Université Paris Cité

École Doctorale n°386 Sciences Mathématiques de Paris Centre
Laboratoire Jacques-Louis Lions

Processus stochastiques, transport de mesures et inégalités fonctionnelles

Stochastic processes, mass transportation, and functional inequalities

Par Pablo LÓPEZ RIVERA

Thèse de doctorat de MATHÉMATIQUES

Dirigée par Max FATHI

Soutenue publiquement le 12 juin 2025 devant le jury composé de :

Martin HUESMANN	PR à l'Université de Münster	Rapporteur
Christian LÉONARD	PR à l'Université Paris Nanterre	Rapporteur
Nathaël GOZLAN	PR à l'Université Paris Cité	Président
Hélène HALCONRUY	MCF à Télécom SudParis	Examinatrice
Maxime LABORDE	MCF à l'Université Paris Cité	Examinateur
Max FATHI	PR à l'Université Paris Cité	Directeur

Titre : Processus stochastiques, transport de mesures et inégalités fonctionnelles.

Mots clefs : inégalités fonctionnelles ; transport de mesures ; processus stochastiques ; application de transport ; inégalités de Sobolev logarithmiques ; inégalités de Sobolev logarithmiques modifiées ; transport optimal entropique.

Résumé : Dans cette thèse, nous nous intéressons aux interactions de trois sujets : les processus stochastiques, le transport optimal et les inégalités fonctionnelles. Le manuscrit comporte deux parties : la première est consacrée aux notions préliminaires concernant les contributions originales de la thèse et la deuxième présente ces contributions.

La première partie est divisée en deux chapitres. Dans le premier chapitre, nous commençons par une introduction aux fondamentaux de la théorie du transport optimal, en mettant l'accent sur le cas quadratique euclidien, la structure de ses solutions et leur régularité. Dans le deuxième chapitre, nous offrons un panorama de la théorie des inégalités fonctionnelles en commençant par les inégalités géométriques et leurs versions fonctionnelles. Nous nous intéressons notamment aux inégalités isopérimétriques et au phénomène de concentration de la mesure qui occupent une place centrale dans la théorie. Finalement, nous étudions différentes familles d'inégalités fonctionnelles qui entraînent la concentration de la mesure : les inégalités de Poincaré, de Sobolev logarithmiques, de Sobolev logarithmiques modifiées et de transport-entropie.

La deuxième partie, constituée des quatre derniers chapitres du manuscrit, présente les contributions originales de la thèse. Dans le troisième chapitre, nous montrons que pour toute variété riemannienne à poids et à courbure contrôlée aux premier et deuxième ordres dans le sens de Bakry-Émery, l'application de Kim-Milman, issue du processus de Langevin associée à la variété, envoyant la mesure de poids sur toute perturbation log-Lipschitz est alors lipschtizienne ; ce résultat permet le transfert d'inégalités fonctionnelles.

Dans le quatrième chapitre, nous construisons une application de transport entre le processus de Poisson ponctuel et des mesures ultra-log-concaves sur les entiers naturels et nous montrons qu'elle est contractante. Cette approche permet de surmonter les difficultés qui entravent le transfert d'inégalités fonctionnelles dans le cadre discret en utilisant des applications de transport. En conséquence, nous obtenons de nouvelles inégalités fonctionnelles pour des mesures ultra-log-concaves. En particulier, notre approche permet d'améliorer la constante connue pour l'inégalité de Sobolev logarithmique modifiée pour les mesures ultra-log-concaves.

Dans le cinquième chapitre, nous exhibons une preuve alternative de l'inégalité de Sobolev logarithmique modifiée de Wu pour la mesure de Poisson via une formulation stochastique variationnelle de l'entropie. Cette technique nous permet de caractériser les cas d'égalité et de montrer un résultat de stabilité quantitative pour l'inégalité sous des hypothèses de convexité.

Dans le sixième et dernier chapitre, dans le contexte de la régularisation entropique du transport optimal, nous exhibons une borne pour le taux de la convergence uniforme sur des ensembles compacts pour les potentiels entropiques et leurs gradients vers le potentiel de Brenier et son gradient, respectivement. Ces résultats sont valides dans le cadre quadratique euclidien, pour des mesures absolument continues sous des hypothèses de convexité.

Title: Stochastic processes, mass transportation, and functional inequalities.

Keywords: functional inequalities; mass transportation; stochastic processes; transport maps; logarithmic Sobolev inequalities; modified logarithmic Sobolev inequalities; entropic optimal transport.

Abstract: In this thesis, we are interested in the interaction between stochastic processes, optimal transport, and functional inequalities. The manuscript is divided into two parts: the first is devoted to the preliminary notions necessary for the understanding of the original contributions of this thesis, and the second presents original contributions to these fields.

In the first part, which is divided into two chapters, we start by providing an introduction to the basic theory of optimal transport in the first chapter, with an accent on the quadratic Euclidean problem, the structure of its solutions, and their regularity. In the second chapter, we survey the theory of functional inequalities, starting from geometric inequalities and their functional counterparts, where isoperimetric inequalities and the concentration of measure phenomenon constitute a central part of the theory, to then review different families of functional inequalities that yield concentration of measure: Poincaré, logarithmic Sobolev, modified logarithmic Sobolev, and transport-entropy inequalities.

The second part, which comprises the last four chapters of the manuscript, concerns the original contributions of the thesis. In the third chapter, our main result states that for any weighted Riemannian manifold that has bounded curvature at first and second order in the sense of Bakry-Émery, then the Kim-Milman transport map, which arises from the Langevin diffusion associated with the manifold, between the weighted measure and any log-Lipschitz perturbation of it is Lipschitz. This result allows the transfer of a number of functional inequalities.

In the fourth chapter, which is based on joint work with Yair Shenfeld, we construct a transport map from Poisson point processes onto ultra-log-concave measures over the natural numbers, and show that it is a contraction. This approach overcomes the known obstacles to transferring functional inequalities using transport maps in discrete settings, thus yielding new functional inequalities for ultra-log-concave measures. In particular, we provide the currently best known constant in modified logarithmic Sobolev inequalities for ultra-log-concave measures.

In the fifth chapter, which is based on joint work with Shrey Aryan and Yair Shenfeld, we provide an alternative proof to Wu's modified logarithmic Sobolev inequality for the Poisson measure using a stochastic variational formula for the entropy. We show that this approach leads to the identification of the extremizers of the inequality, as well as a quantitative stability result under convexity assumptions.

In the sixth and final chapter, in the context of the entropic regularization of the optimal transport problem, we exhibit a bound on the rate of uniform convergence in compact sets for both entropic potentials and their gradients towards the Brenier potential and its gradient. Both results hold in the quadratic Euclidean setting for absolutely continuous measures satisfying some convexity assumptions.

A mi familia, a los que están y a los que ya partieron:
Rosa; Alejandra y Hugo; Ella y Gissella; Mateo, Simona y Victoria.

I'll move to Paris [...]
I'll miss the playgrounds and the animals and diggin' up worms
I'll miss the comfort of my mother and the weight of the world
I'll miss my sister, miss my father, miss my dog and my home
Yeah, I'll miss the boredom and the freedom and the time spent alone.

MGMT
Time to Pretend

Fueron jornadas de grandes dudas [...] los números me habían abandonado. Si los números no venían a mí, yo iría hasta la guarida de los números y los sacaría de allí con zalamerías o a patadas.
[...]
Una mañana, tal como esperaba, volvieron los números. Las secuencias, al principio, eran endemoniadas, pero no tardé en encontrarles su lógica. El secreto consistía en plegarse.

Roberto Bolaño
Los detectives salvajes

Je soutins devant le jury de l'université Paris IV – Sorbonne ma thèse de doctorat : *Joris-Karl Huysmans, ou la sortie du tunnel.* Dès le lendemain matin (ou peut-être dès le soir même, je ne peux pas l'assurer, le soir de ma soutenance fut solitaire, et très alcoolisé), je compris qu'une partie de ma vie venait de s'achever, et que c'était probablement la meilleure.

Michel Houellebecq
Soumission

Contents

Remerciements

> Al final del camino me dirán:
> —¿Has vivido? ¿Has amado?
> Y yo, sin decir nada, abriré el corazón lleno de nombres.
>
> PEDRO CASALDÁLIGA

Dans presque toute thèse que j'ai lue, je remarque qu'il y a une sorte d'accord tacite lors de la rédaction de la section de remerciements, le voici : on commence souvent par remercier son directeur de thèse, les rapporteurs et les membres du jury ; ensuite, on mentionne ses collaborateurs, les professeurs et les chercheurs avec lesquels on a échangé, ainsi que ses collègues thésards ; puis, on conclut en rendant hommage à sa famille, à ses amis et à sa copine (si l'on en a une), et souvent, on remercie aussi l'« adhérence topologique »[1] de la réunion de ces groupes de personnes. Je ne suis pas un iconoclaste, mais j'aimerais rompre avec cette structure. Plus précisément, je vais vous raconter ces précieuses quatre années que j'ai vécues en France, en suivant l'ordre dans lequel j'ai rencontré les personnes qui m'ont permis de réaliser mon rêve. Je vais tout de même remercier toutes les personnes mentionnées ci-dessus dans la structure, disons, « traditionnelle », mais je ne vais pas suivre l'ordre typique imposé par cette convention.

Je tiens tout d'abord à remercier la **Fondation Sciences Mathématiques de Paris (FSMP)** pour avoir financé mes études de master et doctorat en France, merci d'avoir fait confiance à moi et mon projet mathématique. C'est grâce à vous que j'ai pu commencer cette aventure et qu'elle s'est si bien déroulée, donc c'est le point de départ de ce récit. Après avoir gagné la bourse de M2 de la FSMP, je suis arrivé à Paris le 30 août 2021, après 14 heures de vol depuis Santiago du Chili, dans un monde où le Covid était toujours présent, mais où la vie quotidienne reprenait peu à peu sa normalité, heureusement. J'étais très content de commencer cette aventure, le rêve de

[1] Bien entendu, il faut également choisir une topologie.

ma vie, mais, d'un autre côté, j'étais très triste, car je quittais ma famille et mes amis, auxquels je remercierai par la suite.

Agradezco a mi tan querida madre, **Alejandra Rivera Neumann**, por todo. Mamá, quizá no me ayudaste directamente a demostrar el teorema 4.17, pero sin tu perpetuo apoyo y sacrificio, esto habría sido imposible (no solamente demostrar el teorema 4.17, sino más bien el haber podido terminar esta tesis). Gracias por haber creído siempre en mí y por haberme apoyado en esta locura que es estudiar matemáticas, por darme ánimo todas las veces en que tuve ganas de tirar la esponja, por recibirme como un rey cada vez que volví a Chile (y con inmunidad diplomática total en la casa), por las reconfortantes llamadas de más de dos horas en las que nos ponemos al día y por tus lindas postales que me hacían llorar a mares de emoción cada vez que llegaban. Esta tesis también es tuya.

Pieza clave de mi familia son mis dos hermanas: gracias por siempre haberse preocupado por mí y por haber estado disponibles ante el más mínimo problema que se me hubiese podido poner por delante, desde un «¿cómo lavo un chaleco de lana?», pasando por un «me duele el estómago, ¿es que tengo apendicitis?», hasta un «ya no doy más». De manera más particular, gracias, **Ella López Rivera**, por ser mi cable a tierra (o mi «cable a Chile») y mi confidente en todos estos años, siempre has estado ahí para ayudarme, escucharme, aconsejarme (y retarme también si corresponde, aunque no me guste). Gracias también por todos los videos de golden retrievers y border collies que me mandas a diario. Gracias, **Gissella López Rivera**, por tu constante ayuda estos años y por haberme salvado en Barcelona de la vorágine que me estaba consumiendo; de no ser por esa pequeña pausa, este manuscrito de tesis no hubiese visto la luz. También, es absolutamente gracias a ti que he vuelto a encantarme con la literatura.

También quiero dar las gracias a mis sobrinos: muchas gracias por aportar juventud a nuestra familia, por su cariño, por todas las risas y por (hasta el día de hoy) dejarme seguir enseñándoles estupideces. Hace cuatro años me fui de Chile y eran unos niños; es bonito ver cómo han crecido tanto, no solamente de porte. Gracias, **Mateo Dujovne López**, por tu sensibilidad, liviandad de sangre y por los excelentes stickers de WhatsApp que siempre me mandas. Gracias, **Simona Dujovne López**, por tu aguda visión de las cosas, por tu cariño incondicional y por saber apreciar el temazo que es la canción de la lavadora. Gracias, **Victoria Rosas López**, por tu eterno buen humor, por todas las videollamadas en las que criticábamos los discos de Bad Bunny recién salidos y hablábamos del futuro y de cómo nos trataba la vida. Sin embargo, Victoria, más allá de eso, gracias por haberte convertido en la hermana chica que nunca tuve.

Por supuesto que quiero agradecer a todos los amigos que dejé en Chile, sobre todo a aquellos que han estado más presentes en estos años fuera del terruño.[2] Sé que es difícil mantener el contacto a la distancia, pero ustedes han estado ahí conmigo, incluso

[2]He dejado de mencionar a algunos nombres. Créanme que no lo hago como una declaración de guerra rencorosa: si hemos perdido el contacto en estos años, sepan que el cariño siempre estará ahí y que la amistad se podrá siempre retomar. Simplemente, en este espacio quise reconocer a aquellos que estuvieron más cerca de mí en estos últimos años.

viniendo a verme a París: todas sus visitas fueron reconfortantes y siempre lo pasé joya con ustedes. Es bonito que, incluso, en algunos casos, ciertas amistades se hayan profundizado en estos años, sobre todo después de una visita o encuentro acá en Europa. Por otro lado, cada vez que volví a Chile me hicieron sentir como si el tiempo no hubiese pasado; gracias por no haberme olvidado, que fue (es) mi eterno miedo. De manera más particular, gracias, **Catalina «Cata» Albert**, **Franco Alfero**, **Matías «Mati» Altamirano**, **Maximiliano «Max» De Amesti**, **Magdalena «Magda» Badal**, **Cloé Blanch**, **Carlos Contreras**, **Felipe «Chap» Contreras**, **Juan d'Etigny**, **Claudia Flores**, **Freddy Flores**, **José «Chapa» Fuenzalida**, **Fernando «Feña» García**, **Jorge Godoy**, **Baldomero «Baldo» Gómez**, **Hernán «Nano» Gómez**, **Fernanda «Feña» Guzmán**, **Magdalena «Mane» Hewstone**, **Catalina «Cata» Huerta**, **Lucas «Jefe» Ibarra**, **Agustín «Chivo» Laborde**, **Sergio Leiva**, **Sofía León**, **Alonso Letelier**, **Maximiliano «Max» Lira**, **Vicente «Mortero» Lobo**, **Joaquín «Joaco» Lozano**, **Rodrigo «Loyan» Lozano**, **Benjamín «Benja» Lustig**, **Sebastián «Coloro» Margozzini**, **Florencia «Flo» Miranda**, **Martín Miranda**, **Matías «Tito» Pablo**, **Eduardo «Pelao» Parker**, **Ignacio Pérez**, **Ignacio Pino**, **Pedro Pablo «Pepe» Pinto**, **Alejandro «Ale» Poblete**, **José Tomás «Pipa» Puelma**, **Lukas «Pugato» Puga**, **Diego Ramírez**, **Andrés «Papitas» Riveros**, **Rolando «Rola» Rogers**, **Santiago Rojas**, **Juan Pedro Ross**, **Benjamín «Benja» Sánchez**, **Fabián Sepúlveda**, **José Miguel «Torro» Torrico**, **Diego Valdivieso** y **Francisco «Pancho» Venegas**.

Pour ma première année, je m'installais au 45 Boulevard Diderot, dans la glorieuse résidence Crous Cîteaux, dans de modiques 14 m^2 qui ne m'ont pas empêché de profiter de ma première année à Paris (ni d'y faire d'épiques soirées). Je remercie les amis que j'ai rencontrés là-bas, surtout grâce au fait que nous étions tous des boursiers FSMP. Merci, **Hedong «Ben» Hou**, **Siiri Kivimäki**, **Francesca Rizzo** et **Katarzyna «Kasia» Szczerba**.

Me gustaría hacer un pequeño alto para agradecer a dos personas que fueron fundamentales en mi llegada a Francia, dos viejos amigos chilenos que me apañaron y ayudaron desde mis primeros días en Francia. Gracias, **Nicolás Bitar**, por tu amistad que se remonta al primer año de universidad en Chile, donde ni nos imaginábamos que íbamos a estudiar matemáticas en París. Gracias por tu constante ayuda y por tus mezclas de whisky con cerveza a las 3 de la mañana. Gracias, **Sebastián «Berto» Hurtado**, también por tu amistad que se remonta a muchos años atrás, a los tiempos donde chocábamos camiones en trabajos de invierno. Gracias por tu eterna buena voluntad y generosidad.

En septembre, je commençais mon M2 à Sorbonne Université le master de mathématiques, mention apprentissage et algorithmes, car après avoir terminé mes études en mathématiques au Chili, j'avais envie de m'orienter thématiquement vers un domaine un peu différent : celui du machine learning. Je remercie **Maxime Sangnier**, qui dans son rôle de co-directeur du master et tuteur FSMP, m'a toujours aidé, encouragé et orienté. J'en profite également pour remercier les professeurs **Eddie Aamari**, **Anna Ben-Hamou** et **Gérard Biau** pour leurs cours qui m'ont beaucoup inspiré et pour leur bienveillance envers moi. Je remercie aussi mes collègues du master, qui dès le début ont été très accueillants et patients avec moi et mon français débutant. Merci pour les sorties au bar et pour le soutien collectif dans les différents cours. Plus particulièrement, je remercie **Paul Liautaud** et **Patrick Lutz** pour leur amitié qui perdure jusqu'à aujourd'hui (soit que j'aille à Jussieu, soit que j'aille à Boston).

Malheureusement, dès la première semaine des cours, je me suis rendu compte que je n'aimais pas le sujet du machine learning et que je voulais retourner aux mathématiques pures, et plus précisément, aux probabilités. Je ne savais pas quoi faire, mais au moment, je savais qu'il y avait une personne qui pouvait m'aider : **Joaquín Fontbona**. Joaquín, aprovecho este espacio para agradecerte: en el curso de probabilidades me enseñaste qué es una medida de probabilidad y en cálculo estocástico me enseñaste qué es el movimiento browniano; fue gracias a esto que me enamoré de las probabilidades y decidí dedicarme a ellas como matemático. Más allá de eso, siempre has estado atento a mis pasos desde que terminé mis estudios en Chile y siempre me has aconsejado y motivado. Para mí es un gran honor que hoy en día incluso estemos colaborando en un proyecto. Gracias por tu humanidad, simpatía y excelente gusto musical (¡repitamos pronto el karaoke!).

Quiero también agradecer a todos los profesores que tuve en Chile. Sin lugar a dudas, es también gracias a la excelente formación que recibí en la Universidad de Chile que pude proseguir con mis estudios en París. En particular, quiero agradecer a los profesores **Hugo F. Arellano**, **Roberto Cortez**, **Aris Daniilidis**, **Raúl Gormaz**, **Alejandro Maass**, **Martín Matamala**, **Gonzalo A. Palma**, **Daniel Remenik**, **Jaime San Martín**, **Jorge San Martín** y **Raimundo Undurraga** por sus inspiradores cursos. Nunca fue mi profesora, mas agradezco también a **Natacha Astromujoff** por su simpatía, preocupación y por todos los cigarros que hemos fumado, incluso cada vez que he vuelto a mi querido Beauchef en estos últimos años. Me gustaría también agradecer a todos los miembros de la comunidad de probabilidades en Chile, a la que he tenido la suerte de irme integrando más durante la preparación de esta tesis, sobre todo en las escuelas de verano en Chile. En particular, y aparte de quienes ya he mencionado arriba, gracias, **Mauricio Duarte**, **Héctor Olivero**, **Santiago Saglietti** y **Avelio Sepúlveda**.

Après avoir discuté avec Joaquín des possibilités qui s'offraient à moi après ne pas vouloir continuer cette brève excursion dans le monde des applications, une option a retenu mon attention : le transport optimal. Cette idée semblait intéressante, car la théorie du transport optimal était belle et permettait d'établir de jolis liens entre différents domaines. Je commençais alors à chercher des professeurs pour encadrer mon mémoire de M2 et peut-être une thèse. C'est ainsi que je suis tombé sur le nom de **Max Fathi** à qui j'ai demandé si je pouvais assister à son cours sur le transport optimal et s'il serait disponible pour encadrer mon mémoire de M2 et une éventuelle thèse. La page de garde de cette thèse fait le spoiler : il a dit oui, et c'est comme ça que c'est parti.

Max, je tiens à vous remercier de m'avoir encadré ces années ; très concrètement, sans votre soutien et votre direction, cette thèse ne se serait pas achevée. D'abord, vous avez toujours su répondre à mes (très floues) questions, en donnant toujours de bonnes pistes et des références très précises. C'est grâce à vous que j'ai pu approfondir ma culture mathématique ces années, et je vous remercie d'avoir toujours poussé en

avant ma curiosité mathématique ; je mesure la chance que j'ai eue de travailler sous votre supervision. Au-delà des mathématiques, merci d'avoir toujours été très soucieux de ma formation et de mon avenir, merci pour toutes les précieuses opportunités que vous m'avez données tout au long de ma thèse. Merci d'avoir été très patient, gentil, généreux et sympathique avec moi. Merci de m'avoir appris des détails de la langue française comme la différence entre « courbure » et « courbature » ou le fait que l'on dit plutôt « un Coca-Cola » et non « une Coca-Cola ». Merci pour la recette de kimchi, je vous dirai comment c'était quand j'aurai essayé.

Au deuxième semestre, quand l'hiver et les mesures anti-Covid s'assouplissaient, j'ai fait de manière progressive la rencontre d'un groupe d'hispanophones à Jussieu avec lesquels j'allais passer le crépuscule de ma jeunesse, los templarios: **Martín Azón**, **Bruno Gálvez**, **Rodrigo «Rodri» Íñigo**, **Pau Maestre** y **Carlos «Charly» Rodríguez**. Gracias por la bella amistad que me han regalado y que perdura hasta hoy en día (no saben la tremenda alegría que me da que vengan a mi defensa), por todos los momentos vividos en aquel mítico semestre, y en particular, por ese verano de 2022, mi último verano de juventud, que tuvo como soundtrack el disco *Un verano sin ti* de Bad Bunny. Yo pensaba que la pandemia me había arrebatado definitivamente la juventud, mas no: ustedes le dieron un último suspiro. Nuestro hábitat natural era Jussieu: cómo olvidar esas sesiones de estudio en la MIR y sus respectivas pausas para ir a comprar el segundo almuerzo de Martín y echarnos un poco de nicotina encima, o esos almuerzos con larga sobremesa en el Crous. Dicho esto, otro lugar natural era el mítico *Salsero*, al cual siempre nos costaba entrar, pero si no lo lográbamos, éramos felices yendo afuera del mítico templo[3] para hacer homenaje a la Santísima Trinidad: *Guillermo*, *Pepe* y *Ricardo*. De manera más personal, gracias, Martín, por todas esas buenas videollamadas y conversaciones sobre la vida y el futuro. Gracias por siempre considerarnos e invitarnos a cada plan que organizabas, por presentarnos a tus amigos, por presentarnos a don Guillermo y por tu extenso prontuario, que siempre nutrió nuestros chistes e historia colectiva templaria. Gracias, Bruno, por ser ese tipo de amigo con el que comparto infinitas cosas: valores, gustos, odios, música, basadería, etc. (aunque discrepemos cuando se trata del uso de cartas locales en una variedad). Gracias por siempre haberme bancado, por el tiempo en que fuimos vecinos y por los tópicos recurrentes que nutren nuestras conversaciones. Gracias, Rodri, por estar siempre presente y haber continuado la amistad, aunque hayas partido de Francia. Gracias por ser (a veces) la voz de la razón entre todos nosotros, por tu amor al transporte óptimo, por el épico viaje a Napoli y por los Uber (mas no te pienso agradecer jamás por la ida a *Vin et Whisky*). Gracias, Pau, por presentarnos a don Pepe, por los puritos, por tu motivación y por el concepto de chavalería. Gracias, Charly, por tu eterna alegría, motivación e inagotable energía durante el M2 que siempre nos condujeron al Salsero. Gracias por tu devoción por el reggaeton, por presentarme a don Ricardo y por ser el único no chileno que no me da cringe que use nuestras expresiones. También quiero agradecer a varias personas que nos rodearon en esos tiempos y después; gracias, **Yvon Bossut**, **Rania Bouhaouita Haddad**, **Mar Derqui**, **Luis Ivorra**, **Paul Martin**, **Francesca Pratali** e **Irlanda Rodríguez**. No fuiste parte de la época dorada del M2, mas fue gracias a este grupo que te conocí: gracias, **Mercè «Merche»**

[3] C'est l'église Saint-Julien-le-Pauvre, qui se trouve au 5^{e} arrondissement.

Sellarès, por tu eterna disposición a escucharme y aconsejarme, por entender también lo que es tener piedras en los riñones, por esas idas espontáneas al Nouvel, por hacerme sentir joven y por haberme llevado a comer las mejores bravas de Barcelona.

Après le M2, j'ai commencé ma thèse à l'université Paris Cité, dans le 13e arrondissement. Dans le roman *Extension du domaine de la lutte*, on trouve une description très précise du quartier :

> « Nous travaillons dans un quartier complètement dévasté, évoquant vaguement la surface lunaire. C'est quelque part dans le treizième arrondissement. Quand on arrive en bus, on se croirait vraiment au sortir d'une troisième guerre mondiale. Pas du tout, c'est juste un plan d'urbanisme.
>
> Nos fenêtres donnent sur un terrain vague, pratiquement à perte de vue, boueux, hérissé de palissades. Quelques carcasses d'immeubles. Des grues immobiles. L'ambiance est calme et froide. »

Oui, le quartier est moche, mais les gens que j'y ai rencontrés sont formidables. Je tiens tout d'abord à remercier les permanents du 5e étage pour leur bienveillance lors de parler de mathématiques ou de la vie, leurs conseils, leur sympathie et leur encouragement. Plus particulièrement, merci, **Yves Achdou**, **Yves Capdeboscq**, **Sylvain Delattre**, **Romain Ducasse**, **Bastien Fernandez**, **Michael Goldman**, **Raphaël Lefevere** (c'était un vrai plaisir assurer les travaux dirigés de tes cours et je te remercie pour la lettre de recommandation que tu as écrite !), **Céline Lévy-Leduc**, **Mathieu Merle** et **Éric Vernier**. Dans toute l'hostile bureaucratie administrative, j'ai toujours eu l'efficace et patiente aide de **Nathalie Bergame** et **Amina Hariti**, à qui j'en remercie beaucoup. Je tiens à remercier mes collègues stagiaires, doctorants, ATERs et postdocs, avec lesquels j'ai eu le plaisir de partager beaucoup de moments lors de la préparation de ma thèse. Ces presque quatre années, je me suis toujours rendu avec grand plaisir au 5e étage du bâtiment Sophie Germain ; merci pour les longues pauses déjeuner, pour les parties de Pédantix et Cémantix et pour les discussions animées sur de nombreux sujets (l'origine des sauces dans la gastronomie, les conséquences de manger le dessert avant les fruits, etc.). Merci, **Aziz Ben Nejma**, **Nathan De Carvalho**, **Orphée Collin**, **Ali Ellouze**, **Dounia Essaket**, **Pierre Faugère**, **Eleanor Gemida**, **Marina Gomtsyan**, **Alexis Houssard**, **Lucas Ketels**, **Lamia Lamrani**, **Łukasz Mądry**, **Hoang-Dung Nguyen**, **Yihao Pang**, **Justin Ruelland** et **Arthur Stéphanovitch**. Mais, de manière spéciale, je voudrais remercier mes amis les plus intimes dans le 5e étage, mes chers « bâtards ». À chaque fois que je n'allais pas bien et que je ne voulais pas continuer, ou quand il y avait une petite chose à fêter, vous étiez toujours là pour moi. Merci, **Archit Chaturvedi**, d'avoir été mon confident au labo, d'avoir écouté mes (toujours infructueuses) histoires et de m'avoir toujours donné ton précieux avis ; merci pour ton sens de l'humour et pour ton exceptionnel goût littéraire. Merci, **Alessandro « Ale » Cosenza**, pour tous les moments où tu m'as fait beaucoup rire (« aïe, Alessandro ! »), d'avoir été le meilleur collègue de bureau, d'avoir toujours bien toléré toutes mes blagues, de m'avoir rappelé le rasoir d'Ockham au moment précis et merci pour toutes les fois où tu m'as sauvé d'ouvrir une boîte de sauce pesto en m'invitant manger chez toi. Merci, **Anna De Crescenzo**, d'être l'une des personnes qui m'ont le plus encouragé le plus dès le début de la thèse (en croyant en mes capacités plus que moi-même), surtout dans les moments où j'avais perdu l'espoir. Merci d'avoir toujours ri de mes blagues et bêtises (et même de les avoir amplifiées), pour les apéros sur la Seine et pour la jolie tasse que tu m'as offerte. Merci, **Léo Daures**, d'avoir été mon premier ami français intime, de t'être donné le temps et la patience de nous expliquer les références et mèmes français, pour ton humanité, de m'avoir présenté la chartreuse, d'avoir été le greffier officiel lors du Pédantix et pour toutes les fois où l'on a dit qu'il aurait été magnifique d'être à Bercy pour *Alive 2007*. Merci, **Maxime Guellil**, d'être mon meilleur « collègue de pot », d'avoir toujours montré de l'intérêt pour mes mathématiques (alors qu'elles ne servent à rien) et de m'avoir fait confiance pour me demander mon avis lorsque tu en avais besoin. Merci, **Ons Rameh**, d'avoir été une magnifique sœur de thèse, pour nos discussions sur la vie (surtout à Rennes) et de m'avoir toujours encouragé lorsque je me sentais nul. Tu ne fais pas partie du 5e étage, mais c'est grâce à Ale et Anna que je t'ai rencontré, **Florin Suciu** : merci de m'avoir toujours accompagné fumer dehors, pour les épiques soirées que l'on a passées ensemble et d'avoir toléré mes cris quand je supportais Nicolás Jarry quand on l'a vu jouer contre Carlos Alcaraz au tournoi de Bercy.

Je n'oublie pas mes collègues doctorants de Jussieu ; merci, **Zhe Chen**, **Nicolaï Gouraud**, **Ruikang Liang**, **Cristóbal Loyola**, **Robin Roussel** et **Fabrice Serret**. Je remercie particulièrement aux « coolest of LJLL », **Federica Padovano**, **Lucia Tessarolo** et **Aleksandra Tomaszek**, pour tous les dîners et sorties où l'on a toujours beaucoup ri (malgré le fait que vous n'aimiez pas la langue de Molière).

Vers la fin de ma première année de thèse, Max m'a invité à une école d'été au SLMath Institute de Berkeley, en Californie, où il allait être l'un des enseignants. Elle portait sur la concentration de la mesure, les inégalités fonctionnelles et les techniques de localisation ; c'est-à-dire, une très bonne opportunité pour consolider mes connaissances dans ces domaines, mais pas seulement, car j'y ai aussi pu rencontrer des personnes formidables. I would like to thank **Dan Mikulincer** for being a great teacher at the summer school and for all our subsequent discussions. Thank you for your insightful questions, your kind words, and for paying such close attention to my work. I also thank **Arianna Piana** and **Shay Sadovsky** for being excellent teaching assistants and for all the discussions we had. Finally, I thank all the other PhD students I met in those two amazing weeks, where we could discuss mathematics, hang out, and visit the Bay Area. Thank you, **Michael Albert**, **Shrey Aryan** (special thanks to you for our collaboration with Yair, reflected in Chapter 5, that I very much enjoyed!), **Sabyasachi Basu**, **Ratul Biswas**, **Shabarish Chenakkod**, **Lorenz Frühwirth**, **Aldo García Guinto**, **Marcel Hudiani**, **Zhen-Chuan Liu**, **Uriel Martínez León**, **Jacob McErlean**, and **Pegah Mohammadipour** (thank you for staying in touch after the summer school; next time, I will not miss the opportunity to go to the sake place!).

Un autre événement marquant dans ma thèse a été ma visite de deux mois à l'université de Brown aux États-Unis pour travailler avec **Yair Shenfeld** vers la fin de la deuxième année. Yair, without a doubt, my visit to Brown was fundamental to completing my PhD. Beyond the fact that Chapters 4 and 5 are based on our collaborations, those two months gave me fresh air at a time when I was questioning myself whether if it was worthwhile continuing with my PhD. We did beautiful mathematics

together, and it was a real pleasure working and hanging out with you. I will never forget all the help you gave me when I was applying for postdoc positions, helping me polishing my application and writing a recommendation letter for me. Thank you for the times you invited me to dinner; I really appreciated your knowledge of the best Asian restaurants in Providence! I would also like to thank **Ramon van Handel**, **Kevin Hu**, **Oanh Nguyen**, and **Kavita Ramanan**, all of whom I had the pleasure of meeting. Gracias, **Molly Wagschal**, por todo: en particular, por hacer que el tiempo que pasé en Providence fuese muy ameno, por presentarme a tus amigos y por tus mac and cheese.

Au long de ces années, j'ai eu la chance de rencontrer des chercheurs dans le cadre de différentes opportunités, qui ont été très bienveillants, patients et sympathiques avec moi lors de nos échanges. De manière particulière, merci, **Aymeric Baradat** (merci beaucoup de m'avoir invité à Lyon !), **Alessandra Bianchi**, **Léonard Cadilhac**, **Giovanni Conforti** (merci beaucoup de m'avoir invité à Padoue !), **Thomas Courtade Alexandros Eskenazis**, **Matthieu Fradelizi**, **Ivan Gentil** (je te remercie de m'avoir fait réfléchir pour la première fois au cours de ma thèse sur les aspects éthiques de la recherche en mathématiques), **Ronan Herry** (merci beaucoup de m'avoir invité à Rennes !), **Flavien Léger**, **Joseph Lehec**, **Pierre Monmarché**, **Loucas Pillaud-Vivien**, **Cyril Roberto**, **Paul-Marie Samson** et **Nikita Simonov** (¡muchas gracias por la invitación a Granada!).

Ces années m'ont aussi permis de rencontrer de nombreux collègues doctorants et jeunes chercheurs lors de plusieurs conférences et rencontres, et d'échanger avec eux autour des mathématiques (ou pas nécessairement). Merci, **Nicolás Agote**, **Pablo Araya**, **Esther Bou-Dagher**, **Katharina Eichinger**, **Felipe Espinosa**, **Marta Gentiloni Silveri**, **David Heredia**, **Dylan Langharst**, **Hugo Malamut**, **Francisco Marín Sola**, **Richard Medina**, **Saliou Ndiaye**, **Giacomo Passuello**, **Jordan Serres** et **Maxime Sylvestre**.

Quiero agradecer a mis amigos compatriotas que me han acompañado en esta aventura acá en París, que sin lugar a dudas han hecho muy amena esta travesía, gracias, **Andrés Contreras**, **Felipe Gambardella**, **Felipe Garrido**, **Moira Mac Auliffe**, **Martín Rapaport**, **Pablo Uribe** y **Nicolás Zalduendo**. De manera más particular, quiero agradecer a dos chilenos con los que he tenido el placer de compartir bastante en la última parte de mi doctorado. Gracias, **Arie Wortsman**, por haber sido una suerte de psicólogo para mí, escuchando mis problemas, siempre llamando a la calma y a la cordura cuando las cosas no andaban bien y quería tirar todo (y a todos) por la borda. Gracias por llevarle a Alejandra mis libros y por haberme mostrado a Roberto Bolaño, la babka y *Culture Rapide*. Gracias, **Pablo Zúñiga**, por ser apañador, por tu ecléctico sentido del humor, por el viaje a Normandía y Bretaña (donde conocimos a **Ricardo Félix**, a quien agradezco por su amistad), por tu excelente gusto musical del cual compartimos bastante y por recordarme la existencia de *La Brígida Orquesta*.

Je suis honoré que **Martin Huesmann** et **Christian Léonard** aient accepté de rapporter ma thèse ; je voudrais les remercier pour l'attentive lecture qu'ils ont fait de mon manuscrit et pour leurs rapports très détaillés. De manière particulière,

merci, Christian, pour tout : j'ai vraiment beaucoup apprécié nos discussions. Notre rencontre a marqué un moment important : tu m'as encouragé quand je n'avais pas trop d'espoir pour l'avenir et je sentais que ma thèse ne servait à rien. Je te remercie aussi pour la jolie lettre de recommandation que tu as écrite. Je tiens également à remercier aussi **Nathaël Gozlan**, **Hélène Halconruy** et **Maxime Laborde** d'avoir accepté de faire partie de mon jury, ce qui me fait très plaisir. En particulier, merci, Maxime, d'avoir été présent tout au long de la préparation de cette thèse. Ton aide, ta sympathie, tes encouragements et les nombreuses discussions que nous avons eues ont été fondamentales durant toutes ces années.

Cette histoire aura une belle continuation : je partirai à Los Angeles, en Californie, pour faire mon postdoc avec **Georg Menz** à l'UCLA. Thank you, Georg, for your kindness from the very beginning, for believing in me, and for supporting my application. I am quite sure that I will enjoy working with you over the next three years. I look forward to learning many things from you and applying what I have already learned.

Enfin, je crois que c'est tout ce que je voulais dire. Désolé si c'était trop long, mais c'est quelque chose que je voulais faire, car ces presque quatre années en France m'ont appris l'importance des autres personnes dans ma vie ; sans toutes celles qui sont mentionnées dans ce texte, cela n'aurait pas été possible. *¡Gracias totales!*

Notations and conventions

La Nature est un temple où de vivants piliers
Laissent parfois sortir de confuses paroles ;
L'homme y passe à travers des forêts de symboles
Qui l'observent avec des regards familiers.

CHARLES BAUDELAIRE
Correspondances

Sets

If $\mathcal{X}$ is a set, we denote the identity function on $\mathcal{X}$ by $\mathrm{id}_{\mathcal{X}} \colon \mathcal{X} \to \mathcal{X}$. The cardinal of $\mathcal{X}$ is written as $\mathrm{card}(\mathcal{X})$. For any subset $A \subset \mathcal{X}$, the function $\mathbb{1}_A \colon \mathcal{X} \to \{0, 1\}$ represents the indicator function of A; that is, for every $x \in \mathcal{X}$, $\mathbb{1}_A(x) = 1$ if and only if $x \in A$. In particular, we define $\mathbb{1} \coloneqq \mathbb{1}_{\mathcal{X}}$, meaning that $\mathbb{1}$ is the constant function equal to 1 on the set $\mathcal{X}$.

We denote the natural numbers by $\mathbb{N} \coloneqq \{0, 1, 2, \dots\}$, the positive natural numbers by $\mathbb{N}^* \coloneqq \mathbb{N} \setminus \{0\}$, the integers by $\mathbb{Z}$, and the real numbers by $\mathbb{R}$. We use $\mathbb{R}_+$ or $\mathbb{R}_{\geqslant 0}$ to denote the nonnegative real numbers, and $\mathbb{R}_{>0}$ for the positive real numbers.

Given two sets $\mathcal{X}$ and $\mathcal{Y}$, we denote their Cartesian product by $\mathcal{X} \times \mathcal{Y}$; $\mathrm{proj}_{\mathcal{X}} \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ and $\mathrm{proj}_{\mathcal{Y}} \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Y}$ are the associated canonical projections. For real-valued functions $f \colon \mathcal{X} \to \mathbb{R}$ and $g \colon \mathcal{Y} \to \mathbb{R}$, we define their sum $f \oplus g \colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ as follows:

$$\forall (x, y) \in \mathcal{X} \times \mathcal{Y}, \quad (f \oplus g)(x, y) \coloneqq f(x) + g(y).$$

Euclidean structure and matrices

For any $d \in \mathbb{N}^*$, and any $x \in \mathbb{R}^d$, we write $x = (x_1, \dots, x_d)$. For $x, y \in \mathbb{R}^d$, we denote the Euclidean inner product between x and y by $x \cdot y \coloneqq \sum_{i=1}^d x_i y_i$, and the Euclidean norm of x by $|x| \coloneqq \sqrt{x \cdot x}$. When we refer to $\mathbb{R}^d$, it is immediately assumed that $d \in \mathbb{N}^*$ unless otherwise specified.

For any $d \in \mathbb{N}^*$, we denote the set of square matrices with real entries of dimension $d \times d$ by $\mathcal{M}_d(\mathbb{R})$. I_d is the identity matrix of dimension $d \times d$. For $M \in \mathcal{M}_d(\mathbb{R})$, we define its Hilbert-Schmidt as $|M|_{\mathrm{HS}} \coloneqq \left(\sum_{i,j=1}^d |M_{ij}|^2\right)^{1/2}$, and its operator norm as $|M|_{\mathrm{op}} \coloneqq \sup_{|x|=1} |Mx|$. We denote the trace of M by $\operatorname{tr}(M)$ and its determinant by $\det(M)$. We write the transpose of M as $M^\intercal$, and we say that M is symmetric if $M = M^\intercal$.

Let $d \in \mathbb{N}^*$, and let $M \in \mathcal{M}_d(\mathbb{R})$ be a symmetric matrix. If for every $x \in \mathbb{R}^d$, $x^\intercal M x \geqslant 0$, we say that M is positive semidefinite, and write $M \succcurlyeq 0$; if for every $x \in \mathbb{R}^d \setminus \{0\}$, $x^\intercal M x > 0$, we say that M is positive definite, and write $M \succ 0$. For two symmetric matrices $M, N \in \mathcal{M}_d(\mathbb{R})$, we say that $M \succcurlyeq N$ if $(M - N) \succcurlyeq 0$. If M is positive semidefinite, we denote its square root by $M^{\frac{1}{2}}$, which is the only symmetric matrix such that $M^{\frac{1}{2}} M^{\frac{1}{2}} = M$.

Metric spaces

Let $(\mathcal{X}, d)$ be a metric space, and let $A \subset \mathcal{X}$. We denote the interior, the closure, and the boundary of A by $\operatorname{int}(A)$, $\bar{A}$, and ∂A, respectively. For any $x \in \mathcal{X}$ and $r > 0$, we write the closed ball with center x and radius r as $\overline{\mathrm{B}}_{\mathcal{X}}(x, r)$ or just $\overline{\mathrm{B}}(x, r)$ if the base space is clear from the context. Similarly, we denote the open ball with center x and radius r by $\mathrm{B}_{\mathcal{X}}(x, r)$ or $\mathrm{B}(x, r)$.

If both $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$ are metric spaces, we say that a map $T \colon \mathcal{X} \to \mathcal{Y}$ is Lipschitz if there exists a constant $L > 0$ such that

$$\forall x_1, x_2 \in \mathcal{X}, \quad d_{\mathcal{Y}}(T(x_1), T(x_2)) \leqslant L \, d_{\mathcal{X}}(x_1, x_2).$$

In the special case when $L \leqslant 1$, we say that T is a contraction.

Convexity, continuity, and differentiability

Let $\mathcal{X}$ be a topological vector space, denote its topological dual by $\mathcal{X}^*$, and the duality pairing between $\mathcal{X}^*$ and $\mathcal{X}$ by $\langle \cdot, \cdot \rangle$. A subset $A \subset \mathcal{X}$ is said to be convex if

$$\forall t \in [0,1], \forall x, y \in A, \quad (tx + (1-t)y) \in A.$$

For a function $f \colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, we define its domain as the set

$$\operatorname{Dom}(f) \coloneqq \{x \in \mathcal{X} : f(x) < +\infty\},$$

and we say that f is proper if $\mathrm{Dom}(f) \neq \varnothing$. A function $f\colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is convex if $\mathrm{Dom}(f)$ is a convex set and

$$\forall t \in [0,1], \forall x, y \in \mathrm{Dom}(f), \quad f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y).$$

For any $A \subset \mathcal{X}$, we define its characteristic function $\chi_A\colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ as

$$\forall x \in \mathcal{X}, \quad \chi_A(x) \coloneqq \begin{cases} 0, & \text{if } x \in A \\ +\infty, & \text{otherwise.} \end{cases}$$

For $f\colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper, its convex conjugate or Legendre transform is the function $f^*\colon \mathcal{X}^* \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\forall x^* \in \mathcal{X}^* \quad f^*(x^*) \coloneqq \sup_{x \in \mathcal{X}} \{\langle x^*, x\rangle - f(x)\}.$$

If $A \subset \mathcal{X}$ and $f\colon A \to \mathbb{R} \cup \{+\infty\}$ is a function, we define its extension $\tilde{f}\colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ as

$$\forall x \in \mathcal{X}, \quad \tilde{f}(x) \coloneqq \begin{cases} f(x), & \text{if } x \in A \\ +\infty, & \text{otherwise;} \end{cases}$$

we say that f is convex if $\tilde{f}$ is convex, and doing an abuse of notation, we define f^* as the function $(\tilde{f})^*$ restricted to the set A.

Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces. We denote the set of continuous functions between $\mathcal{X}$ and $\mathcal{Y}$ by $\mathcal{C}(\mathcal{X}, \mathcal{Y})$. If $\mathcal{Y} = \mathbb{R}$, we simply write $\mathcal{C}(\mathcal{X})$. We write $\mathcal{C}_{\mathrm{b}}(\mathcal{X})$ for the set of continuous and bounded real-valued functions defined on $\mathcal{X}$, which is a Banach space if we equip it with the uniform norm $\|\cdot\|_\infty$, and we denote the set of continuous and compactly supported real-valued functions by $\mathcal{C}_{\mathrm{c}}(\mathcal{X})$. If we have a sequence $(f_n)_{n\in\mathbb{N}}$ of continuous functions $f_n\colon \mathcal{X} \to \mathbb{R}$, we say that it converges uniformly on compact sets to a continuous function $f\colon \mathcal{X} \to \mathbb{R}$ if for each compact set $K \subset \mathcal{X}$, $\|f_n - f\|_{K,\infty} \coloneqq \sup_{x\in K}|f_n(x) - f(x)| \to 0$ as $n \to \infty$. If $f\colon \mathcal{X} \to \mathbb{R}^d$ or $f\colon \mathcal{X} \to \mathcal{M}_d(\mathbb{R})$, we can extend the definition of $\|\cdot\|_{K,\infty}$ by doing the following abuse of notation: $\|f\|_{K,\infty} \coloneqq \||f|\|_{K,\infty}$, where $|\cdot|$ denotes the Euclidean or Hilbert-Schmidt norm, depending on the codomain of f.

If $f\colon \mathbb{R}^d \to \mathbb{R}$ is differentiable, we denote its gradient at $x \in \mathbb{R}^d$ by $\nabla f(x) = (\partial_i f(x))_{i=1}^d \in \mathbb{R}^d$, where for $1 \leqslant i \leqslant d$, ∂_i is the i-th partial derivative; if $d = 1$, we denote its derivative by $f'(x)$. If $f\colon \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ is differentiable, we also denote its Jacobian or differential matrix at $x \in \mathbb{R}^{d_1}$ by $\nabla f(x)$. If $f\colon \mathbb{R}^d \to \mathbb{R}$ is twice differentiable, we denote its Hessian matrix at $x \in \mathbb{R}^d$ by $\nabla^2 f(x)$. If $F\colon \mathbb{R}^d \to \mathbb{R}^d$ is a differentiable vector field with $F = (F_1, \ldots, F_d)$, we define its divergence at $x \in \mathbb{R}^d$ by $\nabla \cdot F(x) \coloneqq \sum_{i=1}^d \frac{\partial F_i}{\partial x_i}(x)$. If $f\colon \mathbb{R}^d \to \mathbb{R}$ is twice differentiable, we define its Laplacian at $x \in \mathbb{R}^d$ by $\Delta f(x) \coloneqq \nabla \cdot (\nabla f)(x) = \mathrm{tr}(\nabla^2 f(x))$.

Measure theory on topological spaces

Let $\mathcal{X}$ be a topological space, and let $\mathcal{B}(\mathcal{X})$ be its Borel σ-algebra. A measure defined on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is said to be a Borel measure, and we say that a real-valued function defined on $\mathcal{X}$ is Borel if it is $\mathcal{B}(\mathcal{X})$-measurable. For any nonnegative Borel measure μ defined on $\mathcal{X}$, we define its support as the set

$$\operatorname{supp}(\mu) \coloneqq \{x \in \mathcal{X} : \text{every open neighborhood of } x \text{ has positive measure}\}.$$

For any nonnegative Borel measure μ on $\mathcal{X}$ and a suitable Borel function $f\colon \mathcal{X} \to \mathbb{R}$, we denote the integral of f with respect to μ by $\int_{\mathcal{X}} f \,\mathrm{d}\mu$, $\int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x)$, or $\langle \mu, f\rangle$.

Let $\mathcal{X}$ be a metrizable topological space. We denote by $\mathcal{M}(\mathcal{X})$ the set of finite signed Borel measures on $\mathcal{X}$, and we equip it with the weak topology, i.e., the one induced by the duality with $\mathcal{C}_{\mathrm{b}}(\mathcal{X})$: a sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}(\mathcal{X})$ converges weakly to $\mu \in \mathcal{M}(\mathcal{X})$ if

$$\forall f \in \mathcal{C}_{\mathrm{b}}(\mathcal{X}), \quad \langle \mu_n, f\rangle \xrightarrow[n\to\infty]{} \langle \mu, f\rangle.$$

We denote by $\mathcal{P}(\mathcal{X})$ the set of Borel probability measures on $\mathcal{X}$. We define $\mathbb{B}(\mathcal{X})$ as the class of bounded and measurable real-valued functions, which is a Banach space if we equip it with the uniform norm $\|\cdot\|_\infty$. If $x \in \mathcal{X}$, δ_x denotes the Dirac measure on x. If $\mathcal{X} = \mathbb{R}^d$, we denote the d-dimensional Lebesgue measure by Leb, $\mathrm{d}x$, or Vol_d.

Given two Borel probability measures μ and ν on a nonempty topological space $\mathcal{X}$, we define the total variation distance between them by

$$\|\mu - \nu\|_{\mathrm{TV}} \coloneqq 2\sup\{|\mu(A) - \nu(A)| : A \in \mathcal{B}(\mathcal{X})\}.$$

For any nonnegative Borel measure μ on a topological space $\mathcal{X}$, and any $p \in [1, +\infty]$, we will write $L^p(\mathcal{X}, \mu)$ or $L^p(\mu)$ for the classical p-Lebesgue space, which is a Banach space if we equip it with the p norm, $\|\cdot\|_{L^p(\mu)}$. If $\mathcal{X} = \mathbb{R}^d$ and $\mu = \mathrm{Leb}$, we write the associated p-norm as $\|\cdot\|_p$. For $f\colon \mathcal{X} \to \mathbb{R}^d$ or $f\colon \mathcal{X} \to \mathcal{M}_d(\mathbb{R})$ Borel measurable, we can extend the definition of $\|\cdot\|_p$ by doing the following abuse of notation: $\|f\|_p \coloneqq \||f|\|_p$, where $|\cdot|$ denotes the Euclidean or Hilbert-Schmidt norm, depending on the codomain of f.

We say that the triple $(\mathcal{X}, d, \mu)$ is a metric measure space if $(\mathcal{X}, d)$ is a metric space and μ is a nonnegative Borel measure on $\mathcal{X}$. In the particular case when $\mu \in \mathcal{P}(\mathcal{X})$, we say that $(\mathcal{X}, d, \mu)$ is a metric probability space. Let $(\mathcal{X}, d, \mu)$ be a metric measure space, and let $p \geqslant 1$. If there exists $x_0 \in \mathcal{X}$ such that the function $x \mapsto d(x_0, x)$ belongs to $L^p(\mathcal{X}, \mu)$, we say that μ has a finite moment of order p. We denote by $\mathcal{P}_p(\mathcal{X})$ the set of probability measures on $\mathcal{X}$ with a finite moment of order p.

For two Polish spaces $\mathcal{X}$ and $\mathcal{Y}$ (i.e., separable and completely metrizable topological spaces), let μ and ν be two nonnegative Borel measures defined on $\mathcal{X}$ and $\mathcal{Y}$, respectively. We write $\mu\otimes\nu$ for their product, which is defined on $(\mathcal{X}\times\mathcal{Y}, \mathcal{B}(\mathcal{X})\otimes\mathcal{B}(\mathcal{Y}))$. Note that since both $\mathcal{X}$ and $\mathcal{Y}$ are separable, then $\mathcal{B}(\mathcal{X} \times \mathcal{Y}) = \mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$.

Given two nonempty topological spaces $\mathcal{X}$ and $\mathcal{Y}$, a Borel probability measure μ on $\mathcal{X}$, and a Borel map $T\colon \mathcal{X} \to \mathcal{Y}$, we define the pushforward measure on $\mathcal{Y}$, denoted by $T_\#\mu$, $T_*\mu$, or $\mu \circ T^{-1}$, as

$$\forall B \in \mathcal{B}(\mathcal{Y}), \quad T_\#\mu(B) \coloneqq \mu(T^{-1}(B)).$$

Equivalently, it is characterized by

$$\forall f \in \mathbb{B}(\mathcal{Y}), \quad \int_{\mathcal{Y}} f(y)\,\mathrm{d}(T_{\#}\mu)(y) = \int_{\mathcal{X}} (f \circ T)(x)\,\mathrm{d}\mu(x).$$

Riemannian manifolds

Let (M, g) be a smooth Riemannian manifold of dimension $\dim(M) \in \mathbb{N}$. We denote its tangent bundle by TM. Usually, the action of the metric g on two vector fields $X, Y\colon M \to TM$ is succinctly written as $X \cdot Y \coloneqq g(X, Y)$, omitting the dependence on $x \in M$ unless necessary. The metric g induces the geodesic distance d_g on M, which is compatible with the topology on M and generates the Riemannian volume measure dVol on M, that is, the $\dim(M)$-dimensional Hausdorff measure associated to the metric space (M, d_g)). We denote by $\mathcal{C}^\infty(M)$ the set of smooth functions on M and by $\mathcal{C}^\infty_{\mathrm{c}}(M)$ the set of compactly supported smooth functions. We denote the Ricci curvature tensor by Ric.

Let (M, g) be a smooth Riemannian manifold, and let ∇ be the Riemannian gradient, which acts on smooth functions $f\colon M \to \mathbb{R}$ as a vector field by $\nabla f = (\nabla^i f)_{i=1}^d$, where $\nabla^i f = g^{ij}\partial_j f$, where we are using Einstein's summation convention to handle operations on tensors. We denote by ∇^2 the Hessian operator acting on smooth functions $f\colon M \to \mathbb{R}$ via $\nabla^2 f = (\nabla^i\nabla^j f)_{i,j=1}^d$. The symbol $\nabla\cdot$ denotes the divergence operator on (M, g); its action on a vector field $Z\colon M \to TM$ is characterized by

$$\forall \varphi \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \int_M \varphi\,\nabla\cdot Z\,\mathrm{dVol} = -\int_M \nabla\varphi \cdot Z\,\mathrm{dVol}.$$

The Laplace-Beltrami operator, denoted by Δ, acts on smooth functions $f\colon M \to \mathbb{R}$, and this action can be characterized by

$$\forall g \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \int_M g\Delta f\,\mathrm{dVol} = -\int_M \nabla f \cdot \nabla g\,\mathrm{dVol}.$$

Let (M, g) be a smooth Riemannian manifold, and let dVol be the associated volume measure. Let $W \in \mathcal{C}^\infty(M)$, and define $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$. Then we say that the triple (M, g, μ) is a weighted Riemannian manifold with weight measure μ.

Probability theory

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{X}$ be a Polish space. We say that a function $X\colon \Omega \to \mathcal{X}$ is a random variable if it is $\mathcal{B}(\mathcal{X})$-$\mathcal{F}$-measurable, and we define the law or distribution of X by $\mathrm{Law}(X) \coloneqq X_{\#}\mathbb{P}$; we write $X \sim \mu$ if X has law μ. Let $X\colon \Omega \to \mathcal{X}$ be a random variable with $\mu \coloneqq \mathrm{Law}(X)$, and let $f\colon \mathcal{X} \to \mathbb{R}$ be a Borel function. We set $\mathbb{E}[f(X)] \coloneqq \int_\Omega f \circ X\,\mathrm{d}\mathbb{P}$, or alternatively we can write $\mathbb{E}_{X\sim\mu}[f(X)]$, which of course is equivalent to $\langle \mu, f\rangle$. We define the variance of any $f \in L^2(\mu)$ as $\mathrm{Var}_\mu(f) \coloneqq \mathbb{E}_{X\sim\mu}[f(X)^2] - \mathbb{E}_{X\sim\mu}[f(X)]^2 = \langle \mu, f^2\rangle - \langle \mu, f\rangle^2$.

If μ is a Borel probability measure on $\mathbb{R}$, we define its distribution function $F_\mu\colon \mathbb{R} \to \mathbb{R}$ by

$$\forall x \in \mathbb{R}, \quad F_\mu(x) \coloneqq \mu((-\infty, x]).$$

For $d \in \mathbb{N}^*$, $a \in \mathbb{R}^d$ and $A \in \mathcal{M}_d(\mathbb{R})$ symmetric and positive semidefinite, we denote by $\mathcal{N}(a, A)$ the Gaussian distribution of mean a and covariance matrix A. If $a = 0$ and $A = I_d$, we say it is the standard d-dimensional Gaussian measure and denote it by γ_d.

For $T > 0$, we denote by π_T the Poisson distribution of parameter T on $\mathbb{N}$.

Let $\mathcal{X}$ be a Polish space, and let μ be a nonnegative Borel measure on $\mathcal{X}$. For any $f\colon \mathcal{X} \to \mathbb{R}_+$, we define the relative entropy of f with respect to μ as

$$\mathrm{Ent}_\mu(f) \coloneqq \int_{\mathcal{X}} f \log f \,\mathrm{d}\mu - \left(\int_{\mathcal{X}} f \,\mathrm{d}\mu \right) \log \left(\int_{\mathcal{X}} f \,\mathrm{d}\mu \right),$$

where we use the convention $0 \log 0 \coloneqq 0$. Let $\nu \in \mathcal{P}(\mathcal{X})$. We define the relative entropy of ν with respect to μ by

$$\mathrm{H}(\nu|\mu) \coloneqq \begin{cases} \mathrm{Ent}_\mu\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right), & \text{if } \nu \ll \mu \\ +\infty, & \text{otherwise,} \end{cases}$$

where $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ is the Radon-Nikodým derivative of ν with respect to μ. If $\mathcal{X} = \mathbb{R}^d$ and $\mu = \mathrm{Leb}$, we write $\mathrm{H}(\nu) \coloneqq \mathrm{H}(\nu|\mu)$ and call it the differential entropy of ν.

Introduction (français)

Attention, mesdames et messieurs, dans un instant, ça va commencer
Nous vous demandons évidemment d'être indulgents
Le spectacle n'est pas bien rôdé, laissez-nous encore quelques années
Il ne pourrait que s'améliorer au fil du temps.

Michel Fugain et Le Big Bazar
Attention, Mesdames et Messieurs

Dès le début de ma thèse, en octobre 2022, ma recherche a porté sur la théorie des probabilités et ses liens avec l'analyse et la géométrie. Plus précisément, j'ai cherché à comprendre comment ces trois sujets dialoguaient par le biais des processus stochastiques, du transport optimal, des inégalités fonctionnelles et de leurs interactions. Les résultats originaux présentés dans ce manuscrit de thèse s'inscrivent dans ce cadre.

Évidemment, ce manuscrit de thèse a pour finalité de condenser et de présenter mes contributions originales dès le début de ma thèse ; cependant, j'ai quand même décidé de les placer dans un contexte plus large, en présentant des préliminaires essentiels à leur compréhension. Cette décision a été motivée par deux raisons. D'abord, je voulais rédiger un manuscrit autosuffisant qui ne présupposait que les fondamentaux de la théorie des probabilités, de l'analyse et de la géométrie. Deuxièmement, une raison plus personnelle a motivé cette décision : j'ai la ferme conviction que la rédaction, qui est l'une des plus précieuses capacités humaines, est une grande alliée pour comprendre, synthétiser et permettre une vision globale d'un sujet.

Dans cette introduction, je vais résumer de manière succinte la structure de ce manuscrit, en soulignant les éléments de chaque chapitre qui révèlent la trame de cette thèse. Le manuscrit est divisé en deux parties : la première contient les résultats préliminaires sur lesquels la deuxième partie est basée, et où je présente mes contributions originales depuis le début de ma thèse.

Préliminaires (Preliminaries)

La première partie de ce manuscrit vise à réviser les résultats fondamentaux des théories du transport optimal et des inégalités fonctionnelles qui concernent les résultats originaux de cette thèse dans deux chapitres différents.

Transport optimal (Optimal transport)

Dans le chapitre 1 nous allons réviser les principes fondamentaux de la théorie du transport optimal en commençant par ses racines historiques et la théorie générale, où nous allons définir les formulations de Monge et Kantorovich du problème de transport optimal de mesures. Enfin, nous allons nous concentrer sur le cadre quadratique euclidien, où le problème, écrit sous sa formulation de Kantorovich, pour μ et ν, deux mesures de probabilité boréliennes définies sur $\mathbb{R}^d$, vise à calculer

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x-y|^2 \, \mathrm{d}\pi(x,y), \tag{0.0.1}$$

où $\Pi(\mu,\nu)$ désigne l'ensemble des plans de transport entre μ et ν.

Dans le cadre quadratique euclidien le théorème de Brenier-McCann révèle la riche structure des solutions au problème (0.0.1) qui, à son tour, coïncident avec les solutions du problème de Monge associé.

Théorème 0.1 (Brenier-McCann)**.** *Soient μ et ν deux mesures de probabilité boréliennes sur $\mathbb{R}^d$ et supposons que μ est absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Alors il existe une fonction $\varphi_0\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ propre semi-continue convexe telle que ν soit la mesure image de μ par l'application $T_0 \coloneqq \nabla\varphi_0$. De plus, T_0 est la seule application étant le gradient d'une fonction convexe envoyant μ sur ν.*

Supposons que μ et ν admettent des moments d'ordre 2. Alors $\mathcal{C}_0(\mu,\nu) < +\infty$ et il existe un unique plan de transport $\pi_0 \in \Pi(\mu,\nu)$ optimal pour le problème (0.0.1) *donné le plan de transport induit par T_0 : $\pi_0 = (\mathrm{id}_{\mathbb{R}^d}, T_0)_{\#}\mu$.*

De plus, si $\psi_0 \coloneqq \varphi_0^*$ désigne la conjuguée de Legendre de φ_0, le couple $(f_0, g_0) \coloneqq \left(\frac{1}{2}|\cdot|^2 - \varphi_0, \frac{1}{2}|\cdot|^2 - \psi_0\right)$ est une solution du problème dual à (0.0.1)*, qui est donné par*

$$\sup_{(f,g)\in\Phi_2} \int_{\mathbb{R}^d} f \, \mathrm{d}\mu + \int_{\mathbb{R}^d} g \, \mathrm{d}\nu, \tag{0.0.2}$$

où Φ_2 désigne l'ensemble

$$\Phi_2 \coloneqq \left\{ (f,g) \in L^1(\mu) \times L^1(\nu) : f \oplus g \leqslant \frac{1}{2}|\cdot - \cdot|^2 \right\}.$$

Nous allons ensuite nous concentrer sur la théorie autour de la régularité de l'application de Brenier, car elle joue un rôle très important dans la théorie des inégalités fonctionnelles, et surtout dans cette thèse. Plus concrètement, pour illustrer cette situation, soit μ une mesure de probabilité satisfaisant une inégalité de Sobolev logarithmique. Alors, pour chaque application lipschitzienne T, la mesure image associée $T_{\#}\mu$ vérifie aussi une inégalité de Sobolev logarithmique. À cet égard, le théorème de contraction de Caffarelli garantit la régularité lipschitzienne de l'application de Brenier envoyant la mesure gaussienne sur une perturbation log-concave quelconque.

Théorème 0.2 (Caffarelli). *Soit γ_d la mesure gaussienne sur $\mathbb{R}^d$, soit $\nu \in \mathcal{P}(\mathbb{R}^d)$ absolument continue par rapport à γ_d et supposons que ν a la forme $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\gamma_d$, avec $V\colon \mathbb{R}^d \to \mathbb{R}$. Si V est convexe alors l'application de Brenier envoyant γ_d sur ν est 1-lipschitzienne.*

Inégalités fonctionnelles (Functional inequalities)

Dans le chapitre 2, nous ferons un panorama de la théorie des inégalités fonctionnelles. Le point de départ sera les inégalités géométriques et leurs versions fonctionnelles, où l'un des exemples les plus significatifs de cette dualité est constitué par les inégalités de Brunn-Minkowski et de Prékopa-Leindler, la dernière étant une version fonctionnelle de la première, qui est une inégalité géométrique.

Théorème 0.3 (Brunn-Minkowski). *Soient $A, B \subset \mathbb{R}^d$ deux corps convexes. Alors*

$$\mathrm{Vol}_d(A+B)^{1/d} \geqslant \mathrm{Vol}_d(A)^{1/d} + \mathrm{Vol}_d(B)^{1/d}. \tag{0.0.3}$$

Théorème 0.4 (Prékopa-Leindler). *Soient $f, g, h\colon \mathbb{R}^d \to \mathbb{R}_+$ trois fonctions boréliennes positives telles que f et g soient Lebesgue-intégrables et soit $\lambda \in (0,1)$. Supposons*

$$\forall x, y \in \mathbb{R}^d, \quad h(\lambda x + (1-\lambda)y) \geqslant f(x)^\lambda g(y)^{1-\lambda}. \tag{0.0.4}$$

Alors

$$\int_{\mathbb{R}^d} h\,\mathrm{d}x \geqslant \left(\int_{\mathbb{R}^d} f\,\mathrm{d}x\right)^\lambda \left(\int_{\mathbb{R}^d} g\,\mathrm{d}x\right)^{1-\lambda}. \tag{0.0.5}$$

Un corollaire remarquable du théorème de Brunn-Minkowski est l'inégalité isopérimétrique euclidienne : les boules sont, à volume fixé, les corps convexes ayant le plus petit périmètre.

Théorème 0.5 (Inégalité isopérimétrique euclidienne). *Pour tout corps convexe $K \subset \mathbb{R}^d$,*

$$\mathrm{Vol}_{d-1}(\partial K) \geqslant d\,\mathrm{Vol}_d(K)^{(d-1)/d}\,\mathrm{Vol}_d(B)^{1/d}, \tag{0.0.6}$$

où $B \coloneqq \overline{\mathrm{B}}(0,1)$.

Le théorème 0.5 nous mène à chercher d'autres phénomènes isopérimétriques dans des contextes plus généraux. À un niveau plus abstrait, dans un espace métrique mesuré $(\mathcal{X}, d_{\mathcal{X}}, \mu)$ le problème isopérimétrique consiste, étant donné une valeur positive fixe $\alpha > 0$, à identifier les parties boréliennes B avec $\mu(B) = \alpha$ ayant un périmètre $\mu^+(B)$ minimal. De manière équivalente, on cherche à identifier le profil isopérimétrique $\mathcal{I}_\mu$ associé à μ.

L'espace gaussien $(\mathbb{R}^d, |\cdot|, \gamma_d)$ est un exemple fondamental, car les solutions au problème isopérimetrique correspondant sont totalement caractérisées.

Théorème 0.6 (Inégalité isopérimétrique gaussienne)**.** *Soient $\alpha \in (0,1)$ et $(\theta_0, t_0) \in \mathbb{R}^d \times \mathbb{R}$ tel que le demi-espace $H = \{x \in \mathbb{R}^d : \theta_0 \cdot x \leqslant t_0\}$ ait pour mesure $\gamma_d(H) = \alpha$. Alors, pour chaque partie borélienne $A \subset \mathbb{R}^d$ avec $\gamma_d(A) = \alpha$ et tout $r > 0$,*

$$\gamma_d(A_r) \geqslant \gamma_d(H_r), \tag{0.0.7}$$

où H_r désigne le r-voisinage de H. En particulier,

$$\gamma_d^+(A) \geqslant \gamma_d^+(H). \tag{0.0.8}$$

De plus, le profil isopérimétrique de $(\mathbb{R}^d, |\cdot|, \gamma_d)$ est donné par la fonction

$$\mathcal{I}_{\gamma_d} = \mathcal{I}_\gamma \coloneqq \Phi' \circ \Phi^{-1}, \tag{0.0.9}$$

où $\Phi \colon \mathbb{R} \to \mathbb{R}_+$ désigne la fonction définie par

$$\forall r \in \mathbb{R}, \quad \Phi(r) = \gamma_1((-\infty, r)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r} e^{-x^2/2}\,\mathrm{d}x.$$

L'une des conséquences les plus remarquables du théorème 0.6 est le phénomène de concentration de la mesure dans l'espace gaussien.

Corollaire 0.7. *Soit $A \subset \mathbb{R}^d$ une partie borélienne avec $\gamma_d(A) = 1/2$. Alors*

$$\forall r > 0, \quad \gamma_d(A_r) \geqslant 1 - \frac{1}{2}\exp\left(-r^2/2\right),$$

où A_r désigne le r-voisinage de l'ensemble A.

L'inégalité isopérimétrique gaussienne a également une version fonctionnelle.

Théorème 0.8 (Inégalité isopérimétrique gaussienne fonctionnelle)**.** *La mesure gaussienne sur $\mathbb{R}^d$ satisfait*

$$\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d, [0,1]), \quad \mathcal{I}_\gamma \left(\int_{\mathbb{R}^d} f \,\mathrm{d}\gamma_d \right) \leqslant \int_{\mathbb{R}^d} \sqrt{(\mathcal{I}_\gamma \circ f)^2 + |\nabla f|^2}\,\mathrm{d}\gamma_d.$$

L'inégalité isopérimétrique gaussienne fonctionnelle est préservée par des images lipschitziennes. En particulier, d'après le théorème 0.2, on en déduit que toute perturbation log-concave de la mesure gaussienne satisfait une inégalité isopérimétrique comme celle du théorème 0.8 en remplaçant la mesure gaussienne γ_d par la perturbation. Comme corollaire, on obtient pour la perturbation une inégalité de concentration similaire à celle du corollaire 0.7.

La motivation principale de notre étude des inégalités isopérimétriques est son lien avec le phénomène de concentration de la mesure, comme nous l'avons vu dans le corollaire 0.7 dans le cas gaussien. Par conséquent, nous allons nous intéresser dans la suite aux propriétés de concentration des espaces métriques mesurés $(\mathcal{X}, d_\mathcal{X}, \mu)$ qui peuvent être étudiées grâce à leur fonction de concentration $\alpha_\mu \colon \mathbb{R}_{>0} \to [0,1]$:

$$\forall r > 0, \quad \alpha_\mu(r) \coloneqq \sup\{1 - \mu(A_r) : \mu(A) \geqslant 1/2\}.$$

Par exemple, le corollaire 0.7 donne une borne supérieure de la fonction de concentration de l'espace gaussien :

$$\alpha_{\gamma_d}(r) \leqslant \frac{1}{2} \exp\left(-r^2/2\right). \tag{0.0.10}$$

Nous pouvons utiliser la borne de concentration (0.0.10) comme référence pour étudier la concentration dans des espaces plus généraux. Plus précisément, nous dirons qu'une mesure arbitraire possède la concentration sous-gaussienne si sa fonction de concentration est bornée supérieurement par un terme du même ordre que (0.0.10). De même, nous pouvons également utiliser les lois exponentielle et de Poisson pour modéliser de différents types de concentration ce qui donnera lieu aux concentrations sous-exponentielle et sous-poissonienne, respectivement.

