



HAL
open science

Une introduction à la convergence d'espaces métriques mesurés

Frédéric Paulin

► **To cite this version:**

| Frédéric Paulin. Une introduction à la convergence d'espaces métriques mesurés. 2009. hal-00628357

HAL Id: hal-00628357

<https://hal.science/hal-00628357>

Preprint submitted on 2 Oct 2011

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.

Une introduction à la convergence d'espaces métriques mesurés

Frédéric Paulin

Version préliminaire, 2 octobre 2011

1 Introduction

Ce texte est une introduction élémentaire à la convergence d'espaces métriques compacts munis de mesures de probabilité, en suivant les textes de [Gro2, Fuk1, Vil, Stu, Jac]. Nous ne considérerons pas dans ces notes le cas d'espaces non compact et de mesures infinies, en renvoyant à ces références.¹

Notations. Notons Dis_1 l'ensemble des classes d'isométries d'espaces métriques (X, d) compacts non vide de diamètre au plus 1.

Notons $\mathcal{M}\mathcal{M}_1$ l'ensemble des classes d'équivalence des triplets (X, d, μ) d'espaces métriques compacts de diamètre au plus 1, munis d'une mesure borélienne de probabilité, modulo isométries préservant la mesure.

Nous noterons $\mathcal{V}_\epsilon(A)$ le ϵ -voisinage ouvert d'une partie A d'un espace métrique. Une partie A d'un espace métrique X est ϵ -dense si $\mathcal{V}_\epsilon(A) = X$, i.e. si tout point de x est à distance strictement inférieure à ϵ d'au moins un point de A .

Si A et B sont deux parties d'un espace métrique (X, δ) , la distance de Hausdorff entre A et B , notée $d_{\text{Hau}}(A, B)$ (ou $d_{\text{Hau}, \delta}(A, B)$ pour indiquer la dépendance en la distance δ), est la borne inférieure des $\epsilon > 0$ tels que $A \subset \mathcal{V}_\epsilon(B)$ et $B \subset \mathcal{V}_\epsilon(A)$.

Nous noterons π_X, π_Y les projections canoniques du produit $X \times Y$ dans X et Y .

Le reste de l'introduction consiste en des remarques sur la définition des ensembles Dis_1 et $\mathcal{M}\mathcal{M}_1$.

Remarques. (1) Nous noterons de la même manière un élément de Dis_1 et l'un de ses représentants. Lorsque le besoin s'en fera sentir, nous noterons d_X la distance d'un élément X de Dis_1 . L'hypothèse de diamètre au plus égal à 1 peut être facilement obtenue en normalisant la distance d'un espace métrique compact non vide.

(2) L'inverse d'une isométrie préservant la mesure est encore une isométrie préservant la mesure. Nous noterons de la même manière un élément de $\mathcal{M}\mathcal{M}_1$ et l'un de ses représentants. Lorsque le besoin s'en fera sentir, nous noterons $d_X, \mu_X, \mathcal{B}_X$ la distance, la mesure et la tribu des boréliens d'un élément X de $\mathcal{M}\mathcal{M}_1$. L'hypothèse de masse totale égale à 1 peut être facilement obtenue en normalisant une mesure finie non nulle.

1. Je remercie les organisateurs V. Beffara, G. Miermeont, S. Sheffield, S. Smirnov de la conférence "Random planar geometry" aux Diablerets du 8 au 14 février 2009, pour des conditions inoubliables de neige et de ski pendant lesquelles une partie de ces notes ont été rédigées.

(3) La collection des classes d'isomorphismes d'espaces de probabilité n'a guère d'intérêt, puisqu'il n'existe qu'un seul espace de probabilité standard non atomique, à isomorphisme d'espaces mesurés près.

(4) Il est difficile de traiter de manière symétrique mesures et distances, ces objets n'ayant pas le même sens de covariance.

En effet, les distances ont tendance à être tirées en arrière : si X, Y sont des ensembles, $f : X \rightarrow Y$ une application, et d une pseudo-distance (vérifiant toutes les propriétés des distances sauf peut-être l'axiome de séparation) sur Y , alors $f^*d : X \times X \rightarrow \mathbb{R}$, définie par

$$f^*d(x, x') = d(f(x), f(x')) ,$$

est une pseudo-distance sur X (qui, même si d est une distance, n'est pas une distance si f n'est pas injective).

Les tribus se tirent en arrière : si X, Y sont des ensembles, $f : X \rightarrow Y$ une application, et \mathcal{A} une tribu (σ -algèbre) sur Y , alors

$$f^{-1}\mathcal{A} = \{f^{-1}(A) : A \in \mathcal{A}\}$$

est une tribu sur X (la plus petite rendant f mesurable).

Mais les mesures ont tendance à être poussées en avant : si X, Y sont des ensembles, $f : X \rightarrow Y$ une application, et μ une mesure extérieure sur X , alors $f_*\mu : \mathcal{P}(Y) \rightarrow \mathbb{R}$, définie par

$$\forall A \subset Y, \quad f_*\mu(A) = \mu(f^{-1}(A)) ,$$

est une mesure extérieure sur Y . Mais si f n'est pas injective, alors f n'est pas forcément mesurable pour les tribus des parties mesurables de X et de Y pour les mesures extérieures μ et $f_*\mu$ respectivement. Bien sûr, si X et Y sont munis de tribus rendant f mesurable, si μ est une mesure sur X , alors $f_*\mu$ est une mesure sur Y .

(5) Rappelons que si X est un espace topologique, et si μ une mesure borélienne sur X , alors le *support* $\text{Supp } \mu$ de X est l'intersection de tous les fermés de X de complémentaire de mesure nulle (qui est le plus petit fermé de complémentaire de mesure nulle si X est à base dénombrable d'ouverts, car alors la réunion des ouverts de mesure nulle est égale à la réunion des éléments d'une base dénombrable d'ouverts contenus dans un ouvert de mesure nulle, donc est de mesure nulle). Si Y est un autre espace topologique, et si $f : X \rightarrow Y$ est une application continue, alors

$$f(\text{Supp } \mu) \subset \text{Supp}(f_*\mu) , \quad (*)$$

avec égalité si f est un homéomorphisme (mais pas en général) : en effet, si F est un fermé de Y tel que $f_*\mu({}^cF) = 0$, alors $f^{-1}(F)$ est un fermé de X tel que $\mu({}^c(f^{-1}(F))) = 0$, donc $f(\text{Supp } \mu)$ est contenu dans $f(f^{-1}(F)) \subset F$.

On a parfois envie de ne considérer que les supports des mesures, et d'oublier les comportements des distances en dehors de ceux-ci. On considère alors l'ensemble $\overline{\mathcal{M}\mathcal{M}_1}$ des classes d'équivalence d'espaces (X, d, μ) métriques compacts de diamètre au plus 1 muni d'une mesure borélienne de probabilité de support total, modulo isométries préservant la mesure. Par définition, $\overline{\mathcal{M}\mathcal{M}_1}$ est un sous-ensemble de $\mathcal{M}\mathcal{M}_1$.

Mais c'est aussi un quotient de $\mathcal{M}\mathcal{M}_1$, car l'application de $\mathcal{M}\mathcal{M}_1$ dans $\overline{\mathcal{M}\mathcal{M}_1}$ qui à (X, d, μ) associe le support de μ muni des restrictions de d et de μ à ce support, est

surjective (notons qu'elle est bien définie : $\text{Supp } \mu$ est fermé donc compact, de diamètre au plus 1; et si $f : (X, d, \mu) \rightarrow (X', d', \mu')$ est une isométrie préservant la mesure, alors $f(\text{Supp } \mu) = \text{Supp } \mu'$ par $(*)$, et donc f induit une isométrie préservant la mesure de $\text{Supp } \mu$ dans $\text{Supp } \mu'$).

Comme $\overline{\mathcal{M}\mathcal{M}_1}$ est à la fois un sous-ensemble et un ensemble quotient de $\mathcal{M}\mathcal{M}_1$, et que nous pouvons considérer à la fois les distances induites et les distances quotients sur $\overline{\mathcal{M}\mathcal{M}_1}$ suivant le problème considéré, nous nous concentrerons dans ces notes sur l'ensemble $\mathcal{M}\mathcal{M}_1$.

(6) Introduisons une classe d'espaces métriques, dans laquelle des arguments de chemins donnent des informations intéressantes sur la distance. La terminologie ci-dessous, introduite par Gromov, est la plus fréquemment utilisée (voir par exemple [BH], mais ce n'est pas celle de [BBI]). Un espace métrique (X, d) est *géodésique* si pour tous x, y dans X , il existe une application isométrique $f : [a, b] \rightarrow X$ telle que $f(a) = x$ et $f(b) = y$ (en particulier, $d(x, y) = |b - a|$). Pour tout $\epsilon > 0$, un espace métrique (X, d) est ϵ -*géodésique* si pour tous x, y dans X , il existe des points $x_0 = x, x_1, \dots, x_n, x_{n+1} = y$ tels que $d(x_i, x_{i+1}) \leq \epsilon$ pour $0 \leq i \leq n$ et $\sum_{i=0}^n d(x_i, x_{i+1}) \leq d(x, y) + \epsilon$. Bien sûr, un espace géodésique est ϵ -géodésique pour tout $\epsilon > 0$, et tout espace métrique de diamètre au plus 1 est 1-géodésique. Le lemme suivant est classique.

