THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris-Dauphine
Dans le cadre d'une cotutelle avec l'Universitá Milano-Bicocca, l'Universitá di Pavia, l'INdAM et le CIAFM

Equations cinétiques, inégalités fonctionelles et distances dans l'espace des mesures de probabilité

Kinetic equations, functional inequalities, and distances in the space of probability measures

Soutenue par
Giovanni BRIGATI
Le 25 mai 2023

École doctorale nº543
École doctorale SDOSE

Spécialité
Mathématiques

Composition du jury :

Ivan GENTIL Professeur de l'Université, Université Claude Bernard Lyon 1	*Rapporteur*
Clément MOUHOT Professeur de l'Université, University of Cambridge	*Rapporteur*
Giuseppe TOSCANI Professeur émérite, Università di Pavia	*Rapporteur*
Matteo BONFORTE Professeur de l'Université, Universidad Autónoma de Madrid	*Examinateur*
Michael LOSS Professeur de l'Université, Georgia Institute of Technology	*Examinateur*
Maria J. ESTEBAN Directrice de Recherches émérite CNRS, Université Paris-Dauphine PSL	*Présidente*
Giuseppe SAVARÉ Professeur de l'Université, Università Commericale Luigi Bocconi	*Co-Directeur*
Jean DOLBEAULT Directeur de Recherches CNRS, Université Paris-Dauphine PSL	*Directeur de thèse*
Max FATHI Professeur de l'Université, Sorbonne Université, Paris	*Invité*
Gabriel STOLTZ Professeur de l'Université, Ecole des Ponts ParisTech	*Invité*

To Ceci ☽*, and Mari* ✳

Contents

Acknowledgements

It is nearly impossible to translate into words the gratitude I feel for my director, Jean Dolbeault. The investment of time, attention, and energy, both in quantity and quality, he has been making on myself since 2019 is a true one of a kind. It is the greatest luck and honour to be a student of him. Not only I have been taught beautiful mathematical ideas from his own research, but I have been given the instruments and the space to grow and to be able to express myself trough mathematics. I am still very glad to have been working in some of his projects, and to have made some progress as a student thanks to them. In addition, Jean Dolbeault is an example to follow for me, educating me towards the ethics of our profession, helping me to integrate the French community, and guiding me with his presence and precious insights in many different – and sometimes complicated – situations.

I thank Professor Giuseppe Savaré, for co-directing my PhD thesis, for advice, and support along all my life as a student. I learnt so much under the supervision of him, and my first contact with actual research took place during the preparation of my MSc thesis (supervised by Professor Savaré himself). So, my approach to mathematics cannot be separated from the influence and the mentorship of him.

To Gabriel Stoltz, I feel grateful for our collaboration, and for all the time you spent exchanging ideas with me, which has been crucial for my training. It is an authentic pleasure to work with you. Thank you.

I thank Nikita Simonov, for teaching me the values of precision and scepticism, and pushing me to work harder. Your role, halfway between a teacher and a more expert colleague, has been unique in my education, for your wit and your dedication. Sometimes it has been hard, but, if I have improved on my weakest sides, it is also because of you.

Thanks to Ivailo Hartarsky, for his genius and elegance in chasing a tricky problem with me, which led to our joint paper.

I thank Ivan Gentil, Clément Mouhot, and Giuseppe Toscani for their reviews, and for their valuable remarks improving the quality of this thesis. I thank Jan Maas for inviting me to Vienna, where a part of this work has been presented. Thanks to Franz Achleitner, Anton Arnold, Emeric Bouin, Matteo Bonforte, Pierre Cardaliaguet, Max Fathi, and Lihan Wang for very interesting discussions.

I acknowledge the FSMP foundation for funding my PhD thesis, and for the unique chance to be part of their Cofund:MathInParis program, managed by Ariela Briani.

Working for almost four years in the excellent environment of CEREMADE has been a great honour for me, and an occasion to learn so much from the interactions with my colleagues and the senior laboratory members, to whom I am very grateful. I would like to mention in particular Katharina Eichinger, Claudia Fonte Sanchez, Gregoire Clarté, Lorenzo Croissant, Donato Scarcella, Luca Ziviani, Chanqing Fu, *Théo Bertand, Umberto Morellini, and Daniele de Gennaro*[1]. I thank the laboratory for funding my participation to conferences and research schools, and all the staff. I thank the doctoral school SDOSE, and Béatrice Taupinart de Tilière, for didactic and financial support. I thank Pierluigi Colli, director of the doctoral school in Pavia. I thank two colleagues and friends from abroad: Esther Bou Dagher, and Davide Manini.

[1] co-organising the YRS and the YRD has been so nice!

I acknowledge my family for all the love and the support I have received along my life. My sisters (this work is dedicated to you, see page i) Cecilia and Maria Sole; my grandparents Ernestina, Edda, Roberto, and Angelo; my parents Margherita and Marco. My desire of committing with science and starting a path in academia owes a lot to the values of knowledge, culture, and education I have been taught by you.

I thank Samuele Gallinari, Ilaria Baroli, Ivonne Spano, Fabio Moroni, Matteo Pati, Alice Marveggio, Flavia Deprati Stroppa, Nicolò Gambardella, and Andrea Panella for the true friendship you have shown me day after day. I also thank all my friends from Rotaract Club Paris Ouest, Rotaract Club Pavia, and Tennis Club Pavia.

Finally, I am grateful to the French academic community and the city of Paris for the welcoming, beautiful, and kind environment I could spend my French years in. *See you soon!*

This thesis has been funded by the European Unions Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 754362. Partial support has been received by the Project EFI (ANR-17-CE40-0030) of the French National Research Agency (ANR).

Résumé

Cette thèse est consacrée aux équations cinétiques de Fokker-Planck, à la stabilité des inégalités fonctionnelles et aux formes de Dirichlet non linéaires. Des taux de convergence vers l'équilibre sont estimés via un cadre analytique fonctionnel basé sur les normes faibles des solutions. L'équation de Vlasov-Fokker-Planck, avec variable de position confinée dans un tore, est analysée comme modèle de référence. La même stratégie est ensuite etendue à une large classe de modèles cinétiques. Nous considérons également des inégalités de Gagliardo–Nirenberg sur la sphère qui interpolent entre les inégalités de Poincaré et les inégalités de Sobolev. Nous prouvons des résultats constructifs de stabilité, dans la norme la plus forte possible, avec des exposants optimaux. L'estimation de stabilité dégénère sur un sous-espace de dimension finie, ce qui nécessite des précautions supplémentaires. Notre technique combine des développements de Taylor, l'analyse harmonique et des methodes parabòliques. Nous prouvons rigoureusement la convergence de la famille de Gagliardo-Nirenberg sur la sphère vers les inégalités de Beckner gaussiennes dans la limite de grandes dimensions. Ensuite, nous donnons des résultats constructifs de stabilité, en utilisant des diffusions non linéaires sur l'espace gaussien. Enfin, nous traitons l'inégalité de Sobolev logarithmique gaussienne comme cas limite. Nous trouvons des estimations explicites de stabilité pour des densités log-concaves ou à support compact par un argument de log-concavité deduit du flot d'Ornstein–Uhlenbeck et de la méthode du *carré du champ*. Nous contribuons à la théorie des formes de Dirichlet non-linéaires en etendant la propriété de contraction normale. La preuve adopte une nouvelle stratégie, basée sur l'approximation des fonctions Lipschitz réelles par des compositions répétées de fonctions linéaires par morceaux simples.

Sommario

Questa tesi analizza le equazioni cinetiche di Fokker–Planck, la stabilità di disuguaglianze funzionali e le forme di Dirichlet non lineari. Tassi costruttivi di convergenza all'equilibrio per equazioni cinetiche sono calcolati all'interno di un quadro analitico funzionale basato sulle norme deboli delle soluzioni. L'equazione di Vlasov–Fokker–Planck, con la variabile spaziale confinata in un toro, viene analizzata come caso modello. La stessa strategia viene poi generalizzata a un'ampia classe di modelli cinetici. Si considerano anche le disuguaglianze di Gagliardo–Nirenberg sulla sfera, interpolanti tra la disuguaglianza di Poincaré e quella critica di Sobolev. Si dimostrano risultati di stabilità costruttiva, nella norma più forte possibile, con esponenti ottimali nel termine di stabilità. Quest'ultimo termine è, però, degenere su un sottospazio di dimensione finita, che richiede quindi ulteriore precisione. La tecnica proposta combina sviluppi di Taylor, analisi armonica e flussi di diffusione (non) lineari. La convergenza della famiglia di Gagliardo–Nirenberg alle disuguaglianze gaussiane di Beckner, per la dimensione della sfera che tende all'infinito, viene dimostrata in modo rigoroso. Vengono poi dati risultati di stabilità costruttiva per tali disuguaglianze di Beckner, utilizzando flussi di diffusione non lineare sullo spazio gaussiano. Infine si studia la disuguaglianza logaritmica di Sobolev come caso limite. Si danno stime di stabilità esplicite per densità log-concave o a supporto compatto, grazie all'interazione tra log-concavità e flusso di Ornstein–Uhlenbeck, migliorando il metodo del *carré du champ*. Si portano contributi alla teoria delle forme di Dirichlet non lineari, estendendo la proprietà di contrazione normale al contesto non lineare. La dimostrazione adotta una nuova strategia, basata sull'approssimazione di funzioni lipschitziane reali mediante composizioni ripetute di funzioni elementari lineari a tratti.

Abstract

This thesis concerns kinetic Fokker-Planck equations, stability of functionals inequalities, and nonlinear Dirichlet forms. Constructive convergence rates to equilibrium for kinetic equations are computed via a functional analytic framework based on weak norms of solutions. The Vlasov-Fokker-Planck equation, with the space variable confined in a torus, is analysed as a benchmark. Then, the same strategy is generalised to a wide class of kinetic Fokker-Planck models. We also consider Gagliardo-Nirenberg inequalities on the sphere, interpolating between the Poincaré and the Sobolev inequalities. We prove constructive stability results, in the strongest possible norm, with sharp exponents in the distance from optimisers. This term degenerates on a finite-dimensional subspace requiring additional care. Our technique combines Taylor expansions, harmonic analysis, and (non)linear diffusion flows. We rigorously prove convergence of the Gagliardo-Nirenberg family of inequalities, for the dimension of the sphere approaching infinity, to the Gaussian Beckner inequalities. Then, we give constructive stability results for those, using nonlinear diffusion flows on the Gaussian space. Finally, we treat the Gaussian logarithmic Sobolev inequality as a limit case. We find explicit stability estimates for log-concave or compactly-supported densities, thanks to the interplay between log-concavity and the Ornstein-Uhlenbeck flow, using the *carré du champ* method. We contribute to the theory of nonlinear Dirichlet forms. We extend the normal contraction property to the nonlinear setting. The proof adopts a new strategy, based on the approximation of real Lipschitz functions with repeated compositions of elementary piecewise linear functions.

Part I

Introduction and main results

Chapter 1

Introduction

Introduction en français

Cette thèse traite des équations aux dérivées partielles décrivant l'évolution temporelle de systèmes composés de nombreuses particules, qui sont soumises aux lois de la mécanique comprenant des interactions entre particules. La dynamique à l'échelle microscopique est réversible en temps, mais l'irréversibilité apparaît dans le système à la limite thermodynamique : l'évolution de l'entropie donne une direction au temps. Dans le langage de la physique, l'entropie augmente avec le temps, conduisant le système à des états d'entropie maximale. Ces états sont stables et donc appelés équilibres. Nous quantifions la vitesse à laquelle les équilibres sont approchés. Pour ce faire, nous trouvons des entropies adaptées et établissons une inégalité entre l'augmentation de l'entropie et l'entropie elle-même. Lorsque l'inégalité est optimale, le taux de convergence vers l'équilibre est optimal. Des autres améliorations dans ces inégalités permettent de saisir des propriétés plus fines de l'évolution temporelle. Ces idées sont appliquées à diverses équations cinétiques et de diffusion.

Nous considérons des systèmes complexes composés d'un grand nombre de particules ou d'agents interagissant ensemble et avec un espace ambiant, dont l'échelle caractéristique est supposée beaucoup plus grande que celle d'une particule typique. Ces systèmes apparaissent dans une large variété de phénomènes naturels, qui vont de la dynamique des molécules dans un gaz ou d'une population de bactéries aux interactions entre étoiles à l'intérieur d'une galaxie, et entre les galaxies elles-mêmes, en passant par les vols d'oiseaux, les bancs de poissons et les comportements socio-économiques. Ces systèmes, et leur évolution dans le temps, peuvent être étudiés à trois niveaux de détails différents, généralement appelés niveaux microscopiques, niveau mésoscopique (ou cinétique) et niveau macroscopique (diffusif où déduit par la dynamique des fluides). Nous renvoyons à [254] pour une discussion beaucoup plus large du sujet.

Le niveau microscopique est le plus détaillé. À ce niveau, toutes les interactions entre les particules sont prises en compte. Individuellement, chaque particule obéit aux lois élémentaires de la physique, typiquement aux lois de Newton ou au principe de moindre action. En outre, nous avons des interactions entre les différentes particules du système, qui peuvent être à longue portée, comme c'est le cas pour les charges électriques et la gravité, ou à courte portée. Ces dernières sont généralement modélisées comme des collisions, régies par une loi d'interaction microscopique. À ce niveau, toutes les dynamiques sont réversibles en temps.

En théorie, la description du système est complète, mais elle n'est jamais accessible en pratique. En effet, dans les systèmes d'intérêt physique, l'ordre de grandeur du nombre de particules est de 10^{9-10} particules pour les phénomènes biologiques et pour la dynamique stellaire, et 10^{23} – l'ordre du nombre d'Avogadro – pour les gaz. Il n'est donc pas possible de suivre toutes les particules et toutes les interactions. Il faut donc envisager un autre niveau de description, plus facile à traiter, au prix d'une perte d'information.

C'est le cas du niveau mésoscopique ou cinétique, où l'objet d'étude est la fonction de distribution des particules en fonction de deux variables, généralement la position et la vitesse, et son évolution dans le temps. L'histoire de chaque particule n'est pas reconstituée et le phénomène est décrit comme l'évolution d'une densité de probabilité sur un espace de phase. Une équation cinétique est alors écrite pour la fonction de distribution. Cette approche remonte à Maxwell, voire [205].

Le passage de l'échelle microscopique à l'échelle mésoscopique se fait généralement en prenant la limite pour un nombre N de particules dans le système qui tend à l'infini, connue sous le nom de *limite thermodynamique*. Un choix d'échelles doit être fait, donc resultant en différentes équations cinétiques. Par exemple, on peut se concentrer sur les interactions faibles et à longue portée, en prenant une limite de champ moyen, et obtenir une équation de Vlasov, voire [163, 174]. Alternativement, les limites de Boltzmann-Grad peuvent être considérées, afin de capturer des interactions intenses mais à courte portée, [153]. C'est la façon de dériver l'équation de Boltzmann à partir de la dynamique microscopique. Dans cette thèse, nous nous concentrerons sur les équations cinétiques de type Boltzmann, avec des noyaux du collision simplifiés.

Dans [62], il est prouvé (de manière formelle) que l'entropie thermodynamique des solutions de l'équation de Boltzmann augmente, au fur et à mesure que le temps s'écoule. L'irréversibilité apparaît à la limite thermodynamique, alors même que toutes les dynamiques au niveau microscopique sont réversibles. Grâce à la monotonie de l'entropie au cours de l'évolution, l'équation de Boltzmann prend en compte l'irréversibilité des grands systèmes dans une description statistique. L'analyse de [62] donne un autre type de résultat. La production d'entropie le long de l'équation est explicitement calculée, et il est prouvé qu'elle disparaît si et seulement si la fonction de distribution est une distribution maxwellienne (ou gaussienne). Ce dernier point, combiné avec la monotonie de l'entropie, prouve que les solutions de l'équation de Boltzmann convergent vers telles solutions en temps long. Les états d'entropie maximal (mesurés par l'entropie) sont stables et attirent toutes les solutions dans la limite des temps grands, quelles que soient les conditions initiales. Le problème de l'estimation quantitative du taux de convergence vers les états d'équilibre est un sujet de recherche majeure en théorie cinétique.

En se concentrant sur un régime spécial de l'équation de Boltzmann, correspondant aux collisions rasantes, on peut obtenir l'équation de Landau. Cette thèse traite d'une simplification linéaire de l'équation de Landau : l'équation de Vlasov–Fokker–Planck. L'équation de Vlasov–Fokker–Planck (VFP) ou équation cinétique de Fokker–Planck s'écrit comme une équation aux dérivées partielles portant sur la fonction de distribution écrite sur l'espace des phases, typiquement l'espace des positions et des vitesses. Cette équation est pilotée par un opérateur de collision de type diffusion dans les variables de vitesse uniquement et un opérateur de transport, reliant les variables de position et de vitesse. L'équation VFP présente un intérêt indépendant. Premièrement, ses solutions sont la loi d'un processus aléatoire appelé *dynamique de Langevin*, [44], qui est une description mésoscopique d'un système de particules soumises à une force aléatoire. L'équation VFP est aussi bien adaptée à la description des plasmas et à la dynamique stellaire, [28, 92]. Certaines de ses généralisations sont également présentes en statistique, ou encore dans les simulations moléculaires et divers autres modèles de la physique [243, 152, 44]. Enfin, certaines applications socio-économiques ont èté proposeés plus rècenement et sont exposeés par exemple dans [228].

Comme pour l'équation de Boltzmann, les solutions de l'équation VFP convergent vers des fonctions maxwelliennes en temps long. La différence est que l'équation VFP est beaucoup plus simple et que les taux de convergence peuvent être quantifiés à l'aide de différentes techniques. Dans le cas le plus simple, une fonction de Green est disponible et les solutions sont totalement explicites [183], ce qui permet des calculs directs. Dans des cas plus géneéraux, l'état de l'art consiste à utiliser une technique développée au cours des vingt dernières années, connue sous le nom d'*hypocoercivité*. Nous en donnons une brève description, dans le cas spécifique de l'équation VFP. Le calcul des taux de convergence pour l'équation VFP n'est pas aussi facile que pour une EDP parabolique, car l'équation VFP est une équation sur une fonctionne inconnue dependant de deux variables de position et vitesse et la relaxation est induite par un opérateur de diffusion seulement dans les variables de vitesse. Les variables spatiales sont affectées indirectement par la diffusion, via l'opérateur de transport. L'opérateur de diffusion a un noyau non trivial, dont les éléments sont appelés *équilibres locaux*. Cependant, la fonction de distribution ne peut pas rester dans un état d'équilibre local qui ne soit pas un équilibre global, du fait de l'opérateur de transport. Cet effet combiné se poursuit jusqu'au point que l'équilibre global soit atteint.

Le générateur de l'équation VFP appartient à une classe d'opérateurs de diffusion dégénérés de la forme $A^\star A + B$, où A et B sont des opérateurs différentiels du premier ordre, et $A^\star A$ n'est pas uniformément elliptique (il n'agit pas sur la position). Ces operateurs A, B ont été analysés dans [169], avec pour objectif príncipal l'étude des propriétés de régularité (*hypoellipticité*).

L'intuition de s'inspirer par l'hypoellipticité pour les taux de convergence est utilisée dans [255, 221, 165], à la suite de [169, 204, 162]. Cette théorie a été baptisée *hypocoercivité* par Gallay, et améliore les résultats antérieurs des années 90, concernant le comportement à long terme des EDP cinétiques, en quantifiant les taux de convergence. Du fait que l'opérateur de diffusion dans l'équation VFP est dégénéré, il ne va pas être coercif par rapport à des normes standard comme la norme H^1. Les techniques d'hypocoercivité résolvent le problème en considérant des normes *tordues* équivalentes qui capturent mieux l'effet combiné de la diffusion et du transport. Dans une telle norme modifiée, des taux de décroissance exponentiels peuvent être calculés. Ils sont ensuite transférés à une norme de référence par équivalence. La même idée a été utilisée pour la norme L^2 dans [127], voir aussi le Chapitre 2. Le cadre de cet article a l'avantage supplémentaire d'être compatible avec les *limites de diffusion*, correspondent au passage de l'échelle cinétique à la dynamique des fluides.

Le niveau de description cinétique est en fait un niveau intermédiaire, et une échelle macroscopique peut être introduite. Le niveau macroscopique correspond à la dynamique des fluides. Bien que moins détaillée que l'échelle cinétique, la dynamique des fluides est généralement plus facile à observer et plus pratique aussi bien pour les simulations numériques et les applications en physique et en ingénierie. Dans le passage entre les deux échelles, le libre parcours moyen des particules est ramené à zéro, tandis que le nombre moyen d'interactions par particule et par unité de temps est envoyé à l'infini. Les variables de vitesse convergent vers l'équilibre à une échelle de temps beaucoup plus rapide que les variables spatiales, de sorte que l'évolution est décrite dans la limite par un nombre fini des moments en fonction du temps et de la variable spatiale. Voir [59, 163] pour des résultats de convergence rigoureux dans le contexte des équations de Boltzmann ou de Vlasov, respectivement. Dans le cas de l'équation VFP, la limite de diffusion donne l'équation de la chaleur ou de Fokker–Planck au premier ordre, [127].

Les contributions dans cette direction participent du sixième problème d'Hilbert : «*Ainsi, les travaux de Boltzmann sur les principes de la mécanique suggèrent le problème de développer mathématiquement les processus limites, là simplement indiqués, qui conduisent de la vue atomistique aux lois du mouvement des* *continuums*», voir [166, 151]. Dans cette perspective, les taux de convergence à l'échelle cinétique devraient être cohérents avec ceux que l'on retrouve dans la limite de diffusion. Avec la motivation de contribuer à cette théorie, nous fournissons de nouvelles techniques pour construire des estimations quantitatives pour les deux niveaux de description, cinétique et macroscopique.

La deuxième partie de cette thèse étudie en effet les dynamiques de diffusion (parmi lesquels nous trouvons l'équation de la chaleur) sur la sphère et dans l'espace gaussien. Le but est de dériver des propriétés fines des solutions des équations de diffusion et de les relier aux inégalités fonctionnelles.

Certaines inégalités fonctionnelles sont reliées aux EDP de diffusion, voir [128, 63], par la *méthode de l'entropie*. Cela signifie que ces inégalités fonctionnelles relient une *entropie* et une *information de Fisher*, dont la différence - dans les bons cas seulement - diminue le long du flot d'une EDP d'évolution appropriée. Un pas en avant consiste à trouver des fonctions optimales (et donc aussi la constante optimale) dans l'inégalité fonctionnelle. Du point de vue de l'EDP, cela revient à décrire le comportement en temps long des solutions. Dans les bons cas, ápres avoir recalé les variables, les fonctions optimales deviendront des équilibres et les solutions convergeront vers eux, avec des taux déterminés par la constante optimale dans l'inégalité fonctionnelle. Cette approche remonte au moins à [24], avec certaines idées déjà présentes dans [220, 106], et elle a été réinterprétée dans un langage plus proche des EDP dans [19].

À ce stade, on peut se demander si la vitesse de convergence est plus élevée pour des donneés initiales bien choisies. Une telle amélioration est possible si une inégalité fonctionnelle améliorée est obtenue. Ce type de resultat peut être réecrit comme une estimation de stabilité dans le cas des inégalités de Gagliardo–Nirenberg sur la sphère. En d'autres termes, le *déficit* associé à l'inégalité contrôlerait une distance entre n'importe quelle fonction et l'ensemble des optimiseurs. Si le contrôle est explicite, l'amélioration des taux de convergence l'est également. Les résultats de stabilité sont souvent prouvés en deux étapes. Tout d'abord, une estimation de la stabilité locale est établie, par le biais d'un développement asymptotique autour des fonctions optimales. Ensuite, un argument est nécessaire pour rendre l'estimation globale. Habituellement, cela se fait de manière non constructive par exemple par la méthode de la concentration-compacité, [47, 144, 142]. Récemment, une méthode constructive est apparue, dans laquelle un flot symétrisant est utilisé pour se rapprocher des optimiseurs, [119]. Dans [63], et dans cette thèse, la même chose est faite avec des flots de diffusion non linéaires. Le problème de la stabilité de l'inégalité de Sobolev a été soulevé par Brezis et Lieb dans [73], et une première réponse non constructive est apparue dans [47]. Cependant, des versions antérieures de la stabilité (par exemple, le long des séquences) peuvent être retracées dans des travaux antérieurs, par exemple [199].

Les inégalités fonctionnelles que nous considérons sont des inégalités de Gagliardo–Nirenberg–Sobolev (GNS) sur la sphère, provenant de [240, 147, 222] et explicites dans le cas de la sphere dans [48]. Les inégalités GNS interpolent les inégalités de Poincaré et les inégalités critiques de Sobolev. En outre, nous couvrons une famille d'inégalités sur l'espace gaussien, qui ont l'inégalité de Poincaré gaussienne et l'inégalité de Sobolev logarithmique comme extrémités. Elles ont été introduites par Beckner dans [35]. Les inégalités de Beckner peuvent être vue comme une version en dimension infinie de GNS, comme nous le montrons dans [P3]. Notre résultat est une version pour les inégalités fonctionnelles du lemme dit de Maxwell–Poincaré, *«qui stipule que la limite faible de la séquence des mesures de Lebesgue sur des sphères de dimension finie avec une normalisation appropriée est la mesure gaussienne standard dans un espace de dimension infinie»*, voir pour exemple [252]. Cette affirmation est connue depuis [209], et elle a été telle quelle utilisée très probablement par Maxwell en mécanique statistique, voir [82]. Les calculs avec les harmoniques sphériques sur des sphères de grandes dimensions convergeant formellement vers les polynômes d'Hermite gaussiens sont en effet présents dans la littérature depuis longtemps, ils ont été revisités par McKean [208]. Le lecteur peut consulter [252] pour une discussion historique approfondie.

Enfin, nous considérons l'inégalité de Sobolev logarithmique gaussienne, qui est à la fois le point extreme des inégalités de Beckner et la limite en grande dimension de l'inégalité de Sobolev sur la sphère. Elle trouve son origine en théorie de l'information, avec les travaux de Shannon [239], et a été explicitement exposée dans [155, 261]. Voir [53, 241, 253] pour plus de détails et [12, 90] pour un compte-rendu historique détaillé. L'inégalité de Sobolev logarithmique est également étroitement liée aux problèmes de géométrie convexe et à l'isopérimétrie dans l'espace gaussien, voire [64, 214, 135, 143]. Notre contribution est de prouver les résultats de stabilité en information de Fisher relative, avec une approche basée sur le flot de diffusion d'Ornstein–Uhlenbeck et la méthode du carré du champ.

Les équations de diffusion ont un sens dans des cadres beaucoup plus abstraits que les varietés ou les espaces euclidiens à poids. On peut considérer le flot de chaleur (ou de diffusion non linéaires) dans des espaces de mesure métriques à la suite de [8, 238, 94], où une structure différentielle a été établie, [149], sous conditions pour la courbure de Ricci. Cependant, dans la dérniere partie de ce travail, notre choix est d'étudier des semigroupes de diffusion non linéaires sur des espaces sans structure métrique ou topologique. Les équations d'évolutions correspondentes généralisent les semi-groupes de Markov linéaires et la théorie des formes de Dirichlet qui leur est associé. Les formes de Dirichlet sont des fonctionnelles bilinéaires satisfaisant une propriété de contraction, voire [46]. De nos jours, les formes de Dirichlet sont un sujet bien établi, lié à la théorie des processus de Markov linéaires et des semi-groupes, [70, 146, 203], et aux équations de diffusion linéaires [24, 128]. Parmi les différentes extensions non linéaires, nous suivons celle de [99], qui contient comme cas particuliers les semi-normes de Sobolev non quadratiques, liées aux géométries de Finsler [223]. Nous étendons les propriétés de contraction qui sont valables dans le cadre bilinéaire aux formes de Dirichlet non linéaires, et nous les caractérisons en matière d'inégalités fonctionnelles uniquement. Nos résultats constituent une contribution à la classification des formes de Dirichlet non linéaires et des semi-groupes de Markov non linéaires.

English introduction

This thesis deals with partial differential equations describing the time evolution of systems of many particles, which are subject to the laws of mechanics and mutual interactions. Every interaction is reversible in time, but irreversibility arises in the system as a whole. Such phenomenon, yielding a direction of time, is measured via an entropy. Entropy increases along time, driving the system to the states of maximal disorder. Those states are stable, hence called equilibria. We quantify the speed at which equilibria are approached. To this extent, we find adapted entropies and establish an inequality between the increment of the entropy and the entropy itself. When the inequality is optimal, the sharp rate of convergence to equilibrium is found. Further improvements in the inequalities capture finer properties of the time evolution. These ideas are applied to various kinetic and diffusion equations.

We consider complex systems composed by a large numbers of particles or agents interacting together and with an ambient space, whose physical scale is assumed much larger than that of a typical particle. These systems arise in a variety of natural phenomena, ranging from the dynamics of molecules in a gas or a population of bacteria to the interactions between stars inside a galaxy, and between the galaxies themselves, passing through flock of birds, shoal of fish, and socio-economic behaviours. These systems, and the time evolution of those, can be investigated at three different levels of detail, usually referred to as the microscopic level, the mesoscopic (or kinetic) level, and the macroscopic (or fluid dynamics) level. We refer to [254] for a much broader discussion of the subject.

The microscopic level is the most detailed one. There, the trajectory of every particle is tracked thoroughly. To this extent, all interactions between particles are taken into account. Individually, each particle obeys elementary laws of physics, as Newton's laws or Hamilton's least action principle. In addition, we have interactions between the different particles of the system, which can be long-range, as it happens for electric charges and gravity, or short-range. The latter are usually modeled as collisions, ruled by a microscopic interaction law. We highlight that at this level all dynamics are reversible in time. The knowledge of the phenomenon is complete, but it is never accessible in practice. Indeed, in systems of physical interest, the typical magnitude is 10^{9-10} particles for biological phenomena and for stellar dynamics, and 10^{23} – in the order of Avogadro's number – for gasses. Then, tracking all particles and interactions is unfeasible. Therefore, another level of description shall be considered, which is more treatable, at the cost of losing some information.

This is the case for the mesoscopic or kinetic level, where the object of study is the distribution function of particles with respect to two variables, typically position and velocity, and the time evolution of it. The history of every single particle is not reconstructed, and the phenomenon is described as the evolution of a probability density over a phase space. A *kinetic equation* is then written for the distribution function. This goes back to Maxwell, see [205].

The passage from the microscopic to the mesoscopic scale is usually done by taking a limit for the number N of particles in the system going to infinity, known as the *thermodynamic limit*. A choice shall be made, resulting in different kinetic equations originating from the same system of particles. For example, one can focus on weak, long-range interactions, by taking a mean-field limit, and obtain a Vlasov equation, see [163, 174]. On the other hand, Boltzmann-Grad limits can be considered, in order to capture intense and short-range interactions, [153]. This way one derives Boltzmann's equation from microscopic dynamics. In this thesis we will focus on kinetic equations of Boltzmann type, with simplified collision kernels.

In [62], the thermodynamic entropy is proved to increase along solutions to the Boltzmann equation, as the time progresses. Irreversibility appears, while all dynamics at the microscopic level are reversible. One cause of this phenomenon can be found in the Boltzmann–Grad, where re-collisions of particles are neglected. Via the monotonicity of the entropy along the evolution, Boltzmann's equation explains the irreversibility of large systems in a statistical description. The analysis of [62] yields another outcome. The entropy production along the equation is explicitly computed, and it is proved to vanish if and only if the distribution function is a Maxwellian (or Gaussian) distribution. The last point, combined with the monotonicity of the entropy, proves that solutions to Boltzmann's equation converge to the Maxwellian for long times. The states of maximal chaos (measured via the entropy) are stable and they attract all solutions in the limit $t \to \infty$, whatever the initial conditions. The problem of deriving quantitative estimates of the rate of convergence to equilibria (i.e. Maxwellians) is a major research subject in kinetic theory.

By focusing on a special regime of Boltzmann's equation, corresponding to grazing collisions, one can find Landau's equation. This thesis deals with a linear simplification of Landau's equation: the Vlasov–Fokker–Planck equation. The Vlasov–Fokker–Planck (VFP) or kinetic Fokker–Planck equation is written as a partial differential equation on the distribution function of space and velocity of particles, driven by a diffusion-type collision operator in velocity variables only and a transport operator, linking spatial and velocity variables. The VFP equation has an independent interest, aside its derivation from Boltzmann's equation. Firstly, its solutions are the law of a random process called *Langevin dynamics*, [44], which is a mesoscopic description of a system of particles subject to a random background force. Secondly, the VFP equation is well-suited for describing gas in plasma form, and stellar dynamics, [28, 92]. Some generalisations of it are also present in statistics, molecular simulations, and physical models [243, 152, 44]. Finally, some recent socio-economic applications are collected in [228].

