

Transport optimal en théorie des processus et en géométrie

Nicolas JUILLET
Université de Strasbourg

Habilitation à diriger des recherches

*préparée à l'Institut de Recherche Mathématique Avancée
UMR 7501 – CNRS et Université de Strasbourg*

Soutenue le 7 décembre 2018 devant la commission d'examen

Guillaume CARLIER
Jacques FRANCHI
Nicola GIGLI
Christian LEONARD
Vlada LIMIC
Ludovic RIFFORD
Anton THALMAIER

Université Paris-Dauphine, examinateur
Université de Strasbourg, garant
SISSA Trieste, rapporteur
Université Paris Nanterre, rapporteur
CNRS et Université de Strasbourg, examinatrice
Université Nice Sophia Antipolis, examinateur
Université du Luxembourg, rapporteur

Remerciements

Lorsque j'ai créé cette page de remerciements au printemps dernier, je l'ai tout d'abord tapissée d'un pseudo-texte en pseudo latin. Ce fut un des rares passages où ma maman, que je remercie, ne trouva pas de faute d'orthographe : « *Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum* ». Je me suis dit que ça ferait l'affaire en attendant les propos sensés :

Je remercie vivement Nicola Gigli, Christian Léonard et Anton Thalmaier d'avoir accepté d'être mes rapporteurs et d'avoir consacré le temps nécessaire à la lecture et la compréhension de mon mémoire. J'ai été très heureux qu'ils acceptent de se pencher sur mon travail. Je remercie également les autres membres de mon jury, en premier lieu Jacques Franchi qui fut un garant d'une grande efficacité, rassurante. Guillaume Carlier, Vlada Limic et Ludovic Rifford m'ont fait le plaisir de prendre part à ce qui est et restera une étape importante de ma carrière. Merci !

Au moment des remerciements et de l'introspection que nécessite le retour sur une période de 10 ans, je pense naturellement à mes anciens directeurs de thèse Hervé Pajot et Karl Theodor Sturm. Chacun à sa manière m'a transmis des savoirs et des méthodes essentiels dans mon exercice ainsi que son point de vue sur le métier. Également, j'ai énormément appris auprès de mes coauteurs. Je les remercie de m'avoir ouvert à leurs mondes mathématiques ; ce fut à chaque fois un nouvel enrichissement. D'autres personnes, des professeurs notamment M. Tosel, ont joué un rôle déterminant dans mon apprentissage. Merci !

Je souhaite ensuite remercier les membres de l'Institut de Recherche Mathématique Avancée et de l'UFR Mathématiques-Informatique pour certains moments sympathiques vécus au cours des dernières années, comme il y en eut souvent au salon de l'IRMA ou dans les bureaux à discuter de nos étudiants, de nos enseignements et de mathématiques, ou lorsque nous nous croisions dans les couloirs ou l'ascenseur avec juste le temps de prendre des nouvelles, lors des repas au RIA ou au restaurant, ou encore au service informatique, à la bibliothèque, à l'IREM ou dans les secrétariats. Nos lieux, notre petit monde. Ils se transforment – nous sommes cernés de travaux ! – mais les meilleurs souvenirs restent. J'ai particulièrement apprécié l'accueil bienveillant des membres de l'équipe « Probabilités ». C'est avec un plaisir intact que je me rends au séminaire pour y partager des mathématiques variées. Les discussions entre nous, mathématiques ou autres, ont toujours été intéressantes et agréables et le restent alors même qu'elles changent au gré des arrivées. Pour finir, je ne voudrais surtout pas oublier une personne très importante pour moi, Emmanuel Opshtein, avec qui je partage le bureau I-606 depuis mon arrivée. Une telle cohabitation en tête à tête n'aurait pas eu la même saveur s'il s'était agi d'une autre personne. Nous avons trouvé, je trouve, le bon accord entre étude, confidences et détente. Merci Emmanuel !

Enfin, la vie professionnelle serait bien fade sans La Vie et il se trouve que ma petite famille, Anne et les stucks, y occupe une place primordiale. Je souhaite donc conclure en remerciant mon épouse qui supporte patiemment mon inattention et autres imperfections lorsque mes pensées me retiennent malgré moi au travail.

Table des matières

Remerciements	5
Table des matières	8
Table des figures	9
Résumé et liste des travaux présentés	11
Introduction	15
1 Transports déterminés par leurs marges	19
1.1 Le problème du transport optimal	19
1.2 Le transport quantile	21
1.3 Le transport entre graphes	27
1.4 Le transport montagne	32
1.5 Le problème martingale et le transport rideau	36
2 Transports ombrés	43
2.1 Ordres stochastiques	43
2.2 Ombre d'une mesure	44
2.3 Définition des couplages ombrés	51
3 Autour du théorème de Kellerer	57
3.1 Mesures martingale et mesures markoviennes	57
3.2 Améliorations dans le théorème de Kellerer	60
3.3 Théorème de Kellerer en dimensions ≥ 2	64
3.4 Martingales sur un ordre partiel et peacocks	66
3.5 Processus rideau gauche limites	67
3.6 La mesure Markov-quantile $\mathfrak{M}\Omega$	70
4 L'équation de continuité	77
4.1 Chaleur dans le groupe de Heisenberg	78
4.2 Énergie finie et équation de continuité	81
4.3 Processus $\mathfrak{M}\Omega$ et équation de continuité	83
4.4 Considérations sur les inégalités de Kuwada–Li	86
4.5 Déformation de Gigli et Mantegazza	89
Bibliographie	91

Table des figures

1.1	Le transport quantile Ω	25
1.2	Transport entre deux graphes lipschitziens	27
1.3	Le transport montagne	35
1.4	Trois routes de transport formant la configuration interdite.	38
2.1	Ombre de la somme $\mu = \gamma_1 + \gamma_2$ sur ν	47
2.2	Triangle de désintégration	52
2.3	Projection de Kellerer	53
2.4	Transports rideau et soleil entre mesures uniformes	55
3.1	La martingale associée à $(\mu_t)_{t \in (0,1,2)}$ au théorème 3.3.1 (i)	65
3.2	Noyaux de transitions de $(X_t)_{t \in \{0,1,2\}}$ du théorème 3.3.1.(ii)	65
3.3	Les martingales du théorème 3.3.1(iii).	65
3.4	Composition de transports rideau entre mesures uniformes	69
3.5	Processus rideau limite non-markovien	69
3.6	Composition de noyaux ℓ_r	73
3.7	Évolution d'un Dirac sous l'action de noyaux ℓ_r	74
3.8	Évolution d'une mesure $\frac{1}{x} \mathbf{1}_{]0,x[}$ sous l'action de noyaux ℓ_r	75
4.1	Les processus Ω , $\mathfrak{M}\Omega$ et le processus markovien Γ de l'exemple 4.3.6	86

Résumé et liste des travaux présentés

Je donne ici la liste des articles présentés dans ce mémoire ainsi qu'un résumé de chaque chapitre.

Les appels bibliographiques sont ceux utilisés dans le reste du mémoire où ils se distinguent par une écriture en caractères gras. Les marques '16+' et '18++' signalent que la date mentionnée (ici 2016) est la date de dépôt, pour les prépublications, respectivement que l'article est en cours d'écriture.

Chapitre 1

Les éléments classiques d'un problème de transport sont mis en place, faisant apparaître l'équivalence pour une mesure π de marges μ et ν données entre trois types de propriétés : le transport est optimal, le transport est monotone, le transport est défini par une formule ou construit par une méthode identifiée. On étudie trois variantes du problème en dimension 1, donnant lieu à des solutions respectant les propriétés citées, sous une forme adaptée : le transport quantile et sa version « graphe » [Jui11], le transport montagne (projet en cours dans [Jui18++]) et le transport rideau gauche \mathfrak{RG} [BeiJ16]. La spécificité de ces problèmes de dimension 1 est que l'existence et l'unicité de solutions se produisent sans hypothèses supplémentaires sur les marges μ et ν du plan de transport. Notre travail [Jui11] alimente aussi ce qui est connu en dimension supérieure au sujet des mesures $\mu \in \mathcal{P}(\mathbb{R}^d)$ dites régulières, c'est-à-dire les mesures auxquelles s'appliquent les conclusions du théorème d'existence et unicité de Brenier (contrairement à la dimension 1, cette existence et unicité ne s'appliquent pas à toutes les mesures).

- [BeiJ16] M. Beiglböck et N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1) :42–106, 2016. 11, 15, 19, 22, 37, 38, 39, 43, 45, 46, 55
- [Jui18++] N. Juillet. Optimal transport on the line for power costs close to but smaller than one. *Article en préparation*, 2018. 11, 13, 19, 32, 34
- [Jui11] N. Juillet. On displacement interpolation of measures involved in Brenier's theorem. *Proc. Amer. Math. Soc.*, 139(10) :3623–3632, 2011. 11, 19, 27, 28, 29, 30, 31
- [Jui16b] N. Juillet. Stability of the shadow projection and the left-curtain coupling. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4) :1823–1843, 2016. 11, 19, 40, 43, 44, 49, 50, 51

Chapitre 2

Dans ce chapitre on complète l'étude du couplage rideau \mathfrak{RG} en centrant la présentation sur la *projection d'ombres*, une opération qui intervient dans la définition de \mathfrak{RG} . On met en évidence les propriétés de stabilité métrique [Jui16b] de cette projection, ainsi que la possibilité de définir de nouveaux plans de transport martingale, notamment le couplage soleil \mathfrak{Sol} et le couplage rideau centré \mathfrak{RC} . On présente aussi les liens entre cette famille de transports dits *ombrés* et le problème de Skorhokhod d'une part, et le problème de transport généralisé d'autre part [BeiJ16+].

- [BeiJ16+] M. Beiglböck et N. Juillet. Shadow couplings. *ArXiv e-prints*, septembre 2016. 11, 16, 43, 54
- [BeiJ16] Voir plus haut, chapitre 1, 11, 15, 19, 22, 37, 38, 39, 43, 45, 46, 55
- [Jui16b] Voir plus haut, chapitre 1, 11, 19, 40, 43, 44, 49, 50, 51

Chapitre 3

Ce chapitre présente le théorème de Kellerer (1972). La démonstration en est revue et enrichie en faisant judicieusement figurer le couplage soleil $\mathfrak{S}\mathfrak{ol}$. On présente ensuite certaines solutions au problème des peacocks, dont seule l'existence est garantie par le théorème de Kellerer. Il est ainsi vu comment, d'après [Jui18], le couplage rideau $\mathfrak{R}\mathfrak{C}$ permet, dans des cas particuliers, de fournir des solutions *explicites* à ce problème. On détaille les difficultés rencontrées lorsqu'on souhaite généraliser le théorème de Kellerer aux dimensions supérieures [Jui16a]. Finalement le théorème de Kellerer est généralisé de l'ordre convexe à l'ordre stochastique [BouJ18⁺]. C'est une possibilité offerte par le processus Markov-quantile $\mathfrak{M}\mathfrak{Q}$, introduit à la fin du chapitre. Une démonstration indépendante est aussi proposée.

- [BouJ18⁺] C. Boubel et N. Juillet. The Markov-quantile process attached to a family of Marginals. *ArXiv e-prints*, avril 2018. 12, 16, 57, 61, 62, 63, 70, 71, 73, 74, 76, 82, 84, 85, 86
- [Jui16a] N. Juillet. Peacocks parametrised by a partially ordered set. In *Séminaire de Probabilités XLVIII*, volume 2168 of *Lecture Notes in Math.*, pages 13–32. Springer, Cham, 2016. 12, 57, 64, 66, 67
- [Jui18] N. Juillet. Martingales associated to peacocks using the curtain coupling. *Electron. J. Probab.*, 23 :Paper No. 8, 29, 2018. 12, 57, 68, 69
- [BeiJ16⁺] *Voir plus haut, chapitre 2*, 11, 16, 43, 54

Chapitre 4

L'équation de continuité peut être vue comme un problème de transport infinitésimal. Dans ce chapitre, il est expliqué que le processus $\mathfrak{M}\mathfrak{Q}$, lorsque $(\mu_t)_{t \in \mathbb{R}}$ est une courbe absolument continue (d'ordre 2), est une représentation lagrangienne de l'équation de continuité qui possède, de plus, la propriété de Markov. Dans un sens à préciser, c'en est aussi l'*unique* solution markovienne. On rend compte de situations rencontrées dans [ErbJ18] où il convient d'étudier l'équation de continuité pour des mesures gaussiennes (on désigne ainsi des mesures obtenues comme la distribution d'un mouvement brownien partant d'un point à un instant donné). Cette étude est justifiée par celle du flot de Ricci par Gigli et Mantegazza conduite à l'aide du transport optimal et de l'équation de la chaleur. On expose ainsi des résultats relatifs au groupe de Heisenberg sous-riemannien. D'autres résultats sont aussi présentés [BonJ18⁺] en relation avec la diffusion de la chaleur dans le groupe de Heisenberg et le couplage des mouvements browniens apparentés.

- [BonJ18⁺] M. Bonnefont et N. Juillet. Couplings in L^p distance of two Brownian motions and their Lévy area. *ArXiv e-prints*, janvier 2018 12, 78, 80, 86, 87, 88
- [ErbJ18] M. Erbar et N. Juillet. Smoothing and non-smoothing via a flow tangent to the Ricci flow. *J. Math. Pures Appl. (9)*, 110 :123–154, 2018. 12, 77, 78, 90, 91
- [BouJ18⁺] *Voir plus haut, chapitre 3*, 12, 16, 57, 61, 62, 63, 70, 71, 73, 74, 76, 82, 84, 85, 86

Autres travaux

J'indique ici la liste des travaux *non présentés* dans le mémoire. Hormis [Jui18⁺⁺], ils y sont au plus évoqués rapidement.

Travaux récents

- [BérJ18⁺] J. Bérard et N. Juillet. Reconciliation of probability measures. *À paraître dans European J. Appl. Math., numéro spécial “Applied Optimal Transport”, février 2018.*
- [GozJ18⁺] N. Gozlan et N. Juillet. On a mixture of Brenier and Strassen theorems. *ArXiv e-prints*, août 2018.
- [Jui18⁺] N. Juillet. SubRiemannian structures do not satisfy Riemannian Brunn–Minkowski inequalities. *Prépublication HAL*, octobre 2018.
- [JSig17] N. Juillet et M. Sigalotti. Pliability, or the Whitney extension theorem for curves in Carnot groups. *Anal. PDE*, 10(7) :1637–1661, 2017. 80

Travaux en cours

- [BHJ18⁺⁺] M. Brückhoff-Plückelmann, M. Huesmann, et N. Juillet. A left-monotone solution to the peacock problem. *Article en préparation*, 2018. 67
- [Jui18⁺⁺] *Voir plus haut, chapitre 1*, 11, 13, 19, 32, 34

Travaux initiés durant la thèse (soutenue en 2008)

L'article [Jui14], qui a connu une période de jachère, a été grandement amélioré depuis sa première version.

- [FigJ08] A. Figalli et N. Juillet. Absolute continuity of Wasserstein geodesics in the Heisenberg group. *J. Funct. Anal.*, 255(1) :133–141, 2008. 31
- [Jui07] N. Juillet. Transport de mesure et courbure de Ricci synthétique dans le groupe de Heisenberg. *Actes du Séminaire de Théorie Spectrale et Géométrie de l'Institut Fourier*, 25 :85–104, 2006–2007.
- [Jui09] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *International Mathematical Research Notices*, 2009, 2009. 87
- [Jui10a] N. Juillet. A counterexample for the geometric traveling salesman problem in the Heisenberg group. *Rev. Mat. Iberoam.*, 26 (3) :1035–1056, 2010.
- [Jui10b] N. Juillet. On a method to disprove generalized Brunn-Minkowski inequalities. In *Probabilistic approach to geometry*, volume 57 of *Adv. Stud. Pure Math.*, pages 189–198. Math. Soc. Japan, Tokyo, 2010.
- [Jui14] N. Juillet. Diffusion by optimal transport in Heisenberg groups. *Calc. Var. Partial Differential Equations*, 50(3-4) :693–721, 2014. 13, 79, 83

Introduction

Dans ce mémoire je souhaite rendre compte de certains résultats que j'ai obtenus, seul ou en collaboration, ces dernières années. Généralement il s'agit de résultats faisant intervenir des probabilités en transport optimal ou, peut-être plus justement, de résultats sur des problèmes partiellement nouveaux apportés aux probabilités par le transport optimal.

Plutôt qu'un véritable sous-domaine des mathématiques, le transport optimal semble davantage être une grille de lecture, un point de vue, un langage mais aussi une famille de questions mathématiques et un outil intervenant dans les grands domaines que sont la géométrie, l'analyse et les probabilités. Généralement la pluralité des champs qu'investit le transport optimal, ou plus concrètement la facilité avec laquelle des mathématiciens d'horizons divers s'approprient les techniques du transport optimal à des fins disparates, surprend les spécialistes eux-mêmes. On peut le remarquer dès les origines de la théorie avec le problème de Monge (« des déblais et des remblais ») [Mon81] qui, peu après avoir été formulé en 1776 a généré des développements inattendus en géométrie différentielle (sur les congruences de droite et la géométrie des surfaces). Dans les années 1940, le problème du transport était intimement lié à l'émergence de la programmation linéaire et de la recherche opérationnelle notamment sous l'impulsion de Kantorovich (qui transforma au passage le problème de Monge en le relaxant) [Kan42]. Avec la mécanique des fluides le transport a ensuite trouvé un excellent terrain d'application. De leur rencontre on retient notamment le célèbre théorème de décomposition polaire de Brenier [Bre91]. Une nouvelle percée spectaculaire s'est faite au cours des années 2000 ouvrant la voie vers des sujets analytico-géométriques tels que la courbure de Ricci ou les équations de diffusion dans les espace métriques. Elle prolonge la vague des inégalités fonctionnelles, de la courbure-dimension de Bakry et Émery et de la concentration de la mesure dans laquelle le transport opérait déjà (voir [BGL14] ou bien le survol [GL10]). Des travaux plus récents concernent quant à eux le pendant numérique du transport avec des applications notamment au traitement d'image, aux équations d'évolution, à l'économie ou à l'analyse statistique. En conséquence le domaine jouit d'une grande reconnaissance. Il est par exemple fait mention du transport optimal dans la citation de la médaille Fields décernée à Alessio Figalli¹ cet été, et de façon plus diffuse dans celle décernée à Cédric Villani en 2010. Pour une présentation complète des théories et applications du transport optimal, nous invitons le lecteur à consulter les ouvrages et survols de référence ([RR98a, RR98b, Vil03, AGS08, Vil09, San15, PC18⁺] et [MG13, AG08, BK12]).

À l'orée de la période moderne, au cours la période 1970–2000, des chercheurs, comme Kellerer, Rachev et Rüschendorf, ont utilisé le transport optimal en théorie des processus, statistique théorique et probabilités appliquées notamment pour quantifier les fluctuations de phénomènes aléatoires ou en relation avec les ordres stochastiques. Ce type de développements semble être ensuite passé au second plan. À la fois à rebours et dans le prolongement de l'évolution actuelle j'ai contribué à fonder le « transport martingale »² pour lequel on ajoute au problème de Monge–Kantorovich la contrainte de ne considérer que des plans de transport qui sont des martingales. Ce nouveau sous-domaine des probabilités est né au début des années 2010 lorsque des membres de la communauté des mathématiques financières ont trouvé de l'intérêt à investir la théorie du transport. Cette connexion est établie dans [BHLP13] et [GHLT14] en relation avec le *robust* ou *model-independent hedging* de Hobson et de ses coauteurs [HN12, HK15]. Le principe en est rappelé en quelques pages non-spécialisées dans la section 1.7.5 du livre de Santambrogio [San15]³. L'article [BeiJ16] est vraisemblablement le troisième sur le sujet ; il a reçu un très bon accueil. Nous y proposons une étude classique et approfondie, entièrement théorique et émancipée

1. pour son apport à l'analyse des ÉDP mais aussi à la géométrie métrique et aux probabilités

2. parfois abrégé MOT : Martingale Optimal Transport

3. Essayons en quelques lignes non-spécialisées ! La valeur d'un titre évolue de façon aléatoire en suivant une loi de martingale. On ne connaît pas cette loi, cependant on sait que c'est une martingale en vertu du *Fundamental Theorem of Asset Pricing* qui stipule que la dynamique des prix dans un modèle sans opportunité d'arbitrage doit en être une. De plus, on suppose que la connaissance du marché donne accès aux marges de la martingale à différents instants. Le juste prix d'une option vendue à l'instant initial est la valeur espérée du gain qu'elle engendrera. Or cette espérance dépend de la loi de martingale. On peut au mieux obtenir un encadrement de ce prix (le *pricing*) et ce sont précisément deux problèmes de transport martingale (multi-marges) que de déterminer

de la modélisation. On s’y concentre sur le problème pour une période, c’est-à-dire pour deux marges, avec des coûts de transport généraux, certaines familles de coûts étant cependant étudiées plus attentivement. Le sous-sujet *transport martingale* compte désormais plusieurs dizaines d’articles, certains de mathématiques financières, d’autres entièrement généralistes. La tendance actuelle est, assez naturellement, à la généralisation aux dimensions supérieures des résultats connus en dimension 1 pour deux marges, qu’il s’agisse de transport martingale multi-marges ou de transport martingale à deux marges pour des loi multi-dimensionnelles.

Je complète à présent le préambule *Résumé et travaux présentés* par quelques mots sur l’organisation de ce mémoire. Le chapitre 1 est un peu plus détaillé et pédagogique que les autres. J’introduis les concepts du transport autour de l’exemple le plus simple imaginable : le problème de transport quadratique sur la droite réelle. J’insiste sur le fait peut-être parfois méconnu qu’on peut obtenir des théorèmes de monotonie cyclique (appelés ici « lemmes d’échange ») de façon directe sans faire appel à la dualité. Je décline cette méthode pour traiter trois variantes du problème de transport à chaque fois en dimension 1. Plus que de simples illustrations il s’agit d’études inspirées par la littérature : une question de Gigli autour de la réciproque du théorème de Brenier, un couplage de mesures aléatoires issu d’une question pour des temps aléatoires de mouvements browniens bi-infinis et finalement le problème de transport martingale. L’enchaînement des chapitres 1–3 suit une progression logique. Je pars au chapitre 1 de problèmes de transport généraux de dimension 1, considérant donc des mesures de couplages sur \mathbb{R}^2 pour, au chapitre 2, me concentrer sur les mesures martingales (toujours sur \mathbb{R}^2) seulement et leur construction par la méthode des ombres [BeiJ16⁺]. Au chapitre 3 ce sont enfin les lois de martingale (des mesures martingale sur $\mathbb{R}^{\mathcal{T}}$ où le cardinal de \mathcal{T} est supérieur à trois) qui sont au centre de l’attention. Ce troisième chapitre est lié au problème de Processus Croissants pour l’Ordre Convexe (les *PCOC* ou *peacocks*) de Hirsch et Yor [HY10]. Étant donnée une famille de mesures $(\mu_t)_{t \in \mathbb{R}}$, il s’agit de construire une martingale markovienne $(X_t)_{t \in \mathbb{R}}$ telle que $X_t \sim \mu_t$ pour tout $t \in \mathbb{R}$. Pour cela, $(\mu_t)_{t \in \mathbb{R}}$ doit nécessairement être un peacock. Un célèbre théorème de Kellerer stipule que, pour des mesures réelles, cette condition est aussi suffisante lorsque $(\mu_t)_{t \in \mathbb{R}}$ est un peacock. J’aurais aimé construire une telle martingale markovienne qui soit complètement comprise et construite par une procédure identifiée, la même pour tout jeu de données. Il faut savoir qu’une telle prouesse a déjà été réalisée par Lowther [Low09, HRY14] sur la base des « diffusions presque continues » obtenues par des techniques de convolution et du calcul stochastique classique. On souhaiterait faire apparaître une autre solution qui serait, elle, fondée sur les techniques et constructions du transport martingale, dans la continuité du chapitre 2. Dans une direction un peu différente le summum serait de généraliser le théorème de Kellerer aux dimensions supérieures. Nous n’atteindrons malheureusement aucun de ces deux objectifs dans ce mémoire. Toutefois le processus Markov-quantile \mathfrak{MQ} construit dans [BouJ18⁺] représente pour un problème très proche (impliquant l’ordre de domination stochastique plutôt que l’ordre convexe) une solution très satisfaisante offrant ainsi un aboutissement aux efforts engagés dans la recherche de martingales particulières, telle que je viens de l’évoquer. La présentation de \mathfrak{MQ} couronne donc le chapitre 3. Le chapitre 4 est un peu différent des précédents puisqu’on sort des processus sur \mathbb{R} pour investir des questions plus géométriques dans le groupe de Heisenberg davantage dans la lignée des travaux de ma thèse⁴. L’équation de continuité fait le lien entre les chapitres 3 et 4. En effet d’une part \mathfrak{MQ} en est une solution intégrale distinguée, d’autre part cette équation entre également en jeu dans notre contribution à l’approche du flot de Ricci par Gigli et Mantegazza. Cette dernière se faisant aussi par l’entremise de la diffusion de la chaleur, c’est là une occasion naturelle de présenter certaines de mes tentatives pour coupler

les bornes de l’intervalle optimal, en considérant l’espérance pour toutes les martingales dont les marges coïncident avec les données. Notons finalement que le problème de transport dual trouve ici une interprétation éclairante du point de vue de la couverture du titre. Le prix maximal de l’option est couvert par les combinaisons des prix de certaines options (dites *européennes*) liées au titre que l’on suppose être toutes disponibles à l’achat et la vente. La borne supérieure du prix est atteinte en minimisant sur ces combinaisons : le prix maximal de l’option correspond au coût minimal de sur-réplication et, de même, le prix minimal correspond au coût maximal de sous-réplication.

4. *Optimal transport and geometric analysis in the Heisenberg groups*, Universités de Bonn et Grenoble 1, 2008

deux mouvements browniens du groupe de Heisenberg. C'est à ce sujet que dans les dernières pages, nous retrouverons un problème de dimension 1 du type de ceux présentés au cours du premier chapitre.

Exemples de transports optimaux inconditionnellement déterminés par leurs marges et ensemble des mesures régulières au sens du théorème de Brenier

Les résultats de ce chapitre sont parus ou à paraître dans les articles suivants : [Jui11], [Jui18⁺⁺], [BeiJ16], [Jui16b]. Ils s'insèrent à la racine de la théorie puisqu'il y est question d'existence et d'unicité d'un transport optimal. Contrairement à ce qui est attendu lorsqu'on connaît le théorème de Brenier, le grand résultat de ce type, nous obtiendrons le diptyque existence-unicité sans aucune hypothèse sur les mesures déblai μ ou remblai ν à transporter l'une sur l'autre (l'hypothèse la plus répandue pour obtenir le théorème de Brenier est d'imposer à μ d'être absolument continue). C'est l'axe directeur et l'originalité de ce chapitre. Les dites mesures seront toutes deux arbitraires et pourront en particulier présenter une partie atomique. Ce « miracle », tout apparent, possède une explication simple : il tient au fait que les hypothèses arrangeantes sont transférées de celles faites sur les mesures à celle concernant l'espace où se joue le problème de transport qui sera, dans les trois cas que nous envisagerons, soit la droite réelle soit la réunion de deux courbes lipschitziennes. Nous verrons ainsi apparaître comme solutions à nos problèmes de transport *le couplage quantile, le couplage montagne et le couplage rideau*.

Paradoxalement c'est en partant de la dimension 1 qui nous seront renseignés, en section 1.3, sur la structure de l'ensemble des mesures μ régulières, qui sont les mesures de probabilité de \mathbb{R}^d (avec $d \geq 2$) satisfaisant précisément aux hypothèses nécessaires et suffisantes à ce que le théorème de Brenier s'applique de la même manière qu'il le fait pour les mesures μ absolument continues. La structure géométrique de l'ensemble des mesures régulières dans l'espace de Wasserstein sera étudiée avec en filigrane et comme point de comparaison un thème géométrique, celui des points réguliers des espaces d'Alexandrov à courbure sectionnelle positive.

Tandis qu'il ne sera question du couplage montagne (section 1.4) que dans ce chapitre, les chapitres à venir exprimeront plus en détails les possibilités offertes par le couplage rideau (introduit en section 1.5).

1.1 Rappels et définitions sur le problème de transport

Nous noterons $\mathcal{P}(\mathcal{X})$ l'espace des mesures de probabilité sur un espace polonais¹ \mathcal{X} muni de sa tribu borélienne et par $\mathcal{M}_+(\mathcal{X})$ l'espace des mesures positives finies. L'espace $\mathcal{M}(\mathcal{X})$ sera celui des mesures de signées et $\mathcal{M}_0(\mathcal{X})$ le sous-espace des mesures de masse zéro. La partie positive d'une mesure $\mu \in \mathcal{M}(\mathcal{X})$ sera μ_+ , sa partie négative $\mu_- \in \mathcal{M}_+(\mathcal{X})$. On notera $|\mu|$ la mesure $\mu_+ + \mu_-$ et $\text{Masse}(\mu) = |\mu|(\mathcal{X})$ sera appelée la masse de μ . La comparaison $\mu_0 \preceq_+ \mu_1$ signifiera $\mu_1 - \mu_0 \in \mathcal{M}_+(\mathcal{X})$.

Pour $p \in [1, +\infty[$, $\mathcal{P}_p(\mathcal{X})$ sera le sous-espace des mesures de probabilité de p^{e} moment fini, classiquement nommé espace de Wasserstein d'ordre p . Pour des mesures positives de masse autre que 1 on notera $\mathcal{M}(\mathcal{X})_{+,p}$ l'espace correspondant.

Un *plan de transport*, ou simplement *transport* entre des mesures μ et ν prises dans $\mathcal{M}_+(\mathcal{X})$ est une mesure π de $\mathcal{M}_+(\mathcal{X} \times \mathcal{X})$ dont la première marge est μ et la seconde est ν . Ces affirmations se résument aux équations $\text{proj}_{\#}^1 \pi = \mu$ et $\text{proj}_{\#}^2 \pi = \nu$ où proj^i désigne, pour $i = 1, 2$, la projection définie par $\text{proj}^i : (x_1, x_2) \in \mathcal{X}^2 \mapsto x_i$ et le symbole $\#$ signale l'opération de « pousser en avant ». Quand le contexte s'y prêtera on utilisera aussi proj^t lorsqu'une variable, ici t mais souvent x ou y , sera associée à un facteur de l'espace produit (ici \mathcal{X}^2 mais plus loin souvent \mathcal{X}^T).

La désintégration de π selon μ induit une famille $(\pi_{x,\cdot})_{x \in \mathcal{X}}$ de mesures dans $\mathcal{P}(\mathcal{X})$ définie presque partout pour μ (de masse finie). Nous noterons $\pi(dx, dy) = \mu(dx)\pi_{x,\cdot}(dy)$. Deux cas spéciaux seront $\pi_{x,\cdot} = \nu/\nu(\mathbb{R})$ correspondant au transport indépendant, et $\pi_{x,\cdot} = \delta_{T(x)}$ lorsque le plan de transport est déterministe, c'est-à-dire que la destination est déterminée par T . Si

1. c'est-à-dire une espace topologique métrisable de façon à être complet et séparable. De façon abusive nous pourrions considérer qu'il s'agit directement d'un espace métrique.

$T_{\#}\mu = \nu$ on parle pour T d'*application de transport*. Des compléments relatifs à la désintégration des mesures, leurs paramétrages, ainsi qu'à la composition et la concaténation des plans de transport et des noyaux seront donnés aux pages 51 et 59.

Précisons l'interprétation donnée à μ , ν et π en transport optimal : la mesure μ représente chez Monge un déblai qu'il s'agit de déplacer pour lui donner la forme d'un remblai distribué selon ν . La masse transportée de $A \subset \mathcal{X}$ vers $B \subset \mathcal{X}$ est de mesure $\pi(A \times B)$.

Nous utiliserons le terme peu usité de *route de transport* pour les couples de points $(x, y) \in \mathcal{X}^2$. Définir un plan de transport consiste donc à pondérer l'ensemble de ces routes de transport. On notera $\text{Marg}(\mu, \nu)$ l'ensemble des plans de transports entre μ et ν . Lorsque, plus loin au chapitre 3, nous introduirons des processus stochastiques, des notations similaires seront employées pour $\text{Marg}((\mu_t)_{t \in \mathcal{T}})$ l'espace des mesures à marges prescrites sur $\mathcal{X}^{\mathcal{T}}$.

Nous pouvons maintenant formuler le problème de Monge–Kantorovich associé à deux mesures μ et ν , à un ensemble contraignant $M \subset \text{Marg}(\mu, \nu)$ et à une fonction (mesurable) de coût $c : (x, y) \in \mathcal{X}^2 \mapsto \mathbb{R}$.

Définition 1.1.1. Le problème de Monge–Kantorovich consiste à

$$\text{Minimiser } C : \pi \mapsto \iint c(x, y) d\pi(x, y) \quad \text{pour } \pi \in M \subset \text{Marg}(\mu, \nu),$$

et, le cas échéant, à prouver l'existence et l'unicité des solutions ainsi qu'à décrire ou caractériser ces dernières.

Nous noterons fréquemment M_c^* (ou Marg_c^* si $\text{Marg} = M$) l'ensemble des minimiseurs et π^* leurs éléments.

Remarque 1.1.2. *Le problème peut s'avérer non déterminé de différentes manières. Il peut arriver que M soit l'ensemble vide, par exemple si μ et ν ne sont pas de même masse. Il peut aussi se faire que c ne soit pas intégrable sous certaines mesures $\pi \in M$. Même bien définie, l'intégrale pourrait prendre la valeur $+\infty$ sur tout M . Il y a donc des préalables à l'étude du problème associé à \mathcal{X} , μ , ν , M et c . Tout au long du texte, nous nous placerons sous des hypothèses permettant d'éviter les cas dégénérés. Nous supposons $M \neq \emptyset$, la partie négative de c sera intégrable pour tout élément $\pi \in \text{Marg}(\mu, \nu)$. Il sera pour cela suffisant de supposer $c(x, y) \geq a(x) + b(y)$ où a et b sont intégrables, respectivement par rapport à μ et ν . Finalement nous supposons qu'il existe au moins un élément $\pi \in M$ tel que $\iint c d\pi < +\infty$.*

Présentons maintenant les trois situations principales pour lesquelles nous étudierons le problème de transport dans le présent chapitre.

Coût quadratique Dans un espace métrique, (\mathcal{X}, d) pour se donner une notation, on parle de coût quadratique lorsque la fonction de coût est $c_2 : (x, y) \mapsto d(x, y)^2$. Dans ce chapitre nous aurons $\mathcal{X} = \mathbb{R}^d$ et $|x|$ représentera la norme euclidienne de x . On considère μ et ν concentrées sur² deux courbes et $c : (x, y) \mapsto |y - x|^2$. Le cas de deux droites identiques, qu'on peut identifier à \mathbb{R} , est extrêmement basique et fondamental en mathématique, pour commencer en statistique mathématique puisqu'il fait apparaître le couplage par quantiles et correspond à la corrélation maximale d'une variable aléatoire (X, Y) de lois marginales μ et ν . Nous verrons que lorsque les courbes sont certains graphes lipschitziens, l'existence, l'unicité et la caractérisation par quantile sont conservées.

Coût sous-linéaire, limite en s'approchant du coût linéaire Ici $\mathcal{X} = \mathbb{R}$, μ et ν sont éléments de $\mathcal{P}_1(\mathbb{R})$ et $c : (x, y) \mapsto |y - x|^{1^-}$ où le sens à donner à 1^- est à préciser. En peu de mots on considère les ensembles optimaux pour $c = c_p : (x, y) \mapsto |y - x|^p$. Ceux-ci possèdent un unique point d'accumulation lorsque $p < 1$ tend vers 1 par valeurs inférieures. C'est une première façon de voir apparaître l'existence et l'unicité du plan de transport que nous appellerons *transport montagne* entre μ et ν . La deuxième façon qui rentre, elle, véritablement dans le cadre de la définition 1.1.1, consiste à voir ce transport montagne

2. c'est-à-dire affectant la masse totale à

comme l'unique transport optimal du problème de transport dont l'ensemble contraignant est $M = \text{Marg}_1^*(\mu, \nu)$ (celui des solutions du problème de coût c_1) et le coût c_p , où tout $p \in]0, 1[$ convient.

Coût c du type $\partial_x \partial_y^2 c < 0$ sous contrainte martingale De nouveau \mathcal{X} sera \mathbb{R} . On se donnera la contrainte supplémentaire de ce que π doit être un *plan de transport martingale* plus simplement nommé *transport ou couplage martingale*. Nous noterons $\text{Mart}(\mu, \nu) = M$ l'ensemble des transports martingale de $\text{Marg}(\mu, \nu)$. Alors que le coût quadratique et certains autres satisfont à $\partial_x \partial_y c < 0$, les coûts analogues du problème martingale satisfont à $\partial_x \partial_y^2 c < 0$, une condition satisfaite par exemple par $c : (x, y) \mapsto (y - x)^3$ ou, pour parler d'un coût positif, $c : (x, y) \mapsto \exp(y - x)$, ou bien encore par $c : (x, y) \mapsto \sqrt{1 + y^2} \frac{e^{-x}}{e^{-x} + e^x}$ qui présente l'avantage d'être à la fois positif et intégrable en y par toute mesure de $\mathcal{P}_1(\mathbb{R})$.

1.2 Rappels sur le problème à coût puissance sur-linéaire et sur le transport quantile

Nous présentons tout d'abord brièvement l'archétype du problème de transport sur \mathbb{R} . Les étapes de sa résolution seront les mêmes que pour les problèmes qui viendront.

À l'aide d'un lemme d'échange on repère, pour un transport optimal donné, les combinaisons de routes interdites ou autorisées, on définit un plan de transport spécifique qui respecte ces interdictions puis on démontre que c'est l'unique plan de transport qui les respecte. Comme il existe au moins un plan de transport optimal (ce qui est en général démontré par le théorème de Prokhorov) on a trouvé une équivalence entre transport optimal, transport respectant des interdictions et le transport spécifique. Les théorèmes 1.3.2, 1.4.7 et 1.5.8 seront de ce type et concerneront, respectivement, les transports quantiles sur des graphes, les transports montagne et les transports riveaux gauches.

Nous rappelons tout d'abord le théorème de monotonie cyclique, fondamental en transport optimal. Nous l'énonçons comme modèle dans un cadre général (celui de l'énoncé d'une équivalence pour un coût borélien) au théorème 1.2.3. Toutefois, dans ce chapitre des versions élémentaires (implication seule et coût continu), que nous nommerons *lemmes d'échange* seront suffisantes. Elles présentent l'avantage de posséder des démonstrations élémentaires. Leur modèle sera le corollaire 1.2.4. Nous aurons aussi des lemmes d'échanges pour le problème à mesures concentrées sur des graphes dans \mathbb{R}^d (corollaire 1.3.1), un lemme d'échange pour un problème secondaire (lemme 1.4.4) et un lemme d'échange dans le cadre du transport martingale (lemme 1.5.4).

Définition 1.2.1 (Ensemble finiment optimal). On dit qu'un borélien S est *finiment optimal* pour c si tout plan de transport $\alpha \in \mathcal{M}_+(\mathcal{X}^2)$ dont le support est fini et inclus dans S est aussi un transport optimal sur $\text{Marg}(\text{proj}_\#^x \alpha, \text{proj}_\#^y \alpha)$. On appelle par ailleurs *compétiteur* tout élément α' de $\text{Marg}(\text{proj}_\#^x \alpha, \text{proj}_\#^y \alpha)$. L'ensemble S est donc finiment optimal si

$$\iint c \, d\alpha \leq \iint c \, d\alpha'$$

pour tout α avec $\text{Spt}(\alpha) \subset S$ fini et tout compétiteur α' (non nécessairement concentré sur S).

Remarque 1.2.2 (Notion équivalente classique). *Un ensemble est finiment optimal si et seulement si il est cycliquement monotone tel que défini dans les traités de référence. Pour rappel, on dit qu'un borélien $S \subset \mathcal{X} \times \mathcal{X}$ est cycliquement monotone (pour c) si*

$$\sum_{k=1}^n c(x_k, y_k) \leq \sum_{k=1}^n c(x_k, y_{k+1}) \quad (1.1)$$

est satisfaite pour toute famille de routes $(x_k, y_k)_{k=1, \dots, n}$ dans S et tout $n \in \mathbb{N}^$. Ici par convention $y_{n+1} = y_1$ ce qui clôt donc un cycle $y_1, x_1, y_2, x_2, \dots, y_n, x_n, y_1$ dont les paires d'éléments*

consécutifs constituent les routes d'un des deux transports pour celle commençant par \mathbf{y}_k , de son compétiteur, pour celle commençant par \mathbf{y}_k .

Alors que $\alpha' = \sum_{k=1}^n \delta_{x_k, y_{k+1}}$ (où $y_{n+1} = y_1$) est un compétiteur de $\alpha = \sum_{k=1}^n \delta_{x_k, y_k}$ et qu'il en va de même des combinaisons linéaires à coefficients positifs impliquant de telles mesures ($\sum_k \alpha'_k$ est un compétiteur de $\sum_k \alpha_k$), on montre assez simplement que toute paire (α, α') de la définition 1.2.1 se représente de cette manière. Il s'ensuit que, comme annoncé, S est cycliquement monotone (pour un coût c donné) si et seulement si il est finiment optimal.

Théorème 1.2.3 (Théorème fondamental de monotonie). *Soit c une fonction de coût borélienne telle que les plans de transport de $\text{Marg}(\mu, \nu)$ sont intégrables. Un transport π est optimal (pour le problème sans contraintes) si et seulement si il existe un borélien $S \subset \mathcal{X} \times \mathcal{X}$ finiment optimal pour c tel que $\pi(S) = 1$.*

Ainsi un plan de transport est optimal si pour un problème discrétisé aucun déroutage n'amène un coût strictement moindre. Un corollaire est le suivant.

Corollaire 1.2.4 (Lemme d'échange). *Soit π un plan de transport optimal. Alors il existe un borélien S de mesure pleine tel que tout (x, y) et (x', y') pris dans S satisfait à*

$$c(x, y) + c(x', y') \leq c(x, y') + c(x', y). \quad (1.2)$$

Si de plus c est un coût continu, tout S tel que $\bar{S} \subset \text{Spt}(\pi)$ et $\pi(S) = 1$ convient.

Le théorème 1.2.3 est une équivalence et concerne des coûts généraux. Sa démonstration nécessite un arsenal spécial, désormais classique, mettant en jeu des théorèmes de dualité (voir notamment [BS11, BC10, Bei15, Kel84, BGMS09] pour des résultats de dualité et monotonie pour c parmi les plus avancés dans les espaces polonais). Toutefois, dans cette partie, nous n'allons utiliser qu'une seule implication et, de plus, seulement pour des coûts continus. La démonstration est alors extrêmement simple et se résume à ceci : on prend S , c et des routes (x, y) , (x', y') contredisant (1.2). On peut supposer $S \subset \text{Spt}(\pi)$ et $\bar{S} = \text{Spt}(\pi)$. Ainsi, pour tout $\varepsilon > 0$ suffisamment petit il est possible de trouver des mesures positives γ et γ' satisfaisant à

- $\gamma \preceq_+ \pi$ (c'est-à-dire $\pi - \gamma \in \mathcal{M}_+(\mathcal{X})$) et γ est concentrée sur $[\mathcal{B}(x, \varepsilon) \times \mathcal{B}(y, \varepsilon)] \cup [\mathcal{B}(x', \varepsilon) \times \mathcal{B}(y', \varepsilon)]$, où $\mathcal{B}(x, \varepsilon)$ est la boule de centre x et rayon ε .
- $\gamma' \in \text{Marg}(\text{proj}_{\#}^1 \gamma, \text{proj}_{\#}^2 \gamma)$, et γ' est concentrée sur $[\mathcal{B}(x, \varepsilon) \times \mathcal{B}(y', \varepsilon)] \cup [\mathcal{B}(x', \varepsilon) \times \mathcal{B}(y, \varepsilon)]$.

Alors, du fait de la continuité de c , pour ε suffisamment petit on trouve $\int c d\gamma \geq \int c d\gamma'$ et le transport $\pi - \gamma + \gamma' \in \mathcal{M}_+(\mathcal{X}^2)$ est un élément de $\text{Marg}(\mu, \nu)$ qui porte un coût total strictement inférieur à celui associé à π .

Remarque 1.2.5. *Le corollaire 1.2.4 trouvera dans la partie sur le couplage montagne un pendant avec le lemme 1.4.4, également simple à démontrer. Dans la partie concernant le transport martingale, on trouvera le même type de résultat (une seule implication pour un coût continu) avec le lemme 1.5.4.*

À noter qu'un théorème élaboré correspondant au théorème 1.2.3 sera aussi cité avec le théorème 1.5.3. Ce dernier est démontré dans [BeiJ16].

Le problème puissance sur-linéaire et le transport quantile

Ici, nous traitons de $\mathcal{X} = \mathbb{R}$, $M = \text{Marg}(\mu, \nu)$ et $c = c_p : (x, y) \mapsto |y - x|^p$ où $p > 1$.

Corollaire 1.2.6 (Lemme d'échange). *Soit π un plan de transport optimal sur \mathbb{R} et $c = c_p$ pour $p > 1$. Alors il existe un borélien S vérifiant $\pi(S) = 1$ et $\bar{S} \subset \text{Spt} \pi$ tel que tout couple de routes (x, y) et (x', y') pris dans S vérifie $|y - x|^p + |y' - x'|^p \leq |y' - x|^p + |y - x'|^p$ et donc*

$$(x' - x)(y' - y) \geq 0. \quad (1.3)$$

Démonstration. Nous appliquons simplement le lemme d'échange (corollaire 1.2.4) à $c = c_p$. La condition $(y - x)(y' - x') \geq 0$ résulte d'une étude de fonction élémentaire. On cherche en effet le signe de $f(y') - f(y)$ où la fonction f est $y \mapsto |y - x'|^p - |y - x|^p$. Sa dérivée en y vaut $f'(y) = h'(y - x') - h'(y - x)$ où $h : y \mapsto |y|^p$. La fonction h étant strictement convexe, sa dérivée h' est strictement croissante. Ainsi, $f(y') - f(y)$ est du même signe que $(y' - y)(x' - x)$. \square

Nous avons donc identifié une interdiction dans les combinaisons de routes : si (x, y) et (x', y') sont des routes de transports, il n'y a pas de croisement dans l'ordre des destinations par rapport à celui des origines. Nous représenterons cette interdiction dans le plan en matérialisant les routes par des segments entre un point x sur une droite réelle de \mathbb{R}_x où est distribuée μ et un point y sur une droite réelle de variable \mathbb{R}_y , parallèle à la première, où est distribuée ν . Les segments $[xy]$ représentant les routes peuvent avoir des extrémités communes mais ne doivent pas se croiser. Voir la figure 1.1 page 25 pour un exemple.

Pour définir le transport quantile nous faisons tout d'abord quelques rappels, utiles non seulement dans ce chapitre mais dans tout le mémoire, sur les fonctions de répartition et les fonctions quantiles.

Définition 1.2.7 (Fonction de répartition). Soit $\mu \in \mathcal{M}_+(\mathbb{R}^d)$. Sa fonction de répartition F_μ , que nous noterons parfois $F[\mu]$ est définie par

$$F_\mu : x \in \mathbb{R}^d \mapsto \mu(] - \infty, x])$$

où $] - \infty, x] = \prod_{i=1}^d] - \infty, x_i]$ avec $x = (x_1, \dots, x_k)$.

Rappelons que $F : \mu \mapsto F[\mu]$ est injective ; la fonction de répartition caractérise la mesure. Avec les fonctions de répartition nous pouvons définir l'ordre \preceq_{lo} dont nous ferons usage à plusieurs reprises dans les chapitres suivants, comme à la page 71.

Définition 1.2.8 (Ordres stochastiques). Soit μ et ν deux mesures de $\mathcal{M}_+(\mathbb{R}^d)$ de même masse. On note $\mu \preceq_{lo} \nu$ si $F[\mu] \geq F[\nu]$. Pour $k = 1$, on note $\mu \preceq_{sto} \nu$ et dit que ν est plus grand que μ dans l'ordre stochastique

Remarque 1.2.9. Comme on peut le constater, \preceq_{lo} prolonge l'ordre stochastique \preceq_{sto} aux dimensions supérieures à deux. Il existe pour autant un ordre stochastique distinct pour les espaces partiellement ordonnés dont la définition, plus cohérente, s'accorde au théorème de Strassen (énoncé à la proposition 2.1.4 pour \mathbb{R}).

Définition 1.2.10 (Fonction quantile et barycentre). Pour $\mu \in \mathcal{M}_+(\mathbb{R})$, la fonction quantile $G_\mu : [0, \text{Masse}(\mu)] \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$, qu'on pourra aussi noter $G[\mu]$, est la fonction croissante continue à gauche telle que $(G_\mu)_\# \lambda = \mu$, c'est-à-dire l'inverse généralisé de la fonction de répartition F_μ :

$$G_\mu(\alpha) := \inf\{x \in \mathbb{R} : F_\mu(x) \geq \alpha\}.$$

Si G_μ est intégrable, on définit le *barycentre* comme

$$\text{Bary}(\mu) = \text{Masse}(\mu)^{-1} \int_{\mathbb{R}} x d\mu(x) = \text{Masse}(\mu)^{-1} \int_0^1 G_\mu(\alpha) d\alpha.$$

Si $\mu \in \mathcal{P}(\mathbb{R})$, la fonction quantile G_μ est une variable aléatoire définie sur $([0, 1], \mathcal{B}, \lambda)$, de loi μ et d'espérance (si celle-ci est bien définie) $\text{Bary}(\mu)$.

L'interprétation graphique de la fonction quantile consiste, pour α donné, à considérer le point d'intersection du graphe de F_μ avec la droite horizontale d'ordonnée α . Les sauts correspondant aux atomes sont complétés par des segments verticaux et lorsque la droite verticale croise un palier horizontal (correspondant à une absence de masse) on prend le point le plus en bas (voir de nouveau la figure 1.1).

Définition 1.2.11 (Transport quantile et inégalité de Hoeffding–Fréchet). Soit μ et ν des mesures de probabilité. Le transport quantile $\pi \in \text{Marg}(\mu, \nu)$ de μ et ν est la loi de la variable aléatoire

$$(G_\mu, G_\nu) : \alpha \in ([0, 1], \lambda) \mapsto (G_\mu(\alpha), G_\nu(\alpha)) \in (\mathbb{R} \cup \{\pm\infty\})^2.$$

Il est également caractérisé par le fait que l’inégalité de Hoeffding–Fréchet $F_\pi(x, y) \leq \min(F_\mu(x), F_\nu(y))$ est une égalité. Le transport quantile s’étend naturellement aux mesures de (mêmes) masses finies. Nous le noterons $\mathfrak{Q}(\mu, \nu)$.

En prenant $S = (G_\mu, G_\nu)(]0, 1[)$, les deux fonctions G_μ et G_ν étant croissantes, il est bien clair que le transport quantile respecte la condition $(y' - y)(x' - x) \geq 0$ du corollaire 1.2.6. Nous montrons maintenant que c’est l’unique transport de marges μ et ν vérifiant cela.

Lemme 1.2.12. *Soit μ, ν des mesures de probabilités sur \mathbb{R} et $\pi \in \text{Marg}(\mu, \nu)$ tels qu’il existe $S \subset \mathbb{R} \times \mathbb{R}$ avec $\pi(S) = 1$ et $(y' - y)(x' - x) \geq 0$ pour tout $(x, y, x', y') \in S \times S$. Alors π est entièrement déterminé par la formule, liant les fonctions de répartition*

$$F_\pi(x, y) = \pi(]-\infty, x[\times]-\infty, y]) = \min(F_\mu(x), F_\nu(y)).$$

On trouve donc $\pi = \mathfrak{Q}(\mu, \nu)$.

La démonstration n’a rien d’original. Elle est extrêmement classique. Nous la donnons tout de même comme modèle pour celles, analogues, que nous ne ferons pas qui concernent les autres problèmes de transport de ce chapitre.