D'autre part, nous souhaiterions disposer de bornes de concentration indépendantes de la dimension comme dans (0.0.10) : notons que la borne supérieure ne dépend pas de la dimension intrinsèque d. En particulier, cette propriété du phénomène de concentration gaussien est très utile pour des applications dans lesquelles nous voudrions étudier la concentration d'un grand nombre de variables aléatoires indépendantes et identiquement distribuées, en souhaitant obtenir une borne indépendante de la quantité de variables aléatoires. Malheureusement, en général, les inégalités de concentration ne se tensorisent pas d'une manière indépendante de la dimension ce qui entrave l'obtention d'une telle borne.

Afin de remédier au problème signalé ci-dessous, la suite du chapitre 2 sera consacrée à l'étude de trois familles importantes d'inégalités fonctionnelles qui entraînent des bornes de concentration indépendantes de la dimension et qui peuvent être établies grâce à des critères directs basés sur des propriétés de convexité, donc plus faciles à obtenir que les inégalités isopérimétriques. Il s'agit des inégalités de Poincaré, des inégalités de Sobolev logarithmiques et des inégalités de transport-entropie. Nous allons réviser leurs propriétés concernant leur tensorisation, leur concentration, leur stabilité par image, etc. ; la hiérarchie entre elles ; et leurs exemples les plus importants. Par exemple, la mesure gaussienne satisfait une inégalité de Sobolev logarithmique.

Théorème 0.9 (Inégalité de Sobolev logarithmique gaussienne)**.** *La mesure gaussienne sur $\mathbb{R}^d$ satisfait une inégalité de Sobolev logarithmique :*

$$\forall f \in \mathcal{C}_c^\infty(\mathbb{R}^d), \quad \mathrm{Ent}_{\gamma_d}(f^2) \leqslant 2 \int_{\mathbb{R}^d} |\nabla f|^2 \, \mathrm{d}\gamma_d. \tag{0.0.11}$$

Nous allons finir le chapitre 2 en révisant la célèbre condition de courbure-dimension de Bakry-Émery dans le cadre riemannien qui entraîne la validité des inégalités fonctionnelles précitées. Il s'agit d'une condition géométrique simple qui peut être exprimée en termes algébriques en utilisant les opérateurs Γ et Γ_2. Les conséquences de la condition de Bakry-Émery vont au-delà de la théorie des inégalités fonctionnelles ; par exemple, elle fournit des bornes des noyaux de la chaleur qui vont jouer un rôle capital dans le chapitre 3.

Dans le chapitre 2, nous allons également étudier des inégalités fonctionnelles dans le cadre discret. Plus précisément, on sait que pour chaque $T > 0$ la loi de Poisson π_T ne satisfait pas une inégalité de Sobolev logarithmique dans la forme (0.0.11). Cependant, elle vérifie des inégalités plus faibles, à savoir les inégalités de Sobolev logarithmiques modifiées. Elles jouent un rôle important dans la théorie, car elles permettent de retrouver la concentration poissonienne indépendante de la dimension. L'inégalité la plus forte dans cette classe est l'inégalité de Wu, qui sera très importante dans les chapitres 4 et 5.

Théorème 0.10 (Inégalité de Sobolev logarithmique modifiée de Wu)**.** *Soit $T > 0$ et soit π_T la loi de Poisson de paramètre T sur $\mathbb{N}$. Alors π_T satisfait l'inégalité de Sobolev logarithmique modifiée suivante :*

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \mathrm{Ent}_{\pi_T}(f) \leqslant T\,\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)], \tag{0.0.12}$$

où

$$\forall u > 0, \forall u + v > 0, \quad \Psi(u, v) \coloneqq (u + v)\log(u + v) - u \log u - (1 + \log u)v,$$

et $\mathrm{D}f$ *désigne la dérivée discrète :*

$$\forall k \in \mathbb{N}, \quad \mathrm{D}f(k) \coloneqq f(k + 1) - f(k).$$

Contributions originales (Original contributions)

La deuxième partie de ce manuscrit de thèse présentera les contributions originales que j'ai réalisées seul ou en collaboration, et qui correspondent aux quatre articles suivants, présentés ci-dessous par ordre d'apparition :

- Pablo López-Rivera. A Bakry-Émery Approach to Lipschitz Transportation on Manifolds. *Potential Anal.*, 62(2):331–353, 2025
- Pablo López-Rivera and Yair Shenfeld. The Poisson transport map. *J. Funct. Anal.*, 288(10):Paper No. 110864, 2025
- Shrey Aryan, Pablo López-Rivera, and Yair Shenfeld. The stability of Wu's logarithmic Sobolev inequality via the Poisson-Föllmer process. *arXiv preprint arXiv:2410.06117*, 2024
- Pablo López-Rivera. A uniform rate of convergence for the entropic potentials in the quadratic Euclidean setting. *arXiv preprint arXiv:2502.00084*, 2025

En particulier, chaque chapitre de cette partie correspond à une adaptation de chacun des articles mentionnés ci-dessus, en suivant l'ordre chronologique.

L'application de transport de diffusion (The diffusion transport map)

La première contribution de cette thèse s'inscrit dans l'esprit du théorème 0.2, le théorème de contraction de Caffarelli. Dans le cadre riemannien, il existe des obstacles qui empêchent d'en obtenir des généralisations, c'est-à-dire, de trouver des applications lipschitziennes entre une mesure source et des perturbations log-concaves, voir par exemple [FFGZ24]. Néanmoins, Fathi, Mikulincer et Shenfeld [FMS24] ont pu montrer un tel résultat pour des perturbations log-lipschitziennes d'une mesure définie sur une variété lisse riemannienne en utilisant l'application de transport Kim-Milman [KM12], issue du processus de Langevin associé à la variété, sous une majoration du tenseur de courbure de Riemann de la variété.

Dans ce contexte, dans le résultat principal du chapitre 3 nous montrons que pour toute variété à poids et à courbure contrôlée aux premier et deuxième ordres dans le sens de Bakry-Émery, l'application de Kim-Milman poussant en avant la mesure de poids et toute perturbation log-lipschitzienne est alors lipschitzienne, voir le théorème 3.21.

Théorème 0.11. *Soit* (M, g, μ) *une variété riemannienne complète connexe à poids avec* $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ *pour* $W \in \mathcal{C}^\infty(M)$ *et* $\mu \in \mathcal{P}(M)$. *Soit* $\mathrm{L} = \Delta - \nabla W \cdot \nabla$, *et soient* Γ, Γ_2 *et* Γ_3 *son carré du champ et ses deux premières itérations dans le sens de Bakry-Émery. Supposons qu'il existe des constantes réelles* $\rho_1, \rho_2 > 0$ *tels que*

(i) $\forall f \in \mathcal{C}_\mathrm{c}^\infty(M), \Gamma_2(f) \geqslant \rho_1 \Gamma(f)$ *; et*

(ii) $\forall f \in \mathcal{C}_\mathrm{c}^\infty(M), \Gamma_3(f) \geqslant \rho_2 \Gamma_2(f)$.

Soit $V \in \mathcal{C}^\infty(M)$ *et supposons qu'il est* K*-lipschitzienne pour* $K > 0$. *Définissons* $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\mu$ *et supposons que* $\nu \in \mathcal{P}(M)$. *Alors il existe une application* $\exp\left(\sqrt{\frac{2\pi}{\rho_2}} K e^{\frac{K^2}{2\rho_1}}\right)$*-lipschitzienne* $T\colon M \to M$ *qui envoie la mesure* μ *sur* ν.

Ce résultat permet le transfert d'inégalités fonctionnelles ; voir le corollaire 3.24 pour le cas particulier des inégalités de Sobolev logarithmiques.

Corollaire 0.12. *Dans le contexte du théorème 0.11, soit* $\nu \in \mathcal{P}(M)$ *une perturbation* K*-log-lipschitzienne de la mesure* μ. *Alors* ν *satisfait une inégalité de Sobolev logarithmique avec*

$$C_{\mathrm{LS}}(\nu) \leqslant \frac{2\exp\left(2\sqrt{\frac{2\pi}{\rho_2}} K e^{\frac{K^2}{2\rho_1}}\right)}{\rho}.$$

La sphère $\mathbb{S}^d$ et le générateur de Laguerre sur $\mathbb{R}_{>0}$ sont des exemples où l'on peut appliquer les résultats précédents. Dans le dernier cas, nous exhibons une estimation pour la croissance de l'application de Brenier en dimension dans le cas gamma ; voir la proposition 3.28.

Proposition 0.13. *Soient* μ_p *la loi gamma sur* $\mathbb{R}_{>0}$ *et* $V\colon \mathbb{R}_{>0} \to \mathbb{R}$ *un potentiel lipschitzien pour la métrique* $x \mapsto \frac{1}{x}$. *Considérons* $T\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ *l'application de Brenier envoyant* μ_p *sur* $e^{-V}\mu_p$. *Alors il existe une constante* $C > 0$ *telle que pour tout* $x > 0$,

$$0 < T(x) \leqslant Cx. \tag{0.0.13}$$

De plus, T *est lipschitzienne pour la métrique euclidienne sur* $\mathbb{R}_{>0}$ *; c'est-à-dire, il existe une constante* $C' > 0$ *telle que pour tout* $x > 0$,

$$0 \leqslant T'(x) \leqslant C'. \tag{0.0.14}$$

L'application de transport de Poisson (The Poisson transport map)

Dans le cadre discret nous pouvons considérer la loi de Poisson π_T, pour chaque paramètre $T > 0$, comme un analogue de la mesure gaussienne. Dans ce sens, nous nous posons la question suivante : est-il possible d'obtenir un résultat similaire au théorème 0.2 en remplaçant la mesure gaussienne par π_T, le gradient d'une fonction par sa version discrète D et la log-concavité par une version adaptée au cadre discret ? En revanche, la classe des mesures obtenues comme image de π_T par une application est très limitée, contrairement au cas continu, où le théorème 0.1 garantit l'existence des applications de transport sous de faibles hypothèses. D'autre part, l'une des applications les plus importantes du théorème de contraction de Caffarelli est le transfert d'inégalités fonctionnelles. Pour accomplir cette tâche, la règle de la chaîne satisfaite par l'opérateur ∇ joue un rôle fondamental, mais l'opérateur discret D ne la satisfait pas.

Dans le chapitre 4, qui est basé sur l'article [LRS25], écrit en collaboration avec Yair Shenfeld, nous allons construire une application qui envoie des processus de Poisson ponctuels sur des mesures ultra-log-concaves sur les entiers naturels, que nous appellerons l'application de transport de Poisson. Nour remarquons que les mesures ultra-log-concaves sont les mesures plus log-concaves que la loi de Poisson dans le cadre discret. De plus, nous allons montrer que cette application est une contraction. Sa construction est basée sur un processus qui minimise l'entropie que nous appelons le processus de Poisson-Föllmer, qui a été introduit précédemment par Klartag et Lehec [KL19]. Le résultat suivant énonce la propriété contractive de l'application de transport de Poisson, voir le corollaire 4.18.

Théorème 0.14. *Soient $T > 0$ et $\mu = f\pi_T$ une mesure de probabilité ultra-log-concave sur $\mathbb{N}$, notons $M := f(1)/f(0)$. Soit X_T l'application de transport de Poisson envoyant $\mathbb{P}$ sur μ. Alors, $\mathbb{P}$-presque sûrement,*

$$\forall (t,z) \in [0,T] \times [0,M], \quad \mathrm{D}_{(t,z)} X_T \in \{0,1\},$$

où $\mathrm{D}_{(t,z)}$ désigne la dérivée de Malliavin au point $(t,z) \in [0,T] \times [0,M]$.

Cette approche nous permet de surmonter les difficultés qui entravent le transfert d'inégalités fonctionnelles dans le cadre discret en utilisant des applications de transport. Nous obtiendrons ainsi de nouvelles inégalités fonctionnelles pour les mesures ultra-log-concaves. En particulier, notre approche permet d'améliorer la constante connue pour l'inégalité de Sobolev logarithmique modifiée pour les mesures ultra-log-concaves.

Théorème 0.15. *Soit μ une mesure de probabilité ultra-log-concave sur $\mathbb{N}$. Alors, pour tout $g \in L^2(\mathbb{N}, \mu)$ strictement positive,*

$$\mathrm{Ent}_\mu(g) \leqslant |\log \mu(0)|\, \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)], \tag{0.0.15}$$

où $\Psi(u,v) := (u+v)\log(u+v) - u\log u - (\log u + 1)v$.

Le théorème 0.15 n'est qu'une conséquence d'un résultat plus général : notre approche nous permet de transporter les inégalités de Φ-Sobolev de Chafaï pour des mesures ultra-log-concaves, voir le théorème 4.24.

Théorème 0.16. *Soit μ une mesure de probabilité ultra-log-concave sur $\mathbb{N}$. Soient $\mathcal{I} \subset \mathbb{R}$ un intervalle fermé et $\Phi\colon \mathcal{I} \to \mathbb{R}$ une fonction lisse convexe. Supposons que la fonction*

$$\{(u,v) \in \mathbb{R}^2 : (u, u+v) \in \mathcal{I} \times \mathcal{I}\} \ni (u,v) \quad \mapsto \quad \Psi(u,v) := \Phi(u+v) - \Phi(u) - \Phi'(u)v$$

est positive et convexe. Alors, pour chaque $g \in L^2(\mathbb{N}, \mu)$ tel que μ-p.s. $g, g + \mathrm{D}g \in \mathcal{I}$,

$$\mathrm{Ent}_\mu^\Phi(g) \leqslant |\log \mu(0)|\, \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)]. \tag{0.0.16}$$

Enfin, nous obtenons aussi une inégalité de transport-entropie α-T_1 pour les mesures ultra-log-concaves, voir le théorème 4.27.

Théorème 0.17. *Soit $\mu = f\pi_T$ une mesure de probabilité ultra-log-concave sur $\mathbb{N}$ et $M := \frac{f(1)}{f(0)}$. Alors, pour chaque mesure de probabilité ν sur $\mathbb{N}$ absolument continue par rapport à μ admettant un moment d'ordre* 1,

$$\alpha_{TM}\left(\mathrm{W}_{1,|\cdot|}(\nu, \mu)\right) \leqslant \mathrm{H}(\nu|\mu), \tag{0.0.17}$$

où

$$\alpha_c(r) := c\left[\left(1 + \frac{r}{c}\right)\log\left(1 + \frac{r}{c}\right) - \frac{r}{c}\right].$$

Stabilité de l'inégalité de Wu (Stability of Wu's inequality)

Le chapitre 5 est basé sur l'article [ALRS24], qui a été écrit en collaboration avec Shrey Aryan et Yair Shenfeld. D'abord, nous exhibons une nouvelle preuve stochastique de l'inégalité de Wu (théorème 0.10) en utlisant une formulation variationnelle stochastique pour l'entropie qui généralise la propriété d'entropie minimale du processus de Poisson-Föllmer, qui était utilisé dans le chapitre 4. De plus, cette approche nous permet d'identifier les cas d'égalité, voir la proposition 5.12.

Proposition 0.18. *Rappelons l'inégalité de Wu :*

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \mathrm{Ent}_{\pi_T}(f) \leqslant T\, \mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)]. \tag{0.0.18}$$

Soit $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ avec $\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] < \infty$. Alors f atteint l'égalité en (0.0.18) *si et seulement s'il existe $a, b \in \mathbb{R}$ tels que $f(k) = e^{ak+b}$ pour tout $k \in \mathbb{N}$.*

Notre approche stochastique nous permet également d'obtenir un résultat de stabilité quantitative pour l'inégalité de Wu. Plus précisément, nous minorons le déficit de l'inégalité sous des hypothèses de convexité, voir le théorème 5.17. Pour $T > 0$ et $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ in $L^1(\pi_T)$ nous définissons son déficit (par rapport à l'inégalité de Wu) par

$$\delta(f) := T\, \mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] - \mathrm{Ent}_{\pi_T}(f).$$

Théorème 0.19. *Soit $T > 0$. Soient $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ ultra-log-concave dans $L^1(\pi_T)$ et $\mu := \frac{f\pi_T}{\int f \,\mathrm{d}\pi_T}$. Alors*

$$\delta(f) \geqslant \frac{T^2}{2}\, \Theta_{\frac{f(0)}{f(1)}}\left(\frac{\mathbb{E}[\mu]}{T}\right),$$

où, pour $c > 0$,

$$\Theta_c(z) := \frac{z^2}{1+cz}\log\left(\frac{1}{1+cz}\right) - \frac{z^2}{1+cz} + z^2, \quad z \geqslant 0.$$

Sur la convergence des potentiels entropiques (Convergence of the entropic potentials)

Les dernières contributions de cette thèse s'écartent des sujets traités dans les trois chapitres précédents et concernent le problème de transport optimal et sa régularisation entropique. Pour chaque $\varepsilon > 0$, il est possible de régulariser le problème (0.0.1) en ajoutant une entropie :

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x-y|^2 \,\mathrm{d}\pi(x,y) + \varepsilon\,\mathrm{H}(\pi|\mu\otimes\nu), \tag{0.0.19}$$

où $\mathrm{H}(\cdot|\mu\otimes\nu)$ désigne la fonctionnelle d'entropie relative par rapport à la mesure $\mu\otimes\nu$. Lorsque $\varepsilon\to 0$, le problème (0.0.19) converge vers le problème (0.0.1) au sens large : par exemple, les solutions au problème dual à (0.0.19), c'est-à-dire, les potentiels entropiques $(\varphi_\varepsilon,\psi_\varepsilon)$, convergent vers les potentiels de Brenier (φ_0,ψ_0) [GT21, NW22, CCGT23].

Dans le chapitre 6, qui est basé sur l'article [LR25b], nous obtenons une borne supérieure pour le taux de convergence uniforme sur des ensembles compacts pour les potentiels entropiques et leurs gradients vers le potentiel de Brenier et son gradient, respectivement. Ces résultats sont valides dans le cadre quadratique euclidien, pour des mesures absolument continues satisfaisant les suivantes hypothèses :

(A1) Les mesures $\mu,\nu\in\mathcal{P}(\mathbb{R}^d)$ ont la forme $\mathrm{d}\mu(x)=e^{-V(x)}\,\mathrm{d}x$ et $\mathrm{d}\nu(y)=e^{-W(y)}\,\mathrm{d}y$, avec $V,W\colon\mathbb{R}^d\to\mathbb{R}$ lisses et telles qu'il existe $\alpha,\beta>0$ tels que

$$\forall x\in\mathbb{R}^d,\quad \nabla^2 V(x)\preccurlyeq \alpha I_d \tag{0.0.20}$$

et

$$\forall y\in\mathbb{R}^d,\quad \nabla^2 W(y)\succcurlyeq \beta I_d, \tag{0.0.21}$$

où I_d désigne la matrice identité de dimension d et $\preccurlyeq$ l'ordre de Löwner dans l'ensemble des matrices semi-définies positives.

(A2) La mesure μ satisfait une inégalité de Poincaré : il existe $C_{\mathrm{P}}(\mu)>0$ tel que pour chaque $h\colon\mathbb{R}^d\to\mathbb{R}$ lisse avec $\int_{\mathbb{R}^d} h\,\mathrm{d}\mu=0$,

$$\|h\|^2_{L^2(\mu)}\leqslant C_{\mathrm{P}}(\mu)\|\nabla h\|^2_{L^2(\mu)}.$$

(A3) La mesure μ a une entropie finie :

$$-\infty<\mathrm{H}(\mu)\coloneqq-\int_{\mathbb{R}^d} V(x)e^{-V(x)}\,\mathrm{d}x<+\infty.$$

Les théorèmes suivants sont les résultats principaux du chapitre 6, voir les théorèmes 6.2 et 6.3, respectivement.

Théorème 0.20. *Soient μ et ν deux mesures de probabilité sur $\mathbb{R}^d$ absolument continues par rapport à la mesure de Lebesgue qui satisfont les hypothèses (A1), (A2) et (A3). Alors, pour chaque compact $K\subset\mathbb{R}^d$, il existe une constante calculable $C_{\mathrm{grad}}=C_{\mathrm{grad}}(K,\mu,\nu,d)>0$ telle que pour tout $\varepsilon>0$,*

$$\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|_{K,\infty}\leqslant C_{\mathrm{grad}}\,\varepsilon^{\frac{1}{d+4}}.$$

Théorème 0.21. *Soient μ et ν deux mesures de probabilité sur $\mathbb{R}^d$ absolument continues par rapport à la mesure de Lebesgue qui satisfont les hypothèses (A1), (A2) et (A3). Supposons que les potentiels satisfont la normalisation suivante : pour tout $\varepsilon > 0$,*

$$\int_{\mathbb{R}^d} \varphi_\varepsilon \,\mathrm{d}\mu = \int_{\mathbb{R}^d} \varphi_0 \,\mathrm{d}\mu = 0. \tag{0.0.22}$$

Alors, pour chaque compact connexe $K \subset \mathbb{R}^d$, il existe une constante calculable $C_{\mathrm{pot}} = C_{\mathrm{pot}}(K, \mu, \nu, d) > 0$ telle que pour tout $\varepsilon > 0$,

$$\|\varphi_\varepsilon - \varphi_0\|_{K,\infty} \leqslant C_{\mathrm{pot}} \left(\varepsilon^{\frac{1}{d+4}} + \varepsilon\right).$$

Introduction (English)

Attention, mesdames et messieurs, dans un instant, ça va commencer
Nous vous demandons évidemment d'être indulgents
Le spectacle n'est pas bien rôdé, laissez-nous encore quelques années
Il ne pourrait que s'améliorer au fil du temps.

Michel Fugain et Le Big Bazar
Attention, Mesdames et Messieurs

Since the beginning of my doctoral studies in October 2022, I have focused my research on probability theory and its connections with analysis and geometry. More specifically, I have sought to understand how these fields interact through the lens of the theories of stochastic processes, optimal transport, functional inequalities, and their interplay. The results exhibited in this thesis belong to that framework.

Of course, the main objective of this thesis manuscript is to present my original mathematical contributions since I started my PhD. Nevertheless, to do so, I have also decided to put those results in a bigger context, thus reviewing and surveying the essential preliminaries for their understanding. I decided to do so for two main reasons: the first and most evident one concerns the readability of this manuscript; I wanted to produce a self-contained text just assuming a basic knowledge of probability theory, analysis, and geometry. The second reason is part of a more "personal" exercise: I firmly believe that writing, one of the most precious capacities we human beings have, constitutes a powerful ally for understanding, synthesizing, and getting the global picture of a subject.

In this introduction, I summarize the structure of this manuscript, highlighting the elements of each chapter that reveal the fundamental storyline of this thesis. I divided the manuscript into two parts: the first contains the preliminaries and basic results that provide context for the second part, where I present the original contributions I have made from the beginning of my PhD.

Preliminaries

The first part of the manuscript aims to review the essential results of the theories of optimal transport and functional inequalities pertinent to this thesis in two dedicated chapters.

Optimal transport

We will commence in Chapter 1 by reviewing the essential elements concerning the theory of optimal transport, starting with its historical roots and then delving into the general theory, where we will define the Monge and Kantorovich formulations for the optimal transport problem. After that, we will focus on the quadratic Euclidean setting, where the optimal transport problem in its Kantorovich formulation reads as follows: given two Borel probability measures μ and ν defined on $\mathbb{R}^d$, it corresponds to the variational problem

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x-y|^2 \,\mathrm{d}\pi(x,y), \tag{0.0.23}$$

where $\Pi(\mu,\nu)$ denotes the set of transport plans between the measures μ and ν.

In the quadratic Euclidean setting, the Brenier-McCann theorem reveals the rich structure of the solutions of (0.0.23), which in turn coincide with the solutions to the associated Monge problem.

Theorem 0.1 (Brenier-McCann)**.** *Let μ and ν be two Borel probability measures on $\mathbb{R}^d$, and assume that μ is absolutely continuous with respect to the d-dimensional Lebesgue measure. Then there exists a lower semicontinuous proper convex function $\varphi_0\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ such that the map $T_0 \coloneqq \nabla\varphi_0$ pushes forward μ towards ν. Furthermore, T_0 is the unique gradient of a convex function sending μ to ν.*

Additionally, assume now that μ and ν have finite moments of order 2. Then $\mathcal{C}_0(\mu,\nu) < +\infty$, and there exists a unique optimal plan $\pi_0 \in \Pi(\mu,\nu)$ for (0.0.23), given by the transport plan induced by T_0: $\pi_0 = (\mathrm{id}_{\mathbb{R}^d}, T_0)_{\#}\mu$.

Moreover, if we define $\psi_0 \coloneqq \varphi_0^*$ as the convex conjugate of φ_0, then $(f_0,g_0) \coloneqq \big(\frac{1}{2}|\cdot|^2 - \varphi_0, \frac{1}{2}|\cdot|^2 - \psi_0\big)$ is a pair of Kantorovich potentials, that is, it solves the dual problem of (0.0.23), which is given by

$$\sup_{(f,g)\in\Phi_2} \int_{\mathbb{R}^d} f\,\mathrm{d}\mu + \int_{\mathbb{R}^d} g\,\mathrm{d}\nu, \tag{0.0.24}$$

where Φ_2 is the set

$$\Phi_2 \coloneqq \left\{(f,g)\in L^1(\mu)\times L^1(\nu) : f\oplus g \leqslant \frac{1}{2}|\cdot - \cdot|^2\right\}.$$

Then we will focus on the regularity theory for the Brenier map since it plays a vital role in the theory of functional inequalities, particularly in this thesis. More precisely, if we have a measure μ that satisfies certain functional inequalities, then for any Lipschitz map T, the associated pushforward measure $T_{\#}\mu$ verifies the same inequalities as μ. In this regard, Caffarelli's contraction theorem provides this desired regularity for the Brenier map pushing forward the Gaussian measure towards any log-concave perturbation of it.

Theorem 0.2 (Caffarelli's contraction theorem). *Let γ_d be the d-dimensional Gaussian measure on $\mathbb{R}^d$. Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ with $\nu \ll \gamma_d$, and suppose that it has the form $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\gamma_d$, where $V\colon \mathbb{R}^d \to \mathbb{R}$. If V is convex, the Brenier map pushing forward γ_d towards ν is 1-Lipschitz.*

Functional inequalities

In Chapter 2, we will focus on the theory of functional inequalities. More precisely, we will start by introducing some geometric inequalities and their functional counterparts. The main example corresponds to the Brunn-Minkowski and Prékopa-Leindler inequalities, the latter being a functional version of the former, which is a purely geometric inequality.

Theorem 0.3 (Brunn-Minkowski). *Let $A, B \subset \mathbb{R}^d$ be two convex bodies. Then*

$$\mathrm{Vol}_d(A+B)^{1/d} \geqslant \mathrm{Vol}_d(A)^{1/d} + \mathrm{Vol}_d(B)^{1/d}. \tag{0.0.25}$$

Theorem 0.4 (Prékopa-Leindler). *Let $f, g, h\colon \mathbb{R}^d \to \mathbb{R}_+$ be nonnegative Borel functions with f and g Lebesgue-integrable, and let $\lambda \in (0,1)$. Assume that*

$$\forall x, y \in \mathbb{R}^d, \quad h(\lambda x + (1-\lambda)y) \geqslant f(x)^\lambda g(y)^{1-\lambda}. \tag{0.0.26}$$

Then

$$\int_{\mathbb{R}^d} h\,\mathrm{d}x \geqslant \left(\int_{\mathbb{R}^d} f\,\mathrm{d}x\right)^\lambda \left(\int_{\mathbb{R}^d} g\,\mathrm{d}x\right)^{1-\lambda}. \tag{0.0.27}$$

Then we will see that a remarkable corollary of the Brunn-Minkowski theorem is the Euclidean isoperimetric inequality, which states that for a prescripted volume, the sets with minimal perimeter are balls.

Theorem 0.5 (Euclidean isoperimetric inequality). *For any convex body $K \subset \mathbb{R}^d$,*

$$\mathrm{Vol}_{d-1}(\partial K) \geqslant d\,\mathrm{Vol}_d(K)^{(d-1)/d}\,\mathrm{Vol}_d(B)^{1/d}, \tag{0.0.28}$$

where $B \coloneqq \overline{\mathrm{B}}(0,1)$.

The validity of Theorem 0.5 motivates the study of isoperimetric phenomena on more general spaces, which will be the next step in our journey through geometric inequalities. At an abstract level, in the context of a metric measure space $(\mathcal{X}, d_{\mathcal{X}}, \mu)$, the isoperimetric problem consists of, given a fixed positive value $\alpha > 0$, identifying the Borel sets B with $\mu(B) = \alpha$ having minimal perimeter $\mu^+(B)$. Equivalently, one aims to identify the isoperimetric profile $\mathcal{I}_\mu$ associated with μ.

The Gaussian space $(\mathbb{R}^d, |\cdot|, \gamma_d)$ is a key example since the solution to the associated isoperimetric problem is completely determined.

Theorem 0.6 (Gaussian isoperimetric inequality)**.** *Let $\alpha \in (0,1)$, and let $(\theta_0, t_0) \in \mathbb{R}^d \times \mathbb{R}$ be such that the halfspace $H = \{x \in \mathbb{R}^d : \theta_0 \cdot x \leqslant t_0\}$ has mass α: $\gamma_d(H) = \alpha$. Then for every $A \subset \mathbb{R}^d$ Borel with $\gamma_d(A) = \alpha$, and every $r > 0$, we have*

$$\gamma_d(A_r) \geqslant \gamma_d(H_r), \tag{0.0.29}$$

where H_r is the r-enlargement of the set H. In particular,

$$\gamma_d^+(A) \geqslant \gamma_d^+(H). \tag{0.0.30}$$

Moreover, the isoperimetric profile of $(\mathbb{R}^d, |\cdot|, \gamma_d)$ is given by the function

$$\mathcal{I}_{\gamma_d} = \mathcal{I}_\gamma := \Phi' \circ \Phi^{-1}, \tag{0.0.31}$$

where $\Phi\colon \mathbb{R} \to \mathbb{R}_+$ is the function defined by

$$\forall r \in \mathbb{R}, \quad \Phi(r) = \gamma_1((-\infty, r)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r} e^{-x^2/2}\,\mathrm{d}x.$$

In particular, Theorem 0.6 has a remarkable consequence, namely the phenomenon of concentration of measure for the Gaussian space.

Corollary 0.7. *Let $A \subset \mathbb{R}^d$ Borel with $\gamma_d(A) = 1/2$. Then*

$$\forall r > 0, \quad \gamma_d(A_r) \geqslant 1 - \frac{1}{2}\exp\left(-r^2/2\right),$$

where A_r is the r-enlargement of the set A.

Another notable feature of the Gaussian isoperimetric phenomenon is that it admits a (equivalent) functional version.

Theorem 0.8 (Functional Gaussian isoperimetric inequality)**.** *The d-dimensional Gaussian measure satisfies*

$$\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d, [0,1]), \quad \mathcal{I}_\gamma\left(\int_{\mathbb{R}^d} f\,\mathrm{d}\gamma_d\right) \leqslant \int_{\mathbb{R}^d} \sqrt{(\mathcal{I}_\gamma \circ f)^2 + |\nabla f|^2}\,\mathrm{d}\gamma_d.$$

The functional Gaussian isoperimetric inequality is preserved by Lipschitz push-forwards. In particular, by applying Theorem 0.2, we deduce that any log-concave perturbation of the Gaussian measure satisfies an isoperimetric inequality in the fashion of Theorem 0.8 just replacing the Gaussian measure γ_d by the perturbation, which in turn yields a concentration bound similar to the one provided by Corollary 0.7 for the perturbation.

For us, the main motivation behind isoperimetric inequalities is the fact that they imply concentration of measure, as Corollary 0.7 reveals in the Gaussian case. Hence, the next topic reviewed in Chapter 2 will be the concentration properties of an abstract metric probability space $(\mathcal{X}, d_{\mathcal{X}}, \mu)$, which can be studied by its associated concentration function $\alpha_\mu \colon \mathbb{R}_{>0} \to [0,1]$:

$$\forall r > 0, \quad \alpha_\mu(r) \coloneqq \sup\{1 - \mu(A_r) : \mu(A) \geqslant 1/2\}.$$

For example, for the Gaussian space, Corollary 0.7 says that

$$\alpha_{\gamma_d}(r) \leqslant \frac{1}{2} \exp\left(-r^2/2\right), \tag{0.0.32}$$

so we can model concentration on more general spaces based on this bound. More precisely, we say that a measure has the subgaussian concentration property if its concentration function is upper bounded by a term of the same order of (0.0.32). Similarly, we can also take the exponential and Poisson distributions as models to define the subexponential and subpoissonian concentration types, respectively.

A desirable property for a concentration bound is to be dimension-free as in (0.0.32): note that the right-hand side of the inequality does not depend on the intrinsic dimension of the space d. This feature of the Gaussian concentration phenomenon is helpful for many applications where one wants to analyze the concentration of a large number of independent and identically distributed random variables, aiming to obtain a bound that does not depend on the number of random variables. However, the big problem is that, by themselves, concentration inequalities do not generally tensorize in a dimension-free way, which hinders getting such a nice bound.

To address the issue pointed out above, we will continue the exposition in Chapter 2 by reviewing three prominent families of functional inequalities that will yield dimension-free concentration bounds and for which we have straightforward criteria that ensure their validity, mainly relying on convexity properties, thus easier to obtain than an isoperimetric inequality. These correspond to Poincaré, logarithmic Sobolev, and transport-entropy inequalities. We will review their basic properties concerning tensorization, concentration, stability by pushforwards, etc.; the hierarchy between them; and their most representative examples. For example, concerning the logarithmic Sobolev family, we have that the Gaussian measure satisfies it.

Theorem 0.9 (Gaussian logarithmic Sobolev inequality)**.** *The d-dimensional standard Gaussian measure satisfies a logarithmic Sobolev inequality:*

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}^d), \quad \mathrm{Ent}_{\gamma_d}(f^2) \leqslant 2 \int_{\mathbb{R}^d} |\nabla f|^2 \, \mathrm{d}\gamma_d. \tag{0.0.33}$$

We finish Chapter 2 by reviewing the well-known Bakry-Émery curvature-dimension condition in the smooth setting, which ensures the validity of the aforementioned functional inequalities. It corresponds to a simple geometrical condition that can be equivalently expressed in algebraic terms using the Γ and Γ_2 operators. The consequences of the validity of the Bakry-Émery condition go beyond functional inequalities since it also provides useful heat kernel bounds, which, for example, will play a key role in Chapter 3.

Along Chapter 2, we will also focus on the functional inequalities panorama in the discrete setting. More precisely, it is known that for any $T > 0$, the Poisson measure π_T does not satisfy a full logarithmic Sobolev inequality in the fashion of (0.0.33). However, weaker types of inequalities do hold for π_T, namely modified logarithmic Sobolev inequalities. They are relevant since they allow the retrieval of the dimension-free Poissonian concentration. The strongest inequality in this class is Wu's inequality, which will be of key importance in Chapters 4 and 5.

Theorem 0.10 (Wu's modified logarithmic Sobolev inequality)**.** *Let $T > 0$, and let π_T be the Poisson distribution of parameter T on $\mathbb{N}$. Then π_T satisfies the following modified logarithmic Sobolev inequality:*

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \mathrm{Ent}_{\pi_T}(f) \leqslant T\, \mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)], \tag{0.0.34}$$

where

$$\forall u > 0, \forall u + v > 0, \quad \Psi(u, v) := (u + v)\log(u + v) - u \log u - (1 + \log u)v,$$

and $\mathrm{D}f$ *is the discrete derivative:*

$$\forall k \in \mathbb{N}, \quad \mathrm{D}f(k) := f(k + 1) - f(k).$$

Original contributions

The second part of the thesis devotes its attention to the original contributions made by myself and in collaboration, based on the following four articles, which are listed in chronological order of appearance:

- Pablo López-Rivera. A Bakry-Émery Approach to Lipschitz Transportation on Manifolds. *Potential Anal.*, 62(2):331–353, 2025
- Pablo López-Rivera and Yair Shenfeld. The Poisson transport map. *J. Funct. Anal.*, 288(10):Paper No. 110864, 2025
- Shrey Aryan, Pablo López-Rivera, and Yair Shenfeld. The stability of Wu's logarithmic Sobolev inequality via the Poisson-Föllmer process. *arXiv preprint arXiv:2410.06117*, 2024
- Pablo López-Rivera. A uniform rate of convergence for the entropic potentials in the quadratic Euclidean setting. *arXiv preprint arXiv:2502.00084*, 2025

In particular, each chapter in this part corresponds to an adapted version of one of the articles mentioned above, following the same chronological order.

The diffusion transport map

The first contribution of this thesis goes in the spirit of Theorem 0.2, the Caffarelli contraction theorem. In the smooth setting, there are obstructions to obtaining an exact generalization of Theorem 0.2, i.e., finding Lipschitz maps between a source measure and log-concave perturbations; see [FFGZ24]. Nevertheless, Fathi, Mikulincer, and Shenfeld [FMS24] were able to obtain such a result in the smooth setting for log-Lipschitz perturbations of a sufficiently well-behaved source measure on a smooth Riemannian manifold using the diffusion transport map, construction based on Kim-Milman's reverse heat-flow map [KM12], under boundedness assumptions on the Riemann curvature tensor of the manifold.

In the above context, in the main result of Chapter 3, we show that if we have a weighted Riemannian manifold that has bounded curvature at first and second order in the sense of Bakry-Émery, then the Kim-Milman transport map between the weighted measure and any log-Lipschitz perturbation of it is Lipschitz; see Theorem 3.21.

Theorem 0.11. *Let (M, g, μ) be a complete and connected weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and assume that $\mu \in \mathcal{P}(M)$. Consider the diffusion operator $\mathrm{L} = \Delta - \nabla W \cdot \nabla$, let Γ be its associated carré du champ, and let Γ_2 and Γ_3 be its respective iterations in the Bakry-Émery sense. Assume that there exist constants $\rho_1, \rho_2 > 0$ such that*

(i) $\forall f \in \mathcal{C}_c^\infty(M), \Gamma_2(f) \geqslant \rho_1 \,\Gamma(f)$; and

(ii) $\forall f \in \mathcal{C}_c^\infty(M), \Gamma_3(f) \geqslant \rho_2 \,\Gamma_2(f)$.

Let $V \in \mathcal{C}^\infty(M)$, and assume that it is K-Lipschitz for some $K > 0$. Define $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\mu$ and assume that $\nu \in \mathcal{P}(M)$. Then there exists a Lipschitz map $T\colon M \to M$ pushing forward μ towards ν which is $\exp\left(\sqrt{\frac{2\pi}{\rho_2}} K e^{\frac{K^2}{2\rho_1}}\right)$-Lipschitz.

The main application of this result is the transfer of functional inequalities; see Corollary 3.24 for the particular case of logarithmic Sobolev inequalities.

Corollary 0.12. *In the context of Theorem 0.11, for $K > 0$, let $\nu \in \mathcal{P}(M)$ be a K-log-Lipschitz perturbation of the measure μ. Then ν satisfies a logarithmic Sobolev inequality with constant*

$$C_{\mathrm{LS}}(\nu) \leqslant \frac{2\exp\left(2\sqrt{\frac{2\pi}{\rho_2}} K e^{\frac{K^2}{2\rho_1}}\right)}{\rho}.$$

The sphere $\mathbb{S}^d$ and the Laguerre generator on $\mathbb{R}_{>0}$ are examples of the applicability of these results. In the last case, as another application, we provide an estimate for the growth of the Brenier map in dimension one in the gamma case; see Proposition 3.28.

Proposition 0.13. *Let μ_p be the gamma distribution on $\mathbb{R}_{>0}$, let $V\colon \mathbb{R}_{>0} \to \mathbb{R}$ be a Lipschitz potential (for the metric $x \mapsto \frac{1}{x}$), and let $T\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be the Monge map pushing forward μ_p towards $e^{-V}\mu_p$. Then there exists a constant $C > 0$ such that for any $x > 0$,*

$$0 < T(x) \leqslant Cx. \tag{0.0.35}$$

Moreover, T is Lipschitz for the Euclidean metric on $\mathbb{R}_{>0}$, that is, there exists $C' > 0$ such that for any $x > 0$,

$$0 \leqslant T'(x) \leqslant C'. \tag{0.0.36}$$

The Poisson transport map

In the discrete setting, one can say that the Poisson measure π_T, for any parameter $T > 0$, is an analog of the Gaussian measure. In that sense, one may wonder if it is possible to obtain a result similar to Theorem 0.2 if one replaces the Gaussian measure by π_T, the standard gradient ∇ by its discrete version D, and log-concavity by an appropriate discrete version. However, it is known that the class of measures that arise as a pushforward of π_T is very limited, in contrast with the Euclidean setting where Theorem 0.1 provides the existence of transport maps under mild assumptions. On the other hand, one of the main applications of Caffarelli's contraction theorem is getting new functional inequalities, but in order to do so, the standard chain rule for the operator ∇ is fundamental, and does not hold for the discrete operator D.

In Chapter 4, which is based on the article [LRS25], done in collaboration with Yair Shenfeld, we will construct a transport map from Poisson point processes onto ultra-log-concave measures over the natural numbers, which are the measures that are "discretely" more log-concave than the Poisson distribution, and show that this map is a contraction. We call this construction the Poisson transport map and it is based on an entropy-minimizing stochastic process that we call the Poisson-Föllmer process, introduced previously by Klartag and Lehec [KL19]. The following is the contraction result for the Poisson transport map; see Corollary 4.18.

Theorem 0.14. *Fix a real number $T > 0$, let $\mu = f\pi_T$ be an ultra-log-concave probability measure over $\mathbb{N}$, and let $M := f(1)/f(0)$. Let X_T be the Poisson transport map from $\mathbb{P}$ to μ. Then, $\mathbb{P}$-almost-surely,*

$$\forall (t,z) \in [0,T] \times [0,M], \quad \mathrm{D}_{(t,z)} X_T \in \{0,1\},$$

where $\mathrm{D}_{(t,z)}$ denotes the Malliavin derivative operator at the point $(t,z) \in [0,T]\times[0,M]$.

This approach overcomes the aforementioned obstacles to transferring functional inequalities using transport maps in discrete settings and will allow us to deduce a number of functional inequalities for ultra-log-concave measures. In particular, we provide the currently best known constant in modified logarithmic Sobolev inequalities for ultra-log-concave measures.

Theorem 0.15. *Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Then, for any positive $g \in L^2(\mathbb{N},\mu)$,*

$$\mathrm{Ent}_\mu(g) \leqslant |\log \mu(0)|\, \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)], \tag{0.0.37}$$

where $\Psi(u,v) := (u+v)\log(u+v) - u\log u - (\log u + 1)v$.

Theorem 0.15 is actually a consequence of a more general result, namely the transfer of Chafaï's Φ-Sobolev inequalities for ultra-log-concave measures; see Theorem 4.24.

Theorem 0.16. *Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Let $\mathcal{I} \subset \mathbb{R}$ be a closed interval, not necessarily bounded, and let $\Phi\colon \mathcal{I} \to \mathbb{R}$ be a smooth convex function. Suppose that the function*

$$\{(u,v) \in \mathbb{R}^2 : (u, u+v) \in \mathcal{I}\times\mathcal{I}\} \ni (u,v) \quad \mapsto \quad \Psi(u,v) := \Phi(u+v) - \Phi(u) - \Phi'(u)v$$

is nonnegative and convex. Then, for any $g \in L^2(\mathbb{N}, \mu)$, such that μ-a.s. $g, g + \mathrm{D}g \in \mathcal{I}$,

$$\mathrm{Ent}_\mu^\Phi(g) \leqslant |\log \mu(0)| \, \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)]. \tag{0.0.38}$$

We show as well the following α-T_1 transport-entropy inequality for ultra-log-concave measures; see Theorem 4.27.

Theorem 0.17. *Let $\mu = f\pi_T$ be an ultra-log-concave probability measure on $\mathbb{N}$, and let $M := \frac{f(1)}{f(0)}$. Then, for any probability measure ν on $\mathbb{N}$ which is absolutely continuous with respect to μ, and has a finite first moment, we have*

$$\alpha_{TM}\left(\mathrm{W}_{1,|\cdot|}(\nu, \mu)\right) \leqslant \mathrm{H}(\nu|\mu), \tag{0.0.39}$$

where

$$\alpha_c(r) := c\left[\left(1 + \frac{r}{c}\right)\log\left(1 + \frac{r}{c}\right) - \frac{r}{c}\right].$$

Stability of Wu's inequality

Chapter 5 is based on the article [ALRS24], written in collaboration with Shrey Aryan and Yair Shenfeld. A stochastic proof of Wu's inequality, Theorem 0.10, is given, using a stochastic variational formula for the entropy that generalizes the entropy-minimizing property of the aforementioned Poisson-Föllmer process. Moreover, this new proof leads to the identification of the extremizers of the inequality; see Proposition 5.12.

Proposition 0.18. *Recall Wu's inequality:*

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \mathrm{Ent}_{\pi_T}(f) \leqslant T\, \mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)]. \tag{0.0.40}$$

If $\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] < \infty$, then equality in (0.0.40) is attained if and only if there exist $a, b \in \mathbb{R}$ such that $f(k) = e^{ak+b}$ for all $k \in \mathbb{N}$.

The proof also leads to a quantitative stability result of the inequality that provides a lower bound for the deficit of Wu's inequality under convexity assumptions; see Theorem 5.17. For $T > 0$ and $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ in $L^1(\pi_T)$, we define its deficit (with respect to Wu's inequality) as

$$\delta(f) := T\, \mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] - \mathrm{Ent}_{\pi_T}(f).$$

Theorem 0.19. *Fix $T > 0$. Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be $L^1(\pi_T)$ integrable and ultra-log-concave, and let $\mu := \frac{f\pi_T}{\int f\, \mathrm{d}\pi_T}$. Then,*

$$\delta(f) \geqslant \frac{T^2}{2}\, \Theta_{\frac{f(0)}{f(1)}}\left(\frac{\mathbb{E}[\mu]}{T}\right),$$

where, for $c > 0$,

$$\Theta_c(z) := \frac{z^2}{1 + cz}\log\left(\frac{1}{1 + cz}\right) - \frac{z^2}{1 + cz} + z^2, \quad z \geqslant 0.$$

Convergence of the entropic potentials

The final contributions in this thesis digress from the previous three chapters and concern the optimal transport problem and its entropic regularization: for any $\varepsilon > 0$, it is possible to regularize problem (0.0.23) and add an entropy:

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x-y|^2 \,\mathrm{d}\pi(x,y) + \varepsilon\,\mathrm{H}(\pi|\mu\otimes\nu), \tag{0.0.41}$$

where $\mathrm{H}(\cdot|\mu\otimes\nu)$ denotes the relative entropy functional with respect to the measure $\mu\otimes\nu$. Problem (0.0.41) converges towards problem (0.0.23) as $\varepsilon\to 0$ in many senses. For example, the dual optimizers of (0.0.41), namely the pair of entropic potentials $(\varphi_\varepsilon,\psi_\varepsilon)$, converges towards the Brenier potentials (φ_0,ψ_0) [GT21, NW22, CCGT23].

In Chapter 6, which is based on the article [LR25b], we will provide a bound on the rate of uniform convergence in compact sets for both entropic potentials and their gradients towards the Brenier potential and its gradient. Both results hold in the quadratic Euclidean setting for absolutely continuous measures satisfying the following set of assumptions:

(A1) The measures $\mu,\nu\in\mathcal{P}(\mathbb{R}^d)$ have the form $\mathrm{d}\mu(x) = e^{-V(x)}\,\mathrm{d}x$ and $\mathrm{d}\nu(y) = e^{-W(y)}\,\mathrm{d}y$, where $V,W\colon\mathbb{R}^d\to\mathbb{R}$ are smooth functions and there exist $\alpha,\beta>0$ such that

$$\forall x\in\mathbb{R}^d,\quad \nabla^2 V(x) \preccurlyeq \alpha I_d \tag{0.0.42}$$

and

$$\forall y\in\mathbb{R}^d,\quad \nabla^2 W(y) \succcurlyeq \beta I_d, \tag{0.0.43}$$

where I_d is the identity matrix of dimension d and $\preccurlyeq$ denotes the Löwner order on the set of positive semidefinite matrices.

(A2) The measure μ satisfies a Poincaré inequality: there exists $C_{\mathrm{P}}(\mu)>0$ such that for any $h\colon\mathbb{R}^d\to\mathbb{R}$ smooth with $\int_{\mathbb{R}^d} h\,\mathrm{d}\mu = 0$,

$$\|h\|^2_{L^2(\mu)} \leqslant C_{\mathrm{P}}(\mu)\|\nabla h\|^2_{L^2(\mu)}.$$

(A3) The measure μ has finite differential entropy:

$$-\infty < \mathrm{H}(\mu) := -\int_{\mathbb{R}^d} V(x)e^{-V(x)}\,\mathrm{d}x < +\infty.$$

The following statements correspond to the main results of Chapter 6; see Theorems 6.2 and 6.3, respectively.

Theorem 0.20. *Let μ and ν be two probability measures on $\mathbb{R}^d$ that are absolutely continuous with respect to the d-dimensional Lebesgue measure and satisfy the assumptions (A1), (A2), and (A3). Then, for any $K\subset\mathbb{R}^d$ compact, there exists a computable constant $C_{\mathrm{grad}} = C_{\mathrm{grad}}(K,\mu,\nu,d) > 0$ such that for any $\varepsilon>0$,*

$$\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|_{K,\infty} \leqslant C_{\mathrm{grad}}\,\varepsilon^{\frac{1}{d+4}}.$$

Theorem 0.21. *Let μ and ν be two probability measures on $\mathbb{R}^d$ that are absolutely continuous with respect to the d-dimensional Lebesgue measure and satisfy the assumptions (A1), (A2), and (A3). In addition, suppose that the following normalization holds: for every $\varepsilon > 0$,*

$$\int_{\mathbb{R}^d} \varphi_\varepsilon \, \mathrm{d}\mu = \int_{\mathbb{R}^d} \varphi_0 \, \mathrm{d}\mu = 0. \tag{0.0.44}$$

Then, for any $K \subset \mathbb{R}^d$ compact and connected, there exists a computable constant $C_{\mathrm{pot}} = C_{\mathrm{pot}}(K, \mu, \nu, d) > 0$ such that for any $\varepsilon > 0$,

$$\|\varphi_\varepsilon - \varphi_0\|_{K,\infty} \leqslant C_{\mathrm{pot}} \left(\varepsilon^{\frac{1}{d+4}} + \varepsilon \right).$$

Part I

Preliminaries

Chapter 1

Optimal transport

> No era un problema de gente ni muchísimo menos, señor Bermúdez, sino de transporte, como ya le había explicado al Pulga Heredia, risas y él automáticamente abrió la boca y arrugó la cara.
>
> Mario Vargas Llosa
> *Conversación en La Catedral*

In this first chapter, we will introduce the fundamentals of the theory of optimal transport, which will be of crucial importance in this thesis, especially in Chapter 6, since it contains new contributions to the field. Besides, some elements of the theory will appear in Chapters 2, 3, and 5, more precisely, in their interplay with the theory of functional inequalities.

After an introductory informal discussion about the optimal transport problem and its roots, we formulate it as a variational problem for a general cost function and state the main results concerning the existence of solutions. We then focus on the quadratic Euclidean setting, the rich structure of its solutions, and their regularity theory. Finally, we will review the method of characteristics for the continuity equation to end with the dynamic formulation of optimal transport. From this point onward, the exposition is mainly based on the well-known references [Vil03, Vil09, San15], where the reader can find complete proofs of the results mentioned in this chapter and further insights.

1.1 The origins of optimal transport

In 1781, Gaspard Monge gave birth to the theory of optimal transport, unaware of the impact it would have in our time. This theory, which lies at the intersection of analysis, probability theory, and geometry, has proven to be a dynamic field with many applications in both pure and applied mathematics. For example, it plays a key role in the synthetic characterization of bounded Ricci curvature for metric spaces [Stu06a, Stu06b, LV09]. It allows the study of some partial differential equations as natural gradient flows in the space of probability measures via the Otto calculus formalism [Ott01], which has many further applications and is a valuable source of intuition [OV00, GLR20]. Wasserstein distances, which are defined through the optimal transport problem, provide a robust way to compare two probability measures and are, therefore, a valuable tool in statistics and related fields [PC19].

Beyond the examples mentioned in the above paragraph, we highlight the interplay of this field with the theory of geometric and functional inequalities since this interaction is the topic of some contributions of this thesis. The optimal transport theory helps to get new inequalities [CEMS01, FMS24], new proofs of already known inequalities [CE02, CENV04], and reinforced versions or stability-type results [FMP09, FMP10, FIL16]. We will deeply elaborate on this point later on.

In his seminal paper [Mon81], Monge formulated the optimal transport problem as follows: Suppose we have a deposit of material that, after being mined, must be transported to another destination. Suppose further that the cost of transporting the material is equal to the product of its mass times the distance between its origin and its destination. How must we transport all the resources to minimize the total transportation cost?

We can model Monge's problem in the following way: Let μ and ν be two probability distributions representing the problem's source and target in the problem, respectively. We want to find a map T that pushes forward μ towards ν, which informally means that T is an assignment rule that tells us that the mass lying at the position $x \in \operatorname{supp}(\mu)$ must be transported to a single point $y = T(x) \in \operatorname{supp}(\nu)$ without splitting it to other positions, see Figure 1.1. In addition, we want this map T to minimize the aggregated transport cost, considering that the cost of transporting one unit of mass from the position x to y is given by the usual Euclidean distance, $(x, y) \mapsto |x - y|$. Mathematically, we can model this situation by the following variational problem:

$$\inf_{T_\# \mu = \nu} \int |x - T(x)| \, \mathrm{d}\mu(x), \tag{1.1.1}$$

where the infimum is taken among all maps T that push forward μ towards ν. Unfortunately, this is a hard problem to solve because of the rigidity of the constraint $T_\# \mu = \nu$, an issue that raises the following questions: is it possible to relax the push-forward constraint in a meaningful sense to ensure the existence of a solution? Does the situation improve if we change the L^1-type cost function in (1.1.1), that is, are there any situations where we can ensure the existence of a solution T? If there is a solution, does it have any particular structure?

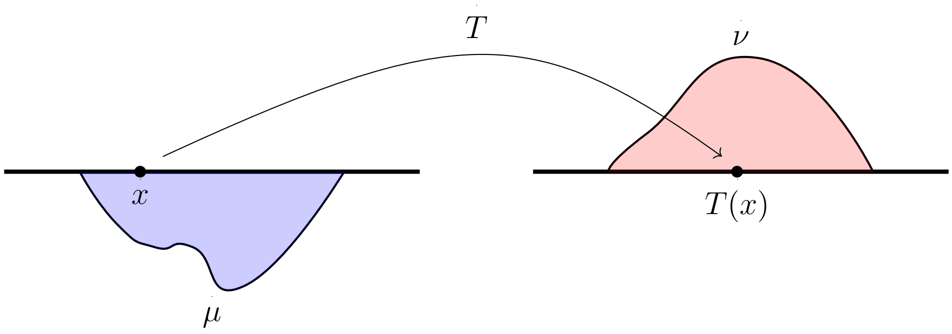

Figure 1.1: Monge's problem between μ and ν.[1]

After this informal discussion, in the next section, we will start reviewing the basic theory of optimal transport, which provides a rich and sufficiently general framework that satisfactorily addresses the previously mentioned questions.

1.2 General theory

Our first step is to generalize the formulation of Monge's problem given in (1.1.1) for an arbitrary cost function defined on the product of two Polish spaces. We chose this level of generality because these spaces provide a convenient setting for doing measure theory.

Definition 1.1 (Optimal transport, Monge's formulation)**.** Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty Polish spaces, and let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. Let $c\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be a nonnegative Borel function. The Monge formulation of the optimal transport problem associated with the measures μ and ν, and the cost function c is the following minimization problem:

$$\inf_{\substack{T\colon \mathcal{X}\to\mathcal{Y};\\ T_{\#}\mu=\nu}} \int_{\mathcal{X}} c(x, T(x))\, \mathrm{d}\mu(x), \tag{OT Monge}$$

where the infimum is taken among all Borel maps $T\colon \mathcal{X} \to \mathcal{Y}$ pushing forward μ towards ν. A minimizer is called a Monge map or an optimal transport map, and we will generally denote it as T_{Mon} or T_0, depending on the context.

In dimension one, Monge's problem is fully understood for strictly convex cost functions. A unique solution exists under mild conditions, and we have access to a very simple closed-form expression for the optimal map depending only on the original data μ and ν.

Proposition 1.2 (One-dimensional Monge's problem)**.** *Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, and let $\mu \in \mathcal{P}(\mathbb{R})$ and $\nu \in \mathcal{P}(\mathbb{R})$ be two absolutely continuous probability measures on $\mathbb{R}$ with full supports and finite second moments. If the cost function is of the form $c(x, y) = h(x-y)$, where $h\colon \mathbb{R} \to \mathbb{R}_+$ is strictly convex, then the associated problem* (OT Monge) *has a unique solution given by the map $T_{\mathrm{Mon}}\colon \mathbb{R} \to \mathbb{R}$ defined by*

$$\forall x \in \mathbb{R}, \quad T_{\mathrm{Mon}}(x) = F_\nu^{-1}(F_\mu(x)).$$

[1] Ce ne sont pas des serpents boas qui digèrent des éléphants.

Monge's problem is also well understood in the Gaussian case under quadratic cost; see, for example, [Gel90].

Proposition 1.3 (Monge's problem for Gaussian measures). *Suppose that $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$, and let $\mu = \mathcal{N}(0, A)$ and $\nu = \mathcal{N}(0, B)$ be two non-degenerate Gaussian measures with $A, B \succ 0$. If the cost function is given by the quadratic cost $c(x, y) = \frac{1}{2}|x - y|^2$, then the problem* (OT Monge) *has a unique solution given by the linear map $T_{\mathrm{Mon}} \colon \mathbb{R}^d \to \mathbb{R}^d$ defined by*

$$\forall x \in \mathbb{R}^d, \quad T_{\mathrm{Mon}}(x) = A^{-\frac{1}{2}}(A^{\frac{1}{2}} B A^{\frac{1}{2}})^{\frac{1}{2}} A^{-\frac{1}{2}} x.$$

Unfortunately, the two situations described in Propositions 1.2 and 1.3 are quite exceptional: these are essentially the only cases where we know closed-form expressions for the optimal transport map.[2] Now, if we are merely inquiring about the existence of a solution, we can say that Monge's problem is challenging to solve, even in very uncomplicated situations. The following examples will illustrate this fact.

Example 1.4 (Transporting a Dirac Mass). For $\mathcal{X}$ and $\mathcal{Y}$ Polish and nonempty, let $x \in \mathcal{X}$, and take $\mu = \delta_x$. For any Borel map $T \colon \mathcal{X} \to \mathcal{Y}$, we have that $T_{\#}\mu = \delta_{T(x)}$, so Monge's problem has no solution unless ν is a Dirac mass.

Example 1.5 (Transport in the two-point space). Let $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, and take $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ and $\nu = \frac{1}{4}\delta_0 + \frac{3}{4}\delta_1$. We claim that there are no maps $T \colon \mathcal{X} \to \mathcal{Y}$ such that $T_{\#}\mu = \nu$. Indeed, if we had a map $T \colon \mathcal{X} \to \mathcal{Y}$ pushing forward μ towards ν, we would have $\nu(\{i\}) = \mu(\{S^{-1}(\{i\}))$ for $i \in \{0, 1\}$, which is impossible by the construction of both measures.

Both examples 1.4 and 1.5 show that the condition $T_{\#}\mu = \nu$ is very restrictive. We can relax Monge's problem by minimizing over a larger set of constraints that are not as rigid as the pushforward condition, namely the set of transport plans between μ and ν.

Definition 1.6 (Transport plan). Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty Polish spaces. If we fix two Borel probability measures $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$, we define $\Pi(\mu, \nu)$, the associated set of transport plans between μ and ν, by

$$\Pi(\mu, \nu) := \{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : (\mathrm{proj}_{\mathcal{X}})_{\#}\pi = \mu \text{ and } (\mathrm{proj}_{\mathcal{Y}})_{\#}\pi = \nu\},$$

where $\mathrm{proj}_{\mathcal{X}} \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ and $\mathrm{proj}_{\mathcal{Y}} \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Y}$ denote the canonical projections.

Remark 1.7.

(i) If $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$, note that $\pi \in \Pi(\mu, \nu)$ if and only if

$$\forall f \in \mathbb{B}(\mathcal{X}), \forall g \in \mathbb{B}(\mathcal{Y}), \int_{\mathcal{X} \times \mathcal{Y}} (f \oplus g)(x, y) \,\mathrm{d}\pi(x, y) = \int_{\mathcal{X}} f(x) \,\mathrm{d}\mu(x) + \int_{\mathcal{Y}} g(y) \,\mathrm{d}\nu(y).$$

(ii) We observe that $\mu \otimes \nu$ is a transport plan between μ and ν, so $\Pi(\mu, \nu)$ is always nonempty.

[2] Another example where a closed-form expression is available for the solution of the optimal transport problem, yet in its Kantorovich formulation, which will be introduced later in Definition 1.8, is when $\mathcal{X} = \mathcal{Y} = T$, where $T = (V, E, w)$ is a (possibly countable infinite) metric tree [MPV23].

We are ready to state the Kantorovich formulation of the optimal transport problem, a relaxation of problem (OT Monge), which minimizes the transport cost among the set of transport plans.

Definition 1.8 (Optimal transport, Kantorovich's formulation)**.** Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty Polish spaces, and let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. Let $c\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be a nonnegative Borel function. The Kantorovich formulation of the optimal transport problem associated with the measures μ and ν, and the cost function c is the following minimization problem:

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\,\mathrm{d}\pi(x,y). \tag{OT}$$

Remark 1.9. We can observe that the problem (OT) is a relaxation of (OT Monge): Indeed, if $T\colon \mathcal{X} \to \mathcal{Y}$ is a Borel map and if we consider the identity map in $\mathcal{X}$, $\mathrm{id}_{\mathcal{X}}\colon \mathcal{X} \to \mathcal{X}$, we can define the measure $\pi_T \coloneqq (\mathrm{id}_{\mathcal{X}}, T)_{\#}\mu$ on $\mathcal{X} \times \mathcal{Y}$, and we say that π_T is the transport plan induced by the map T. If $T_{\#}\mu = \nu$, then $\pi_T \in \Pi(\mu,\nu)$. On the other hand, we have that

$$\int_{\mathcal{X}} c(x, T(x))\,\mathrm{d}\mu(x) = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\,\mathrm{d}\pi_T(x,y),$$

so the value of (OT) will always be less than or equal to the value of the problem (OT Monge).

Remark 1.10 (Couplings and the probabilistic formulation)**.** In probabilistic language, if we have two Borel distributions $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$ on some nonempty Polish spaces $\mathcal{X}$ and $\mathcal{Y}$, we say that a coupling between μ and ν is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with two random variables $X\colon \Omega \to \mathcal{X}$ and $Y\colon \Omega \to \mathcal{Y}$ such that they have laws μ and ν, respectively. Then we have that the problem (OT) is equivalent to the following one, called the probabilistic formulation of the optimal transport problem:

$$\inf_{\substack{(X,Y)\text{ coupling}\\ X\sim\mu,\, Y\sim\nu}} \mathbb{E}[c(X,Y)],$$

where the infimum is taken among all the possible couplings $(\Omega, \mathcal{F}, \mathbb{P}, X, Y)$ between μ and ν, and $\mathbb{E}$ is the expectation operator associated to the measure $\mathbb{P}$. Indeed, if $(\Omega, \mathcal{F}, \mathbb{P}, X, Y)$ is a coupling, we can define the measurable map $S\colon \Omega \to \mathcal{X} \times \mathcal{Y}$ given by $\omega \mapsto (X(\omega), Y(\omega))$. Then, if we take $\pi = S_{\#}\mathbb{P}$, it is easy to check that it belongs to $\Pi(\mu,\nu)$. In other words, couplings always factor through the product space $\mathcal{X} \times \mathcal{Y}$, so the following diagram commutes:

$$\begin{tikzcd}
 & (\Omega, \mathcal{F}, \mathbb{P}) \arrow[dl, "X"'] \arrow[d, "S"] \arrow[dr, "Y"] & \\
(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu) & (\mathcal{X}\times\mathcal{Y}, \mathcal{B}(\mathcal{X})\otimes\mathcal{B}(\mathcal{Y}), \pi) \arrow[l, "\mathrm{proj}_{\mathcal{X}}"] \arrow[r, "\mathrm{proj}_{\mathcal{Y}}"'] & (\mathcal{Y}, \mathcal{B}(\mathcal{Y}), \nu).
\end{tikzcd}$$

On the other hand,

$$\mathbb{E}[c(X,Y)] = \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\,\mathrm{d}\pi(x,y);$$

that is, the formulation (OT) coincides with the probabilistic one. In the particular case that there exists a map $T\colon \mathcal{X} \to \mathcal{Y}$ such that $T_{\#}\mu = \nu$, we say that the coupling induced by T is deterministic: for any random variable $X \sim \mu$, then $T(X) \sim \nu$.

The set $\Pi(\mu, \nu)$ has remarkable topological properties. First, we observe that it is a closed set within $\mathcal{M}(\mathcal{X} \times \mathcal{Y})$ for the weak topology. Furthermore, it is a tight family of measures, which, thanks to Prokhorov's theorem, implies that it is sequentially compact. We recall the definition of tightness and Prokhorov's theorem; see, for example, [Bil99, Chapter 1, Section 5] for an extensive reference on the topic.

Definition 1.11 (Tightness). Let $\mathcal{X}$ be a nonempty Polish space, and let $M \subset \mathcal{P}(\mathcal{X})$ be a collection of Borel probability measures. We say that M is tight if for every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \mathcal{X}$ such that

$$\forall \mu \in M, \quad \mu(K_\varepsilon) \geqslant 1 - \varepsilon.$$

Theorem 1.12 (Prokhorov). *Let $\mathcal{X}$ be a nonempty Polish space, and let $M \subset \mathcal{P}(\mathcal{X})$ be a collection of Borel probability measures. Then M is compact in the weak topology of $\mathcal{M}(\mathcal{X})$ if and only if it is closed and tight.*

Proposition 1.13. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty Polish spaces, and fix two Borel probability measures $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. Then the set of transport plans $\Pi(\mu, \nu)$ is tight and thus compact.*

Recalling the classical Weirstraß criterion, we know that if we minimize a lower semicontinuous function on a compact set, then we have the existence of at least one minimizer. On the other hand, we can show that if we assume that the cost function c is lower semicontinuous, then the map $\pi \mapsto \int_{\mathcal{X}\times\mathcal{Y}} c \,\mathrm{d}\pi$ will be lower semicontinuous. Summing up all the above facts, we conclude that the problem (OT) admits a solution under very reasonable conditions.

Theorem 1.14. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty Polish spaces, and let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. Let $c\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be a nonnegative lower semicontinuous function. Then there exists $\pi_0 \in \Pi(\mu, \nu)$ solving* (OT). *Furthermore, if we assume that $c \in L^1(\mu \otimes \nu)$, then the value of* (OT) *is also finite.*

Now, we exhibit some concrete choices of c that will be useful in the sequel.

Example 1.15 (Total variation distance). Let $\mathcal{X}$ be a nonempty Polish space, and define its diagonal $\Delta \subset \mathcal{X} \times \mathcal{X}$ by $\Delta \coloneqq \{(x, x) : x \in \mathcal{X}\}$. Let $c\colon \mathcal{X} \times \mathcal{X} \to \{0, 1\}$ be the Hamming cost, which is given by

$$\forall (x, y) \in \mathcal{X} \times \mathcal{X}, \quad c(x, y) \coloneqq \mathbb{1}_{\mathcal{X}\times\mathcal{X}\setminus\Delta}(x, y).$$

For any fixed marginals $\mu, \nu \in \mathcal{P}(\mathcal{X})$, in the light of Remark 1.10, we can write their associated optimal transport problem with cost function c as

$$\inf_{\substack{(X,Y)\text{ coupling}\\ X\sim\mu, Y\sim\nu}} \mathbb{E}[c(X, Y)] = \inf_{\substack{(X,Y)\text{ coupling}\\ X\sim\mu, Y\sim\nu}} \mathbb{P}(X \neq Y) = \frac{1}{2}\|\mu - \nu\|_{\mathrm{TV}},$$

where $\|\mu - \nu\|_{\mathrm{TV}}$ denotes the total variation distance between μ and ν. That is, we can formulate this distance in the language of optimal transport.

Example 1.16 (p-distances)**.** Let $(\mathcal{X}, d)$ be a nonempty complete and separable metric space. For $p \geqslant 1$, we can take the cost function given by the p-distance associated to d:

$$\forall (x, y) \in \mathcal{X} \times \mathcal{X}, \quad c(x, y) := d(x, y)^p.$$

If μ and ν are Borel probability measures on $\mathcal{X}$ with finite moments of order p, then $c \in L^1(\mu \otimes \nu)$. Furthermore, every transport plan has finite cost: let $\pi \in \Pi(\mu, \nu)$, and let $z_0 \in \mathcal{X}$:

$$\begin{aligned}\int_{\mathcal{X}\times\mathcal{X}} d(x, y)^p \, \mathrm{d}\pi(x, y) &\leqslant 2^{p-1} \int_{\mathcal{X}\times\mathcal{X}} (d(x, z_0)^p + d(y, z_0)^p) \, \mathrm{d}\pi(x, y) \\ &= 2^{p-1} \int_{\mathcal{X}} d(x, z_0)^p \, \mathrm{d}\mu(x) + 2^{p-1} \int_{\mathcal{X}} d(y, z_0)^p \, \mathrm{d}\nu(y) < +\infty,\end{aligned}$$

so Theorem 1.14 grants a finite value for the Kantorovich problem associated to μ and ν under the cost c.

Example 1.16 motivates the definition of the p-Wasserstein distances on $\mathcal{P}_p(\mathcal{X})$, the set of Borel probability measures on $\mathcal{X}$ with finite moments of order p.

Definition 1.17 (p-Wasserstein distances)**.** Let $(\mathcal{X}, d)$ be a nonempty complete and separable metric space, and let $p \geqslant 1$. For any $\mu, \nu \in \mathcal{P}_p(\mathcal{X})$, we define the p-Wasserstein distance between μ and ν by

$$\mathrm{W}_p(\mu, \nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} d(x, y)^p \, \mathrm{d}\pi(x, y) \right)^{\frac{1}{p}}.$$

Remark 1.18.

(i) W_p is a metric on $\mathcal{P}_p(\mathcal{X})$. Furthermore, $(\mathcal{P}_p(\mathcal{X}), \mathrm{W}_p)$ is a Polish space.

