mathbf{C}$  formée des bijections continues de  $\mathbb{R}^+$  dans  $\mathbb{R}^+$ .

Donnons nous une suite  $(k_n)_n$  dans  $\mathbb{N}^*$  telle que  $\lim_n k_n/a_n = 1$ . On a

$$G_{\eta_n}(\omega)(t) = \frac{S_{[\eta_n(\omega)t](\omega)}}{\eta_n^\alpha(\omega)} \underset{\mathbf{D}}{\sim} \frac{S_{[\frac{\eta_n(\omega)}{a_n} k_n t](\omega)}}{\left(\frac{\eta_n(\omega)}{a_n}\right)^\alpha k_n^\alpha} = \Gamma(F_n(\omega), G_{k_n}(\omega), 1/\theta_n^\alpha(\omega)),$$

où  $\Gamma$  est l'application

$$\begin{cases} \mathbf{C}_0^\infty \times \mathbf{D} \times \mathbb{R} & \rightarrow \mathbf{D} \\ (F, G, \rho) & \mapsto \rho G \circ F. \end{cases}$$

On vérifie [CRdFV04, Lemma 9.4.7] que l'application  $\Gamma$  ainsi définie est continue.<sup>3</sup> Or la suite  $(\theta_n)_n$  converge en probabilité vers  $\theta$  et  $(F_n)_n$  converge en probabilité vers l'élément aléatoire  $F$  de  $\mathbf{C}_0^\infty \subset \mathbf{D}(\mathbb{R}^+, \mathbb{R}^+)$  défini par

$$F(\omega, t) = \theta(\omega)t.$$

On en déduit, à l'aide du théorème du produit fibré (théorème 1.5.11), que la suite  $(\Gamma(F_n, G_{k_n}, 1/\theta_n^\alpha))_{n \geq 1}$  converge de manière W-stable vers la mesure de Young  $\gamma' \in \mathcal{Y}(\Omega, \mathcal{F}, \mathbf{P}; \mathbf{D})$  telle que, pour presque tout  $\omega \in \Omega$ ,

$$\gamma'_\omega = \Gamma_\#(\delta_{F(\omega)} \otimes \gamma_\omega \otimes \delta_{\theta^{-\alpha}(\omega)}) = (\Gamma_{\theta(\omega)})_\#(\gamma_\omega),$$

où, pour tout  $\rho > 0$ ,  $\Gamma_\rho : \mathbf{D} \rightarrow \mathbf{D}$  est l'application définie par

$$\Gamma_\rho(G)(t) = \frac{1}{\rho^\alpha} G(\rho t)$$

tout  $G \in \mathbf{D}$  et tout  $t \in \mathbb{R}^+$ .

---

<sup>3</sup> L'intérêt de se limiter à  $\mathbf{C}_0^\infty$  pour première variable est que, sur cet espace, la topologie de la convergence uniforme sur les parties bornées de  $\mathbb{R}^+$  (donc la topologie  $J_1$ ) coïncide avec la topologie de la convergence simple [CRdFV04, Lemma 9.4.6]. De plus, si une suite  $(F_n)$  converge dans  $\mathbf{C}_0^\infty$  vers une limite  $F \in \mathbf{C}_0^\infty$ , la suite  $(F_n^{-1})$  des applications réciproques converge vers  $F^{-1}$  dans  $\mathbf{C}_0^\infty$ .

Pour prouver le théorème, il reste seulement à démontrer que l'on a  $\gamma' = \gamma$ . Ceci provient [CRdFV04, Lemma 9.4.8] de la forme particulière des coefficients normalisateurs  $C_n = n^\alpha$ . En effet, soit  $F_\rho \in \mathbf{C}_0^\infty$  défini par  $F_\rho(t) = \rho t$  pour tout  $t \in \mathbb{R}^+$ . Avec les notations précédentes, on a  $\Gamma_\rho(G) = \Gamma(F_\rho, G, 1/\rho^\alpha)$  pour tout  $G \in \mathbf{D}$ . Donc  $\Gamma_\rho$  est une application continue. Soit

$$\underline{\Gamma}_\rho : \begin{cases} \Omega \times \mathbf{D} & \rightarrow \Omega \times \mathbf{D} \\ (\omega, G) & \mapsto (\omega, \Gamma_\rho(G)). \end{cases}$$

L'application  $(\underline{\Gamma}_\rho)_\# : \mathcal{Y}(\Omega, \mathcal{F}, \mathbf{P}; \mathbf{D}) \rightarrow \mathcal{Y}(\Omega, \mathcal{F}, \mathbf{P}; \mathbf{D})$  est donc continue pour la convergence W-stable. Supposons que  $\rho = N$  soit un entier. On a, pour tout  $(\omega, t) \in \Omega \times \mathbb{R}^+$  et pour tout  $n \in \mathbb{N}$ ,

$$\Gamma_N(G_n(\omega))(t) = \frac{1}{N^\alpha} \frac{S_{[nNt]}(\omega)}{n^\alpha} = G_{Nn}(\omega)(t).$$

Donc

$$\gamma = \lim_n \underline{\delta}_{G_n} = \lim_n \underline{\delta}_{G_{Nn}} = \lim_n (\underline{\Gamma}_N)_\# (\underline{\delta}_{G_n}) = (\underline{\Gamma}_N)_\# (\gamma).$$

où les convergences ont lieu dans la convergence W-stable. On a donc, lorsque  $\rho \in \mathbb{N}^*$ ,

$$(1.6.7) \quad (\Gamma_\rho)_\# (\gamma_\omega) = \gamma_\omega \quad \text{p.s.}$$

Il reste à montrer que (1.6.7) reste vrai si l'on prend  $\rho$  quelconque dans  $]0, +\infty[$ . On le vérifie d'abord pour  $\rho = 1/N$  avec  $N \in \mathbb{N}^*$ , puis pour  $\rho$  rationnel, puis on obtient le cas général par continuité de  $\rho \mapsto \Gamma_\rho(G)$  pour tout  $G \in \mathbf{D}$ .

□

# Chapitre 2

## Autour de la loi forte des grands nombres

Dans l'étude des mesures de Young, les seules intégrales que nous avons envisagées étaient les intégrales de fonctions à valeurs réelles. Nous nous intéressons ici, à travers l'étude de la loi des grands nombres, à des intégrales plus exotiques : intégrale de fonctions à valeurs dans un espace métrique (section 2.1), puis intégrale de Pettis pour des fonctions à valeurs dans un espace vectoriel topologique localement convexe (section 2.2).

### 2.1 Moyenne et loi des grands nombres dans un espace métrique

#### Barycentre au sens de Herer

Plusieurs définitions ont été proposées pour l'espérance d'une variable aléatoire à valeurs dans un espace métrique  $\mathbb{M}$  : Fréchet [Fré48] en 1948, Doss [Dos49] en 1949, (voir aussi [Dos62], [Ben62], [Her83], [Her91] et [BHL93]), Herer [Her88] en 1988 pour les espaces dits "à courbure négative" (voir aussi [Her] et [Her92]), plus récemment Heinich et Essahib [ESH99] puis K. T. Sturm [Stu02]. Une espérance pour des variables aléatoires à valeurs dans une variété différentielle  $\mathbb{V}$  munie d'une connexion a été définie par Émery et Mokobodzki [EM91], puis une autre par Picard [Pic94]. Dans ces définitions (sauf [Pic94, ESH99, Stu02]), l'espérance d'un élément aléatoire est un fermé de  $\mathbb{M}$  ou de  $\mathbb{V}$ , non nécessairement réduit à un point.

Lorsque l'on sort du cadre habituel des espaces vectoriels, on peut se demander quelles propriétés doit vérifier un objet mathématique associé à une classe d'éléments aléatoires d'un espace métrique  $\mathbb{M}$  pour "mériter" le

nom d’“espérance mathématique”. Une condition que l’on pourrait exiger est certainement que cet objet coïncide avec l’espérance mathématique usuelle (disons, de Bochner) lorsque  $\mathbb{M}$  est un espace vectoriel normé. D’autres conditions intéressantes sont que cet objet vérifie le théorème de convergence dominée, le théorème de Fubini ou la loi des grands nombres. Toutes les notions de barycentre évoquées plus haut vérifient la première condition. La situation est plus compliquée au regard des autres conditions. Le problème de savoir si l’espérance de Doss vérifie la loi des grands nombres est toujours ouvert.

Nous nous intéressons ici à l’espérance de Herer et à la loi des grands nombres. Commençons par la construction de cet objet. La présentation qui suit est un peu différente de celle de Herer [Her88, Her92] et de celle de [RdF97]. Dans ce qui suit,  $\mathbb{M}$  est un espace métrique, et on notera  $d(x, y) = [x, y]$  la distance entre deux points  $x$  et  $y$  de  $\mathbb{M}$ . Si  $A$  et  $B$  sont deux parties de  $\mathbb{M}$ , on notera également  $[A, B]$  la *distance de Hausdorff* entre  $A$  et  $B$ , c’est-à-dire

$$\begin{aligned} [A, B] &= \max \left\{ \sup_{x \in A} d(x, B), \sup_{y \in B} d(y, A) \right\} \\ &= \sup_{z \in \mathbb{M}} |d(z, A) - d(z, B)| \end{aligned}$$

([Cor73], § 2, déf. 1.1. et prop. 1.2.).

Habituellement, la construction de l’espérance mathématique d’une variable aléatoire (disons, comme intégrale de Riemann, ou de Lebesgue ou de Bochner) se fait en deux étapes :

- une étape algébrique, où l’on définit l’espérance d’une variable ne prenant qu’un nombre fini de valeurs,
- une étape topologique, c’est-à-dire un passage à la limite.

Dans le cas d’un espace métrique sans structure vectorielle, l’idée de Herer [Her88, Her92] suit grosso modo ce schéma. La première étape est réalisée d’une manière itérative très naturelle et suppose peu d’hypothèses sur la structure géométrique de  $\mathbb{M}$  : il suffit que, pour tous points  $a$  et  $b$  de  $\mathbb{M}$  et pour tout  $t \in [0, 1]$ , on puisse trouver  $c \in \mathbb{M}$  tel que  $[b, c] = t[a, b]$  et  $[a, c] = (1 - t)[a, b]$ , on dit alors que  $\mathbb{M}$  est *convexe*. C’est la deuxième étape qui fera appel à la condition géométrique de “courbure négative”.

*Première étape* Commençons par définir la notion de combinaison convexe. Pour tous  $x$  et  $y$  dans  $\mathbb{M}$ , et pour tous  $p_1$  et  $p_2$  dans  $[0, 1]$ , tels que  $p_1 + p_2 = 1$ , on appelle *combinaison convexe* ([Her88], def.1) de  $x$  et  $y$  de poids  $p_1$  et  $p_2$  l’ensemble (éventuellement vide)

$$\{z \in \mathbb{M}; [z, x] = p_2[x, y] \quad \text{et} \quad [z, y] = p_1[x, y]\},$$

que l'on notera  $p_1x + p_2y$ . Cette définition s'étend par récurrence à un système de  $n$  éléments (distincts ou non)  $x_1, \dots, x_n$  de  $\mathbb{M}$  ( $n \in \mathbb{N}, n \geq 2$ ) et de  $n$  poids  $p_1, \dots, p_n$  positifs ou nuls tels que  $\sum_{1 \leq i \leq n} p_i = 1$  : on écrira  $z \in \sum_{1 \leq i \leq n} p_i x_i$  s'il existe une partition  $(I_1, I_2)$  de  $\{1, \dots, n\}$  telle que

$$z \in \left( \sum_{i \in I_1} p_i \right) \sum_{j \in I_1} \frac{p_j}{\sum_{i \in I_1} p_i} x_j + \left( \sum_{i \in I_2} p_i \right) \sum_{j \in I_2} \frac{p_j}{\sum_{i \in I_2} p_i} x_j.$$

Ainsi, dans la figure 2.1, la combinaison convexe de  $(x, p_1)$ ,  $(y, p_2)$  et  $(z, p_3)$  est formée des points

$$\begin{aligned} x' &= p_1x + (1 - p_1) \left( \frac{p_2}{1 - p_1} y + \frac{p_3}{1 - p_1} z \right), \\ y' &= p_2y + (1 - p_2) \left( \frac{p_1}{1 - p_2} x + \frac{p_3}{1 - p_2} z \right), \\ z' &= p_3z + (1 - p_3) \left( \frac{p_1}{1 - p_3} x + \frac{p_2}{1 - p_3} y \right) \end{aligned}$$

(on a supposé dans la figure que la combinaison convexe de deux éléments est toujours réduite à un seul point).