Likewise the Boltzmann equation, solutions to the VFP equation converge to Maxwellian functions for long times. The difference is that the VFP equation is much simpler, and convergence rates can be quantified with different techniques. In the simplest case, a Green function is available and solutions are fully explicit [183], allowing for direct computations. More in general, the state of the art is using a technique developed in the last twenty years, going under the name of *hypocoercivity*. We give a brief description of it, in the specific case of the VFP equation. Computing convergence rates for the VFP equation is not as easy as it would be for a parabolic PDE, as the VFP equation is set in the phase space of positions and velocities, but relaxation is induced by a diffusion operator in velocity variables only. Spatial variables are affected indirectly by diffusion, via the transport operator which links them to velocity variables. The diffusion operator has a nontrivial kernel, whose elements are called *local equilibria*. However, the distribution function cannot lie in the space of local equilibria for long, as the transport operator would push it away, turning on again the dissipation induced by the diffusion operator. This combined effect carries on along solutions the VFP equation until global equilibrium is reached, which means that the distribution lies in the kernel both of transport and diffusion operators. Hence, it is forced to be a Maxwellian.

The generator of the VFP equation belongs to a class of degenerate diffusion operators of the form $A^\star A + B$, where A, B are first-order differential operators, and $A^\star A$ is not uniformly elliptic. These have been analysed in [169]. There, they main point is studying regularity properties (in order to prove *hypoellipticity*).

The intuition of taking inspiration from hypoellipticity for convergence rates is used in [255, 221, 165], after [169, 204, 162]. This theory was dubbed *hypocoercivity* by Gallay, and improves on previous results for the long-time behaviour of kinetic PDEs, which started appearing in the 90s, by quantifying convergence rates. Due to the fact that the diffusion operator in the VFP equation is degenerate, it will not be coercive in standard norms as the H^1 norm. Hypocoercivity techniques solve the issue by considering equivalent twisted norms which better capture the combined effect of diffusion and transport. In the modified norm decay rates are computed. Then, those are transferred to the reference norm by equivalence. The same idea has been used for the L^2 norm in [127]. The framework of this paper has the additional advantage of being consistent with *diffusion limits* from the kinetic to the fluid dynamics scale.

The kinetic level of description is actually an intermediate one, and a macroscopic scale can be introduced. The macroscopic level corresponds to fluid dynamics. Albeit less detailed than the kinetic scale, fluid dynamics is observable and convenient for numerical simulations, and applications to physics and engineering. In the passage between the two scales, the mean free path of particles is pushed to zero, while the average number of interactions for a particle in a unit of time is sent to infinity. Velocity variables converge to equilibrium at a much faster time scale than the spatial variables, so the evolution is described in the limit by the spatial marginal only, called *density function*. See [59, 163] for rigorous convergence results in the context of Boltzmann's or Vlasov's equations, respectively. In the case of the VFP equation, the diffusion limit yields the heat equation as a first-order correction, [127].

Contribution in this direction point towards Hilbert's sixth problem: *«Thus Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, there merely indicated, which lead from the atomistic view to the laws of motion of continua»*, see [166, 151]. To this extent, convergence rates in the kinetic scale should be consistent with those recovered in the diffusion limit. With the motivation to contribute to this theory, we provide new techniques to construct quantitative estimates for both levels of description.

The second part of this thesis indeed studies diffusion flows (among which we find the heat equation) on the sphere and the Gaussian space. The goal is to derive fine properties of solutions to diffusion equations and to apply them to functional inequalities.

Some functional inequalities are linked with diffusion PDEs, see [128, 63], through the *entropy method*. This means that the two sides of a functional inequality are an *entropy* and a *Fisher information*, whose difference – in good cases only – is decreasing along the flow of a certain evolution PDE. One step forward is finding optimal functions (hence the optimal constant) in the functional inequality. On the PDE side, this corresponds to the long-time behaviour of solutions. In good cases, after rescaling the variables, optimal functions will become equilibria and solutions will converge towards them, with sharp rates determined by the optimal constant in the functional inequality. This approach goes back to [24] at least, with some ideas already present in [220, 106], and it has been reinterpreted in a language closer to PDEs in [19].

At this point, one can ask whether the speed of convergence is higher for well-chosen initial data. For GNS inequalities on the sphere, this would correspond to an improved inequality, which can be rewritten as a *stability result*. The *deficit* associated with the inequality would control a distance between any function and the set of the optimisers. If the control is explicit, also the improvement over the convergence rates is explicit. Stability results are often proved in two steps. First, a local stability estimate is established, via an asymptotic expansion around optimisers. Then, an argument is needed to make the estimate global. Usually this is non-constructively done with concentration-compactness, [47, 144, 142]. Recently, a constructive method appeared, where a symmetrising flow is used to get close to optimisers, [119]. In [63], and in this thesis, the same is done with nonlinear diffusion flows. The problem of stability for Sobolev's inequality has been raised by Brezis and Lieb in [73], and a first non-constructive answer appeared in [47]. However, earlier versions of stability (for instance, along sequences) can be traced back to previous works on, e.g. , concentration compactness methods [199].

The functional inequalities we consider are Sobolev–Gagliardo–Nirenberg (GNS) inequalities on the sphere, originating from [240, 147, 222] and written in [48]. The GNS inequalities interpolate between the Poincaré and the critical Sobolev inequalities. In addition, we cover a family of inequalities on the Gaussian space, which have the Gaussian Poincaré and the logarithmic Sobolev inequality as the lower and the upper endpoint. These have been studied by Beckner in [35]. Beckners's inequalities are indeed an infinite-dimensional version of the Sobolev–Gagliardo–Nirenberg family, as we show rigorously in [P3]. Our result is a version for functional inequalities and diffusion flows of the so-called Maxwell–Poincaré Lemma, *«which states that the weak limit of the sequence of Lebesgue measures on finite-dimensional spheres with an appropriate normalization is the standard Gaussian measure in an infinite-dimensional space»* [252]. This statement is informally known since [209], and it has been used by Maxwell in statistical mechanics, see [82]. Computations with spherical harmonics on large-dimensional spheres converging formally to the Gaussian Hermite polynomials are present in the literature, and were revisited by McKean [208]. The reader can check [252] for a complete historical discussion.

Finally, we consider the Gaussian logarithmic Sobolev inequality, which is both the critic point of Beckner's inequalities, and the large dimensional limit of Sobolev's inequality on the sphere. Its very origin lies in communication engineering, with the work of Shannon [239], but it has been written out explicitly in [155, 261]. See [53, 241, 253, 249] for more links with information theory, also in view of log-concavity. See also [12, 90] for a detailed historical account. The logarithmic Sobolev inequality is also closely related with convex geometry problems and isoperimetry in the Gaussian space, see [64, 214, 135, 143]. Our contribution is to prove stability results in relative Fisher information, with an approach purely based on the Ornstein–Uhlenbeck diffusion flow.

Diffusion equations make sense in much more abstract settings than manifolds or weighted Euclidean spaces. One can consider the heat flow (or nonlinear diffusion flows) in metric measure spaces after [8, 238, 94], where a differential structure can still be established, [149] under Ricci lower bounds. However, in the third part of this work, our choice is to study nonlinear diffusion semigroups on spaces without any metric or topological structure. These evolutions generalise linear Markov semigroup and the related Dirichlet forms theory. Dirichlet forms are bilinear functionals satisfying a contraction property, [46]. Nowadays, Dirichlet forms are a well-established topic, related to the theory of linear Markov processes and semigroups, [70, 146, 203], and with linear diffusion equations [24, 128]. Among different nonlinear extension, we follow the one of [99], which contains as particular cases the key–examples of nonquadratic Sobolev seminorms, related to Finsler geometries [223]. We extend contraction properties which are valid in the bilinear setting to nonlinear Dirichlet forms, and we characterise them in terms of functional inequalities only. Our results can contribute to classify nonlinear Dirichlet forms and nonlinear Markov semigroups.

Chapter 2

Main results

2.1 Kinetic Equations

This section is devoted to results of Chapters 3-4 about hypocoercivity for kinetic equations.

Our main results are the following. Time averages of the norms of solutions to kinetic Fokker–Planck equations verify a constructive entropy-entropy production estimate along the flow. The entropy-entropy production estimate is based on modified functional inequalities of Lions–Poincaré type, involving weak norms of solutions. Exponential or polynomial decay rates are derived, according to the shape of local equilibria. In the case of algebraic decay rates, extra moments must be controlled via a Lyapunov condition.

2.1.1 On the kinetic Fokker–Planck equation

Let x, v be variables designating *position*, and *velocity*, respectively. Note that the variable v may correspond to actual physical velocity or to a momentum in case of Hamiltonian Langevin dynamics. We call $\mathscr{X}$ the space of all positions x, and $\mathscr{V}$ the space of all velocities v. The product $\mathscr{X} \times \mathscr{V}$ is called *phase space*. Let us take $\mathscr{V} = \mathbb{R}^d$. For the space $\mathscr{X}$ our two main examples are $\mathscr{X} = \mathbb{R}^d$ or $\mathscr{X} = [0, L]^d$, for any $L > 0$, with periodic boundary conditions (i.e. a flat torus of length $L > 0$). We introduce a time variable $t \geq 0$. Our main object of study is functions

$$f = f(t, x, v) : [0, \infty) \times \mathscr{X} \times \mathscr{V} \to \mathbb{R},$$

modeling the time evolution of a distribution of particles with respect to space and velocity. Let $\phi : \mathscr{X} \to \mathbb{R}$ be a smooth function, called *spatial potential*, which represents potential energy. Let $\psi : \mathscr{V} \to \mathbb{R}$, be a smooth function, standing for kinetic energy. The motion of a single particle subject to the action of ϕ, ψ is modeled by Hamilton dynamics. The same dynamics on a distribution of particles are encoded by the transport equation

$$\partial_t f + \nabla_v \psi \cdot \nabla_x f - \nabla_x \phi \cdot \nabla_v f = 0. \tag{2.1}$$

We introduce a diffusion operator on the velocity variable $v \in \mathscr{V}$, corresponding to collisions between particles. In Chapter 1, we motivate our choice of considering the Fokker-Planck operator (as a simplification of Boltzmann's operator) we define below. Let $\gamma(v) = \mathrm{e}^{-\psi(v)}$ and consider the operator

$$L_{\mathrm{FP}}(f) = \nabla_v \cdot \left(\gamma \nabla_v \left(\frac{f}{\gamma} \right) \right).$$

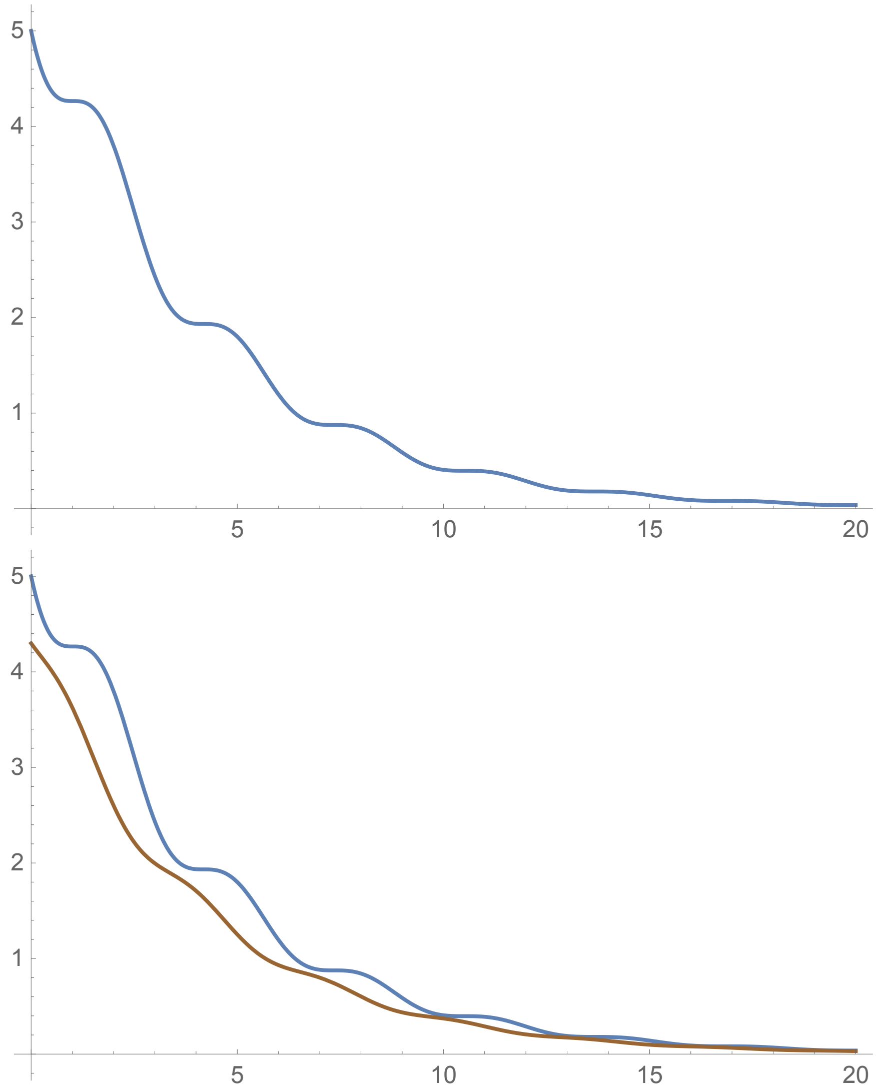

Figure 2.1: Hypocoercivity. Consider the ODE system $x' = v$, $v' = -x - \frac{1}{4}v$, which is a classical *toy model*. This system combines a symmetric, dissipative operator on the variable v, with a skew-symmetric operator in x, v, and these operators are a caricature of the collision and the transport operators of (VFP), respectively. On the top: decay in time of $x^2 + v^2$ along a solution to the system. Note the *local equilibrium* points, where the slope is zero. These correspond to points where the dissipative term in the system vanishes. On the bottom: decay in time of $x^2 + v^2$ (in blue) *vs.* decay of its time average (in brown). Some dissipation is present even at the points of *local equilibrium*.

Let $f_0 = f_0(x,v)$ be any distribution of particles subject to potential energy, kinetic energy and collisions (modeled with the Fokker-Planck operator). Then, the distribution function of particles evolves in time according to the kinetic Fokker-Planck equation

$$\begin{cases} \partial_t f + \nabla_v \psi \cdot \nabla_x f - \nabla_x \phi \cdot \nabla_v f = \nabla_v \cdot \left(\gamma \nabla_v \left(\frac{f}{\gamma} \right) \right), \\ f(t=0,x,v) = f_0(x,v). \end{cases} \tag{2.2}$$

Chapters 3 and 4 deal with the long-time behaviour of solutions to (2.2). Note that the well posedness of the equation is classical. See [169] for regularisation properties of the equation. We assume that the total *mass* of the system is finite, which means that

$$\iint_{\mathscr{X}\times\mathscr{V}} f_0(x,v)\,dx\,dv \in \mathbb{R}.$$

Without loss of generality, since (2.2) is linear and mass-preserving, we will consider only solutions such that

$$\iint_{\mathscr{X}\times\mathscr{V}} f(t,x,v)\,dx\,dv = \iint_{\mathscr{X}\times\mathscr{V}} f_0(x,v)\,dx\,dv = 1.$$

2.1.2 The kinetic Fokker-Planck equation on the torus

We present the results contained in Chapter 3. Let $\mathscr{X} = [0,L]^d$, and let $\phi = 0$. Let $v \in \mathbb{R}^d$. Consider $\psi = \frac{1}{2}|v|^2$ so that the kinetic energy of the system is the standard one, [20]. With this choice we have that

$$\gamma(v) = \mathrm{e}^{-\frac{|v|^2}{2}},$$

the standard Gaussian. Moreover, (2.2) recasts into

$$\partial_t f + v \cdot \nabla_x f = \nabla_v \cdot (\nabla_v f + v f). \tag{2.3}$$

One fact is that $f = \gamma$ is a stationary solution to (2.3), and indeed it is the only one. Our main goal is to prove that γ is also a global attractive equilibrium. In other words, solutions to (2.3), for well-prepared initial data, are such that

$$f(t,\cdot,\cdot) \to \gamma, \qquad \text{as} \quad t \to \infty.$$

We also aim at making such limit quantitative in a proper functional norm. One standard choice is to consider

$$\|g\|^2_{\mathrm{L}^2(dx\gamma^{-1})} = \int_{\mathscr{X}\times\mathscr{V}} g^2(x,v)\gamma^{-1}(v)\,dx\,dv.$$

Hence, the term we investigate is

$$\|f(t,\cdot,\cdot) - \gamma\|^2_{\mathrm{L}^2(dx\gamma^{-1})}, \qquad \text{for } t \to \infty.$$

Computing the evolution of such norm along (2.3) we obtain

$$\frac{d}{dt}\|f(t,\cdot,\cdot) - \gamma\|^2_{\mathrm{L}^2(dx\gamma^{-1})} = -2\int_{\mathscr{X}\times\mathscr{V}} |\nabla_v (f\gamma^{-1}\,dv)|^2 \gamma\,dx\,dv. \tag{2.4}$$

We see that $\|f(t,\cdot,\cdot) - \gamma\|^2_{\mathrm{L}^2(dx\gamma^{-1})}$ is non-expansive along (2.3). However, the right-hand side in (2.4) is zero whenever $f = g(x)\gamma(v)$, such functions being called local equilibria, for any function $g(x)$. Then, at this level we cannot derive a decay estimate for the L^2-norm. This led the authors in [127] to twist the usual L^2–norm, to obtain an equivalent one with a better (i.e. non-degenerate) time derivative along (2.3). This idea was used for H^1–norms in [256] and [215].

We adopt the point of view of [5], that is, we study a time average of the usual norm. The heuristics is that the right-hand side of (2.4) cannot be degenerate too often, so *on average* it must be strictly negative. Therefore, on average, some entropy-entropy production estimate should hold. Let $\tau > 0$ be small enough and consider

$$\mathscr{H}_\tau(t) = \int_0^\tau \|f(t+s,\cdot,\cdot) - \gamma\|^2_{L^2(dx\gamma^{-1}dv)} \frac{ds}{\tau}.$$

The main result of [P1] is the following.

Theorem 1. *For any $\tau > 0$, small enough, there exists an explicit and constructive constant $\lambda(\tau)$ such that, for all solutions to (2.3) with $f_0 \in L^2(dx\gamma^{-1})$, we have*

1. *for all $t \geq 0$,*

$$\frac{d}{dt}\mathscr{H}_\tau(t) = -2\int_0^\tau \|\nabla_v(f(t+s,\cdot,\cdot)\gamma^{-1})\|^2_{L^2(dx,\gamma dv)}\tau^{-1}\,ds \leq -\lambda(\tau)\mathscr{H}_\tau(t),$$

2. *for all $t \geq 0$,*

$$\mathscr{H}_\tau(t) \leq \|f_0 - \gamma\|^2_{L^2(dx\gamma^{-1}dv)} e^{-\lambda(\tau)t}.$$

Moreover,

$$\frac{1}{\lambda} = \frac{1}{\tau}\left(\tau + \sqrt{d L^2 + \tau^2}\right)c,$$

for an explicitly determined constant c depending only on the dimension d.

This results makes the strategy of [5] constructive, tracking the dependence of the constants on all parameters of the model. As a corollary, we have that

$$\|f(t,\cdot,\cdot) - \gamma\|^2_{L^2(dx\gamma^{-1})} \leq e^{\lambda(\tau)\tau} e^{-\lambda(\tau)t} \|f_0 - \gamma\|^2_{L^2(dx\gamma^{-1})}. \tag{2.5}$$

Let us comment that the estimate of Theorem 1 implies usual hypocoercivity estimates, which – like (2.5) – are of the form

$$\|f(t,\cdot,\cdot) - \gamma\|^2_{L^2(dx\gamma^{-1}dv)} \leq C e^{-\lambda t} \|f_0 - \gamma\|^2_{L^2(dx\gamma^{-1}dv)},$$

with a constant $C > 1$ as a pre-factor. We have the parameters λ, τ, while in more standard hypocoercive approaches λ and C cannot be dealt with at once. In particular, having good values of both is complicated (see the discussion in [4]). In a special case: $d = 1$ and $L = 2\pi$ we see in Table 3.1 that the value of λ we determine is off by an order 5 from the actual rate, although our method is not tailored on the particular equation we consider.

The hardest step in our method – which is proved in a simpler and more quantitative way compared to [5, 79] – is the point 1. of Theorem 1. Heuristically, it means that the average in time of the velocity gradient controls the average in time of the solution f to (2.3). Of course, this statement cannot hold true point-wise in time, which is consistent with the degeneracy of $\lambda(\tau)$ as τ becomes very small. Let us rewrite it as

$$\tau^{-1}\int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{L^2(dx,\gamma dv)}\,ds \geq \frac{\lambda(\tau)}{2}\tau^{-1}\int_0^\tau \|h(t+s,\cdot,\cdot)\|^2_{L^2(dx,\gamma dv)}\,ds, \tag{2.6}$$

where $h = f/\gamma$. Denote with ds the uniform measure on $[0,\tau]$. Moreover, let

$$\Pi f = \gamma(v) \int_{\mathscr{V}} f(\cdot,\cdot,v)\, dv,$$

be the local projection of f over the space of local equilibria of (2.3).

Estimates as (2.6) are achieved after three functional inequalities, which *do not* use the fact that f is a solution to (2.3).

1. A Poincaré–Lions inequality in time and space variables:

$$\|g\|^2_{\mathrm{H}^{-1}(ds\,dx)} \lesssim \|\nabla_{t,x} g\|^2_{\mathrm{H}^{-1}(ds\,dx)} \tag{2.7}$$

   for all smooth functions g depending on s, x only, with zero average in $ds\,dx$. Here in particular we take $g = \Pi f - \int_0^\tau \int_{\mathscr{X}} \Pi f\, ds\,dx$.

2. A time-space-velocity averaging lemma , [13]:

$$\|\nabla_{t,x}\Pi h\|^2_{\mathrm{H}^{-1}(ds\,dx)} \lesssim \|h-\Pi h\|^2_{\mathrm{L}^2(ds\,dx\,\gamma\,dv)} + \|(\partial_t + v\cdot\nabla_x)h\|^2_{\mathrm{L}^2(ds\,dx;\mathrm{H}^{-1}(\gamma\,dv))}. \tag{2.8}$$

3. A Poincaré inequality in velocity only:

$$\|h-\Pi h\|^2_{\mathrm{L}^2(\gamma\,dv)} \lesssim \|\nabla_v h\|^2_{\mathrm{L}^2(\gamma\,dv)}. \tag{2.9}$$

Then, combining these estimates, and using the equation (2.3), one arrives to (2.6). Note that (2.9) is the Gaussian Poincaré inequality and it is sharp, see for instance [219].

Indeed, this proof does not use the fact that γ is *actually* the Gaussian. The same can be done for any local equilibrium γ which decays fast enough at infinity to admit (2.9). To this extent, in [P1] we studied also (2.3) in the same setting with

$$\psi(v) = \left(\sqrt{1+|v|^2}\right)^{\alpha}, \qquad \alpha \geq 1.$$

Then, the following holds. We always indicate with $\gamma := \mathrm{e}^{-\psi}$, which is not necessarily a Maxwellian anymore.

Theorem 2. *Let $\alpha \geq 1$ and fix $\psi = (\sqrt{1+|v|^2})^{\alpha}$. For any $\tau > 0$, small enough, there exists an explicit and constructive constant $\lambda(\tau,\alpha)$ such that, for all solutions to (2.2) with $f_0 \in \mathrm{L}^2(dx\,\gamma^{-1})$, we have*

1. *for all $t \geq 0$,*

$$\frac{d}{dt}\mathscr{H}_\tau(t) = -2\int_0^\tau \|\nabla_v(f(t+s,\cdot,\cdot)\gamma^{-1})\|^2_{\mathrm{L}^2(dx,\gamma\,dv)}\tau^{-1}\, ds \leq -\lambda(\tau,\alpha)\,\mathscr{H}_\tau(t),$$

2. *for all $t \geq 0$,*

$$\mathscr{H}_\tau(t) \leq \|f_0 - \gamma\|^2_{\mathrm{L}^2(dx\,\gamma^{-1}\,dv)}\, \mathrm{e}^{-\lambda(\tau,\alpha)\,t}.$$

Moreover,

$$\frac{1}{\lambda} = \frac{1}{\tau}\left(\tau + \sqrt{d\,L^2+\tau^2}\right) c,$$

for an explicitly determined constant c depending only on d, α.

A natural question is the following: what happens if $0 < \alpha < 1$? This problem has been considered in [66] for example.

The strategy described above works except for step (2.9). Indeed, this type of inequality holds for kinetic energies of the form $\psi(v) = (\sqrt{1+|v|^2})^\alpha$ exactly if and only if $\alpha \geq 1$. However, we can replace (2.9) with the weighted inequality

$$\|h - \Pi h\|^2_{\mathrm{L}^2(\gamma dv)} \lesssim \|\nabla_v h\|^{\frac{1}{1+\sigma}}_{\mathrm{L}^2(\gamma dv)} \left(\int_{\mathscr{V}} (1+|v|^2)^{(1-\alpha)\sigma} h^2 \gamma(dv) \right)^{\frac{\sigma}{1+\sigma}}, \tag{2.10}$$

for any $\sigma > 0$. The price to pay is twofold. First, we have to bound the term

$$\int_{\mathscr{V}} (1+|v|^2)^{(1-\alpha)\sigma} f^2 \gamma^{-1}(dv)$$

uniformly in time along the flow of (2.2), which is possible after [66]. Secondly, the type of control we achieve is weaker than that of Theorem 2. The entropy-entropy production inequality of Theorem 2, triggering Gronwall's lemma, is replaced by a nonlinear differential inequality. Hence, algebraic decay rates are recovered as shown.

Theorem 3. *Let $\alpha \in (0,1)$ and fix $\psi = (\sqrt{1+|v|^2})^\alpha$. For any $\tau > 0$, small enough, and all $\sigma > 0$ there exists an explicit positive and increasing function $\Xi : \mathbb{R} \to \mathbb{R}$ with $\Xi^{-1}(y) \approx y + y^{\frac{\sigma}{\sigma+1}}$, such that, for all solutions to (2.2) with $f_0 \in \mathrm{L}^2(dx\gamma^{-1})$, we have*

1. *for all $t \geq 0$,*
$$\frac{d}{dt}\mathscr{H}_\tau(t) = -2\int_0^\tau \|\nabla_v (f(t+s,\cdot,\cdot)\gamma^{-1})\|^2_{\mathrm{L}^2(dx,\gamma)} \tau^{-1}\, ds \leq -\Xi(\mathscr{H}_\tau(t)),$$
2. *for all $t \geq 0$,*
$$\mathscr{H}_\tau(t) \leq c(1+t)^{-\frac{\sigma}{2(1-\alpha)}} \int_{\mathscr{X}} \int_{\mathscr{V}} (1+|v|^2)^{(1-\alpha)\sigma} f_0^2 \gamma^{-1}(dv)\, dx.$$

for an explicitly determined constant c depending only on d, α.

In this result, as it is common for kinetic equations with sub-exponential local equilibria, we assess the phenomenon of momentum loss. This means that to control a momentum of the solution to (2.2), we shall suppose that the initial datum admits some momentum of higher order. Here we need a momentum of order $2(1-\alpha)\sigma$ for f_0 to ensure a time control of the L^2–norm of the corresponding solution to (2.2).

Note that [P1] is the first time where the strategy of [5] is applied to non-Maxwellian local equilibria. Our result relates with [66], where $\mathscr{X} = \mathbb{R}^d$. Our setting is compact in x, and this allows a precise control on the structure of the constant c in Theorem 3. The strategy is indeed different, as we use time averages instead of twisting the L^2–norm. Finally, in [P1], we notice that the threshold at $\alpha = 1$ is continuous between Theorem 2 and Theorem 3. So that, the estimates of Theorem 3 for $\alpha \to 1^-$ converge to those of Theorem 2 at $\alpha = 1$.

An alternative to (2.10) is to consider L^∞ solutions to (2.2) and apply a weak Poincaré inequality as in [158]. There, exponential decay rates are recovered. Since we enlarge the space of initial data (having weaker regularity assumptions), we find a slower decay rate.

A possible question one could deal with is the meaning one can give (and the subsequent theory one finds) to $\alpha = 0$. This is the case of ψ growing logarithmically at infinity, and it is covered by the results of the next section.

2.1.3 General kinetic Fokker-Planck equations

We explain the results of Chapter 4. In this section let $\mathscr{X} = \mathbb{R}^d$ or $[0, L]$ with periodic boundary conditions. Unlike [P1], in [P6], we consider a potential ϕ which is nonzero. Indeed, our works [P6] is the first one where constructive decay rates are given for (2.2) with general potentials ϕ, ψ. Some structural assumptions on ϕ are required, let us sketch them and refer to the paper for details. We denote with $\mu = \mathrm{e}^{-\phi(x)}$, and we allow a slight confusion between the density μ and the measure $\mu(dx)$.

Assumption 1. *Let $\mu(dx)$ be a finite measure. In addition, let*

$$\liminf_{|x|\to\infty} \frac{\phi(x)}{|x|} > 0.$$

Our model case is

$$\phi(x) = \left(\sqrt{1+|x|^2}\right)^p,$$

for $p \geq 1$. On the other hand, we also have some requirements on ψ. As those are complicated, we refer the reader to the corresponding assumptions in [P6]. For the sake of presentation, we consider three concrete cases.

- **Case A.** The function ψ is such that (2.9) holds true. As we have seen, a typical choice is

$$\psi(v) = (\sqrt{1+|v|^2})^\alpha, \qquad \alpha \geq 1.$$

- **Case B.**

$$\psi(v) = (\sqrt{1+|v|^2})^\alpha, \qquad 0 < \alpha < 1.$$

- **Case C.**

$$\psi(v) = \beta \log(\sqrt{1+|v|^2}), \qquad d+2 < \beta.$$

Let us define

$$\Theta = \mu(x)\gamma(v) = \mathrm{e}^{-(\phi+\psi)}.$$

Our goal is to show that solutions to (2.2) with unitary mass converge to Θ for $t \to \infty$. Here we perform a convenient change of variable, taking $h = f\,\Theta^{-1}$. Solutions to (2.2) are in one-to-one correspondence with solutions to the following kinetic Ornstein-Uhlenbeck equation

$$\begin{cases} \partial_t h + \nabla_v \psi \cdot \nabla_x h - \nabla_x \phi \cdot \nabla_v h = \Delta_v h - \nabla_v \psi \cdot \nabla_v h, \\ h(t=0,\cdot,\cdot) = h_0 \in \mathrm{L}^2(\Theta), \end{cases} \tag{VOU}$$

In this context, the role of Θ is played by the constant 1. Let us denote by

$$Th := \nabla_v \psi \cdot \nabla_x h - \nabla_x \phi \cdot \nabla_v h.$$

Following the same strategy as in the previous section, we are able to construct explicit decay rates for solutions to (VOU). However, none of the three steps (2.7)-(2.8)-(2.9) is actually easy. The first result of [P6] is a version of (2.8), which goes as follows.

Theorem 4. *Under suitable hypotheses on ϕ, ψ there exists a constructive constant $K_{\mathrm{avg}}(d, \phi, \psi) \in \mathbb{R}^+$ such that all functions $h : [0,\tau] \times \mathscr{X} \times \mathscr{V} \to \mathbb{R}$ regular enough satisfy*

$$\|\nabla_{t,x} \Pi h\|^2_{\mathrm{H}^{-1}(ds\,\mu\,dx)} \leq K_{\mathrm{avg}} \left(\|h - \Pi h\|^2_{\mathrm{L}^2(ds\,\mu\,dx\,\gamma\,dv)} + \|(\partial_t + T)h\|^2_{\mathrm{L}^2(ds\,\mu\,dx;\mathrm{H}^{-1}(\gamma\,dv))} \right). \tag{2.11}$$

Both this result and its proof are new, up to the best of our knowledge. Moreover, Theorem 42 is a functional inequality, i.e. it does not depend on (VOU).

The second result is a weighted Poincaré–Lions inequality, in the spirit of (2.7). A similar one is given in [84], but it is not suitable for a geometry as that of $[0,\tau]\times\mathscr{X}$. Another related result – implying Theorem 5 – below appears in [117]. However, its proof is not completely constructive and the control on the constant is not detailed. Another version of Theorem 5 is [79, Lemma 2.6]. We improve on such result, with fully transparent dimensional dependence in the constants and less restrictive hypotheses on the potential ϕ.

Theorem 5. *Let Assumption 1 hold true. Let ds be the measure on the interval $[0,\tau]$ for any $\tau>0$. Then, all functions g regular enough verify*

$$\|g\|^2_{\mathrm{L}^2(ds\mu dx)} \leq C(\phi)\left(1+\sqrt{d}\right)\left(\tau^2+\tau^{-2}\right)^2\|\nabla_{t,x}g\|^2_{\mathrm{H}^{-1}(ds\mu dx)}, \tag{2.12}$$

for a dimension-free constant $C(\phi)$ depending only on ϕ in an explicit way.