Lemme 1.1 *Si x et y sont deux points d'un espace métrique X , tels qu'il existe des points $x_0 = x, x_1, \dots, x_n, x_{n+1} = y$ de X tels que $d(x_i, x_{i+1}) \leq \epsilon$ et $\sum_{i=0}^n d(x_i, x_{i+1}) \leq d(x, y) + \epsilon$, alors il existe une application $f : [0, d(x, y)] \rightarrow X$ telle que $f(0) = x$, $f(d(x, y)) = y$ et pour tous s, t dans $[0, d(x, y)]$,*

$$|d(f(s), f(t)) - |s - t|| \leq 2\epsilon .$$

Preuve. Notons tout d'abord que pour $0 \leq i \leq j \leq n + 1$, nous avons $d(x_j, x_i) \leq d(x_j, x) - d(x_i, x) + \epsilon$. En effet, par l'inégalité triangulaire, $d(x, x_i) + d(x_i, x_j) + d(x_j, y) \leq \sum_{i=0}^n d(x_i, x_{i+1}) \leq d(x, y) + \epsilon$. Donc $d(x, x_i) + d(x_i, x_j) \leq d(x, y) - d(x_j, y) + \epsilon \leq d(x, x_j) + \epsilon$, ce qui montre le résultat.

Pour tout $t \in [0, d(x, y)]$, posons $i_t = \min\{i \in \{0, \dots, n + 1\} : d(x, x_i) \geq t\}$, et $f(t) = x_{i_t}$. Notons que $i_t = 0$ si et seulement si $t = 0$. Soient $0 \leq t \leq s \leq 1$. Alors $i_t \leq i_s$. Si $i_s = 0$, alors $s = t = 0$, et le résultat est clair. Si $i_s \neq 0$, alors par définition de i_t et de i_s ,

$$\begin{aligned} d(x_{i_s}, x_{i_t}) &\leq d(x_{i_s}, x) - d(x_{i_t}, x) + \epsilon \leq d(x_{i_s}, x) - t + \epsilon \\ &\leq d(x_{i_s}, x_{i_s-1}) + d(x_{i_s-1}, x) - t + \epsilon \leq s - t + 2\epsilon . \end{aligned}$$

Réciproquement, si $i_t = 0$, alors $d(x_{i_s}, x_{i_t}) = d(x_{i_s}, x) \geq s = s - t$. Si $i_t \neq 0$, alors

$$\begin{aligned} d(x_{i_s}, x_{i_t}) &\geq d(x_{i_s}, x) - d(x_{i_t}, x) \geq s - d(x_{i_t}, x_{i_t-1}) - d(x_{i_t-1}, x) \\ &\geq s - t - \epsilon , \end{aligned}$$

ce qui montre le résultat. □

2 Distance de Hausdorff-Gromov

Soient X, Y deux espaces métriques, et $\epsilon > 0$. Appelons *couplage* (terminologie de K.-T. Sturm [Stu]) entre X et Y toute distance δ sur $X \amalg Y$ telle que $\delta|_{X \times X} = d_X$ et

$\delta|_{Y \times Y} = d_Y$. Notons $\Delta(X, Y)$ l'ensemble de ces couplages. Appelons ϵ -couplage entre X et Y tout couplage δ tel que $d_{\text{Hau}, \delta}(X, Y) < \epsilon$ (c'est-à-dire tel que, pour tout x dans X (respectivement y dans Y), il existe y dans Y (respectivement x dans X) tel que $\delta(x, y) < \epsilon$).

Par exemple, si $\delta(x, y) = \delta(y, x) = 1$ pour tous x dans X et y dans Y , alors δ , étendue par d_X et d_Y sur $X \times X$ et $Y \times Y$ respectivement, est un 1-couplage. Pour tout $\epsilon > 0$, si $f : X \rightarrow Y$ est une éventuelle isométrie, si $\delta(x, y) = \delta(y, x) = d(f(x), y) + \epsilon$ pour tous x dans X et y dans Y , alors δ , étendue par d_X et d_Y sur $X \times X$ et $Y \times Y$ respectivement, est un ϵ -couplage.

Appelons *approximation* entre X et Y toute relation $\mathcal{R} \subset X \times Y$ telle que $\pi_X(\mathcal{R}) = X$ et $\pi_Y(\mathcal{R}) = Y$. Appelons ϵ -approximation entre X et Y toute approximation $\mathcal{R} \subset X \times Y$ telle que, pour tous x, x' dans X , pour tous y, y' dans Y , si $x \mathcal{R} y$ et $x' \mathcal{R} y'$, alors $|d_X(x, x') - d_Y(y, y')| < \epsilon$. Par exemple, pour tout $\epsilon > 0$, le produit $X \times Y$ tout entier est une $(1 + \epsilon)$ -approximation, et le graphe d'une éventuelle isométrie de X dans Y est une ϵ -approximation.

Appelons ϵ -isométrie de X dans Y toute application $f : X \rightarrow Y$, qui est ϵ -isométrique (ou $(1, \epsilon)$ -quasi-isométrique), i.e. telle que

$$\forall x, x' \in X, \quad |d_Y(f(x), f(x')) - d_X(x, x')| < \epsilon$$

et dont l'image est ϵ -dense, i.e. telle que

$$\forall y \in Y, \exists x \in X \quad d_Y(f(x), y) < \epsilon.$$

Par exemple, pour tout $\epsilon > 0$, toute application constante de X dans Y est une $(1 + \epsilon)$ -isométrie de X dans Y ; une éventuelle isométrie de X dans Y est une ϵ -approximation de X dans Y ; l'inclusion d'une partie ϵ -dense dans un espace métrique est une ϵ -isométrie.

Remarquons que le lemme 1.1 dit que si un espace métrique X est ϵ -géodésique, alors pour tous x et y dans X , il existe une application 2ϵ -isométrique de $[0, d(x, y)]$ dans X valant x et y respectivement en 0 et $d(x, y)$.

Appelons ϵ -bisométrie entre X et Y tout couple (f, g) d'applications ϵ -isométriques $f : X \rightarrow Y$ et $g : Y \rightarrow X$, telles que

$$\forall x \in X, d_X(g \circ f(x), x) < \epsilon \quad \text{et} \quad \forall y \in Y, d_Y(f \circ g(y), y) < \epsilon.$$

Remarques. (1) (Voir par exemple [Pau1, Lemme 1.1]) Si δ est un ϵ -couplage entre X et Y , alors la relation \mathcal{R} définie par $x \mathcal{R} y$ si et seulement si $\delta(x, y) < \epsilon$ est une 2ϵ -approximation entre X et Y . Réciproquement, si \mathcal{R} est une 2ϵ -approximation entre X et Y , si, pour tous x dans X et y dans Y , on pose

$$\delta(x, y) = \delta(y, x) = \inf_{x' \mathcal{R} y'} d_X(x, x') + d_Y(y', y) + \epsilon,$$

alors l'application δ , étendue par d_X et d_Y sur $X \times X$ et $Y \times Y$ respectivement, est un ϵ' -couplage entre X et Y , pour tout $\epsilon' > \epsilon$.

(2) Le graphe de l'identité de X dans X est une ϵ -approximation entre X et X pour tout $\epsilon > 0$. Si \mathcal{R} est une ϵ -approximation entre X et Y , alors la relation \mathcal{R}^{-1} entre Y et X définie par $y \mathcal{R}^{-1} x$ si et seulement si $x \mathcal{R} y$ est une ϵ -approximation. Si \mathcal{R} est une ϵ -approximation entre X et Y , et \mathcal{R}' est une ϵ' -approximation entre Y et Z , alors la relation

\mathcal{R}'' entre X et Z définie par $x \mathcal{R}'' z$ si et seulement si il existe $y \in Y$ tel que $x \mathcal{R} y$ et $y \mathcal{R}' z$ est une $(\epsilon + \epsilon')$ -approximation.

(3) S'il existe une ϵ -isométrie f entre X et Y , alors toute application $g : Y \rightarrow X$ telle que $d(f(g(y)), y) < \epsilon$ pour tout y dans Y est une 3ϵ -isométrie entre Y et X , et le couple (f, g) est une 3ϵ -biséométrie. Puisque f est d'image ϵ -dense, et par l'axiome du choix, une telle application g existe (en utilisant la séparabilité des espaces métriques compacts, et quitte à remplacer 3ϵ par 4ϵ , il n'y a pas besoin de l'axiome du choix). Réciproquement, si (f, g) est une ϵ -biséométrie, alors f et g sont des ϵ -isométries.

(4) Si $f : X \rightarrow Y$ est une ϵ -isométrie, alors la relation \mathcal{R} définie par $x \mathcal{R} y$ si et seulement si $d(f(x), y) < \epsilon$ est une 3ϵ -approximation entre X et Y . En particulier, s'il existe une ϵ -isométrie entre X et Y , alors il existe un $3\epsilon'/2$ -couplage entre X et Y , pour tout $\epsilon' > \epsilon$. Réciproquement, si \mathcal{R} est une ϵ -approximation entre X et Y , alors toute application $f : X \rightarrow Y$, telle que $x \mathcal{R} f(x)$ pour tout x dans X , est une ϵ -isométrie. Une telle application existe, par l'axiome du choix (mais là encore, quitte à affaiblir la conclusion en 2ϵ -isométrie, un argument de séparabilité permet d'éviter cet axiome).