Démonstration. Soit $\pi \in \text{Marg}(\mu, \nu)$ et S de masse pleine telle que $(y' - y)(x' - x) \geq 0$ pour tout $(x, y, x', y') \in S \times S$. Nous souhaitons montrer $F_\pi(s, t) = \min(F_\mu(s), F_\nu(t))$. L’inégalité \leq est celle de Hoeffding–Fréchet. Elle vient simplement de $F_\mu(s) = \pi(]-\infty, s] \times \mathbb{R})$ et $F_\nu(t) = \pi(\mathbb{R} \times]-\infty, t])$. Pour l’autre inégalité il suffit de démontrer que $F_\mu(s) > F_\pi(s, t)$ entraîne $F_\pi(s, t) = F_\nu(t)$. ($F_\nu(t) > F_\pi(s, t)$ entraînera $F_\pi(s, t) = F_\mu(s)$ pour les mêmes raisons). Supposons donc $F_\mu(s) > F_\pi(s, t)$. On a alors $\pi(]-\infty, s] \times]t, +\infty[) > 0$ et il existe ainsi une route $(x, y) \in S$ telle que $x \leq s$ et $y > t$. On déduit ensuite de $(y' - y)(x' - x) \geq 0$ qu’aucun point (x', y') de $]s, +\infty[\times]-\infty, t[\subset]x, +\infty[\times]-\infty, y[$ ne peut être une route de S . On trouve alors

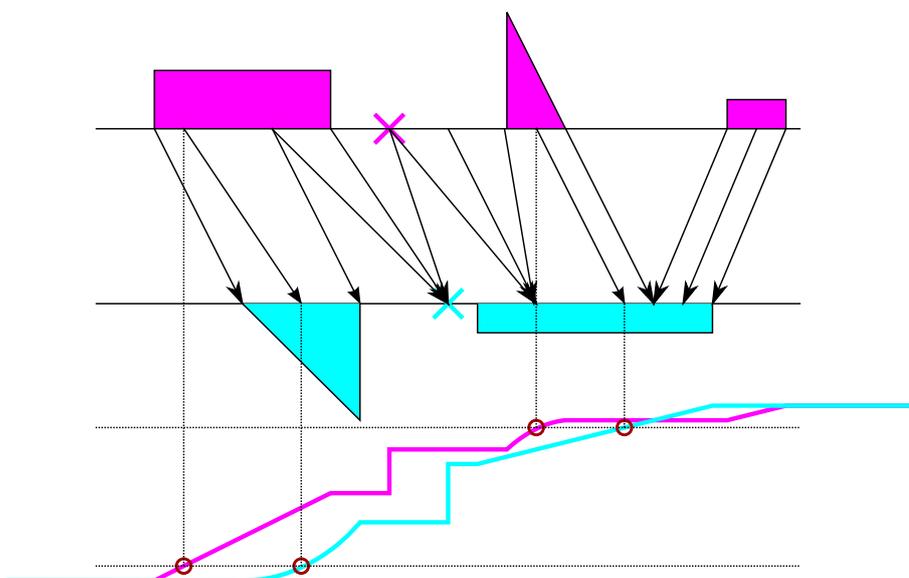
$$\begin{aligned} F_\nu(t) - F_\pi(s, t) &= \pi(\mathbb{R} \times]-\infty, t[\setminus]-\infty, s] \times]-\infty, t]) \\ &= \pi(]s, +\infty[\times]-\infty, t]) = 0. \end{aligned}$$

□

Comme le théorème de Prokhorov garantit l’existence d’un transport optimal, avec le corollaire 1.2.6 et le lemme 1.2.12 on peut conclure au théorème suivant, qui sera le modèle des théorèmes à venir dans ce chapitre.

Théorème 1.2.13. *Soit $p > 1$ et $\mu, \nu \in \mathcal{P}_p(\mathbb{R})$. Soit π un transport de $\text{Marg}(\mu, \nu)$. Alors les énoncés suivants sont équivalents :*

- π est un transport optimal pour le coût c_p .
- π est monotone au sens où il existe S de masse pleine avec $(y' - y)(x' - x) \geq 0$ sur $S \times S$.
- π est le transport quantile $\mathfrak{Q}(\mu, \nu)$ entre μ et ν .

FIGURE 1.1 – Le transport quantile Ω

Distances de Wasserstein. Cas des mesures sur la droite réelle.

Nous rappelons la définition de la distance de Wasserstein, directement liée à l'étude du problème à coût puissance surlinéaire. Le théorème 1.2.13 nous donnera l'occasion de rappeler certaines formules valables lorsque $\mathcal{X} = \mathbb{R}$.

Définition 1.2.14. Soit (\mathcal{X}, d) un espace métrique complet et séparable. Pour tout $p \geq 1$, on appelle *espace de Wasserstein* l'espace des mesures de probabilité boréliennes possédant un p^e moment fini. On note $\mathcal{P}_p(\mathcal{X})$ cet espace. Ainsi

$$\mathcal{P}_p(\mathcal{X}) = \left\{ \mu \in \mathcal{P}(\mathcal{X}) : \int_{\mathcal{X}} d(x, x_0)^p d\mu(x) \right\}$$

où x_0 est un point quelconque de \mathcal{X} (cela n'influence pas la définition).

Étant donné $\mu, \nu \in \mathcal{M}_+(\mathbb{R})$, on introduit la *distance de Wasserstein* $W_p(\mu, \nu)$ comme suit :

$$W_p(\mu, \nu) = \inf_{\pi \in \text{Marg}(\mu, \nu)} \left(\iint d(x, y)^p d\pi(x, y) \right)^{1/p}.$$

Sur $\mathcal{P}_p(\mathcal{X})$, W_p est effectivement une distance. Elle vérifie en particulier l'inégalité triangulaire et prend ses valeurs dans $[0, \infty[$. Pour deux mesures de masses différentes, la distance est infinie ($\text{Marg}(\mu, \nu)$ est d'ailleurs vide). La topologie induite par W_p sur $\mathcal{P}_p(\mathcal{X})$ est la topologie faible avec p^e moment telle que décrite dans [Vil03, Theorem 7.12].

Les propriétés suivantes, classiques, concernent le transport sur \mathbb{R} .

Proposition 1.2.15. Soit μ et ν des mesures de $\mathcal{P}_p(\mathcal{X})$. On a

$$W_p(\mu, \nu) = \left(\int |G_\nu - G_\mu|^p(\alpha) d\alpha \right)^{1/p} \geq |\text{Bary}(\mu) - \text{Bary}(\nu)|.$$

L'égalité est obtenue exactement dans les cas suivants :

- si μ et ν sont des translatées l'une de l'autre,

— de plus, seulement dans le cas $p = 1$: si $G_\mu \leq G_\nu$ ou $G_\nu \leq G_\mu$ (c'est-à-dire si $\mu \preceq_{sto} \nu$ ou $\nu \preceq_{sto} \mu$, voir les définitions 1.2.8 et 1.2.10).

Démonstration. On utilise du théorème 1.2.13 le fait que (G_μ, G_ν) a pour loi le transport optimal. Cela donne

$$\begin{aligned} W_p(\mu, \nu)^p &= \int |G_\nu - G_\mu|^p(\alpha) \, d\alpha \\ &\leq \left| \int (G_\nu - G_\mu)(\alpha) \, d\alpha \right|^p = |\text{Bary}(\mu) - \text{Bary}(\nu)|^p. \end{aligned} \quad (1.4)$$

Le cas d'égalité repose sur l'analyse de l'inégalité de Jensen, sur le fait que $x \rightarrow x^p$ est strictement convexe pour $p > 1$ et sur celui que pour $p = 1$, l'application $x \mapsto |x|$ est linéaire par morceaux sur \mathbb{R}^- et \mathbb{R}^+ . \square

Un résultat trahissant un comportement différent pour la puissances $p = 1/2$, sera énoncé au lemme 4.4.7 pour être utilisé dans le cadre du transport de mesures sur le groupe de Heisenberg.

Autres caractérisations et propriétés du transport quantile

En préparation aux chapitres suivants nous évoquons ici une autre caractérisation du transport quantile.

Définition 1.2.16. Soit ν et $\alpha \in [0, \text{Masse}(\nu)]$. Nous nommons *mesure de masse α la plus à gauche sous ν* la mesure $\eta = (G_\nu)_\# \lambda_{[0, \alpha]}$. Elle possède en effet la propriété suivante : si ν' vérifie à la fois $\text{Masse}(\eta') = \text{Masse}(\eta) = \alpha$ et $\eta' \preceq_+ \nu$, alors $\eta \preceq_{sto} \eta'$.

Proposition 1.2.17. Soit μ et ν des mesures réelles et $\pi = \Omega(\mu, \nu)$. Alors pour tout $x \in \mathbb{R}$ la mesure $\nu_x^\pi = (\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}}$ est un élément η de

$$S_x = \{\eta : \eta \preceq_+ \nu \text{ et } \text{Masse}(\eta) = \mu([-\infty, x])\}$$

qui vérifie de plus $\eta \preceq_{sto} \eta'$ pour tout autre élément $\eta' \in S_x$.

En particulier, en termes probabilistes, si $(X, Y) \sim \Omega(\mu, \nu)$, alors $\eta = \text{Loi}(Y|X \leq x)$ est inférieure pour \preceq_{sto} à toute mesure $\eta' = \text{Loi}(Y'|X' \leq x)$ où $\text{Loi}(X', Y') \in \text{Marg}(\mu, \nu)$.

Réciproquement, si pour tout x la mesure $(\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}}$ est le minimum pour \preceq_{sto} de S_x , alors la mesure π est le transport quantile $\Omega(\mu, \nu)$.

Remarque 1.2.18 (Signification de $(\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}}$). Nous rencontrerons à d'autres occasions des mesures $(\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}}$ associées à π et x . Leur interprétation peut être malaisée. Elles sont le résultat du transport par π de la mesure $\mu_{[-\infty, x]}$. On serait tenter de les noter « $\pi_\# \mu_{[-\infty, x]}$ » puisque dans le cas déterministe $\pi = (\text{Id}, T)_\# \mu$ on a vraiment $(\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}} = T_\# \mu_{[-\infty, x]}$.

Par ailleurs la mesure π peut se noter $\mu \cdot (\text{id}, k)$ où $k : x \in \mathbb{R} \mapsto k(x, \cdot) \in \mathcal{P}(\mathbb{R})$ est un noyau (voir à la page 59 pour des rappels sur les noyaux). En fait $k(x) = \pi_x$, presque sûrement pour μ , avec $(\pi_x, \cdot)_{x \in \mathbb{R}}$ comme à la page 20). Alors $(\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}} = \mu_{[-\infty, x]} \cdot k$. Cette deuxième écriture présente l'avantage de s'étendre puisque $\theta \cdot k$ a du sens pour toute mesure $\theta \ll \mu$.

Démonstration de la proposition 1.2.17. Partons de (G_μ, G_ν) dont la loi est $\Omega(\mu, \nu)$. La condition $\{G_\mu \leq x\}$ désigne une partie $[0, \alpha] \subset [0, \text{Masse}(\mu)]$. Ainsi $(\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}}$ est exactement $(G_\nu)_\# \lambda_{[0, \alpha]}$, c'est à dire la mesure de masse α la plus à gauche sous ν .

Réciproquement, soit π telle que pour tout $x \in \mathbb{R}$, $\pi_{[-\infty, x] \times \mathbb{R}}$ soit la mesure de masse $F_\mu(x) = \mu([-\infty, x])$ la plus à gauche sous ν . Sa fonction de répartition est donc donnée par $F_\pi(x, y) = (\text{proj}^y)_\# \pi_{[-\infty, x] \times \mathbb{R}}([-\infty, y]) = (G_\nu)_\# \lambda_{[0, \alpha]}([-\infty, y])$ où $\alpha = F_\mu(x)$. Si $y \geq G_\nu(\alpha)$ l'ensemble est de mesure pleine, à savoir de mesure $\alpha = F_\mu(x)$. Sinon $(G_\nu)_\# \lambda_{[0, \alpha]}([-\infty, y]) = \nu([-\infty, y]) = F_\nu(y)$. En tout état de cause $F_\pi(x, y) = \min(F_\mu(x), F_\nu(y))$, si bien que $\pi = \Omega(\mu, \nu)$ (voir la définition 1.2.11). \square

Remarque 1.2.19. Alors qu'il est possible de définir un transport $\pi = \text{Loi}(X, Y)$ par la famille $\text{Loi}(Y|X = x)_{x \in \mathbb{R}}$ la proposition 1.2.17 caractérise la loi jointe (X, Y) de façon univoque par la famille $\text{Loi}(Y|X \leq x)_{x \in \mathbb{R}}$.

De façon générale une famille de mesures de probabilités $(\gamma_x)_{x \in \mathbb{R}}$ définit une loi jointe (X, Y) par le biais de l'équation $\gamma_x = \text{Loi}(Y|X \leq x)$ à la condition nécessaire et suffisante suivante que $x \mapsto \mathbb{P}(X \leq x) \times \gamma_x$ soit croissante pour \preceq_+ . Il est en effet nécessaire que $x \mapsto (\text{proj}^2)_\#(\pi_{[-\infty, x]})$ croisse pour \preceq_+ .

1.3 Le problème à coût puissance sur-linéaire, étude des mesures régulières pour le théorème de Brenier

Dans cette section nous démontrons le théorème 1.3.11 tiré de [Jui11]. Ce résultat concerne la géométrie et les géodésiques de $\mathcal{P}(\mathbb{R}^d)$ vu comme un espace d'Alexandrov ainsi que la notion d'interpolation par déplacement, classique en transport. Alors que $\mathcal{P}(\mathbb{R}^d)$ est de dimension infinie la démonstration de notre théorème passe par un problème de transport de dimension 1 dont la solution est résumée au théorème 1.3.2.

Le problème quadratique et le transport quantile entre des mesures concentrées sur des graphes

Ici, nous traitons de $\mathcal{X} = \mathbb{R}^d$, $M = \text{Marg}(\mu, \nu)$ et $c = c_2 : (x, y) \mapsto \|y - x\|^2$ où la norme est euclidienne. Dans cette situation le lemme d'échange (corollaire 1.2.4) s'applique et s'adapte aisément, donnant l'énoncé suivant dont la démonstration est aussi simple que celle du corollaire 1.2.6.

Corollaire 1.3.1 (Lemme d'échange). Soit π un plan de transport optimal sur \mathbb{R}^d et $c = c_2$. Alors il existe un borélien S vérifiant $\pi(S) = 1$, tel que tout couple de routes (x, y) et (x', y') pris dans S vérifie

$$\langle x' - x, y' - y \rangle \geq 0. \quad (1.5)$$

Par ailleurs, tout S tel que $\pi(S) = 1$ et $\text{Spt}(\pi) \subset \bar{S}$ convient.

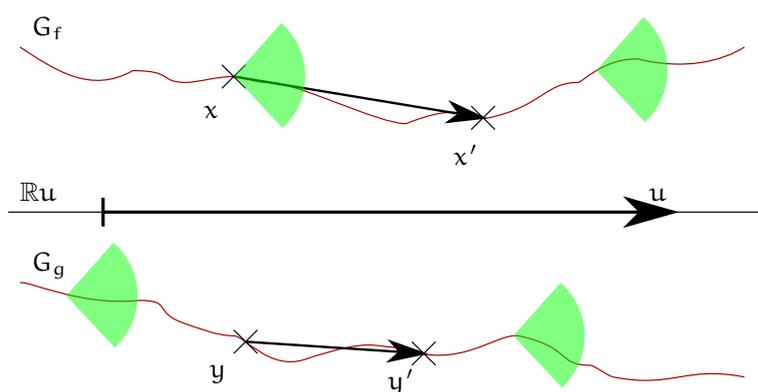


FIGURE 1.2 – Transport entre deux graphes lipschitziens.

Nous allons maintenant appliquer ce résultat au transport optimal entre deux mesures concentrées sur des graphes lipschitziens. Cela mènera au théorème 1.3.2, un analogue de notre théorème modèle 1.2.13.

Théorème 1.3.2 ([Jui11, voir Lemma 1.6]). *Soit $d \geq 2$ un entier et μ et ν des mesures de $\mathcal{P}_2(\mathbb{R}^d)$ concentrées, respectivement, sur G_f et G_g les graphes des fonctions lipschitziennes $f, g : \mathbb{R}u \rightarrow (\mathbb{R}u)^\perp$ de constantes optimales < 1 , où u est un vecteur unitaire. Soit $\pi \in \text{Marg}(\mu, \nu)$. Alors, les énoncés suivants sont équivalents*

- π est un transport optimal pour le coût quadratique c_2 .
- π est concentrée sur un ensemble S de routes de transport vérifiant $\langle x' - x, y' - y \rangle \geq 0$ pour tout $(x, y) \in S$ et $(x', y') \in S$.
- π s'exprime sous la forme $\pi = (f \otimes g)_{\#} \gamma$ où γ est le couplage quantile entre $\text{proj}_{\#}^u \mu$ et $\text{proj}_{\#}^u \nu$ sur l'axe $\mathbb{R}u \cong \mathbb{R}$.

En particulier nous avons l'existence et l'unicité (le transport quantile, f et g sont bien définis) d'un tel transport π .

La démonstration suivante suit le plan logique annoncé en page 21.

Démonstration. Remarques préalables : dans le cadre des hypothèses on dispose de u , un vecteur unitaire de \mathbb{R}^d et de $G_f \subset \mathbb{R}^d$ le graphe défini par $G_f = \{\tilde{f}(x) = (x, f(x)) \in \mathbb{R}u \times (\mathbb{R}u)^\perp : x \in \mathbb{R}u\}$ ou $f : \mathbb{R}u \cong \mathbb{R} \rightarrow (\mathbb{R}u)^\perp$ est une application k_f -lipschitzienne avec $k_f < 1$. Le même u est pris pour G_g avec g lipschitzienne de constante $k_g < 1$ et $\tilde{g} := (\text{Id}, g)$. Alors, la projection orthogonale $\text{proj}^u : \mathbb{R}^d \rightarrow \mathbb{R}u$ induit une bijection, d'inverse $\tilde{f} := (\text{Id}, f)$, entre G_f et $\mathbb{R}u$ d'une part, et, entre les mesures concentrées sur G_f et $\mathbb{R}u$ d'autre part.

Considérons maintenant les mesures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ concentrées l'une sur G_f , l'autre sur G_g . Alors, de nouveau, comme $\text{proj}^u \otimes \text{proj}^u$ constitue une bijection mesurable entre $G_f \times G_g$ et $(\mathbb{R}u)^2$, elle induit aussi une bijection entre $\text{Marg}(\mu, \nu)$ et $\text{Marg}(\text{proj}_{\#}^u \mu, \text{proj}_{\#}^u \nu)$. Finalement, dû au choix de prendre des constantes de Lipschitz < 1 , le signe de $\langle \tilde{f}(x') - \tilde{f}(x), \tilde{g}(y') - \tilde{g}(y) \rangle$, c'est-à-dire > 0 ou $= 0$ ou < 0 , est le même que celui de $\langle x' - x, y' - y \rangle$. Soit $\pi^* \in \text{Marg}^*(\mu, \nu)$ un transport optimal pour le coût quadratique. En vertu du corollaire 1.3.1, son support vérifie (1.5). Ainsi $S = (\text{proj}^u \otimes \text{proj}^u)(\text{Spt}(\pi^*)) \subset (\mathbb{R}u)^2$ est un ensemble de mesure 1 pour $\pi^u := (\text{proj}^u \otimes \text{proj}^u)_{\#} \pi^*$ et on a bien $\langle y' - y, x' - x \rangle \geq 0$ pour tout couple de routes $(x, y), (x', y')$ dans S . La mesure π^u est donc la mesure quantile $\Omega(\text{proj}_{\#}^u \mu, \text{proj}_{\#}^u \nu)$, ce qui finit de caractériser π^* puisque c'est alors $(\tilde{f} \otimes \tilde{g})_{\#} \Omega$.

Notons finalement qu'il existe effectivement un transport optimal π^* : le théorème de Prokhorov assure que $\text{Marg}^*(\mu, \nu)$ est non vide. \square

Les mesures « régulières » pour lesquelles le théorème de Brenier s'applique

Contrairement à ce que nous avons vu pour le coût de transport c_2 , et ce que nous verrons encore pour d'autres problèmes en dimension 1, dans les dimensions supérieures l'ensemble $\text{Marg}^*(\mu, \nu)$ des transports optimaux n'est pas réduit à un unique élément. Le théorème de Brenier est un résultat fondamental qui précise une condition suffisante sur μ pour que cela se produise avec, de plus, un transport déterministe comme solution. L'hypothèse en est que μ doit être *absolument continue*. Le sous-espace des mesures absolument continues présente par la suite l'avantage d'être stable par interpolation de type déplacement, ce qui sera important vis-à-vis du théorème 1.3.11, le résultat principal de cette section.

On trouve toutefois des versions plus élaborées du théorème de Brenier. Elles s'appuient sur des hypothèses plus faibles, par exemple impliquant les « petits ensembles » (voir [Vil03, Theorem 2.12], [San15, Theorem 1.22] et l'exemple 1.3.7 ii). Tout ces théorèmes sont communément et à raison appelés théorème (d'existence et d'unicité) de Brenier [Bre87, Bre91]. On pourra toutefois utilement consulter [MG13, §1.3] au sujet de résultats proches démontrés à la même époque par d'autres auteurs. Une mesure μ de $\mathcal{P}_2(\mathbb{R}^d)$ est *régulière* (pour le théorème de Brenier) si elle vérifie les conclusions classiquement attribuées dans (la plupart des versions de) ce théorème aux mesures absolument continues. La notion a été introduite par Gigli dans [Gig11, Definition 2.3] en commençant par une définition équivalente, celle suivant la proposition 1.3.6.

Définition 1.3.3 (Mesure régulière). Une mesure déblai $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ est dite *régulière*, ou *régulière pour le théorème de Brenier* si pour tout remblai $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ les plans de transport optimaux $\pi \in \text{Marg}^*(\mu, \nu)$ s'écrivent sous la forme $\pi = (\text{Id}, T)_\# \mu$ où T est une application mesurable de \mathbb{R}^d dans lui-même.

Une conséquence immédiate est l'unicité de T et donc de π . Rappelons en effet que si T_1 et T_2 sont des applications de transport amenant le minimum alors $\frac{1}{2}((\text{Id}, T_1)_\# \mu + (\text{Id}, T_2)_\# \mu)$ est également un transport optimal. Puisque tout transport optimal est de la forme $\pi = (\text{Id}, T)_\# \mu$, on en déduit $T_1 = T_2$ presque sûrement. Avec la théorie de la dualité on sait en réalité bien plus. On connaît par exemple la « forme » de T , basée sur les fonctions convexes (voir la proposition 1.3.6).

Il n'est pas essentiel d'énoncer ici le théorème de Brenier sous sa forme la plus complète. Nous nous contenterons donc d'une version courte.

Théorème 1.3.4 (Théorème de Brenier). Soit $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ une mesure absolument continue. Alors pour toute mesure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ il existe un unique plan de transport optimal $\pi \in \text{Marg}^*(\mu, \nu)$. Celui-ci s'écrit sous la forme $\pi = (\text{Id}, \nabla \varphi)_\# \mu$ où $\varphi : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$ est une fonction convexe, U un ouvert convexe et $\nabla \varphi : U \subset \mathbb{R}^d \mapsto \mathbb{R}^d$ est définie presque sûrement pour μ .

En particulier, en des termes anachroniques, les mesures absolument continues de $\mathcal{P}_2(\mathbb{R}^d)$ sont régulières pour le théorème de Brenier.

Comme préalable à la proposition 1.3.6 voyons que tous les transports du type $\nabla \varphi$ peuvent être considérés lorsque μ est une mesure régulière.

Proposition 1.3.5 ([Gig11]). Une mesure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ est régulière si et seulement si pour toute application convexe $\varphi : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$ où U est un ouvert convexe, le lieu de non-différentiabilité de φ est de mesure nulle pour μ .

On peut donc maintenant considérer des applications de transport de la forme $\nabla \varphi : U \subset \mathbb{R}^d \mapsto \mathbb{R}^d$ avec φ convexe car elles sont bien définies pour les mesures régulières.

Proposition 1.3.6 (Voir [Gig11, §2], Théorème de Brenier pour une mesure régulière). Si μ est régulière, pour toute mesure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, il existe (pour le coût quadratique) un unique plan de transport $\pi \in \text{Marg}^*(\mu, \nu)$. Celui-ci s'exprime sous la forme $\pi = (\text{Id}, \nabla \varphi)_\# \mu$ où $\varphi : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$ est une fonction convexe et U un ouvert convexe.

Partant de la proposition 1.3.5 on sait exactement caractériser les mesures régulières (voir [GM96, Gig12] et [Zaj79]). Cette caractérisation est l'énoncé de théorie géométrique de la mesure que voici : une mesure μ est régulière si et seulement si $\pi(G) = 0$ est vérifiée pour tout ensemble G qui s'écrit sous la forme d'un graphe $G = \{(x, f - g(x)) \in (\mathbb{R}u)^\perp \times \mathbb{R}u : x \in (\mathbb{R}u)^\perp\}$, où f et g sont des fonctions convexes. Le travail de Zajíček [Zaj79] indique en effet que le lieu de non-différentiabilité de $\varphi : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$ est recouvert par une réunion dénombrable $\cup_n G_n$ de tels ensembles (pour des vecteurs $(u_n)_n$ variables).

Donnons maintenant des exemples de mesures régulières. Nous donnerons à chaque fois une caractérisation duale en spécifiant les ensembles F auxquels les mesures de la classe en question doivent affecter la valeur zéro.

Exemple 1.3.7 ([Jui11, Exemples 1.9 and 1.10]). *i Les mesures absolument continues. Elles sont régulières en vertu du théorème de Rademacher. Ici nous considérons comme ensembles d'annulation les ensembles F dont la mesure (de Lebesgue \mathcal{L}^d ou, de façon équivalente, Hausdorff \mathcal{H}^d) est zéro.*

ii Les mesures ne donnant pas de masse aux petits ensembles, ceux de dimension de Hausdorff $\leq d - 1$. Ici nous considérons donc comme ensembles d'annulation les F tels que $\mathcal{H}^{d-1+\varepsilon}(F) = 0$ pour tout $\varepsilon > 0$. Un exemple de telle mesure est la mesure 'uniforme' sur un ensemble de dimension de Hausdorff comprise strictement entre $k-1$ et k (et de mesure

finie). On peut par exemple prendre sur \mathbb{R}^2 le produit de deux ensembles triadiques de Cantor, ou, plus précisément la mesure de Hausdorff $\mathcal{H}^{2 \log(2)/\log(3)}$ restreinte à cet ensemble. En effet $2 \log(2)/\log(3) > 1$.

iii Les mesures $k-1$ -purement non-rectifiables au sens de la théorie géométrique de la mesure. Ici F est de type $F_1 = f(\mathbb{R}^{d-1})$ où $f : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^d$ est lipschitzienne, ou de type F_2 à savoir vérifiant $\mathcal{H}^{k-1}(F_2) = 0$. Un exemple sur \mathbb{R}^2 est la mesure uniforme sur l'ensemble de Cantor 4 coins, dont la construction est analogue à celle de l'ensemble du précédent exemple avec, dans la construction, un rapport $1/4$ au lieu de $1/3$ (voir [Mat95, Paj02] au sujet de l'ensemble de Cantor 4 coins).

iv Les mesures $k-1$ -purement non rectifiables pour la régularité \mathcal{C}^2 . Ici, F est de type $F_1 = f(\mathbb{R}^{d-1})$ où $f : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^d$ est de classe $\mathcal{C}^2(\mathbb{R}^{d-1}, \mathbb{R}^d)$ ou de type F_2 , à savoir vérifie $\mathcal{H}^{d-1}(F_2) = 0$. Le fait que ces mesures sont régulières provient d'un résultat d'Alberti [Alb94], que nous utilisons dans [Jui11, Proposition 1.8].

C'est dans la classe iv que nous puiserons notre exemple du théorème 1.3.11. La raison, qui prendra du sens à la remarque 1.3.12, en est que les applications lipschitziennes n'envoient pas les ensembles du type $f(\mathbb{R}^{k-1})$ où f est de classe \mathcal{C}^2 sur des ensembles \mathcal{C}^2 -rectifiable. C'est-à-dire que $f(\mathbb{R}^{k-1})$ peut n'être recouvrable par aucun ensemble $(\cup_n f_n(\mathbb{R}^{k-1}) \cup N$ où $\mathcal{H}^{k-1}(N) = 0$ et les f_n sont de classe \mathcal{C}^2).

Structure géodésique de l'espace d'Alexandrov $\mathcal{P}(\mathbb{R}^d)$

Nous avons donc vu comment la régularité d'une mesure pour le théorème de Brenier implique finalement des considérations de théorie géométrique de la mesure. La caractérisation suivante est peut-être encore plus inattendue puisqu'elle concerne les espaces d'Alexandrov à courbure positive tels que définis, par exemple, dans [BGP92, BBI01]. Rappelons que si \mathcal{X} est un espace d'Alexandrov à courbure positive il en va de même de $\mathcal{P}_2(\mathcal{X})$ muni de sa distance de Wasserstein W_2 . Cela s'applique ainsi à \mathbb{R}^d .

Proposition 1.3.8 ([LV09, Appendix A][Stu06, Proposition 2.10 (iv)][Gig11, Corollary 6.6]). *L'espace de Wasserstein $\mathcal{P}_2(\mathbb{R}^d)$ muni de la distance de Wasserstein W_2 est un espace géodésique qui est à courbure sectionnelle positive au sens d'Alexandrov.*

Toutefois, alors que $\mathcal{P}_2(\mathbb{R}^d)$ est de dimension infinie, la théorie des espaces d'Alexandrov est essentiellement développée pour les espaces de dimension (topologique ou de Hausdorff, cela revient au même) *finie*. Dans cette théorie, un point est dit *régulier* si son cône tangent (au sens de la Définition 12.4.3 dans [AGS08]) est isométrique à un espace euclidien. En étendant le sens de euclidien à hilbertien pour la dimension infinie, on dispose de l'équivalence suivante.

Proposition 1.3.9 (« Régulier égale régulier », [Gig11, Corollary 6.6]). *Un point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ est régulier au sens (géométrique) de la théorie des espaces d'Alexandrov si et seulement si cette mesure μ est régulière pour le théorème de Brenier.*

Le théorème 1.3.11 sera directement en lien avec ce résultat :

Proposition 1.3.10 ([Pet98, Corollary of Theorem 1.2(A), p. 132]). *Les points réguliers constituent un ouvert dense géodésiquement convexe.*

Rappelons la structure géodésique de l'espace de Wasserstein : soit $t \in [0, 1] \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ une courbe de l'espace de Wasserstein. Les deux propriétés suivantes concernant $(\mu_t)_{t \in [0, 1]}$ sont équivalentes :

- C'est une géodésique paramétrée à vitesse constante entre μ_0 et μ_1 , c'est-à-dire que pour tout s, t dans $[0, 1]$ nous avons

$$W_2(\mu_s, \mu_t) = |t - s|W_2(\mu_0, \mu_1).$$

- Il existe $(X_0, X_1) : \Omega \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ un couple de variables aléatoires tel que
 - $\text{Loi}(X_0, X_1)$ est un transport optimal de $\text{Marg}_{\mathbb{C}_2}^*(\mu_0, \mu_1)$
 - $\mu_t = \text{Loi}(tX_1 + (1-t)X_0)$ pour tout $t \in [0, 1]$.

Cette procédure porte aussi le nom d'*interpolation par déplacement*. Le sous-espace $\mathcal{P}^{\text{ac}}(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ des mesures absolument continues de \mathbb{R}^d présente l'intérêt d'être géodésiquement convexe, ou, en d'autres termes, stable pour l'interpolation par déplacement.

Le résultat principal de [Jui11] (voir le théorème 2.5) est le suivant.

Théorème 1.3.11 ([Jui11, §2]). *Soit $d \geq 2$. Il existe μ_0 et μ_1 dans $\mathcal{P}_2(\mathbb{R}^d)$ toutes deux régulières, telles que l'unique géodésique $(\mu_t)_{t \in [0, 1]}$ les reliant passe à l'instant $t = 1/2$ par une mesure $\mu_{1/2}$ qui n'est pas régulière.*

Remarque 1.3.12. *i Le théorème 1.3.11 indique donc qu'il existe un espace d'Alexandrov à courbure positive, et deux points réguliers, qui sont interpolés par une géodésique traversant un point non régulier. Cela fournit une réponse négative à la question posée dans [Gig11, Remark 2.12].*

ii Nous avons rappelé que $\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ est géodésiquement convexe. C'est également le cas des mesures apparaissant aux exemples 1.3.7 ii et iii (mais pas de celle en iv). La raison en est que pour $t \in]0, 1[$, il existe une application de transport T lipschitzienne qui transporte μ_t sur $\mu_0 = T_{\#}\mu_t$ (voir [Vil09, Chapitre 8]). Ainsi les ensembles d'annulation des exemples 1.3.7 i, ii, iii (par exemple ceux vérifiant $\dim_{\mathcal{H}}(F) \leq d - 1$ pour ii) sont-ils préservés lors du transport.

iii Dans le groupe de Heisenberg (introduit au chapitre 4), nous avons démontré avec Alessio Figalli ([FigJ08] résultat de la thèse de doctorat) que $\mathcal{P}_2^{\text{ac}}(\mathbb{H})$ est géodésiquement convexe. On répondait ainsi à une question énoncée dans [AR04] le premier article mêlant transport optimal et géométrie sous-riemannienne. D'autres, dont [FR10] ont suivi. On pourra consulter l'ouvrage [Rif14] pour un panorama.

iv Pour que μ_t (avec $t < 1$) soit dans une des classes i, ii ou iii de l'exemple 1.3.7 il est en fait suffisant que μ_0 en fasse partie. Rien n'est exigible concernant μ_1 . Cette remarque est aussi valable pour $\mathcal{P}_2^{\text{ac}}(\mathbb{H})$.

Démonstration. Nous démontrons ce résultat pour $d = 2$. En tensorisant $\tilde{\mu}_t = \mu_t \otimes \lambda^{\otimes d-2}$ pour tout $t \in [0, 1]$ on obtient une géodésique de $\mathcal{P}(\mathbb{R}^d)$. Le principe est alors le même que ce qui va suivre avec recours au théorème de Fubini pour conclure.

Considérons une application $f : [0, 1] \rightarrow \mathbb{R}$ qui vérifie ceci

- f est lipschitzienne de constante $k < 1$,
- Pour toute fonction $g : [0, 1] \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 , le lieu d'intersection $\{x \in [0, 1] : f(x) = g(x)\}$ est de mesure nulle.

Une telle fonction existe; elle peut être construite comme la primitive d'une trajectoire brownienne à la condition que celle-ci reste dans $[-k, k]$. On peut aussi utiliser une série de fonctions pour la définir.

Nous définissons maintenant $\mu_0 = (\text{Id}, f)_{\#}\lambda$ et $\mu_1 = (\text{Id}, (-f))_{\#}\lambda$. D'après ce que nous avons expliqué au théorème 1.3.2, il existe un unique plan de transport entre ces deux mesures, c'est celui induit par la symétrie dont l'axe est celui des abscisses. On trouve donc $\mu_t = (\text{Id}, ((1-2t)f))_{\#}\lambda$ et en particulier $\mu_{1/2} = \lambda \otimes \delta_0$. Il est très clair que cette mesure n'est pas régulière, par la proposition 1.3.5, par exemple. En revanche les mesures μ_0 et μ_1 le sont car elles entrent dans la classe iv de l'exemple 1.3.7 : elle affecte la masse 0 aux courbes de classe \mathcal{C}^2 , ce qu'on obtient avec les hypothèses faites sur f en appliquant le théorème de Sard. \square

1.4 Le problème à coût puissance sous-linéaire et le transport montagne

Ce travail est en cours d'écriture [Jui18⁺⁺]. Nous avons souhaité en faire état parce qu'il s'insère naturellement dans le chapitre 1 de ce mémoire.

Ici, nous traitons de $\mathcal{X} = \mathbb{R}$, $M = \text{Marg}(\mu, \nu)$ et $c = c_p : (x, y) \mapsto |y - x|^p$ où $p \in]0, 1[$. L'objectif de cette section est de faire apparaître un transport vérifiant certaines contraintes. Pour la source d'inspiration, voir la remarque 1.4.6 au sujet des mesures aléatoires.

Les mesures μ et ν considérées sont dans $\mathcal{P}_1(\mathbb{R})$ si bien que, pour $p \in]0, 1[$, le coût $c_p : (x, y) \mapsto |y - x|^p$ est intégrable pour tout élément π de $\text{Marg}(\mu, \nu)$. Alors que pour $p > 1$ (et des marges dans $\mathcal{P}_p(\mathbb{R})$) nous savons que le transport optimal est le transport quantile, ce, indépendamment de la valeur de p , la situation pour $p < 1$ est bien moins nette. L'exemple suivant, central, témoigne d'une part de la *non-unicité* des solutions au problème, d'autre part du fait que, lorsque elle est unique, cette solution *dépend du paramètre* p .

Exemple 1.4.1. On pose $\mu = \frac{1}{2}(\delta_0 + \delta_5)$ et $\nu = \frac{1}{2}(\delta_4 + \delta_9)$. L'ensemble des plans de transport est alors clairement

$$\text{Marg}(\mu, \nu) = \{\pi = \lambda\pi' + (1 - \lambda)\pi'' : \lambda \in [0, 1]\}$$

où $\pi' = \frac{1}{2}(\delta_{0,4} + \delta_{5,9})$ et $\pi'' = \frac{1}{2}(\delta_{0,9} + \delta_{5,4})$. Pour la valeur $p = 1/2$ chacun des $\pi \in \text{Marg}(\mu, \nu)$ donne le même coût global $C_p(\pi) = \lambda \frac{1}{2}(\sqrt{4} + \sqrt{4}) + (1 - \lambda) \frac{1}{2}(\sqrt{9} + \sqrt{1}) = 2$, ce qui témoigne de la non-unicité; $\text{Marg}(\mu, \nu) = \text{Marg}_{1/2}^*(\mu, \nu)$. Si maintenant $p < 1/2$ ou $p > 1/2$, la mesure π' ou π'' , respectivement, est l'unique plan de transport optimal. Ainsi la solution du problème dépend de la valeur de p .

Fixons donc $p < 1$. Le coût c_p étant continu, on peut comme dans le cas $p > 1$ disséquer les informations fournies par le corollaire 1.2.4 pour quatre points distincts. Nous en donnons ici une version spécifique au coût c_p avec $p < 1$:

Corollaire 1.4.2 (Lemme d'échange). Soit $\pi \in \text{Marg}_p^*(\mu, \nu)$ un plan de transport optimal pour le coût c_p avec $p \in]0, 1[$. Alors il existe S de mesure pleine tel que tout (x, y) et (x', y') pris dans S satisfait à

$$|y - x|^p + |y' - x'|^p \leq |y' - x|^p + |y - x'|^p. \quad (1.6)$$

Comme de plus c_p est un coût continu, tout S tel que $\bar{S} \subset \text{Spt}(\pi)$ et $\pi(S) = 1$ convient.

Le premier constat est classique et concerne le fait qu'autant de masse que possible restera sur place. Ceci se démontre assez simplement comme dans [GM96] où la démonstration est bâtie sur la comparaison de $|y - x|^p + |y' - x'|^p$ et $|y' - x|^p + |y - x'|^p$ lorsque $x = y$. Quels que soient les emplacements de x' et y' , pour peu que l'un des deux au moins est différent de $x = y$, on aura $|y - x|^p + |y' - x'|^p < |y' - x|^p + |y - x'|^p$. La réduction du problème à des mesures étrangères est illustrée sur le haut de la figure 1.3.

Lemme 1.4.3. Soit μ et ν dans $\mathcal{P}_1(\mathbb{R})$ et $\mu' = (\mu - \nu)_+ = \mu - (\mu \wedge \nu)$ et $\nu' = (\nu - \mu)_+ = \nu - (\mu \wedge \nu)$, les mesures μ et ν réduites de la partie commune à μ et ν . Pour $p \in]0, 1[$ tout plan de transport optimal $\pi \in \text{Marg}_p^*(\mu, \nu)$ s'écrit sous la forme $(\text{Id} \otimes \text{Id})_{\#}(\mu \wedge \nu) + \pi'$ où π' est un élément (optimal) de $\text{Marg}_p^*(\mu', \nu')$.

Réciproquement, si π est construit de cette manière il est optimal.

Exploitions maintenant dans le corollaire 1.4.2 le cas de quatre points distincts dont on connaît seulement les emplacements respectifs. En représentant les configurations des quatre points $(x, y) \in S$ et $(x', y') \in S$ par $abcd$ lorsque $a < b < c < d$ il apparaît que l'ordre $xx'y'y$ est préféré à $xx'yy'$ (qui ne peut donc pas être généré par S). De même pour $xyy'x'$, à la faveur de $xy'yx'$. Les symétries du problème nous permettent de nous épargner le détail de toutes les configurations : nous avons $(x, y; x', y') \leftrightarrow (x', y'; x, y)$ et $(x, y; x', y') \leftrightarrow (y, x; y', x')$ car

$c_p(x, y) = c_p(y, x)$. Nous nous contentons donc d'examiner les mots commençant par x . Avant de comparer $xyx'y'$ et $xy'x'y$ (les derniers mots à comparer), nous allons commenter une représentation graphique commode, celle de la figure 1.3. Celle-ci contient toutes les informations extraites jusqu'à présent du corollaire 1.4.2 et même davantage. Elle fait donc apparaître graphiquement que $xx'y'y$ est préféré à $xx'yy'$ et $xyy'x'$ l'est à $xy'yx'$. On représente μ et ν sur une unique droite réelle et les routes de transport $(x, y) \in S$ sont représentées par des arches orientées de x à y , au-dessus de la droite. Les interdictions que nous avons précisées nous indiquent exactement qu'il est impossible que deux arches se coupent. (On pourra supposer que les arches décrivent des demi-cercles afin d'empêcher que deux arches enchâssées ne se coupent.)

L'exemple 1.4.1 illustre le fait que la condition de non-intersection des arches ne suffit pas à déterminer le transport optimal. Nous allons donc ajouter une préférence entre les configurations $xyx'y'$ et $xy'x'y$ à la faveur de la première. En terme d'arche, cela se lira de la façon suivante : deux arches enchâssées sont orientées de la même manière. Une façon de faire apparaître cette préférence combinatoire, et, on le verra³, ainsi de contraindre entièrement le transport, est d'introduire un problème secondaire de coût c_p pour les solutions $\text{Marg}_1^*(\mu, \nu)$ du problème primaire de coût c_1 . Le lemme suivant se démontre en peu de lignes d'une manière très analogue au corollaire 1.2.4. Si la démonstration est plus délicate du point de vue technique, elle reste plus simple que celle du lemme 1.5.4. Conformément à nos notations, $\text{Marg}_{1,p}^{**} = (\text{Marg}_1^*)_p^*$ désigne l'ensemble de minimiseurs du problème de transport dit *secondaire* de coût c_p et de contrainte $M = \text{Marg}_1^*(\mu, \nu)$.

Lemme 1.4.4 (Lemme d'échange secondaire). *Soit $\pi \in \text{Marg}_{1,p}^{**}(\mu, \nu)$ où $p < 1$ et $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$. Alors, il existe un ensemble $S \subset \mathbb{R}^2$ avec $\pi(S) = 1$ tel que pour tout couples (x, y) et (x', y') de S nous avons*

$$|y - x| + |y' - x'| \leq |y - x'| + |y' - x| \quad (1.7)$$

et, dans le cas où $|y - x| + |y' - x'| = |y - x'| + |y' - x|$, également :

$$|y - x|^p + |y' - x'|^p \leq |y - x'|^p + |y' - x|^p. \quad (1.8)$$

Remarquons qu'avec (1.7) $xx'y'y$ et $xx'yy'$ donnent le même coût pour c_1 mais que le second est ensuite préféré au premier dans (1.8). Pour conclure la discussion suivant le lemme 1.4.3, les configurations $xx'y'y$ et $xyy'x'$ sont donc autorisées au détriment, respectivement, de $xx'yy'$ et $xy'yx'$. À celles-ci s'ajoute donc une troisième, à savoir $xyx'y'$ qui est préférée à $xy'x'y$.

Avant de présenter le transport montagne, expliquons encore le contexte mathématique qui a prévalu à son introduction. Dans [LMT14], Last, Mörters et Thorisson ont observé l'existence d'(une sorte de) plan de transport satisfaisant à une condition de stabilité. Cette condition se traduisant sur un ensemble S analogue à celui des conditions de monotonie il était naturel de s'interroger sur la possibilité de retrouver le transport semi-stable comme la solution d'un problème de transport adapté.

Remarque 1.4.5 (Stabilité). *La stabilité d'un plan de transport/mariage, au sens du problème des mariages (hétérosexuels) de Gale et Shapley [GS62] est la propriété d'existence d'un ensemble $S \subset \mathcal{X}^2$ de masse pleine $\pi(S) = 1$ tel que*

$$\min(c(x, y), c(x', y')) \leq \min(c(x, y'), c(x', y)) \quad (1.9)$$

pour tout $(x, y) \in S$ et $(x', y') \in S$. La signification en est que pour deux femmes x, x' et deux hommes y, y' les individus les plus « proches », où c mesure la « distance », sont mariés.

Si (\mathcal{X}, \leq) est ordonné, un transport π est stable à droite au sens de [LMT14, §7] si il est concentré sur $T := \{(x, y) \in \mathcal{X}^2 : x \leq y\}$ et qu'il existe un $S \subset T$ de mesure pleine pour π avec (1.9) dès que $\{(x, y), (x', y')\} \subset S$ et $\{(x, y'), (x', y)\} \subset T$.

On pourra vérifier que $\pi \in \text{Marg}_{1,p}^{**}$ n'est pas nécessairement stable pour $c = c_1$. Toutefois en vertu du lemme 1.4.4 si $\pi(S) = 1$ (c'est un transport croissant selon la définition 3.2.6) il est stable à droite.

3. ou, pour être prudent « on le suppose » puisque l'article est en cours d'écriture.

Remarque 1.4.6 (Mesures aléatoires). *La section 1.4 s'inspire de [LMT14] au sujet du problème de Skorohod pour un mouvement Brownien éternel $(B_t)_{t \in \mathbb{R}}$. On y recherche un temps aléatoire T nommé décalage non-biaisé tel que B_T possède à la fois la loi $\nu \in \mathcal{P}(\mathbb{R})$ demandée et soit aussi indépendante de $(B_{T+t} - B_T)_{t \in \mathbb{R}}$. Pour la résolution les auteurs transportent les unes sur les autres les mesures d'occupation de 0 et de ν . À peu de choses près, ce sont des mesures aléatoires sur \mathbb{R} invariantes par translation et donc de masse infinie transportées à l'aide d'« applications de transport infinies et aléatoires », nommées allocations, ou par des plans de transport du même type (induits par des noyaux de transport). Une solution minimale est exhibée [LMT14, §7] pour laquelle la masse n'est transportée que vers la droite (ce qui dans notre construction du transport montagne qui s'en inspire ne se produit que si $\mu \preceq_{sto} \nu$). La solution minimale jouit de plus de la propriété de stabilité à droite au sens de la remarque 1.4.5. Notons encore que le même type de problématique est repris dans [LTT16⁺].*

La formule qui gouverne ce transport de masse est celle du temps local d'un mouvement brownien et concerne les processus à variation quadratique finie. Comme l'ont constaté Bertoin et Yor [BY14], une analogie forte peut être établie avec une version non stochastique pour les fonctions à variation finie. C'est cette version, faisant intervenir l'indicatrice de Banach que nous exploitons pour définir le transport montagne.

Présentation du transport montagne

Nous donnons une présentation heuristique du transport montagne. On se référera à [Jui18⁺⁺], en cours d'écriture pour plus de détails. Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ càdlàg et à variation bornée, telle que $\lim_{x \rightarrow \pm\infty} F(x) = 0$. La fonction F correspond à une mesure signée $\sigma \in \mathcal{M}_0$ dont elle est la fonction de répartition. Les variations signées de F permettent donc de retrouver σ , ses variations absolues $|\sigma|$, ses variations positives σ_+ et ses variations négatives σ_- . Nous allons faire apparaître un plan de transport entre σ_+ et σ_- . L'appellation « transport montagne » est une référence sans subtilité au fait que le graphe de F peut faire songer à un profil montagneux dont les façades montantes et descendantes décrivent σ_+ et σ_- respectivement.

L'application F permet de transporter $|\sigma|$ de l'axe des abscisses à celui des ordonnées pour définir une mesure ζ dont nous aurons besoin. Toutefois $F_{\#}|\sigma| = \zeta$ est légèrement faux. Plutôt que de transporter les atomes $\alpha \delta_x$ de σ (avec $\alpha \in \mathbb{R}$) de façon déterministe en $F(x)$ on choisit de répartir leur masse uniformément sur tout l'intervalle $[F(x^-), F(x)] = [F(x^-), F(x^-) \pm \alpha]$. Soit donc ζ la mesure définie ainsi. Il est assez intuitif que ζ est absolument continu par rapport à la mesure de Lebesgue et admet pour densité une fonction i_F à valeur entière. Cette fonction, appelée parfois *indicatrice de Banach* indique en y_0 le nombre d'intersections entre le graphe de F et la droite horizontale d'équation $y = y_0$, autrement dit le nombre de solutions de l'équation $F = y_0$. Il convient de préciser que le graphe de F est complété par des parties verticales à chaque discontinuité, correspondant comme on le sait aux atomes de $|\sigma|$. La variation totale de F étant finie, il est facile d'admettre que presque sûrement, pour ζ , les intersections seront nettes : de cardinal fini et pair, alternant entre intersections montantes et descendantes (à commencer par les intersections montantes si $y > 0$, descendantes si $y < 0$).

Partant de l'observation suivante nous allons pouvoir définir le transport montagne de la façon illustrée sur la figure 1.3 : voyant ζ comme la loi d'une variable aléatoire Z (ce qui, quitte à normaliser, est possible), nous choisissons, conditionnellement à Z et uniformément une valeur $I \in \{1, \dots, i_F(Z)/2\}$ et retournons la i^e intersection montante avec le graphe (généralisé) de F . Cette procédure construit une variable aléatoire X de loi σ_+ (à normalisation près). On peut construire simultanément une variable aléatoire Y en procédant de la même manière avec le même Z et le même I mais en comptant cette fois les intersections descendantes. Nous simulons ainsi σ_+ et σ_- de façon couplée. Le transport $\pi \in \text{Marg}(\sigma_+, \sigma_-)$ résultant du couplage (X, Y) est notre définition du transport montagne entre σ_+ et σ_- . Un lemme généralisant le théorème des valeurs intermédiaires nous permet de vérifier les conditions sur $(x, y) = (X, Y)(\omega)$ et $(x', y') = (X, Y)(\omega')$: leurs arches ne s'intersectent pas et si elles sont enchâssées elles sont orientées dans le même sens.