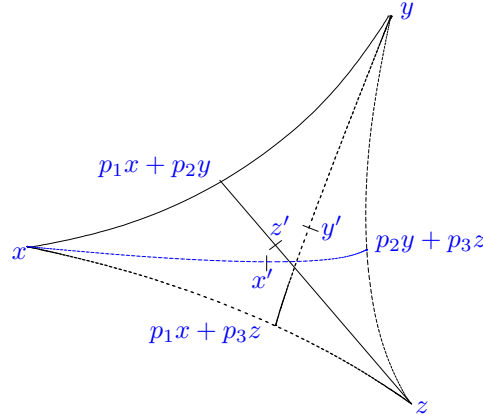


FIG. 2.1 – combinaison convexe de  $(x, p_1)$ ,  $(y, p_2)$  et  $(z, p_3)$

On voit que l'on a, par exemple,

$$\frac{p_1}{2}x + \frac{p_1}{2}x + p_2y + p_3z \supset p_1x + p_2y + p_3z.$$

On appelle *barycentre* du système de points pondérés  $(x_i, p_i)_{1 \leq i \leq n}$  et on note  $\sum_{1 \leq i \leq n}^* p_i x_i$  l'adhérence dans  $\mathbb{M}$  de la réunion de toutes les combinaisons



convexes de la forme  $\sum_{i \in I} q_j y_j$  où  $J$  est un ensemble fini,  $y_j$  prend ses valeurs dans  $\{x_1, \dots, x_n\}$  et, pour tout  $i \in \{1, \dots, n\}$ ,  $\sum_{y_j=x_i} q_j = p_i$ . Si  $\mu$  est la probabilité  $\sum_{1 \leq i \leq n} p_i \delta_{x_i}$  (où  $\delta_{x_i}$  désigne la masse de Dirac en  $x_i$ ),  $\sum_{1 \leq i \leq n}^* p_i x_i$  est appelé le *barycentre de  $\mu$* , ce qui s'écrit

$$\text{bar } \mu = \sum_{1 \leq i \leq n}^* p_i x_i.$$

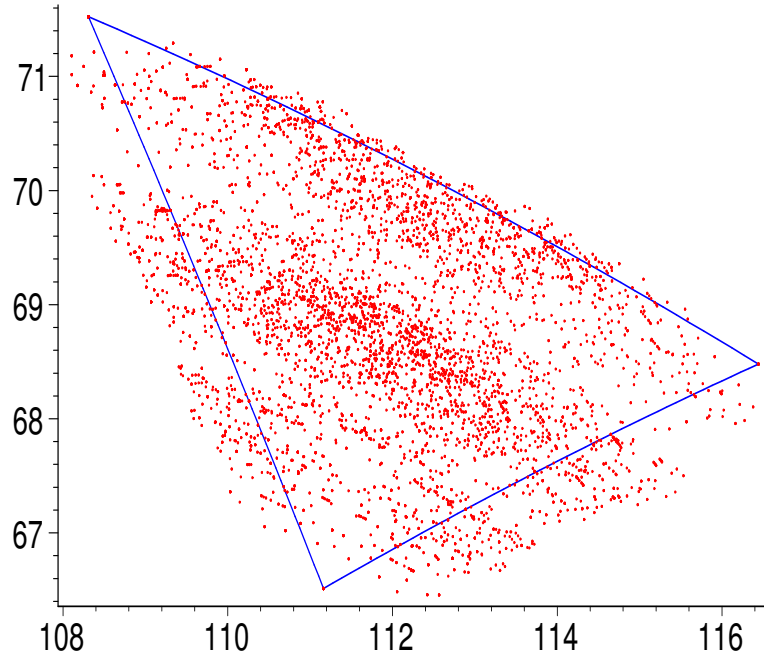


FIG. 2.2 – barycentre au sens de Herer de trois points du plan hyperbolique

La figure 2.2 donne un exemple de calcul approché de l'isobarycentre des points  $x = (100, 10)$ ,  $y = (100, 200)$  et  $z = (200, 100)$  dans le modèle de Poincaré (demi-plan supérieur) du plan hyperbolique. Les sommets du triangle hyperbolique bleu sont les trois éléments de la combinaison convexe de  $(x, 1/3)$ ,  $(y, 1/3)$  et  $(z, 1/3)$ . Les points rouges sont les éléments de la combinaison convexe

$$\frac{1}{9}x + \frac{1}{9}x + \frac{1}{9}x + \frac{1}{9}y + \frac{1}{9}y + \frac{1}{9}y + \frac{1}{9}z + \frac{1}{9}z + \frac{1}{9}z.$$

Cette combinaison convexe contient 8483 points : chaque point est associé à un arbre à 9 feuilles, mais de nombreux arbres sont associés au même point, car le barycentre de deux points est commutatif et, de plus, le barycentre de  $(x, \alpha)$  et de  $(x, \beta)$  vaut  $x$ .

Comme le barycentre est un espace convexe pour la métrique induite, il serait intéressant de pouvoir déterminer ses points frontière ou ses points extrêmes pour diminuer la taille des calculs.

*Deuxième étape* Il reste à étendre cette notion de barycentre aux mesures diffuses sur  $\mathbb{M}$ . On dit que l'espace métrique convexe  $\mathbb{M}$  est à *courbure négative* ([Her88], [Her92], voir aussi les définitions de Busemann [Bus48], [Bus55] et de Gromov [Gro75, Gro99]) si pour tous  $x, y, x', y' \in \mathbb{M}$ ,  $p \in [0, 1]$ ,  $z \in px + (1 - p)y$  et  $z' \in px' + (1 - p)y'$ , on a

$$(2.1.1) \quad [z, z'] \leq p[x, x'] + (1 - p)[y, y'].$$

On vérifie facilement que la région simplement connexe du plan représentée sur la figure 2.3, munie de la *distance géodésique* est un espace à courbure négative (rappelons que la distance géodésique entre deux points d'une partie  $B$  du plan euclidien est égale à la longueur pour la distance euclidienne du plus court chemin contenu dans  $B$  reliant ces deux points) .

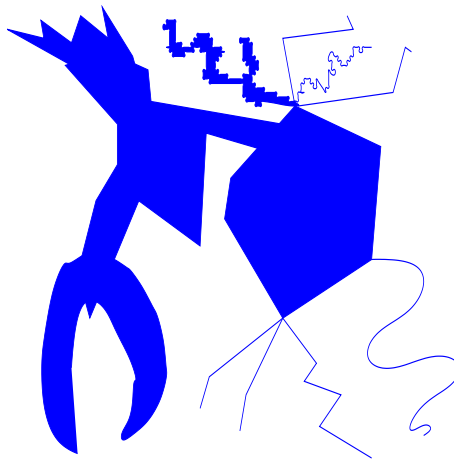


FIG. 2.3 – un espace à courbure négative

Busemann [Bus48] a prouvé en 1948 que les variétés riemanniennes simplement connexes, munies de leur distance géodésique, sont des espaces métriques à courbure négative (au sens de la définition ci-dessus) si et seulement si leur courbure sectionnelle est négative. D'autre part, les espaces de Banach strictement convexes sont également des espaces métriques à courbure négative. Or, d'après un résultat de Clarkson [Cla36], tout espace de Banach séparable peut être muni d'une norme équivalente qui le rend strictement convexe. De plus, dans un espace de Banach strictement convexe, la combinaison convexe (au sens défini ci-dessus)  $\sum_{i \in I} p_i x_i$  a pour unique élément

la combinaison linéaire habituelle (au sens algébrique) des  $x_i$  avec les coefficients  $p_i$ , qui, elle, ne dépend pas de la norme. Il s'ensuit que les résultats que nous présentons ici pour les espaces métriques séparables complets à courbure négative généralisent les résultats analogues pour les espaces de Banach séparables, obtenus en prenant (par exemple)  $a = 0$  et en remplaçant, pour tout  $x \in \mathbb{M}$ , la distance  $[a, x]$  par la norme  $\|x\|$ .

On supposera désormais que  $\mathbb{M}$  est à courbure négative. Cette propriété permet de prolonger par densité la définition du barycentre. On remarque d'abord que, si  $(x_i)_{i \in I}$  et  $(x'_i)_{i \in I}$  sont deux familles finies de points de  $\mathbb{M}$  et si  $(p_i)_{i \in I}$  est une famille d'éléments de  $[0, 1]$  telle que  $\sum_{i \in I} p_i = 1$ , alors

$$(2.1.2) \quad \left[ \sum_{i \in I} p_i x_i, \sum_{i \in I} p_i x'_i \right] \leq \sum_{i \in I} p_i [x_i, x'_i].$$

Pour  $\mu, \nu \in P(\mathbb{M})$ , rappelons (voir page 23) que l'on note

$$D(\mu, \nu) = \{ \pi \in P(\mathbb{M} \times \mathbb{M}); \pi(\cdot \times \mathbb{M}) = \mu \text{ et } \pi(\mathbb{M} \times \cdot) = \nu \}.$$

Appelons *probabilité intégrable* sur  $\mathbb{M}$  toute mesure de probabilité  $\mu \in P(\mathbb{M})$  vérifiant

$$\int_{\mathbb{M}} [a, \cdot] d\mu < +\infty$$

pour un  $a \in \mathbb{M}$  (ou, ce qui revient au même, pour tout  $a \in \mathbb{M}$ ). On notera  $P^1(\mathbb{M})$  l'ensemble des probabilités intégrables sur  $\mathbb{M}$ . On munit  $P^1(\mathbb{M})$  de la *distance de Wasserstein* définie par

$$\Delta_W(\mu, \nu) = \inf_{\pi \in D(\mu, \nu)} \int_{\mathbb{M} \times \mathbb{M}} [x, y] d\pi(x, y).$$

Notons  $P^e(\mathbb{M})$  les éléments de  $P(\mathbb{M})$  qui sont des combinaisons convexes d'un nombre fini de masses de Dirac. On a donc  $P^e(\mathbb{M}) \subset P^1(\mathbb{M})$ , et de plus  $P^e(\mathbb{M})$  est dense dans  $P^1(\mathbb{M})$  pour  $\Delta_W$ . Or on déduit facilement de la formule (2.1.2) que

$$(\forall \mu, \nu \in P^e(\mathbb{M})) \quad [\text{bar } \mu, \text{bar } \nu] \leq \Delta_W(\mu, \nu).$$

L'application  $\mu \mapsto \text{bar } \mu$  définie sur  $P^e(\mathbb{M})$  étant lipschitzienne pour  $\Delta_W$  et la distance de Hausdorff, on peut donc la prolonger par continuité à  $P^1(\mathbb{M})$ . C'est ce prolongement que l'on appellera *barycentre (au sens de Herer)*. On obtient ainsi immédiatement la propriété fondamentale

$$(2.1.3) \quad (\forall \mu, \nu \in P^1(\mathbb{M})) \quad [\text{bar } \mu, \text{bar } \nu] \leq \Delta_W(\mu, \nu).$$

## Lois des grands nombres dans un espace métrique

**Loi forte des grands nombres à la Cuesta et Matrán** Cuesta et Matrán ont prouvé de deux manières différentes, dans [CM88] et [CM92], une loi forte des grands nombres très générale pour des éléments aléatoires d'un espace de Banach séparable. Chacune de leurs deux méthodes s'adapte au cas de points aléatoires de  $\mathbb{M}$ , moyennant quelques changements mineurs, dont le principal est le remplacement de  $\|x\|$  par  $[a, x]$ ... La première méthode utilise une généralisation par Blackwell et Dubins [BD83] du théorème de Skorokhod<sup>1</sup>, tandis que la deuxième fait appel aux propriétés de la distance de Wasserstein. On obtient ainsi le théorème très général ci-dessous.

### **Théorème 2.1.1 (Loi des grands nombres à la Cuesta et Matrán)**

Soit  $(X_n)_{n \geq 1}$  une suite de points aléatoires intégrables de  $\mathbb{M}$ . On se donne arbitrairement un point  $a$  de  $\mathbb{M}$  et on suppose que les trois conditions suivantes sont réalisées.

- (i) La suite  $\left( \sum_{1 \leq i \leq n} \frac{1}{n} L(X_i) \right)_n$  est tendue (on note  $L(X)$  la loi d'une variable aléatoire  $X$ );
- (ii) la suite  $([a, X_n])_n$  vérifie la loi forte des grands nombres :

$$\text{p.s.} \quad \sum_{1 \leq i \leq n} \frac{1}{n} [a, X_i(\omega)] \xrightarrow{n \rightarrow \infty} 0;$$

- (iii) l'application  $x \mapsto [a, x]$  est uniformément intégrable par rapport à

$$\left( \sum_{1 \leq i \leq n} \frac{1}{n} L(X_i) \right)_n, \text{ c'est-à-dire}$$

$$\lim_{\alpha \rightarrow +\infty} \sup_n \int_{\{[a, \cdot] > \alpha\}} [a, x] d \left( \sum_{1 \leq i \leq n} \frac{1}{n} L(X_i) \right) (x) = 0$$

$$\text{(ce qui revient à } \lim_{\alpha \rightarrow +\infty} \sup_n \int_{\{[a, X_i] > \alpha\}} [a, X_i] dP = 0 \text{)}.$$

---

<sup>1</sup>La démonstration de [BD83] comporte une faille, remarquée par P.A. Meyer, voir [RdF97]. Toutefois le résultat effectivement démontré par Blackwell et Dubins suffit aux besoins de la méthode de Cuesta et Matrán. Par ailleurs, une preuve complète du résultat annoncé dans [BD83] a été apportée par Fernique [Fer88] par une méthode complètement différente.