(5) Bien que nous ne nous en servions pas, remarquons enfin que si (f, g) est une ϵ -biséométrie entre X et Y , si, pour tous x dans X et y dans Y , on pose

$$\delta(x, y) = \delta(y, x) = \min \left\{ \inf_{x' \in X} d_X(x, x') + d_Y(f(x'), y), \inf_{y' \in Y} d_X(x, g(y')) + d_Y(y', y) \right\} + 2\epsilon,$$

alors l'application δ , étendue par d_X et d_Y sur $X \times X$ et $Y \times Y$ respectivement, est un $2\epsilon'$ -couplage entre X et Y , pour tout $\epsilon' > \epsilon$.

Pour tous X, Y dans Dis_1 , appelons *distance de Hausdorff-Gromov* entre X et Y , et notons $d_{HG}(X, Y)$, la borne inférieure des $\epsilon > 0$ tels qu'il existe un ϵ -couplage entre X et Y , i.e.

$$d_{HG}(X, Y) = \inf_{\delta \in \Delta(d_X, d_Y)} d_{\text{Hau}, \delta}(X, Y).$$

Pour tous X, Y dans Dis_1 , notons $d'_{HG}(X, Y)$ la borne inférieure des $\epsilon > 0$ tels qu'il existe une ϵ -approximation entre X et Y , i.e., en notant $\|\cdot\|_\infty$ la norme uniforme sur les applications de $\mathcal{R} \times \mathcal{R}$ dans \mathbb{R} ,

$$d'_{HG}(X, Y) = \inf_{\mathcal{R} \text{ approximation entre } X \text{ et } Y} \|\pi_X^* d_X - \pi_Y^* d_Y\|_\infty.$$

Proposition 2.1 *Le couple (Dis_1, d_{HG}) est un espace métrique séparable complet, et $d_{HG} = \frac{1}{2} d'_{HG}$.*

Certes appelée distance de Gromov-Hausdorff dans [BBI], et plus connue sous ce nom, cette distance a été introduite par Gromov, qui l'appelait distance de Hausdorff. Dans la fin des années 80, elle était connue sous le nom de distance de Hausdorff-Gromov (voir par exemple [Pau1]), et nous continuerons de l'appeler ainsi.

Preuve. Nous renvoyons aux références citées pour les vérifications de détail de la preuve, en nous contentant des indications suivantes (quasi-complètes quand même!).

- Le fait que d_{HG} et d'_{HG} soient des distances homothétiques découle des remarques (1) et (2) ci-dessus.

- Tout espace métrique compact contient une partie finie ϵ -dense, pour tout $\epsilon > 0$, et est donc à distance de Hausdorff-Gromov au plus ϵ d'un espace métrique fini de diamètre au plus 1. Notons \mathcal{D}_N l'ensemble des distances sur $\{1, \dots, N\}$, de diamètre au plus 1, muni de la distance uniforme. En tant que fermé borné dans l'espace vectoriel normé de dimension finie $(\mathbb{R}^{N^2}, \|\cdot\|_\infty)$, il est métrisable compact, donc séparable. D'autre part, l'application évidente (qui à un espace métrique associe sa classe d'isométrie) de \mathcal{D}_N dans Dis_1 est continue.

- Toute suite de Cauchy dans Dis_1 est en fait uniformément précompacte, au sens ci-dessous, donc admet une sous-suite convergente par le théorème 2.3 ci-dessous, donc converge. \square

La caractérisation suivante de la convergence pour cette distance découle des remarques (3) et (4) ci-dessus, qui montrent que

- s'il existe une ϵ -isométrie entre X et Y , alors $d_{HG}(X, Y) < \frac{3\epsilon}{2}$;
- si $d_{HG}(X, Y) < \epsilon$, alors il existe une 2ϵ -isométrie entre X et Y .

Proposition 2.2 *Soient $((X_i, d_i))_{i \in \mathbb{N}}$ une suite dans Dis_1 , et (X, d) un élément de Dis_1 . Les assertions suivantes sont équivalentes :*

- (1) *la suite $((X_i, d_i))_{i \in \mathbb{N}}$ converge vers (X, d) pour la distance de Hausdorff-Gromov ;*
- (2) *pour tout i dans \mathbb{N} , il existe une ϵ_i -isométrie de (X_i, d_i) dans (X, d) avec ϵ_i convergent vers 0 ;*
- (3) *pour tout i dans \mathbb{N} , il existe une ϵ_i -isométrie de (X, d) dans (X_i, d_i) avec ϵ_i convergent vers 0 ;*
- (4) *pour tout i dans \mathbb{N} , il existe une ϵ_i -bisométrie entre (X_i, d_i) et (X, d) , avec $\lim_{i \rightarrow +\infty} \epsilon_i = 0$. \square*

Vu l'importance de la notion de compacité, il est naturel de se demander quelles sont les parties compactes de l'espace métrique (Dis_1, d_{HG}) .

Un ensemble (ou une famille) \mathcal{M} d'espaces métriques non vides est dit *uniformément précompact* si leur diamètre est uniformément majoré et si l'une des deux conditions équivalentes suivantes est satisfaite :

- (1) pour tout $\epsilon > 0$, il existe $N \in \mathbb{N}$ tel que tout X dans \mathcal{M} admette un recouvrement par N boules de rayon ϵ ;
- (2) pour tout $\epsilon > 0$, il existe $N \in \mathbb{N}$ tel que tout X dans \mathcal{M} admette une partie ϵ -dense de cardinal N ;

Cette notion a été introduite par Gromov [Gro1] lors de sa célèbre preuve de son théorème sur les groupes à croissance polynomiale. Le résultat suivant dit ainsi que les parties compactes de (Dis_1, d_{HG}) sont exactement les parties fermées uniformément précompactes.

Théorème 2.3 (Critère de précompacité de Gromov) *Toute partie uniformément précompacte de Dis_1 est relativement compacte pour la distance de Hausdorff-Gromov.*

Le sous-espace de Dis_1 des espaces métriques géodésiques est fermé.

Le but de la fin de cette partie est de donner une preuve originale de ce critère de précompacité de Gromov (voir [Gro1, Gro2] pour les preuves originelles, ainsi que [BBI]), comme conséquence du théorème 2.6.

Dans le reste de cette partie, nous fixons un ultrafiltre ω sur \mathbb{N} plus fin que le filtre de Fréchet (voir [Bou, Chap. I, §6][Dix, §4.3]).

Remarque. Un ultrafiltre sur \mathbb{N} plus fin que le filtre de Fréchet est un filtre maximal sur \mathbb{N} , plus fin que le filtre des complémentaires des parties finies de \mathbb{N} . Il peut être défini comme une partie ω de $\mathcal{P}(\mathbb{N})$ vérifiant les conditions suivantes : pour tous $A, A' \in \omega$, pour tous $B \in \mathcal{P}(\mathbb{N})$,

- (1) A est non vide, $A \cap A' \in \omega$ et si $A \subset B$, alors $B \in \omega$;
- (2) $B \in \omega$ ou ${}^c B \in \omega$;
- (3) si B est finie, alors $B \notin \omega$.

Son existence découle de l'axiome du choix. Nous dirons qu'une propriété est *vraie pour ω -presque tout $i \in \mathbb{N}$* lorsque l'ensemble des $i \in \mathbb{N}$ pour lesquels la propriété est vérifiée appartient à l'ultrafiltre ω .

Cette terminologie est justifiée par la raison suivante. Par les propriétés (1) à (3) ci-dessus, l'application $\mu : \mathcal{P}(\mathbb{N}) \rightarrow \{0, 1\}$, définie par $\mu(A) = 1$ si et seulement si $A \in \omega$, est une mesure finiment additive du tout ou rien, nulle sur les parties finies. De plus, l'application $\mu \mapsto \mu^{-1}(1)$ est une bijection de l'ensemble des mesures finiment additives du tout ou rien sur \mathbb{N} nulles sur les parties finies, à valeurs dans l'ensemble des ultrafiltres sur \mathbb{N} plus fins que le filtre de Fréchet.

Il faut penser à ω comme à un moyen magique de choisir une valeur d'adhérence de toute suite relativement compacte (i.e. dont l'adhérence de l'image est compacte — et en particulier séparée! —). Il n'est pas nécessaire de s'en servir pour obtenir les résultats de cette note (voir les preuves originales). Mais cela simplifie tellement les arguments de passages à des suites extraites et autres extractions diagonales, que c'est en un bonheur! Voir par exemple [Gro3, Pau2, Dru] pour des exemples d'applications. Le point clef concernant ω est en effet le résultat suivant (dont la preuve est tout simplement, outre l'unicité des limites de filtres dans un espace séparé, que le filtre image d'un ultrafiltre est un ultrafiltre, et que tout ultrafiltre sur un espace topologique compact converge, voir par exemple [Bou, Dix]).

Proposition 2.4 *Pour tout espace topologique T , pour toute suite $(x_i)_{i \in \mathbb{N}}$ dans T relativement compacte, il existe une et une seule valeur d'adhérence de cette suite, notée $x_\omega = \lim_\omega x_i$, telle que, pour tout $\epsilon > 0$, pour ω -presque tout i , on ait $|x_i - x_\omega| < \epsilon$. \square*

Fixons-nous une suite $((X_i, d_i))_{i \in \mathbb{N}}$ dans Dis_1 .