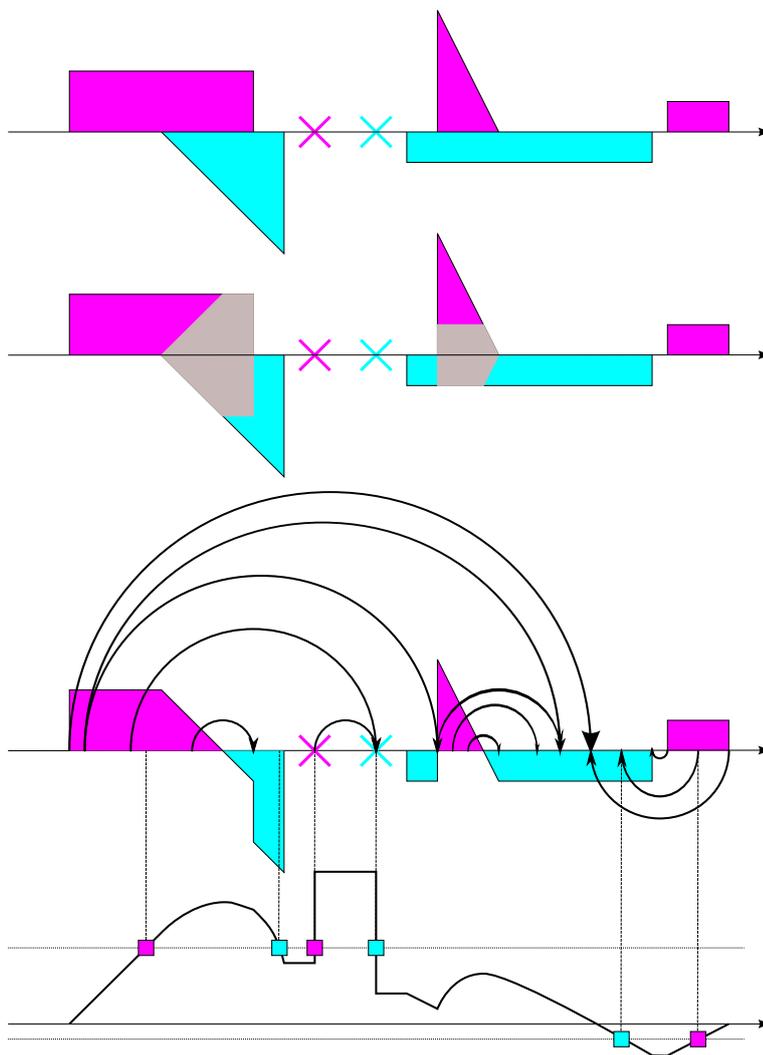


FIGURE 1.3 – Le transport montagne

Le théorème suivant est à paraître dans un article en préparation, la partie la plus délicate duquel est sans doute l'implication (iii) \Rightarrow (iv). Pour des mesures μ et ν générales, c'est-à-dire non nécessairement mutuellement singulières, le transport montagne est défini comme $(\text{Id} \otimes \text{Id})_{\#}(\mu \wedge \nu) + \pi'$ où π' est le transport montagne associé à $\sigma = \mu - \nu$ de la manière décrite ci-dessus.

Théorème 1.4.7. *Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R})$ et $\pi \in \text{Marg}(\mu, \nu)$. Soit de plus $p \in]0, 1[$. Les assertions suivantes sont équivalentes :*

- i (Solution du problème de transport L^{1^-} , première version) Il existe une suite $(\pi_n, p_n)_n \in \mathbb{N}$ avec, pour tout n , $p_n < 1$ et $\pi_n \in \text{Marg}_{p_n}^*(\mu, \nu)$, telle que $(\pi_n, p_n) \rightarrow_n (\pi, 1)$.*
- ii (Solution du problème de transport L^{1^-} , seconde version) π est un élément de $\text{Marg}_1^*(\mu, \nu)$ qui sur cet ensemble minimise $\gamma \in \text{Marg}_1^*(\mu, \nu) \mapsto \iint |y - x|^p d\pi(x, y)$.*
- iii (Monotonie) π is monotone au sens des arches (pas d'intersection et même orientation quand les arches sont enchâssées).*
- iv (Transport montagne) π est le transport montagne de μ vers ν .*

1.5 Le problème martingale et le transport rideau

Ici, nous traitons de $\mathcal{X} = \mathbb{R}$, $\mathcal{M} = \text{Mart}(\mu, \nu)$ et c vérifiant $\partial_{xy}c < 0$ (par exemple $c : (x, y) \mapsto (y - x)^3$). Introduisons toutefois $\text{Mart}(\mu, \nu)$ dans un cadre un peu plus général.

Définition 1.5.1. (a) Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R}^d)$. Un transport de $\text{Marg}(\mu, \nu)$ est un *transport martingale* si pour toute fonction mesurable positive bornée $f : \mathbb{R}^d \rightarrow \mathbb{R}$:

$$\iint f(x)(y - x)d\pi(x, y) = 0. \quad (1.10)$$

(b) Pour $d = 1$, π est un *transport sous-martingale* si l'intégrale dans (1.10) est positive pour toute fonction positive bornée $f : \mathbb{R} \rightarrow \mathbb{R}$.

Ces définitions se généralisent simplement aux mesures de $\mathcal{M}_{+,1}$ (plutôt que de \mathcal{P}_1 seulement). On notera $\text{Mart}(\mu, \nu)$ l'ensemble des transports martingale et $\text{SousMart}(\mu, \nu)$ celui des transports sous-martingale.

Nous présentons maintenant le *problème de transport martingale* que nous allons poursuivre dans les chapitres à venir et nous nous plaçons dans $\mathcal{P}_1(\mathbb{R})$. Alors que nous parlons de martingales, il semble indiqué de rappeler la formulation probabiliste du problème de transport et de faire le lien avec les martingales (pour rappel, voir la définition 3.1.1). Des lois marginales μ et ν , de probabilités, étant données il s'agit de trouver des variables aléatoires $X \sim \mu$ et $Y \sim \nu$ sur le même espace probabilisé (Ω, \mathbb{P}) telle que $C : \pi \mapsto \mathbb{E}(c(X, Y))$ atteigne son minimum, où $\pi := \text{Loi}(X, Y) \in \text{Marg}(\mu, \nu)$. Le calcul de cette espérance ne dépend que de π . Comme la contrainte *martingale* $\pi \in \text{Mart}(\mu, \nu)$ s'exprime aussi $\mathbb{E}(Y|X) = X$ on s'aperçoit qu'elle n'est vérifiée que si le processus $(X_i)_{i=1,2}$ à deux instants $X_1 = X$ et $X_2 = Y$ est une martingale dans sa propre filtration.

Suivant le formalisme de la définition 1.1.1, le problème de transport martingale consiste à minimiser C pour la contrainte $\mathcal{M} = \text{Mart}(\mu, \nu)$.

Nous ne nous intéresserons au problème martingale que sous l'hypothèse que $\text{Mart}(\mu, \nu)$ est non-vide. Notons que cette non-vacuité implique $\int \varphi d\mu \leq \int \varphi d\nu$ pour toute fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexe. En effet partant de $\text{Loi}(X, Y) \in \text{Mart}(\mu, \nu)$, nous trouvons à l'aide de l'inégalité de Jensen,

$$\int \varphi d\nu = \mathbb{E}(\varphi(Y)) = \mathbb{E}(\mathbb{E}(\varphi(Y)|X)) \geq \mathbb{E}(\varphi(X)) = \int \varphi d\mu,$$

ce, d'ailleurs, que ces quantités soit finies ou valent $+\infty$. Un cas particulier du théorème de Strassen (voir la proposition 2.1.4) nous permet d'affirmer qu'il s'agit en fait d'une condition nécessaire et suffisante (que les mesures soient de probabilité ou non). Elle se notera $\mu \preceq_C \nu$ et nous dirons que μ et ν sont dans l'ordre convexe.

Un lemme d'échange pour le problème martingale.

Nous présentons maintenant l'analogie de la notion de compétiteur introduite à la définition 1.2.1.

Définition 1.5.2. Soit $\alpha \in \mathcal{M}_+(\mathbb{R} \times \mathbb{R})$ une mesure dont le premier moment est fini pour la deuxième variable. On dit que α' est un *compétiteur martingale* de α si α' a les mêmes marges que α et pour presque tout $x \in \mathbb{R}$, par rapport à la mesure $(\text{proj}_{\#}^x \alpha)$, on trouve

$$\int y d\alpha_{x,\cdot}(y) = \int y d\alpha'_{x,\cdot}(y),$$

pour les familles $(\alpha_{x,\cdot})_{x \in \mathbb{R}}$ et $(\alpha'_{x,\cdot})_{x \in \mathbb{R}}$ de mesures de $\mathcal{P}_1(\mathbb{R})$ apparaissant dans la désintégration de α et α' conditionnellement à $\text{proj}_{\#}^x \alpha$.

Le théorème suivant est l'analogie du théorème fondamental de monotonie 1.2.3, celui pour lequel la démonstration n'est pas élémentaire.

Théorème 1.5.3 (Théorème de monotonie martingale). *Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R})$ prises dans l'ordre convexe et $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction borélienne intégrable pour les éléments de $\text{Mart}(\mu, \nu)$. Soit $\pi \in \text{Mart}^*(\mu, \nu)$ un plan de transport optimal.*

Alors, il existe un borélien S de mesure 1, pour π , qui vérifie ceci : si α est une mesure de $\mathbb{R} \times \mathbb{R}$ de support fini $|\text{Spt}(\alpha)| < \infty$ et $\text{Spt}(\alpha) \subseteq S$, alors on a $\iint c d\alpha \leq \iint c d\alpha'$ pour tout compétiteur α' de α .

Si, de plus, c est continue et bornée la réciproque est vérifiée : tout transport martingale concentré sur un borélien S finiment optimal (pour le problème martingale) est un transport optimal.

La démonstration de la partie directe [BeiJ16, Lemma 1.11], qui concerne les coûts boréliens, s'appuie sur une technologie sophistiquée qui passe par un résultat de Kellerer [Kel84] s'appuyant lui-même sur le théorème de capacibilité de Choquet. Le principe s'étend au problème martingale de dimension supérieure et est utilisé également dans [BGMS09, BCH17]. La partie réciproque s'appuie sur des résultats de dualité aménagés au problème martingale [BeiJ16, Lemma A.2] (prolongés dans [BNT17]). Toutefois, comme précédemment dans les sections 1.2, 1.3 et 1.4 il est possible de démontrer la partie directe seule avec des arguments élémentaires de théorie (géométrique) de la mesure. Le lemme d'échange suivant s'obtient ainsi de nouveau en prolongeant la technique de démonstration du corollaire 1.2.4.

Lemme 1.5.4 (Lemme d'échange, [BeiJ16, Theorem B.4]). *Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R})$ prises dans l'ordre convexe et $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction continue intégrable pour les éléments de $\text{Mart}(\mu, \nu)$. Soit $\pi \in \text{Mart}_c^*(\mu, \nu)$ un plan de transport optimal.*

Alors, il existe un borélien S de mesure 1, pour π , qui vérifie ceci : si α est une mesure de $\mathbb{R} \times \mathbb{R}$ de support fini $|\text{Spt}(\alpha)| < \infty$ et $\text{Spt}(\alpha) \subseteq S$, alors on a $\iint c d\alpha \leq \iint c d\alpha'$ pour tout compétiteur α' de α .

L'ensemble S peut être choisi sous une forme spécifique spécifiée dans [BeiJ16, Appendix B].

Le lemme 1.5.4 permet d'extraire de nombreuses informations sur les transports optimaux $\pi \in \text{Mart}^*(\mu, \nu)$, en particulier dans les cas suivants :

- i Si μ est diffuse et $c = c_4$ alors $\text{card}(\text{Spt } \pi_x) \leq 3$, presque sûrement pour μ . L'exemple décrit dans [BeiJ16, §7.3.1] illustre que la valeur 3 peut effectivement être atteinte presque sûrement.
- ii Plus généralement, si μ est diffuse et c est de la forme $(x, y) \rightarrow h(y - x)$ où $h \in \mathcal{C}^1(\mathbb{R})$ est telle que h' possède au maximum k points d'intersection avec les fonctions affines, alors $\text{card}(\text{Spt } \pi_{x,\cdot}) \leq k$, presque sûrement pour μ . ([BeiJ16, Theorem 7.1])
- iii If μ est diffuse et $c(x, y) = -|y - x|$, alors il existe un *unique* transport optimal $\pi \in \text{Mart}^*(\mu, \nu)$. De plus, $\text{card}(\text{Spt } \pi_{x,\cdot}) \leq 2$, presque sûrement pour μ (démontré en premier par Hobson et Neuberger [HN12] et dans [BeiJ16, Theorem 7.3] avec notre méthode)
- iv Si μ est diffuse et $c(x, y) = |y - x|$, alors il existe un *unique* transport optimal $\pi \in \text{Mart}^*(\mu, \nu)$. De plus, $\text{card}(\text{Spt } \pi_{x,\cdot}) \leq 3$ et $\text{card}(\text{Spt } \pi_{x,\cdot} \setminus \{x\}) \leq 2$, presque sûrement pour μ (voir Hobson et Klimmek [HK15] et [BeiJ16, Theorem 7.4])

Nous remarquons que les énoncés suivants donnent une idée précise de la forme des transports optimaux pour des coût variés, au point même d'obtenir l'unicité du transport optimal pour $c(x, y) = \pm|y - x|$. Toutefois, ces résultats sont obtenus sous l'hypothèse que μ est *diffuse*. En ce qui concerne le transport rideau, le transport optimal pour le coût $(x, y) \mapsto (y - x)^3$ (ainsi que pour d'autres coûts), qui nous accompagnera à la fin de ce chapitre et dans les chapitres suivants, cette hypothèse sera superflue.

Présentation du transport rideau

Nous appliquons le lemme d'échange 1.5.4 aux coûts vérifiant $\partial_{xy} c < 0$.

Lemme 1.5.5 (Lemme d'échange). Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R})$ prises dans l'ordre convexe et $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction continue intégrable pour les éléments de $\text{Mart}(\mu, \nu)$ telle que $\partial_{xy}c < 0$. Soit $\pi \in \text{Mart}_c^*(\mu, \nu)$ un plan de transport optimal.

Alors, il existe un borélien S de mesure 1, pour π , qui vérifie ceci : pour tout $(x, y^-), (x, y^+), (x', y') \in S$ la configuration suivante (voir la figure 1.4, à gauche) est impossible

$$x < x' \quad \text{et} \quad y^- < y' < y^+. \quad (1.11)$$

Remarque 1.5.6. Nous donnerons plus loin la définition 1.5.12, celle d'un ensemble S , construit de façon systématique à partir de π , qui convient.

Définition 1.5.7. Un plan de transport martingale π est dit *martingale-monotone (gauche)* si il existe un borélien $S \subseteq \mathbb{R} \times \mathbb{R}$ de masse 1 tel que pour tout $(x, y^-), (x, y^+), (x', y') \in S$ la configuration suivante est impossible

$$x < x' \quad \text{et} \quad y^- < y' < y^+. \quad (1.12)$$

De façon analogue π est *martingale-monotone droite* si il existe S tel que pour des éléments $(x, y^-), (x, y^+)$ et (x', y') de S on n'a pas

$$x > x' \quad \text{et} \quad y^- < y' < y^+.$$

On parlera pour S d'un *ensemble de monotonie martingale* de π .

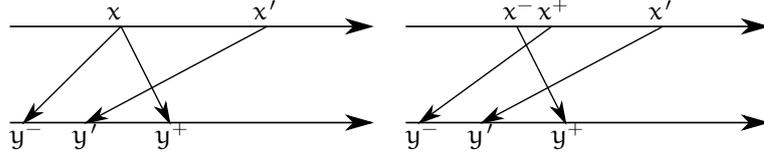


FIGURE 1.4 – Configurations de routes interdites d'après la proposition 1.5.13.

Voici l'équivalent martingale des théorèmes 1.2.13, 1.3.2 et 1.4.7.

Théorème 1.5.8 ([BeiJ16, Theorem 1.9]). Soit $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ des mesures dans l'ordre convexe et c un coût intégrable pour les éléments des $\text{Mart}(\mu, \nu)$ tel que $\partial_{xy}^3 c < 0$. Soit $\pi \in \text{Mart}(\mu, \nu)$ un plan de transport martingale. Les énoncés suivants sont équivalents

- Le transport π est optimal.
- Le transport π est martingale-monotone gauche.
- Le transport π est $\mathfrak{RG} = \mathfrak{RG}(\mu, \nu)$, le transport rideau (gauche), défini ainsi : pour tout $x \in \mathbb{R}$, la mesure $\nu_x^\pi = (\text{proj}^2)_\# \pi|_{[-\infty, x] \times \mathbb{R}}$ est, parmi les éléments de $S_x = \{\eta \in \mathcal{M}_+ : \mu|_{[-\infty, x]} \preceq_c \eta \text{ et } \eta \preceq_+ \nu\}$ celui qui est le plus petit pour l'ordre convexe.

Notre théorème suppose, de façon anticipée, que l'on peut toujours définir un couplage rideau gauche $\mathfrak{RG}(\mu, \nu)$ et que celui-ci est défini de manière univoque. Des détails quant à la définition et sur la façon de la généraliser seront donnés au cours du chapitre 2. Pour lors, nous rappelons que la signification de $(\text{proj}^2)_\# \pi|_{[-\infty, x] \times \mathbb{R}}$ a été commentée à la remarque 1.2.18 et faisons également remarquer que la définition de \mathfrak{RG} est du type de celle du couplage quantile, voir la remarque 1.2.19. La fin du chapitre est dédiée, d'une part, à une présentation des problèmes triviaux en transport martingale, d'autre part, à une définition alternative du couplage rideau gauche obtenue par un passage à la limite en s'appuyant sur le fait que c'est un transport martingale-monotone gauche.

Remarque 1.5.9. Dans le même cadre que celui du théorème 1.5.8, si μ est diffuse, nous pouvons extraire, comme dans les exemples suivant le lemme 1.5.4, une information concernant le cardinal de $\pi_{x,\cdot}$. Il existe un borélien $A \subset \mathbb{R}$ et $T_i : A \rightarrow \mathbb{R}$ pour $i \in \{1, 2\}$ des fonctions mesurables telles que

- le transport \mathfrak{RG} est concentré sur la réunion des graphes de T_1 et T_2 ,
- pour tout $x \in A$, $T_1(x) \leq x \leq T_2(x)$,
- Pour tout $x < x' \in \mathbb{R}$ on a $T_2(x) \leq T_2(x')$ et $T_1(x') \notin]T_1(x), T_2(x)[$.

Ceci est démontré dans [BeiJ16, Corollary 1.6].

Discussions sur le coût c dans le problème martingale

Fonctions de coût triviales Alors que c_2 est un des coûts les plus importants de la théorie du transport, il ne l'est plus dans la variante martingale. Supposons $\text{Loi}(X, Y) = \pi \in \text{Mart}(\mu, \nu)$ et admettons que μ et ν possèdent un second moment fini. Alors $\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY|X]] = \mathbb{E}[X^2]$, ce qui amène

$$\iint c_2(x, y) d\pi(x, y) = \iint (y - x)^2 d\pi(x, y) = \mathbb{E}[(Y - X)^2] = \mathbb{E}[Y^2] - \mathbb{E}[X^2].$$

Ceci est une constante sur $\text{Mart}(\mu, \nu)$, et donc $\text{Mart}(\mu, \nu) = \text{Mart}_2^*(\mu, \nu)$.

On peut en tirer la conséquence suivante : soit c une fonction de coût et

$$\tilde{c}(x, y) = c(x, y) + p \cdot (y - x)^2 + q \cdot (y - x)$$

où p et q sont des constantes réelles. Alors les solutions au problème sont les mêmes pour c et \tilde{c} . En particulier, si $c(x, y) = h(y - x)$, une hypothèse de monotonie ou de croissance sur h n'apporte rien.

Autres types de coûts triviaux Il convient aussi de constater, comme dans [BeiJ16, §7.3.1], que les coûts proposés ne sont pas du type $c(x, y) = h(|y - x|) = c(y, x)$. En fait pour $\mu = 1/2(\delta_{-1} + \delta_1)$ et $\nu = 1/3(\delta_{-2} + \delta_0 + \delta_2)$, tous les plans de transport donnent le même coût global à savoir $h(1) + (h(3) - h(1))/12$. Notons, pour le voir que $\text{Mart}(\mu, \nu)$ est un segment. Ces éléments sont les combinaisons linéaires de deux mesures, celle transportant paritairement la masse de -1 vers $\{-2, 0\}$ et celle le faisant pour 1 vers $\{0, 2\}$.

Exemples de coûts associés au transport rideau \mathfrak{RG} Quant bien même une grande partie de nos résultats sur le problème de transport martingale sont généraux, une partie concerne le couplage rideau, le transport optimal pour un coût spécifique. Or, celui-ci est obtenu comme solution de chacun des problèmes associés à des coûts c qui vérifient la condition du troisième ordre $\partial_{xy}^3 c < 0$. C'est le cas, par exemple, de $c : (x, y) \mapsto (y - x)^3$ ou de $c : (x, y) \mapsto \exp(y - x)$ qui sont des fonctions dans la classe $c : (x, y) \mapsto h(y - x)$ avec $h^{(3)} > 0$, ou bien de $(x, y) \mapsto \sqrt{1 + y^2} \frac{e^{-x}}{e^x + e^{-x}}$ qui est du type $\varphi(x)\psi(y)$ avec $\varphi' < 0$ et $\psi'' > 0$. Ces deux types de fonction de coût sont apparus dans [BeiJ16]. La généralisation du propos à $\partial_{xy}^3 c < 0$ est venue dans [HLT16]. C'est dès lors la condition retenue par les chercheurs du domaine, avec même une interprétation faible (si $x < x'$ alors $c(x, \cdot) - c(x', \cdot)$ est convexe) dans [NST17+] ou bien des conditions supplémentaires du second ordre, telles que $\partial_{xy}^2 c < 0$ dans le problème de transport surmartingale présenté dans [NS16]. Notons finalement que l'intégrabilité requise dans le théorème 1.5.8 n'est pas vérifiée pour toutes ces fonctions de manière équivalente. Pour le coût $c(x, y) = (y - x)^3$, une condition suffisante est par exemple $\nu \in \mathcal{P}_3(\mathbb{R})$.

Ensembles de monotonie martingale et stabilité du transport rideau \mathfrak{RG}

Une différence entre les lemmes d'échanges du transport optimal classique (corollaire 1.2.4), celui conduisant au transport montagne (lemme 1.4.4) et le lemme d'échange du transport martingale conduisant au transport rideau (lemmes 1.5.4 et 1.5.5) est le fait que l'ensemble de monotonie S ne soit pas aussi bien connu dans les trois cas. En particulier on sait que $S = \text{Spt}(\pi)$ est un choix valide dans le cas de référence. Voyons maintenant qu'il ne l'est pas dans les deux autres cas.

Exemple 1.5.10. Soit μ la mesure de probabilité uniforme sur $[-2, -1] \cup [0, 1]$ et ν celle uniforme sur $[-1, 0] \cup [1, 2]$. Il n'est alors pas très difficile de voir que le transport montagne, c'est-à-dire le transport optimal π auquel s'applique le lemme 1.4.4 est donné par l'application de transport

$$\Gamma = \begin{cases} -2 - x & \text{si } x \in [-2, -1] \\ 2 - x & \text{si } x \in [0, 1]. \end{cases}$$

Ainsi les routes $(x, y) = (-2, 0)$ et $(x', y') = (0, 2)$ sont-elles dans le support de π et, donc, $\text{Spt}(\pi)$ n'est pas un ensemble de monotonie au sens du lemme 1.4.4.

Exemple 1.5.11. [Jui16b, Exemple 2.11] Prenons pour μ la mesure uniforme sur $[-1, 1]$ et $\nu = \frac{\delta_{-1} + 2\delta_0 + \delta_1}{4}$. L'ensemble des transports martingale $\text{Mart}(\mu, \nu)$ ne contient alors qu'un seul élément π , nécessairement optimal pour tous les coût c . Celui-ci s'écrit $\pi = \mu \otimes \pi_{x, \cdot}$ avec

$$\pi_{x, \cdot} = \begin{cases} \frac{1}{2}(-x\delta_{-1} + (x-1)\delta_0) & \text{si } x \in [-1, 0[\\ \frac{1}{2}((1-x)\delta_0 + x\delta_1) & \text{si } x \in]0, 1]. \end{cases}$$

Ainsi $(x, y^-) = (0, -1)$ et (x, y^+) et $(x', y') = (1, 0)$ sont-elles trois routes de transport contenues dans $\text{Spt}(\pi)$. Elle contredisent les conclusions qui s'appliquent à S dans le lemme 1.5.5. Le support du couplage rideau n'est donc pas systématiquement un ensemble de monotonie martingale.

En réponse à l'exemple 1.5.11 il est possible de définir une notion de *support réduit* $\text{Spt}^*(\pi)$ qui soit un ensemble de monotonie.

Définition 1.5.12. Soit $\pi \in \mathcal{M}_+(\mathbb{R}^2)$ s'écrivant $\pi = \mu \otimes \pi_{x, \cdot}$. On introduit d'abord $A = \{x \in \mathbb{R} : \pi(]-\infty, x[\times]x, +\infty[) = 0$ et A^- le sous-ensemble des points isolés à droite dans A , c'est-à-dire $A^- = \{x \in A : \exists \varepsilon > 0, A \cap]x, x + \varepsilon[= \emptyset\}$. Alors le support réduit est défini par la formule

$$\text{Spt}^*(\pi) = (\text{Spt}(\pi) \setminus (A^- \times \mathbb{R})) \cup \bigcup_{\mu(x) > 0} \{x\} \times \text{Spt}(\pi_{x, \cdot}).$$

On peut donner pour exemple $\pi = (\delta_0 + \lambda)^{\otimes 2}$ dont le support est $[0, 1]^2$ mais qui a $\{(0, 0)\} \cup ([0, 1] \times [0, 1])$ pour support réduit.

Ce qui suit est la proposition 2.14 et le corollaire 2.15 de [Jui16b].

Proposition 1.5.13. Un plan de transport martingale $\pi \in \text{Mart}(\mu, \nu)$ est martingale-monotone (gauche) si et seulement si il vérifie

$$x < x' \implies y' \notin]y^-, y^+[. \quad (1.13)$$

pour tout triplet de routes (x', y') , (x, y^-) et (x, y^+) contenues dans $\text{Spt}^*(\pi)$.

De plus, un plan de transport martingale $\pi \in \text{Mart}(\mu, \nu)$ est martingale-monotone si et seulement si il vérifie

$$x^- \leq x^+ < x' \quad \text{et} \quad y^- \leq y^+ \implies y' \notin]y^-, y^+[. \quad (1.14)$$

pour tout triplet de routes (x', y') , (x^-, y^+) et (x^+, y^-) contenues dans $\text{Spt}^*(\pi)$.

Cette caractérisation est plus précise que le lemme 1.5.5 car elle décrit, avec le support réduit, un ensemble de monotonie martingale. Nous pouvons dès lors obtenir le résultat de stabilité suivant en nous basant sur la seconde forme du lemme (illustrée à droite de la figure 1.4).

Théorème 1.5.14 (Theorem 2.16 de [Jui16b]). Dans l'espace $\mathcal{D} := \{(\mu, \nu) \in \mathcal{P}_1(\mathbb{R})^2 : \mu \preceq_c \nu\}$ muni de la topologie faible, le transport rideau, vu comme l'application

$$(\mu, \nu) \in \mathcal{D} \mapsto \mathfrak{RG}(\mu, \nu)$$

est une application continue.

Le plan de la démonstration est assez simple et bien identifié puisqu'il suit un schéma classique (il correspond à [Vil09, Theorem 5.20]) où c'est le support $\text{Spt}(\pi)$ qui joue le rôle que nous ferons jouer au support réduit. Soit $(\mu_n, \nu_n) \in \mathcal{D}$ convergeant vers $(\mu, \nu) \in \mathcal{D}$ et soit π une valeur d'adhérence de la suite $(\pi_n)_{n \in \mathbb{N}}$ avec $\pi_n = \mathfrak{R}\mathfrak{G}(\mu_n, \nu_n)$. On va montrer que π est monotone gauche. En utilisant le théorème 1.5.8 c'est alors l'unique transport rideau $\mathfrak{R}\mathfrak{G}(\mu, \nu)$. On suppose par l'absurde qu'il existe un vecteur $\mathbf{v} = (x^-, y^-, x^+, y^+, x', y')$ qui à la fois est dans $(\text{Spt}^*(\pi))^3$ mais qui est aussi dans l'ensemble interdit

$$E = \{(x^-, y^+, x^-, y^+, x', y') \in \mathbb{R}^6 : x^+ \leq x^- < x' \text{ et } y^- < y' < y^+\}.$$

Comme $(\pi_n)_n$ tend vers π , il en va de même de $((\pi_n)^{\otimes 3})_n$ vers $\pi^{\otimes 3}$ (à une sous-suite près pour les deux suites). Comme $\pi_n(\text{Spt}^*(\pi_n)) = 1$ il existe donc une suite \mathbf{v}_n d'éléments de $(\text{Spt}^*(\pi_n))^3$ tendant vers \mathbf{v} . Il est alors suffisant de démontrer que $\mathbf{v}_n \in E$ pour n suffisamment grand. Si E était ouvert l'analyse serait rapide. il ne l'est pas mais le théorème se démontre tout de même rapidement. À noter que pour le transport quantile l'ensemble E serait $E = \{(x, y, x', y') \in \mathbb{R}^4 : x < x' \text{ et } y' < y\}$, un ensemble ouvert.

Une construction du couplage rideau $\mathfrak{R}\mathfrak{G}(\mu, \nu)$ entre μ et ν

Nous présentons maintenant une construction de $\mathfrak{R}\mathfrak{G}(\mu, \nu)$. Puisque celle-ci passe par le théorème 1.5.14 elle s'appuie aussi sur le théorème 1.5.8 et donc sur l'équivalence entre martingale-monotone gauche et rideau. Sans cette équivalence le principe de démonstration du théorème 1.5.14 permettrait seulement d'établir qu'il existe un transport martingale-monotone de marges μ et ν . Ainsi, notre construction ne peut elle pas logiquement se substituer à la définition du transport rideau.

Les étapes de la construction sont

- i Approximation de μ par une suite $(\mu_n)_{n \in \mathbb{N}^*}$ convergeant vers μ et telle que μ_n est atomique de support fini.
- ii Expression du couplage rideau $\mathfrak{R}\mathfrak{G}(\mu_n, \nu)$ entre une mesure de support fini μ_n et ν .
- iii Utilisation du théorème 1.5.14.

(i) On approche la mesure μ par une suite de mesures $(\mu_n)_{n \in \mathbb{N}^*}$ telle que μ_n est la somme de 2^n atomes de masse $1/2^n$. Pour ce faire on pose $\mu_n = \sum_{k=1}^{2^n} \frac{1}{2^n} \delta_{x_k}$ avec

$$x_k = 2^n \int_{k/2^n}^{(k+1)/2^n} G_\mu(t) d\lambda(t).$$

La fonction quantile associée à μ_n est constante sur chaque intervalle $]k/2^n, (k+1)/2^n]$; sa moyenne sur cet intervalle est celle de G_μ . Comme μ possède un premier moment, on reconnaît avec $(G_{\mu_n})_{n \in \mathbb{N}^*}$ une martingale L^1 fermée sur la filtration des intervalles dyadiques. D'après la proposition 1.2.15 on a $\|G_{\mu_n} - G_\mu\| = W_1(\mu_n, \mu)$, si bien que $\mu_n \rightarrow \mu$. Notons au passage que $\mu_m \preceq_C \mu_n \preceq_C \mu$ pour tout $m \leq n$.

(ii) Nous allons maintenant décrire le couplage rideau $\pi = \mathfrak{R}\mathfrak{G}(\mu_n, \nu)$ où μ_n est à support fini et $\mu_n \preceq_C \nu$. Nous notons $\mu_n = \sum_{k=1}^N \alpha_k \delta_{x_k}$ où $x_1 < \dots < x_N$. Les faits que $N = 2^n$ et $\alpha_k = 1/N$ ne sont pas utiles à l'étude. Nous oublierons donc cette information pour permettre une description plus générale. Nous pouvons écrire π sous la forme $\pi = \sum_{k=1}^N \delta_{x_k} \otimes \nu_k$ où $\sum_{k=1}^N \nu_k = \nu$. Notons que cette forme implique que $\pi_k := \delta_{x_k} \otimes \nu_k$ et ν_k ont pour masse α_k et que $\pi_k \in \text{Mart}(\alpha_k \delta_{x_k}, \nu_k)$. À partir de la propriété de martingale-monotonie il est facile de déduire que ν_1 est de la forme $(G_\nu)_\# \lambda_{]s_1, t_1[}$: aucune masse venant de x_k pour un $k > 1$ ne peut se glisser entre y_- et y_+ dans le support de ν_1 . Cette forme, $\nu_1 = (G_\nu)_\# \lambda_{]s_1, t_1[}$, laisse deux degrés de liberté qui sont ensuite fixés par les équations de masse et de barycentre, déterminant entièrement ν_1 . Il reste donc à déterminer $\pi - \pi_1 = \sum_{k=2}^N \delta_{x_k} \otimes \nu_k \in \text{Mart}(\mu - \alpha_1 \delta_{x_1}, \nu - \nu_1)$. Ce transport martingale est bien évidemment lui-même martingale-monotone. Il s'ensuit que l'on connaît ν_2 et ainsi de suite toutes les mesures ν_k jusqu'à ν_N . Pour résumer, la mesure ν_k est

la mesure de masse α_k qui s'écrit sous la forme $\nu_k = (G_{\nu - (\nu_1 + \dots + \nu_{k-1})} \# \lambda)_{s_k, t_k}$ et qui satisfait aux équations de barycentre (à savoir x_k) et de masse (à savoir α_k).

(iii) On utilise le théorème 1.5.14 pour caractériser $\mathfrak{R}\mathfrak{G}(\mu, \nu)$ comme la limite de la suite $(\mathfrak{R}\mathfrak{G}(\mu_n, \nu))_{n \in \mathbb{N}}$.

Plans de transport définis par les ombres

Ce chapitre est principalement construit à partir de certains résultats de [Jui16b] et [BeiJ16⁺] autour de la notion d'ombre $S^\nu(\mu)$ d'une mesure μ sur ν . Nous étudions en particulier l'aspect métrique de cette opération de projection, étendant le théorème 1.5.14, celui stipulant que l'application $(\mu, \nu) \mapsto \mathfrak{R}\mathfrak{G}(\mu, \nu)$ est continue. D'autre part, à l'aide des ombres, on rend le principe de construction du transport rideau plus explicite, et on donne naissance à la famille des *transports ombrés* avec en particulier, en plus $\mathfrak{R}\mathfrak{G}$ et $\mathfrak{R}\mathfrak{D}$, le transport rideau centré $\mathfrak{R}\mathfrak{C}$ et le transport soleil $\mathfrak{S}\mathfrak{ol}$.

2.1 Ordres stochastiques

Les ordres stochastiques jouent un rôle important dans ce mémoire. Nous avons déjà rencontré \preceq_{sto} et \preceq_{lo} à la définition 1.2.8, \preceq_+ aux pages 19, 22 et \preceq_{C} après la définition 1.5.1. Ceux définis maintenant seront des ordres composés des précédents. D'autres ordres seront considérés plus tard ; voir la définition 2.2.15 pour \preceq_{CD} .

Définition 2.1.1. La lettre E est une variable représentant un ensemble de fonctions à croissance au plus linéaire en $\pm\infty$. Soit, ainsi, E_+ l'ensemble des fonctions positives, E_{sto} celui des fonctions croissantes et E_{C} celui des fonctions convexes, tous trois, donc, restreints aux fonctions à croissance au plus linéaire. Pour $\mu, \nu \in \mathcal{M}_+$ nous posons $P(E)$ la propriété suivante :

$$P(E) : \quad \forall \phi \in E, \quad \int \phi \, d\mu \leq \int \phi \, d\nu.$$

Alors, pour $\mu, \nu \in \mathcal{M}_+$,

- Si $P(E_+)$ est vérifiée, nous noterons $\mu \preceq_+ \nu$ (ordre usuel),
- si $P(E_{\text{sto}})$ est vérifiée, nous noterons $\mu \preceq_{\text{sto}} \nu$ (ordre stochastique, 1^{er} ordre de domination de Choquet),
- si $P(E_{\text{C}})$ est vérifiée, nous noterons $\mu \preceq_{\text{C}} \nu$ (ordre convexe, 2^e ordre de domination de Choquet),
- si $P(E_{\text{C}} \cap E_+)$ est vérifiée, nous noterons $\mu \preceq_{\text{C},+} \nu$,
- si $P(E_+ \cap E_{\text{sto}})$ est vérifiée, nous noterons $\mu \preceq_{+, \text{sto}} \nu$,
- si $P(E_{\text{C}} \cap E_{\text{sto}})$ est vérifiée, nous noterons $\mu \preceq_{\text{C}, \text{sto}} \nu$,
- si $P(E_{\text{C}} \cap E_+ \cap E_{\text{sto}})$ est vérifiée, nous noterons $\mu \preceq_{\text{C},+, \text{sto}} \nu$.

Pour illustrer $\mu \preceq_{\text{C}} \nu$ nous donnons quelques exemples simples.

Exemple 2.1.2 (Exemples pour $\mu \preceq_{\text{C}} \nu$ [BeiJ16, §2.2]). — Soit δ un atome de masse $\mathbf{a} > 0$ concentré en $x : \delta = \mathbf{a}\delta_x$, alors $\delta \preceq_{\text{C}} \nu$ est vérifiée si et seulement si $\text{Masse}(\nu) = \mathbf{a}$ et $\text{Bary}(\nu) = x$.

- Si $\mu_i \preceq_{\text{C}} \nu_i$ pour $i = 1, \dots, n$, alors $\sum_{i=1}^n \mu_i \preceq_{\text{C}} \sum_{i=1}^n \nu_i$.
- Soit μ et ν de même barycentre et de même masse telles que μ est concentrée sur $[a, b]$ et ν sur $\mathbb{R} \setminus]a, b[$. Alors $\mu \preceq_{\text{C}} \nu$.
- Soit μ et ν de même barycentre et de même masse telles que $\mu - (\mu \wedge \nu)$ est concentrée sur $[a, b]$ et $\nu - (\mu \wedge \nu)$ est concentrée sur $\mathbb{R} \setminus]a, b[$. Alors $\mu \preceq_{\text{C}} \nu$.

L'ordre $\preceq_{\text{C},+}$ sera essentiel pour définir l'ombre $S^\nu(\mu)$.

Remarque 2.1.3 (Remarques au sujet de $\preceq_{\text{C},+}$). — Chacune des relations $\mu \preceq_{\text{C}} \nu$ et $\mu \preceq_+ \nu$ implique $\mu \preceq_{\text{C},+} \nu$.

- Sous la condition $\mu \preceq_{\text{C},+} \nu$, les mesures μ et ν n'ont, à priori, ni la même masse ni le même barycentre.

- Sous $\mu \preceq_{C,+} \nu$ on a $\text{Masse}(\mu) \leq \text{Masse}(\nu)$. De plus, dans le cas d'égalité où $\text{Masse}(\mu) = \text{Masse}(\nu)$, la relation $\mu \preceq_{C,+} \nu$ équivaut à $\mu \preceq_C \nu$.

Pour les ordres \preceq_C et \preceq_{sto} la comparaison entre μ et ν est équivalente à l'existence d'un plan de transport spécial. Ce sont des résultats attribués à Strassen bien que ces résultats soient en partie antérieurs et en fait bien plus généraux chez Strassen. Se basant sur [Str65, Théorème 7], l'article de Lindvall [Lin99] donne les détails sur la façon dont se démontrent les deux lignes concernant \preceq_C après la démonstration du Théorème 11 dans [Str65]. Comme nous le verrons plus tard à la remarque 3.1.4, [Str65, Theorem 8] concerne les martingales de dimension k indexées sur \mathbb{N} . On rappelle que nous avons introduit SousMart à la définition 1.5.1.

Proposition 2.1.4 (Attribués à Strassen d'après [Str65]). *Soit μ et ν deux mesures de $\mathcal{P}_1(\mathbb{R})$.*

- i On a $\mu \preceq_{sto} \nu$ si et seulement si il existe (X, Y) sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ tel que $\text{Loi}(X, Y) \in \text{Marg}(\mu, \nu)$ et $X \leq Y$, presque sûrement pour \mathbb{P} . Ceci vaut également pour $\mu, \nu \in \mathcal{P}(\mathbb{R}) \supset \mathcal{P}_1(\mathbb{R})$.*
- ii On a $\mu \preceq_C \nu$ si et seulement si il existe (X, Y) sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ tel que $\text{Loi}(X, Y) \in \text{Mart}(\mu, \nu)$.*
- iii On a $\mu \preceq_{C,sto} \nu$ si et seulement si il existe (X, Y) sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ tel que $\text{Loi}(X, Y) \in \text{SousMart}(\mu, \nu)$.*

On observe que si, dans (iii), on note $V = \mathbb{E}(Y|X)$, on se retrouve avec une mesure $\eta = \text{Loi}(V)$ telle que $\mu \preceq_{sto} \eta$ et $\eta \preceq_C \nu$ permettant ainsi une décomposition de $\preceq_{C,sto}$ en « \preceq_{sto} puis \preceq_C ». Mieux encore, le vecteur (X, V, Y) réalise cette décomposition sur un unique espace de probabilité. Le théorème suivant illustre le même comportement pour les ordres de la définition 2.1.1. Notons seulement que dans son énoncé, nous nous permettons de changer l'ordre des indices. L'ordre $\preceq_{+,sto,C}$ aura par exemple exactement la même signification que $\preceq_{C,+,sto}$.

Théorème 2.1.5 (Chaîne de relations, [Jui16b, Theorem 1.7]). *Les sept relations d'ordre apparaissant dans la définition 2.1.1 sont transitives et antisymétriques; ce sont des ordres partiels.*

De plus pour toute suite $(\mu_i)_{i=0,\dots,n}$ (avec $n = 2$ ou 3) vérifiant les relations $\mu_{i-1} \preceq_{r_i} \mu_i$ pour $i = 1, \dots, n$ on a $\mu_0 \preceq_{r_1, \dots, r_n} \mu_n$.

Réciproquement si $\mu_0 \preceq_{r_1, \dots, r_n} \mu_n$ on trouvera une suite $(\mu_i)_{i=0,\dots,n}$ telle que $\mu_{i-1} \preceq_{r_i} \mu_i$ pour tout $i \geq 1$.

Pour la suite, il est important de détailler le cas de $\preceq_{C,+}$.

Corollaire 2.1.6. *Soit μ, ν et η des mesures telles que $\mu \preceq_C \eta$ et $\eta \preceq_+ \nu$. Alors $\mu \preceq_{C,+} \nu$. Réciproquement, supposons $\mu \preceq_{C,+} \nu$. Alors, il existe une mesure η telle que $\mu \preceq_C \eta$ et $\eta \preceq_+ \nu$.*

Remarque 2.1.7. *Le théorème 2.1.5 s'applique à $\preceq_{C,+,sto}$ et se prolonge au couplage de variables aléatoires. Comme établi au corollaire 1.8 de [Jui16b], les quatres mesures $\mu = \mu_0 \preceq_+ \mu_1 \preceq_C \mu_2 \preceq_{sto} \mu_3 = \nu$ peuvent être couplées sur une même espace $(\Omega, \mathcal{F}, \mathbb{P})$ notamment à l'aide d'un ensemble $A \in \mathcal{F}$, utilisé pour restreindre la masse afin de réaliser $\mu_0 \preceq_+ \mu_1$.*

2.2 Définition et propriétés de l'ombre d'une mesure sur une autre

Fonctions potentiel

Nous introduisons les *fonctions potentiel* qui jouent, pour \preceq_C , un rôle proche de celui joué pour \preceq_{sto} par les fonctions de répartition. Une fonction potentiel U_μ , ou $U[\mu]$, est associée à toute mesure $\mu \in \mathcal{M}_{+,1}(\mathbb{R})$ de la façon suivante

$$U_\mu : x \in \mathbb{R} \mapsto \int_{-\infty}^{\infty} |y - x| d\mu(y).$$

Proposition 2.2.1. *Soit μ une mesure de $\mathcal{M}_{+,1}$ et $k = \text{Masse}(\mu)$, $m = \text{Bary}(\mu)$. Alors U_μ vérifie les propriétés suivantes :*

i U_μ est convexe,

ii $\lim_{x \rightarrow -\infty} U_\mu(x) - k|x - m| = 0$ et $\lim_{x \rightarrow +\infty} U_\mu(x) - k|x - m| = 0$.

Réciproquement, si une fonction f vérifie ces propriétés pour certains $m > 0$ et $k > 0$, alors il existe une mesure $\mu \in \mathcal{M}_+$ dont f est la fonction potentiel. La mesure μ est la moitié de f'' , la dérivé seconde de f , au sens des distributions.

Démonstration. On pourra consulter [HR12, Proposition 2.1]. \square

Nous listons ici les propriétés importantes des fonctions potentiel, en lien avec l'ordre convexe.

Proposition 2.2.2. *Soit μ et ν des éléments de \mathcal{M}_+ .*

i Si μ et ν ont même masse, alors $U[\mu] \leq U[\nu]$ équivaut à $\mu \preceq_C \nu$.

ii On a $\mu \preceq_+ \nu$ si et seulement si U_μ est moins courbé que U_ν . Plus précisément, $\mu \preceq_+ \nu$ équivaut à la convexité de $U_\nu - U_\mu$.

iii Une suite de mesures $(\mu_n)_n$ uniformément intégrables de masse k et barycentre m converge dans \mathcal{M}_+ vers une mesure $\mu \in \mathcal{M}_+$ si et seulement si $(U[\mu_n])_n$ converge simplement vers la fonction potentiel $U[\mu']$ d'une mesure $\mu' \in \mathcal{M}_+$. Dans ce cas-là, $\mu = \mu'$.

Démonstration. Voir [HPRY11, Exercice 1.7] pour (i), et [HR12, Proposition 2.3] pour (iii). La propriété (ii) se déduit de la proposition 2.2.1. En effet 2μ et 2ν sont les dérivées secondes de U_μ et U_ν . \square

Définition des ombres

La définition qui suit contient en fait un résultat : il faut reconnaître que l'ombre d'une mesure est correctement définie. Ceci est démontré dans [BeiJ16] au Lemma 4.6 où la démonstration s'appuie sur les propriétés des fonctions potentiel.

Définition 2.2.3 (Ombre d'une mesure sur une autre). Soit $\mu, \nu \in \mathcal{M}_+$ telles que $\mu \preceq_{C,+} \nu$. Alors, il existe une mesure $S^\nu(\mu) \in \mathcal{M}_+$, appelée l'ombre de μ sur ν , telle que

i $S^\nu(\mu) \preceq_+ \nu$.

ii $\mu \preceq_C S^\nu(\mu)$.

iii Si une mesure η satisfait aux conditions (i) et (ii) à la place de $S^\nu(\mu)$, alors on a $S^\nu(\mu) \preceq_C \eta$.

En conséquence de (iii), la mesure $S^\nu(\mu)$ est unique.

Avec cette définition nous pouvons reformuler la définition du couplage rideau $\pi = \mathfrak{R}\mathfrak{G}(\mu, \nu)$, donnée dans l'énoncé du théorème 1.5.8, de la façon suivante : la mesure π est l'unique mesure telle que pour tout $x \in \mathbb{R}$ la mesure $(\text{proj}^2)_\# \pi|_{[-\infty, x] \times \mathbb{R}}$ soit $S^\nu(\mu|_{[-\infty, x]})$. Il convient de remarquer que cette définition ne sera entièrement éclaircie que lorsqu'on aura vu que $\theta \preceq_+ \theta'$ implique $S^\nu(\theta) \preceq_+ S^\nu(\theta')$ (voir la remarque 1.2.19 à ce sujet). Ce sera le cas avec le théorème 2.2.7 sur l'associativité des ombres, dont c'est une conséquence.

De plus, il faut noter la similarité entre cette définition et celle de $\Omega(\mu, \nu)$ présentée à la page 26. Dans le cas de Ω les mesures $(\mu|_{[-\infty, x]})_{x \in \mathbb{R}}$ sont transportées sur la mesure de masse $\mu|_{[-\infty, x]}$ la plus à gauche sous ν , à savoir le minimum pour \preceq_{sto} de $S_x = \{\eta \in \mathcal{M}_+ : \eta \preceq_+ \nu \text{ et } \text{Masse}(\eta) = \mu|_{[-\infty, x]}\}$. Dans le cas de $\mathfrak{R}\mathfrak{G}$ les mêmes mesures sont transportées sur le minimum pour \preceq_C de $S'_x = \{\eta : \eta \preceq_+ \nu \text{ et } \mu|_{[-\infty, x]} \preceq_C \eta \text{ et } \text{Masse}(\eta) = \mu|_{[-\infty, x]}\}$ (où la troisième condition est en fait une conséquence de la seconde).

Premiers exemples de calculs d'ombres

Il est souvent difficile de déterminer quelle est l'ombre d'une mesure μ sur une autre ν . On y parvient parfois en observant qu'il s'agit d'une mesure η dont le barycentre est celui de μ , comme c'est le cas de toutes les mesures de $\{\eta : \mu \preceq_C \eta \preceq_+ \nu\}$, et qui minimise le second moment (correspondant à la variance pour la variable aléatoire canoniquement associée) parmi celles de cet ensemble. En effet, $\eta \preceq_C \eta'$ implique $\int x^2 d\eta \leq \int x^2 d\eta'$ avec, puisque $x \mapsto x^2$ est strictement convexe, égalité dans l'inégalité de Jensen si et seulement si $\eta = \eta'$ ou bien $\int x^2 d\eta = +\infty$. Suivant ce principe on peut aisément reconnaître, par exemple, l'ombre d'un atome.

Exemple 2.2.4 (Exemple trivial). *Si $\mu \preceq_+ \nu$, alors $S^\nu(\mu) = \nu$.*

Exemple 2.2.5 (Ombre d'un atome, Exemple 4.7 dans [BeiJ16]). *Soit δ un atome de masse \mathbf{a} concentré en $x : \delta = \mathbf{a}\delta_x$. Supposons $\delta \preceq_{C,+} \nu$. Alors $S^\nu(\delta)$ est la restriction de ν entre deux quantiles, c'est-à-dire $S^\nu(\delta) = f(\alpha_d, \alpha_u)$ où $f(\alpha_d, \alpha_u) = (G_\nu)_{\#\lambda_{[\alpha_d, \alpha_u]}}$ avec, pour déterminer (α_d, α_u) les équations $\text{Masse}(f(\alpha_d, \alpha_u)) = \alpha_u - \alpha_d = \mathbf{a}$, et $\text{Bary}(f(\alpha_d, \alpha_u)) = x$.*

Dans le même ordre d'idée on reconnaît facilement que $\delta \preceq_{C,+} \nu$ est réalisé si et seulement si x est encadré par les barycentres des mesures de masse \mathbf{a} les plus à gauche et les plus à droite sous ν (cette notion est introduite à la définition 1.2.16).

Exemple 2.2.6. *Soit μ et ν telles que μ est concentrée sur $[a, b]$ et ν sur $\mathbb{R} \setminus]a, b[$. On suppose $\mu \preceq_{C,+} \nu$. Alors $S^\nu(\mu) = f(\alpha_d, \alpha_u)$ où $f(\alpha_d, \alpha_u) = (G_\nu)_{\#\lambda_{[\alpha_d, \alpha_u]}}$ avec, pour déterminer (α_d, α_u) les équations $\text{Masse}(f(\alpha_d, \alpha_u)) = \alpha_u - \alpha_d = \text{Masse}(\mu)$ et $\text{Bary}(f(\alpha_d, \alpha_u)) = \text{Bary}(\mu)$.*

Associativité des ombres

Le théorème suivant est fondamental dans la compréhension des ombres. Ainsi que nous l'avons expliqué à la suite de la définition 2.2.3, en relation avec la remarque 1.2.19, c'est aussi l'élément manquant pour définir le transport rideau. Plus généralement l'implication « $\theta \preceq_+ \theta' \Rightarrow S^\nu(\theta) \preceq_+ S^\nu(\theta')$ » aura son importance dans la définition des transports ombrés.