(ii) For any $p \geqslant 1$, the topology induced by the metric W_p is finer than the weak one. More precisely, convergence in the p-Wasserstein distance can be characterized as follows: let $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}_p(\mathcal{X})$ and $\mu \in \mathcal{P}_p(\mathcal{X})$. Then

$$\mathrm{W}_p(\mu_n, \mu) \to 0 \iff \begin{cases} \mu_n \to \mu \text{ weakly, and} \\ \forall x_0 \in \mathcal{X}, \displaystyle\int_{\mathcal{X}} d(x, x_0)^p \, \mathrm{d}\mu_n(x) \to \int_{\mathcal{X}} d(x, x_0)^p \, \mathrm{d}\mu(x). \end{cases}$$

(iii) As a consequence of the previous item, if $\mathcal{X}$ is compact, then for any $p \geqslant 1$, the convergence in p-Wasserstein distance characterizes the weak convergence, since for any $x_0 \in \mathcal{X}$, the function $x \mapsto d(x, x_0)^p$ is bounded and continuous.

Given an optimization problem, we are almost always interested in its dual formulation, hoping to recover the same amount of information contained in the original problem; in other words, we aim to obtain strong duality results. For problem (OT), such a result holds under mild hypotheses, and the following theorem due to Kantorovich [Kan42] accounts for this fact.

Theorem 1.19 (Kantorovich duality). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty Polish spaces, and let $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$. Let $c\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be a nonnegative lower semicontinuous function and assume that $c \in L^1(\mu \otimes \nu)$. Define*

$$\Phi_c := \left\{(f, g) \in L^1(\mu) \times L^1(\nu) : f \oplus g \leqslant c\right\}. \tag{1.2.1}$$

Then

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c \,\mathrm{d}\pi = \sup_{(f,g)\in\Phi_c} \left(\int_{\mathcal{X}} f \,\mathrm{d}\mu + \int_{\mathcal{Y}} g \,\mathrm{d}\nu \right), \tag{1.2.2}$$

and both values are finite. If there exists a pair $(f_0, g_0) \in \Phi_c$ attaining the supremum on the right-hand side of (1.2.2), we call it a pair of Kantorovich potentials.

Remark 1.20. For any pair of Kantorovich potentials (f_0, g_0), and any $a \in \mathbb{R}$, we observe that $(f_0 + a, g_0 - a)$ is also a pair of Kantorovich potentials. However, under some technical assumptions and if they exist, the potentials are unique if we impose a normalization criterion; see [BGN22, Appendix B]. In practice, two widely used normalization rules are $\int_{\mathcal{X}} f_0 \,\mathrm{d}\mu = \int_{\mathcal{Y}} g_0 \,\mathrm{d}\nu$ or $\int_X f_0 \,\mathrm{d}\mu = 0$.

1.3 The quadratic Euclidean case

From this point onward, we will focus our attention on the quadratic Euclidean case; that is, unless we state otherwise, we will assume that $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ and $c(x, y) = \frac{1}{2}|x - y|^2$.

Definition 1.21 (Quadratic Euclidean optimal transport). Given two Borel probability measures μ and ν on $\mathbb{R}^d$, we define their associated quadratic Euclidean optimal transport problem, in its Kantorovich formulation, as

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x - y|^2 \,\mathrm{d}\pi(x, y), \tag{1.3.1}$$

and we denote the value of (1.3.1) by $\mathcal{C}_0(\mu, \nu)$.

1.3.1 Structure of the solution: the Brenier map

The quadratic Euclidean setting represents a very convenient framework for the optimal transport problem since under mild assumptions on the first marginal, the Kantorovich problem (1.3.1) has a unique solution π_0 with a specific structure, where convexity plays a key role: the transport plan π_0 is induced by a pushforward map T_0 that solves the associated Monge problem. In addition, T_0 is the gradient of a convex function φ_0 that plays an essential role in the dual problem to (1.3.1). These assertions are the content of the Brenier-McCann theorem [Bre91, McC95].

Theorem 1.22 (Brenier-McCann). *Let μ and ν be two Borel probability measures on $\mathbb{R}^d$, and assume that μ is absolutely continuous with respect to the d-dimensional Lebesgue measure. Then there exists a lower semicontinuous proper convex function $\varphi_0\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that the map $T_0 := \nabla\varphi_0$ pushes forward μ towards ν. Furthermore, T_0 is the unique gradient of a convex function sending μ to ν.*

Additionally, assume now that μ and ν have finite moments of order 2. Then $\mathcal{C}_0(\mu,\nu) < +\infty$, and there exists a unique optimal plan $\pi_0 \in \Pi(\mu,\nu)$ for (1.3.1)*, given by the transport plan induced by T_0: $\pi_0 = (\mathrm{id}_{\mathbb{R}^d}, T_0)_{\#}\mu$.*

Moreover, if we define $\psi_0 := \varphi_0^*$ as the convex conjugate of φ_0, then $(f_0, g_0) := \left(\frac{1}{2}|\cdot|^2 - \varphi_0, \frac{1}{2}|\cdot|^2 - \psi_0\right)$ is a pair of Kantorovich potentials, that is, it solves the dual problem of (1.3.1)*, which is given by*

$$\mathcal{C}_0(\mu,\nu) = \sup_{(f,g)\in\Phi_2} \int_{\mathbb{R}^d} f\,\mathrm{d}\mu + \int_{\mathbb{R}^d} g\,\mathrm{d}\nu, \tag{1.3.2}$$

where Φ_2 is the set

$$\Phi_2 := \left\{(f,g) \in L^1(\mu) \times L^1(\nu) : f \oplus g \leqslant \frac{1}{2}|\cdot - \cdot|^2\right\}.$$

We give some remarks on Theorem 1.22.

Remark 1.23.

(i) The map T_0 is also the unique solution to the associated Monge problem. In the following, we will call it the Brenier, Monge, or optimal transport map, depending on the context.

(ii) The second-moment assumption on μ and ν ensures that $\mathcal{C}_0(\mu,\nu) < +\infty$, it is not necessary for the existence of the map T_0, as we see in the first part of the theorem.

(iii) If we additionally assume that ν is absolutely continuous with respect to Leb, then the map T_0 has an inverse that pushes ν towards μ and solves the inverse Monge problem, i.e., from ν towards μ. Moreover, $T_0^{-1} = \nabla\psi_0$.

(iv) We can relax the absolute continuity of μ: it just suffices to assume that μ vanishes on all Borel sets of Hausdorff dimension $d-1$ [McC95]. The result holds even under a weaker condition [GM96], which was shown to be sharp in [Gig11b].

(v) Theorem 1.22 can be generalized to the Riemannian setting; see [McC01].

1.3.2 Regularity of the Brenier map

Theorem 1.22 is a strong result since it ensures the existence of an optimal transport map under mild assumptions on the first marginal, which is very practical for some applications. Nevertheless, in some cases, the sole existence of the Brenier map is not enough; some additional regularity properties may be required.

For example, in Chapter 2, we will see that given two measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, if μ satisfies certain types of functional inequalities (e.g., concentration bounds, Poincaré, logarithmic Sobolev, etc.), and if there exists a Lipschitz regular map pushing forward μ towards ν, then the same inequalities are inherited by ν. Hence, it is natural to inquire about the regularity of the Brenier map in terms of the inputs μ and ν.

One of the main tools to address the regularity of the optimal transport map is the analysis of its associated Monge-Ampère equation, using techniques from the theory of partial differential equations, which we will introduce. Let μ and ν be two absolutely continuous measures on $\mathbb{R}^d$ with full support, and let $T_0 \colon \mathbb{R}^d \to \mathbb{R}^d$ be the associated Brenier map. Since $(T_0)_{\#}\mu = \nu$, we have that for any bounded Borel function $\psi \colon \mathbb{R}^d \to \mathbb{R}$,

$$\int_{\mathbb{R}^d} \psi(T_0(x))\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} \psi(y)\,\mathrm{d}\nu(y).$$

Both μ and ν are absolutely continuous and have full support; denote their positive densities by f and g, respectively, and rewrite the last equation as

$$\int_{\mathbb{R}^d} \psi(T_0(x)) f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \psi(y) g(y)\,\mathrm{d}y.$$

Let us assume that the map T_0 is nice enough: suppose it is a smooth diffeomorphism. By the classical change of variables theorem, we can rewrite the integral on the right-hand side in terms of T_0 so that we get

$$\int_{\mathbb{R}^d} \psi(T_0(x)) f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \psi(T_0(x)) g(T_0(x)) |\det(\nabla T_0(x))|\,\mathrm{d}x.$$

If we assume that f and g are smooth, since ψ was arbitrary, we deduce that

$$\forall x \in \mathbb{R}^d, \quad |\det(\nabla T_0(x))| = \frac{f(x)}{g(T_0(x))},$$

which is a first-order PDE. So far, we have not used the fact that T_0 is the Brenier map. By Theorem 1.22, there exists a convex function φ_0 such that $T_0 = \nabla\varphi_0$, so we get the following equation:

$$\forall x \in \mathbb{R}^d, \quad \det(\nabla^2 \varphi_0(x)) = \frac{f(x)}{g(\nabla\varphi_0(x))}. \tag{1.3.3}$$

Equation (1.3.3) is the Monge-Ampère equation, which is a second-order nonlinear partial differential equation.

In an ex-post analysis, let us note that to derive the Monge-Ampère equation, we assumed that the Brenier map was a diffeomorphism. Fortunately, this is not an issue since McCann [McC97] proved that (1.3.3) holds Lebesgue-almost everywhere under mild assumptions on the marginals.

Regularity theory for (1.3.3) is a very active subject; the reader may consult [Fig19] for a more detailed exposition of the subject. On the qualitative side, Caffarelli's contributions [Caf92b, Caf92a, Caf96] provide a positive answer concerning the regularity of the Brenier map. The following result condensates his results. Before stating the theorem, recall that if $\Omega \subset \mathbb{R}^d$, we say that $f \colon \Omega \to \mathbb{R}$ belongs to $\mathcal{C}^{k,\alpha}(\Omega)$ if there exist $k \in \mathbb{N}^*$ and $\alpha \in (0,1]$ such that f is k times differentiable and its kth derivative is α-Hölder.

Theorem 1.24 (Caffarelli's qualitative regularity theorem). *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^d$ be two connected and bounded open sets, let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ be absolutely continuous with respect to* Leb *with $\mathrm{d}\mu(x) = f(x)\,\mathrm{d}x$ and $\mathrm{d}\nu(y) = g(y)\,\mathrm{d}y$, and suppose that both f and g are bounded away from zero and infinity and zero outside of Ω_1 and Ω_2, respectively. Let $T_0 = \nabla\varphi_0 \colon \Omega_1 \to \Omega_2$ be the Brenier map pushing forward μ towards ν and assume that Ω_2 is convex. If there exist $k \in \mathbb{N}^*$ and $\alpha \in (0,1]$ such that $f \in \mathcal{C}^{k,\alpha}(\Omega_1)$ and $g \in \mathcal{C}^{k,\alpha}(\Omega_2)$, then $\varphi_0 \in \mathcal{C}^{k+2,\alpha}(\Omega_1)$.*

Remark 1.25. In Theorem 1.24, the convexity assumption on Ω_2 is necessary [Vil09, Theorem 12.3]. Moreover, the assumption needed on the domain Ω_2 is of a geometric nature, but not of a topological or regularity nature [Caf92b, p. 100].

On the quantitative side of the theory, one searches for explicit estimates for the derivatives of the Brenier map. A very remarkable result, again by Caffarelli [Caf00], provides 1-Lipschitz regularity for the Brenier map pushing forward the d-dimensional Gaussian measure towards any log-concave perturbation of it.

Theorem 1.26 (Caffarelli's contraction theorem). *Let γ_d be the d-dimensional Gaussian measure on $\mathbb{R}^d$. Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ with $\nu \ll \gamma_d$, and suppose that it has the form $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\gamma_d$, where $V \colon \mathbb{R}^d \to \mathbb{R}$. If V is convex, the Brenier map pushing forward γ_d towards ν is 1-Lipschitz.*

Theorem 1.26 can be refined and generalized; see Theorem 6.5 in Chapter 6. Originally, Theorem 1.26 was proved by Caffarelli using a maximum principle-like argument for (1.3.3). Nevertheless, we can find simpler proofs based on the entropic regularization of optimal transport in [FGP20, CP23].

Caffarelli's contraction theorem is of vital importance in the theory of functional inequalities, as it allows the transport of the many functional inequalities satisfied by the Gaussian measure towards log-concave perturbations of it. This consequence is the starting point and inspiration for some results and directions developed in this thesis. We will elaborate on this point later in Chapter 2.

1.3.3 About the continuity equation

In this subsection, we will make a slight detour in the exposition to review the method of characteristics for the continuity equation, which will be crucial for the following subsection and Chapter 3.

Let (M, g) be a smooth Riemannian manifold that we assume to be complete and connected. In particular, the following results are valid for the (flat) Euclidean space $\mathbb{R}^d$. Let $A_\bullet \colon [0, +\infty) \times M \to TM$ be a time-dependent vector field on M. Given a fixed probability measure ν on M, we are interested in solving the continuity equation on M with velocity $A_\bullet$ and initial condition ν:

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t A_t) = 0 \\ \rho_0 = \nu, \end{cases} \tag{1.3.4}$$

where $\nabla\cdot$ denotes the divergence operator on (M, g), which acts on a vector field

$Z\colon M \to TM$ by

$$\forall \varphi \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \int_M \varphi \, \nabla \cdot Z \, \mathrm{dVol} = -\int_M \nabla \varphi \cdot Z \, \mathrm{dVol}.$$

We will be interested in measure-valued solutions to (1.3.4). More precisely, we say that $(\rho_t)_{t \geqslant 0}$, a sequence of probability measures on M with $\rho_0 = \nu$, satisfies the continuity equation (1.3.4) in the distributional sense if

$$\forall \varphi \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \frac{\mathrm{d}}{\mathrm{d}t} \int_M \varphi \, \mathrm{d}\rho_t = \int_M \nabla \varphi \cdot A_t \, \mathrm{d}\rho_t.$$

Now we introduce the flow of diffeomorphisms $(S_t)_{t \geqslant 0}$ induced by the vector field $A_\bullet$, which is given by

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} S_t(x) = A_t(S_t(x)) \\ \quad S_0(x) = x. \end{cases} \tag{1.3.5}$$

It is well defined if $A_\bullet$ is at least locally Lipschitz.

The following classical result summarizes the method of characteristics for the continuity equation, characterizing the weak solutions of the system (1.3.4) in terms of its associated flow $(S_t)_{t \geqslant 0}$ defined by (1.3.5).

Theorem 1.27. *Let (M, g) be a complete and connected Riemannian manifold, and let $A_\bullet\colon [0, +\infty) \times M \to TM$ be a locally Lipschitz time-dependent vector field on M. Let ν be a Borel probability measure on M, and let $(\rho_t)_{t \geqslant 0}$ be a flow of Borel probability measures on M continuous on $[0, +\infty)$, with $\rho_0 = \nu$, and such that*

$$\int_0^{+\infty} \int_M |A_t(x)| \, \mathrm{d}\rho_t \, \mathrm{d}t < +\infty. \tag{1.3.6}$$

Then $(\rho_t)_{t \geqslant 0}$ is a distributional solution of the continuity equation

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t A_t) = 0 \\ \qquad\qquad\quad \rho_0 = \nu \end{cases}$$

if and only if $\rho_t = S_{t\#}\nu$, where $(S_t)_{t \geqslant 0}$ is the flow of diffeomorphisms associated to $A_\bullet$, which is defined by (1.3.5).

1.3.4 Lagrangian and Eulerian formulations of optimal transport

We will keep the following discussion at an informal level, as it is not essential for the rest of the thesis. Let μ and ν be two absolutely continuous measures on $\mathbb{R}^d$ with finite second moments. By Theorem 1.22, we can express the squared 2-Wasserstein distance between μ and ν as

$$\mathrm{W}_2^2(\mu, \nu) = \inf \left\{ \int_{\mathbb{R}^d} |x - S(x)|^2 \, \mathrm{d}\mu(x) : S_{\#}\mu = \nu \right\}, \tag{1.3.7}$$

where the infimum is attained for the Monge map $T_{\mathrm{Mon}} = \nabla\varphi_0$. Now let us look at the following problem, which is called the Lagrangian time-dependent formulation of the Monge problem:

$$\inf\left\{\int_0^1 \int_{\mathbb{R}^d} |S_t(x)|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}t : S_0 = \mathrm{id}_{\mathbb{R}^d}, (S_1)_\#\mu = \nu\right\}, \tag{1.3.8}$$

where the infimum is taken among flows $(S_t)_{t\in[0,1]}$ of maps $S_t\colon \mathbb{R}^d \to \mathbb{R}^d$. One can show that (1.3.8) is compatible with the static formulation (1.3.7), in the sense that if $(S_t)_{t\in[0,1]}$ is optimal for (1.3.8), then the map S_1 is optimal for (1.3.7), which is equivalent to saying that $S_1 = T_{\mathrm{Mon}}$. One may wonder about the converse: does the map T_{Mon} induce a flow of maps $(T_t)_{t\in[0,1]}$ optimal for (1.3.8)? Indeed, if we define

$$\forall t \in [0,1], \quad T_t \coloneqq (1-t)\,\mathrm{id}_{\mathbb{R}^d} + t\,T_{\mathrm{Mon}}, \tag{1.3.9}$$

then $(T_t)_{t\in[0,1]}$ solves (1.3.8). We call the flow $(T_t)_{t\in[0,1]}$ the displacement interpolation between μ and ν.

Now consider the flow of measures $(\rho_t)_{t\in[0,1]}$ defined by

$$\forall t \in [0,1], \quad \rho_t \coloneqq (T_t)_\#\mu. \tag{1.3.10}$$

Note that $\rho_0 = \mu$ and $\rho_1 = \nu$, that is, $(\rho_t)_{t\in[0,1]}$ interpolates between μ and ν. Observe also that $T_t = \nabla\left[(1-t)\frac{1}{2}|\cdot|^2 + t\,\varphi_0\right]$; i.e., it is the gradient of a convex function, so it must be the optimal transport map that pushes μ toward ρ_t. In particular,

$$\begin{aligned}\mathrm{W}_2^2(\mu,\rho_t) &= \int_{\mathbb{R}^d} |x - T_t(x)|^2 \,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} |x - [(1-t)x - t\nabla\varphi_0(x)]|^2 \,\mathrm{d}\mu(x)\\ &= t^2 \int_{\mathbb{R}^d} |x - \nabla\varphi_0(x)|^2 \,\mathrm{d}\mu(x) = t^2\,\mathrm{W}_2^2(\mu,\nu).\end{aligned}$$

Furthermore, one can show that

$$\forall s,t \in [0,1], \quad \mathrm{W}_2(\rho_s,\rho_t) = |s-t|\,\mathrm{W}_2(\mu,\nu),$$

that is, W_2 is a geodesic distance.

By Theorem 1.27, we can transform the Lagrangian time-dependent formulation (1.3.8) into its equivalent Eulerian formulation. More precisely, we have that (1.3.10) is equivalent to the validity of the following continuity equation for the flow $(\rho_t)_{t\in[0,1]}$:

$$\begin{cases}\partial_t\rho_t + \nabla\cdot(v_t\rho_t) = 0, \forall t \in (0,1),\\ \rho_0 = \mu,\end{cases}$$

where $(v_t)_{t\in[0,1]}$ is the velocity field given by $v_t\colon \mathbb{R}^d \to \mathbb{R}^d$, $v_t \coloneqq (T_{\mathrm{Mon}} - \mathrm{id}_{\mathbb{R}^d}) \circ T_t^{-1}$. If we combine this with the Lagrangian formulation (1.3.8), we get the Benamou-Brenier formula [BB00], which is the time-dependent Eulerian formulation of optimal transport:

$$\mathrm{W}_2^2(\mu,\nu) = \inf\left\{\int_0^1 \int_{\mathbb{R}^d} |v_t(x)|^2 \,\mathrm{d}\rho_t(x)\,\mathrm{d}t : \partial_t\rho_t + \nabla\cdot(v_t\rho_t) = 0, \rho_0 = \mu, \rho_1 = \nu\right\}, \tag{1.3.11}$$

where the infimum is taken among all pairs (ρ, v), where $\rho = (\rho_t)_{t\in[0,1]}$ and $v = (v_t)_{t\in[0,1]}$ together satisfy the continuity equation.

Given $\rho = (\rho_t)_{t\in[0,1]}$, we write $\dot{\rho}_t \coloneqq \partial_t \rho_t$ and define

$$|\dot{\rho}_t|^2_{\rho_t} \coloneqq \inf \left\{ \int_{\mathbb{R}^d} |v|^2 \,\mathrm{d}\rho_t : v = (v_t)_{t\in[0,1]}, \partial_t \rho_t + \nabla \cdot (v_t \rho_t) = 0 \right\}.$$

Then the Benamou-Brenier formula (1.3.11) can be recast as

$$\mathrm{W}_2^2(\mu, \nu) = \inf \left\{ \int_0^1 |\dot{\rho}_t|^2_{\rho_t} \,\mathrm{d}t : \rho = (\rho_t)_{t\in[0,1]}, \rho_0 = \mu, \rho_1 = \nu \right\},$$

an expression that has a strong Riemannian flavor. It allows us to formally view $\mathcal{P}_2(\mathbb{R}^d)$ as an infinite dimensional manifold equipped with a geodesic distance induced by a "Riemannian" structure: by polarization, we can define an inner product compatible with $|\cdot|_{\rho_t}$ on "the tangent space at ρ_t", $T_{\rho_t}\mathcal{P}_2(\mathbb{R}^d)$, which we represent by all the velocities v_t tangent to ρ_t.

The so-called Otto calculus is the formalism that exploits this "Riemannian structure" on the Wasserstein space, allowing formal computations that provide deep insights and intuition on the structure of the Wasserstein space. Otto introduced it in [Ott01], although it traces its roots back to the seminal contribution [JKO98], where these ideas were used to consider some partial differential equations as gradient flows of some functional defined on Wasserstein space. A good reference on this formalism is [Vil09, Chapter 15]. We provide one of the most important examples: the heat equation corresponds to the gradient flow of the entropy functional in the Otto sense.

Example 1.28 (Gradient flow of the entropy)**.** Consider the functional $\mathcal{F}\colon \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{-\infty, +\infty\}$ given by $\rho \mapsto \mathcal{F}(\rho) \coloneqq \mathrm{H}(\rho)$, where $\mathrm{H}(\rho)$ represents the differential entropy of ρ. Then

$$\nabla_\rho \mathcal{F} = -\Delta\rho,$$

where $\nabla_\rho \mathcal{F}$ is the "Riemannian gradient" of the functional $\mathcal{F}$ in the Otto sense. Thus, the gradient flow associated with the functional $\mathcal{F}$ is exactly the heat equation:

$$\partial_t \rho = \Delta\rho.$$

A similar interpretation connects the Fokker-Planck equation and the relative entropy functional.

Example 1.29 (Gradient flow of the relative entropy)**.** Now fix $\mu \in \mathcal{P}(\mathbb{R}^d)$ of the form $\mathrm{d}\mu = \exp(-W)\,\mathrm{d}x$ for some $W\colon \mathbb{R}^d \to \mathbb{R}$, and consider the functional $\mathcal{G}\colon \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}_+ \cup \{+\infty\}$ given by $\rho \mapsto \mathcal{G}(\rho) \coloneqq \mathrm{H}(\rho|\mu)$, where $\mathrm{H}(\rho|\mu)$ represents the relative entropy of ρ with respect to μ. Then

$$\nabla_\rho \mathcal{G} = -(\Delta f - \nabla W \cdot \nabla f)\nu,$$

where f is the density of ρ with respect to μ. Thus, the gradient flow associated with the functional $\mathcal{G}$ is exactly the Fokker-Planck equation:

$$\partial_t \rho = \Delta\rho + \nabla \cdot (\rho \nabla W).$$

Chapter 2

Functional inequalities

With a little bit of concentration
And a little bit of helping hands, yeah
And a little bit of raving madness, hey
You know it makes me feel, baby
Both my feet are back on the ground.

LED ZEPPELIN
South Bound Saurez

Functional inequalities have proven to be a powerful tool with applications in many areas of mathematics. For example, the concentration properties of a random variable, which play a ubiquitous role in many subfields within probability and statistics, are intimately connected with this theory; even beyond that, concentration phenomena constitute one of the central objects and motivations in the subject, as we will see in this chapter. In the theory of stochastic processes, they represent one of the main tools for studying the long-time behavior of a Markov process and getting quantitative estimates; we will comment briefly on this in the following sections. Their applications go beyond probability theory: the Gaussian logarithmic Sobolev inequality was crucial in Perelman's resolution of the Poincaré conjecture via the Ricci flow [Per02]. A final example, which we previously mentioned in Chapter 1, is the synthetic characterization of Ricci curvature bounds for metric spaces [Stu06a, Stu06b, LV09]; here, the interplay between both theories was a key piece in the puzzle.

The original contributions contained in Chapters 3, 4, and 5 belong to the theory of functional inequalities. Hence, this chapter aims to briefly motivate, introduce, and survey the aspects of the theory of functional inequalities pertinent to this thesis, thus putting the aforementioned new results in a larger context and, at the same time, providing the essential preliminaries for their basic understanding.

We will devote the first three sections of this chapter to studying geometric inequalities, isoperimetric inequalities, and the phenomenon of concentration of measure. After that, we will focus our attention on different families of functional inequalities intimately connected with concentration phenomena: Poincaré, logarithmic Sobolev, and transport-entropy. In these sections, apart from those inequalities mentioned above, there will be other related topics that we will take the opportunity to delve a little deeper into, such as the theory of Markov semigroups, log-concave measures, functional inequalities in the discrete setting, and the Bakry-Émery theory.

2.1 Geometric inequalities

A geometric inequality compares quantities intrinsic to purely geometric objects. For example, for any convex body in an Euclidean space, we can compute its volume, the measure of its boundary, the volume of its polar set, etc. On the other hand, a functional inequality is a relation verified by a large family of functions that compares some observables associated with them, such as the variance, the entropy, and the integrated squared norm of their gradient, all computed with respect to a given measure on the space.

One of the motivations of the theory of functional inequalities is to get functional versions of geometric inequalities that contain the original geometric inequality. As we will see later, the Brunn-Minkowski and Prékopa-Leindler inequalities provide a good example: the latter is a functional version of the former, which is purely geometric. In this case, the embedding $A \mapsto \mathbb{1}_A$ from $\mathcal{B}(\mathbb{R}^d)$ to $\mathbb{B}(\mathbb{R}^d)$ allows recovering the geometric inequality.

Isoperimetric inequalities constitute an important family of geometric inequalities. They characterize, among a large class of sets with equal volume, the ones having minimal perimeter. A first example is the Euclidean isoperimetric inequality, which states that among convex sets with equal volume, the ones with minimal perimeter are balls. Besides, this particular inequality is a direct consequence of the Brunn-Minkowski inequality. Isoperimetric inequalities, which we will review in more detail in Section 2.2, are key in the theory since they are intimately linked with the phenomenon of concentration of measure, one of the main concepts that motivate the study of functional inequalities, as we will see in Section 2.3.

This section is organized as follows: We introduce both the Brunn-Minkowski and the Prékopa-Leindler inequalities to highlight the interplay between geometric inequalities and their functional counterparts. Then, in order to provide a first example of an isoperimetric phenomenon, we state the Euclidean isoperimetric inequality, which is a direct consequence of the Brunn-Minkowski theorem. We used as main references both [Gar02, AAGM15], where the reader can find complete proofs and further development of the results presented here.

2.1.1 Brunn-Minkowski and Prékopa-Leindler

Our starting point is the Brunn-Minkowski theorem [Bru87, Bru89, Min10], which provides an inequality that compares the sum of the volumes of two convex bodies and the volume of their sum. We chose this inequality as a commencement for many reasons beyond its intrinsic importance. First, it admits a functional version, namely the Prékopa-Leindler theorem, which will be our first example of a functional inequality. Second, it admits a proof based on a transportation argument, so from the very beginning we can illustrate the interplay between the theories of optimal transport and functional inequalities. Finally, one of its direct corollaries is the Euclidean isoperimetric inequality, which will be our first example of isoperimetric phenomena, which will be discussed in detail in the next section.

We start by defining what are convex bodies, the essential geometric objects that appear in the Brunn-Minkowski inequality. We say that $K \subset \mathbb{R}^d$ is a convex body if it is a convex and compact set that has nonempty interior. Now, in order to state the inequality, we need to define some basic operations on sets, namely additions and dilations.

Definition 2.1. For any $A, B \subset \mathbb{R}^d$, and any $\lambda, \mu \in \mathbb{R}$, we define the set

$$\lambda A + \mu B \coloneqq \{\lambda x + \mu y : x \in A, y \in B\}.$$

In particular, we define the Minkowski addition between A and B as $A+B \coloneqq 1A+1B$.

Remark 2.2.

(i) We note that if $\lambda, \mu \geqslant 0$ and $A, B \subset \mathbb{R}^d$ are convex sets, then $\lambda A + \mu B$ is also convex.

(ii) In terms of measurability, we have that for any $\lambda, \mu \in \mathbb{R}$, and any $A, B \subset \mathbb{R}^d$ Borel, the set $\lambda A + \mu B$ is Borel.

We are ready to state the Brunn-Minkowski inequality.

Theorem 2.3 (Brunn-Minkowski)**.** *Let $A, B \subset \mathbb{R}^d$ be two convex bodies. Then*

$$\mathrm{Vol}_d(A+B)^{1/d} \geqslant \mathrm{Vol}_d(A)^{1/d} + \mathrm{Vol}_d(B)^{1/d}. \tag{2.1.1}$$

Remark 2.4.

(i) Theorem 2.3 admits some generalizations: inequality (2.1.1) holds for any nonvoid Borel sets A and B. Furthermore, (2.1.1) is also valid if A and B are Lebesgue-measurable and nonempty, but in that case, we have to additionally assume that $A+B$ is Lebesgue-measurable as well.

(ii) Theorem 2.3 is equivalent to the following statement: for any Borel sets $A, B \subset \mathbb{R}^d$, and any $\lambda \in (0,1)$,

$$\mathrm{Vol}_d(\lambda A + (1-\lambda)B) \geqslant \mathrm{Vol}_d(A)^{\lambda}\mathrm{Vol}_d(B)^{1-\lambda}. \tag{2.1.2}$$

(iii) An arbitrary Borel probability measure on $\mathbb{R}^d$ that satisfies inequality (2.1.2) is said to be log-concave. These measures play an important role in the theory of functional inequalities; we will elaborate on this in Section 2.4.4.

Now, as an appetizer, we will exhibit a proof of Theorem 2.3 that relies entirely on optimal transport theory, using the nice properties of the Brenier map between the normalized uniform measures on the sets A and B. The argument we present here is due to Figalli, Maggi, and Pratelli [FMP09]. However, the first who gave a transport proof of the inequality was Knothe [Kno57].

Proof of Theorem 2.3. Define $\mathrm{d}\mu := \frac{1}{\mathrm{Vol}_d(A)}\mathbb{1}_A\,\mathrm{d}x$ and $\mathrm{d}\nu := \frac{1}{\mathrm{Vol}_d(B)}\mathbb{1}_B\,\mathrm{d}x$, that is, μ and ν are the normalized uniform measures on A and B, respectively. Note that both are absolutely continuous probability measures supported A and B, respectively, so Theorem 1.22 in Chapter 1 ensures the existence of $T_0\colon A \to B$, the Brenier map pushing forward μ towards ν, which is of the form $T_0 = \nabla\varphi_0$, for some convex function $\varphi_0\colon A \to \mathbb{R}\cup\{+\infty\}$. Note that $(\mathrm{id}_A + T_0)(A) \subset A + B$, so

$$\mathrm{Vol}_d(A+B) = \int_{\mathbb{R}^d} \mathbb{1}_{A+B}(x)\,\mathrm{d}x \geqslant \int_{\mathbb{R}^d} \mathbb{1}_{(\mathrm{id}_A + T_0)(A)}(x)\,\mathrm{d}x = \int_A |\det(\nabla(\mathrm{id}_A + T_0)(x))|\,\mathrm{d}x,$$

where we used the classical change of variables theorem. Its usage is justified by the fact that T_0 is indeed a diffeomorphism thanks to Caffarelli's qualitative regularity theory; see Theorem 1.24 in Chapter 1.

Let $x \in A$, and let $(\lambda_i(x))_{i=1}^n$ be the eigenvalues of $\nabla^2\varphi_0(x)$, which are all nonnegative. Then note that

$$\begin{aligned}|\det(\nabla(\mathrm{id}_A + T_0)(x))| &= \left|\det\left(I_d + \nabla^2\varphi_0(x)\right)\right| = \prod_{i=1}^d (1+\lambda_i(x)) \\ &\geqslant \left(1 + \left(\prod_{i=1}^d \lambda_i(x)\right)^{1/d}\right)^{1/d},\end{aligned}$$

where we used the arithmetic-geometric means inequality.

Finally, we have that the associated Monge-Ampère equation, see equation (1.3.3) in Chapter 1, holds everywhere since T_0 is regular enough, so for any $x \in A$,

$$\prod_{i=1}^d \lambda_i(x) = \det(\nabla^2\varphi_0(x)) = \frac{\mathrm{Vol}_d(B)}{\mathrm{Vol}_d(A)}.$$

After blending up all the above arguments we obtain

$$\mathrm{Vol}_d(A+B) \geqslant \mathrm{Vol}_d(A)\left(1 + \frac{\mathrm{Vol}_d(B)^{1/d}}{\mathrm{Vol}_d(A)^{1/d}}\right)^d = (\mathrm{Vol}_d(A)^{1/d} + \mathrm{Vol}_d(B)^{1/d})^d,$$

which yields the desired inequality (2.1.1). □

Now we introduce the Prékopa-Leindler inequality [Pré71, Pré73, Lei72], which is a functional version of Theorem 2.3.

Theorem 2.5 (Prékopa-Leindler). *Let $f, g, h\colon \mathbb{R}^d \to \mathbb{R}_+$ be nonnegative Borel functions with f and g Lebesgue-integrable, and let $\lambda \in (0,1)$. Assume that*

$$\forall x, y \in \mathbb{R}^d, \quad h(\lambda x + (1-\lambda)y) \geqslant f(x)^\lambda g(y)^{1-\lambda}. \tag{2.1.3}$$

Then

$$\int_{\mathbb{R}^d} h \, \mathrm{d}x \geqslant \left(\int_{\mathbb{R}^d} f \, \mathrm{d}x \right)^\lambda \left(\int_{\mathbb{R}^d} g \, \mathrm{d}x \right)^{1-\lambda}. \tag{2.1.4}$$

We see that Theorem 2.5 is indeed a functional version of the Brunn-Minkowski inequality, in the sense that it allows its retrieval: indeed, for $A, B \subset \mathbb{R}^d$ Borel, it suffices to take, for a fixed $\lambda \in (0,1)$, $f := \mathbb{1}_A, g := \mathbb{1}_B$ and $h := \mathbb{1}_{\lambda A + (1-\lambda)B}$. This trio of functions verifies (2.1.3), so (2.1.4) yields directly the geometric inequality (2.1.2). The converse is also true: we can deduce the Prékopa-Leindler theorem from the Brunn-Minkowski inequality; see, for example, [DG80].

It is valuable to mention that optimal transport leads as well to a proof of Theorem 2.5, as it was proved by McCann [McC97]: if we identify an absolutely continuous measure ρ on $\mathbb{R}^d$ with its density, we may define the functional $\rho \mapsto -\int_{\mathbb{R}^d} \rho^{(d-1)/d} \, \mathrm{d}x$. McCann proved its convexity along the displacement interpolations (1.3.10) defined in Chapter 1, a fact that allows to recover the Prékopa-Leindler inequality.

2.1.2 Euclidean isoperimetry

One of the most remarkable consequences of the Brunn-Minkowski theorem is the Euclidean isoperimetric inequality, which says that among convex bodies, those with minimal perimeter are balls.

To state the result, we need a notion of perimeter. If $K \subset \mathbb{R}^d$ is a convex body, let ∂K be its boundary. We define the perimeter of K as the $(d-1)$-dimensional volume of ∂K. It is well known that

$$\mathrm{Vol}_{d-1}(\partial K) = \lim_{r \to 0^+} \frac{\mathrm{Vol}_d(K + r\mathrm{B}(0,1)) - \mathrm{Vol}_d(K)}{r}. \tag{2.1.5}$$

Theorem 2.6 (Euclidean isoperimetric inequality). *For any convex body $K \subset \mathbb{R}^d$,*

$$\mathrm{Vol}_{d-1}(\partial K) \geqslant d \, \mathrm{Vol}_d(K)^{(d-1)/d} \, \mathrm{Vol}_d(B)^{1/d}, \tag{2.1.6}$$

where $B := \overline{\mathrm{B}}(0,1)$.

Remark 2.7. We note that $d \, \mathrm{Vol}_d(K)^{(d-1)/d} \, \mathrm{Vol}_d(B)^{1/d}$ is no more than the perimeter of a d-dimensional ball with volume $\mathrm{Vol}_d(K)$.

We exhibit a direct proof of the isoperimetric inequality based on the Brunn-Mikowski theorem.

Proof of Theorem 2.6. Fix a convex body $K \subset \mathbb{R}^d$ and let $r > 0$. In the light of (2.1.5), we have that

$$\mathrm{Vol}_d(K + rB) = \mathrm{Vol}_d(K + r\mathrm{B}(0,1)) = \mathrm{Vol}_d(K) + r \, \mathrm{Vol}_{d-1}(\partial K) + o(r). \tag{2.1.7}$$

On the other hand, by Brunn-Minkowski's inequality, we get

$$\begin{aligned}\mathrm{Vol}_d(K+rB) &\geqslant \left(\mathrm{Vol}_d(K)^{1/d}+\mathrm{Vol}_d(rB)^{1/d}\right)^d\\ &= \left(\mathrm{Vol}_d(K)^{1/d}+r\,\mathrm{Vol}_d(B)^{1/d}\right)^d\\ &= \mathrm{Vol}_d(K)+dr\,\mathrm{Vol}_d(B)^{1/d}\,\mathrm{Vol}_d(K)^{(d-1)/d}+o(r).\end{aligned}$$

If we combine this inequality with (2.1.7), and then let $r \to 0$, we get (2.1.6). □

We comment on the fact that Theorem 2.6 admits a transport proof. This idea traces back to Gromov [MS86, Appendix I], who used Knothe's transport map in the spirit of the proof of the Brunn-Minkowski theorem given by Knothe [Kno57]. Figalli, Maggi, and Pratelli [FMP10] provided a similar argument that employs the Brenier map, which is the starting point of their striking stability theorem for the Euclidean isoperimetric inequality.

2.2 Isoperimetric inequalities

In the last section, we introduced the Euclidean isoperimetric inequality as a corollary of the Brunn-Minkowski inequality in the context of geometric inequalities. Despite the lack of a rigorous proof until the XIXth century, humankind has known (or at least intuited) the solution to the isoperimetric problem since the times of the ancient Greeks. The symmetries of the problem are a helpful factor.

Therefore, we may wonder if there are other "symmetric" situations where an isoperimetric principle may hold, guided by our intuition. Some examples could be the sphere $\mathbb{S}^d$, the hyperbolic space $\mathbb{H}^d$, a Riemannian manifold with some control on its curvature, the discrete hypercube $\{0,1\}^d$, etc. But beyond these highly symmetric situations, we may still wonder whether obtaining similar versions of Theorem 2.6 for more abstract spaces is possible; a priori, isoperimetry is only a metric notion. Which could be a good notion of an isoperimetric problem on a general metric measure space? In an abstract setting, what is the proper generalization of the concept of the perimeter of a set? Does it make sense to say to solve only partially an isoperimetric problem? Is that helpful in applications? Is it possible to get a functional version of an isoperimetric inequality?

Besides answering these questions just because they are beautiful, one of the main reasons to study isoperimetric phenomena is that they are an essential part of the theory of functional inequalities. More precisely, they are the starting point of the study of the concentration of measure phenomenon. One may say that the theory as we know it today began in the 70s from Milman's observation of the concentration phenomenon on the sphere $\mathbb{S}^d$, which stems directly from the spherical isoperimetric inequality. Generally speaking, we can translate good isoperimetric properties into good concentration properties.

After the previous introductory discussion, we organize the exposition of this section as follows: we provide the proper elements that generalize the isoperimetric question for abstract metric measure spaces. Then, based on these preliminaries, we review some classic examples focusing on two: the spherical and Gaussian cases. Good references in the subject that inspired the exposition of the topic are [LT91, Gro99, AAGM15, AAGM21].

2.2.1 Abstract isoperimetric principles

If we want to inquire about isoperimetric phenomena in an abstract setting, say, in a metric space, we need to generalize the Euclidean notion of the perimeter of a set. Let us look closer at the Euclidean case, more precisely at equation (2.1.5). One of the main ingredients in the formula is the notion of enlargement of a set, which approximates the measure of the original set in order to study its infinitesimal variation and hence recover the associated perimeter. In the Euclidean case, for any $r > 0$, the r-enlargement of a set A is given by the set $A + rB$, which coincides with the set of points that are at a distance less than or equal to r from the original set A. The latter characterization will be our proxy to define the r-enlargement of a set, modulo the strict inequality.

Definition 2.8 (r-enlargement)**.** Let $(\mathcal{X}, d)$ be a metric space. For $A \subset \mathcal{X}$ Borel and $r > 0$, we define the r-enlargement of A as the set

$$A_r \coloneqq \{x \in \mathcal{X} : \operatorname{dist}(x, A) < r\},$$

where $\operatorname{dist}(x, A) \coloneqq \inf_{y\in\mathcal{X}} d(x, y)$ is the distance between $x \in \mathcal{X}$ and the set A.

We are ready to define the perimeter of a set that lies in a metric space.

Definition 2.9 (Perimeter of a set)**.** Let $(\mathcal{X}, d, \mu)$ be a metric measure space, and let $A \in \mathcal{B}(\mathcal{X})$ with $\mu(A) < +\infty$. We define the perimeter of A as

$$\mu^+(A) \coloneqq \liminf_{r\to 0^+} \frac{\mu(A_r) - \mu(A)}{r}.$$

Now, we have the necessary ingredients to set up the isoperimetric question on any abstract metric measure space of interest.

Definition 2.10 (Isoperimetric problem)**.** Let $(\mathcal{X}, d, \mu)$ be a metric measure space. The isoperimetric problem for $(\mathcal{X}, d, \mu)$ is the following: given a fixed $\alpha > 0$, determine

$$\inf\{\mu^+(A) : A \in \mathcal{B}(\mathcal{X}), \mu(A) = \alpha\}$$

and identify its extremizers if they exist.

Remark 2.11. The isoperimetric problem on a metric measure space $(\mathcal{X}, d, \mu)$ is equivalent to the following formulation: we define the isoperimetric profile of the metric measure space $(\mathcal{X}, d, \mu)$ as the largest function $\mathcal{I}_\mu \colon [0, \mu(\mathcal{X})] \to \mathbb{R}_+$ such that

$$\forall A \in \mathcal{B}(\mathcal{X}) \text{ with } \mu(A) < +\infty, \quad \mu^+(A) \geqslant \mathcal{I}_\mu(\mu(A)).$$

Solving the isoperimetric problem is equivalent to identifying the associated isoperimetric profile and the sets $B \in \mathcal{B}(\mathcal{X})$ with $\mu(B) < +\infty$ satisfying

$$\mu^+(B) = \mathcal{I}_\mu(\mu(B)),$$

namely the optimal sets. Note that we can interpret a lower bound on $\mathcal{I}_\mu$ as a partial solution to the isoperimetric problem.

Example 2.12. Recalling Theorem 2.6, note that for $(\mathbb{R}^d, |\cdot|, \mathrm{Vol}_d)$, we can explicitly identify its isoperimetric profile: let $v\colon \mathbb{R}_+ \to \mathbb{R}_+$ be the function defined by

$$\forall r \geqslant 0, \quad v(r) \coloneqq \mathrm{Vol}_d(\mathrm{B}(0,r)) = r^d \mathrm{Vol}_d(\mathrm{B}(0,1)),$$

which is smooth and bijective. Then

$$\forall r \geqslant 0, \quad \mathcal{I}_{\mathrm{Vol}_d}(r) = (v' \circ v^{-1})(r) = d\mathrm{Vol}_d(\mathrm{B}(0,1))^{1/d} r^{(d-1)/d}.$$

Given these previous definitions, we are ready to review some key examples of isoperimetric phenomena.

2.2.2 Isoperimetry on the sphere

We start with the isoperimetric inequality on the sphere, proven independently by Lévy [Lév51] and Schmidt [Sch48]. For $d \geqslant 2$ integer, we consider the d-dimensional sphere as the set

$$\mathbb{S}^d \coloneqq \{x = (x_1, \dots, x_{d+1}) \in \mathbb{R}^{d+1} : |x| = 1\} \subset \mathbb{R}^{d+1},$$

endowed with the geodesic distance ρ. Let σ be the normalized surface measure on $\mathbb{S}^d$, i.e., the only probability measure on $\mathcal{B}(\mathbb{S}^d)$ invariant by rotations. That is, the metric measure space of interest here is $(\mathbb{S}^d, \rho, \sigma)$.

Let $N \in \mathbb{S}^d$ be the north pole, $N = (0, \dots, 0, 1)$, which will act as a reference point without loss of generality. We note that balls for the geodesic distance ρ are just spherical caps: for $r > 0$, the ball with center N and radius r can be written as

$$\mathrm{B}_{\mathbb{S}^d}(N, r) = \{x \in \mathbb{S}^d : x_{d+1} > \cos(r)\},$$

see Figure 2.1. Balls are remarkable subsets because they solve the isoperimetric problem in $\mathbb{S}^d$, as in the Euclidean case.

Theorem 2.13 (Spherical isoperimetric inequality)**.** *For $d \geqslant 2$, let $(\mathbb{S}^d, \rho, \sigma)$ be the sphere endowed with the geodesic distance and the unique rotationally invariant Borel probability measure. Let $N \in \mathbb{S}^d$ be the north pole. Fix $\alpha \in (0,1)$, and let $t_0 > 0$ be such that $\sigma(\mathrm{B}_{\mathbb{S}^d}(N, t_0)) = \alpha$. Then for every $A \subset \mathbb{S}^d$ Borel with $\sigma(A) = \alpha$,*

$$\forall r > 0, \quad \sigma(A_r) \geqslant \sigma(\mathrm{B}_{\mathbb{S}^d}(N, t_0)_r) = \sigma(\mathrm{B}_{\mathbb{S}^d}(N, t_0 + r)), \tag{2.2.1}$$

so in particular,

$$\sigma^+(A) \geqslant \sigma^+(\mathrm{B}_{\mathbb{S}^d}(N, t_0)). \tag{2.2.2}$$

That is, the isoperimetric profile of $(\mathbb{S}^d, \rho, \sigma)$ is given by the function

$$\mathcal{I}_\sigma = v' \circ v^{-1},$$

where $v\colon [0,\pi] \to [0,1]$ is the function defined by

$$\forall r \in [0,\pi], \quad v(r) = \sigma(\mathrm{B}_{\mathbb{S}^d}(N, r)). \tag{2.2.3}$$

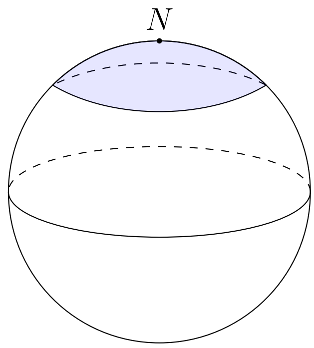

Figure 2.1: $\mathbb{S}^2$ and a spherical cap $\mathrm{B}_{\mathbb{S}^2}(N, r)$ (in blue).

Theorem 2.13 says that the isoperimetric problem on the sphere has a full solution, in the sense of Definition 2.10. The fact that spherical caps are the solution to it could not be very surprising, given the symmetries of the problem, at least from an intuitive point of view.

Considering the case $\alpha = 1/2$, we see that minimal-perimeter sets are precisely half-spheres as $\mathrm{B}_{\mathbb{S}^d}(N, \pi/2)$. Then, for any $A \subset \mathbb{S}^d$ with $\sigma(A) = 1/2$, and for any $r > 0$, we can estimate the right-hand side term in (2.2.1), thus getting the following result, which provides a lower bound for A_r.

Corollary 2.14. *For $d \geqslant 2$, let $A \subset \mathbb{S}^d$ be a Borel set with $\sigma(A) = 1/2$. Then*

$$\forall r > 0, \quad \sigma(A_r) \geqslant 1 - \sqrt{\frac{\pi}{8}} \exp\left(-(d-1)r^2/2\right).$$

The bound provided in Corollary 2.14 is the exact expression of what we have been calling the concentration of measure phenomenon on the sphere. It is a remarkable result in many ways. First, the obvious interpretation is that at fixed dimension $d \geqslant 2$, the area of A_r increases very fast as we let r grow. Now fix $r > 0$ and vary the dimension d: if we let $d \to +\infty$, the area of A_r also goes to 1 exponentially fast. If we apply this to the particular case when A is a half-sphere, we deduce that most of the area of a high-dimensional sphere concentrates around equatorial regions.

The spherical concentration of measure phenomenon traces back its roots to Milman [Mil71] as a key piece in his simplified proof of Dvoretzky's theorem [Dvo61], which in turn was the answer to a question of Grothendieck [Gro53] in the context of the local theory of Banach spaces, where one studies the structure of Banach spaces by the properties of their finite-dimensional spaces. That is how the theory of concentration of measure was born.

2.2.3 Isoperimetry in the Gaussian space

Now we consider the Gaussian space, that is, the metric probability space $(\mathbb{R}^d, |\cdot|, \gamma_d)$, where γ_d denotes the d-dimensional standard Gaussian measure. The solution to the isoperimetric problem is also known for this space, i.e., the function $\mathcal{I}_{\gamma_d}$ is explicit, and we can characterize the extremizers for the problem.

First, let us remark that the Gaussian can be approximated by the surface measure on a high-dimensional sphere. More precisely, fix $d \in \mathbb{N}^*$, and for $N \geqslant d$ integer, let $\sqrt{N}\,\mathbb{S}^{N-1}$ be the $(N-1)$-dimensional sphere dilated by a factor of $\sqrt{N}$, endowed with σ_N, the normalized uniform measure on $\sqrt{N}\,\mathbb{S}^{N-1}$. Let $\pi_{N,d}\colon \sqrt{N}\,\mathbb{S}^{N-1} \to \mathbb{R}^d$ be the projection of the first d coordinates of $\sqrt{N}\,\mathbb{S}^{N-1}$ onto $\mathbb{R}^d$. Then the sequence of measures $(\pi_{N,d})_\#\sigma_N \in \mathcal{P}(\mathbb{R}^d)$ converges weakly to γ_d as $N \to +\infty$. This result makes part of mathematical folklore and is commonly known as the "Poincaré-Maxwell lemma"; see, for example, [Gro99, Chapter $3\frac{1}{2}$.22].

With the help of the Poincaré-Maxwell lemma, Borell [Bor75a], and Sudakov and Tsirel'son [ST74] proved independently the Gaussian isoperimetric inequality as a consequence of Theorem 2.13. In this case, the optimizers for the isoperimetric problem are halfspaces, i.e., sets of the form

$$\{x \in \mathbb{R}^d : \theta_0 \cdot x \leqslant t_0\},$$

for some $(\theta_0, t_0) \in \mathbb{R}^d \times \mathbb{R}$.

Theorem 2.15 (Gaussian isoperimetric inequality)**.** *Let $\alpha \in (0,1)$, and let $(\theta_0, t_0) \in \mathbb{R}^d \times \mathbb{R}$ be such that the halfspace $H = \{x \in \mathbb{R}^d : \theta_0 \cdot x \leqslant t_0\}$ has mass α: $\gamma_d(H) = \alpha$. Then for every $A \subset \mathbb{R}^d$ Borel with $\gamma_d(A) = \alpha$, and every $r > 0$, we have*

$$\gamma_d(A_r) \geqslant \gamma_d(H_r). \tag{2.2.4}$$

In particular,

$$\gamma_d^+(A) \geqslant \gamma_d^+(H). \tag{2.2.5}$$

Moreover, the isoperimetric profile of $(\mathbb{R}^d, |\cdot|, \gamma_d)$ is given by the function

$$\mathcal{I}_{\gamma_d} = \mathcal{I}_\gamma := \Phi' \circ \Phi^{-1}, \tag{2.2.6}$$

where $\Phi\colon \mathbb{R} \to \mathbb{R}_+$ is the function defined by

$$\forall r \in \mathbb{R}, \quad \Phi(r) = \gamma_1((-\infty, r)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r} e^{-x^2/2}\,\mathrm{d}x.$$

Remark 2.16. We note that the Gaussian isoperimetric inequality is dimension-free: its isoperimetric profile $\mathcal{I}_{\gamma_d}$, given by (2.2.6), does not depend on the parameter d; that is why we denote it by $\mathcal{I}_\gamma$.

Similarly to Corollary 2.14, we can get an analog concentration result for the Gaussian space.

Corollary 2.17. *Let $A \subset \mathbb{R}^d$ Borel with $\gamma_d(A) = 1/2$. Then*

$$\forall r > 0, \quad \gamma_d(A_r) \geqslant 1 - \frac{1}{2}\exp\left(-r^2/2\right).$$

The Gaussian concentration exhibited in Corollary 2.17 resembles Corollary 2.14, but there is an important difference: the right-hand side term in the inequality does not depend on the dimension d; i.e., the rate at which halfspaces concentrate is the same for every dimension, which is consistent with Remark 2.16.

Theorem 2.15 admits a functional version, proven by Bobkov [Bob96, Bob97].

Theorem 2.18 (Functional Gaussian isoperimetric inequality)**.** *The d-dimensional Gaussian measure satisfies*

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}^d, [0,1]), \quad \mathcal{I}_\gamma \left(\int_{\mathbb{R}^d} f \, \mathrm{d}\gamma_d \right) \leqslant \int_{\mathbb{R}^d} \sqrt{(\mathcal{I}_\gamma \circ f)^2 + |\nabla f|^2} \, \mathrm{d}\gamma_d.$$

Remark 2.19. We can show that Theorem 2.18 recovers the geometric isoperimetric inequality, Theorem 2.15: for any $A \subset \mathbb{R}^d$ Borel, let $r > 0$ and define the function $f_r \colon \mathbb{R}^d \to [0,1]$ by

$$\forall x \in \mathbb{R}^d, \quad f_r(x) = \max\left\{1 - \mathrm{dist}(x, A_r)/r, 0\right\},$$

where $\mathrm{dist}(x, A_r)$ is the Euclidean distance from the point $x \in \mathbb{R}^d$ to the set A_r. If we apply Theorem 2.18 to f_r and let $r \to 0$, we recover the inequality $\gamma_d^+(A) \geqslant \mathcal{I}_\gamma(\gamma_d(A))$.

Another remarkable property of the functional Gaussian isoperimetric inequality is that it is preserved by Lipschitz pushforwards, modulo a constant. This fact was mentioned by Caffarelli [Caf00] as an application of his contraction result, Theorem 1.26 in Chapter 1; for a direct proof, see [CE02, Section 3]. Historically, we can trace back the usage of these stability properties to getting new functional inequalities to Pisier [Pis86, p. 181]. Many functional inequalities are stable by Lipschitz pushforwards, as we will see in the following sections. Concerning this thesis, this kind of property will be of crucial importance later in Chapters 3 and 4.

Proposition 2.20. *Let $T \colon \mathbb{R}^d \to \mathbb{R}^d$ be a differentiable map, and assume that there exists $K > 0$ such that T is K-Lipschitz. Let μ be the pushforward of γ_d by the map T, i.e., $\mu = T_\# \gamma_d$. Then*

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}^d, [0,1]), \quad \mathcal{I}_\gamma \left(\int_{\mathbb{R}^d} f \, \mathrm{d}\mu \right) \leqslant \int_{\mathbb{R}^d} \sqrt{(\mathcal{I}_\gamma \circ f)^2 + K^2 |\nabla f|^2} \, \mathrm{d}\mu. \tag{2.2.7}$$

The previous result motivates the following definition.

Definition 2.21 (Functional Gaussian-type isoperimetric inequality)**.** Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a Borel probability measure on $\mathbb{R}^d$. We say that μ satisfies a functional Gaussian-type isoperimetric inequality if there exists a constant $C_{\mathrm{isop}}(\mu) > 0$ such that

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}^d, [0,1]), \quad \mathcal{I}_\gamma \left(\int_{\mathbb{R}^d} f \, \mathrm{d}\mu \right) \leqslant \int_{\mathbb{R}^d} \sqrt{(\mathcal{I}_\gamma \circ f)^2 + C_{\mathrm{isop}}(\mu)^2 |\nabla f|^2} \, \mathrm{d}\mu. \tag{2.2.8}$$

Remark 2.22. A functional Gaussian-type isoperimetric inequality is a powerful device, as it implies many other functional inequalities and is one of the strongest in the large hierarchy of functional inequalities; we will elaborate on that point later in Section 2.5.2. As an appetizer, let us remark the following: let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a Borel probability measure satisfying inequality (2.2.8) for some $C_{\mathrm{isop}}(\mu) > 0$. By an argument similar to the one sketched in Remark 2.19, one can prove the following geometric isoperimetric-type inequality:

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \quad \mu^+(A) \geqslant C_{\mathrm{isop}}(\mu) \; \mathcal{I}_\gamma(\mu(A)). \tag{2.2.9}$$

In other words, we have deduced that if an inequality of the type (2.2.8) holds for μ, then we can lower bound its isoperimetric profile $\mathcal{I}_\mu$ by $C_{\text{isop}}(\mu)\,\mathcal{I}_\gamma$. In turn, note that inequality (2.2.9) yields concentration bounds similar to the one exhibited in Corollary 2.17.

2.2.4 Other isoperimetric-type inequalities

For the sake of completeness, we mention that the examples reviewed here do not constitute an exhaustive list of isoperimetric phenomena. For instance, in the discrete setting, there is an isoperimetric inequality on the hypercube $\{0,1\}^d$ endowed with its normalized uniform measure, which was proven by Harper [Har66]. It is a striking result because of the discrete nature of $\{0,1\}^d$; see, for example, [AAGM15, Section 3.1.5].

In the smooth setting, we have the Lévy-Gromov inequality, which was conjectured by Lévy [Lév51] and proved by Gromov [Gro80]; see for further reference [GHL04, Section 4.H] or [Gro99, Appendix C]. More precisely, let (M,g) be a Riemannian manifold of dimension $d \geqslant 2$ and Ricci curvature bounded from below by $K > 0$, i.e., $\mathrm{Ric} \geqslant Kg$. Then, as a consequence of the Bonnet-Myers theorem, M is compact; thus, we can consider the normalized volume measure Vol on M. The Lévy-Gromov inequality says that the isoperimetric profile associated with the rescaled sphere $\frac{K}{d-1}\,\mathbb{S}^d$ (i.e., a rescaled version of (2.2.3)) is a lower bound for the isoperimetric profile of the manifold (M,g,Vol). In other words, the model space $\frac{K}{d-1}\,\mathbb{S}^d$, which has dimension d and Ricci curvature bounded from below by K (in fact, equal to K), provides as well a "model" bound on the isoperimetric profile of all such manifolds.

In the context of the synthetic characterization of bounded Ricci curvature for metric spaces, Cavelletti and Mondino [CM17] accomplished a generalization of the Lévy-Gromov inequality by adapting Klartag's Riemannian needle decomposition [Kla17], which is in turn based on the geometry of the optimal transport for the L^1 cost. See [Vil19] for a comprehensive exposition of the result.

We mention that in the Euclidean setting, log-concave measures have nice isoperimetric properties; we will elaborate on this point in Section 2.4.4. This is a highly active research topic nowadays [AAGM21].

Finally, we mention that the isoperimetric problem remains open in a number of situations [Ros05]. For example, it is conjectured that the extremal sets for the problem on the flat torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ endowed with the uniform measure are of the form $B \times \mathbb{T}^{d-k}$, where $1 \leqslant k \leqslant d$ and $B \subset \mathbb{T}^k$ is a ball for the geodesic distance on $\mathbb{T}^k$.

2.3 Concentration of measure

In the last section, we had our first encounter with the phenomenon of concentration of measure in Corollaries 2.14 and 2.17, which state that the volume of an r-enlarged set tends very fast to 1 as r grows; both are a direct consequence of their respective isoperimetric inequalities. In turn, a consequence of these results is that almost all the mass is condensed in a specific region of the space. This observation is what we call concentration of measure. But, what is this, more precisely?

This section aims to introduce in a general framework the phenomenon of concentration of measure, in the same way as we did for abstract isoperimetric inequalities in Section 2.2.1. We start by defining the concentration function associated with a metric measure space, which is the device that allows us to quantify precisely the rate of concentration for a measure. After that, we provide a functional characterization of concentration written in terms of the deviation of Lipschitz functions. Then we define the most essential types of concentration pertinent to this thesis: subgaussian, subexponential, and subpoissonian. Finally, we end with a discussion about dimension-free concentration. This section's exposition was mainly based on the well-known references [LT91, Led01, BLM13].

2.3.1 Abstract concentration

We start by defining the concentration function of a metric measure space, which quantifies how fast a measure concentrates. We will suppose in this section that every metric measure space is a probability space unless otherwise stated.

Definition 2.23 (Concentration function)**.** Let $(\mathcal{X}, d, \mu)$ be a metric probability space. We define the associated concentration function, $\alpha_\mu \colon \mathbb{R}_{>0} \to [0,1]$, by

$$\forall r > 0, \quad \alpha_\mu(r) \coloneqq \sup\{1 - \mu(A_r) : \mu(A) \geqslant 1/2\}.$$

We note that without any further assumption, $\lim_{r\to+\infty} \alpha_\mu(r) = 0$, since μ is a probability measure. Then, a natural question is whether it is possible to quantify its decay, which is the focus of concentration theory. The theory found its origins after the seminal contribution of Milman [Mil71], namely Corollary 2.14 in the previous section. After that, the theory started to build up in a systematic framework, notably after contributions by Amir and Milman [AM80], Gromov and Milman [GM83], Milman and Schechtman [MS86] and Gromov [Gro99].

A first property of concentration functions is that Lipschitz pushforwards preserve them in the following sense.

Proposition 2.24. *Let $(\mathcal{X}, d_{\mathcal{X}}, \mu)$ be metric probability space and let $(\mathcal{Y}, d_{\mathcal{Y}})$ be a metric space. Let $T \colon \mathcal{X} \to \mathcal{Y}$ be a K-Lipschitz map for some $K > 0$, and let $\nu \in \mathcal{P}(\mathcal{Y})$ be the pushforward of μ by T, i.e., $\nu = T_\#\mu$. Let α_μ and α_ν be the concentration functions associated to $(\mathcal{X}, d_{\mathcal{X}}, \mu)$ and $(\mathcal{Y}, d_{\mathcal{Y}}, \nu)$, respectively. Then*

$$\forall r > 0, \quad \alpha_\nu(r) \leqslant \alpha_\mu(r/K).$$

Concentration phenomena admit a functional version in the following sense. Let $(\mathcal{X}, d, \mu)$ be a metric probability space. One can study how observables $f \colon \mathcal{X} \to \mathbb{R}$ belonging to a large class deviate from their median. In this case, the suitable class of observables to obtain such a statement is the one of Lipschitz functions. Morally, if a function $f \colon \mathcal{X} \to \mathbb{R}$ is Lipschitz, that means that the oscillation of f between two points x and y, i.e., $|f(x) - f(y)|$, is proportional to the distance $d(x, y)$ of the points in the base space. Hence, it is reasonable to expect that if a measure μ on $\mathcal{X}$ concentrates at rate α_μ, then the mass of the set of points where a Lipschitz observable is far from its median with respect to μ should decay at a rate similar to α_μ; the following proposition makes this intuition precise. Recall that if $(\mathcal{X}, d, \mu)$ is a metric probability space, we say that a value $m \in \mathbb{R}$ is a median for $f \colon \mathcal{X} \to \mathbb{R}$ with respect to μ if

$$1/2 = \mu\left(\{x \in \mathcal{X} : f(x) \leqslant m\}\right) = \mu\left(\{x \in \mathcal{X} : f(x) \geqslant m\}\right).$$

We denote a median for f by $m_\mu(f) \in \mathbb{R}$.

Proposition 2.25. *Let $(\mathcal{X}, d, \mu)$ be a metric probability space, and let α_μ be its concentration function. Then α_μ is the smallest function such that:*

(i) For every 1-Lipschitz function $f \colon \mathcal{X} \to \mathbb{R}$, and any median $m_\mu(f)$,

$$\forall r > 0, \quad \mu(\{x \in \mathcal{X} : f - m_\mu(f) > r\}) \leqslant \alpha_\mu(r).$$

(ii) For every L-Lipschitz function $f \colon \mathcal{X} \to \mathbb{R}$, and any median $m_\mu(f)$,

$$\forall r > 0, \quad \mu(\{x \in \mathcal{X} : f - m_\mu(f) > r\}) \leqslant \alpha_\mu(r/L).$$

In particular, it holds the following concentration bound: for every L-Lipschitz function $f \colon \mathcal{X} \to \mathbb{R}$, and for any median $m_\mu(f)$,

$$\mu(\{x \in \mathcal{X} : |f - m_\mu(f)| > r\}) \leqslant 2\alpha_\mu(r/L).$$

2.3.2 Subgaussian and subexponential concentration

Concentration of measure can occur at different rates. In particular, we are interested in concentration functions that decay exponentially fast in the following sense.

Definition 2.26. Let $(\mathcal{X}, d, \mu)$ be a metric probability space, let α_μ be its concentration function, and let $p \geqslant 1$. We say that the measure μ has p-exponential concentration of measure if there exist constants $C_1, C_2 > 0$ such that

$$\forall r > 0, \quad \alpha_\mu(r) \leqslant C_1 \exp\left(-C_2 r^p\right).$$

In particular, if $p = 1$, we say that μ has subexponential concentration; if $p = 2$, we say that μ has subgaussian concentration.

Remark 2.27. We immediately note that subgaussian concentration is stronger than subexponential.

Example 2.28. From Corollaries 2.14 and 2.17 we know that both $(\mathbb{S}^d, \rho, \sigma)$ and $(\mathbb{R}^d, |\cdot|, \gamma_d)$ have subgaussian concentration. More generally, Remark 2.22 shows that every measure satisfying a Gaussian-type isoperimetric inequality concentrates at a subgaussian rate.

Example 2.29. Let μ be the exponential measure on $\mathbb{R}_{>0}$, that is, $\mathrm{d}\mu(x) \coloneqq \exp(-x)\,\mathrm{d}x$. Then the μ concentrates at a subexponential rate but not at a subgaussian rate.

When a measure is concentrated in the sense of Definition 2.26, we can replace in Proposition 2.25 the median of the observable by its mean.

Proposition 2.30. *Let $(\mathcal{X}, d, \mu)$ be a metric probability space, let α_μ be its concentration function, and let $p \geqslant 1$. Then the following statements are equivalent:*

(i) There exist constants $C_1, C_2 > 0$ such that

$$\forall r > 0, \quad \alpha_\mu(r) \leqslant C_1 \exp\left(-C_2 r^p\right).$$

(ii) There exist constants $C_3, C_4 > 0$ such that for any 1-Lipschitz function $f \colon \mathcal{X} \to \mathbb{R}$, and any median $m_\mu(f)$ of f,

$$\forall r > 0, \quad \mu\left(\{x \in \mathcal{X} : |f(x) - m_\mu(f)| > r\}\right) \leqslant C_3 \exp\left(-C_4 r^p\right).$$

(iii) There exist constants $C_5, C_6 > 0$ such that for any 1-Lipschitz function $f \colon \mathcal{X} \to \mathbb{R}$,

$$\forall r > 0, \quad \mu\left(\{x \in \mathcal{X} : |f(x) - \mathbb{E}_{X\sim\mu}[f(X)]| > r\}\right) \leqslant C_5 \exp\left(-C_6 r^p\right).$$

2.3.3 Poissonian concentration

If we now work on $\mathcal{X} = \mathbb{N}$ endowed with the graph distance, a particular type of concentration arises naturally, the subpoissonian phenomenon, which is modeled after the folkloric concentration bound satisfied by the Poisson distribution on $\mathbb{N}$, a consequence of Bennett's inequality. First, we define Bennett's function $h \colon \mathbb{R}_+ \to \mathbb{R}_+$ as the function

$$\forall r > 0, \quad h(r) \coloneqq (1 + r)\log(1 + r) - r,$$

which appears naturally in the Poissonian case.

Example 2.31 (Poisson distribution)**.** For $T > 0$, let π_T be the Poisson distribution of parameter T on $\mathbb{N}$, and let α_{π_T} be its concentration function. Then

$$\forall r > 0, \quad \alpha_{\pi_T}(r) \leqslant \exp\left(-Th(r/T)\right),$$

and the bound is asymptotically sharp when $r \to +\infty$; in particular, π_T is not subgaussian.

We model abstract subpoissonian concentration on Example 2.31.

Definition 2.32. Let $\mu \in \mathcal{P}(\mathbb{N})$, and let α_μ be its concentration function. We say that μ has subpoissonian concentration if there exist constants $C_1, C_2 > 0$ such that

$$\forall r > 0, \quad \alpha_\mu(r) \leqslant C_1 \exp\left(-C_2 h(r/C_2)\right),$$

where h is the Bennett function.

Remark 2.33. Based on the analytical properties of h, we see that a subpoissonian measure μ concentrates at a slightly better rate than a subexponential measure. More precisely, since $h(r) \geqslant \frac{r^2}{2(1+r/3)}$ for all $r > 0$, then

$$\forall r > 0, \quad \alpha_\mu(r) \leqslant C_1 \exp\left(-\frac{r^2}{2\left(C_2 + r/3\right)}\right),$$

where $C_1, C_2 > 0$.

As in the former examples seen in Section 2.3.2, subpoissonian concentration also controls the deviation of Lipschitz observables. For $f\colon \mathbb{N} \to \mathbb{R}$, note that f is 1-Lipschitz if and only if $\sup_{k\in\mathbb{N}}|\mathrm{D}f(k)| \leqslant 1$, where $\mathrm{D}f$ is the discrete gradient on $\mathbb{N}$, which we define by

$$\forall k \in \mathbb{N}, \quad \mathrm{D}f(k) \coloneqq f(k+1) - f(k).$$

Proposition 2.34. *Let $\mu \in \mathcal{P}(\mathbb{N})$ be subpoissonian with constants C_1 and C_2. Then for any 1-Lipschitz function $f\colon \mathcal{X} \to \mathbb{R}$,*

$$\forall r > 0, \quad \mu\left(\{k \in \mathbb{N} : f(k) - \mathbb{E}_{X\sim\mu}[f(X)] > r\}\right) \leqslant C_1 \exp\left(-C_2 h(r/C_2)\right),$$

where h is the Bennett function.