Construisons une pseudo-distance \tilde{d}_ω sur l'ensemble produit $\prod_{i \in \mathbb{N}} X_i$, par

$$\tilde{d}_\omega((x_i)_{i \in \mathbb{N}}, (y_i)_{i \in \mathbb{N}}) = \lim_\omega d_i(x_i, y_i).$$

Par les propriétés de limites de filtres, \tilde{d}_ω est bien une pseudo-distance, de diamètre au plus 1. Notons (X_ω, d_ω) l'espace métrique quotient de $(\prod_{i \in \mathbb{N}} X_i, \tilde{d}_\omega)$. En particulier X_ω est l'ensemble quotient de $\prod_{i \in \mathbb{N}} X_i$ par la relation d'équivalence $u \sim v$ si et seulement si $\tilde{d}_\omega(u, v) = 0$, et $d_\omega([u], [v]) = \tilde{d}_\omega(u, v)$ pour n'importe quels représentants u, v de classes $[u], [v]$ respectivement. Nous noterons $[x_i]$ la classe dans X_ω d'une suite $(x_i)_{i \in \mathbb{N}}$.

Si $((X'_i, d'_i))_{i \in \mathbb{N}}$ est une autre suite dans Dis_1 , et si $f_i : X_i \rightarrow X'_i$ est une application telle que

$$\forall (x_i)_{i \in \mathbb{N}}, (y_i)_{i \in \mathbb{N}} \in \prod_{i \in \mathbb{N}} X_i, \quad \text{si } \lim_\omega d_i(x_i, y_i) = 0, \quad \text{alors } \lim_\omega d'_i(f_i(x_i), f_i(y_i)) = 0,$$

alors l'application $f_\omega : X_\omega \rightarrow X'_\omega$ définie par

$$f_\omega([x_i]) = [f_i(x_i)]$$

est bien définie.

Remarque 2.5 Soient X et Y deux espaces métriques, et $\lambda \geq 1, c \geq 0$. Une application f de X dans Y est dite (λ, c) -quasi-isométrie si

$$\forall x, x' \in X, \quad \frac{1}{\lambda} d_X(x, x') - c \leq d_Y(f(x), f(x')) \leq \lambda d_X(x, x') + c.$$

Si $f_i : X_i \rightarrow X'_i$ est une application (λ_i, ϵ_i) -quasi-isométrique, si f_ω est bien définie, et si $\lambda = \lim_\omega \lambda_i$ et $\epsilon = \lim_\omega \epsilon_i$ sont finis, alors f_ω est une application (λ, ϵ) -quasi-isométrique. Notons que si $f_i : X_i \rightarrow X'_i$ est une application (λ_i, ϵ_i) -quasi-isométrique, et si $\lambda = \lim_\omega \lambda_i$ est fini et $\lim_\omega \epsilon_i = 0$, alors f_ω est bien définie; si de plus f_i est d'image ϵ_i -dense pour ω -presque tout i , alors f_ω est un homéomorphisme λ -bilipschitzien. En particulier, si f_i est une ϵ_i -isométrie pour ω -presque tout i , et si $\lim_\omega \epsilon_i = 0$, alors f_ω est une isométrie.

Théorème 2.6 *L'espace métrique (X_ω, d_ω) est complet, de diamètre au plus 1.*

Si X_i est ϵ_i -géodésique avec $\lim_\omega \epsilon_i = 0$, alors X_ω est géodésique.

Si la suite $((X_i, d_i))_{i \in \mathbb{N}}$ d'espaces métriques est uniformément précompacte, alors l'espace métrique (X_ω, d_ω) est compact, et

$$\lim_\omega (X_i, d_i) = (X_\omega, d_\omega)$$

dans l'espace métrique (Dis_1, d_{HG}) .

En particulier, puisque dans un espace métrique, les valeurs d'adhérence de suites sont les limites des sous-suites, si $((X_i, d_i))_{i \in \mathbb{N}}$ est uniformément précompacte, alors il existe une sous-suite de cette suite qui converge vers (X_ω, d_ω) . Le théorème 2.3 en découle.

Preuve du théorème 2.6. • Montrons que le diamètre est au plus 1 : pour tous $[x_i], [y_i]$ dans X_ω , nous avons $d_\omega([x_i], [y_i]) = \lim_\omega d_i(x_i, y_i) \leq 1$, car X_i est de diamètre au plus 1 (et par passage à la limite des inégalités larges).

Remarque. En fait, si le diamètre de X_i est égal à 1 pour (ω -presque) tout i , alors le diamètre de X_ω est aussi égal à 1 : si pour tout i , il existe x_i, y_i dans X_i tels que $d_i(x_i, y_i) = 1$, alors $d_\omega([x_i], [y_i]) = 1$.

- Nous renvoyons à [DW] pour une preuve de la complétude.
- Supposons que X_i soit géodésique (pour ω -presque tout i). Soient $x = [x_i]$ et $y = [y_i]$ deux points distincts de X_ω . Puisque X_i est géodésique, il existe une application isométrique $f_i : [0, d_i(x_i, y_i)] \rightarrow X_i$ telle que $f_i(0) = x_i$, $f_i(d_i(x_i, y_i)) = y_i$. Alors l'application $f : [0, d_\omega(x, y)] \rightarrow X_i$ définie par $t \mapsto [f_i(t \frac{d_i(x_i, y_i)}{d_\omega(x, y)})]$ est une géodésique entre x et y .

Pour traiter le cas où X_i est ϵ_i -géodésique avec $\lim_\omega \epsilon_i = 0$, on raisonne de la même manière, en utilisant le lemme 1.1.

- Supposons dans la fin de cette preuve que la suite $((X_i, d_i))_{i \in \mathbb{N}}$ soit uniformément précompacte. Montrons que (X_ω, d_ω) est compact.

Pour tout $\epsilon > 0$, soit $N = N(\epsilon) \in \mathbb{N}$ et pour tout $i \in \mathbb{N}$, soit $(y_{i,k})_{1 \leq k \leq N}$ une partie ϵ -dense dans X_i . Montrons que la famille finie $([y_{i,k}])_{1 \leq k \leq N}$ est 2ϵ -dense dans X_ω . Puisque (X_ω, d_ω) est complet, et par le théorème de Bolzano-Weierstrass, ceci conclura.

Pour tout $[x_i]$ dans X_ω , pour tout i dans \mathbb{N} , soit $k_i \in \{1, \dots, N\}$ tel que $d(x_i, y_{i,k_i}) < \epsilon$. Puisque k_i varie dans un ensemble (discret) fini, soit $k_\omega = \lim_\omega k_i \in \{1, \dots, N\}$, qui vérifie que $k_i = k_\omega$ pour ω -presque tout i . Alors

$$d_\omega([x_i], [y_{i,k_\omega}]) = \lim_\omega d_i(x_i, y_{i,k_\omega}) = \lim_\omega d_i(x_i, y_{i,k_i}) \leq \epsilon < 2\epsilon,$$

ce qui conclut.

• Enfin, montrons que les espaces (X_i, d_i) convergent vers (X_ω, d_ω) pour la distance de Hausdorff-Gromov.

Fixons $\epsilon > 0$. Par compacité de (X_ω, d_ω) , soit $\{x_k\}_{1 \leq k \leq N}$ une partie finie $\frac{\epsilon}{3}$ -dense dans X_ω . Pour tout $k \in \{1, \dots, N\}$, choisissons un représentant $(x_{i,k})_{i \in \mathbb{N}}$ de x_k . Pour tout x dans X_ω , soit k_x tel que $d_\omega(x, x_{k_x}) < \frac{\epsilon}{3}$. Considérons l'application $f_i : X_\omega \rightarrow X_i$ définie par $x \mapsto x_{i,k_x}$. Montrons que pour ω -presque tout i , l'application f_i est une ϵ -isométrie. Par la proposition 2.2, ceci montrera le résultat.

Pour tous $k, \ell \in \{1, \dots, N\}$, nous avons $d_\omega(x_k, x_\ell) = \lim_\omega d_i(x_{i,k}, x_{i,\ell})$. Donc par finitude, pour ω -presque tout i , pour tous $k, \ell \in \{1, \dots, N\}$, nous avons $|d_\omega(x_k, x_\ell) - d_i(x_{i,k}, x_{i,\ell})| < \frac{\epsilon}{3}$. Pour cet ensemble de mesure pleine de i , nous avons donc, pour tous x, y dans X_ω , en posant $k = k_x$ et $\ell = k_y$,

$$|d_\omega(x, y) - d_i(f_i(x), f_i(y))| < \frac{2\epsilon}{3} + |d_\omega(x_k, x_\ell) - d_i(x_{i,k}, x_{i,\ell})| < \epsilon,$$

donc f_i est ϵ -isométrique.

Si l'image de f_i n'est pas ϵ -dense pour ω -presque tout i , alors pour ω -presque tout i , il existe y_i dans X_i tel que pour tout x dans X_ω , nous avons $d_i(f_i(x), y_i) \geq \epsilon$. Posons $y = [y_i] \in X_\omega$ (en complétant la suite des y_i par n'importe quel point, pour i en dehors de l'ensemble de ω -mesure pleine considéré). Notons $k = k_y$. Alors nous avons, par passage à la limite des inégalités,

$$\epsilon \leq \lim_\omega d_i(f_i(y), y_i) = \lim_\omega d_i(x_{i,k}, y_i) = d_\omega(x_k, y) < \frac{\epsilon}{3},$$

contradiction. □

Remarque. L'espace métrique X_ω n'est en général pas localement compact, ni même séparable. En particulier, la théorie de la mesure sur X_ω peut être délicate à considérer, voire à définir.