Théorème 2.2.7 (Ombre d'une somme / associativité des ombres). *Soit γ_1, γ_2 et ν des éléments de \mathcal{M}_+ tels que $\mu = \gamma_1 + \gamma_2 \preceq_{C,+} \nu$. On a alors $\gamma_2 \preceq_{C,+} \nu - S^\nu(\gamma_1)$ et*

$$S^\nu(\gamma_1 + \gamma_2) = S^\nu(\gamma_1) + S^{\nu - S^\nu(\gamma_1)}(\gamma_2).$$

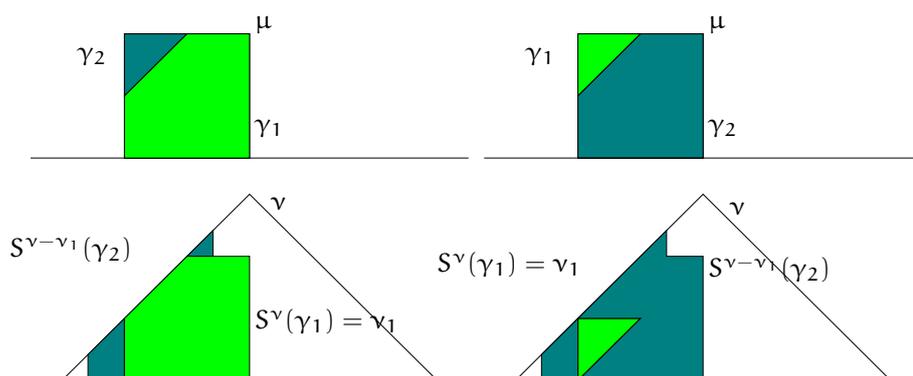
En particulier lorsque $\gamma_1 \preceq_+ \mu$ et $\mu \preceq_{C,+} \nu$, on a $S^\nu(\gamma_1) \preceq_+ S^\nu(\mu)$.

Dans [BeiJ16], la démonstration de ce théorème (voir Theorem 4.8) repose sur une série de lemmes pour lesquels γ_1 et γ_2 sont successivement des atomes, des sommes d'atomes où des mesures générales obtenues comme des limites de sommes d'atome (voir, entre autres, Lemma 4.12 et Lemma 4.16). Les démonstrations s'appuient sur les propriétés d'ordre liées aux fonctions quantile et sur des raisonnements élémentaires du type de ceux rencontrés dans les exemples précédents.

La figure 2.1 illustre le calcul visuel de l'ombre de $\mu = \gamma_1 + \gamma_2$ sur ν en jouant sur les ordres possibles dans la décomposition de μ . Comme $\gamma_1 \preceq_+ \nu$ il est bien clair que $S^\nu(\gamma_1) = \gamma_1$. Le calcul de la seconde ombre sur la partie gauche de la figure est assez simple (voir l'exemple 2.2.6). À partir du théorème 2.2.7 on déduit facilement, à droite, l'ombre de γ_2 sur $S^{\nu - S^\nu(\gamma_1)}$. C'est bien entendu la mesure $S^\nu(\mu) - S^\nu(\gamma_1)$ où $S^\nu(\mu)$ est connue par la construction à gauche.

De nouveaux exemples de calculs d'ombre résultent de l'emploi du théorème 2.2.7. L'exemple suivant enrichit l'exemple 2.2.6 et peut s'appliquer pour un calcul direct de l'ombre $S^{\nu - \nu_1}(\gamma_2)$ sur la partie droite de la figure 2.1 (avec $(\gamma_2, \nu - \nu_1)$ à la place de (μ, ν)).

Exemple 2.2.8. *Soit μ et ν telles que $\mu - (\mu \wedge \nu)$ est concentrée sur $[a, b]$ et $\nu - (\mu \wedge \nu)$ sur $\mathbb{R} \setminus]a, b[$. On suppose $\mu \preceq_{C,+} \nu$. Alors $S^\nu(\mu) = (\mu \wedge \nu) + f(\alpha_d, \alpha_u)$ où $f(\alpha_d, \alpha_u) = (G_{\nu - \nu \wedge \mu})_{\#\lambda_{[\alpha_d, \alpha_u]}}$ avec, pour déterminer (α_d, α_u) les équations $\text{Masse}(f(\alpha_d, \alpha_u)) = \alpha_u - \alpha_d = \text{Masse}(\mu - \mu \wedge \nu)$ et $\text{Bary}(f(\alpha_d, \alpha_u)) = \text{Bary}(\mu - \mu \wedge \nu)$.*

FIGURE 2.1 – Ombre de la somme $\mu = \gamma_1 + \gamma_2$ sur ν .

Exemple 2.2.9 (Ombre d'une somme d'atomes). *Pour une mesure μ à support fini, du type $\sum_{k=1}^N a_k \delta_{x_k}$ donc, et $\nu \succeq_{C,+} \mu$, l'ombre de μ sur ν peut se calculer de façon itérative en utilisant le théorème 2.2.7 pour $\mu_i = \sum_{k=1}^i a_k \delta_{x_k}$ et $a_{i+1} \delta_{x_{i+1}}$, respectivement dans le rôle de γ_1 et γ_2 , ainsi que l'exemple 2.2.5 au sujet de l'ombre d'un atome. On peut naturellement indexer les atomes dans l'ordre que l'on souhaite, mais arrêtons-nous sur le choix $x_1 < x_2 < \dots < x_N$. Dans ce cas on est en effet très près des N étapes relatives à la construction d'un transport martingale-monotone gauche expliquée à la fin du chapitre 1. Il n'y a pas de hasard à cela, les ombres sont le moyen de définir le transport rideau gauche \mathfrak{RG} , par un parcours de la gauche vers la droite, et \mathfrak{RG} est aussi l'unique transport martingale-monotone gauche.*

Le lemme suivant concerne la queue des ombres sur ν au-delà des limites du support de μ . On prêtera spécialement attention aux mesures μ dont le support se finit par un atome.

Lemme 2.2.10. *Soit μ, ν des éléments de \mathcal{M}_+ avec $\mu \preceq_{C,+} \nu$. Supposons que le support de μ est borné supérieurement et posons $y = \sup[\text{Spt } \mu] < +\infty$. Alors $S^\nu(\mu)$ est sur $[y, +\infty[$ la mesure la plus à gauche sous ν parmi celles de même masse. Plus précisément, la restriction de $(S^\nu(\mu) - \mu)_+$ à $[y, +\infty[$ est une mesure θ , la plus petite mesure pour l'ordre stochastique parmi celles de même masse satisfaisant à $\theta \preceq_+ (\nu - \mu)_+ \llbracket [y, +\infty[$.*

L'énoncé équivalent est naturellement valable si le support de μ est borné inférieurement.

Utilisation des ombres pour définir le transport rideau centré

Nous allons définir un transport martingale particulièrement simple, le transport rideau centré $\mathfrak{RC}(\mu, \nu)$. Toutefois, la définition n'est vraiment simple que lorsque μ et ν sont dans l'ordre diatomique $\mu \preceq_{CD} \nu$, un ordre que nous allons définir et qui implique $\mu \preceq_C \nu$. Plus loin, nous étendrons cette définition à tous les couples $\mu \preceq_C \nu$ dans le cadre des transports ombrés. Pour simplifier la présentation, nous supposons que μ et ν sont des mesures de probabilité.

Nouveau paramétrage du couplage quantile et du couplage rideau gauche

Nous avons indiqué un parallèle entre $\Omega(\mu, \nu)$ et $\mathfrak{RG}(\mu, \nu)$ (voir les propositions 1.2.17 et théorème 1.5.8). Ces transports se présentent en effet comme une loi jointe $\pi = \text{Loi}(X, Y)$ définie à partir de la famille $\text{Loi}(Y|X \leq x)_{x \in \mathbb{R}}$. Nous revisitons ici ce principe en paramétrant désormais μ à l'aide des niveaux de quantile $\alpha \in [0, 1]$ plutôt qu'avec $x \in \mathbb{R}$. Notons tout d'abord

$$\mu = \int_0^1 \delta_{G_\mu(\alpha)} d\alpha, \quad \nu = \int_0^1 \delta_{G_\nu(\alpha)} d\alpha, \quad \Omega(\mu, \nu) = \int_0^1 \delta_{G_\mu(\alpha)} \otimes \delta_{G_\nu(\alpha)} d\alpha. \quad (2.1)$$

La famille $(\gamma_u)_{u \in [0,1]}$ avec $\gamma_u := \int_0^u \delta_{G_\mu(\alpha)} \otimes \delta_{G_\nu(\alpha)} d\alpha$ est croissante pour \preceq_+ dans $\mathcal{M}_+(\mathbb{R}^2)$. Pour $u = F_\mu(x)$ la mesure γ_u possède pour marges $\int_0^{F_\mu(x)} \delta_{G_\mu(\alpha)} d\alpha = \mu \llbracket]-\infty, x]$ et $(G_\nu)_\# \lambda_{[0, F_\mu(x)]}$,

les mesures apparaissant dans la proposition 1.2.17 au sujet d'une caractérisation de $\mathfrak{Q}(\mu, \nu)$. Voici, sans démonstration, une nouvelle caractérisation qui servira de modèle à celle du transport rideau centré.

Lemme 2.2.11. *Pour μ et ν , il existe une unique mesure $\pi \in \text{Marg}(\mu, \nu)$ pour laquelle on peut trouver une famille croissante (pour \preceq_+) de mesures $(\gamma_u)_{u \in]0,1]}$ tendant vers $\pi = \gamma_1$ avec $\gamma_u \in \text{Marg}(\int_0^u \delta_{G_\mu(\alpha)}, \int_0^u \delta_{G_\nu(\alpha)})$. Les deux marges de γ_u sont les mesures de masse u les plus à gauche, sous μ et ν respectivement. Cette mesure est le transport quantile $\mathfrak{Q}(\mu, \nu)$.*

De même, voici une caractérisation du transport rideau gauche :

Lemme 2.2.12. *Soit $\mu \preceq_C \nu$. Il existe une unique mesure π et une unique famille $(\gamma_u)_{u \in]0,1]}$ tendant vers π avec $\gamma_u \in \text{Mart}(\int_0^u \delta_{G_\mu(\alpha)}, S^\nu(\int_0^u \delta_{G_\nu(\alpha)}))$, telle que $u \mapsto \gamma_u$ soit croissante pour \preceq_+ . C'est le transport rideau gauche $\mathfrak{RG}(\mu, \nu)$.*

De nouveau, cette caractérisation est légèrement différente de celle du théorème 1.5.8. À cet endroit on ne considère que les mesures $\int_0^{F_\mu(x)} \delta_{G_\mu(\alpha)} = \mu_{] -\infty, x]}$ plutôt que toutes les mesures $\int_0^u \delta_{G_\mu(\alpha)}$.

Remarque 2.2.13. *Le transport rideau centré répond au schéma de construction que nous venons d'exposer dans les lemmes 2.2.11 et 2.2.12. Toutefois, l'unicité de la caractérisation nécessitera une analyse approfondie qui sera faite en section 2.3 quand nous définirons les transports ombrés. Ce cadre passe par des transports paramétrés $\hat{\pi} \in \widehat{\text{Marg}}(\hat{\mu}, \hat{\nu})$. La mesure γ_u y apparaît sous l'écriture $\hat{\pi}_{[0, u]}$. Par rapport aux deux lemmes, $\delta_{G_\mu(u)}$ et $\int_0^u \delta_{G_\mu(\alpha)} d\alpha$ seraient dans ce cadre respectivement notés $\hat{\mu}_u$ et $\hat{\mu}_{[0, u]}$.*

Le transport rideau centré

Soit $\mu \preceq_C \nu$. Nous supposons toujours qu'il s'agit de probabilités et notons $\mathfrak{m} = \text{Bary}(\mu) = \text{Bary}(\nu)$. Alors les deux équations de gauche dans (2.1) peuvent être remplacées par

$$\mu = \int_0^1 \theta[f^\mu(\alpha), g^\mu(\alpha)] d\alpha \quad \text{et} \quad \nu = \int_0^1 \theta[f^\nu(\alpha), g^\nu(\alpha)] d\alpha \quad (2.2)$$

où $\theta[a, b]$ est la seule mesure de probabilité concentrée sur $\{a, b\}$ dont le barycentre est \mathfrak{m} . Les fonctions f^μ, g^μ vont être définies dans quelques lignes. On pourra vérifier, par exemple à l'aide du théorème de Douglas [Dou64], que les mesures du type $\theta[a, b]$ sont les points extrêmes de l'ensemble des mesures de masse 1 et de barycentre \mathfrak{m} , tout comme dans (2.1) on avait les points extrêmes de l'espace des mesures de probabilité. L'équation (2.2) constitue une représentation de Choquet, tout comme c'était le cas de (2.1). En la matière il n'y a pas d'unicité comme le montre l'exemple suivant

Exemple 2.2.14. *Supposons $\mathfrak{m} = 0$ si bien que $\theta[a, b] = \frac{b}{b-a} \delta_a + \frac{-a}{b-a} \delta_b$. Voici alors deux décompositions différentes de la même mesure :*

$$(2/3)\theta[-1, 1] + (1/3)\theta[-2, 2] = (1/2)\theta[-1, 2] + (1/2)\theta[-2, 1].$$

Toutefois on peut supposer qu'à gauche dans (2.2) les fonctions f^μ et g^μ sont monotones, donnant ainsi un moyen de faire émerger une décomposition de Choquet canonique. Nous définissons pour ce faire les fonctions f^μ, g^μ (et f^ν, g^ν) de la façon suivante :

$$f^\mu : \alpha \mapsto \max(\text{Spt}(S^\mu(\alpha \delta_{\mathfrak{m}}))) \quad \text{et} \quad g^\mu : \alpha \mapsto \min(\text{Spt}(S^\mu(\alpha \delta_{\mathfrak{m}}))).$$

On utilisera aussi les notations $f[\mu]$ et $g[\mu]$ pour les mêmes quantités.

Les propriétés des ombres, en particulier l'ombre d'un atome (exemple 2.2.5) nous indiquent que

- la mesure $S^\mu(\alpha\delta_m)$ est de masse α et barycentre m .
- elle est du type $(G_\mu)_\# \lambda_{[\alpha_d, \alpha_u]}$ c'est-à-dire du type $\mu_{[\cdot]f^\mu(\alpha), g^\mu(\alpha)[\cdot] + \alpha\delta_{f^\mu(\alpha)} + \beta\delta_{g^\mu(\alpha)}$ (où $\alpha, \beta \geq 0$ sont nuls lorsqu'il n'y a pas d'atome en $f^\mu(\alpha)$ ou $g^\mu(\alpha)$).
- l'intervalle $[\alpha_d, \alpha_u]$ est un intervalle croissant pour \subset en α
- la fonction g^μ est continue à gauche et croissante,
- la fonction f^μ est continue à gauche et décroissante.

Le cas d'une relation spéciale μ et ν doit maintenant être exposé.

Définition 2.2.15 (Ordre \preceq_{CD}). Soit μ et ν deux mesures de probabilité. On dit que μ et ν sont dans l'ordre convexe diatomique, et note $\mu \preceq_{CD} \nu$ lorsqu'une des assertions suivantes est vérifiée :

- Pour tout $\alpha \in]0, 1]$, $[f^\mu(\alpha), g^\mu(\alpha)] \subset [f^\nu(\alpha), g^\nu(\alpha)]$,
- pour tout $\alpha \in]0, 1]$, $\theta[f^\mu(\alpha), g^\mu(\alpha)] \preceq_C \theta[f^\nu(\alpha), g^\nu(\alpha)]$,
- pour tout $\alpha \in]0, 1]$, $S^\mu(\alpha\delta_m) \preceq_C S^\nu(\alpha\delta_m)$,
- pour tout $\alpha \in]0, 1]$, $S^\nu(S^\mu(\alpha\delta_m)) = S^\nu(\alpha\delta_m)$.

De plus ces assertions sont équivalentes.

En prenant $\alpha = 1$ dans la troisième assertion on trouve que $\mu \preceq_{CD} \nu$ implique $\mu \preceq_C \nu$.

Nous sommes maintenant en mesure de définir le transport rideau centré $\mathfrak{RC}(\mu, \nu)$ pour $\mu \preceq_C \nu$ et même de le décrire sous l'hypothèse supplémentaire $\mu \preceq_{CD} \nu$. C'est cependant seulement en section 2.3 que nous pourrons vérifier la consistance de cette définition.

Définition 2.2.16 (Transport rideau centré). Soit μ et ν dans l'ordre convexe. Le transport rideau centré $\mathfrak{RC}(\mu, \nu)$ est l'unique mesure qui est la limite en 1 d'une suite $(\gamma_u)_{u \in [0, 1]}$, croissante pour \preceq_+ dans $\mathcal{M}_+(\mathbb{R}^2)$, de mesures de masse $u \in]0, 1]$, telle que γ_u est un élément de $\text{Mart}(S^\mu(u\delta_m), S^\nu(S^\mu(u\delta_m)))$.

Si de plus $\mu \preceq_{CD} \nu$, on a

$$\mathfrak{RC}(\mu, \nu) = \int_0^1 \theta[f^\mu(\alpha), g^\mu(\alpha); f^\nu(\alpha), g^\nu(\alpha)] d\alpha.$$

où $\theta[a, b; c, d]$ est l'unique élément de $\text{Mart}(\theta[a, b], \theta[c, d])$ (non vide si $\theta[a, b] \preceq_C \theta[c, d]$, ce qui est garanti ici). L'expression en est

$$\frac{b-m}{b-a} \left(\frac{d-a}{d-c} \delta_{(a,c)} + \frac{a-c}{d-c} \delta_{(a,d)} \right) + \frac{m-a}{b-a} \left(\frac{d-b}{d-c} \delta_{(b,c)} + \frac{b-c}{d-c} \delta_{(b,d)} \right).$$

Stabilité métrique des ombres

Nous allons ici établir le théorème 2.2.17 qui est un résultat quantifié de continuité de l'application $\mu \mapsto S^\nu(\mu)$. Pour ce faire nous citerons un à un les éléments nécessaires pour aboutir, en page 50, à la démonstration du théorème. Ce théorème est important pour la construction des transports ombrés qui viendra ensuite. Nous l'avons établi dans [Jui16b] ainsi que ses prolongements, le théorème 2.2.25 et son corollaire [Jui16b, Corollary 2.32]. Ce dernier est une version *quantitative* de la continuité de $(\mu, \nu) \mapsto \mathfrak{RC}(\mu, \nu)$, déjà obtenue au théorème 1.5.14 en s'appuyant sur la propriété de monotonie gauche.

Théorème 2.2.17 (Stabilité des ombres). Soit $\mu, \mu' \in \mathcal{M}_+$ deux mesures de même masse. Supposons que $\nu \in \mathcal{M}_+$ vérifie $\mu \preceq_{C,+} \nu$ et que, de même, $\mu' \preceq_{C,+} \nu$. Alors,

$$W_1(S^\nu(\mu), S^\nu(\mu')) \leq W_1(\mu, \mu'). \quad (2.3)$$

Pour ce lemme non trivial, il faut plus que la compréhension attentive de l'exemple 2.2.9. Donnons les ingrédients menant à la démonstration page 50.

Lemme 2.2.18 ([Jui16b, Lemma 2.33]). *Soit n un entier et μ, μ' deux mesures obtenues comme la somme de n atomes de même masse. Supposons de plus $\mu \preceq_{sto} \mu'$. Alors, si $\nu \in \mathcal{M}_+$ vérifie $\mu \preceq_{C,+} \nu$ et $\mu' \preceq_{C,+} \nu$ on trouve encore l'ordre stochastique pour les ombres : $S^\nu(\mu) \preceq_{sto} S^\nu(\mu')$.*

Définition 2.2.19. Soit $\mu, \nu \in \mathcal{M}_+$ des mesures de même masse, $m > 0$. On définit les minimums et maximums stochastiques de μ et ν de la façon suivante :

$$\max_{sto}(\mu, \nu) = \max(G_\mu, G_\nu) \# \lambda_{[0, m]} \quad \text{et} \quad \min_{sto}(\mu, \nu) = \min(G_\mu, G_\nu) \# \lambda_{[0, m]}.$$

Les exemples suivants résultent de ce que $(G_\mu, G_\nu) \sim \Omega(\mu, \nu)$.

Exemple 2.2.20. *i) Si μ et ν sont des probabilités et $(X, Y) \sim \Omega(\mu, \nu)$, alors $\max_{sto}(\mu, \nu)$ and $\min_{sto}(\mu, \nu)$ sont les lois de $\max(X, Y)$ et $\min(X, Y)$.*

ii) Si $\mu = \sum_{i=1}^n \delta_{x_i}$ et $\nu = \sum_{i=1}^n \delta_{y_i}$ où les suites $(x_i)_i, (y_i)_i$ sont croissantes, alors $\max_{sto}(\mu, \nu) = \sum_{i=1}^n \delta_{\max(x_i, y_i)}$ et $\min_{sto}(\mu, \nu) = \sum_{i=1}^n \delta_{\min(x_i, y_i)}$.

Le lemme suivant résulte simplement de la proposition 1.2.15.

Lemme 2.2.21. *Soit μ et ν des éléments de $\mathcal{P}_1(\mathbb{R})$. On a*

$$\begin{aligned} W_1(\mu, \nu) &= W_1(\mu, \min_{sto}(\mu, \nu)) + W_1(\min_{sto}(\mu, \nu), \nu) \\ &= W_1(\mu, \max_{sto}(\mu, \nu)) + W_1(\max_{sto}(\mu, \nu), \nu). \end{aligned}$$

Lemme 2.2.22 (Ajout de masse en $\pm\infty$). *Soit des mesures μ, ν satisfaisant à $\mu \preceq_{C,+} \nu$. Pour tout $a \in \mathbb{R}$ il existe ν' où ν' est concentrée sur $] -\infty, a]$ et $\mu \preceq_{C,+} \nu + \nu'$. De plus, sans aucune hypothèse sur μ , pour tout $a, b \in \mathbb{R}$, il existe ν' avec $\nu'([a, b]) = 0$ et $\mu \preceq_{C,+} \nu'$.*

Lemme 2.2.23 (Rôle joué par la portion de ν qui est proche de $\pm\infty$). *Soit μ et ν des éléments de \mathcal{M}_+ tels que $\mu \preceq_{C,+} \nu$. Soit $(\nu_n)_n$ telle que $\inf(\text{Spt } \nu_n) \rightarrow_n +\infty$. Alors, la suite $S^{\nu+\nu_n}(\mu)$ tend vers $S^\nu(\mu) \in \mathcal{M}_+$.*

Les hypothèses $\sup(\text{Spt } \nu_n) \rightarrow_n -\infty$ ou $\nu_n([a_n, b_n]) = 0$ avec $-a_n, b_n \rightarrow_n +\infty$ conduisent à la même conclusion.

Proposition 2.2.24. *Supposons que $(\mu_n)_n$ est croissante dans l'ordre convexe avec $\mu_n \preceq_C \mu \preceq_{C,+} \nu$ pour tout $n \in \mathbb{N}$. Alors les suites $(\mu_n)_n$ et $(S^\nu(\mu_n))_n$ convergent toutes les deux dans (\mathcal{M}_+, W_1) . Si μ_∞ et S_∞ désignent leurs limites, la mesure S_∞ est alors l'ombre de μ_∞ sur ν : $S^\nu(\lim_n \mu_n) = \lim_n S^\nu(\mu_n)$.*

Démonstration du théorème 2.2.17. 1. On suppose tout d'abord que μ et μ' sont des sommes d'atomes de même masse. On suppose également $\min_{sto}(\mu, \mu') \preceq_{C,+} \nu$ et on note $\tilde{\mu} = \min_{sto}(\mu, \mu')$. La mesure $\tilde{\mu}$ est du même type que μ et μ' . On applique maintenant Lemma 2.2.18 séparément aux couples $(\mu, \tilde{\mu})$ et $(\mu', \tilde{\mu})$. À l'aide de la proposition 1.2.15 et du lemme 2.2.21 on trouve ainsi

$$\begin{aligned} W_1(\mu, \mu') &= W_1(\mu, \tilde{\mu}) + W_1(\mu', \tilde{\mu}) \\ &= W_1(S^\nu(\mu), S^\nu(\tilde{\mu})) + W_1(S^\nu(\mu'), S^\nu(\tilde{\mu})) \\ &\geq W_1(S^\nu(\mu), S^\nu(\mu')). \end{aligned}$$

2. Nous allons maintenant nous défaire de l'hypothèse $\tilde{\mu} \preceq_{C,+} \nu$ faite au-dessus. On a toujours $\tilde{\mu} \preceq_{C,+} \nu$ si bien qu'il existe une suite $(\nu_n)_n$ comme dans le lemme 2.2.22, c'est-à-dire telle $\tilde{\mu} \preceq_{C,+} \nu + \nu_n$ et $\sup \text{Spt } \nu_n$ tend vers $-\infty$. Le calcul fait en (1.) reste valable et donne $W(S^{\nu+\nu_n}(\mu), S^{\nu+\nu_n}(\mu')) \leq W(\mu, \mu')$. Finalement à l'aide du lemme 2.2.23, on obtient (2.3).

3. La seconde hypothèse à lever est celle sur le type, atomique, des mesures μ et ν . Pour ce faire on approche les mesures par des suites de terme $\mu_n = \sum_{k=1}^{2^n} \frac{1}{2^n} \delta_{x_k}$ et $\mu'_n = \sum_{k=1}^{2^n} \frac{1}{2^n} \delta_{x'_k}$ telles que celle construite à la page 41. Pour ces mesures on a, avec (2.) l'inégalité $W(S^\nu(\mu_n), S^\nu(\mu'_n)) \leq W(\mu_n, \mu'_n)$. Puisque les suites sont croissantes pour \preceq_C et convergent vers μ et μ' , on peut finalement obtenir (2.3) avec la proposition 2.2.24 en passant à la limite. \square

Le théorème suivant est une continuation (démontrée au Theorem 2.31 de [Jui16b]) du théorème 2.2.17. Une majoration de la distance entre $\mathfrak{R}\mathfrak{G}(\mu, \nu)$ et $\mathfrak{R}\mathfrak{G}(\mu', \nu')$ est également possible en introduisant une (semi-)distance ad hoc Z (voir [Jui16b, Definition 2.28 et Corollary 2.32] pour cet énoncé).

Théorème 2.2.25. *Soit μ, μ', ν, ν' des mesures de \mathcal{M}_+ . On suppose $\mu \preceq_{C,+} \nu$ et $\mu' \preceq_{C,+} \nu'$ ainsi que $\text{Masse}(\mu) = \text{Masse}(\mu')$ et $\text{Masse}(\nu) = \text{Masse}(\nu')$. On a alors $W_1(S^\nu(\mu), S^{\nu'}(\mu')) \leq W_1(\mu, \mu') + 2W_1(\nu, \nu')$.*

2.3 Définition des couplages ombrés

Le principe de la remarque 1.2.19 a valeur de définition pour Ω et $\mathfrak{R}\mathfrak{G}$: un couplage (X, Y) est entièrement défini par la famille de mesures $(\text{Loi}(Y|X \leq x))_{x \in \mathbb{R}}$. Informellement, cela permet de définir $\pi_{x,\cdot} = \text{Loi}(Y|X = x)$ pour presque tout $x \in \mathbb{R}$ (pour μ) et ainsi de retrouver π . En fait, les mesures $\pi_{]-\infty, x], \cdot} = \mathbb{P}(X \leq x) \text{Loi}(Y|X \leq x)$ sont des mesures croissantes pour \preceq_+ dont la masse est $\mathbb{P}(X \leq x)$ et on peut assimiler $\pi_{x,\cdot}$ au taux d'accroissement limite $h^{-1}(\pi_{]-\infty, x], \cdot} - \pi_{]-\infty, x-h], \cdot})$ lorsque h tend vers zéro.

Le même principe fonctionne pour le couplage rideau droit $\mathfrak{R}\mathfrak{D}$ en utilisant $(\text{Loi}(Y|X \geq x))_{x \in \mathbb{R}}$ et $(\text{Loi}(Y|X \geq x))_{x \in \mathbb{R}}$. On semble donc libre de paramétrer μ de la façon que l'on souhaite. Cette intuition est à l'origine des *couplages ombrés*. Pour ces derniers on parcourt μ de façon croissante pour \preceq_+ à l'aide d'un *paramétrage* $(\hat{\mu}_{[0, \alpha], \cdot})_{\alpha \in [0, 1]}$ de limite $\mu = \hat{\mu}_{[0, 1], \cdot}$, et on y associe l'ombre de ces mesures sur ν , constituant ainsi un deuxième paramétrage, de ν cette fois, $(\hat{\nu}_{[0, \alpha], \cdot})_{\alpha \in [0, 1]}$. Il s'avère, ce qui n'est pas évident, que ces deux paramétrages amènent une unique mesure $\pi = \hat{\pi}_{[0, 1], \cdot, \cdot}$, elle-même paramétrée par $(\hat{\pi}_{[0, \alpha], \cdot, \cdot})_{\alpha \in [0, 1]}$ où $\hat{\pi}_{[0, \alpha], \cdot, \cdot}$ a pour marges $\hat{\mu}_{[0, \alpha], \cdot}$ et $\hat{\nu}_{[0, \alpha], \cdot}$. Ces notations paraissent certainement barbares au lecteur ; le système conduisant à ce choix est expliqué à la suite de cette introduction.

Revisitons maintenant le transport rideau par notre nouvelle approche, donnant de la consistance à notre pré-approche de la page 47. On considère pour tout $\alpha \in [0, 1]$ la mesure la plus à gauche (voir la définition 1.2.16) de masse α sous μ que nous notons $\hat{\mu}_{[0, \alpha], \cdot}$. Nous introduisons ensuite $\hat{\nu}_{[0, \alpha], \cdot} = S^\nu(\hat{\mu}_{[0, \alpha], \cdot})$. On trouve alors une unique famille croissante $(\hat{\pi}_{[0, \alpha], \cdot, \cdot})_{\alpha \in [0, 1]}$ avec $\hat{\pi}_{[0, \alpha], \cdot, \cdot} \in \text{Mart}(\hat{\mu}_{[0, \alpha], \cdot}, \hat{\nu}_{[0, \alpha], \cdot})$. La mesure $\hat{\pi}_{[0, 1], \cdot, \cdot}$ est alors le transport rideau gauche $\mathfrak{R}\mathfrak{G}(\mu, \nu)$. Notons qu'avec $\hat{\mu}_{[0, \alpha], \cdot} = S^\mu(\alpha \delta_m)$ on retrouve, à la place, la construction du transport rideau centré comme à la définition 2.2.16. Un des paramétrages naturels considérés plus tard est $\hat{\mu}_{[0, \alpha], \cdot} = \alpha \mu$ qui donne naissance au *couplage soleil* (voir la définition 2.3.6).

Les avantages de travailler avec α plutôt qu'avec x sont les suivants : on connaît la masse des mesures, c'est α , et on n'a pas besoin de traiter les atomes différemment des autres points, on peut envisager de nouveaux paramétrages. Notre nouveau paradigme est donc : « $\text{Loi}(U, X, Y)$ est, sous certaines conditions, entièrement définie par les familles $\text{Loi}(X|U \leq \alpha)$ et $\text{Loi}(Y|U \leq \alpha)$ ».

Fonctionnement des notations On continuera à considérer μ et ν comme les marges de $\pi \in \text{Marg}(\mu, \nu)$. La plupart du temps il s'agira d'un transport martingale mais nous aurons aussi à l'esprit le transport quantile comme modèle. Par ailleurs μ, ν et π seront la plupart du temps des mesures de probabilité. Les mesures accentuées $\hat{\mu}, \hat{\nu}$ et $\hat{\pi}$ signaleront respectivement des éléments de $\text{Marg}(\lambda, \mu), \text{Marg}(\lambda, \nu)$ et $\text{Marg}(\lambda, \pi)$ où λ est la mesure de Lebesgue restreinte à $[0, 1]$. Ainsi $\hat{\pi}$ est elle une mesure sur $\mathbb{R} \times \mathbb{R}^2$.

On utilisera en principe le triplet de lettres $(\alpha, x, y) \in [0, 1] \times \mathbb{R} \times \mathbb{R}$ avec parfois u à la place de α . Rappelons, comme premier exemple que $(\pi_{x,\cdot})_{x \in \mathbb{R}}$ est la famille de mesures sur \mathbb{R} correspondant à la désintégration $\pi(dx dy) = \mu(dx) \pi_{x,\cdot}(dy)$. De façon analogue on a $(\hat{\mu}_{\alpha,\cdot})_{\alpha \in [0, 1]}$, $(\hat{\mu}_{\cdot, x})_{x \in \mathbb{R}}$ et $(\hat{\nu}_{\cdot, y})_{y \in \mathbb{R}}$ avec $\hat{\mu}(d\alpha dx) = \lambda(d\alpha) \hat{\mu}_{\alpha,\cdot}(dx)$, $\hat{\mu}(d\alpha dx) = \mu(dx) \hat{\mu}_{\cdot, x}(d\alpha)$ et $\hat{\nu}(d\alpha dy) = \nu(dy) \hat{\nu}_{\cdot, y}(d\alpha)$. On utilise aussi $(\hat{\pi}_{\alpha,\cdot,\cdot})_{\alpha \in [0, 1]}$ et $(\hat{\pi}_{\alpha, x,\cdot})_{\alpha, x \in [0, 1] \times \mathbb{R}}$ pour la désintégration $\hat{\pi}(d\alpha dx dy) = \hat{\mu}(d\alpha dx) \hat{\pi}_{\alpha, x,\cdot}(dy) = \lambda(d\alpha) \hat{\pi}_{\alpha,\cdot,\cdot}(dx dy)$.

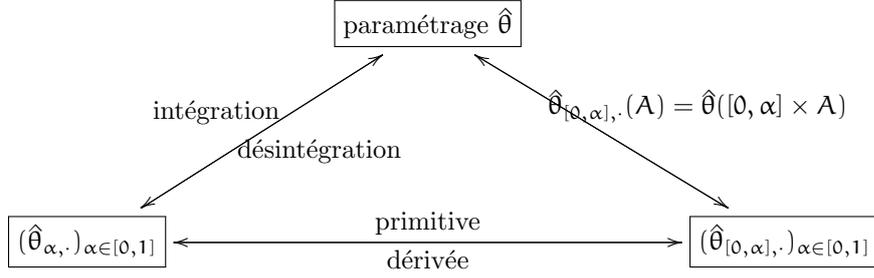


FIGURE 2.2 – Triangle de désintégration

Nous avons également une notation pour les mesures où le paramètre de conditionnement est partiellement intégré, donnant lieu à des « marginales partielles ». Des exemples en sont $\hat{\pi}_{[0,\alpha], \cdot, \cdot}$ et $\hat{\mu}_{[0,\alpha], \cdot, \cdot}$ qui sont des mesures de masse α . Pour $\alpha = 1$ on retrouve $\pi = \hat{\pi}_{[0,1], \cdot, \cdot}$ et $\mu = \hat{\mu}_{[0,1], \cdot, \cdot}$.

On appelle *paramétrage en pseudo-quantiles* de μ tout élément de $\hat{\mu} \in \text{Marg}(\lambda, \mu)$. L'ensemble des paramétrages en pseudo-quantile des transports de $\text{Mart}(\mu, \nu)$ respectant le paramétrage en pseudo-quantile $\hat{\mu}$ de μ est alors

$$\widehat{\text{Mart}}(\hat{\mu}, \nu) = \left\{ \hat{\pi} \in \text{Marg}(\hat{\mu}, \nu) : \int y \, d\hat{\pi}_{\alpha, x, \cdot} = x, \text{ pour } \hat{\mu}\text{-presque tout } (\alpha, x) \right\}.$$

Le triangle de la figure 2.2 indique les trois manières équivalentes de paramétrer θ , où θ représente l'une des mesures μ, ν ou π . Si $\theta = \pi$ alors $A \subset \mathbb{R}^2$, sinon $A \subset \mathbb{R}$.

Le théorème des couplages ombrés

Le théorème suivant étend le théorème 1.5.8, sur le transport rideau.

Théorème 2.3.1. *Soit $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ des mesures de probabilité dans l'ordre convexe. On se donne de plus une paramétrage en pseudo-quantiles $\hat{\mu} \in \text{Marg}(\lambda, \mu)$ de μ . Alors il existe un unique transport paramétré $\hat{\pi} \in \widehat{\text{Mart}}(\hat{\mu}, \nu) \subset \text{Marg}(\lambda, \mu, \nu)$ vérifiant chacune des propriétés suivantes (qui le caractérisent donc chacune)*

i La mesure $\hat{\pi}$ minimise sur $\widehat{\text{Mart}}(\hat{\mu}, \nu)$ la fonction

$$\hat{\gamma} \mapsto \iiint_{[0,1] \times \mathbb{R} \times \mathbb{R}} (1 - \alpha) \sqrt{1 + y^2} \, d\hat{\gamma}(\alpha, x, y).$$

ii Il existe un ensemble borélien $\hat{S} \subset \mathbb{R}^3$, tel que $\hat{\pi}(\hat{S}) = 1$, qui vérifie la contrainte suivante : si $\alpha, \alpha', x, x', y^-, y^+$ et y' sont des réels tels que $\alpha < \alpha'$ et $(\alpha, x, y^-), (\alpha, x, y^+), (\alpha', x', y')$ sont dans \hat{S} , alors y' n'appartient pas à $]y^-, y^+[$. Cela peut s'écrire de façon synthétique :

$$\alpha < \alpha', (\alpha, x, y^-), (\alpha, x, y^+), (\alpha', x', y') \in \hat{S} \Rightarrow y' \notin]y^-, y^+[.$$

iii Pour tout $\alpha \in [0, 1]$ la mesure $\hat{\pi}_{[0,\alpha], \cdot, \cdot}$ a pour première marge $\hat{\mu}_{[0,\alpha], \cdot}$ et pour seconde marge $S^\nu(\hat{\mu}_{[0,\alpha], \cdot})$, l'ombre de la première marge sur ν .

Une démonstration de ce théorème peut se faire en suivant assez fidèlement celle du théorème 1.5.8. Toutefois un passage délicat concerne le fait que dans le point (iii) il n'est pas établi que les deux familles $(\hat{\mu}_{[0,\alpha], \cdot})_{\alpha \in [0,1]}$ et $(\hat{\nu}_{[0,\alpha], \cdot})_{\alpha \in [0,1]}$ où $\hat{\nu}_{[0,\alpha], \cdot} = S^\nu(\hat{\mu}_{[0,\alpha], \cdot})$ permettent de définir totalement une unique famille $(\hat{\pi}_{[0,\alpha], \cdot, \cdot})_{\alpha \in [0,1]}$ (et donc $\hat{\pi}$). La stratégie de notre démonstration consiste à remplacer les trois « courbes » de mesures $(\hat{\pi}_{[0,\alpha], \cdot, \cdot})_{\alpha \in [0,1]}$, $(\hat{\mu}_{[0,\alpha], \cdot, \cdot})_{\alpha \in [0,1]}$ et $(\hat{\nu}_{[0,\alpha], \cdot, \cdot})_{\alpha \in [0,1]}$ de masse α par leurs dérivées dans (\mathcal{M}_+, W_1) à savoir les trois familles

$(\hat{\pi}_{\alpha,\cdot})_{\alpha \in [0,1]}$, $(\hat{\mu}_{\alpha,\cdot})_{\alpha \in [0,1]}$ et $(\hat{\nu}_{\alpha,\cdot})_{\alpha \in [0,1]}$, qui sont des familles de mesures de probabilités avec $\hat{\pi}_{\alpha,\cdot} \in \text{Mart}(\hat{\mu}_{\alpha,\cdot}, \hat{\nu}_{\alpha,\cdot})$. C'est à cet endroit que l'équation $\hat{\nu}_{[0,\alpha],\cdot} = S^\nu(\hat{\mu}_{[0,\alpha],\cdot})$ se transforme et donne

$$\hat{\nu}_{\alpha,\cdot} = P^{\text{Spt}(\nu - \hat{\nu}_{[0,\alpha],\cdot})}(\hat{\mu}_{\alpha,\cdot}), \quad (2.4)$$

où P est la projection de Kellerer, une notion présentée à la définition 2.3.2. Nous détaillerons ensuite au théorème 2.3.3 les étapes permettant le passage à l'équation (2.4). Insistons seulement encore sur le fait que (2.4) détermine $(\hat{\pi}_{\alpha,\cdot})_{\alpha \in [0,1]}$, $(\hat{\pi}_{[0,\alpha],\cdot})_{\alpha \in [0,1]}$, et donc finalement $\hat{\pi}$.

Soit $\mathcal{F}(\mathbb{R})$ l'espace des fermés de \mathbb{R} , et \mathcal{I} le sous-espace des ensembles $T \in \mathcal{F}(\mathbb{R})$ tels que $\sup T = -\inf T = +\infty$. L'espace $\mathcal{F}(\mathbb{R})$ est muni de la topologie la plus grossière telle que $T \in \mathcal{F}(\mathbb{R}) \mapsto d(x, T) \in [0, \infty]$ est continue. D'après [Kel73, Satz 13] l'espace est ainsi compact et métrisable, et d'après [Kel73, Satz 14], \mathcal{I} est un sous-ensemble de type G_δ dans $\mathcal{F}(\mathbb{R})$.

Définition 2.3.2 (Projection et dilatation de Kellerer, voir la figure 2.3). Soit T un élément de \mathcal{I} . Pour tout $x \in T$, soit $x_T^- = \sup(T \cap (-\infty, x])$ et $x_T^+ = \inf(T \cap [x, +\infty))$. Nous définissons tout d'abord un noyau, la dilatation¹ de Kellerer ([Kel73, Definition 16]) :

$$p_T(x, \cdot) = \begin{cases} \delta_x & \text{si } x \in T, \\ (x_T^+ - x_T^-)^{-1} [(x_T^+ - x)\delta_{x_T^-} + (x - x_T^-)\delta_{x_T^+}] & \text{sinon.} \end{cases}$$

Ainsi pour $\mu \in \mathcal{M}_+$, la projection de μ sur T est $\nu = \mu \cdot p_T$ et le transport associé entre μ et ν est $\pi = \mu \cdot (\text{id}, p_T)$.

Par similarité avec S^ν , la notation des ombres, on notera P^T la projection de Kellerer, à savoir l'opérateur défini par $P^T(\mu) = \mu \cdot p_T$.

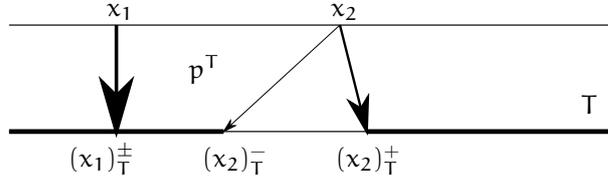


FIGURE 2.3 – La dilatation de Kellerer.

Notons que si $T \notin \mathcal{I}$ mais que $\text{Spt}(\mu) \subset [\inf T, \sup T]$ il est possible de préserver le sens de $P^T(\mu)$ en remplaçant T par $T^* =]-\infty, \inf T] \cup \bar{T} \cup [\sup T, +\infty[$. Cette possibilité est toujours offerte par $T = \text{Spt}(\nu)$ lorsque $\mu \preceq_{C,+} \nu$.

Conformément à la discussion précédente, on s'intéresse aux dérivées de $\alpha \mapsto \hat{\theta}_{[0,\alpha],\cdot}$.

Théorème 2.3.3. Soit $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ telles que $\mu \preceq_C \nu$ et soit $\hat{\mu} \in \text{Marg}(\lambda, \mu)$ un paramétrage en pseudo-quantiles de μ . Supposons que $\alpha \mapsto \hat{\mu}_{[0,\alpha],\cdot}$ admette une dérivée à droite $\hat{\mu}_{\alpha_0,\cdot}$ en α_0 et soit $\hat{\nu}_{[0,\alpha],\cdot}$ l'ombre $S^\nu(\hat{\mu}_{[0,\alpha],\cdot})$ de $\hat{\mu}_{[0,\alpha],\cdot}$ sur ν . Alors, $\alpha \mapsto (\hat{\nu}_{[0,\alpha],\cdot})$ possède une dérivée à droite en α_0 . Cette dérivée est donnée par $\hat{\mu}_{\alpha_0,\cdot} \cdot p_T$ où T est le support de $\hat{\nu}_{[\alpha_0,1],\cdot} := \nu - \hat{\nu}_{[0,\alpha_0],\cdot}$. La projection $P^T(\hat{\mu}_{\alpha_0,\cdot})$ est par ailleurs bien définie car $\text{Spt}(\hat{\mu}_{\alpha_0,\cdot}) \subset [\inf T, \sup T]$.

Démonstration. À l'aide du théorème d'associativité 2.2.7 on peut travailler l'expression du taux d'accroissement. Ainsi

$$\begin{aligned} \sigma_h &= h^{-1}(\hat{\nu}_{[0,\alpha_0+h],\cdot} - \hat{\nu}_{[0,\alpha_0],\cdot}) = h^{-1}(S^\nu(\hat{\mu}_{[0,\alpha_0+h],\cdot}) - S^\nu(\hat{\mu}_{[0,\alpha_0],\cdot})) \\ &= h^{-1}S^{\hat{\nu}_{[\alpha_0,1],\cdot}}(\hat{\mu}_{[\alpha_0,\alpha_0+h],\cdot}) = S^{h^{-1}\hat{\nu}_{[\alpha_0,1],\cdot}}(h^{-1}\hat{\mu}_{[\alpha_0,\alpha_0+h],\cdot}) \end{aligned}$$

1. Une dilatation est un noyau de martingale. Les définitions concernant les noyaux sont rappelées au début du chapitre 3.

où on a posé $\hat{\mu}_{[\alpha, \beta], \cdot} = \hat{\mu}_{[0, \beta], \cdot} - \hat{\mu}_{[0, \alpha], \cdot}$. Or $h^{-1}\hat{\mu}_{[\alpha_0, \alpha_0+h], \cdot}$ tend vers $\hat{\mu}_{\alpha_0, \cdot}$ lorsque $h \downarrow 0$. À la limite nous sommes « moralement » presque en train de considérer l'ombre de $\hat{\mu}_{\alpha_0, \cdot}$ sur une mesure infinie $\infty \cdot \hat{\nu}_{[u_0, 1], \cdot}$, faisant apparaître la projection de Kellerer. Techniquement, cela se règle avec le lemme suivant.

Lemme 2.3.4. *Soit $(\eta_n)_n$ une suite de $\mathcal{P}_1(\mathbb{R})$ et $(H_n)_n$ une suite de réels positifs. Soit de plus ν une mesure. Supposons $\eta_n \rightarrow \eta$ dans $(\mathcal{P}_1(\mathbb{R}), W_1)$ et $H_n \rightarrow \infty$. Supposons de plus que pour tout $n \geq 1$, $\eta_n \preceq_{C,+} H_n \nu$. Alors, $S^{H_n \nu}(\eta_n) \rightarrow P^T(\eta)$ avec $T = \text{Spt}(\nu)$ et $\text{Spt}(\eta) \subset [\inf T, \sup T]$.*

Ce lemme est démontré dans [BeiJ16⁺]. La démonstration étudie successivement les cas $\eta_n = \delta_x$, $\eta_n = \sum_{k=1}^N a_k \delta_{x_k}$ (en utilisant le théorème d'associativité 2.2.7), puis $\eta_n = \eta$ une mesure constante (en passant à la limite comme à la fin du chapitre 1) pour finir par le cas général. Le théorème de stabilité des ombres 2.2.17 est employé aux deux dernières étapes. \square

Représentation graphique et problème de Skorohod

À la figure 2.4, nous avons représenté trois couplages ombrés, correspondant aux mêmes mesures μ et ν mais à des paramétrages $\hat{\mu}$ différents (voir à la page 55). Puisque le mécanisme repose sur trois coordonnées (α, x, y) il bien faut comprendre comment celles-ci se distribuent sur deux axes : le premier axe correspond aux pseudo-quantiles $\alpha \in [0, 1]$; le second axe est là pour x ainsi que pour y . La mesure en vert est $\hat{\mu} \in \text{Marg}(\lambda, \mu)$. Sa projection sur le second axe est μ . Pour $\alpha \in [0, 1]$ fixé, la mesure $\hat{\mu}_{\alpha, \cdot}$ se lit comme la mesure de dimension 1 à l'intersection de la droite verticale d'abscisse α et la tâche verte représentant $\hat{\mu}$. On sait, avec le théorème 2.3.3, que cette mesure est projetée sur $\hat{\nu}_{\alpha, \cdot} = P^{T_\alpha}(\hat{\mu}_{\alpha, \cdot})$ à l'aide de la projection de Kellerer sur $T_\alpha = \text{Spt}(\nu - \hat{\nu}_{[0, \alpha], \cdot})$. Ici T_α est représenté par l'intersection entre la droite verticale et la région hachurée par les lignes sinueuses. C'est ainsi qu'on reconnaît $\hat{\nu}$ comme la projection de $\hat{\nu}$ sur l'ensemble en question, que l'on nomme *barrière*, en référence, comme on le verra, au problème de Skorohod. Finalement on retrouve ν en projetant sur le second axe. Si on souhaite comprendre directement la loi jointe $\pi \in \text{Mart}(\mu, \nu)$ on peut s'intéresser aux noyaux $\pi_{x, \cdot}$ dont la lecture consiste d'abord à prendre la partie de $\hat{\mu}$ sur la droite horizontale d'ordonnée x , à projeter ensuite sur la barrière, pour finalement projeter le résultat sur l'axe des ordonnées.

On rappelle que le problème de Skorohod consiste en la recherche d'un temps d'arrêt τ d'un mouvement Brownien $(B_t)_{t \geq 0}$ tel que $B_0 \sim \mu$ afin de réaliser $B_\tau \sim \nu$. De nombreuses méthodes existent ; un survol sur le sujet est celui d'Obłój [Obł04]. On consultera aussi à profit [BCH17], un article récent qui prend l'angle des problèmes et des méthodes du transport optimal pour retrouver, à l'aide de lemmes d'échanges, de nombreuses solutions connues précédemment et en créer de nouvelles. Nos couplages ombrés s'inscrivent dans cette mouvance car la projection de Kellerer $\hat{\nu}_{\alpha, \cdot} = P^{T_\alpha}(\hat{\mu}_{\alpha, \cdot})$ et surtout la loi jointe qui en résulte, s'obtiennent en considérant un mouvement Brownien distribué selon $\hat{\mu}_{\alpha, \cdot}$, arrêté lorsqu'il atteint T_α pour finir distribué selon $\hat{\nu}_{\alpha, \cdot}$.

Nous pouvons donc revisiter sous cette perspective la représentation graphique de la figure 2.4 : un point de départ (U, X) est choisi selon $\hat{\mu}$. On commence un mouvement Brownien vertical jusqu'à atteindre la barrière $\text{Barr} = \{(\alpha, T_\alpha) \in \mathbb{R}^2 : \alpha \in [0, 1]\}$ en un point (U, Y) . Alors, le couplage ombré est la loi de (X, Y) . Le nom « barrière », très classique en transport de Skorohod, fait notamment référence aux solutions de Rost et Root, pour lesquelles on arrête le processus lorsque $(t, B_t)_{t \geq 0}$ rencontre une barrière Barr qui est, dans un cas, saturée sur la droite (c'est-à-dire $(t, y) \in \text{Barr}$ et $t \leq t' \Rightarrow (t', y) \in \text{Barr}$), dans l'autre cas, saturée sur la gauche. Notre barrière est saturée sur la gauche. Sur figure T_α apparaît avec la notation $\text{Barr}[\alpha]$.

Remarque 2.3.5 (Unicité de la barrière). *Plusieurs barrières peuvent donner lieu au même transport ombré de $\widehat{\text{Mart}}(\hat{\mu}, \nu)$, car on peut en effet « trouver » la barrière aux endroits où $\hat{\mu}$ n'attribue aucune masse. Toutefois un transport ombré entre $\hat{\mu}$ et ν qui est obtenu par une barrière saturée sur la gauche sera en fait l'unique transport ombré de $\widehat{\text{Mart}}(\hat{\mu}, \nu)$. Cela est démontré de deux façons dans [BeiJ16⁺]. On peut s'appuyer sur le théorème 2.3.1 en évoquant*

le fait que les transports ainsi obtenus sont monotones au sens du deuxième point. On peut aussi s'appuyer sur l'argument classique de Loynes [Loy70] utilisé pour le problème de Skorohod. C'est ce que nous faisons dans la partie 4 de [BeiJ16].