Alors

$$\text{p.s.} \quad \left[ \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i(\omega), \text{ bar} \left( \sum_{1 \leq i \leq n} \frac{1}{n} L(X_i) \right) \right] \xrightarrow{n \rightarrow \infty} 0.$$

Il est à noter que ce théorème contient la plupart des lois des grands nombres connues pour des vecteurs aléatoires Bochner intégrables d'un espace de Banach séparable (sans condition géométrique sur l'espace), en particulier la loi des grands nombres d'Ettemadi pour des vecteurs aléatoires équadistribués deux à deux indépendants.

La démonstration du théorème 2.1.1 se fait en deux étapes. La première repose sur une version du lemme de Glivenko Cantelli [CM88, Theorem 2], dont la forme la plus générale (ci-dessous) est donnée dans [CRdF00]. On dit que deux suites  $(x_n)$  et  $(y_n)$  d'un espace topologique  $\mathbb{T}$  sont (*topologiquement équivalentes*) si, pour toute sous-suite  $(x_{n_k})$  qui converge dans  $\mathbb{T}$  vers une limite  $x$ , la sous-suite  $(y_{n_k})$  ayant les mêmes indices converge vers  $x$  et si, pour toute sous-suite  $(y_{n_k})$  qui converge dans  $\mathbb{T}$  vers une limite  $y$ , la sous-suite  $(x_{n_k})$  converge vers  $y$ . On dira en particulier que deux suites  $(P_n)$  et  $(Q_n)$  dans  $\mathcal{P}(\mathbb{T})$  sont *faiblement équivalentes* si elles sont équivalentes pour la topologie faible sur  $\mathcal{P}(\mathbb{T})$ . De même, deux suites  $(P_n)$  et  $(Q_n)$  dans  $\mathcal{P}^1(\mathbb{M})$  sont dites  $\Delta_W$ -*équivalentes* si  $(P_n)$  et  $(Q_n)$  si elles sont équivalentes pour la topologie induite par la distance  $\Delta_W$  de Wasserstein.

**Lemme 2.1.2 (Glivenko–Cantelli)** *Soit  $\mathbb{T}$  un espace topologique tel qu'il existe un ensemble dénombrable  $\mathcal{H}$  de fonctions continues bornées de  $\mathbb{T}$  dans  $\mathbb{R}$  qui sépare les points de  $\mathbb{T}$  (c'est le cas si, par exemple,  $\mathbb{T}$  est un espace cosmique régulier). Soit  $(X_n)$  une suite d'éléments aléatoires deux à deux indépendants de  $\mathbb{T}$  définis sur  $(\Omega, \mathcal{F}, \mathbb{P})$ . On suppose que la suite  $(P_n)_n = (1/n \sum_{i=1}^n L(X_i))_n$  est tendue. Alors,  $\mathbb{P}$ -presque sûrement, la suite  $(Q_n^\omega)_n = (1/n \sum_{i=1}^n \delta_{X_i(\omega)})_n$  est tendue et  $(P_n)_n$  et  $(Q_n^\omega)_n$  sont faiblement équivalentes.*

C'est en fait la seule étape probabiliste. Dans la démonstration du lemme 2.1.2 intervient la loi des grands nombres pour des variables réelles (celle d'Ettemadi [Ete83] ou celle de Csörgő–Tandori–Totik [CTT83]), appliquée à des fonctions indicatrices de compacts.

La deuxième étape utilise une caractérisation bien connue [Rac82] de la convergence pour la distance de Wasserstein : on sait qu'une suite  $(P_n)$  dans  $\mathcal{P}^1(\mathbb{M})$  converge pour la distance  $\Delta_W$  vers une probabilité  $P$  si et seulement si  $(P_n)$  converge faiblement vers  $P$  et, pour un point  $a \in \mathbb{M}$  (ou pour tout point  $a \in \mathbb{M}$ ), la fonction  $[a, \cdot]$  est uniformément intégrable par rapport à  $(P_n)$ . D'après les hypothèses (ii) et (iii) du théorème 2.1.1, la fonction

$[a, \cdot]$  est uniformément intégrable par rapport à  $(P_n)_n = (1/n \sum_{i=1}^n L(X_i))_n$ , et presque sûrement uniformément intégrable par rapport à  $(Q_n^\omega)_n = (1/n \sum_{i=1}^n \delta_{X_i(\omega)})_n$ . On en déduit que  $(P_n)$  et  $(Q_n^\omega)$  sont presque sûrement  $\Delta_W$ -équivalentes. Le théorème en découle immédiatement par application de la propriété fondamentale (2.1.3).

**Autre approche : transport dans un espace de type  $\ell^1$**  Pour démontrer des lois fortes des grands nombres, on peut également commencer par approcher de manière uniforme les points aléatoires de  $\mathbb{M}$  par des points aléatoires d'une partie dénombrable dense  $\mathbb{D} = \{e_0, \dots, e_k, \dots\}$  de  $\mathbb{M}$  : on construit facilement pour tout entier  $N \geq 1$  une application  $\psi^N : \mathbb{M} \rightarrow \mathbb{D}$  telle que  $[\psi^N(x), x] \leq 1/N$  pour tout  $x \in \mathbb{M}$ . D'autre part, toute probabilité  $\mu \in \mathcal{P}(\mathbb{D})$  s'identifie à un élément de l'espace

$$\mathcal{S} = \{(p_k) \in (\mathbb{R}^+)^{\mathbb{N}}; \sum_{k=0}^{\infty} p_k = 1\}$$

via la bijection  $\varphi : \mathcal{P}(\mathbb{D}) \rightarrow \mathcal{S}$  définie par  $\varphi(\mu) = (\mu(e_k))_k$ . Donnons nous une fois pour toutes un point  $a$  de  $M$ , que l'on supposera pour simplifier dans  $\mathbb{M} \setminus \mathbb{D}$ . Pour toute suite  $(p_k) \in \mathbb{R}^{\mathbb{N}}$ , on note

$$\|(p_k)\|_a = \sum_{k \in \mathbb{N}} |p_k| [a, e_k].$$

Soit  $\ell_a^1$  l'ensemble des suites  $(p_k)$  telles que  $\|(p_k)\|_a < +\infty$ . Comme  $a \notin \mathbb{D}$ , l'espace  $\ell_a^1$  est un espace de Banach pour la norme  $\|\cdot\|_a$ . De plus, on a

$$\mathcal{P}^1(\mathbb{D}) = \varphi(\ell_a^1 \cap \mathcal{S}).$$

On montre que l'espace  $\ell_a^1 \cap \mathcal{S}$  ainsi que sa topologie induite par  $\ell_a^1$  sont indépendants du choix de  $a \notin \mathbb{D}$  [RdF97, Proposition 2.1 et Remarque 2.1].

L'application  $\varphi$  permet de transporter sur  $\mathcal{P}^1(\mathbb{D})$  la métrique de  $\ell_a^1$  : pour  $\mu = \sum_k p_k \delta_{e_k}$  et  $\nu = \sum_k p'_k \delta_{e_k}$ , on pose

$$[\mu, \nu]_a = \|(p_k - p'_k)\|_a = \sum_k |p_k - p'_k| [a, e_k].$$

Le lemme ci-dessous montre que le barycentre est une application lipschitzienne sur  $\mathcal{P}^1(\mathbb{D})$  pour  $[\cdot, \cdot]_a$ .

**Lemme 2.1.3** *Soient  $(p_k)_k$  et  $(p'_k)_k$  deux éléments de  $\mathcal{S}$  et soit  $a$  un point de  $\mathbb{M}$ . On a*

$$\left[ \sum_k^* p_k e_k, \sum_k^* p'_k e_k \right] \leq \sum_k |p_k - p'_k| [a, e_k].$$

Montrons sur un exemple comment on peut prouver dans notre cadre métrique, à l'aide des applications  $\psi^N$  et  $\varphi$  une loi forte des grands nombres connue dans le cadre des espaces de Banach séparables (celle ci-dessous provient de [BG70]).

**Théorème 2.1.4 (lois fortes des grands nombres à la Beck et Giesy)**

Soit  $(X_n)_{n \geq 1}$  une suite indépendante de points aléatoires intégrables de  $\mathbb{M}$  vérifiant l'une des conditions suivantes :

$$(i) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq i \leq n} (\text{Var}[a, X_i])^{1/2} = 0 \quad \text{et} \quad \sum_{n \geq 1} \frac{\text{Var}[a, X_n]}{n^2} < +\infty$$

(où, pour une variable aléatoire réelle intégrable  $f$ ,  $\text{Var } f = \mathbf{E}(f - \mathbf{E} f)^2$ ),

$$(ii) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq i \leq n} \text{ess sup}[a, X_i] = 0.$$

Alors

$$\text{p.s.} \quad \left[ \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i(\omega), \text{bar} \left( \sum_{1 \leq i \leq n} \frac{1}{n} L(X_i) \right) \right] \xrightarrow{n \rightarrow \infty} 0.$$

**Idée de la preuve.**

Première étape : on suppose les  $X_n$  à valeurs dans  $\mathbb{D}$ . Pour tout  $\omega \in \Omega$ , tout  $n \in \mathbb{N}^*$  et tout  $k \in \mathbb{N}$  on note

$$Y_{n,k}(\omega) = \delta_{e_k}(X_n(\omega)) \quad (\text{i.e. } Y_{n,k}(\omega) = 1 \text{ si } X_n(\omega) = e_k, 0 \text{ sinon}),$$

$$Z_{n,k}(\omega) = \frac{1}{n} \sum_{i=1}^n Y_{i,k}(\omega) = \frac{1}{n} \text{card}\{i \in \{1, \dots, n\} / X_i(\omega) = e_k\},$$

puis, pour tout  $\omega \in \Omega$  et tout  $n \in \mathbb{N}^*$ ,

$$Y_n(\omega) = (Y_{n,k}(\omega))_k \in \ell_a^1 \cap \mathcal{S}$$

$$Z_n(\omega) = (Z_{n,k}(\omega))_k = \frac{1}{n} \sum_{i=1}^n Y_i(\omega) \in \ell_a^1 \cap \mathcal{S}.$$

Notons encore, pour tout  $n \in \mathbb{N}^*$  et tout  $k \in \mathbb{N}$ ,

$$p_{n,k} = \mathbf{P}(X_n = e_k) = \mathbf{E} Y_{n,k},$$

$$q_{n,k} = \mathbf{E} \left( \sum_{1 \leq i \leq n} \frac{1}{n} Y_{i,k} \right) = \mathbf{E} \left( \frac{1}{n} \text{card}\{i \in \{1, \dots, n\} / X_i = e_k\} \right),$$

et, pour tout  $n \in \mathbb{N}^*$ ,

$$p_n = (p_{n,k})_k = \varphi(\mathbf{L}(X_n)) \quad \text{et} \quad q_n = (q_{n,k})_k = \varphi \left( \sum_{1 \leq i \leq n} \frac{1}{n} \mathbf{L}(X_i) \right) = \mathbf{E} Z_n.$$

On vérifie alors que  $(Y_n)$  satisfait la loi des grands nombres de Beck et Giesy [BG70], c'est-à-dire

$$\text{p.s.} \quad \|Z_n - q_n\|_a \xrightarrow[n \rightarrow \infty]{} 0.$$

Or, d'après le lemme 2.1.3, on a

$$\begin{aligned} \left[ \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i(\omega), b \left( \sum_{1 \leq i \leq n} \frac{1}{n} \mathbf{L}(X_i) \right) \right] &= \left[ \sum_{k \in \mathbb{N}}^* Z_{n,k}(\omega) e_k, \sum_{k \in \mathbb{N}}^* q_{n,k} e_k \right] \\ &\leq \|Z_n - q_n\|_a, \end{aligned}$$

d'où le résultat.

*Deuxième étape : cas général.* On montre que la suite  $(X_n^n)_n = (\psi^n(X_n))_n$  vérifie les hypothèses du théorème. D'après la première étape, on a donc

$$\text{p.s.} \quad \left[ \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i^i(\omega), \text{bar} \left( \sum_{1 \leq i \leq n} \frac{1}{n} \mathbf{L}(X_i^i) \right) \right] \xrightarrow[n \rightarrow \infty]{} 0.$$

Or on a

$$\begin{aligned} \left[ \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i(\omega), \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i^i(\omega) \right] \\ \leq \sum_{1 \leq i \leq n} \frac{1}{n} [X_i(\omega), X_i^i(\omega)] \leq \frac{1}{n} \sum_{1 \leq i \leq n} \frac{1}{i} \xrightarrow[n \rightarrow \infty]{} 0 \end{aligned}$$

et

$$\begin{aligned} \left[ \text{bar} \left( \sum_{1 \leq i \leq n} \frac{1}{n} \mathbf{L}(X_i^i) \right), \text{bar} \left( \sum_{1 \leq i \leq n} \frac{1}{n} \mathbf{L}(X_i) \right) \right] \\ \leq \sum_{1 \leq i \leq n} \frac{1}{n} \mathbf{E} [X_i^i, X_i] \leq \frac{1}{n} \sum_{1 \leq i \leq n} \frac{1}{i} \xrightarrow[n \rightarrow \infty]{} 0, \end{aligned}$$

d'où l'on déduit immédiatement le résultat.  $\square$



**Loi forte des grands nombres i.i.d. et théorème ergodique** Dans le cas de variables aléatoires équidistribuées, la méthode précédente permet d'obtenir un résultat faisant intervenir les combinaisons convexes  $\sum_{1 \leq i \leq n} \frac{1}{n} X_i(\omega)$  au lieu des barycentres  $\sum_{1 \leq i \leq n}^* \frac{1}{n} X_i(\omega)$ . L'intérêt pratique dans les calculs est évident, voir page 42.