Par exemple, si (X, d) est l'arbre régulier de degré $k + 1$ avec $k \geq 2$ muni de la distance géodésique rendant chaque arête isométrique à $[0, 1]$ (ou le plan hyperbolique réel (à courbure constante -1) muni de sa distance riemannienne, et si $X_i = (B_X(x_0, i), \frac{1}{i}d)$, qui appartient à Dis_1 pour tout point base x_0 dans X , alors X_ω est un arbre réel de diamètre 1, dans lequel se plonge isométriquement tout arbre réel séparable de diamètre au plus 1, mais il n'est ni séparable, ni localement compact (voir par exemple [Gro3, DP, MO]).

Pour un exemple plus élémentaire, pour tout $i \in \mathbb{N}$, notons X_i le cube $[0, 1]^i$ muni de la distance uniforme (induite par la norme $\|(x_j)_{0 \leq j \leq i-1}\|_\infty = \max_{0 \leq j \leq i-1} |x_j|$ sur \mathbb{R}^i). Considérons l'espace métrique $\ell_\infty([0, 1])$ des suites à valeurs dans $[0, 1]$ muni de la distance uniforme $d_\omega((x_j)_{j \in \mathbb{N}}, (y_j)_{j \in \mathbb{N}}) = \sup_{j \in \mathbb{N}} |x_j - y_j|$. Il n'est pas séparable (voir par

exemple [Bré, Lem. IV.3], l'ensemble des boules ouvertes de rayon $1/2$ et de centre les suites ne prenant que 0 ou 1 comme valeurs, qui sont deux à deux disjointes, est non dénombrable). Il se plonge isométriquement dans X_ω par l'application $(x_j)_{j \in \mathbb{N}} \mapsto [(x_j)_{0 \leq j \leq i-1}]$, car $\lim_\omega \max_{0 \leq j \leq i-1} |x_j| = \sup_{j \in \mathbb{N}} |x_j|$. Comme tout sous-espace d'un espace métrique séparable est séparable, X_ω n'est pas séparable.

3 Distance de Hausdorff-Gromov mesurée

Soient (X, d) un espace métrique compact (dont la tribu des boréliens sera notée \mathcal{B}_X quand nécessaire), et $\text{Prob}(X)$ l'ensemble des mesures boréliennes de probabilité sur X .

L'espace $\mathcal{M}_\mathbb{R}(X)$ des mesures signées sur X est le dual topologique de l'espace de Banach $\mathcal{C}(X, \mathbb{R})$ des fonctions continues réelles sur X (pour la norme uniforme). L'espace vectoriel normé $\mathcal{C}(X, \mathbb{R})$ étant séparable, la boule unité (pour la norme duale $\mu \mapsto |\mu|(X)$) de $\mathcal{M}_\mathbb{R}(X)$, et donc son sous-ensemble $\text{Prob}(X)$, est métrisable, par exemple, en notant $(f_n)_{n \in \mathbb{N}}$ une suite dense dans $\mathcal{C}(X, \mathbb{R})$, par

$$d(\mu, \nu) = \sum_{n \in \mathbb{N}} 2^{-n} \max\{1, |\mu(f_n) - \nu(f_n)|\}.$$

Mais cette distance est peu pratique à manipuler, et ne reflète pas la géométrie de la distance de X .

Toutefois (voir par exemple [Bil, Par, Doo]), il existe de nombreuses distances géométriquement pertinentes sur $\text{Prob}(X)$, de topologie induite égale à la topologie *étroite* (ou *faible-étoile* ou *vague*, ces trois topologies, respectivement les moins fines rendant continues les évaluations $\mu \mapsto \mu(f)$ des mesures en les applications f continues bornées, continues nulles à l'infini, continues à support compact, coïncidant par compacité de X). Nous n'en donnons que deux.

La *distance de Prokhorov* sur $\text{Prob}(X)$ est

$$d_{\text{Pro}}(\mu, \nu) = \inf \{ \epsilon > 0 : \forall A \in \mathcal{B}_X, \mu(A) < \nu(\mathcal{V}_\epsilon(A)) + \epsilon \text{ et } \nu(A) < \mu(\mathcal{V}_\epsilon(A)) + \epsilon \}.$$

Bien sûr, la distance de Prokhorov sur $\text{Prob}(X)$ dépend de la distance d sur X , et la nous noterons $d_{\text{Pro}, d}$ lorsque nous voulons préciser d .

Soient μ et ν deux mesures boréliennes de probabilité sur des espaces métriques compacts X et Y . Un *couplage* de μ et ν est une mesure (borélienne, de probabilité) λ sur l'espace borélien $X \times Y$ telle que, en notant π_X et π_Y les projections de $X \times Y$ sur X et sur Y respectivement,

$$(\pi_X)_* \lambda = \mu \text{ et } (\pi_Y)_* \lambda = \nu.$$

Notons $\Lambda(\mu, \nu)$ l'ensemble des couplages de μ et ν .

Remarques. (1) Par exemple, la mesure produit $\mu \otimes \nu$ sur $X \times Y$ est un couplage (rarement optimal!), appelé le *couplage produit* de μ et ν (et en particulier, $\Lambda(\mu, \nu)$ est non vide). Si $X = Y$ et $\mu = \nu$, alors le *couplage diagonal*, image de μ par l'application $x \mapsto (x, x)$, est un couplage, de support la diagonale de $X \times X$.

(2) L'image directe des mesures par l'application de symétrie $(x, y) \mapsto (y, x)$ est une bijection de $\Lambda(\mu, \nu)$ sur $\Lambda(\nu, \mu)$.

(3) Rappelons le lemme suivant.

Lemme 3.1 *L'image directe des mesures par une application continue f de X dans Y est une application continue de $\text{Prob}(X)$ dans $\text{Prob}(Y)$ pour les topologies étroites.*

Preuve. Si $\varphi : Y \rightarrow \mathbb{R}$ est continue (bornée), alors $\varphi \circ f : X \rightarrow \mathbb{R}$ est continue (bornée).
□

Puisque les projections π_X et π_Y sont continues, $\Lambda(\mu, \nu)$ est donc un convexe fermé de $\text{Prob}(X \times X)$. Puisque $\text{Prob}(X \times X)$ est compact pour la topologie faible-étoile (par le théorème de Banach-Alaoglu), $\Lambda(\mu, \nu)$ est donc convexe compact.

(4) Soit $f : X \times Y \rightarrow [0, +\infty[$ une application continue. Un couplage λ_0 de μ et ν est dit *optimal pour f* si

$$\int_{X \times Y} f d\lambda_0 = \inf_{\lambda \in \Lambda(\mu, \nu)} \int_{X \times Y} f d\lambda .$$

Par continuité de l'évaluation des mesures sur $X \times Y$ en l'application continue bornée f pour la topologie étroite, et par compacité de $\Lambda(\mu, \nu)$ (voir la remarque (3) ci-dessus), il existe au moins un couplage optimal. Le résultat suivant est démontré par exemple dans [Vil, Theo. 5.20, page 77].

Lemme 3.2 *Pour tout $p \in [1, +\infty[$, si $(\mu_n)_{n \in \mathbb{N}}, (\nu_n)_{n \in \mathbb{N}}$ sont deux suites dans $\text{Prob}(X)$ convergeant pour la topologie étroite vers μ et ν respectivement, si λ_n est un couplage optimal de μ_n et ν_n pour l'application d_X^p , convergeant pour la topologie étroite vers $\lambda \in \text{Prob}(X \times X)$, alors λ est un couplage optimal entre μ et ν pour d_X^p . □*

(5) Pour $i = 1, 2, 3$, soit (X_i, μ_i) un espace métrique compact muni d'une mesure borélienne de probabilité. Nous avons une application $\Lambda(\mu_1, \mu_2) \times \Lambda(\mu_2, \mu_3) \rightarrow \Lambda(\mu_1, \mu_3)$ définie par

$$(\lambda_{12}, \lambda_{23}) \mapsto \lambda_{12} * \lambda_{23} ,$$

où $d\lambda_{12}(x, y) = d\lambda_{12|y}(x)d\mu_2(y)$ et $d\lambda_{23}(y, z) = d\lambda_{23|y}(z)d\mu_2(y)$ sont les désintégrations de λ_{12} et λ_{23} par les deux projections π_{X_2} , de sorte que $y \mapsto \lambda_{12|y}$ et $y \mapsto \lambda_{23|y}$ sont des applications mesurables de X_2 dans $\text{Prob}(X_1)$ et $\text{Prob}(X_3)$ respectivement, et

$$d(\lambda_{12} * \lambda_{23})(x, z) = \int_{y \in X_2} d\lambda_{12|y}(x)d\lambda_{23|y}(z) d\mu_2(y) .$$

i.e. pour tout $f \in \mathcal{C}(X \times Z, \mathbb{R})$, nous avons

$$\lambda_{12} * \lambda_{23}(f) = \int_{y \in X_2} \int_{(x, z) \in X_1 \times X_3} f(x, z) d(\lambda_{12|y} \otimes \lambda_{23|y})(x, z) d\mu_2(y) .$$

Notons que le support de $\lambda_{12} * \lambda_{23}$ est contenu dans l'ensemble des couples (x, z) tels qu'il existe un point y appartenant au support de μ_2 tel que (x, y) et (y, z) appartiennent au support de λ_{12} et λ_{23} respectivement.