The proof of Theorem 5 is done in three steps.

1. We prove that (2.12) is equivalent to the following property. For all g regular enough with zero average, there exists a vector field in $d+1$ components $Z:[0,\tau]\times\mathscr{X}\to\mathbb{R}^{d+1}$ such that

$$\begin{cases}(\partial_t,\nabla_x-\nabla_x\phi)\cdot Z=g,\\ \|Z\|^2_{\mathrm{H}^1_0(ds\mu dx)}\lesssim\|g\|^2_{\mathrm{L}^2(\mathrm{ds}\mu\mathrm{dx})}.\end{cases} \tag{2.13}$$

2. We solve

$$-\partial^2_{tt}u-\Delta_x u+\nabla_x\phi\cdot\nabla_x u=g,$$

   with Neumann's boundary conditions, with the help of Lax-Milgram's theorem [72]. We pose

$$W=\nabla_{t,x}u.$$

   A standard regularity theory [44] ensures $W\in\mathrm{H}^1(ds\mu)$.

3. If $W\in\mathrm{H}^1_0(ds\mu)$, we simply pose $Z=W$. Otherwise, we build Z by performing a constructive correction to W, achieved by solving a sort of waves' equation. Hence, Z is not a gradient vector field in general.

The proof is different from the more general one of [117], but it allows a better control on the constant for our choice of the space-time cylinder.

We pass to decay estimates, as the necessary functional inequalities are established.

Case A - exponential decay rates. Let $\gamma=\mathrm{e}^{-\psi}$, be such that (2.9) is satisfied. After Theorems 4 and 5, this case easily follows. Recall

$$\mathscr{H}(t)=\int_0^\tau\|h(t+s,\cdot,\cdot)-1\|^2_{\mathrm{L}^2(\Theta dx dv)}\tau^{-1}ds.$$

Theorem 6. *For any $\tau>0$, small enough, there exists an explicit and constructive constant $\lambda(\tau)$ such that, for all solutions to (VOU) we have*

1. *for all $t\geq 0$,*

$$\frac{d}{dt}\mathscr{H}_\tau(t)=-2\int_0^\tau\|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta dx dv)}\tau^{-1}ds\leq-\lambda(\phi,\psi,d,\tau)\mathscr{H}_\tau(t),$$

2. *for all* $t \geq 0$,
$$\mathcal{H}_\tau(t) \leq \|h_0 - 1\|^2_{\mathrm{L}^2(\Theta\,dx\,dv)}\, \mathrm{e}^{-\lambda(\tau)\,t}.$$

Case B – sub-exponential local equilibria. After Theorems 4 and 5, this case is analogous to Theorem 3, with the attention that we should consider, for all $\sigma > 0$,

$$\|h - \Pi h\|^2_{\mathrm{L}^2(\gamma\,dv)} \lesssim \|\nabla_v h\|^{\frac{1}{1+\sigma}}_{\mathrm{L}^2(\gamma)} \left(\int_{\mathcal{V}} (1+|v|^2)^{(1-\alpha)\sigma} h^2\, \gamma(dv)\right)^{\frac{\sigma}{1+\sigma}}, \tag{2.14}$$

The missing point is that we should uniformly control a momentum

$$\int_{\mathcal{X}} \int_{\mathcal{V}} (1+|v|^2)^{(1-\alpha)\sigma} h^2\, \Theta(dx\,dv)$$

along the flow of (VOU). This was unknown until a personal communication [69], inspired by [66, 78], proving that

$$\int_{\mathcal{X}\times\mathcal{V}} (1+|v|^2)^{(1-\alpha)\sigma} h(t,x,v)^2\, \Theta(dx\,dv) \lesssim \int_{\mathcal{X}\times\mathcal{V}} (1+|x|+|v|)^\eta h_0^2\, \Theta(dx\,dv), \tag{2.15}$$

for all $t \geq 0$ and for an explicit $\eta > \sigma$, depending only on α, ϕ, σ. A simpler option is taking $h_0 \in \mathrm{L}^\infty$ as we do in Chapter 4. Then, our next theorem goes as follows.

Theorem 7. *Let* $1 > \alpha > 0$ *and fix* $\psi = (\sqrt{1+|v|^2})^\alpha$. *For any* $\tau > 0$, *small enough, and all* $\sigma > 0$ *there exists an explicit positive and increasing function* $\Xi : \mathbb{R} \to \mathbb{R}$, *as in Theorem 2, such that, for all solutions to* (VOU), *we have*

1. *for all* $t \geq 0$,
$$\frac{d}{dt}\mathcal{H}_\tau(t) = -2\int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\, \tau^{-1}\, ds \leq -\Xi\left(\mathcal{H}_\tau(t)\right),$$

2. *for all* $t \geq 0$,
$$\mathcal{H}_\tau(t) \leq c\,(1+t)^{-\frac{\sigma}{2(1-\alpha)}} \int_{\mathcal{X}} \int_{\mathcal{V}} (1+|x|+|v|)^\eta h_0^2\, \Theta(dx\,dv),$$

for an explicitly determined constant c.

This results compares with [212]. Our choice is to take the largest possible space for initial data at the cost of having a worse decay rate. On the other hand, the authors in [212] find exponential decay estimates in a space of the form $\mathrm{L}^2\left(\mu\,dx, \mathrm{e}^{\psi^\sigma}\gamma\,dv\right)$.

Case C – fat-tail local equilibria. This case is very similar to the previous one. The only difference is that we

consider, for all $\sigma > 0$,

$$\|h - \Pi h\|^2_{\mathrm{L}^2(\gamma\,dv)} \lesssim \|\nabla_v h\|^{\frac{1}{1+\sigma}}_{\mathrm{L}^2(\gamma)} \left(\int_{\mathcal{V}} (1+|v|^2)^{\sigma} h^2\, \gamma(dv)\right)^{\frac{\sigma}{1+\sigma}}, \tag{2.16}$$

instead of (2.14), as in [54]. The only available option here is taking $h_0 \in \mathrm{L}^\infty$, as we lack an analogous of (2.15). Finally, the decay estimate.

Theorem 8. *Let* $\beta > d+2$ *and fix* $\psi = \beta\log(\sqrt{1+|v|^2})$. *For any* $\tau > 0$, *small enough, and all* $\sigma > 0$ *there exists an explicit positive and increasing function* $\Xi : \mathbb{R} \to \mathbb{R}$ *such that, for all solutions to* (VOU) *with* $h_0 \in \mathrm{L}^\infty(\Theta)$, *we have*

1. *for all* $t \geq 0$,
$$\frac{d}{dt}\mathscr{H}_\tau(t) = -2\int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta\,dx\,dv)}\tau^{-1}\,ds \leq -\Xi(\mathscr{H}_\tau(t)),$$

2. *for all* $t \geq 0$,
$$\mathscr{H}_\tau(t) \leq c\,(1+t)^{-\frac{\sigma}{2(1-\alpha)}}\int_{\mathscr{X}}\int_{\mathscr{V}}(1+\mathrm{e}^{|x|}+|v|)^\eta h_0^2\,\Theta(dx\,dv),$$

for an explicitly determined constant c.

Theorems 7 and 8 are new up to the best of our knowledge, when $\nabla_v\psi \neq v$ [212]. The same is true for the estimates (2.15) by [69], achieved via an argument based on nonlinear Lyapunov conditions.

We complete [P6] with a careful analysis of the constants in decay estimates, with respect to d, τ, and a parameter ξ which is used to tune the intensity of the diffusion operator in (VOU). Those explicit estimates are useful in molecular dynamics and statistics, [243]. We notice that the ideas collected here are not specific for the model of interest and could be adapted to other situations.

2.2 Functional inequalities

This section is devoted to the results of Chapters 5-6-7. We start with a brief introduction of the setting in Section 2.2.1.

Constructive stability estimates, in H^1 norms, with sharp exponents and explicit constants, are derived for Gagliardo–Nirenberg inequalities on the sphere. The proof combines nonlinear diffusion flows to approach the manifold of optimisers, with a local Taylor expansion. Additional care is paid to low-degree spherical harmonics, as those of degree one induce a cancellation in the Taylor expansions. In the large dimensional limit, we find Beckner's inequalities on the Gaussian space. Stability estimates, based on diffusion flows and local Taylor expansions are derived for these Beckner's inequalities. The Gaussian logarithmic Sobolev inequality is the critical case in the large-dimensional limit. Stability in relative Fisher information (defined below) is discussed. The arguments rely on the Ornstein–Uhlenbeck flow and its interplay with log-concavity.

2.2.1 Interpolation inequalities on the sphere and the Gaussian space

In Chapters 5, 6, and 7 we deal with some Sobolev-like functional inequalities. The main problem we consider is their stability, and making the results constructive.

Let $d \geq 3$ (actually $d = 1, 2$ is analogous with a slightly different notation), and let $\mathbb{S}^d$ be the unitary sphere of dimension d embedded in $\mathbb{R}^{d+1}$:

$$\mathbb{S}^d = \{x = (x_1, \cdots, x_{d+1}) : |x| = 1\}.$$

We denote by ω the intrinsic coordinates on the sphere. Let μ_d be the uniform probability measure on $\mathbb{S}^d$. We define also the Laplace-Beltrami operator Δ as the only self-adjoint operator with respect to μ_d such that

$$\int_{\mathbb{S}^d}\nabla f\cdot\nabla g\,d\mu_d = \int_{\mathbb{S}^d} -f\,\Delta g\,d\mu_d,$$

for all regular functions $f, g : \mathbb{S}^d \to \mathbb{R}$. We denote by ∇ the intrinsic gradient on $\mathbb{S}^d$. Finally, let $2^\star = \frac{2d}{d-2}$.

We start with a standard fact, known since [240, 147, 222]. The inclusion

$$\mathrm{H}^1(\mathbb{S}^d, d\mu_d) \hookrightarrow L^p(\mathbb{S}^d, d\mu_d),$$

is continuous if and only if $p \in [1, 2^\star]$, with

$$\mathrm{H}^1(\mathbb{S}^d, d\mu_d) = \left\{F \in \mathrm{L}^2(\mathbb{S}^d, d\mu_d) : \nabla F \in \mathrm{L}^2(\mathbb{S}^d, d\mu_d)^d\right\}.$$

Then, a quantitative way to state such inclusion is the following *Gagliardo-Nirenberg-Sobolev inequalities.*

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d \geq \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d) \tag{GNS}$$

for any $p \in [1,2) \cup (2, 2^*]$. For $p = 2$, (GNS) does not make any sense. However, by taking the derivative with respect to p, we obtain the logarithmic Sobolev inequality for the measure μ_d :

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d \geq \frac{d}{2}\int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d) \tag{LS}$$

In this context, the value $p = 2$ has nothing special. The only distinction is between the critical regime – i.e. $p = 2^\star$ – and the subcritical regime $p \in [1, 2^\star)$, with the attention that for $p = 2$ we have to refer to (LS) instead of (GNS). We also note that, for $p = 1$, (GNS) reduces to Poincaré's inequality.

All inequalities (LS)–(GNS) are given with optimal constants. Hence, one could hope to find extremal functions for them (i.e. functions which saturate the inequality).

- For $p = 2^\star$, (GNS) is equivalent to the Euclidean Sobolev's inequality via stereographic projection. Hence, the family of optimisers is given by the Aubin-Talenti family, [21, 248].

- For $p \in [1, 2^\star)$, the optimisers are constant functions. Note that in this case there is no known correspondence with the Gagliardo-Nirenberg inequalities on the Euclidean space (nor with the Euclidean logarithmic Sobolev inequality for $p = 2$).

Henceforth, we focus on the subcritical regime only, as results in the case $p = 2^\star$ correspond to those already obtained in [47, 119]. Introduce the deficit functional (i.e. the difference between the two sides of the inequality) associated with (GNS):

$$\delta_{p,d}(F) = \int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d - \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right),$$

and its analogous $\delta_{2,d}$ for the logarithmic Sobolev inequality (LS):

$$\delta_{2,d}(F) := \int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d - \frac{d}{2}\int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d.$$

We know [49] that

$$\delta_{p,d}(F) \geq 0, \qquad \delta_{p,d}(F) = 0 \iff F \equiv c, \quad c \in \mathbb{R},$$

for all d, and all $p \in [1, 2^\star)$.

Having a stability result means that the deficit $\delta_{p,d}(F)$ is a good measure of how much F is *not* a constant. In other words, we would like that the deficit functional controls some distance between F and constant functions. Let

$$\mathscr{F}^\star = \{F \in \mathrm{H}^1(\mathbb{S}^d) : F \equiv c,\, c \in \mathbb{R}\}.$$

So, a stability estimate for (GNS)–(LS) would look like

$$\delta_{p,d}(F) \ge C \inf_{F^\star \in \mathscr{F}^\star} \mathrm{d}^\theta(F, F^\star), \tag{2.17}$$

for some constant $C > 0$, some reference distance d (fixed a priori), and some exponent θ. In [P2] we provide estimates as (2.17) with the strongest known distance, explicit constant C, and sharp exponent θ. Note that our strategy differs completely from that of [144, 47], using concentration-compactness arguments [199], where the constant C is never explicit. The natural distance we consider is the homogeneous Sobolev seminorm $\mathrm{d} = \|\nabla F\|_{\mathrm{L}^2(\mathbb{S}^d)}$, and variations of it. Chapter 5 deals with the proofs of constructive estimates of the form (2.17) for (GNS)–(LS).

Let γ be the n−dimensional Gaussian density. It is general understanding, [252], that $(\mathbb{S}^d, \mu_d)$ behaves like a Gaussian measure space $(\mathbb{R}^n, \gamma)$ as $d \to \infty$. Hence, one question would be how (GNS)–(LS) behave as the dimension becomes bigger and bigger, and what happens to the stability estimates of Chapter 5 in the limit. In Chapter 6 we clarify this limit.

Note that $2^\star \to 2$ as $d \to \infty$. Hence, the range $p \in [2, 2^\star]$ collapses onto the point 2 as $d \to \infty$. As we will see, for $p \in [2, 2^\star]$, inequalities (GNS)–(LS) become the Gaussian logarithmic Sobolev inequality, in the limit $d = \infty$:

$$\delta_{2,\infty}(v) = \|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n, d\gamma)} - \frac{1}{2} \int_{\mathbb{R}^n} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n, d\gamma)}}\right) d\gamma \ge 0. \tag{GLS}$$

Then, the logarithmic Sobolev inequality (corresponding to $p = 2$) in the large dimensional limit inherits the role of upper endpoint of the family played by the critical Sobolev inequality ($p = 2^\star$) in finite dimension.

On the other hand, the range $p \in [1, 2)$ makes perfectly sense also in the large dimensional limit. For $p \in [1, 2)$, (GNS) converges to the family of Beckner's inequalities

$$\delta_{p,\infty}(v) = \|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n, d\gamma)} - \frac{1}{2-p}\left(\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n, d\gamma)} - \|v\|^2_{\mathrm{L}^p(\mathbb{R}^n, d\gamma)}\right) \ge 0, \tag{PLS}$$

for $d \to \infty$. Hence, in the large dimensional limit we are left with a limit family of inequalities, parameterised by $p \in [1, 2]$, interpolating between the Gaussian Poincaré inequality, $p = 1$, and the Gaussian logarithmic Sobolev inequality, $p = 2$, which is the upper endpoint of the family. Unlike in finite d, in the large dimensional limit, the logarithmic Sobolev inequality is critical and somewhat distinguished from the range $p \in [1, 2)$. This makes its analysis more complicated, and we put it off until Chapter 7, where we collect our partial results. The range $p \in [1, 2)$ in the large dimensional limit shares many properties with the same range in finite d. In this case, we have a rather complete theory, see [P3], exposed in Chapter 6.

2.2.2 Stability of functional inequalities on the sphere

We give the main results of Chapter 5. A way to prove (GNS)-(LS), after [145], is the Funk-Hecke formula we give an account of below. Recall that the Laplace-Beltrami operator $-\Delta$ on $(\mathbb{S}^d, \mu_d)$ admits a discrete spectrum, given by

$$\lambda_j = j(j + d - 1), \qquad \forall\, j \in \mathbb{N}.$$

The eigenfunctions associated with λ_j are called *spherical harmonics* of degree j, and all of them are polynomials of degree j in the coordinates x. We indicate with $\mathcal{H}_j$ the eigenspace associated with λ_j and with $H_{j,k}$ a basis of eigenfunctions for $\mathcal{H}_j$, for $j \in \mathbb{N}$, and k ranging between 1 and the dimension of $\mathcal{H}_j$, which amounts to $\begin{pmatrix} d+j \\ j \end{pmatrix}$. We have the orthogonal decomposition $\mathrm{L}^2(\mathbb{S}^d, \mu_d) = \bigoplus_j \mathcal{H}_j$. Note that $\mathcal{H}_0$ corresponds to constant functions and $\mathcal{H}_1 = \mathrm{span}(x_1, \cdots, x_{d+1})$. Then, we can write any function $F \in \mathrm{H}^1(\mathbb{S}^d, \mu_d)$ as

$$F = \sum_j F_j, \qquad F_j \in \mathcal{H}_j, \quad j \in \mathbb{N}.$$

The Funk-Hecke formula reads

$$\frac{1}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \le \sum_{j=1}^{\infty} \zeta_j(p) \int_{\mathbb{S}^d} |F_j|^2 \, d\mu_d \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d) \tag{2.18}$$

for any $p \in (1,2) \cup (2, 2^*)$ with

$$\zeta_j(p) := \frac{\gamma_j\left(\frac{d}{p}\right) - 1}{p-2} \quad \text{and} \quad \gamma_j(x) := \frac{\Gamma(x)\Gamma(j+d-x)}{\Gamma(d-x)\Gamma(x+j)}.$$

An analogous formula holds for the logarithmic entropy of (LS):

$$\int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \le 2 \sum_{j=1}^{\infty} \zeta_j(2) \int_{\mathbb{S}^d} |F_j|^2 \, d\mu_d \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d).$$

The fact that

$$\zeta_j(p) \le \frac{\lambda_j}{d}, \qquad \forall\, j \in \mathbb{N},$$

proves (GNS) and (LS). More precisely, we have that

$$\zeta_j(p) < \frac{\lambda_j}{d}, \qquad p \neq 2^\star, \qquad j \neq 1.$$

Exploiting this strict gap, we can improve on (GNS)-(LS), every time that $F_1 = 0$ and $p \neq 2^\star$.

Theorem 9. *Assume that $d \ge 1$ and $p \in (1,2) \cup (2, 2^*)$. For any function $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d)$ such that*

$$\int_{\mathbb{S}^d} x\, F \, d\mu_d = 0\,, \tag{2.19}$$

we have

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \ge \mathscr{C}_{p,d} \int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \tag{2.20}$$

with $\mathscr{C}_{p,d} = \frac{2d - p(d-2)}{2(d+p)}$.

Theorem 10. *Let $d \ge 1$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d)$ such that the orthogonality condition (2.19) holds, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \ge \mathscr{C}_{2,d} \int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \tag{2.21}$$

with $\mathscr{C}_{2,d} = \frac{2}{d+2}$.

If we have additional orthogonality constraints, i.e. $F_j = 0$ for $1 \le j \le k$, we have better constants in

Theorems 9-10.

Estimates (2.20)–(2.21) can be seen as stability results in the spirit of (2.17):

$$\delta_{p,d}(F) \geq \mathscr{C}_{p,d} \, \|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \, .$$

The constants are fully explicit, the distance is the strongest possible one, and the exponent *two* is sharp, as one can easily test by taking $F = 1 + \epsilon(x_1^2 - (d+1))$, as $\epsilon \to 0$.

The next point is describing what happens if a function F has a nonzero component F_1, since the Funk-Hecke formula does not provide any improvement along this component. Rupert Frank in [144] shows that if $p \in (2, 2^\star)$, there exists an unknown constant $\mathrm{c}_\star > 0$ such that

$$\delta_{p,d}(F) \geq \mathrm{c}_\star \, \|\nabla F\|^4_{\mathrm{L}^2(\mathbb{S}^d)} \, \mathrm{N}^{-1}(F), \qquad \forall \, F \in \mathrm{H}^1(\mathbb{S}^d, \mu_d). \tag{2.22}$$

In (2.22), the term $N(F) := \|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ is a normalising factor whose role is to preserve the homogeneity of the inequality. We have that $N(F) \approx 1$ if $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ is small. Moreover, the exponent *four* is sharp, as one can check by taking $F = 1 + \epsilon \, x_1$, as $\epsilon \to 0$. Note that this sequence is not admissible in Theorems 9-10.

To sum up, we know that the exponent *two* in the stability estimate is attained as long as the component F_1 of F, given by

$$F_1 = \sum_{i=1}^{d+1} x_i \, \frac{\int_{\mathbb{S}^d} F(x) \, x_i \, d\mu_d}{(d+1)^2},$$

is zero. However, a concrete example of functions F, such that $F = 1 + F_1$, provides the degeneracy to exponent *four*. So, we conjecture that such kind of functions must be the only responsible for this phenomenon. In other words, we expect the deficit $\delta_{p,d}$ to be flatten along the component of F along the $\mathcal{H}_1$ subspace. Indeed, we can prove the following, which answers completely the question.

Theorem 11. *Let $d \geq 1$ and $p \in (1,2) \cup (2, 2^*)$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{p-2} \left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \right) \geq \mathscr{S}_{d,p} \left(\frac{\|\nabla F_1\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{N(F)} + \|\nabla (F - F_1)\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \right),$$

*for some **constructive** stability constant $\mathscr{S}_{d,p} > 0$.*

Theorem 12. *Let $d \geq 1$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} \right) d\mu_d \geq \mathscr{S}_d \left(\frac{\|\nabla F_1\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{N(F)} + \|\nabla (F - F_1)\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \right),$$

*for some **constructive** stability constant $\mathscr{S}_d > 0$.*

The constants can be tracked along the proofs of the two theorems.

The sequences used above show sharpness of exponents in both components. For completeness, the normalisation factor is given by

$$N(F) = \|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \frac{d}{2} \, \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \, .$$

We give some details of the proofs. As it is classical in stability results, the proof is split into two parts:

1. Local regime. We assume that $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ is small (i.e. F is approximately a constant). Then, a Taylor expansion with explicit uniform control of the remainders provides the desired result. At this level, additional care is necessary to deal with the components of F along low-degree spherical harmonics.

2. Far-away regime. We show that we can always reduce to the local regime via a nonlinear diffusion flow.

The second step replaces usual local-to-global arguments, which are obtained by concentration-compactness as in [47, 144], and are not constructive.

The strategy we adopt is the following. For all $p \in (1, 2^\star)$ and all $d \geq 1$, there exist $0 < m_-(p,d) < m_+(p,d)$, such that, if $m_-(p,d) \leq m \leq m_+(p,d)$, then, along the flow of the nonlinear PDE,

$$\begin{cases} \frac{\partial u}{\partial t} = u^{-p(1-m)} \left(\Delta u + (m\,p-1) \frac{|\nabla u|^2}{u} \right), \\ u_0 = F. \end{cases} \tag{2.23}$$

we have that

$$\frac{d}{dt} \delta_{p,d}(u_t) \leq 0, \qquad \|\nabla u_t\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \xrightarrow{t\to\infty} 0, \tag{2.24}$$

see [113, 121]. We come back to this point in the next section, while here we just stick to what is necessary for the purposes of stability. We highlight that, the sphere being a compact manifold, there is no difficulty in proving global existence of a positive bounded solution. The case of a non-negative solution is also covered by a standard approximation of the initial datum. See for instance [111, Section 5] for a proof in a similar setting. The key of the proof is to combine L^1-contraction and the Maximum Principle as in the classical theory of nonlinear diffusions: see for instance [251] for a reference in this area. We remark that (2.24) provides another proof of (GNS)–(LS) after time integration. Indeed, more careful computations prove that

$$\frac{d}{dt} \delta_p(u_t) \leq -\gamma \beta^2 \int_{\mathbb{S}^d} \frac{|\nabla v_t|^4}{v_t^2} \, d\mu_d, \tag{2.25}$$

where $v = u^{\frac{1}{\beta}}$, and γ, β are explicit positive numbers depending only on d, m. Since (2.25) is an improved version of (2.24), which proves (GNS)–(LS), we expect that (2.25) provides some improved inequalities. This is the case, and, after an interpolation argument based on the analysis of an ODE, one can show that (2.25) entails

$$\delta_{p,d}(F) \geq d\, \|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\, \psi_{p,d}\left(\frac{1}{d} \frac{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}} \right), \tag{2.26}$$

for an explicit convex function $\psi_{p,d} : \mathbb{R} \to \mathbb{R}$ such that $\psi_{p,d}(0) = 0 = \psi'_{p,d}(0)$, and $\psi''(0) > 0$, given in Chapter 5. Equation (2.26) is a sort of nonlinear stability result for (GNS)–(LS). The convexity of $\psi_{p,d}$ proves Theorems 11 and 12 in the far-away regime, as it allows a control of the stability term for $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ big. Finally, the actual shape of $\psi_{p,d}$ gives a constructive estimate of $c_\star$ in (2.22).

2.2.3 Subcritical Gaussian inequalities

Here we present the results of Chapter 6. In particular, we are concerned with the behaviour of estimates in the previous section as $d \to \infty$. See Chapter 1 for some historical comments. The setting is the following. Let n be a positive integer, and let $d \geq n$. Let $F_d : \mathbb{S}^d \to \mathbb{R}$ be a function depending on the first n coordinates only. Let γ be the Gaussian density of dimension n. Then, we prove a limit theorem for the deficit functionals $\delta_{p,d}$ as $d \to \infty$.

Theorem 13. *Let n be a positive integer, $p \in [1,2)$ and consider a function $v \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$ with compact support. For any $d \geq n$, large enough, if*

$$F_d(x) = v\big(x_1/\sqrt{d}, x_2/\sqrt{d}, \ldots, x_n/\sqrt{d}\big)$$

where $x=(x_1,x_2,\dots,x_{d+1})\in\mathbb{R}^{d+1}\supset\mathbb{S}^d$ *is such that* $|x|=1$*, then*

$$\lim_{d\to+\infty} d\left(\|\nabla F_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)}-\frac{d}{2-p}\left(\|F_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)}-\|F_d\|^2_{\mathrm{L}^p(\mathbb{S}^d,d\mu_d)}\right)\right)$$
$$=\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\frac{1}{2-p}\left(\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right).$$

Note that Theorem 13 covers the range $[1,2)$ only and it can be reformulated as

$$\lim_{d\to\infty} d\,\delta_{p,d}(F_d)=\delta_{p,\infty}(v).$$

As we anticipated, the range $p\in[2,2^*]$ for (GNS)–(LS) collapses on the point $p=2$ as $d\to\infty$, corresponding to the logarithmic Sobolev inequality.

Proposition 14. *In the same setting as Theorem 13, we have that*

$$\lim_{d\to+\infty} d\,\delta_{2,d}(F_d)=\delta_{2,\infty}(v).$$

Moreover, if $(p_d)_{d\geq 3}$ *is a sequence of real numbers such that* $p_d\in(1,2)\cup(2,2^*)$*, with* $2^*=2d/(d-2)$*, and* $\lim_{d\to+\infty}p_d=2$*, then*

$$\lim_{d\to+\infty} d\,\delta_{p_d,d}(F_d)=\delta_{2,\infty}(v).$$

We show that this large-dimension limit is compatible with diffusion flows. We introduce the Ornstein–Uhlenbeck operator Δ_γ on $\mathrm{H}^1(\mathbb{R}^n,d\gamma)$ defined by

$$\Delta_\gamma v=\Delta v-\nabla v\cdot x.$$

The Ornstein-Uhlenbeck operator Δ_γ in the large dimensional limit plays the role of the Laplace-Beltrami operator on $\mathbb{S}^d$.

In the previous Section we considered the nonlinear diffusion flow (2.23) on $\mathbb{S}^d$. We give here a more precise statement. Let us define

$$m_\pm(p,d):=\frac{1}{(d+2)\,p}\left(d\,p+2\pm\sqrt{d\,(p-1)\,\big(2\,d-(d-2)\,p\big)}\right). \tag{2.27}$$

Proposition 15 ([113, 121])**.** *Assume that* $d\geq 1$*, with either* $p\in[1,2)\cup(2,+\infty)$ *if* $d=2$ *or* $p\in[1,2)\cup(2,2^*]$ *if* $d\geq 3$*, and let*

$$m\in[m_-(p,d),m_+(p,d)].$$

If $u>0$ *solves* (2.23)*, with an initial datum* F *in* $\mathrm{L}^2\cap\mathrm{L}^p(\mathbb{S}^d,d\mu_d)$*, then*

$$\frac{d}{dt}\left(\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)}-\frac{d}{p-2}\left(\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d,d\mu_d)}-\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)}\right)\right)\leq 0\quad\forall\, t>0,$$
$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)}\to 0,\quad t\to\infty.$$

We established a similar picture in the Gaussian space. The kind of flow we shall consider is

$$\frac{\partial v_t}{\partial t}=v_t^{-p(1-m)}\left(\Delta_\gamma v_t+(m\,p-1)\,\frac{|\nabla v_t|^2}{v_t}\right). \tag{2.28}$$

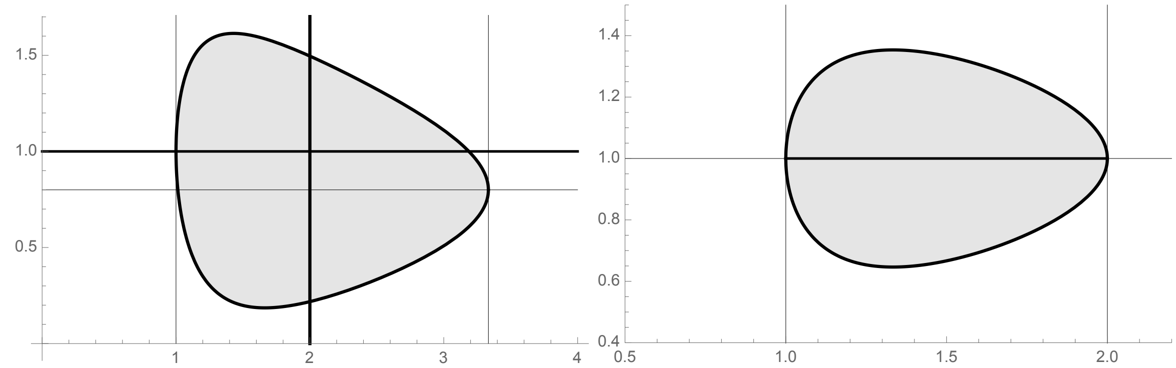

Figure 2.2: Finite VS infinite dimension interpolation inequalities.
On the left: the Gagliardo-Nirenberg-Sobolev range $p \in [1, 2^\star]$ on the horizontal axis, and the exponents of nonlinear diffusion flows $m(p,d)$ which are admissible in Proposition 15 on the vertical axis, for $d = 5$.
On the right: Beckner's Gaussian inequalities, for $p \in [1,2]$ on the horizontal axis, and the admissible exponents $m(p,\infty)$ of nonlinear diffusion flows in the sense of Theorem 16. The figure on the left converges to the figure on the right as $d \to \infty$, while $2^\star = 2d/(d-2) \to 2$.

Then, if we define

$$m_\pm(p,\infty) = 1 \pm \frac{1}{p}\sqrt{(p-1)(2-p)}. \tag{2.29}$$

we have entropy-entropy production estimates as follows.

Theorem 16. *Assume that* $n \geq 1$, $p \in [1,2]$. *If* $v_t > 0$ *solves* (2.28) *with*

$$m \in [m_-(p,\infty), m_+(p,\infty)]$$

for an initial datum in $\mathrm{L}^2 \cap \mathrm{L}^p(\mathbb{R}^n, d\gamma)$, *then*

$$\frac{d}{dt}\delta_{p,\infty}(v_t) \leq 0 \quad \forall\, t > 0,$$
$$\|\nabla v_t\|_{\mathrm{L}^2(\mathbb{R}^n, d\gamma)} \to 0, \quad t \to \infty.$$

The result for $d = \infty$ is new. Finally, one can notice the consistency between a finite d and $d = \infty$, as

$$\lim_{d\to\infty} m_\pm(p,d) = m_\pm(p,\infty),$$

for every $p \in [1,2]$. Note that $m_+(2,\infty) = m_-(2,\infty) = 1$, so only this choice – corresponding to the heat flow acting on v_t^2 – is compatible with (GLS).

As for the case of finite d we can get some nonlinear improvement for (PLS)–(GLS) via the flow (2.28). More precise computations along the flow (2.28) show that

$$\frac{d}{dt}(\delta_{p,\infty}(v_t)) \leq -\kappa^2 \int_{\mathbb{R}^n} \frac{|\nabla w|^4}{w^2}\, d\gamma,$$

for an explicit constant κ. The right-hand-side of the last inequality can be exploited as in the case of the sphere and produce an analogous of (2.26).