2.3.4 Dimension-free concentration

Let us recall the Gaussian concentration bound given by Corollary 2.17. It is a remarkable result because it is a dimension-free bound; see Remark 2.16. This property is very useful in practice: for example, suppose that we have a collection of n independent real-valued standard Gaussian random variables $(X_i)_{i=1}^n$, so their joint law is γ_n. Proposition 2.30 implies that there exist two universal constants $C_1, C_2 > 0$, independent of d, such that for any 1-Lipschitz observable $f\colon \mathbb{R}^n \to \mathbb{R}$,

$$\forall r > 0, \quad \mathbb{P}\left(|f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)]| > r\right) \leqslant C_1 \exp\left(-C_2 r^2\right);$$

i.e., we obtain a bound for the deviation of $f(X_1, \ldots, X_n)$ from its mean that is independent of the size of the sample $(X_i)_{i=1}^n$.

Concerning the last point, there is a big issue: concentration inequalities generally do not tensorize in a dimension-free way. More precisely, suppose that $(\mathcal{X}, d, \mu)$ is a metric probability space with the subgaussian concentration property. For $n \in \mathbb{N}^*$, we can equip the product space $\mathcal{X}^n$ with the metric $d^n \coloneqq \bigoplus_{i=1}^n d$ and the product measure $\mu^n \coloneqq \bigotimes_{i=1}^n \mu$. It is not generally true that $(\mathcal{X}^n, d^n, \mu^n)$ is subgaussian with constants that do not depend on n.

In the last section, we saw that isoperimetric inequalities are an ally if we want to obtain a good concentration rate. However, such an inequality is a very strong property that is generally difficult to prove. In the sequel, we will study other functional inequalities that are easier to establish and entail dimension-free concentration bounds.

2.4 Poincaré inequalities and Markov semigroups

In this section, we will introduce Poincaré inequalities, the first example of a functional inequality that yields concentration bounds, more precisely, of subexponential type. One of its main features is the fact that Poincaré inequalities tensorize in a dimension-free fashion, which addresses the issue raised at the end of the previous section. As a warm-up, let us give the first example of a Poincaré inequality, the Gaussian one, which traces its origins back to Nash [Nas58], Chernoff [Che81], and Chen [Che82].

Theorem 2.35 (Gaussian Poincaré inequality). *The d-dimensional standard Gaussian measure satisfies a Poincaré inequality:*

$$\forall f \in \mathcal{C}_c^\infty(\mathbb{R}^d), \quad \mathrm{Var}_{\gamma_d}(f) \leqslant \int_{\mathbb{R}^d} |\nabla f|^2 \,\mathrm{d}\gamma_d. \tag{2.4.1}$$

Remark 2.36.

(i) The bound is valid for every dimension $d \in \mathbb{N}^*$; i.e., it is a dimension-free bound.

(ii) The class of functions for which (2.4.1) is verified actually is bigger: it holds for any locally Lipschitz function.

Let us take an arbitrary Borel probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$. It makes sense to inquire if an inequality of the form (2.4.1) holds for μ: it just suffices to replace γ_d by μ at both sides of the inequality and maybe admit a constant premultiplying the integrated squared norm of the gradient different than one; this would be a Poincaré-type inequality for μ. But, what happens in other structures? For example, if (M, g) is a Riemannian manifold endowed with a measure $\mu \in \mathcal{P}(M)$, does the validity of a Poincaré inequality in the form (2.4.1) still have meaning if we replace ∇ by the Riemannian gradient? What happens if we now work on a discrete space, e.g., the natural numbers, the two-point space, or, more generally, a graph? In those situations, the concept of what is a Poincaré inequality is less straightforward.

The following subsection addresses this issue via the theory of Markov semigroups. This framework provides a general structure where we can recognize the essential objects playing a role in a Poincaré inequality, thus opening the path towards a more general definition of the inequality for other settings.

We continue this section by reviewing the basic objects and elements of the Markov semigroup theory. After that, we will define in more generality Poincaré inequalities and study their essential properties and most remarkable examples. One of the best references on the interplay between Markov semigroups and functional inequalities is the book [BGL14], which inspired the exposition of those topics here. Finally, in order to provide a sufficient condition ensuring a Poincaré inequality on the Euclidean case, we introduce the class of log-concave measures and review their isoperimetric properties. Excellent references on the topic are [AAGM21, KL24b].

2.4.1 Markov semigroups

Here, we will review the basic theory of Markov semigroups, a more general setting to study (2.4.1). We start by defining Markov semigroups on Polish spaces.

Definition 2.37 (Markov semigroup). Let $\mathcal{X}$ be a Polish space, and let μ be a nonnegative σ-finite Borel measure on $\mathcal{X}$. Let $(\mathrm{P}_t)_{t\geqslant 0}$ be a family of bounded linear operators on $\mathbb{B}(\mathcal{X})$. We say that $(\mathrm{P}_t)_{t\geqslant 0}$ is a Markov semigroup if it satisfies the following properties:

(i) $\mathrm{P}_0 = \mathrm{id}_{L^2(\mu)}$.

(ii) For every $t \geqslant 0$, $\mathrm{P}_t \mathbb{1} = \mathbb{1}$.

(iii) For every $t \geqslant 0$, the operator P_t is positive: if $f \in \mathbb{B}(\mathcal{X})$ is such that $f \geqslant 0$, then $\mathrm{P}_t f \geqslant 0$.

(iv) For every $t, s \geqslant 0$, $\mathrm{P}_{t+s} = \mathrm{P}_t \circ \mathrm{P}_s$.

(v) For every $p \geqslant 1$ and every $t \geqslant 0$, P_t can be extended to a bounded operator $\mathrm{P}_t \colon L^p(\mu) \to L^p(\mu)$ which is a contraction; that is, for every $f \in L^p(\mu)$, $\|\mathrm{P}_t f\|_p \leqslant \|f\|_p$.

(vi) For every $f \in L^2(\mu)$, it holds that $\mathrm{P}_t f \to f$ in $L^2(\mu)$, as $t \to 0^+$.

Additionally, we say that μ is an invariant measure for $(\mathrm{P}_t)_{t\geqslant 0}$ if

$$\forall t \geqslant 0, \forall f \in \mathbb{B}(E), \quad \int_{\mathcal{X}} \mathrm{P}_t f \,\mathrm{d}\mu = \int_{\mathcal{X}} f \,\mathrm{d}\mu. \tag{2.4.2}$$

Remark 2.38 (Markov processes)**.** One of the greatest motivations for defining Markov semigroups is Markov processes. Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ and a Polish space $\mathcal{X}$, we say that an adapted process $(X_t)_{t\geqslant 0}$ with values in $\mathcal{X}$ is a time-homogeneous Markov process if

$$\forall t > s \geqslant 0, \forall f \in \mathcal{C}_\mathrm{b}(\mathcal{X}), \quad \mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|X_s] = \mathbb{E}_{X_s}[f(X_{t-s})],$$

where for $x \in \mathcal{X}$, we denote by $(X_t^x)_{t\geqslant 0}$ the process conditioned to $X_0 = x$, so $\mathbb{P}_x$ is the probability measure on Ω induced by $(X_t^x)_{t\geqslant 0}$. Assume that there exists a measure μ such that $X_0 \sim \mu$ implies $X_t \sim \mu$ for every $t \geqslant 0$; i.e., μ is a stationary law for the process $(X_t)_{t\geqslant 0}$. Then, under mild regularity assumptions,[1] there exists a Markov semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ on $\mathcal{X}$ with invariant measure μ that characterizes the law of $(X_t)_{t\geqslant 0}$, that is,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_\mathrm{b}(\mathcal{X}), \forall x \in \mathcal{X}, \quad \mathbb{E}_x[f(X_t)] = \mathrm{P}_t f(x).$$

Now, returning to the general framework, under the assumptions given at Definition 2.37, Hille-Yosida theory [Yos80] guarantees the existence of $\mathcal{D} \subset L^2(\mu)$, a dense linear subspace where $(\mathrm{P}_t)_{t\geqslant 0}$ is differentiable at $t = 0$ with derivative in $L^2(\mu)$:

$$\forall f \in \mathcal{D}, \quad \lim_{t\to 0^+} \frac{1}{t}(\mathrm{P}_t f - f) \in L^2(\mu). \tag{2.4.3}$$

Therefore, the existence of this limit for functions in $\mathcal{D}$ induces a linear operator $\mathrm{L} \colon \mathcal{D} \subset L^2(\mu) \to L^2(\mu)$ with domain $\mathrm{Dom}(\mathrm{L}) \coloneqq \mathcal{D}$, via (2.4.3). We say that L is the infinitesimal generator, or just the generator, associated with the semigroup $(\mathrm{P}_t)_{t\geqslant 0}$.

Observe that by the semigroup property, the operator L satisfies the following equation on $\mathcal{D}$:

$$\forall t \geqslant 0, \quad \partial_t \mathrm{P}_t = \mathrm{P}_t \mathrm{L} = \mathrm{L}\mathrm{P}_t. \tag{2.4.4}$$

This is the heat equation associated with $(\mathrm{P}_t)_{t\geqslant 0}$.

[1] Namely, assume that there exists a dense subset of $L^2(\mu)$ of bounded functions $\mathcal{A}$ such that for each $f \in \mathcal{A}$, the real-valued process $t \mapsto f(X_t^x)$ is càdlàg for any initial value $x \in \mathcal{X}$, and for any $t_0 > 0$, the function $x \mapsto \sup_{t\in[0,t_0]} |\mathbb{E}_x[f(X_t)]|$ belongs to $L^2(\mu)$ [BGL14, p. 11].

Working with the domain $\mathcal{D}$ can be difficult since it will often not be explicit. However, in almost every situation, there is an explicit class of well-behaved functions belonging to $\mathcal{D}$ that allow for explicit computations with L. Moreover, this class satisfies an appropriate density property that permits working with those functions instead of the functions in $\mathcal{D}$ without any loss of generality. This motivates the definition of core algebra.

Definition 2.39 (Core algebra)**.** Let $\mathcal{X}$ be a Polish space, and let μ be a nonnegative σ-finite Borel measure on $\mathcal{X}$. Let $\mathrm{L}\colon \mathrm{Dom}(\mathrm{L}) \subset L^2(\mu) \to L^2(\mu)$ be a linear operator with dense domain $\mathrm{Dom}(\mathrm{L})$. Let $\mathcal{A}_0$ be a linear subspace of $\mathrm{Dom}(\mathrm{L})$ that is an algebra, i.e., it is stable by the pointwise product operation between real-valued functions. We say that $\mathcal{A}_0$ is a core algebra for L if for any $f \in \mathrm{Dom}(\mathrm{L})$, there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{A}_0$ such that $f_n \to f$ and $\mathrm{L}f_n \to \mathrm{L}f$, where both convergences hold in the $L^2(\mu)$ sense.

In all the examples we will treat here, the semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ will satisfy the following self-adjointness property named symmetry.

Definition 2.40 (Symmetry)**.** Let $\mathcal{X}$ be a Polish space, and let μ be a nonnegative σ-finite Borel measure on $\mathcal{X}$. We say that a Markov semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ defined on $\mathcal{X}$ with invariant measure μ is symmetric with respect to μ, or μ is reversible for $(\mathrm{P}_t)_{t\geqslant 0}$, if

$$\forall t \geqslant 0, \forall f, g \in L^2(\mu), \qquad \int_{\mathcal{X}} g\, \mathrm{P}_t f \,\mathrm{d}\mu = \int_{\mathcal{X}} f\, \mathrm{P}_t g \,\mathrm{d}\mu. \tag{2.4.5}$$

Remark 2.41 (Integration by parts)**.** Note that the symmetry property is equivalent to the following integration by parts formula:

$$\forall f, g \in \mathrm{Dom}(\mathrm{L}), \qquad \int_{\mathcal{X}} g\, \mathrm{L} f \,\mathrm{d}\mu = \int_{\mathcal{X}} f\, \mathrm{L} g \,\mathrm{d}\mu. \tag{2.4.6}$$

Remark 2.42. We highlight the fact that the infinitesimal generator L of a symmetric Markov semigroup is a self-adjoint operator on $L^2(\mu)$, i.e., $\mathrm{L} = \mathrm{L}^*$ and $\mathrm{Dom}(\mathrm{L}) = \mathrm{Dom}(\mathrm{L}^*)$.

We state all the basic regularity conditions for a Markov semigroup that will allow us to work in the sequel under the name of the usual conditions.

Definition 2.43 (Usual conditions)**.** Let $\mathcal{X}$ be a Polish space, let μ be a nonnegative σ-finite Borel measure on $\mathcal{X}$, and let $(\mathrm{P}_t)_{t\geqslant 0}$ be a Markov semigroup defined on $\mathcal{X}$ with invariant measure μ. We say that $(\mathrm{P}_t)_{t\geqslant 0}$ satisfies the usual conditions if it is symmetric with respect to μ and its infinitesimal generator admits a core algebra.

We proceed with the definition of the carré du champ operator, which comes originally from potential theory. It plays a fundamental role in the interplay between the theory of Markov semigroups and functional inequalities.

Definition 2.44 (Carré du champ operator)**.** Let $\mathcal{X}$ be a Polish space, let μ be a nonnegative σ-finite Borel measure on $\mathcal{X}$, and let $(\mathrm{P}_t)_{t\geqslant 0}$ be a Markov semigroup defined on $\mathcal{X}$ with generator L satisfying the usual conditions. Let $\mathcal{A}_0$ be the core algebra associated to L. We define the carré du champ operator $\Gamma\colon \mathcal{A}_0 \times \mathcal{A}_0 \to \mathcal{A}_0$ by

$$\forall f, g \in \mathcal{A}_0, \quad \Gamma(f,g) \coloneqq \frac{1}{2}(\mathrm{L}(fg) - g\,\mathrm{L}f - f\,\mathrm{L}g).$$

By abuse of notation, for $f \in \mathcal{A}_0$, we write $\Gamma(f) \coloneqq \Gamma(f,f)$.

Remark 2.45.

(i) One can show that the carré du champ is a symmetric and bilinear operator. Moreover, it is nonnegative: for any $f \in \mathcal{A}_0$, $\Gamma(f) \geqslant 0$.

(ii) Observe that Γ reflects how L is far from being a derivation: for example, in dimension one, if $\mathrm{L}f = f'$, then $\Gamma(f) = 0$.

(iii) Note that

$$\forall f, g \in \mathcal{A}_0, \quad \int_{\mathcal{X}} \Gamma(f,g)\,\mathrm{d}\mu = -\int_{\mathcal{X}} g\,\mathrm{L}f\,\mathrm{d}\mu = -\int_{\mathcal{X}} f\,\mathrm{L}g\,\mathrm{d}\mu$$

because of the integration by parts formula (2.4.6).

Finally, Markov triples will be the main setting for functional inequalities.

Definition 2.46 (Markov triple)**.** We say that $(\mathcal{X}, \mathrm{L}, \mu)$ is a Markov triple if $\mathcal{X}$ is a Polish space, μ is a nonnegative σ-finite Borel measure on $\mathcal{X}$ and L is the infinitesimal generator associated to a Markov semigroup satisfying the usual conditions.

We give some examples of Markov triples, all induced by Markov processes; recall Remark 2.38. For some of them, the semigroup does not have an explicit expression. Nevertheless, in Section 3.2 of Chapter 3, we will see that this is not an obstacle, since we can translate all the desired properties in terms of the generator L, which we will always know explicitly.

Example 2.47 (Brownian motion)**.** Let $\mathcal{X} = \mathbb{R}^d$, and let $(B_t)_{t\geqslant 0}$ be the standard d-dimensional Brownian motion. Define for each $t \geqslant 0$, $X_t \coloneqq \sqrt{2}B_t$, where the rescaling factor of $\sqrt{2}$ is only cosmetic and has the purpose of avoiding a $1/2$ factor in the infinitesimal generator. Its only invariant and reversible measure is Leb. The semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ associated to $(X_t)_{t\geqslant 0}$ is the heat semigroup:

$$\forall t > 0, \forall f \in \mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}^d), \forall x \in \mathbb{R}^d, \quad \mathrm{P}_t f(x) = \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} f(y) e^{-\frac{|x-y|^2}{4t}}\,\mathrm{d}y.$$

Its infinitesimal generator is the standard Laplacian operator on $\mathbb{R}^d$, $\mathrm{L}f = \Delta f$, acting on functions $f \in \mathcal{C}^2(\mathbb{R}^d)$. It admits $\mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}^d)$ as a core algebra. Its carré du champ is given by

$$\forall f, g \in \mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}^d), \quad \Gamma(f,g) = \nabla f \cdot \nabla g.$$

Example 2.48 (Reflected Brownian motion in a domain)**.** Let $\mathcal{X} = \Omega \subset \mathbb{R}^d$ be a nonempty open bounded and convex domain, and let $(X_t)_{t\geqslant 0}$ be the reflected Brownian motion in the domain Ω rescaled by a factor of $\sqrt{2}$. Its only invariant and reversible measure is the normalized uniform measure on Ω, $\mu = \mathrm{Leb}(\cdot \cap \Omega)/\mathrm{Leb}(\Omega)$. Its infinitesimal generator is the Laplacian operator on Ω, $\mathrm{L}f = \Delta f$, acting on functions $f \in \mathcal{C}^2(\Omega)$ satsfying Neumann boundary conditions on $\partial\Omega$. We remark that it admits $\mathcal{C}_{\mathrm{c}}^\infty(\Omega)$ as a core algebra. Its carré du champ is given by

$$\forall f, g \in \mathcal{C}_{\mathrm{c}}^\infty(\Omega), \quad \Gamma(f,g) = \nabla f \cdot \nabla g.$$

Example 2.49 (Reflected Brownian motion with negative drift)**.** Let $\mathcal{X} = \mathbb{R}_{>0}$, and let $(B_t)_{t\geqslant 0}$ be the standard Brownian motion on $\mathbb{R}$. For $t \geqslant 0$, define $Y_t := B_t - t$, and let $(X_t)_{t\geqslant 0}$ be the reflection at 0 of the process $(Y_t)_{t\geqslant 0}$. Its only invariant and reversible measure is the exponential measure on $\mathbb{R}_{>0}$, $\mathrm{d}\mu(x) = \exp(-x)\,\mathrm{d}x$. Its infinitesimal generator is the operator $\mathrm{L}f = f'' - f'$ acting on functions $f \in \mathcal{C}^2(\mathbb{R}_{>0})$ with $f'(0) = 0$. We remark that it admits $\mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}_{>0})$ as a core algebra. Its carré du champ is given by

$$\forall f, g \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}_{>0}), \quad \Gamma(f,g) = f'g'.$$

Example 2.50 (Ornstein-Uhlenbeck process)**.** Let $\mathcal{X} = \mathbb{R}^d$, and let $(B_t)_{t\geqslant 0}$ be the standard d-dimensional Brownian motion. The Ornstein-Uhlenbeck process $(X_t)_{t\geqslant 0}$ is the solution to the following stochastic differential equation:

$$\mathrm{d}X_t = \sqrt{2}\,\mathrm{d}B_t - X_t\,\mathrm{d}t, \quad X_0 = x.$$

Its only invariant and reversible measure is the standard Gaussian γ_d. Its semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ admits the following explicit representation for $t > 0$, known as the Mehler formula:

$$\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d), \forall x \in \mathbb{R}^d, \quad \mathrm{P}_t f(x) = \int_{\mathbb{R}^d} f\big(e^{-t}x + \sqrt{1 - e^{-2t}}y\big)\,\mathrm{d}\gamma_d(y). \tag{2.4.7}$$

Its infinitesimal generator is given by the elliptic operator $\mathrm{L}f = \Delta f - x \cdot \nabla f$, acting on functions $f \in \mathcal{C}^2(\mathbb{R}^d)$. It admits $\mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d)$ as a core algebra, and its carré du champ is given by

$$\forall f, g \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d), \quad \Gamma(f,g) = \nabla f \cdot \nabla g.$$

Examples 2.47 and 2.50 are particular cases of the following more general example.

Example 2.51 (Langevin diffusion on a manifold)**.** Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and let $(B_t)_{t\geqslant 0}$ be a standard Brownian motion on M. We have that μ is the invariant and reversible measure of the Langevin diffusion $(X_t)_{t\geqslant 0}$, which is the solution to the following stochastic differential equation:

$$\mathrm{d}X_t = \sqrt{2}\,\mathrm{d}B_t - \nabla W(X_t)\,\mathrm{d}t.$$

In general, its semigroup does not admit a closed-form expression, but that is not a limitation to compute its generator $\mathrm{L}f = \Delta f - \nabla W \cdot \nabla f$, where Δ and ∇ are the Laplace-Beltrami operator and the Riemannian gradient, respectively, acting on functions $f \in \mathcal{C}^\infty(M)$. It admits $\mathcal{C}_{\mathrm{c}}^\infty(M)$ as a core algebra. Its carré du champ is given by

$$\forall f, g \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \Gamma(f,g) = \nabla f \cdot \nabla g.$$

Example 2.52 (Markov chains)**.** Let $\mathcal{X}$ be a countable set endowed with the discrete topology. Let $\mathrm{L} = (\mathrm{L}_{ij})_{i,j\in\mathcal{X}}$ be a Q-matrix, i.e., for each $i, j \in \mathcal{X}$ with $i \neq j$, $\mathrm{L}_{ij} \geqslant 0$, and for every $i \in \mathcal{X}$, $\mathrm{L}_{ii} = -\sum_{j\neq i} \mathrm{L}_{ij}$. The matrix L is the generator of a continuous-time Markov chain $(X_t)_{t\geqslant 0}$ on the space $\mathcal{X}$, with semigroup given by the matrix exponential of L: for every $t \geqslant 0$, $\mathrm{P}_t = e^{t\,\mathrm{L}}$. An invariant measure μ for L, which can be identified with an array $(\mu_i)_{i\in\mathcal{X}}$ of nonnegative numbers, is such that for any $j \in \mathcal{X}$, $\mu_j = \sum_{i\in\mathcal{X}} \mu_i \,\mathrm{L}_{ij}$, and it is unique if the associated Markov chain is irreducible and recurrent. The measure μ is reversible if for every $i, j \in \mathcal{X}$, $\mu_i \,\mathrm{L}_{ij} = \mu_j \,\mathrm{L}_{ij}$. We can take as a core algebra the set of finitely-supported sequences:

$$\mathcal{A}_0 = \left\{(f_i)_{i\in\mathcal{X}} \in \mathbb{R}^{\mathcal{X}} : \mathrm{card}(\{i \in \mathcal{X} : f_i \neq 0\}) < +\infty\right\}.$$

The carré du champ is given by

$$\forall f, g \in \mathcal{A}_0, \forall i \in \mathcal{X}, \quad \Gamma(f, g)(i) = \sum_{j\in\mathcal{X}} \mathrm{L}_{ij}(f_i - f_j)(g_i - g_j).$$

Example 2.53 (M/M/∞ queue)**.** This is a particular case of the previous example. Let $\mathcal{X} = \mathbb{N}$ and fix a parameter $T > 0$. For $f \colon \mathbb{N} \to \mathbb{R}$, we adopt the convention $f(-1) = 0$. We define the operators D and D^* by

$$\forall k \in \mathbb{N}, \quad \mathrm{D}f(k) \coloneqq f(k+1) - f(k),$$

and

$$\forall k \in \mathbb{N}, \quad \mathrm{D}^*f(k) \coloneqq f(k-1) - f(k).$$

The M/M/∞ queue is the Markov chain given by the generator

$$\forall f \colon \mathbb{N} \to \mathbb{R}, \forall k \in \mathbb{N}, \quad \mathrm{L}f(k) \coloneqq \mathrm{D}f(k) + \frac{k}{T}\,\mathrm{D}^*f(k).$$

Then π_T, the Poisson distribution with parameter T, is the invariant and reversible measure. The carré du champ operator is given by

$$\forall f, g \colon \mathbb{N} \to \mathbb{R}, \quad \Gamma(f, g) = \frac{1}{2}\,\mathrm{D}f \cdot \mathrm{D}g + \frac{k}{2T}\,\mathrm{D}^*f \cdot \mathrm{D}^*g. \tag{2.4.8}$$

Now we define the diffusion property for a Markov triple.

Definition 2.54 (Diffusive Markov triple)**.** Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple with core algebra $\mathcal{A}_0$ and carré du champ operator Γ. We say that it is a diffusive Markov triple if for every $k \in \mathbb{N}^*$, and every smooth function $\Psi \colon \mathbb{R}^k \to \mathbb{R}$ with $\Psi(0) = 0$,

$$\forall f_1, \ldots, f_k, g \in \mathcal{A}_0, \quad \Gamma(\Psi(f_1, \ldots, f_k), g) = \sum_{i=1}^{k} \partial_i\Psi(f_1, \ldots, f_k)\Gamma(f_i, g).$$

In particular,

$$\forall f, g, h \in \mathcal{A}_0, \quad \Gamma(fg, h) = f\Gamma(g, h) + g\Gamma(f, h). \tag{2.4.9}$$

Remark 2.55.

(i) The Markov triples appearing in Examples 2.47, 2.48, 2.49, 2.50, and 2.51 satisfy the diffusion property.

(ii) Since the chain rule (2.4.9) is verified, we observe that the Markov triples introduced in Examples 2.52 and 2.53 do not satisfy the diffusive property in general.

2.4.2 Poincaré inequalities and notable examples

Now that we have reviewed all the theory of Markov semigroups, we can adequately define in the context of Markov triples what a Poincaré inequality is.

Definition 2.56 (Poincaré inequality)**.** Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple such that μ is a probability measure. Let $\mathcal{A}_0$ be the core algebra associated with L, and let Γ be its carré du champ operator. We say that the measure μ satisfies a Poincaré inequality if there exists a constant $C_{\mathrm{P}}(\mu) > 0$ such that

$$\forall f \in \mathcal{A}_0, \quad \mathrm{Var}_\mu(f) \leqslant C_{\mathrm{P}}(\mu) \int_{\mathcal{X}} \Gamma(f)\,\mathrm{d}\mu.$$

Now we will prove Theorem 2.35, the Gaussian Poincaré inequality, using a semigroup argument. Recall from example 2.50 that γ_d is the invariant and symmetric measure associated with the Ornstein-Uhlenbeck semigroup.

Proof of Theorem 2.35. Let $(\mathrm{P}_t)_{t\geqslant 0}$ be the Ornstein-Uhlenbeck semigroup in d dimensions, and let $f \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d)$. Define $\alpha\colon \mathbb{R}_+ \to \mathbb{R}$ by

$$\forall t \geqslant 0, \quad \alpha(t) := \mathrm{Var}_{\gamma_d}(\mathrm{P}_t f).$$

If we differentiate α, we obtain

$$\alpha'(t) = -2\int_{\mathbb{R}^d} \Gamma(\nabla \mathrm{P}_t f)\,\mathrm{d}\gamma_d = -2\int_{\mathbb{R}^d} |\nabla \mathrm{P}_t f|^2\,\mathrm{d}\gamma_d.$$

From Mehler's formula (2.4.7), we see that for all $x \in \mathbb{R}^d$, $\lim_{t\to+\infty} \mathrm{P}_t f(x) = \int_{\mathbb{R}^d} f\,\mathrm{d}\gamma_d$. Then, if we integrate in $t \geqslant 0$ the expression for α', we get

$$\mathrm{Var}_{\gamma_d}(f) = 2\int_0^{+\infty}\int_{\mathbb{R}^d} |\nabla \mathrm{P}_t f|^2\,\mathrm{d}\gamma_d.$$

On the other hand, again from (2.4.7), it can be shown that for $1 \leqslant i \leqslant d$,

$$\partial_i \mathrm{P}_t f(x) = e^{-t}\,\mathrm{P}_t(\partial_i f)(x).$$

Hence,

$$|\nabla \mathrm{P}_t f(x)|^2 \leqslant e^{-2t}\,\mathrm{P}_t(|\nabla f|^2). \tag{2.4.10}$$

Thus,

$$\mathrm{Var}_{\gamma_d}(f) \leqslant 2\int_0^{+\infty}\int_{\mathbb{R}^d} e^{-2t}\,\mathrm{P}_t(|\nabla f|^2)\,\mathrm{d}\gamma_d = 2\int_0^{+\infty}\int_{\mathbb{R}^d} e^{-2t}|\nabla f|^2\,\mathrm{d}\gamma_d = \int_{\mathbb{R}^d} |\nabla f|^2\,\mathrm{d}\gamma_d,$$

where we used the fact that μ is invariant for the semigroup. $\square$

Remark 2.57. Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$. From Example 2.51, we know that the Langevin generator $\mathrm{L} = \Delta - \nabla W \cdot \nabla$ admits μ as its reversible and invariant measure. The carré du champ is given by $\Gamma(f) = |\nabla f|^2$ for $f \in \mathcal{C}_c^\infty(M)$. Therefore, a Poincaré inequality for (M, L, μ) reads as

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \mathrm{Var}_\mu(f) \leqslant C_{\mathrm{P}}(\mu) \int_M |\nabla f|^2 \,\mathrm{d}\mu \tag{2.4.11}$$

for some $C_{\mathrm{P}}(\mu) > 0$. In simple words, for a given probability measure absolutely continuous with respect to the volume measure on a Riemannian manifold and with positive smooth density, there always exists a Markov generator L that allows us to match the functional inequality (2.4.11) with the Definition 2.56.

We provide some classical examples apart from Theorem 2.35. We start with the original type of inequality proven by Poincaré [Poi87, Poi90], which names this class of inequalities. The following sharp version was proven by Payne and Weinberger [PW60]. In this case, the associated Markov triple corresponds to the one associated with the reflected Brownian motion on a domain, recall Example 2.48.

Theorem 2.58 (Poincaré inequality on a convex domain)**.** *Let $\Omega \subset \mathbb{R}^d$ be a nonempty open bounded and convex domain, and denote by μ the normalized uniform measure on Ω. Define the diameter of Ω by* $\mathrm{diam}(\Omega) \coloneqq \sup_{x,y\in\Omega}|x-y|$*. Then the measure μ satisfies a Poincaré inequality:*

$$\forall f \in \mathcal{C}_c^\infty(\Omega), \quad \mathrm{Var}_\mu(f) \leqslant \frac{\mathrm{diam}(\Omega)^2}{\pi^2} \int_\Omega |\nabla f|^2 \,\mathrm{d}\mu.$$

Talagrand [Tal91] proved that the one-dimensional exponential distribution satisfies a Poincaré inequality. The respective Markov triple is the one introduced in Example 2.49.

Theorem 2.59 (Exponential Poincaré inequality)**.** *Let* $\mathrm{d}\mu(x) = \exp(-x)\,\mathrm{d}x$ *be the exponential measure on $\mathbb{R}_{>0}$. Then μ satisfies a Poincaré inequality:*

$$\forall f \in \mathcal{C}_c^\infty(\mathbb{R}_{>0}), \quad \mathrm{Var}_\mu(f) \leqslant 4 \int_{\mathbb{R}_{>0}} |f'|^2 \,\mathrm{d}\mu.$$

The following Poissonian Poincaré inequality is due to Klaassen [Kla85], where the associated Markov triple is the one induced by the M/M/∞ queue, see Example 2.53.

Theorem 2.60 (Poissonian Poincaré inequality)**.** *Fix $T > 0$, and let π_T be the Poisson measure of parameter T and consider the carré du champ given by* (2.4.8)*. Then π_T satisfies a Poincaré inequality with respect to Γ:*

$$\forall f \in L^2(\pi_T), \quad \mathrm{Var}_{\pi_T}(f) \leqslant T \sum_{k=0}^{+\infty} \Gamma(f)(k)\, \pi_T(k).$$

Remark 2.61. Sometimes, in the Poissonian setting, it is more natural to consider the bilinear form given by

$$\forall f, g\colon \mathbb{N} \to \mathbb{R}, \quad \tilde{\Gamma}(f,g) = \mathrm{D}f \cdot \mathrm{D}g \tag{2.4.12}$$

as an alternative to Γ. Note that

$$\forall f, g \colon \mathbb{N} \to \mathbb{R}, \quad \mathbb{E}_{\pi_T}[\Gamma(f,g)] = \mathbb{E}_{\pi_T}[\tilde{\Gamma}(f,g)],$$

so the Poincaré inequality of Theorem 2.60 is actually equivalent to

$$\forall f \in L^2(\pi_T), \quad \mathrm{Var}_{\pi_T}(f) \leqslant T \sum_{k=0}^{+\infty} |\mathrm{D} f(k)|^2 \, \pi_T(k).$$

2.4.3 Poincaré inequalities and their properties

We promised from the very beginning of this section that tensorization was one of the main features of Poincaré inequalities.

Proposition 2.62 (Tensorization)**.** *Let $n \in \mathbb{N}^*$. For $1 \leqslant i \leqslant n$, let $(\mathcal{X}_i, \mathrm{L}_i, \mu_i)$ be a Markov triple, assume that μ_i is a probability measure, and denote by $\mathcal{A}_0^i$ and Γ_i the core algebra and the carré du champ operator, respectively. Define the Markov triple $(\mathcal{X}, \mathrm{L}, \mu)$ by $\mathcal{X} \coloneqq \prod_{i=1}^n \mathcal{X}_i$, $\mu \coloneqq \bigotimes_{i=1}^n \mu_i$, and $\mathrm{L} \coloneqq \bigoplus_{i=1}^n \mathrm{L}_i$, which has a carré du champ operator given by $\Gamma \coloneqq \bigoplus_{i=1}^n \Gamma_i$. If each μ_i satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu_i) > 0$, then $(\mathcal{X}, \mathrm{L}, \mu)$ satisfies a Poincaré inequality with constant*

$$C_{\mathrm{P}}(\mu) \leqslant \max_{1 \leqslant i \leqslant n} C_{\mathrm{P}}(\mu_i).$$

We now arrive at another promised result, namely that a Poincaré inequality yields subexponential concentration. This result is due to Gromov and Milman [GM83].

Theorem 2.63 (Gromov-Milman)**.** *Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple. Assume that μ is a probability measure satisfying a Poincaré inequality with constant $C_{\mathrm{P}}(\mu) > 0$. Then the measure μ has subexponential concentration.*

Remark 2.64. If we blend together Proposition 2.62 and Theorem 2.63, we obtain the following result: if μ is a measure satisfying a Poincaré inequality, then for any $n \in \mathbb{N}^*$, the measure $\mu^n \coloneqq \bigotimes_{i=1}^n \mu$ is subexponential with a constant that does not depend on n. This is the dimension-free concentration property granted by a Poincaré inequality.

Suppose that the base space $\mathcal{X}$ is a metric space, so we may consider Lipschitz maps $T \colon \mathcal{X} \to \mathcal{X}$ with respect to the given metric. In the diffusive case, Lipschitz pushforwards preserve poincaré inequalities. Unfortunately, the diffusion property is crucial for the validity of the result. For simplicity, we state the following result in the smooth context.

Proposition 2.65. *Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and assume that $\mu \in \mathcal{P}(M)$. Let $T \colon M \to M$ be a K-Lipschitz map for some $K > 0$, and let $\nu \in \mathcal{P}(\mathcal{X})$ be the pushforward of μ by T, i.e., $\nu = T_\# \mu$. Let $\Gamma(f) = |\nabla f|^2$ be the usual carré du champ. If μ satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu)$, then ν satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\nu) \leqslant K^2 \, C_{\mathrm{P}}(\mu)$.*

Poincaré inequalities contain spectral information of the associated infinitesimal generator L. Namely, the validity of such an inequality means that the operator L has a spectral gap, so the related semigroup is ergodic, and we may quantify this property.

Theorem 2.66. *Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple, assume that μ is a probability measure and let $(\mathrm{P}_t)_{t\geqslant 0}$ be the associated semigroup. Moreover, suppose that μ satisfies a Poincaré inequality with constant $C_\mathrm{P}(\mu) > 0$. Then $(\mathrm{P}_t)_{t\geqslant 0}$ is ergodic:*

$$\forall f \in L^2(\mu), \quad \lim_{t\to+\infty} \mathrm{P}_t f = \int_{\mathcal{X}} f \,\mathrm{d}\mu \text{ in } L^2(\mu).$$

Furthermore,

$$\forall f \in L^2(\mu), \forall t \geqslant 0, \quad \mathrm{Var}_\mu(\mathrm{P}_t f) \leqslant e^{-2t/C_\mathrm{P}(\mu)} \,\mathrm{Var}_\mu(f). \tag{2.4.13}$$

A significant consequence of Theorem 2.66 is the following: suppose that we have a Markov process $(X_t)_{t\geqslant 0}$ with invariant and reversible measure μ, in the sense of Remark 2.38. If the measure μ satisfies a Poincaré inequality, then the process $(X_t)_{t\geqslant 0}$ will be ergodic, which in simple terms means that for any initial law $X_0 \sim \nu$, then the law of the random variable X_t will converge towards the invariant measure μ as t goes to infinity. Furthermore, (2.4.13) quantifies this convergence to the equilibrium as an exponentially fast one.

2.4.4 Log-concave measures

Up to this point, we have only given particular examples of measures satisfying a Poincaré inequality. In the Euclidean setting, there is a sufficient condition ensuring a Poincaré inequality based on a convexity criterion, namely log-concavity. In both sections 2.1.1 and 2.2.4 of this chapter, we have referred to log-concave measures, particularly in terms of their good isoperimetric properties; the fact that log-concavity implies a Poincaré inequality reflects this. As promises are made to keep, this is the moment to introduce this class of measures.

Definition 2.67 (Log-concave measures)**.** Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a Borel probability measure. We say that μ is log-concave if

$$\forall A, B \in \mathcal{B}(\mathbb{R}^d), \forall \lambda \in (0,1), \quad \mu(\lambda A + (1-\lambda)B) \geqslant \mu(A)^\lambda \mu(B)^{1-\lambda}. \tag{2.4.14}$$

The Brunn-Minkowski inequality, i.e., Theorem 2.3, in its equivalent form (2.1.2) provides the first examples of log-concave measures.

Example 2.68 (Uniform measure on a convex body)**.** If $K \subset \mathbb{R}^d$ is a convex body, let $\mu_K \in \mathcal{P}(\mathbb{R}^d)$ be the normalized uniform measure on K:

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \quad \mu_K(A) := \mathrm{Vol}_d(A \cap K)/\mathrm{Vol}_d(K).$$

Then μ_K is log-concave.

We also define the class of log-concave functions.

Definition 2.69 (Log-concave functions)**.** Let $f\colon \mathbb{R}^d \to \mathbb{R}_+$ be a nonnegative function. We say that f is log-concave if there exists $V\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ convex such that $f = e^{-V}$, under the convention $e^{-\infty} = 0$.

Indeed, both definitions 2.67 and 2.69 are intimately connected. The following theorem accounts for this fact.

Theorem 2.70. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$. If $\mu \ll \mathrm{Leb}$ and its density $f\colon \mathbb{R}^d \to \mathbb{R}_+$ is a log-concave function, then μ is a log-concave measure. Conversely, if μ is a log-concave measure and $\mu \ll \mathrm{Leb}$, then its density $f\colon \mathbb{R}^d \to \mathbb{R}_+$ is a log-concave function.*

The first assertion follows from the Prékopa-Leindler inequality (Theorem 2.5), while the second was proved by Borell [Bor75b].

From this moment, we will delve into the isoperimetric properties of log-concave measures. As we saw in Section 2.2, we can interpret a lower bound on the isoperimetric profile of a measure as a partial solution to the isoperimetric problem in the sense of Definition 2.10 and Remark 2.11. One of the simplest bounds for the isoperimetric profile one can imagine is a linear one. Cheeger [Che70] introduced this type of isoperimetric bound in the study of the spectral properties of the Laplace-Beltrami operator associated with a Riemannian manifold.

Definition 2.71 (Cheeger inequality)**.** Let $\mu \in \mathcal{P}(\mathbb{R}^d)$, and let $\mathcal{I}_\mu$ be its isoperimetric profile. We say that μ satisfies a Cheeger inequality if there exists a constant $\chi(\mu) > 0$ such that

$$\forall r \in [0,1], \quad \mathcal{I}_\mu(r) \geqslant \chi(\mu) \min\{r, 1-r\}. \tag{2.4.15}$$

Beyond the intuition of a simple linear bound for the isoperimetric profile of a measure, the exponential measure on the real line is shown to be a natural example of a measure having exactly an isoperimetric profile of the form (2.4.15), as we see in the following example. For further details, we refer to [Fou05, p. 114].

Example 2.72 (Exponential measure isoperimetry)**.** Let μ be the double-sided exponential measure on $\mathbb{R}$, that is, $\mathrm{d}\mu(x) = \frac{1}{2}\exp(-|x|)\,\mathrm{d}x$, and let $\mathcal{I}_\mu$ be its isoperimetric profile. It can be shown that half lines solve the isoperimetric problem associated with μ, that is, sets of the form $(-\infty, x]$ or $[x, +\infty)$ for $x \in \mathbb{R}$. On the other hand,

$$\mathcal{I}_\mu = F_\mu' \circ F_\mu^{-1},$$

where F_μ is the distribution function of μ. In this case, F_μ is explicit, so we get

$$\forall r \in [0,1], \quad \mathcal{I}_\mu(r) = \min\{r, 1-r\}.$$

Then Example 2.72 shows that we may consider Definition 2.71 as a model isoperimetric inequality based on the exponential case. In Section 2.2, we obtained the concentration inequalities in Corollaries 2.14 and 2.17 from the spherical and Gaussian isoperimetric inequalities, respectively, so we may wonder if it is possible to do the same but from Cheeger's inequality. Indeed, Cheeger [Che70] proved that the validity of inequality (2.4.15) entails the validity of a Poincaré inequality, which by Gromov-Milman's theorem implies subexponential concentration.

Theorem 2.73. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$, suppose that μ satisfies Cheeger's inequality with constant $\chi(\mu) > 0$, and define Cheeger's reciprocal constant as $\psi(\mu) := 1/\chi(\mu)$. Then μ satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu) \leqslant 4\,\psi(\mu)^2$.*

The good isoperimetric properties of log-concave measures stem from the fact that they satisfy Cheeger's inequality. This fact was proven by Kannan, Lovász, and Simonovits [KLS95], and Bobkov [Bob99].

Theorem 2.74. *Let μ be a Borel log-concave probability measure. Then μ satisfies a Cheeger inequality. In particular, μ satisfies a Poincaré inequality.*

In the log-concave case, Milman [Mil09] proved that the converse of Theorem 2.73 holds on the class of log-concave measures. That is, the log-concave Poincaré inequality is an isoperimetric property as strong as the log-concave Cheeger inequality.

Theorem 2.75 (Milman)**.** *There exists a universal constant $c > 0$ such that for any Borel log-concave probability measure μ, $c\,\psi(\mu)^2 \leqslant C_{\mathrm{P}}(\mu)$.*

Let us remark that Theorem 2.75 provides a universal constant, which does not depend on the dimension, to compare both constants. That is, studying the isoperimetric properties of log-concave measures is equivalent to studying the behavior of their Poincaré constants.

One may wonder what is the influence of the dimension of the ambient space on the isoperimetric properties of a log-concave measure. More precisely, the question is if it is possible to study the order of magnitude of the Poincaré constant with respect to the dimension of the ambient space. To do so, we need to normalize our log-concave measures in order to compare them: for example, for $\alpha > 0$ and $d \in \mathbb{N}^*$, let $\gamma_{d,\alpha} := \mathcal{N}(0, \alpha I_d)$, which is log-concave. By a rescaling argument, it is possible to show that $\gamma_{d,\alpha}$ satisfies a sharp Poincaré inequality with constant $C_{\mathrm{P}}(\gamma_{d,\alpha}) = 1/\alpha$, which explodes as α goes to 0. To avoid this issue, we will normalize measures with respect to their covariance matrix.

Definition 2.76 (Isotropic measure)**.** We say that a Borel probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ is isotropic if $\int_{\mathbb{R}^d} x \, \mathrm{d}\mu(x) = 0$ and for every $\theta \in \mathbb{R}^d$ with $|\theta| = 1$, $\int_{\mathbb{R}^d} |\theta \cdot x|^2 \, \mathrm{d}\mu(x) = 1$.

If $\mu \in \mathcal{P}(\mathbb{R}^d)$ is isotropic and log-concave, then we have the following direct lower bound on its Poincaré constant that follows from the Poincaré inequality applied to the linear forms $x \mapsto \theta \cdot x$:

$$1 \leqslant C_{\mathrm{P}}(\mu). \tag{2.4.16}$$

One of the biggest open questions in the field is the KLS conjecture, which roughly states that there is also an adimensional matching upper bound in (2.4.16).

Conjecture 2.77 (KLS conjecture)**.** *There exists a universal constant $c > 0$ such that for any isotropic Borel log-concave probability measure, its Poincaré constant is upper bounded by c.*

The currently best bound is of order $O(\log d)$, and it was found by Klartag [Kla23]. The KLS conjecture is a stronger statement than Bourgain's slicing problem [Bou86, Bou87], which was solved recently by Klartag and Lehec [KL24a].

2.5 Logarithmic Sobolev inequalities

Poincaré inequalities were the functional inequalities we used as a proxy to obtain dimension-free subexponential concentration. Logarithmic Sobolev inequalities play the same role for the subgaussian phenomenon. In this section, we define them and study their most essential properties and some examples. Finally, we introduce modified logarithmic Sobolev inequalities, a weaker kind of functional inequality that plays a central role in the discrete setting due to the lack of a full logarithmic Sobolev inequality in this context. The content of this section is based on the references [ABC$^+$00, BGL14, CL23].

2.5.1 Basic properties

As we have already introduced the language of Markov triples, we define logarithmic Sobolev inequalities in those terms.

Definition 2.78 (Logarithmic Sobolev inequality)**.** Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple such that μ is a probability measure. Let $\mathcal{A}_0$ be the core algebra associated with L, and let Γ be the carré du champ operator. We say that the measure μ satisfies a logarithmic Sobolev inequality if there exists a constant $C_{\mathrm{LS}}(\mu) > 0$ such that

$$\forall f \in \mathcal{A}_0, \quad \mathrm{Ent}_\mu(f^2) \leqslant C_{\mathrm{LS}}(\mu) \int_{\mathcal{X}} \Gamma(f)\, \mathrm{d}\mu.$$

As a first property, we have that logarithmic Sobolev inequalities are stronger than Poincaré inequalities, as proved by Rothaus [Rot81].

Proposition 2.79. *Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple such that μ is a probability measure. If μ satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) > 0$, then μ satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu) \leqslant C_{\mathrm{LS}}(\mu)$.*

These inequalities tensorize in the same fashion as Poincaré inequalities.

Proposition 2.80 (Tensorization)**.** *Let $n \in \mathbb{N}^*$. For $1 \leqslant i \leqslant n$, let $(\mathcal{X}_i, \mathrm{L}_i, \mu_i)$ be a Markov triple, assume that μ_i is a probability measure, and denote by $\mathcal{A}_0^i$ and Γ_i the core algebra and the carré du champ operator, respectively. Define the Markov triple $(\mathcal{X}, \mathrm{L}, \mu)$ by $\mathcal{X} := \prod_{i=1}^n \mathcal{X}_i$, $\mu := \bigotimes_{i=1}^n \mu_i$, and $\Gamma := \bigoplus_{i=1}^n \Gamma_i$. If each μ_i satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu_i) > 0$, then $(\mathcal{X}, \mathrm{L}, \mu)$ satisfies a logarithmic Sobolev inequality with constant*

$$C_{\mathrm{LS}}(\mu) \leqslant \max_{1 \leqslant i \leqslant n} C_{\mathrm{LS}}(\mu_i).$$

We arrive at one of the quintessential properties of these inequalities, namely the Herbst argument, which yields subgaussian concentration whenever a measure satisfies a logarithmic Sobolev inequality. Ira Herbst communicated this unpublished result in a letter addressed to Leonard Gross.

Theorem 2.81 (Herbst's argument)**.** *Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple. Assume that μ is a probability measure satisfying a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) > 0$. Then μ has subgaussian concentration.*

Remark 2.82. We can state an analog conclusion to the one commented in Remark 2.64 when a logarithmic Sobolev inequality holds for a measure μ: for any $n \in \mathbb{N}^*$, the product measure $\mu^n = \bigotimes_{i=1}^n \mu$ is subgaussian with a dimension-free constant.

Similarly to Proposition 2.65, logarithmic Sobolev inequalities are stable by Lipschitz pushforwards in the smooth case.

Proposition 2.83. *Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and assume that $\mu \in \mathcal{P}(M)$. Let $T\colon M \to M$ be a K-Lipschitz map for some $K > 0$ and let $\nu \in \mathcal{P}(\mathcal{X})$ be the pushforward of μ by T, i.e., $\nu = T_\#\mu$. Let $\Gamma(f) = |\nabla f|^2$ be the usual carré du champ. If μ satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu)$, then ν satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\nu) \leqslant K^2\, C_{\mathrm{LS}}(\mu)$.*

In the fashion of Theorem 2.66, logarithmic Sobolev inequalities yield the decay of the entropy along the semigroup exponentially fast in time.

Theorem 2.84. *Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple, assume that μ is a probability measure and let $(\mathrm{P}_t)_{t\geqslant 0}$ be the associated semigroup. Moreover, suppose that μ satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) > 0$. Then $(\mathrm{P}_t)_{t\geqslant 0}$ is ergodic:*

$$\forall f \in L^2(\mu), \quad \lim_{t\to+\infty} \mathrm{P}_t f = \int_{\mathcal{X}} f\,\mathrm{d}\mu \ \text{in } L^2(\mu).$$

Furthermore,

$$\forall f \in L^2(\mu), \forall t \geqslant 0, \quad \mathrm{Ent}_\mu(\mathrm{P}_t f) \leqslant e^{-2t/C_{\mathrm{LS}}(\mu)}\, \mathrm{Ent}_\mu(f). \tag{2.5.1}$$

Recall that any Markov semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ with invariant measure μ, by definition, is a contraction in $L^p(\mu)$, i.e., for all $p \geqslant 1$ real, and every $f \in L^p(\mu)$, then $\|\mathrm{P}_t f\|_p \leqslant \|f\|_p$. In the case when μ satisfies a logarithmic Sobolev inequality, the semigroup satisfies a stronger regularizing property denominated hypercontractivity: for every $f \in L^p(\mu)$, we have $\|\mathrm{P}_t f\|_q \leqslant \|f\|_p$ for some $q > p$, thus going beyond plain contractivity. Nelson [Nel66] realized that the Ornstein-Uhlenbeck semigroup satisfies the property, and then Gross [Gro75] proved the equivalence between a logarithmic Sobolev inequality and the hypercontractivity of the associated semigroup.

Theorem 2.85 (Gross)**.** *Let $(\mathcal{X}, \mathrm{L}, \mu)$ be a Markov triple, assume that μ is a probability measure, and let $(\mathrm{P}_t)_{t\geqslant 0}$ be the associated semigroup. Then μ satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) > 0$ if and only if $(\mathrm{P}_t)_{t\geqslant 0}$ is hypercontractive:*

$$\forall p > 1, \forall t \geqslant 0, \forall f \in L^p(\mu), \quad \|\mathrm{P}_t f\|_{q(t)} \leqslant \|f\|_p,$$

where $q(t) = 1 + (p-1)e^{4t/C_{\mathrm{LS}}(\mu)}$.

2.5.2 Examples

We provide some examples of logarithmic Sobolev inequalities, starting with the Gaussian case, proven independently by Stam [Sta59] and Gross [Gro75].

Theorem 2.86 (Gaussian logarithmic Sobolev inequality)**.** *The d-dimensional standard Gaussian measure satisfies a logarithmic Sobolev inequality:*

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}^d), \quad \operatorname{Ent}_{\gamma_d}(f^2) \leqslant 2 \int_{\mathbb{R}^d} |\nabla f|^2 \, \mathrm{d}\gamma_d. \tag{2.5.2}$$

There are many known proofs of this inequality; we will do a semigroup argument similar to how we proved the Gaussian Poincaré inequality of Theorem 2.35.

Proof of Theorem 2.86. First, note that by the chain rule and an approximation argument, the Gaussian logarithmic Sobolev inequality (2.5.2) is equivalent to

$$\forall f \in \mathcal{C}_{>0}^{\infty}(\mathbb{R}^d), \quad \operatorname{Ent}_{\gamma_d}(f) \leqslant \frac{1}{2} \int_{\mathbb{R}^d} \frac{|\nabla f|^2}{f} \, d\gamma_d, \tag{2.5.3}$$

where $\mathcal{C}_{>0}^{\infty}(\mathbb{R}^d)$ is the set of positive smooth functions. We will establish (2.5.3).

Let $(\mathrm{P}_t)_{t\geqslant 0}$ be the Ornstein-Uhlenbeck semigroup on $\mathbb{R}^d$, and let $f \in \mathcal{C}_{>0}^{\infty}(\mathbb{R}^d)$. If we differentiate $t \mapsto \operatorname{Ent}_{\gamma_d}(\mathrm{P}_t f)$, we obtain the so-called entropy production formula, or de Bruijn's identity:

$$\frac{\mathrm{d}}{\mathrm{d}t} \operatorname{Ent}_{\gamma_d}(\mathrm{P}_t f) = -\int_{\mathbb{R}^d} \frac{\Gamma(\mathrm{P}_t f)}{\mathrm{P}_t f} \, \mathrm{d}\gamma_d = -\int_{\mathbb{R}^d} \frac{|\nabla \mathrm{P}_t f|^2}{\mathrm{P}_t f} \, \mathrm{d}\gamma_d,$$

which integrated in $t \geqslant 0$ yields

$$\operatorname{Ent}_{\gamma_d}(f) = \int_0^{+\infty} \int_{\mathbb{R}^d} \frac{|\nabla \mathrm{P}_t f|^2}{\mathrm{P}_t f} \, \mathrm{d}\gamma_d.$$

Note that the Mehler formula (2.4.7) yields

$$\frac{|\nabla \mathrm{P}_t f|^2}{\mathrm{P}_t f} \leqslant e^{-2t} \, \mathrm{P}_t \left(\frac{|\nabla f|^2}{f} \right), \tag{2.5.4}$$

so inequality (2.5.3) follows by the same argument we used at the end of the proof of the Gaussian Poincaré inequality. □

Remark 2.87. Alternatively, a very short proof of Theorem 2.86 can be given from the functional Gaussian isoperimetric inequality, Theorem 2.18, thanks to an argument due to Beckner [Led00] based on a Taylor expansion. More generally, if a Borel measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ satisfies a functional Gaussian-type isoperimetric inequality with constant $C_{\mathrm{isop}}(\mu) > 0$, then it satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) \leqslant 2C_{\mathrm{isop}}(\mu)$. Thus, we confirm what we stated in Remark 2.22: isoperimetric inequalities are one of the strongest in the hierarchy of functional inequalities.

In the discrete setting, one of the most important examples is the two-point space $\mathcal{X} = \{0, 1\}$ endowed with the Bernoulli uniform measure $\mu(\{0\}) = \mu(\{1\}) = 1/2$. In this case, let us consider the following Q-matrix:

$$\mathrm{L} := \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix},$$

which is the generator of the symmetric continuous-time random walk on $\mathcal{X}$, a particular case of Example 2.52. For every function, the carré du champ is constant:

$$\forall f\colon \mathcal{X} \to \mathbb{R}, \forall i \in \mathcal{X}, \quad \Gamma(f)(i) = (f(0) - f(1))^2,$$

so the energy term reads

$$\forall f\colon \mathcal{X} \to \mathbb{R}, \quad \int_{\mathcal{X}} \Gamma(f)\,\mathrm{d}\mu = (f(0) - f(1))^2.$$

We have the following logarithmic Sobolev inequality on $\mathcal{X}$.

Theorem 2.88 (Logarithmic Sobolev inequality on $\{0,1\}$)**.** *Let μ be the uniform measure on $\mathcal{X} = \{0,1\}$. Then μ satisfies a logarithmic Sobolev inequality:*

$$\forall f\colon \mathcal{X} \to \mathbb{R}, \quad \mathrm{Ent}_\mu(f^2) \leqslant \frac{1}{2} \int_{\mathcal{X}} \Gamma(f)\,\mathrm{d}\mu.$$

The logarithmic Sobolev inequality on the two-point space was proven originally by Gross [Gro75], who combined it with Proposition 2.80 to extend the inequality to the hypercube $\{0,1\}^n$ and then obtain Theorem 2.86 using the central limit theorem.

Unfortunately, the validity of a logarithmic Sobolev inequality is not trivial and is not always the case. For example, it is not satisfied by the exponential measure.

Example 2.89. Recall the exponential measure on $\mathbb{R}_{>0}$. It does not concentrate at a subgaussian level, so a logarithmic Sobolev inequality is impossible, as it would contradict Theorem 2.81.

2.5.3 Modified logarithmic Sobolev inequalities

We recall from Section 2.3.3 that in the discrete setting, a remarkable concentration model was the Poissonian one. One of the main reasons why we have studied Poincaré and logarithmic Sobolev inequalities in this chapter is to obtain functional inequalities that have, at the same time, good tensorization properties and entail the phenomenon of concentration of measure so that we can obtain dimension-free bounds for tensorized measures. Then we wonder, is there any such model inequality regarding the Poissonian case?

We first recall from Remark 2.33 that the Poissonian concentration is slightly stronger than the subexponential one, so if we want to recover the sharp bound via a functional inequality, a Poincaré inequality is useless (recall Gromov-Milman's theorem). On the other hand, from Example 2.31, we see that the Poisson distribution does not satisfy a subgaussian concentration bound, so a logarithmic Sobolev inequality is impossible in the light of Herbst's argument. Despite these limitations, Bobkov and Ledoux [BL98] found an appropriate inequality between Poincaré and logarithmic Sobolev.

Theorem 2.90 (Bobkov and Ledoux's modified logarithmic Sobolev inequality)**.** *Let $T > 0$, and let π_T be the Poisson distribution of parameter T on $\mathbb{N}$. Then π_T satisfies* *the following modified logarithmic Sobolev inequality:*

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \operatorname{Ent}_{\pi_T}(f) \leqslant T\,\mathbb{E}_{\pi_T}\!\left[\frac{|\mathrm{D}f|^2}{f}\right]. \tag{2.5.5}$$

Remark 2.91. Note that in the smooth setting, a full logarithmic Sobolev inequality in the sense of Definition 2.78 is equivalent to an inequality of the form (2.5.5) thanks to the chain rule, which is not satisfied by the discrete derivative D. For example, in the Gaussian case, (2.5.5) holds with a constant of 1/2 for every positive and smooth function, see equation (2.5.3).

The inequality by Bobkov and Ledoux tensorizes properly and yields the sharp concentration for the Poisson distribution. However, Wu [Wu00] improved on Theorem 2.90, getting a stronger modified inequality that has the same desirable properties as the one by Bobkov and Ledoux.

Theorem 2.92 (Wu's modified logarithmic Sobolev inequality)**.** *Let $T > 0$, and let π_T be the Poisson distribution of parameter T on $\mathbb{N}$. Then π_T satisfies the following modified logarithmic Sobolev inequality:*

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \operatorname{Ent}_{\pi_T}(f) \leqslant T\,\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)], \tag{2.5.6}$$

where

$$\forall u > 0, \forall u+v > 0, \quad \Psi(u,v) \coloneqq (u+v)\log(u+v) - u\log u - (1+\log u)v.$$

Remark 2.93. Note that Wu's inequality is stronger than the one by Bobkov and Ledoux since

$$\forall u > 0, \forall u+v > 0, \quad \Psi(u,v) \leqslant \frac{|v|^2}{u}.$$

There is a large family of modified logarithmic Sobolev inequalities; see [BT06] for more information on the subject. However, for the Poisson distribution, Wu's inequality is the strongest; therefore, it will be our model inequality in the discrete setting.

Definition 2.94. Let $\mu \in \mathcal{P}(\mathbb{N})$. We say that μ satisfies Wu's modified logarithmic Sobolev inequality if there exists a constant $C_{\mathrm{MLS}}(\mu) > 0$ such that

$$\forall f\colon \mathbb{N} \to \mathbb{R}_{>0}, \quad \operatorname{Ent}_{\mu}(f) \leqslant C_{\mathrm{MLS}}(\mu)\,\mathbb{E}_{\mu}[\Psi(f, \mathrm{D}f)],$$

where

$$\forall u > 0, \forall u+v > 0, \quad \Psi(u,v) \coloneqq (u+v)\log(u+v) - u\log u - (1+\log u)v.$$

Now, we review the essential properties associated with Wu's inequality. First of all, it implies a Poincaré inequality.

Proposition 2.95. *Let $\mu \in \mathcal{P}(\mathbb{N})$ and suppose that it satisfies Wu's modified logarithmic Sobolev inequality with constant $C_{\mathrm{MLS}}(\mu) > 0$. Then it satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu) \leqslant C_{\mathrm{MLS}}(\mu)$ with respect to the carré du champ given by $\Gamma(f) = |\mathrm{D}f|^2$.*

Wu's inequality tensorizes in the same way as full logarithmic Sobolev inequalities do.

Proposition 2.96 (Tensorization)**.** *Let $n \in \mathbb{N}^*$. For $1 \leqslant i \leqslant n$, let $\mu_i \in \mathcal{P}(\mathbb{N})$. Define the probability measure $\mu := \bigotimes_{i=1}^n \mu_i$ on $\mathbb{N}^n$ and $\Gamma := \bigoplus_{i=1}^n \Gamma_i$, where $\Gamma_i(f) = |\mathrm{D}f|^2$, for $f\colon \mathbb{N} \to \mathbb{R}$. If each μ_i satisfies Wu's modified logarithmic Sobolev inequality with constant $C_{\mathrm{MLS}}(\mu_i) > 0$, then the measure μ satisfies Wu's modified logarithmic Sobolev inequality with constant*

$$C_{\mathrm{MLS}}(\mu) \leqslant \max_{1 \leqslant i \leqslant n} C_{\mathrm{MLS}}(\mu_i).$$

Herbst's argument can be adapted for this inequality in order to obtain concentration inequalities, which will be of subpoissonian type.

Theorem 2.97 (Discrete Herbst's argument)**.** *Let $\mu \in \mathcal{P}(\mathbb{N})$ and suppose that it satisfies Wu's modified logarithmic Sobolev inequality with constant $C_{\mathrm{MLS}}(\mu) > 0$. Then μ has subpoissonian concentration.*

Finally, it yields spectral properties for a semigroup having an invariant measure satisfying Wu's inequality, in the same spirit as Theorem 2.84.

Theorem 2.98. *Let $\mu \in \mathcal{P}(\mathbb{N})$, and let $(\mathrm{P}_t)_{t \geqslant 0}$ be a Markov semigroup defined on $\mathbb{N}$ with invariant measure μ and that satisfies the usual conditions. Suppose that μ satisfies Wu's modified logarithmic Sobolev inequality with constant $C_{\mathrm{MLS}}(\mu) > 0$. Then $(\mathrm{P}_t)_{t \geqslant 0}$ is ergodic:*

$$\forall f \in L^2(\mu), \quad \lim_{t \to +\infty} \mathrm{P}_t f = \mathbb{E}_\mu[f] \text{ in } L^2(\mu).$$

Furthermore, for every $f \in L^2(\mu)$,

$$\forall t \geqslant 0, \quad \mathrm{Ent}_\mu(\mathrm{P}_t f) \leqslant e^{-2t/C_{\mathrm{MLS}}(\mu)} \,\mathrm{Ent}_\mu(f). \tag{2.5.7}$$

Remark 2.99. Unfortunately, Wu's inequality is not preserved by Lipschitz pushforwards since the discrete derivative D does not satisfy the chain rule. We will elaborate on this issue and address it in Chapter 4.

2.6 Transport-entropy inequalities

Another class of functional inequalities related to the concentration phenomenon is constituted by transport-entropy inequalities. It is a diverse family of inequalities, but all its members satisfy an archetypal structure: given a fixed measure μ, they bound, for any measure ν, a function of a transport-based cost between μ and ν by the relative entropy of ν with respect to μ. Their power resides in the fact that they characterize certain concentration bounds for the measure μ, and sometimes, even dimension-free concentration. Good references in the subject are [Vil09, Chapter 22] and [GL10].

We start by defining the general structure of a transport-entropy inequality.

Definition 2.100 (Transport-entropy inequalities)**.** Let $\mathcal{X}$ be a nonempty Polish space, let $c\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ be a Borel function, let $\mu \in \mathcal{P}(\mathcal{X})$, and let $\alpha\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function. We say that μ satisfies an α-T_c transport-entropy inequality if

$$\forall \nu \ll \mu, \quad \alpha(\mathcal{T}_c(\nu,\mu)) \leqslant \mathrm{H}(\nu|\mu),$$

where $\mathcal{T}_c(\nu,\mu)$ is the value of the optimal transport problem in its Kantorovich formulation associated with μ, ν, and c, in the sense of Definition 1.8, and $\mathrm{H}(\cdot|\mu)$ denotes the relative entropy with respect to μ.

In the particular case when $(\mathcal{X}, d)$ is a nonempty complete and separable metric space, we say that μ satisfies an α-T_1 transport-entropy inequality if c is the 1-distance cost, introduced in Example 1.16:

$$\forall \nu \ll \mu, \quad \alpha(\mathrm{W}_1(\nu,\mu)) \leqslant \mathrm{H}(\nu|\mu);$$

and if, specifically, $\alpha = \frac{1}{C}(\cdot)^2$ for some $C > 0$, we just denote it by $\mathrm{T}_1(C)$:

$$\forall \nu \ll \mu, \quad \mathrm{W}_1^2(\nu,\mu) \leqslant C\,\mathrm{H}(\nu|\mu).$$

Now, if if c is the 2-distance cost, and $\alpha = \frac{1}{C}(\cdot)$ for some $C > 0$, we just denote it by $\mathrm{T}_2(C)$:

$$\forall \nu \ll \mu, \quad \mathrm{W}_2^2(\nu,\mu) \leqslant C\,\mathrm{H}(\nu|\mu).$$

The first example of a transport-entropy inequality we provide is the Csiszár-Kullback-Pinsker inequality [Csi67, Kul67, Pin64].

Theorem 2.101 (Csiszár-Kullback-Pinsker)**.** *Let $\mathcal{X}$ be a nonempty Polish space. Then*

$$\forall \mu, \nu \in \mathcal{P}(\mathcal{X}), \quad \|\mu - \nu\|_{\mathrm{TV}}^2 \leqslant \frac{1}{2}\,\mathrm{H}(\nu|\mu). \tag{2.6.1}$$

Inequality (2.6.1) corresponds to an α-T_c transport-entropy inequality for $\alpha(r) \coloneqq 8r^2$ and c being the Hamming cost, see Example 1.15 in Chapter 1.

Concerning $\mathrm{T}_1(C)$ inequalities, Bobkov and Götze [BG99] proved that they characterize the subgaussian concentration phenomenon.

Theorem 2.102 (Bobkov-Götze)**.** *Let $(\mathcal{X}, d)$ be a nonempty complete and separable metric space, fix $\mu \in \mathcal{P}_1(\mathcal{X})$, and let $C > 0$. Then the following are equivalent:*

(i) The measure μ satisfies a $\mathrm{T}_1(C)$ inequality.

(ii) The measure μ is subgaussian.

About the family of α-T_1 inequalities. Gozlan and Léonard [GL07] proved that they can be characterized as a bound of the log-Laplace transform of any Lipschitz observable. We remark that this kind of bound helps get concentration inequalities.

Theorem 2.103 (Gozlan-Léonard)**.** *Let $(\mathcal{X}, d)$ be a nonempty complete and separable metric space, fix $\mu \in \mathcal{P}_1(\mathcal{X})$, and let $\alpha\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function. Then the following are equivalent:*

(i) The measure μ satisfies an α-T_1 transport-entropy inequality.

(ii) For every 1*-Lipschitz function* $f\colon \mathcal{X} \to \mathbb{R}$*,*

$$\forall r > 0, \quad \log\left(\int_{\mathcal{X}} \exp\left(r(f - \langle \mu, f\rangle)\right) \mathrm{d}\mu\right) \leqslant \alpha^*(r),$$

where α^* *is the Legendre transform of* α*.*

An example of a measure satisfying an α-T_1 inequality is the Poisson distribution on $\mathbb{N}$, as it was shown by Liu [Liu11].

Theorem 2.104. *Fix* $T > 0$ *and let* $\alpha\colon \mathbb{R}_+ \to \mathbb{R}_+$ *be the function defined as* $\alpha(r) := (r+T)\log((r+T)/T) - r$ *for* $r \geqslant 0$*. Then the Poisson distribution* π_T *of parameter* $T > 0$ *on* $\mathbb{N}$ *satisfies an* α*-*T_1 *inequality.*

Remark 2.105. If we combine Theorems 2.103 and 2.104, we recover the sharp Poissonian concentration seen in Example 2.31.

Now, note that as a consequence of Hölder's inequality, a $\mathrm{T}_2(C)$ inequality is stronger than $\mathrm{T}_1(C)$, so we may wonder if under $\mathrm{T}_2(C)$ it is possible to obtain an improvement on Theorem 2.102. Indeed, Gozlan [Goz09] proved that these inequalities characterize dimension-free concentration. Recall the notation $\mu^n := \bigotimes_{i=1}^n \mu$.

Theorem 2.106 (Gozlan)**.** *Let* $(\mathcal{X}, d)$ *be a nonempty complete and separable metric space, fix* $\mu \in \mathcal{P}_2(\mathcal{X})$*, and let* $C > 0$*. Then the following are equivalent:*

(i) The measure μ *satisfies a* $\mathrm{T}_2(C)$ *inequality.*

(ii) For every $n \in \mathbb{N}^*$*, the measure* μ^n *satisfies* $\mathrm{T}_1(C)$*.*

(iii) For every $n \in \mathbb{N}^*$*, the measure* μ^n *is subgaussian with constants that do not depend on* n*.*

An example of a measure satisfying a $\mathrm{T}_2(C)$ inequality is the Gaussian measure, as shown by Talagrand [Tal96].

Theorem 2.107 (Talagrand)**.** *The* d*-dimensional Gaussian measure satisfies a* $\mathrm{T}_2(2)$ *inequality.*

We remark that T_1 does not imply T_2.

Example 2.108. Let $\mathcal{X} = \{0, 1\}$ endowed with the trivial distance and the Bernoulli uniform measure μ. Then μ satisfies $\mathrm{T}_1(1/2)$. However, μ does not satisfy a T_2 inequality [Gen01].

Recall from Theorem 2.81 that a logarithmic Sobolev inequality yields subgaussian tails, so in the light of Theorem 2.106, it will imply, in turn, a T_2 inequality. On the other hand, T_2 inequalities yield Poincaré inequalities. These two statements were proved by Otto and Villani [OV00], so we call them the Otto-Villani theorem. We state them in the smooth setting for simplicity, although more general versions and different proofs are available [BGL01, GL13, GRS14, FGJ17].

Theorem 2.109 (Otto-Villani)**.** *Let* (M, g, μ) *be a connected and complete weighted Riemannian manifold, and assume that* $\mu \in \mathcal{P}(M)$*.*

(i) If μ satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu)$, then it satisfies a $\mathrm{T}_2(C)$ inequality with constant $C \leqslant C_{\mathrm{LS}}(\mu)$.