Pour tout $p \in [1, +\infty]$, la *distance de Wasserstein* $d_{\text{Was}, p}$ d'ordre p sur $\text{Prob}(X)$ (introduite, au moins pour $p = 1$, par Kantorovich) est la borne inférieure, sur tous les couplages λ des deux mesures, de la norme de l'application continue d dans l'espace $\mathbb{L}^p(X \times X, \lambda)$. Si $p \neq +\infty$, alors

$$d_{\text{Was}, p}(\mu, \nu) = \inf_{\lambda \in \Lambda(\mu, \nu)} \left(\int_{X \times X} d(x, y)^p d\lambda(x, y) \right)^{1/p} .$$

Bien sûr, la distance de Wasserstein d'ordre p sur $\text{Prob}(X)$ dépend de la distance d sur X , et la nous noterons $d_{\text{Was},p,d}$ lorsque nous voulons préciser d .

Un couplage λ est dit *optimal* pour la distance de Wasserstein d'ordre p s'il est optimal pour l'application d_X^p , i.e. si

$$d_{\text{Was},p}(\mu, \nu) = \|d\|_{\mathbb{L}^p(\lambda)}.$$

Si $p \neq +\infty$, par la remarque (4) ci-dessus, il existe au moins un couplage optimal.

Proposition 3.3 *Les couples $(\text{Prob}(X), d_{\text{Was},p})$ et $(\text{Prob}(X), d_{\text{Pro}})$ sont des espaces métriques compacts, de topologie induite égale à la topologie étroite si $p \neq +\infty$.*

Preuve. • Montrons que d_{Pro} est une distance (voir par exemple [Bill]).

En prenant $\epsilon = 1$, l'application d_{Pro} est bien définie. Elle est clairement symétrique, et nulle sur la diagonale. Les mesures étant croissantes, et puisque $\mathcal{V}_\epsilon(\mathcal{V}_\eta(A)) \subset \mathcal{V}_{\epsilon+\eta}(A)$, la distance de Prokhorov vérifie l'inégalité triangulaire. Puisqu'une mesure de probabilité sur un espace métrique compact X est régulière, et qu'un fermé est l'ensemble des points à distance nulle de lui-même, pour tout fermé F de X et pour tout $\mu \in \text{Prob}(X)$, nous avons

$$\mu(F) = \lim_{\epsilon \rightarrow 0^+} \mu(\mathcal{V}_\epsilon(F)). \quad (-1-)$$

Puisque deux éléments de $\text{Prob}(X)$ qui coïncident sur les fermés coïncident, l'axiome de séparation de la distance de Prokhorov en découle.

• Montrons que $d_{\text{Was},p}$ est une distance.

La distance de Wasserstein d'ordre p est bien définie par la remarque (1) ci-dessus. Elle est symétrique par la remarque (2) ci-dessus (et la symétrie de la distance d). Si $p \neq \infty$, elle vérifie l'inégalité triangulaire par la remarque (5) ci-dessus (et l'inégalité triangulaire de d) : pour tous $\lambda_{12} \in \Lambda(X, X)$ et $\lambda_{23} \in \Lambda(X, X)$,

$$\begin{aligned} d_{\text{Was},p}(\mu_1, \mu_3) &\leq \left(\int_{X \times X} d^p(x, z) d(\lambda_{12} * \lambda_{23}) \right)^{1/p} \\ &\leq \left(\int_{X \times X \times X} (d(x, y) + d(y, z))^p d\lambda_{12|y}(x) d\lambda_{23|y}(z) d\mu_2(y) \right)^{1/p} \\ &\leq \|d\|_{\mathbb{L}^p(\lambda_{12})} + \|d\|_{\mathbb{L}^p(\lambda_{23})}, \end{aligned}$$

la dernière inégalité découlant de l'inégalité de Minkovsky, et du fait que $\lambda_{12|y}$ and $\lambda_{23|y}$ sont des mesures de probabilité ; en prenant la borne inférieure sur les couplages $\lambda_{12} \in \Lambda(X, X)$ et $\lambda_{23} \in \Lambda(X, X)$, le résultat en découle.

Si $p \neq \infty$, et si $d_{\text{Was},p}(\mu, \nu) = 0$, alors tout couplage optimal λ pour μ, ν est à support sur la diagonale, par la propriété de séparation de d . Les restrictions des deux projections π_1 et π_2 à la diagonale de $X \times X$ étant des homéomorphismes tels que $\pi_1 \circ \pi_2^{-1} = \text{id}_X$, on a

$$\mu = (\pi_1)_* \lambda = (\pi_1 \circ \pi_2^{-1})_* \nu = \nu,$$

ce qui montre l'axiome de séparation de $d_{\text{Was},p}$. L'inégalité triangulaire de $d_{\text{Was},\infty}$ découle du fait que, pour tout couplage λ , comme d est continue,

$$\|d\|_{\mathbb{L}^\infty(\lambda)} = \max_{(x,z) \in \text{Supp}(\lambda)} d(x, z),$$

ainsi que de la propriété du support de $\lambda_{12} * \lambda_{23}$ décrite à la fin de la remarque (5), et de l'inégalité triangulaire de d . Enfin, l'axiome de séparation de $d_{\text{Was},\infty}$ découle de la croissance en p des distances de Wassertein d'ordre p : pour tous $p, q \in [1, +\infty[$,

$$d_{\text{Was},p} \leq d_{\text{Was},q} \leq d_{\text{Was},\infty} .$$

• Montrons que la topologie induite par d_{Pro} est la topologie étroite. Soient $(\mu_i)_{i \in \mathbb{N}}$ une suite dans $\text{Prob}(X)$ et $\mu \in \text{Prob}(X)$.

Il est bien connu que $(\mu_i)_{i \in \mathbb{N}}$ converge étroitement vers μ si et seulement si pour tout borélien B tel que $\mu(\partial B) = 0$, la suite $(\mu_i(B))_{i \in \mathbb{N}}$ converge vers $\mu(B)$ (voir par exemple [LeG]).

Supposons tout d'abord que μ_i converge vers μ pour la distance de Prokhorov, et soit $B \in \mathcal{B}_X$ tel que $\mu(\partial B) = 0$. Soit $\eta > 0$. Par l'égalité (-1-), nous avons, pour tout $\epsilon \in]0, \eta]$ suffisamment petit,

$$\mu(\mathcal{V}_\epsilon(B)) \leq \mu(\overline{B}) + \eta = \mu(B \cup \partial B) + \eta = \mu(B) + \eta ,$$

et

$$\mu(\mathcal{V}_{2\epsilon}(\partial B)) \leq \mu(\partial B) + \eta = \eta .$$

Donc pour i suffisamment grand, $\mu_i(B) \leq \mu(\mathcal{V}_\epsilon(B)) + \epsilon \leq \mu(B) + 2\eta$.

Soit $\epsilon' \in]0, \eta]$ suffisamment petit tel que $\mathcal{V}_{\epsilon'}(B) \subset B \cup \mathcal{V}_\epsilon(\partial B)$. Un tel ϵ' existe. Sinon, il existerait une suite $(x_i)_{i \in \mathbb{N}}$ dans X vérifiant les trois propriétés : $d(x_i, B) \leq 1/i$, $x_i \notin B$ et $d(x_i, \partial B) \geq \epsilon$, pour tout $i \in \mathbb{N}$. Quitte à extraire, x_i convergerait donc vers un point x , qui est dans ∂B par les deux premières propriétés, ce qui contredirait la troisième.

Donc pour i suffisamment grand,

$$\mu(B) \leq \mu_i(\mathcal{V}_{\epsilon'}(B)) + \epsilon' \leq \mu_i(B) + \mu_i(\mathcal{V}_\epsilon(\partial B)) + \epsilon' \leq \mu_i(B) + \mu(\mathcal{V}_{2\epsilon}(\partial B)) + \epsilon + \epsilon' \leq \mu_i(B) + 3\eta ,$$

ce qui montre que $(\mu_i)_{i \in \mathbb{N}}$ converge étroitement vers μ .

Supposons réciproquement que $(\mu_i)_{i \in \mathbb{N}}$ converge étroitement vers μ . Supposons par l'absurde que la distance de Prokhorov entre μ_i et μ ne tende pas vers 0. Il existe donc $\epsilon > 0$, une sous-suite $(\mu_{i_k})_{k \in \mathbb{N}}$ et une suite $(B_k)_{k \in \mathbb{N}}$ dans \mathcal{B}_X tels que, pour tout $k \in \mathbb{N}$,

$$\mu(B_k) \geq \mu_{i_k}(\mathcal{V}_\epsilon(B_k)) + \epsilon \quad \text{ou} \quad \mu_{i_k}(B_k) \geq \mu(\mathcal{V}_\epsilon(B_k)) + \epsilon . \quad (-2-)$$

En particulier, B_k est non vide. Puisque X est compact, quitte à extraire, la suite de fermés non vides $(\overline{B_k})_{k \in \mathbb{N}}$ converge vers un fermé non vide B pour la distance de Hausdorff. Donc si k est assez grand, $B_k \subset \mathcal{V}_{\epsilon/2}(B)$.