Proposition 17. *Let $p \in [1,2]$ and n be a positive integer. For any $v \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$ we have*

$$\delta_{p,\infty}(v) \geq \|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} \psi_p\left(\frac{\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}}\right),$$

for a convex function ψ_p, such that $\psi_p(0) = 0 = \psi_p'(0)$, and $\psi_p''(0) > 0$:

$$\psi_p(s) = s - (2-p)\left(1 + s - (1+s)^{p-1}\right)^{-1}.$$

The last Proposition can be used again in the proof of constructive stability estimates for (PLS).

Let us focus on the subcritical range for $d = \infty$, which is $p \in (1,2)$. The critical case $p = 2$, namely (GLS), is put off to the next section. The very same strategy as in Section 2.2.2, provides us with constructive stability results for the Gaussian interpolation inequalities (PLS). The structure is the same as Theorem 11. However, spherical harmonics are replaced by the eigenfunctions of Δ_γ. Indeed, the eigenvalues of $-\Delta_\gamma$ are given by $\mathbb{N}$. The corresponding eigenfunctions are the well-known Hermite polynomials. More precisely,

$$\Delta_\gamma v = 0 \iff v \equiv c, \qquad -\Delta_\gamma v = v \iff v \in \operatorname{span}(x_1, \cdots, x_n).$$

We start with an improved inequality under orthogonality constraint, which can be also read as a stability result (2.17), with optimal exponent *two*.

Proposition 18. *Let $n \geq 1$ and $p \in [1,2)$. For any $v \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$, such that*

$$\int_{\mathbb{R}^n} v\, x\, d\gamma = 0,$$

we have

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{p-2}\left(\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} - \|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\right) \geq \frac{1}{2}(2-p)\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}.$$

The proof is different from that of Theorem 9, as the spectral improvement of the Funk-Hecke formula (2.18), is replaced by Nelson's hypercontractivity estimates of the Ornstein-Uhlenbeck flow on the Gaussian space.

On the first eigenspace of Δ_γ, which is excluded in Proposition 18, we have the same degeneracy as in (2.22). Let $v_\epsilon := 1 + \epsilon x_1$, then

$$\delta_{p,\infty}(v_\epsilon) \approx \epsilon^4, \quad \epsilon \to 0,$$

for all $p \in (1,2)$. Then, comparing this fact with Proposition 18, as in Section 2.2.2, we prove a global constructive stability estimate. As in Theorem 11, we have two different orders along the component of the function which is parallel to x, and the orthogonal one. Let v_1 be the projection of v on the subspace of $\mathrm{H}^1(\mathbb{R}^n, d\gamma)$ corresponding to $\operatorname{span}(x_1, \cdots, x_n)$.

Theorem 19. *For all $n \geq 1$, and all $p \in (1,2)$, there is an explicit constant $c_{n,p} > 0$ such that, for all $v \in \mathrm{H}^1(d\gamma)$, it holds*

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{p-2}\left(\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} - \|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\right)$$
$$\geq c_{n,p}\left(\|\nabla(v - v_1)\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} + \frac{\|\nabla v_1\|^4_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} + \|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right).$$

The strategy of proof goes as in Theorem 11.

1. Local stability. If we assume v to be approximately a constant, then an explicit Taylor expansion, with uniform control of the remainder terms and careful treatment of low-degree Hermite polynomials does the job.

2. Far-away regime. If v is far away from the manifold of constants, we can approach it using the heat flow, as in Theorem 11. Notice that the choice $m = 1$ is always feasible here. The key-point is the improved inequality contained in Proposition 17, giving stability in the far-away regime by convexity of ψ_p.

The local Taylor expansion and Proposition 17 are not available for the case $p = 2$. Different strategies should then be considered. This is the subject of the next section.

2.2.4 The Gaussian logarithmic Sobolev inequality

We end our survey with (partial) results for the critical case in $d = \infty$, collected in Chapter 7. Stability results for the logarithmic Sobolev inequality are a very active research trend with a huge literature and many applications in various areas. In particular, obtaining stability in H^1–norm for (GLS) is still an open, and very tricky, question. However, with different methods with respect to the subcritical regime, we are able to achieve explicit stability estimates for some special classes of measures. A nice feature of our method is that it is fully flow-based, without any Taylor expansion. The main ingredient is an improvement of entropy-entropy production inequalities along the heat flow. Let $v \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$ be such that

$$\int_{\mathbb{R}^n} (1, x, |x|^2)\, v^2\, d\gamma = (1, 0, d), \tag{2.30}$$

which means that we restrict to functions with the same momenta as the standard Gaussian. Then, by a careful treatment of the terms given by the *carré du champ* method, we have the following.

Proposition 20. *For all $v \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$ verifying (2.30), we have*

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n, d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^n} |v|^2 \log|v|^2\, d\gamma \geq \frac{1}{2\,n}\left(\int_{\mathbb{R}^d} |v|^2 \log|v|^2\, d\gamma\right)^2. \tag{2.31}$$

One can relax the hypotheses and have the same with $\int_{\mathbb{R}^n} |x|^2\, v^2\, d\gamma \leq d$.
Combining equation (2.31) with classical inequalities one can control other distances between v and 1, as the Wasserstein distance or the L^1–norm of the difference. Proposition 20 provides an exponent *four* in the stability term (cf (2.17)), which is not fully satisfying, since the results for Sobolev's inequality in finite dimension would suggest an optimal exponent equal to *two*. This is the case for two distinguished families of functions.

Theorem 21. *For all $v \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ such that $\log v$ is a concave function and*

$$\int_{\mathbb{R}^d} \left(1, x, |x|^2\right) |v|^2\, d\gamma = (1, 0, d) = \int_{\mathbb{R}^d} \left(1, x, |x|^2\right) d\gamma = (1, 0, d), \tag{2.32}$$

with $\mathscr{C} = 1 + \frac{1}{1728} \approx 1.0005787$ we have

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)} - \frac{\mathscr{C}}{2}\int_{\mathbb{R}^d} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)}}\right) d\gamma \geq 0\,. \tag{2.33}$$

Inequality (2.33) with improved constant $\mathscr{C} > 1$ compared to (GLS) can be recast in the form of a stability inequality of type (2.17) with sharp exponent *two*. Log-concavity might appear as a rather restrictive assumption, but it is not the case because a function which is compactly supported at time $t = 0$ evolves through the heat flow into a logarithmically concave function for some finite time that can be estimated by [192]. This is enough to produce a stability result with an explicit constant. It turns out that the compact support assumption can be relaxed.

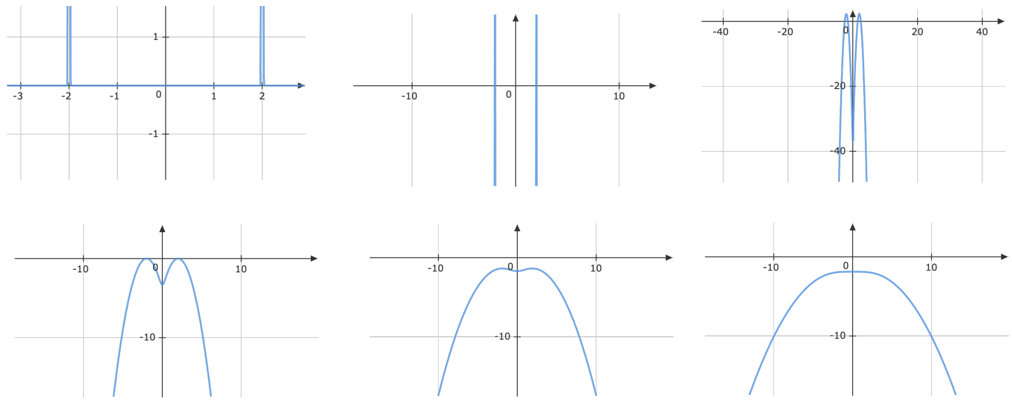

Figure 2.3: Creation of log-concavity in finite time along the heat equation.
The first picture corresponds to the initial datum $f_0 = \frac{1}{2}(\delta_{-2} + \delta_2)$. The others pictures plot the log of the corresponding solution at $t = 0.01$, $t = 0.1$, $t = 1$, $t = 2$, $t = 8$, respectively.

Theorem 22. *Let $d \geq 1$. For any $\epsilon > 0$, there is some explicit $\mathscr{C} > 1$ depending only on ϵ such that*

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)} \geq \frac{\mathscr{C}}{2} \int_{\mathbb{R}^d} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^d)}}\right) d\gamma \tag{2.34}$$

holds for any $v \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ satisfying (2.32) with

$$\int_{\mathbb{R}^d} |v|^2 e^{\epsilon |x|^2} d\gamma < \infty .$$

2.3 Nonlinear Dirichlet forms

Let (X, m) be any sigma-finite measure space. Consider the space $\mathrm{L}^2(X, m)$ and let $(T_t)_{t \geq 0}$ be a one-parameter family of bounded linear operators over $\mathrm{L}^2(X, m)$ such that

1. $T_0 = \mathrm{Id}$,
2. the mapping $t \mapsto T_t(f)$ is continuous for all $f \in \mathrm{L}^2(X, m)$,
3. for all $t, s \geq 0$, we have $T_t \circ T_s = T_s \circ T_t = T_{t+s}$,
4. for all $t \geq 0$, the operator T_t is 1-Lipschitz in the L^2–norm,
5. for all $f \in \mathrm{L}^2(X, m)$ we have $0 \leq f \leq 1 \implies 0 \leq T_t f \leq 1$.

These properties define a classical Markov semigroup. Then, one can consider the generator of T_t, which is an unbounded operator defined by

$$A = \lim_{t \to 0^+} \frac{T_t - \mathrm{Id}}{t}.$$

We usually write $T_t = e^{tA}$. In addition, we have

$$\forall t > 0, \forall f \in L^2(X,m), \qquad \partial_t T_t f = A T_t f. \tag{2.35}$$

The emblematic example of such structure is the semigroup associated with the heat flow, see (2.24) or (2.28) for $m = 1$, where $A = \Delta$ or $A = \Delta_\gamma$, respectively. However, such structure adapts to much more general settings, such as the heat flow on Riemannian manifolds or RCD metric measure spaces, [8]. Also some flows induced by non-local operators are covered, see [146] for a complete survey.

If the operator A is self-adjoint, then, one can define a symmetric bilinear form

$$\mathcal{E}(f,g) = \left(A^{\frac{1}{2}} f, A^{\frac{1}{2}} g\right).$$

Polarising the bilinear form, one can obtain a quadratic form, we always call $\mathcal{E}$. If T_t is a Markov semigroup, then $\mathcal{E}$ is called a Dirichlet form. In the examples above, $\mathcal{E}$ is exactly the H^1 semi-norm.

A fundamental feature of Dirichlet forms is the *normal contraction property*. Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-Lipschitz function such that $\phi(0) = 0$, also called a normal contraction. Let Φ the collection of all *normal contractions* ϕ. It is classical [146, Chapter 1] that if $\mathcal{E}$ is a Dirichlet form, then $\mathcal{E}$ verifies the following normal contraction property

$$\mathcal{E}(\phi(f)) \le \mathcal{E}(f), \quad \forall f \in L^2(X,m), \quad \forall \phi \in \Phi. \tag{2.36}$$

In [99] the authors consider nonlinear Markov semigroups and nonlinear Dirichlet forms.
Let $(S_t)_t$ be a one-parameter family of 1−Lipschitz (possibly nonlinear) operators over $L^2(X,m)$, verifying the properties 1.-4. , and, in addition

6. for all $f, g \in L^2(X,m)$, and all $t \ge 0$, we have $f \le g \implies S_t f \le S_t g$,

7. for all $f, g \in L^2(X,m)$, for all $t \ge 0$, and for $m-$almost all $x \in X$, we have $|S_t f(x) - S_t g(x)| \le |f(x) - g(x)|$.

Then, $(S_t)_t$ is a nonlinear Markov semigroup. Some examples are (2.23) and (2.28), for $m \neq 1$. Note that the last definition generalises that of linear Markov semigroups, as property 5. implies properties 6.-7. for linear operators. Assume that the semigroup $(S_t)_t$ is also the gradient flow of some lower-semicontinuous (not necessarily quadratic) functional $\mathcal{E}$ on $L^2(X,m)$, i.e. for all $f \in L^2(X,m)$, the curve $t \mapsto S_t f$ is the only global solution to

$$\begin{cases} \partial_t S_t f \in -\partial\mathcal{E}(S_t f), \\ S_0 f = f, \end{cases} \tag{2.37}$$

where ∂E is the subdifferential of $\mathcal{E}$:

$$\partial\mathcal{E}(f) = \left\{ v \in L^2(X,m) : \mathcal{E}(g) - \mathcal{E}(f) \ge \int_X v(g-f)\, dm, \quad \forall g \in L^2(X,m) \right\}.$$

In this case, according to [99], the functional $\mathcal{E}$ is a nonlinear Dirichlet form. As a byproduct, $\mathcal{E}$ is necessarily convex over $L^2(X,m)$. We remark that (2.37) is a generalised version of (2.35) for the linear case, as $-\partial\mathcal{E} = A$ if $\mathcal{E}$ is a quadratic form.

The first result of our work [P5] is a more efficient characterisation of the nonlinear Dirichlet forms, which is purely based on functional inequalities.

For all $f, g \in L^2(X,m)$, and $\alpha \in [0,\infty)$ we denote by $f \vee g$ and $f \wedge g$ the pointwise maximum and minimum and set $H_\alpha(f,g) = (g-\alpha) \vee f \wedge (g+\alpha)$ (see Figure 8.1), that is,

$$H_\alpha(f,g)(x) = \begin{cases} g(x)-\alpha & \text{if } f(x)-g(x) < -\alpha, \\ f(x) & \text{if } f(x)-g(x) \in [-\alpha,\alpha], \\ g(x)+\alpha & \text{if } f(x)-g(x) > \alpha. \end{cases} \tag{2.38}$$

Theorem 23. *Let $\mathcal{E} : \mathrm{L}^2(X,m) \to [0,\infty]$ be a l.s.c. functional. Then, $\mathcal{E}$ is a non-bilinear Dirichlet form if and only if, for all $f, g \in \mathrm{L}^2(X,m)$, and $\alpha \in [0,\infty)$, $\mathcal{E}$ verifies*

$$\mathcal{E}(f \vee g) + \mathcal{E}(f \wedge g) \le \mathcal{E}(f) + \mathcal{E}(g), \tag{2.39}$$

$$\mathcal{E}(H_\alpha(f,g)) + \mathcal{E}(H_\alpha(g,f)) \le \mathcal{E}(f) + \mathcal{E}(g). \tag{2.40}$$

Thanks to Theorem 23, we can prove our main result, that is (2.36) for the nonlinear Dirichlet forms of [99] under minimal assumptions.

Theorem 24. *Let $\mathcal{E}$ be a non-bilinear Dirichlet form. Then $\mathcal{E}$ has the normal contraction property* (2.36) *if and only if*

$$\mathcal{E}(-f) \le \mathcal{E}(f) \quad \forall f \in \mathrm{L}^2(X,m). \tag{2.41}$$

The normal contraction property is fundamental for the classification of bilinear Dirichlet forms. Then, we see Theorem 24 as a first step in this direction for the nonlinear setting.

Let us conclude by giving an account of our methodology for the proof, as it is original.

- The literature on classical Dirichlet forms would suggest to prove (2.36) for indicator functions f and proceed by linear combination to cover the whole space $\mathrm{L}^2(X,m)$. In our context this is unfeasible, though.
- Our idea is to approximate the normal contraction ϕ instead of f in (2.36).
- We show that it is sufficient to prove (2.36) for piecewise linear contractions ϕ such that $\phi' = \pm 1$ almost-everywhere in $\mathbb{R}$.
- We prove that (2.36) holds for three-piece piecewise linear contractions ϕ.
- We conclude, by induction, using the lemma below.

Lemma 25. *Let G be the set of all normal contractions $\phi \in \Phi$ such that $|\phi'| = 1$ and ϕ' has at most two points of discontinuity. Let $\langle G \rangle$ be the collection of all finite compositions of elements in G. Then, $\langle G \rangle$ is dense in Φ for the pointwise convergence on $\mathbb{R}$.*

The last result has a constructive proof of independent interest.

2.4 Open problems

Time averages for more classes of kinetic equations

The approach we develop in [P6] is not equation-specific. For example, it makes very limited use of the linearity of Vlasov–Fokker–Planck models. Moreover, the functional setting we develop involves general inequalities which are later specialised to solutions of the PDEs.

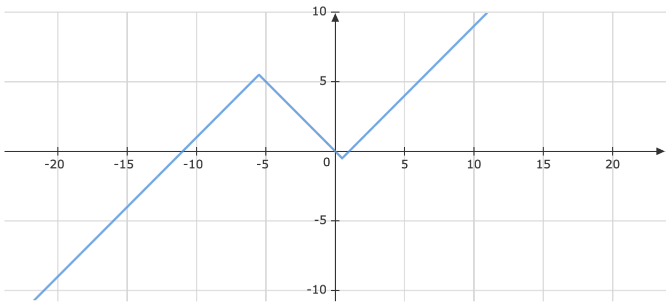

Figure 2.4: A typical *irreducible* element in the decomposition of Lemma 25.

Problem 26. *Compute global constructive decay rates for Generalised Langevin Dynamics [229, 211], Vlasov–Poisson systems, and Langevin dynamics with periodic forcing in time.*

- For the first class of models, our strategy is effective, except for Lemma 4, which should be adapted.
- For Vlasov–Poissons systems, the keypoint should be proving a Poincaré inequality like (4.11) for the spatial confinement driving the model, which depends on the solution itself.
- For the periodic case, all functional inequalities should be reworked, but the presence of the forcing should make solutions to be close to periodic for long times. Then, functional inequalities would reduce to spectral problems in time, space, and velocity modes, at least in certain ranges of the parameters.

Fully flow-based stability of functional inequalities

In [P2, P3] we provide constructive stability results based on the combinations of nonlinear flows with Taylor expansions. Still, the information collected from the nonlinear diffusion flows is not completely exploited. In addition, the knowledge of optimal constants in the stability results (or at least estimates of them scaling well with the dimension of the problem) would be a decisive advance in the theory.

Problem 27. *Compute explicit optimal constants in stability estimates for functional inequalities of Sobolev type, on $\mathbb{R}^d$ and $\mathbb{S}^d$ by pure nonlinear diffusion equations and entropy methods.*

The technique would rely on diffusion flows, with accurate estimates for the variation of the inequalities' deficit along solutions. The paradigmatic example to be treated is the critical Sobolev inequality, in view of the recent results of [119]. However, this is expected to be a very difficult task. As an easier entry point, we would consider subcritical inequalities on the sphere (or the Gaussian space).

Stability for the logarithmic Sobolev inequality

Many stability results and counterexamples are available for the Gaussian logarithmic Sobolev inequality. However, the picture is not complete, yet. Sharp exponents in the stability terms are mostly unknown even for distances inducing the narrow convergence. In the H^1–distance, we know stability results of order two for special classes of functions, but the least restrictive conditions to have those are an open issue. The same is true for order-four exponents, which hold under more general hypotheses.

Problem 28. *Find minimal hypotheses of stability results for the Gaussian logarithmic Sobolev inequality, in the distance induced by the* H^1 *norm.*

First, one should assess how many finite moments of the density are necessary for a uniform stability estimate in the square of the Fisher information w.r.t. the standard Gaussian. In [171] the problem is solved with a fourth-order moment. On the other side, in [P4] a similar estimate is achieved for densities with just a second-order moment, provided that it is small enough. A natural question is to investigate densities with an arbitrarily big, yet finite, second order-moment, but with no finite fourth-order moment.

A much more ambitious point is to give necessary and sufficient conditions for order-two stability estimates, not only in Fisher information, but in the H^1–distance from the set of optimisers for (GLS). The final (and very difficult) goal would be to have sharp exponents and constructive constants in the stability estimates.

As an intermediate step, we plan to use the Bianchi–Egnell strategy, which should provide the optimal exponent in the stability term. Due to the criticality of the problem, Taylor expansions are especially difficult. Then, we suggest using either an approach based on flows or some selection principle in the fashion of [98].

Gradient flows for kinetic equations

The fruitful connections between diffusion equations, gradient flows, and optimal transport have been a leading research theme in the last twenty years, after [6, 176, 225]. Such a framework still lacks in the context of kinetic equations, but diffusion limits, a part of Hilbert's sixth problem, point in this direction.

Problem 29. *Find distances between probability measures, in the spirit of optimal transport capturing the geometric structure of spatially inhomogeneous kinetic PDEs. Recast dissipative kinetic PDEs as gradient flows in such distances. The new metric structures shall be consistent with diffusion limits.*

The geometric structure behind kinetic models is complicated, due to the asymmetry between *positions* and *velocities*. Diffusion limits are often formal, so more tools and new methods are needed. Results in the sense of Problem 29 will hopefully contribute to clarify these points. The scientific benefit of any contribution to this topic would impact various related fields: mathematical physics, probability theory, and applications of those.

Practically, some steps in this directions are:

1. After the variational schemes of [133, 87] and the *splines* of [40], provide a right metric setting for Vlasov-Fokker-Planck equations, which is compatible with diffusion limits.

2. Generalise the work [86] to the spatially inhomogeneous setting.

Nonlinear Dirichlet forms and metric spaces

It is known that all Cheeger's energies in the sense of [8] are non-bilinear Dirichlet forms. Hence, an abstract framework to study metric spaces whose Cheeger's energy is non-quadratic may emerge.

Problem 30. *Represent local, non-bilinear Dirichlet forms as non-quadratic Cheeger's energies.*
Study geometric conditions using the formalism of non-bilinear Dirichlet forms.

Ricci curvature lower bounds are useful for diffusion equations, functional inequalities, and geometric measure theory. The simplest spaces where the Cheeger's energy is a non-bilinear Dirichlet form are Finsler manifolds. After [223, 224], the two approaches towards Ricci curvature bounds discussed in [10] are shown to be equivalent also in this setting. However, only one of them – namely that of [201, 245] – is well-posed in the general framework of metric measure spaces. The language of non-bilinear Dirichlet could contribute to close this gap. The first steps one could envision in the direction of Problem 30:

1. Rewrite [223, 224] with the language of non-bilinear Dirichlet forms.
2. Study the heat flow and its linearised version on metric measure spaces, taking inspiration from [223, 149, 136].
3. Represent abstract non-bilinear Dirichlet forms as non-quadratic Cheeger's energies, following [10, 51].

2.5 Bibliographical comments

We recall hereby some of the most important references for this thesis, which were not introduced before for the sake of simplicity.

Kinetic equations and hypocoercivity

- The H^1 framework for hypocoercivity has been developed in [255, 256, 115], after [221, 165]. We refer to [126] for direct applications to the Fokker–Planck equations. See also [33, 34, 32] for a H^1 approach to hypocoercivity reminiscent of [24].
- Hypocoercivity methods based on couplings are used in [134, 116].
- The L^2–hypocoercivity framework is due to [127], after [164]. More models are analysed in [67, 65, 66]. An alternative approach based on Lyapunov's inequality for matrices is established in [18] for general Fokker–Planck equations. More detailed calculations are carried out in [2, 4, 17], for specific examples where computations can be done mode-by-mode. An improvement of those is given in [16]. An interpretation in terms of Shur's complements of [127], with careful computations of the constants, appears in [44]. For an approach to hypocoercivity based on Lyapunov's functionals we refer to [78]. See also [243, 68] for hypocoercivity results for a general class of Langevin dynamics, and [45] for recent developments in presence of boundary conditions. In [212] the authors recover exponential decay for well-chosen initial data in case of weak confinement. See also [117] for a different point of view covering most known cases.
- An alternative approach to hypocoercivity, based on negative Sobolev norms of solutions, appeared in [5]. There, instead of twisting the reference norm as in [256, 127], the strategy is to prove a modified space-time-velocity Poincaré inequality. Then, the same inequality is used to establish a convergence rate in the L^2 norm for solutions to the equations towards Maxwellians, via a time-integrated estimate. More quantitative results are given in [79].
- We refer to [11, 117, 79, 84] for up-to-date results on Lions-Poincaré inequalities, inspiring [P6].
- Time integrated functionals of solution to kinetic PDEs have also been considered in [161, 244].

Functional inequalities

- The family of Gagliardo–Nirenberg (GN) inequalities we consider in Chapter 5, which interpolate between the Poincaré and the critical Sobolev inequalities for the uniform probability measure on the unit sphere, are due to [48]. Further historical reference in the subject are [196, 37]. Improved versions of GN inequalities on the sphere, obtained with the carré du champ technique, are given in [123]. The extra term found in the last reference has been fruitfully exploited in [112]. In [145] a proof of GN inequalities by spectral methods appears, which can be adapted to prove a local stability result. Then, in [144], a first global stability result for a subfamily of GN inequalities on the sphere is given. The result is not constructive, but it is optimal in the order of the stability term. For the critical case, corresponding to Sobolev's inequality, stability issues have a longer history we discussed above. We mention [119] and [185] as the latest advances in the subject.

- By taking the large-dimensional limit of spheres equipped with the uniform probability measure, the Gaussian space is recovered, see [252] for an historic account on the subject. This idea is also used in [37]. Gaussian subcritical inequalities, interpolating between the Poincaré and the logarithmic Sobolev inequalities are due to Beckner [37]. Optimal constants for the inequalities we consider are derived in [14]. See also [187] for an alternative interpolation family. Finally, [220], provides an hypercontractivity property we directly use to establish improved inequalities under orthogonality constraint.

- As the critical case in the family of Gaussian interpolation inequalities, we find the the Gaussian logarithmic Sobolev inequality (LSI). See the previous section and [12] for a complete historical account. Among the different approaches to LSI, we follow more closely that of [24, 19]. Stability for the LSI is a research trend under development, whose interest is motivated also by problems in physics, statistics, and engineering. See [138, 61, 172] for an account on recent stability results in Wasserstein-like distances. Stability in L^1–norm is established in [58], and all available results in L^p–norms are collected in [172], except the very delicate L^2–estimate of [119]. Some counterexamples, which show that a control on the second order moment of the distribution in necessary, are given in [182, 135, 171]. In [119] the problem of stability in H^1 norm is stated. Two partial contributions in this direction are [138, 171]. However, a complete solution to the problem is not known, yet. To the extent of our results, two important instruments are preservation/creation of log-concavity along the heat flow [192, 236], and isoperimetry for log-concave measures [56].

Nonlinear Dirichlet Forms

- Linear Dirichlet forms are a classical subject, see the references quoted in the previous section. The first contribution in the sense of a nonlinear extension of the theory is [99]. The theory is continued in [101, 100], with a nonlinear theory of capacity and a partial result on normal contractions. Important examples of nonlinear Dirichlet forms are the Cheeger's energies of [8] and the Finsler–Sobolev seminorms in [223].

Part II

Kinetic equations

Chapter 3

Time averages for kinetic Fokker-Planck equations

This Chapter corresponds to [P1], published in Kinetic and Related Models (2022).

Abstract

We consider kinetic Fokker-Planck equations on the torus with Maxwellian or fat tail local equilibria. Results based on weak norms have recently been achieved by S. Armstrong and J.-C. Mourrat in case of Maxwellian local equilibria. Using adapted Poincaré and Lions-type inequalities, we develop a constructive method for estimating the decay rate of time averages of norms of the solutions, which covers various regimes corresponding to subexponential, exponential and superexponential (including Maxwellian) local equilibria. As a consequence, we also derive hypocoercivity estimates, which are compared to similar results obtained by other techniques.

3.1 Introduction

Let us consider the *kinetic Fokker-Planck equation*

$$\partial_t f + v\cdot\nabla_x f = \nabla_v\cdot\left(\nabla_v f + \alpha\,\langle v\rangle^{\alpha-2}\,v\,f\right),\qquad f(0,\cdot,\cdot) = f_0. \tag{3.1}$$

where f is a function of time $t\ge 0$, position x, velocity v, and α is a positive parameter. Here we use the notation

$$\langle v\rangle = \sqrt{1+|v|^2},\quad \forall\, v\in\mathbb{R}^d.$$

We consider the spatial domain $\mathbb{T} := (0,L)^d \ni x$, with *periodic boundary conditions*, and define $\Omega_t := (t,t+\tau)\times\mathbb{T}$, for some $\tau>0$, $t\ge 0$ and $\Omega = \Omega_0$. The normalized *local equilibrium*, that is, the equilibrium of the spatially homogeneous case, is

$$\gamma_\alpha(v) = \frac{1}{Z_\alpha}\,\mathrm{e}^{-\langle v\rangle^\alpha},\quad \forall\, v\in\mathbb{R}^d,$$

where Z_α is a non-negative normalization factor, so that $d\gamma_\alpha := \gamma_\alpha(v)\,dv$ is a probability measure. We shall distinguish a sublinear regime if $\alpha\in(0,1)$, a linear regime if $\alpha = 1$ and a superlinear regime if $\alpha\ge 1$. The superlinear regime covers the Maxwellian case $\alpha = 2$. The threshold case $\alpha = 1$ corresponds to a linear growth of $\langle v\rangle^{\alpha}$. The estimates in the linear case are similar to the ones of the superlinear regime. In the literature, γ_α is said to be *subexponential*, *exponential* or *superexponential* depending whether the regime is sublinear, linear or superlinear.

The mass

$$M := \int_{\mathbb{T}\times\mathbb{R}^d} f(\cdot,x,v)\,dx\,dv$$

is conserved under the evolution according to the kinetic Fokker-Planck equation (3.1). We are interested in the convergence of the solution to the stationary solution $M\,L^{-d}\,\gamma_\alpha$. By linearity, we can assume from now on that $M = 0$ with no loss of generality. The function

$$h = \frac{f}{\gamma_\alpha}$$

solves the *kinetic-Ornstein-Uhlenbeck* equation

$$\partial_t h + v\cdot\nabla_x h = \Delta_\alpha h, \qquad h(0,\cdot,\cdot) = h_0, \tag{3.2}$$

with

$$\Delta_\alpha h := \Delta_v h - \alpha\, v\,\langle v\rangle^{\alpha-2}\cdot\nabla_v h,$$

and *zero-average initial datum* in the sense that

$$\int_{\mathbb{T}\times\mathbb{R}^d} h_0(x,v)\,dx\,d\gamma_\alpha = 0.$$

By mass conservation, solutions to (3.2) are zero-average for any $t > 0$.

Therefore, we consider the *time average* defined as

$$⨍_t^{t+\tau} g(s)\,ds := \frac{1}{\tau}\int_t^{t+\tau} g(s)\,ds$$

without specifying the τ dependence when not necessary. Our first result is devoted to the decay rate of $h(t,\cdot,\cdot)\to 0$ as $t\to\infty$ using time averages.

Theorem 31. *Let $\alpha \ge 1$. Then, for all $L > 0$ and $\tau > 0$, there exists a constant $\lambda > 0$ such that, for all $h_0 \in \mathrm{L}^2(dx\,d\gamma_\alpha)$ with zero-average, the solution to* (3.2) *satisfies*

$$⨍_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \le \|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,e^{-\lambda t}, \quad \forall\, t\ge 0. \tag{3.3}$$

The expression of λ as a function of τ and L is given in Section 3.4. To deal with large-time asymptotics in kinetic equations, it is by now standard to use hypocoercivity methods. Although not being exactly a hypocoercive method in the usual sense, Theorem 31 provides us with a hypocoercivity estimate.

Corollary 32. *Under the assumptions of Theorem 31, there exists an explicit constant $C > 1$ such that all solutions h to* (3.2) *fulfill*

$$\|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)} \le C\,\|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,e^{-\lambda t}, \quad \forall\, t\ge 0. \tag{3.4}$$

A typical feature of hypocoercive estimates is the factor $C > 1$ in (3.4). The prefactor $C > 1$ cannot be avoided. Otherwise, inequality (3.4) would be equivalent to a Poincaré inequality where the L^2–norm of a function is controlled with the velocity gradient only. We can see explicitly that for $\alpha = 2$ the Green function of (3.1), computed in [183], has a built-in delay. In particular,

$$\|h_0\|^2_{L^2(dx\,d\gamma_2)} - \|h(t,\cdot,\cdot)\|^2_{L^2(dx\,d\gamma_2)} = O(t^3),$$

as $t \to 0^+$. Note that there is no such a constant in (3.3).

Now, let us turn our attention to the subexponential case $0 < \alpha < 1$.