(ii) If μ satisfies a $\mathrm{T}_2(C)$ inequality, then it satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu) \leqslant C/2$.

Example 2.110. An example of a measure satisfying $\mathrm{T}_2(C)$ for some $C > 0$ that does not satisfy a logarithmic Sobolev inequality is given by the measure $\mu_\beta \in \mathcal{P}(\mathbb{R})$ defined as $\mathrm{d}\mu_\beta(x) = \exp(-V_\beta(x))\,\mathrm{d}x$, where

$$V_\beta(x) = x^3 + 3x^2 \sin^2(x) + x^\beta$$

for $\beta \in (2, 5/2)$, see [CG06].

Similarly to the previously introduced families of functional inequalities, T_p inequalities are stable by Lipschitz pushforwards [DGW04].

Proposition 2.111. *Let $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$ be two nonempty complete and separable metric spaces. Let $\mu \in \mathcal{P}(M)$. Let $T\colon \mathcal{X} \to \mathcal{Y}$ be a K-Lipschitz map for some $K > 0$ and let $\nu \in \mathcal{P}(\mathcal{X})$ be the pushforward of μ by T, i.e., $\nu = T_\#\mu$. If μ satisfies $\mathrm{T}_p(C)$ for some $p \in \{1, 2\}$ and $C > 0$, then ν satisfies $\mathrm{T}_p(K^2C)$.*

2.7 The Bakry-Émery criterion

In the last three sections of this chapter, we have reviewed three prominent families of functional inequalities that are deeply connected to the phenomenon of concentration of measure. In particular, Section 2.4 was devoted to Poincaré inequalities, where log-concavity was a straightforward sufficient condition to ensure their validity in the Euclidean setting. However, in the case of logarithmic Sobolev and transport-entropy inequalities, we have not presented any similar result yet. In this section, we will review the Bakry-Émery criterion, which provides sufficient conditions in the smooth setting for the validity of those functional inequalities. Besides that, the Bakry-Émery criterion will play a fundamental role in Chapter 3. Concerning this topic, the primary reference we used is [BGL14].

For simplicity, we will discuss the smooth setting. Still, obtaining the same results in more general structures is possible; see [BGL14, Chapter 3]. The main object appearing in the seminal contribution [BE85] by Bakry and Émery is the iterated carré du champ operator.

Definition 2.112 (Iterated carré du champ)**.** Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$. Let L be its associated Langevin diffusion operator, and Γ be the associated carré du champ. We define Γ_2, the iterated carré du champ operator, as

$$\forall f, h \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \Gamma_2(f, h) \coloneqq \frac{1}{2}\left(\mathrm{L}\Gamma(f, h) - \Gamma(h, \mathrm{L}f) - \Gamma(f, \mathrm{L}h)\right). \tag{2.7.1}$$

By abuse of notation, similarly to Γ, we define $\Gamma_2(f) \coloneqq \Gamma_2(f, f)$ for any $f \in \mathcal{C}^\infty_{\mathrm{c}}(M)$.

Recall that if $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$, then the Langevin diffusion operator takes the form $\mathrm{L} = \Delta - \nabla W \cdot \nabla$. Since we are in the Riemannian setting, Bochner's formula allows for an explicit computation of Γ_2 in terms of differential operators on (M, g) and the potential W. For further reference on Bochner's formula, we refer to [GHL04, Section 4.B].

Theorem 2.113 (Bochner's formula)**.** *Let (M, g) be a Riemannian manifold. Then*

$$\forall f \in \mathcal{C}^\infty(M), \quad \frac{1}{2}\Delta(|\nabla f|^2) = |\nabla^2 f|^2_{\mathrm{HS}} + \nabla f \cdot \nabla(\Delta f) + \mathrm{Ric}(\nabla f, \nabla f).$$

Corollary 2.114. *Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$. Let Γ_2 be its associated iterated carré du champ. Then*

$$\forall f \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \Gamma_2(f) = |\nabla^2 f|^2_{\mathrm{HS}} + \nabla^2 W(\nabla f, \nabla f) + \mathrm{Ric}(\nabla f, \nabla f).$$

Now we introduce the key idea appearing in Bakry and Émery's work: the curvature-dimension condition, which is a minoration of the operator Γ_2 by Γ.

Definition 2.115 (Curvature-dimension condition)**.** Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$. Let L be its associated Langevin diffusion operator, and let Γ and Γ_2 be its associated carré du champ and iterated carré du champ. Let $\rho \in \mathbb{R}$. We say that (M, g, μ) satisfies the curvature-dimension condition $\mathrm{CD}(\rho, \infty)$ if

$$\forall f \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \Gamma_2(f) \geqslant \rho\,\Gamma(f). \tag{2.7.2}$$

Remark 2.116. For the weighted manifold $(M, g, \exp(-W)\,\mathrm{dVol})$, we define the Bakry-Émery-Ricci tensor $\mathrm{Ric}_W := \mathrm{Ric} + \nabla^2 W$. Then, via Corollary 2.114, we see that $\mathrm{CD}(\rho, \infty)$ holds if and only if $\mathrm{Ric}_W \succeq \rho g$. More generally, the Bakry-Émery-Ricci tensor is an object of interest in itself in the context of weighted manifolds as it encodes the structure of the weight measure $e^{-W}\,\mathrm{dVol}$ into a tensor. Assuming bounds on Ric_W also implies topological and geometrical properties; see, for example, [Lot03].

The curvature-dimension condition is actually equivalent to some heat kernel estimates and commutations for the associated semigroup. That is, a bound on the curvature of a weighted Riemannian manifold entails analytical properties for the heat kernel.

Theorem 2.117. *Let (M, g, μ) be a connected and complete weighted Riemannian manifold, and let $(\mathrm{P}_t)_{t \geqslant 0}$ be its associated Markov semigroup. Let $\rho \in \mathbb{R}$. Then the following assertions are equivalent:*

(i) $\mathrm{CD}(\rho, \infty)$ *holds.*

(ii) The following commutation holds:

$$\forall t \geqslant 0, \forall f \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \Gamma(\mathrm{P}_t f) \leqslant e^{-2\rho t}\,\mathrm{P}_t(\Gamma(f)).$$

(iii) The following strong commutation holds:

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \sqrt{\Gamma(\mathrm{P}_t f)} \leqslant e^{-\rho t}\, \mathrm{P}_t(\sqrt{\Gamma(f)}).$$

(iv) The local Poincaré inequality holds: if $\rho \neq 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \mathrm{P}_t(f^2) - (\mathrm{P}_t f)^2 \leqslant \frac{1 - e^{-2\rho t}}{\rho}\, \mathrm{P}_t(\Gamma(f));$$

if $\rho = 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \mathrm{P}_t(f^2) - (\mathrm{P}_t f)^2 \leqslant 2t\, \mathrm{P}_t(\Gamma(f)).$$

(v) The local logarithmic Sobolev inequality holds: if $\rho \neq 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \mathrm{P}_t(f^2 \log f^2) - \mathrm{P}_t(f^2) \log \mathrm{P}_t(f^2) \leqslant 2\frac{1 - e^{-2\rho t}}{\rho}\, \mathrm{P}_t\left(\Gamma(f)\right);$$

if $\rho = 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(M), \quad \mathrm{P}_t(f^2 \log f^2) - \mathrm{P}_t(f^2) \log \mathrm{P}_t(f^2) \leqslant 4t\, \mathrm{P}_t\left(\Gamma(f)\right).$$

Recall the proofs of Theorems 2.35 and 2.86: both used a semigroup argument based on the properties of the Ornstein-Uhlenbeck semigroup. We can easily note that in both proofs the specific structure of the semigroup, namely Mehler's formula (2.4.7), was used only to get inequalities (2.4.10) and (2.5.4), respectively; the rest of both proofs is valid for any Markov semigroup. Now note that those specific bounds actually correspond to the commutations (ii) and (iii) in Theorem 2.117. That is, the Bakry-Émery criterion implies both Poincaré and logarithmic Sobolev inequalities. Note that the theorem holds under the $\mathrm{CD}(\rho, \infty)$ condition for some $\rho > 0$.

Theorem 2.118 (Bakry-Émery)**.** *Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^{\infty}(M)$, and suppose additionally that $\mu \in \mathcal{P}(M)$. If there exists $\rho > 0$ such that the curvature-dimension condition $\mathrm{CD}(\rho, \infty)$ holds, then:*

(i) μ satisfies a Poincaré inequality with constant $C_{\mathrm{P}}(\mu) \leqslant 1/\rho$.

(ii) μ satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) \leqslant 2/\rho$.

Actually, there is a more general inequality in the context of a weighted Riemannian manifold, the HWI inequality, proven originally by Otto and Villani [OV00]. It interpolates between the entropy H, the Fisher information I, and the Wasserstein distance W; that is the reason why for its name. If $\mu \in \mathcal{P}(M)$ and $\nu \ll \mu$ with a positive density $f \colon M \to \mathbb{R}_{>0}$, we define the relative Fisher information of ν with respect to μ as

$$\mathrm{I}(\nu|\mu) := \int_M \frac{|\nabla f|^2}{f}\, \mathrm{d}\mu. \tag{2.7.3}$$

Theorem 2.119 (HWI inequality)**.** *Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and suppose additionally that $\mu \in \mathcal{P}(M)$. If there exists $\rho \in \mathbb{R}$ such that the curvature-dimension condition $\mathrm{CD}(\rho, \infty)$ holds, then*

$$\forall \nu \ll \mu, \quad \mathrm{H}(\nu|\mu) \leqslant \mathrm{W}_2(\nu, \mu)\sqrt{\mathrm{I}(\nu|\mu)} - \frac{\rho}{2}\,\mathrm{W}_2^2(\nu, \mu).$$

Remark 2.120. Note that if $\rho > 0$ and the $\mathrm{CD}(\rho, \infty)$ condition holds, then if we apply Young's inequality together with Theorem 2.119, we recover Theorem 2.118.

In the context of Otto calculus, see Section 1.3.4 in Chapter 1, the HWI inequality can be considered as an expression of the strong convexity of the relative entropy functional $\mathcal{G}$ appearing in Example 1.29 under the curvature-dimension condition, along the geodesics of the formal Riemannian structure given to $\mathcal{P}_2(M)$. This unrigorous argument was given by Otto and Villani in [OV00] on top of the true proof to justify the heuristics of the proof of both the HWI inequality and the Otto-Villani theorem. We remark that these ideas were implemented rigorously by Gigli and Ledoux [GL13] in the nonsmooth setting. On the other hand, for the same reason, the HWI inequality plays a key role in the synthetic characterization of bounded Ricci curvature for metric spaces.

Part II

Original contributions

Chapter 3

The diffusion transport map

> Los niños habían de recordar el resto de su vida la augusta solemnidad con que su padre se sentó a la cabecera de la mesa, temblando de fiebre, devastado por la prolongada vigilia y por el encono de su imaginación, y les reveló su descubrimiento:
> —La tierra es redonda como una naranja.
>
> Gabriel García Márquez
> *Cien años de soledad*

The first original contribution of this thesis belongs to the theory of functional inequalities, and is an adaptation of the article [LR25a]. More precisely, the main result of this chapter is that if we have a weighted Riemannian manifold that has bounded curvature at first and second order in the sense of Bakry-Émery, then the Kim-Milman transport map between the weighted measure and any log-Lipschitz perturbation of it is Lipschitz, result which in turn allows the transfer of many families of functional inequalities, as we saw in Chapter 2. The interplay between the theories of functional inequalities and optimal transport plays a crucial role here.

After the motivation we will provide in Section 3.1, we introduce all the necessary preliminary notions in Section 3.2, mainly concerning diffusion operators on Riemannian manifolds. After that, in Section 3.3, we study the consequences of a second-order version of Bakry and Émery's curvature-dimension condition, namely $\Gamma_3 \geqslant \rho\,\Gamma_2$. In Section 3.4, we detail the construction of the diffusion transport map and prove that it is Lipschitz under our assumptions, which is the main result of this chapter, Theorem

3.21. Finally, we provide some applications of the main result in Section 3.5.

3.1 Introduction

We saw in the previous chapter that the d-dimensional Gaussian space $(\mathbb{R}^d, |\cdot|, \gamma_d)$ is of central importance in the theory of functional inequalities since it satisfies fundamental model inequalities. For example, recall Theorem 2.86, the Gaussian logarithmic Sobolev inequality:

$$\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d), \quad \mathrm{Ent}_{\gamma_d}(f^2) \leqslant 2 \int_{\mathbb{R}^d} |\nabla f|^2 \, \mathrm{d}\gamma_d.$$

Given a probability measure ν on $\mathbb{R}^d$ and a C-Lipschitz map $T\colon \mathbb{R}^d \to \mathbb{R}^d$ with $T_\# \gamma_d = \nu$, Proposition 2.83 says that also ν satisfies a logarithmic Sobolev inequality. We did not exhibit a proof, but we can demonstrate the result in a few lines: let $g \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d)$ be a test function, and apply the Gaussian inequality to $f := g \circ T$:

$$\mathrm{Ent}_\nu(g) = \mathrm{Ent}_{\gamma_d}(g \circ T) \leqslant 2 \int_{\mathbb{R}^d} |\nabla (g \circ T)|^2 \, \mathrm{d}\mu \leqslant 2C^2 \int_{\mathbb{R}^d} |\nabla g|^2 \, \mathrm{d}\nu, \tag{3.1.1}$$

where in the last inequality we used the chain rule, the Cauchy-Schwarz inequality, and the fact that T is a C-Lipschitz map. Note that the constant does not depend on the dimension d, so the adimensional nature of the Gaussian logarithmic Sobolev inequality is also preserved for the new inequality. We remark that the same argument applies to isoperimetric, concentration, and Poincaré inequalities; see Propositions 2.20, 2.24, and 2.65, respectively.

Recall Theorem 1.26, the Caffarelli contraction theorem, in the context of the quantitative regularity theory of optimal transport: it states that the Brenier map pushing forward the Gaussian towards any log-concave perturbation of it is 1-Lipschitz. Thus, Caffarelli's theorem is helpful if we want to obtain novel functional inequalities; see [CE02, Har04] for some examples beyond the classical families of functional inequalities mentioned before. There are some extensions to Caffarelli's theorem. For example, in [CFJ17], it was proven that the Brenier map between a uniformly log-concave measure and a compactly supported perturbation of it is Lipschitz, again with a constant that does not depend on the dimension. Unfortunately, we point out that in the Riemannian setting, there are obstructions to getting a similar result [FFGZ24].

As it was seen in the transport of the logarithmic Sobolev inequality (3.1.1), the optimal character of the transport map in the sense of Monge's problem (OT Monge) does not play any role in the proof. In particular, in this note, we prove the existence of a Lipschitz transport map, not necessarily optimal [Tan21, LS22], between a measure on a manifold and a log-Lipschitz perturbation of it, using a construction originally developed in [KM12] in the Euclidean setting, which traces its origins back to [OV00], adapted to the smooth setting in [FMS24], that we call the Kim-Milman or the diffusion transport map since its definition relies on the interpolation between two measures induced by an ad hoc diffusion process. This transport map was also revisited recently in [KP23, Nee22, MS23, CE25]. Similar schemes that use Polchinski's flow and multiscale Bakry-Émery criteria have been recently explored in [Ser24] and [She24].

Let us explain briefly the construction of the diffusion transport map: consider a weighted Riemannian manifold (M, g, μ) with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W : M \to \mathbb{R}$, and assume that $\mu \in \mathcal{P}(M)$. Let $V : M \to \mathbb{R}$ be a K-Lipschitz potential, so we define the log-perturbation of μ by V as the measure $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\mu$, and suppose that $\nu \in \mathcal{P}(M)$. Recall Example 2.51 in Chapter 2: we may define the Langevin diffusion $(X_t)_{t\geqslant 0}$, which is the solution of the following stochastic differential equation

$$\mathrm{d}X_t = \sqrt{2}\,\mathrm{d}B_t - \nabla W(X_t)\,\mathrm{d}t, \quad X_0 \sim \nu, \tag{3.1.2}$$

where $(B_t)_{t\geqslant 0}$ is the Brownian motion on the manifold (M, g). The method of characteristics applied to the Fokker-Planck equation associated with the law of the process $(X_t)_{t\geqslant 0}$ permits the construction of a flow of diffeomorphisms that in turn provides a map $T : M \to M$ such that $T_{\#}\mu = \nu$, so T is the desired diffusion transport map. Moreover, if one is able to provide appropriate second-order bounds on the heat kernel associated with (3.1.2), then the map T is C-Lipschitz, for some explicit $C > 0$; see Lemma 3.20 for the precise statement. That is, the regularity of the diffusion transport map can be studied via the analysis of the associated heat kernel, which offers a straightforward alternative to the Brenier map and its regularity theory, recall Section 1.3.2 of Chapter 1.

In order to provide systematic criteria ensuring the Lipschitz regularity of the diffusion transport map, we exploit the classical Bakry-Émery criterion and a further iteration that we will define precisely in Section 3.3. These ingredients will provide the second-order heat kernel bounds mentioned above that grant the desired Lipschitz regularity. More precisely, let L be the infinitesimal generator associated with (3.1.2). Invoking Section 2.4.1 of Chapter 2, we recall the respective carré du champ operator Γ and its iteration Γ_2, see Section 2.7 of the same chapter. We may reiterate this scheme and define Γ_3 inductively. Our main result states that if there exist positive constants $\rho_1, \rho_2 > 0$ such that $\Gamma_2 \geqslant \rho_1\,\Gamma$ and $\Gamma_3 \geqslant \rho_2\,\Gamma_2$, then the diffusion transport map is C-Lipschitz, where $C = C(\rho_1, \rho_2, K)$. Note that the constant does not depend on the dimension of the manifold M. The following theorem is the precise statement of these ideas, corresponding to Theorem 3.21 in this note.

Theorem. *Let (M, g, μ) be a complete and connected weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and assume that $\mu \in \mathcal{P}(M)$. Consider the diffusion operator $\mathrm{L} = \Delta - \nabla W \cdot \nabla$, let Γ be its associated carré du champ, and let Γ_2 and Γ_3 be its respective iterations in the Bakry-Émery sense. Assume that there exist constants $\rho_1, \rho_2 > 0$ such that*

(i) $\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \Gamma_2(f) \geqslant \rho_1\,\Gamma(f)$; and

(ii) $\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \Gamma_3(f) \geqslant \rho_2\,\Gamma_2(f)$.

Let $V \in \mathcal{C}^\infty(M)$, and assume that it is K-Lipschitz for some $K > 0$. Define $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\mu$ and assume that $\nu \in \mathcal{P}(M)$. Then there exists a Lipschitz map $T : M \to M$ pushing forward μ towards ν which is $\exp\left(\sqrt{\frac{2\pi}{\rho_2}} K e^{\frac{K^2}{2\rho_1}}\right)$-Lipschitz.

The presence of Γ_3 is not new in the literature: it has been explored previously in [Led93, Led95, Bak94, LNP15]. In particular, in [LNP15], the condition $\Gamma_3(f) \geqslant \rho_2 \Gamma_2(f)$ was employed to obtain quantitative regularity second order estimates for the solutions of the partial differential equation

$$\Delta f - \nabla W \cdot \nabla f = V - \int_M V \,\mathrm{d}\mu \tag{3.1.3}$$

for Lipschitz data V. Note that (3.1.3) corresponds to the linearization of the Monge-Ampère equation associated with the quadratic optimal transport problem between μ and ν. On the other hand, we also remark that (3.1.3) is the linearization of the diffusion transport map [FMS24].

3.2 Preliminaries and notations

In this section, we introduce all the main objects on which this note is based and their preliminary properties. We first review the basics of diffusion operators on weighted manifolds and then recall their probabilistic interpretation.

3.2.1 Markov diffusion generators on a manifold

Let (M, g) be a d-dimensional smooth Riemannian manifold that we assume to be complete and connected unless otherwise stated, and let us denote its tangent bundle by TM. Let d_g be the geodesic distance, let Vol be the volume measure, and let Ric be the Ricci curvature tensor. We denote by ∇, ∇^2, and Δ the Riemannian gradient, the Hessian, and the Laplace-Beltrami operator, respectively.

We equip (M, g) with a probability measure $\mu \in \mathcal{P}(M)$ of the form

$$\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$$

for some $W \in \mathcal{C}^\infty(M)$ so we get a weighted Riemannian manifold (M, g, μ), and we define the associated Langevin elliptic diffusion operator L by

$$\forall f \in \mathcal{C}^\infty(M), \quad \mathrm{L}f \coloneqq \Delta f - \nabla W \cdot \nabla f. \tag{3.2.1}$$

In Section 2.4.1 of Chapter 2, we introduced the language of Markov semigroups and their main properties. We started with a semigroup and built the theory on top of it; in particular, given a semigroup, we could construct its infinitesimal generator. In what follows, we will proceed inversely: our starting point will be the operator L defined by (3.2.1), and we will see how we can express the same properties that we had in the former case, now in terms of L. This is practical since, in general, we will not have access to a concrete expression for the semigroup associated with L.

First, we note that the measure μ is invariant for L:

$$\forall f \in \mathcal{C}_\mathrm{c}^\infty(M), \quad \int_M \mathrm{L}f \,\mathrm{d}\mu = 0. \tag{3.2.2}$$

Moreover, we can see that the following integration by parts formula holds: for all $f, h \in \mathcal{C}_c^\infty(M)$,

$$\forall f, h \in \mathcal{C}_c^\infty(M), \quad \int_M f \, \mathrm{L} h \, \mathrm{d}\mu = -\int_M \nabla f \cdot \nabla h \, \mathrm{d}\mu, \tag{3.2.3}$$

which in turn implies that L is symmetric or reversible with respect to the measure μ:

$$\forall f, h \in \mathcal{C}_c^\infty(M), \quad \int_M f \, \mathrm{L} h \, \mathrm{d}\mu = \int_M h \, \mathrm{L} f \, \mathrm{d}\mu. \tag{3.2.4}$$

From the perspective of operator theory, L can be regarded as an unbounded linear operator with domain $\mathcal{D}(\mathrm{L})$. Note that $\mathcal{C}_c^\infty(M) \subset \mathcal{D}(\mathrm{L}) \subset L^2(\mu)$. Moreover, our assumptions entail its essential self-adjointness on the algebra $\mathcal{C}_c^\infty(M)$, see [BGL14, Proposition 3.2.1]; in particular, $\mathcal{C}_c^\infty(M)$ is a core algebra for the operator L in the sense of Definition 2.39 in Chapter 2.

Remark 3.1. Regarding the discussion about the essential self-adjointness of L, we point out two possible issues that could happen in practice:

(i) Sometimes the manifold M will not be complete, which is the case if, for example, M is diffeomorphic to an open and connected domain $\mathcal{O} \neq \mathbb{R}^d$. Fortunately, essential self-adjointness is not an exclusive property of diffusion operators on complete manifolds.

(ii) The operator L might be essentially self-adjoint with respect to a different class of functions, not $\mathcal{C}_c^\infty(M)$, but, again, this makes no essential difference for its treatment.

Following Hille-Yosida theory [Yos80], since L is essentially self-adjoint with respect to $\mathcal{C}_c^\infty(M)$, then there exists a unique semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ on M with invariant measure μ such that L is its infinitesimal generator, and the respective heat equation is verified, see equations (2.4.3) and (2.4.4) in Chapter 2, respectively. Note that the invariance and symmetry properties that were stated in terms of L in equations (3.2.2) and (3.2.3) are equivalent to their semigroup counterparts (2.4.2) and (2.4.5) in Chapter 2, respectively.

With respect to the definition we gave in Chapter 2, the only property that could fail for $(\mathrm{P}_t)_{t\geqslant 0}$ is the conservation of mass or Markovianity, that is, point (ii) in Definition 2.37. However, the assumptions we will make later will ensure that the semigroup is indeed Markovian, see Remark 3.5 below, so unless stated otherwise, we assume that $(\mathrm{P}_t)_{t\geqslant 0}$ is Markovian.

Finally, we see that L is ergodic; i.e., if $f \in \mathcal{C}_c^\infty(M)$ and $\mathrm{L}f = 0$, then f has to be identically constant. Indeed, if $\mathrm{L}f = 0$, then by (3.2.3) we have that $|\nabla f|^2 = 0$ on M, which in turn yields that f is constant since M was assumed to be connected. The following proposition translates ergodicity into a property for the semigroup $(\mathrm{P}_t)_{t\geqslant 0}$; see, for example, [BGL14, Proposition 3.1.13].

Proposition 3.2. *Let (M, g, μ) be a connected and complete weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, assume that $\mu \in \mathcal{P}(M)$,* *and let* L *be its associated Langevin diffusion operator. Suppose that the associated semigroup* $(\mathrm{P}_t)_{t\geqslant 0}$ *is Markovian. If* L *is ergodic, then*

$$\forall f \in L^2(\mu), \quad \lim_{t\to+\infty} \mathrm{P}_t f = \int_M f \,\mathrm{d}\mu \text{ in } L^2(\mu).$$

To end this part, we justify that the choice of L given by (3.2.5) does not entail any loss of generality with respect to a general elliptic diffusion operator.

Remark 3.3. Let $\widetilde{\mathrm{L}}$ be a given diffusion operator on M defined in local coordinates by

$$\forall f \in \mathcal{C}^\infty(M), \quad \widetilde{\mathrm{L}} f \coloneqq a^{ij}\partial^2_{ij} f + b^i \partial_i f, \tag{3.2.5}$$

where its coefficients $x \mapsto a(x) \coloneqq (a^{ij}(x))_{i,j=1}^d$ and $x \mapsto b(x) \coloneqq (b^i(x))_{i=1}^d$ are smooth, and assume that $\widetilde{\mathrm{L}}$ is elliptic: that is, for each $x \in M$, the matrix $a(x)$ is symmetric and positive definite. Then the matrix $a^{-1} = (a_{ij}(x))_{i,j=1}^d$ induces a Riemannian metric on M that we denote by $\widetilde{g}$. If we define $Z \coloneqq b^i\partial_i \colon M \to TM$, then we can write $\widetilde{\mathrm{L}}$ as

$$\forall f \in \mathcal{C}^\infty(M), \quad \widetilde{\mathrm{L}} f = \Delta_{\widetilde{g}} f + Zf,$$

where $\Delta_{\widetilde{g}}$ denotes the Laplace-Beltrami operator with respect to the new metric $\widetilde{g}$. Moreover, let us suppose that $\widetilde{\mathrm{L}}$ is invariant with respect to a measure μ of the form $d\mu = e^{-W}\, \mathrm{dVol}_{\widetilde{g}}$, where $\mathrm{Vol}_{\widetilde{g}}$ denotes the Riemannian volume measure on $(M, \widetilde{g})$ and $W \colon M \to \mathbb{R}$ is a smooth function. Then $Zf = -\nabla_{\widetilde{g}} W \cdot \nabla_{\widetilde{g}} f$, so we get

$$\forall f \in \mathcal{C}^\infty(M), \quad \widetilde{\mathrm{L}} f = \Delta_{\widetilde{g}} f - \nabla_{\widetilde{g}} W \cdot \nabla_{\widetilde{g}} f. \tag{3.2.6}$$

That is, if we start with a linear diffusion operator in the form (3.2.5), we can always express it in the Langevin form (3.2.6) for a Riemannian structure that is natural for the operator $\widetilde{\mathrm{L}}$.

3.2.2 The probabilistic counterpart of the generator

The semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ allows the study of the Langevin operator L using probabilistic techniques. Let $(\mathrm{P}^*_t)_{t\geqslant 0}$ be the dual semigroup: for each $t \geqslant 0$, P^*_t acts on every nonnegative Borel measure ν on M by

$$\forall f \in \mathcal{C}_{\mathrm{b}}(M), \quad \int_M f \,\mathrm{d}(\mathrm{P}^*_t \nu) = \int_M \mathrm{P}_t f \,\mathrm{d}\nu.$$

Reciprocally to what we did in Remark 2.38 in Chapter 2, for the semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ and a given an initial law $\nu \in \mathcal{P}(M)$, one can construct a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ such that there is an adapted Markov process $(X_t)_{t\geqslant 0}$ defined on Ω with values on M such that $X_0 \sim \nu$ and whose law is characterized by the dual semigroup:

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{b}}(M), \quad \mathbb{E}[f(X_t)] = \int_M f \,\mathrm{d}(\mathrm{P}^*_t \nu). \tag{3.2.7}$$

In particular, if $\nu = \delta_x$ for $x \in M$, then

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{b}}(M), \quad \mathbb{E}_x[f(X_t)] = \mathrm{P}_t f(x).$$

The Markov process $(X_t)_{t\geqslant 0}$ constructed above is a diffusion process on M with random initial condition $\nu \in \mathcal{P}(M)$ satisfying the following stochastic differential equation:

$$\mathrm{d}X_t = \sqrt{2}\,\mathrm{d}B_t - \nabla W(X_t)\,\mathrm{d}t, \quad X_0 \sim \nu, \tag{3.2.8}$$

where $(B_t)_{t\geqslant 0}$ is the Brownian motion on (M,g); see, for example, [Éme89] for a succinct introduction to stochastic calculus on manifolds.

For each $t \geqslant 0$, define $\rho_t \coloneqq \mathrm{Law}(X_t)$. Then the flow of measures $(\rho_t)_{t\geqslant 0}$ satisfies the Fokker-Planck equation, which is adjoint of the heat equation with respect to the volume measure on M:

$$\partial_t \rho = \Delta\rho + \nabla\cdot(\rho\nabla W) \tag{3.2.9}$$

in the distributional sense, which is equivalent to

$$\forall t \geqslant 0, \forall \varphi \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \frac{\mathrm{d}}{\mathrm{d}t}\int_M \varphi\,\mathrm{d}\rho_t = \int_M \mathrm{L}\varphi\,\mathrm{d}\rho_t = \int_M (\Delta\varphi - \nabla W\cdot\nabla\varphi)\,\mathrm{d}\rho_t.$$

3.3 Revisiting the Bakry-Émery condition

In this section, we revisit the classic Bakry-Émery Γ-calculus, first recalling the basic definitions of Γ, its iteration Γ_2, and the curvature-dimension condition $\mathrm{CD}(\rho,\infty)$, which we have already introduced in Section 2.7 of Chapter 2. Then, we study further iterations of the carré du champ operator that give birth to the operators Γ_n for $n \in \mathbb{N}$. In particular, we will be interested in higher analogs of the curvature-dimension condition and their consequences. We will see that they yield some estimates involving the associated semigroup that will be useful in the next section to prove this chapter's main theorem.

We continue working under the same context as the last section, that is, (M,g) is a complete and connected Riemannian manifold with weight $\mathrm{d}\mu = e^{-W}\,\mathrm{dVol} \in \mathcal{P}(M)$ for some $W \in \mathcal{C}^\infty(M)$, with its associated Langevin diffusion operator $\mathrm{L} = \Delta - \nabla W\cdot\nabla$ and semigroup $(\mathrm{P}_t)_{t\geqslant 0}$.

3.3.1 The first-order curvature-dimension condition

Recall the carré du champ operator Γ associated with the Langevin generator L. From Example 2.51 in Chapter 2, we know that

$$\forall f \in \mathcal{C}^\infty_{\mathrm{c}}(M), \quad \Gamma(f) = |\nabla f|^2. \tag{3.3.1}$$

A direct consequence of (3.3.1) is that the Riemannian distance d_g can be written in terms of the operator Γ [BGL14, Appendix C]. More precisely,

$$\forall x,y \in M, \quad d_g(x,y) = \sup_{\substack{f\in\mathcal{C}^\infty_{\mathrm{c}}(M),\\ \Gamma(f)\leqslant 1}} (f(x)-f(y))\,.$$

This leads to the following characterization of Lipschitz functions, which will play an essential role in this note.

Proposition 3.4. *In the above context, let $f\colon M \to \mathbb{R}$ be a smooth function. Then f is K-Lipschitz for the distance d_g if and only if $\sqrt{\Gamma(f)} \leqslant K$ uniformly on M.*

Now recall the iterated carré du champ Γ_2. From Corollary 2.114 in Chapter 2, we know that

$$\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \Gamma_2(f) = |\nabla^2 f|_{\mathrm{HS}}^2 + \nabla^2 W(\nabla f, \nabla f) + \mathrm{Ric}(\nabla f, \nabla f). \tag{3.3.2}$$

We recall Definition 2.115 in Chapter 2, the Bakry-Émery curvature-dimension condition. Let $\rho \in \mathbb{R}$. We say that the diffusion operator L satisfies $\mathrm{CD}(\rho, \infty)$ if

$$\forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \Gamma_2(f) \geqslant \rho\, \Gamma(f),$$

which holds if and only if the Bakry-Émery-Ricci tensor Ric_W is lower bounded by ρ: $\mathrm{Ric}_W \coloneqq \mathrm{Ric} + \nabla^2 W \succeq \rho g$; see Remark 2.116 in Chapter 2.

Remark 3.5. The curvature-dimension condition $\mathrm{CD}(\rho, \infty)$ grants lots of properties for L and its associated semigroup $(\mathrm{P}_t)_{t \geqslant 0}$, even in the qualitative side. For example, in the above context, if there exists $\rho > 0$ such that $\mathrm{CD}(\rho, \infty)$ holds, then $(\mathrm{P}_t)_{t \geqslant 0}$ is Markovian [BGL14, Theorem 3.2.6]. Equivalently, if ζ denotes the explosion time of the process $(X_t)_{t \geqslant 0}$ generated by $(\mathrm{P}_t)_{t \geqslant 0}$, solution to the stochastic differential equation (3.2.8), then $\mathbb{P}(\zeta = +\infty) = 1$.

We discussed in Section 2.7 of Chapter 2 some of the consequences of the Bakry-Émery condition. The most notable are the commutations and local inequalities provided by Theorem 2.117, and the validity of both Poincaré and logarithmic Sobolev inequalities if the Bakry-Émery-Ricci tensor is lower bounded by a positive constant, namely Theorem 2.118. We recall the statement of both results for the sake of completeness.

Theorem 3.6. *Let (M, g, μ) be a connected and complete weighted Riemannian manifold, and let $(\mathrm{P}_t)_{t \geqslant 0}$ be its associated Markov semigroup. Let $\rho \in \mathbb{R}$. Then the following assertions are equivalent:*

(i) $\mathrm{CD}(\rho, \infty)$ *holds.*

(ii) The following commutation holds:

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \Gamma(\mathrm{P}_t f) \leqslant e^{-2\rho t}\, \mathrm{P}_t(\Gamma(f)).$$

(iii) The following strong commutation holds:

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \sqrt{\Gamma(\mathrm{P}_t f)} \leqslant e^{-\rho t}\, \mathrm{P}_t(\sqrt{\Gamma(f)}).$$

(iv) The local Poincaré inequality holds: if $\rho \neq 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \mathrm{P}_t(f^2) - (\mathrm{P}_t f)^2 \leqslant \frac{1 - e^{-2\rho t}}{\rho}\, \mathrm{P}_t(\Gamma(f));$$

if $\rho = 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_{\mathrm{c}}^\infty(M), \quad \mathrm{P}_t(f^2) - (\mathrm{P}_t f)^2 \leqslant 2t\, \mathrm{P}_t(\Gamma(f)).$$

(v) *The local logarithmic Sobolev inequality holds:* $\rho \neq 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_c^\infty(M), \quad \mathrm{P}_t(f^2 \log f^2) - \mathrm{P}_t(f^2) \log \mathrm{P}_t(f^2) \leqslant 2\frac{1-e^{-2\rho t}}{\rho} \mathrm{P}_t(\Gamma(f));$$

if $\rho = 0$,

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_c^\infty(M), \quad \mathrm{P}_t(f^2 \log f^2) - \mathrm{P}_t(f^2) \log \mathrm{P}_t(f^2) \leqslant 4t\, \mathrm{P}_t(\Gamma(f)).$$

Theorem 3.7 (Bakry-Émery)**.** *Let* (M, g, μ) *be a connected and complete weighted Riemannian manifold with* $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ *for some* $W \in \mathcal{C}^\infty(M)$, *and suppose additionally that* $\mu \in \mathcal{P}(M)$. *If there exists* $\rho > 0$ *such that the curvature-dimension condition* $\mathrm{CD}(\rho, \infty)$ *holds, then:*

(i) μ *satisfies a Poincaré inequality with constant* $C_{\mathrm{P}}(\mu) \leqslant 1/\rho$.

(ii) μ *satisfies a logarithmic Sobolev inequality with constant* $C_{\mathrm{LS}}(\mu) \leqslant 2/\rho$.

3.3.2 Higher order iterations

The operator Γ_2 was constructed by iterating Γ. Following the same idea, we construct Γ_n inductively for $n \geqslant 3$ integer [Led93, Led95, Bak94]. If we set $\Gamma_0(f, h) \coloneqq fh$ for $f, h \in \mathcal{C}_c^\infty(M)$, then we define, for $n \in \mathbb{N}$,

$$\forall f, h \in \mathcal{C}_c^\infty(M), \quad \Gamma_{n+1}(f, h) \coloneqq \frac{1}{2}\left(\mathrm{L}\Gamma_n(f, h) - \Gamma_n(f, \mathrm{L}h) - \Gamma_n(h, \mathrm{L}f)\right), \qquad (3.3.3)$$

and as usual, we set $\Gamma_n(f) \coloneqq \Gamma_n(f, f)$.

Remark 3.8. The higher iterations given by (3.3.3) are consistent with the classical Γ-calculus since $\Gamma_1 = \Gamma$. In what follows, we will always assume that the iterations start from Γ_0: that is, Γ_1 is the first iteration, Γ_2 the second, and so on. For the purposes of this note, we are concerned only up to the third iteration, that is, Γ_3. However, the results in this section will be stated and proved in more generality.

The goal of this section is to obtain analytical bounds for the semigroup $(\mathrm{P}_t)_{t \geqslant 0}$ assuming that $\Gamma_{n+1}(f) \geqslant \rho_n \Gamma_n(f)$, in the spirit of Theorem 3.6. We start with a technical lemma that will be at the heart of our computations. It corresponds to a classical result when $n = 1$ that is used to prove Theorem 3.6 (see, for example, [BGL14]). For the sake of completeness, we provide a full proof.

Lemma 3.9. *In the above context, let* $n \in \mathbb{N}$, $f \in \mathcal{C}_c^\infty(M)$, $t \geqslant 0$, *and* $x \in M$. *We define* $\Lambda_n \colon \mathbb{R}_+ \to \mathbb{R}_+$ *by*

$$\Lambda_n(s) \coloneqq \mathrm{P}_s(\Gamma_n(\mathrm{P}_{t-s} f))(x).$$

Then

$$\forall t \geqslant s, \quad \frac{\mathrm{d}}{\mathrm{d}s}\Lambda_n(s) = 2\, \mathrm{P}_s(\Gamma_{n+1}(\mathrm{P}_{t-s} f)) = 2\Lambda_{n+1}(s).$$

Proof. Let us define the auxiliary functions

$$\begin{aligned}
&\varphi_1 \colon [0, t] \to \mathcal{C}^\infty(M) \cap L^\infty(\mu), && s \mapsto \varphi_1(s) \coloneqq \mathrm{P}_{t-s} f;\\
&\varphi_2 \colon \mathcal{C}^\infty(M) \cap L^\infty(\mu) \to \mathcal{C}^\infty(M) \cap L^\infty(\mu), && u \mapsto \varphi_2(u) \coloneqq \Gamma_n(u);\\
&\varphi_3 \colon \mathbb{R}_+ \times \mathcal{C}^\infty(M) \cap L^\infty(\mu) \to \mathcal{C}^\infty(M) \cap L^\infty(\mu), && (s, u) \mapsto \varphi_3(s, u) \coloneqq \mathrm{P}_s u.
\end{aligned}$$

We can note that $\Lambda_n(s) = \varphi_3(s, \varphi_2(\varphi_1(s)))$. Using the chain rule, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\Lambda_n(s) = \partial_s\varphi_3(s, \Gamma_n(\mathrm{P}_{t-s}f)) + \nabla_u\varphi_3(s, \Gamma_n(\mathrm{P}_{t-s}f))\nabla\varphi_2(\mathrm{P}_{t-s}f)(\varphi_1'(s)), \tag{3.3.4}$$

On the other hand, we observe that

$$\begin{aligned}\partial_s\varphi_3(s,u) &= \lim_{\varepsilon\to 0}\frac{\varphi_3(s+\varepsilon,u)-\varphi_3(s,u)}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{\mathrm{P}_{s+\varepsilon}u - \mathrm{P}_s u}{\varepsilon} = \lim_{\varepsilon\to 0}\mathrm{P}_s\left(\frac{\mathrm{P}_\varepsilon u - \mathrm{P}_0 u}{\varepsilon}\right)\\ &= \mathrm{P}_s\left(\lim_{\varepsilon\to 0}\frac{\mathrm{P}_\varepsilon u - \mathrm{P}_0 u}{\varepsilon}\right)\\ &= \mathrm{P}_s\mathrm{L}u.\end{aligned}$$

For the second variable, we observe that φ_3 is linear in its second variable, so

$$\nabla_u\varphi_3(s,u)(h) = \varphi_3(s,h).$$

We know that $\Gamma_n(f) = \Gamma_n(f,f)$ with $\Gamma_n(\cdot,\cdot)$ being a symmetric bilinear form, so

$$\nabla\varphi_2(u)(h) = 2\Gamma_n(u,h).$$

For $\frac{\mathrm{d}}{\mathrm{d}s}\varphi_1$, we note that

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi_1(s) = \frac{\mathrm{d}}{\mathrm{d}s}\left(\varphi_3(t-s,f)\right) = -\partial_s\varphi_3(t-s,f) = -\,\mathrm{P}_{t-s}\mathrm{L}f.$$

Finally, if we put everything into (3.3.4), we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\Lambda_n(s) = \mathrm{P}_s(\mathrm{L}(\Gamma_n(\mathrm{P}_{t-s}f))) - 2\,\mathrm{P}_s(\Gamma_n(\mathrm{P}_{t-s}f, \mathrm{P}_{t-s}\,\mathrm{L}f)) = \mathrm{P}_s(\Gamma_{n+1}(\mathrm{P}_{t-s}f)).$$

□

The following proposition goes in the spirit of the commutation granted by the curvature-dimension condition, namely point (ii) in Theorem 3.6.

Proposition 3.10. *Let $n \in \mathbb{N}$, and suppose that there exists a constant $\rho_n \in \mathbb{R}$ such that for all $f \in \mathcal{C}_c^\infty(M)$, $\Gamma_{n+1}(f) \geqslant \rho_n\,\Gamma_n(f)$. Then*

$$\forall t \geqslant 0, \forall f \in \mathcal{C}_c^\infty(M), \quad \Gamma_n(\mathrm{P}_t f) \leqslant e^{-2\rho_n t}\,\mathrm{P}_t(\Gamma_n(f)).$$

Proof. If we apply Lemma 3.9 and use $\Gamma_{n+1} \geqslant \rho_n\,\Gamma_n$, we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\Lambda_n(s) = 2\,\mathrm{P}_s(\Gamma_{n+1}(\mathrm{P}_{t-s}f)) \geqslant 2\rho_n\,\mathrm{P}_s(\Gamma_n(\mathrm{P}_{t-s}f)) = \Lambda_n(s).$$

Then by Grönwall's inequality, we obtain

$$\Lambda_n(0) \leqslant e^{-2\rho_n t}\Lambda_n(t),$$

so

$$\Gamma_n(\mathrm{P}_t f) \leqslant e^{-2\rho_n t}\,\mathrm{P}_t(\Gamma_n(f)).$$

□

We present another useful bound.

Proposition 3.11. *Let $n \in \mathbb{N}^*$, and assume that for any $h \in \mathcal{C}_c^\infty(M)$, both $\Gamma_{n-1}(h)$ and $\Gamma_{n+1}(h)$ are non-negative. Then*

$$\forall t > 0, \forall f \in \mathcal{C}_c^\infty(M), \quad \Gamma_n(\mathrm{P}_t f) \leqslant \frac{1}{2t}\mathrm{P}_t(\Gamma_{n-1}(f)).$$

Proof. Using Lemma 3.9, we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\Lambda_n(s) = 2\,\mathrm{P}_s(\Gamma_{n+1}(\mathrm{P}_{t+s}f)) \geqslant 0,$$

so Λ_n is non-decreasing. Then, for $t \geqslant s \geqslant 0$,

$$\Gamma_n(\mathrm{P}_t f) = \mathrm{P}_0(\Gamma_n(\mathrm{P}_t f)) = \Lambda_n(0) \leqslant \Lambda_n(s) = \mathrm{P}_s(\Gamma_n(\mathrm{P}_{t-s}f)).$$

Thus integrating and applying again Lemma 3.9, we see that

$$\begin{aligned}
t\Gamma_n(\mathrm{P}_t f) &= \int_0^t \Gamma_n(\mathrm{P}_t f)\,ds \leqslant \int_0^t \mathrm{P}_s(\Gamma_n(\mathrm{P}_{t-s})f)\,ds = \frac{1}{2}\int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\Lambda_{n-1}(s)\,ds \\
&= \frac{1}{2}\left(\Lambda_{n-1}(t) - \Lambda_{n-1}(0)\right) \\
&= \frac{1}{2}\left(\mathrm{P}_t(\Gamma_{n-1}(f)) - \Gamma_{n-1}(\mathrm{P}_t f)\right),
\end{aligned}$$

from where we get

$$\Gamma_n(\mathrm{P}_t f) \leqslant \frac{1}{2t}\left(\mathrm{P}_t(\Gamma_{n-1}(f)) - \Gamma_{n-1}(\mathrm{P}_t f)\right) \leqslant \frac{1}{2t}\,\mathrm{P}_t(\Gamma_{n-1}(f)).$$

□

By mixing the two previous propositions, we deduce the following inequality.

Proposition 3.12. *Let $n \in \mathbb{N}^*$, and suppose that for any $h \in \mathcal{C}_c^\infty(M)$, both $\Gamma_{n-1}(h)$ and $\Gamma_{n+1}(h)$ are non-negative. Additionally, let us assume that there exists a constant $\rho_n \in \mathbb{R}$ such that for each $f \in \mathcal{C}_c^\infty(M)$, $\Gamma_{n+1}(f) \geqslant \rho_n\,\Gamma_n(f)$. Then*

$$\forall t > 0, \forall f \in \mathcal{C}_c^\infty(M), \quad \Gamma_n(\mathrm{P}_t f) \leqslant \frac{1}{t}e^{-\rho_n t}\,\mathrm{P}_t(\Gamma_{n-1}f).$$

Proof. Let $t > 0$. By Proposition 3.10:

$$\Gamma_n(\mathrm{P}_t f) = \Gamma_n(\mathrm{P}_{t/2}(\mathrm{P}_{t/2}f)) \leqslant e^{-\rho_n t}\,\mathrm{P}_{t/2}(\Gamma_n(\mathrm{P}_{t/2}f)). \tag{3.3.5}$$

Then, by applying Proposition 3.11, we know that

$$\Gamma_n(\mathrm{P}_{t/2}f) \leqslant \frac{1}{t}\,\mathrm{P}_{t/2}(\Gamma_{n-1}(f)). \tag{3.3.6}$$

Therefore, if we mix up inequalities (3.3.5) and (3.3.6), we finally obtain

$$\Gamma_n(\mathrm{P}_t f) \leqslant \frac{1}{t}e^{-\rho_n t}\,\mathrm{P}_t(\Gamma_{n-1}f).$$

□

Remark 3.13. Let us note that if $n = 1$ and if there exist some constants $\rho_1, \rho_2 > 0$ such that for all $f \in \mathcal{C}_c^\infty(M)$, $\Gamma_3(f) \geqslant \rho_2\Gamma_2(f)$ and $\Gamma_2(f) \geqslant \rho_1\Gamma_1(f)$, then the non-negativeness hypotheses in Proposition 3.12 are satisfied.

3.3.3 Examples

Before the end of the section, we provide some illustrative examples of diffusion operators satisfying the higher-order Bakry-Émery criterion $\Gamma_3 \geqslant \rho\,\Gamma_2$ for some $\rho \in \mathbb{R}$.

Example 3.14 (Laplace-Beltrami operator on $\mathbb{S}^d$)**.** For any $d \geqslant 2$, consider the d-dimensional sphere $\mathbb{S}^d$, which was previously introduced in Section 2.2.2 of Chapter 2. The canonical Riemannian metric g on $\mathbb{S}^d$ is such that the Ricci tensor is constant: $\mathrm{Ric} = (d-1)g$. We equip $\mathbb{S}^d$ with the unique normalized measure invariant by translations $\mu = \mathrm{Vol}$ (in Section 2.2.2 we denoted this measure by σ, but to be consistent with the notation of this chapter, we prefer to rename it). Let $\mathrm{L} = \Delta_{\mathbb{S}^d}$ be the Laplace-Beltrami operator on $\mathbb{S}^d$, which is essentially self-adjoint on $\mathcal{C}_c^\infty(\mathbb{S}^d) = \mathcal{C}^\infty(\mathbb{S}^d)$, and has μ as invariant and reversible measure. Since the Ricci tensor is constant, we have that

$$\forall f \in \mathcal{C}^\infty(\mathbb{S}^d), \quad \Gamma_2(f) = |\nabla^2 f|_{\mathrm{HS}}^2 + (d-1)|\nabla f|^2,$$

so the operator L satisfies the curvature-dimension condition $\mathrm{CD}(d-1,\infty)$. Now let us compute Γ_3. Let $f \in \mathcal{C}^\infty(\mathbb{S}^d)$.

$$\begin{aligned}\Gamma_3(f) &= \frac{1}{2}\Delta\Gamma_2(f) - \Gamma_2(f,\Delta f)\\ &= \left(\frac{1}{2}\Delta\left|\nabla^2 f\right|_{\mathrm{HS}}^2 + \frac{1}{2}(d-1)\Delta|\nabla f|^2\right)\\ &\qquad - \left(\langle\nabla^2 f, \nabla^2(\Delta f)\rangle_{\mathrm{HS}} + (d-1)\nabla f\cdot\nabla(\Delta f)\right).\end{aligned}$$

Using Bochner's formula (Theorem 2.113 in Chapter 2) and again the constant Ricci curvature, we obtain

$$\frac{1}{2}(d-1)\Delta|\nabla f|^2 = (d-1)\left|\nabla^2 f\right|_{\mathrm{HS}}^2 + (d-1)\nabla f\cdot\nabla(\Delta f) + (d-1)^2|\nabla f|^2,$$

so

$$\begin{aligned}\Gamma_3(f) &= \frac{1}{2}\Delta\left|\nabla^2 f\right|_{\mathrm{HS}}^2 - \langle\nabla^2 f, \nabla^2(\Delta f)\rangle_{\mathrm{HS}} + (d-1)\left|\nabla^2 f\right|_{\mathrm{HS}}^2 + (d-1)^2|\nabla f|^2\\ &= \frac{1}{2}\Delta\left|\nabla^2 f\right|_{\mathrm{HS}}^2 - \langle\nabla^2 f, \nabla^2(\Delta f)\rangle_{\mathrm{HS}} + (d-1)\Gamma_2(f),\end{aligned}$$

so if we use again Bochner's formula and the constant Ricci curvature for the first term, we get the bound

$$\Gamma_3(f) \geqslant (d-1)\Gamma_2(f).$$

Example 3.15 (Laguerre generator on $\mathbb{R}^d_{>0}$)**.** Fix $p \in \mathbb{R}^d_{>0}$, and define on $\mathbb{R}^d_{>0}$ the Laguerre operator by

$$\forall f \in \mathcal{C}^\infty(\mathbb{R}^d_{>0}), \quad \mathrm{L}_p f \coloneqq \sum_{i=1}^d x_i \frac{\partial^2 f}{\partial x_i^2} + \sum_{i=1}^d (p_i - x_i)\frac{\partial f}{\partial x_i},$$

and let μ_p be the multivariate gamma law of parameter p on $\mathbb{R}^d_{>0}$:

$$\mathrm{d}\mu_p = \bigotimes_{i=1}^d \frac{1}{\gamma(p_i)} x_i^{p_i-1} e^{-x_i}\,\mathrm{d}x_i,$$

where γ denotes the usual gamma function. Note that L_p has μ_p as invariant and reversible measure. Following Remark 3.3, note that the function $x \mapsto (\frac{1}{x_i}\delta_{ij})$ defines a Riemannian metric g on $\mathbb{R}^d_{>0}$. The associated diffusion process $(X_t)_{t\geqslant 0}$ is non-explosive as soon as $p_i \geqslant 1$ for $1 \leqslant i \leqslant d$ [GJY03]. On the other hand, it is known [Kra82] that if $p_i \geqslant \frac{3}{2}$ for $1 \leqslant i \leqslant d$, then L_p is essentially self-adjoint on $\mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}^d_{>0})$, so we will restrict our attention towards this case. The Γ_n operators have an explicit representation in terms of the derivatives of f [LNP15]:

$$\begin{aligned}
\Gamma(f) &= \sum_{i=1}^d x_i \left(\frac{\partial f}{\partial x_i}\right)^2 \\
\Gamma_2(f) &= \sum_{i,j=1}^d x_i x_j \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)^2 + \sum_{i=1}^d x_i \frac{\partial f}{\partial x_i}\frac{\partial^2 f}{\partial x_i^2} + \frac{1}{2}\sum_{i=1}^d (p_i + x_i)\left(\frac{\partial f}{\partial x_i}\right)^2 \\
\Gamma_3(f) &= \sum_{i,j,k=1}^d x_i x_j x_k \left(\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}\right)^2 + 3\sum_{i,j=1}^d x_i x_j \frac{\partial^2 f}{\partial x_i \partial x_j}\frac{\partial^3 f}{\partial x_i^2 \partial x_j} \\
&\quad + \frac{3}{2}\sum_{i,j=1}^d (p_i + x_i) x_j \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)^2 + \frac{3}{2}\sum_{i=1}^d x_i \left(\frac{\partial^2 f}{\partial x_i^2}\right)^2 \\
&\quad + \frac{3}{2}\sum_{i=1}^d x_i \frac{\partial f}{\partial x_i}\frac{\partial^2 f}{\partial x_i^2} + \frac{1}{4}\sum_{i=1}^d (3p_i + x_i)\left(\frac{\partial f}{\partial x_i}\right)^2 .
\end{aligned}$$

As $p_i \geqslant \frac{3}{2}$, we can easily observe that $\Gamma_2(f) \geqslant \frac{1}{2}\Gamma(f)$ and that $\Gamma_3(f) \geqslant \frac{1}{2}\Gamma_2(f)$.

Remark 3.16. In the special case when $d = 1$ and $p = 1$, then μ_1 corresponds to the exponential measure on $\mathbb{R}_{>0}$. L_1 is not essentially self-adjoint on $\mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}_{>0})$, however, it has this property with respect to the algebra of functions $\mathcal{C}^\infty_{\mathrm{c,\,Neu}}(\mathbb{R}_{>0})$ defined by

$$\mathcal{C}^\infty_{\mathrm{c,\,Neu}}(\mathbb{R}_{>0}) := \left\{ f \in \mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}_{>0}) : \lim_{x\to 0} x e^{-x} f'(x) = 0 \right\} \subset \mathcal{C}^\infty_{\mathrm{c}}(\mathbb{R}_{>0}),$$

which we can interpret as imposing a Neumann boundary condition at 0 for the operator L_p. On the other hand, it can be easily seen that in this case, both $\Gamma_2(f) \geqslant \frac{1}{2}\Gamma(f)$ and $\Gamma_3(f) \geqslant \frac{1}{2}\Gamma_2(f)$ are verified for every $f \in \mathcal{C}^\infty_{\mathrm{c,\,Neu}}(\mathbb{R}_{>0})$.

3.4 The diffusion transport map on smooth manifolds

In this section, we prove the main result of this note, Theorem 3.21, which employs the Kim-Milman heat flow transport map [KM12] on a weighted manifold setting to obtain a Lipschitz transport map between the prescripted weight measure and a log-Lipschitz perturbation of it.

The construction will be detailed, taking care of the details that appear when passing from the Euclidean setting towards the manifold one. After that, using the Γ-calculus machinery exhibited in Section 3.3, we will prove that the heat flow transport map is Lipschitz, giving an explicit bound on its Lipschitz constant.

Here the context will always be given by a complete and connected smooth Riemannian manifold (M, g) with weight $\mathrm{d}\mu = e^{-W}\,\mathrm{dVol} \in \mathcal{P}(M)$ for some $W \in \mathcal{C}_c^\infty(M)$. Recall the the natural Langevin diffusion operator L given by $\mathrm{L} = \Delta - \nabla W \cdot \nabla$ and its associated semigroup $(\mathrm{P}_t)_{t\geqslant 0}$. On the other hand, we assume that there exist positive constants $\rho_1, \rho_2 > 0$ such that for any $f \in \mathcal{C}_c^\infty(M)$, $\Gamma_2(f) \geqslant \rho_1\,\Gamma(f)$ and $\Gamma_3(f) \geqslant \rho_2\,\Gamma_2(f)$. We emphasize that under these assumptions, L is essentially self-adjoint on the core algebra $\mathcal{C}_c^\infty(M)$ and $(\mathrm{P}_t)_{t\geqslant 0}$ is Markovian.

3.4.1 Construction of the transport map

Let $V\colon M \to \mathbb{R}$ be a smooth function which is K-Lipschitz for the metric d_g, so in the light of Proposition 3.4, V is such that $\sqrt{\Gamma_1(V)} \leqslant K$ uniformly on M. Let us define $f := e^{-V}$. The following result justifies the integrability of f with respect to μ under $\mathrm{CD}(\rho, \infty)$ for $\rho > 0$, and is related to the Herbst argument Theorem 2.81 in Chapter 2; see [BGL14, Proposition 5.4.1]

Lemma 3.17. *In the above context, let $V\colon M \to \mathbb{R}$ be a smooth function such that $\sqrt{\Gamma_1(V)} \leqslant K$. If there exists $\rho > 0$ such that $\mathrm{CD}(\rho, \infty)$ holds, then both V and e^{sV} are μ-integrable for any $s \in \mathbb{R}$. In particular, $e^{-V} \in L^p(\mu)$ for each $p \geqslant 1$.*

Without loss of generality, we will assume $\int_M f\,\mathrm{d}\mu = \int_M e^{-V}\,\mathrm{d}\mu = 1$ so we can define the probability measure $\mathrm{d}\nu := f\,\mathrm{d}\mu$. For each $t \geqslant 0$, set $\mathrm{d}\rho_t = \mathrm{P}_t f\,\mathrm{d}\mu$, which is a probability measure since μ is invariant for $(\mathrm{P}_t)_{t\geqslant 0}$. In particular, we have that the flow $(\rho_t)_{t\geqslant 0}$ is such that $\rho_0 = \nu$ and $\rho_\infty := \lim_{t\to+\infty} \rho_t = \mu$ in distribution since L is ergodic and $(\mathrm{P}_t)_{t\geqslant 0}$ is Markovian; thus, $(\rho_t)_{t\geqslant 0}$ interpolates between ν and μ.

Now let us define, for each $t \geqslant 0$ and $x \in M$, $V_t(x) := -\log \mathrm{P}_t f(x)$, so that ∇V_t is a vector field for each $t \geqslant 0$. The following proposition shows that actually that $\nabla V_\bullet$ and $(\rho_t)_{t\geqslant 0}$ verify together the continuity equation, recall Section 1.3.3 of Chapter 1.

Proposition 3.18. *The flow $(\rho_t)_{t\geqslant 0}$ satisfies the continuity equation with velocity field $\nabla V_\bullet$ and initial condition ν:*

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t \nabla V_t) = 0 \\ \rho_0 = \nu \end{cases}$$

in the distributional sense.

Proof. Let $t > 0$ and $\varphi \in \mathcal{C}_c^\infty(M)$. We shall prove that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_M \varphi\,\mathrm{d}\rho_t = \int_M \nabla\varphi \cdot \nabla V_t\,\mathrm{d}\rho_t.$$

Indeed, as $(\rho_t)_{t\geqslant 0}$ is solution of the SDE (3.2.8), then it satisfies the Fokker-Planck equation (3.2.9) in the distributional sense. Thus, we can write

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_M \varphi\,\mathrm{d}\rho_t &= \int_M (\Delta\varphi - \nabla W\cdot\nabla\varphi)\,\mathrm{d}\rho_t = \int_M \Delta\varphi e^{-W}\,\mathrm{P}_t f\,\mathrm{dVol} - \int_M \nabla W\cdot\nabla\varphi\,\mathrm{d}\rho_t\\ &= -\int_M \nabla\varphi\cdot\nabla(e^{-W}\,\mathrm{P}_t f)\,\mathrm{dVol} - \int_M \nabla W\cdot\nabla\varphi\,\mathrm{d}\rho_t\\ &= -\int_M \nabla\varphi\cdot\nabla\,\mathrm{P}_t f\,\mathrm{d}\mu = -\int_M \nabla\varphi\cdot\frac{\nabla\,\mathrm{P}_t f}{\mathrm{P}_t f}\,\mathrm{d}\rho_t = \int_M \nabla\varphi\cdot\nabla V_t\,\mathrm{d}\rho_t.\end{aligned}$$

□

Let us consider the flow of diffeomorphisms $(S_t)_{t\geqslant 0}$ induced by the vector field ∇V_t:

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}S_t(x) = \nabla V_t(S_t(x)), \quad t>0\\ \quad S_0(x) = x.\end{cases} \tag{3.4.1}$$

Our goal will be to characterize the flow $(\rho_t)_{t\geqslant 0}$ in terms of $(S_t)_{t\geqslant 0}$ using Theorem 1.27 in Chapter 1. First of all, note that $\nabla V_\bullet$ is smooth since the semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ preserves smoothness because L is an elliptic diffusion operator. It is also clear that $t\mapsto\rho_t$ is continuous for the weak topology. Finally, we have to justify the integrability condition (1.3.6), which will hold under the curvature-dimension condition $\mathrm{CD}(\rho_1,\infty)$.

Lemma 3.19. *Let $V_\bullet$ be defined as above. If there exists $\rho_1>0$ such that $\mathrm{CD}(\rho_1,\infty)$ holds, then*

$$\int_0^{+\infty}\int_M |\nabla V_t(x)|\,\mathrm{d}\rho_t\,\mathrm{d}t < +\infty.$$

Proof. Under $\mathrm{CD}(\rho_1,\infty)$ for $\rho_1>0$, we may use the strong commutation between Γ and P_t, item (iii) in Theorem 3.6, so we have that for each $t\geqslant 0$,

$$\sqrt{\Gamma(\mathrm{P}_t f)} \leqslant e^{-\rho_1 t}\,\mathrm{P}_t(\sqrt{\Gamma(f)}).$$

Note that

$$\mathrm{P}_t(\sqrt{\Gamma(f)}) = \mathrm{P}_t(|\nabla f|) = \mathrm{P}_t(|\nabla V| f).$$

V is a K-Lipschitz potential, so $|\nabla V|\leqslant K$. Then

$$|\nabla\,\mathrm{P}_t f| = \sqrt{\Gamma(\mathrm{P}_t f)} \leqslant e^{-\rho_1 t}K\,\mathrm{P}_t f.$$

It follows that $|\nabla V_t(x)|\leqslant Ke^{-\rho_1 t}$; hence,

$$\int_0^{+\infty}\int_M |\nabla V_t(x)|\,d\rho_t\,dt \leqslant K\int_0^{+\infty} e^{-\rho_1 t}\,dt < +\infty.$$

□

As we have verified the hypotheses of Theorem 1.27 in Chapter 1, then for each $t \geqslant 0$, $S_{t\#}\nu = \rho_t$. Now, for each $t \geqslant 0$, define $T_t \coloneqq S_t^{-1}$, so that $T_{t\#}\rho_t = \nu$.

If we suppose that for each $t \geqslant 0$ there exists a constant $K_t > 0$ such that T_t is K_t-Lipschitz, and that $K \coloneqq \limsup_{t\to+\infty} K_t < +\infty$, then Lemma 1 in [MS23] states that, modulo subsequence, for each $x \in M$, the limit $\lim_{t\to+\infty} T_t(x)$ exists and, moreover, it defines a K-Lipschitz mapping $T\colon M \to M$ such that $T_{\#}\mu = \nu$, as a consequence of Arzelà-Ascoli's theorem. The following lemma exhibits a condition on V_t to ensure that the mapping T is Lipschitz, with a quantitative estimate on the Lipschitz constant.

Lemma 3.20. *In the above context, suppose that there exists $\lambda\colon \mathbb{R}_+ \to \mathbb{R}_+$ integrable such that*

$$\forall t \geqslant 0, \forall x \in M, \quad -\nabla^2 V_t(x) \preceq \lambda(t) g_x.$$

Then T is $\exp\left(\int_0^{+\infty} \lambda(t)\,\mathrm{d}t\right)$-Lipschitz.

We give just a formal proof (see, for example, [FMS24, Proposition 1] or [Nee22, Lemma 2.1] for a complete argument): if we look at the flow (3.4.1), then for any $x, y \in M$, we may define $\alpha(t) \coloneqq d_g^2(S_t(x), S_t(y))$. If we differentiate α with respect to t and recall that $\frac{\mathrm{d}}{\mathrm{d}t} S_t(x) = \nabla V_t(S_t(x))$ and $\frac{\mathrm{d}}{\mathrm{d}t} S_t(y) = \nabla V_t(S_t(y))$, then the uniform bound $-\nabla^2 V_t(x) \preceq \lambda(t) g_x$ yields the following estimate:

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha(t) \geqslant -2\lambda(t)\alpha(t).$$

Therefore, Grönwall's inequality implies that

$$d_g(S_t(x), S_t(y)) \geqslant d_g(x, y) \exp\left(-\int_0^t \lambda(s)\,\mathrm{d}s\right).$$

As $T_t = S_t^{-1}$, the last inequality yields

$$d_g(x, y) \exp\left(\int_0^t \lambda(s)\,\mathrm{d}s\right) \geqslant d_g(T_t(x), T_t(y));$$

thus, T_t is $\exp\left(\int_0^t \lambda(s)\,\mathrm{d}s\right)$-Lipschitz, so the limiting map T is $\exp\left(\int_0^{+\infty} \lambda(s)\,\mathrm{d}s\right)$-Lipschitz.

We are ready to state the main result of this chapter.

Theorem 3.21. *Let (M, g, μ) be a complete and connected weighted Riemannian manifold with $\mathrm{d}\mu = \exp(-W)\,\mathrm{dVol}$ for some $W \in \mathcal{C}^\infty(M)$, and assume that $\mu \in \mathcal{P}(M)$. Consider the diffusion operator $\mathrm{L} = \Delta - \nabla W \cdot \nabla$, let Γ be its associated carré du champ, and let Γ_2 and Γ_3 be its respective iterations in the Bakry-Émery sense. Assume that there exist constants $\rho_1, \rho_2 > 0$ such that*

(i) $\forall f \in \mathcal{C}_\mathrm{c}^\infty(M), \Gamma_2(f) \geqslant \rho_1 \Gamma(f)$; and

(ii) $\forall f \in \mathcal{C}_\mathrm{c}^\infty(M), \Gamma_3(f) \geqslant \rho_2 \Gamma_2(f)$.

Let $V \in \mathcal{C}^\infty(M)$, and assume that it is K-Lipschitz for some $K > 0$. Define $\mathrm{d}\nu = e^{-V}\,\mathrm{d}\mu$ and assume that $\nu \in \mathcal{P}(M)$. Then there exists a Lipschitz map $T\colon M \to M$ pushing forward μ towards ν which is $\exp\left(\sqrt{\frac{2\pi}{\rho_2}}Ke^{\frac{K^2}{2\rho_1}}\right)$-Lipschitz.

Before proceeding with the proof of the theorem, we will make a few remarks.

Remark 3.22.

(i) First of all, we observe that the estimate for the Lipschitz constant given by Theorem 3.21 is intrinsically independent of the dimension $\dim(M)$ of the manifold (M, g). This is relevant for applications: in the next section, we will see how Theorem 3.21 can be applied to transport functional inequalities from the measure μ towards log-Lipschitz perturbations of it, so if we have a dimension-free inequality for μ, then the transported inequality for ν is dimension-free as well; recall the initial motivation given in Section 3.1, and more precisely, the argument given in (3.1.1).

(ii) Theorem 5 in [FMS24] provides another estimate for the Lipschitz constant of T in the smooth setting as well. The assumptions in their result are similar to those of Theorem 3.21, but instead of supposing that $\Gamma_3 \geqslant \rho_2\,\Gamma_2$, they assume a uniform bound on R, the Riemann tensor of curvature associated with (M, g). The estimates are of the same type, in the sense that both are of the form $O\left(\exp(\exp(K^2))\right)$, with K the Lipschitz constant of the potential V.

Before starting with the proof of Theorem 3.21, we recall the following lemma, which again is a consequence of Herbst's argument; see [Mil23, p. 298].

Lemma 3.23. *Let (E, δ, m) be a metric probability measure space. If m satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(m) > 0$, then for $-\infty < q < p < +\infty$, and for any $g\colon E \to \mathbb{R}$ K-Lipschitz,*

$$\|e^g\|_{L^p(m)} \leqslant \exp\left(K^2\frac{(p-q)}{2C_{\mathrm{LS}}(m)}\right)\|e^g\|_{L^q(m)}.$$

Proof of Theorem 3.21. Given what we have discussed in this section, and more precisely, thanks to Lemma 3.20, we only have to prove that we can find an integrable upper bound λ for $\nabla^2 \log \mathrm{P}_t f$. We note that for any $t > 0$, $x \in M$, and $Y \in T_xM$ with $|Y| = 1$,

$$\begin{aligned}\nabla^2 \log \mathrm{P}_t f(x)(Y,Y) &= \frac{\nabla^2\,\mathrm{P}_t f(x)(Y,Y)}{\mathrm{P}_t f(x)} - |\nabla \log \mathrm{P}_t f(x)\cdot Y|^2 \\ &\leqslant \frac{\nabla^2\,\mathrm{P}_t f(x)(Y,Y)}{\mathrm{P}_t f(x)} \\ &\leqslant \frac{|\nabla^2\,\mathrm{P}_t f(x)|}{\mathrm{P}_t f(x)};\end{aligned}$$

that is,

$$\nabla^2 \log \mathrm{P}_t f(x)(Y,Y) \leqslant \frac{|\nabla^2\,\mathrm{P}_t f(x)|}{\mathrm{P}_t f(x)}. \tag{3.4.2}$$

On the one hand, thanks to equation (3.3.2), we note that $|\nabla^2 \mathrm{P}_t f| \leqslant \sqrt{\Gamma_2(\mathrm{P}_t f)}$. On the other hand, as both ρ_2 and ρ_1 are positive, and in the light of Remark 3.13, we may use Proposition 3.12; thus, we have that

$$\Gamma_2(\mathrm{P}_t f) \leqslant \frac{1}{t} e^{-\rho_2 t} \, \mathrm{P}_t(\Gamma_1(f));$$

thus

$$|\nabla^2 \mathrm{P}_t f|^2 \leqslant \frac{1}{t} e^{-\rho_2 t} \, \mathrm{P}_t(\Gamma_1(f)). \tag{3.4.3}$$

Now, let us note that

$$\begin{aligned} \mathrm{P}_t(\Gamma_1(f)) = \mathrm{P}_t(|\nabla f|^2) = \mathrm{P}_t(|f \nabla V|^2) \leqslant K^2 \, \mathrm{P}_t(f^2) \leqslant K^2 \exp\left(K^2 \frac{1 - e^{-2\rho_1 t}}{\rho_1}\right) (\mathrm{P}_t f)^2 \\ \leqslant K^2 \exp\left(K^2/\rho_1\right) (\mathrm{P}_t f)^2 \end{aligned}$$

where we employed Lemma 3.23 for $p = 2, q = 1$, and the measure $m = \mathrm{P}_t^* \delta_x$ since it satisfies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mathrm{P}_t^* \delta_x) \leqslant \frac{1 - e^{-2\rho_1 t}}{\rho_1}$, see item (iv) in Theorem 3.6. If we blend up this with both (3.4.2) and (3.4.3), we finally get

$$\nabla^2 \log \mathrm{P}_t f(x)(Y, Y) \leqslant K e^{\frac{K^2}{2\rho_1}} \frac{1}{\sqrt{t}} e^{-\frac{1}{2}\rho_2 t},$$

so the Lipschitz estimate follows from the fact that

$$\int_0^{+\infty} \frac{1}{\sqrt{t}} e^{-\frac{1}{2}\rho_2 t} \, \mathrm{d}t = \sqrt{\frac{2}{\rho_2}} \int_0^{+\infty} \frac{1}{\sqrt{s}} e^{-s} \, \mathrm{d}s = \sqrt{\frac{2\pi}{\rho_2}}.$$

□

3.5 Applications

In this section we provide different applications of Theorem 3.21.