Rappelons que, pour toute mesure borélienne de probabilité ν sur X , en dehors d'un ensemble dénombrable de $\epsilon > 0$, nous avons $\nu(\partial \mathcal{V}_\epsilon(B)) = 0$ (en effet, puisque $\partial \mathcal{V}_\epsilon(B)$ et $\partial \mathcal{V}_{\epsilon'}(B)$ sont disjoints si $\epsilon \neq \epsilon'$, pour tout $k \in \mathbb{N} - \{0\}$, il existe au plus k valeurs de ϵ telles que $\nu(\partial \mathcal{V}_\epsilon(B)) \geq \frac{1}{k}$).

Par ce rappel, quitte à diminuer ϵ , nous pouvons donc supposer que $\mu(\partial \mathcal{V}_{\epsilon/2}(B)) = 0$. Dans le premier cas de l'alternative (-2-) ci-dessus, (et puisque $\mathcal{V}_\eta(\mathcal{V}_{\eta'}(A)) \subset \mathcal{V}_{\eta+\eta'}(A)$), nous avons donc

$$\mu(\mathcal{V}_{\epsilon/2}(B)) \geq \mu(B_k) \geq \mu_{i_k}(\mathcal{V}_\epsilon(B_k)) + \epsilon \geq \mu_{i_k}(\mathcal{V}_{\epsilon/2}(B)) + \epsilon .$$

Par passage à la limite quand k tend vers $+\infty$, nous avons $\mu(\mathcal{V}_{\epsilon/2}(B)) \geq \mu(\mathcal{V}_{\epsilon/2}(B)) + \epsilon$, une contradiction. Dans le second cas de l'alternative (-2-), une contradiction est obtenue de manière similaire. Nous avons donc montré que la distance de Prokhorov induit la topologie étroite.

- Montrons que la topologie induite par $d_{\text{Was},p}$ est la topologie étroite, si $p \in [1, +\infty[$. Soient $(\mu_i)_{i \in \mathbb{N}}$ une suite dans $\text{Prob}(X)$ et $\mu \in \text{Prob}(X)$.

Supposons tout d'abord que μ_i converge vers μ pour la distance de Wasserstein d'ordre p . Soit μ' une valeur d'adhérence de $(\mu_i)_{i \in \mathbb{N}}$ pour la topologie étroite, limite de $(\mu_{i_k})_{k \in \mathbb{N}}$. Soit λ_i un couplage de μ_i et de μ tel que $\int_{X \times X} d^p d\lambda_i$ tende vers 0 quand i tend vers $+\infty$. Quitte à extraire, λ_{i_k} converge étroitement vers λ , qui, par le lemme 3.1, est un couplage de μ' et de μ . Puisque d^p est continue (bornée), nous avons $\int_{X \times X} d^p d\lambda = 0$. Donc la distance de Wasserstein d'ordre p entre μ' et μ est nulle, et $\mu' = \mu$. Dans l'espace compact $\text{Prob}(X)$, la suite $(\mu_i)_{i \in \mathbb{N}}$, qui admet μ comme seule valeur d'adhérence, converge donc vers μ pour la topologie étroite.

Réciproquement, supposons que μ_i converge vers μ pour la topologie étroite. Soit λ_i un couplage optimal entre μ_i et μ . Alors quitte à extraire, et par le lemme 3.2, λ_i converge vers un couplage optimal λ entre μ et μ . En particulier, $\int_{X \times X} d^p d\lambda = 0$. Puisque d^p est continue (bornée), nous avons donc $\int_{X \times X} d^p d\lambda_i$ tend vers 0 quand i tend vers $+\infty$. Donc μ_i converge vers μ pour la distance de Wasserstein d'ordre p . \square

Remarque. Nous renvoyons aux commentaires de [Vil, page 109] pour des références sur le fait que la topologie induite par $d_{\text{Was},+\infty}$ n'est pas toujours la topologie étroite.

Soit $\|\cdot\|$ une norme sur \mathbb{R}^2 (par exemple la norme $\|(x, y)\|_q = (|x|^q + |y|^q)^{1/q}$ pour $q \in [1, +\infty[$ ou $\|(x, y)\|_\infty = \max\{|x|, |y|\}$). Nous identifions X et Y avec leurs images dans $X \amalg Y$ par les inclusions canoniques, et toute mesure sur X ou sur Y avec sa mesure image dans $X \amalg Y$ par ces inclusions canoniques.

La *distance de Hausdorff-Gromov-Prokhorov* entre deux éléments X et Y de $\mathcal{M}\mathcal{M}_1$ est

$$d_{\text{HGP}}(X, Y) = \inf_{\delta \in \Delta(d_X, d_Y)} \left\| (d_{\text{Hau}, \delta}(X, Y), d_{\text{Pro}, \delta}(\mu_X, \mu_Y)) \right\|.$$

Cette distance dépend du choix de la norme, mais deux choix différents donnent deux distances équivalentes. De même, la *distance de Hausdorff-Gromov-Wasserstein* d'ordre p entre deux éléments X et Y de $\mathcal{M}\mathcal{M}_1$ est

$$d_{\text{HGW}, p}(X, Y) = \inf_{\delta \in \Delta(d_X, d_Y)} \left\| (d_{\text{Hau}, \delta}(X, Y), d_{\text{Was}, p, \delta}(\mu_X, \mu_Y)) \right\|.$$

Nous appellerons *distance de Hausdorff-Gromov mesurée* l'une de ces deux distances. Dans de nombreuses applications, la forme de la distance sur $\mathcal{M}\mathcal{M}_1$ n'a pas d'importance, car on utilise seulement la seconde assertion de la proposition suivante, qui donne une caractérisation de la convergence pour la distance de Hausdorff-Gromov mesurée, par une notion de convergence due à Fukaya [Fuk1].

Proposition 3.4 (1) Pour $p \in [1, +\infty[$, les couples $(\mathcal{M}\mathcal{M}_1, d_{\text{HGP}})$ et $(\mathcal{M}\mathcal{M}_1, d_{\text{HGW}, p})$ sont des espaces métriques complets séparables, induisant la même topologie.

(2) Une suite $(X_i, d_i, \mu_i)_{i \in \mathbb{N}}$ dans $\mathcal{M}\mathcal{M}_1$ converge vers (X, d, μ) pour une distance de Hausdorff-Gromov mesurée si et seulement si pour tout i dans \mathbb{N} , il existe $f_i : X_i \rightarrow X$

une ϵ_i -isométrie, telles que ϵ_i converge vers 0 et $(f_i)_*\mu_i$ converge vers μ pour la topologie étroite dans $\text{Prob}(X)$ quand i tend vers $+\infty$.

(3) L'application d'oubli de la mesure $\Psi : \mathcal{M}\mathcal{M}_1 \rightarrow \text{Dis}_1$ est lipschitzienne, et propre.

En particulier, une suite $(X_i, d_i, \mu_i)_{i \in \mathbb{N}}$ dans $\mathcal{M}\mathcal{M}_1$ est relativement compacte (on dit parfois *tendue*, et cela signifie que toute sous-suite admet une sous-suite convergente) si et seulement si la suite $(X_i, d_i)_{i \in \mathbb{N}}$ l'est dans Dis_1 , et par le critère de compacité de Gromov 2.3, c'est le cas si et seulement si la suite $(X_i, d_i)_{i \in \mathbb{N}}$ est uniformément compacte.

Preuve. (1) et (2) Nous renvoyons par exemple à [Vil, Part III, §27].

(3) L'application d'oubli de la mesure $\Psi : \mathcal{M}\mathcal{M}_1 \rightarrow \text{Dis}_1$ est l'application qui à (la classe d'isométrie préservant la mesure de) (X, d, μ) associe (la classe d'isométrie de) (X, d) . Elle est bien définie. Comme deux distances de Hausdorff-Gromov-Prokhorov sont équivalentes si l'on change la norme sur \mathbb{R}^2 , pour montrer que Ψ est lipschitzienne, il suffit de montrer que Ψ est 1-lipschitzienne si on utilise la norme $\|\cdot\|_q$ sur \mathbb{R}^2 , ce qui est immédiat. La propriété de Ψ (i.e. le fait que l'image réciproque de tout compact soit compact) découle de l'assertion (2) et du théorème de Prokhorov (ou de Banach-Alaoglu) de compacité de $\text{Prob}(X)$ pour X métrique compact. \square

Remarque. Soit $(X_i, d_i, \mu_i)_{i \in \mathbb{N}}$ une suite dans $\mathcal{M}\mathcal{M}_1$, et supposons, en considérant de nouveau l'ultrafiltre ω sur \mathbb{N} , qu'il existe $(X_\omega, d_\omega) \in \text{Dis}_1$ tel que $\lim_\omega (X_i, d_i) = (X_\omega, d_\omega)$ dans (Dis_1, d_{HG}) (ce qui est bien sûr le cas si $(X_i, d_i, \mu_i)_{i \in \mathbb{N}}$ converge vers (X_ω, d_ω) au sens usuel dans (Dis_1, d_{HG}) , et, par le théorème 2.6, si la suite est uniformément compacte). Par la proposition 2.2 (2), il existe alors une ϵ_i -isométrie $f_i : X_i \rightarrow X_\omega$ telle que $\lim_\omega \epsilon_i = 0$. Posons $\mu_\omega = \lim_\omega (f_i)_*\mu_i$, qui existe dans l'espace compact $\text{Prob}(X_\omega)$. Alors μ_ω ne dépend pas, modulo isométries de (X_ω, d_ω) , du choix des f_i par la remarque 2.5, et, par la proposition 3.4,

$$\lim_\omega (X_i, d_i, \mu_i) = (X_\omega, d_\omega, \mu_\omega)$$

dans $(\mathcal{M}\mathcal{M}_1, d_{HGP})$.