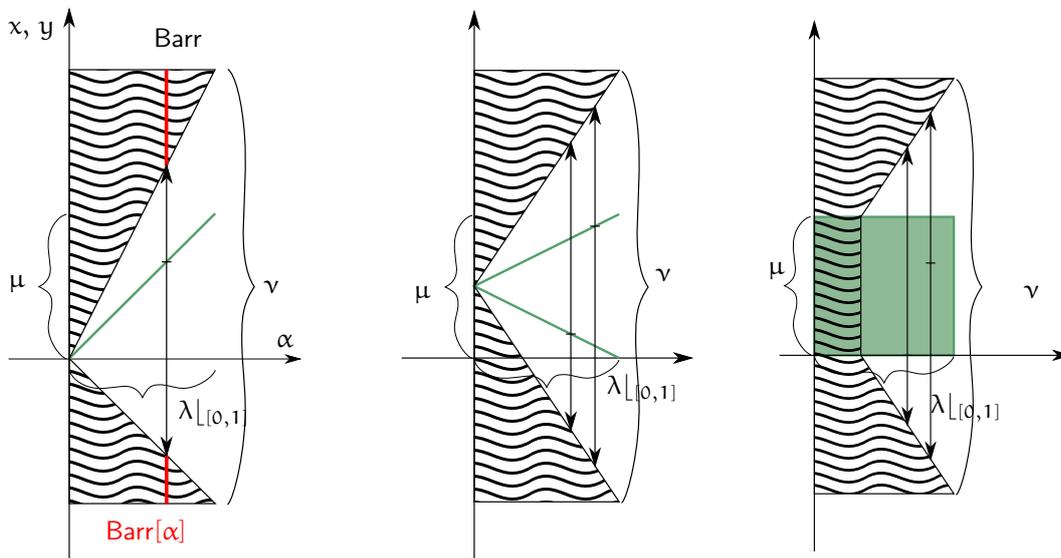


FIGURE 2.4 – La mesure $\hat{\mu}$, en vert, se projette sur $Ox = Oy$ pour former μ , ou bien, pour former ν après avoir diffusé verticalement et heurté la barrière. De gauche à droite : transport rideau gauche, centré, transport soleil. Les mesures μ et ν sont uniformes sur $[0, 1]$ et $[-1, 2]$.

Exemples de transports ombrés

Nous complétons ici la présentation des transports rideau gauche \mathfrak{RG} (ou droit \mathfrak{RD}) et rideau centré \mathfrak{RC} faite aux lemmes 2.2.12, remarque 2.2.13 et à la définition 2.2.16. Ce sont les transports ombrés correspondant aux paramétrages

- $\hat{\mu}_{[0,\alpha],\cdot} = (G_\mu) \# \lambda|_{[0,\alpha]}$ pour le couplage rideau gauche ;
- $\hat{\mu}_{[0,\alpha],\cdot} = (G_\mu) \# \lambda|_{[1-\alpha,1]}$ pour le couplage rideau droit ;
- $\hat{\mu}_{[0,\alpha],\cdot} = S^\mu(\alpha, \delta_m)$ où $m = \int x d\mu(x)$ pour le couplage rideau centré.

Nous introduisons maintenant le transport soleil

Définition 2.3.6 (Transport soleil \mathfrak{Sol}). Soit $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ deux mesures dans l'ordre convexe. Nous appelons transport soleil $\mathfrak{Sol}(\mu, \nu)$ le transport ombré dont le paramétrage est donné par $\hat{\mu}_{[0,\alpha],\cdot} = \alpha \cdot \mu$ (ou, de façon équivalente, par $\hat{\mu} = \lambda \otimes \mu$ ou encore par $\hat{\mu}_{\alpha,\cdot} = \mu$).

Notons que dans l'exemple de la figure 2.4, on a $\mu \preceq_{CD} \nu$ si bien \mathfrak{RC} suit la description faite à la définition 2.2.16.

Les transports ombrés vus comme solutions de problèmes de transport généralisés

Alors que certains transports ombrés, comme \mathfrak{RG} sont l'unique solution d'un problème de transport martingale, ce n'est pas le cas de tous les transports ombrés, loin s'en faut. Nous allons illustrer ce propos par un exemple puis montrer comment tout transport ombré est l'unique solution d'un problème de transport *généralisé*, au sens de [GRST17] où ceux-ci d'ailleurs nommés *transports faibles*.

Exemple 2.3.7. Soit μ et ν dans l'ordre convexe telles que

- i μ est concentrée sur $[a, b]$ et ν sur $\mathbb{R} \setminus]a, b[$ avec $a < b \in \mathbb{R}$.
- ii μ, ν sont à support fini.

On suppose que \mathfrak{Sol} est optimal pour le problème martingale associé à un coût de transport c . Alors, les éléments de $\text{Mart}(\mu, \nu)$ sont tous optimaux pour ce problème de transport.

Démonstration. Le point (i) nous situe dans le cadre de l'exemple 2.2.6. On peut facilement vérifier que $\mathfrak{Sol}(\mu, \nu)$ a pour support $\text{Spt}(\mu) \times \text{Spt}(\nu)$. Tout transport martingale dont le support est inclus dans cet ensemble est optimal entre ses marges. Puisque, avec (ii), le problème est discret cela résulte de la dualité classique en programmation linéaire. Sans (ii), l'énoncé reste vrai pour c continue et bornée, mais il faut invoquer la réciproque du théorème 1.5.3, ce dont nous nous sommes passé jusqu'à présent. \square

Le problème de transport généralisé consiste en la recherche du minimum pour

$$\pi \in \text{Marg}(\mu, \nu) \mapsto \int C(x, \pi_x, \cdot) d\mu(x), \quad (2.5)$$

où $C : \mathbb{R} \times \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ est une fonction de coût généralisée.

Proposition 2.3.8. Soit μ et ν dans l'ordre convexe, et $\hat{\mu} \in \text{Marg}(\lambda, \mu)$ un paramétrage. Soit encore la fonction de coût généralisée

$$C^{\hat{\mu}} : (x, \eta) \mapsto \inf_{\hat{\gamma} \in \widehat{\text{Mart}}(\hat{\mu}_{\cdot, x} \otimes \delta_x, \eta)} \int (1 - \alpha) \sqrt{1 + y^2} d\hat{\gamma}(\alpha, x', y). \quad (2.6)$$

définie à partir de $\hat{\mu}$. Alors, le couplage ombré $\pi \in \text{Mart}(\mu, \nu)$ associé au paramétrage $\hat{\mu}$ est l'unique minimiseur du problème de transport généralisé (2.5) associé à $C^{\hat{\mu}}$.

Remarque 2.3.9. — On a $C^{\hat{\mu}}(x, \cdot) < +\infty$ si et seulement si $\text{Bary}(\eta) = x$. En effet $\widehat{\text{Mart}}(\hat{\mu}_{\cdot, x} \otimes \delta_x, \eta)$ n'est non vide qu'à cette condition.

- L'équation définissant en (2.6) le coût $C^{\hat{\mu}}$ constitue un problème de transport martingale paramétré du type de celui résolu par $\hat{\pi}$ dans le théorème 2.3.1. Le paramétrage est différent puisque la première marge $\hat{\mu}_{\cdot, x}$ n'est pas λ mais $\hat{\mu}_{x, \cdot}$. La solution s'obtient toutefois de la même manière, par les ombres, amenant ici une construction s'appuyant sur les ombres $S^\eta(\hat{\mu}_{\cdot, x}([0, \alpha])\delta_x)$. En particulier

$$C^{\hat{\mu}}(x, \eta) = \int_0^1 \left(\int \sqrt{1 + y^2} S^\eta(\hat{\mu}_{\cdot, x}([0, \alpha])\delta_x)(dy) \right) d\alpha,$$

avec $C^{\hat{\mu}}(x, \eta) = +\infty$ si une des ombres n'est pas définie, c'est-à-dire si $x \neq \text{Bary}(\eta)$.

- Pour distinguer un transport ombré π^* comme transport optimal, on pourrait très bien définir un coût $C(x, \eta)$ qui ne serait fini que pour $\eta = \pi_{x, \cdot}^*$, et infini sinon. La définition du coût serait alors indexée sur μ et ν , ce qui n'est pas souhaitable. Dans notre problème $C^{\hat{\mu}}$ dépend encore de μ ce qui reste dérangeant. Toutefois pour $\hat{\mu} = \lambda \otimes \mu$, le paramétrage du transport soleil, on a $\hat{\mu}_{\cdot, x} = \lambda$ et donc

$$\begin{aligned} C^{\hat{\mu}}(x, \eta) &:= \inf_{\hat{\gamma} \in \widehat{\text{Mart}}(\lambda \otimes \delta_x, \eta)} \int (1 - \alpha) \sqrt{1 + y^2} d\hat{\gamma}(\alpha, x', y) \\ &= \begin{cases} \int_0^1 \left(\int \sqrt{1 + y^2} S^\eta(\alpha \delta_x)(dy) \right) d\alpha & \text{si } \text{Bary}(\eta) = x, \\ +\infty & \text{sinon} \end{cases} \end{aligned}$$

ne dépend pas de μ . On a donc un coût $C^{\mathfrak{Sol}}$, indépendant de μ , pour lequel le problème de transport martingale généralisé a une unique solution, à savoir \mathfrak{Sol} . Cela détone avec l'exemple 2.3.7 au sujet de \mathfrak{Sol} , pour lequel tout transport martingale est un minimiseur.

Autour du théorème de Kellerer, utilisation dynamique des transports martingale et quantile pour définir des processus particuliers

Dans ce nouveau chapitre l'objectif est de construire des processus markoviens en s'appuyant sur le transport martingale, tel que présenté au chapitre 2. Pour ce faire nous nous appuyerons principalement sur les articles [Jui18], [Jui16a] et [BouJ18⁺]. Nous passerons ainsi des mesures π à deux marges à des mesures \mathbb{P} possédant une infinité de marges $(\mu_t)_{t \in \mathcal{T}}$. Ici μ_t sera le plus souvent une mesure de probabilité sur \mathbb{R} ou \mathbb{R}^d , et \mathcal{T} un ensemble totalement ordonné tel que, par exemple, $[0, 1]$, \mathbb{R}^+ ou \mathbb{R} .

Définition 3.0.1. Soit $(\mu_t)_{t \in \mathcal{T}}$ une famille de mesures de $\mathcal{M}_+(\mathcal{X})$ où \mathcal{X} est un espace métrique complet séparable muni de sa tribu borélienne. On note $\text{Marg}((\mu_t)_{t \in \mathcal{T}})$ l'ensemble des mesures \mathbb{P} sur $\mathcal{X}^{\mathcal{T}}$ muni de la tribu cylindrique qui possèdent de plus les mesures μ_t pour marges :

$$\text{Marg}((\mu_t)_{t \in \mathcal{T}}) = \{\mathbb{P} \in \mathcal{P}(\mathcal{X}^{\mathcal{T}}) : \forall t \in \mathcal{T}, (\text{proj}^t)_\# \mathbb{P} = \mu_t\}.$$

Comme précédemment pour $\text{Marg}(\mu, \nu)$ nous chercherons dans ce chapitre à distinguer des éléments particuliers de $\text{Marg}((\mu_t)_{t \in \mathcal{T}})$. Aucune cohérence entre les mesures, aucune régularité de $t \mapsto \mu_t$ n'est requise a priori. Des éléments particuliers de $\text{Marg}((\mu_t)_{t \in \mathcal{T}})$ sont la mesure produit $\mathfrak{Ind} = \bigotimes_{t \in \mathcal{T}} \mu_t$ et la mesure quantile $\mathfrak{Q} = (\prod_{t \in \mathcal{T}} G_{\mu_t})_\# \lambda$ qui sont définies quels que soient \mathcal{T} et $(\mu_t)_{t \in \mathcal{T}}$. Ces mesures jouissent de propriétés les caractérisant. Nous sommes intéressés par la recherche de mesures canoniques dans $\text{Mart}((\mu_t)_{t \in \mathcal{T}})$, le sous-espace des mesures martingale (voir la définition 3.1.1 plus loin).

Nous appellerons *processus (au sens large)* toute famille de variables aléatoires $(X_t)_{t \in \mathcal{T}}$ définies sur un même espace probabilisé. Aucune mesurabilité de $(t, \omega) \mapsto X_t(\omega)$ n'est requise. Rappelons que le processus canonique associé à $\mathbb{P} \in \mathcal{P}(\mathcal{X}^{\mathcal{T}})$ est la famille des X_t définis par $X_t \equiv \text{proj}^t : \omega \in \mathcal{X}^{\mathcal{T}} \mapsto \text{proj}^t(\omega)$. La loi de $(X_t)_{t \in \mathcal{T}}$ est naturellement \mathbb{P} . Rappelons que cette dernière est entièrement déterminée par $\{\text{Loi}(X_t)_{t \in \mathcal{S}} : \mathcal{S} \text{ fini } \subset \mathcal{T}\}$. C'est une partie du théorème de Kolmogorov–Daniell. Au cours de ce chapitre nous noterons $\mathbb{P}^{\mathcal{S}}$ la projection $(\text{proj}^{\mathcal{S}})_\# \mathbb{P}$.

Théorème 3.0.2 (Kolmogorov–Daniell). *Soit $(\mu_S)_S$ une famille de mesures de probabilité où S est un sous-ensemble fini de \mathcal{T} telles que $\mu_S \in \mathcal{P}(\mathcal{X}^S)$. Si pour tout $S' \subset S$ on a $(\text{proj}^{S'})_\# \mu_S = \mu_{S'}$, alors il existe une unique mesure $\mathbb{P} \in \mathcal{P}(\mathcal{X}^{\mathcal{T}})$ telle que $\mathbb{P}^S = \mu_S$ pour tout S .*

3.1 Mesures martingale et mesures markoviennes

La propriété de Markov ou bien celle d'être une martingale font partie des plus importantes en théorie des processus. Comme on le sait, leur définition fait intervenir une filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$ où \mathcal{T} est partiellement ordonné. Dans notre contexte, \mathcal{F}_t sera systématiquement $\sigma((X_s)_{s \leq t})$, la tribu engendrée par les variables aléatoires dont l'indice s est, pour l'ordre partiel de \mathcal{T} , plus petit que t . Ce n'est en particulier pas nécessairement la « filtration naturelle » de la théorie des processus.

Mesures martingale

Définition 3.1.1 (Martingale à indices partiellement ordonnés). Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné. Un processus $(X_t)_{t \in \mathcal{T}}$ à valeur dans \mathbb{R}^d est appelé *martingale* si $\mathbb{E}(\|X_t\|) < +\infty$ pour tout $t \in \mathcal{T}$ et si pour tout $s \leq t$ on a $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$, où \mathcal{F}_s est la tribu définie par $\mathcal{F}_s = \sigma(X_r | r \leq s)$.

De façon équivalente $(X_t)_{t \in \mathcal{T}}$ est une martingale si pour tout $s \leq t$, et toute famille $\{s_1, \dots, s_n\}$ finie d'indices inférieurs à s on a $\mathbb{E}(X_t | X_s, X_{s_1}, \dots, X_{s_n}) = X_s$.

Une mesure P de marges $\mu_t \in \mathcal{P}_1(\mathbb{R}^d)$ à premier moment fini est dite *mesure martingale* si c'est la loi d'une martingale. Cela se traduit de façon équivalente par

$$\int f(x_1, \dots, x_n, x) \times (y - x) dP^S(x_1, \dots, x_n, x, y) = 0 \quad (3.1)$$

pour toute famille d'indice $S = \{s_1, \dots, s_n \leq s < t\} \subset \mathcal{T}$ telle que précédemment et toute fonction f mesurable, bornée et positive.

Nous désignerons par $\text{Mart}((\mu_t)_{t \in \mathcal{T}})$ l'ensemble des mesures martingale de $\text{Marg}((\mu_t)_{t \in \mathcal{T}})$.

Définition 3.1.2 (Sous-martingales). Pour $\mathbb{R}^d = \mathbb{R}$, les mêmes définitions sont opérantes pour les sous-martingales en remplaçant $=$ par \geq , notamment dans (3.1). On notera $\text{SousMart}((\mu_t)_{t \in \mathcal{T}})$ l'ensemble des mesures sous-martingale.

Remarque 3.1.3. Lorsque $\mathcal{T} = \{1, 2\}$ est un ensemble totalement ordonné à deux indices, nous retrouvons les transports martingale et sous-martingale de la définition 1.5.1. Alors, $\text{Mart}((\mu_t)_{t \in \mathcal{T}}) = \text{Mart}(\mu_1, \mu_2)$ et $\text{SousMart}((\mu_t)_{t \in \mathcal{T}}) = \text{SousMart}(\mu_1, \mu_2)$.

Remarque 3.1.4. Pour $\mathcal{T} = \{0, \dots, n\}$, l'ensemble $\text{Mart}((\mu_t)_{t \in \mathcal{T}})$ est non-vide si et seulement si $\mu_i \preceq_C \mu_{i+1}$. C'est un théorème de Strassen [Str65, Theorem 8], qui part de la version de dimension d de la proposition 2.1.4 puis se démontre par concaténation (voir le paragraphe suivant). Pour $\mathcal{T} = \mathbb{N}$ le théorème 3.0.2 permet d'aboutir à la même conclusion.

Mesures markoviennes

Sur \mathbb{R}^m on définit la transposition par ${}^t x = (x_m, \dots, x_1)$ où $x = (x_1, \dots, x_m)$. L'opération s'étend à $\text{Marg}(\mu_1, \dots, \mu_m)$ de façon à ce que ${}^t P \in \text{Marg}(\mu_m, \dots, \mu_1)$ lorsque $P \in \text{Marg}(\mu_1, \dots, \mu_m)$.

L'opération de concaténation markovienne est définie par Kellerer dans [Kel72] p. 111. Elle est centrale dans cet article, duquel sont extraits les théorèmes 3.1.14, 3.2.1 et 3.3.2, et correspond simplement, en probabilités, à la loi obtenue lorsqu'on met bout à bout deux processus de façon à ce que le passé soit indépendant du futur conditionnellement au présent.

Définition 3.1.5 (Concaténation de mesures partageant une marge). Soit $m, n \geq 2$ des entiers et $(\mu_t)_{t \in \{1, \dots, m+n-1\}}$ des mesures de $\mathcal{P}(\mathcal{X})$. Étant donnés $P \in \text{Marg}(\mu_1, \dots, \mu_m)$ et $Q \in \text{Marg}(\mu_m, \dots, \mu_{m+n-1})$, la concaténée $P \circ Q$ des deux mesures est l'unique élément de $\text{Marg}(\mu_1, \dots, \mu_{m+n-1})$ tel que

$$(P \circ Q)(B_1 \times B_2 \times B_3) = \int_{y \in B_2} \left(\iint_{{}^t x \in B_1, z \in B_3} {}^t P_{y, \cdot} (dx) Q_{y, \cdot} (dz) \right) d\mu_m(y).$$

pour tout boréliens $B_1 \subset \mathcal{X}^{m-1}$, $B_2 \subset \mathcal{X}$ et $B_3 \subset \mathcal{X}^{n-1}$.

La concaténation est une opération associative. Définissons maintenant la propriété de Markov.

Définition 3.1.6 (Mesure markovienne). Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné. Un processus $(X_t)_{t \in \mathcal{T}}$ est appelé *markovien* si pour tout $t \in \mathcal{T}$ et $s < t$ on a égalité entre $\text{Loi}(X_t | \mathcal{F}_s)$ et $\text{Loi}(X_t | X_s)$, où, comme précédemment, \mathcal{F}_s est la tribu $\sigma(X_r)_{r \leq s}$.

Une mesure $P \in \mathcal{P}(\mathcal{X}^{\mathcal{T}})$ est *markovienne* si elle est la loi d'un processus markovien. Cela se traduit de façon équivalente par l'équation

$$P^{s_1, \dots, s_n} = P^{s_1, s_2} \circ \dots \circ P^{s_{n-1}, s_n}$$

pour tout ensemble fini d'indice totalement ordonné $s_1 < \dots < s_n$.

Remarque 3.1.7. Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné et $(\mu_t)_{t \in \mathcal{T}}$ une famille de mesure de probabilité dans $\mathcal{P}_1(\mathbb{R}^d)$. Une mesure $\mathbb{P} \in \mathcal{P}(\mathcal{X}^{\mathcal{T}})$ est une mesure (ou loi) de martingale markovienne si pour tout ensemble fini d'indices totalement ordonnés $s_1 < \dots < s_n$ on a

$$\begin{cases} \mathbb{P}^{s_i, s_{i+1}} \in \text{Mart}(\mu_{s_i}, \mu_{s_{i+1}}), & \forall i \leq n-1, \\ \mathbb{P}^{s_1, \dots, s_n} = \mathbb{P}^{s_1, s_2} \circ \dots \circ \mathbb{P}^{s_{n-1}, s_n}. \end{cases}$$

l'énoncé suivant est un corollaire classique du théorème 3.0.2

Corollaire 3.1.8. Soit \mathcal{T} un ensemble totalement ordonné et $(\mu_t)_{t \in \mathcal{T}}$ une famille de probabilités sur \mathcal{X} . Soit encore $(\mu_{s,t})_{s < t}$ une famille de plans de transport avec $\mu_{s,t} \in \text{Marg}(\mu_s, \mu_t)$. Si

$$\mu_{s,u} = (\text{proj}^{s,u})_{\#}(\mu_{s,t} \circ \mu_{t,u}) \quad (3.2)$$

pour tout triplet ordonné $s < t < u$, alors il existe une unique mesure markovienne $\mathbb{P} \in \mathcal{P}(\mathcal{X}^{\mathcal{T}})$ telle que $\mathbb{P}^{s,t} = \mu_{s,t}$ pour tout $s < t$.

Définition 3.1.9. On appelle *famille consistente* toute famille $(\mu_s)_s$ ou $(\mu_{s,t})_{s < t}$ telle que celles apparaissant dans le théorème 3.0.2 ou le corollaire 3.1.8.

Composition de noyaux et de plans de transport.

Rappelons que deux noyaux se composent de la façon suivante :

$$(k \cdot k')(x, A) = \int k'(y, A) k(x, dy).$$

Nous savons désintégrer un plan de transport $\pi \in \text{Marg}(\mu, \nu)$ par rapport à sa première marge μ , produisant la famille de lois conditionnelles $(\pi_{x,\cdot})_{x \in \mathbb{R}}$. Nous noterons k_π le noyau k_π défini par $k_\pi(x, \cdot) = \pi_{x,\cdot}$. Notons que $x \mapsto k_\pi(x, \cdot) = \pi_{x,\cdot}$ est seulement presque sûrement défini, par rapport à μ . L'opération inverse à la désintégration, celle de jonction d'une mesure et d'un noyau sera notée Joint. Elle est définie par $\text{Joint}(\mu, k) = \mu \cdot (\text{id}, k) \in \text{Marg}(\mu, \mu \cdot k)$.

Pour $\pi \in \text{Marg}(\mu, \mu')$ et $\pi' \in \text{Marg}(\mu', \mu'')$, nous définissons la composition des deux plans de transport par :

$$\pi \cdot \pi' := \text{Joint}(\mu, k_\pi \cdot k_{\pi'}) \in \text{Marg}(\mu, \mu'').$$

On observe que $k_{\pi \cdot \pi'} = k_\pi \cdot k'_{\pi'}$. Il y a une parenté claire entre les relations de composition et de concaténation. Celle-ci s'illustre par exemple dans l'équation (3.2) lorsqu'on la complète ainsi : $\mu_{s,u} = (\text{proj}^{s,u})_{\#}(\mu_{s,t} \circ \mu_{t,u}) = \mu_{s,t} \cdot \mu_{t,u}$

PCOC et théorème de Kellerer.

Le concept de *Processus Croissant pour l'Ordre Convexe (PCOC)* est apparu pour la première fois dans [HY10]. Un jeu de mots est à l'origine de l'intitulé *peacock*, en anglais. Le terme a évolué et peut prendre des significations diverses selon les auteurs. En particulier ce que nous appellerons PCOC n'est pas un processus.

Définition 3.1.10 (PCOC). Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné et $d \geq 1$ un entier. Une famille $(\mu_t)_{t \in \mathcal{T}}$ de mesures dans $\mathcal{P}_1(\mathbb{R}^d)$ est un *PCOC*, ou *peacock*, si pour tout couple (s, t) satisfaisant à $s \leq t$ on a $\mu_s \preceq_C \mu_t$.

Du fait de l'inégalité de Jensen conditionnelle, si $(X_t)_{t \in \mathcal{T}}$ est une martingale, la famille $(\mu_t)_{t \in \mathcal{T}}$ définie par $\mu_t = \text{Loi}(X_t)$ est un PCOC. On dira que la martingale $(X_t)_{t \in \mathcal{T}}$ est *associée* au PCOC $(\mu_t)_{t \in \mathcal{T}}$. C'est l'objet de ce chapitre et du thème des PCOC que de trouver des martingales associées à un PCOC. Nous sommes en particulier intéressé par les martingales markoviennes. Le concept suivant est spécialement important à cet égard.

Définition 3.1.11 (Noyau (1-)lipschitzien). Un noyau $k : x \mapsto k(x, \cdot) \in \mathcal{P}_1(\mathbb{R}^d)$ transportant μ sur $\nu = \mu \cdot k$ est dit *1-lipschitzien* ou, plus souvent, *lipschitzien* si il existe un borélien $A \subset \mathbb{R}^d$ tel que $\mu(A) = 1$ et pour lequel $k|_A$ est lipschitzien de constante ≤ 1 entre $(A, \|\cdot\|_{\mathbb{R}^d})$ et $(\mathcal{P}(\mathbb{R}^d), W_1)$, c'est-à-dire :

$$W_1(k(x, \cdot), k(x', \cdot)) \leq \|x - x'\| \quad (3.3)$$

pour tout $x, x' \in A \subset \mathbb{R}^d$. On notera $\text{Marg}^{\text{NL}}(\mu, \nu)$ l'ensemble des plans de transport $\pi \in \text{Marg}(\mu, \nu)$ tels que k_π est un noyau lipschitzien.

Remarque 3.1.12 (Extension du noyau). *Pour tout noyau lipschitzien k il existe un noyau \tilde{k} , coïncidant avec k presque sûrement (pour μ) tel que la relation (3.3) est vérifiée sur tout \mathbb{R} (voir le théorème d'extension des fonction lipschitziennes de Lang et Schroeder [LS97]. Il s'applique du fait que $\mathcal{P}(\mathbb{R}^d)$ est un espace d'Alexandrov). La même remarque vaut pour la définition 3.2.7 des noyaux croissants.*

Définition 3.1.13 (Martingale markovienne à noyaux lipschitziens). Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné. Un processus $(X_t)_{t \in \mathcal{T}}$ sera appelé *martingale markovienne à noyaux lipschitziens* si c'est une martingale markovienne dont les lois de transitions $\text{Loi}(X_s, X_t)$ sont des noyaux lipschitziens (au sens de la définition 3.1.11) pour tout $s \leq t$.

Notre énoncé du Theorem 3 de Kellerer dans [Kel72] (celui que, en l'absence d'autre précision, nous appelons « le théorème de Kellerer ») est formulé en terme de peacocks :

Théorème 3.1.14 (Théorème de Kellerer). *Soit $(\mu_t)_{t \in \mathcal{T}}$ une famille de mesures de $\mathcal{P}_1(\mathbb{R})$ indexée sur un ensemble totalement ordonné \mathcal{T} . Les assertions suivantes sont équivalentes :*

- i $(\mu_t)_{t \in \mathcal{T}}$ est un PCOC,*
- ii $(\mu_t)_{t \in \mathcal{T}}$ est associé à une martingale $(X_t)_{t \in \mathcal{T}}$,*
- iii $(\mu_t)_{t \in \mathcal{T}}$ est associé à une martingale markovienne $(X_t)_{t \in \mathcal{T}}$,*
- iv $(\mu_t)_{t \in \mathcal{T}}$ est associé à une martingale markovienne à noyaux lipschitziens $(X_t)_{t \in \mathcal{T}}$.*

L'analogie vaut avec une sous-martingale à noyaux lipschitziens si $(\mu_t)_{t \in \mathcal{T}}$ est croissante pour $\preceq_{C,sto}$.

On notera que les implications $4 \Rightarrow 3 \Rightarrow 2 \Rightarrow 1$ sont évidentes.

3.2 Améliorations de la démonstration et du théorème de Kellerer

Nous donnons le plan de la démonstration de Kellerer au sujet des martingales markoviennes [Kel72] ou des sous-martingales markoviennes [Kel73]. Notre premier objectif est ici d'expliquer comment nous étendons ce théorème aux processus croissants.

La démonstration de Kellerer de [Kel72] est construite sur le Theorem 1 (reproduit ici dans le théorème 3.2.1) au sujet de l'existence d'un processus markovien dont les marges sont données. Ce théorème peut être vu comme une prolongation du théorème de Kolmogorov–Daniell et de son corollaire (théorème 3.0.2 et corollaire 3.1.8). Vient ensuite, chez Kellerer, Theorem 2 qui (avec Satz 20) amène l'existence d'un transport martingale à noyau lipschitzien (voir théorème 3.3.2). Le pendant sous-martingale est démontré dans [Kel73]. Au théorème 3.2.14 nous reviendrons sur l'existence d'un tel transport martingale à noyau lipschitzien en lui attribuant un choix canonique, tout à fait dans l'esprit du travail de Kellerer, celui du transport soleil. Finalement, Kellerer conclut avec le Theorem 3, celui que nous appelons théorème de Kellerer dans ce mémoire et que nous avons cité, en le reformulant légèrement, au théorème 3.1.14.

Théorème 3.2.1 ([Kel72, Theorem 1]). *Soit $(\mu_t)_{t \in \mathbb{R}}$ une famille de mesures de probabilité sur un espace métrique polonais \mathcal{X} , et, pour tout couple $s < t$, soit $\mathcal{N}_{s,t} \subset \mathcal{P}(\mathcal{X}^2)$ un ensemble de plans de transport. On suppose que :*

- (1) pour tout s, t , $\mathcal{N}_{s,t}$ est non-vide,
- (2) pour tout s, t , $\mathcal{N}_{s,t} \subset \text{Marg}(\mu_s, \mu_t)$,
- (3) pour tout s, t , $\mathcal{N}_{s,t}$ est fermé pour la topologie faible,
- (4) pour $r < s < t$ et $(\pi, \pi') \in \mathcal{N}_{r,s} \times \mathcal{N}_{s,t}$, le produit $\pi \cdot \pi'$ est dans $\mathcal{N}_{r,t}$,
- (5) pour tout entier d et $t_1 < \dots < t_d$, et des suites $(\pi_{t_i, t_{i+1}}^n)_n \in \mathcal{N}_{t_i, t_{i+1}}$ convergeant faiblement vers des limites $\pi_{t_i, t_{i+1}}$, la suite $(\pi_{t_1, t_2}^n \circ \dots \circ \pi_{t_{d-1}, t_d}^n)_n$ tend faiblement vers $\pi_{t_1, t_2} \circ \dots \circ \pi_{t_{d-1}, t_d}$.

Alors, il existe une mesure $\mathbf{P} \in \text{Marg}((\mu_t)_t)$, markovienne, telle que $(\text{proj}^{s,t})_{\#} \mathbf{P} \in \mathcal{N}_{s,t}$ pour tout couple $s < t$.

Le lemme suivant est central. C'est celui sur la continuité de la concaténation \circ lorsque les noyaux sont lipschitziens. C'était un élément utilisé d'une manière où d'une autre dans toutes les démonstrations connues du théorème de Kellerer [Kel72, Kel73, Low08⁺, BHS16, HRY14].

Lemme 3.2.2 ([Kel72, Sätze 14 and 15]). *Avec les notations du théorème 3.2.1 et en prenant $\mathcal{N}_{s,t} = \text{Marg}^{\text{NL}}(\mu_s, \mu_t)$ pour tout $s < t$, la propriété (5) du même théorème est satisfaite par $(\mathcal{N}_{s,t})_{s < t}$.*

Théorème 3.2.3 ([Kel72, Theorem 2, Satz 20] et [Kel73]). *Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R})$ dans l'ordre convexe. Alors l'ensemble*

$$\text{Mart}^{\text{NL}}(\mu, \nu) := \text{Mart}(\mu, \nu) \cap \text{Marg}^{\text{NL}}(\mu, \nu)$$

est non-vide.

Si maintenant μ et ν sont dans l'ordre convexe croissant, l'ensemble

$$\text{SousMart}^{\text{NL}}(\mu, \nu) := \text{SousMart}(\mu, \nu) \cap \text{Marg}^{\text{NL}}(\mu, \nu)$$

est également non-vide.

Du fait du lemme 3.2.2, Kellerer démontre que les cinq propriétés de son Theorem 1 (c'est-à-dire le théorème 3.2.1) sont satisfaites par la famille $(\text{Mart}^{\text{NL}}(\mu_s, \mu_t))_{s < t \in \mathbb{R}}$ (sous réserve, bien sûr, que $(\mu_t)_{t \in \mathbb{R}}$ est un PCOC). La remarque 3.1.7 met alors en évidence le fait que le théorème de Kellerer (c'est-à-dire théorème 3.1.14) est démontré.

Extension du théorème de Kellerer à l'ordre stochastique

Le théorème de Kellerer concerne les martingales et les sous-martingales, en relation avec les ordres $\preceq_{\mathbb{C}}$ et $\preceq_{\mathbb{C}, \text{sto}}$. Il semble naturel, notamment quand on pense à la décomposition de Doob–Meyer, de se pencher sur le cas des processus croissants. Nous démontrons ainsi l'extension du théorème de Kellerer suivante :

Théorème 3.2.4 (Théorème « parapluie » [BouJ18⁺, Theorem C]). *Soit $(\mu_t)_{t \in \mathbb{R}}$ une famille de mesures de $\mathcal{P}_1(\mathbb{R})$ croissante pour $\preceq_{\mathbb{C}}$, $\preceq_{\mathbb{C}, \text{sto}}$ ou pour \preceq_{sto} . Alors, il existe une mesure markovienne $\mathbf{P} \in \text{Marg}((\mu_t)_t)$ qui est, respectivement, la loi d'une martingale, d'une sous-martingale ou d'un processus croissant.*

Remarque 3.2.5 (Remarques sur le théorème). — *Notre démonstration épouse les contours de celle de Kellerer à un détail important près, que nous allons souligner, et qui nous permet d'inclure l'ordre stochastique. Il s'agira de remplacer les noyaux lipschitziens apparaissant dans le lemme 3.2.2 et utilisés dans toutes les démonstrations connues du théorème de Kellerer par des « noyaux croissants » (voir la définition 3.2.7).*

- « \mathbf{P} est la loi d'un processus croissant » se comprend d'abord dans un sens large comme : $\mathbf{P} \in \text{Crois}((\mu_t)_{t \in \mathbb{R}})$ (voir définition 3.2.6). Le corollaire 3.2.12 apparaissant plus loin donne toutefois un fondement plus sûr à cette dénomination.

— Avec le processus Markov-quantile, nous donnerons au chapitre 4 une seconde démonstration de ce que nous obtenons ici pour \preceq_{sto} .

Définition 3.2.6 (Transport croissant). Soit μ et ν des mesures de $\mathcal{P}(\mathbb{R})$. Une mesure $\pi \in \text{Marg}(\mu, \nu)$ est appelée *transport croissant* si $\pi(\{(x, y) \in \mathbb{R}^2 : x \leq y\}) = 1$, c'est-à-dire si $\pi = \text{Loi}((X_i)_{i \in \{1, 2\}})$ où $X_1 \leq X_2$ presque sûrement.

Nous noterons $\text{Crois}(\mu, \nu)$ l'ensemble des transports croissants de marge μ et ν . De même $\text{Crois}((\mu_t)_{t \in \mathcal{T}})$ sera l'ensemble des mesures \mathbb{P} telle que $\mathbb{P}^{s, t} \in \text{Crois}(\mu_s, \mu_t)$ pour tout $s \leq t$.

La notion suivante est différente de celle de transport croissant.

Définition 3.2.7 (Noyau croissant). Un noyau $k : x \mapsto k(x, \cdot) \in \mathcal{P}(\mathbb{R}^d)$ transportant μ sur $\nu = \mu \cdot k$ est dit *croissant* si il existe un borélien $A \subset \mathbb{R}^d$ tel que $\mu(A) = 1$ et pour lequel $k|_A$ est croissant de $(A \subset \mathbb{R}, \leq)$ dans $(\mathcal{P}(\mathbb{R}), \preceq_{sto})$, c'est-à-dire :

$$x \leq x' \implies k(x, \cdot) \preceq_{sto} k(x', \cdot)$$

pour tout couple (x, x') de points de A . On notera $\text{Marg}^{\text{NC}}(\mu, \nu)$ l'ensemble des plans de transport $\pi \in \text{Marg}(\mu, \nu)$ tels que k_π est un noyau croissant.

Ce lemme démontré dans [BouJ18⁺] se substitue au lemme 3.2.2.

Lemme 3.2.8 (Continuité de \circ lorsque les noyaux sont croissants). Avec les notations du théorème 3.2.1 et en prenant $\mathcal{N}_{s, t} = \text{Marg}^{\text{NC}}(\mu_s, \mu_t)$ pour tout $s < t$, la propriété (5) du même théorème est satisfaite par $(\mathcal{N}_{s, t}^{\text{NC}})_{s < t}$.

Remarque 3.2.9. Le lemme 3.2.8 signifie en particulier que lorsqu'elles sont définies les limites suivantes commutent :

$$\lim_{n \rightarrow \infty} \pi_{s, t}^{(n)} \cdot \pi_{s, t}^{(n)} = \left(\lim_{n \rightarrow \infty} \pi_{s, t}^{(n)} \right) \cdot \left(\lim_{n \rightarrow \infty} \pi_{s, t}^{(n)} \right).$$

Dans ce lemme, l'hypothèse $\pi_{s, t}^{(n)} \in \text{Marg}(\mu_s, \mu_t)$ est indispensable. Ce n'est pas le cas dans le lemme 3.2.8 dont l'énoncé peut être étendu à des transports dont les marges varient avec n .

Théorème 3.2.10 (Ersatz du Theorem 2 de Kellerer). Soit μ et ν des mesures de $\mathcal{P}_1(\mathbb{R})$ dans l'ordre convexe. Alors l'ensemble

$$\text{Mart}^{\text{NC}}(\mu, \nu) := \text{Mart}(\mu, \nu) \cap \text{Marg}^{\text{NC}}(\mu, \nu)$$

est non-vide.

Si maintenant μ et ν sont dans l'ordre convexe croissant, l'ensemble

$$\text{SousMart}^{\text{NC}}(\mu, \nu) := \text{SousMart}(\mu, \nu) \cap \text{Marg}^{\text{NC}}(\mu, \nu)$$

est également non-vide.

Finalement si μ et ν sont des mesures de $\mathcal{P}(\mathbb{R})$ dans l'ordre stochastique, l'ensemble

$$\text{Crois}^{\text{NC}}(\mu, \nu) := \text{Crois}(\mu, \nu) \cap \text{Marg}^{\text{NC}}(\mu, \nu)$$

est non-vide lui aussi.

Démonstration. On peut tout d'abord constater $\text{Mart}^{\text{NC}}(\mu, \nu) = \text{Mart}^{\text{NL}}(\mu, \nu)$. Ceci se déduit simplement des définitions de noyau croissant, de noyau lipschitzien et de la proposition 1.2.15 appliquée, pour $p = 1$, aux mesures $k(x, \cdot)$ et $k(x', \cdot)$ dont les barycentres sont x et x' . Ainsi le résultat de non-vacuité est-il une conséquence du théorème 3.3.2.

Il en va de même pour le fait que $\text{SousMart}^{\text{NC}}(\mu, \nu)$ est non vide (pour peu que $\mu \preceq_{C, sto} \nu$). Toutefois c'est seulement en consultant la démonstration, par Kellerer, du fait que $\text{SousMart}^{\text{NL}}(\mu, \nu)$ est non vide, que l'on s'aperçoit que l'élément exhibé est aussi dans $\text{SousMart}^{\text{NC}}(\mu, \nu)$. Il n'y a en particulier pas égalité des deux ensembles en général, contrairement au cas martingale.

Finalement si $\mu \preceq_{sto} \nu$, le transport quantile $\mathcal{Q}(\mu, \nu)$ est un élément de $\text{Crois}^{\text{NC}}(\mu, \nu)$. En effet c'est un transport croissant au sens de la définition 3.2.6. Le noyau correspondant est aussi croissant au sens de la définition 3.2.7. Nous retiendrons ce résultat pour plus tard.

Lemme 3.2.11. *Le transport quantile $\mathcal{Q}(\mu, \nu)$ est à noyau croissant. De plus $\text{Crois}(\mu, \nu)$ est non vide si et seulement si $\mu \preceq_{sto} \nu$, si et seulement si $\mathcal{Q}(\mu, \nu) \in \text{Crois}(\mu, \nu)$.*

□

Faisant suite au théorème 3.2.4 on peut finalement parler de processus croissant au sens du corollaire suivant, démontré au Corollary 1.14 de [BouJ18⁺].

Corollaire 3.2.12. *Soit $(\mu_t)_{t \in \mathbb{R}} \in \mathcal{P}(\mathbb{R})^{\mathbb{R}}$ une famille de mesures croissantes pour \preceq_{sto} . Alors, il existe un processus markovien $X = (X_t)_{t \in \mathbb{R}} : \Omega \rightarrow \mathbb{R}^{\mathbb{R}}$ défini sur un espace probabilisé tel que $\text{Loi}(X) \in \text{Marg}((\mu_t)_t)$ et les trajectoires $t \mapsto X_t(\omega)$ sont croissantes pour tout $\omega \in \Omega$.*

La démonstration du corollaire 3.2.12 se base sur le fait qu'une courbe $(\mu_t)_{t \in \mathbb{R}}$ croissante pour \preceq_{sto} possède au maximum un nombre dénombrable de discontinuités.

Le transport soleil comme transport canonique dans [Kel73] par Kellerer

Nous allons décrire le principe de la démonstration par Kellerer du théorème 3.3.2, le théorème qui stipule que pour $\mu \preceq_C \nu$ l'ensemble $\text{Mart}^{\text{NL}}(\mu, \nu)$ est non-vidé. Le cas de $\text{SousMart}^{\text{NL}}(\mu, \nu) \neq \emptyset$, similaire, ne sera pas détaillé. Nous avons déjà mentionné $\text{Mart}^{\text{NC}}(\mu, \nu) = \text{Mart}^{\text{NL}}(\mu, \nu)$ si bien que nous montrerons à la fois l'existence d'un transport martingale à noyau lipschitzien et à noyau croissant. Alors que la démonstration de Kellerer est basée sur l'existence d'une *représentation de Choquet* de laquelle on extrait l'existence d'un élément dans $\text{Mart}^{\text{NL}}(\mu, \nu)$, nous allons voir que le transport soleil $\mathfrak{Sol}(\mu, \nu)$ correspond à cette représentation de Choquet, fournissant donc un choix canonique.

Principe de la démonstration du théorème 3.3.2. Dans [Kel73], Kellerer exhibe un noyau lipschitzien de martingale entre μ et ν en décrivant ν comme la combinaison d'une famille de mesures extrêmes dans l'ensemble convexe $E(\mu) = \{\omega \in \mathcal{P}_1(\mathbb{R}) : \mu \preceq_C \omega\}$. Il prouve que si ω est une mesure extrême de cet ensemble, alors $\text{Mart}(\mu, \omega)$ est réduit à un unique transport martingale π pour lequel k_π est un noyau lipschitzien. Or la propriété d'être un noyau lipschitzien est conservée pour les combinaisons convexes. On obtient donc pour tout $\nu \in E(\mu)$ l'existence d'un noyau lipschitzien k avec $\nu = \mu \cdot k$. Notons tout de même que $E(\mu)$ n'est pas compact, ce qui nécessite des détours dans l'emploi du théorème de Choquet. □

Comparons cette construction à celle de $\mathfrak{Sol} \in \text{Mart}(\mu, \nu)$. D'après les théorèmes 2.3.1, 2.3.3, il est aussi défini comme une combinaison convexe :

$$\mathfrak{Sol} = \int_0^1 \hat{\pi}_{\alpha, \cdot, \cdot} d\alpha.$$

Ici $\hat{\pi}_{\alpha, \cdot, \cdot} = \text{Loi}((B_0, B_\tau) \mid U = \alpha)$, la variable U est de loi uniforme sur $[0, 1]$, (B_t) est un mouvement brownien indépendant de U tel que $B_0 \sim \mu$, et τ (conditionné à $U = \alpha$) le temps d'atteinte de la section $\text{Barr}[\alpha] := \{y \in \mathbb{R} : (\alpha, y) \in \text{Barr}\}$ avec la barrière (voir les figures 2.3 et 2.4). Nous avons

$$\nu = \int_0^1 p^{\text{Barr}[\alpha]}(\mu) d\alpha \quad \text{et} \quad \mathfrak{Sol} = \int_0^1 \underbrace{\mu \cdot (\text{id}, p^{\text{Barr}[\alpha]})}_{\hat{\pi}_{\alpha, \cdot, \cdot} \in \text{Mart}(\mu, p^{\text{Barr}[\alpha]}(\mu))} d\alpha$$

où on rappelle que $p^{\text{Barr}[\alpha]}$ est la projection de Kellerer et $p_{\text{Barr}[\alpha]}$ la dilatation de Kellerer comme à la définition 2.3.2.

Or il a été établi par Kellerer que les mesures extrêmes de $E(\mu)$ sont précisément les mesures $\omega = P^T(\mu)$ obtenues par projection de Kellerer. Le principe de démonstration que nous venons de rappeler plus haut précisait que les dilatations correspondantes étaient des noyaux lipschitziens. On a donc

Lemme 3.2.13. *Les dilatations de Kellerer p_T sont des noyaux lipschitziens. Il en va de même de $\int_0^1 p_{T_\alpha} d\alpha$ pour tout $\alpha \mapsto T_\alpha \in \mathcal{I}$ mesurable.*

Autre démonstration. Rappelons qu'une dilatation est un noyau lipschitzien si et seulement si c'est un noyau croissant, ce qui a été expliqué dans la démonstration du théorème 3.2.10. La définition 2.3.2 de p_T ainsi que la figure 2.3 le décrivant nous permettent aisément de conclure sur ce point. La deuxième propriété se voit par exemple à partir de la formule $F[(\int_0^1 p_{T_\alpha} d\alpha)(x)] = \int_0^1 F[p_{T_\alpha}(x)]d\alpha$ et de la définition 1.2.8 de l'ordre stochastique. \square

Pour compléter ce qui a été signalé plus haut $\hat{\pi}_{\alpha, \cdot, \cdot} = \text{Joint}(\mu, p_{\text{Barr}[\alpha]}) = \mu \cdot (\text{id}, p_{\text{Barr}[\alpha]})$ est en particulier l'unique élément de $\text{Mart}(\mu, p_{\text{Barr}[\alpha]})(\mu)$ et la dilatation de Kellerer $p_{\text{Barr}[\alpha]}$ est l'unique noyau lipschitzien correspondant. Le théorème suivant récapitule et insiste sur le caractère canonique du transport soleil en rappelant des faits déjà évoqués à la section 2.3 et la remarque 2.3.5.

Théorème 3.2.14 (Transport soleil et noyau lipschitzien). *Soit $\mu, \nu \in \mathcal{P}_1$ telles que $\mu \preceq_C \nu$. Il existe une mesure de probabilité χ sur $\mathcal{I} = \{T \subset \mathbb{R} : T \text{ est fermé, } \inf T = -\infty, \sup T = +\infty\}$ qui représente ν dans le sens où pour tout borélien A*

$$\nu(A) = \int_{\mathcal{I}} P^T(\mu)(A) d\chi(T) =: P^X(\mu)(A). \quad (3.4)$$

L'opération P^X s'obtient à l'aide de $p_X = \int_{\mathcal{I}} dp_T \chi(T)$ si bien que $\mu \cdot (\text{id}, p_X) = \text{Joint}(\mu, p_X) \in \text{Mart}^{\text{NL}}(\mu, \nu)$.

Un choix pour χ est celui de la mesure uniforme sur $(R_u)_{u \in [0,1]}$ où R_u est une famille décroissante pour \subset . Si de plus χ' est une autre mesure du même type, c'est-à-dire associée à une famille décroissante $(R'_u)_{u \in [0,1]}$ avec $\nu = P^{\chi'}(\mu)$ et $\mu \cdot (\text{id}, p_{\chi'}) \in \text{Mart}(\mu, \nu)$, alors $\text{Joint}(\mu, p_X) = \text{Joint}(\mu, p_{\chi'})$ et ce transport est le transport soleil $\mathfrak{Sol}(\mu, \nu)$.

Remarque 3.2.15. *Les autres démonstrations de ce que $\text{Mart}^{\text{NL}} \neq \emptyset$ reposent sur des plongements de Skorohod (celui de Root dans [BHS16], celui de Hobson [Hob98] dans [Low08⁺]). Avec \mathfrak{Sol} , notre point de vue sur la démonstration de Kellerer construit finalement un pont entre la démonstration de Kellerer et celles par plongement de Skorohod.*

3.3 Généralisation du théorème de Kellerer aux dimensions supérieures

Les problèmes 7a et 7b de [HPRY11] interrogent sur la possibilité de généraliser le théorème 3.1.14 au cas de PCOC, soit indexés sur \mathbb{R}^d , soit constitués de mesures vectorielles indexées sur \mathbb{R} . Par exemple, citons le problème 7a : “Let $(X_{t,\lambda}; t, \lambda \geq 0)$ be a two-parameter peacock. Does there exist an associated two-parameter martingale $(M_{t,\lambda}; t, \lambda \geq 0)$?”. Ici *peacock* est un processus plutôt qu'une mesure : il faut remplacer $(X_{t,\lambda})_{t,\lambda}$ par $\text{Loi}(X_{t,\lambda})_{t,\lambda}$ pour permettre une lecture conforme à la définition 3.1.10.

Problème 7a de [HPRY11]

Au théorème 3.3.1 nous répondons négativement au problème 7a, en niant une à une les implications du théorème 3.1.14.

Théorème 3.3.1 ([Jui16a, Theorem 2]). *Soit (\mathcal{T}, \leq) l'un des ensembles $\{0, 1\}^2$, \mathbb{R}_+^2 ou \mathbb{R}^2 muni de son ordre partiel. Pour chacun de ces choix de \mathcal{T} :*

- i Il existe un PCOC indexé par \mathcal{T} qui n'est associé à aucune martingale,*
- ii il existe un PCOC indexé par \mathcal{T} qui est bien associé à une martingale mais à aucune martingale markovienne,*

iii il existe un PCOC indexé par \mathcal{T} qui est bien associé à une martingale markovienne mais à aucune martingale markovienne à noyaux lipschitziens.

Démonstration. Il suffit de traiter du cas $\mathcal{T} = \{0, 1, 1', 2\} \equiv \{0, 1\}^2 \equiv \{0, 1\}^2$ avec 1, 1' comme indices intermédiaires, $0 \equiv (0, 0)$ et $2 \equiv (1, 1)$. Les cas $\mathcal{T} = \mathbb{R}_+^2$ et $\mathcal{T} = \mathbb{R}^2$ s'obtiennent en prenant des peacocks constantes sur quatre morceaux ordonnés de \mathcal{T} .

Les trois constructions sont illustrées sur les figures 3.1, 3.2 et 3.3 sur lesquelles nous préférons indiquer des entiers proportionnels aux probabilités plutôt que les probabilités elles-mêmes (facteur 6 sur la figure 3.1 et 3.2, et 12 sur la figure 3.3). Les choix de marges que nous faisons, ainsi que les propriétés (martingale, martingale markovienne) déterminent entièrement les martingales de $\text{Mart}((\mu_t)_{t \in \{0, 1, 2\}})$ et $\text{Mart}((\mu_t)_{t \in \{0, 1', 2\}})$. En particulier les dilatations utilisées sont parfaitement connues. Nous les représentons aux figures 3.1, 3.2 et 3.3.

Le premier point du théorème 3.3.1 est démontré par le PCOC $(\mu_t)_{t \in \mathcal{T}}$ de la figure 3.1.

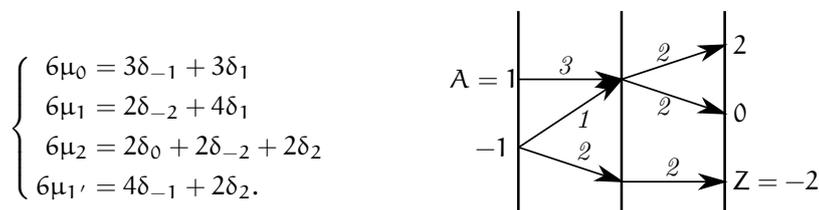


FIGURE 3.1 – La martingale associée à $(\mu_t)_{t \in \{0, 1, 2\}}$ au théorème 3.3.1 (i). Aucune masse ne transite de A vers Z.