3.5.1 Transfer of functional inequalities

As it was commented in Section 2, in the smooth setting, the existence of Lipschitz maps between measures allows the transfer of functional inequalities from the source measure towards the target, recall (3.1.1) and, more generally, Proposition 2.83 in Chapter 2 for the logarithmic Sobolev inequality.

Let us recall Theorem 3.7, which says that under $\mathrm{CD}(\rho, \infty)$ for $\rho > 0$, the measure μ verifies a logarithmic Sobolev inequality with constant $C_{\mathrm{LS}}(\mu) \leqslant 2/\rho$. If we combine this result with Proposition 2.83 in Chapter 2, then we obtain the following corollary to Theorem 3.21.

Corollary 3.24. *In the context of Theorem 3.21, for $K > 0$, let $\nu \in \mathcal{P}(M)$ be a K-log-Lipschitz perturbation of the measure μ. Then ν satisfies a logarithmic Sobolev inequality with constant*

$$C_{\mathrm{LS}}(\nu) \leqslant \frac{2 \exp\left(2\sqrt{\frac{2\pi}{\rho_2}} K e^{\frac{K^2}{2\rho_1}}\right)}{\rho}.$$

Remark 3.25. Transfer of functional inequalities for perturbations is a kind of result that has existed in the literature for a long time; see, for example, [AS94]. Nevertheless, our method allows us to transfer not just logarithmic Sobolev inequalities but isoperimetric, concentration, and Poincaré inequalities; recall Propositions 2.20, 2.24, and 2.65, respectively. To complement what we just said, we point out that under the Bakry-Émery condition, a functional Gaussian-type isoperimetric inequality holds, in the sense of Definition 2.21 in Chapter 2; see [BGL14, Corollary 8.5.4]

3.5.2 The sphere

Recall Example 3.14. We saw that the spherical Laplacian is such that $\Gamma_2 \geqslant (d-1)\,\Gamma$ and $\Gamma_3 \geqslant (d-1)\,\Gamma_2$, so we can apply Theorem 3.21. If we do so, we obtain that for each $V\colon \mathbb{S}^d \to \mathbb{R}$ 1-Lipschitz, then there exists a Lipschitz mapping T pushing forward Vol towards e^{-V} Vol with Lipschitz constant $\exp\left(\sqrt{\frac{2\pi}{d-1}}e^{\frac{1}{2(d-1)}}\right)$.

Now let $\sqrt{d}\,\mathbb{S}^d$ be the sphere rescaled by a factor of $\sqrt{d}$, and let μ_d be the uniform measure on it. Then its Ricci tensor is given by $\mathrm{Ric}_{\mathbb{S}^d(\sqrt{d})} = \frac{d-1}{d} g_{\mathbb{S}^d(\sqrt{d})}$, so it will satisfy $\Gamma_2 \geqslant \frac{d-1}{d}\,\Gamma$ and $\Gamma_3 \geqslant \frac{d-1}{d}\,\Gamma_2$; thus we can apply Theorem 3.21 to get a transport map with Lipschitz constant $\exp\left(\sqrt{2\pi\frac{d-1}{d}}e^{\frac{1}{2\frac{d-1}{d}}}\right)$ for 1-Lipschitz log-perturbations of μ_d; note that this quantity converges to $\exp\left(\sqrt{2\pi}e^{\frac{1}{2}}\right)$ as $d \to +\infty$.

Remark 3.26. The Lipschitz constant obtained for $\mathbb{S}^d$ is worse than the one provided in [FMS24, Theorem 2], where the authors stress the constant curvature of $\mathbb{S}^d$ to obtain an ad hoc bound. Despite this, the spherical example illustrates the applicability of Theorem 3.21.

3.5.3 Laguerre generator

We saw in Example 3.15 that for any $p \in \mathbb{R}^d_{>0}$ with $p_i \geqslant \frac{3}{2}$ for $1 \leqslant i \leqslant d$, the Laguerre operator L_p is essentially self-adjoint, and its associated semigroup conservative, so we may still apply Theorem 3.21 to its invariant measure μ_p. Thus, we have a Lipschitz map pushing forward the multivariate gamma distribution towards its log-Lipschitz perturbations.

In the particular case when $d = 1$ and $p = 1$, where the invariant measure μ_1 corresponds to the exponential distribution on $\mathbb{R}_{>0}$, from Remark 3.16 we know that L_1 is essentially self-adjoint with respect to the class of functions $\mathcal{C}^\infty_{\mathrm{c,\,Neu}}(\mathbb{R}_{>0})$. This makes no difference in the arguments, so we may also apply Theorem 3.21.

In the latter case, the operator $-\mathrm{L}_1$ has a discrete spectrum equal to $\mathbb{N}$; thus, it satisfies Poincaré's inequality with a better constant than the one given by $\mathrm{CD}(1/2, \infty)$: for each $f\colon \mathbb{R}_{>0} \to \mathbb{R}$ smooth,

$$\mathrm{Var}_{\mu_1}(f) \leqslant \int_{\mathbb{R}_+} x(f'(x))^2 \,\mathrm{d}\mu_1(x).$$

In particular, in the light of Corollary 3.24, it is possible to transfer this inequality for any log-Lipschitz perturbation of the exponential measure, where Lipschitzianity has to be regarded with respect to the metric generated by the Laguerre generator L_1. That is, if ν is a log-Lipschitz perturbation, then there exists $C_{\mathrm{P}}(\nu) > 0$ such that

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}_{>0}), \quad \mathrm{Var}_{\nu}(f) \leqslant C_{\mathrm{P}}(\nu) \int_{\mathbb{R}_+} y(f'(y))^2 \, \mathrm{d}\nu(y).$$

Remark 3.27. The exponential measure is a rich object for the theory of functional inequalities; see, for example, [BL97, BK08]. In particular, recall Theorem 2.59 in Chapter 2: the exponential measure verifies Poincaré's inequality with respect to the usual carré du champ operator $f \mapsto |f'|^2$:

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}_{>0}), \quad \mathrm{Var}_{\mu_1}(f) \leqslant 4 \int_{\mathbb{R}_{>0}} |f'|^2 \, \mathrm{d}\mu_1. \tag{3.5.1}$$

At first glance, we cannot use Corollary 3.24 to extend inequality (3.5.1) towards log-Lipschitz perturbations (for the metric generated by the Laguerre generator L_1, namely $x \mapsto \frac{1}{x}$) of μ_1 since the Poincaré inequality (3.5.1) is stated in terms of the carré du champ operator $f \mapsto |f'|^2$, which is not compatible with the metric $x \mapsto \frac{1}{x}$. However, we observe that the second conclusion in Proposition 3.28 below allows us to extend the same kind of functional inequality for its log-Lipschitz (with respect to the metric $x \mapsto \frac{1}{x}$) perturbations. That is, if ν is such a perturbation, then there exists $C_{\mathrm{P}}(\nu) > 0$ such that

$$\forall f \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}_{>0}), \quad \mathrm{Var}_{\nu}(f) \leqslant C \int_{\mathbb{R}_+} (f')^2 \, \mathrm{d}\nu. \tag{3.5.2}$$

It is known that in dimension one, the diffusion transport map T coincides with the Monge map; see Proposition 1.2 in Chapter 1 for the structure of Monge's map in dimension one. The following result provides a new estimate of the growth of Monge's map for $d = 1$ and $p \geqslant \frac{3}{2}$ or $p = 1$, pushing forward the measure μ_p towards a log-Lipschtitz perturbation of it. Let us recall that T is non-decreasing and positive.

Proposition 3.28. *Let μ_p be the gamma distribution on $\mathbb{R}_{>0}$, let $V \colon \mathbb{R}_{>0} \to \mathbb{R}$ be a Lipschitz potential (for the metric $x \mapsto \frac{1}{x}$), and let $T \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be the Monge map pushing forward μ_p towards $e^{-V}\mu_p$. Then there exists a constant $C > 0$ such that for any $x > 0$,*

$$0 < T(x) \leqslant Cx. \tag{3.5.3}$$

Moreover, T is Lipschitz for the Euclidean metric on $\mathbb{R}_{>0}$, that is, there exists $C' > 0$ such that for any $x > 0$,

$$0 \leqslant T'(x) \leqslant C'. \tag{3.5.4}$$

Proof. Let us denote by g the metric $x \mapsto \frac{1}{x}$. Using Theorem 3.21, we obtain a constant $M > 0$ such that $\sup_{x>0} |T'(x)|_{\mathrm{op}} \leqslant M$. Therefore, for each $x > 0$,

$$\begin{aligned} M \geqslant |T'(x)|_{\mathrm{op}} &= \sup_{|v|_{g_x}=1} |T'(x) \cdot v|_{g_{T(x)}} = \sup_{|v|_{g_x}=1} \sqrt{\frac{1}{T(x)} \, (T'(x)v)^2} \\ &= \sup_{|v|_{g_x}=1} \sqrt{\frac{x}{T(x)} \, (T'(x))^2 \, |v|_{g_x}^2} = \sqrt{\frac{x}{T(x)}} T'(x), \end{aligned}$$

which yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sqrt{T(x)}\right) \leqslant M\frac{1}{2\sqrt{x}}.$$

Then, if we integrate the last inequality and use the fact that $\lim_{y\to 0} T(y) = 0$, we obtain that for each $x > 0$, $\sqrt{T(x)} \leqslant M\sqrt{x}$, which yields (3.5.3) for $C = M^2$

Now, to bound T', we just use the bound on T and the first inequality, namely

$$M \geqslant \sqrt{\frac{x}{T(x)}}T'(x),$$

thus getting the desired conclusion (3.5.4). □

Remark 3.29. In [CF21, Theorem 1.3] provides a growth estimate for the derivative of the Monge map pushing forward the Gaussian measure γ_d onto a log-concave measure on $\mathbb{R}^d$ which moreover is log-Lipschitz, under certain technical bounds for the Hessian of its log-density: more precisely, $|\nabla T(x)|_{\mathrm{op}} = O(1 + |x|^2)$ as $|x| \to \infty$. Proposition 3.28 states that for the family of gamma distributions, the growth for the derivative of Monge's map pushing forward the measure towards a log-Lipschitz perturbation of it is $O(1)$ as $x \to \infty$ for the Euclidean metric. Similar bounds under different assumptions have been found in [CFS24, Fat24].

Chapter 4

The Poisson transport map

> I keep pushin' forwards, but he keeps pullin' me backwards
> (Nowhere to turn, no way, nowhere to turn, no)
> Now I'm standin' back from it, I finally see the pattern
> (I never learn, I never learn).
>
> DUA LIPA
> *New Rules*

This chapter is based on the article [LRS25], written in collaboration with Yair Shenfeld. The contributions appearing in this chapter belong to the theory of functional inequalities in the discrete setting. More concretely, we construct a transport map from Poisson point processes onto ultra-log-concave measures over the natural numbers, and show that this map is a contraction. Our approach overcomes the known obstacles to transferring functional inequalities using transport maps in discrete settings, and allows us to deduce a number of functional inequalities for ultra-log-concave measures. In particular, we provide the currently best known constant in modified logarithmic Sobolev inequalities for ultra-log-concave measures.

In Section 4.1, we motivate and introduce the context for our main results. In Section 4.2, we review some of the basics of ultra-log-concave measures, as well as the basics of the Poisson semigroup. Section 4.3 provides the construction of the Poisson transport map, as well as some of its properties. In Section 4.4 we prove our contraction theorem (Theorem 4.2). In addition, in Section 4.4.1, we compare and contrast the Brownian transport map and the Poisson transport map. Finally, in Section 4.5 we prove our functional inequalities (Theorem 4.5, Theorem 4.7, and Theorem 4.8).

4.1 Introduction

A classical way to establish functional inequalities for a given probability measure is to find a Lipschitz transport map from a source measure, for which the inequality is known, onto the target measure of interest as we saw in Chapter 3. If we wish to apply this method for discrete measures then we face a number of obstacles. Consider for instance the problem of constructing a Lipschitz transport map $X\colon \mathbb{N} \to \mathbb{N}$ between the Poisson measure (with intensity 1) π_1 on $\mathbb{N}$, and another probability measure μ on $\mathbb{N}$. The fact that the map X cannot split the mass of π_1 at any position in $\mathbb{N}$ severely restricts the type of measures μ that can arise as the pushforward of π_1 under X. This is not a particular problem of the natural numbers; it also emerges in other discrete structures: Example 1.5 in Chapter 1 shows that the existence of transport maps in the two-point space is limited as well. In addition, even if we can construct a Lipschitz transport map X between π_1 and μ, the lack of chain rule in the discrete setting hinders the argument.

4.1.1 The Poisson transport map

In this work we show that these obstacles can be overcome by transporting the Poisson point processes $\mathbb{P}$ onto probability measures μ on $\mathbb{N}$. In the notation above, $d = 1$ and $n = \infty$. In addition, we will show that in the setting considered in this work, the chain rule issue can be avoided. Let us describe informally our transport map. Fix a time $T > 0$ and $M > 0$, and consider a Poisson point process over $[0, T] \times [0, M]$:

- The numbers of points that fall in disjoint regions of $[0, T] \times [0, M]$ are independent.
- Given $B \subset [0, T] \times [0, M]$, the number of points that fall into B is distributed like a Poisson measure on $\mathbb{N}$ with intensity $\mathrm{Leb}(B)$.

Now let $\lambda\colon [0, T] \to [0, M]$ be a regular curve, and define the counting process $(X_t^\lambda)_{t\in[0,T]}$ by letting

$$X_t^\lambda := \text{number of points in } [0, t] \times [0, M] \text{ that fall below the curve } \lambda, \quad \text{(Figure 4.1)}. \tag{4.1.1}$$

Given a measure μ on $\mathbb{N}$, we can choose λ in a stochastic way so that $X_T^\lambda \sim \mu$. We call X_T^λ the Poisson transport map as it transports the Poisson point process $\mathbb{P}$ onto μ.

The Poisson transport map can be viewed as the discrete analog of the Brownian transport map of Mikulincer and Shenfeld [MS24], which transports the Wiener measure on path space onto probability measures over $\mathbb{R}^d$. The Brownian transport map is based on the Föllmer process [Föl85, Föl86, Leh13], and, analogously, the Poisson transport map is based on the process $(X_t^\lambda)_{t\in[0,T]}$, which is the discrete analogue of the Föllmer process. The process $(X_t^\lambda)_{t\in[0,T]}$ was constructed by Klartag and Lehec [KL19] (specializing and elaborating on earlier work of Budhiraja, Dupuis,

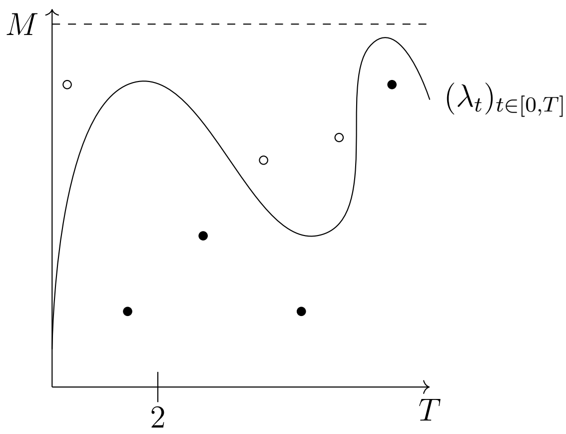

Figure 4.1: The points in $[0, T] \times [0, M]$ are generated according to a standard Poisson process (7 points in this case). At time $t \in [0, T]$ the value of the process X_t^λ is equal to the number of points under the curve (filled circles). In the figure, $X_2^\lambda = 1$ and $X_T^\lambda = 4$.

and Maroulas [BDM11]), who used it to prove functional inequalities. In Section 4.4.1, we discuss the similarities and differences between the Brownian transport map and the Poisson transport map.

4.1.2 Ultra-log-concave measures

Just as in the continuous case, we cannot expect to have a Lipschitz transport map (with good constants) from $\mathbb{P}$ onto any probability measure μ on $\mathbb{N}$ since the existence of such map will imply functional inequalities for μ. The classical result on the existence of Lipschitz transport maps in the continuous setting is Caffarelli's contraction theorem, see Theorem 1.26 in Chapter 1, who showed that if $n = d$, and $\mu = f\gamma_d$ with $f : \mathbb{R}^d \to \mathbb{R}_{\geqslant 0}$ log-concave, then there exists a 1-Lipschitz transport map between γ_d and μ. Closer to our setting, it was shown in [MS24] that the Brownian transport map is 1-Lipschitz when the target measure over $\mathbb{R}^d$ is of the form $\mu = f\gamma_d$, with f log-concave.

In the discrete setting, the analogue of a measure μ being "more log-concave than the Gaussian" is that the measure is ultra-log-concave. To define this notion we recall that a positive function $f : \mathbb{N} \to \mathbb{R}_{>0}$ is log-concave if

$$\forall k \in \mathbb{N}^*, \quad f^2(k) \geqslant f(k-1)f(k+1). \tag{4.1.2}$$

Definition 4.1. A probability measure μ on $\mathbb{N}$ is ultra-log-concave if $\mu = f\pi_\lambda$, where π_λ is the Poisson measure with intensity λ, and $f : \mathbb{N} \to \mathbb{R}_{>0}$ is a positive log-concave function[1].

Ultra-log-concave measures form an important class of discrete probability measures as it possesses desirable properties such as closure under convolution [SW14, Theorem 4.1(b)]. They are also ubiquitous and show up in fields outside of probability

[1]In Section 4.2 we recall equivalent definitions of ultra-log-concave measures.

such as combinatorics and convex geometry. We refer to the introduction of [AMM22] for more information.

Our first main result is that the Poisson transport map from the Poisson point process $\mathbb{P}$ onto ultra-log-concave measures μ is 1-Lipschitz. We will formulate this condition in terms of the Malliavin derivative $\mathrm{D}_{(t,z)}$ of X_T, which captures the effect of adding a point at (t, z) to the point process on the value of X_T (see Section 4.3.1).

Theorem 4.2. *Fix a real number $T > 0$, let $\mu = f\pi_T$ be an ultra-log-concave probability measure over $\mathbb{N}$, and let $M := f(1)/f(0)$. Let X_T be the Poisson transport map from $\mathbb{P}$ to μ. Then, $\mathbb{P}$-almost-surely,*

$$\forall (t, z) \in [0, T] \times [0, M], \quad \mathrm{D}_{(t,z)} X_T \in \{0, 1\}.$$

The fact that $\mathrm{D}_{(t,z)} X_T$ is integer-valued follows from the definition of the Malliavin derivative $\mathrm{D}_{(t,z)}$, and since X_T is integer-valued. However, a priori, saying that X_T is 1-Lipschitz could have implied $\mathrm{D}_{(t,z)} X_T \in \{-1, 0, 1\}$. Theorem 4.2 shows that $\mathrm{D}_{(t,z)} X_T \geqslant 0$, which will be important to tackle the chain rule issue when transporting functional inequalities from $\mathbb{P}$ to μ.

4.1.3 Functional inequalities for ultra-log-concave measures

The absence of the chain rule in the discrete setting complicates the study of functional inequalities for measures on $\mathbb{N}$. For example, Poisson measures π_λ over $\mathbb{N}$, the discrete analogues of Gaussians, do not satisfy logarithmic Sobolev inequalities. Rather, they satisfy modified logarithmic Sobolev inequalities as was first developed by Bobkov and Ledoux [BL98]. In the discrete setting there are various choices for modified logarithmic Sobolev inequalities, and in the context of the Poisson measure Wu's inequality is the strongest; see Section 2.5.3 of Chapter 2 for a more extensive discussion. For example, to recover the Gaussian logarithmic Sobolev inequality via the combination of the Poisson modified logarithmic Sobolev inequality and the central limit theorem one needs Wu's inequality [CL23, Remark on page 75]. To introduce Wu's inequality let D be the discrete derivative of a function $g \colon \mathbb{N} \to \mathbb{R}$:

$$\forall k \in \mathbb{N}, \quad \mathrm{D}g(k) := g(k+1) - g(k).$$

Theorem 4.3. *[Wu00, Theorem 1.1]. Let π_T be the Poisson measure over $\mathbb{N}$ with intensity T. Then, for any positive $g \in L^2(\mathbb{N}, \pi_T)$,*

$$\mathrm{Ent}_{\pi_T}(g) \leqslant T\, \mathbb{E}_{\pi_T}[\Psi(g, \mathrm{D}g)], \tag{4.1.3}$$

where $\Psi(u, v) := (u + v)\log(u + v) - u \log u - (\log u + 1)v$.

In the continuous setting, as a consequence of the existence of 1-Lipschitz transport maps, measures which are more log-concave than Gaussians satisfy logarithmic Sobolev inequalities. Thus, in the discrete setting we can expect ultra-log-concave measures to satisfy modified logarithmic Sobolev inequalities. Indeed, such inequalities for ultra-log-concave measures were obtained by Caputo, Dai Pra, and Posta [CDPP09, Theorem 3.1], but the stronger Wu-type modified logarithmic Sobolev inequality for ultra-log-concave measures was only obtained later by Johnson in [Joh17].

Theorem 4.4. *[Joh17, Theorem 1.3 and Lemma 5.1]. Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Then, for any positive $g \in L^2(\mathbb{N}, \mu)$,*

$$\mathrm{Ent}_\mu(g) \leqslant \frac{\mu(1)}{\mu(0)} \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)], \tag{4.1.4}$$

where $\Psi(u, v) \coloneqq (u + v)\log(u + v) - u \log u - (\log u + 1)v$.

Note that $\frac{\mu(1)}{\mu(0)} = T$ when $\mu = \pi_T$, so (4.1.3) and (4.1.4) agree in this case. Our second main result shows that we can in fact improve the constant in the strongest modified logarithmic Sobolev inequalities for ultra-log-concave measures.

Theorem 4.5. *Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Then, for any positive $g \in L^2(\mathbb{N}, \mu)$,*

$$\mathrm{Ent}_\mu(g) \leqslant |\log \mu(0)|\, \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)], \tag{4.1.5}$$

where $\Psi(u, v) \coloneqq (u + v)\log(u + v) - u \log u - (\log u + 1)v$.

It will follow from our work (Corollary 4.16) that $|\log \mu(0)| \leqslant \frac{\mu(1)}{\mu(0)}$, so that (4.1.5) improves on (4.1.4). (Note however that (4.1.4) holds, with constant $1/c$, for the larger class of c-log-concave measures [Joh17].) Again, when $\mu = \pi_T$, we have $|\log \mu(0)| = T$.

Remark 4.6 (The optimal constant)**.** Theorem 4.5 raises the question of what is the optimal constant in modified logarithmic Sobolev inequalities for ultra-log-concave measures. Daly and Johnson [DJ13, Corollary 2.4] showed that the Poincaré inequality for ultra-log-concave measures holds with a constant at least as good as $\mathbb{E}[\mu] \coloneqq \mathbb{E}_{Z\sim\mu}[Z]$. On the other hand, it will follow from our work (Corollary 4.16) that

$$\mathbb{E}[\mu] \leqslant |\log \mu(0)| \leqslant \frac{\mu(1)}{\mu(0)},$$

which begs the question of whether (4.1.5) holds with constant $\mathbb{E}[\mu]$. As evidence for an affirmative answer, it was shown by Aravinda, Marsiglietti, and Melbourne [AMM22, Theorem 1.1] that ultra-log-concave measures satisfy concentration inequalities with Poisson tail bounds. On the other hand, if the modified logarithmic Sobolev inequalities were to hold for ultra-log-concave measures with constant $\mathbb{E}[\mu]$, the result [AMM22, Theorem 1.1] could be deduced from the usual Herbst argument.

Theorem 4.5 is in fact a corollary of the following more general result, namely, the validity of Chafaï's Φ-Sobolev inequalities for ultra-log-concave measures; see Section 4.5.1 for the precise definitions.

Theorem 4.7. *Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Let $\mathcal{I} \subset \mathbb{R}$ be a closed interval, not necessarily bounded, and let $\Phi\colon \mathcal{I} \to \mathbb{R}$ be a smooth convex function. Suppose that the function*

$$\{(u, v) \in \mathbb{R}^2 : (u, u + v) \in \mathcal{I} \times \mathcal{I}\} \ni (u, v) \quad \mapsto \quad \Psi(u, v) \coloneqq \Phi(u + v) - \Phi(u) - \Phi'(u)v$$

is nonnegative and convex. Then, for any $g \in L^2(\mathbb{N}, \mu)$, such that μ-a.s. $g, g + \mathrm{D}g \in \mathcal{I}$,

$$\mathrm{Ent}^\Phi_\mu(g) \leqslant |\log \mu(0)|\, \mathbb{E}_\mu[\Psi(g, \mathrm{D}g)]. \tag{4.1.6}$$

We conclude with the following α-T_1 transport-entropy inequality for ultra-log-concave measures; see Section 4.5.2 for the precise definitions and recall Definition 2.100 from Chapter 2.

Theorem 4.8. *Let $\mu = f\pi_T$ be an ultra-log-concave probability measure on $\mathbb{N}$, and let $M := \frac{f(1)}{f(0)}$. Then, for any probability measure ν on $\mathbb{N}$ which is absolutely continuous with respect to μ, and has a finite first moment, we have*

$$\alpha_{TM}\left(\mathrm{W}_{1,|\cdot|}(\nu,\mu)\right) \leqslant \mathrm{H}(\nu|\mu), \tag{4.1.7}$$

where

$$\alpha_c(r) := c\left[\left(1+\frac{r}{c}\right)\log\left(1+\frac{r}{c}\right)-\frac{r}{c}\right].$$

The constant TM in (4.1.7) can in fact be improved; cf. Remark 4.28.

4.2 Ultra-log-concave measures

In this section we establish some of the properties of ultra-log-concave measures that will be used throughout the paper. We say that a positive function $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ is log-concave if

$$\forall k \in \mathbb{N}^*, \quad f^2(k) \geqslant f(k-1)f(k+1). \tag{4.2.1}$$

Equivalently, $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ is log-concave if the function

$$\mathbb{N}^* \ni k \mapsto \frac{f(k)}{f(k-1)} \text{ is non-increasing.} \tag{4.2.2}$$

The following definition captures the intuition of a probability measure being more log-concave than a Poisson measure.

Definition 4.9. A probability μ on $\mathbb{N}$ is ultra-log-concave if there exists $\lambda > 0$, and a positive log-concave function f, such that $\mu(k) = f(k)\pi_\lambda(k)$ for all $k \in \mathbb{N}$.

The intensity λ in Definition 4.9 does not in fact play any role. It is readily verified from the definition that μ is ultra-log-concave, with respect to any intensity $\lambda > 0$, if and only if

$$\forall k \in \mathbb{N}^*, \quad \mu^2(k) \geqslant \frac{k+1}{k}\mu(k+1)\mu(k-1). \tag{4.2.3}$$

In other words, once μ is more log-concave than π_λ for some λ, it is in fact more log-concave than π_λ for all λ.

The Poisson semigroup $(\mathrm{P}_t)_{t\geqslant 0}$ will play an important role in our work. Given a function $g\colon \mathbb{N} \to \mathbb{R}$ we define, for $t \geqslant 0$,

$$\mathrm{P}_0 g := g, \quad \text{and} \quad \forall k \in \mathbb{N}, \forall t > 0, \quad \mathrm{P}_t g(k) := \sum_{n=0}^{\infty} g(k+n)\pi_t(n).$$

The Poisson semigroup satisfies the identity

$$\forall k \in \mathbb{N}, \quad \partial_t(\mathrm{P}_t g)(k) = \mathrm{D}(\mathrm{P}_t g)(k), \tag{4.2.4}$$

where

$$\mathrm{D}h(k) := h(k+1) - h(k),$$

for any $h\colon \mathbb{N} \to \mathbb{R}$. Fix a time $T > 0$. For future reference, given nonnegative $f\colon \mathbb{N} \to \mathbb{R}$, we set

$$F(t,k) := \log \mathrm{P}_{T-t} f(k), \text{ which satisfies } \forall t \in [0,T], \forall k \in \mathbb{N}, \quad \partial_t F(t,k) = -e^{\mathrm{D}F(t,k)} + 1. \tag{4.2.5}$$

Our next result shows that the Poisson semigroup preserves log-concavity. While a number of proofs are available, our proof will mimic the proof of the fact that the heat semigroup preserves log-concavity. The latter is a consequence of the Prékopa-Leindler inequality, so we will use a discrete analogue of the Prékopa-Leindler inequality proven by Klartag and Lehec.

Proposition 4.10. *$f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be a log-concave function. Then, for any $t \geqslant 0$, $\mathrm{P}_t f$ is a log-concave function.*

Proof. Let $V := \log f$. Our goal is to show that

$$\forall k \in \mathbb{N}^*, \quad (\mathrm{P}_t e^V(k))^2 \geqslant \mathrm{P}_t e^V(k+1) \mathrm{P}_t e^V(k-1),$$

which, by definition, is equivalent to

$$\left(\sum_{n=0}^{\infty} e^{V(k+n)} \pi_t(n)\right)^2 \geqslant \left(\sum_{n=0}^{\infty} e^{V(k+1+n)} \pi_t(n)\right)\left(\sum_{n=0}^{\infty} e^{V(k-1+n)} \pi_t(n)\right). \tag{4.2.6}$$

The discrete Prékopa-Leindler inequality [KL19, Proposition 5.1] implies that for all functions $W, Y, Z\colon \mathbb{N} \to \mathbb{R}$,

$$\begin{aligned} &\forall \ell, m \in \mathbb{N}, \quad W(\ell) + Y(m) \leqslant Z\left(\left\lfloor \frac{\ell+m}{2} \right\rfloor\right) + Z\left(\left\lceil \frac{\ell+m}{2} \right\rceil\right) \\ &\Longrightarrow \\ &\left(\sum_{n=0}^{\infty} e^{Z(n)} \pi_t(n)\right)^2 \geqslant \left(\sum_{n=0}^{\infty} e^{W(n)} \pi_t(n)\right)\left(\sum_{n=0}^{\infty} e^{Y(n)} \pi_t(n)\right). \end{aligned} \tag{4.2.7}$$

To apply (4.2.7) we fix $k \in \mathbb{N}$ and define $Z(n) := V(k+n)$, $W(n) := V(k+1+n)$, and $Y(n) := V(k-1+n)$, so that to establish (4.2.6) it suffices to show

$$\forall \ell, m \in \mathbb{N}, \quad V(k+1+\ell) + V(k-1+m) \leqslant V\left(k + \left\lfloor \frac{\ell+m}{2} \right\rfloor\right) + V\left(k + \left\lceil \frac{\ell+m}{2} \right\rceil\right). \tag{4.2.8}$$

To verify (4.2.8) we note that the log-concavity of f implies that

$$\forall m \in \mathbb{N}^*, \quad 2V(m) \geqslant V(m+1) + V(m-1), \tag{4.2.9}$$

which is in fact equivalent to

$$\forall p, q \in \mathbb{N}, \quad V(p) + V(q) \leqslant V\left(\left\lfloor \frac{p+q}{2} \right\rfloor\right) + V\left(\left\lceil \frac{p+q}{2} \right\rceil\right). \tag{4.2.10}$$

Taking $p = k+1+\ell$ and $q = k-1+m$ in (4.2.10) yields

$$\begin{aligned} V(k+1+\ell) + V(k-1+m) &\leqslant V\left(\left\lfloor \frac{2k+\ell+m}{2} \right\rfloor\right) + V\left(\left\lceil \frac{2k+\ell+m}{2} \right\rceil\right) \\ &= V\left(k + \left\lfloor \frac{\ell+m}{2} \right\rfloor\right) + V\left(k + \left\lceil \frac{\ell+m}{2} \right\rceil\right), \end{aligned}$$

thus deducing (4.2.8). □

As a consequence of the preservation of log-concavity let us deduce a number of corollaries which will be helpful later on.

Corollary 4.11. *Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be a log-concave function. Fix $T > 0$ and $k \in \mathbb{N}$. The map*

$$[0,T] \ni t \mapsto \frac{\mathrm{P}_{T-t} f(k+1)}{\mathrm{P}_{T-t} f(k)}$$

is non-decreasing.

Proof. Define $\theta\colon [0,T] \to \mathbb{R}$ by $\theta(t) := \frac{\mathrm{P}_{T-t} f(k+1)}{\mathrm{P}_{T-t} f(k)}$, so we need to show that $\theta'(t) \geqslant 0$. Indeed, by (4.2.4),

$$\begin{aligned} \theta'(t) &= \frac{-\partial_t(\mathrm{P}_{T-t} f)(k+1)}{\mathrm{P}_{T-t} f(k)} - \frac{\mathrm{P}_{T-t} f(k+1)(-\partial_t(\mathrm{P}_{T-t} f)(k))}{(\mathrm{P}_{T-t} f(k))^2} \\ &= \frac{\mathrm{P}_{T-t} f(k+1)(\mathrm{D}(\mathrm{P}_{T-t} f)(k))}{(\mathrm{P}_{T-t} f(k))^2} - \frac{\mathrm{D}(\mathrm{P}_{T-t} f)(k+1)}{\mathrm{P}_{T-t} f(k)} \\ &= \frac{1}{(\mathrm{P}_{T-t} f(k))^2}\{\mathrm{P}_{T-t} f(k+1)[\mathrm{P}_{T-t} f(k+1) - \mathrm{P}_{T-t} f(k)] \\ &\qquad - \mathrm{P}_{T-t} f(k)[\mathrm{P}_{T-t} f(k+2) - \mathrm{P}_{T-t} f(k+1)]\} \\ &= \frac{1}{(\mathrm{P}_{T-t} f(k))^2}\left\{(\mathrm{P}_{T-t} f)^2(k+1) - \mathrm{P}_{T-t} f(k) \mathrm{P}_{T-t} f(k+2)\right\} \geqslant 0, \end{aligned}$$

where the last inequality holds by Proposition 4.10. □

Corollary 4.12. *Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be a log-concave function. Fix $T > 0$. Then, for any $t \in [0,T]$ and $k \in \mathbb{N}$,*

$$\frac{\mathrm{P}_{T-t} f(k+1)}{\mathrm{P}_{T-t} f(k)} \leqslant \frac{f(1)}{f(0)}. \tag{4.2.11}$$

Proof. Fix $k \in \mathbb{N}$. By Corollary 4.11 the function $[0,T] \ni t \mapsto \frac{\mathrm{P}_{T-t} f(k+1)}{\mathrm{P}_{T-t} f(k)}$ is non-decreasing, so

$$\forall\, t \in [0,T], \quad \frac{\mathrm{P}_{T-t} f(k+1)}{\mathrm{P}_{T-t} f(k)} \leqslant \frac{f(k+1)}{f(k)}.$$

On the other hand, by (4.2.2), the function $\mathbb{N} \ni k \mapsto \frac{f(k+1)}{f(k)}$ is non-increasing, so

$$\forall\, t \in [0,T], \quad \frac{\mathrm{P}_{T-t} f(k+1)}{\mathrm{P}_{T-t} f(k)} \leqslant \frac{f(k+1)}{f(k)} \leqslant \frac{f(1)}{f(0)}.$$

□

4.3 The Poisson transport map

In this section we construct the Poisson transport map. In Section 4.3.1 we recall the construction of the canonical space for the Poisson point process, as well as the basics of the Malliavin calculus on this space. We will use [Las16, BP16] as our references. In Section 4.3.2 we describe the process $(X_t^\lambda)_{t\in[0,T]}$ constructed by Klartag and Lehec [KL19], which we interpret as a transport map from the Poisson measure on the canonical space onto probability measures over $\mathbb{N}$.

4.3.1 The Poisson space

Fix a real number $T > 0$. Let $\mu = f\pi_T$ be an ultra-log-concave probability measure on $\mathbb{N}$, where $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ is a positive log-concave function. Set $M := \frac{f(1)}{f(0)}$, and let $\mathbb{X} := [0,T] \times [0,M]$. We let $\mathcal{X}$ be the σ-algebra generated by the Borel sets of $\mathbb{X}$ endowed with the product topology, and we let Leb be the Lebesgue measure on $\mathcal{X}$. We define the Poisson space $(\Omega, \mathcal{F}, \mathbb{P})$ over $(\mathbb{X}, \mathcal{X}, \mathrm{Leb})$ by letting the probability space be

$$\Omega := \left\{ \omega : \omega = \sum_i \delta_{(t_i,z_i)},\ (t_i, z_i) \in \mathbb{X} \text{ (at most countable)} \right\},$$

the σ-algebra be

$$\mathcal{F} := \sigma\left(\Omega \ni \omega \mapsto \omega(B) : B \in \mathcal{X}\right),$$

and defining the probability measure $\mathbb{P}$ by

$$\forall B \in \mathcal{X},\ \forall k \in \mathbb{N}, \quad \mathbb{P}(\{\omega(B) = k\}) = \pi_{\mathrm{Leb}(B)}(k),$$
$$\forall n \in \mathbb{N}^*, \quad \omega(B_1), \dots, \omega(B_n) \text{ are } \mathbb{P}\text{-independent if } B_1, \dots, B_n \in \mathcal{X} \text{ are disjoint.}$$

Given a measurable function $G\colon \Omega \to \mathbb{R}$, we define the Malliavin derivative D of G as the function $\mathrm{D}G\colon \Omega \times \mathbb{X} \to \mathbb{R}$ given by

$$\forall (t,z) \in \mathbb{X}, \forall \omega \in \Omega, \quad \mathrm{D}_{(t,z)}G(\omega) := G(\omega + \delta_{(t,z)}) - G(\omega). \tag{4.3.1}$$

Of particular importance to us will be binary Malliavin derivatives, for which one has the following chain rule.

Lemma 4.13. *Let $G\colon \Omega \to \mathbb{N}$ be a measurable function such that $\mathrm{D}_{(t,z)}G \in \{0,1\}$ for all $(t,z) \in \mathbb{X}$. Then, for any $g\colon \mathbb{N} \to \mathbb{R}$,*

$$\forall (t,z) \in \mathbb{X}, \quad \mathrm{D}_{(t,z)}(g \circ G) = \mathrm{D}g(G) \cdot \mathrm{D}_{(t,z)}G. \tag{4.3.2}$$

Proof. Fix $\omega \in \Omega$ and $(t,z) \in \mathbb{X}$. If $\mathrm{D}_{(t,z)}G(\omega) = G(\omega + \delta_{(t,z)}) - G(\omega) = 0$, then

$$\mathrm{D}_{(t,z)}(g \circ G(\omega)) = g(G(\omega + \delta_{(t,z)})) - g(G(\omega)) = 0,$$

since $G(\omega+\delta_{(t,z)}) = G(\omega)$, which establishes (4.3.2). Suppose then that $\mathrm{D}_{(t,z)}G(\omega) = 1$, so that $G(\omega + \delta_{(t,z)}) = G(\omega) + 1$. Then,

$$\begin{aligned}\mathrm{D}_{(t,z)}(g \circ G(\omega)) &= g(G(\omega + \delta_{(t,z)})) - g(G(\omega)) = g(G(\omega) + 1) - g(G(\omega)) = \mathrm{D}g(G(\omega)) \\ &= \mathrm{D}g(G(\omega))\mathrm{D}_{(t,z)}G(\omega),\end{aligned}$$

where in the last equality we used $\mathrm{D}_{(t,z)}G(\omega) = 1$. □

4.3.2 The Poisson transport map

Our construction of the Poisson transport map is based on the stochastic process used by Klartag and Lehec in [KL19] (whose origin can be found in Budhiraja, Dupuis, and Maroulas [BDM11]). Let the notation and assumptions of Section 4.3.1 hold. Given $t \in [0,T]$ let $\mathcal{X}_t$ be the σ-algebra generated by the Borel sets of $[0,t] \times [0,M]$ endowed with the product topology, and define the σ-algebra $\mathcal{F}_t$ on Ω by

$$\mathcal{F}_t := \sigma\left(\Omega \ni \omega \mapsto \omega(B) : B \in \mathcal{X}_t\right).$$

We say that a stochastic process $(\lambda_t)_{t\in[0,T]}$, where $\lambda_t \colon \Omega \to \mathbb{R}$, is predictable if the function $(t,\omega) \mapsto \lambda_t(\omega)$ is measurable with respect to $\sigma\left(\{(s,t] \times B : s \leqslant t \leqslant T,\ B \in \mathcal{F}_s\}\right)$. Given a predictable nonnegative stochastic process $\lambda := (\lambda_t)_{t\in[0,T]}$, such that $\lambda_t \leqslant M$ for all $t \in [0,T]$, we define the stochastic counting process $(X_t^\lambda)_{t\in[0,T]}$ by

$$X_t^\lambda(\omega) = \omega\left(\{(s,x) \in \mathbb{X} : s < t,\ x \leqslant \lambda_s(\omega)\}\right), \quad \text{(see Figure 4.1).} \tag{4.3.3}$$

Note that $(X_t^\lambda)_{t\in[0,T]}$ is a non-decreasing integer-valued left-continuous process such that X_t^λ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$, and hence $(X_t^\lambda)_{t\in[0,T]}$ is predictable. In addition, almost-surely, there are only finitely many jumps of $(X_t^\lambda)_{t\in[0,T]}$, each of which is of size 1. Thus, the process $(X_t^\lambda)_{t\in[0,T]}$ is a Poisson process with stochastic intensity λ. We will work with a specific stochastic intensity λ, namely, we will take the stochastic intensity λ^* defined by the equation

$$\lambda_t^*(\omega) = \frac{\mathrm{P}_{T-t} f(X_t^{\lambda^*(\omega)}(\omega)+1)}{\mathrm{P}_{T-t} f(X_t^{\lambda^*(\omega)}(\omega))} = e^{\mathrm{D}F\left(t, X_t^{\lambda^*(\omega)}(\omega)\right)}, \tag{4.3.4}$$

where we recall (4.2.5). The existence of a solution to (4.3.4) was given in [KL19, Lemma 4.3] via a fixed-point argument, which requires the function $k \mapsto \frac{\mathrm{P}_{T-t}f(k+1)}{\mathrm{P}_{T-t}f(k)}$ to be bounded for each $t \in [0,T]$. In [KL19], f itself is assumed to be bounded, which gives the necessary condition, but in our case the log-concavity of f suffices by Corollary 4.12. Note that $\mathbb{P}$-a.s. $t \mapsto \lambda_t^*$ is continuous except at finitely many points, and in addition, by Corollary 4.12,

$$\lambda_t^* \leqslant \frac{f(1)}{f(0)} = M. \tag{4.3.5}$$

To ease the notation, from here on, we will denote

$$X := (X_t)_{t\in[0,T]} := (X_t^{\lambda^*})_{t\in[0,T]} \quad \text{and} \quad \lambda := (\lambda_t)_{t\in[0,T]} := (\lambda_t^*)_{t\in[0,T]}. \tag{4.3.6}$$

The next lemma provides the time marginals of X.

Lemma 4.14. *Let X be the process defined by* (4.3.6)*. For every $t \in [0,T]$, the law of X_t is $(\mathrm{P}_{T-t}f)\pi_t$.*

Proof. Let $h \colon \mathbb{N} \to \mathbb{R}$ be any bounded function and fix $\omega \in \Omega$. By construction, $[0,T] \ni t \mapsto h(X_t(\omega))$ is $\mathbb{P}$-a.s. piecewise constant with jumps of size 1 at $t_1 < t_2 < \cdots$, where $\omega = \sum_i \delta_{(t_i,z_i)}$. Hence, for each $t \in [0,T]$,

$$h(X_t(\omega)) = h(0) + \int_0^t \mathrm{D}h(X_s(\omega))\,\mathrm{d}X_s := h(0) + \sum_{t_i \leqslant t} \mathrm{D}h(X_{t_i}(\omega)). \tag{4.3.7}$$

Taking expectation in (4.3.7), and applying [KL19, Lemma 4.1], we get

$$\mathbb{E}[h(X_t)] = h(0) + \mathbb{E}\left[\int_0^t \mathrm{D}h(X_s)\lambda_s \,\mathrm{d}s\right]. \tag{4.3.8}$$

Let μ_t be the law of X_t. Differentiating (4.3.8) in t, and using (4.3.4), we can apply summation by parts to get

$$\sum_{j=0}^{\infty} h(j)\partial_t\mu_t(j) = \sum_{j=0}^{\infty} \mathrm{D}h(j)e^{\mathrm{D}F(t,j)}\mu_t(j) \tag{4.3.9}$$

$$= -h(0)e^{\mathrm{D}F(t,0)}\mu_t(0) - \sum_{j=0}^{\infty} h(j+1)\,\mathrm{D}\left[e^{\mathrm{D}F(t,j)}\mu_t(j)\right]. \tag{4.3.10}$$

Equation (4.3.9) holds for all bounded h, so fix a non-zero $k \in \mathbb{N}$ and let $h(j) = 1_{j=k}$ to get the discrete Fokker-Planck equation,

$$\forall k \in \mathbb{N}, \quad \partial_t\mu_t(k) = -\,\mathrm{D}\left[e^{\mathrm{D}F(t,k-1)}\mu_t(k-1)\right]. \tag{4.3.11}$$

To get an equation at $k = 0$, take $h(j) = 1_{j=0}$ and use (4.3.9) to deduce

$$\partial_t\mu_t(0) = -e^{\mathrm{D}F(t,0)}\mu_t(0). \tag{4.3.12}$$

Using the convention $e^{\mathrm{D}F(t,-1)} = \mu_t(-1) = 0$, we can combine (4.3.11) and (4.3.12) to get

$$\forall k \in \mathbb{N}, \quad \partial_t\mu_t(k) = -\,\mathrm{D}\left[e^{\mathrm{D}F(t,k-1)}\mu_t(k-1)\right]. \tag{4.3.13}$$

One can check that (4.3.13) is uniquely solved by

$$\forall k \in \mathbb{N}, \quad \mu_t(k) = \mathrm{P}_{T-t}f(k)\pi_t(k). \tag{4.3.14}$$

$\square$

An immediate corollary of Lemma 4.14 is that X_T is distributed as μ. We call the map $X_T\colon \Omega \to \mathbb{N}$ the Poisson transport map as it transports $\mathbb{P}$ to μ.

4.3.3 Properties of the Poisson transport map and ultra-log-concave measures

Let us prove a number of properties of the processes X, λ, which we will use later.

Lemma 4.15. *Let X be the process defined by (4.3.6). Then λ is a $\mathbb{P}$-martingale, i.e., the process*

$$[0,T] \ni t \mapsto \frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)}$$

is a $\mathbb{P}$-martingale. Further, the common mean of λ is $\mathrm{P}_Tf(1)$.

Proof. Let $h\colon [0,T]\times\mathbb{N}\to\mathbb{R}$ be such that the function $[0,T]\ni t\mapsto h(t,k)$ is continuous for all $k\in\mathbb{N}$. Then the function $[0,T]\ni t\mapsto h(t,X_t)$ is piecewise absolutely-continuous function in t, so

$$h(t,X_t)=h(0,0)+\int_0^t \mathrm{D}h(s,X_s)\,\mathrm{d}X_s+\int_0^t \partial_s h(s,X_s)\,\mathrm{d}s. \tag{4.3.15}$$

Take $h(t,k):=\frac{\mathrm{P}_{T-t}f(k+1)}{\mathrm{P}_{T-t}f(k)}$, and note that it satisfies the continuity condition. Then, using (4.2.5), we get

$$\begin{aligned}\frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)}-\frac{\mathrm{P}_Tf(1)}{\mathrm{P}_Tf(0)}&=\int_0^t \mathrm{D}(e^{\mathrm{D}F(s,X_s)})\,\mathrm{d}X_s+\int_0^t\partial_s(e^{\mathrm{D}F(s,X_s)})\,\mathrm{d}s\\&=\int_0^t \mathrm{D}(1-\partial_sF(s,X_s))\,\mathrm{d}X_s+\int_0^t\partial_s(e^{\mathrm{D}F(s,X_s)})\,\mathrm{d}s\\&=-\int_0^t \mathrm{D}(\partial_sF(s,X_s))\,\mathrm{d}X_s+\int_0^t\partial_s(e^{\mathrm{D}F(s,X_s)})\,\mathrm{d}s.\end{aligned}$$

On the other hand, for every $k\in\mathbb{N}$,

$$\partial_s(e^{\mathrm{D}F(s,k)})=e^{\mathrm{D}F(s,k)}\partial_s\,\mathrm{D}F(s,k)=e^{\mathrm{D}F(s,k)}\,\mathrm{D}\partial_sF(s,k),$$

so by (4.3.4),

$$\partial_s(e^{\mathrm{D}F(s,X_s)})=\mathrm{D}(\partial_sF(s,X_s))\,\lambda_s.$$

We conclude that

$$\frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)}-\frac{\mathrm{P}_Tf(1)}{\mathrm{P}_Tf(0)}=-\int_0^t \mathrm{D}(\partial_sF(s,X_s))[\mathrm{d}X_s-\lambda_s\,\mathrm{d}s].$$

The process $\left(X_t-\int_0^t\lambda_s\,\mathrm{d}s\right)_{t\in[0,T]}$ is called the compensated process, and is a martingale. Hence, the process $\frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)}$ is a stochastic integral with respect to a martingale, and hence a martingale [KL19, §4].

To compute the common mean of λ note that since $X_T\sim\mu$ (cf. Lemma 4.14),

$$\mathbb{E}_{\mathbb{P}}[\lambda_T]=\mathbb{E}_{\mathbb{P}}\left[\frac{f(X_T+1)}{f(X_T)}\right]=\sum_{j=0}^{\infty}\frac{f(j+1)}{f(j)}\mu(j)=\sum_{j=0}^{\infty}f(j+1)\pi_T(j)=\mathrm{P}_Tf(1).$$

□

The fact that λ is a martingale allows us give a representation of the mean of μ in terms of the Poisson semigroup, as well as an upper bound.

Corollary 4.16.

$$\mathbb{E}[\mu]\overset{(1)}{=}T\,\mathrm{P}_Tf(1)\overset{(2)}{\leqslant}\int_0^T\frac{\mathrm{P}_{T-t}f(1)}{\mathrm{P}_{T-t}f(0)}\,\mathrm{d}t\overset{(3)}{=}T-\log f(0)\overset{(4)}{=}|\log\mu(0)|\overset{(5)}{\leqslant}\frac{\mu(1)}{\mu(0)}.$$

Proof. To prove identity (1), take $h(j) = j$ and $t = T$ in (4.3.8) to get

$$\mathbb{E}[X_T] = \mathbb{E}\left[\int_0^T \lambda_s \,\mathrm{d}s\right] = T\mathbb{E}[\lambda_T] = T\,\mathrm{P}_T f(1),$$

where we used Lemma 4.15. For the inequality (2), note that $\mathrm{P}_T f(0) = 1$, since $\mu = f\pi_T$ is a probability measure, and use Corollary 4.11 to get $\frac{\mathrm{P}_T f(1)}{\mathrm{P}_T f(0)} \leqslant \frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)}$ for all $t \in [0,T]$. For the identity (3), use (4.2.5) to compute

$$\int_0^T \frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)} \,\mathrm{d}t = \int_0^T e^{\mathrm{D}F(t,0)} \,\mathrm{d}t = \int_0^T [1 - \partial_t F(t,0)] \,\mathrm{d}t = T - [F(T,0) - F(0,0)].$$

The result follows since $F(T,0) = \log f(0)$, and $F(0,0) = \log \mathrm{P}_T f(0) = 0$ (because $\mathrm{P}_T f(0) = 1$ as $\mu = f\pi_T$ is a probability measure). The identity (4) follows from $\mu = f\pi_T$. Finally, the inequality (5) holds since, by Corollary 4.11, $\frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)} \leqslant \frac{f(1)}{f(0)}$ so, by (3)-(4),

$$|\log \mu(0)| = \int_0^T \frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)} \,\mathrm{d}t \leqslant T\frac{f(1)}{f(0)} = \frac{\mu(1)}{\mu(0)}.$$

□

4.4 Contraction of the Poisson transport map

The main result of this section is that the Poisson transport map is a contraction when the target measures are ultra-log-concave. This result will follow from the following more general theorem, showing that the Malliavin derivative of X is binary and nonnegative.

Theorem 4.17. *Fix a real number $T > 0$ and let $\mu = f\pi_T$ be an ultra-log-concave probability measure over $\mathbb{N}$. Let X be the process defined by (4.3.6). Then, $\mathbb{P}$-almost-surely, for every $s \in [0,T]$,*

$$\forall (t,z) \in \mathbb{X}, \quad \mathrm{D}_{(t,z)} X_s \in \{0,1\}. \tag{4.4.1}$$

An immediate corollary of Theorem 4.17 is that the Poisson transport map is a contraction, thus proving Theorem 4.2.

Corollary 4.18. *Fix a real number $T > 0$ and let $\mu = f\pi_T$ be an ultra-log-concave probability measure over $\mathbb{N}$. Let X_T be the Poisson transport map from $\mathbb{P}$ to μ. Then, $\mathbb{P}$-almost-surely,*

$$\forall (t,z) \in \mathbb{X}, \quad \mathrm{D}_{(t,z)} X_T \in \{0,1\}. \tag{4.4.2}$$

Let us turn to the proof of Theorem 4.17.

Proof of Theorem 4.17. Fix $(t,z) \in \mathbb{X}$ and $\omega \in \Omega$. Then $\mathbb{P}$-a.s., there exists $n \in \mathbb{N}$ such that $\omega = \sum_{i=1}^n \delta_{(t_i,z_i)}$ for $(t_i,z_i) \in \mathbb{X}$, with $i \in [n] \coloneqq \{1,\dots,n\}$, $0 < t_1 < \cdots < t_n < T$, and $t \neq t_i$ for all $i \in [n]$. Fix $s \in [0,T]$. We need to show that

$$\mathrm{D}_{(t,z)} X_s(\omega) = X_s(\omega + \delta_{(t,z)}) - X_s(\omega) \in \{0,1\}. \tag{4.4.3}$$

Let us first explain the intuition why (4.4.3) holds, and then turn to its rigorous verification. There are three cases to consider. The first two are easy, and the third one is the interesting one.

- **Case 1**. $s \leqslant t$: Then the contribution of the atom (t,z) is not captured by either $X_s(\omega+\delta_{(t,z)})$ or $X_s(\omega)$, so both processes behave identically, and hence $\mathrm{D}_{(t,z)}X_s(\omega)=0$.
- **Case 2**. $t<s$ and z lies above the curve $\lambda(\omega+\delta_{(t,z)})$: Then the atom (t,z) is not counted by the process $X(\omega+\delta_{(t,z)})$, so the processes $X(\omega+\delta_{(t,z)})$ and $X(\omega)$ are equal, and hence $\mathrm{D}_{(t,z)}X_s(\omega)=0$.
- **Case 3**. The interesting case is $t<s$ and z lies below the curve $\lambda(\omega+\delta_{(t,z)})$, so the processes $X(\omega+\delta_{(t,z)})$ and $X(\omega)$ can in fact differ. Our goal is show that when the two processes differ, $X(\omega+\delta_{(t,z)})$ is always greater than $X(\omega)$, but by no more than 1. The key to prove this is to use the log-concavity of f. Using the explicit expression of λ (4.3.4), we can reason about the relation between $\lambda(\omega+\delta_{(t,z)})$ and $\lambda(\omega)$, and hence about the relation between $X(\omega+\delta_{(t,z)})$ and $X(\omega)$.

Let us now turn to the actual proof of the theorem.

Case 1. $s \leqslant t$: We will show

$$\mathrm{D}_{(t,z)}X_s(\omega)=X_s(\omega+\delta_{(t,z)})-X_s(\omega)=0. \tag{4.4.4}$$

From the definition of $X(\omega+\delta_{(t,z)})$, we know that the atom (t,z) is not counted by $X(\omega+\delta_{(t,z)})$. So to verify (4.4.4) it suffices to show that each atom (t_i,z_i) is either counted by both $X(\omega+\delta_{(t,z)})$ and $X(\omega)$, or by neither. If $t_i\geqslant s$ for all $i\in[n]$, then (4.4.4) holds since both $X(\omega+\delta_{(t,z)})$ and $X(\omega)$ start at 0, and neither adds any atom by time s.

If there exists $i\in[n]$ such that $t_i<s$, let us denote $i_{\max}:=\max\{i\in[n]:t_i<s\}$. Since the processes are left-continuous, starting at 0, $X_{t_1}(\omega+\delta_{(t,z)})=X_{t_1}(\omega)=0$. Hence, by (4.3.4),

$$\lambda_{t_1}(\omega)=\frac{\mathrm{P}_{T-t_1}f(X_{t_1}(\omega)+1)}{\mathrm{P}_{T-t_1}f(X_{t_1}(\omega))}=\frac{\mathrm{P}_{T-t_1}f(X_{t_1}(\omega+\delta_{(t,z)})+1)}{\mathrm{P}_{T-t_1}f(X_{t_1}(\omega+\delta_{(t,z)}))}=\lambda_{t_1}(\omega+\delta_{(t,z)}).$$

It follows that

$$z_1\leqslant\lambda_{t_1}(\omega+\delta_{(t,z)})\quad\Longleftrightarrow\quad z_1\leqslant\lambda_{t_1}(\omega). \tag{4.4.5}$$

Hence, for each $r\in(t_1,t_2\wedge s]$, (if $n=1$ then for each $r\in(t_1,s]$), $X_r(\omega+\delta_{(t,z)})=X_r(\omega)$. If $i_{\max}=1$, we are done. Otherwise, if $i_{\max}\geqslant 2$, we can repeat the above argument inductively for $i\in\{2,\dots,i_{\max}\}$ to conclude that (4.4.4) holds.

Case 2. $t<s$ and $z>\lambda_t(\omega+\delta_{(t,z)})$: We will show

$$\mathrm{D}_{(t,z)}X_s(\omega)=X_s(\omega+\delta_{(t,z)})-X_s(\omega)=0. \tag{4.4.6}$$

The argument of Case 1 shows that $X_t(\omega+\delta_{(t,z)}) = X_t(\omega)$. Since $z > \lambda_t(\omega+\delta_{(t,z)})$, the atom (t,z) is not counted by $X(\omega+\delta_{(t,z)})$. It remains to analyze the atoms (t_i, z_i) for $i \in [n]$. If there exist no t_i such that $t < t_i < s$, then it is clear that $X_s(\omega+\delta_{(t,z)}) = X_s(\omega)$, so (4.4.6) holds. Suppose then that there exist t_i such that $t < t_i < s$, and let $i_{\min} \coloneqq \min\{i \in [n] : t < t_i < s\}$. Similar to Case 1, for $r \in (t, t_{i_{\min}}]$ we have $X_r(\omega+\delta_{(t,z)}) = X_r(\omega)$. In particular, $X_{t_{i_{\min}}}(\omega+\delta_{(t,z)}) = X_{t_{i_{\min}}}(\omega)$ so, as in Case 1,

$$z_{i_{\min}} \leqslant \lambda_{t_{i_{\min}}}(\omega+\delta_{(t,z)}) \quad \Longleftrightarrow \quad z_{i_{\min}} \leqslant \lambda_{t_{i_{\min}}}(\omega). \tag{4.4.7}$$

Hence, for each $r \in (t_{i_{\min}}, t_{i_{\min}+1} \wedge s]$, (if $i_{\min} = n$ then for $r \in (t_{i_{\min}}, s]$), $X_r(\omega+\delta_{(t,z)}) = X_r(\omega)$. We may repeat the argument above inductively for all $i \in [n]$ satisfying $t < t_i < s$ to conclude that $X_s(\omega+\delta_{(t,z)}) = X_s(\omega)$, so (4.4.6) holds.

Case 3. $t < s$ and $z \leqslant \lambda_t(\omega+\delta_{(t,z)})$: In contrast to Cases 1 and 2 we will show that

$$\mathrm{D}_{(t,z)}X_s(\omega) = X_s(\omega+\delta_{(t,z)}) - X_s(\omega) \in \{0,1\}. \tag{4.4.8}$$

Again, the argument of Case 1 shows that $X_t(\omega+\delta_{(t,z)}) = X_t(\omega)$. In contrast to Case 2, since $z \leqslant \lambda_t(\omega+\delta_{(t,z)})$, the atom (t,z) is counted by $X(\omega+\delta_{(t,z)})$. Let us analyze the possible values of $X_s(\omega+\delta_{(t,z)})$ and $X_s(\omega)$. If there exist no t_i such that $t < t_i < s$, then

$$X_s(\omega+\delta_{(t,z)}) = X_t(\omega+\delta_{(t,z)}) + 1 \quad \text{and} \quad X_s(\omega) = X_t(\omega),$$

so $\mathrm{D}_{(t,z)}X_s(\omega) = 1$, and hence (4.4.8) holds.

Suppose then that there exists t_i such that $t < t_i < s$, and as in Case 2, let $i_{\min} \coloneqq \min\{i \in [n] : t < t_i < s\}$. For $r \in (t, t_{i_{\min}}]$, we have

$$X_r(\omega+\delta_{(t,z)}) = X_t(\omega) + 1 \quad \text{and} \quad X_r(\omega) = X_t(\omega).$$

In particular,

$$X_{t_{i_{\min}}}(\omega+\delta_{(t,z)}) = X_{t_{i_{\min}}}(\omega) + 1,$$

so by Proposition 4.10, and (4.2.1),

$$\begin{aligned}\lambda_{t_{i_{\min}}}(\omega+\delta_{(t,z)}) &= \frac{\mathrm{P}_{T-t_{i_{\min}}} f(X_{t_{i_{\min}}}(\omega+\delta_{(t,z)}) + 1)}{\mathrm{P}_{T-t_{i_{\min}}} f(X_{t_{i_{\min}}}(\omega+\delta_{(t,z)}))} = \frac{\mathrm{P}_{T-t_{i_{\min}}} f(X_{t_{i_{\min}}}(\omega) + 2)}{\mathrm{P}_{T-t_{i_{\min}}} f(X_{t_{i_{\min}}}(\omega) + 1)}\\ &\leqslant \frac{\mathrm{P}_{T-t_{i_{\min}}} f(X_{t_{i_{\min}}}(\omega) + 1)}{\mathrm{P}_{T-t_{i_{\min}}} f(X_{t_{i_{\min}}}(\omega))} = \lambda_{t_{i_{\min}}}(\omega).\end{aligned} \tag{4.4.9}$$

Let us record then the one-direction analogue of (4.4.7),

$$z_{i_{\min}} \leqslant \lambda_{t_{i_{\min}}}(\omega+\delta_{(t,z)}) \quad \Longrightarrow \quad z_{i_{\min}} \leqslant \lambda_{t_{i_{\min}}}(\omega). \tag{4.4.10}$$

We now have a three sub-cases to consider.

Case 3.1. $z_{i_{\min}} \leqslant \lambda_{t_{i_{\min}}}(\omega+\delta_{(t,z)})$: Applying (4.4.10) we can deduce that for each $r \in (t_{i_{\min}}, t_{i_{\min}+1} \wedge s]$, (if $i_{\min} = n$ then for each $r \in (t_{i_{\min}}, s]$), $\mathrm{D}_{(t,z)}X_r(\omega) = 1$, since both $X(\omega)$ and $X(\omega+\delta_{(t,z)})$ count $(t_{i_{\min}}, z_{i_{\min}})$.

Case 3.2. $z_{i_{\min}} > \lambda_{t_{i_{\min}}}(\omega+\delta_{(t,z)})$ and $z_{i_{\min}} > \lambda_{t_{i_{\min}}}(\omega)$: For each $r \in (t_{i_{\min}}, t_{i_{\min}+1} \wedge s]$, (if $i_{\min} = n$ then for each $r \in (t_{i_{\min}}, s]$), we have $\mathrm{D}_{(t,z)}X_r(\omega) = 1$, since both $X(\omega)$ and $X(\omega+\delta_{(t,z)})$ did not count $(t_{i_{\min}}, z_{i_{\min}})$.

Case 3.3. $z_{i_{\min}} > \lambda_{t_{i_{\min}}}(\omega+\delta_{(t,z)})$ and $z_{i_{\min}} \leqslant \lambda_{t_{i_{\min}}}(\omega)$: For each $r \in (t_{i_{\min}}, t_{i_{\min}+1} \wedge s]$, (if $i_{\min} = n$ then for each $r \in (t_{i_{\min}}, s]$), we have $\mathrm{D}_{(t,z)}X_r(\omega) = 0$, since $X(\omega)$ counted $(t_{i_{\min}}, z_{i_{\min}})$, but $X(\omega+\delta_{(t,z)})$ did not.

If there exists no t_i, for $i > i_{\min}$, such that $t < t_i < s$, then Cases 3.1-3.3 verify (4.4.8). Suppose then that there exist $i > i_{\min}$ such that $t < t_i < s$. We will proceed inductively. From Cases 3.1-3.3 we have that $\mathrm{D}_{(t,z)}X_{t_{i_{\min}+1}}(\omega) \in \{0,1\}$. If $\mathrm{D}_{(t,z)}X_{t_{i_{\min}+1}}(\omega) = 1$, then arguing as in (4.4.9), we have

$$\lambda_{t_{i_{\min}+1}}(\omega+\delta_{(t,z)}) \leqslant \lambda_{t_{i_{\min}+1}}(\omega). \tag{4.4.11}$$

We now repeat Cases 3.1-3.3, replacing $i_{\min}$ by $i_{\min}+1$. If $\mathrm{D}_{(t,z)}X_{t_{i_{\min}+1}}(\omega) = 0$, then $\lambda_{t_{i_{\min}+1}}(\omega+\delta_{(t,z)}) = \lambda_{t_{i_{\min}+1}}(\omega)$, so (4.4.11) holds, and again we repeat Cases 3.1-3.3, replacing $i_{\min}$ by $i_{\min}+1$. Continuing in this manner we deduce (4.4.8). $\square$

The proof of Theorem 4.17 yields the following necessary condition for the Malliavin derivative being 1.

Corollary 4.19. *Fix $(t,z) \in \mathbb{X}$ and $\omega \in \Omega$. Then $\mathbb{P}$-a.s., given $s \in [0,T]$, if $\mathrm{D}_{(t,z)}X_s(\omega) = 1$, we must have $t < s$ and $z \leqslant \lambda_t(\omega+\delta_{(t,z)}) = \lambda_t(\omega)$.*

Proof. The conditions $t < s$ and $z \leqslant \lambda_t(\omega+\delta_{(t,z)})$ hold because the proof of Theorem 4.17 showed that $\mathrm{D}_{(t,z)}X_s(\omega) = 0$ in Cases 1-2. The condition $\lambda_t(\omega+\delta_{(t,z)}) = \lambda_t(\omega)$ holds since in Case 3 we have shown $X_t(\omega+\delta_{(t,z)}) = X_t(\omega)$, so the result follows from (4.3.4). $\square$

4.4.1 The Brownian transport map vs. the Poisson transport map

Let us elaborate on the similarities and dissimilarities between the Brownian transport map [MS24] and the Poisson transport map. For simplicity, let us take $T = 1$. We begin with a sketch of the Brownian transport map. Let μ be a probability measure on $\mathbb{R}^d$ of the form $\mu = f\gamma_d$. Denote by $(\mathrm{Q}_t)_{t\in[0,1]}$ the heat semigroup on $\mathbb{R}^d$, and consider the stochastic differential equation,

$$\mathrm{d}Y_t = \nabla \log \mathrm{Q}_{1-t} f(Y_t)\,\mathrm{d}t + \mathrm{d}B_t, \quad Y_0 = 0, \tag{4.4.12}$$

where $(B_t)_{t\in[0,1]}$ is a standard Brownian motion in $\mathbb{R}^d$. The process $Y := (Y_t)_{t\in[0,1]}$ is known as the Föllmer process [Föl85, Föl86, Leh13], and can be seen as Brownian motion conditioned to be distributed like μ at time 1. Alternatively, the process Y is a solution to an entropy minimization problem over the Wiener space. In [MS24], Y_1

is called the Brownian transport map, as it transports the Wiener measure (the law of $(B_t)_{t\in[0,1]}$) onto μ.

Now suppose that $\mu = f\gamma_d$ is such that $f\colon \mathbb{R}^d \to \mathbb{R}_{\geqslant 0}$ is log-concave. It was shown in [MS24, Theorem 1.1] that, in such setting, the Brownian transport map Y_1 is 1-Lipschitz, in the sense that the Malliavin derivative D of Y_1 is bounded in absolute value by 1. The proof of this result proceeds by differentiating (4.4.12) (with D) to get [MS24, Proposition 3.10],

$$\partial_s \mathrm{D}Y_s = \nabla^2 \log \mathrm{Q}_{1-s} f(Y_s) \mathrm{D}Y_s. \tag{4.4.13}$$

So, to show that Y_1 is 1-Lipschitz, i.e., $|\mathrm{D}Y_1| \leqslant 1$, it suffices to control $\nabla^2 \log \mathrm{Q}_{1-s} f(Y_s)$, and then use Grönwall's inequality. In particular, when f is log-concave, $\mathrm{Q}_{1-s}f$ is also log-concave (consequence of the Prékopa-Leindler inequality), i.e.,

$$\nabla^2 \log \mathrm{Q}_{1-s} f(Y_s) \leqslant 0, \tag{4.4.14}$$

so (4.4.13) and Grönwall's inequality yield $|\mathrm{D}Y_1| \leqslant 1$.

The analogue in the discrete setting of the Föllmer process (4.4.12) is the process X defined in (4.3.6). Indeed, it was shown in [KL19] that the process X is the solution to the corresponding entropy minimization problem on the Poisson space. Unlike the continuous setting, here we do not have an analogue of (4.4.13), but we do have an analogue of (4.4.14). The process $\lambda_s = \frac{\mathrm{P}_{1-s}f(X_s+1)}{\mathrm{P}_{1-s}f(X_s)}$ plays the role of $\nabla \log \mathrm{Q}_{1-s} f(Y_s)$, and the next result is the analogue of (4.4.14).