Theorem 33. *Let $\alpha \in (0,1)$ Then, for all $L > 0$ and $\tau > 0$, for all $\sigma > 0$, there is a constant $K > 0$ such that all solutions to* (3.2) *decay according to*

$$\fint_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{L^2(dx\,d\gamma_\alpha)}\,ds \le K\,(1+t)^{-\frac{\sigma}{2(1-\alpha)}} \int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^\sigma h_0^2\,dx\,d\gamma_\alpha, \quad \forall\, t \ge 0. \tag{3.5}$$

Further details will be given in Section 3.5. The constants λ, C in Corollary 32 and K in Theorem 33 depend on $L > 0$ and $\tau > 0$ and their values are discussed later. The rate of Theorem 33 is the same as in the spatially-homogeneous case of [67, Proposition 11]. In the spatially-inhomogenous case, rates are known, see [180, 66]. Finally, see Section 3.5 for a discussion of the limit $\alpha \to 1^-$.

Equation (3.2) is used in physics to describe the distribution function of a system of particles interacting randomly with some background, see for instance [28]. The *kinetic Fokker-Planck equation* is the Kolmogorov forward equation of Langevin dynamics

$$\begin{cases} dx_t = v_t\,dt, \\ dv_t = -v_t + \sqrt{2}\,dW_t, \end{cases}$$

where W_t is a standard Brownian motion. See [44, Introduction] for further details on connections with probability theory. The kinetic Fokker-Planck equation (3.1) is a simple kinetic equation which has a long history in mathematics that we will not retrace in details here. Mathematical results go back at least to [183] and are at the basis of the theory of L. Hörmander (see, *e.g.*, [169]), at least in the case $\alpha = 2$. For the derivation of the kinetic-Fokker-Planck equation from underlying stochastic ODEs, particularly in the context of astrophysics, we can refer to [92, eq. (328)]. The hypocoercivity theory emerged from [165, 114] and was built up in a fully developed theory in [256] with important contributions in [164, 215]. Existence theory for solutions to the Vlasov-Fokker-Planck equation was discussed also in [108, Appendix A].

The word *hypocoercivity* was coined by T. Gallay, in analogy with the already quoted *hypoelliptic* theory of Hörmander in [169]. In [256], C. Villani distinguishes the regularity point of view for elliptic and parabolic problems driven by degenerate elliptic operators from the issue of the long-time behaviour of solutions, which is nowadays attached to the word hypocoercivity. The underlying idea is to twist the reference norm, in order to carry properties (as the coercivity of the operator driving (3.2)) from velocity direction to space directions, thanks to commutators. Twisting the H^1-norm creates equivalent norms, which are exponentially decaying along the evolution. So works the H^1 framework, see [256, 255, 126]. The H^1-framework has been connected to the *carré du champ* method of D. Bakry and M. Emery in [24] by F. Baudoin, who proved decay also with respect to the Wasserstein distance, as shown in [32, 33, 34]. We report also the works [134, 116], where accurate convergence rates in the Wasserstein distance for (3.2) are computed trough a coupling argument.

The H^1 hypocoercivity implies a decay rate for the L^2 norm [215], but the corresponding estimates turn out to be sub-optimal. Moreover, kinetic equations driven by non-regularising operators are not well suited for the H^1–framework. This motivates the development of direct L^2 techniques based on a perturbation of the L^2 norm. Such an approach can be found in [127] and [67], which is consistent with diffusion limits, after [164]. In [66], the authors extend the technique to the subexponential case. Another possibility is to perform rotations in the phase space and use a Lyapunov inequality for matrices as in [18]. This approach gives optimal rates, but it is less general as it requires further algebraic properties for the diffusion operator and a detailed knowledge of its spectrum. The core of [18] is a spectral decomposition, that was originally understood via a toy model exposed in [127]. In a domain with periodic boundary conditions and no confining potential, the problem is reduced to an infinite set of ODEs corresponding to spatial modes. See [2, 3, 16] for details and extensions. Other techniques related to hypocoercivity – involving time-integrated functionals and the application of the so called *kinetic-fluid decomposition*, appear in [244, 161] and subsequent papers.

A new hypocoercivity theory, involving Sobolev norms with negative exponents of the transport operator, was recently proposed by S. Armstrong and J.-C. Mourrat in [5]. Using space-time adapted Poincaré inequalities they derive qualitative hypocoercive estimates in the case $\alpha = 2$. The constants appearing there are not quantified. One of the difficulties lies in controlling the constant in Lions' Lemma, which is done in our Section 3.2. An extension to the whole space in presence of a confining potential can be found in [79]. Note that the strategy of using time-integrated functionals of the solutions to kinetic equations is present also in [250, 132], after [161].

Adopting the strategy of [5], in this paper we study the convergence to equilibrium of solutions to (3.1) and (3.2), as it is a simple benchmark in kinetic theory, [127, 255, 5], and a simplified model of the Boltzmann equation when collisions become grazing, see [109].

Our original contribution lies in making the strategy of [5] effective, and to generalise it to kinetic Fokker-Planck equations where local equilibria are not necessarily Maxwellians. First, we are able to track the Lions' constant in terms of the parameters (see Lemma 38). Moreover, we achieve a fully constructive proof of the averaging Lemma 42. This allows both for an explicit estimate of the constant and for an adaptation to more general models. One important point is the control in terms of the offset of the solution from the velocity average, without explicitly using gradients, see Proposition 43. So, we compute explicit and accurate decay rates of time averages of solutions to (3.2). Hypocoercivity estimates are obtained as a consequence of these decay rates, see Corollary 32. We perform an analysis for all positive values of α, which is consistent in the threshold case $\alpha = 1$. Since the estimates are explicit, we are able to compare the strategy of [5] to other L^2–hypocoercivity methods.

This document is organized as follows. In Section 3.2 we collect some preliminary results: Poincaré and weighted Poincaré inequalities (Propositions 35 and 36), adapted Lions' inequality (Lemmas 37 and 38). In Section 3.3 we introduce an averaging lemma (Lemma 42), which is then used to prove the generalized Poincaré inequality of Proposition 43, at the core of the method. In Section 3.4 we use Proposition 43 and a Grönwall estimate to prove Theorem 31 and compute an explicit formula for λ (Proposition 45). Section 3.5 is devoted to the proof of Theorem 33, with additional details, and to the limit $\alpha \to 1^-$. Finally, in Section 3.6, we derive the hypocoercive estimates of Corollary 32. On the benchmark case $\alpha = 2$ in one spatial dimension, we also compare our results with those obtained by more standard methods.

3.2 Preliminaries

Let us start with some preliminary results.

3.2.1 Weighted spaces

For functions g of the variable v only, that is, of the so-called homogeneous case, we define the *weighted Lebesgue* and *Sobolev spaces*

$$\mathrm{L}^2_\alpha := \mathrm{L}^2(\mathbb{R}^d, d\gamma_\alpha) \quad \text{and} \quad \mathrm{H}^1_\alpha := \left\{ g \in \mathrm{L}^2_\alpha : \nabla_v g \in \left(\mathrm{L}^2_\alpha\right)^d \right\}.$$

We equip L^2_α with the scalar product

$$(g_1, g_2) = \int_{\mathbb{R}^d} g_1(v)\, g_2(v)\, d\gamma_\alpha \tag{3.6}$$

and consider on H^1_α the norm defined by

$$\|h\|^2_{\mathrm{H}^1_\alpha} := \left(\int_{\mathbb{R}^d} h\, d\gamma_\alpha \right)^2 + \int_{\mathbb{R}^d} |\nabla_v h|^2\, d\gamma_\alpha$$

as in [5]. The duality product between $\mathrm{H}^{-1}_\alpha \ni z$ and $\mathrm{H}^1_\alpha \ni g$ is given by

$$\langle z, g \rangle := \int_{\mathbb{R}^d} \nabla_v w_z \cdot \nabla_v g\, d\gamma_\alpha,$$

where w_z is the weak solution in H^1_α to

$$-\Delta_\alpha w_z = z - \int_{\mathbb{R}^d} z\, d\gamma_\alpha, \quad \int_{\mathbb{R}^d} w\, d\gamma_\alpha = 0.$$

Here we write $\int_{\mathbb{R}^d} z\, d\gamma_\alpha$ for functions which are integrable w.r.t. $d\gamma_\alpha$ and, up to a little abuse of notations, this quantity has to be understood in the distribution sense for more general measures. As a consequence and with the above notations, we define

$$\|z\|^2_{\mathrm{H}^{-1}_\alpha} := \left(\int_{\mathbb{R}^d} z\, d\gamma_\alpha \right)^2 + \|w_z\|^2_{\mathrm{H}^1_\alpha}.$$

With these notation, the key property of the operator Δ_α, is

$$\langle g_1, \Delta_\alpha g_2 \rangle = -\int \nabla_v g_1 \cdot \nabla_v g_2\, d\gamma_\alpha$$

for any functions g_1, $g_2 \in \mathrm{H}^1_\alpha$.

We recall that $\Omega_t = (t, t+\tau) \times \mathbb{T} \subset \mathbb{R}^+_t \times \mathbb{R}^d_x$ and that x-periodic boundary conditions are assumed. Consider next functions h of $(t, x, v) \in \mathbb{R}^+ \times \mathbb{T} \times \mathbb{R}^d$ and define the space

$$\mathrm{H}_{\mathrm{kin}} := \left\{ h \in \mathrm{L}^2\left((t, t+\tau) \times \mathbb{T}; \mathrm{H}^1_\alpha \right) : \partial_t h + v \cdot \nabla_x h \in \mathrm{L}^2\left(\Omega_t; \mathrm{H}^{-1}_\alpha \right) \forall\, t \geq 0 \right\}.$$

The dependence of the space on t, τ is implicit for readability purposes. We can equip $\mathrm{H}_{\mathrm{kin}}$ with the norm

$$\|h\|^2_{\mathrm{kin}} := \|h\|^2_{\mathrm{L}^2(\Omega_t; \mathrm{L}^2_\alpha)} + |h|^2_{\mathrm{kin}}$$

where the *kinetic semi-norm* is given by

$$|h|^2_{\mathrm{kin}} := \|\nabla_v h\|^2_{\mathrm{L}^2(\Omega_t; \mathrm{L}^2_\alpha)} + \|\partial_t h + v \cdot \nabla_x h\|^2_{\mathrm{L}^2(\Omega_t; \mathrm{H}^{-1}_\alpha)}.$$

We refer to [5, Section 6] for the proof of following result.

Proposition 34. *The embedding* $\mathrm{H}^1_{\mathrm{kin}} \hookrightarrow \mathrm{L}^2\left(\Omega_t \times \mathbb{T}; \mathrm{L}^2_\alpha\right)$ *is continuous and compact for any* $t \geq 0$.

3.2.2 Poincaré inequalities

In this subsection, we consider functions g depending only on the variable v. Let $\alpha \geq 1$. We can state some *Poincaré inequalities.*

Proposition 35. *If* $\alpha \geq 1$, *there exists a constant* $P_\alpha > 0$ *such that, for all functions* $g \in \mathrm{H}^1_\alpha$, *we have*

$$\int_{\mathbb{R}^d} |g - \rho_g|^2 \, d\gamma_\alpha \leq P_\alpha \int_{\mathbb{R}^d} |\nabla_v g|^2 \, d\gamma_\alpha \quad \textit{with} \quad \rho_g := \int_{\mathbb{R}^d} g \, d\gamma_\alpha. \tag{3.7}$$

With $\alpha \geq 1$, the operator Δ_α admits a compact resolvent on $\mathrm{L}^2(d\gamma_\alpha)$. Then, (3.7) holds by the standard results of [71, Chapter 6]. The best constant P_α is such that P_α^{-1} is the minimal positive eigenvalue of $-\Delta_\alpha$. See [89] and the references quoted therein for estimates on P_α. In the case of the Gaussian Poincaré inequality, it is shown in [219] that $P_2 = 1$, although the result was probably known before.

3.2.3 Weighted Poincaré inequalities

Here we consider again functions depending only on v. For $\alpha \in (0,1)$, inequality (3.7) has to be replaced by the following *weighted Poincaré inequality.*

Proposition 36. *If* $\alpha \in (0,1)$, *there exists a constant* $P_\alpha > 0$ *such that, for all functions* $g \in \mathrm{H}^1_\alpha$, *we have*

$$\int_{\mathbb{R}^d} \langle v \rangle^{2(\alpha-1)} |g - \rho_g|^2 \, d\gamma_\alpha \leq P_\alpha \int_{\mathbb{R}^d} |\nabla_v g|^2 \, d\gamma_\alpha \quad \textit{with} \quad \rho_g := \int_{\mathbb{R}^d} g \, d\gamma_\alpha. \tag{3.8}$$

For more details, we refer for instance to [66, Appendix A]. Notice that the average in the l.h.s. is taken w.r.t. $d\gamma_\alpha$, not w.r.t. $\langle v \rangle^{2(\alpha-1)} \, d\gamma_\alpha$

3.2.4 Lions' Lemma

Let $\mathcal{O}$ be an open, bounded and Lipschitz-regular subset of $\mathbb{R}^{d+1} \approx \mathbb{R}_t \times \mathbb{R}^d_x$. We recall that

$$\mathcal{H}^{-1}(\mathcal{O}) = \left\{ w \in \mathcal{D}^*(\mathcal{O}) : |\langle w, u \rangle_{\mathcal{O}}| \leq C \, \|u\|_{\mathcal{H}^1_0(\mathcal{O})}, C > 0 \right\},$$

where $\mathcal{D}^*(\mathcal{O})$ denotes the space of distributions over $\mathcal{O}$, equipped with the weak$*$ topology, and $\langle w, u \rangle_{\mathcal{O}}$ is the duality product between $\mathcal{H}^{-1}$ and $\mathcal{H}^1_0$. The norm on $\mathcal{H}^1_0(\mathcal{O})$ is as usual $u \mapsto \|\nabla u\|_{\mathrm{L}^2(\mathcal{O})}$. On $\mathcal{H}^1(\mathcal{O})$, we introduce the norm

$$\|u\|^2_{\mathrm{H}^1} = \left| \int_{\mathcal{O}} u \, dt \, dx \right|^2 + \|\nabla u\|^2_{\mathrm{L}^2(\mathcal{O})}.$$

The norm induced on $\mathrm{H}^{-1}(\mathcal{O})$ is then

$$\|w\|^2_{\mathrm{H}^{-1}(\mathcal{O})} = \langle w, 1 \rangle^2_{\mathcal{O}} + \|z_w\|^2_{\mathrm{H}^1(\mathcal{O})},$$

where z_w is the solution to

$$-(\partial_{tt} + \Delta_x) \, z_w = w - \langle w, 1 \rangle_{\mathcal{O}}, \quad \int_{\mathcal{O}} z_w \, dt \, dx = 0.$$

Lions' Lemma gives a sufficient condition for a distribution to be an L^2 function. The following statement is taken from [11] in the case without external potential.

Lemma 37. *Let $\mathcal{O}$ be a bounded, open and Lipschitz-regular subset in $\mathbb{R}^{d+1}$. Then, for all $u \in \mathcal{D}^*(\mathcal{O})$, we have that $u \in \mathrm{L}^2(\mathcal{O})$ if and only if the weak gradient ∇u belongs to $\mathrm{H}^{-1}(\mathcal{O})$. Moreover, there exists a constant $C_L(\mathcal{O})$ such that*

$$\left\| u - \int_{\mathcal{O}} u \, dx \, dt \right\|^2_{\mathrm{L}^2(\mathcal{O})} \le C_L \| \nabla u \|^2_{H^{-1}(\mathcal{O})},$$

for any $u \in \mathrm{L}^2(\mathcal{O})$.

According to [77, 60, 104], if $\mathcal{O}$ is star-shaped w.r.t. a ball, then the constant C_L has the following structure:

$$C_L = 4|S^d| \frac{\mathrm{D}(\mathcal{O})}{\mathrm{d}(\mathcal{O})}, \tag{3.9}$$

where D is the diameter of $\mathcal{O}$, while $\mathrm{d}(\mathcal{O})$ is the diameter of the largest ball one can include in $\mathcal{O}$. See in particular [104, Remark 9.3] and [60, Lemma 1]. As a consequence, we have the following explicit expression of C_L when $\mathcal{O} = \Omega$.

Lemma 38. *Let $L > 0$, $\tau \in (0, L)$ and $\Omega = (0, \tau) \times (0, L)^d$. Lemma 37 holds with*

$$C_L = 4|S^d| \frac{\sqrt{d\,L^2 + \tau^2}}{\tau}. \tag{3.10}$$

3.2.5 The kinetic Ornstein-Uhlenbeck equation

We consider solutions to (3.2) in the weak sense, *i.e.*, functions h in the space $\mathrm{C}(\mathbb{R}^+; \mathrm{L}^2(dx\, d\gamma_\alpha))$ with initial datum $h_0 = h(0,\cdot,\cdot)$ in $\mathrm{L}^2(dx\, d\gamma_\alpha)$ such that (3.2) holds in the sense of distributions on $(0,\infty) \times \mathbb{R}^d_x \times \mathbb{R}^d_v$. The following result is taken from [5] if $\alpha = 2$. The extension to $\alpha \neq 2$ is straightforward as follows from a careful reading of the proof in [5, Proposition 6.10].

Proposition 39. *Let $L > 0$ and $\alpha > 0$. With $\Omega = (0,\tau) \times (0, L)^d$, for all zero-average initial datum $h_0 \in \mathrm{L}^2(dx\, d\gamma_\alpha)$, there exists a unique solution h to (3.2) such that $h \in \mathrm{H}_{\mathrm{kin}}$ for all $\tau > 0$.*

Regularity properties for (3.2) are collected in [5, Section 6]. In the special case $\alpha = 2$, some fractional regularity along all directions of the phase space are known. Also see [231] for further result on regularity theory for kinetic Fokker-Planck equations.

3.2.6 A priori estimates

We state two estimates for solutions to (3.2).

Lemma 40. *Let $L > 0$, $\tau > 0$, $\Omega = (0, \tau) \times (0, L)^d$, and $\alpha > 0$. If h is a solution to (3.2), then we have*

$$\| (\partial_t + v \cdot \nabla_x) h \|_{\mathrm{L}^2(\Omega; \mathrm{H}^{-1}_\alpha)} \le \| \nabla_v h \|_{\mathrm{L}^2(\Omega; \mathrm{L}^2_\alpha)}. \tag{3.11}$$

Proof. Take a test function $\phi \in \mathrm{L}^2(\mathrm{H}^1_\alpha)$, and write

$$\int_{\mathbb{T}} \langle (\partial_t + v \cdot \nabla_x) h, \phi \rangle \, dx = \int_{\mathbb{T}} \langle \Delta_\alpha h, \phi \rangle \, dx = - \int_{\mathbb{T}} (\nabla_v h, \nabla_v \phi) \, dx,$$

from which (3.11) easily follows, after maximizing over $\| \nabla_v \phi \|_{\mathrm{L}^2_\alpha} \le 1$. □

For completeness, let us recall the classical L^2 decay estimate for solutions to (3.2).

Lemma 41. *Let $L > 0$, $\tau > 0$, $\Omega = (0,\tau) \times (0,L)^d$, and $\alpha > 0$. If h is a solution to (3.2), then we have*

$$\frac{d}{dt}\|h\|^2_{L^2(dx\,d\gamma_\alpha)} = -2\,\|\nabla_v h\|^2_{L^2(dx\,d\gamma_\alpha)}.$$

3.3 An averaging lemma and a generalized Poincaré inequality

For all functions $h \in H_{kin}$, we define the *spatial density*

$$\rho_h := \int_{\mathbb{R}^d} h(\cdot,\cdot,v)\,d\gamma_\alpha.$$

Notice that $\int_Q \rho_h\,dx = 0$ whenever h is a zero-average function.

3.3.1 Averaging lemma

Inspired by [5, Proposition 6.2], the following *averaging lemma* provides a norm of the spatial density, as for instance in [231].

Lemma 42. *Let $L > \tau > 0$, $\Omega = (0,\tau) \times (0,L)^d$, and $\alpha > 0$. For all $h \in H_{kin}$, we have*

$$\|\nabla_{t,x}\rho_h\|^2_{H^{-1}(\Omega)} \le d_\alpha\left(\|h-\rho_h\|^2_{L^2(dt\,dx\,d\gamma_\alpha)} + \|\partial_t h + v\cdot\nabla_x h\|^2_{L^2(\Omega;H^{-1}_\alpha)}\right) \tag{3.12}$$

with

$$d_\alpha = 2\left(\|v_1|v|^2\|^2_{L^2_\alpha} + \left(1+\tfrac{L^2}{4\pi^2}\right)\||v|^2\|^2_{L^2_\alpha} + \tfrac{d^2L^2}{4\pi^2}\,\|v\|^2_{L^2_\alpha}\right). \tag{3.13}$$

Inequality (3.12) can be extended to any measure $d\gamma$ such that $\int_{\mathbb{R}^d}|v|^4\,d\gamma < \infty$ and $\int_{\mathbb{R}^d} v\,d\gamma = 0$. The proof of Lemma 42 is technical, but follows in a standard way from the time-independent case, as it is common in averaging lemmas: see [231]. See also the more general Lemma 52 of Chapter 4.

Proof of Lemma 42. Assume that $h \in H_{kin}$ does not depend on t. Let $\phi \in \mathcal{D}(\mathbb{T})^d$ be a smooth test-vector field with zero average on each component. We write

$$-\int_{\mathbb{T}}\rho_h\nabla_x\cdot\phi\,dx = \int_{\mathbb{T}}(\nabla_x\rho_h)\cdot\phi\,dx,$$

with a slight abuse of notation, since the integral of the r.h.s. is in fact a duality product. Using $\int_{\mathbb{R}^d} v_i\,v_j\,d\gamma_\alpha = d^{-1}\,\|v\|^2_{L^2_\alpha}\,\delta_{ij}$, we obtain

$$\int_{\mathbb{T}}\nabla_x\rho_h\cdot\phi\,dx = d\,\|v\|^{-2}_{L^2_\alpha}\int_{\mathbb{T}\times\mathbb{R}^d} v\cdot\nabla_x\rho_h\,\phi\cdot v\,dx\,d\gamma_\alpha.$$

By adding and subtracting h, and then integrating by parts, still at formal level, we obtain

$$\begin{aligned}
&\int_{\mathbb{T}\times\mathbb{R}^d} v\cdot\nabla_x\rho_h\,\phi\cdot v\,dx\,d\gamma_\alpha\\
&\qquad = -\int_{\mathbb{T}\times\mathbb{R}^d} v\cdot(h-\rho_h)\,\nabla_x\phi\cdot v\,dx\,d\gamma - \int_{\mathbb{T}\times\mathbb{R}^d} v\cdot\nabla_x h\,v\cdot\phi\,dx\,d\gamma_\alpha\\
&\qquad \le \|h-\rho_h\|_{L^2(dx\,d\gamma_\alpha)}\,\|\nabla_x\phi\|_{L^2(dx)}\,\left\|v^2\right\|_{L^2_\alpha}\\
&\qquad\qquad + \|v\cdot\nabla_x h\|_{L^2(H^{-1}_\alpha)}\,\|\phi\|_{L^2(dx)}\,\|v\|_{H^1_\alpha}
\end{aligned}$$

using Cauchy-Schwarz inequalities and duality estimates. By the Poincaré inequality, we know that

$$\frac{4\pi^2}{dL^2}\|\phi\|^2_{L^2(dx)} \leq \|\nabla\phi\|^2_{L^2(dx)}$$

Maximizing the r.h.s. on ϕ such that $\|\nabla_x\phi\|_{L^2(dx)} \leq 1$ completes the proof of the t-independent case. When h additionally depends on t, the same scheme can be applied with $v\cdot\nabla_x$ replaced by $\partial_t + v\cdot\nabla_x$. Let $\Omega := [0,\tau]\times\mathbb{T}$. Let $\phi \in H^1_0(\Omega)$ and compute

$$\begin{aligned}
&\int_0^\tau\int_{\mathbb{T}} \partial_t\rho_h\phi\, dx\, dt = \\
&= \int_0^\tau\int_{\mathbb{T}} \partial_t\rho_h\phi\int |v|^2\, d\gamma_\alpha\, dx\, dt + \int_0^\tau\int_Q \nabla_x\rho_h\cdot\int v|v|^2\, d\gamma_\alpha\phi\, dx\, dt = \\
&= \int_0^\tau\int_{\mathbb{T}} \phi\Big(\int (\partial_t + v\nabla_x)\rho_h|v|^2\, d\gamma_\alpha\Big)\, dx\, dt = \\
&= \int_0^\tau\int_{\mathbb{T}} \phi\Big(\int (\partial_t + v\nabla_x)h|v|^2\, d\gamma_\alpha\Big)\, dx\, dt + \\
&\quad + \int_0^\tau\int_{\mathbb{T}} \phi\Big(\int (\partial_t + v\nabla_x)(\rho_h - h)|v|^2\, d\gamma_\alpha\Big)\, dx\, dt \leq \\
&\leq \|(\partial_t + v\nabla_x)h\|_{L^2(\Omega;H^{-1}_\alpha)}\||v|^2\|_{H^1_\alpha}\|\phi\|_{L^2(\Omega)} + \\
&\quad + \int_0^\tau\int_{\mathbb{T}} \Big(\int (\rho_h - h)|v|^2\, d\gamma_\alpha\Big)(-\partial_t\phi)\, dx\, dt + \\
&\quad + \int_0^\tau\int_{\mathbb{T}} \Big(\int v\cdot\nabla_x(\rho_h - h)|v|^2\, d\gamma_\alpha\Big)(\phi)\, dx\, dt \leq \\
&\leq \|(\partial_t + v\nabla_x)h\|_{L^2(\Omega;H^{-1}_\alpha)}\||v|^2\|_{H^1_\alpha}\|\phi\|_{L^2(\Omega)} + \\
&\quad + \|h - \rho\|_{L^2(\Omega;d\gamma_\alpha)}\||v|^2\|_{L^2_\alpha}\|\partial_t\phi\|_{L^2(\Omega)} + \\
&\quad + \int_0^\tau\int_{\mathbb{T}} -\nabla_x\phi\cdot\int v|v|^2(\rho_h - h)\, d\gamma_\alpha \leq \\
&\leq \|(\partial_t + v\nabla_x)h\|_{L^2(\Omega;H^{-1}_\alpha)}\||v|^2\|_{H^1_\alpha}\|\phi\|_{L^2(\Omega)} + \\
&\quad + \|h - \rho_h\|_{L^2({L^2}_\alpha)}\||v|^2\|_{L^2_\alpha}\|\partial_t\phi\|_{L^2(\Omega)} + \\
&\quad + d\|\nabla_x\phi\|_{L^2}\|v_1|v|^2\|_{L^2_\alpha}\|h - \rho_h\|_{L^2({L^2}_\alpha)}.
\end{aligned}$$

Hence, $\|\partial_t\rho_h\|_{H^{-1}(\Omega)}$ is controlled by

$$\|(\partial_t + v\nabla_x)h\|_{L^2(\Omega;H^{-1}_\alpha)}\||v|^2\|_{H^1_\alpha} + \|h - \rho_h\|_{L^2({L^2}_\alpha)}(\||v|^2\|_{L^2_\alpha} + d\|v_1|v|^2\|_{L^2_\alpha}).$$

A very similar argument gives

$$\|\nabla_x\rho_h\|_{H^{-1}} \leq \|v\|_{H^1_\alpha} + (d^2+1)\|v\|_{L^2_\alpha}\|h - \rho_h\|_{L^2(\Omega;L^2_\alpha)}.$$

The two estimates together yield the thesis with

$$d_\alpha = \big(2(d\||v|^2 v_1\|_{L^2_\alpha} + \||v|^2\|_{L^2_\alpha})^2 + (d+1)^2\|v\|^2_{L^2_\alpha}\big).$$

□

3.3.2 A generalized Poincaré inequality

The next *a priori* estimate is at the core of the method. It is a modified Poincaré inequality in t, x and v which relies on Lemma 42 and involves derivatives of various orders.

Proposition 43. *Let $L>0$, $\tau>0$, $\Omega=(0,\tau)\times(0,L)^d$, and $\alpha>0$. Then, for all $h\in\mathrm{H}_{\mathrm{kin}}$ with zero average, we have that*

$$\|h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)}\le C\Big(\|h-\rho_h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)}+\|\partial_t h+v\cdot\nabla_x h\|^2_{\mathrm{L}^2(\Omega;H_\alpha^{-1})}\Big) \tag{3.14}$$

with $C=1+C_L\,d_\alpha$, where C_L and d_α are given respectively by (3.10) *and* (3.13).

Proof. By orthogonality in $\mathrm{L}^2(\Omega;\mathrm{L}^2_\alpha)$ and because $d\gamma_\alpha$ is a probability measure, we have the decomposition

$$\|h\|^2_{\mathrm{L}^2(\Omega;\mathrm{L}^2_\alpha)}=\|h-\rho_h\|^2_{\mathrm{L}^2(\Omega;\mathrm{L}^2_\alpha)}+\|\rho_h\|^2_{\mathrm{L}^2(\Omega)}.$$

The function ρ_h has zero average on Ω by construction, so that

$$\|h\|^2_{\mathrm{L}^2(\Omega;\mathrm{L}^2_\alpha)}\le\|h-\rho_h\|^2_{\mathrm{L}^2(\Omega;\mathrm{L}^2_\alpha)}+C_L\,\|\nabla_{x,t}\rho\|^2_{\mathcal{H}^{-1}(\Omega)}$$

by Lemma 38. Hence

$$\|h\|^2_{\mathrm{L}^2(\Omega;\mathrm{L}^2_\alpha)}\le(1+C_L\,d_\alpha)\,\|h-\rho_h\|^2_{\mathrm{L}^2(\Omega;d\gamma_\alpha)}+C_L\,d_\alpha\,\|\partial_t h+v\cdot\nabla_x h\|^2_{\mathrm{L}^2(\Omega;H_\alpha^{-1})}$$

by Lemma 42. This concludes the proof. □

3.4 (Super)linear local equilibria: exponential decay rate

In this Section, we consider the case $\alpha\ge1$ and the domain $\Omega=(t,t+\tau)\times(0,L)^d$, for an arbitrary $t\ge0$. Let us define $\kappa_\alpha:=(1+C_L\,d_\alpha)(P_\alpha+1)$ where P_α is the Poincaré constant in (3.7) and where C_L and d_α are given respectively by (3.10) and (3.13).

Lemma 44. *Let $L>0$, $\tau>0$, $t\ge0$, $\Omega_t=(t,t+\tau)\times(0,L)^d$, and $\alpha\ge1$. Then, for all $h\in\mathrm{H}_{\mathrm{kin}}$ with zero average which solve* (3.2)*, we have that*

$$\|h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)}\le\kappa_\alpha\,\|\nabla_v h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)}. \tag{3.15}$$

Proof. We know that

$$\|h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)}\le(1+C_L\,d_\alpha)\Big(P_\alpha\,\|\nabla_v h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha))}+\|\partial_t h+v\cdot\nabla_x h\|^2_{\mathrm{L}^2(H_\alpha^{-1})}\Big)$$

as a consequence of (3.7) and (3.14). Then (3.15) follows from Lemma 40. □

We are ready to prove Theorem 31 with an explicit estimate of the constant λ.

Proof of Theorem 31. Inequality (3.15) – on the interval $(t,t+\tau)$ – gives

$$\int_t^{t+\tau}\|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds\le\kappa_\alpha\int_t^{t+\tau}\|\nabla_v h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds.$$

With $\lambda = 2/\kappa_\alpha$, we deduce from Lemma 41 that

$$\frac{d}{dt}\int_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds = -2\int_t^{t+\tau} \|\nabla_v h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds$$
$$\le -\lambda \int_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds.$$

Grönwall's Lemma and the monotonicity of $t \mapsto \|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}$ imply

$$\int_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \le \int_0^{\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds\, e^{-\lambda t} \le \tau\, \|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\, e^{-\lambda t}$$

for ant $t \ge 0$, which proves (3.3), that is, Theorem 31. □

Indeed, the estimate for λ is explicit, as we state in the following.

Proposition 45. *For any $\alpha \ge 1$, Theorem 31 holds true with*

$$\frac{1}{\lambda} = \frac{1}{\tau}\left(\tau + \sqrt{d\,L^2 + \tau^2}\right)\left(2\,d_\alpha\,|\mathbb{S}^{d-1}|\,(P_\alpha + 1)\right).$$

Notice that the r.h.s. vanishes as $\tau \downarrow 0$, which is expected because of the degeneracy of Δ_α: an exponential decay rate of $\|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}$ cannot hold. The section is concluded showing how the result above yields the classical hypocoercivity estimate of Corollary 32.