Nous utiliserons également une notion de convergence plus faible : on dira qu'une suite  $(F_n)$  de fermés de  $\mathbb{M}$  converge *au sens de Wijsman* vers  $F \in \mathfrak{F}(\mathbb{M})$  si, pour tout  $a \in \mathbb{M}$ ,

$$\lim_{n \rightarrow \infty} |d(a, F_n) - d(a, F)| = 0.$$

On écrira alors

$$F_n \xrightarrow[n \rightarrow \infty]{\text{WIJSMAN}} F$$

(voir [Bee94] ou [Cor73] pour les propriétés de cette convergence). Écrivons également

$$F_n \xrightarrow[n \rightarrow \infty]{\text{HAUSDORFF}} F$$

si  $(F_n)$  converge vers  $F$  pour la distance de Hausdorff.

**Théorème 2.1.5 (théorème ergodique)** *Soit  $(X_n)$  une suite stationnaire de points aléatoires intégrables de  $\mathbb{M}$ . Soit  $\mathcal{I}$  la tribu des événements invariants par  $(X_n)$ . Alors*

$$\text{p.s.} \quad \sum_{1 \leq i \leq n}^* \frac{1}{n} X_i(\omega) \xrightarrow[n \rightarrow \infty]{\text{HAUSDORFF}} \text{bar}(\text{L}(X_1/\mathcal{I})(\omega)),$$

$$\text{p.s.} \quad \sum_{1 \leq i \leq n} \frac{1}{n} X_i(\omega) \xrightarrow[n \rightarrow \infty]{\text{WIJSMAN}} \text{bar}(\text{L}(X_1/\mathcal{I})(\omega)).$$

*Si, de plus,  $\mathbb{M}$  est à boules fermées compactes,*

$$\text{p.s.} \quad \sum_{1 \leq i \leq n} \frac{1}{n} X_i(\omega) \xrightarrow[n \rightarrow \infty]{\text{HAUSDORFF}} \text{bar}(\text{L}(X_1/\mathcal{I})(\omega)).$$

Le dernier résultat, pour  $\mathbb{M}$  à boules fermées compactes, a été obtenu par Herer [Her92] dans le cas où  $(X_n)$  est i.i.d. (dans ce cas la limite est bien sûr  $\text{bar} \text{L}(X_1)$ ).

La preuve du théorème 2.1.5 repose sur un raffinement de la méthode précédemment décrite d'approximation uniforme sur  $\mathbb{D}$  (par les applications  $\psi^N$ ) et transport dans  $\ell_a^1$  (par l'application  $\varphi$ ).

## 2.2 Loi des grands nombres pour des fonctions Pettis intégrables

Il est bien connu qu'une suite indépendante équadistribuée  $(X_n)_n$  de vecteurs aléatoires d'un espace de Banach vérifie la loi des grands nombres si et seulement si les  $X_n$  sont Bochner intégrables. Pour une suite  $(X_n)_n$  non nécessairement équadistribuée, la plupart des lois fortes des grands nombres connues supposent également l'intégrabilité au sens de Bochner des  $X_n$ . La loi des grands nombres présentée dans cette section [CRdF00] échappe à cette règle, puisque la condition d'intégrabilité exigée est celle plus faible au sens de Pettis. Il s'agit d'une généralisation de la loi des grands nombres de Cuesta et Matrán [CM88], dont nous avons donné plus haut une version pour les espaces métriques à courbure négative (théorème 2.1.1 page 45).

Énonçons dans le théorème ci-dessous un résultat équivalent à la loi des grands nombres de Cuesta et Matrán, sous une forme qui nous sera commode pour mener à bien une généralisation aux variables Pettis intégrables.

**Théorème ([CM88])** *Soit  $(X_n)_n$  une suite de vecteurs aléatoires deux à deux indépendants Bochner intégrables d'un espace de Banach séparable  $\mathbb{E}$ , définis sur un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ , tels que*

- (i) *la suite  $(P_n)_n = (1/n \sum_{i=1}^n L(X_i))_n$  est tendue (on note  $L(X)$  la loi d'une variable aléatoire  $X$ ),*
- (ii) *la fonction  $\|\cdot\|$  est uniformément intégrable par rapport à  $(P_n)_n$ , c'est-à-dire*

$$\lim_{a \rightarrow +\infty} \sup_{n \in \mathbb{N}^*} \int_{\{\|\cdot\| > a\}} \|x\| dP_n(x) = 0,$$

- (iii) *pour presque tout  $\omega \in \Omega$ ,  $\|\cdot\|$  est uniformément intégrable par rapport à la suite de lois empiriques  $(1/n \sum_{i=1}^n \delta_{X_i(\omega)})_n$  (où  $\delta_x$  désigne la masse de Dirac au point  $x$ ),*

*Alors  $(X_n)_n$  vérifie la loi des grands nombres, c'est-à-dire*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n (X_i - \mathbb{E}X_i) = 0 \text{ p.s.}$$

**Première généralisation** Les conditions (ii) et (iii) sont du type

$$\|\cdot\| \text{ est uniformément intégrable par rapport à } (Q_n)$$

pour une certaine suite  $(Q_n)$  de probabilités sur  $\mathbb{E}$ , c'est-à-dire

$$\lim_{a \rightarrow +\infty} \sup_{n \in \mathbb{N}} \int_{\mathbb{E}} \sup_{\substack{x' \in \mathbb{E}' \\ \|x'\| \leq 1}} |\langle x, x' \rangle| \mathbf{1}_{\{\|\cdot\| > a\}}(x) dQ_n(x) = 0.$$

On obtient une condition plus faible en permutant le deuxième supremum avec l'intégrale :

$$\lim_{a \rightarrow +\infty} \sup_{n \in \mathbb{N}} \sup_{\substack{x' \in \mathbb{E}' \\ \|x'\| \leq 1}} \int_{\mathbb{E}} |\langle x, x' \rangle| \mathbf{1}_{\{|\langle x', \cdot \rangle| > a\}}(x) dQ_n(x) = 0,$$

ce qui revient à

$$(2.2.1) \quad B_{\mathbb{E}'}(0, 1) \text{ est uniformément intégrable par rapport à } (Q_n),$$

où  $B_{\mathbb{E}'}(0, 1)$  désigne la boule unité du dual  $\mathbb{E}'$  de  $\mathbb{E}$ . Lorsque les  $Q_n$  sont les lois des  $X_n$  (ou des combinaisons convexes de ces lois, comme dans la condition (ii)), on voit que la nouvelle condition (2.2.1) ne nécessite plus que les  $X_n$  soient Bochner intégrables : cette condition peut être réalisée si les  $X_n$  sont seulement scalairement intégrables. En fait, l'intégrabilité scalaire et la condition d'uniforme intégrabilité de  $B_{\mathbb{E}'}(0, 1)$  par rapport à chaque  $P(X_n)$  entraîne que chaque  $X_n$  est *Pettis intégrable*, c'est-à-dire que, pour tout  $A \in \mathcal{F}$ , il existe un élément de  $\mathbb{E}$ , noté  $\int_A X_n dP$ , tel que, pour tout  $x' \in \mathbb{E}'$ ,

$$\langle x', \int_A X_n dP \rangle = \int_A \langle x', X_n \rangle dP.$$

(voir [Mus87, Mus91]). On note alors  $\mathbf{E}X = \int_{\Omega} X dP$ . Si  $\mathbb{E}$  est de dimension infinie (et si  $\mathcal{F}$  est suffisamment riche), on peut toujours définir un vecteur aléatoire Pettis intégrable non Bochner intégrable de  $\mathbb{E}$ . En effet, d'après un théorème de Dvoretzky et Rogers [DR50], il existe dans tout espace de Banach  $\mathbb{E}$  de dimension infinie une série  $\sum_n x_n$  qui converge inconditionnellement (c'est-à-dire indépendamment de l'ordre des termes) mais pas absolument (c'est-à-dire  $\sum_n \|x_n\| = +\infty$ ). Par exemple, si  $\mathbb{E}$  est un espace de Hilbert et  $(e_n)$  un système orthonormé, il suffit de poser  $x_n = \frac{1}{n} e_n$ . Soit alors  $(p_n)$  une suite de nombres réels tels que  $p_n > 0$  pour tout  $n$  et  $\sum_n p_n = 1$ . Il suffit de construire  $X$  tel que  $X(\omega) = \frac{1}{p_n} x_n$  avec probabilité  $p_n$ . On montre d'ailleurs [Mus91, Theorem 5.1] que tout vecteur aléatoire Pettis intégrable non Bochner intégrable  $U$  est de la forme  $U = X + Y$ , avec  $X$  de la forme ci-dessus et  $Y$  borné.

**Deuxième généralisation** *On suppose désormais que  $\mathbb{E}$  est un espace vectoriel topologique localement convexe séparable. Soit  $\mathfrak{S}$  l'ensemble des parties équicontinues de  $\mathbb{E}'$ . On sait que la topologie de  $\mathbb{E}$  est celle de la convergence uniforme sur les éléments de  $\mathfrak{S}$ , déterminée par les seminormes*

$$N_A(x) = \sup_{x' \in A} |\langle x', x \rangle| \quad (A \in \mathfrak{S}).$$

Pour définir la topologie de  $\mathbb{E}$ , on peut bien entendu restreindre  $\mathfrak{S}$  : par exemple, si  $\mathbb{E}$  est un espace de Banach, il suffit de prendre  $\mathfrak{S} = \{B_{\mathbb{E}'}(0, 1)\}$ . Ainsi, on généralise naturellement (2.2.1) en remplaçant  $B_{\mathbb{E}'}(0, 1)$  par les éléments de  $\mathfrak{S}$ . On dira qu'un ensemble  $\mathfrak{M}$  de probabilités sur  $\mathbb{E}$  est  $\mathfrak{S}$ -uniformément scalairement intégrable si chaque élément de  $\mathfrak{S}$  est uniformément intégrable par rapport à  $\mathfrak{M}$ . Nous obtenons la loi des grands nombres suivante.

**Théorème 2.2.1 (loi des grands nombres pour des vecteurs Pettis intégrables)** *On suppose qu'il existe une suite de fonctions continues de  $\mathbb{E}$  dans  $[0, 1]$  qui sépare les points de  $\mathbb{E}$ . Soit  $(X_n)_n$  une suite de vecteurs aléatoires de  $\mathbb{E}$  deux à deux indépendants et Pettis intégrables, définis sur l'espace probabilisé  $(\Omega, \mathcal{F}, P)$ . On suppose que*

- (i) *la suite  $(1/n \sum_{i=1}^n L(X_i))_n$  est tendue,*
- (ii) *la suite  $(P_n)_n = (1/n \sum_{i=1}^n L(X_i))_n$  est  $\mathfrak{S}$ -uniformément scalairement intégrable.*
- (iii) *pour presque tout  $\omega \in \mathbb{T}$ , la suite  $(Q_n^\omega)_n = (1/n \sum_{i=1}^n \delta_{X_i(\omega)})_n$  est  $\mathfrak{S}$ -uniformément scalairement intégrable.*

*Alors  $(X_n)_n$  vérifie la loi forte des grands nombres, c'est-à-dire*

$$p.s. \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n (X_i - \mathbf{E} X_i) = 0.$$

La preuve du théorème 2.2.1 repose sur lemme 2.1.2 (Glivenko-Cantelli) déjà rencontré, ainsi que sur le lemme suivant.

**Lemme 2.2.2 (convergence faible et uniforme intégrabilité)**

1. *Soit  $\mathbb{T}$  un espace topologique séparable. Soit  $(P_\alpha)_{\alpha \in \mathbb{A}}$  une suite généralisée de probabilités sur  $\mathbb{T}$ , convergeant faiblement vers une probabilité  $P \in \mathcal{P}(\mathbb{T})$ . Soit  $\mathcal{H}$  un ensemble de fonctions continues de  $\mathbb{T}$  dans  $\mathbb{R}$  et supposons que  $\mathcal{H}$  est uniformément intégrable par rapport à  $(P_\alpha)_\alpha$ . Alors  $\mathcal{H}$  est uniformément intégrable par rapport à  $P$  et on a, pour tout  $f \in \mathcal{H}$ ,*

$$(2.2.2) \quad P_\alpha(f) \xrightarrow{\text{faiblement}} P(f).$$

2. *Si de plus  $\mathcal{H}$  est équicontinu, la convergence dans (2.2.2) est uniforme par rapport à  $\mathcal{H}$ .*

**Preuve du théorème 2.2.1.** D'après (i) et le lemme 2.1.2, il existe un ensemble presque sûr  $\Omega_0 \subset \Omega$  tel que, pour tout  $\omega \in \Omega_0$ , les suites  $(P_n)$  et  $(Q_n^\omega)$  soient presque faiblement équivalentes.