Lemma 4.20. *For every $s \in [0,1]$, $\mathbb{P}$-almost-surely,*

$$\forall (t,z) \in \mathbb{X}, \quad \mathrm{D}_{(t,z)}\lambda_s \leqslant 0.$$

Proof. Fix $\omega \in \Omega$. By definition,

$$\mathrm{D}_{(t,z)}\lambda_s(\omega) = \lambda_s(\omega + \delta_{(t,z)}) - \lambda_s(\omega) = \frac{\mathrm{P}_{1-s}f(X_s(\omega+\delta_{(t,z)})+1)}{\mathrm{P}_{1-s}f(X_s(\omega+\delta_{(t,z)}))} - \frac{\mathrm{P}_{1-s}f(X_s(\omega)+1)}{\mathrm{P}_{1-s}f(X_s(\omega))}.$$

By Theorem 4.17, $X_s(\omega+\delta_{(t,z)}) \in \{X_s(\omega), X_s(\omega)+1\}$. If $X_s(\omega+\delta_{(t,z)}) = X_s(\omega)$, then $\mathrm{D}_{(t,z)}\lambda_s(\omega) = 0$. If $X_s(\omega+\delta_{(t,z)}) = X_s(\omega)+1$, then

$$\mathrm{D}_{(t,z)}\lambda_s(\omega) = \frac{\mathrm{P}_{1-s}f(X_s(\omega)+2)}{\mathrm{P}_{1-s}f(X_s(\omega)+1)} - \frac{\mathrm{P}_{1-s}f(X_s(\omega)+1)}{\mathrm{P}_{1-s}f(X_s(\omega))} \leqslant 0,$$

where the inequality holds by Proposition 4.10 and (4.2.1). $\square$

Let us conclude with a few remarks on some other differences between the Brownian and Poisson transport maps.

Remark 4.21.

(i) In the Brownian transport map setting, the source measure is always the Wiener measure on the Wiener space $C([0,1])$ of continuous functions $[0,1] \to \mathbb{R}$, independent of the target measure μ. In contrast, the transported Poisson measure $\mathbb{P}$ depends on μ, because the space $\mathbb{X} = [0,T] \times [0,M]$ depends on μ via M. This difference is not material since the functional inequalities satisfied by $\mathbb{P}$ do not depend on M.

(ii) The fact that the Brownian transport map is 1-Lipschitz, when μ is more log-concave than the Gaussian, means that the functional inequalities which hold for the Gaussian also hold for μ with the same constants. In contrast, the constants in the functional inequalities for ultra-log-concave measures $\mu = f\pi_T$, obtained from the Poisson transport map, are different from those satisfied by π_T. This is not a deficiency of the Poisson transport map, but rather a manifestation of the discrete nature of the probability measures under consideration.

4.5 Functional inequalities

In this section we show how Corollary 4.18 can be used to deduce functional inequalities for ultra-log-concave measures. In particular, the results of this section verify Theorem 4.5, Theorem 4.7, and Theorem 4.8. The proofs of all of the results below proceed by using an appropriate functional inequality for $\mathbb{P}$ (cf. Section 4.3.1), and then, using Corollary 4.18, transporting these inequalities to ultra-log-concave measures.

4.5.1 Φ-Sobolev inequalities

In this section we prove both Theorem 4.7 and Theorem 4.5.

Definition 4.22. Let $\mathcal{I} \subset \mathbb{R}$ be a closed interval, not necessarily bounded, and let $\Phi\colon \mathcal{I} \to \mathbb{R}$ be a smooth convex function. Let $(E, \mathcal{E}, Q)$ be a probability Borel space. The Φ-entropy functional Ent_Q^Φ is defined on the set of Q-integrable functions $G\colon (E, \mathcal{E}) \to (\mathcal{I}, \mathcal{B}(\mathcal{I}))$, where $\mathcal{B}(\mathcal{I})$ stands for the Borel σ-algebra of $\mathcal{I}$, by

$$\mathrm{Ent}_Q^\Phi(G) := \int_E \Phi(G)\,\mathrm{d}Q - \Phi\left(\int_E G\,\mathrm{d}Q\right). \tag{4.5.1}$$

As shown by Chafaï, the Poisson measure $\mathbb{P}$ satisfies Φ-Sobolev inequalities:

Theorem 4.23. *[Cha04, Eq. (61)]. Let $\mathcal{I} \subset \mathbb{R}$ be a closed interval, not necessarily bounded, and let $\Phi\colon \mathcal{I} \to \mathbb{R}$ be a smooth convex function. Suppose that the function*

$$\{(u,v) \in \mathbb{R}^2 : (u, u+v) \in \mathcal{I} \times \mathcal{I}\} \ni (u,v) \quad \mapsto \quad \Psi(u,v) := \Phi(u+v) - \Phi(u) - \Phi'(u)v$$

is nonnegative and convex. Then, for any $G \in L^2(\Omega, \mathbb{P})$, such that $\mathbb{P}$-a.s. $G, G + \mathrm{D}G \in \mathcal{I}$,

$$\mathrm{Ent}_{\mathbb{P}}^\Phi(G) \leqslant \mathbb{E}_{\mathbb{P}}\left[\int_{\mathbb{X}} \Psi(G, \mathrm{D}_{(t,z)}G)\,\mathrm{d}t\,\mathrm{d}z\right]. \tag{4.5.2}$$

Let us now transport the inequality (4.5.2) to ultra-log-concave measures, using the Poisson transport map, thus proving Theorem 4.7.

Theorem 4.24. *Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Let $\mathcal{I} \subset \mathbb{R}$ be a closed interval, not necessarily bounded, and let $\Phi\colon \mathcal{I} \to \mathbb{R}$ be a smooth convex function. Suppose that the function*

$$\{(u,v) \in \mathbb{R}^2 : (u, u+v) \in \mathcal{I} \times \mathcal{I}\} \ni (u,v) \quad \mapsto \quad \Psi(u,v) := \Phi(u+v) - \Phi(u) - \Phi'(u)v$$

is nonnegative and convex. Then, for any $g \in L^2(\mathbb{N},\mu)$, such that μ-a.s. $g, g+\mathrm{D}g \in \mathcal{I}$,

$$\mathrm{Ent}_\mu^\Phi(g) \leqslant |\log \mu(0)|\,\mathbb{E}_\mu[\Psi(g,\mathrm{D}g)]. \tag{4.5.3}$$

Proof. Define $G \in L^2(\Omega,\mathbb{P})$ by $G(\omega) \coloneqq g(X_T(\omega))$, and apply (4.5.2) to get

$$\begin{aligned}\mathrm{Ent}_\mu^\Phi(g) = \mathrm{Ent}_\mathbb{P}^\Phi(G) &\leqslant \mathbb{E}_\mathbb{P}\left[\int_\mathbb{X} \Psi(G, \mathrm{D}_{(t,z)}G)\,\mathrm{d}t\,\mathrm{d}z\right]\\ &= \mathbb{E}_\mathbb{P}\left[\int_\mathbb{X} \Psi(g\circ X_T, (\mathrm{D}g\circ X_T)\cdot \mathrm{D}_{(t,z)}X_T)\,\mathrm{d}t\,\mathrm{d}z\right],\end{aligned} \tag{4.5.4}$$

where the last equality holds by Corollary 4.18 and Lemma 4.13. Since $\mathrm{D}_{(t,z)}X_T \in \{0,1\}$ by Corollary 4.18, we have that $\mathbb{P}$-a.s.,

$$\Psi(g\circ X_T, (\mathrm{D}g\circ X_T)\cdot \mathrm{D}_{(t,z)}X_T) = \Psi(g\circ X_T, (\mathrm{D}g\circ X_T))1_{\{\mathrm{D}_{(t,z)}X_T=1\}}. \tag{4.5.5}$$

On the other hand, by Corollary 4.19, we have, $\mathbb{P}$-a.s., $1_{\{\mathrm{D}_{(t,z)}X_T=1\}} \leqslant 1_{\{z\leqslant\lambda_t\}}$. Since Ψ is nonnegative, we conclude from (4.5.5) that

$$\Psi(g\circ X_T, (\mathrm{D}g\circ X_T)\cdot \mathrm{D}_{(t,z)}X_T) \leqslant \Psi(g\circ X_T, (\mathrm{D}g\circ X_T))1_{\{z\leqslant\lambda_t\}}. \tag{4.5.6}$$

It follows from (4.5.4) and (4.5.6) that

$$\begin{aligned}\mathrm{Ent}_\mu^\Phi(g) &\leqslant \mathbb{E}_\mathbb{P}\left[\int_\mathbb{X} \Psi(g\circ X_T, \mathrm{D}g\circ X_T)1_{\{z\leqslant\lambda_t\}}\,\mathrm{d}t\,\mathrm{d}z\right]\\ &= \mathbb{E}_\mathbb{P}\left[\Psi(g\circ X_T, \mathrm{D}g\circ X_T)\int_\mathbb{X} 1_{\{z\leqslant\lambda_t\}}\,\mathrm{d}t\,\mathrm{d}z\right]\\ &= \mathbb{E}_\mathbb{P}\left[\Psi(g\circ X_T, \mathrm{D}g\circ X_T)\int_0^T \lambda_t\,\mathrm{d}t\right].\end{aligned}$$

By (4.3.4), $\int_0^T \lambda_t\,\mathrm{d}t = \int_0^T \frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)}\,\mathrm{d}t$. On the other hand, $\frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)} \leqslant \frac{\mathrm{P}_{T-t}f(1)}{\mathrm{P}_{T-t}f(0)}$ by Proposition 4.10 and (4.2.2). The proof is complete by Corollary 4.16(3). □

Taking $\Phi(r) = r\log r$ we deduce a modified logarithmic Sobolev inequality, thus proving Theorem 4.5.

Corollary 4.25. *Let μ be an ultra-log-concave probability measure over $\mathbb{N}$. Then, for any positive $g \in L^2(\mathbb{N},\mu)$,*

$$\mathrm{Ent}_\mu(g) \leqslant |\log \mu(0)|\,\mathbb{E}_\mu[\Psi(g,\mathrm{D}g)],$$

where $\Psi(u,v) \coloneqq (u+v)\log(u+v) - u\log u - (\log u + 1)v$.

4.5.2 Transport-entropy inequalities

In this section we prove Theorem 4.8. We fix an ultra-log-concave probability measure $\mu = f\pi_T$ on $\mathbb{N}$, and recall the definition of the associated Poisson space from Section 4.3.1. The starting point is a transport-entropy inequality for the Poisson measure $\mathbb{P}$ by Ma, Shen, Wang, and Wu (a special case of their more general result) that generalizes Theorem 2.104 in Chapter 2, the Poissonian α-T_1 inequality. To state it, we require the following definitions: Let d be the total variation distance on Ω given by $d(\omega,\omega') := |\omega-\omega'|(\mathbb{X})$ [MSWW11, Remark 2.4]. Given two probability measures Q, P on $(\Omega,\mathcal{F})$, with finite first moments, let the Wasserstein 1-distance between them be given by

$$\mathrm{W}_{1,d}(Q,P) := \inf_{\Pi}\int_{\Omega\times\Omega} d(\omega,\omega')\,\mathrm{d}\Pi(\omega,\omega'),$$

where the infimum is taken over all couplings Π of (Q,P). If Q is absolutely continuous with respect to P, let the relative entropy between them be

$$\mathrm{H}(Q|P) := \int_{\Omega}\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right)\mathrm{d}Q.$$

Finally, given $c>0$, let

$$\alpha_c(r) := c\left[\left(1+\frac{r}{c}\right)\log\left(1+\frac{r}{c}\right)-\frac{r}{c}\right].$$

Theorem 4.26. *[MSWW11, Eq. (2.4)]. For any probability measure Q on $(\Omega,\mathcal{F})$ which is absolutely continuous with respect to $\mathbb{P}$, and has a finite first moment, we have*

$$\alpha_{TM}\left(\mathrm{W}_{1,d}(Q,\mathbb{P})\right) \leqslant \mathrm{H}(Q|\mathbb{P}), \tag{4.5.7}$$

where $M=\frac{f(1)}{f(0)}$.

Let us now transport the inequality (4.5.7), thus proving Theorem 4.8. To do so, we define the Wasserstein 1-distance between two probability measures ν,ρ on $\mathbb{N}$, with finite first moments, by

$$\mathrm{W}_{1,|\cdot|}(\nu,\rho) := \inf_{\Pi}\int_{\mathbb{N}\times\mathbb{N}} |x-y|\,\mathrm{d}\Pi(x,y),$$

where the infimum is taken over all couplings Π of (ν,ρ).

Theorem 4.27. *Let $\mu=f\pi_T$ be an ultra-log-concave probability measure on $\mathbb{N}$ with $M=\frac{f(1)}{f(0)}$. Then, for any probability measure ν on $\mathbb{N}$ which is absolutely continuous with respect to μ, and has a finite first moment, we have*

$$\alpha_{TM}\left(\mathrm{W}_{1,|\cdot|}(\nu,\mu)\right) \leqslant \mathrm{H}(\nu|\mu). \tag{4.5.8}$$

Proof. We follow the proof of [DGW04, Lemma 2.1]. Fix a probability measure ν on $\mathbb{N}$ which is absolutely continuous with respect to μ, and has a finite first moment. By [DGW04, Eq. (2.1)],

$$\mathrm{H}(\nu|\mu) = \inf_{Q}\{\mathrm{H}(Q|\mathbb{P}) : Q\circ X_T^{-1}=\nu\}. \tag{4.5.9}$$

Hence, by (4.5.7), it suffices to show that

$$\alpha_{TM}\left(\mathrm{W}_{1,|\cdot|}(\nu,\mu)\right) \leqslant \inf_{Q}\left\{\alpha_{TM}\left(\mathrm{W}_{1,d}(Q,\mathbb{P})\right) : Q\circ X_T^{-1}=\nu\right\}. \tag{4.5.10}$$

Since α_{TM} is monotonic, (4.5.10) is equivalent to

$$\mathrm{W}_{1,|\cdot|}(\nu,\mu) \leqslant \inf_{Q} \left\{ \mathrm{W}_{1,d}(Q,\mathbb{P}) : Q \circ X_T^{-1} = \nu \right\}. \tag{4.5.11}$$

To establish (4.5.11), note that by Corollary 4.18, and [MSWW11, Lemma 2.3], we have that $X_T \colon (\Omega, d) \to (\mathbb{N}, |\cdot|)$ is 1-Lipschitz. Fix Q such that $Q \circ X_T^{-1} = \nu$, and let Π be the coupling attaining the minimum in the definition of $\mathrm{W}_{1,d}(Q,\mathbb{P})$. Note that $\Pi \circ X_T^{-1}$ is a coupling of $(Q \circ X_T^{-1}, \mathbb{P} \circ X_T^{-1}) = (\nu, \mu)$. Hence,

$$\begin{aligned} \mathrm{W}_{1,|\cdot|}(\nu,\mu) &\leqslant \int_{\Omega\times\Omega} |X_T(\omega) - X_T(\omega')| \,\mathrm{d}\Pi(\omega,\omega') \\ &\leqslant \int_{\Omega\times\Omega} d(\omega,\omega') \,\mathrm{d}\Pi(\omega,\omega') \\ &= \mathrm{W}_{1,d}(Q,\mathbb{P}), \end{aligned}$$

which establishes (4.5.11) by taking the infimum over Q. $\square$

Remark 4.28. It is possible in principle to improve the constant TM to $|\log \mu(0)|$ as follows. Instead of working with $\mathbb{X} = [0,T] \times [0,M]$, we can work with

$$\tilde{\mathbb{X}} := \left\{ (t,z) \in [0,T] \times \mathbb{R}_{\geqslant 0} : z \leqslant \frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)} \right\},$$

since $\lambda_t \leqslant \frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)}$ $\mathbb{P}$-a.s. (cf. (4.3.5)). Then the volume of $\tilde{\mathbb{X}}$ is $\int_0^T \frac{\mathrm{P}_{T-t} f(1)}{\mathrm{P}_{T-t} f(0)} = |\log \mu(0)|$, where the equality holds by Corollary 4.16(3). This approach, however, requires a modification of the formulation we used in this paper, with minor benefits, so we do not pursue this improvement.

Chapter 5

Stability of Wu's inequality

> Alors, grand capitaine, tu verras venir le devin
> qui te dira ta route et les mesures de ta route
> et comment revenir par la mer poissonneuse.
>
> Homère
> *L'Odyssée (traduction de Philippe Jaccottet)*

This chapter is based on the article [ALRS24], written in collaboration with Shrey Aryan and Yair Shenfeld. We provide an alternative proof to Wu's logarithmic Sobolev inequality for the Poisson measure on the nonnegative integers (see Theorem 2.92 in Chapter 2) using a stochastic variational formula for relative entropy. In addition, we characterize the extremizers of Wu's inequality and show how the stochastic approach leads to quantitative stability of the inequality under convexity assumptions.

We start by motivating the content of this chapter and providing the essential notations in Section 5.1. In Section 5.2 we introduce the Poisson-Föllmer process. In Section 5.3 we introduce the entropy representation formula, and derive from it Wu's inequality as well as a characterization of its equality cases. In Section 5.4 we prove our stability result Theorem 5.2. In addition, we compare in Section 5.4.3 between entropy stochastic representation formulas for the Gaussian and the Poisson measures, as well as their implications for logarithmic Sobolev inequalities.

5.1 Introduction

The classical logarithmic Sobolev inequality for the Gaussian measure, Theorem 2.86 in Chapter 2, has numerous applications in probability, analysis, and geometry, as we have seen in detail in Chapter 2. On the other hand, recall the discussion of Section 2.5.3 in Chapter 2: a logarithmic Sobolev inequality cannot hold for the discrete analogue of the Gaussian, namely the Poisson measure on the nonnegative integers. However, it was shown by Bobkov and Ledoux that modified logarithmic Sobolev inequalities do hold for the Poisson measure, which in turn imply concentration of measure properties for the Poisson. The sharpest form of these inequalities is due to Wu [Wu00] (who in fact proved them for more general Poisson point processes); see Section 2.5.3 of Chapter 2 for information on modified logarithmic Sobolev inequalities. We recall Theorem 2.92 in Chapter 2.

Theorem 5.1 (Wu)**.** *Fix $T > 0$, and let π_T be the Poisson measure on $\mathbb{N}$ with intensity T. Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be $L^1(\pi_T)$-integrable, and define*

$$\mathrm{Ent}_{\pi_T}[f] := \sum_{k=0}^{\infty} f(k)\log f(k)\pi_T(k) - \left(\sum_{k=0}^{\infty} f(k)\pi_T(k)\right)\log\left(\sum_{k=0}^{\infty} f(k)\pi_T(k)\right).$$

Then,

$$\mathrm{Ent}_{\pi_T}[f] \leqslant T\sum_{k=0}^{\infty} f(k+1)\left\{\log\left(\frac{f(k+1)}{f(k)}\right) - 1 + \frac{f(k)}{f(k+1)}\right\}\pi_T(k). \tag{5.1.1}$$

In the Gaussian setting, there is a beautiful proof of the classical logarithmic Sobolev inequality due to Lehec [Leh13], who showed how to deduce the inequality from a stochastic representation formula for the relative entropy with respect to the Gaussian. Our first result is to show that using a stochastic representation formula for the relative entropy with respect to the Poisson measure, due to Klartag and Lehec [KL19] who built on the earlier work of Budhiraja, Dupuis, and Maroulas [BDM11], we can give a precise analogue of this proof in the discrete setting to prove Wu's inequality (5.1.1). Our main interest in this proof technique is to deduce stability results for Wu's inequality. Concretely, in Proposition 5.12 we show that equality is attained in (5.1.1), if and only if there exist $a, b \in \mathbb{R}$ such that

$$f(k) = e^{ak+b} \quad \text{for all } k \in \mathbb{N}. \tag{5.1.2}$$

Denoting by $\delta(f)$ the deficit in (5.1.1),

$$\delta(f) := T\sum_{k=0}^{\infty} f(k+1)\left\{\log\left(\frac{f(k+1)}{f(k)}\right) - 1 + \frac{f(k)}{f(k+1)}\right\}\pi_T(k) - \mathrm{Ent}_{\pi_T}[f],$$

a stability result for Wu's inequality should lower bound $\delta(f)$ by some notion of "distance" between f and the family of equality cases (5.1.2). In the continuous setting, the study of the stability of functional inequalities is a highly active research area [Fig13]. In contrast, much fewer results seem to exist in the discrete setting. In this work, we show that, under convexity-type assumptions on f, we can use the entropy representation formula to obtain stability estimates. In the Gaussian setting, this program was carried out by Eldan, Lehec, and Shenfeld [ELS20], but the discrete setting raises new challenges which, as we show in this work, can nonetheless be (partially) overcome. We refer the reader to Section 5.4.3 for a discussion comparing the Gaussian and Poisson settings.

Our stability result will hold under the assumption that f is ultra-log-concave:

$$kf(k)^2 \geqslant (k+1)f(k+1)f(k-1) \quad \text{for all } k \in \mathbb{N}, \quad \text{with} \quad f(-1) := 0. \tag{5.1.3}$$

The equality cases (5.1.2) of Wu's inequality are not ultra-log-concave, so if f is ultra-log-concave we should be able to lower bound $\delta(f)$. Our next result provides such lower bound in terms of $\mathbb{E}[\mu]$, the mean of $\mu := f\pi_T$, and the values of $f(0), f(1)$. (These parameters naturally appear in functional inequalities for ultra-log-concave measure; cf. Remark 4.6 in Chapter 4.)

Theorem 5.2 (Stability under ultra-log-concavity)**.** *Fix $T > 0$. Let $f\colon \mathbb{N} \to (0,\infty)$ be $L^1(\pi_T)$-integrable, ultra-log-concave, and satisfying $\int f \,\mathrm{d}\pi_T = 1$. Let $\mu := f\pi_T$. Then,*

$$\delta(f) \geqslant \frac{T^2}{2}\,\Theta_{\frac{f(0)}{f(1)}}\left(\frac{\mathbb{E}[\mu]}{T}\right), \tag{5.1.4}$$

where, for $c > 0$,

$$\Theta_c(z) := \frac{z^2}{1+cz}\log\left(\frac{1}{1+cz}\right) - \frac{z^2}{1+cz} + z^2, \quad z \geqslant 0.$$

Note that the function Θ_c is nonnegative for $z \geqslant 0$, and in fact strictly positive for $z > 0$, so the right-hand side of (5.1.4) is strictly positive.

Remark 5.3. We can relate Θ_c to the relative entropy between Poisson measures of different intensities:

$$\Theta_c(z) = z^2 \mathrm{H}\left(\pi_{(1+cz)^{-1}} | \pi_1\right). \tag{5.1.5}$$

On the other hand, as we show in Proposition 5.13, the deficit $\delta(f)$ is equal to a (random) weighted sum of relative entropies between Poisson measures of different intensities. Thus, expressing stability in the sense of relative entropy distance between Poisson measures of different intensities is natural in this setting.

Remark 5.4. The function Θ_c has the correct scaling with respect to T. If we rewrite $\mu = f\pi_T$ as $\mu = g\pi_t$, with $g = f\frac{\pi_T}{\pi_t}$, then (5.1.4) becomes

$$\delta(g) \geqslant \frac{T^2}{2}\,\Theta_{\frac{T}{t}\frac{g(0)}{g(1)}}\left(\frac{t}{T}\frac{\mathbb{E}[\mu]}{t}\right) = \frac{t^2}{2}\,\Theta_{\frac{g(0)}{g(1)}}\left(\frac{\mathbb{E}[\mu]}{t}\right),$$

where we used the identity $\Theta_{rc}\left(\frac{z}{r}\right) = \frac{1}{r^2}\Theta_c(z)$ for any $r > 0$. This scaling behavior is consistent with the scaling behavior of Wu's inequality.

Remark 5.5. The question of stability under log-concavity is delicate. For example, when f is ultra-log-concave, we have that $\mu = f\pi_T$ is β-log-concave:

$$\frac{\mu(k+1)^2 - \mu(k+2)\mu(k)}{\mu(k+1)\mu(k+2)} \geqslant \beta \quad \text{for all } k \in \mathbb{N},$$

with $\beta = \frac{1}{T}\frac{f(0)}{f(1)}$ [Joh17]. On the other hand, in this setting, $\mathbb{E}[\mu] \leqslant \frac{1}{\beta}$ [Joh17, Lemma 5.3]. One might wonder whether we have the following stability estimate (which will be stronger than (5.1.4) as $z \mapsto \Theta_c(z)$ is increasing): For all β-log-concave measures,

$$\delta(f) \overset{?}{\geqslant} \frac{T^2}{2}\Theta_{T\beta}\left(\frac{1}{T\beta}\right) = \frac{1-\log(2)}{4}\frac{1}{\beta^2}. \tag{5.1.6}$$

However, taking $\beta \downarrow 0$ shows that (5.1.6) is not possible. Indeed, take $\mu = f\pi_T$ with f as in (5.1.2), so that μ is at best 0-log-concave. Then, as $\beta \downarrow 0$, the left-hand side of (5.1.6) vanishes (since f is an equality case of Wu's inequality), while the right-hand side diverges.

5.2 Proof of Wu's inequality

5.2.1 Preliminaries

For $t > 0$, we say that a function $f\colon \mathbb{N} \to \mathbb{R}$ is $L^1(\pi_t)$-integrable if

$$\int |f| \,\mathrm{d}\pi_t \coloneqq \sum_{n\in\mathbb{N}} |f(n)|\pi_t(n) < \infty.$$

Fix $t > 0$ and a function $f\colon \mathbb{N} \to \mathbb{R}$ which is $L^1(\pi_t)$-integrable. We define

$$\forall k \in \mathbb{N}, \quad \mathrm{P}_t f(k) \coloneqq \sum_{n\in\mathbb{N}} f(k+n)\pi_t(n). \tag{5.2.1}$$

For $t = 0$ we set $\mathrm{P}_0 f = f$. The Poisson semigroup is the collection of operators $(\mathrm{P}_t)_{t\geqslant 0}$. The Poisson semigroup satisfies a heat equation of the form

$$\forall k \in \mathbb{N}, \quad \partial_t \mathrm{P}_t f(k) = \mathrm{D}\mathrm{P}_t f(k) = \mathrm{P}_t\,\mathrm{D} f(k), \tag{5.2.2}$$

where

$$\forall k \in \mathbb{N}, \quad \mathrm{D} f(k) \coloneqq f(k+1) - f(k). \tag{5.2.3}$$

Fix $T > 0$. Given $f\colon \mathbb{N} \to \mathbb{R}_+$ which is $L^1(\pi_t)$-integrable we will denote

$$\forall k \in \mathbb{N}, \quad F(t,k) \coloneqq \log \mathrm{P}_{T-t} f(k), \tag{5.2.4}$$

and

$$\forall k \in \mathbb{N}, \quad G(t,k) \coloneqq e^{\mathrm{D}F(t,k)}. \tag{5.2.5}$$

Lemma 5.6. *We have*

$$\forall k \in \mathbb{N}, \quad \partial_t F(t,k) = -e^{\mathrm{D}F(t,k)} + 1, \tag{5.2.6}$$

and

$$\forall k \in \mathbb{N}, \quad \partial_t G(t,k) = -G(t,k)\,\mathrm{D}G(t,k). \tag{5.2.7}$$

Proof. Equation (5.2.6) follows from (5.2.2). For equation (5.2.7) we use (5.2.5) and (5.2.6),

$$\begin{aligned}\partial_t G(t,k) = \partial_t e^{\mathrm{D}F(t,k)} &= e^{\mathrm{D}F(t,k)}\partial_t(\mathrm{D}F(t,k)) = G(t,k)\,\mathrm{D}(\partial_t F(t,k)) \\ &= G(t,k)\,\mathrm{D}\left(-e^{\mathrm{D}F(t,k)}+1\right) = -G(t,k)\,\mathrm{D}G(t,k).\end{aligned}$$

□

5.2.2 The Poisson-Föllmer process

Fix $T > 0$, and let $\mu := f\pi_T$ be a positive probability measure on $\mathbb{N}$ with f bounded or log-concave. Klartag and Lehec [KL19], building on and specializing the work of Budhiraja, Dupuis, and Maroulas [BDM11], constructed a stochastic counting process $(X_t)_{t\in[0,T]}$ such that $X_T \sim \mu$. We will describe the process briefly and refer to [KL19] for a complete description[1]. We let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space on which the following random variables are defined. We let N be a Poisson point process on $[0,T]\times\mathbb{R}_{\geqslant 0}$ with Lebesgue intensity measure. Then $N(F)$, for a Borel set $F \subset [0,T]\times\mathbb{R}_{\geqslant 0}$, is a Poisson random variable with intensity equal to the Lebesgue measure of F. For $t\in[0,T]$ we let $\mathcal{F}_t$ be the sigma-algebra generated by the following collection of random variables,

$$\mathcal{F}_t := \sigma\left(\{N(F) : F \subset [0,t]\times\mathbb{R}_{\geqslant 0} \text{ is a Borel set}\}\right).$$

The collection $(\mathcal{F}_t)_{t\in[0,T]}$ is a filtration, and we say that a stochastic process $(\lambda_t)_{t\in[0,T]}$, where $\lambda_t : \Omega \to \mathbb{R}$, is predictable, if the function $(t,\omega)\mapsto\lambda_t(\omega)$ is measurable with respect to the sigma-algebra $\sigma(\{(s,t]\times B : s\leqslant t\leqslant T, B\in\mathcal{F}_s\})$. Let $(\lambda_t)_{t\in[0,T]}$ be a predictable, nonnegative, and bounded stochastic process. Define

$$X_t^{\lambda}(\omega) := N\left(\{(s,x)\in[0,T]\times\mathbb{R}_{\geqslant 0} : s<t,\ x\leqslant\lambda_s(\omega)\}\right), \quad \text{(see Figure 5.1)}. \tag{5.2.8}$$

Klartag and Lehec [KL19] showed that we can choose a particular density $(\lambda_t)_{t\in[0,T]}$ given by

$$\lambda_t := \frac{\mathrm{P}_{T-t}f(X_t+1)}{\mathrm{P}_{T-t}f(X_t)}, \tag{5.2.9}$$

where (P_t) is the Poisson semigroup, and that the resulting process $(X_t)_{t\in[0,T]} := (X_t^{\lambda})_{t\in[0,T]}$ is well-defined. Moreover, $X_T\sim\mu$. We call this process the Poisson-Föllmer process, since it is the discrete analogue of the Föllmer process in continuous setting; see Section 5.4.3. Let us establish some useful properties of $(X_t)_{t\in[0,T]}$ and $(\lambda_t)_{t\in[0,T]}$. First, we recall Lemma 4.14 from Chapter 4.

Lemma 5.7. *Let $(X_t)_{t\in[0,T]}$ be the Poisson-Föllmer process. Then, $X_t\sim(\mathrm{P}_{T-t}f)\pi_t$.*

Lemma 5.8.

(1) Denote the compensated process $(\tilde{X}_t)_{t\in[0,T]}$ as

$$\tilde{X}_t := X_t - \int_0^t \lambda_s\,\mathrm{d}s. \tag{5.2.10}$$

Then, $(\tilde{X}_t)_{t\in[0,T]}$ is a martingale.

[1] Recall Chapter 4 for the justification of replacing the boundedness assumption by log-concavity.

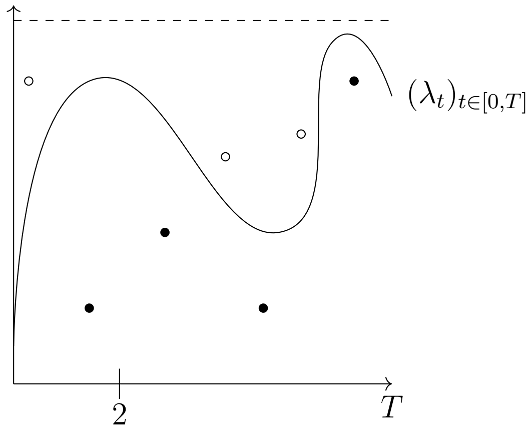

Figure 5.1: The points in $[0,T]\times\mathbb{R}_{\geqslant 0}$ are generated according to a standard Poisson process (7 points in this case). At time $t\in[0,T]$ the value of the process X_t^λ is equal to the number of points under the curve (filled circles). In the figure $X_2^\lambda = 1$ and $X_T^\lambda = 4$.

(2) The process $(\lambda_t)_{t\in[0,T]}$ is a martingale.

(3) For every $t\in[0,T]$,

$$\mathbb{E}[X_t] = t\,\mathbb{E}[\lambda_t] = \frac{t}{T}\mathbb{E}[\mu]. \tag{5.2.11}$$

Proof. Item (1) can be found in [KL19, p. 100], while item (2) can be found [LRS25, Lemma 3.3] (see also Lemma 5.9 below). Taking expectation in (5.2.10), and using item (2), we find that $\mathbb{E}[X_t] = t\,\mathbb{E}[\lambda_t]$. Using $\mathbb{E}[X_t] = t\,\mathbb{E}[\lambda_t]$ with $t=T$ gives, since $X_T\sim\mu$, $\mathbb{E}[\mu] = \mathbb{E}[X_T] = T\,\mathbb{E}[\lambda_T] = T\,\mathbb{E}[\lambda_t]$, where the last equality holds by item (2). It follows that $\mathbb{E}[\lambda_t] = \frac{\mathbb{E}[\mu]}{T}$, and hence $t\mathbb{E}[\lambda_t] = \frac{t}{T}\mathbb{E}[\mu]$. □

As we saw in Lemma 5.8, the process $(\lambda_t)_{t\in[0,T]}$ is a martingale. The next result gives an explicit expression for this martingale. To this end we recall (5.2.5) and note that

$$\lambda_t = G(t, X_t). \tag{5.2.12}$$

We recall that a stochastic integral of the form $\int\cdot\,\mathrm{d}X_t$ is simply the sum of the integrand at the jump points of the process $(X_t)_{t\in[0,T]}$; cf. [KL19, LRS25]. The integral against the compensated process $(\tilde{X}_t)_{t\in[0,T]}$ (Lemma 5.8) is defined as the sum of the integrals $\int\cdot\,\mathrm{d}X_t$ and $\int\cdot\,\mathrm{d}t$.

Lemma 5.9. *We have*

$$\mathrm{d}\lambda_t = \mathrm{D}G(t,X_t)\,\mathrm{d}\tilde{X}_t. \tag{5.2.13}$$

Proof.

$$\begin{aligned}\lambda_t = G(t,X_t) &= \int_0^t \partial_s G(s,X_s)\,\mathrm{d}s + \int_0^t \mathrm{D}G(s,X_s)\,\mathrm{d}X_s\\ &= \int_0^t [\partial_s G(s,X_s) + \mathrm{D}G(s,X_s)G(s,X_s)]\,\mathrm{d}s + \int_0^t \mathrm{D}G(s,X_s)\,\mathrm{d}\tilde{X}_s,\end{aligned}$$

and the first integrand vanishes by (5.2.7). □

5.3 The entropy representation formula and Wu's inequality

5.3.1 The entropy representation formula

Given a positive measure μ on $\mathbb{N}$ we define its relative entropy with respect to π_T by

$$\mathrm{H}(\mu|\pi_T) := \sum_{k=0}^{\infty} \log\left(\frac{\mu(k)}{\pi_T(k)}\right)\mu(k), \tag{5.3.1}$$

and note that if $\mu = f\pi_T$ then

$$\mathrm{H}(\mu|\pi_T) = \mathrm{Ent}_{\pi_T}[f]. \tag{5.3.2}$$

Let

$$\Phi(r) := r\log r \quad \text{for } r \geqslant 0, \tag{5.3.3}$$

and define the function $\Psi(u, v)$, for $u > 0$ and $u + v > 0$, by

$$\begin{aligned}\Psi(u, v) &:= \Phi(u + v) - \Phi(u) - \Phi'(u)v \\ &= (u + v)\log(u + v) - u\log u - (\log u + 1)v.\end{aligned} \tag{5.3.4}$$

Note that Ψ is nonnegative and convex when $u > 0$ and $u + v > 0$ [Wu00, §1.4]. With the above notation, Wu's inequality (5.1.1) can be written as,

$$\mathrm{H}(\mu|\pi_T) \leqslant T\,\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] \quad \text{for all probability measures } \mu = f\pi_T \text{ on } \mathbb{N}. \tag{5.3.5}$$

We now come to our main tool which is a representation formula for $\mathrm{H}(\mu|\pi_T)$. This representation formula, which already appears as a special case of [Wu00, Equation (1.7)], and also follows from the variational formula in [KL19, Remark 3, p. 102], will be used to give an elementary proof of Wu's inequality, characterize its equality cases, and also yield a stability-type result. In the above works the entropy representation formula appears under stronger assumptions that are not sufficient for the characterization of the equality cases of Wu's inequality, so we will provide a more elementary proof which is valid under weaker regularity assumptions. The following result does not use the Poisson-Föllmer process, but we will show how it can be obtained (under further regularity) using the stochastic approach.

Proposition 5.10 (Entropy representation formula). *Let $\mu = f\pi_T$ be a probability measure on $\mathbb{N}$ such that $\mathrm{H}(\mu|\pi_T) < \infty$ and $\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] < \infty$. Then,*

$$\mathrm{H}(\mu|\pi_T) = \int_0^T \mathbb{E}_{\pi_t}[\Psi(\mathrm{P}_{T-t}f, \mathrm{D}\mathrm{P}_{T-t}f)]\,\mathrm{d}t. \tag{5.3.6}$$

Proof. First note that since Ψ is convex [Wu00, §1.4], Jensen's inequality yields

$$\Psi(\mathrm{P}_{T-t}f, \mathrm{D}\mathrm{P}_{T-t}f) \leqslant \mathrm{P}_{T-t}\Psi(f, \mathrm{D}f),$$

and hence the integrand on the right-hand side of (5.3.6) is always finite,

$$\mathbb{E}_{\pi_t}[\Psi(\mathrm{P}_{T-t}f, \mathrm{D}\mathrm{P}_{T-t}f)] \leqslant \mathbb{E}_{\pi_t}[\mathrm{P}_{T-t}\Psi(f, \mathrm{D}f)] = \mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] < \infty. \tag{5.3.7}$$

Fix $k \in \mathbb{N}$ and define $\alpha\colon [0,T] \to \mathbb{R}$ by

$$\alpha(t) \coloneqq \mathbb{E}_{\pi_t}[\mathrm{P}_{T-t}f \log \mathrm{P}_{T-t}f], \tag{5.3.8}$$

and note that the convexity of $r \mapsto r \log r$ and Jensen's inequality yield

$$\alpha(t) = \mathbb{E}_{\pi_t}[\mathrm{P}_{T-t}f \log \mathrm{P}_{T-t}f] \leqslant \mathbb{E}_{\pi_t}\mathrm{P}_{T-t}[f \log f] = \mathrm{H}(\mu|\pi_T) < \infty.$$

Since $\pi_0 = \delta_0$ and $\mathrm{P}_T f(0) = \mathbb{E}_{\pi_T}[f] = 1$ we have

$$\mathrm{H}(\mu|\pi_T) = \int_0^T \partial_t \alpha(t)\,\mathrm{d}t. \tag{5.3.9}$$

Using (5.2.2), (5.2.6), and $\partial_t \pi_t(k) = \pi_t(k-1) - \pi_t(k)$, with $\pi_t(-1) \coloneqq 0$, we get

$$\begin{aligned}
\partial_t \alpha(t) = &\sum_{k=0}^{\infty} [\mathrm{P}_{T-t}f(k) - \mathrm{P}_{T-t}f(k+1)]\pi_t(k) \\
&+ \sum_{k=0}^{\infty} \log \mathrm{P}_{T-t}f(k)[\mathrm{P}_{T-t}f(k) - \mathrm{P}_{T-t}f(k+1)]\pi_t(k) \\
&+ \sum_{k=0}^{\infty} \mathrm{P}_{T-t}f(k) \log \mathrm{P}_{T-t}f(k)[\pi_t(k-1) - \pi_t(k)].
\end{aligned} \tag{5.3.10}$$

The second and third lines in (5.3.10) can be written as, using $\pi_t(-1) = 0$,

$$\begin{aligned}
&\sum_{k=0}^{\infty} \log \mathrm{P}_{T-t}f(k)[\mathrm{P}_{T-t}f(k) - \mathrm{P}_{T-t}f(k+1)]\pi_t(k) \\
&\quad + \sum_{k=0}^{\infty} \mathrm{P}_{T-t}f(k) \log \mathrm{P}_{T-t}f(k)[\pi_t(k-1) - \pi_t(k)] \\
&= -\sum_{k=0}^{\infty} \mathrm{P}_{T-t}f(k+1) \log \mathrm{P}_{T-t}f(k)\pi_t(k) + \sum_{k=1}^{\infty} \mathrm{P}_{T-t}f(k) \log \mathrm{P}_{T-t}f(k)\pi_t(k-1) \\
&= -\sum_{k=0}^{\infty} \mathrm{P}_{T-t}f(k+1) \log \mathrm{P}_{T-t}f(k)\pi_t(k) + \sum_{k=0}^{\infty} \mathrm{P}_{T-t}f(k+1) \log \mathrm{P}_{T-t}f(k+1)\pi_t(k),
\end{aligned}$$

so

$$\begin{aligned}
\partial_t \alpha(t) &= \sum_{k=0}^{\infty} \{\mathrm{P}_{T-t}f(k+1) \log \mathrm{P}_{T-t}f(k+1) - \mathrm{P}_{T-t}f(k+1) \log \mathrm{P}_{T-t}f(k) \\
&\qquad - [\mathrm{P}_{T-t}f(k) - \mathrm{P}_{T-t}f(k+1)]\}\pi_t(k) \\
&= \mathbb{E}_{\pi_t}[\Psi(\mathrm{P}_{T-t}f, \mathrm{D}\mathrm{P}_{T-t}f)].
\end{aligned} \tag{5.3.11}$$

□

When the Poisson-Föllmer process exist we can rewrite the entropy representation formula as follows.

Corollary 5.11 (Stochastic entropy representation formula)**.** *Let $\mu = f\pi_T$ be a positive probability measure for which the Poisson-Föllmer process $(X_t)_{t\in[0,T]}$ associated to μ exists. Let $(\lambda_t)_{t\in[0,T]}$ be as in* (5.2.9)*. Then,*

$$\mathrm{H}(\mu|\pi_T) = \int_0^T \mathbb{E}[\lambda_t \log \lambda_t - \lambda_t + 1]\,\mathrm{d}t. \tag{5.3.12}$$

Proof. Equation (5.3.12) is a rewrite of Equation (5.3.6) where we used the definition of λ_t and the fact that $X_t \sim (\mathrm{P}_{T-t}f)\pi_t$ (Lemma 5.7). □

5.3.2 Wu's inequality and its equality cases

We can now prove Wu's inequality (5.3.5) and also characterize its extremizers. For the proof of the inequality we can in principle use either Proposition 5.10 or Corollary 5.11, but to get the characterization of the equality cases without imposing unnecessary assumptions we will need to use Proposition 5.10.

For the sake of exposition, let us begin however by deriving Wu's inequality (5.3.5) from Corollary 5.11. It suffices to note that $(\lambda_t)_{t\in[0,T]}$ is a martingale (Lemma 5.8(2)), and that the function $(0,\infty) \ni r \mapsto r\log r - r + 1$ is convex, to get from (5.3.12),

$$\mathrm{H}(\mu|\pi_T) = \int_0^T \mathbb{E}[\lambda_t \log \lambda_t - \lambda_t + 1]\,\mathrm{d}t \leqslant T\,\mathbb{E}[\lambda_T \log \lambda_T - \lambda_T + 1] = T\,\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)]. \tag{5.3.13}$$

To get the inequality for all $L^1(\pi_T)$-integrable functions we can use approximation as in [Wu00].

We now turn to the proof of Wu's inequality and the characterization of its equality cases under minimal assumptions.

Proposition 5.12 (Wu's inequality and its equality cases)**.** *Fix $T > 0$, and let π_T be the Poisson measure on $\mathbb{N}$ with intensity T. Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be $L^1(\pi_T)$-integrable and let $\mu \coloneqq f\pi_T$. Then,*

$$\mathrm{H}(\mu|\pi_T) \leqslant T\,\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)]. \tag{5.3.14}$$

Further, if $\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] < \infty$ then equality holds if and only if $f(k) = e^{ak+b}$ for some $a, b \in \mathbb{R}$.

Proof. We start with proof of (5.3.14). Without loss of generality we may assume that $\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] < \infty$ or else (5.3.14) holds trivially. Further, we assume that $\mathrm{H}(\mu|\pi_T) < \infty$ since otherwise we can use approximation as in [Wu00]. By Proposition 5.10, and arguing as in (5.3.7), we get

$$\mathrm{H}(\mu|\pi_T) = \int_0^T \mathbb{E}_{\pi_t}[\Psi(\mathrm{P}_{T-t}f, \mathrm{D}\mathrm{P}_{T-t}f)]\,\mathrm{d}t \leqslant T\,\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)], \tag{5.3.15}$$

which establishes (5.3.14). To obtain the equality cases, first note that equality holds if $f(k) = e^{ak+b}$ for some $a, b \in \mathbb{R}$. For the reverse direction, we note that equality in (5.3.15) implies that $t \mapsto \mathbb{E}_{\pi_t}[\Psi(\mathrm{P}_{T-t}f, \mathrm{D}\mathrm{P}_{T-t}f)]$ is a constant, and hence

$$\mathbb{E}_{\pi_T}[\Psi(f, \mathrm{D}f)] = \mathbb{E}_{\pi_0}[\Psi(\mathrm{P}_T f, \mathrm{D}\mathrm{P}_T f)] = \Psi(\mathrm{P}_T f, \mathrm{D}\mathrm{P}_T f)(0) = \Psi(\mathbb{E}_{\pi_T}[(f, \mathrm{D}f)]). \tag{5.3.16}$$

Using the relation

$$\Psi(u,v) = u\phi\left(\frac{u+v}{u}\right) \qquad \text{for all } u>0 \text{ and } u+v>0, \tag{5.3.17}$$

with

$$\phi(r) := r\log r - r + 1,$$

we have

$$\mathbb{E}_{\pi_T}[\Psi(f,\mathrm{D}f)] = \mathbb{E}_\mu\left[\phi\left(\frac{f+\mathrm{D}f}{f}\right)\right] = \sum_{k=0}^\infty \phi\left(\frac{f(k+1)}{f(k)}\right)\mu(k),$$

and, as $\mathbb{E}_{\pi_T}[f]=1$,

$$\begin{aligned}\Psi\left(\mathbb{E}_{\pi_T}[(f,\mathrm{D}f)]\right) = \mathbb{E}_{\pi_T}[f]\phi\left(\frac{\mathbb{E}_{\pi_T}[\mathrm{D}f]+\mathbb{E}_{\pi_T}[f]}{\mathbb{E}_{\pi_T}[f]}\right) &= \phi\left(\mathbb{E}_{\pi_T}[f+\mathrm{D}f]\right)\\ &= \phi\left(\sum_{k=0}^\infty \frac{f(k+1)}{f(k)}\mu(k)\right).\end{aligned}$$

Since ϕ is strictly convex on $(0,\infty)$, we conclude by the equality cases of Jensen's inequality that there exists a constant c such that

$$\forall k\in\mathbb{N}, \quad f(k+1) = cf(k), \tag{5.3.18}$$

which shows that $f(k)=e^{ak+b}$ for some $a,b\in\mathbb{R}$. □

5.4 Stability of Wu's inequality

5.4.1 An identity for the deficit in Wu's inequality

Recall that our stability result (Theorem 5.2) is stated for ultra-log-concave functions f, for which the entropy representation formula (5.3.12) holds. The proof of Theorem 5.2 will rely on an identity for the deficit in Wu's inequality (5.1.1),

$$\delta(f) := T\sum_{k=0}^\infty f(k+1)\left\{\log\left(\frac{f(k+1)}{f(k)}\right) - 1 + \frac{f(k)}{f(k+1)}\right\}\pi_T(k) - \mathrm{Ent}_{\pi_T}[f]. \tag{5.4.1}$$

To state the deficit identity denote $\mathrm{D}^2 := \mathrm{D}\circ\mathrm{D}$, and also recall the definition (5.2.4) of $F(t,k)$.

Proposition 5.13. *Let $f\colon\mathbb{N}\to\mathbb{R}_{>0}$ be ultra-log-concave such that $\int f\,\mathrm{d}\pi_T=1$, and let $\mu := f\pi_T$. Let $(X_t)_{t\in[0,T]}$ be the Poisson-Föllmer process associated with μ, with $(\lambda_t)_{t\in[0,T]}$ defined by* (5.2.9)*. Then,*

$$\delta(f) = \int_0^T\int_t^T \mathbb{E}\left[\lambda_s^2\,\mathrm{H}\left(\pi_{e^{\mathrm{D}^2F(s,X_s)}}\middle|\pi_1\right)\right]\mathrm{d}s\,\mathrm{d}t.$$

The proof of Proposition 5.13 requires a number of preliminary results. Recall that by Corollary 5.11,

$$\mathrm{H}(\mu|\pi_T) = \int_0^T \mathbb{E}[\phi(\lambda_t)]\,\mathrm{d}t,$$

where $\phi : (0,\infty) \to [0,\infty)$ is a convex function given by

$$\phi(r) \coloneqq r\log r - r + 1. \tag{5.4.2}$$

The next result uses Taylor expansion to re-express $\mathbb{E}[\phi(\lambda_t)]$, and hence $\mathrm{H}(\mu|\pi_T)$.

Lemma 5.14. *Let G be as in* (5.2.5)*. Then,*

$$\mathbb{E}[\phi(\lambda_t)] = \int_0^t \lambda_s[\phi(\lambda_s + \mathrm{D}G(s,X_s)) - \phi(\lambda_s) - \phi'(\lambda_s)\,\mathrm{D}G(s,X_s)]\,\mathrm{d}s.$$

Proof. Recall the martingale $(\tilde{X}_t)_{t\in[0,T]}$ from Lemma 5.8, and recall Equation (5.2.12). We have

$$\begin{aligned}
\phi(\lambda_t) = \phi(G(t,X_t)) &= \int_0^t \partial_s(\phi\circ G)(s,X_s)\,\mathrm{d}s + \int_0^t \mathrm{D}(\phi\circ G)(s,X_s)\,\mathrm{d}X_s \\
&= \int_0^t \phi'(G(s,X_s))\partial_s G(s,X_s)\,\mathrm{d}s + \int_0^t [\phi(G(s,X_s+1)) - \phi(G(s,X_s))]\,\mathrm{d}X_s \\
&\overset{\text{Lemma 5.6}}{=} \int_0^t [-\phi'(\lambda_s)\lambda_s\,\mathrm{D}G(s,X_s)]\,\mathrm{d}s \\
&\quad + \int_0^t [\phi(G(s,X_s) + \mathrm{D}G(s,X_s)) - \phi(G(s,X_s))]\,\mathrm{d}X_s \\
&= \int_0^t \lambda_s[\phi(G(s,X_s) + \mathrm{D}G(s,X_s)) - \phi(G(s,X_s)) - \phi'(\lambda_s)\,\mathrm{D}G(s,X_s)]\,\mathrm{d}s \\
&\quad + \int_0^t [\phi(G(s,X_s) + \mathrm{D}G(s,X_s)) - \phi(G(s,X_s))]\,\mathrm{d}\tilde{X}_s \\
&= \int_0^t \lambda_s[\phi(\lambda_s + \mathrm{D}G(s,X_s)) - \phi(\lambda_s) - \phi'(\lambda_s)\,\mathrm{D}G(s,X_s)]\,\mathrm{d}s \\
&\quad + \int_0^t [\phi(G(s,X_s) + \mathrm{D}G(s,X_s)) - \phi(G(s,X_s))]\,\mathrm{d}\tilde{X}_s.
\end{aligned}$$

Since $(\tilde{X}_s)$ is a martingale, taking expectation completes the proof. □

We need two more results. The first is an identity for the relative entropy between Poisson measures with different intensities, whose proof is immediate.

Lemma 5.15.

$$\forall \alpha,\beta > 0, \quad \mathrm{H}(\pi_\alpha|\pi_\beta) = \beta\,\mathrm{H}\left(\pi_{\frac{\alpha}{\beta}}|\pi_1\right) = \alpha\left(\frac{\beta}{\alpha} - \log\left(\frac{\beta}{\alpha}\right) - 1\right).$$

The second result shows that the integrand in Lemma 5.14 can be written as relative entropy between Poisson measures. The proof is again immediate.

Lemma 5.16. *Let $\phi(r) = r\log r - r + 1$ for $r \geqslant 0$. Then, for any $x, y > 0$,*

$$\phi(y) - \phi(x) - \phi'(x)(y-x) = \mathrm{H}(\pi_y|\pi_x).$$

Combining the above results we can now prove the deficit identity.

Proof of Proposition 5.13. By Lemma 5.14, Lemma 5.15, and Lemma 5.16 we have

$$\begin{aligned}\partial_s\mathbb{E}[\phi(\lambda_s)] &\overset{\text{Lemma 5.14}}{=} \mathbb{E}\left[\lambda_s\{\phi(\lambda_s + \mathrm{D}G(s,X_s)) - \phi(\lambda_s) - \phi'(\lambda_s)\,\mathrm{D}G(s,X_s)\}\right] \\ &\overset{\text{Lemma 5.16}}{=} \mathbb{E}\left[\lambda_s\,\mathrm{H}\left(\pi_{\lambda_s+\mathrm{D}G(s,X_s)}|\pi_{\lambda_s}\right)\right] \overset{\text{Lemma 5.15}}{=} \mathbb{E}\left[\lambda_s^2\,\mathrm{H}\left(\pi_{\frac{\lambda_s+\mathrm{D}G(s,X_s)}{\lambda_s}}\middle|\pi_1\right)\right] \\ &= \mathbb{E}\left[\lambda_s^2\,\mathrm{H}\left(\pi_{e^{\mathrm{D}^2F(s,X_s)}}\middle|\pi_1\right)\right],\end{aligned}$$

where the last equality used

$$\mathrm{D}G(t,k) = \frac{\mathrm{P}_{T-t}f(k+2)}{\mathrm{P}_{T-t}f(k+1)} - \frac{\mathrm{P}_{T-t}f(k+1)}{\mathrm{P}_{T-t}f(k)} = \frac{\mathrm{P}_{T-t}f(k+2)}{\mathrm{P}_{T-t}f(k+1)} - G(t,k),$$

so

$$\frac{G(t,k) + \mathrm{D}G(t,k)}{G(t,k)} = \frac{\mathrm{P}_{T-t}f(k+2)\mathrm{P}_{T-t}f(k)}{\mathrm{P}_{T-t}f(k+1)^2} = e^{\mathrm{D}^2\log\mathrm{P}_{T-t}f(k)}.$$

Finally, the proof is complete since by Proposition 5.10 and the proof of Theorem 5.1,

$$\delta(f) = \int_0^T \mathbb{E}[\phi(\lambda_T) - \phi(\lambda_t)]\,\mathrm{d}t = \int_0^T\int_t^T \partial_s\mathbb{E}[\phi(\lambda_s)]\,\mathrm{d}s\,\mathrm{d}t. \tag{5.4.3}$$

$\square$

Proposition 5.13 is our main tool to obtain a deficit estimate by almost-surely lower bounding $\mathrm{H}\left(\pi_{e^{\mathrm{D}^2F(s,X_s)}}|\pi_1\right)$. Specifically, Lemma 5.15 shows that the map $\alpha \mapsto \mathrm{H}(\pi_\alpha|\pi_1)$ is decreasing on $(0,1]$, so to lower bound $\mathrm{H}\left(\pi_{e^{\mathrm{D}^2F(s,X_s)}}|\pi_1\right)$ we can show that $e^{\mathrm{D}^2F(s,X_s)} \leqslant 1$, and then upper bound $e^{\mathrm{D}^2F(s,X_s)}$. We will show this can indeed be done when f is ultra-log-concave.

5.4.2 Stability of Wu's inequality under ultra-log-concavity

In this section we prove Theorem 5.2:

Theorem 5.17. *Fix $T > 0$. Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be $L^1(\pi_T)$ integrable and ultra-log-concave, and let $\mu := \frac{f\pi_T}{\int f\,\mathrm{d}\pi_T}$. Then,*

$$\delta(f) \geqslant \frac{T^2}{2}\,\Theta_{\frac{f(0)}{f(1)}}\left(\frac{\mathbb{E}[\mu]}{T}\right),$$

where, for $c > 0$,

$$\Theta_c(z) := \frac{z^2}{1+cz}\log\left(\frac{1}{1+cz}\right) - \frac{z^2}{1+cz} + z^2, \quad z \geqslant 0.$$

Our first task is to show the preservation of ultra-log-concavity under the Poisson flow.

Lemma 5.18. *Let $f\colon \mathbb{N} \to \mathbb{R}_+$ be an ultra-log-concave function. Then, for every $t \geqslant 0$, $\mathrm{P}_t f\colon \mathbb{N} \to \mathbb{R}_+$ is also ultra-log-concave.*

Proof. The lemma follows from the closability under convolutions of ultra-log-concave functions [Wal76]: If $\{a_k\}_{k\in\mathbb{Z}}, \{b_k\}_{k\in\mathbb{Z}}$ are ultra-log-concave then $\{(a*b)_k\}_{k\in\mathbb{Z}}$ is ultra-log-concave as well where

$$(a*b)_k := \sum_{n=-\infty}^{+\infty} a_n b_{k-n} = \sum_{n=-\infty}^{+\infty} a_{k-n} b_n. \tag{5.4.4}$$

To apply (5.4.4) in our setting we extend f and π_t to $\mathbb{Z}$ by setting them to zero on $\{k \in \mathbb{Z} : k < 0\}$, and note that f and π_t remain ultra-log-concave as functions on $\mathbb{Z}$. Next we define $\tilde{\pi}_t : \mathbb{Z} \to \mathbb{R}_+$ by $\tilde{\pi}_t(k) := \pi_t(-k)$ for $k \in \mathbb{Z}$, and note that $\tilde{\pi}_t$ is also ultra-log-concave on $\mathbb{Z}$. The lemma now follows since $\mathrm{P}_t f = (f * \tilde{\pi}_t)$. ☐

The next result gives a useful bound for ultra-log-concave functions.

Lemma 5.19. *Let $f\colon \mathbb{N} \to \mathbb{R}_{>0}$ be an ultra-log-concave function. Then, for every $k \in \mathbb{N}$,*

$$\frac{f(k+2)f(k)}{f^2(k+1)} \leqslant \frac{1}{1 + c\frac{f(k+1)}{f(k)}} < 1 \quad \textit{with} \quad c := \frac{f(0)}{f(1)}.$$

Proof. By [Joh17, Lemma 5.1] the fact that f is ultra-log-concave means that f is c-log-concave with $c := \frac{f(0)}{f(1)}$:

$$\frac{f(k+1)^2 - f(k+2)f(k)}{f(k+1)f(k+2)} \geqslant c.$$

The result follows by rearrangement. ☐

We are now ready for the proof of stability in Wu's inequality under ultra-log-concavity.

Proof of Theorem 5.17. By approximation we may assume that f is bounded. The function $\mathrm{P}_{T-t}f$ is ultra-log-concave by Lemma 5.18, so Lemma 5.19 can be applied to give

$$e^{\mathrm{D}^2 F(s,X_s)} = \frac{\mathrm{P}_{T-s}f(X_s+2)\mathrm{P}_{T-s}f(X_s)}{(\mathrm{P}_{T-s}f)^2(X_s+1)} \leqslant \frac{1}{1 + c_s\frac{\mathrm{P}_{T-s}f(X_s+1)}{\mathrm{P}_{T-s}f(X_s)}} = \frac{1}{1 + c_s\lambda_s}, \tag{5.4.5}$$

with $c_s := \frac{\mathrm{P}_{T-s}f(0)}{\mathrm{P}_{T-s}f(1)}$. Next we will lower bound c_s. We claim that

$$c_s = \frac{\mathrm{P}_{T-s}f(0)}{\mathrm{P}_{T-s}f(1)} \geqslant \frac{f(0)}{f(1)}. \tag{5.4.6}$$

Indeed, it suffices to show that the function $[0,T] \ni s \mapsto \eta(s) := \frac{\mathrm{P}_{T-s}f(0)}{\mathrm{P}_{T-s}f(1)}$ is non-increasing since the right-hand side above is equal to $\eta(T)$. The latter holds since, using (5.2.2), we have

$$\partial_s \eta(s) = -\frac{1}{(\mathrm{P}_s f)^2(1)} \left\{ (\mathrm{P}_s f)^2(1) - \mathrm{P}_s f(2) \mathrm{P}_s f(0) \right\} \leqslant 0,$$

where the inequality holds as $\mathrm{P}_{T-s} f$ is ultra-log-concave (Lemma 5.18). Combining (5.4.5) and (5.4.6) we conclude that

$$e^{\mathrm{D}^2 F(s,X_s)} \leqslant \frac{1}{1 + \frac{f(0)}{f(1)}\lambda_s} < 1. \tag{5.4.7}$$

Since $\alpha \mapsto \mathrm{H}(\pi_\alpha|\pi_1)$ is decreasing on $(0,1]$ we get from (5.4.5) that

$$\mathrm{H}\left(\pi_{e^{\mathrm{D}^2 F(s,X_s)}} \middle| \pi_1\right) \geqslant \mathrm{H}\left(\pi_{(1+\frac{f(0)}{f(1)}\lambda_s)^{-1}} \middle| \pi_1\right).$$

Hence, it follows from Proposition 5.13 that

$$\delta(f) \geqslant \int_0^T \int_t^T \mathbb{E}\left[\lambda_s^2 \mathrm{H}\left(\pi_{(1+\frac{f(0)}{f(1)}\lambda_s)^{-1}} \middle| \pi_1\right)\right] \mathrm{d}s\, \mathrm{d}t = \int_0^T \int_t^T \mathbb{E}\left[\Theta_{\frac{f(0)}{f(1)}}(\lambda_s)\right] \mathrm{d}s\, \mathrm{d}t,$$

where the last equality follows from the definition of Θ_c and Lemma 5.15. The function $z \mapsto \Theta_c(z)$ can be verified to be convex, so by Jensen's inequality,

$$\delta(f) \geqslant \int_0^T \int_t^T \Theta_{\frac{f(0)}{f(1)}}(\mathbb{E}[\lambda_s])\, \mathrm{d}s\, \mathrm{d}t \overset{(5.2.11)}{=} \int_0^T \int_t^T \Theta_{\frac{f(0)}{f(1)}}\left(\frac{\mathbb{E}[\mu]}{T}\right) \mathrm{d}s\, \mathrm{d}t = \frac{T^2}{2} \Theta_{\frac{f(0)}{f(1)}}\left(\frac{\mathbb{E}[\mu]}{T}\right).$$

□

5.4.3 Comparison with the Gaussian setting

The analogue of Wu's inequality in the Gaussian setting is the logarithmic Sobolev inquality: Let γ be the standard Gaussian measure on $\mathbb{R}^n$. Then, for any $\mu = f\gamma$ (for which the quantities below are well-defined),

$$\mathrm{H}(\mu|\gamma) \leqslant \frac{1}{2} \int_{\mathbb{R}^n} |\nabla \log f|^2\, \mathrm{d}\mu. \tag{5.4.8}$$

Let us now present Lehec's proof [Leh13] of (5.4.8) using the Gaussian analogue of the entropy representation formula (5.3.12). In the continuous setting (taking $T = 1$ for simplicity), the Poisson-Föllmer process is replaced by the Föllmer process [Föl85, Föl86, Leh13], which is the solution of the following stochastic differential equation,

$$\mathrm{d}Y_t = v_t\, \mathrm{d}t + \mathrm{d}B_t, \quad Y_0 = 0, \tag{5.4.9}$$

where $(B_t)_{t\geqslant 0}$ is a standard Brownian motion in $\mathbb{R}^n$,

$$v_t := \nabla \log \mathrm{Q}_{1-t} f(Y_t), \tag{5.4.10}$$

with (Q_t) the heat semigroup,

$$\mathrm{Q}_t f(x) \coloneqq \int_{\mathbb{R}^n} f(x+\sqrt{t}z)\,\mathrm{d}\gamma(z). \tag{5.4.11}$$

The process $(Y_t)_{t\in[0,1]}$ satisfies $Y_1 \sim \mu = f\pi$, and we have the entropy representation formula

$$\mathrm{H}(\mu|\gamma) = \frac{1}{2}\int_0^1 \mathbb{E}[\varphi(v_t)]\,\mathrm{d}t, \tag{5.4.12}$$

where

$$\varphi(x) \coloneqq \frac{x^2}{2}. \tag{5.4.13}$$

The representation (5.4.12) is the analogue of (5.3.12). In Table 5.1 we summarize the comparisons of the stochastic constructions in the Poisson and Gaussian settings.

	Poisson	**Gaussian**
Process	X_t (5.2.8)	Y_t (5.4.9)
Control	λ_t (5.2.9)	v_t (5.4.10)
Semigroup	P_t (5.2.1)	Q_t (5.4.11)
Rate function	ϕ (5.4.2)	φ (5.4.13)
Entropy representation formula	(5.3.12)	(5.4.12)

Table 5.1: Comparison between Poisson and Gaussian

Let us now present Lehec's proof of (5.4.8) using (5.4.12). It can be shown that the process (v_t) is a martingale, so since φ is convex, it follows from Jensen's inequality that

$$\mathrm{H}(\mu|\gamma) = \frac{1}{2}\int_0^1 \mathbb{E}[\varphi(v_t)]\,\mathrm{d}t \leqslant \frac{1}{2}\mathbb{E}[\varphi(v_1)] = \frac{1}{2}\int_{\mathbb{R}^n} |\nabla \log f|^2\,\mathrm{d}\mu. \tag{5.4.14}$$

Thus, we see that our proof of Wu's inequality is the exact discrete analogue of Lehec's proof of the Gaussian logarithmic Sobolev inequality. Turning to the question of stability estimates, Eldan, Lehec, and Shenfeld [ELS20] used (5.4.12) to get stability estimates for the Gaussian logarithmic Sobolev inequality. Denote the deficit in the Gaussian logarithmic Sobolev inequality as

$$\delta(f) \coloneqq \frac{1}{2}\int_{\mathbb{R}^n} |\nabla \log f|^2\,\mathrm{d}\mu - \mathrm{H}(\mu|\gamma).$$

Then, the analogue of (5.4.3) is the identity [ELS20, Proposition 10],

$$\delta(f) \coloneqq \frac{1}{2}\int_0^1 \mathbb{E}[|v_1 - v_t|^2]\,\mathrm{d}t. \tag{5.4.15}$$

However, from this point on the analogies begin to break. In [ELS20] it is shown that

$$\mathbb{E}[\varphi(v_t)] \geqslant \int_0^t \mathbb{E}[\varphi(v_s)]^2\,\mathrm{d}s, \tag{5.4.16}$$

so differentiating (5.4.16) (more precisely a matrix version of this inequality) yields a differential inequality for $t \mapsto \mathbb{E}[\varphi(v_t)]$, which is a key point in some of the stability estimates of [ELS20]. On a high-level we can view (5.4.16) as the analogue of Lemma 5.14. However, while in (5.4.16) the expression $\mathbb{E}[\varphi(v_t)]$ appears on both sides of the inequality, in our setting we must deal with discrete derivatives which hinders such differential inequalities. Instead, we make crucial use of the observation in Lemma 5.16 that we can express the right-hand side in Lemma 5.14 in terms of relative entropy, which then leads to our stability result.

Chapter 6

Convergence of the entropic potentials

So now I gotta add you to my list of people to try and forget about
It could've been magic, nearly had ya, can you imagine? Nearly had ya
'Til it becomes another one of the things that I just can't talk about
I'm gonna have to keep you on my list of people to try and forget about.

Tame Impala
List of People (To Try and Forget About)

This chapter is based on the article [LR25b]. In the context of the entropic regularization of the optimal transport problem, we provide a bound on the rate of uniform convergence in compact sets for both entropic potentials and their gradients towards the Brenier potential and its gradient. Both results hold in the quadratic Euclidean setting for absolutely continuous measures satisfying some convexity assumptions.

We start by introducing the entropic regularization of optimal transport and its connections with its unregularized counterpart in Section 6.1; we also motivate the problem and introduce the main results of this chapter, Theorems 6.2 and 6.3. In Section 6.2, we will present all the necessary preliminaries on optimal transport and its entropic regularization and detail the assumptions for our main results. Then, in

Section 6.3, we prove both results. Finally, in Section 6.4, we prove Proposition 6.1, which corresponds to the Gaussian case.

6.1 Introduction

This section aims to motivate and introduce the main contributions of this chapter. In particular, we introduce the entropic regularization of the optimal transport problem and its convergence towards the unregularized problem in the small noise limit.

6.1.1 Optimal transport

This chapter concerns the quadratic Euclidean optimal transport problem between absolutely continuous measures, which has already been introduced in Section 1.3 of Chapter 1: given two Borel probability measures μ and ν on $\mathbb{R}^d$ that are absolutely continuous with respect to the d-dimensional Lebesgue measure, we define the quadratic optimal transport problem associated to them, in its Kantorovich formulation, as

$$\mathcal{C}_0(\mu,\nu) \coloneqq \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x-y|^2 \,\mathrm{d}\pi(x,y), \tag{6.1.1}$$

where $\Pi(\mu,\nu)$ denotes the set of transport plans between μ and ν. Theorem 1.22 in Chapter 1, the Brenier-McCann theorem, states that there exists a convex function $\varphi_0\colon \mathbb{R}^d \to \mathbb{R}$ such that the map $T_0\colon \mathbb{R}^d \to \mathbb{R}^d$ defined by $T_0 \coloneqq \nabla\varphi_0$ pushes forward the measure μ towards ν and induces the unique optimal coupling $\pi_0 \in \Pi(\mu,\nu)$ for the variational problem (6.1.1). We say that T_0 is the Brenier or the optimal transport map and φ_0 is a Brenier potential. Moreover, if we define $\psi_0 \coloneqq \varphi_0^*$ as the convex conjugate of φ_0, the pair $(f_0, g_0) \coloneqq \left(\frac{1}{2}|\cdot|^2 - \varphi_0, \frac{1}{2}|\cdot|^2 - \psi_0\right)$ solves the dual problem to (6.1.1):

$$\mathcal{C}_0(\mu,\nu) = \sup_{\substack{(f,g)\in L^1(\mu)\times L^1(\nu),\\ f\oplus g\leqslant \frac{1}{2}|\cdot-\cdot|^2}} \int_{\mathbb{R}^d} f\,\mathrm{d}\mu + \int_{\mathbb{R}^d} g\,\mathrm{d}\nu. \tag{6.1.2}$$

6.1.2 Entropic optimal transport

We can regularize Problem (6.1.1) by adding an entropy to the objective function: for $\varepsilon > 0$, we define the entropic regularization of the optimal transport problem (equivalent as well to the Schrödinger bridge [Léo14]) as

$$\mathcal{C}_\varepsilon(\mu,\nu) \coloneqq \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|x-y|^2 \,\mathrm{d}\pi(x,y) + \varepsilon\,\mathrm{H}(\pi|\mu\otimes\nu), \tag{6.1.3}$$

where $\mathrm{H}(\cdot|\mu\otimes\nu)$ denotes the relative entropy functional with respect to the measure $\mu\otimes\nu$. An excellent introductory reference to the subject is [Nut21].