Proof of Corollary 32. For any $t \ge 0$, we know from Theorem 31 that

$$\|h(t+\tau,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)} \le \fint_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \le \|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\, e^{-\lambda t},$$

as a consequence of the monotonicity of the L^2 norm, according to Lemma 41. This proves that

$$\|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)} \le C\,\|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\, e^{-\lambda t}$$

with $C = e^{\lambda\tau}$ for any $t \ge \tau$. However, if $t \in [0,\tau)$, it turns out that $C\,e^{-\lambda t} \ge 1$ so that the inequality is also true by Lemma 41. This concludes the proof. □

3.5 Sublinear equilibria: algebraic decay rates

3.5.1 Proof of the decay estimate

Assume that $\alpha \in (0,1)$. Let us define the parameter $\beta = 2\,(1-\alpha)/p$ where p, $q > 1$ are Hölder conjugate exponents, *i.e.*, $\frac{1}{p} + \frac{1}{q} = 1$ and define

$$Z_h(t) := \int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\,|h - \rho_h|^2\,dx\,d\gamma_\alpha. \tag{3.16}$$

The following estimates replace Proposition 43.

Proposition 46. *Let $L > 0$, $\tau > 0$, $t \ge 0$, $\Omega_t = (t,t+\tau)\times(0,L)^d$, and $\alpha \ge 1$. With the above notations, for all $h \in \mathrm{H}_{\mathrm{kin}}$ with zero average, we have that*

$$\|h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)} \le C\,P_\alpha^{\frac{1}{p}}\,\|\nabla_v h\|^{\frac{2}{p}}_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)}\left(\int_t^{t+\tau} Z_h(s)\,ds\right)^{\frac{1}{q}} + C\,\|\partial_t h + v\cdot\nabla_x h\|^2_{\mathrm{L}^2(\Omega;H_\alpha^{-1})}$$

where $C = 1 + C_L\, d_\alpha$ is as in Proposition 43 and P_α denotes the constant in the weighted Poincaré inequality (3.8).

Proof. Using (3.14) and Hölder's inequality w.r.t. the variable v, we find that

$$\|h-\rho_h\|^2_{\mathrm{L}^2_\alpha} \le \left(\int_{\mathbb{R}^d} \langle v\rangle^{-\beta p}\, |h-\rho_h|^2\, d\gamma_\alpha\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^d} \langle v\rangle^{\beta q}\, |h-\rho_h|^2\, d\gamma_\alpha\right)^{\frac{1}{q}}.$$

The weighted Poincaré inequality (3.8) with $\beta\, p = 2\,(1-\alpha)$ and an additional Hölder inequality w.r.t. the variables t and x allow us to complete the proof. □

Lemma 47. *Let $L>0$, $\tau>0$, $t\ge 0$, $\Omega_t = (t, t+\tau)\times(0,L)^d$, and $\alpha\in(0,1)$. There is a constant $W>0$ such that, for all solution $h\in\mathrm{H}_{\mathrm{kin}}$ to (3.2) with an initial datum h_0 with zero average, using the notation (3.16) as in Proposition 46, we have*

$$Z_h(t) \le W \int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, h_0^2\, dx\, d\gamma_\alpha, \quad \forall\, t\ge 0.$$

Proof. An elementary computation shows that

$$\begin{aligned}\int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, |h-\rho_h|^2\, dx\, d\gamma_\alpha &\le 2\int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, \big(h^2+\rho_h^2\big)\, dx\, d\gamma_\alpha \\ &\le 2\left(1+\textstyle\int_{\mathbb{R}^d} \langle v\rangle^{\beta q}\, d\gamma_\alpha\right) \int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, h^2\, dx\, d\gamma_\alpha\end{aligned}$$

because $\rho_h^2 = \left(\int_{\mathbb{R}^d} h\, d\gamma_\alpha\right)^2 \le \int_{\mathbb{R}^d} h^2\, d\gamma_\alpha \le \int_{\mathbb{R}^d} \langle v\rangle^{\beta q}\, h^2\, d\gamma_\alpha$. According to [66, Proposition 4], there is a constant $\mathcal{K}_{\beta q} > 1$ such that

$$\int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, |h(t,x,v)|^2\, dx\, d\gamma_\alpha \le \mathcal{K}_{\beta q} \int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, h_0^2\, dx\, d\gamma_\alpha, \quad \forall\, t\ge 0.$$

The result follows with $W = 2\left(1+\int_{\mathbb{R}^d} \langle v\rangle^{\beta q}\, d\gamma_\alpha\right) \mathcal{K}_{\beta q}$. □

Assume that $h\in\mathrm{H}_{\mathrm{kin}}$ solves (3.2) with an initial datum h_0 with zero average and let us collect our estimates. With Proposition 40, Proposition 46, and Lemma 47, the estimate of Lemma 44 is replaced by

$$\|h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)} \le A\, \|\nabla_v h\|^{\frac{2}{p}}_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)} + C\, \|\nabla_v h\|^2_{\mathrm{L}^2(dt\,dx\,d\gamma_\alpha)} \tag{3.17}$$

with $A = C\, P_\alpha^{\frac{1}{p}}\, (\tau\, W)^{1/q} \left(\int_{\mathbb{T}\times\mathbb{R}^d} \langle v\rangle^{\beta q}\, h_0^2\, dx\, d\gamma_\alpha\right)^{1/q}$.

The main result of the section is a technical version of Theorem 33. Let

$$\mathsf{x}(t) := \fint_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\, ds \quad\text{and}\quad \mathsf{y}(t) := \fint_t^{t+\tau} \|\nabla_v h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\, ds$$

where norms are taken on $\mathbb{T}\times\mathbb{R}^d$. We know from Lemma 41 and (3.17) that

$$\mathsf{x}' = -2\,\mathsf{y} \quad\text{and}\quad \mathsf{x} \le \varphi(\mathsf{y}) := A\,\mathsf{y}^{1/p} + C\,\mathsf{y}.$$

Finally, let us denote by φ^{-1} the inverse of $\mathsf{y}\mapsto\varphi(\mathsf{y})$ and consider

$$\psi(\mathsf{z}) := \int_{\mathsf{z}}^{\mathsf{x}_0} \frac{dz}{2\,\varphi^{-1}(z)} \quad\text{with}\quad \mathsf{x}_0 = \|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}.$$

Theorem 48. *Let $L > 0$, $\tau > 0$, $t \geq 0$, $\Omega_t = (t, t+\tau) \times (0, L)^d$, and $\alpha \in (0,1)$. With the above notations, for all solution $h \in \mathrm{H}_{\mathrm{kin}}$ to (3.2) with an initial datum h_0 with zero average, we have*

$$\mathsf{x}(t) = \fint_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \leq \psi^{-1}(t), \quad \forall\, t \geq 0.$$

Proof. The strategy goes as in [198, 66]. Everything reduces to the differential inequality

$$\mathsf{x}' \leq -2\,\varphi^{-1}(\mathsf{x})$$

using the monotonicity of $\mathsf{y} \mapsto \varphi(\mathsf{y})$. From by the elementary Bihari-Lasalle inequality, see [50, 186], which is obtained by a simple integration, we obtain

$$\mathsf{x}(t) = \fint_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \leq \psi^{-1}\big(t + \psi(\mathsf{x}(0))\big).$$

Since, on the one hand

$$\mathsf{x}(0) = \fint_0^{\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \leq \|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)} = \mathsf{x}_0$$

because $s \mapsto \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}$ is nonincreasing according to Lemma 41, and ψ is nonincreasing on the other hand, then

$$\psi^{-1}\big(t + \psi(\mathsf{x}(0))\big) \leq \psi^{-1}(t),$$

which concludes the proof. Notice that the dependence on h_0 enters in A and x_0, and henceforth in φ and ψ. □

Proof of Theorem 33. Since $\lim_{t\to+\infty} \mathsf{y}(t) = 0$, we have that $\varphi(\mathsf{y}(t)) \sim A\,\mathsf{y}(t)^{1/p}$ as $t \to +\infty$, which heuristically explains the role played by p in (3.5). This can be made rigorous as follows. Notice that

$$\varphi(\mathsf{y}) = A\,\mathsf{y}^{1/p} + C\,\mathsf{y} \leq A_0\,\mathsf{y}^{1/p}, \quad \forall\,\mathsf{y} \leq \mathsf{y}_0, \quad \text{with} \quad A_0 = A + C\,\mathsf{y}_0^{1-1/p}.$$

With A replaced by A_0 and C replaced by 0, the computation of the proof of Theorem 48 is now explicit. With the choice $\mathsf{y}_0 = \varphi^{-1}(\mathsf{x}_0)$, we know that $\mathsf{y}(t) \leq \mathsf{y}_0$ for any $t \geq 0$ and obtain

$$\mathsf{x}(t) \leq \left(\mathsf{x}_0^{1-p} + 2\,(p-1)\,A_0^{-p}\,t\right)^{-\frac{1}{p-1}}, \quad \forall\, t \geq 0. \tag{3.18}$$

Using $\mathsf{x}_0 = A\,\mathsf{y}_0^{1/p} + C\,\mathsf{y}_0 \geq C\,\mathsf{y}_0$, we know that $A_0 = \mathsf{x}_0\,\mathsf{y}_0^{-1/p} \leq A + C^{1/p}\,\mathsf{x}_0^{1-1/p}$, which proves (3.5) with

$$K = \max\left\{1, \big(2\,(p-1)\big)^{1/(1-p)}\left(C\,P_\alpha^{1/p}\,(\tau\,W)^{1-1/p} + C^{1/p}\right)\right\}.$$

The conclusion holds using $\sigma = \beta\,q = 2\,(1-\alpha)/(p-1)$. □

3.5.2 The linear threshold: from algebraic to exponential rates

A very natural question arises: *is the result Theorem 45 (corresponding to $\alpha \in (0,1)$ consistent with the result of Theorem 48 (which covers any $\alpha \geq 1$) ?* A first observation is that we can vary α in the assumptions concerning the initial data.

Lemma 49. *If $h_0 \in \mathrm{L}^2(\mathbb{T}; \mathrm{L}^2_{\alpha_0})$ for some $\alpha_0 \in (0,1)$, then $\langle v\rangle^{\sigma/2}\,h_0 \in \mathrm{L}^2(\mathbb{T}; \mathrm{L}^2_\alpha)$ for any $\alpha > \alpha_0$ and any $\sigma > 0$.*

The proof is a simple consequence of the fact that $v \mapsto \langle v\rangle^\sigma \exp\langle v\rangle^{\alpha-\alpha_0}$ is uniformly bounded. For any $\alpha > \alpha_0$, let us denote the corresponding solution of (3.2) with initial datum h_0, of zero average, by $h^{(\alpha)}$.

If $\alpha \in (\alpha_0, 1)$, then (3.18) can be rewritten as

$$\fint_t^{t+\tau} \|h^{(\alpha)}(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,ds \le \|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\big(1+(p-1)\,\ell(\alpha)\,t\big)^{-\frac{1}{p-1}}.$$

By passing to the limit as $\alpha \to 1^-$, we recover (3.3) with

$$\lambda = \lim_{\alpha\to1^-} \ell(\alpha),$$

where $\ell(\alpha) = 2\,\|h_0\|^{2(p-1)}_{\mathrm{L}^2(dx\,d\gamma_\alpha)}\,A_0^{-p}$ and $A_0 = A_0(\alpha)$ as above. The Poincaré constant P_α in inequality (3.8) admits a limit as $\alpha \to 1^-$, according to [67, Appendix A].

The limit of $\lim_{\alpha\to1^-}\ell(\alpha)$ is certainly not optimal. By working directly on the Bihari-Lasalle estimate of Theorem 48, we can recover the value of λ in Theorem 45. Notice here that $\sigma > 0$ plays essentially no role and can be taken arbitrarily small, even depending on α, but such that $p = 2\,(1-\alpha)/\sigma \to 1$ as $\alpha \to 1^-$.

As in [85, 180], it is possible to obtain improved decay rates in (3.5) by picking the initial datum in a smaller space. Typically, the control of additional norms or moments is asked. However, the strategy in the current paper is in the opposite direction. If $\alpha \in (0,1)$ we are interested in taking the initial data in a space as large as possible so that we can compute decay rates. The additional conditions to be imposed have been shown to vanish as $\alpha \to 1^-$.

3.6 Hypocoercivity and comparison with some other methods

3.6.1 An explicit hypocoercivity result

Theorem 31 implies an L^2–hypocoercivity result in the linear and superlinear regimes $\alpha \ge 1$, see Corollary 32. The remainder of this section is devoted to a comparison with earlier hypocoercivity results in a simple benchmark case: let $\alpha = 2$, $d = 1$ and $L = 2\pi$. In this case, for the choice $\tau = 2\pi$, Theorem 31 amounts to

$$\|h\|^2_{\mathrm{L}^2(dx\,d\gamma_2)} \le e^{\frac{\pi}{8\sqrt{3}}}\,\|h_0\|^2_{\mathrm{L}^2(dx\,d\gamma_2)}\,e^{-\frac{t}{8\sqrt{3}}}, \quad \forall\, t \ge 0.$$

For sake of comparison, notice that $\lambda = 1/(8\sqrt{3}) \approx 0.0721688$. Even if we are aware of explicit or sharp results in other metrics than L^2 for (3.2), as [116, 215], we restrict our discussion to L^2 hypocoercivity methods. The paper [117] covers several cases, but it would be not trivial to extract an explicit rate to compare with our method.

3.6.2 The L^2 hypocoercivity method

The first comparison is with the abstract twisted L^2 hypocoercivity method of [127, 67]. Let $\|\cdot\|$ be the norm of $\mathrm{L}^2(dx\,d\gamma_2)$ and $(\cdot,\cdot)$ the associated scalar product. We consider the evolution equation

$$\partial_t h + T h = \mathcal{L} h. \tag{3.19}$$

Theorem 50. *Let h be a solution of (3.19) with initial datum $h_0 \in \mathrm{L}^2(Q;\mathrm{L}^2_2)$ and assume that T and $\mathcal{L}$ are respectively anti-self-ajoint and self-adjoint operators on $\mathrm{L}^2(Q;\mathrm{L}^2_2)$ such that, for some positive constants λ_m,*

λ_M and C_M, we have

(A1) $(-\mathcal{L}h, h) \geq \lambda_m \|(1-\Pi)\, h\|^2$ *for all* $h \in D(\mathcal{L})$,

(A2) $\|T\,\Pi\, h\|^2 \geq \lambda_M \|\Pi\, h\|^2$ *for all* $h \in D(T\,\Pi)$,

(A3) $\Pi\, T\,\Pi\, h = 0$,

(A4) $\|A\,T\,(1-\Pi)\,h\| + \|A\,\mathcal{L}\,h\| \leq C_M \|(\mathrm{Id}-\Pi)\,h\|$ *for all admissible* $h \in \mathrm{L}^2(Q;\mathrm{L}^2_2)$

where $A := \big(\mathrm{Id} + (T\,\Pi)^*\,T\,\Pi\big)^{-1}\,(T\,\Pi)^*$ *and* Π *is the projection in* L^2_2 *onto the kernel of* $\mathcal{L}$*. Then we have*

$$\|h(t,\cdot,\cdot)\|^2 \leq C\,\|h_0\|^2\,e^{-\lambda\,t} \quad \forall\, t \geq 0$$

with $C = (1+\delta)/(1-\delta)$, $\delta = \frac{1}{2}\,\min\left\{1, \lambda_m, \frac{\lambda_m\,\lambda_M}{(1+\lambda_M)\,C_M^2}\right\}$ *and* $\lambda = \frac{2\,\delta\,\lambda_M}{3\,(1+\lambda_M)}$.

This result is taken from [67, Proposition 4]. According to [67, Corollary 9], we have the estimate $\lambda = 1/24 \approx 0.041667$. A minor improvement is obtained as follows. Using Theorem 50 applied with $d = 1$, $L = 2\pi$, $T = v\,\partial_x$ and $\mathcal{L} = \partial_v^2 - v\,\partial_v$, in Fourier variables, we obtain $\lambda_m = \lambda_M = 1$ and $C_M = (1+\sqrt{3})/2$ according to [16, Section II.1.3.2], so that $\lambda = 1/(12+6\sqrt{3}) \approx 0.0446582$. Using Fourier modes, a slightly better estimate is obtained from [16, Section II.1.2] with $\lambda \approx 0.176048$.

3.6.3 Direct spectral methods

In a series of papers, F. Achleitner, A. Arnold, E. Carlen and several other collaborators use direct spectral methods. We refer in particular to [18, 2, 3, 4] and also [16] for an introduction to the method, which can be summarized as follows.

Let us consider (3.19) written after a Fourier transform in x, so that $T = i\,\xi\cdot v$, and acting on $\mathrm{L}^2(dx;\mathrm{L}^2_2)$ now considered as a space of complex valued functions. Assume that for some positive definite bounded Hermitian operator P and some constant $\lambda \in (0,+\infty)$, we have

$$(L-T)^*\,P + P\,(L-T) \geq 2\,\lambda\,P.$$

Let us consider the *twisted* norm $\|\hat f\|_P^2 := \int_Q (\hat f, P\,\hat f)\,dx$ where $(\cdot,\cdot)$ is the natural extension of the scalar product as defined in (3.6). From

$$\frac{d}{dt}\|\hat f\|_P^2 = -\langle \hat f, ((L-T)^*\,P + P\,(L-T))\,\hat f\rangle \leq -2\,\lambda\,\|\hat f\|_P^2,$$

for some $C > 1$, we deduce that

$$C^{-1}\,\|f(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_\alpha)} \leq \|\hat f(t,\cdot,\cdot)\|_P^2 \leq e^{-2\,\lambda\,t}\,\|\hat f_0\|_P^2 \quad \forall\, t \geq 0.$$

To our knowledge, μ has not yet been computed in the case of (3.1). The spectral decomposition

$$h(t,x,v) = \sum_{\xi\in\mathbb{Z}^d}\sum_{k\in\mathbb{N}^d} a_{\xi,k}(t)\,H_k(v)\,\mathrm{e}^{-i\,\frac{2\pi}{L}\,\xi\cdot x}$$

provides an easy framework for finite dimensional approximations using the basis of Hermite functions $(H_k)_{k\in\mathbb{N}}$ and the numerical value $\mu \approx 0.4$ has been obtained according to [1].

Proposition 45	$\lambda \approx 0.07$
DMS [127]	$\lambda \approx 0.04$
ADSW [16]	$\lambda \approx 0.17$
Achleitner (numerics) [1]	$\lambda \approx 0.4$

Table 3.1: Comparison among different L^2 hypocoercivity methods

3.6.4 Comparison for decay rates in limit regimes

Let $\alpha \geq 1$. Corollary 32 provides us with a decay estimate depending on the parameter L, which represents the length of the spatial domain $\mathbb{T}$. Note that (3.10) is meaningful if $0 < \tau < L$. We shall now consider two situations, corresponding to $L \to \infty$, where spatial diffusion dominates, and to $L \to 0$, where the dominant term is the collision operator Δ_α. In the first case, we have that the decay exponent

$$\lambda \approx \frac{\tau}{L^3} \to 0, \qquad \text{as } L \to \infty.$$

The hypocoercivity constant $C \approx 1$. Hence, exponential decay is lost in the limit. On the other hand, for $L \to 0$, we have

$$\lambda \approx \frac{1}{4(P_\alpha + 1)\,\|v_1|v|^2\|^2_{L^2_\alpha}} \approx 0.04,$$

if $\alpha = 1$. Moreover,

$$C \approx 1.$$

This rate has the wrong order once compared to [16], where the authors recover the value

$$\lambda \approx 1 - \sqrt{3/7}.$$

Our inaccuracy is mainly due to the incompatibility between (3.10) and Lemma 42. Moreover, the value of the Lions constant in (3.10) is just an estimate and it is not expected to be as accurate as something achieved by a spectral method (even if its scaling is correct).

Chapter 4

How to construct decay rates for kinetic Fokker-Planck equations?

This chapter corresponds to [P6], in collaboration with Gabriel Stoltz.

Abstract

This is a preliminary work. An update to a definitive version will follow. We study time averages for the norm of the solution to kinetic Fokker–Planck equations associated with general Hamiltonians. We provide explicit and constructive decay estimates, allowing fat-tail, sub-exponential and (super-)exponential local equilibria, which also include the classical Maxwellian case. The evolution is subject to a transport operator induced by an external space-confining force. The key step in our estimates is a modified Poincaré inequality, obtained via a Lions–Poincaré inequality and an averaging lemma.

4.1 Introduction

This work presents an extension of the hypocoercive approach studied in [P1] to the case of a non-zero potential energy function, and also considers the case of general kinetic energies beyond the standard quadratic one. Such generalizations can be relevant to enhance the performance of certain sampling methods in molecular dynamics and Bayesian statistics; see for instance [243, 200]. The more general framework also allows to make more transparent the structural assumptions behind the algebraic computations of [79, 5].

Kinetic models describe interacting multi-agent or multi-particle systems at an intermediate level between full microscopic and macroscopic scales [28]. More precisely, a kinetic partial differential equation (PDE) encodes the time evolution of the *distribution function* of particles with respect to *position* and some *momentum* variable. Since the XIXth century, kinetic PDEs are formulated as a coupling between a transport and a collision/relaxation term. This is the case of the celebrated Boltzmann equation, for example. A general theory for Boltzmann's equation is still open. However, due to its physical validity, various simplified versions of it have been studied. A remarkable one is the Vlasov–Fokker–Planck equation, derived from Boltmann's equation in a limiting regime corresponding to grazing collisions [109]. Vlasov–Fokker–Planck equations date back at least to [183], and they are useful models in statistical physics [28] and stellar dynamics [92]. In general, Vlasov–Fokker–Planck equations should be understood as models for the time evolution of a distribution of particles subject to a potential and random background force inducing relaxation towards equilibrium. This paper is concerned with a rather general class of Vlasov–Fokker–Planck equations we introduce below.

We consider systems described by their configuration (x, v), where $x \in \mathcal{X}$ is a spatial variable representing the positions of the particles composing the system, and $v \in \mathcal{V}$ their velocities. Typical choices for the position space are $\mathcal{X} = \mathbb{R}^d$ for unconfined systems on the full space, or $\mathcal{X} = (L\mathbb{T})^d$ (with $\mathbb{T} = \mathbb{R}\backslash\mathbb{Z}$ the unit torus) for systems enclosed in a cubic box of size $L > 0$ with periodic boundary conditions, a situation typical in molecular dynamics. The velocity space usually is $\mathcal{V} = \mathbb{R}^d$, but here as well one could consider $\mathcal{V} = \mathbb{T}^d$. In the latter case, and more generally when considering non-quadratic kinetic energies, one should rather think of v as being a momentum variable, the variable conjugated to the positions, rather than a velocity. The Hamiltonian associated with the system is supposed to be separable, and is therefore the sum $\phi(x) + \psi(v)$ of the potential energy function $\phi : \mathcal{X} \to \mathbb{R}$ and the kinetic energy $\psi : \mathcal{V} \to \mathbb{R}$. We introduce the following measures:

$$\mu(dx) := \mathrm{e}^{-\phi(x)}\,dx, \qquad \gamma(v) = \mathrm{e}^{-\psi(v)}\,dv.$$

In order to have an invariant Boltzmann–Gibbs probability measure

$$\Theta(dx\,dv) = \mu(dx)\gamma(dv).$$

we shall assume that $\mathrm{e}^{-\phi} \in L^1(\mathcal{X})$ and $\mathrm{e}^{-\psi} \in L^1(\mathcal{V})$, and, without loss of generality, that $\mathrm{e}^{-\phi}$ and $\mathrm{e}^{-\psi}$ integrate to 1, upon adding a constant to to ϕ, ψ. The Vlasov–Fokker–Planck equation we study reads

$$\begin{cases} \partial_t f + \nabla_v \psi \cdot \nabla_x f - \nabla_x \phi \cdot \nabla_v f = \xi \nabla_v \cdot \left(\gamma \nabla_v \left(\dfrac{f}{\gamma}\right)\right), \\ f(t=0,\cdot,\cdot) = f_0 \in \mathrm{L}^2(\Theta^{-1}), \end{cases} \tag{VFP}$$

where $\xi > 0$ is a parameter regulating the intensity of the collision/relaxation term. We call

$$Tf := \nabla_v \psi \cdot \nabla_x f - \nabla_x \phi \cdot \nabla_v f, \qquad \mathscr{S} f := \nabla_v \cdot \left(\gamma \nabla_v \left(\frac{f}{\gamma}\right)\right),$$

the transport and the collision operator, respectively.

In order to highlight the role of ξ, we briefly discuss a microscopic derivation of (VFP), seen as the PDE governing the evolution of the law of Langevin dynamics. The Langevin dynamics associated with the Hamiltonian $\phi(x) + \psi(v)$ read

$$\begin{cases} dX_t = \nabla_v \psi(V_t)\,dt, \\ dV_t = -\nabla_x \phi(X_t)\,dt - \xi \nabla_v \psi(V_t)\,dt + \sqrt{2\xi}\,dW_t, \end{cases} \tag{4.1}$$

where $\xi > 0$ is the friction coefficient and $(W_t)_{t \geq 0}$ is a standard d-dimensional Brownian motion. Let us motivate that, as for systems with quadratic kinetic energies, we expect the relaxation time of (4.1) to the stationary state to be of order $\max(\xi, \xi^{-1})$. In the limit $\xi \to 0$, Langevin dynamics reduce to Hamiltonian dynamics. For small frictions, the average energy evolves on timescales of order ξ^{-1} since, by Itô calculus,

$$\frac{d}{dt}\mathbb{E}[H(X_t, V_t)] = \mathbb{E}[(\mathcal{L}_\xi H)(X_t, V_t)] = \xi \mathbb{E}[(\mathcal{S} H)(X_t, V_t)],$$

where $\mathcal{L}_\xi = T + \xi \mathcal{S}$ is the generator of (4.1). The other limit of interest is the overdamped regime $\xi \to +\infty$, where one recovers the overdamped Langevin dynamics upon rescaling time by a factor ξ:

$$d\overline{X}_t = -\nabla_x \phi(\overline{X}_t)\,dt + \sqrt{2}\,dB_t,$$

where $(B_t)_{t\geq 0}$ is also a standard d-dimensional Brownian motion. The formal proof of this limit relies on the observation that solutions of (4.1) satisfy (upon replacing the term $-\xi\nabla_v\psi(V_t)\,dt$ in the second line of (4.1) by dX_t and integrating in time from 0 to ξt)

$$\begin{aligned} X_{\xi t} - X(0) &= \frac{V_0 - V_{\xi t}}{\xi} - \frac{1}{\xi}\int_0^{\xi t} \nabla_x\phi(X_s)\,ds + \sqrt{\frac{2}{\xi}}\,dW_{\xi s} \\ &= \frac{V_0 - V_{\xi t}}{\xi} - \int_0^{t} \nabla_x\phi(X_{\xi s})ds + \sqrt{\frac{2}{\xi}}\,dW_{\xi s}, \end{aligned}$$

where $(\xi^{-1/2}W_{\xi t})_{t\geq 0}$ still is a standard d-dimensional Brownian motion. This suggests to introduce the limiting process $(\overline{X}_t)_{t\geq 0}$ obtained from $(X_{\xi t})_{t\geq 0}$ as $\xi \to +\infty$, and highlights the fact that the relaxation time is of order ξ in this regime.

It is in fact more convenient to perform a change of unknown function and write $f = h\Theta$, in which case we expect h to converge in the longtime limit to the constant function with values 1. The function h satisfies the following kinetic Ornstein–Uhlenbeck equation:

$$\begin{cases} \partial_t h + \nabla_v\psi\cdot\nabla_x h - \nabla_x\phi\cdot\nabla_v h = -\xi\,\nabla_v^\star\nabla_v h, \\ h(t=0,\cdot,\cdot) = h_0 \in \mathrm{L}^2(\Theta), \end{cases} \tag{VOU}$$

where $A^\star$ denotes the adjoint of a closed operator on $\mathrm{L}^2(\Theta)$. Equation (VOU) is linear and mass-preserving, hence we can assume the initial data to be of average zero with respect to Θ and study the long-time behaviour of the corresponding solutions to (VOU). We expect that

$$\lim_{t\to\infty} \|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} = 0.$$

The goal of this work is to make this limit quantitative by obtaining effective decay rates.

In the recent literature, the problem of making the last limit quantitative has been widely investigated. Equations such as (VOU) fall inside the abstract theory of [169]. From this founding paper, a recent theory originated, specialising on decay rates rather than regularity issues. This approach goes under the name of hypocoercivity. We briefly review the related literature, see [44] for a much more detailed account. Quantitative decay rates for some degenerate (but still hypoelliptic) diffusion equations appeared in [255, 256, 165]. The strategy there relied on twisting the H^1-norm in order to get an equivalent Lyapunov functional showing exponential decay in time along the flow of the equations under investigation. Then, exponential decay (at the same rate) could be recovered for the H^1-norm, with a pre-factor $C > 1$ in front arising from the twist between the reference norm and the Lyapunov functional.

Later, a L^2–setting for hypocoercivity was established, [164, 127]. The strategy of [127] was to twist the standard L^2-norm with a well-suited term in order to find an equivalent norm with a better dissipation estimate along kinetic PDEs. In addition, in this setting, one can study consistency with hydrodynamic limits and non-regularising kinetic PDEs with general initial data. The L^2-framework is general enough to treat a large class of models [67, 65, 66], and it can provide sharp estimates for some specific models [18, 2, 4, 3, 16]. Precise algebraic features of the framework and estimates with respect to the parameters defining the dynamics were studied in [44]. In particular, the scaling with respect to ξ has been also treated in [157, 152].

In order to complete this short review, let us mention also a few other works related to hypocoercive techniques, namely [134, 116] for an approach based on couplings, [33] for a version of the Bakry–Émery methods for hypocoercive dynamics, and [117] which recently studied a model with further degeneracy.

In the present paper, we focus on the approach of [5], whose techniques work directly in the standard L^2-norm. This is possible thanks to an adapted space-time-velocity Poincaré inequality. We notice that time-integrated functionals of the solutions to kinetic equations are in use since [161, 244] at least. The framework of [5] was qualitative, although, constructive estimates can be derived in special cases [P1, 79]. One idea common to both references is to exploit a version of a space–time Poincaré-Lions inequality [11, 84, 117], coupled with an explicit version of the computations in [5, Proposition 6.2].

In this paper we further generalise the techniques of [P1, 79], still keeping all estimates fully constructive. The main original contributions are:

1. Adapting the Armstrong–Mourrat technique, we get fully explicit convergence estimates for (VOU) with general potentials ϕ, ψ. In [P1] the spatial variable was confined in a torus with $\phi = 0$, while [79] only considers quadratic kinetic energies. Decay rates are shown to be exponential if ψ grows fast enough at infinity. Otherwise, algebraic decay rates are obtained.

2. The potentials ϕ we consider are rather general. In particular, we remove the condition that the embedding $H^1(\mu) \hookrightarrow L^2(\mu)$ is compact, see [79, Assumption 3].

3. Our estimates for Theorem 59 should be compared with those of [117]. While performed in a simpler geometry, our result is fully constructive, as we actually build a solution for the elliptic equation involved to obtain the result .

4. We establish a theory for Lyapunov functionals for (VOU), depending on the behaviour of ψ at infinity.

5. Decay estimates are fully explicit in the fluctuation–collision parameter ξ.

Some key-points of our method are the following.

- We use time averages of the L^2-norm of the solutions to (VOU) as entropies. Time averages allow for an effective entropy-entropy production estimate along the flow of (VOU), where the key step in the proof is an adapted time-space-velocity Poincaré inequality.

- If ψ grows at least linearly at infinity, the entropy production is strong enough to allow the use of a Gronwall lemma. Otherwise, the dissipation controls just a power of the time average of the norm. Then, an algebraic decay is obtained via a Bihari–Lasalle argument, combined with a uniform bound in time for a momentum of the solution.

- A fundamental ingredient is a space-time weighted Lions' inequality, whose proof is fully constructive. Indeed, the problem is reduced to a regularity estimate for some elliptic PDE, which is solved either by explicit computations or with the Lax–Milgram theorem, depending on the nature of the source term in the elliptic PDE.

Extensions.

As mentioned earlier, our aim in this work is to present a general methodology to obtain decay rates for Fokker–Planck type equations which are fully explicit in the various parameters of the dynamics, in particular the friction and the dimension. Our presentation is as general as possible, so that extensions and adaptations to various other dynamics can be worked out. We have in mind here the linear Boltzmann equation, the linearised Landau equation [232], Fokker–Planck equations with fractional Laplacians, generalized Langevin dynamics [227], Adaptive Langevin dynamics [193], etc. In fact, as our approach makes only a limited use of the linearity of the Fokker–Planck equation at hand, generalizations to nonlinear equations can be envisioned.

Outline of the paper.

This work is organized as follows. We start in Section 4.2 by presenting the problem we consider and stating the main results we establish. We give in Section 4.3 some preliminary results which are useful for the proofs of the main results. Section 4.4 is devoted to the proof of exponential decay estimates where the kinetic energy ψ grows linearly or faster at infinity. Section 4.5 proves algebraic decay rates for solutions to (VOU) driven by kinetic energies ψ which grow slower than linearly at infinity. Finally, weighted Lions estimates are proved in Section 4.6. Complementary results are collected in the Appendix.