Fixons  $\omega \in \Omega_0$  et  $A \in \mathfrak{G}$ . D'après (ii), (iii) et le lemme 2.2.2, on a, en raisonnant sur les sous-suites de  $(P_n)$  et  $(Q_n^\omega)$  qui convergent vers une même limite,

$$\text{p.s.} \quad \sup_{x' \in A} \left| \int_{\mathbb{E}} \langle x', x \rangle dP_n(x) - \int_{\mathbb{E}} \langle x', x \rangle dQ_n^\omega(x) \right| \rightarrow 0,$$

c'est-à-dire

$$\text{p.s.} \quad N_A \left( \frac{1}{n} \sum_{i=1}^n \mathbf{E} X_i - \frac{1}{n} \sum_{i=1}^n X_i(\omega) \right) \rightarrow 0.$$

□

On donne dans [CRdF00, Exemple 1] un exemple qui prouve que le théorème 2.2.1 peut s'appliquer à des vecteurs aléatoires Pettis intégrables non Bochner intégrables.

Donnons quelques exemples d'application du théorème 2.2.1.

Soit  $\mathbb{E}$  un espace de Banach séparable et soit  $\mathbb{E}'_c$  son dual muni de la topologie  $\tau_c$  de la convergence uniforme sur les parties compactes de  $\mathbb{E}$ . On montre facilement [CRdF00] que  $\mathbb{E}'_c$  est lusinien et qu'un vecteur aléatoire de  $\mathbb{E}'_c$  est Pettis intégrable si et seulement si il est scalairement intégrable. On note  $L_{\mathbb{E}'}^1[\mathbb{E}]$  l'ensemble des vecteurs aléatoires Pettis intégrables  $X$  de  $\mathbb{E}'_c$  définis sur  $(\Omega, \mathcal{F}, P)$  tels que  $\int_{\Omega} \|X\| dP < +\infty$  ( $X$  n'est pas nécessairement mesurable pour la topologie forte sur  $\mathbb{E}'$ , mais la mesurabilité de  $\|X\|$  provient de la mesurabilité scalaire de  $X$  et de la séparabilité de  $\mathbb{E}$ ). L'espace  $L_{\mathbb{E}'}^1[\mathbb{E}]$  est étudié dans [ITIT69].

Le corollaire ci-dessous contient la loi forte des grands nombres pour des éléments deux à deux indépendants équidistribués de  $L_{\mathbb{E}'}^1[\mathbb{E}]$ .

**Corollaire 2.2.3 (loi des grands nombres pour des éléments de  $L_{\mathbb{E}'}^1[\mathbb{E}]$ )**

Soit  $\mathbb{E}$  un espace de Banach séparable. Soit  $(X_n)_n$  une suite d'éléments deux à deux indépendants de  $L_{\mathbb{E}'}^1[\mathbb{E}]$ . On suppose que

- (i) la suite  $(P_n)_n = (1/n \sum_{i=1}^n L(X_i))_n$  est tendue dans  $P(\mathbb{E}'_c)$ ,
- (ii)' la fonction  $\|\cdot\|$  est uniformément intégrable par rapport à la suite  $(1/n \sum_{i=1}^n L(X_i))_n$ ,
- (iii)' la suite de variables aléatoires réelles  $(\|X_n\|)_n$  vérifie la loi forte des grands nombres :

$$\text{p.s.} \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n (\|X_i\| - \mathbf{E} \|X_i\|) = 0.$$

Alors  $(X_n)_n$  vérifie la loi forte des grands nombres dans  $\mathbb{E}'_c$ , c'est-à-dire qu'il existe  $\Omega_0 \in \mathcal{F}$  tel que  $P(\Omega_0) = 1$  et que, pour tout  $\omega \in \Omega_0$  et toute partie

compacte  $K$  de  $\mathbb{E}$ , on a

$$\lim_{n \rightarrow +\infty} \sup_{x \in K} \left| \frac{1}{n} \sum_{i=1}^n \langle x, X_i(\omega) - \mathbf{E} X_i \rangle \right| = 0.$$

Dans le cas d'un espace de Banach muni de sa topologie faible, on peut affaiblir légèrement la condition de tension des lois.

**Corollaire 2.2.4 (loi forte des grands nombres dans la topologie  $\sigma(\mathbb{F}, \mathbb{F}')$  d'un espace de Banach  $\mathbb{F}$ )** Soit  $\mathbb{F}$  un espace de Banach séparable et posons  $\mathbb{E} = (\mathbb{F}, \sigma(\mathbb{F}, \mathbb{F}'))$  (dans ce cas  $\mathbb{E}$  et  $\mathbb{F}$  ont les mêmes parties bornées, les mêmes boréliens et les mêmes vecteurs aléatoires Pettis intégrables). Soit  $(X_n)_n$  une suite de vecteurs aléatoires Pettis intégrables deux à deux indépendants de  $\mathbb{E}$ , définis sur l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ . On suppose que

- (i) la suite  $(P_n)_n = (1/n \sum_{i=1}^n \mathbf{L}(X_i))_n$  est tendue par rapport aux parties bornées, c'est-à-dire que, pour tout  $\epsilon > 0$ , il existe une partie bornée  $M$  de  $\mathbb{E}$  telle que, pour tout entier  $n \geq 1$ , on ait  $P_n(M) > 1 - \epsilon$ ,
- (ii) chaque élément de  $\mathbb{E}' = \mathbb{F}'$  est uniformément intégrable par rapport à la suite  $(P_n)_n$ ,
- (iii) pour presque tout  $\omega \in \Omega$ , chaque élément de  $\mathbb{E}' = \mathbb{F}'$  est uniformément intégrable par rapport à la suite  $(1/n \sum_{i=1}^n \delta_{X_i(\omega)})_n$ .

Alors  $(X_n)_n$  vérifie la loi forte des grands nombres dans  $\mathbb{E}$ , c'est-à-dire

$$p.s. \quad (\forall x' \in \mathbb{E}') \quad \lim_{n \rightarrow +\infty} \langle x', \frac{1}{n} \sum_{i=1}^n (X_i(\omega) - \mathbf{E} X_i) \rangle = 0.$$

La démonstration consiste à appliquer le théorème 2.2.1 dans l'espace sous-linien  $(\mathbb{F}'', \sigma(\mathbb{F}'', \mathbb{F}'))$ .

## 2.3 La loi des grands nombres comme résultat de compacité

**La topologie  $\sharp(\mathfrak{S})$  sur  $\mathbf{MU}_E^1$**  Comme on l'a vu dans la démonstration du théorème 2.1.1, la loi des grands nombres de Cuesta et Matrán repose sur l'inégalité fondamentale (2.1.3) qui montre que le barycentre est une application lipschitzienne relativement à la distance de Wasserstein. On va voir que cette propriété et la condition de tension sur les lois permettent d'interpréter la loi des grands nombres de Cuesta et Matrán comme un résultat de compacité séquentielle dans une topologie appropriée.

Cette approche s'étend également à la loi des grands nombres pour des vecteurs Pettis intégrables (théorème 2.2.1). Nous allons d'abord définir une distance de Wasserštejn généralisée.

Soit  $\mathbb{E}$  un espace vectoriel topologique localement convexe séparé. On note  $\text{MU}_E^1$  l'ensemble des probabilités  $P$  de Radon sur  $\mathbb{E}$  qui sont  $\mathfrak{S}$ -uniformément scalairement intégrable, c'est-à-dire telles que chaque  $A \in \mathfrak{S}$  est uniformément intégrable par rapport à  $P$ . On note  $\text{MP}_{\mathbb{E}}^1$  l'ensemble des probabilités de Radon sur  $\mathbb{E}$  qui sont Pettis intégrables, c'est-à-dire les probabilités  $P$  sur  $\mathbb{E}$  qui sont  $\mathfrak{S}$ -uniformément scalairement intégrable et telles que, pour tout  $B \in \mathfrak{B}(\mathbb{E})$ , il existe un élément de  $\mathbb{E}$ , noté  $\int_B x dP(x)$ , tel que, pour tout  $x' \in \mathbb{E}'$ , on ait

$$\langle x', \int_B x dP(x) \rangle = \int_B \langle x', x \rangle dP(x)$$

(si  $X$  est un vecteur aléatoire Pettis intégrable, on a donc  $L(X) \in \text{MP}_{\mathbb{E}}^1$ , la réciproque est vraie lorsque  $\mathbb{E}$  est quasi-complet, voir [CRdF00, page 101]). Rappelons que, pour  $P, Q \in \text{P}(\mathbb{E})$ , on note

$$D(P, Q) = \{ \pi \in \text{P}(\mathbb{M} \times \mathbb{M}); \pi(\cdot \times \mathbb{M}) = P \text{ et } \pi(\mathbb{M} \times \cdot) = Q \}.$$

Soit  $\mathfrak{S}$  comme dans la section précédente. On pose, pour  $P, Q \in \text{MU}_E^1$  et  $A \in \mathfrak{S}$ ,

$$d_A(P, Q) = \inf_{\pi \in D(P, Q)} \sup_{x' \in A} \int_{\mathbb{E} \times \mathbb{E}} |\langle x', x - y \rangle| d\pi(x, y).$$

On montre facilement que l'infimum dans la définition de  $d_A(P, Q)$  est toujours atteint et que  $d_A$  est une semidistance (de Wasserštejn généralisée) sur  $\text{MU}_E^1$ . De plus, sur le sous-espace  $\text{MP}_{\mathbb{E}}^1$ , on a l'inégalité fondamentale

$$(2.3.1) \quad (\forall P, Q \in \text{MP}_{\mathbb{E}}^1) \quad N_A \left( \int_{\mathbb{E}} x dP(x), \int_{\mathbb{E}} x dQ(x) \right) \leq d_A(P, Q).$$

La différence principale entre la topologie définie par les semidistances  $d_A$  (appelons la provisoirement  $\tau_{\text{WP}}$ ) et celle définie par la distance usuelle de Wasserštejn (disons, dans le cas où  $\mathbb{E}$  est un espace de Banach) est que cette dernière est plus fine que la trace sur  $\text{P}^1(\mathbb{E})$  de la topologie faible (voir page 46 la deuxième étape de la démonstration du théorème 2.1.1), tandis que l'on peut trouver des suites dans  $\text{MU}_E^1$  convergeant pour la topologie faible mais pas pour  $\tau_{\text{WP}}$ , ou bien pour  $\tau_{\text{WP}}$  mais pas pour la topologie faible [CRdF00, Example 2 et Example 3]. Notons  $\flat(\mathfrak{S})$  la topologie faible sur  $\text{MU}_E^1$ . Pour retrouver le schéma de la preuve du théorème 2.1.1, on est ainsi amené à considérer sur  $\text{MU}_E^1$  la topologie  $\sharp(\mathfrak{S})$ , supremum des topologies  $\tau_{\text{WP}}$  et  $\flat(\mathfrak{S})$ , qui permet la caractérisation ci-dessous.

**Théorème 2.3.1** *Une partie  $\mathfrak{M}$  de  $MU_E^1$  est séquentiellement relativement compacte pour  $\sharp(\mathfrak{S})$  si et seulement si  $\mathfrak{M}$  est  $\mathfrak{S}$ -uniformément scalairement intégrable et séquentiellement relativement compacte pour  $\flat(\mathfrak{S})$ .*

À l'aide de ce théorème, il est facile de réécrire la preuve du théorème 2.2.1 comme une légère variante de celle du théorème 2.1.1.

**Un principe général** Soit  $\mathbb{T}$  un espace uniforme vérifiant les hypothèses du lemme 2.1.2, c'est-à-dire qu'il existe un ensemble dénombrable  $\mathcal{H}$  de fonctions continues bornées de  $\mathbb{T}$  dans  $\mathbb{R}$  qui sépare les points de  $\mathbb{T}$ . Soit  $(d_\alpha)_{\alpha \in \mathfrak{U}}$  une famille de semidistances qui définit la structure uniforme de  $\mathbb{T}$ . On définit la *structure uniforme de Hausdorff* sur  $\mathfrak{F}(\mathbb{T})$  à l'aide des écarts

$$[A, B]_\alpha = \max \left\{ \sup_{x \in A} d_\alpha(x, B), \sup_{y \in B} d_\alpha(y, A) \right\}$$

(cette structure uniforme est indépendante du choix des écarts  $d_\alpha$  pour définir la structure uniforme de  $\mathbb{T}$ , voir [CV77, Chapter II]).