Pour le second cas du théorème 3.3.1 on effectue une petite modification.

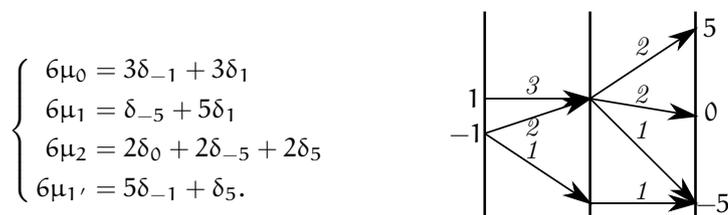


FIGURE 3.2 – Noyaux de transitions de $(X_t)_{t \in \{0, 1, 2\}}$ du théorème 3.3.1.(ii)

Pour le dernier point, les peacocks $(\mu_t)_{t \in \{0, 1, 2\}}$ et $(\mu_t)_{t \in \{0, 1', 2\}}$ ne sont plus symétriques l'une par rapport à l'autre.

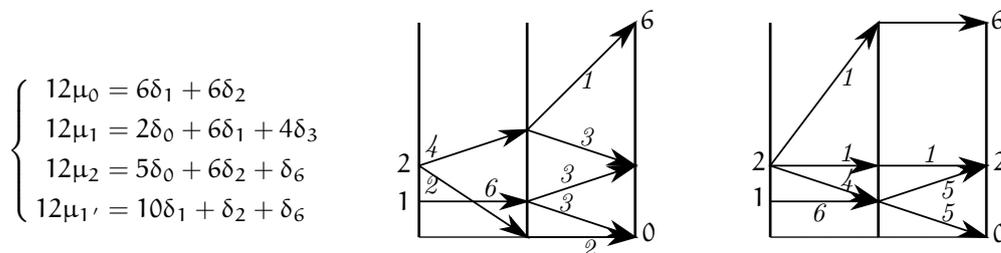


FIGURE 3.3 – Les martingales associées à $(\mu_t)_{t \in \{0, 1, 2\}}$ et $(\mu_t)_{t \in \{0, 1', 2\}}$ pour le théorème 3.3.1(iii).

□

Problème 7b de [HPRY11]

Le problème 7b est le suivant “*Is a \mathbb{R}^n -valued peacock a \mathbb{R}^n -valued 1-martingale ?*” En d’autres termes : Peut-on à tout PCOC sur \mathbb{R}^d associer une martingale à valeur dans \mathbb{R}^d ? Hirsch et Roynette y ont répondu positivement dans [HR13].

Théorème 3.3.2 ([HR13]). *Soit $(\mu_t)_{t \in \mathcal{T}}$ une famille de mesures de $\mathcal{P}_1(\mathbb{R}^d)$ où (\mathcal{T}, \leq) est totalement ordonné. Les énoncés suivants sont équivalents.*

i μ_t est un PCOC,

ii μ_t est associé à une martingale $(X_t)_{t \in \mathcal{T}}$.

Il reste le problème ouvert de déterminer si tout PCOC peut être associé à une martingale markovienne. Cette question était signalée par Kellerer dès l’introduction de [Kel72] : “[...] während die Übertragung der im zweiten Teil enthaltenen Ergebnisse etwa auf den mehrdimensionalen Fall ein offenes Problem darstellt”. Cette équivalence serait celle de (1) et (3) dans le théorème 3.1.14. Nous savons avec le théorème 3.3.2 que (1) et (2) sont équivalents et montrons maintenant dans la proposition 3.3.3 que (1) et (4) ne sont pas équivalents.

Proposition 3.3.3 ([Jui16a, Proposition 1]). *Il existe un PCOC $(\mu_t)_{t \in \mathcal{T}}$ indexé sur $\mathcal{T} = \{0, 1\}$ avec $\mu_t \in \mathcal{P}(\mathbb{R}^2)$ tel qu’aucune martingale markovienne à noyaux lipschitziens ne lui soit associée. Ceci est également valable pour $\mathcal{T} = \mathbb{R}_+$, par exemple (en prenant un PCOC constant par morceaux sur deux morceaux.)*

Démonstration. Soit $\mu_0 = \lambda_{[0,1]} \otimes \delta_0 \in \mathcal{P}(\mathbb{R}^2)$ et k_f le noyau $(x, 0) \mapsto \frac{1}{2}(\delta_{(x, f(x))} + \delta_{(x, -f(x))})$. Soit $\mu_1 := \mu_0 \cdot k_f \in \mathcal{P}_1(\mathbb{R}^2)$. Si $\mu_1 = \mu_0 \cdot k$ pour une autre dilatation $k : \mathbb{R}^2 \rightarrow \mathcal{P}_1(\mathbb{R}^2)$, la projection de k sur l’axe des abscisses est une dilatation envoyant $\lambda_{[0,1]}$ sur elle-même. C’est donc l’identité $\text{id} : x \in \mathbb{R} \mapsto \delta_x \in \mathcal{P}_1(\mathbb{R})$. On a donc $k = k_f$. En choisissant f discontinue comme, par exemple $f = \mathbb{1}_{[1/2, 1]}$, on obtient le résultat : k n’est pas un noyau lipschitzien et $\text{Mart}^{\text{NL}}(\mu_0, \mu_1) = \emptyset$. \square

Le problème suivant reste ouvert. C’est la grande motivation pour trouver de nouvelles idées pour construire des processus markoviens dont les marges sont prescrites, telles que celle de remplacer noyau « lipschitzien » par « croissant » (lemme 3.2.8) ou celle derrière la construction du processus Markov-quantile $\mathfrak{M}\Omega$.

Problème 3.3.4 (Théorème de Kellerer dans \mathbb{R}^d). Peut-on, dans le théorème 3.3.2, demander additionnellement que $(X_t)_{t \in \mathcal{T}}$ soit markovienne ?

3.4 Construction d’une martingale construite à partir de $\mathfrak{R}\mathcal{C}$ pour des peacocks partiellement indexés

Dans cette partie nous considérons un PCOC $(\mu_t)_{t \in \mathcal{T}}$ indexé sur un ensemble partiellement ordonné (\mathcal{T}, \leq) et cherchons une martingale qui lui est associée. Nous avons vu au théorème 3.3.1 que cela ne se fait pas de façon générale. Nous nous restreindrons ici aux PCOC tels que $t \leq t'$ implique $\mu_t \preceq_{\text{CD}} \mu_{t'}$. Rappelons que l’ordre convexe diatomique \preceq_{CD} est apparu à la définition 2.2.15. À la page 48 nous avons également défini $\theta[a, b]$ comme la mesure de probabilité de barycentre \mathfrak{m} concentrée sur $\{a, b\}$. Nous rappelons que si $\theta[a, b] \preceq_{\text{C}} \theta[c, d]$, l’ensemble $\text{Mart}(\theta[a, b], \theta[c, d])$ contient un unique élément, noté $\theta[a, b; c, d]$. Notre résultat est le suivant

Théorème 3.4.1 ([Jui16a, Theorem 4]). *Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné et $(\mu_t)_{t \in \mathcal{T}}$ une famille de mesures dans $\mathcal{P}_1(\mathbb{R})$ croissant pour l’ordre convexe diatomique \preceq_{CD} . Alors il existe une martingale $(X_t)_{t \in \mathcal{T}}$, de loi \mathbb{P} , associée à $(\mu_t)_{t \in \mathcal{T}}$. On peut de plus imposer que les lois jointes $\mathbb{P}^{t, t'}$ soient des transports rideau centrés.*

De plus, si toutes les mesures sont du type $\theta[a, b]$ et \mathcal{T} est totalement ordonné, la loi de $(\mu_t)_{t \in \mathcal{T}}$ est entièrement déterminée.

Principe de la démonstration. Pour simplifier la présentation nous considérons que le barycentre des mesures μ_t est 0.

De la même façon qu'à la page 48 nous décomposons chaque mesure μ_t sous la forme $\mu_t^\alpha := \theta[f^{\mu_t}(\alpha), g^{\mu_t}(\alpha)]$. Du fait de notre hypothèse sur l'ordre diatomique nous obtenons ainsi une famille à un paramètre $\alpha \in]0, 1[$ de peacocks $(\mu_t^\alpha)_{t \in \mathcal{T}}$. Nous allons définir la martingale $(X_t)_{t \in \mathcal{T}}$ comme le mélange des martingales $(X_t^\alpha)_{t \in \mathcal{T}}$ associées à $(\mu_t^\alpha)_{t \in \mathcal{T}}$ pour tout $\alpha \in]0, 1[$. Le processus $(X_t)_{t \in \mathcal{T}}$ de loi $\mathbb{P} = \int_0^1 \mathbb{P}_\alpha d\alpha$ avec $\mathbb{P}_\alpha = \text{Loi}((X_t^\alpha)_{t \in \mathcal{T}})$ est une martingale associée à $(\mu_t)_{t \in \mathcal{T}}$.

Il nous reste donc à associer une martingale $(X_t^\alpha)_{t \in \mathcal{T}}$ à $(\mu_t^\alpha)_{t \in \mathcal{T}}$. Cela se fait en considérant un mouvement brownien standard $(B_t)_{t \geq 0}$ et en définissant les variables aléatoires en fonction de $(B_t)_{t \geq 0}$ par

$$X_t^\alpha = f^{\mu_t}(\alpha) \mathbb{1}_{\tau(f[\mu_t(\alpha)]) > \tau(g[\mu_t(\alpha)])} + g^{\mu_t}(\alpha) \mathbb{1}_{\tau(f[\mu_t(\alpha)]) < \tau(g[\mu_t(\alpha)])}$$

où, pour tout $x \in \mathbb{R}$, $\tau(x)$ est le temps d'atteinte de x par le mouvement brownien. \square

Définition 3.4.2. Soit (\mathcal{T}, \leq) un ensemble partiellement ordonné et $(\mu_t)_{t \in \mathcal{T}}$ une famille de mesures dans $\mathcal{P}_1(\mathbb{R})$ croissant pour l'ordre convexe diatomique \preceq_{CD} . Nous appelons processus rideau centré $\mathfrak{RC}((\mu_t)_{t \in \mathcal{T}})$, le processus défini dans la démonstration du théorème 3.4.1 dont les marges de dimension deux sont des transports rideau centrés.

Remarque 3.4.3. Dans [BHJ18⁺⁺], en préparation, à l'aide d'une notion d'ombres généralisées, nous construisons des martingales définies de façon univoque dont fait partie \mathfrak{RC} . Dans l'article en question, \mathcal{T} est totalement ordonné et le PCOC $(\mu_t)_{t \in \mathcal{T}}$ arbitraire. Le principe de définition est proche de celui des transports ombrés.

Processus quantiles Ω markoviens

En raison du théorème de Kellerer et du problème 3.3.4 les constructions aboutissant à des martingales markoviennes présentent un intérêt certain. Dans [Jui16a, Theorem 5] nous exprimons les conditions nécessaires et suffisantes pour que $\mathfrak{RC}((\mu_t)_{t \in \mathcal{T}})$ soit markovien. L'énoncé est compliqué et difficilement mémorisable si bien que nous préférons présenter seulement le résultat analogue pour les processus quantile, à savoir le théorème 3.4.4 paru dans le même article.

Pour $(\mu_t)_{t \in \mathcal{T}}$ donné, la mesure quantile $\Omega = \Omega((\mu_t)_{t \in \mathcal{T}})$ est la loi du processus $(G_{\mu_t})_{t \in \mathcal{T}}$ défini sur $(]0, 1[, \lambda)$. Nous observons que, pour tout $s < t$ la mesure $\Omega^{s,t}$ est le transport quantile $\Omega(\mu_s, \mu_t)$. Dans certains cas seulement cette mesure est markovienne.

Théorème 3.4.4 (Ω markovien et « X », [Jui16a, Proposition 3]). *La mesure quantile Ω est markovienne si et seulement si pour tout $s \leq t \leq u$ et $\alpha < \alpha' \in]0, 1[$, la conjonction des conditions suivantes*

$$\begin{cases} G_{\mu_s}(\alpha) < G_{\mu_s}(\alpha') \\ G_{\mu_t}(\alpha) = G_{\mu_t}(\alpha') \end{cases}$$

implique $G_{\mu_u}(\alpha) = G_{\mu_u}(\alpha')$. Autrement dit, la condition peut se traduire comme « $G_{\mu_t}(\alpha) = G_{\mu_t}(\alpha')$ implique $\{G_{\mu_s}(\alpha) = G_{\mu_s}(\alpha') \text{ ou } G_{\mu_u}(\alpha) = G_{\mu_u}(\alpha')\}$ ».

La condition pour que Ω soit markovien peut se mémoriser en disant que deux trajectoires associées à $\alpha < \alpha'$ ne forment jamais de « X ». Voir la figure 4.1 pour une illustration d'un processus Ω non-markovien : sur le schéma de gauche les trajectoires bleues rejoignent les trajectoires rouges puis s'en séparent.

3.5 Martingales construites à l'aide du couplage rideau gauche \mathfrak{RG}

Il est tentant de penser que la construction canonique d'un transport martingale que nous avons obtenue avec \mathfrak{RG} au chapitre 1 se décline dans une version (loi de) processus canonique.

En rapport avec le théorème de Kellerer, cela l'est d'autant plus qu'on peut espérer obtenir une martingale markovienne. Dans [Jui18], nous avons fait l'essai de concaténer des transports rideau gauches à partir de marges extraites du PCOC $(\mu_t)_{t \in \mathcal{T}}$ auquel on souhaite associer une martingale. Nous rendons compte ici des théorèmes principaux. Une démarche analogue a été proposée par Henry-Labordère, Tan et Touzi dans [HTT16].

Dans les théorèmes à venir les peacocks seront des familles de mesures de $\mathcal{P}_1(\mathbb{R})$ indexées sur $[0, 1]$. On note $\text{Subdiv}([a, b])$ l'ensemble des *subdivisions* de $[a, b]$. Le *pas* de \mathbb{R} de $\mathbf{R} = \{\mathbf{a} = r_0 < r_1 < \dots < r_m < r_{m+1} = \mathbf{b}\}$ est $\max_{k=0}^m |r_{k+1} - r_k|$. Pour une subdivision $\mathbf{R} \in \text{Subdiv}([0, 1])$ donnée nous associons à $(\mu_t)_{t \in [0, 1]}$ une mesure $\mathbb{P}_{\{\mathbf{R}\}} \in \text{Mart}((\mu_{\lfloor t \rfloor_{\mathbf{R}}}))_{t \in [0, 1]}$ où $\lfloor t \rfloor_{\mathbf{R}} = \max\{r_k \in \mathbf{R} : r_k \leq t, k \leq m\}$. C'est l'unique mesure qui, à la fois

i est markovienne,

ii vérifie $\mathbb{P}_{\{\mathbf{R}\}}^{r_k, r_{k+1}} = \mathfrak{RG}(\mu_{r_k}, \mu_{r_{k+1}})$, où on précise que $\mathbb{P}_{\{\mathbf{R}\}}^{r_k, r_{k+1}}$ est le transport $(\mathbb{P}_{\{\mathbf{R}\}})^{r_k, r_{k+1}} = (\text{proj}^{\{r_k, r_{k+1}\}})_{\#} \mathbb{P}_{\{\mathbf{R}\}}$.

Puisque $\mathbb{P}_{\{\mathbf{R}\}}$ est une loi de martingale dont les marges sont constantes par morceaux, on peut noter que le processus le sera aussi sur chacun des intervalles $[r_k, r_{k+1}[$ et $[r_m, 1]$.

Nous appelons *mesure rideau gauche limite* toute valeur d'adhérence obtenue lorsque $|\mathbf{R}|$ tend vers zéro, c'est-à-dire toute limite d'une suite $(\mathbb{P}_{\{\mathbf{R}_n\}})_{n \in \mathbb{N}^*}$ où $|\mathbf{R}_n| \rightarrow 0$. Cette notion dépend de la topologie choisie, dans cette section la topologie fini-dimensionnelle ou celle de Skorohod.

Remarque 3.5.1. *Des façons différentes d'associer une mesure à une famille $(\mu_t)_{t \in [0, 1]}$ et une subdivision \mathbf{R} sont introduites plus loin. Par exemple, le processus $\mathbb{P}_{\{\mathbf{R}\}}$ introduit à la page 70 n'est pas constant par morceaux.*

On peut aisément montrer qu'une martingale càdlàg est associée à un PCOC continu à droite et possédant une limite à gauche (pour la topologie faible). Notons que tous les PCOC possèdent une limite à gauche en tout point. Pour voir cela on peut considérer la fonction croissante $t \mapsto \int \sqrt{1 + x^2} d\mu_t(x)$ ou la famille $(\mathbf{U}[\mu_t])_{t \in [0, 1]}$ des fonctions potentielles.

Théorème 3.5.2 ([Jui18, Theorem A]). *Soit $(\mu_t)_{t \in [0, 1]}$ un PCOC continu à droite. Pour la topologie fini-dimensionnelle l'ensemble des mesures rideau gauche limites est non vide. Cet ensemble contient au moins une mesure \mathbb{P} pertinente, c'est à dire vérifiant $\mathbb{P} \in \text{Mart}((\mu_t)_{t \in [0, 1]})$ et associée à un processus càdlàg.*

Pour la topologie de Skorohod les mesures rideau gauche limites sont toutes pertinentes (mais il n'est pas démontré que l'ensemble est non-vide).

Le second résultat concerne les PCOC dont les marges sont des mesures uniformes. La figure 3.4 illustre les étapes de construction du processus de loi $\mathbb{P}_{\{\mathbf{R}\}}$ pour f linéaire et \mathbf{R} équirépartie.

Théorème 3.5.3 ([Jui18, Theorem B]). *Soit μ_t la mesure uniforme sur $[-f(t), f(t)]$, où f est strictement positive, croissante et continue sur $[0, 1]$. Il existe une unique mesure rideau gauche limite associée à $(\mu_t)_{t \in [0, 1]}$ et celle-ci est markovienne.*

Le processus associé est continu par morceaux, croissant entre les sauts (à incréments négatifs) vers $-f(t)$ à l'instant t . Les sauts dépendent seulement du temps et non de la position. Ils se produisent avec intensité $2^{-1} d \log(f(t))/dt$.

Pour $f(t) = \exp(2t)/2$, les sauts se produisent en suivant un processus de Poisson standard.

Le troisième résultat concerne des mesures atomiques. On obtient de nouveau l'unicité du processus limite.

Théorème 3.5.4 ([Jui18, Theorem C]). *Soit $(\mu_t)_{t \in [0, 1]}$ un PCOC du type*

$$\mu_t = \sum_{i=1}^n a_i(t) \delta_{x_i(t)}$$

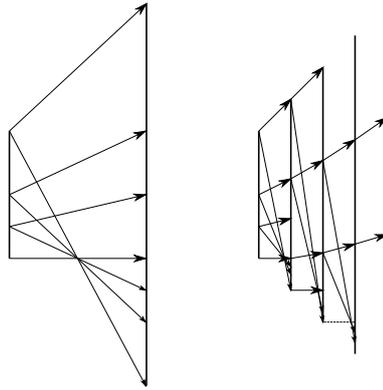


FIGURE 3.4 – Composition de transports rideau gauches entre mesures uniformes.

où $t \mapsto (x_1, \dots, x_n, a_1, \dots, a_n)(t)$ est analytique réelle et satisfait de plus à

$$a_1(t), \dots, a_n(t) > 0 \quad \text{et} \quad x_1(t) < \dots < x_n(t).$$

Alors, il existe une unique mesure rideau gauche limite associée à $(\mu_t)_{t \in [0,1]}$ et celle-ci est markovienne.

Comme énoncé dans le théorème suivant, il n’y a pas unicité du processus limite et la limite peut être markovienne ou non. La figure 3.5 représente le processus contre-exemple décrit dans [Jui18, §6]. Il présente la particularité de posséder des marges μ_t toutes continues. Sur la figure, les trajectoires x_1 et x_2 partent séparément, puis fusionnent, et enfin se séparent de nouveau.

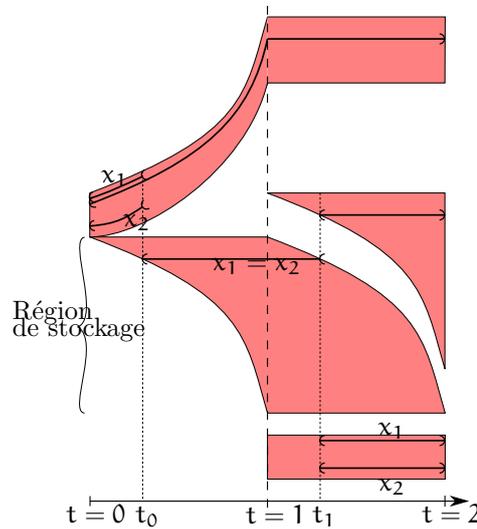


FIGURE 3.5 – Processus rideau limite non-markovien associé à un PCOC de marges absolument continues.

Théorème 3.5.5 ([Jui18, Theorem D]). *Il existe un PCOC $(\mu_t)_{t \in [0,1]}$ et un processus rideau gauche limite $(X_t)_{t \in [0,1]}$ associé à $(\mu_t)_{t \in [0,1]}$ tel que $\text{Loi}(X_t) = \mu_t$ pour tout t , mais $(X_t)_{t \in [0,1]}$ n’est pas markovien. De plus, nous connaissons deux exemples où les mesures de la famille $(\mu_t)_{t \in [0,1]}$ peuvent être choisies toutes discrètes ou toutes continues.*

3.6 La mesure Markov-quantile $\mathfrak{M}\mathfrak{Q}$

Nous avons vu au théorème 3.4.4 que $\mathfrak{Q}(\mu_t)_{t \in \mathbb{R}}$ n'est pas (la mesure d'un processus markovien. Dans cette section nous rendons compte de travaux effectués dans [BouJ18⁺] faisant apparaître $\mathfrak{M}\mathfrak{Q}$, la version *markovifiée* du processus quantile. Le processus $\mathfrak{M}\mathfrak{Q}$ conserve des propriétés de \mathfrak{Q} mais il est en plus markovien. Commençons par considérer la *markovification* d'un processus en général.

Markovification d'un processus à noyaux croissants ou lipschitziens

Dans la section précédente nous avons expliqué notre objectif de créer un processus markovien canonique, dont les marges sont prescrites, à l'aide de transports canoniques (dans ladite section, le transport rideau gauche). Nous poursuivons ici avec cette même visée en prenant comme plans de transport de référence les lois jointes 2-dimensionnelles d'une loi processus P qui nous est donnée. L'objectif est de bâtir une famille compatible de transports, donnant ainsi une mesure « markovifiée » P' , la mesure P rendue markovienne. La markovifiée d'une mesure P markovienne est P elle-même. Nous nous préoccupons plus tard de la markovifiée de $\mathfrak{Q}((\mu_t)_{t \in \mathbb{R}})$ qui sera unique et nommée *mesure Markov-quantile*. Si nous ne connaissons pas d'autres résultats généraux d'unicité sur la markovifiée, voici toutefois un résultat d'existence général.

Théorème 3.6.1 (Markovification). *Soit $(\mu_t)_{t \in \mathbb{R}}$ une famille de mesures dans $\mathcal{P}(\mathbb{R})$ et $P \in \text{Marg}((\mu_t)_{t \in \mathbb{R}})$. Si P est à noyaux croissants ou bien à noyaux lipschitziens, il existe une mesure markovienne $P' \in \text{Marg}((\mu_t)_{t \in \mathbb{R}})$ dont tous les transports $(P'^{s,t})_{s < t}$ sont du type*

$$P'^{s,t} = \lim_{n \rightarrow \infty} P_{[\mathbb{R}_n^{s,t}]}^{s,t} = \lim_{n \rightarrow \infty} P^{s,r_1^{(n)}} \cdot \dots \cdot P^{r_m^{(n)},t}$$

où, pour tout $s < t$ et n , $\mathbb{R}_n^{s,t} = \{s = r_0^{(n)} < r_1^{(n)} < \dots < r_m^{(n)} < r_{m+1}^{(n)} = t\}$ est une subdivision de $[s, t]$. On peut de plus exiger de la suite $(\mathbb{R}_n)_{n \in \mathbb{N}}$ qu'elle satisfasse à $|\mathbb{R}_n| \rightarrow 0$.

Ce théorème, du moins pour les noyaux lipschitziens, aurait pu être écrit par Kellerer puisque ce n'est pas bien plus qu'un corollaire de son travail. La démonstration s'appuie sur les lemmes 3.2.2 et 3.2.8.

L'idée du nom « markovification » est que les transitions infinitésimales de P' ressemblent à celles de P , si bien qu'on estime que P' conserve quelque chose de P , tout en étant markovienne.

Markovification du processus quantile

La notation $P_{[R]}$ est apparue dans le théorème 3.6.1 par le biais du transport $P_{[\mathbb{R}_n^{s,t}]}^{s,t}$ associé à la mesure P et à la subdivision $\mathbb{R}_n^{s,t} \subset [s, t]$ sans que l'on ait besoin de définir le sens de la mesure $P_{[\mathbb{R}_n^{s,t}]}$ elle-même. Le voici donc. Pour une subdivision $R = \{r_1 < \dots < r_m\}$ donnée nous associons à $(\mu_t)_{t \in \mathbb{R}}$ la mesure $P_{[R]} \in \text{Mart}((\mu_t)_{t \in \mathbb{R}})$ telle que

$$P_{[R]} = P^{]-\infty, r_1]} \circ P^{[r_1, r_2]} \circ \dots \circ P^{[r_{m-1}, r_m]} \circ P^{[r_m, +\infty[}$$

où on généralise la concaténation \circ , apparue en définition 3.1.5, de la seule façon envisageable. Plus précisément la mesure P satisfait à

$$(\text{proj}^S)_{\#} P_{[R]} = P^{s_1^0, \dots, s_{n_0}^0, r_1} \circ P^{r_1, s_1^1, \dots, s_{n_1}^1, r_2} \circ \dots \circ P^{r_m, s_1^m, \dots, s_{n_m}^m}$$

où $S = \{s_1^0, \dots, s_{n_0}^0, r_1, s_1^1, \dots, s_{n_1}^1, r_2, \dots, r_m, s_1^m, \dots, s_{n_m}^m\}$.

En terme de processus, le principe est que la loi de $P_{[R]}$ est celle de P à cela près qu'aux instants $r_k \in R$ le passé est rendu indépendant du futur conditionnellement au présent. Les instants de R peuvent être considérés comme des « coupes » du temps dans lesquelles on injecte et impose de l'indépendance au processus.

Notons spécifiquement que si P vaut Ω et $R = \{r_1, \dots, r_m\} \subset]s, t[$, on a $\Omega_{[R]}^{s,t} = \Omega^{s,r_1} \cdot \Omega^{r_1,r_2} \cdot \dots \cdot \Omega^{r_{m-1},r_m} \cdot \Omega^{r_m,t}$.

Pour toute mesure quantile $\Omega \in \text{Marg}((\mu_t)_{t \in \mathcal{T}})$ indexée par un ensemble arbitraire \mathcal{T} l'inégalité de Hoeffding-Fréchet, rappelée en définition 1.2.11 pour deux indices, se prolonge : si t_1, \dots, t_m sont des éléments de \mathcal{T} alors pour tout $x \in \mathbb{R}^m$, $\Omega^{t_1, \dots, t_m}]-\infty, x] \geq P^{t_1, \dots, t_m}]-\infty, x]$. On note cela $\Omega^{t_1, \dots, t_m} \preceq_{\text{lo}} P^{t_1, \dots, t_m}$ à l'aide de \preceq_{lo} , l'ordre du plus petit orthant, conformément à la définition 1.2.8. De façon générale si $(X_s, X_t) \sim P$ et $(Y_s, Y_t) \sim Q$ sont des éléments du même espace de transports $\text{Marg}(\mu, \nu)$ alors $P \preceq_{\text{lo}} Q$ est équivalent à

$$\mathbb{P}(X_t \leq y | X_s \leq x) \geq \mathbb{P}(Y_t \leq y | Y_s \leq x) \quad \text{pour tout } (x, y) \in \mathbb{R}^2.$$

Théorème 3.6.2 (Existence et propriétés de $\mathfrak{M}\Omega$, [BouJ18⁺, Theorem A]). (a) Soit $(\mu_t)_{t \in \mathbb{R}}$ une famille de mesures de probabilités. Alors, il existe une unique mesure $\mathfrak{M}\Omega \in \text{Marg}((\mu_t)_{t \in \mathbb{R}})$ telle que :

- (i) $\mathfrak{M}\Omega$ est markovienne,
- (ii) $\mathfrak{M}\Omega$ est à noyaux croissants,
- (iii) Les transports $(\mathfrak{M}\Omega^{s,t})_{s < t}$ de $\mathfrak{M}\Omega$ sont des éléments minimaux pour \preceq_{lo} dans le sous-ensemble des mesures $P \in \text{Marg}((\mu_t)_{t \in \mathbb{R}})$ satisfaisant à (i) et (ii), au sens où $\mathfrak{M}\Omega^{s,t} \preceq_{\text{lo}} P^{s,t}$ pour tout $s \leq t$.

Nous appellerons cette mesure mesure Markov-quantile.

(b) La mesure $\mathfrak{M}\Omega$ est aussi l'unique mesure markoviniée de Ω .

Enfinement :

(c) Si $(\mu_t)_{t \in \mathbb{R}}$ est croissante pour \preceq_{sto} , il existe un processus $X = (X_t)_{t \in \mathbb{R}} : \Omega \rightarrow \mathbb{R}^{\mathbb{R}}$ de loi $\mathfrak{M}\Omega$ dont les trajectoires sont croissantes, c'est-à-dire tel que $t \mapsto X_t(\omega)$ est croissante pour tout $\omega \in \Omega$.

Remarque 3.6.3. — Rappelons que l'énoncé (b) sur la markovification de Ω signifie que $\mathfrak{M}\Omega$ est l'unique mesure markovienne telle que pour tout $s < t$, il existe une suite $(R_n^{s,t})_{n \in \mathbb{N}}$ de subdivisions de $]s, t]$ pour laquelle $\mathfrak{M}\Omega^{s,t}$ est la limite de la suite $(\Omega_{[R_n^{s,t}]}^{s,t})_{n \in \mathbb{N}}$.

- L'énoncé (c) contient notre extension du théorème de Kellerer à l'ordre stochastique (théorème 3.2.4) et son corollaire, le 3.2.12. Il donne une précision supplémentaire en exhibant $\mathfrak{M}\Omega$ comme processus markovien croissant. Sous (c) on conserve ainsi la propriété de croissance du processus quantile en lui ajoutant la propriété de Markov.
- Toujours par rapport au théorème 3.2.4 et à (c), nous ne connaissons en revanche pas de processus qui pour \preceq_C (et un PCOC $(\mu_t)_{t \in \mathbb{R}}$) se construit d'une façon analogue à celle dont nous construisons $\mathfrak{M}\Omega$.

Nous avons en fait non seulement une convergence vers les transports $(\mathfrak{M}\Omega^{s,t})_{s < t}$ comme au point (a) du théorème mais aussi une convergence des mesures, impliquant en particulier une unique suite $(R_n)_n$ plutôt qu'une suite $(R_n^{s,t})_n$ par couple (s, t) .

Théorème 3.6.4. Il existe une suite $(R_n)_{n \in \mathbb{N}}$ d'ensembles finis de \mathbb{R} , croissante pour l'inclusion, telle que $\Omega_{[R_n]} \in \text{Marg}((\mu_t)_{t \in \mathbb{R}})$ converge vers $\mathfrak{M}\Omega$.

De plus, si pour une suite de subdivision $(R_n)_{n \in \mathbb{N}}$, la suite $(\Omega_{[R_n]})_{n \in \mathbb{N}}$ possède une limite qui est markovienne, alors cette limite est $\mathfrak{M}\Omega$.

Le résultat des deux théorèmes précédents apparaît comme un aboutissement de nos tentatives des sections 3.4 et 3.5. Aucune restriction n'est faite sur la famille $(\mu_t)_{t \in \mathbb{R}}$ (sauf le fait que \mathbb{R} est totalement ordonné), la mesure limite existe toujours et elle conserve certaines propriétés de Ω (comme au théorème 3.6.2(c) et (a)(iii) ou dans le chapitre 4 au théorème 4.3.4). Par rapport au théorème de Kellerer qui est un pur résultat d'existence, on a même l'unicité de la mesure limite.

Construction du processus Markov-quantile

Nous allons expliquer comment définir $\mathfrak{M}\Omega$. Cela requiert des étapes et des démonstrations que nous ne précisons pas toutes. Pour simplifier les notations nous notons $G_t := G[\mu_t]$ la fonction quantile de μ_t et F_t sa fonction de répartition. La construction de $\mathfrak{M}\Omega$ est fondée sur une famille de transports $L_{[s,t]} \in \text{Marg}(\lambda, \lambda)$ où $\lambda = \lambda|_{[0,1]}$ est la mesure de Lebesgue restreinte à $[0, 1]$. Ce sont les marges de dimension 2 d'une mesure $\text{Lev} \in \text{Marg}(\lambda_t)_{t \in \mathbb{R}}$ (voir définition 3.6.13) que nous nommerons *la mesure sur les niveaux de quantiles*. Ici chaque λ_t est une copie de λ indexée par $t \in \mathbb{R}$. Le lien entre Lev et $\mathfrak{M}\Omega$ est le suivant : G_t transporte $([0, 1], \lambda_t)$ sur $(\mathbb{R} \cup \{\pm\infty\}, \mu_t)$. On pose $G = \prod_{t \in \mathbb{R}} G_t$ si bien que $G_{\#}$ transporte $\text{Marg}((\lambda_t)_{t \in \mathbb{R}})$ sur $\text{Marg}((\mu_t)_{t \in \mathbb{R}})$. Le processus Markov-quantile sera donc défini à partir de celui sur les niveaux de quantile comme suit :

$$G_{\#} \text{Lev} = \mathfrak{M}\Omega. \quad (3.5)$$

Noyaux de transitions sur $[0, 1]$

Les transports quantiles entre λ et μ_t nous permettent de passer de l'espace des niveaux de quantile à celui des quantiles.

Notation 3.6.5 ($q_r, k_r, {}^t k_r$). Pour tout $r \in \mathbb{R}$, nous posons $q_r = \Omega(\lambda, \mu_r)$; et, donc, ${}^t q_r = \Omega(\mu_r, \lambda)$. La désintégration de ces transports donnent les noyaux

$$k_r : (\alpha, B) \mapsto \delta_{G_r(\alpha)}(B)$$

et

$${}^t k_r(x, \cdot) = \begin{cases} \delta_{F_r(x)} & \text{si } \mu_r(x) = 0 \\ (\mu_r(x))^{-1} \lambda|_{]F_r(x^-), F_r(x)[} & \text{si } \mu_r(x) > 0. \end{cases}$$

Remarque 3.6.6. (a) On a ${}^t q_s \cdot q_t = \text{Joint}(\mu_s, {}^t k_s \cdot k_t) = \Omega^{s,t} = \Omega(\mu_s, \mu_t)$.

(b) En revanche, $q_r \cdot {}^t q_r \neq \text{Id}_{\lambda,2} \in \text{Marg}(\lambda, \lambda)$. En effet $k_r \cdot {}^t k_r$ envoie chaque niveau de quantile $\alpha \in]0, 1[$ sur lui-même sauf si $G_r(\alpha)$ est un atome de μ_r . Plus précisément on a presque sûrement

$$k_r \cdot {}^t k_r(\alpha) = \begin{cases} \delta_{\alpha} & \text{if } \mu_r(G_r(\alpha)) = 0 \\ (\alpha^+ - \alpha^-)^{-1} \lambda|_{\alpha^-, \alpha^+[} & \text{if } \alpha \in]\alpha^-, \alpha^+[, \end{cases}$$

où $]\alpha^-, \alpha^+[$ est un des intervalles $A_{r,x}$ défini tel que suit.

Notation 3.6.7 (Niveaux atomiques, ℓ_r). (a) On pose $A_{r,x}$ l'intervalle ouvert $]F_r(x^-), F_r(x)[$ dont les niveaux de quantiles sont réunis par G_r sur un atome x de μ_r . L'ensemble A_r est la réunion $\bigcup_{\mu_r(x) > 0} A_{r,x} \subset]0, 1[$ des « niveaux atomiques » de μ_r .

(b) On pose $\ell_r := k_r \cdot {}^t k_r$.

Remarque 3.6.8. Les mesures $\theta \ll \lambda$ sont transportées par ℓ_r de la façon suivante : $\theta \cdot \ell_r$ coïncide avec θ sur $]0, 1[\setminus A_r$, et sur $A_{r,x}$ sa densité est constante et déterminée par la masse, $\theta(A_{r,x})$, qui est conservée. Cette densité vaut donc $(\alpha^+ - \alpha^-)^{-1} \theta(A_{r,x})$. En ce qui concerne les fonctions de répartition, $F[\theta \cdot \ell_r]$ est continue, vaut $F[\theta]$ sur $]0, 1[\setminus A_{r,x}$ et est affine sur chaque composante connexe $A_{r,x}$ de A_r .

Le produit $\Omega_{[R]}^{s,t} = \Omega^{s,r_1} \cdot \Omega^{r_1,r_2} \cdot \dots \cdot \Omega^{r_{m-1},r_m} \cdot \Omega^{r_m,t}$ qui apparait dans les théorèmes 3.6.2 et 3.6.4 peut maintenant être analysé plus finement. Son noyau est

$$\begin{aligned} & ({}^t k_s \cdot k_{r_1}) \cdot ({}^t k_{r_1} \cdot k_{r_2}) \cdot \dots \cdot ({}^t k_{r_{m-1}} \cdot k_{r_m}) \cdot ({}^t k_{r_m} \cdot k_t) \\ &= {}^t k_s \cdot (k_{r_1} \cdot {}^t k_{r_1}) \cdot \dots \cdot (k_{r_m} \cdot {}^t k_{r_m}) \cdot k_t \\ &= {}^t k_s \cdot \ell_{r_1} \cdot \dots \cdot \ell_{r_m} \cdot k_t. \end{aligned} \quad (3.6)$$

Suite à (3.6) on introduit ℓ_R et L_R pour $R = \{r_1, \dots, r_m\}$:

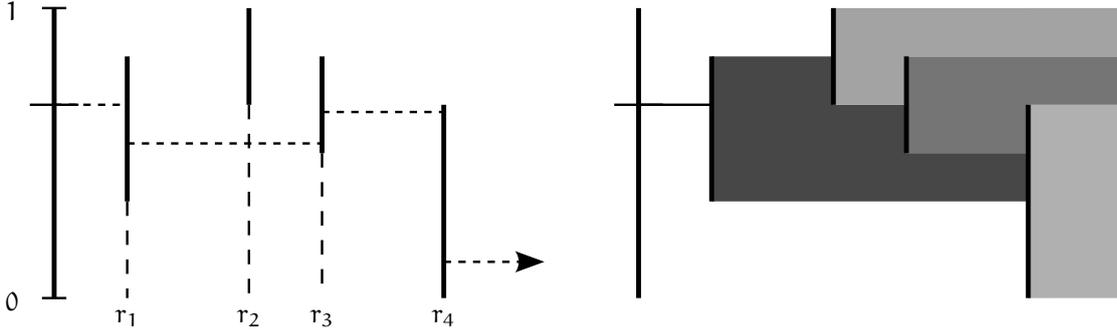


FIGURE 3.6 – Composition de noyaux ℓ_r

Notation 3.6.9. Soit $R = \{r_1, \dots, r_m\} \subset \mathbb{R}$ une subdivision. On notera ℓ_R le noyau $\ell_{r_1} \cdot \ell_{r_2} \dots \ell_{r_m}$ et L_R le transport $\text{Joint}(\lambda, \ell_R) \in \text{Marg}(\lambda, \lambda)$.

On remarque $\ell_{\{r\}} = \ell_r$ et que pour tout R , $\lambda \cdot \ell_R = \lambda$. Il apparait de plus que ℓ_R ne dépend que de $(A_r)_{r \in R}$ et non pas des mesures $(\mu_r)_{r \in R}$. Le lemme suivant étudie le comportement des noyaux ℓ_R par rapport à l'ordre stochastique.

Lemme 3.6.10. Soit $R \subset \mathbb{R}$ un ensemble fini et μ, ν des mesures absolument continues par rapport à λ .

- (a) Si $\theta \preceq_{sto} \theta'$, alors $\theta \cdot \ell_R \preceq_{sto} \theta' \cdot \ell_R$, c'est-à-dire que L_R est à noyaux croissants.
- (b) Si θ possède une densité décroissante, alors la densité de $\theta \cdot \ell_R$ est également décroissante. De plus, $\theta \preceq_{sto} \theta \cdot \ell_R$.

Nous rappelons que l'ordre du plus petit orthant \preceq_{lo} a été introduit à la définition 1.2.8 en même temps que \preceq_{sto} . Nous spécifions maintenant une équation qui les relie.

Remarque 3.6.11. On peut remarquer que pour $\pi = \text{Joint}(\mu, k)$ on a $F[\pi](x, y) = F[\mu]_{[-\infty, x]} \cdot k(y)$. Ainsi

$$\forall (x, y) \in \mathbb{R}^2, \mu_{[-\infty, x]} \cdot k \preceq_{sto} \mu_{[-\infty, x]} \cdot k'$$

est équivalent à

$$\text{Joint}(\mu, k) \preceq_{lo} \text{Joint}(\mu, k').$$

Plus spécifiquement pour $L_R = \text{Joint}(\lambda, \ell_R)$,

$$\forall (x, y) \in [0, 1]^2, \lambda_{[0, x]} \cdot \ell_R \preceq_{sto} \lambda_{[0, x]} \cdot \ell_{R'}$$

est équivalent à

$$L_R \preceq_{lo} L_{R'}.$$

Nous sommes maintenant en mesure d'énoncer une définition plus générale de L_R et ℓ_R . Celle-ci nécessite des démonstrations faites dans [BouJ18⁺]. Le lemme 3.6.10 et la remarque 3.6.11 en sont des éléments.

Définition 3.6.12. Soit R une partie quelconque de \mathbb{R} . Nous notons L_R le supremum pour \preceq_{lo} de l'ensemble

$$\{L_{R'} \in \text{Marg}(\lambda, \lambda) : R' \text{ fini } \subset R\}.$$

En d'autres termes il existe un élément M de $\text{Marg}(\lambda, \lambda)$ tel que

- $\forall R' \text{ fini } \subset R, L_{R'} \preceq_{lo} M$,
- Si M' vérifie la ligne précédente, alors $M \preceq_{lo} M'$.

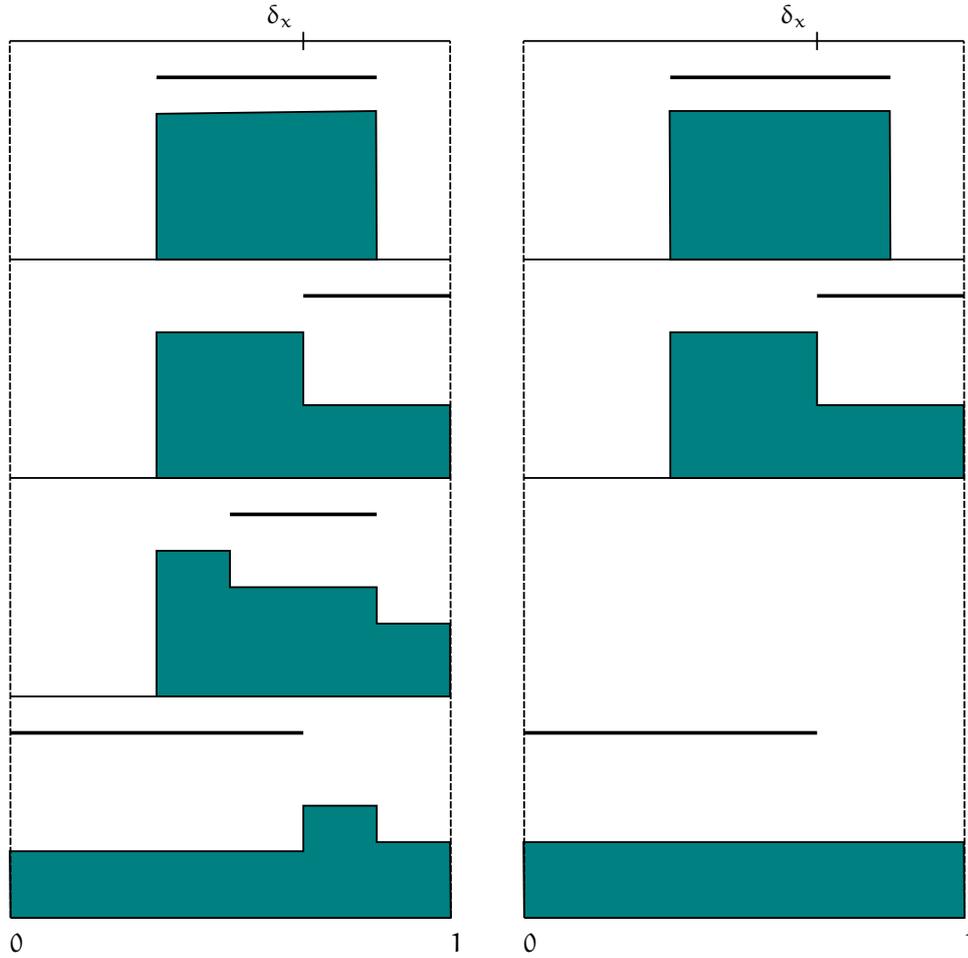


FIGURE 3.7 – Composition des noyaux ℓ_r , agissant sur une mesure de Dirac δ_x . Sur chaque colonne, l'intervalle horizontal est l'espace des niveaux de quantile $]0, 1[$. De haut en bas, les barres horizontales représentent les intervalles A_{r_i} successifs et après chacun la densité de la mesure de δ_x à laquelle a été appliqué $\ell_{r_1}, \dots, \ell_{r_i}$. Sur le côté droit A_{r_3} est enlevé.

Nous notons ce transport L_R et ℓ_R le noyau k_{L_R} associé.

Cette définition est en cohérence avec la notation 3.6.9. Nous concluons par la définition de $\mathfrak{M}\Omega$. Celle-ci aussi requiert des étapes de démonstration qu'on trouve dans [BouJ18⁺]. Elle repose sur le corollaire 3.1.8.

Définition 3.6.13. La famille $(L_{[s,t]})_{s<t}$ est consistante. Nous appelons *mesure sur les niveaux de quantile* et notons Lev la mesure dont les marges de dimension 2 sont les transports $(L_{[s,t]})_{s<t}$.

La mesure Markov-quantile est définie par l'équation 3.5, que nous rappelons ici :

$$G_{\#}\text{Lev} = \mathfrak{M}\Omega. \quad (3.7)$$

Éclaircissements et compléments

À l'aide des figures 3.6–3.8, nous illustrons le principe de la construction de Lev (et donc de $\mathfrak{M}\Omega$). Nous n'allons pas directement définir $\mathfrak{M}\Omega^{s,t}$ comme la limite de transports $\Omega_{[R_n]}^{s,t}$ pour une suite $(R_n)_n$ choisie. Nous procédons différemment en définissant $\mathfrak{M}\Omega^{s,t}$ comme un supremum

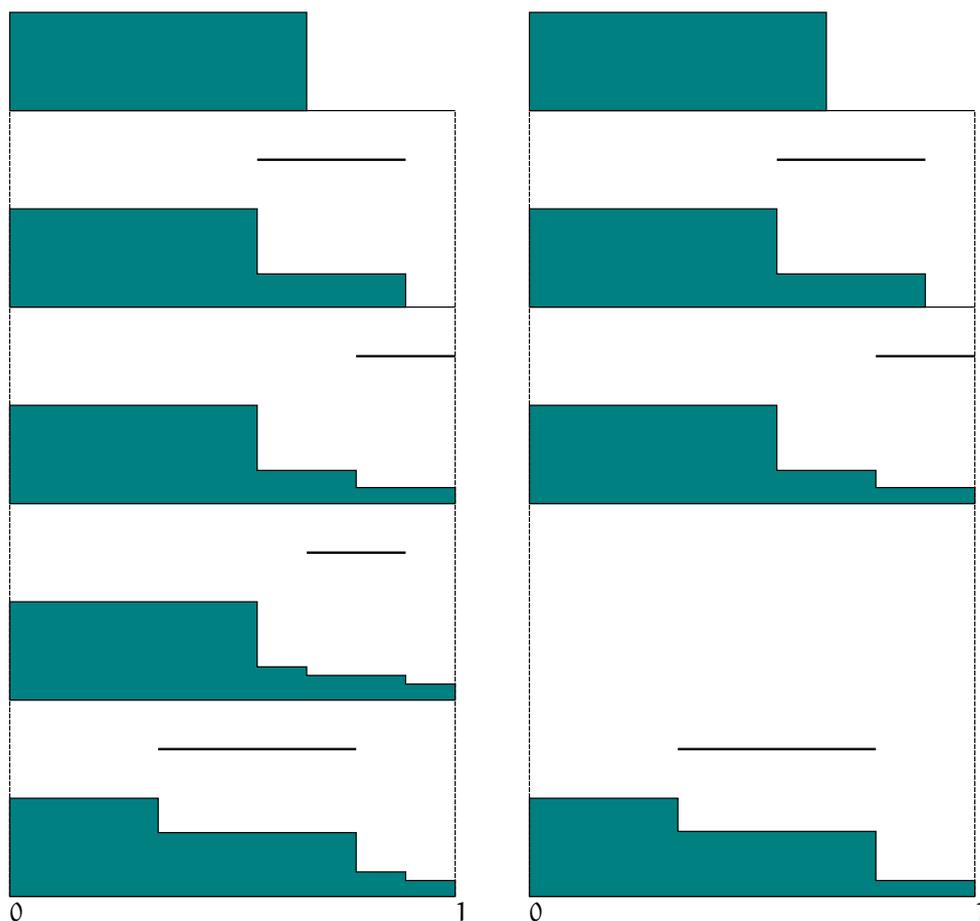


FIGURE 3.8 – Les noyaux l_r de la figure 3.7 agissant sur la mesure de densité $\frac{1}{x} \mathbb{1}_{]0,x[}$. Le cadre est le même que celui de la figure 3.7.

défini sur l'ensemble $\{\Omega_{\mathbb{R}}^{s,t} : \mathbb{R} \text{ fini et } \mathbb{R} \subset]s, t[\}$. Comme nous l'avons vu, cela est techniquement réalisé sur l'espace des niveaux de quantile où on cherche le supremum des $\{\ell_{\mathbb{R}} : \mathbb{R} \text{ fini et } \mathbb{R} \subset]s, t[\}$ pour un ordre approprié. Comme nous l'avons vu il s'agit de \preceq_{sto} pour les mesures (et les noyaux) sur \mathbb{R} , ce qui se transfère à \preceq_{lo} pour les transports.