4.2 Main results

This section contains the main assumptions (Section 4.2.1) and results (Ssections 4.2.2, 4.2.3, 4.2.4, and 4.2.5) of the paper.

4.2.1 Structural assumptions

In order to present our main convergence result, we introduce a time $\tau > 0$, and the associated uniform probability measure on the time interval $[0, \tau]$:

$$\mathrm{U}_\tau(dt) = \frac{1}{\tau}\mathbf{1}_{[0,\tau]}(t)\,dt.$$

With some abuse of notation we sometimes denote probability measures and their densities with the same symbol. We state the main assumptions we need for the potentials ϕ, ψ.

Assumption 2. *The function ϕ is smooth and $\mathrm{e}^{-\phi} \in \mathrm{L}^1(dx)$, with $\int_{\mathscr{X}} \mathrm{e}^{-\phi}\,dx = 1$. Moreover, for any time $\tau > 0$, there exists a constant $C_\tau^{\mathrm{Lions}} > 0$ such that the following Lions inequality holds true:*

$$\forall g \in \mathrm{L}^2(\mathrm{U}_\tau \otimes \mu), \quad \left\| g - \int_{[0,\tau]\times\mathscr{X}} g(t,x)\,\mathrm{U}_\tau(dt)\otimes\mu(dx)\right\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \leq C_\tau^{\mathrm{Lions}} \|\nabla_{t,x} g\|^2_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}. \tag{4.2}$$

Sufficient conditions for (4.2) to hold with explicit estimates of C_τ^{Lions} are discussed in Section 4.2.5.

Assumption 3. *The function ψ is smooth and $\mathrm{e}^{-\psi} \in \mathrm{L}^1(dv)$ with $\int_{\mathscr{V}} \mathrm{e}^{-\psi}\,dv = 1$. Moreover,*

$$\int_{\mathcal{V}} |\nabla_v \psi(v)|^4\,\gamma(dv) + \int_{\mathcal{V}} \left|\nabla_v^2 \psi(v)\right|^2 \gamma(dv) < +\infty, \quad \limsup_{R\to\infty} \int_{\mathscr{V}\setminus B(0,R)} |\nabla\psi|^{-2}\,\gamma(dv) = 0, \tag{4.3}$$

and the symmetric positive matrix

$$\mathscr{M} = \int_{\mathcal{V}} \nabla_v \psi \otimes \nabla_v \psi\,d\gamma \tag{4.4}$$

is definite.

The condition (4.4) is satisfied when $\mathcal{V} = \mathbb{R}^d$, see [243, Appendix A] for a proof. More generally, this condition holds as soon as there is no direction $\rho \in \mathbb{R}^d$ such that $\rho^\top \nabla\psi(v) = 0$ for all $v \in \mathcal{V}$. The integrability conditions (4.3) are rather easy to satisfy. The estimates hold for instance if $\mathcal{V}$ is bounded, or ψ increases polynomially or logarithmically at infinity. In the last case ψ shall grow fast enough so that γ is integrable and has a finite momentum of order two.

Remark 51. *The analysis we perform can straightforwardly be extended, at least from an algebraic viewpoint, to other types of degenerate symmetric dissipation operators, such as the one appearing in the linear Boltzmann equation, or involving fractional Laplacians, see [127, 65].*

Assumption 4. *The Lie algebra generated by ∇_v and $\nabla_v\psi \cdot \nabla_x - \nabla_x\phi \cdot \nabla_v$ is of dimension $2d$ at all points $(x, v) \in \mathcal{X} \times \mathcal{V}$.*

4.2.2 An averaging lemma

Averaging lemmata are a typical tool in kinetic theory to recover increased regularity on the velocity average of the distribution function. This is the idea behind our first result, which will be fundamental for the main theorems in Sections 4.2.3 and 4.2.4.

We make a further assumption on ϕ, which is satisfied (with explicit control on the constant) in many cases of interest, see [256, Lemma A.24], as well as [79, Lemma 2.2] and [44, Lemma 3.7] for a precise quantification of the dependence of the constant L_ϕ on the dimension, under certain structural assumptions on the potential ϕ. This is recalled in Lemma 61 below.

Assumption 5. *There exists a constant $L_\phi \in \mathbb{R}_+$ such that*

$$\|z\nabla\phi\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \leq L_\phi \|z\|_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}, \qquad \forall z \in \mathrm{H}^1(\mathrm{U}_\tau \otimes \mu). \tag{4.5}$$

A functional space which will be useful in the proofs is the kinetic space, defined as follows for a positive time $\tau > 0$:

$$\mathrm{H}^1_{\mathrm{kin}} = \left\{ h \in \mathrm{L}^2\left(\mathrm{U}_\tau \otimes \mu, \mathrm{H}^1(\gamma)\right) \,\middle|\, (\partial_t + T)h \in \mathrm{L}^2\left(\mathrm{U}_\tau \otimes \mu, \mathrm{H}^{-1}(\gamma)\right) \right\}.$$

We introduce the following projector on $\mathrm{L}^2(\mathrm{U}_\tau \otimes \Theta)$:

$$\Pi h = \int_{\mathcal{V}} h(\cdot,\cdot,v)\,\gamma(dv).$$

We denote by $\mathrm{L}^2_0(\mathrm{U}_\tau \otimes \Theta)$ the space of functions in $\mathrm{L}^2(\mathrm{U}_\tau \otimes \Theta)$ with zero average. Similar zero-average spaces are defined in the same way. We are now in position to state the following useful averaging lemma for (VOU).

Lemma 52. *Under Assumptions 3 and 5, there exists an explicit constant $K_{\mathrm{avg}} \in \mathbb{R}_+$ such that, for any $h \in \mathrm{H}^1_{\mathrm{kin}} \cap \mathrm{L}^2_0(\mathrm{U}_\tau \otimes \Theta)$,*

$$\|\nabla_{t,x}\Pi h\|^2_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)} \leq K_{\mathrm{avg}} \left(\|(\mathrm{Id} - \Pi)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} + \|(\partial_t + T)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^{-1}(\gamma))} \right).$$

An explicit upper bound on K_{avg} can be obtained, see (4.25).

4.2.3 Exponential convergence rates

We consider the longtime behaviour of solutions to (VOU). In particular, our main goal is to prove that

$$h(t,\cdot,\cdot) \xrightarrow[t\to\infty]{} \int_{\mathscr{X}\times\mathscr{V}} h_0(x,v)\,\Theta(dx\,dv),$$

for all solutions to (VOU), with quantitative and explicit convergence rates. As explained in Section 4.1, this reduces to studying the decay to 0 for solutions to (VOU) with initial condition

$$h_0 \in \mathrm{L}^2_0(\Theta) = \left\{ g \in \mathrm{L}^2(\Theta) \,\middle|\, \int_{\mathcal{X}\times\mathcal{V}} g\,d\Theta = 0 \right\}.$$

Other sub-spaces such as $\mathrm{L}^2_0(\gamma)$ are defined in a similar way. In order to achieve exponential decay rates, we need the following.

Assumption 6. *The following Poincaré inequality for γ holds: there exists $c_\psi > 0$ such that*

$$\forall g \in \mathrm{H}^1(\gamma)\cap \mathrm{L}^2_0(\gamma), \qquad \|g\|_{\mathrm{L}^2(\gamma)} \le c_\psi^{-1/2}\|\nabla_v g\|_{\mathrm{L}^2(\gamma)}. \tag{4.6}$$

A sufficient condition on ψ for the latter inequality to hold is that there exists a constant $a \in (0,1)$ such that

$$\liminf_{|v|\to\infty}\left(a|\nabla_v\psi(v)|^2 - \Delta_v\psi(v)\right) > 0,$$

see [230, 22]. The inequality (4.8) is always true when $\mathcal{V} = \mathbb{T}^d$ and ψ is smooth.

As in [P1], following up on ideas used in [161, 5] we introduce, for any $h \in \mathrm{L}^2([0,\infty);\mathrm{L}^2(\Theta))$, the following functional:

$$\mathscr{H}_\tau(t) = \int_t^{t+\tau} \|h(s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\,ds. \tag{4.7}$$

Then, our main result goes as follows.

Theorem 53. *Under Assumptions 2-3-4-5-6, there exists a constant $\lambda > 0$ such that all functions h solving (VOU) satisfy*

$$\forall\, t \ge 0, \qquad \mathscr{H}_\tau(t) \le \mathrm{e}^{-2\lambda t}\mathscr{H}_\tau(0).$$

Moreover,

$$\lambda \ge \bar\lambda\left(\frac{1}{\xi c_\psi} + \xi\right)^{-1}, \qquad \bar\lambda = \left(1 + C_\tau^{\mathrm{Lions}} K_{\mathrm{avg}}\right)^{-1}.$$

Theorem 53 implies a standard hypocoercivity estimate (see [256]) for solutions to (VOU). Note that there is no prefactor C in the exponential decay estimate of Theorem 53. Pointwise in time decay estimates can be deduced from Theorem 53 upon adding such a prefactor.

Corollary 54. *Under the same assumptions as Theorem 53, all solutions to (VOU) satisfy*

$$\forall t \ge 0, \qquad \|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \le C\,\mathrm{e}^{-\lambda t}\,\|h_0\|^2_{\mathrm{L}^2(\Theta)},$$

with $C = \mathrm{e}^{\lambda\tau}$.

4.2.4 Algebraic convergence rates

Condition (4.6) is equivalent to the fact that the symmetric part $\mathcal{S}$ is coercive, in the following sense:

$$\mathcal{S} \geq c_\psi \mathrm{Id} \quad \text{on} \quad \mathrm{L}^2_0(\gamma). \tag{4.8}$$

Typical cases where the last condition (i.e. Assumption 6) does not hold correspond to a kinetic energy ψ growing less than linearly at infinity, such as

$$\psi = (1+|v|^2)^{\frac{\alpha}{2}}, \quad \alpha \in (0,1), \tag{4.9}$$

or

$$\psi(v) = \frac{\beta}{2}\log\left(1+|v|^2\right), \quad \beta > d+2, \tag{4.10}$$

so that γ is heavy tailed at infinity. However, these classes of examples satisfy the following weighted Poincaré inequality (see [54, 66]) that we state in a general form in Assumption 7 below.

Assumption 7. *There exist a constant $P_\psi \in \mathbb{R}_+$, and a function $\mathscr{G} \in \mathrm{L}^1(\gamma)$ such that*

$$\forall g \in \mathrm{H}^1(\gamma) \cap \mathrm{L}^2_0(\gamma), \qquad \int_{\mathbb{R}^d} \mathscr{G}^{-1}(v)\, g^2(v)\, \gamma(dv) \leq P_\psi \|\nabla_v g\|^2_{\mathrm{L}^2(\gamma)}. \tag{4.11}$$

For the classes of examples (4.9)-(4.10), we can respectively consider $\mathscr{G}(v) = (1+|v|^2)^{1-\alpha}$ and $\mathscr{G}(v) = 1+|v|^2$, see [54, 66].

In the cases covered by Assumption 7, we recover constructive algebraic decay rates for $\mathscr{H}_\tau$ as follows.

Theorem 55. *Suppose that Assumptions 2-3-4-5-7 hold, and consider $\sigma > 0$ such that $\mathscr{G}^\sigma \in \mathrm{L}^1(\gamma)$. If all solutions h to (VOU) satisfy*

$$\int_{\mathscr{X}\times\mathscr{V}} \mathscr{G}(v)^\sigma\, h^2(t,x,v)\, \Theta(dx\,dv) \leq A\,\mathscr{U}(h_0), \qquad \forall t \geq 0, \tag{4.12}$$

for a functional $\mathscr{U} : \mathrm{L}^2(\Theta) \to [0,\infty]$ and a constant $A(\psi,\sigma,\phi) > 0$, then there exist two explicit constants $c_1(h_0), c_2(h_0) \in \mathbb{R}_+$ depending only on $\|h_0\|^2_{\mathrm{L}^2(\Theta)}, \mathscr{U}(h_0), \phi, \psi, \sigma$, and d such that

$$\forall \xi > 0, \quad \forall t \geq 0, \qquad \mathscr{H}_\tau(t) \leq \frac{\mathscr{H}_\tau(0)}{\left(1+\left(\xi^{-\frac{\sigma}{\sigma+1}}c_1+\xi c_2\right)^{-\frac{\sigma+1}{\sigma}} t\right)^{\sigma}},$$

We have that (4.12) holds true if $h \in \mathrm{L}^\infty(\Theta)$, with $\mathscr{U}(h_0) = \|h_0\|^2_{\mathrm{L}^\infty(\Theta)}$.

Explicit expressions for c_1, c_2 can be obtained, see Section 4.4.2. As in Section 4.2.4, a decay estimate for $\mathscr{H}$ implies a point-wise decay estimate in time. We state a result analogous to the one of Corollary 54 in the present situation.

Corollary 56. *Under the same assumptions as Theorem 55, all solutions to (VOU) satisfy*

$$\forall t \geq 0, \qquad \|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \leq \frac{\|h_0\|^2_{\mathrm{L}^2(\Theta)}}{\left(1+\left(\xi^{-\frac{\sigma}{\sigma+1}}c_1+\xi c_2\right)^{-\frac{\sigma+1}{\sigma}}(t+\tau)\right)^{\sigma}}.$$

Remark 57. *Note that Theorems 53 and 55 are qualitatively consistent. Indeed, the first one can be recovered from the second one in the regime corresponding to a sequence of sub-exponential local equilibria γ becoming* *exponential in the limit. In this case,* (4.11) *holds with* $\mathscr{G} \to 1$*, so that the choice* $\sigma \to \infty$ *is allowed, yielding a decay which is faster than any inverse power of time, hence no longer algebraic.*

Remark 58. *An alternative approach to obtaining a convergence result similar to Theorem 55 is that of weak Poincaré inequalities, see [158], when the initial data of* (VOU) *are* L^∞ *functions. This works thanks to the maximum principle for hypoelliptic operators, see [247].*

Theorem 55 crucially relies on the weighted L^2 bound (4.12), which we now motivate and discuss. When the local equilibrium γ decays slower than exponentially at infinity, the phenomenon of *loss of velocity moments* occurs. This means that the L^2–norm of the solution cannot be controlled without assuming the control of extra velocity moments for the initial data. In addition, we have to ensure that these moments are propagated by (VOU). Such an issue is present even in the spatially-homogeneous case. This has already been studied in [66, 85] and [P1]. In order to follow the strategy of proof of [P1], the term we need to bound along the flow of (VOU) reads

$$\int_{\mathscr{X}\times\mathscr{V}} \mathscr{G}(v)^\sigma\, h^2(t,x,v)\,\Theta(dx\,dv). \tag{4.13}$$

However, obtaining a direct uniform control is tricky in practice. Then, our choice is taking $h_0 \in \mathrm{L}^\infty(\Theta)$ and $\mathscr{U}(h_0) = \|h_0\|^2_{\mathrm{L}^\infty(\Theta)}$.

Another strategy is using an argument based on Lyapunov functionals, which would allow more general initial data. This is possible, after [69].

4.2.5 Weighted Lions' inequalities

Inequality (4.2) is a generalised version of Lions' inequality [11]. Such an inequality with a weighted measure has been first studied in [84]. We give here a simple sufficient condition on the potential ϕ for (4.2) to hold. We also obtain an explicit estimate of C_τ^{Lions}, under some growth conditions on the spatial potential ϕ.

Theorem 59. *Assume that* ϕ *is a smooth potential such that* $\int_{\mathscr{X}} \mathrm{e}^{-\phi}\,dx = 1$*, that its associated probability measure* μ *satisfies the Poincaré inequality:*

$$\|\nabla_x z\|_{\mathrm{L}^2(\mu)} \geq c_\phi^{-\frac{1}{2}} \|z\|_{\mathrm{L}^2(\mu)}, \qquad \forall z \in \mathrm{H}^1(\mu)\cap \mathrm{L}^2_0(\mu), \tag{4.14}$$

for some $c_\phi \in \mathbb{R}_+$*, and that*

$$|\nabla^2\phi|^2 \leq (c'_\phi)^2\left(d + |\nabla_x\phi|^2\right), \qquad \Delta\phi \leq c''_\phi\left(d + |\nabla_x\phi|^2\right). \tag{4.15}$$

for some constants $c'_\phi, c''_\phi \in \mathbb{R}_+$*. Then, Assumption 2 is satisfied, with*

$$C_\tau^{\mathrm{Lions}} = C_\phi\left(1+\sqrt{d}\right)\left(\tau^2+\tau^{-2}\right)^2,$$

for a constructive constant C_ϕ*, depending only on* $c_\phi, c'_\phi, c''_\phi$*, but not directly on* d*.*

The explicit dependence on d in the estimate of C_τ^{Lions}, which is exactly of order $\sqrt{d}$ if the Poincaré constant c_ϕ does not depend on d, is important for the case of a Boltzmann-Gibbs measure obtained as the tensorisation of several 1-dimensional ones (as in the context of sampling methods, see [243, 200]). Dimension independent Poincaré constants can also be obtained for systems with *finite number of interactions*, see [44]. The scaling in the second inequality of (4.15) is natural, as it is the correct one for separable potentials $\phi = \sum_{i=1}^d \bar\phi(x_i)$, with $\bar\phi : \mathbb{R} \to \mathbb{R}$ such that $\bar\phi'' \leq c''_\phi\left(1 + |\bar\phi'|^2\right)$. Sufficient conditions for (4.14) to hold true are given in [230, 22]. A noteworthy case in the theorem above is when the injection $\mathrm{H}^1(\mu) \hookrightarrow \mathrm{L}^2(\mu)$ is compact. In this situation, the estimates can be inferred from [79, Lemma 2.6]; in particular C_τ^{Lions} is bounded by the constant in [79, Equation (27)].

4.3 Notation and preliminary technical results

We denote by $A^\star$ for the adjoint of a closed operator A on the Hilbert space $\mathrm{L}^2(\Theta)$, which, for any smooth functions h_1, h_2 with compact supports, satisfies

$$\int_{\mathcal{X}\times\mathcal{V}} (Ah_1) h_2 \, d\Theta = \int_{\mathcal{X}\times\mathcal{V}} h_1 \left(A^\star h_2\right) d\Theta.$$

The action of $A^\star$ is obtained by integration by parts for differential operators. In particular, for $1 \le i \le d$,

$$\partial_{x_i}^\star = -\partial_{x_i} + \partial_{x_i}\phi, \qquad \partial_{v_i}^\star = -\partial_{v_i} + \partial_{v_i}\psi. \tag{4.16}$$

This allows to rewrite the operators appearing in (VOU) as the sum of a skew-symmetric operator

$$T =: \nabla_v \psi \cdot \nabla_x - \nabla_x \phi \cdot \nabla_v = \sum_{i=1}^d \partial_{v_i}^\star \partial_{x_i} - \partial_{x_i}^\star \partial_{v_i},$$

and a symmetric one

$$\mathcal{S} = -\nabla_v^\star \nabla_v = -\sum_{i=1}^d \partial_{v_i}^\star \partial_{v_i}, \tag{4.17}$$

so that (VOU) rewrites

$$\partial_t h + Th = \xi \mathcal{S} h.$$

The right hand side of (4.20) is indeed always finite. Solutions to (VOU) live in the Bochner space

$$\mathrm{L}^2(0,\infty; \mathrm{L}^2(\mu, \mathrm{H}^1(\gamma))),$$

as we show thanks to the following inequality for solutions to (VOU):

$$\frac{d}{dt}\|h(t)\|_{\mathrm{L}^2(\Theta)}^2 = 2\left\langle h(t), Th(t)\right\rangle_{\mathrm{L}^2(\Theta)} + 2\xi\left\langle h(t), \mathcal{S}h(t)\right\rangle_{\mathrm{L}^2(\Theta)} = -2\xi\|\nabla_v h\|_{\mathrm{L}^2(\Theta)}^2 \le 0, \tag{4.18}$$

in view of the skew-symmetry of T and (4.17).

We also denote by $\mathrm{H}^{-1}(\mathrm{U}_\tau \otimes \mu)$ the dual of the space

$$\mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau \otimes \mu) = \left\{ h \in \mathrm{H}^1(\mathrm{U}_\tau \otimes \mu) \,\middle|\, h(0,\cdot) = h(\tau,\cdot) = 0 \right\}, \tag{4.19}$$

where the subscript DC stands for *Dirichlet conditions.* Note that the boundary conditions in time are well defined because the trace of functions $h \in \mathrm{H}^1(\mathrm{U}_\tau \otimes \mu)$ makes sense on the boundary $\{0,\tau\} \times \mathcal{E}$ of the domain. Under Assumption 4, the operator $-T + \xi\mathcal{S}$ is hypoelliptic (see [169, 256]), hence the solutions to (VOU) are smooth.

The first result provides estimates on the skew-symmetric part of the space-time Fokker–Planck operator $\partial_t + T$ in terms of its symmetric part $\mathcal{S}$, for solutions of (VOU).

Lemma 60. *Consider a solution of* (VOU)*. Then,*

$$\|(\partial_t + T)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu;\mathrm{H}^{-1}(\gamma))} \leq \xi \|\nabla_v h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}. \tag{4.20}$$

In particular, solutions to (VOU) *belong to* $\mathrm{H}^1_{\mathrm{kin}}$.

Proof. Fix a test function $u \in \mathrm{L}^2(\mathrm{U}_\tau\otimes\mu;\mathrm{H}^1(\gamma))$ such that $\|u\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu;\mathrm{H}^1(\gamma))} \leq 1$. Then, denoting by $\langle\cdot,\cdot\rangle_{\mathrm{L}^2(\mathrm{H}^{-1}),\mathrm{L}^2(\mathrm{H}^1)}$ the duality bracket between $\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu;\mathrm{H}^{-1}(\gamma))$ and $\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu;\mathrm{H}^1(\gamma))$,

$$\langle(\partial_t + T)h, u\rangle_{\mathrm{L}^2(\mathrm{H}^{-1}),\mathrm{L}^2(\mathrm{H}^1)} = \xi\langle \mathcal{S}h, u\rangle_{\mathrm{L}^2(\mathrm{H}^{-1}),\mathrm{L}^2(\mathrm{H}^1)} = -\int_0^\tau \int_{\mathcal{X}\times\mathcal{V}} \nabla_v h\cdot\nabla_v u \, d\mathrm{U}_\tau \, d\Theta.$$

The desired estimate follows from first bounding the right hand side via the Cauchy-Schwarz inequality, and then taking the supremum over u. □

The next lemma, whose proof is omitted, provides estimates similar to [79, Lemma 2.3], obtained with an adaptation of the approach of [44, Lemma 3.6 and 3.7] to spacetime operators. It provides a fully explicit dependence of the estimates on the dimension d.

Lemma 61. *Assume that ϕ is a smooth potential such that $\int_{\mathcal{X}} \mathrm{e}^{-\phi}\,dx = 1$, and* (4.14)*-*(4.15) *hold true. Then, Assumption 5 holds true with*

$$L_\phi^2 = \max\left(16, 4\,d\,c''_\phi\right).$$

Moreover, for any $z \in \mathrm{H}^2(\mathrm{U}_\tau\otimes\mu)$ such that $\nabla_{t,x} z \in \mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu;\mathbb{R}^d)$,

$$\|\nabla^2_{t,x} z\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \leq 2\|\nabla^\star_{t,x}\nabla_x z\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} + 2c'_\phi\left(\sqrt{d} + 2\max\left(8c'_\phi, \sqrt{c''_\phi\, d}\right)\right)\|\nabla_{t,x} z\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}. \tag{4.21}$$

It would be possible to have better constants in (4.21) under additional assumptions, for example if ϕ is strongly convex, see [44, 79].

4.4 Modified Poincaré inequalities and exponential decay rates

We start by proving Lemma 52 in Section 4.4.1, using a Poincaré-type inequality. We turn to the proof of Theorem 53 in Section 4.4.2. We finally prove Corollary 54 in Section 4.4.3.

4.4.1 Proof of the averaging lemma and of a Poincaré-like inequality

The important Poincaré-type inequality which is used to prove Theorem 53 is the following. Its proof crucially relies on Lemma 52.

Proposition 62. *Under Assumptions 2-3-5, it holds*

$$\forall h \in \mathrm{H}^1_{\mathrm{kin}} \cap \mathrm{L}^2_0(\mathrm{U}_\tau\otimes\Theta), \qquad \bar\lambda\|h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} \leq \|(\partial_t + T)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu;\mathrm{H}^{-1}(\gamma))} + \|(\mathrm{Id}-\Pi)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}, \tag{4.22}$$

with

$$\bar\lambda = \left(1 + C_\tau^{\mathrm{Lions}} K_{\mathrm{avg}}\right)^{-1},$$

where K_{avg} is the constant appearing in Lemma 52.

Proof of Proposition 62. Since Π is an orthogonal projector,

$$\|h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} = \|(\mathrm{Id}-\Pi)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} + \|\Pi h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}. \tag{4.23}$$

We apply (4.2) to the second term on the right hand side in (4.23), and then use Lemma 52 to write

$$\begin{aligned}\|\Pi h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} &\leq C_\tau^{\mathrm{Lions}} \|\nabla_{t,x}\Pi h\|^2_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}\\ &\leq C_\tau^{\mathrm{Lions}}\,\mathrm{K}_{\mathrm{avg}}\left(\|(\mathrm{Id}-\Pi)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} + \|(\partial_t+T)h\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^{-1}(\gamma))}\right).\end{aligned}$$

Plugging the last estimate into (4.23) provides the claimed result. □

We conclude this section with the proof of Lemma 52.

Proof of Lemma 52. Some integrals in the calculation below are formal and represent duality products. In order to estimate the H^{-1} norm of $\nabla_{t,x}\Pi h$, we first compute the H^{-1} norm of $\partial_t\Pi h$, and then turn to $\nabla_x\Pi h$. To bound $\|\partial_t\Pi h\|_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}$, we fix a test function $z\in\mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu)$ with $\|z\|^2_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}\leq 1$; while the estimation of $\|\nabla_x\Pi h\|_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}$ can be performed by considering a test function $Z=(Z_1,\dots,Z_d)\in\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)^d$ with $\|Z\|^2_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}\leq 1$.

For $\partial_t\Pi h$, we start by noticing that $\nabla_v\Pi h = 0$ and

$$\int_{\mathcal{V}}\nabla\psi(v)\,\gamma(dv) = 0,$$

so that, using the skew-symmetry of $\partial_t + T$ and the Dirichlet boundary conditions for the second term,

$$\begin{aligned}\int_0^\tau\int_{\mathcal{X}}(\partial_t\Pi h)\,z\,d\mathrm{U}_\tau\,d\mu &= \int_0^\tau\int_{\mathcal{X}}\int_{\mathcal{V}}((\partial_t+T)\Pi h)(t,x,v)\,z(t,x)\,\mathrm{U}_\tau(dt)\,\mu(dx)\,\gamma(dv)\\ &= \int_0^\tau\int_{\mathcal{X}}\int_{\mathcal{V}}[(\partial_t+T)h]\,z\,d\mathrm{U}_\tau\,d\Theta + \int_0^\tau\int_{\mathcal{X}}\int_{\mathcal{V}}[(\mathrm{Id}-\Pi)h]\,(\partial_t+T)z\,d\mathrm{U}_\tau\,d\Theta\\ &\leq \|(\partial_t+T)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^{-1}(\gamma))}\|z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^1(\gamma))} + \|(\mathrm{Id}-\Pi)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}\|(\partial_t+T)z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}.\end{aligned}$$

The last factor in the last inequality can be bounded as

$$\|(\partial_t+T)z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} \leq \|\partial_t z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} + \|Tz\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} \leq 1 + \|\nabla_v\psi\|_{\mathrm{L}^2(\gamma)},$$

where we used $\|z\|_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}\leq 1$ and a Cauchy–Schwarz inequality to write

$$\|Tz\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} = \|\nabla\psi\cdot\nabla_x z\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)} \leq \|\nabla_x z\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}\int_0^\tau\int_{\mathcal{X}}\int_{\mathcal{V}}|\nabla\psi(v)|^2\,\mathrm{U}_\tau(dt)\,\mu(dx)\,\gamma(dv).$$

By taking the supremum over functions $z\in\mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu)$ with $\|z\|^2_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}\leq 1$, we finally obtain

$$\|\partial_t\Pi h\|_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)} \leq \|(\partial_t+T)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^{-1}(\gamma))} + \left(1+\|\nabla_v\psi\|_{\mathrm{L}^2(\gamma)}\right)\|(\mathrm{Id}-\Pi)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}.$$

We next turn to the estimation of $\|\nabla_x\Pi h\|_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}$. We introduce

$$G = \frac{\nabla_v\psi}{\|\nabla_v\psi\|_{\mathrm{L}^2(d\gamma)}}.$$

In view of the definition (4.4), it holds by construction that

$$\int_{\mathcal{V}} G(v) \otimes MG(v)\gamma(dv) = \mathrm{Id}_{d\times d}, \qquad M = \|\nabla_v\psi\|^2_{\mathrm{L}^2(d\gamma)}\mathscr{M}^{-1}.$$

The quantity to estimate is

$$\begin{aligned}
\int_0^\tau \int_{\mathcal{X}} \nabla_x \Pi h\cdot Z\, d\mathrm{U}_\tau\, d\mu &= \int_0^\tau \int_{\mathcal{X}}\int_{\mathcal{V}} [G(v)\cdot\nabla_x\Pi h(t,x)][MG(v)\cdot Z(t,x)]\mathrm{U}_\tau(dt)\Theta(dx\,dv)\\
&= \|\nabla_v\psi\|^{-1}_{\mathrm{L}^2(\gamma)} \int_0^\tau \int_{\mathcal{X}}\int_{\mathcal{V}} [(\partial_t+T)\Pi h](t,x,v)[MG(v)\cdot Z(t,x)]\mathrm{U}_\tau(dt)\Theta(dx\,dv)\\
&= \|\nabla_v\psi\|^{-1}_{\mathrm{L}^2(\gamma)} \int_0^\tau \int_{\mathcal{X}}\int_{\mathcal{V}} [(\partial_t+T)h](t,x,v)[MG(v)\cdot Z(t,x)]\mathrm{U}_\tau(dt)\Theta(dx\,dv)\\
&\quad - \|\nabla_v\psi\|^{-1}_{\mathrm{L}^2(\gamma)} \int_0^\tau \int_{\mathcal{X}}\int_{\mathcal{V}} (T(1-\Pi)h)(t,x,v)[MG(v)\cdot Z(t,x)]\mathrm{U}_\tau(dt)\Theta(dx\,dv)\\
&\quad - \|\nabla_v\psi\|^{-1}_{\mathrm{L}^2(\gamma)} \int_0^\tau \int_{\mathcal{X}}\int_{\mathcal{V}} (\partial_t(1-\Pi)h)(t,x,v)[MG(v)\cdot Z(t,x)]\mathrm{U}_\tau(dt)\Theta(dx\,dv),
\end{aligned} \tag{4.24}$$

where the second equality follows from the fact that $\nabla_v\Pi h = 0$ and $\int_{\mathcal{V}} G\, d\gamma = 0$. Since

$$\|MZ\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le \|M\|_{\mathcal{B}(\ell^2)}\|Z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le \|M\|_{\mathcal{B}(\ell^2)}\|Z\|_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)} \le \|M\|_{\mathcal{B}(\ell^2)},$$

where $\|M\|_{\mathcal{B}(\ell^2)}$ denotes the matrix norm induced by the Euclidean norm, and $\|G\|_{\mathrm{H}^1(\gamma)} < +\infty$ (in view of the second condition in (4.3)), the first integral in the last equality can be bounded by $\|G\cdot MZ\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^1(\gamma))}\|(\partial_t + T)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^{-1}(\gamma))} \le \|M\|_{\mathcal{B}(\ell^2)}\|G\|_{\mathrm{H}^1(\gamma)}\|(\partial_t+T)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu,\mathrm{H}^{-1}(\gamma))}$. Since T is antisymmetric, the second integral in the last equality can be controlled as

$$\left|\int_0^\tau \int_{\mathcal{X}}\int_{\mathcal{V}} [(1-\Pi)h][T(MG\cdot Z)]\,d\mathrm{U}_\tau\, d\Theta\right| \le \|(1-\Pi)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}\|T(MG\cdot Z)\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}.$$

Introducing $G^M = (G_1^M,\dots,G_d^M) = MG$ to simplify the notation,

$$\begin{aligned}
[T(MG\cdot Z)](t,x,v) &= T\left[\sum_{i=1}^d G_i^M Z_i\right](t,x,v)\\
&= \sum_{i=1}^d G_i^M(v)\nabla\psi(v)\nabla_x Z_i(t,x) - \nabla G_i^M(v)\cdot\nabla\phi(x)Z_i(t,x),
\end{aligned}$$

where $G_i^M\nabla\psi \in \mathrm{L}^2(\gamma)$ and $\nabla G_i^M \in \mathrm{L}^2(\gamma)$ in view respectively of the first and second condition in (4.3); while $Z_i\nabla\phi \in \mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)$ with $\|Z_i\nabla\phi\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le L_\phi\|Z_i\|_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}$ by Assumption 5. The last integral in (4.24) can be bounded, after transferring the partial derivative in time to the test function, by $\|(1-\Pi)h\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}\|MG\cdot\partial_t Z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}$. The second factor tensorizes as $\|MG\|_{\mathrm{L}^2(\gamma)}\|\partial_t Z\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le \|M\|_{\mathcal{B}(\ell^2)}\|G\|_{\mathrm{L}^2(\gamma)} = \|M\|_{\mathcal{B}(\ell^2)}$.