On se donne une partie convexe  $\mathfrak{M}$  de  $P(\mathbb{T})$  contenant les masses de Dirac  $\delta_x$  ( $x \in \mathbb{T}$ ), une application

$$\text{bar} : \begin{cases} \mathfrak{M} & \rightarrow \mathfrak{F}(\mathbb{T}) \\ P & \mapsto \text{bar } P \end{cases}$$

(que l'on appellera *barycentre*) et une topologie  $\tau_{\mathfrak{M}}$  sur  $\mathfrak{M}$  telles que

(H) si  $(P_n)_{n \in \mathbb{N}}$  et  $(Q_n)_{n \in \mathbb{N}}$  sont deux suites faiblement équivalentes dans  $\mathfrak{M}$  telles que  $\{P_n; n \in \mathbb{N}\}$  et  $\{Q_n; n \in \mathbb{N}\}$  sont séquentiellement relativement compact pour  $\tau_{\mathfrak{M}}$ , alors les suites  $(\text{bar } P_n)$  et  $(\text{bar } Q_n)$  sont équivalentes pour la structure uniforme de Hausdorff, c'est-à-dire

$$(\forall \alpha \in \mathfrak{U}) \lim_{n \rightarrow +\infty} [\text{bar } P_n, \text{bar } Q_n]_\alpha = 0.$$

L'hypothèse (H) est vérifiée en particulier si on a

(H')  $\tau_{\mathfrak{M}}$  provient d'une structure uniforme  $\mathfrak{u}_{\mathfrak{M}}$  et si l'application  $\text{bar}$  est uniformément continue pour  $\mathfrak{u}_{\mathfrak{M}}$  et la structure uniforme de Hausdorff sur  $\mathfrak{F}(\mathbb{T})$ .

Dans les lois exemples précédents (théorèmes 2.1.1 et 2.2.1), la propriété (H') est assurée par les inégalités (2.1.3) et (2.3.1).

**Théorème 2.3.2** *Soient  $\mathbb{T}$ ,  $\text{bar} : \mathbb{T} \rightarrow \mathfrak{F}(\mathbb{T})$  et  $\tau_{\mathfrak{M}}$  comme ci-dessus, vérifiant l'hypothèse (H). Soit  $(X_n)_{n \in \mathbb{N}^*}$  une suite d'éléments aléatoires deux à deux indépendants de  $\mathbb{T}$ , tels que  $L(X_n) \in \mathfrak{M}$  pour tout  $n \in \mathbb{N}^*$ . On note  $P_n = 1/n \sum_{i=1}^n L(X_i)$  et  $Q_n^\omega = 1/n \sum_{i=1}^n \delta_{X_i}$  ( $n \in \mathbb{N}^*$ ,  $\omega \in \Omega$ ). On suppose que*



- (i) la suite  $(P_n)_n$  est tendue,
- (ii) l'ensemble  $\{P_n; n \in \mathbb{N}^*\}$  est une partie séquentiellement relativement compacte de  $(\mathfrak{M}, \tau_{\mathfrak{M}})$ ,
- (iii) pour presque tout  $\omega \in \Omega$ , l'ensemble  $\{Q_n^\omega; n \in \mathbb{N}^*\}$  est une partie séquentiellement relativement compacte de  $(\mathfrak{M}, \tau_{\mathfrak{M}})$ .

Alors  $(X_n)$  vérifie la loi forte des grands nombres, c'est-à-dire

$$p.s. \quad (\forall \alpha \in \mathfrak{A}) \quad \lim_{n \rightarrow +\infty} [\text{bar } Q_n^\omega, \text{bar } P_n]_\alpha = 0.$$

**Preuve.** D'après le lemme 2.1.2, il existe  $\Omega_0 \in \mathcal{F}$  tel que  $P(\Omega_0) = 1$  et que, pour tout  $\omega \in \Omega_0$ , la suite  $(Q_n^\omega)_n$  soit tendue et faiblement équivalente à  $(P_n)_n$ . D'après (ii) et (iii), on peut supposer de plus que, pour tout  $\omega \in \Omega_0$ , les ensembles  $\{P_n; n \in \mathbb{N}^*\}$  et  $\{Q_n^\omega; n \in \mathbb{N}^*\}$  sont séquentiellement relativement compacts dans  $(\mathfrak{M}, \tau_{\mathfrak{M}})$ . La conclusion découle alors de (H).  $\square$

On remarquera que la partie probabiliste de la preuve du théorème 2.3.2 réside dans le lemme 2.1.2 (Glivenko-Cantelli), qui fait appel à une loi forte des grands nombres pour des variables aléatoires réelles. Le reste de la démonstration utilise la compacité séquentielle relative de  $(P_n)$  et  $(Q_n^\omega)$  pour  $\tau_{\mathfrak{M}}$  ainsi que la continuité de l'application  $\text{bar}$  pour  $\tau_{\mathfrak{M}}$ . La compacité séquentielle pour  $\tau_{\mathfrak{M}}$  dans les théorèmes 2.1.1 et 2.2.1 provient de la compacité séquentielle pour la topologie faible et d'une condition d'uniforme intégrabilité.

# Références

- [AGK76] Adamski (Wolfgang), Gänszler (Peter) et Kaiser (Sigurd). – On compactness and convergence in spaces of measures. *Math. Ann.*, vol. 220, n° 3, 1976, pp. 193–210.
- [AKP<sup>+</sup>92] Akhmerov (R. R.), Kamenskii (M. I.), Potapov (A. S.), Rodkina (B. N.) et Sadovskii (B. N.). – *Measures of Noncompactness and Condensing Operators*. – Basel, Boston, Berlin, Birkhäuser, 1992, *Operator Theory : Advances and Applications*. revised and translated edition from the Russian 1986 original.
- [Ald78] Aldous (David J.). – Weak convergence of randomly indexed sequences of random variables. *Math. Proc. Camb. Phil. Soc.*, vol. 83, 1978, pp. 117–126.
- [AP03] Artstein (Zvi) et Popa (Cristian Constantin). – Convexity and the natural best approximation in spaces of integrable Young measures. *J. Convex Anal.*, vol. 10, n° 1, 2003, pp. 169–184.
- [AP04] Artstein (Zvi) et Popa (Cristian Constantin). – Orlicz-Young structures of Young measures. *Acta Appl. Math.*, vol. 80, n° 1, 2004, pp. 1–33.
- [Bal84a] Balder (Erik J.). – A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control and Optimization*, vol. 22, 1984, pp. 570–598.
- [Bal84b] Balder (Erik J.). – A general denseness result for relaxed control theory. *Bull. Austral. Math. Soc.*, vol. 30, 1984, pp. 463–475.
- [Bal88] Balder (Erik J.). – Generalized equilibrium results for games with incomplete information. *Math. Oper. Res.*, vol. 13, n° 2, 1988, pp. 265–276.
- [Bal89a] Balder (Erik). – On Prohorov’s theorem for transition probabilities. *Sém. Anal. Convexe*, vol. 19, 1989, pp. 9.1–9.11.
- [Bal89b] Ball (J. M.). – A version of the fundamental theorem for Young measures. *In : PDEs and continuum models of phase transitions*

- (*Nice 1988*), éd. par Serre (D.). pp. 207–215. – Berlin, New York, 1989.
- [Bal90] Balder (E. J.). – New sequential compactness results for spaces of scalarly integrable functions. *J. Math. Anal. Appl.*, vol. 151, 1990, pp. 1–16.
- [Bal01] Balder (Erik J.). – On ws-convergence of product measures. *Math. Oper. Res.*, vol. 26, n° 3, 2001, pp. 494–518.
- [Bar98] Barbu (D.). – Local and global existence for mild solutions of stochastic differential equations. *Portugal. Math.*, vol. 55, n° 4, 1998, pp. 411–424.
- [BB02] Barbu (Dorel) et Bocşan (Gheorghe). – Approximations to mild solutions of stochastic semilinear equations with non-Lipschitz coefficients. *Czechoslovak Math. J.*, vol. 52(127), n° 1, 2002, pp. 87–95.
- [BD83] Blackwell (D.) et Dubins (L. E.). – An extension of Skorohod’s almost sure representation theorem. *Proc. Amer. Math. Soc.*, vol. 89, 1983, pp. 691–692.
- [Bee94] Beer (G.). – Wijsman convergence : a survey. *Set-Valued Anal.*, vol. 2, 1994, pp. 77–94.
- [Ben62] Beneš (V. E.). – Martingales on metric spaces. *Teor. Veroyatnost. i Primenen.*, vol. 7, 1962, pp. 82–83.
- [BG70] Beck (A.) et Giesy (D. P.). –  $p$ -uniform convergence and a vector-valued strong law of large numbers. *Trans. Amer. Math. Soc.*, vol. 147, 1970, pp. 541–559.
- [BHL93] Bru (Bernard), Heinich (Henri) et Lootgieter (Jean Claude). – Distances de Lévy et extensions des théorèmes de la limite centrale et de Glivenko-Cantelli. *Pub. Inst. Stat. Univ. Paris*, vol. XXXVII, n° 3–4, 1993, pp. 29–42.
- [Bil68] Billingsley (Patrick). – *Convergence of Probability Measures*. – New York, London, J. Wiley, 1968.
- [Bog98] Bogachev (Vladimir I.). – Measures on topological spaces. *J. Math. Sci.*, vol. 91, n° 4, 1998, pp. 3033–3156. – Functional analysis, 1.
- [Bou02] Bouziad (Ahmed). – Coincidence of the upper Kuratowski topology with the co-compact topology on compact sets, and the Prohorov property. *Topology Appl.*, vol. 120, 2002, pp. 283–299.
- [Bus48] Busemann (Herbert). – Spaces with non-positive curvature. *Acta Math.*, vol. 80, 1948, pp. 259–310.

- [Bus55] Busemann (Herbert). – *Theory of Geodesics*. – New York, Academic Press, 1955.
- [Cal82] Calbrix (Jean). – Une propriété des espaces topologiques réguliers, images continues d’espaces métrisables séparables. *C. R. Acad. Sci. Paris Sér. I*, vol. 295, 1982, pp. 81–82.
- [Cal84] Calbrix (Jean). – Plongement dans les espaces lusiniens (d’après Amer Bešlagić) et sur un théorème d’Hurewicz. *Pub. Math. Univ. Pierre et Marie Curie, Sémin. Initiation Anal. 23ème année - 1983/84*, vol. 66, 1984, pp. 11.1–11.10.
- [Chr74] Christensen (J. P. R.). – *Topology and Borel structure*. – Amsterdam, London, North-Holland, 1974.
- [Cla36] Clarkson (J. A.). – Uniformly convex spaces. *Trans. Amer. Math. Soc.*, vol. 40, 1936, pp. 396–414.
- [CM88] Cuesta (Juan A.) et Matrán (Carlos). – Strong convergence of weighted sums of random elements through the equivalence of sequences of distributions. *J. Multivariate Anal.*, vol. 25, 1988, pp. 311–322.
- [CM92] Cuesta (Juan A.) et Matrán (Carlos). – A review on strong convergence of weighted sums of random elements based on Wasserstein metrics. *J. of Statistical Planning and Inference*, vol. 30, 1992, pp. 359–370.
- [Cor73] Cornet (B.). – *Topologies sur les fermés d’un espace métrique*. – Cahiers de Mathématiques de la Décision n° 7309, Université Paris-Dauphine, 1973. Thèse de 3ème cycle, Paris.
- [CRdF00] Castaing (Charles) et Raynaud de Fitte (Paul). –  $\mathfrak{S}$ -Uniform scalar integrability and strong laws of large numbers for Pettis integrable functions with values in a separable locally convex space. *J. Theor. Probab.*, vol. 13, n° 1, 2000, pp. 93–134.
- [CRdF04] Castaing (Charles) et Raynaud de Fitte (Paul). – On the fiber product of Young measures with application to a control problem with measures. In : *Kusuoka, Shigeo (ed.) et al., Advances in mathematical economics. Vol. 6. Tokyo : Springer. Adv. Math. Econ. 6, 1-38*. – 2004.
- [CRdFV04] Castaing (Charles), Raynaud de Fitte (Paul) et Valadier (Michel). – *Young measures on Topological Spaces. With Applications in Control Theory and Probability Theory*. – Dordrecht, Kluwer Academic Publishers, 2004.