Nous concluons ce survol par un critère de précompacité pour les distances de Hausdorff-Gromov mesurées.

Soit $\psi :]0, +\infty[\rightarrow]0, +\infty[$ une application. Une mesure borélienne positive μ sur un espace métrique X est dite ψ -bien répartie si pour tout $\epsilon > 0$ et tout x dans X ,

$$\mu(B(x, \epsilon)) \geq \psi(\epsilon).$$

Le support d'une telle mesure est égal à tout X . Si X est compact et si μ est de support total, alors il existe une application ψ telle que μ soit ψ -bien répartie (il suffit de prendre $\psi(\epsilon) = \inf_{x \in X} \mu(B(x, \epsilon))$, un argument de compacité montrant qu'elle est strictement positive). Mais l'important sera d'obtenir des ψ uniformes. Le lemme suivant donne des exemples.

Soit $D > 0$, une mesure borélienne positive μ sur un espace métrique X est dite D -doublante si pour tout $\epsilon > 0$ et tout x dans X ,

$$\mu(B(x, 2\epsilon)) \leq D \mu(B(x, \epsilon)).$$

Voir par exemple [Hei, Vil] pour de très nombreuses références sur les mesures D -doublantes. Remarquons qu'être ψ -bien répartie pour une certaine fonction ψ est bien moins contraignant qu'être D -doublante pour un certain D . En particulier, sur tout espace métrique

compact sans point isolé, toute mesure de support total ayant un atome est ψ -bien répartie pour une certaine application ψ , mais n'est D -doublante pour aucun $D > 0$.

Lemme 3.5 *Pour tous $\alpha, \beta, D > 0$, il existe une application $\psi :]0, +\infty[\rightarrow]0, +\infty[$ telle que toute mesure D -doublante de masse au moins α sur un espace métrique de diamètre au plus β soit ψ -bien répartie.*

Preuve. Pour tous $\alpha, \beta, D > 0$, posons $\psi : \epsilon \mapsto \alpha D^{-E[\log_2 \frac{\beta}{\epsilon}] + 1}$, où $E[\cdot]$ est la partie entière.

Pour tous $\epsilon > 0$ et $x \in X$, si $n = E[\log_2 \frac{\beta}{\epsilon}] + 1$ nous avons $2^n \epsilon \geq \beta$, donc $B(x, 2^n \epsilon) = X$ et par récurrence

$$\alpha \leq \mu(B(x, 2^n \epsilon)) \leq D^n \mu(B(x, \epsilon)) ,$$

ce qui montre le résultat. □

Le résultat suivant donne un critère pour qu'une partie de $\mathcal{M}\mathcal{M}_1$ soit d'adhérence compacte : il suffit que les mesures soient uniformément bien réparties.

Théorème 3.6 *Soit $\psi :]0, +\infty[\rightarrow]0, +\infty[$ une application. L'ensemble des (X, d, μ) dans $\mathcal{M}\mathcal{M}_1$ tels que μ soit ψ -bien répartie est relativement compact dans $\mathcal{M}\mathcal{M}_1$ pour toute distance de Hausdorff-Gromov mesurée.*

En particulier (voir par exemple [Vil, Theo. 27.32]), par le lemme précédent et ce théorème, l'ensemble des éléments de $\mathcal{M}\mathcal{M}_1$, dont la mesure est D -doublante, est relativement compact dans $\mathcal{M}\mathcal{M}_1$, pour tout $D > 0$.

Preuve. Notons P cet ensemble. Par la proposition 3.4 (3) et le commentaire qui la suit, il suffit de montrer que pour tout $\epsilon > 0$, il existe $N(\epsilon) \in \mathbb{N}$ tel que pour tout (X, d, μ) dans P et tout x dans X , on puisse recouvrir X par N boules de rayon ϵ .

Soit $\epsilon > 0$, posons $N(\epsilon) = E[\frac{1}{\psi(\frac{\epsilon}{2})}] + 1$. Pour tout (X, d, μ) dans P , soit $k \in \mathbb{N}$ maximal tel qu'il existe x_1, \dots, x_k dans X tels que les boules ouvertes $B(x_i, \frac{\epsilon}{2})$ soient deux à deux disjointes. Alors les boules ouvertes $B(x_i, \epsilon)$ recouvrent, sinon pour x_{k+1} un point qui n'est pas dans leur réunion, les boules $B(x_i, \frac{\epsilon}{2})$ pour $1 \leq i \leq k+1$ seraient deux à deux disjointes, ce qui contredirait la maximalité de k . Mais alors puisque la mesure μ est de probabilité

$$1 \geq \mu\left(\bigcup_{i=1}^k B(x_i, \frac{\epsilon}{2})\right) = \sum_{i=1}^k \mu(B(x_i, \frac{\epsilon}{2})) \geq k \psi\left(\frac{\epsilon}{2}\right) ,$$

ce qui montre le résultat. □

Références

- [Bil] P. Billingsley, *Convergence of probability measures*, Wiley, 1968.
- [Bou] N. Bourbaki, *Topologie générale*, chap. 1 à 4, Hermann, Paris, 1971.
- [Bré] H. Brézis, *Analyse fonctionnelle*, Masson, 1983.
- [BH] M.R. Bridson, A. Haefliger, *Metric spaces with non-positive curvature*, Grund. math. Wiss. **319**, Springer Verlag (1998).
- [BBI] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Grad. Stud. Math. **33**, Amer. Math. Soc. (2001).

- [Dix] J. Dixmier, *Topologie générale*, PUF, 1981.
- [Doo] J. L. Doob, *Measure theory*, Grad. Text Math. **143**, Springer Verlag, 1994.
- [DW] L. van den Dries, A. Wilkie, *On Gromov's theorem concerning groups of polynomial growth and elementary logic*, J. Alg. **89** (1984), 349-374.
- [Dru] C. Drutu, *Quasi-isometry rigidity of groups*, Sémin. Congrès **18**, (L. Bessières, A. Parreau, B. Rémy eds), 323-374, Soc. Math. France, 2009.
- [DP] A. Dyubina, I. Polterovich, *Structures at infinity of hyperbolic spaces*, Uspekhi Mat. Nauk **53** (1998), no. 5(323), 239–240.
- [Fuk1] K. Fukaya, *Collapsing of Riemannian manifolds and eigenvalues of Laplace operator*, Invent. Math. **87** (1987) 517-547.
- [Fuk2] K. Fukaya, *Hausdorff convergence of Riemannian manifolds and applications*, in “Recent Topics in Differential and Analytic Geometry”, pp. 143-238, T. Ochiai ed., Adv. Stud. Pure Math. **18** (1990).
- [Gro1] M. Gromov, *Groups of polynomial growth and expanding maps*, Pub. I.H.E.S. **53** (1981), 53-78.
- [Gro2] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Prog. Math. **152**, Birkhäuser, 1999.
- [Gro3] M. Gromov, *Asymptotics invariants of infinite groups*, Vol. 2 de “Geometric group theory”, A. Niblo, M. Roller eds, LMS LNS **182**, Cambridge Univ. Press, 1993.
- [Hei] J. Heinonen, *Lectures on analysis on metric spaces*, Universitext, Springer Verlag, 2001.
- [Jac] K. Jacob, *Convergence d'espaces métriques mesurés*, mémoire de maîtrise, ENS Ulm, Juin 2006. <http://www.fimfa.ens.fr/spip.php?article14>
- [LeG] J.-F. Le Gall, *Intégration, probabilités et processus aléatoires*, notes de cours, ENS Ulm, Sep. 2006. <http://www.fimfa.ens.fr/>
- [MO] J. Mayer, J. Nikiel, L. Oversteegen, *Universal spaces for \mathbb{R} -trees*, Trans. Amer. Math. Soc. **334** (1992) 411-432.
- [Par] K. R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, 1967.
- [Paul1] F. Paulin, *Topologie de Gromov équivariante, structures hyperboliques et arbres réels*, Invent. Math. **94** (1988), 53-80.
- [Paul2] F. Paulin, *Sur les automorphismes extérieurs des groupes hyperboliques*, Ann. Ecole Norm. Sup. **30** (1997) 147-167.
- [Stu] K.-T. Sturm, *On the geometry of metric measure spaces I, II*, Acta Math. **196** (2006) 65-131 et 133-177.
- [Vil] C. Villani, *Optimal transport, old and new*, Grund. math. Wissen. **338**, Springer Verlag, 2009.

Département de Mathématique et Applications, UMR 8553 CNRS
 École Normale Supérieure, 45 rue d'Ulm
 75230 PARIS Cedex 05, FRANCE
 e-mail : Frederic.Paulin@ens.fr