Observons d'abord la figure 3.6 où est illustrée la façon dont $\ell_{\mathbb{R}}$ agit sur les mesures θ concentrée sur $[0, 1]$. Nous le faisons pour $\mathbb{R} = \{r_1, r_2, r_3, r_4\}$ et $\theta = \delta_x$ une masse de Dirac. Le segment vertical à gauche est l'espace des niveaux de quantile. On suppose que chaque μ_{r_i} comporte exactement un unique atome qui est représenté à l'abscisse r_i par une barre verticale, l'intervalle A_{r_i} correctement positionnée, de longueur la masse de cet atome. Puisque $x \in A_{r_1}$, $\delta_x \cdot \ell_{r_1}$ est simplement la mesure uniforme sur A_{r_1} . À la suite ℓ_{r_2} ne modifie pas cette nouvelle mesure en dehors de A_{r_2} et étale uniformément le reste de la masse sur cet intervalle. La figure de gauche représente une trajectoire possible d'un élément de masse partant de x . Celui-ci est déplacé selon une chaîne de Markov inhomogène dont les transitions sont les ℓ_{r_i} . Dans un premier temps x est remplacé à l'instant r_1 par x' choisi uniformément sur A_{r_1} . Comme $x' \notin A_{r_2}$ l'obstacle suivant est A_{r_3} à l'instant r_3 . On choisit donc alors le nouveau niveau x'' uniformément sur A_{r_3} . Le noyau ℓ_{r_4} le déplace sur $x''' \in A_{r_4}$. Le second dessin montre successivement, à l'aide de niveaux de gris proportionnels à la densité, les mesures δ_x , puis $\delta_x \cdot \ell_{r_1}$, puis $\delta_x \cdot \ell_{r_1} \cdot \ell_{r_2} \dots$

Nous observons ainsi que chaque ℓ_{r_i} étale encore plus la masse $\delta_x \cdot \ell_{r_1} \dots \ell_{r_{i-1}}$ déjà étalée au cours des étapes précédentes. Quantitativement, la variation totale de la densité, par exemple,

est réduite à chaque étape. On ne pourra jamais étaler plus si on obtient λ la mesure sur laquelle l'action des noyaux est triviale. Une idée naturelle est de voir si pour $\mathbf{R}' \subsetneq \mathbf{R}$ et $\theta \ll \lambda$, la densité de $\theta \cdot \ell_{\mathbf{R}}$ pourrait s'approcher de $\mathbb{1}_{[0,1]}$, pour une distance adéquate, plus près que ne l'est $\theta \cdot \ell_{\mathbf{R}'}$.

Malheureusement, si c'est bien le cas lorsque \mathbf{R}' est le « début » de \mathbf{R} , ce n'est pas vrai en général. À la figure 3.7 nous montrons du haut vers le bas la mesure δ_x et les densités de

- $\delta_x \cdot \ell_{r_1}$, $\delta_x \cdot \ell_{r_1} \cdot \ell_{r_2}$, $\delta_x \cdot \ell_{r_1} \cdot \ell_{r_2} \cdot \ell_{r_3}$, et $\delta_x \cdot \ell_{r_1} \cdot \ell_{r_2} \cdot \ell_{r_3} \cdot \ell_{r_4}$ (sur la gauche),
- $\delta_x \cdot \ell_{r_1}$, $\delta_x \cdot \ell_{r_1} \cdot \ell_{r_2}$, et $\delta_x \cdot \ell_{r_1} \cdot \ell_{r_2} \cdot \ell_{r_4}$ (sur la droite),

Si $x = \frac{1}{2}$, $A_{r_1} =]\frac{1}{3}, \frac{5}{6}[$, $A_{r_2} =]\frac{2}{3}, 1[$, $A_{r_3} =]\frac{1}{2}, \frac{5}{6}[$, $A_{r_4} =]0, \frac{2}{3}[$, alors $\delta_x \cdot \ell_{\{r_1, r_2, r_3, r_4\}} = \lambda$, c'est-à-dire que sa densité est $\mathbb{1}_{]0,1]}$, quand bien même $\delta_x \cdot \ell_{\mathbf{R}} \neq \lambda$.

Notre solution est d'observer l'action de $\ell_{\mathbf{R}}$ sur les mesures de densité $\theta = \frac{1}{x} \mathbb{1}_{[0,x]}$. Quels que soient les éléments de \mathbf{R}' prélevés dans \mathbf{R} , on aura $\theta \cdot \ell_{\mathbf{R}'} \preceq_{\text{sto}} \theta \cdot \ell_{\mathbf{R}}$. Dans l'exemple de la figure 3.8, avec les mêmes noyaux ℓ_{r_i} qu'à la figure 3.7, on agit sur une telle mesure θ pour $x = \frac{2}{3}$. On obtient bien $\theta \cdot \ell_{r_1} \cdot \ell_{r_2} \cdot \ell_{r_3} \cdot \ell_{r_4} \succeq_{\text{sto}} \theta \cdot \ell_{r_1} \cdot \ell_{r_2} \cdot \ell_{r_4}$.

Le théorème 3.6.4 est une amélioration du théorème 3.6.2. Dans [BouJ18⁺, Theorem 4.21] nous disposons d'une version plus forte encore. En voici quelques aspects.

Remarque 3.6.14 (Précisions sur $(\mathbf{R}_n)_{n \in \mathbb{N}}$ dans le théorème 3.6.4). *(a) Il s'avère que pour $(\mathbf{R}_n)_n$ une suite croissante, la convergence de $\mathfrak{Q}_{[\mathbf{R}_n]}$ vers $\mathfrak{M}\mathfrak{Q}$ ne dépend que de $\mathbf{R} = \cup_n \mathbf{R}_n$. Peu importe la façon dont \mathbf{R} est décrite comme une réunion, on obtiendra $\mathfrak{m}\mathfrak{q}$ à la limite.*

(b) Dans le calcul du transport $\mathfrak{Q}_{[\mathbf{R}]}^{s,t}$ de nos théorèmes ainsi que du noyau $\ell_{\mathbf{R}}$, l'influence des instants r_i pour lesquels μ_{r_i} ne possède pas d'atome est nulle. Ainsi nous pouvons nous contenter de prendre des instants « atomiques » dans \mathbf{R} . Maintenant il s'avère que tous les instants atomiques n'ont pas le même statut. Cela est important car $\mathbf{R} = \cup_n \mathbf{R}_n$ ne peut contenir qu'un ensemble dénombrable d'instants alors que les instants atomiques forment en général un ensemble indénombrable.

– On distingue dans [BouJ18⁺] des instants atomiques essentiels que l'on sait caractériser (voir [BouJ18⁺, Definition 4.25]). La suite $\mathfrak{Q}_{[\mathbf{R}_n]}$ ne tendra vers $\mathfrak{M}\mathfrak{Q}$ que si tous ces instants sont dans \mathbf{R} .

– Les autres instants atomiques sont inessentiels. Pour tout ensemble fini de ceux-ci il existe un ensemble \mathbf{R} convenable qui évite chacun d'entre eux. Ainsi, l'intersection de tous les ensembles \mathbf{R} satisfaisant à la propriété de convergence est exactement l'ensemble des instants atomiques essentiels.

La partie 6 de [BouJ18⁺] fournit des exemples de processus Markov-quantile dans des situations où l'ensemble $\{(t, \alpha) \in \mathbb{R} \times [0, 1] : \alpha \in A_t\}$ est propice à l'étude directe. Pour ne donner qu'un exemple, pour $\mu_t = \text{Poisson}(t)$ le processus Markov-quantile est simplement le processus de comptage de Poisson.

Développements autour de l'équation de continuité

Comme l'ont démontré Jordan, Kinderlehrer et Otto [JKO98], Benamou et Brenier [BB00], puis d'autres dont Ambrosio, Gigli et Savaré [AGS08], l'équation de continuité est le bon cadre technique pour étudier le transport optimal, non pas entre deux mesures éloignées, mais au niveau infinitésimal, c'est-à-dire entre une mesure μ et une mesure mobile μ_h dont la distance à μ est proportionnelle à h lorsque h tend vers zéro. Ce type d'affirmation peut faire penser aux transports infinitésimaux que nous avons vus apparaître en dérivant les ombres, selon le modèle de la section 2.3, il s'agit toutefois d'une chose bien différente.

Le principe général est le suivant. On dispose d'une courbe $(\mu_s)_{s \in \mathbb{R}}$. Celle-ci doit être vue comme l'analogie des marges à transporter dans le problème initial. Il y a autant de problèmes de transports que d'instant s . Les plans de transports sont remplacés par des champs de vecteurs $(V_s)_{s \in \mathbb{R}}$. Un vecteur à l'instant s remplace une route pour le transport infinitésimal de l'instant s entre μ_s et cette même mesure déplacée de façon infinitésimale. La contrainte des marges est traduite par

$$\frac{d\mu_s}{ds} + \operatorname{div}(\mu_s V_s) = 0. \quad (4.1)$$

Le vecteur V_s est optimal lorsque $\int \|V_s\|^2 d\mu_s$ atteint sa borne inférieure sur l'ensemble des vecteurs admissibles pour (4.1). Enfin on retrouve une sorte de théorème de Brenier dans le fait qu'un champ de vecteurs est optimal si et seulement si il est « de type gradient », dans un sens précis rappelé à la proposition 4.2.4 (dans le cas du groupe de Heisenberg \mathbb{H}).

Dans ce chapitre nous proposons de présenter certains de nos résultats autour de l'équation de continuité, dans un contexte probabiliste d'une part, en lien avec le processus Markov-quantile, dans un contexte géométrique d'autre part, en lien avec la transformation de Gigli et Mantegazza introduite dans [GM14].

Notre *développement probabiliste* poursuit notre étude des chapitres précédents. Il est possible de définir un processus, au sens large, qui à la fois est associé à $(\mu_s)_{s \in \mathbb{R}}$ et, de plus, est optimal. Dans le contexte de l'équation de continuité il s'agit d'une mesure de probabilité sur $\mathcal{C}([0, 1], \mathcal{X})$ (où dans les cas de base, \mathcal{X} est une variété différentielle) qui est concentrée sur un ensemble de courbes intégrales $s \mapsto \gamma_s$, au sens où $\dot{\gamma}_s = V_s(\gamma_s)$. Il n'y a pas unicité de ce processus. Cependant, avec le processus Markov-quantile, nous parvenons à trouver un tel processus qui soit en plus markovien, mais pour $\mathcal{X} = \mathbb{R}$ seulement puisque c'est l'espace où évolue le processus Markov-quantile. Le champ de recherches reste cependant entièrement ouvert à d'autres espaces.

Notre *développement géométrique* concerne les courbes du type $s \mapsto \mu^{a_s}$ où $s \mapsto a_s$ évolue dans un espace métrique (spécialement, pour nous, un cône euclidien où le groupe de Heisenberg). La fonction $a \in \mathcal{X} \mapsto \mu^a \in \mathcal{P}(\mathcal{X})$ désigne la mesure obtenue par diffusion de la chaleur en partant de a et s'arrêtant à un instant $t > 0$ convenu, le même pour tout a . Plutôt qu'une seule famille $(\mu^a)_{a \in \mathcal{X}}$ de mesures plongées dans $\mathcal{P}(\mathcal{X})$ on a donc de fait pour tout $t > 0$ une famille différente $(\mu_t^a)_{a \in \mathcal{X}}$ et, pour a fixé, la courbe $t \mapsto \mu_t^a$ est celle de la diffusion de la chaleur. Pour tout $t > 0$, il est possible de définir une nouvelle distance sur \mathcal{X} par

$$\tilde{d}_t(a, a') = W_2(\mu_t^a, \mu_t^{a'}) \quad (4.2)$$

puis, dans un second temps de prendre sa version intrinsèque d_t dans l'optique de disposer avec (\mathcal{X}, d_t) , d'un espace géodésique. Cette transformation $d \rightarrow \tilde{d}_t \rightarrow d_t$ des distances sur \mathcal{X} due à Gigli et Mantegazza [GM14] s'avère intéressante lorsque \mathcal{X} est une variété riemannienne compacte car la déformation $t \mapsto (\mathcal{X}, d_t)$ reste dans cette classe et, encore mieux, est *tangente* en $t = 0$ au flot de Ricci, dans un sens qui est précisé dans l'article en question. Il est ainsi naturel d'examiner la transformation pour des espaces métriques avec l'objectif de définir le flot de Ricci partant d'espaces singuliers. C'est ce que nous avons fait dans [ErbJ18] partant de cônes droits euclidiens et du groupe de Heisenberg.

Le lien avec l'équation de continuité apparaît en ce que les champs de vecteurs optimaux $(V_s)_{s \in \mathbb{R}}$ associés à la courbe $s \mapsto \mu^{a_s}$ définissent une métrique L^2 au point μ_t^a en tout $a \in \mathcal{X}$, une métrique qui permet de redéfinir d_t différemment.

Dans [ErbJ18, Assumption 2.1 et (2.1)] une étape dans la définition de d_t pour un espace métrique général, est la majoration lipschitzienne suivante

$$d_t(\mathbf{a}, \mathbf{a}') = W_2(\mu_t^{\mathbf{a}}, \mu_t^{\mathbf{a}'}) \leq C(t)d(\mathbf{a}, \mathbf{a}'). \quad (4.3)$$

Dans le groupe de Heisenberg $(\mathcal{X}, d) = (\mathbb{H}, d_{cc})$, les relations du type de (4.3) que nous appelons « inégalités de Kuwada–Li » [Li06, Kuw10] sont devenues un sujet d'intérêt en soi que nous avons continué d'étoffer dans [ErbJ18] et [BonJ18⁺]. Elles restent mystérieuses, car, alors même que

$$W_\infty(\mu_t^{\mathbf{a}}, \mu_t^{\mathbf{a}'}) \leq C d_{cc}(\mathbf{a}, \mathbf{a}') \quad (4.4)$$

est démontré, des transports engendrant directement ces relations n'ont jamais été exhibés. Soulignons que (4.4) constitue bien une amélioration de (4.3) : cela réside non seulement dans le fait que la distance de Wasserstein W_∞ est plus grande que W_2 , mais aussi dans celui que la constante C ne dépend pas du temps. Les relations sont, comme l'a montré Kuwada [Kuw10], duales d'autres relations sur les gradients du semi-groupe de la chaleur dont la plus forte est celle de H.Q. Li. [Li06]

4.1 Introduction rapide au groupe de Heisenberg.

Le groupe de Heisenberg

Le groupe de Heisenberg \mathbb{H} est un groupe de Lie identifié à $\mathbb{C} \times \mathbb{R}$ avec pour loi de composition

$$(z, \mathbf{u}) \cdot (z', \mathbf{u}') = \left(z + z', \mathbf{u} + \mathbf{u}' - \frac{1}{2} \operatorname{Im}(z\bar{z}') \right),$$

où Im représente la partie imaginaire d'un nombre complexe.

Une base de l'algèbre de Lie associée est donnée par les champs de vecteurs invariants par translation à gauche suivants :

$$\mathbf{X} = \partial_x - \frac{y}{2} \partial_u, \quad \mathbf{Y} = \partial_y + \frac{x}{2} \partial_u, \quad \mathbf{U} = \partial_u,$$

où on remarquera comme unique relation non-triviale $[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}] = \mathbf{U}$.

On introduit également les champs de vecteurs invariants à droite suivant

$$\hat{\mathbf{X}} = \partial_x + \frac{y}{2} \partial_u, \quad \hat{\mathbf{Y}} = \partial_y - \frac{x}{2} \partial_u, \quad \hat{\mathbf{U}} = \mathbf{U}.$$

La mesure de Haar, qui est définie à une constante multiplicative près, coïncide avec la mesure de Lebesgue \mathcal{L} . Elle est à la fois invariante par translations à gauche et à droite.

Distances riemanniennes et sous-riemanniennes

À l'aide de la norme de Carnot–Carathéodory on munit \mathbb{H} d'une structure sous-riemannienne

$$\|m\mathbf{X} + n\mathbf{Y} + q\mathbf{U}\|_{cc}^2 = \begin{cases} m^2 + n^2 & \text{si } q = 0, \\ +\infty & \text{sinon.} \end{cases}$$

En minimisant la longueur des courbes, on définit alors la distance d_{cc} : pour tout couple $\mathbf{a}, \mathbf{a}' \in \mathbb{H}$ nous avons ainsi

$$d_{cc}(\mathbf{a}, \mathbf{a}') = \inf \int_0^T \|\dot{\gamma}_s\|_{cc} ds,$$

où l'infimum est obtenu sur la classe des courbes absolument continues¹ $(\gamma_s)_{s \in [0, T]}$ (par rapport à la distance euclidienne) telle que $\gamma_0 = \mathbf{a}$ et $\gamma_T = \mathbf{a}'$. Il apparaît clairement que la longueur

1. D'autres classes comme celle des courbes lipschitziennes donnent des définitions équivalentes.

sous-riemannienne est seulement finie dans le cas où $\dot{\gamma}_s$ est contenue dans le sous-fibré horizontal $\mathbb{H}\mathbb{H} = \text{Vect}(\mathbf{X}, \mathbf{Y}) \subset \mathbb{T}\mathbb{H}$ en presque tout instant.

Comme conséquence de la *condition de Hörmander* – le fait que les crochets de Lie engendrent l'espace tangent en tout point – la distance \mathbf{d}_{cc} est finie : il est possible de joindre un point à n'importe quel autre par une courbe de longueur finie. Une courbe est absolument continue par rapport à \mathbf{d}_{cc} si et seulement si elle est absolument continue au sens euclidien et que son vecteur tangent est presque partout horizontal.

À une constant multiplicative près, la mesure de Lebesgue \mathcal{L} coïncide avec la mesure de Hausdorff de dimension 4 de l'espace métrique $(\mathbb{H}, \mathbf{d}_{\text{cc}})$. Cela souligne si il est besoin que $(\mathbb{H}, \mathbf{d}_{\text{cc}})$ n'est pas riemannien.

Cependant, la pseudo-norme $\|\cdot\|_{\text{cc}}$ est naturellement approchée par une famille de métriques riemanniennes indexées par $\varepsilon > 0$, à savoir

$$\|\mathbf{m}\mathbf{X} + \mathbf{n}\mathbf{Y} + \mathbf{q}\mathbf{U}\|_{\text{Riem}(\varepsilon)}^2 = \mathbf{m}^2 + \mathbf{n}^2 + (\mathbf{q}/\varepsilon)^2.$$

On note $\mathbf{d}_{\text{Riem}(\varepsilon)}$ la distance riemannienne associée. Son volume riemannien est de nouveau \mathcal{L} à une constante près.

Les distances \mathbf{d}_{cc} et $\mathbf{d}_{\text{Riem}} := \mathbf{d}_{\text{Riem}(1)}$ se comparent à l'aide de considérations géométriques simples, liées au problème de Didon et à l'inégalité isopérimétrique planaire.

Proposition 4.1.1 (voir par exemple [Jui14, Lemma1.1]). *On a*

$$\mathbf{d}_{\text{Riem}} \leq \mathbf{d}_{\text{cc}} \leq \mathbf{d}_{\text{Riem}} + 4\pi.$$

et aussi $\mathbf{d}_{\text{cc}}(0_{\mathbb{H}}, (x, y, 0)) = \sqrt{x^2 + y^2}$, $\mathbf{d}_{\text{cc}}((x, y, z), (x, y, z + h)) = 2\sqrt{\pi|h|}$.

On introduite maintenant la quasinorme² homogène

$$\mathbb{H} : \mathbf{a} = (z, \mathbf{u}) \in \mathbb{H} \mapsto \sqrt{|z|^2 + |\mathbf{u}|} \in \mathbb{R},$$

et la quasidistance homogène invariante à gauche qui lui est associée $\mathbf{d}_{\mathbb{H}}(\mathbf{a}, \mathbf{a}') = \mathbb{H}(\mathbf{a}^{-1}\mathbf{a}')$. Alors

Proposition 4.1.2. *Il existe une constante $c > 1$ telle que*

$$\max(|z' - z|, c^{-1}\mathbf{d}_{\mathbb{H}}(\mathbf{a}, \mathbf{a}')) \leq \mathbf{d}_{\text{cc}}(\mathbf{a}, \mathbf{a}') \leq c\mathbf{d}_{\mathbb{H}}(\mathbf{a}, \mathbf{a}') \quad (4.5)$$

pour tout $\mathbf{a} = (z, \mathbf{u})$ et $\mathbf{a}' = (z', \mathbf{u}')$ dans \mathbb{H} .

Isométries

Pour tout $\mathbf{p} \in \mathbb{H}$, on note, respectivement, $\text{trans}_{\mathbf{a}} : \mathbb{H} \rightarrow \mathbb{H}$ et $\widehat{\text{trans}}_{\mathbf{a}} : \mathbb{H} \rightarrow \mathbb{H}$ les translations à gauche et à droite d'élément \mathbf{a} :

$$\text{trans}_{\mathbf{a}}(\mathbf{a}') = \mathbf{a} \cdot \mathbf{a}' = \widehat{\text{trans}}_{\mathbf{a}'}(\mathbf{a}).$$

Par définition, un champ de vecteurs \mathbf{V} est invariant par translation à gauche si et seulement si $\mathbf{D}\text{trans}_{\mathbf{a}}(\mathbf{V}) = \mathbf{V}$ pour tout $\mathbf{a} \in \mathbb{H}$. Ainsi, la translation à gauche $\text{trans}_{\mathbf{a}}$ est-elle une isométrie à la fois de \mathbf{d}_{cc} et des distances $\mathbf{d}_{\text{Riem}(\varepsilon)}$. Ceci n'est satisfait par $\widehat{\text{trans}}_{\mathbf{a}'}$ que lorsque \mathbf{a}' est de la forme $(0, \mathbf{u})$.

D'autres isométries sont

- les rotations $\text{rot}_{\theta} : \mathbb{H} \ni (z, \mathbf{u}) \mapsto (e^{i\theta}z, \mathbf{u})$ définies pour $\theta \in \mathbb{R}$,
- la réflexion $\text{sym} : \mathbb{H} \ni (z, \mathbf{u}) \mapsto (\bar{z}, -\mathbf{u})$,

et, à une constante multiplicative λ près,

2. L'inégalité triangulaire n'est vraie qu'à une constante multiplicative près.

— les dilatations homogènes $\text{dil}_\lambda : \mathbb{H} \ni (z, \mathbf{u}) \mapsto (\lambda z, \lambda^2 \mathbf{u})$ où $\lambda > 0$.

On a $D \text{dil}_\lambda(V) = \lambda V$ si et seulement si V est horizontal. En général on a

$$D \text{dil}_\lambda(m\mathbf{X} + n\mathbf{Y} + q\mathbf{U}) = \lambda(m\mathbf{X} + n\mathbf{Y}) + \lambda^2 q\mathbf{U}.$$

Donc dil_λ est une isométrie entre (\mathbb{H}, d_{cc}) et $(\mathbb{H}, \lambda^{-1} d_{cc})$ et également entre $(\mathbb{H}, d_{\text{Riem}(\varepsilon)})$ et $(\mathbb{H}, \lambda^{-1} d_{\text{Riem}(\varepsilon\lambda)})$. Il s'ensuit que toutes les variétés riemanniennes $(\mathbb{H}, d_{\text{Riem}(\varepsilon)})_{\varepsilon > 0}$ sont isométriques à une constante multiplicative près, ce qui justifie le fait de considérer principalement $(\mathbb{H}, d_{\text{Riem}})$ (notation de $(\mathbb{H}, d_{\text{Riem}(\varepsilon)})$ pour $\varepsilon = 1$).

Remarque 4.1.3. *Bien que moins classiquement utilisées, les dilatations pour $\lambda < 0$ sont aussi des isométries à une constante près. Dans [JSig17] elles ont joué un rôle important pour réaliser une chirurgie de courbes horizontales.*

Noyau de la chaleur

Une autre conséquence importante venant de la condition de Hörmander est l'hypoellipticité des opérateurs $\Delta_{cc} = \mathbf{X}^2 + \mathbf{Y}^2$ et $\Delta_{cc} - \partial_t$. Cela signifie en particulier que les fonctions $\rho : (0, \infty) \times \mathbb{H} \rightarrow \mathbb{R}$ satisfaisant, au sens des distributions, à l'équation de la chaleur

$$\partial_t \rho = \Delta_{cc} \rho,$$

sont lisses (c'est-à-dire de classe C^∞). Soulignons que l'équation de la chaleur est invariante par translation à gauche. Comme il a été montré par Gaveau [Gav77], la solution $\mu_t = \rho_t \mathcal{L}$ de condition initiale $\mu_0 \in \mathcal{P}_2(\mathbb{H})$ peut-être obtenue par convolution avec la solution fondamentale h_t :

$$\rho_t(p) = \int h_t(q^{-1}p) d\mu_0(q),$$

où h_t est explicitement donnée par la formule

$$h_t(z, \mathbf{u}) = \frac{2}{(4\pi t)^2} \int_{\mathbb{R}} \exp\left(\frac{\lambda}{t} \left(iu - \frac{|z|^2}{4} \coth \lambda\right)\right) \frac{\lambda}{\sinh \lambda} d\lambda.$$

En fait h_t est la densité de \mathbf{B}_{2t} où $\mathbf{B}_t = (B_t, L_t)$ avec $(B_t)_{t \geq 0}$ un mouvement brownien planaire $B = B^1 + iB^2$ et $L_t = \frac{1}{2} \int_0^t (B_s^1 dB_s^2 - B_s^2 dB_s^1)$ l'aire de Lévy qui lui est associée.

Nous serons amenés à utiliser les identités suivantes [BBBC08, (14) avec la démonstration de Theorem. 3.1] :

$$\int ((\mathbf{X} \log h_t)^2 + (\mathbf{Y} \log h_t)^2) h_t d\mathcal{L} = \frac{2}{t}, \quad \int (\mathbf{U} \log h_t)^2 h_t d\mathcal{L} < \infty. \quad (4.6)$$

Les mêmes relations sont valables pour les champs de vecteurs $\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{U}}$. Notons encore la relation de mise à l'échelle satisfaite par ces densités :

$$h_t(z, \mathbf{u}) = \frac{1}{t^2} h_1(z/\sqrt{t}, \mathbf{u}/t). \quad (4.7)$$

Pour tout $t \geq 0$ et $\mathbf{a} \in \mathbb{H}$ nous définissons ainsi la mesure $\nu_{\mathbf{a}}^t \in \mathcal{P}(\mathbb{H})$:

$$\nu_{\mathbf{a}}^t = \begin{cases} \delta_{\mathbf{q}} & \text{si } t = 0, \\ (\tau_{\mathbf{q}})_{\#}(h_t \mathcal{L}) = h_t(q^{-1}p) \mathcal{L}(dp) & \text{sinon.} \end{cases}$$

Nous l'appellerons noyau de la chaleur centré en \mathbf{a} au temps t .

Dans la présentation de [BonJ18⁺] où le mouvement brownien joue le premier rôle, c'est en revanche la mesure

$$\mu_t^{\mathbf{a}} = \text{Loi}(\mathbf{B}_t^{\mathbf{a}}) = \nu_{\mathbf{a}}^{t/2}$$

que nous utiliserons préférentiellement.

Inégalités de Kuwada–Li concernant le semi-groupe de la chaleur

L'inégalité de type L^1 suivante concerne le gradient du semi-groupe de la chaleur sur \mathbb{H} . Elle a été établie par H.Q. Li [Li06] (voir aussi [BBBC08] pour d'autres démonstrations) qui généralisait ainsi un inégalité de type L^2 de Driver et Melcher [DM05]. On a ainsi

$$\forall f \in \mathcal{C}_c^\infty(\mathbb{H}_1), \forall t \geq 0, \forall \mathbf{a} \in \mathbb{H}, |\nabla_{\text{cc}} P_t f(\mathbf{a})| \leq K P_t(|\nabla_{\text{cc}} f|)(\mathbf{a}), \quad (4.8)$$

où K est une constante, P_t est le semi-groupe de la chaleur associé à $(1/2)\Delta_{\text{cc}}$ et ∇_{cc} est le gradient horizontal, défini par $\nabla_{\text{cc}} f(x) = (\mathbf{X}f)\mathbf{X} + (\mathbf{Y}f)\mathbf{Y}$. En s'appuyant sur ces résultats et sur un argument de dualité, Kuwada [Kuw10] établit l'existence d'un contrôle lipschitzien pour la distance de Wasserstein L^∞ entre les noyaux de diffusion :

Théorème 4.1.4 (H.Q. Li, Kuwada). *Il existe $C > 0$ telle que pour tout $t > 0$ et $\mathbf{a}, \mathbf{a}' \in \mathbb{H}$*

$$W_\infty(\mu_t^\mathbf{a}, \mu_t^{\mathbf{a}'}) \leq C d_{\text{cc}}(\mathbf{a}, \mathbf{a}'), \quad (4.9)$$

où $\mu_t^\mathbf{a} = \text{Loi}(\mathbf{B}_t^\mathbf{a})$ et $\mu_t^{\mathbf{a}'} = \text{Loi}(\mathbf{B}_t^{\mathbf{a}'})$ et $(\mathbf{B}_s^\mathbf{a})_{s \geq 0}, (\mathbf{B}_s^{\mathbf{a}'})_{s \geq 0}$ sont deux mouvements browniens de \mathbb{H} , partant, respectivement, de \mathbf{a} et \mathbf{a}' . De plus,

$$W_p(\mu_t^\mathbf{a}, \mu_t^{\mathbf{a}'}) \leq C d_{\text{cc}}(\mathbf{a}, \mathbf{a}') \quad (4.10)$$

est vérifiée pour tout $p \in [1, \infty]$.

Remarque 4.1.5 (Constantes optimales). *Du fait des dilatations homogènes, la constante optimale $C_p(t) := \sup_{\mathbf{a} \neq \mathbf{a}'} W_p(\mu_t^\mathbf{a}, \mu_t^{\mathbf{a}'}) / d_{\text{cc}}(\mathbf{a}, \mathbf{a}')$ ne varie pas avec $t > 0$. En particulier la même constante optimale apparaît lorsque les mesures $\nu_\mathbf{a}^t$ sont prises en considération à la place des mesures $\mu_t^\mathbf{a}$.*

Les constantes optimales $(C_p)_{p \in [1, \infty]}$ sont conservées par la dualité de Kuwada. Ainsi, elles interviennent dans les inégalités de gradient (celles dont (4.8) est la version la plus forte) où elles apparaissent sous la forme $K_q = C_p$ avec $1/p + 1/q = 1$. Sous cette forme elles ont été étudiées par Driver et Melcher [DM05, §2.4] comme les meilleures constantes K pouvant vérifier $|\nabla_{\text{cc}} P_t f|^q \leq K^q P_t(|\nabla_{\text{cc}} f|^q)$ où $t > 0$. Il a été démontré $(K_q)^q > 1$ et $(K_2)^2 \geq 2$ ainsi que $K_q \geq K_{2k}$ dès que $q \leq 2k$ où k est un entier naturel non-nul. Cela implique en particulier $C_p \geq C_2 \geq \sqrt{2}$ pour tout $p \in [2, +\infty]$. Dans [BBBC08, Remark 3.2] il est conjecturé $C_2 = \sqrt{2} = C_\infty$.

On sait donc que pour \mathbf{a} et \mathbf{a}' dans \mathbb{H} et $t \geq 0$, il existe un couplage $(\mathbf{B}_s^\mathbf{a}, \mathbf{B}_s^{\mathbf{a}'})_{s \geq 0}$ de deux mouvements browniens de \mathbb{H} tel que

$$d_{\text{cc}}(\mathbf{B}_t^\mathbf{a}, \mathbf{B}_t^{\mathbf{a}'}) \leq C d_{\text{cc}}(\mathbf{a}, \mathbf{a}') \quad \text{presque sûrement.} \quad (4.11)$$

Soyons attentifs au fait que le temps $t \geq 0$ est fixé dans un premier temps et que le couplage ne concerne finalement que $\mathbf{B}_t^\mathbf{a}$ et $\mathbf{B}_t^{\mathbf{a}'}$, les deux processus étant couplés de façon arbitraire aux autres instants. Le problème de Kuwada réside exactement dans l'inversion des quantificateurs \forall et \exists . Il souhaite savoir s'il est possible de définir un couplage entre deux mouvements browniens de \mathbb{H} , $(\mathbf{B}_t^\mathbf{a})_{t \geq 0}$ et $(\mathbf{B}_t^{\mathbf{a}'})_{t \geq 0}$, de façon à ce que (4.11) soit satisfaite pour tout $t \geq 0$.

4.2 Courbes d'énergie finie et équation de continuité

Dans cette section nous introduisons l'énergie d'une courbe. Elle nous sera utile à la fois pour ses courbes $\gamma : [0, T] \rightarrow (\mathcal{X}, d)$ à valeur dans un espace métrique et pour des familles $(\mu_t)_{t \in [0, T]}$, vues comme des courbes $\mu : t \in [0, t] \mapsto (\mathcal{P}(\mathcal{X}), W_2)$. Commençons par des rappels sur les notions de *courbe absolument continue* et de *longueur*.

Rappel/Notation 4.2.1. Soit γ un élément de $\mathcal{C}([0, 1], \mathcal{X})$. (a) Pour $0 \leq a < b \leq 1$, la *longueur* de γ sur $[a, b]$ est définie comme sa variation totale $\mathcal{L}_a^b(\gamma) = \sup_{\mathbf{R} \in \text{Subdiv}([a, b])} \sum_{k=0}^m d(\gamma(r_k), \gamma(r_{k+1}))$, où $\mathbf{R} = \{r_0, r_1, \dots, r_m, r_{m+1}\}$.

(b) La courbe γ est *absolument continue* si pour tout $\delta > 0$ il existe $\varepsilon > 0$ tel que, étant donnée une famille d'intervalles $]a_k, b_k[$ satisfaisant à $\sum (b_k - a_k) \leq \varepsilon$, on ait $\sum d(\gamma(a_k), \gamma(b_k)) \leq \delta$.

L'ensemble $\mathcal{AC}([0, 1], \mathcal{X})$ est celui des courbes absolument continues. Comme il est démontré dans [AGS08], où la définition est légèrement différente mais équivalente, ces courbes admettent une *dérivée métrique*, notée $|\dot{\gamma}|(t)$, en presque tout instant t :

$$|\dot{\gamma}|(t) = \lim_{h \rightarrow 0} \frac{d(\gamma(t+h) - \gamma(t))}{h}.$$

(Notons que pour $\mathcal{X} = \mathbb{R}^n$ et γ dérivable en t la notation $|\dot{\gamma}|(t)$ est cohérente). Alors $\mathcal{L}_a^b(\gamma) = \int_a^b |\dot{\gamma}|(t) dt$. La notion de courbe absolument continue peut, de façon équivalente, être définie en requérant que la courbe $t \in [0, 1] \mapsto \mathcal{L}_0^t(\gamma) \in \mathbb{R}$ soit absolument continue, où qu'il existe une fonction $m : [0, 1] \rightarrow \mathbb{R}^+$, intégrable, telle que $d(\gamma(a), \gamma(b)) \leq \int_a^b m d\lambda$ pour tout $a < b$.

(c) Soit encore $\mathcal{AC}_2([0, 1], \mathcal{X}) \subset \mathcal{AC}([0, 1], \mathcal{X})$ l'espace des courbes γ absolument continue d'ordre deux, c'est-à-dire telles que $\int_0^1 |\dot{\gamma}|^2 < +\infty$. Les courbes lipschitziennes sont, par exemple, des courbes absolument continues d'ordre deux.

Nous introduisons maintenant la notion d'*énergie*. La proposition 4.2.3, qui suit, en rappelle certaines propriétés classiques.

Définition 4.2.2. Soit γ une application de $[0, 1]$ dans un espace métrique (\mathcal{X}, d) . Pour $0 \leq a < b \leq 1$ l'énergie (cinétique) $\mathcal{E}_a^b(\gamma)$ de γ sur $[a, b]$ est définie comme suit :

$$\mathcal{E}_a^b(\gamma) = \sup_{\mathbf{R} \in \text{Subdiv}([a, b])} \mathcal{E}_a^b(\gamma, \mathbf{R}), \quad \text{où} \quad (4.12)$$

$$\mathcal{E}_a^b(\gamma, \{r_0, \dots, r_{m+1}\}) = \sum_{k=0}^m \frac{d(\gamma(r_k), \gamma(r_{k+1}))^2}{r_{k+1} - r_k}.$$

Note. Pour $[a, b] = [0, 1]$ on abrège \mathcal{E}_0^1 et \mathcal{L}_0^1 , respectivement par \mathcal{E} et \mathcal{L} .

Proposition 4.2.3 ([BouJ18⁺, §5]). Soit γ une application de $[0, 1]$ dans \mathcal{X} . Alors :

- (a) Si $\mathcal{E}(\gamma) < \infty$, alors γ est continue, c'est-à-dire $\mathcal{AC}_2 \subset \mathcal{C}$.
- (b) (i) Si $\mathbf{R}' \in \text{Subdiv}([a, b])$ est une subdivision plus fine que \mathbf{R} , alors $\mathcal{E}_a^b(\gamma, \mathbf{R}') \leq \mathcal{E}_a^b(\gamma, \mathbf{R})$.
- (ii) Si γ est continue, la limite $\lim_{|\mathbf{R}| \rightarrow 0} \mathcal{E}(\gamma, \mathbf{R})$ est bien définie et vaut $\mathcal{E}(\gamma)$.
- (iii) $\mathcal{E}(\gamma)$ est finie si et seulement si $\gamma \in \mathcal{AC}_2([0, 1], \mathcal{X})$. On a alors $\mathcal{E}_a^b(\gamma) = \int_a^b |\dot{\gamma}|^2(t) dt$ pour tout $a < b$.
- (c) \mathcal{E} est semi-continue inférieurement pour la convergence uniforme.

Courbes d'énergie finie dans les espaces de Wasserstein

Étant donné, sur $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, une courbe $\mu = (\mu_s)_{s \in [0, T]}$ d'énergie finie, il est possible de lui associer une famille de champs de vecteurs telle que le couple vérifie l'équation de continuité. C'est une des étapes vers la *représentation probabiliste* de μ telle qu'on peut la lire dans [Man07, AGS08]. Il s'agit, comme dans le problème des PCOC, d'associer à la courbe $(\mu_s)_{s \in [0, T]}$ un mesure $P \in \text{Marg}(\mu)$. Plus précisément on cherche une probabilité, non pas sur la tribu cylindrique de $\mathcal{X}^{[0, T]}$, mais sur la tribu borélienne de $\mathcal{C}([0, T], \mathcal{X})$. Étant donné que nous notons γ les éléments $\mathcal{C}([0, T], \mathcal{X})$ les mesures de probabilités sur $\mathcal{C}([0, T], \mathcal{X})$ seront notées Γ plutôt que P . Finalement $\text{Marg}_{\mathcal{C}}(\mu)$ sera le sous-espace de $\mathcal{P}(\mathcal{C}([0, T], \mathcal{X}))$ des mesures de marges μ_s . Évidemment chaque élément Γ de $\text{Marg}_{\mathcal{C}}(\mu)$ est associé à un unique élément P de $\text{Marg}(\mu)$.

Nous parlerons de *représentation eulérienne* lorsqu'il s'agira d'associer des champs de vecteurs à $(\mu_s)_{s \in [0, T]}$. Comme expliqué ci-dessus, c'est une étape vers la représentation probabiliste, l'énoncé *lagrangien*. Ces résultats seront rappelés et enrichis des nôtres concernant $\mathfrak{M}\Omega$ (dans le cas de la dimension $d = 1$) au théorème 4.3.4.

Cette représentation, à commencer par l'énoncé eulérien, se généralise à des espaces $(\mathcal{P}_2(\mathcal{X}), W_2)$ plus généraux (pour peu qu'on sache donner un sens aux champs de vecteurs). En relation avec la

partie §4.5 dont les enjeux sont un peu différents mais engagent l'équation de continuité, nous citons, pour l'exemple, un énoncé de représentation eulérienne pour le groupe de Heisenberg. Pour ce faire on pose W_{cc} la distance de Wasserstein W_2 construite à partir de d_{cc} . On note $\operatorname{div} V$ la divergence d'un champ de vecteurs de V de \mathbb{R}^3 par rapport à la mesure de Lebesgue. Comme les vecteurs de base \mathbf{X} , \mathbf{Y} et \mathbf{U} sont de divergence nulle, nous avons $\operatorname{div}(f\mathbf{X} + g\mathbf{Y} + h\mathbf{U}) = \mathbf{X}f + \mathbf{Y}g + \mathbf{U}h$ pour tout triplet de fonctions f, g, h suffisamment régulière. Rappelons que le gradient horizontal de f est donné par l'expression $\nabla_{cc}f = (\mathbf{X}f)\mathbf{X} + (\mathbf{Y}f)\mathbf{Y}$. Alors, pour toute fonction régulière f à support compact et tout champ de vecteurs V , on trouve, en intégrant par parties

$$\int_{\mathbb{H}} f \operatorname{div} V d\mathcal{L} = - \int_{\mathbb{H}} \langle \nabla_{cc}f, V \rangle_{cc} d\mathcal{L}.$$

Notons maintenant $L_{cc}^2(\mu)$ l'espace de Hilbert des champs de vecteurs V équipé de la norme $\|V\|_{L_{cc}^2(\mu)}^2 = \int \|V\|_{cc}^2 d\mu$. Tout champ de vecteurs de norme finie $V \in L_{cc}^2(\mu)$ est nécessairement horizontal, par rapport à μ . Voici donc la caractérisation des courbes absolument continues d'ordre deux.

Proposition 4.2.4 (Représentation eulérienne, [Jui14, Proposition 3.1]). *Soit $\mu = (\mu_s)_{s \in [0, T]}$ une courbe sur $\mathcal{P}_2(\mathbb{H})$. Alors, cette courbe est d'énergie finie (pour la distance W_{cc}) si et seulement si il existe une famille de champs de vecteurs $(V_s)_{s \in [0, T]}$ vérifiant $\int_0^T \|V_s\|_{L_{cc}^2(\mu_s)}^2 ds < \infty$ d'une part, et l'équation de continuité $\partial_s \mu + \operatorname{div}(\mu_s V_s) = 0$ d'autre part.*

On a alors $|\dot{\mu}_s| \leq \|V_s\|_{L_{cc}^2(\mu_s)}$ en presque tout instant $s \in [0, T]$ et donc aussi $\mathcal{E}(\mu) \leq \int_0^T \|V_s\|_{L_{cc}^2(\mu_s)}^2 ds$.

De plus, l'égalité $\int_0^T \|V_s\|_{L_{cc}^2(\mu_s)}^2 ds = \mathcal{E}(\mu)$ peut être atteinte; elle est réalisée si et seulement si $|\dot{\mu}_s| = \|V_s\|_{L_{cc}^2(\mu_s)}$ en presque tout $s \in [0, T]$, ce qui se produit si et seulement si

$$V_s \in T_{\mu_s} \mathcal{P}_2(\mathbb{H}) := \overline{\{\nabla_{cc}\psi \mid \psi \in C_c^\infty(\mathbb{R}^3)\}}^{L_{cc}^2(\mu_s)}$$

en presque tout $s \in [0, T]$.

Finalement le champ de vecteurs $(V_s)_{s \in [0, T]}$ réalisant l'égalité est unique presque partout (pour $\operatorname{Joint}(\lambda|_{[0, T]}, \mu)$).

4.3 Processus Markov-quantile et équation de continuité

La représentation probabiliste est un processus qui sur $\operatorname{Marg}_{\mathcal{C}}(\mu)$ qui minimise une fonctionnelle d'action \mathcal{A} , que nous définissons maintenant.

Action d'une mesure de probabilité sur $\mathcal{C}([0, 1], \mathbb{R}^d)$

Définition 4.3.1 (Action \mathcal{A}). Soit $\Gamma \in \mathcal{P}((\mathbb{R}^d)^{\mathbb{R}})$ une mesure concentrée sur $\mathcal{C}([0, 1], \mathbb{R}^d)$. Son action, $\mathcal{A}(\Gamma)$, est donnée par

$$\mathcal{A}(\Gamma) = \int_{\mathcal{C}} \mathcal{E}(\gamma) d\Gamma(\gamma).$$

À la proposition 4.3.3 nous allons voir qu'en dimension $d = 1$, $\mathfrak{M}\Omega(\mu)$ minimise l'action sur $\operatorname{Marg}_{\mathcal{C}}(\mu)$. Cette propriété est partagée par $\Omega(\mu)$. En prenant la mesure Markov-quantile on conserve donc des propriétés intéressantes de la mesure quantile. C'est ici le troisième exemple d'une bonne propriété conservée en passant de Ω à $\mathfrak{M}\Omega$, après celle de la croissance (théorème 3.6.2(c)) et celle de minimalité par rapport à \preceq_{lo} (théorème (a)(iii)).

Remarque 4.3.2. (a) *Par le théorème de Fubini, si $\mathcal{A}(\Gamma) < +\infty$, alors Γ est concentrée sur \mathcal{AC}_2 .*

(b) Si Γ est une mesure sur $\mathcal{C}([0, 1], \mathbb{R}^d)$, comme le sont les éléments de $\text{Marg}_{\mathcal{C}}(\mu)$, alors du fait du théorème de convergence monotone on a

$$\mathcal{A}(\Gamma) := \int_{\mathcal{C}} \lim_{|\mathbf{R}| \rightarrow 0} \mathcal{E}(\gamma, \mathbf{R}) \, d\Gamma(\gamma) = \lim_{|\mathbf{R}| \rightarrow 0} \int_{\mathcal{C}} \mathcal{E}(\gamma, \mathbf{R}) \, d\Gamma(\gamma) \quad (4.13)$$

(Prendre une suite monotone de subdivision et utiliser la proposition 4.2.3(b).)

(c) Si $\Gamma \in \text{Marg}_{\mathcal{C}}(\mu)$, alors :

$$\mathcal{A}(\Gamma) \geq \mathcal{E}(\mu). \quad (4.14)$$

Dans le cas $d = 1$, cette inégalité devient une égalité pour $\Gamma = \mathfrak{Q}(\mu)$. En effet :

$$\begin{aligned} \int_{\mathcal{C}} \mathcal{E}(\gamma, \mathbf{R}) \, d\Gamma(\gamma) &= \int_{\mathcal{C}} \sum_{k=1}^m \|\gamma(r_k) - \gamma(r_{k+1})\|^2 / (r_{k+1} - r_k) \, d\Gamma(\gamma) \\ &= \sum_{k=1}^m \left(\int_{\mathcal{C}} \|\gamma(r_k) - \gamma(r_{k+1})\|^2 / (r_{k+1} - r_k) \, d\Gamma(\gamma) \right) \\ &\geq \sum_{k=1}^m W_2(\mu_{r_k}, \mu_{r_{k+1}})^2 / (r_{k+1} - r_k) = \mathcal{E}(\mu, \mathbf{R}). \end{aligned} \quad (4.15)$$

L'inégalité ressort de ce que $(\text{proj}^{r_k, r_{k+1}})_{\#} \Gamma$ est élément de $\text{Marg}(\mu_{r_k}, \mu_{r_{k+1}})$, si bien que $\int_{\mathcal{C}} \|\gamma(r_k) - \gamma(r_{k+1})\|^2 \, d\Gamma(\gamma) \geq W_2(\mu_{r_k}, \mu_{r_{k+1}})^2$. Pour $d = 1$, on obtient une égalité si $(\text{proj}^{r_k, r_{k+1}})_{\#} \Gamma = \mathfrak{Q}(\mu_{r_k}, \mu_{r_{k+1}})$, ce qui est vrai pour $\Gamma = \mathfrak{Q}$. Grâce à (4.13), en faisant tendre le pas de \mathbf{R} vers 0 on trouve $\mathcal{A}(\Gamma) \geq \mathcal{E}(\mu)$, et donc l'inégalité recherchée, ou l'égalité si $\Gamma = \mathfrak{Q}$.

(d) Si l'égalité dans (c) est obtenue pour une mesure Γ , cette égalité sera encore satisfaite pour $\Gamma_{[\mathbf{R}]}$ (voir page 70) — donc, si $d = 1$, par les mesures $\mathfrak{Q}_{[\mathbf{R}]}(\mu)$. Pour voir cela il suffit de considérer (4.15) pour des subdivisions plus fines que \mathbf{R} seulement (par la proposition 4.2.3(b)(i), le minimum dans (4.12) reste inchangé) : on obtient $\mathcal{A}(\Gamma) = \mathcal{A}(\Gamma_{[\mathbf{R}]})$.

La remarque 4.3.2 (d) « passe à la limite » lorsqu'on considère une suite de subdivisions $(\mathbf{R}_n)_n$ pour laquelle $\mathfrak{Q}_{[\mathbf{R}_n]}$ converge. Ainsi :

Proposition 4.3.3 ($\mathfrak{M}\mathfrak{Q}$ minimise l'action, [BouJ18⁺, Proposition 5.11]). Soit $\mu = (\mu_t)_{t \in [0, 1]}$ une courbe d'énergie finie de mesures dans $\mathcal{P}_2(\mathbb{R})$. La mesure Markov-quantile $\mathfrak{M}\mathfrak{Q} \in \text{Marg}(\mu)$ vérifie $\mathcal{A}(\mathfrak{M}\mathfrak{Q}) = \mathcal{E}(\mu)$ au lieu de seulement l'inégalité $\mathcal{A}(\mathfrak{M}\mathfrak{Q}) \geq \mathcal{E}(\mu)$.

De plus, pour une suite $(\mathbf{R}_n)_{n \in \mathbb{N}}$, avec $\mathbf{R}_n \in \text{Subdiv}([0, 1])$, telle que dans le théorème 3.6.4, $(\mathfrak{Q}_{[\mathbf{R}_n]})_n$ tend vers $\mathfrak{M}\mathfrak{Q}$ pour la topologie de $\text{Marg}_{\mathcal{C}}(\mu)$.

Lien entre $\mathfrak{M}\mathfrak{Q}$ et l'équation de continuité

Dans le théorème suivant sont rassemblés des faits bien connus sur la représentation probabiliste, eulérienne ou lagrangienne, de $\mu = (\mu_t)_{t \in [0, 1]}$. L'ajout concerne le caractère markovien en dimension $d = 1$, son existence garantie par $\mathfrak{M}\mathfrak{Q}$ et son unicité.

Théorème 4.3.4 (Existence et unicité des représentations). Soit $\mu = (\mu_t)_{t \in [0, 1]}$ une courbe d'énergie finie $\mathcal{E}(\mu)$ dans l'espace de Wasserstein $\mathcal{P}_2(\mathbb{R})$. Alors :

(a) (Énoncé eulérien.) Il existe un champ de vecteurs (paramétré) $(V_t)_{t \in [0, 1]}$ vérifiant l'équation de continuité et tels que l'inégalité

$$\int_0^1 \int |V_t|^2 \, d\mu_t \, dt \geq \mathcal{E}(\mu)$$

est une égalité. Ce champ de vecteurs est unique presque partout (pour $\text{Joint}(\lambda_{[0, 1]}, \mu)$).

(b) (Énoncé lagrangien.) Il existe $\Gamma \in \text{Marg}_{\mathcal{C}}(\mu)$ telle que :

- (i) L'inégalité (4.14) : $\mathcal{A}(\Gamma) \geq \mathcal{E}(\mu)$ est une égalité,
- (ii) la mesure Γ est markovienne,
- (iii) Γ est la limite d'une suite $(\Omega_{[\mathbb{R}_n]})_n$ dans $\mathcal{P}(\mathcal{C})$.

Un tel Γ est unique dans $\text{Marg}_{\mathcal{C}}(\mu)$; c'est le processus Markov-quantile $\mathfrak{M}\Omega$.

(c) (Relation entre les deux.) Pour Γ minimisant l'action, c'est-à-dire telle que (4.14) est une égalité, la courbe $\gamma \in \mathcal{C}$ est presque sûrement, pour Γ , une solution de l'EDO :

$$\dot{\gamma}(t) = V_t(\gamma_t),$$

pour presque tout temps.

Problème 4.3.5 (Voir aussi [BouJ18⁺, §6]). — Est-il possible d'obtenir un énoncé similaire au théorème 4.3.4 en dimensions supérieures ou pour un espace métrique ? La formulation de ce problème nécessite de remplacer les suites $(\Omega_{[\mathbb{R}_n]})_n$ par des suites de mesures ayant un sens dans ce cadre élargi. Le théorème 4.3.9, ci-dessous, propose un tel cadre *sans conclure* à l'existence d'un processus markovien. Les éléments de la suite sont les minimiseurs de l'action \mathcal{A} sur l'espace des mesures pour lesquelles seules les marges $(\mu_r)_{r \in \mathbb{R}}$ sont prescrites et où le processus est en sus rendu markovien aux instants de \mathbb{R} .

- Peut-on en dimension 1 avoir d'autres énoncés d'unicité d'un processus markovien ? En particulier, l'exemple 4.3.6 illustré à la figure 4.1 montre qu'un processus markovien, différent de $\mathfrak{M}\Omega$ peut être un minimiseur de \mathcal{A} .