The problem (6.1.3) is strictly convex, so there automatically exists a unique $\pi_\varepsilon \in \Pi(\mu,\nu)$ that solves (6.1.3). On the other hand, there exists a pair of functions $(f_\varepsilon, g_\varepsilon) \in$

$L^1(\mu) \times L^1(\nu)$, which is unique up to a constant, that solves the dual problem to (6.1.3):

$$\mathcal{C}_\varepsilon(\mu,\nu) = \sup_{(f,g)\in L^1(\mu)\times L^1(\nu)} \int_{\mathbb{R}^d} f \,\mathrm{d}\mu + \int_{\mathbb{R}^d} g \,\mathrm{d}\nu - \varepsilon \int_{\mathbb{R}^d\times\mathbb{R}^d} e^{\frac{f\oplus g - \frac{1}{2}|\cdot - \cdot|^2}{\varepsilon}} \,\mathrm{d}(\mu\otimes\nu) + \varepsilon. \quad (6.1.4)$$

We define for each $\varepsilon > 0$ the entropic potentials as $(\varphi_\varepsilon, \psi_\varepsilon) := \left(\frac{1}{2}|\cdot|^2 - f_\varepsilon, \frac{1}{2}|\cdot|^2 - g_\varepsilon\right)$, in analogy to their unregularized counterpart (φ_0, ψ_0).

6.1.3 The connection between both problems

It is known that computing φ_0 or $T_0 = \nabla\varphi_0$ is difficult, as one would solve the associated Monge-Ampère equation, which we introduced in Section 1.3.2 of Chapter 1. One of the advantages of the entropic problem is that the potentials $(\varphi_\varepsilon, \psi_\varepsilon)$ are very tractable, numerically speaking, thanks to the Sinkhorn algorithm [Cut13, ANWR17, PC19]. On the other hand, as the regularization parameter ε vanishes, the problem (6.1.3) converges in many senses to (6.1.1), a fact which allows us to approximate the optimal transport through the entropic regularization for small values of the parameter $\varepsilon > 0$.

More precisely, as $\varepsilon \to 0$, the entropic problem Γ-converges towards the unregularized one [Léo12, CDPS17], which yields the convergence of the value functions, $\mathcal{C}_\varepsilon(\mu,\nu) \to \mathcal{C}_0(\mu,\nu)$, and the convergence of the optimal couplings in the weak topology, $\pi_\varepsilon \to \pi_0$. Concerning the dual optimizers, it is known that $\varphi_\varepsilon \to \varphi_0$ (modulo subsequence) both in $L^1(\mu)$ and uniformly on compact sets [GT21, NW22] and $\nabla\varphi_\varepsilon \to \nabla\varphi_0$ (modulo subsequence) in $L^2(\mu)$ [CCGT23]. All of the above results use Γ-convergence or compactness arguments, so a natural question is whether it is possible to quantify the rate at which these convergences happen.

In the case of the convergence of the value functions, several contributions have been made in the continuous setting: the first-order expansion

$$\mathcal{C}_\varepsilon(\mu,\nu) = \mathcal{C}_0(\mu,\nu) - \frac{d}{2}\varepsilon\log\varepsilon + \frac{\varepsilon}{2}\left(\mathrm{H}(\mu) + \mathrm{H}(\nu)\right) + o(\varepsilon)$$

was proven to be true for the quadratic case and as a Γ-limit in dimension one in [ADPZ11, DLR13] and in higher dimensions in [EMR15]. The same expansion was established as a pointwise limit for strictly convex cost functions in [Pal24] (thus generalizing the quadratic case). The second-order expansion

$$\mathcal{C}_\varepsilon(\mu,\nu) = \mathcal{C}_0(\mu,\nu) - \frac{d}{2}\varepsilon\log\varepsilon + \frac{\varepsilon}{2}\left(\mathrm{H}(\mu) + \mathrm{H}(\nu)\right) + \frac{\varepsilon^2}{8}\mathrm{I}(\mu,\nu) + o(\varepsilon^2),$$

where $\mathrm{I}(\mu,\nu)$ denotes the integrated Fisher information on the Wasserstein geodesic between μ and ν, was proven for the Euclidean setting in [CRL+20, CT21]. For more general cost functions, the expansion

$$\mathcal{C}_\varepsilon(\mu,\nu) = \mathcal{C}_0(\mu,\nu) - \frac{d}{2}\varepsilon\log\varepsilon + O(\varepsilon)$$

was established in [CPT23, EN24]. In [MS23], the same formula was demonstrated. Additionally, the authors were able to identify the separate asymptotic behavior of $\mathrm{H}(\pi_\varepsilon|\mu\otimes\nu)$ and $\mathcal{C}_\varepsilon(\mu,\nu)-\mathrm{H}(\pi_\varepsilon|\mu\otimes\nu)$ as ε vanishes.

Concerning the convergence of the optimal plans π_ε, in [MS23] the authors were able to quantify the 2-Wasserstein distance between π_ε and π_0:

$$\mathrm{W}_2(\pi_\varepsilon,\pi_0)=\Theta(\sqrt{\varepsilon}).$$

Up to now, the convergence of potentials in the continuous setting was only quantified in the $L^2(\mu)$ norm in terms of the difference of their gradients: in [PNW21], for the quadratic setting under compactness and convexity assumptions, it was proven that

$$\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|^2_{L^2(\mu)}=O(\varepsilon^2).$$

The rate

$$\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|^2_{L^2(\mu)}=O(-\varepsilon\log\varepsilon)$$

was found in [CPT23] under slightly weaker assumptions. Finally, the rate

$$\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|^2_{L^2(\mu)}=O(\varepsilon)$$

was established in [MS23] for non-necessarily compactly supported measures.

We aim to exhibit a bound for the rate of convergence of uniform convergence on compact sets of the potentials φ_ε and their gradients $\nabla\varphi_\varepsilon$ as $\varepsilon\to 0$, when both measures μ and ν are absolutely continuous with respect to the Lebesgue measure, under the quadratic cost. The following observation is the starting point of our analysis: the entropic potentials φ_ε are convex functions; see Proposition 6.4 below in Section 6.2. As mentioned above, they converge (up to a subsequence) uniformly on compact sets to φ_0. It is a classical fact that convexity yields the convergence of the gradients $\nabla\varphi_\varepsilon$ to $\nabla\varphi_0$ uniformly on compact sets as well; see, for example, [Roc70, Theorem 25.7]. This justifies to search, for a fixed compact set $K\subset\mathbb{R}^d$, an asymptotic rate of convergence as ε goes to 0 for $\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|_{K,\infty}\coloneqq\sup_{x\in K}|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)|$.

If both μ and ν are Gaussian we can explicitly bound $\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|_{K,\infty}$. We present the result of this computation as an appetizer and defer its proof to Section 6.4.

Proposition 6.1. *Let $\mu=\mathcal{N}(0,A)$ and $\nu=\mathcal{N}(0,B)$ be two non-degenerate Gaussian measures with $A,B\succ 0$. For $R>0$, let $K\coloneqq\overline{\mathrm{B}}(0,R)\subset\mathbb{R}^d$ be the Euclidean closed ball of radius R. Then*

$$\|\nabla\varphi_\varepsilon-\nabla\varphi_0\|_{K,\infty}\leqslant R\varepsilon\frac{|A^{-1}|_{\mathrm{op}}}{2}\left(\varepsilon\frac{\left|(A^{\frac12}BA^{\frac12})^{-1}\right|_{\mathrm{op}}^{\frac12}}{4}+\varepsilon^3\frac{\left|(A^{\frac12}BA^{\frac12})^{-1}\right|_{\mathrm{op}}^{\frac32}}{16}+1\right). \tag{6.1.5}$$

Proposition 6.1 reveals that we can quantify uniform convergence on compact sets as $O(\varepsilon)$ as $\varepsilon\to 0$, with a dimension-free bound depending only on the size of the compact set K and the operator norms of the matrices A and B.

6.1.4 Main results

Our first result gives an estimate in the spirit of Proposition 6.1 for measures μ and ν satisfying convexity assumptions; see the end of Section 6.2 for their precise statement.

Theorem 6.2. *Let μ and ν be two probability measures on $\mathbb{R}^d$ that are absolutely continuous with respect to the d-dimensional Lebesgue measure and satisfy the assumptions (A1), (A2), and (A3). Then, for any $K \subset \mathbb{R}^d$ compact, there exists a computable constant $C_{\text{grad}} = C_{\text{grad}}(K, \mu, \nu, d) > 0$ such that for any $\varepsilon > 0$,*

$$\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|_{K,\infty} \leqslant C_{\text{grad}}\, \varepsilon^{\frac{1}{d+4}}.$$

We remark that Theorem 6.2 does not provide an optimal optimal bound; recall Proposition 6.1.

Our next result is a corollary of the previous theorem: we can also quantify the convergence in compact sets of the entropic potentials. For $K \subset \mathbb{R}^d$ compact we define, doing an abuse of notation, $\|\varphi_\varepsilon - \varphi_0\|_{K,\infty} := \sup_{x\in K}|\varphi_\varepsilon(x) - \varphi_0(x)|$.

Theorem 6.3. *Let μ and ν be two probability measures on $\mathbb{R}^d$ that are absolutely continuous with respect to the d-dimensional Lebesgue measure and satisfy the assumptions (A1), (A2), and (A3). In addition, suppose that the following normalization holds: for every $\varepsilon > 0$,*

$$\int_{\mathbb{R}^d} \varphi_\varepsilon \,\mathrm{d}\mu = \int_{\mathbb{R}^d} \varphi_0 \,\mathrm{d}\mu = 0. \tag{6.1.6}$$

Then, for any $K \subset \mathbb{R}^d$ compact and connected, there exists a computable constant $C_{\text{pot}} = C_{\text{pot}}(K, \mu, \nu, d) > 0$ such that for any $\varepsilon > 0$,

$$\|\varphi_\varepsilon - \varphi_0\|_{K,\infty} \leqslant C_{\text{pot}} \left(\varepsilon^{\frac{1}{d+4}} + \varepsilon\right).$$

To the best of our knowledge, this is the first work that addresses this problem for the entropic regularization in the continuous setting. Previously in the literature, this question has been answered both in the discrete [CSM94] and semi-discrete [Del22, SGK24] settings for the entropic regularization. On the other hand, in [GSN24], a rate was found for the quadratic-regularized optimal transport problem in dimension one.

6.2 Preliminaries and assumptions

In this section, we review some properties of both the Brenier and entropic potentials that complement the basic statements previously introduced. Finally, we state our main assumptions and discuss their consequences and some sufficient conditions for them to hold.

6.2.1 Further properties of the potentials

We recall the Brenier potentials (φ_0, ψ_0) and their entropic counterparts $(\varphi_\varepsilon, \psi_\varepsilon)$. Note that they are not unique, since for any constant $\alpha \in \mathbb{R}$, then $(\varphi_0 + \alpha, \psi_0 - \alpha)$ and $(\varphi_\varepsilon + \alpha, \psi_\varepsilon - \alpha)$ are two new pairs of Brenier and entropic potentials, respectively. Therefore, if we need to enforce the uniqueness of the potentials, an additional normalization condition has to be imposed, for example, (6.1.6). Nevertheless, we remark that their gradients are uniquely determined.

Recall that for $\varepsilon > 0$ we have that $(f_\varepsilon, g_\varepsilon) = \left(\frac{1}{2}|\cdot| - \varphi_\varepsilon, \frac{1}{2}|\cdot| - \psi_\varepsilon\right)$. The pair $(f_\varepsilon, g_\varepsilon)$ is intimately linked with π_ε, the optimal coupling for (6.1.3):

$$\frac{\mathrm{d}\pi_\varepsilon}{\mathrm{d}(\mu \otimes \nu)}(x, y) = \exp\left(\frac{f_\varepsilon(x) + g_\varepsilon(y) - \frac{1}{2}|x-y|^2}{\varepsilon}\right). \tag{6.2.1}$$

A direct consequence of (6.2.1) is that the pair $(f_\varepsilon, g_\varepsilon)$ satisfies the so-called Schrödinger system: for any $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$f_\varepsilon(x) = -\varepsilon \log\left(\int_{\mathbb{R}^d} e^{\frac{1}{\varepsilon}\left[g_\varepsilon(y) - \frac{1}{2}|x-y|^2\right]}\,\mathrm{d}\nu(y)\right), \tag{6.2.2}$$

$$g_\varepsilon(y) = -\varepsilon \log\left(\int_{\mathbb{R}^d} e^{\frac{1}{\varepsilon}\left[f_\varepsilon(x) - \frac{1}{2}|y-x|^2\right]}\,\mathrm{d}\mu(x)\right). \tag{6.2.3}$$

In particular, the Schrödinger system allows us to prove fine properties of the pair $(\varphi_\varepsilon, \psi_\varepsilon)$ such as convexity.

Proposition 6.4. *For any $\varepsilon > 0$ the potentials φ_ε and ψ_ε are convex functions.*

Proof. Let $\lambda \in (0,1)$ and $x_1, x_2 \in \mathbb{R}^d$. Then by (6.2.2) we get

$$\begin{aligned}
\lambda\varphi_\varepsilon(x_1) + (1-\lambda)\varphi_\varepsilon(x_2) &= \varepsilon \log\left(\left[\int_{\mathbb{R}^d} e^{\frac{1}{\varepsilon}[x_1\cdot y - \psi_\varepsilon(y)]}\,\mathrm{d}\nu(y)\right]^\lambda \left[\int_{\mathbb{R}^d} e^{\frac{1}{\varepsilon}[x_2\cdot y - \psi_\varepsilon(y)]}\,\mathrm{d}\nu(y)\right]^{1-\lambda}\right) \\
&\geqslant \varepsilon \log\left(\int_{\mathbb{R}^d} e^{\frac{1}{\varepsilon}[(\lambda x_1 + (1-\lambda)x_2)\cdot y - \psi_\varepsilon(y)]}\,\mathrm{d}\nu(y)\right) \\
&= \varphi_\varepsilon(\lambda x_1 + (1-\lambda)x_2),
\end{aligned}$$

where we used Hölder's inequality for $p = 1/\lambda$ and $q = 1/(1-\lambda)$. The convexity of ψ_ε follows similarly from (6.2.3). □

6.2.2 Assumptions

Our main results are stated in terms of two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ that are absolutely continuous with respect to the d-dimensional Lebesgue measure and satisfy the following assumptions.

(A1) The measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ have the form $\mathrm{d}\mu(x) = e^{-V(x)}\,\mathrm{d}x$ and $\mathrm{d}\nu(y) = e^{-W(y)}\,\mathrm{d}y$, where $V, W\colon \mathbb{R}^d \to \mathbb{R}$ are smooth functions and there exist $\alpha, \beta > 0$ such that

$$\forall x \in \mathbb{R}^d, \quad \nabla^2 V(x) \preccurlyeq \alpha I_d \tag{6.2.4}$$

and

$$\forall y \in \mathbb{R}^d, \quad \nabla^2 W(y) \succcurlyeq \beta I_d, \tag{6.2.5}$$

where I_d is the identity matrix of dimension d and $\preccurlyeq$ denotes the Löwner order on the set of positive semidefinite matrices.

(A2) The measure μ satisfies a Poincaré inequality: there exists $C_{\mathrm{P}}(\mu) > 0$ such that for any $h\colon \mathbb{R}^d \to \mathbb{R}$ smooth with $\int_{\mathbb{R}^d} h \,\mathrm{d}\mu = 0$,

$$\|h\|^2_{L^2(\mu)} \leqslant C_{\mathrm{P}}(\mu)\|\nabla h\|^2_{L^2(\mu)}.$$

(A3) The measure μ has finite differential entropy:

$$-\infty < \mathrm{H}(\mu) \coloneqq -\int_{\mathbb{R}^d} V(x)e^{-V(x)}\,\mathrm{d}x < +\infty.$$

The main consequence of (A1) is that we obtain quantitative control on both $\nabla^2\varphi_0$ and $\nabla^2\varphi_\varepsilon$, for every $\varepsilon > 0$; that is, both $\nabla\varphi_0$ and $\nabla\varphi_\varepsilon$ are Lipschitz and we have explicit control on the value of their Lipschitz constants. Theorem 1.26, the Caffarelli contraction theorem, provides global Lipschitz regularity for the Brenier map $\nabla\varphi_0$, which pushes μ towards ν. The version below is a generalized version.

Theorem 6.5 (Caffarelli)**.** *Suppose that both $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ satisfy (A1). Then the Brenier map $\nabla\varphi_0$ is globally Lipschitz. Moreover, the following estimate holds:*

$$\forall x \in \mathbb{R}^d, \quad 0 \preccurlyeq \nabla^2\varphi_0(x) \preccurlyeq \sqrt{\frac{\alpha}{\beta}} I_d.$$

For the entropic counterpart of the Brenier map, namely $\nabla\varphi_\varepsilon$, we can also exhibit bounds on its derivative. These bounds were used in [FGP20, CP23] to give alternative proofs of Theorem 6.5 based on the entropic regularization. Here, we use the ones proven in [CP23, Theorem 4].

Theorem 6.6. *Suppose that both $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ satisfy (A1). Then the entropic Brenier map $\nabla\varphi_\varepsilon$ is globally Lipschitz. Moreover, the following estimate holds:*

$$\forall x \in \mathbb{R}^d, \quad 0 \preccurlyeq \nabla^2\varphi_\varepsilon(x) \preccurlyeq \frac{1}{2}\left(\sqrt{\frac{4\alpha}{\beta} + \varepsilon^2\alpha^2} - \varepsilon\alpha\right) I_d.$$

As we will see in the following remark, the functional inequality provided by (A2) also entails regularity properties for the measure μ itself.

Remark 6.7. A Poincaré inequality implies that the measure has finite moments of all orders [BGL14, Proposition 4.4.2], so under (A2), the measure μ will have this property.

Remark 6.8. Log-concavity is a sufficient condition that entails both (A2) and (A3). Recall Definition 2.67 in Chapter 2: we say that an absolutely continuous probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ is log-concave if its density $f\colon \mathbb{R}^d \to \mathbb{R}_+$ is of the form $f = e^{-V}$ for some $V\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ convex. Since we assumed V to be smooth, this is equivalent to

$$\forall x \in \mathbb{R}^d, \quad \nabla^2 V(x) \succcurlyeq 0.$$

First, log-concavity yields the validity of a Poincaré inequality; recall Theorem 2.74 in Chapter 2. Second, let us see that log-concavity implies (A3): indeed,

$$\mathrm{H}(\mu) = -\int_{\mathbb{R}^d} V(x)e^{-V(x)}\,\mathrm{d}x = -\int_{\mathbb{R}^d} V(x)\,\mathrm{d}\mu(x) \leqslant -V(0) - \nabla V(0)\cdot \mathbb{E}_{X\sim\mu}[X] < +\infty,$$

where we used the convexity of V. For the lower bound, we recall that μ has finite moments of all orders (see Remark 6.7), so

$$-\infty < \mathrm{H}(\mathcal{N}) \leqslant \mathrm{H}(\mu),$$

where $\mathcal{N}$ is the d-dimensional Gaussian measure with the same mean and covariance as μ.

Remark 6.9. Note that under (A1), the measure ν is log-concave; see (6.2.5). In particular, it has finite differential entropy and finite moments of all orders as well.

Under the assumption (A1), we saw that Theorem 6.5 ensures quantitatively the Lipschitz regularity of the Brenier map. A direct consequence of this fact is the following: we can control the difference of the gradients of the potentials in the $L^2(\mu)$ norm by the difference between both costs (see, for example, [MS23, Lemma 3.8] or [CPT23, Proposition 4.5], which are based on an unpublished result of Ambrosio which was reported in [Gig11a]).

Proposition 6.10. *Suppose that the Brenier map is L-Lipschitz. Then*

$$\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|^2_{L^2(\mu)} \leqslant 2L\langle \pi_\varepsilon - \pi_0, \frac{1}{2}|\cdot - \cdot|^2\rangle = 2L\left(\mathcal{C}_\varepsilon(\mu,\nu) - \varepsilon\,\mathrm{H}(\pi_\varepsilon|\mu\otimes\nu) - \mathcal{C}_0(\mu,\nu)\right).$$

That is, if we are able to control $\langle \pi_\varepsilon - \pi_0, \frac{1}{2}|\cdot - \cdot|^2\rangle$, then we can control the L^2 norm of the difference. In [MS23, Theorem 3.7], we can find this control, which can be applied as a consequence of both (A2) and (A3), recall Remarks 6.7 and 6.9.

Theorem 6.11. *([MS23, Theorem 3.7]). Suppose that both μ and ν have finite moments of order $2+\delta$, for some $\delta > 0$, and that both have finite differential entropy. Then*

$$\langle \pi_\varepsilon - \pi_0, \frac{1}{2}|\cdot - \cdot|^2\rangle \leqslant C\varepsilon.$$

That is, under our three assumptions, there exists a constant $C_1 > 0$ depending on d, μ and ν such that

$$\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|^2_{L^2(\mu)} \leqslant C_1\varepsilon. \tag{6.2.6}$$

Now let us assume that the following normalization holds:

$$\forall \varepsilon > 0, \quad \int_{\mathbb{R}^d} \varphi_\varepsilon\,\mathrm{d}\mu = \int_{\mathbb{R}^d} \varphi_0\,\mathrm{d}\mu = 0. \tag{6.2.7}$$

Note that if are working with the gradients of the potentials, they are uniquely determined, so we may assume without loss of generality that (6.2.7) holds. In particular, under this normalization, we can use (A2) to control the L^2 norm of the difference of the potentials:

$$\|\varphi_\varepsilon - \varphi_0\|^2_{L^2(\mu)} \leqslant C_{\mathrm{P}}(\mu)\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|^2_{L^2(\mu)}. \tag{6.2.8}$$

That is, again under all our assumptions, we have that there exists a constant $C_2 > 0$ depending on d, μ and ν such that

$$\|\varphi_\varepsilon - \varphi_0\|^2_{L^2(\mu)} \leqslant C_2\varepsilon. \tag{6.2.9}$$

6.3 Proof of the main results

This section aims to prove Theorem 6.2 and then to obtain as a corollary Theorem 6.3. The starting point of our proof will be the Gagliardo-Nirenberg inequality, which allows us to control the p norm of the derivatives of order j of a function by the r norm of its derivatives of order k and the q norm of the function itself, where the parameters i, j, p, q, and r satisfy some relations. We state the inequality in the following version, found in [SMR18, Theorem 3.1], since it allows the critical values $p = +\infty$, $r = +\infty$. Here the ambient space is $\mathbb{R}^d$ and we recall the notation $\|\cdot\|_s = \|\cdot\|_{L^s(\mathrm{Leb})}$.

Theorem 6.12 (Gagliardo-Nirenberg)**.** *Let $j, k \in \mathbb{N}$ such that $1 \leqslant j < k$. Let $\theta \in [j/k, 1]$, $q \in [1, +\infty]$, $p \in (1, +\infty]$, $r \in [1, +\infty]$ such that*

$$\frac{1}{p} = \frac{j}{d} + \theta\left(\frac{1}{r} - \frac{k}{d}\right) + \frac{1-\theta}{q} \tag{6.3.1}$$

and such that

$$\forall 0 \leqslant i \leqslant k - j - 1, \quad r^{(i)} \neq d, \tag{6.3.2}$$

where $r^{(0)} := r$ and $r^{(i)} := \frac{nr^{(i-1)}}{n-r^{(i-1)}}$, for $i \geqslant 1$. Then there exists a constant $C_{\mathrm{GN}} = C_{\mathrm{GN}}(j, k, d, \theta, p, q, r) > 0$ such that for any $u\colon \mathbb{R}^d \to \mathbb{R}$ sufficiently smooth and integrable

$$\|\nabla^j u\|_p \leqslant C_{\mathrm{GN}}\|\nabla^k u\|_r^\theta \|u\|_q^{1-\theta}. \tag{6.3.3}$$

We are ready to start the proof of Theorem 6.2.

Proof (of Theorem 6.2): The proof is divided into three consecutive steps.

Step 1: The first step will be to choose appropriate parameters to apply the Gagliardo-Nirenberg inequality (6.3.3). From Theorem 6.12 and choosing

$$j = 1, k = 2, p = +\infty, q = 2, r = +\infty,$$

the value of θ is be determined by (6.3.1), so we get

$$\theta = \frac{d+2}{d+4} \in [1/2, 1].$$

On the other hand, we notice that the condition (6.3.2) is trivially verified since $r^{(0)} = +\infty$. Hence the inequality (6.3.3) takes the following form:

$$\|\nabla u\|_\infty \leqslant C_{\mathrm{GN}} \big\|\nabla^2 u\big\|_\infty^{\frac{d+2}{d+4}} \|u\|_2^{\frac{2}{d+4}}. \tag{6.3.4}$$

Now let $K \subset \mathbb{R}^d$ be a compact set, take $R > 0$ such that $K \subset \mathrm{B}(0, R)$, where $\mathrm{B}(0, R)$ denotes the open ball with center 0 and radius R, and such that $\mu(\mathrm{B}(0, R)) > 4/5$, and let $w \colon \mathbb{R}^d \to \mathbb{R}_+$ be a compactly supported smooth function satisfying the following properties:

- $\forall x \in \mathbb{R}^d$, $0 \leqslant w(x) \leqslant 1$;
- $\forall x \in K$, $w(x) = 1$; and
- $\operatorname{supp}(w) \subset \mathrm{B}(0, R)$.

Let $\varepsilon > 0$. Then

$$\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|_{K,\infty} \leqslant \|\nabla[(\varphi_\varepsilon - \varphi_0)w]\|_\infty.$$

If we apply (6.3.4) to $u \coloneqq (\varphi_\varepsilon - \varphi_0)w$, then

$$\|\nabla[(\varphi_\varepsilon - \varphi_0)w]\|_\infty \leqslant C_{\mathrm{GN}} \big\|\nabla^2[(\varphi_\varepsilon - \varphi_0)w]\big\|_\infty^{\frac{d+2}{d+4}} \|(\varphi_\varepsilon - \varphi_0)w\|_2^{\frac{2}{d+4}}. \tag{6.3.5}$$

Step 2: Now we will bound the first term on the right-hand side of (6.3.5):

$$\big\|\nabla^2[(\varphi_\varepsilon - \varphi_0)w]\big\|_\infty^{\frac{d+2}{d+4}}. \tag{6.3.6}$$

Note that

$$\begin{aligned}
\big\|\nabla^2[(\varphi_\varepsilon - \varphi_0)w]\big\|_\infty &\leqslant \big\|w\nabla^2(\varphi_\varepsilon - \varphi_0)\big\|_\infty + \big\|(\varphi_\varepsilon - \varphi_0)\nabla^2 w\big\|_\infty + 2\|\nabla(\varphi_\varepsilon - \varphi_0)\nabla w^\intercal\|_\infty \\
&\leqslant \big\|\nabla^2(\varphi_\varepsilon - \varphi_0)\big\|_\infty + \|\varphi_\varepsilon - \varphi_0\|_{\mathrm{B}(0,R),\infty} \times \big\|\nabla^2 w\big\|_\infty \\
&\quad + 2\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|_{\mathrm{B}(0,R),\infty} \times \|\nabla w\|_\infty.
\end{aligned}$$

For the first term on the right-hand side, (A1) plays a key role: note that for every $\varepsilon > 0$, the upper bound in Theorem 6.6 is such that

$$\frac{1}{2}\left(\sqrt{\frac{4\alpha}{\beta} + \varepsilon^2\alpha^2} - \varepsilon\alpha\right) \leqslant \sqrt{\frac{\alpha}{\beta}}.$$

In particular, using also the estimate granted by Theorem 6.5, we get

$$\big\|\nabla^2\varphi_\varepsilon - \nabla^2\varphi_0\big\|_\infty \leqslant \big\|\nabla^2\varphi_\varepsilon\big\|_\infty + \big\|\nabla^2\varphi_0\big\|_\infty \leqslant 2\sqrt{\frac{\alpha}{\beta}}.$$

In order to bound the two remaining terms on the right-hand side, let us state and prove the following auxiliary lemma.

Lemma 6.13. *In the above context, there exists $M > 0$ such that for any $\varepsilon > 0$, there exists $x_\varepsilon^* \in \mathrm{B}(0, R)$ satisfying both*

$$|\nabla\varphi_\varepsilon(x_\varepsilon^*) - \nabla\varphi_0(x_\varepsilon^*)| < M \tag{6.3.7}$$

and

$$|\varphi_\varepsilon(x_\varepsilon^*) - \varphi_0(x_\varepsilon^*)| < M. \tag{6.3.8}$$

Proof. To start, let us prove that

$$\sup_{\varepsilon>0}\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon-\nabla\varphi_0|^2\,\mathrm{d}\mu<+\infty,$$

so let $\varepsilon>0$. First, note that for any $x\in\mathbb{R}^d$,

$$\nabla\varphi_\varepsilon(x)=\int_{\mathbb{R}^d}y\,\mathrm{d}\pi_\varepsilon^x(y),$$

where for $x\in\mathbb{R}^d$, we define $\mathrm{d}\pi_\varepsilon^x(y):=\frac{\mathrm{d}\pi_\varepsilon}{\mathrm{d}(\mu\otimes\nu)}(x,y)\,\mathrm{d}\nu(y)$. Therefore, by Jensen's inequality and the fact that $\pi_\varepsilon\in\Pi(\mu,\nu)$, we have that

$$\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon(x)|^2\,\mathrm{d}\mu(x)\leqslant\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|y|^2\,\mathrm{d}\pi_\varepsilon^x(y)\,\mathrm{d}\mu(x)=\int_{\mathbb{R}^d\times\mathbb{R}^d}|y|^2\,\mathrm{d}\pi_\varepsilon(x,y)=\int_{\mathbb{R}^d}|y|^2\,\mathrm{d}\nu(y).$$

On the other hand, note that

$$\int_{\mathbb{R}^d}|\nabla\varphi_0(x)|^2\,\mathrm{d}\mu(x)=\int_{\mathbb{R}^d}|y|^2\,\mathrm{d}\nu(y)$$

since the Brenier map $x\mapsto\nabla\varphi_0(x)$ pushes forward μ towards ν. Then

$$\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon-\nabla\varphi_0|^2\,\mathrm{d}\mu\leqslant 2\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon|^2\,\mathrm{d}\mu+2\int_{\mathbb{R}^d}|\nabla\varphi_0|^2\,\mathrm{d}\mu\leqslant 4\int_{\mathbb{R}^d}|y|^2\,\mathrm{d}\nu(y),$$

so $\sup_{\varepsilon>0}\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon-\nabla\varphi_0|^2\,\mathrm{d}\mu<+\infty$ since ν has finite moments of all orders; recall Remarks 6.7 and 6.9.

Let $M_1>0$ to be fixed. By Chebyshev's inequality, we have

$$\begin{aligned}\mu(\{x\in\mathbb{R}^d:|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)|\geqslant M_1\})&\leqslant\frac{1}{M_1^2}\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon-\nabla\varphi_0|^2\,\mathrm{d}\mu\\&\leqslant\frac{4}{M_1^2}\int_{\mathbb{R}^d}|y|^2\,\mathrm{d}\nu(y),\end{aligned}$$

so we may fix M_1 big enough such that $\frac{4}{M_1^2}\int_{\mathbb{R}^d}|y|^2\,\mathrm{d}\nu(y)\leqslant\frac{1}{5}$. Hence,

$$\mu(\{x\in\mathbb{R}^d:|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)|<M_1\})>\frac{4}{5},\tag{6.3.9}$$

so

$$\mu(\{x\in\mathrm{B}(0,R):|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)|<M_1\})>\frac{3}{5}.\tag{6.3.10}$$

Without loss of generality, we may assume that the normalization (6.2.7) holds so that (6.2.8) is verified. Similarly to the previous argument, we deduce that there exists $M_2>0$ such that

$$\mu(\{x\in\mathrm{B}(0,R):|\varphi_\varepsilon(x)-\varphi_0(x)|<M_2\})>\frac{3}{5}.\tag{6.3.11}$$

Let $M:=\max\{M_1,M_2\}$. If we put together (6.3.10) and (6.3.11), we deduce that

$$\mu(\{x\in\mathrm{B}(0,R):|\varphi_\varepsilon(x)-\varphi_0(x)|<M,|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)|<M\})>0,$$

which yields the existence of $x_\varepsilon^*\in\mathrm{B}(0,R)$ such that both (6.3.7) and (6.3.8) hold. $\square$

Recall that we wanted to bound

$$\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|_{\mathrm{B}(0,R),\infty} \tag{6.3.12}$$

and

$$\|\varphi_\varepsilon - \varphi_0\|_{\mathrm{B}(0,R),\infty}. \tag{6.3.13}$$

By Lemma 6.13, there exists $M > 0$ and $x_\varepsilon^* \in \mathrm{B}(0,R)$ verifying both (6.3.7) and (6.3.8). For (6.3.12), note that

$$\begin{aligned}\sup_{x\in\mathrm{B}(0,R)} |\nabla\varphi_\varepsilon(x) - \nabla\varphi_0(x)| &\leqslant \sup_{x\in\mathrm{B}(0,R)} |\nabla\varphi_\varepsilon(x) - \nabla\varphi_\varepsilon(x_\varepsilon^*)| + \sup_{x\in\mathrm{B}(0,R)} |\nabla\varphi_0(x) - \nabla\varphi_0(x_\varepsilon^*)| \\ &\quad + |\nabla\varphi_\varepsilon(x_\varepsilon^*) - \nabla\varphi_0(x_\varepsilon^*)| \\ &\leqslant 4\sqrt{\frac{\alpha}{\beta}}R + M,\end{aligned}$$

where we used the second-order estimates holding under (A1) and (6.3.7). For the remaining term (6.3.13), observe that a Taylor expansion combined with the aforementioned second-order estimates yields

$$\begin{aligned}\sup_{x\in\mathrm{B}(0,R)} |\varphi_\varepsilon(x) - \varphi_0(x)| &\leqslant |\varphi_\varepsilon(x_\varepsilon^*) - \varphi_0(x_\varepsilon^*)| + 2R|\nabla\varphi_\varepsilon(x_\varepsilon^*) - \nabla\varphi_0(x_\varepsilon^*)| + R^2\sqrt{\frac{\alpha}{\beta}} \\ &\leqslant M + 2RM + R^2\sqrt{\frac{\alpha}{\beta}}.\end{aligned}$$

In conclusion, if we combine the three bounds, we know that there exists a constant $C > 0$ depending on K, μ, and ν such that for every $\varepsilon > 0$,

$$\left\|\nabla^2[(\varphi_\varepsilon - \varphi_0)w]\right\|_\infty^{\frac{d+2}{d+4}} \leqslant C. \tag{6.3.14}$$

Step 3: Now we aim to bound the remaining term in the right-hand side of (6.3.5) in an asymptotic way in order to obtain a quantity that converges to 0 as ε vanishes; that is, we want to bound

$$\|(\varphi_\varepsilon - \varphi_0)w\|_2^{\frac{2}{d+4}}.$$

To control this term, the bound (6.2.9) will be crucial. Indeed,

$$\begin{aligned}\|(\varphi_\varepsilon - \varphi_0)w\|_2^2 &= \int_{\mathbb{R}^d} |\varphi_\varepsilon(x) - \varphi_0(x)|^2 w(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} |\varphi_\varepsilon(x) - \varphi_0(x)|^2 e^{V(x)} w(x) e^{-V(x)}\,\mathrm{d}x \\ &\leqslant \int_{\mathbb{R}^d} |\varphi_\varepsilon(x) - \varphi_0(x)|^2 e^{V(x)}\,\mathrm{d}\mu(x) \\ &\leqslant \left\|e^V\right\|_\infty \|\varphi_\varepsilon - \varphi_0\|_{L^2(\mu)}^2,\end{aligned}$$

so if we combine this with (6.2.9), we get

$$\|(\varphi_\varepsilon - \varphi_0)w\|_2^{\frac{2}{d+4}} \leqslant C'\varepsilon^{\frac{1}{d+4}}, \tag{6.3.15}$$

where $C' > 0$ is a constant depending on d, K, μ and ν. Finally, if we put together (6.3.5), (6.3.14), and (6.3.15), we obtain the desired inequality.

□

We end by proving Theorem 6.3.

Proof (of Theorem 6.3): Let $K \subset \mathbb{R}^d$ be a compact and connected set. Since we have assumed that the potentials are normalized as $\int_{\mathbb{R}^d} \varphi_\varepsilon \,\mathrm{d}\mu = \int_{\mathbb{R}^d} \varphi_\varepsilon \,\mathrm{d}\mu = 0$, we can use [Eva10, Theorem 1, p. 290] and the bound (6.3.15) to get

$$\|\varphi_\varepsilon - \varphi_0\|_{K,\infty} \leqslant C'' \left(\|\nabla\varphi_\varepsilon - \nabla\varphi_0\|_{K,\infty} + \varepsilon \right),$$

where $C'' > 0$ is a constant depending on d, K, μ and ν. The conclusion follows from the estimate given by Theorem 6.2. $\square$

6.4 The Gaussian case

In this section, we provide the proof of Proposition 6.1.

Proof (of Proposition 6.1): From Proposition 1.3 in Chapter 1, we know that we can compute explicitly $\nabla\varphi_0$:

$$\forall x \in \mathbb{R}^d, \quad \nabla\varphi_0(x) = A^{-\frac{1}{2}}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}A^{-\frac{1}{2}}x.$$

For $\nabla\varphi_\varepsilon$, we have a similar closed-form expression [Gel90, JMPC20, dBL20, MGM22]:

$$\forall x \in \mathbb{R}^d, \quad \nabla\varphi_\varepsilon(x) = \left(A^{-\frac{1}{2}} \left(A^{\frac{1}{2}}BA^{\frac{1}{2}} + \frac{\varepsilon^2}{4}I_d \right)^{\frac{1}{2}} A^{-\frac{1}{2}} - \frac{\varepsilon}{2}A^{-1} \right) x.$$

Now fix $R > 0$ and let $K = \overline{\mathrm{B}}(0, R) \subset \mathbb{R}^d$, so that

$$\sup_{x\in K} |\nabla\varphi_\varepsilon(x) - \nabla\varphi_0(x)|$$

$$= R \left| \left(A^{-\frac{1}{2}} \left(A^{\frac{1}{2}}BA^{\frac{1}{2}} + \frac{\varepsilon^2}{4}I_d \right)^{\frac{1}{2}} A^{-\frac{1}{2}} - \frac{\varepsilon}{2}A^{-1} \right) - A^{-\frac{1}{2}}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}A^{-\frac{1}{2}} \right|_{\mathrm{op}}.$$

Let us note that we may expand the matrix $\left(A^{\frac{1}{2}}BA^{\frac{1}{2}} + \frac{\varepsilon^2}{4}I_d \right)^{\frac{1}{2}}$, using a Taylor expansion of order one for the matrix square root function around the point $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ (see [DMN18, Theorem 1.1]):

$$\left(A^{\frac{1}{2}}BA^{\frac{1}{2}} + \frac{\varepsilon^2}{4}I_d \right)^{\frac{1}{2}} = (A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}} + \frac{\varepsilon^2}{4} \int_0^{+\infty} e^{-2t(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}} \,\mathrm{d}t + R(A, B, \varepsilon),$$

where $R(A, B, \varepsilon)$ is a matrix such that

$$|R(A, B, \varepsilon)|_{\mathrm{op}} \leqslant \frac{\varepsilon^4}{32} \lambda_{\min}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{-\frac{3}{2}} = \frac{\varepsilon^4}{32} \left| (A^{\frac{1}{2}}BA^{\frac{1}{2}})^{-1} \right|_{\mathrm{op}}^{\frac{3}{2}},$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a positive-definite matrix. On the other hand, note that

$$\left|\int_0^{+\infty} e^{-2t(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}}\,\mathrm{d}t\right|_{\mathrm{op}} \leqslant \int_0^{+\infty} e^{-2t\lambda_{\min}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}}\,\mathrm{d}t$$
$$= \frac{1}{2\lambda_{\min}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}} = \frac{1}{2}\left|(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{-1}\right|_{\mathrm{op}}^{\frac{1}{2}}.$$

In consequence, we have that

$$\begin{aligned}
&\sup_{x\in K}|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)| \\
&= R\left|\left(A^{-\frac{1}{2}}\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}+\frac{\varepsilon^2}{4}I_d\right)^{\frac{1}{2}}A^{-\frac{1}{2}}-\frac{\varepsilon}{2}A^{-1}\right)-A^{-\frac{1}{2}}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}A^{-\frac{1}{2}}\right|_{\mathrm{op}} \\
&\leqslant R\left|A^{-\frac{1}{2}}\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}+\frac{\varepsilon^2}{4}I_d\right)^{\frac{1}{2}}A^{-\frac{1}{2}}-A^{-\frac{1}{2}}(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{\frac{1}{2}}A^{-\frac{1}{2}}\right|_{\mathrm{op}}+R\left|\frac{\varepsilon}{2}A^{-1}\right|_{\mathrm{op}} \\
&\leqslant \varepsilon R\frac{|A^{-1}|_{\mathrm{op}}}{2}\left(\varepsilon\frac{\left|(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{-1}\right|_{\mathrm{op}}^{\frac{1}{2}}}{4}+\varepsilon^3\frac{\left|(A^{\frac{1}{2}}BA^{\frac{1}{2}})^{-1}\right|_{\mathrm{op}}^{\frac{3}{2}}}{16}+1\right),
\end{aligned}$$

so $\sup_{x\in K}|\nabla\varphi_\varepsilon(x)-\nabla\varphi_0(x)| = O(\varepsilon)$ as $\varepsilon\to 0$, with a constant which depends only on R and the matrices A and B.

□

Bibliography

C'est que les marges d'un livre ne sont jamais nettes ni rigoureusement tranchées : par-delà le titre, les premières lignes et le point final, par-delà sa configuration interne et la forme qui l'autonomise, il est pris dans un système de renvois à d'autres livres, d'autres textes, d'autres phrases : nœud dans un réseau.

Michel Foucault
L'Archéologie du savoir

[ADPZ11] Stefan Adams, Nicolas Dirr, Mark A. Peletier, and Johannes Zimmer. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Comm. Math. Phys.*, 307(3):791–815, 2011.

[AS94] Shigeki Aida and Ichiro Shigekawa. Logarithmic Sobolev inequalities and spectral gaps: perturbation theory. *J. Funct. Anal.*, 126(2):448–475, 1994.

[ANWR17] Jason Altschuler, Jonathan Niles-Weed, and Philippe Rigollet. Near-linear time approximation algorithms for optimal transport via sinkhorn iteration. *Advances in neural information processing systems*, 30, 2017.

[AM80] Dan Amir and Vitali D. Milman. Unconditional and symmetric sets in n-dimensional normed spaces. *Israel J. Math.*, 37(1-2):3–20, 1980.

[ABC+00] Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société

Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[AMM22] Heshan Aravinda, Arnaud Marsiglietti, and James Melbourne. Concentration inequalities for ultra log-concave distributions. *Studia Math.*, 265(1):111–120, 2022.

[AAGM15] Shiri Artstein-Avidan, Apostolos Giannopoulos, and Vitali D. Milman. *Asymptotic geometric analysis. Part I*, volume 202 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

[AAGM21] Shiri Artstein-Avidan, Apostolos Giannopoulos, and Vitali D. Milman. *Asymptotic geometric analysis. Part II*, volume 261 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2021.

[ALRS24] Shrey Aryan, Pablo López-Rivera, and Yair Shenfeld. The stability of Wu's logarithmic Sobolev inequality via the Poisson-Föllmer process. *arXiv preprint arXiv:2410.06117*, 2024.

[Bak94] Dominique Bakry. Une suite d'inégalités remarquables pour les opérateurs ultrasphériques. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(2):161–164, 1994.

[BE85] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[BK08] Franck Barthe and Alexander V. Kolesnikov. Mass transport and variants of the logarithmic Sobolev inequality. *J. Geom. Anal.*, 18(4):921–979, 2008.

[BB00] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[BGN22] Espen Bernton, Promit Ghosal, and Marcel Nutz. Entropic optimal transport: geometry and large deviations. *Duke Math. J.*, 171(16):3363–3400, 2022.

[Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[Bob96] Sergey G. Bobkov. A functional form of the isoperimetric inequality for the Gaussian measure. *J. Funct. Anal.*, 135(1):39–49, 1996.

[Bob97] Sergey G. Bobkov. An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. *Ann. Probab.*, 25(1):206–214, 1997.

[Bob99] Sergey G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4):1903–1921, 1999.

[BGL01] Sergey G. Bobkov, Ivan Gentil, and Michel Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 80(7):669–696, 2001.

[BG99] Sergey G. Bobkov and Friedrich Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[BL97] Sergey G. Bobkov and Michel Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3):383–400, 1997.

[BL98] Sergey G. Bobkov and Michel Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.

[BT06] Sergey G. Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.*, 19(2):289–336, 2006.

[Bor75a] Christer Borell. The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30(2):207–216, 1975.

[Bor75b] Christer Borell. Convex set functions in d-space. *Period. Math. Hungar.*, 6(2):111–136, 1975.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities.* Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[Bou86] Jean Bourgain. On high-dimensional maximal functions associated to convex bodies. *Amer. J. Math.*, 108(6):1467–1476, 1986.

[Bou87] Jean Bourgain. Geometry of Banach spaces and harmonic analysis. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*, pages 871–878. Amer. Math. Soc., Providence, RI, 1987.

[BP16] Solesne Bourguin and Giovanni Peccati. The Malliavin-Stein method on the Poisson space. In *Stochastic analysis for Poisson point processes*, volume 7 of *Bocconi Springer Ser.*, pages 185–228. Bocconi Univ. Press, 2016.

[Bre91] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[Bru87] Hermann Brunn. Über Ovale und Eiflächen. Inaugural Dissertation, München, 1887.

[Bru89] Hermann Brunn. Über Curven ohne Wendepunkte. Habilitationsschrift, München, 1889.

[BDM11] Amarjit Budhiraja, Paul Dupuis, and Vasileios Maroulas. Variational representations for continuous time processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(3):725–747, 2011.

[Caf92a] Luis A. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45(9):1141–1151, 1992.

[Caf92b] Luis A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.

[Caf96] Luis A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.

[Caf00] Luis A. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.*, 214(3):547–563, 2000.

[CDPP09] Pietro Caputo, Paolo Dai Pra, and Gustavo Posta. Convex entropy decay via the Bochner-Bakry-Emery approach. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):734–753, 2009.

[CDPS17] Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.*, 49(2):1385–1418, 2017.

[CFS24] Guillaume Carlier, Alessio Figalli, and Filippo Santambrogio. On optimal transport maps between $\frac{1}{d}$-concave densities. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 2024. To appear.

[CPT23] Guillaume Carlier, Paul Pegon, and Luca Tamanini. Convergence rate of general entropic optimal transport costs. *Calc. Var. Partial Differential Equations*, 62(4):Paper No. 116, 28, 2023.

[CG06] Patrick Cattiaux and Arnaud Guillin. On quadratic transportation cost inequalities. *J. Math. Pures Appl. (9)*, 86(4):341–361, 2006.

[CM17] Fabio Cavalletti and Andrea Mondino. Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds. *Invent. Math.*, 208(3):803–849, 2017.

[Cha04] Djalil Chafaï. Entropies, convexity, and functional inequalities: on Φ-entropies and Φ-Sobolev inequalities. *J. Math. Kyoto Univ.*, 44(2):325–363, 2004.

[CL23] Djalil Chafaï and Joseph Lehec. Logarithmic Sobolev Inequalities Essentials. *Available at* `https://djalil.chafai.net/docs/M2/chafai-lehec-m2-lsie-lecture-notes.pdf`, 2023.

[Che70] Jeff Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Sympos. in honor of Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*, pages 195–199. Princeton Univ. Press, Princeton, NJ, 1970.

[Che82] Louis H. Y. Chen. An inequality for the multivariate normal distribution. *J. Multivariate Anal.*, 12(2):306–315, 1982.

[Che81] Herman Chernoff. A note on an inequality involving the normal distribution. *Ann. Probab.*, 9(3):533–535, 1981.

[CP23] Sinho Chewi and Aram-Alexandre Pooladian. An entropic generalization of Caffarelli's contraction theorem via covariance inequalities. *C. R. Math. Acad. Sci. Paris*, 361:1471–1482, 2023.

[CCGT23] Alberto Chiarini, Giovanni Conforti, Giacomo Greco, and Luca Tamanini. Gradient estimates for the Schrödinger potentials: convergence to the Brenier map and quantitative stability. *Comm. Partial Differential Equations*, 48(6):895–943, 2023.

[CRL+20] Lenaic Chizat, Pierre Roussillon, Flavien Léger, François-Xavier Vialard, and Gabriel Peyré. Faster wasserstein distance estimation with the sinkhorn divergence. *Advances in Neural Information Processing Systems*, 33:2257–2269, 2020.

[CF21] Maria Colombo and Max Fathi. Bounds on optimal transport maps onto log-concave measures. *J. Differential Equations*, 271:1007–1022, 2021.

[CFJ17] Maria Colombo, Alessio Figalli, and Yash Jhaveri. Lipschitz changes of variables between perturbations of log-concave measures. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 17(4):1491–1519, 2017.

[CSM94] Roberto Cominetti and Jaime San Martín. Asymptotic analysis of the exponential penalty trajectory in linear programming. *Math. Programming*, 67(2):169–187, 1994.

[CE25] Giovanni Conforti and Katharina Eichinger. A coupling approach to Lipschitz transport maps. *arXiv preprint arXiv:2502.01353*, 2025.

[CT21] Giovanni Conforti and Luca Tamanini. A formula for the time derivative of the entropic cost and applications. *J. Funct. Anal.*, 280(11):Paper No. 108964, 48, 2021.

[CE02] Dario Cordero-Erausquin. Some applications of mass transport to Gaussian-type inequalities. *Arch. Ration. Mech. Anal.*, 161(3):257–269, 2002.

[CEMS01] Dario Cordero-Erausquin, Robert J. McCann, and Michael Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[CENV04] Dario Cordero-Erausquin, Bruno Nazaret, and Cédric Villani. A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities. *Adv. Math.*, 182(2):307–332, 2004.

[Csi67] Imre Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.

[Cut13] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013.

[DJ13] Fraser Daly and Oliver Johnson. Bounds on the Poincaré constant under negative dependence. *Statist. Probab. Lett.*, 83(2):511–518, 2013.

[DG80] Somesh Das Gupta. Brunn-Minkowski inequality and its aftermath. *J. Multivariate Anal.*, 10(3):296–318, 1980.

[dBL20] Eustasio del Barrio and Jean-Michel Loubes. The statistical effect of entropic regularization in optimal transportation. *arXiv preprint arXiv:2006.05199*, 2020.

[DMN18] Pierre Del Moral and Angèle Niclas. A Taylor expansion of the square root matrix function. *J. Math. Anal. Appl.*, 465(1):259–266, 2018.

[Del22] Alex Delalande. Nearly tight convergence bounds for semi-discrete entropic optimal transport. In *International Conference on Artificial Intelligence and Statistics*, pages 1619–1642. PMLR, 2022.

[DGW04] Hacène Djellout, Arnaud Guillin, and Liming Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32(3B):2702–2732, 2004.

[DLR13] Manh Hong Duong, Vaios Laschos, and Michiel Renger. Wasserstein gradient flows from large deviations of many-particle limits. *ESAIM Control Optim. Calc. Var.*, 19(4):1166–1188, 2013.

[Dvo61] Aryeh Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Jerusalem Academic Press, Jerusalem, 1961.

[EN24] Stephan Eckstein and Marcel Nutz. Convergence rates for regularized optimal transport via quantization. *Math. Oper. Res.*, 49(2):1223–1240, 2024.

[ELS20] Ronen Eldan, Joseph Lehec, and Yair Shenfeld. Stability of the logarithmic Sobolev inequality via the Föllmer process. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):2253–2269, 2020.

[Éme89] Michel Émery. *Stochastic calculus in manifolds*. Universitext. Springer-Verlag, Berlin, 1989. With an appendix by P.-A. Meyer.

[EMR15] Matthias Erbar, Jan Maas, and D. R. Michiel Renger. From large deviations to Wasserstein gradient flows in multiple dimensions. *Electron. Commun. Probab.*, 20:no. 89, 12, 2015.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[Fat24] Max Fathi. Growth estimates on optimal transport maps via concentration inequalities. *arXiv preprint arXiv:2407.11951*, 2024.

[FFGZ24] Max Fathi, Matthieu Fradelizi, Nathael Gozlan, and Simon Zugmeyer. Some obstructions to contraction theorems on the half-sphere. *arXiv preprint arXiv:2402.04649*, 2024.

[FGP20] Max Fathi, Nathael Gozlan, and Maxime Prod'homme. A proof of the Caffarelli contraction theorem via entropic regularization. *Calc. Var. Partial Differential Equations*, 59(3):Paper No. 96, 18, 2020.

[FIL16] Max Fathi, Emanuel Indrei, and Michel Ledoux. Quantitative logarithmic Sobolev inequalities and stability estimates. *Discrete Contin. Dyn. Syst.*, 36(12):6835–6853, 2016.

[FMS24] Max Fathi, Dan Mikulincer, and Yair Shenfeld. Transportation onto log-Lipschitz perturbations. *Calc. Var. Partial Differential Equations*, 63(3):Paper No. 61, 25, 2024.

[Fig13] Alessio Figalli. Stability in geometric and functional inequalities. In *European Congress of Mathematics*, pages 585–599. Eur. Math. Soc., Zürich, 2013.

[Fig19] Alessio Figalli. On the Monge-Ampère equation. *Astérisque*, (414):Exp. No. 1148, 477–503, 2019.

[FMP09] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A refined Brunn-Minkowski inequality for convex sets. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 26(6):2511–2519, 2009.

[FMP10] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. *Invent. Math.*, 182(1):167–211, 2010.

[Föl85] Hans Föllmer. An entropy approach to the time reversal of diffusion processes. In *Stochastic differential systems (Marseille-Luminy, 1984)*, volume 69 of *Lect. Notes Control Inf. Sci.*, pages 156–163. Springer, Berlin, 1985.

[Föl86] Hans Föllmer. Time reversal on Wiener space. In *Stochastic processes—mathematics and physics (Bielefeld, 1984)*, volume 1158 of *Lecture Notes in Math.*, pages 119–129. Springer, Berlin, 1986.

[FGJ17] Joaquin Fontbona, Nathael Gozlan, and Jean-François Jabir. A variational approach to some transport inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(4):1719–1746, 2017.

[Fou05] Pierre Fougères. Spectral gap for log-concave probability measures on the real line. In *Séminaire de Probabilités XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 95–123. Springer, Berlin, 2005.

[GHL04] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry.* Universitext. Springer-Verlag, Berlin, third edition, 2004.

[GM96] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[Gar02] Richard J. Gardner. The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)*, 39(3):355–405, 2002.

[Gel90] Matthias Gelbrich. On a formula for the L^2 Wasserstein metric between measures on Euclidean and Hilbert spaces. *Math. Nachr.*, 147:185–203, 1990.

[Gen01] Ivan Gentil. *Inégalités de Sobolev logarithmiques et hypercontractivité en mécanique statistique et en E.D.P.* PhD thesis, Université Paul Sabatier - Toulouse III, 2001.

[GLR20] Ivan Gentil, Christian Léonard, and Luigia Ripani. Dynamical aspects of the generalized Schrödinger problem via Otto calculus—a heuristic point of view. *Rev. Mat. Iberoam.*, 36(4):1071–1112, 2020.

[Gig11a] Nicola Gigli. On Hölder continuity-in-time of the optimal transport map towards measures along a curve. *Proc. Edinb. Math. Soc. (2)*, 54(2):401–409, 2011.

[Gig11b] Nicola Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(\mathscr{P}_2(M), W_2)$. *Methods Appl. Anal.*, 18(2):127–158, 2011.

[GL13] Nicola Gigli and Michel Ledoux. From log Sobolev to Talagrand: a quick proof. *Discrete Contin. Dyn. Syst.*, 33(5):1927–1935, 2013.

[GT21] Nicola Gigli and Luca Tamanini. Second order differentiation formula on $\mathrm{RCD}^*(K, N)$ spaces. *J. Eur. Math. Soc. (JEMS)*, 23(5):1727–1795, 2021.

[GJY03] Anja Göing-Jaeschke and Marc Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[GSN24] Alberto González-Sanz and Marcel Nutz. Sparsity of quadratically regularized optimal transport: Scalar case. *arXiv preprint arXiv:2410.03353*, 2024.

[Goz09] NATHAEL GOZLAN. A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.*, 37(6):2480–2498, 2009.

[GL07] NATHAEL GOZLAN AND CHRISTIAN LÉONARD. A large deviation approach to some transportation cost inequalities. *Probab. Theory Related Fields*, 139(1-2):235–283, 2007.

[GL10] NATHAEL GOZLAN AND CHRISTIAN LÉONARD. Transport inequalities. A survey. *Markov Process. Related Fields*, 16(4):635–736, 2010.

[GRS14] NATHAEL GOZLAN, CYRIL ROBERTO, AND PAUL-MARIE SAMSON. Hamilton Jacobi equations on metric spaces and transport entropy inequalities. *Rev. Mat. Iberoam.*, 30(1):133–163, 2014.

[Gro80] MIKHAEL GROMOV. Paul Lévy's isoperimetric inequality. *Preprint IHES*, 2:398, 1980.

[Gro99] MIKHAEL GROMOV. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[GM83] MIKHAEL GROMOV AND VITALI D. MILMAN. A topological application of the isoperimetric inequality. *Amer. J. Math.*, 105(4):843–854, 1983.

[Gro75] LEONARD GROSS. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[Gro53] ALEXANDRE GROTHENDIECK. Sur certaines classes de suites dans les espaces de Banach et le théorème de Dvoretzky-Rogers. *Bol. Soc. Mat. São Paulo*, 8:81–110, 1953.

[Har04] GILLES HARGÉ. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields*, 130(3):415–440, 2004.

[Har66] LAWRENCE H. HARPER. Optimal numberings and isoperimetric problems on graphs. *J. Combinatorial Theory*, 1:385–393, 1966.

[JMPC20] HICHAM JANATI, BORIS MUZELLEC, GABRIEL PEYRÉ, AND MARCO CUTURI. Entropic Optimal Transport between Unbalanced Gaussian Measures has a Closed Form. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 10468–10479. Curran Associates, Inc., 2020.

[Joh17] OLIVER JOHNSON. A discrete log-Sobolev inequality under a Bakry-Émery type condition. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(4):1952–1970, 2017.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[KLS95] Ravindran Kannan, László Lovász, and Miklós Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13(3-4):541–559, 1995.

[Kan42] Leonid Kantorovich. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.

[KM12] Young-Heon Kim and Emanuel Milman. A generalization of Caffarelli's contraction theorem via (reverse) heat flow. *Math. Ann.*, 354(3):827–862, 2012.

[Kla85] Chris A. J. Klaassen. On an inequality of Chernoff. *Ann. Probab.*, 13(3):966–974, 1985.

[Kla17] Bo'az Klartag. Needle decompositions in Riemannian geometry. *Mem. Amer. Math. Soc.*, 249(1180):v+77, 2017.

[Kla23] Bo'az Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inven. Anal.*, pages Paper No. 4, 17, 2023.

[KL19] Bo'az Klartag and Joseph Lehec. Poisson processes and a log-concave Bernstein theorem. *Studia Math.*, 247(1):85–107, 2019.

[KL24a] Bo'az Klartag and Joseph Lehec. Affirmative Resolution of Bourgain's Slicing Problem using Guan's Bound. *arXiv preprint math/2412.15044*, 2024.

[KL24b] Bo'az Klartag and Joseph Lehec. Isoperimetric inequalities in high-dimensional convex sets. *arXiv preprint math/2406.01324*, 2024.

[KP23] Bo'az Klartag and Eli Putterman. Spectral monotonicity under Gaussian convolution. *Ann. Fac. Sci. Toulouse Math. (6)*, 32(5):939–967, 2023.

[Kno57] Herbert Knothe. Contributions to the theory of convex bodies. *Michigan Math. J.*, 4:39–52, 1957.

[Kra82] Allan M. Krall. On boundary values for the Laguerre operator in indefinite inner product spaces. *J. Math. Anal. Appl.*, 85(2):406–408, 1982.

[Kul67] Solomon Kullback. A lower bound for discrimination information in terms of variation (corresp.). *IEEE Transactions on Information Theory*, 13(1):126–127, 1967.

[Las16] Günter Last. Stochastic analysis for Poisson processes. In *Stochastic analysis for Poisson point processes*, volume 7 of *Bocconi Springer Ser.*, pages 1–36. Bocconi Univ. Press, 2016.

[LS22] Hugo Lavenant and Filippo Santambrogio. The flow map of the Fokker-Planck equation does not provide optimal transport. *Appl. Math. Lett.*, 133:Paper No. 108225, 7, 2022.

[Led93] Michel Ledoux. L'algèbre de Lie des gradients itérés d'un générateur markovien. *C. R. Acad. Sci. Paris Sér. I Math.*, 317(11):1049–1052, 1993.

[Led95] Michel Ledoux. L'algèbre de Lie des gradients itérés d'un générateur markovien—développements de moyennes et entropies. *Ann. Sci. École Norm. Sup. (4)*, 28(4):435–460, 1995.

[Led00] Michel Ledoux. The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math. (6)*, 9(2):305–366, 2000. Probability theory.

[Led01] Michel Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[LNP15] Michel Ledoux, Ivan Nourdin, and Giovanni Peccati. Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.*, 25(1):256–306, 2015.

[LT91] Michel Ledoux and Michel Talagrand. *Probability in Banach spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. Isoperimetry and processes.

[Leh13] Joseph Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):885–899, 2013.

[Lei72] László Leindler. On a certain converse of Hölder's inequality. In *Linear operators and approximation (Proc. Conf., Math. Res. Inst., Oberwolfach, 1971)*, volume Vol. 20 of *Internat. Ser. Numer. Math.*, pages 182–184. Birkhäuser Verlag, Basel-Stuttgart, 1972.

[Léo12] Christian Léonard. Girsanov theory under a finite entropy condition. In *Séminaire de Probabilités XLIV*, volume 2046 of *Lecture Notes in Math.*, pages 429–465. Springer, Heidelberg, 2012.

[Léo14] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[Lév51] Paul Lévy. *Problèmes concrets d'analyse fonctionnelle. Avec un complément sur les fonctionnelles analytiques par F. Pellegrino*. Gauthier-Villars, Paris, 1951. 2d ed.

[Liu11] Wei Liu. Optimal transportation-entropy inequalities for several usual distributions on $\mathbb{R}$. *Acta Math. Appl. Sin. Engl. Ser.*, 27(4):713–720, 2011.

[LR25a] Pablo López-Rivera. A Bakry-Émery Approach to Lipschitz Transportation on Manifolds. *Potential Anal.*, 62(2):331–353, 2025.

[LR25b] Pablo López-Rivera. A uniform rate of convergence for the entropic potentials in the quadratic Euclidean setting. *arXiv preprint arXiv:2502.00084*, 2025.

[LRS25] Pablo López-Rivera and Yair Shenfeld. The Poisson transport map. *J. Funct. Anal.*, 288(10):Paper No. 110864, 2025.

[Lot03] John Lott. Some geometric properties of the Bakry-Émery-Ricci tensor. *Comment. Math. Helv.*, 78(4):865–883, 2003.

[LV09] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[MSWW11] Yutao Ma, Shi Shen, Xinyu Wang, and Liming Wu. Transportation inequalities: from Poisson to Gibbs measures. *Bernoulli*, 17(1):155–169, 2011.

[MS23] Hugo Malamut and Maxime Sylvestre. Convergence rates of the regularized optimal transport: Disentangling suboptimality and entropy. *arXiv preprint arXiv:2306.06940*, 2023.

[MGM22] Anton Mallasto, Augusto Gerolin, and Hà Quang Minh. Entropy-regularized 2-Wasserstein distance between Gaussian measures. *Inf. Geom.*, 5(1):289–323, 2022.

[MPV23] Maxime Mathey-Prevot and Alain Valette. Wasserstein distance and metric trees. *Enseign. Math.*, 69(3-4):315–333, 2023.

[McC95] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.

[McC97] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[McC01] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[MS23] Dan Mikulincer and Yair Shenfeld. On the lipschitz properties of transportation along heat flows. In Ronen Eldan, Bo'az Klartag, Alexander Litvak, and Emanuel Milman, editors, *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2020-2022*, pages 269–290. Springer International Publishing, Cham, 2023.

[MS24] Dan Mikulincer and Yair Shenfeld. The Brownian transport map. *Probab. Theory Related Fields*, 190(1-2):379–444, 2024.

[Mil09] Emanuel Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.*, 177(1):1–43, 2009.

[Mil23] Emanuel Milman. Reverse Hölder inequalities for log-Lipschitz functions. *Pure Appl. Funct. Anal.*, 8(1):297–310, 2023.

[Mil71] Vitali D. Milman. A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. *Funkcional. Anal. i Priložen.*, 5(4):28–37, 1971.

[MS86] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[Min10] Hermann Minkowski. *Geometrie der Zahlen*, volume 1. B.G. Teubner, 1910.

[Mon81] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie royale des Sciences, avec les Mémoires de Mathématique et de Physique*, pages 666–704, 1781.

[Nas58] John Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.

[Nee22] Joe Neeman. Lipschitz changes of variables via heat flow, 2022. arXiv preprint arXiv:2201.03403.

[Nel66] Edward Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles (Proc. Conf., Dedham, Mass., 1965)*, pages 69–73. MIT Press, Cambridge, Mass.-London, 1966.

[Nut21] Marcel Nutz. Introduction to entropic optimal transport, 2021. Lecture notes, Columbia University.

[NW22] Marcel Nutz and Johannes Wiesel. Entropic optimal transport: convergence of potentials. *Probab. Theory Related Fields*, 184(1-2):401–424, 2022.

[Ott01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[OV00] Felix Otto and Cédric Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[Pal24] Soumik Pal. On the difference between entropic cost and the optimal transport cost. *Ann. Appl. Probab.*, 34(1B):1003–1028, 2024.

[PW60] Lawrence E. Payne and Hans F. Weinberger. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.*, 5:286–292, 1960.

[Per02] Grisha Perelman. The entropy formula for the Ricci flow and its geometric applications. *arXiv preprint math/0211159*, 2002.

[PC19] Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[Pin64] Mark S. Pinsker. *Information and information stability of random variables and processes.* Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, 1964.

[Pis86] GILLES PISIER. Probabilistic methods in the geometry of Banach spaces. In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 167–241. Springer, Berlin, 1986.

[Poi87] HENRI POINCARÉ. Sur la théorie analytique de la chaleur. *Comptes rendus de l'Académie des sciences*, 104:1753–1759, 1887.

[Poi90] HENRI POINCARÉ. Sur les Équations aux Dérivées Partielles de la Physique Mathématique. *Amer. J. Math.*, 12(3):211–294, 1890.

[PNW21] ARAM-ALEXANDRE POOLADIAN AND JONATHAN NILES-WEED. Entropic estimation of optimal transport maps. *arXiv preprint arXiv:2109.12004*, 2021.

[Pré71] ANDRÁS PRÉKOPA. Logarithmic concave measures with application to stochastic programming. *Acta Sci. Math. (Szeged)*, 32:301–316, 1971.

[Pré73] ANDRÁS PRÉKOPA. On logarithmic concave measures and functions. *Acta Sci. Math. (Szeged)*, 34:335–343, 1973.

[Roc70] R. TYRRELL ROCKAFELLAR. *Convex analysis*, volume No. 28 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1970.

[Ros05] ANTONIO ROS. The isoperimetric problem. In *Global theory of minimal surfaces*, volume 2 of *Clay Math. Proc.*, pages 175–209. Amer. Math. Soc., Providence, RI, 2005.

[Rot81] OSCAR S. ROTHAUS. Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities. *J. Functional Analysis*, 42(1):102–109, 1981.

[SGK24] RITWIK SADHU, ZIV GOLDFELD, AND KENGO KATO. Approximation rates of entropic maps in semidiscrete optimal transport. *arXiv preprint arXiv:2411.07947*, 2024.

[San15] FILIPPO 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.

[SW14] ADRIEN SAUMARD AND JON A. WELLNER. Log-concavity and strong log-concavity: a review. *Stat. Surv.*, 8:45–114, 2014.

[Sch48] ERHARD SCHMIDT. Die Brunn-Minkowskische Ungleichung und ihr Spiegelbild sowie die isoperimetrische Eigenschaft der Kugel in der euklidischen und nichteuklidischen Geometrie. I. *Math. Nachr.*, 1:81–157, 1948.

[Ser24] JORDAN SERRES. Behavior of the Poincaré constant along the Polchinski renormalization flow. *Commun. Contemp. Math.*, 26(7):Paper No. 2350035, 16, 2024.

[She24] Yair Shenfeld. Exact renormalization groups and transportation of measures. *Ann. Henri Poincaré*, 25(3):1897–1910, 2024.

[SMR18] Filip Soudský, Anastasia Molchanova, and Tomáš Roskovec. Interpolation between Hölder and Lebesgue spaces with applications. *J. Math. Anal. Appl.*, 466(1):160–168, 2018.

[Sta59] Aart Johannes Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, 2:101–112, 1959.

[Stu06a] Karl-Theodor Sturm. On the geometry of metric measure spaces. *Acta Mathematica*, 196(1):65 – 131, 2006.

[Stu06b] Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Mathematica*, 196(1):133 – 177, 2006.

[ST74] Vladimir N. Sudakov and Boris S. Tirel'son. Extremal properties of half-spaces for spherically invariant measures. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 41:14–24, 165, 1974. Problems in the theory of probability distributions, II.

[Tal91] Michel Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric aspects of functional analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 94–124. Springer, Berlin, 1991.

[Tal96] Michel Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

[Tan21] Anastasiya Tanana. Comparison of transport map generated by heat flow interpolation and the optimal transport Brenier map. *Commun. Contemp. Math.*, 23(6):Paper No. 2050025, 7, 2021.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

[Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009. Old and new.

[Vil19] Cédric Villani. Inégalités isopérimétriques dans les espaces métriques mesurés [d'après F. Cavalletti & A. Mondino]. *Astérisque*, (407):Exp. No. 1127, 213–265, 2019. Séminaire Bourbaki. Vol. 2016/2017. Exposés 1120–1135.

[Wal76] David W. Walkup. Pólya sequences, binomial convolution and the union of random sets. *J. Appl. Probability*, 13(1):76–85, 1976.

[Wu00] Liming Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.

[Yos80] Kôsaku Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, sixth edition, 1980.

Al final del viaje está el horizonte
Al final del viaje partiremos de nuevo
Al final del viaje comienza un camino
Otro buen camino que seguir, descalzos, contando la arena.

Silvio Rodríguez
Al final de este viaje en la vida