- [CTT83] Csörgő (S.), Tandori (K.) et Totik (V.). – On the strong law of large numbers for pairwise independent random variables. *Acta Math. Hung.*, vol. 42, n° 3–4, 1983, pp. 319–330.
- [CV77] Castaing (Charles) et Valadier (Michel). – *Convex Analysis and Measurable Multifunctions*. – Berlin, Springer Verlag, 1977, *Lecture Notes in Math*.
- [DM02] Dedecker (Jérôme) et Merlevède (Florence). – Necessary and sufficient conditions for the conditional central limit theorem. *Ann. Probab.*, vol. 30, n° 3, 2002, pp. 1044–1081.
- [Dos49] Doss (Shafik). – Sur la moyenne d’un élément aléatoire dans un espace distancié. *Bull. Sc. Math. 2è série*, vol. 73, 1949, pp. 1–26.
- [Dos62] Doss (Shafik). – Moyennes conditionnelles et martingales dans un espace métrique. *C. R. Acad. Sci. Paris Sér. I*, vol. 254, 1962, pp. 3630–3632.
- [DPF94] Da Prato (G.) et Frankowska (H.). – A stochastic Filippov Theorem. *Stochastic Anal. Appl.*, vol. 12, n° 4, 1994, pp. 409–426.
- [DPRdF] Dedecker (Jérôme), Prieur (Clémentine) et Raynaud de Fitte (Paul). – Parametrized Kantorovich-Rubinstein Theorem and application to the coupling of random variables. – soumis pour publication et preprint mis en ligne sur le site arXiv <http://arxiv.org/pdf/math.PR/0410052>.
- [DPZ92a] Da Prato (Giuseppe) et Zabczyk (Jerzy). – A note on stochastic convolution. *Stochastic Anal. Appl.*, vol. 10, n° 2, 1992, pp. 143–153.
- [DPZ92b] Da Prato (Giuseppe) et Zabczyk (Jerzy). – *Stochastic Equations in Infinite Dimensions*. – Cambridge, Cambridge University Press, 1992, *Encyclopedia of Mathematics and Its Applications*.
- [DR50] Dvoretzky (A.) et Rogers (C. A.). – Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U. S. A.*, vol. 36, 1950, pp. 192–197.
- [Dud66] Dudley (R. M.). – Convergence of Baire measures. *Studia Math.*, vol. 27, 1966, pp. 251–268.
- [EM91] Emery (Michel) et Mokobodzki (G.). – Sur le barycentre d’une probabilité dans une variété. *In : Séminaire de Probabilités XXV*. pp. 220–233. – Berlin, 1991.

- [ESH99] Es-Sahib (Aziz) et Heinich (Henri). – Barycentre canonique pour un espace métrique à courbure négative. *In : Séminaire de Probabilités, XXXIII*, pp. 355–370. – Berlin, Springer, 1999.
- [Ete83] Etemadi (Nasrollah). – On the laws of large numbers for non-negative random variables. *J. Multivariate Anal.*, vol. 13, 1983, pp. 187–193.
- [Fer88] Fernique (Xavier). – Un modèle presque sûr pour la convergence en loi. *C. R. Acad. Sci. Paris Sér. I*, vol. 306, 1988, pp. 335–338.
- [Fer94] Fernique (Xavier). – Une caractérisation des espaces de Fréchet nucléaires. *In : Probability in Banach spaces, 9 : Proceedings from the 9th International Conference on Banach Spaces, Held at Sandbjerg, Denmark, August 16–21, 1993*, éd. par Hoffmann-Jørgensen (J.). pp. 173–181. – Boston, 1994.
- [Fis70] Fischler (Roger). – Suites de bi-probabilités stables. *Annales de la Faculté des Sciences de l'Université de Clermont*, vol. 43, 1970, pp. 159–167.
- [Fis76] Fischler (Roger). – Convergence faible avec indices aléatoires. *Ann. Inst. Henri Poincaré – Sect. B*, vol. XII, n° 4, 1976, pp. 391–399.
- [Fré48] Fréchet (Maurice). – Les éléments aléatoires de nature quelconque. *Ann. Inst. H. Poincaré*, vol. 14, 1948, pp. 215–310.
- [Gän71] Gänsler (Peter). – Compactness and sequential compactness in spaces of measures. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, vol. 17, 1971, pp. 124–146.
- [GP84] Gardner (R. J.) et Pfeffer (W. F.). – Borel measures. *In : Handbook of set-theoretic topology*, pp. 961–1043. – Amsterdam, North-Holland, 1984.
- [Gro75] Gromov (M.). – *Structures Métriques pour les Variétés Riemanniennes*. – Paris, Cedic, 1975.
- [Gro99] Gromov (Misha). – *Metric Structures for Riemannian and Non-Riemannian Spaces*. – Boston, Mass., Birkhäuser, 1999, *Progress in Mathematics*.
- [Har73] Hartman (Philip). – *Ordinary differential equations*. – S. M. Hartman, Baltimore, Md., 1973, xiv+612p. Corrected reprint.
- [Her] Herer (Wojciech). – Martingales à valeurs fermées bornées d'un espace métrique à courbure négative. – non publié.
- [Her83] Herer (Wojciech). – Espérance mathématique au sens de Doss d'une variable aléatoire à valeurs dans un espace métrique. *C. R. Acad. Sci. Paris Sér. I*, vol. 302, 1983, pp. 131–134.

- [Her88] Herer (Wojciech). – Espérance mathématique d’une variable aléatoire à valeurs dans un espace métrique à courbure négative. *C. R. Acad. Sci. Paris Sér. I*, vol. 306, 1988, pp. 681–684.
- [Her91] Herer (Wojciech). – Mathematical expectation and martingales of random subsets of a metric space. *Probability and Mathematical Statistics*, vol. 11, n° 2, 1991, pp. 291–304.
- [Her92] Herer (Wojciech). – Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature. *Probability and Mathematical Statistics*, vol. 13, n° 1, 1992, pp. 59–70.
- [HJ71] Hoffmann-Jørgensen (Jørgen). – Existence of conditional probabilities. *Math. Scand.*, vol. 28, 1971, pp. 257–264.
- [HJ91] Hoffmann-Jørgensen (Jørgen). – *Stochastic Processes on Polish spaces*. – Aarhus, Denmark, Matematisk Institut, Aarhus Universitet, 1991, *Various Publication Series*.
- [HJ98] Hoffmann-Jørgensen (Jørgen). – Convergence in law of random elements and random sets. *In : High dimensional probability (Oberwolfach, 1996)*, pp. 151–189. – Basel, Birkhäuser, 1998.
- [ITIT69] Ionescu Tulcea (A.) et Ionescu Tulcea (C.). – *Topics in the Theory of Liftings*. – Berlin, Springer Verlag, 1969.
- [Jac85] Jacod (Jean). – Théorèmes limite pour les processus. *In : École d’Été de Probabilités de Saint-Flour XIII-1983*, éd. par Hennequin (P. L.). pp. 298–409. – Berlin, 1985.
- [Jac97] Jacod (Jean). – On continuous conditional Gaussian martingales and stable convergence in law. *In : Séminaire de Probabilités XXXI*, éd. par Azéma (Jacques), Emery (Michel) et Yor (Marc). pp. 232–246. – Berlin, 1997.
- [Jak86] Jakubowski (Adam). – On the Skorokhod topology. *Ann. Inst. H. Poincaré Probab. Statist.*, vol. 22, n° 3, 1986, pp. 263–285.
- [Jaw84] Jawhar (Abdelali). – Mesures de transition et applications. *Sém. Anal. Convexe*, vol. 14, 1984, pp. 13.1–13.62.
- [JKRdF05] Jakubowski (Adam), Kamenskiï (Mikhail I.) et Raynaud de Fitte (Paul). – Existence of weak solutions to stochastic differential inclusions. *Stochastic Anal. Appl.*, 2005. – à paraître, 25 pages, preprint mis en ligne sur le site arXiv <http://arxiv.org/pdf/math.PR/0407512>.
- [JM81a] Jacod (Jean) et Mémin (Jean). – Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en

- probabilité. In : *Sémin. de Probabilités XV, Univ. Strasbourg 1979/80*, éd. par Azéma (Jacques) et Yor (Marc). pp. 529–546. – Berlin, 1981.
- [JM81b] Jacod (Jean) et Mémin (Jean). – Weak and strong solutions of stochastic differential equations : Existence and stability. In : *Stochastic Integrals. Proceedings of the LMS Durham Symposium, July 7 – 17, 1980*, éd. par Williams (D.). pp. 169–212. – Berlin, 1981.
- [Kel55] Kelley (John L.). – *General Topology*. – Berlin, Springer Verlag, 1955.
- [KOZ01] Kamenskiĭ (M. I.), Obukhovskĭ (Valeri) et Zecca (Pietro). – *Condensing Multivalued Maps and Semilinear Differential Inclusions in Banach Spaces*. – Berlin, de Gruyter, 2001, *de Gruyter Series in Nonlinear Analysis and Applications*.
- [KP94] Kinderlehrer (David) et Pedregal (Pablo). – Gradient Young measures generated by sequences in Sobolev spaces. *J. Geom. Anal.*, vol. 4, n° 1, 1994, pp. 59–90.
- [Kuc98] Kucia (Anna). – Some results on Carathéodory selections and extensions. *J. Math. Anal. Appl.*, vol. 223, n° 1, 1998, pp. 302–318.
- [Let98] Letta (Giorgio). – Convergence stable et applications. *Atti Sem. Mat Fis. Modena (supplemento)*, vol. 46, 1998, pp. 191–211. – Dedicated to Prof. C. Vinti (Perugia 1996).
- [Lev84] Levin (V. L.). – The problem of mass transfer in a topological space and probability measures with given marginal measures on the product of two spaces. *Dokl. Akad. Nauk SSSR*, vol. 276, n° 5, 1984, pp. 1059–1064. – English translation : *Soviet Math. Dokl.* 29 (1984), no. 3, 638–643.
- [Lin73] Lindvall (Torgny). – Weak convergence of probability measures and random functions in the function space  $D[0, \infty)$ . *J. Appl. Probability*, vol. 10, 1973, pp. 109–121.
- [Man90] Manthey (Ralf). – Convergence of successive approximation for parabolic partial differential equations with additive white noise. *Serdica*, vol. 16, n° 3-4, 1990, pp. 194–200.
- [Mic51] Michael (Ernest). – Topologies on spaces of subsets. *Trans. Amer. Math. Soc.*, vol. 71, 1951, pp. 152–182.
- [Mic66] Michael (E.). –  $\aleph_0$ -spaces. *J. Math. Mech.*, vol. 15, 1966, pp. 983–1002.



- [Mit83] Mitoma (Itaru). – Tightness of probabilities on  $C([0, 1]; \mathcal{S}')$  and  $D([0, 1]; \mathcal{S}')$ . *Ann. Probab.*, vol. 11, n° 4, 1983, pp. 989–999.
- [Mus87] Musiał (Kazimierz). – Vitali and Lebesgue theorems for Pettis integral in locally convex spaces. *Atti Sem. Mat Fis. Modena*, vol. 25, 1987, pp. 159–166.
- [Mus91] Musiał (Kazimierz). – Topics in the theory of Pettis integration. *Rendiconti dell'istituto di matematica dell'Università di Trieste*, vol. 23, 1991, pp. 176–262. – School on Measure Theory and Real Analysis Grado (Italy).
- [OW98] O'Brien (George L.) et Watson (Stephen). – Relative compactness for capacities, measures, upper semicontinuous functions and closed sets. *J. Theor. Probab.*, vol. 11, 1998, pp. 577–588.
- [Pel81a] Pellaumail (J.). – Weak solutions for semimartingales. *Canad. J. Math.*, vol. 33, n° 5, 1981, pp. 1165–1181.
- [Pel81b] Pellaumail (Jean). – Solutions faibles et semi-martingales. *In : Séminaire de Probabilités XV, 1979/80. Université de Strasbourg*, éd. par Azéma (Jacques) et Yor (Marc), pp. 561–586. – Berlin, Springer Verlag, 1981.
- [Pic94] Picard (Jean). – Barycentres et martingales sur une variété. *Ann. Inst. H. Poincaré*, vol. 30, n° 4, 1994, pp. 647–702.
- [PV95] Piccinini (Laurent) et Valadier (Michel). – Uniform integrability and Young measures. *J. Math. Anal. Appl.*, vol. 195, n° 2, 1995, pp. 428–439.
- [Rac82] Rachev (Svetlozar T.). – Minimal metrics in the random variables space. *Publ. Inst. Stat. Univ. Paris*, vol. 27, 1982, pp. 27–47.
- [Rac91] Rachev (Svetlozar T.). – *Probability Metrics and the Stability of Stochastic Models*. – Chichester, New York, Wiley, 1991.
- [RdF97] Raynaud de Fitte (Paul). – Théorème ergodique ponctuel et lois fortes des grands nombres pour des points aléatoires d'un espace métrique à courbure négative. *The Annals of Probability*, vol. 25, n° 2, 1997, pp. 738–766.
- [RdF03] Raynaud de Fitte (Paul). – Compactness criteria for the stable topology. *Bull. Pol. Acad. Sci., Math.*, vol. 51, n° 4, 2003, pp. 343–363.
- [RdFZ02] Raynaud de Fitte (Paul) et Zięba (Wiesław). – On the construction of a stable sequence with given density. *Ann. Univ. Mariae Curie-Skłodowska Sect. A*, vol. 56, n° 8, 2002, pp. 77–84.