Exemple 4.3.6 (Un atome traverse une mesure diffuse). On considère $\mu = (\mu_t)_{t \in [0,1]}$ où $\mu_t = \frac{1}{2}\lambda_{[t-3/4, t-1/4]} + \frac{1}{2}\delta_0$. C'est la famille des marges de Γ , le processus dont les trajectoires, toutes affines, sont définies par $\Gamma(t \mapsto 0) = 1/2$ et $\Gamma(\{t \mapsto x_0 + t : x_0 \in A\}) = \lambda_{[-3/4, -1/4]}(A)$. Ce processus n'est pas le processus $\mathfrak{M}\Omega$ associé μ mais, cependant, il est markovien et tangent au champ de vecteurs optimal du théorème 4.3.4(a), à savoir :

$$V_t(x) = 0 \quad \text{si } x = 0 \quad \text{et} \quad V_t(x) = 1 \quad \text{sinon.}$$

Ainsi l'action $\mathcal{A}(\Gamma)$ vaut $\mathcal{E}(\mu)$, la valeur minimale de \mathcal{A} sur $\text{Marg}_{\mathcal{C}}(\mu)$.

Le processus Markov-quantile $(X_t)_{t \in [0,1]}$ associé à $(\mu_t)_{t \in [0,1]}$ se comporte comme suit : partant d'un point de départ X_0 est choisi selon μ_0 , les trajectoires sont ensuite affines par morceaux, épousant sur ces morceaux des trajectoires parmi celles décrites au-dessus. Si $X_0 \in [-3/4, -1/4]$, le premier morceau est $X_t = X_0 + t$ sur $[0, \tau]$ avec $-X_0 = \tau$. Sur $[\tau, \min(\tau + \eta, 1)]$ X_t vaut constamment zéro. Ici η est une variable aléatoire exponentielle de paramètre 2, indépendante de X_0 . Sur le troisième morceau, si il y en a un, X est affine de pente 1, c'est-à-dire $X_t = t - (\tau + \eta)$ sur $[\tau + \eta, 1]$.

À l'opposé de Γ , le processus Markov-quantile $(X_t)_{t \in [0,1]}$ est à noyaux croissants et il est fortement markovien.

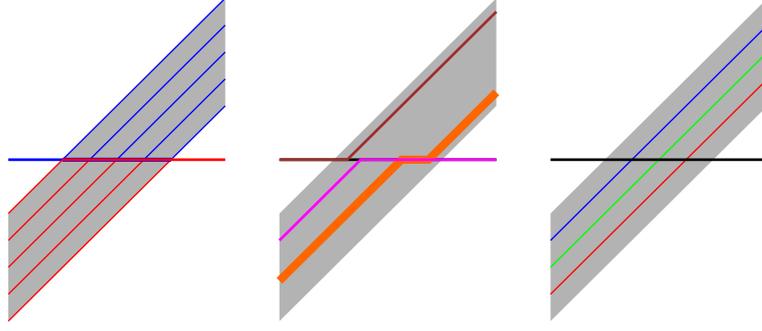
Voici maintenant notre alternative à $\Omega_{[\mathbb{R}]}$ en dimensions supérieures. On consultera, pour comparaison utile, la définition de l'interpolation par déplacement rappelée à la page 31, où, pour une construction très comparable, notre formulation est probabiliste .

Définition 4.3.7. Soit $\mathbb{R} = \{r_0, r_1, \dots, r_m, r_{m+1}\}$ une subdivision de $\text{Subdiv}([0, 1])$. On note $\text{Disp}_{\mathbb{R}}$ l'ensemble des mesures $M \in \mathcal{P}(\mathcal{C})$ qui interpolent linéairement et optimalement les mesures $(\mu_r)_{r \in \mathbb{R}}$ avec composition markovienne aux instants $r \in \mathbb{R}$. Une formulation mathématique en est :

(a) Pour tout $i \in \{0, \dots, m\}$, le transport $M^{r_i, r_{i+1}} \in \text{Marg}(\mu_{r_i}, \mu_{r_{i+1}})$ est optimal de marges μ_{r_i} et $\mu_{r_{i+1}}$,

(b) pour $\{\lambda_1, \dots, \lambda_n\} \subset [0, 1]$ et $i^\lambda : (x, y) \in (\mathbb{R}^d)^2 \mapsto \lambda y + (1 - \lambda)x$, on a :

$$(i^{\lambda_1}, \dots, i^{\lambda_n})_{\#} M^{r_i, r_{i+1}} = M^{\lambda_1 r_i + (1 - \lambda_1) r_{i+1}, \dots, \lambda_n r_i + (1 - \lambda_n) r_{i+1}},$$

FIGURE 4.1 – Les processus Ω , $\mathfrak{M}\Omega$ et le processus markovien Γ de l'exemple 4.3.6

(c) Pour tout S fini contenant $\{r_1, \dots, r_m\}$,

$$(\text{proj}^S)_\# M = M^{s_1^0, \dots, s_{n_0}^0, r_1} \circ M^{r_1, s_1^1, \dots, s_{n_1}^1, r_2} \circ \dots \circ M^{r_m, s_1^m, \dots, s_{n_m}^m},$$

où $S = \{s_1^0, \dots, s_{n_0}^0, r_1, s_1^1, \dots, s_{n_1}^1, r_2, \dots, r_m, s_1^m, \dots, s_{n_m}^m\}$ et où le premier et/ou dernier terme disparaissent si n_0 et/ou n_m sont nuls.

Remarque 4.3.8. — Sur chacun des intervalles on exécute finalement une interpolation géométrique au sens de la page 30.

- Les éléments de $\mathfrak{Disp}_{\mathbb{R}}$ minimisent l'action \mathcal{A} sur $\{M \in \mathcal{P}(\mathcal{C}([0, 1], \mathbb{R}^d)) : \forall r \in \mathbb{R}, M^r = \mu_r\}$.
- On a $\#\mathfrak{Disp}_{\mathbb{R}} = 1$ si et seulement si chacun des ensembles $\text{Marg}(\mu_{r_i}, \mu_{r_{i+1}})$, apparaissant dans (a), contient un unique transport optimal. C'est le cas si $d = 1$, où $\text{Marg}(\mu_{r_i}, \mu_{r_{i+1}})$ est le singleton $\{\Omega(\mu_{r_i}, \mu_{r_{i+1}})\}$.

La démonstration du théorème suivant s'inspire de [Lis07] et [Vil09, Chapter 7]. On a la représentation probabiliste lagrangienne sans passer par un énoncé eulérien, ce qui a aussi du sens pour un espace métrique général.

Théorème 4.3.9 ([BouJ18⁺, Theorem 5.20]). Soit d un entier positif et $\mu = (\mu_t)_{t \in [0, 1]}$ une courbe dans $\mathcal{P}_2(\mathbb{R}^d)$ d'énergie finie. Pour toute suite croissante $(\mathbb{R}_n)_n$ où on suppose $\mathbb{R}_\infty := \cup_n \mathbb{R}_n$ dense dans $[0, 1]$, et $(\Gamma_n)_n$ avec $\Gamma_n \in \mathfrak{Disp}_{\mathbb{R}_n}$ pour tout $n \in \mathbb{N}$, il existe $\Gamma \in \text{Marg}_{\mathcal{C}}(\mu)$ qui est la limite dans $\mathcal{P}(\mathcal{C}([0, 1], \mathbb{R}^d))$ d'une sous-suite de $(\Gamma_n)_n$. De plus, pour chaque Γ l'action $\mathcal{A}(\Gamma)$ est minimale, c'est-à-dire que (4.14) est une égalité.

Finalement, en dimension $d = 1$, une mesure Γ markovienne peut être obtenue pour certaines suites $(\mathbb{R}_n)_n$, et si Γ est une limite markovienne, alors $\Gamma = \mathfrak{M}\Omega$.

4.4 Considérations sur les inégalités de Kuwada–Li

Comme nous l'avons expliqué, les inégalités (4.10) ont été obtenues indirectement par dualité, en s'appuyant sur un résultat de Kuwada, et partant de l'inégalité de H-Q. Li. Dans cette section nous cherchons à donner une démonstration indépendante et directe. On est donc à la recherche de bons transports, nous donnant une borne uniforme, à défaut d'être nécessairement optimaux. Au théorème 4.4.6 nous présenterons un tel transport « statique ». Celui-ci ne fonctionnera malheureusement que pour $p = 1$. Avant cela nous expliquerons au théorème 4.4.1 que la méthode des couplages co-immersés de processus, classique en calcul stochastique (voir par exemple [BACK95, Ken07, Ken10, Cra91, vR04, ATW06, KS07, PP16]), est vouée à l'échec. Ces résultats sont le fruit d'une collaboration avec Michel Bonnefont [BonJ18⁺]. Signalons pour conclure cette introduction, que les résultats se prolongent aux groupes de Heisenberg de dimensions supérieures (voir [BonJ18⁺, §6]).

Théorème 4.4.1. *Soit $(\mathbf{B}_t^\alpha)_t$ et $(\mathbf{B}_t^{\alpha'})_t$ deux mouvements browniens co-immérgés partant respectivement de $\mathbf{a} = (x, y, z)$ et $\mathbf{a}' = (x', y', z')$ avec $(x' - x)^2 + (y' - y)^2 > 0$. Alors, quelque soit $C > 0$,*

$$\mathbb{P} \left(\forall t \geq 0, d_{\text{cc}}(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'}) \leq C \right) \neq 1.$$

De plus on a même

$$\limsup_{t \rightarrow +\infty} \mathbb{E} \left[d_{\text{cc}}(\mathbf{B}_t, \mathbf{B}_t')^2 \right] \rightarrow +\infty.$$

Ce résultat peut être légèrement amplifié en utilisant les symétries en temps et espace du groupe de Heisenberg :

Corollaire 4.4.2. *Soit $p \geq 2$. Pour tout $T > 0$ et $C > 0$ il existe deux points $\mathbf{a}, \mathbf{a}' \in \mathbb{H}$ avec $\mathbf{a} \neq \mathbf{a}'$ tels que pour tout couplage co-immérgé $(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'})_{0 \leq t \leq T}$, il existe $t \in [0, T]$ pour lequel*

$$\mathbb{E} \left[d_{\text{cc}}^p(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'}) \right] > C d_{\text{cc}}(\mathbf{a}, \mathbf{a}')^p.$$

Remarque 4.4.3 (Comparaison avec le cas riemannien). *La situation du théorème est bien différente de celle prévalant pour les variétés riemanniennes. Sur une telle variété M , il est connu (voir, par exemple [Whi34] ou [vRS05]) que la courbure de Ricci est bornée inférieurement par $k \in \mathbb{R}$ si et seulement si il existe un couplage markovien (donc co-immérgé) des mouvements browniens partant de deux points \mathbf{a} et \mathbf{a}' tel que*

$$d(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'}) \leq e^{-(k/2)t} d(\mathbf{a}, \mathbf{a}'). \quad (4.16)$$

La version L^p pour $p \in]1, \infty[$ de cette majoration correspondant à $p = \infty$ est

$$W_p(\mu_t^\alpha, \mu_t^{\alpha'}) \leq e^{-kt/2} d(\mathbf{a}, \mathbf{a}') \text{ pour tout } t \geq 0, \mathbf{a} \in M, \mathbf{a}' \in M.$$

Elle est vérifiée si et seulement si la courbure de Ricci est bornée inférieurement par k . Alors que la motivation originelle pour coupler des mouvements browniens était d'obtenir des inégalités concernant les semi-groupes (voir par exemple [Cra91, Cra92]), nous nous proposons au contraire de voir les inégalités sur le semi-groupe comme des contrôles de la distance de Wasserstein entre des noyaux de la chaleur. À cet égard, historique, les couplages statiques peuvent apparaître comme des curiosités. On notera par ailleurs que

- *le groupe de Heisenberg peut se voir comme le premier espace modèle sous-elliptique de courbure 0. Cependant son comportement vis-à-vis des mouvements browniens est tout à fait différent de celui observé sur les variétés dont la courbure de Ricci est bornée par $k = 0$.*
- *Le groupe de Heisenberg est également notoirement connu pour être la limite d'une suite de variétés riemanniennes sur le même groupe de Lie dont la meilleure borne inférieure pour la courbure de Ricci tend vers $-\infty$ (voir par exemple [Jui09, BG17, GT16]). Ceci est cohérent avec le théorème 4.4.1 si on interprète la constante $C = e^{-kt}$ dans (4.16) avec $k = -\infty$: aucun contrôle n'est possible pour $t > 0$.*

Le théorème 4.4.1 est tout d'abord démontré dans [BonJ18⁺] pour $p = \infty$. La démonstration est facilement étendue à $p \geq 4$ (voir [BonJ18⁺, Remark 3.1]). On écrit les équations régissant le couplage co-immérgée. Une matrice $(2, 2)$ co-immérgée $\begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}$ contrôle les variations infinitésimales stochastiques de $\mathbf{B}_t^{\alpha'}$ par rapport à celles de \mathbf{B}_t^α . Avec une autre matrice \tilde{K} qui correspond aux variations indépendantes, elles vérifient l'équation

$$K^t K + \tilde{K}^t \tilde{K} = I_2.$$

Le couplage parallèle, donné par $K = I_2$ crée une grande différence d'aire entre les trajectoires planaires (car celles-ci sont parallèles), ce qui engendre une grande différence dans la troisième coordonnée de $(\mathbf{B}_t^\alpha)^{-1} \cdot \mathbf{B}_t^{\alpha'}$. Il faut donc s'écarter de ce mode. Mais le fait que $d_{\text{cc}}(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'})$ est

borné engendre également le fait que l'intégrale de $(K_{12})^2 + (K_{21})^2$ en temps et en espérance est elle aussi finie (voir [BonJ18⁺, page 7]). On parvient à une contradiction.

Pour étendre le résultat à $p \geq 2$ on utilise des astuces ad hoc fondées sur l'observation de l'aire relative Z_t , à savoir la troisième coordonnée de $(\mathbf{B}_t^\alpha)^{-1} \cdot \mathbf{B}_t^{\alpha'}$. La démonstration serait simple si le terme de diffusion était constant et la dérive nulle. On peut argumenter quantitativement du fait qu'il est impossible de s'éloigner trop de ce cadre. En résumant et modifiant un peu les faits (voir [BonJ18⁺, (23)] pour les détails) on s'aperçoit que, vues comme des courbes sur $L^1(\Omega)$ (correspondant à L^2 pour le troisième axe $(\mathbb{R}\mathbf{U}, d_{cc})$), la partie martingale de Z_t varie sur $[0, t]$ davantage que $|t|$ alors que la partie à variations bornées varie moins que $\sqrt{|t|}$.

Premier résultat positif

Pour $p \in]0, 1[$ il est possible d'obtenir un contrôle uniforme par un couplage co-immégré. Il s'agit du *couplage par réflexion* obtenu en ne considérant que la partie planaire des deux mouvements browniens. Ceux-ci évoluent d'abord symétriquement par rapport à la médiatrice du segment les reliant à un instant donné (la même médiatrice qu'à l'instant initial). Alors que les mouvements browniens planaires se rencontrent sur la médiatrice (lors d'un temps d'atteinte que nous notons t_0) et continuent de façon synchrone, les mouvements browniens de \mathbb{H} correspondants se déplacent alors de façon telle que $(\mathbf{B}_t^\alpha)^{-1} \cdot \mathbf{B}_t^{\alpha'}$ reste une constante $(0, 0, z) = (\mathbf{B}_{t_0}^\alpha)^{-1} \cdot \mathbf{B}_{t_0}^{\alpha'}$.

Théorème 4.4.4. *Soit $(\mathbf{B}_t^\alpha)_{t \geq 0}$ et $(\mathbf{B}_t^{\alpha'})_{t \geq 0}$ deux mouvements browniens de \mathbb{H} couplés par réflexion. Alors pour tout $p \in]0, 1[$,*

$$\sup_{t \geq 0} \mathbb{E} \left[d_{cc}(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'})^p \right] < +\infty. \quad (4.17)$$

De plus, pour le couplage par réflexion, pour tout $p \in]0, 1[$, on a aussi :

$$\sup_{\alpha \neq \alpha' \in \mathbb{H}} \sup_{t \geq 0} \frac{\mathbb{E}[d_{cc}(\mathbf{B}_t^\alpha, \mathbf{B}_t^{\alpha'})^p]}{d_{cc}(\alpha, \alpha')^p} < +\infty. \quad (4.18)$$

Problème 4.4.5. Alors que les puissances $p \in [2, \infty]$ et $p \in]0, 1[$ ont été traitées dans les théorèmes 4.4.1 et 4.4.4, respectivement, nous ne savons pas, pour $p \in [1, 2[$ si il est possible d'obtenir un couplage co-immégré fournissant les bornes (4.10). Qu'en est-il ?

Le couplage statique, un problème de transport

Le résultat suivant est positif. On donne un transport explicite entre $\mu_t^\alpha = \text{Loi}(\mathbf{B}_t^\alpha)$ et $\mu_t^{\alpha'} = \text{Loi}(\mathbf{B}_t^{\alpha'})$. Ce transport n'est pas le résultat d'un couplage de processus mais un simple transport tels que ceux que nous avons introduits au chapitre 1. On obtient une nouvelle démonstration de (4.10) pour $p = 1$. Ce cas est une conséquence des autres, le plus fort d'entre eux étant le cas $p = \infty$. Toutefois la démonstration est, pour la première fois, directe, c'est-à-dire qu'elle n'est pas obtenue par la dualité de Kuwada. En retour cette même dualité propose une démonstration indirecte de $q = \infty$ concernant les inégalités de gradient du semi-groupe de la chaleur.

Théorème 4.4.6. *Il existe une constante $C > 0$ telle que pour tout $t \geq 0$ et tout $\alpha = (x, y, z), \alpha' = (x', y', z') \in \mathbb{H}$, il existe un vecteur aléatoire (X, Y, Z, X', Y', Z') de marges $\mu_t^\alpha = \text{Loi}(\mathbf{B}_t^\alpha) = \text{Loi}(X, Y, Z)$ et $\mu_t^{\alpha'} = \text{Loi}(\mathbf{B}_t^{\alpha'}) = \text{Loi}(X', Y', Z')$ tel que*

$$W_1(\mu_t^\alpha, \mu_t^{\alpha'}) \leq \mathbb{E}(d_{cc}((X, Y, Z), (X', Y', Z'))) \leq C d_{cc}(\alpha, \alpha')$$

et

$$\begin{cases} X' = X + (x' - x) & \text{presque sûrement,} \\ Y' = Y + (y' - y) & \text{presque sûrement.} \end{cases}$$

Idée de la démonstration. Pour démontrer ce théorème on part d'une idée naïve : traduire la mesure μ_t^a sur $\mu_t^{a'}$ par le vecteur $a^{-1} \cdot a'$. En effet $W_1(\mu_t^a, (\widehat{\text{trans}}_{a^{-1} \cdot a'})_{\#} \mu_t^a) \leq d_{cc}(0, a^{-1} \cdot a') = d_{cc}(a, a')$. Or $(\widehat{\text{trans}}_{a^{-1} \cdot a'})_{\#} \mu_t^a \neq \mu_t^{a'}$; c'est $(\text{trans}_{a^{-1} \cdot a'})_{\#} \mu_t^a = \mu_t^{a'}$ qui est vrai. Toutefois on a fait un déplacement qu'on peut conserver. Il reste donc à transporter $(\widehat{\text{trans}}_{a^{-1} \cdot a'})_{\#} \mu_t^a$ sur $\mu_t^{a'}$ pour un coût contrôlé par $d_{cc}(a, a')$.

Du fait des symétries du groupe de Heisenberg, il est possible de réduire le problème à $a = (0, 0, 0)$, $a' = (-x', 0, 0)$ et $t = 1$. En posant

$$\mu = \mu_1^0 = \text{Loi}(X, Y, Z) \quad (4.19)$$

$$\nu = \mu_1^{(x', 0, 0)} = \text{Loi}[(x', 0, 0) \cdot (X, Y, Z)] = \text{Loi}(X + x', Y, Z + (1/2)x'Y) \quad (4.20)$$

$$\tilde{\mu} = \text{Loi}[(x', 0, 0) \cdot (X, Y, Z) \cdot (-x', 0, 0)] = \text{Loi}(X, Y, Z + x'Y) \quad (4.21)$$

l'approche qu'on vient juste de présenter consiste finalement à chercher une borne uniforme linéaire en $|x'|$ pour $W_1(\mu, \tilde{\mu})$. Le couplage suggéré par les variables aléatoires apparaissant aux lignes (4.19) et (4.21) donne

$$\begin{aligned} W_1(\mu, \tilde{\mu}) &\leq \mathbb{E}[\mathbb{E}[d_{cc}((X, Y, Z), (X, Y, Z + x'Y)) | (X, Y)]] \\ &\leq \mathbb{E}[\mathbb{E}[d_{cc}((0, 0, 0), (0, 0, x'Y)) | (X, Y)]] \\ &\leq \mathbb{E}[2\sqrt{\pi}|Yx'|] = \sqrt{|x'|} \times [2\sqrt{\pi}\mathbb{E}(\sqrt{|Y|})]. \end{aligned} \quad (4.22)$$

L'ordre de grandeur $\sqrt{|x'|}$ ne convient pas puisque $\sqrt{|x'|}/d_{cc}((x', 0, 0), (0, 0, 0)) = |x'|^{-1/2} \rightarrow_{x' \rightarrow 0} \infty$. Nous avons fait la faute de croire que la translation est un transport optimal sur \mathbb{R} pour le coût $c : (x, y) \mapsto |y - x|^{1/2}$. Le lemme suivant constitue un changement par rapport à ce que nous avons vu à la proposition 1.2.15, à savoir que pour $p \geq 1$, l'inégalité

$$\inf_{\pi \in \text{Marg}(\mu, \nu)} \iint |y - x|^p d\pi(x, y) \geq |\text{Bary}(\mu) - \text{Bary}(\nu)|^p$$

est valide avec de plus égalité lorsqu'une mesure est la translatée de l'autre. Comme l'indique le lemme 4.4.7, pour $p = 1/2$ la situation est tout autre, ce qui permet d'obtenir le théorème 4.4.6.

Lemme 4.4.7. *Soit $\eta \in \mathcal{P}(\mathbb{R})$ une mesure possédant une densité $f = d\eta/d\lambda$ à décroissance rapide. Alors, pour tout $t \in \mathbb{R}$, le coût de transport global associé à $c : (x, y) \mapsto |y - x|^{1/2}$, entre η et $(\text{trans}_t^{\mathbb{R}})_{\#} \eta$, la mesure translatée de t peut être majorée linéairement de la façon suivante :*

$$\inf_{\pi \in \text{Marg}(\eta, (\text{trans}_t^{\mathbb{R}})_{\#} \eta)} \iint \sqrt{|y - x|} d\pi(x, y) \leq C_\eta \times |t|,$$

où $C_\eta = \left(\int |f'(x)|\sqrt{|x|} dx \right)$.

□

Remarque 4.4.8. *On a obtenu $C_1 \leq 2\sqrt{\pi} \iiint |y|\sqrt{|z|} \times |\partial_z h(x, y, z)| d\mathcal{L}(x, y, z)$, où C_1 est la constante optimale discutée à la remarque 4.1.5.*

4.5 Déformation du groupe de Heisenberg par la transformation de Gigli et Mantegazza

À l'aide de la diffusion de la chaleur, Gigli et Mantegazza [GM14] ont défini, pour les variétés riemanniennes compactes et certains espaces RCD (espaces riemanniens à courbure et dimension bornées en un sens faisant intervenir le transport optimal et l'énergie de Cheeger–Dirichlet), une transformation remplaçant l'espace métrique par un autre censé être plus régulier ou lisse. C'est en fait une famille à un paramètre de transformations indexées par le temps $t \geq 0$ de la chaleur

qui est définie, donnant naissance à autant d'espaces métriques. Son intérêt, motivant l'étude, est le fait qu'à droite au temps $t = 0$, la famille d'espaces métriques évolue comme le flot de Ricci au premier ordre, dans un sens précis (voir [GM14, Theorem 1.1]). Dans [ErbJ18] nous avons étudié différents espaces avec l'espoir qu'ils soient transformés en variétés riemanniennes.

Le groupe de Heisenberg n'entre dans aucune des catégories envisagées par Gigli et Mantegazza mais le principe, simple, de la transformation se transpose sans difficulté. Nous pouvons définir une distance de la façon suivante :

$$\tilde{d}_t(\mathbf{a}, \mathbf{a}') = W_{cc}(\mathbf{v}_\mathbf{a}^t, \mathbf{v}_{\mathbf{a}'}^t).$$

À cette distance on associe la distance intrinsèque, à savoir la distance de longueur définie par

$$d_t(\mathbf{a}, \mathbf{a}') = \inf \int_0^T |\dot{\mathbf{a}}_s|_t ds = \inf \int_0^T |\dot{\mathbf{v}}_{\mathbf{a}_s}^t| ds.$$

Ici, l'infimum est obtenu en considérant les courbes absolument continues (ou de façon équivalente lipschitziennes) $(\mathbf{a}_s)_{s \in [0, T]}$ dans $(\mathbb{H}, \tilde{d}_t)$ telles que $\mathbf{a}_0 = \mathbf{a}$, $\mathbf{a}_T = \mathbf{a}'$ et les quantités $|\dot{\mathbf{a}}_s|_t$ et $|\dot{\mathbf{v}}_{\mathbf{a}_s}^t|$ sont les dérivées métriques des deux courbes par rapport, respectivement, à \tilde{d}_t et W_{cc} . Enfin nous démontrons dans [ErbJ18] que l'on peut aussi procéder avec les courbes absolument continues *d'ordre deux*. C'est ici que la proposition 4.2.4 sur la représentation eulérienne de l'équation de continuité peut entrer en jeu. Notons que pour des courbes d'énergie minimale paramétrées sur $[0, 1]$, l'énergie est le carré de la longueur.

Nous avons donc une transformation qui remplace (\mathbb{H}, d_{cc}) par (\mathbb{H}, d_t) pour tout $t > 0$. L'étude menée dans [ErbJ18] conduit au théorème suivant.

Théorème 4.5.1. *La distance d_t est la distance riemannienne invariante à gauche induite par un tenseur g_t . Précisément, pour tout $t > 0$,*

$$d_t = K \cdot d_{\text{Riem}(\kappa\sqrt{t})}$$

c'est-à-dire

$$g_t = \begin{pmatrix} K^2 & 0 & 0 \\ 0 & K^2 & 0 \\ 0 & 0 & K^2/\kappa^2 t \end{pmatrix}$$

dans la base $\mathbf{X}, \mathbf{Y}, \mathbf{U}$. De plus, les constantes K et κ satisfont à $K \geq \sqrt{2}$ et $K/\kappa < \sqrt{2}$.

Ainsi, lorsque t tend vers zéro, (\mathbb{H}, d_t) converge vers $(\mathbb{H}, Kd_{cc}) \equiv (\mathbb{H}, d_{cc})$ dans la topologie de Gromov-Hausdorff pointée.

La démonstration de ce théorème passe par l'équation de continuité et par l'identification pour chacune des directions $m\mathbf{X} + n\mathbf{Y} + q\mathbf{U}$ d'un champ de vecteurs optimal $\mathbf{V}_\mathbf{a}^{m,n,q}$, plus précisément un champ de vecteurs dans $T_{\mathbf{v}_{\mathbf{a}_s}^t} \mathcal{P}(\mathbb{H})$ tel que dans la proposition 4.2.4 alors que $(\mathbf{a}_s)_s$ passe par \mathbf{a} dans la direction $m\mathbf{X} + n\mathbf{Y} + q\mathbf{U}$. Peu importe quel est exactement $\mathbf{a}_s \in \mathbb{H}$ puisque l'ensemble des mesures $\mathbf{v}_\mathbf{a}^t$ est invariant par translation à gauche. Ainsi $\mathbf{V}_\mathbf{a}^{m,n,q} = (D \text{trans}_\mathbf{a}) \cdot \mathbf{V}_{0_{\mathbb{H}}}^{m,n,q}$. L'équation de continuité étant linéaire, il s'avère que $(m, n, q) \mapsto \mathbf{V}_{0_{\mathbb{H}}}^{m,n,q}$ est aussi linéaire. Les commentaires suivants concernent les valeurs des constantes numériques dans le théorème 4.5.1.

Remarque 4.5.2 (Estimation de κ). *La différence principale entre d_t et d_{cc} est que d'autres courbes que les courbes horizontales peuvent être de longueur finie. Une même courbe $(\mathbf{a}_s)_s$ peut être de longueur infinie par rapport à d_{cc} et pourtant vérifier que $(\mathbf{v}_{\mathbf{a}_s}^t)_s$ est de longueur finie par rapport à la distance de Wasserstein W_{cc} .*

Prenons un exemple de base, celui de $s \mapsto \mathbf{a}_s = (0, 0, s)$. Cette courbe n'est pas horizontale et sa longueur est infinie car $d_{cc}(\mathbf{q}_s, \mathbf{q}_r) = 2 \cdot \sqrt{\pi|s-r|}$ (voir proposition 4.1.1). Cependant la courbe $\mathbf{v}_{\mathbf{q}_s}^t = \rho_s^t \mathcal{L}$ où $\rho_s(\mathbf{p}) = \mathfrak{h}_t(\mathbf{q}_s^{-1} \mathbf{p})$ vérifie l'équation de continuité

$$\partial_s \rho_s = -\mathbf{U} \rho_s = -[\mathbf{X}, \mathbf{Y}] \rho_s = -\mathbf{X}(\mathbf{Y} \rho_s) + \mathbf{Y}(\mathbf{X} \rho_s) = -\text{div}(\rho_s \mathbf{V}_s)$$

associée aux champs de vecteurs horizontaux (mais non de type gradient) $(V_s)_s$ où $V_s := (\mathbf{Y} \log \rho_s) \mathbf{X} - (\mathbf{X} \log \rho_s) \mathbf{Y}$. Ainsi nous obtenons

$$|\dot{v}_{q_s}| < \|V_s\|_{L^2_{cc}} = \sqrt{\int \frac{(\mathbf{X}h_t)^2 + (\mathbf{Y}h_t)^2}{h_t} d\mathcal{L}} = \sqrt{\frac{2}{t}}$$

comme conséquence de (4.6). En conclusion (v_{q_s}) est bien de longueur finie pour la longueur attachée à W_{cc} et $K/\kappa = g_1(\mathbf{U}) \leq \sqrt{2}$.

Remarque 4.5.3 (Estimation de K). Rappelons la définition de C_2 , la constante optimale discutée à la remarque 4.1.5 :

$$C_2 = \sup_{p \neq q} \frac{W_{cc}(v_a^t, v_{a'}^t)}{d_{cc}(a, a')}$$

Nous pouvons établir $K = C_2$, donnant un nouvel éclairage à cette constante optimale. On a

$$\frac{\tilde{d}_t}{d_{cc}} = \frac{d_t}{d_{cc}} \frac{\tilde{d}_t}{d_t},$$

où $d_t/d_{cc} \leq K$ et $\tilde{d}_t/d_t \leq 1$. Mais en posant $a_s = (s, 0, 0)$ on s'aperçoit que le quotient des distances entre 0 et a_s tend, respectivement, vers K et 1 lorsque s tend vers 0. Ainsi $g_1(\mathbf{X}) = K = C_2 \geq \sqrt{2}$.

Remarque 4.5.4 (Formule concernant C_2). Pour $t = 1$, le vecteur $\mathbf{V}_0^{m,n,q}$ est caractérisé par le fait qu'il vérifie $\operatorname{div}(h\mathbf{V}_0^{m,n,q}) = \operatorname{div}(h(m\hat{\mathbf{X}} + n\hat{\mathbf{Y}} + q\hat{\mathbf{U}}))$.

Au sujet de la constante C_2 , on peut finalement constater que c'est le minimum de $\|\mathbf{V}\|_{L^2(h)}$ pour \mathbf{V} satisfaisant à $\operatorname{div}(h\mathbf{V}) = \operatorname{div}(h\hat{\mathbf{X}})$. De même pour C_2/κ qui est le minimum de $\|\mathbf{V}\|_{L^2(h)}$ pour \mathbf{V} satisfaisant à $\operatorname{div}(h\mathbf{V}) = \operatorname{div}(h\hat{\mathbf{U}}) = \operatorname{div}(h\mathbf{U})$.

Remarque 4.5.5 (Cas des cônes euclidiens de dimension 2). En quotientant $\mathbb{R}^2 \equiv \mathcal{C}$ par la symétrie centrale $R_\pi : z \mapsto -z$ ou bien par la rotation d'un quart de tour $R_{\pi/2} : z \mapsto e^{i\pi/2}z$, on obtient deux cônes euclidiens dont les angles au sommet sont, respectivement π et $\pi/2$. Ces espaces entrent à la fois dans la classe des espaces d'Alexandrov à courbure sectionnelle positive et celle des espaces RCD. Hormis la singularité conique au sommet ce sont des variétés riemanniennes. Dans le cas compact, ce type de variétés avec singularité a été étudié dans [BLM17] sans que les auteurs parviennent à savoir si la transformation de Gigli et Mantegazza pouvait « gommer » la singularité.

Dans [ErbJ18] nous avons calculé le comportement étrange suivant : pour le premier cône, l'angle au sommet passe de π à $\sqrt{2}\pi$; pour le second il passe de $\pi/2$ à 0.

Problème 4.5.6. Étudier la fonction qui à l'angle au sommet d'un cône droit associe l'angle du cône déformé.

Remarque 4.5.7. La transformation laisse les espaces vectoriels normés invariants. La remarque revêt de l'intérêt quand on considère ceux-ci comme les représentants les plus simples des variétés de Finsler, les cônes étant ceux des espaces d'Alexandrov et \mathbb{H} celui des géométries de Carnot (c'est-à-dire sous-riemanniennes).

Remarque 4.5.8. Hormis Gigli et Mantegazza, d'autres auteurs ont exploité les liens entre diffusion de la chaleur et flot de Ricci. On peut notamment citer les travaux sur les surflois de Ricci dont font partie [MT10, ACT08, CP11, GPT15, KS16⁺, HN15⁺].

Bibliographie

- [ACT08] M. Arnaudon, K. A. Coulibaly, et A. Thalmaier. Brownian motion with respect to a metric depending on time : definition, existence and applications to Ricci flow. *C. R. Math. Acad. Sci. Paris*, 346(13-14) :773–778, 2008. 91
- [AG08] L. Ambrosio et N. Gigli. Construction of the parallel transport in the Wasserstein space. *Methods Appl. Anal.*, 15(1) :1–29, 2008. 15
- [AGS08] L. Ambrosio, N. Gigli, et G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008. 15, 30, 77, 82
- [Alb94] G. Alberti. On the structure of singular sets of convex functions. *Calc. Var. Partial Differential Equations*, 2(1) :17–27, 1994. 30
- [AR04] L. Ambrosio et S. Rigot. Optimal mass transportation in the Heisenberg group. *J. Funct. Anal.*, 208(2) :261–301, 2004. 31
- [ATW06] M. Arnaudon, A. Thalmaier, et F.-Y. Wang. Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below. *Bull. Sci. Math.*, 130(3) :223–233, 2006. 86
- [BACK95] G. Ben Arous, M. Cranston, et W. S. Kendall. Coupling constructions for hypoelliptic diffusions : two examples. In *Stochastic analysis (Ithaca, NY, 1993)*, volume 57 of *Proc. Sympos. Pure Math.*, pages 193–212. Amer. Math. Soc., Providence, RI, 1995. 86
- [BB00] J.-D. Benamou et Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3) :375–393, 2000. 77
- [BBBC08] D. Bakry, F. Baudoin, M. Bonnefont, et D. Chafaï. On gradient bounds for the heat kernel on the Heisenberg group. *J. Funct. Anal.*, 255(8) :1905–1938, 2008. 80, 81
- [BBI01] D. Burago, Y. Burago, et S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. 30
- [BC10] S. Bianchini et L. Caravenna. On optimality of c -cyclically monotone transference plans. *C. R. Math. Acad. Sci. Paris*, 348(11-12) :613–618, 2010. 22
- [BCH17] M. Beiglböck, A. M. G. Cox, et M. Huesmann. Optimal transport and Skorokhod embedding. *Invent. Math.*, 208(2) :327–400, 2017. 37, 54
- [Bei15] M. Beiglböck. Cyclical monotonicity and the ergodic theorem. *Ergodic Theory Dynam. Systems*, 35(3) :710–713, 2015. 22
- [BeiJ16] M. Beiglböck et N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1) :42–106, 2016. 11, 15, 19, 22, 37, 38, 39, 43, 45, 46, 55
- [BeiJ16⁺] M. Beiglböck et N. Juillet. Shadow couplings. *ArXiv e-prints*, septembre 2016. 11, 16, 43, 54
- [BG17] F. Baudoin et N. Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *J. Eur. Math. Soc. (JEMS)*, 19(1) :151–219, 2017. 87
- [BGL14] D. Bakry, I. Gentil, et M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014. 15

- [BGMS09] M. Beiglböck, M. Goldstern, G. Maresch, et W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6) :1907–1927, 2009. 22, 37
- [BGP92] Y. Burago, M. Gromov, et G. Perelman. A. D. Aleksandrov spaces with curvatures bounded below. *Uspekhi Mat. Nauk*, 47(2(284)) :3–51, 222, 1992. 30
- [BHHJ18⁺⁺] M. Brücknerhoff-Plückelmann, M. Huesmann, et N. Juillet. A left-monotone solution to the peacock problem. *Article en préparation*, 2018. 67
- [BHL13] M. Beiglböck, P. Henry-Labordère, et F. Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.*, 17(3) :477–501, 2013. 15
- [BHS16] M. Beiglböck, M. Huesmann, et F. Stebegg. Root to Kellerer. In *Séminaire de Probabilités XLVIII*, volume 2168 of *Lecture Notes in Math.*, pages 1–12. Springer, Cham, 2016. 61, 64
- [BK12] V. I. Bogachev et A. V. Kolesnikov. The Monge-Kantorovich problem : achievements, connections, and prospects. *Uspekhi Mat. Nauk*, 67(5(407)) :3–110, 2012. 15
- [BLM17] L. Bandara, S. Lakzian, et M. Munn. Geometric singularities and a flow tangent to the Ricci flow. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 17(2) :763–804, 2017. 91
- [BNT17] M. Beiglböck, M. Nutz, et N. Touzi. Complete duality for martingale optimal transport on the line. *Ann. Probab.*, 45(5) :3038–3074, 2017. 37
- [BonJ18⁺] M. Bonnefont et N. Juillet. Couplings in L^p distance of two Brownian motions and their Lévy area. *ArXiv e-prints*, janvier 2018. 12, 78, 80, 86, 87, 88
- [BouJ18⁺] C. Boubel et N. Juillet. The Markov-quantile process attached to a family of Marginals. *ArXiv e-prints*, avril 2018. 12, 16, 57, 61, 62, 63, 70, 71, 73, 74, 76, 82, 84, 85, 86
- [Bre87] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 305(19) :805–808, 1987. 28
- [Bre91] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4) :375–417, 1991. 15, 28
- [BS11] M. Beiglböck et W. Schachermayer. Duality for Borel measurable cost functions. *Trans. Amer. Math. Soc.*, 363(8) :4203–4224, 2011. 22
- [BY14] J. Bertoin et M. Yor. Local times for functions with finite variation : two versions of Stieltjes change-of-variables formula. *Bull. Lond. Math. Soc.*, 46(3) :553–560, 2014. 34
- [CP11] K. A. Coulibaly-Pasquier. Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(2) :515–538, 2011. 91
- [Cra91] M. Cranston. Gradient estimates on manifolds using coupling. *J. Funct. Anal.*, 99(1) :110–124, 1991. 86, 87
- [Cra92] M. Cranston. A probabilistic approach to gradient estimates. *Canad. Math. Bull.*, 35(1) :46–55, 1992. 87
- [DM05] B. K. Driver et T. Melcher. Hypocoelliptic heat kernel inequalities on the Heisenberg group. *J. Funct. Anal.*, 221(2) :340–365, 2005. 81
- [Dou64] R. G. Douglas. On extremal measures and subspace density. *Michigan Math. J.*, 11 :243–246, 1964. 48

- [ErbJ18] M. Erbar et N. Juillet. Smoothing and non-smoothing via a flow tangent to the Ricci flow. *J. Math. Pures Appl. (9)*, 110 :123–154, 2018. 12, 77, 78, 90, 91
- [FigJ08] A. Figalli et N. Juillet. Absolute continuity of Wasserstein geodesics in the Heisenberg group. *J. Funct. Anal.*, 255(1) :133–141, 2008. 31
- [FR10] A. Figalli et L. Rifford. Mass transportation on sub-Riemannian manifolds. *Geom. Funct. Anal.*, 20(1) :124–159, 2010. 31
- [Gav77] B. Gaveau. Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents. *Acta Math.*, 139(1-2) :95–153, 1977. 80
- [GHLT14] A. Galichon, P. Henry-Labordère, et N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.*, 24(1) :312–336, 2014. 15
- [Gig11] N. Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(\mathcal{P}_2(M), W_2)$. *Methods Appl. Anal.*, 18(2) :127–158, 2011. 28, 29, 30, 31
- [Gig12] N. Gigli. Second order analysis on $(\mathcal{P}_2(M), \mathcal{W}_2)$. *Mem. Amer. Math. Soc.*, 216(1018) :xii+154, 2012. 29
- [GL10] N. Gozlan et C. Léonard. Transport inequalities. A survey. *Markov Process. Related Fields*, 16(4) :635–736, 2010. 15
- [GM96] W. Gangbo et R. J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2) :113–161, 1996. 29, 32
- [GM14] N. Gigli et C. Mantegazza. A flow tangent to the Ricci flow via heat kernels and mass transport. *Adv. Math.*, 250 :74–104, 2014. 77, 89, 90
- [GPT15] H. Guo, R. Philipowski, et A. Thalmaier. An entropy formula for the heat equation on manifolds with time-dependent metric, application to ancient solutions. *Potential Anal.*, 42(2) :483–497, 2015. 91
- [GRST17] N. Gozlan, C. Roberto, P.-M. Samson, et P. Tetali. Kantorovich duality for general transport costs and applications. *J. Funct. Anal.*, 273(11) :3327–3405, 2017. 55
- [GS62] D. Gale et L. S. Shapley. College Admissions and the Stability of Marriage. *Amer. Math. Monthly*, 69(1) :9–15, 1962. 33
- [GT16] E. Grong et A. Thalmaier. Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations : part I. *Math. Z.*, 282(1-2) :99–130, 2016. 87
- [HK15] D. Hobson et M. Klimmek. Robust price bounds for the forward starting straddle. *Finance Stoch.*, 19(1) :189–214, 2015. 15, 37
- [HLT16] P. Henry-Labordère et N. Touzi. An explicit martingale version of the one-dimensional Brenier theorem. *Finance Stoch.*, 20(3) :635–668, 2016. 39
- [HN12] D. Hobson et A. Neuberger. Robust bounds for forward start options. *Math. Finance*, 22(1) :31–56, 2012. 15, 37
- [HN15⁺] R. Haslhofer et A. Naber. Weak solutions for the Ricci flow I. *ArXiv e-prints*, avril 2015. 91
- [Hob98] D. G. Hobson. The maximum maximum of a martingale. In *Séminaire de probabilités XXXII*, pages 250–263. Berlin : Springer, 1998. 64
- [HPRY11] F. Hirsch, C. Profeta, B. Roynette, et M. Yor. *Peacocks and associated martingales, with explicit constructions*, volume 3 of *Bocconi & Springer Series*. Springer, Milan, 2011. 45, 64, 66

- [HR12] F. Hirsch et B. Roynette. A new proof of Kellerer’s theorem. *ESAIM Probab. Stat.*, 16 :48–60, 2012. 45
- [HR13] F. Hirsch et B. Roynette. On \mathbb{R}^d -valued peacocks. *ESAIM Probab. Stat.*, 17 :444–454, 2013. 66
- [HRY14] F. Hirsch, B. Roynette, et M. Yor. Kellerer’s theorem revisited. In Springer, editor, *Asymptotic Laws and Methods in Stochastics. Volume in Honour of Miklos Csorgo*, Fields Institute Communications Series, 2014. 16, 61
- [HTT16] P. Henry-Labordère, X. Tan, and N. Touzi. An explicit martingale version of the one-dimensional Brenier’s theorem with full marginals constraint. *Stochastic Process. Appl.*, 126(9) :2800–2834, 2016. 68
- [HY10] F. Hirsch et M. Yor. Looking for martingales associated to a self-decomposable law. *Electron. J. Probab.*, 15 :no. 29, 932–961, 2010. 16, 59
- [JKO98] R. Jordan, D. Kinderlehrer, et F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1) :1–17, 1998. 77
- [JSig17] N. Juillet et M. Sigalotti. Pliability, or the Whitney extension theorem for curves in Carnot groups. *Anal. PDE*, 10(7) :1637–1661, 2017. 80
- [Jui09] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *International Mathematical Research Notices*, 2009, 2009. 87
- [Jui11] N. Juillet. On displacement interpolation of measures involved in Brenier’s theorem. *Proc. Amer. Math. Soc.*, 139(10) :3623–3632, 2011. 11, 19, 27, 28, 29, 30, 31
- [Jui14] N. Juillet. Diffusion by optimal transport in Heisenberg groups. *Calc. Var. Partial Differential Equations*, 50(3-4) :693–721, 2014. 13, 79, 83
- [Jui16a] N. Juillet. Peacocks parametrised by a partially ordered set. In *Séminaire de Probabilités XLVIII*, volume 2168 of *Lecture Notes in Math.*, pages 13–32. Springer, Cham, 2016. 12, 57, 64, 66, 67
- [Jui16b] N. Juillet. Stability of the shadow projection and the left-curtain coupling. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4) :1823–1843, 2016. 11, 19, 40, 43, 44, 49, 50, 51
- [Jui18] N. Juillet. Martingales associated to peacocks using the curtain coupling. *Electron. J. Probab.*, 23 :Paper No. 8, 29, 2018. 12, 57, 68, 69
- [Jui18⁺⁺] N. Juillet. Optimal transport on the line for power costs close to but smaller than one. *Article en préparation*, 2018. 11, 13, 19, 32, 34
- [Kan42] L. Kantorovitch. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37 :199–201, 1942. 15
- [Kel72] H. G. Kellerer. Markov-Komposition und eine Anwendung auf Martingale. *Math. Ann.*, 198 :99–122, 1972. 58, 60, 61, 66
- [Kel73] H. G. Kellerer. Integraldarstellung von Dilationen. In *Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes (Tech. Univ., Prague, 1971; dedicated to the memory of Antonín Špaček)*, pages 341–374. Academia, Prague, 1973. 53, 60, 61, 63
- [Kel84] H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4) :399–432, 1984. 22, 37
- [Ken07] W. S. Kendall. Coupling all the Lévy stochastic areas of multidimensional Brownian motion. *Ann. Probab.*, 35(3) :935–953, 2007. 86

- [Ken10] W. S. Kendall. Coupling time distribution asymptotics for some couplings of the Lévy stochastic area. In *Probability and mathematical genetics*, volume 378 of *London Math. Soc. Lecture Note Ser.*, pages 446–463. Cambridge Univ. Press, Cambridge, 2010. 86
- [KS07] K. Kuwada et K.-T. Sturm. A counterexample for the optimality of Kendall-Cranston coupling. *Electron. Comm. Probab.*, 12 :66–72, 2007. 86
- [KS16⁺] E. Kopfer et K.-T. Sturm. Heat Flows on Time-dependent Metric Measure Spaces and Super-Ricci Flows. *À paraître dans Comm. Pure Appl. Math.*, November 2016. 91
- [Kuw10] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11) :3758–3774, 2010. 78, 81
- [Li06] H.-Q. Li. Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg. *J. Funct. Anal.*, 236(2) :369–394, 2006. 78, 81
- [Lin99] T. Lindvall. On Strassen’s theorem on stochastic domination. *Electron. Comm. Probab.*, 4 :51–59 (electronic), 1999. Voir aussi le corrigendum. 44
- [Lis07] S. Lisini. Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Partial Differential Equations*, 28(1) :85–120, 2007. 86
- [LMT14] G. Last, P. Mörters, et H. Thorisson. Unbiased shifts of Brownian motion. *Ann. Probab.*, 42(2) :431–463, 2014. 33, 34
- [Low08⁺] G. Lowther. Fitting Martingales To Given Marginals. *ArXiv e-prints*, août 2008. 61, 64
- [Low09] G. Lowther. Limits of one-dimensional diffusions. *Ann. Probab.*, 37(1) :78–106, 2009. 16
- [Loy70] R. M. Loynes. Stopping times on Brownian motion : Some properties of Root’s construction. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 16 :211–218, 1970. 55
- [LS97] U. Lang et V. Schroeder. Kirszbraun’s theorem and metric spaces of bounded curvature. *Geom. Funct. Anal.*, 7(3) :535–560, 1997. 60
- [LTT16⁺] G. Last, W. Tang, et H. Thorisson. Transporting random measures on the line and embedding excursions into Brownian motion. *À paraître dans Ann. Inst. Henri Poincaré Probab. Stat.*, août 2016. 34
- [LV09] J. Lott et C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3) :903–991, 2009. 30
- [Man07] S. Maniglia. Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. *J. Math. Pures Appl. (9)*, 87(6) :601–626, 2007. 82
- [Mat95] P. Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability. 30
- [MG13] R. J. McCann et N. Guillen. Five lectures on optimal transportation : geometry, regularity and applications. In *Analysis and geometry of metric measure spaces*, volume 56 of *CRM Proc. Lecture Notes*, pages 145–180. Amer. Math. Soc., Providence, RI, 2013. 15, 28
- [Mon81] G. Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l’académie Royale des Sciences de Paris*, 1781. 15

- [MT10] R. J. McCann et P. M. Topping. Ricci flow, entropy and optimal transportation. *Amer. J. Math.*, 132(3) :711–730, 2010. 91
- [NS16] M. Nutz et F. Stebegg. Canonical Supermartingale Couplings. *À paraître dans Ann. Probab.*, septembre 2016. 39
- [NST17⁺] M. Nutz, F. Stebegg, et X. Tan. Multiperiod Martingale Transport. *ArXiv e-prints*, mars 2017. 39
- [Obł04] J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1 :321–390, 2004. 54
- [Paj02] H. Pajot. *Analytic capacity, rectifiability, Menger curvature and the Cauchy integral*, volume 1799 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. 30
- [PC18⁺] G. Peyré et M. Cuturi. Computational optimal transport. Prépublication 1803.00567, Arxiv, 2018. 15
- [Pet98] A. Petrunin. Parallel transportation for Alexandrov space with curvature bounded below. *Geom. Funct. Anal.*, 8(1) :123–148, 1998. 30
- [PP16] M. N. Pascu et I. Popescu. Shy and fixed-distance couplings of Brownian motions on manifolds. *Stochastic Process. Appl.*, 126(2) :628–650, 2016. 86
- [Rif14] L. Rifford. *Sub-Riemannian geometry and optimal transport*. SpringerBriefs in Mathematics. Springer, Cham, 2014. 31
- [RR98a] S. T. Rachev et L. Rüschendorf. *Mass transportation problems. Vol. I. Probability and its Applications (New York)*. Springer-Verlag, New York, 1998. Theory. 15
- [RR98b] S. T. Rachev et L. Rüschendorf. *Mass transportation problems. Vol. II. Probability and its Applications (New York)*. Springer-Verlag, New York, 1998. Applications. 15
- [San15] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling. 15, 28
- [Str65] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36 :423–439, 1965. 44, 58
- [Stu06] K.-T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1) :133–177, 2006. 30
- [Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003. 15, 25, 28
- [Vil09] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 2009. 15, 31, 41, 86
- [vR04] M.-K. von Renesse. Intrinsic coupling on Riemannian manifolds and polyhedra. *Electron. J. Probab.*, 9 :no. 14, 411–435, 2004. 86
- [vRS05] M.-K. von Renesse et K.-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7) :923–940, 2005. 87
- [Whi34] H. Whitney. Differentiable functions defined in closed sets. I. *Trans. Amer. Math. Soc.*, 36(2) :369–387, 1934. 87
- [Zaj79] L. Zajíček. On the differentiation of convex functions in finite and infinite dimensional spaces. *Czechoslovak Math. J.*, 29(104)(3) :340–348, 1979. 29