- [Rén63] Rényi (Alfred). – On stable sequences of events. *Sankhyā Ser. A*, vol. 25, 1963, pp. 293–302.
- [Rod84] Rodkina (A. E.). – On existence and uniqueness of solution of stochastic differential equations with heredity. *Stochastics*, vol. 12, n° 3-4, 1984, pp. 187–200.
- [Rou97] Roubíček (Tomáš). – *Relaxation in Optimization Theory and Variational Calculus*. – Berlin, Walter de Gruyter, 1997, *de Gruyter Series in Nonlinear Analysis and Applications*.
- [RR58] Rényi (A.) et Révész (P.). – On mixing sequences of random variables. *Acta Math. Acad. Sci. Hungar.*, vol. 9, 1958, pp. 389–393.
- [RR98] Rachev (S. T.) et Rüschendorf (L.). – *Mass Transportation Problems. Volume I : Theory*. – New York, Berlin, Springer Verlag, 1998, *Probability and its Applications*.
- [Sch73] Schwartz (Laurent). – *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. – London, Oxford University Press, 1973, *Tata Institute of Fundamental Research Studies in Mathematics*.
- [Sch75] Schäl (Manfred). – On dynamic programming : compactness of the space of policies. *Stochastic Processes Appl.*, vol. 3, n° 4, 1975, pp. 345–364.
- [Stu02] Sturm (Karl-Theodor). – Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature. *Ann. Probab.*, vol. 30, n° 3, 2002, pp. 1195–1222.
- [Tan92] Taniguchi (Takeshi). – Successive approximations to solutions of stochastic differential equations. *J. Differential Equations*, vol. 96, n° 1, 1992, pp. 152–169.
- [Tat02] Tateishi (Hiroshi). – On the existence of equilibria of equicontinuous games with incomplete information. *Adv. Math. Econ.*, vol. 4, 2002, pp. 41–59.
- [Top70a] Topsøe (Flemming). – Compactness in spaces of measures. *Studia Mathematica*, vol. 36, 1970, pp. 195–212.
- [Top70b] Topsøe (Flemming). – *Topology and Measure*. – Berlin, Springer Verlag, 1970, *Lecture Notes in Math*.
- [Top74] Topsøe (Flemming). – Compactness and tightness in a space of measures with the topology of weak convergence. *Math. Scand.*, vol. 34, 1974, pp. 187–210.

- [Val72] Valadier (Michel). – Comparaison de trois théorèmes de désintégration. *In : Travaux du Séminaire d'Analyse Convexe, Vol. II, Exp. No. 10*, pp. 10.1–10.21, Secrétariat des Math., Publ. No. 122. – U.E.R. de Math., Univ. Sci. Tech. Languedoc, Montpellier, 1972.
- [Val73] Valadier (Michel). – Désintégration d'une mesure sur un produit. *C. R. Acad. Sci. Paris Sér. I*, vol. 276, 1973, pp. A33–A35.
- [Val90] Valadier (Michel). – Young measures. *In : Methods of Nonconvex Analysis*, éd. par Cellina (A.). pp. 152–158. – Berlin, 1990.
- [Val94] Valadier (Michel). – A course on Young measures. *Rendiconti dell'istituto di matematica dell'Università di Trieste*, vol. 26, suppl., 1994, pp. 349–394. – Workshop di Teoria della Misura et Analisi Reale Grado, 1993 (Italia).
- [Val99] Valadier (Michel). – Analysis of the asymptotic distance between oscillating functions and their weak limit in  $L^2$ . *In : Advances in mathematical economics, Vol. 1 (Tokyo, 1999)*, pp. 99–113. – Tokyo, Springer, 1999.
- [Whe83] Wheeler (Robert F.). – A survey of Baire measures and strict topologies. *Exposition. Math.*, vol. 1, n° 2, 1983, pp. 97–190.
- [Yam81] Yamada (Toshio). – On the successive approximation of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, vol. 21, n° 3, 1981, pp. 501–515.
- [You37] Young (L. C.). – Generalized curves and the existence of an attained absolute minimum in the Calculus of Variations. *C. R. Soc. Sc. Varsovie*, vol. 30, 1937, pp. 212–234.

# Index des définitions

- élément aléatoire, 3
- barycentre, 41, 44, 57
- combinaison convexe, 40
- compact
  - pour les suites généralisées, 10
  - relativement, 11
  - séquentiellement relativement, 11, 18
- complétion universelle, 12
- convergence
  - au sens de Wijsman, 50
  - en probabilité, 19
  - stable, voir topologie stable
- convexe (espace), voir espace métrique convexe
- courbure négative, voir espace à courbure négative
- distance ou semidistance
  - de Hausdorff, 40
  - de Wasserstein, 23, 44
  - de Wasserstein généralisée, 56
  - de Wasserstein paramétrée, 24
  - géodésique, 43
- distance ou semidistance de Dudley, 21
  - paramétrée, 21
- domine
  - (pour des pavages), 12
- $\epsilon$ -réseau, 32
- ensemble aléatoire, 5
- équicontinu
  - (ensemble de mesures), 13
- équivalentes
  - $\Delta_W$ -équivalentes (pour des suites de probabilités), 46
  - faiblement équivalentes (pour des suites de probabilités); 46
  - topologiquement, 46
  - uniformément, 57
- espace
  - à courbure négative, 43
  - cosmique, 5
  - de Prokhorov, 15
  - de Radon, 3
  - de Skorokhod, 35
  - lusinien, 5
  - métrique convexe, 40
  - sous-métrisable, 12
  - souslinien, 5
- intégrable
  - (pour une probabilité), voir probabilité intégrable
  - au sens de Pettis, 52
  - au sens de Pettis (pour une mesure), 56
- intégrande, 4
  - continu ou de Carathéodory, 4
  - semicontinu supérieurement, inférieurement, 4
- mesure, 2
  - de Radon, 3
  - $\tau$ -régulière, 15
- mesure de non compacité, 32

mesure de Young, 3  
   d'ordre  $p$ , 8  
   dégénérée, 4  
   désintégrable, 3  
    $p$ -intégrable, 8

pavage, 11

principe d'invariance, voir théorème  
   de la limite centrale fonctionnel

probabilité intégrable, 44

produit fibré (de mesures de Young),  
   28

règle, 29

semidistance, voir distance ou semidistance

solution d'évolution  
   faible, ou solution-mesure, 31  
   forte, 31

structure uniforme de Hausdorff, 57

suite cofinale de compacts, 16

tendu  
   souplement, 14  
   strictement, 14

théorème  
   de Kantorovich–Rubinštejn, 26  
   de Kantorovich–Rubinštejn avec  
   paramètre, 27  
   de Komlós (pour les mesures de  
   Young), 11, 33  
   de la limite centrale, 34–35  
   de la limite centrale fonctionnel,  
   35–36  
   de semicontinuité, 7, 17  
   du produit fibré, 28  
   “portmanteau”, 6

TLC, voir théorème de la limite centrale

topologie  
   étroite, 3  
   de la convergence en moyenne  
   d'ordre  $p$  (sur  $\mathcal{Y}_D^p$ ), 24  
   de la convergence en probabilité  
   (sur  $\mathcal{Y}$ ), 21  
   faible (sur  $M(\mathbb{T})$ ), 3  
    $J_1$  de Skorokhod, 35  
   s-topologie, 10  
   stable, 4  
   w-topologie, 10

uniformément intégrable  
    $\mathfrak{S}$ -uniformément scalairement  
   intégrable, 53  
    $\mathfrak{S}$ -uniformément scalairement  
   intégrable (pour une mesure),  
   56  
   asymptotiquement uniformément  
    $p$ -intégrable, 8  
   par rapport à une famille de  
   mesures, 45  
   uniformément  $p$ -intégrable, 8

# Index des notations

$\mathfrak{B}(\mathbb{T})$ , 2 $\mathfrak{G}(\mathbb{T})$ , 2 $\mathfrak{F}(\mathbb{T})$ , 2 $\mathfrak{K}(\mathbb{T})$ , 2 $M(\mathbb{T}, \mathcal{F})$ , 3 $P(\mathbb{T}, \mathcal{F})$ , 3 $\pi_\Omega, \pi_\mathbb{T}$ , 3 $(\pi_\Omega)_\#(\mu), (\pi_\mathbb{T})_\#(\mu)$ , 3 $\mathcal{Y}(\mathbb{T})$ , 3 $\mathcal{Y}_d(\mathbb{T})$ , 3 $L^0(\mathbb{T})$ , 3 $\underline{d}_X$ , 4 $\tau_{\mathcal{Y}}^\times$ , ( $\times = S, M, N, W$ ), 4–5 $\underline{\mathfrak{G}}$ , 5 $\underline{\mathfrak{F}}$ , 5 $\underline{\mathfrak{K}}$ , 5 $g \otimes f$ , 6 $BL_1(\mathbb{T}, d)$ , 7 $\underline{BL}_1(\Omega, \mathbb{T}, d)$ , 7 $\underline{BL}'_1(\Omega, \mathbb{T}, d)$ , 7 $\tau_{\mathcal{Y}, \mathcal{D}}^{\times, p}$ , ( $\times = S, M, N, W$ ), 8 $\mathcal{Y}_{\mathcal{D}}^p$ , 8 $L^p_{(\mathbb{T}, \mathcal{D})}$ , 9 $\tau_M^\times$ , ( $\times = S, M, N, W$ ), 10  $\mathcal{F}^*$ , 12 $\mathcal{F}_\mu$ , 12 $\underline{\mathfrak{G}}^*$ , 12 $\tau_M^S$ , 12  $C(\mathbb{T}, [0, 1])$ , 19 $\Delta_{\text{prob}}^{(d)}$ , 20 $\tau_{L^0}^{\text{prob}}$ , 20	$\tau_{\mathcal{Y}_d}^{\text{prob}}$ , 20–21 $\Delta_{\text{prob}}^{(d^{(f)})}$ , 21 $\tilde{\Delta}_{\text{prob}}^{(f)}(\mu, \nu)$ , 21 $\tau_{\mathcal{Y}}^{\text{prob}}$ , 21–22 $\underline{\Delta}_{\text{BL}}^{(d)}(\mu, \nu)$ , 21–22 $D(\mu, \nu)$ , 23, 44, 56 $\Delta_{\text{KR}}^{(c)}(\mu, \nu)$ , 23 $\Delta_{\text{KR}}^{(p, d)}(\mu, \nu)$ , 23 $\underline{D}(\mu, \nu)$ , 23 $\underline{\Delta}_{\text{KR}}^{(c)}(\mu, \nu)$ , 23 $\underline{\Delta}_{\text{KR}}^{(p, d)}(\mu, \nu)$ , 24 $\tau_{\mathcal{Y}, \mathcal{D}}^p$ , 24 $\Delta_L^{(c)}(\mu, \nu)$ , 26 $\text{Lip}_{\mathbb{T}}^{(c)}$ , 27 $\mathcal{Y}_c^1$ , 27 $\Delta_L^{(c)}(\mu, \nu)$ , 27 $\mu \otimes \nu$ , 28  $T$ , 29 $\mathfrak{F}, (\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ , 29 $\mathbb{H}$ , 29 $\mathbb{U}$ , 29 $\mathbb{L}$ , 29 $\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})$ , 29 $\mathfrak{K}_c^*(\mathbb{E})$ , 29 $\text{Hausd}_{\mathbb{E}}$ , 29 $(S(t))_{t \geq 0}$ , 29 $C_{\text{croiss}}$ , 30 $L$ , 30 $C_{\text{Conv}}$ , 30 $\Phi$ , 31 $C([0, T]; \mathbb{H})$ , 31
--	---

$\Psi(\Lambda)$ , 32  
 $(X_n), (S_n), (C_n), (Z_n)$ , 34  
 $\mathbf{D}(\mathbb{R}^+; \mathbb{E})$ , 35–36  
 $\text{Id}_{\mathbb{R}^+}$ , 35  
 $\mathbf{C}(R^+; \mathbb{E})$ , 35–36  
 $G_n(\omega, t)$ , 36  
  
 $\mathbb{M}$ , 40, 44  
 $[x, y], [A, B]$ , 40  
 $\sum_{1 \leq i \leq n}^* p_i x_i$ , 41  
 $\text{bar } \mu$ , 42, 57  
 $P^1(\mathbb{M})$ , 44  
 $\Delta_W(\mu, \nu)$ , 44  
 $P^e(\mathbb{M})$ , 44  
 $L(X)$ , 45  
 $\mathbb{D}$ , 47  
 $\mathcal{S}$ , 47  
 $\psi^N$ , 47  
 $\varphi$ , 47  
 $\|(p_k)\|_a$ , 47  
 $\ell_a^1$ , 47  
 $[\mu, \nu]_a$ , 47  
 $F_n \xrightarrow[n \rightarrow \infty]{\text{WIJSMAN}} F$ , 50  
 $F_n \xrightarrow[n \rightarrow \infty]{\text{HAUSDORFF}} F$ , 50  
  
 $B_{\mathbb{E}'}(0, 1)$ , 52  
 $\mathbf{E} X$ , 52  
 $\mathfrak{S}$ , 52  
 $N_A(x)$ , 52  
 $L_{\mathbb{E}'}^1[\mathbb{E}]$ , 54  
  
 $\text{MU}_E^1$ , 56  
 $\text{MP}_{\mathbb{E}}^1$ , 56  
 $d_A(P, Q)$ , 56  
 $\tau_{\text{WP}}$ , 56  
 $\flat(\mathfrak{S})$ , 56  
 $\sharp(\mathfrak{S})$ , 56  
 $[A, B]_\alpha$ , 57