By gathering the above estimates, the desired result follows with the constant

$$K_{\mathrm{avg}}^{\frac{1}{2}} = \max\left(1 + \|\nabla_v\psi\|_{\mathrm{L}^2(\gamma)}, \frac{\|M\|_{\mathcal{B}(\ell^2)}\left(1+\|G\|_{\mathrm{H}^1(\gamma)}\right) + \sqrt{\sum_{i=1}^d \|G_i^M\nabla_v\psi\|^2_{\mathrm{L}^2(\gamma)}} + L_\phi\sqrt{\sum_{i=1}^d\|\nabla G_i^M\|^2_{\mathrm{L}^2(\gamma)}}}{\|\nabla_v\psi\|_{\mathrm{L}^2(\gamma)}}\right), \tag{4.25}$$

which concludes the proof. □

Remark 63. *The choice of G in the previous calculation is natural. Indeed, it yields an orthogonal decomposi-*

tion in $\mathrm{L}^2(\mathrm{U}_\tau \otimes \Theta)$ *of the term*

$$[(\partial_t + T)\Pi h](t,x,v) = (\partial_t \Pi h)(t,x) + G(v)\cdot \int_{\mathcal{V}} G(v)\,(T\Pi h)(t,x,v')\gamma(dv').$$

One component lies in the image of Π*, while the second one is spanned by* G *itself. Our result compares with [5, Lemma 3.1], which is not quantitative at all, since the authors do not provide an explicit expression to terms like* M *and* G *in our proof. We also compare with the estimates at the beginning of [79, Theorem 2]. In spite of the similarity in the strategy, we make the algebra behind the proof neat and usable also for general kinetic energies* ψ*.*

4.4.2 Exponential decay rates: proof

The exponential convergence is a direct consequence of (4.22). Indeed, note first that (4.18) implies

$$\begin{aligned}\frac{d\mathscr{H}_\tau(t)}{dt} &= \frac{d}{dt}\left(\int_0^\tau \|h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\,ds\right) = \int_0^\tau \frac{d}{dt}\|h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\,ds \\ &= -2\xi \int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\,ds = -2\xi\tau \int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\,\mathrm{U}_\tau(ds).\end{aligned} \tag{4.26}$$

On the other hand, the space-time Poincaré inequality (4.22) implies, together with (4.20) and (4.6), that

$$\bar{\lambda}\int_0^\tau \|h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\mathrm{U}_\tau(ds) \le \left(\frac{1}{c_\psi} + \xi^2\right)\int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)}\mathrm{U}_\tau(ds).$$

The combination of the previous inequality and (4.26) leads to

$$\frac{d\mathscr{H}_\tau(t)}{dt} \le -2\bar{\lambda}\left(\frac{1}{\xi c_\psi} + \xi\right)^{-1}\mathscr{H}_\tau(t).$$

from which Theorem 53 follows by a Gronwall inequality.

4.4.3 Hypocoercivity estimates

Decay estimates on $\mathscr{H}_\tau(t)$ lead to pointwise in time decay estimates on $h(t)$ since $t \mapsto \|h(t)\|^2_{\mathrm{L}^2(\Theta)}$ is nonincreasing, which is a direct consequence of (4.18). Indeed, for any $t \ge 0$,

$$\|h(t+\tau)\|^2_{\mathrm{L}^2(\Theta)} \le \frac{1}{\tau}\mathscr{H}_\tau(t) \le \frac{\mathscr{H}_\tau(0)}{\tau}\mathrm{e}^{-2\lambda t} \le \|h(0)\|^2_{\mathrm{L}^2(\Theta)}\mathrm{e}^{-2\lambda t},$$

so that

$$\forall t \ge 0, \qquad \|h(t)\|_{\mathrm{L}^2(\Theta)} \le \|h(0)\|_{\mathrm{L}^2(\Theta)}\mathrm{e}^{-\lambda(t-\tau)},$$

which gives the claimed result.

4.5 Algebraic decay rates

In this section we prove Theorem 55, see Section 4.5.1. Then, we discuss (4.12) and related conditions.

4.5.1 Algebraic decay rates: proof

The overall strategy of the proof is similar to the one presented in [P1], so that we only sketch the main steps of the argument. Note however that, compared to [P1] (and various other works, for instance [66]), the transport part of the evolution now includes an additional term $\nabla_x\phi\cdot\nabla_v$ and $\nabla_v\psi\cdot\nabla_x$ replaces $v\cdot\nabla_x$.

The overall idea for the proof is to rely on the Poincaré-type estimate (4.22), where the second term is controlled by a combination of the weighted Poincaré inequality (4.11) and a moment bound. More precisely, consider Hölder conjugate exponents, $p,q\geq 1$, with $p=\sigma^{-1}(1+\sigma)>1$. Note that $\frac{q}{p}=\sigma$. Fix $t>0$, and define, for any $s\in[0,\tau]$, the following moment based on $\mathscr{G}$:

$$Z_t(s):=\int_{\mathcal{X}\times\mathcal{V}}\mathscr{G}(v)^\sigma\,|h(t+s,x,v)-\Pi h(t+s,x)|^2\,\mu(dx)\,\gamma(dv). \tag{4.27}$$

The first step of the proof is to bound $\|(\mathrm{Id}-\Pi)h(t+\cdot,\cdot,\cdot)\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\Theta)}$. We start from the following pointwise bound for a given $x\in\mathcal{X}$, obtained with a Hölder inequality:

$$\begin{aligned}&\|h(t+s,x,\cdot)-\Pi h(t+s,x)\|^2_{L^2(\gamma)}\\&\leq\left(\int_{\mathcal{V}}\mathscr{G}(v)^{-1}|h(t+s,x,v)-\Pi h(t+s,x)|^2\gamma(dv)\right)^{\frac{1}{p}}\\&\quad\times\left(\int_{\mathcal{V}}\mathscr{G}(v)^{\frac{q}{p}}|h(t+s,x,v)-\Pi h(t+s,x)|^2\gamma(dv)\right)^{\frac{1}{q}}.\end{aligned}$$

Via (4.11), the last right-hand side of the above inequality can be bounded as

$$P_\Psi^{\frac{1}{p}}\left(\int_{\mathcal{V}}|\nabla_v h(t+s,x,v)|^2\gamma(dv)\right)^{\frac{1}{p}}\left(\int_{\mathcal{V}}\mathscr{G}(v)^{\frac{q}{p}}|h(t+s,x,v)-\Pi h(t+s,x)|^2\gamma(dv)\right)^{\frac{1}{q}}.$$

Integrating with respect to the measure $\mathrm{U}_\tau\otimes\mu$, and applying Hölder's inequality in the variables s,x gives

$$\begin{aligned}&\int_0^\tau\|h(t+s,\cdot,\cdot)-\Pi h(t+s,\cdot)\|^2_{L^2(\Theta)}\,\mathrm{U}_\tau(ds)\\&\leq P_\Psi^{\frac{1}{p}}\left(\int_0^\tau\int_{\mathscr{X}}\int_{\mathscr{V}}|\nabla_v h(t+s,x,v)|^2\,\mathrm{U}_\tau(ds)\,\Theta(dx\,dv)\right)^{\frac{1}{p}}\\&\quad\times\left(\int_0^\tau\int_{\mathscr{X}}\int_{\mathscr{V}}\mathscr{G}(v)^{\frac{q}{p}}|h(t+s,x,v)-\Pi h(t+s,x)|^2\,\mathrm{U}_\tau(ds)\,\Theta(dx\,dv)\right)^{\frac{1}{q}}\\&=P_\Psi^{\frac{1}{p}}\left(\int_0^\tau\|\nabla_v h(t+s,\cdot,\cdot)\|^2_{L^2(\Theta)}\,\mathrm{U}_\tau(ds)\right)^{\frac{1}{p}}\left(\int_0^\tau Z_t(s)\,\mathrm{U}_\tau(ds)\right)^{\frac{1}{q}}\\&\leq P_\Psi^{\frac{1}{p}}\left(\int_0^\tau\|\nabla_v h(t+s,\cdot,\cdot)\|^2_{L^2(\Theta)}\,\mathrm{U}_\tau(ds)\right)^{\frac{1}{p}}W^{\frac{1}{q}}\left(\|h_0\|^2_{\mathrm{L}^2(\Theta)}+\mathscr{U}(h_0)\right)^{\frac{1}{q}},\end{aligned}$$

where the last inequality follows from Lemma 65 below. Combining the last bound with (4.22), and then using (4.20) gives

$$\begin{aligned}
\bar{\lambda} \int_0^\tau & \|h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \mathrm{U}_\tau(ds) \\
&\leq \int_0^\tau \|(\partial_t + T)\, h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(\mu;\mathrm{H}^{-1}(\gamma))} \\
&\quad + W^{\frac{1}{q}} \left(\|h_0\|^2_{\mathrm{L}^2(\Theta)} + \mathscr{U}(h_0)\right)^{\frac{1}{q}} P_\Psi^{\frac{1}{p}} \left(\int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \mathrm{U}_\tau(ds)\right)^{\frac{1}{p}} \\
&\leq \xi^2 \int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \mathrm{U}_\tau(ds) \\
&\quad + W^{\frac{1}{q}} \left(\|h_0\|^2_{\mathrm{L}^2(\Theta)} + \mathscr{U}(h_0)\right)^{\frac{1}{q}} P_\Psi^{\frac{1}{p}} \left(\int_0^\tau \|\nabla_v h(t+s,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \mathrm{U}_\tau(ds)\right)^{\frac{1}{p}}.
\end{aligned}$$

In view of the definitions (4.7) and (4.26), the latter inequality can be rewritten as

$$\mathscr{H}_\tau(t) \leq \frac{\xi}{2\bar{\lambda}} \left(-\frac{d}{dt}\mathscr{H}_\tau(t)\right) + \xi^{-\frac{1}{p}} M_0 \left(-\frac{d}{dt}\mathscr{H}_\tau(t)\right)^{\frac{1}{p}}, \tag{4.28}$$

where

$$M_0 := \tau^{\frac{1}{\sigma+1}} \frac{1}{\bar{\lambda}} \left(\frac{P_\Psi}{2}\right)^{\frac{1}{p}} W^{\frac{1}{q}} \left(\|h_0\|^2_{\mathrm{L}^2(\Theta)} + \mathscr{U}(h_0)\right)^{\frac{1}{q}}.$$

The proof can now be concluded using classical arguments, relying on a generalization of the Gronwall inequality leading to algebraic decay rates. We however carefully keep track of the parameters influencing the convergence rate, by retaining the largest term in (4.28). In order to make this precise, consider

$$y \mapsto \vartheta(y) = \frac{\xi}{2\bar{\lambda}}\, y + M_0 \left(\frac{y}{\xi}\right)^{\frac{1}{p}},$$

which is strictly increasing from $[0,\infty)$ to $[0,\infty)$. Set $y_0 = \vartheta^{-1}(\mathscr{H}_\tau(0))$. Note that

$$\forall y \leq y_0, \quad \vartheta(y) \leq \left(\xi^{-\frac{1}{p}} M_0 + \frac{\xi}{2\bar{\lambda}}\, y_0^{\frac{p-1}{p}}\right) y^{\frac{1}{p}}.$$

Hence, as $t \mapsto \mathscr{H}_\tau(t)$ is nonincreasing, we have

$$\mathscr{H}_\tau(t) \leq \left(\xi^{-\frac{1}{p}} M_0 + \frac{\xi}{2\bar{\lambda}}\, y_0^{\frac{p-1}{p}}\right) \left(-\frac{d}{dt}\mathscr{H}_\tau(t)\right)^{\frac{1}{p}}.$$

Applying the standard Bihari–LaSalle [50, 186] argument, one proves, as in [P1],

$$\mathscr{H}_\tau(t) \leq \left(\mathscr{H}_\tau(0)^{1-p} + (p-1)\left(\xi^{-\frac{1}{p}} M_0 + \frac{\xi}{2\bar{\lambda}}\, y_0^{\frac{p-1}{p}}\right)^{-p} t\right)^{-\frac{1}{p-1}}. \tag{4.29}$$

Equation (4.29) can be rewritten as

$$\mathscr{H}_\tau(t) \leq \frac{\mathscr{H}_\tau(0)}{\left(1 + \left(\xi^{-\frac{1}{p}} c_1 + \xi c_2\right)^{-p} t\right)^{\frac{1}{p-1}}},$$

with $c_1 = (p-1)^{-\frac{1}{p}} M_0 \mathscr{H}_\tau(0)^{p(1-p)}$ and $c_2 = (p-1)^{-\frac{1}{p}} \frac{\xi}{2\bar{\lambda}}\, y_0^{\frac{p-1}{p}}\, \mathscr{H}_\tau(0)^{\frac{1-p}{p}}$. The proof of Theorem 55 is concluded.

Remark 64. *A more accurate estimate could be achieved by rewriting (4.28) as*

$$\frac{d}{dt}\mathscr{H}_\tau(t) \le -\zeta(\mathscr{H}_\tau(t)),$$

with $\zeta = \vartheta^{-1}$, which is indeed invertible, and then integrating the latter inequality.

The proof of Theorem 55 written in Section 4.5.1 relies on moment estimates encoded by $Z_t(s)$ introduced in (4.27) (see Lemma 65 below). When $h \in \mathrm{L}^\infty(\Theta)$, we take

$$\mathscr{U}(h_0) = \|h_0\|^2_{\mathrm{L}^\infty(\Theta)}, \qquad A = \|\mathscr{G}(v)^\sigma\|_{\mathrm{L}^1(d\gamma)}.$$

We next show how to obtain the control of the moment $Z(t)$ defined in (4.27), which appears in the proof of Theorem 55. This can be done in general if (4.12) holds.

Lemma 65. *There is an explicit constant $W > 0$ (not depending on d, ξ), such that, for any solution $h \in \mathrm{H}_{\mathrm{kin}}$ to (VOU) with initial datum h_0 with average zero with respect to Θ and such that (4.12) holds,*

$$\forall\, t \ge 0, \qquad Z(t) \le W\Big(\|h_0\|^2_{\mathrm{L}^2(\Theta)} + \mathscr{U}(h_0)\Big).$$

Proof. We bound $Z(t)$ as

$$Z(t) \le 2\int h(t,x,v)^2 \mathscr{G}(v)^\sigma\,\Theta(dx,dv) + 2\int (\Pi h)^2(t,x)\mathscr{G}(v)^\sigma\,\Theta(dx\,dv).$$

The first term in the right-hand side of the previous inequality is bounded with (4.12) as

$$\int h(t,x,v)^2 \mathscr{G}(v)^\sigma\,\Theta(dx\,dv) \le A\,\mathscr{U}(h_0).$$

On the other hand,

$$\begin{aligned}\int_{\mathscr{X}\times\mathscr{V}} (\Pi h)^2(t,x)\mathscr{G}(v)^\sigma\,\Theta(dx\,dv) &= \int_{\mathscr{X}} (\Pi h)^2(t,x)\mathscr{G}(v)^\sigma\,\mu(dx)\gamma(v)\\ &\le \|h_0\|^2_{\mathrm{L}^2(\Theta)}\,\|\mathscr{G}(v)^\sigma\|_{\mathrm{L}^1(d\gamma)},\end{aligned}$$

where we used a Cauchy–Schwarz inequality in the last step to write

$$\|\Pi h(t,\cdot)\|^2_{\mathrm{L}^2(\mu)} \le \|h(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(\Theta)} \le \|h_0\|^2_{\mathrm{L}^2(\Theta)}.$$

The result therefore follows with $W = 2\max(A, \|\mathscr{G}(v)^\sigma\|_{\mathrm{L}^1(d\gamma)})$. □

4.6 Weighted Poincaré-Lions inequalities

We present a constructive proof, based on a generalization of the manipulations performed in [79, Lemma 2.6]. We extend the approach to the situation when the operator $\nabla_x^*\nabla_x$ has a spectral gap but not necessarily a discrete spectrum. We also present the argument in a more abstract manner, in order to highlight the key steps in the construction.

4.6.1 Reduction to solving a divergence equation

Recall the space $\mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu)$, defined in (4.19). The Lions inequality (4.2) is implied by the following property (in fact, [11] shows the equivalence between the two statements): there exists a positive constant C'_τ such that, for any $f\in\mathrm{L}^2_0(\mathrm{U}_\tau\otimes\mu)$, there exists a solution $Z=(Z_0,Z_1,\dots,Z_d)\in\mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu)^{d+1}$ to

$$-\partial_t Z_0+\sum_{i=1}^d \partial_{x_i}^* Z_i=f, \tag{4.30}$$

which is such that

$$\|Z\|_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}=\sqrt{\sum_{i=0}^d\|Z_i\|^2_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}}\le C'_\tau\,\|f\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}. \tag{4.31}$$

Of course, there may be many solutions to the divergence-like equation (4.30). When such a solution exists,

$$\begin{aligned}\|f\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}&=\int_0^\tau\int_{\mathcal{X}} f(t,x)^2\,\mathrm{U}_\tau(dt)\,\mu(dx)=\int_0^\tau\int_{\mathcal{X}}\left(-\partial_t Z+\sum_{i=1}^d\partial_{x_i}^* Z\right)f\,d\mathrm{U}_\tau\,d\mu\\&=\langle\nabla_{t,x}f,Z\rangle_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu),H^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu)}\le\|Z\|_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}\|\nabla_{t,x}f\|_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}\\&\le C'_\tau\|f\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}\|\nabla_{t,x}f\|_{\mathrm{H}^{-1}(\mathrm{U}_\tau\otimes\mu)}.\end{aligned}$$

Hence, Lions' inequality (4.2) is recovered with $C_\tau^{\mathrm{Lions}}\le(C'_\tau)^2$. We therefore concentrate in this section on obtaining solutions to (4.30) which satisfy (4.31).

4.6.2 Solving the divergence equation

We show in this section how to solve (4.30) in order to have estimates such as (4.31). The precise result is the following.

Theorem 66. *Suppose that (4.14) and (4.15) hold. Then, there exists an explicit $C_\phi\in\mathbb{R}_+$, depending only on $c_\phi, c'_\phi, c''_\phi$ and not explicitly on d, such that, for any $\tau>0$ and $f\in\mathrm{L}^2_0(\mathrm{U}_\tau\otimes\mu)$, the equation posed on the space-time domain $[0,\tau]\times\mathcal{X}$*

$$-\partial_t Z_0+\sum_{i=1}^d \partial_{x_i}^* Z_i=f, \tag{4.32}$$

admits a solution $(Z_0,\dots,Z_d)\in\mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau\otimes\mu)^{d+1}$ which satisfies

$$\sum_{i=0}^d\|Z_i\|^2_{\mathrm{H}^1(\mathrm{U}_\tau\otimes\mu)}\le C_\phi\left(1+\sqrt{d}\right)\left(\frac{1}{\tau^2}+\tau^2\right)^2\|f\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}. \tag{4.33}$$

In fact, considering the positive self-adjoint operator satisfying (in view of (4.14)) the following coercivity inequality on $\mathrm{L}^2_0(\mu)$:

$$\mathcal{L}=\nabla_x^*\nabla_x\ge c_\phi,$$

and its square-root denoted by $L=\mathcal{L}^{1/2}$, we construct an explicit solution as follows by generalising the construction in [79]:

$$\phi=\nabla_{t,x}\mathscr{W}^{-1}P_{\mathcal{N}^\perp}f+\begin{pmatrix}F_0(t,\mathcal{L})\\ \partial_{x_1}F_1(t,\mathcal{L})\\ \vdots\\ \partial_{x_d}F_1(t,\mathcal{L})\end{pmatrix}P_{\mathcal{N},+}f+\begin{pmatrix}F_0(\tau-t,\mathcal{L})\\ \partial_{x_1}F_1(\tau-t,\mathcal{L})\\ \vdots\\ \partial_{x_d}F_1(\tau-t,\mathcal{L})\end{pmatrix}P_{\mathcal{N},-}f, \tag{4.34}$$

where we introduced the following objects:

- the operator $P_{\mathcal{N}^\perp}$ is the projector onto the orthogonal of the vector space $\mathcal{N} \subset \mathrm{L}^2_0(\mathrm{U}_\tau \otimes \mu)$ composed of linear combinations of forward and backward wave-like functions: $g \in \mathcal{N}$ if and only if there exist $g_+, g_- \in \mathrm{L}^2_0(\mu)$ such that

$$g(t,x) = \left(\mathrm{e}^{-tL} g_+\right)(x) + \left(\mathrm{e}^{-(\tau-t)L} g_-\right)(x); \tag{4.35}$$

- the operator $\mathcal{W} = -\partial_t^2 + \mathcal{L}$ is considered on $\mathrm{L}^2_0(\mathrm{U}_\tau \otimes \mu)$ and endowed with Neumann boundary conditions in time;

- the bounded operators $F_0(t,\mathcal{L})$ and $F_1(t,\mathcal{L})$ respectively act as

$$\begin{aligned} F_0(t,\mathcal{L}) &= -L^{-1}\left(1+\mathrm{e}^{-\tau L}\right)\left(1-\mathrm{e}^{-\tau L}\right)^{-2}\left(1-\mathrm{e}^{-tL}\right)\left(1-\mathrm{e}^{-(\tau-t)L}\right)\left(\mathrm{e}^{-tL} - 2\mathrm{e}^{-\tau L}(1+\mathrm{e}^{-\tau L})^{-1}\right), \\ F_1(t,\mathcal{L}) &= 3L^{-2}\left(1+\mathrm{e}^{-\tau L}\right)\left(1-\mathrm{e}^{-\tau L}\right)^{-2}\left(1-\mathrm{e}^{-tL}\right)\left(1-\mathrm{e}^{-(\tau-t)L}\right); \end{aligned}$$

- the projectors $P_+, P_- : \mathrm{L}^2_0(\mathrm{U}_\tau \otimes \mu) \to \mathcal{N} \subset \mathrm{L}^2_0(\mu)$ are defined as

$$\begin{aligned} (P_{\mathcal{N},+} f)(t) &= \mathrm{e}^{-tL}\left(\int_0^\tau G_+(s,L)\mathrm{e}^{-sL}\,\mathrm{U}_\tau(ds)\right)^{-1}\int_0^\tau G_+(s,L) f(s)\,\mathrm{U}_\tau(ds), \\ (P_{\mathcal{N},-} f)(t) &= \mathrm{e}^{-(\tau-t)L}\left(\int_0^\tau G_-(s,L)\mathrm{e}^{-(\tau-s)L}\,\mathrm{U}_\tau(ds)\right)^{-1}\int_0^\tau G_-(s,L) f(s)\,\mathrm{U}_\tau(ds), \end{aligned} \tag{4.36}$$

  with

$$G_+(s,L) = U_\tau^{-1}\mathrm{e}^{-(2\tau-s)L} - \mathrm{e}^{-sL}, \qquad G_-(s,L) = U_\tau^{-1}\mathrm{e}^{-(\tau+s)L} - \mathrm{e}^{-(\tau-s)L}, \qquad U_\tau = \frac{1}{\tau}\int_0^\tau \mathrm{e}^{-2sL}\,ds. \tag{4.37}$$

The form of this solution is motivated in Section 4.6.3, while the estimates (4.33) are proved in Section 4.6.4. Note already that $P_{\mathcal{N},+}$ and $P_{\mathcal{N},-}$ are bounded operators since $P^2_{\mathcal{N},+} = P_{\mathcal{N},+}$ and $P^2_{\mathcal{N},-} = P_{\mathcal{N},-}$. To obtain the estimates (4.33) for the last two terms in (4.34), it is therefore sufficient to obtain uniform-in-time bounds on the operators

$$F_0(t,\mathcal{L}), \partial_t F_0(t,\mathcal{L}), \partial_{x_i} F_0(t,\mathcal{L}), \partial_{x_i} F_1(t,\mathcal{L}), \partial_t\partial_{x_i} F_1(t,\mathcal{L}), \partial^2_{x_i,x_j} F_1(t,\mathcal{L})$$

considered on $\mathrm{L}^2(\mu)$.

An inspection of the proof finally shows that the factor τ^2 on the right hand side of (4.33) comes from the first part of the solution (4.34) (it can be traced back to some coercivity estimate based on the Poincaré inequality for U_τ), while the factor τ^{-2} arises from the part of the solution arising from $P_{\mathcal{N}} f$, because of the operator $\left(1-\mathrm{e}^{-\tau L}\right)^{-2}$ in the expressions of $F_0(t,\mathcal{L})$ and $F_1(t,\mathcal{L})$.

4.6.3 Heuristic construction of the solution

To find one possible solution to (4.32), a natural idea is to look for $(Z_0, \ldots, Z_d)$ of the form

$$(\partial_t u, \partial_{x_1} u, \ldots, \partial_{x_d} u),$$

with u the unique solution of the following equation

$$(-\partial_t^2 + \mathcal{L})u = f, \qquad \partial_t u(0,\cdot) = \partial_t u(\tau,\cdot) = 0. \tag{4.38}$$

The functions $\partial_{x_i} u(0,\cdot)$ and $\partial_{x_i} u(\tau,\cdot)$ are however not 0 in general. They turn out to vanish provided the right hand side of (4.32) is restricted to the orthogonal of the vector space $\mathcal{N}$. Note that functions $g \in \mathcal{N}$ satisfy $(-\partial_t^2 + \mathcal{L})g = 0$ in the weak sense. This equality is satisfied in the strong sense when g_+, g_- are sufficiently regular, namely $g_+, g_- \in D(\mathcal{L})$. The first step in the proof, as done in [79, Lemma 2.4], is to solve (4.32) for the part of the right hand side f which is in $\mathcal{N}^\perp$, using (4.38) and the *Lax–Milgram* lemma see [72, Chapter 5] for instance.

Lemma 67. *Suppose* (4.15) *and* (4.14). *Fix* $\tau > 0$. *For any* $f \in \mathrm{L}_0^2(\mathrm{U}_\tau \otimes \mu)$, *there exists a unique weak solution* $u \in H^2(\mathrm{U}_\tau \otimes \mu) \cap \mathrm{L}_0^2(\mathrm{U}_\tau \otimes \mu)$ *to* (4.38)*: for any* $w \in \mathrm{H}^1(\mathrm{U}_\tau \otimes \mu)$,

$$\int_0^\tau \int_{\mathcal{X}} [(\partial_t u)(\partial_t w) + \nabla u \cdot \nabla w]\, \mathrm{U}_\tau(dt)\,\mu(dx) = \int_0^\tau \int_{\mathcal{X}} f w\, \mathrm{U}_\tau(dt)\,\mu(dx). \tag{4.39}$$

Moreover,

$$\|u\|_{H^2(\mathrm{U}_\tau\otimes\mu)} \le \left(C_{\mathrm{P}}^2 + C_{\mathrm{P}}^4 + 2c_\phi'\left(\sqrt{d} + 2\max\left(8c_\phi', \sqrt{c_\phi'' d}\right)\right)\right)^{\frac{1}{2}} \|f\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)},$$

with $C_{\mathrm{P}} = \min\left(c_\phi, \frac{\pi}{\tau}\right)$.

The second step is to construct a solution for the part of the right hand side in $\mathcal{N}$. This is done in two steps: (i) providing the expression for the orthogonal projection $P_{\mathcal{N}} = P_{\mathcal{N},+} + P_{\mathcal{N},-}$ of $\mathrm{L}_0^2(\mathrm{U}_\tau \otimes \mu)$ onto $\mathcal{N}$; and then (ii) solving the divergence equation (4.30) with right hand side $P_{\mathcal{N},\pm} f$. To identify $P_{\mathcal{N},\pm}$, we write

$$(P_{\mathcal{N}} f)(t,x) = \left(\mathrm{e}^{-tL} Q_+ f\right)(x) + \left(\mathrm{e}^{-(\tau-t)L} Q_- f\right)(x). \tag{4.40}$$

By definition of the orthogonal projection, it holds, for any $w_+, w_- \in \mathrm{L}^2(\mu)$,

$$\int_0^\tau \int_{\mathcal{X}} [f(t,x) - (P_{\mathcal{N}} f)(t,x)]\left(\mathrm{e}^{-tL} w_+\right)(x)\, \mathrm{U}_\tau(dt)\,\mu(dx) = 0,$$
$$\int_0^\tau \int_{\mathcal{X}} [f(t,x) - (P_{\mathcal{N}} f)(t,x)]\left(\mathrm{e}^{-(\tau-t)L} w_-\right)(x)\, \mathrm{U}_\tau(dt)\,\mu(dx) = 0.$$

Therefore, using the self-adjointness of L,

$$\begin{aligned}
\int_{\mathcal{X}} (U_\tau Q_+ f) w_+\, d\mu + \int_{\mathcal{X}} \left(\mathrm{e}^{-\tau L} Q_- f\right) w_+\, d\mu &= \int_{\mathcal{X}} \left[\int_0^\tau \mathrm{e}^{-tL} f(t)\, \mathrm{U}_\tau(dt)\right] w_+\, d\mu, \\
\int_{\mathcal{X}} \left(\mathrm{e}^{-\tau L} Q_+ f\right) w_-\, d\mu + \int_{\mathcal{X}} (U_\tau Q_- f) w_-\, d\mu &= \int_{\mathcal{X}} \left[\int_0^\tau \mathrm{e}^{-(\tau-t)L} f(t)\, \mathrm{U}_\tau(dt)\right] w_-\, d\mu,
\end{aligned} \tag{4.41}$$

where we recall from (4.37) that

$$U_\tau = \int_0^\tau \mathrm{e}^{-2tL}\, \mathrm{U}_\tau(dt) = \frac{1 - \mathrm{e}^{-2\tau L}}{2\tau} L^{-1}.$$

Choosing $w_+ = U_\tau^{-1} \mathrm{e}^{-\tau L} w_-$ and subtracting the second equation of (4.41) from the first one leads to

$$\int_{\mathcal{X}} \left[\left(U_\tau^{-1} \mathrm{e}^{-2\tau \mathrm{L}^{1/2}} - U_\tau\right) Q_- f\right] w_-\, d\mu = \int_{\mathcal{X}} \left[\int_0^\tau \left(U_\tau^{-1} \mathrm{e}^{-(\tau+t)\mathrm{L}^{1/2}} - \mathrm{e}^{-(\tau-t)\mathrm{L}^{1/2}}\right) f(t)\, \mathrm{U}_\tau(dt)\right] w_-\, d\mu.$$

Since the latter equality holds for any $w_\in \mathrm{L}^2(\mu)$, it follows that

$$Q_- f = \left(U_\tau^{-1} \mathrm{e}^{-2\tau \mathrm{L}^{1/2}} - U_\tau\right)^{-1} \int_0^\tau \left(U_\tau^{-1} \mathrm{e}^{-(\tau+t)\mathrm{L}^{1/2}} - \mathrm{e}^{-(\tau-t)\mathrm{L}^{1/2}}\right) f(t)\, \mathrm{U}_\tau(dt),$$

which leads to (4.36) in view of (4.40). The formula for $Q_+ f$ is obtained in a similar manner. The operators

$Q_\pm$ are well-defined, since the pre-factor of the integrals reads

$$\left(U_\tau^{-1}\mathrm{e}^{-2\tau \mathrm{L}^{1/2}} - U_\tau\right)^{-1} U_\tau^{-1} = \mathrm{L}^2\left(\mathrm{L}^2\mathrm{e}^{-2\tau L} - \left(1-\mathrm{e}^{-2\tau L}\right)^2\right)^{-1}.$$

The latter operator is well defined since $\mathrm{L}^2\mathrm{e}^{-2\tau L} - \left(1-\mathrm{e}^{-2\tau L}\right)^2$ is coercive on $\mathrm{L}^2(\mu)$, hence invertible.

Let us next solve the divergence equation (4.30) with right hand side $(P_{\mathcal{N},+}f)(t) = \mathrm{e}^{-tL}f_+$. The manipulations for the right hand side $\mathrm{e}^{-(\tau-t)L}f_-$ are similar, upon replacing t by $\tau - t$ in the operators which appear. We consider the following ansatz, using a decomposition of the self-adjoint operator $\mathcal{L}$ based on functional calculus, for functions $\mathcal{F}_0, \mathcal{F}_1 : [0,\tau]\times[c_\phi, +\infty[\to \mathbb{R}$ to be determined:

$$Z_{0,+}(t) = \mathcal{F}_0(t,\mathcal{L})f_+, \qquad Z_{i,+}(t) = \partial_{x_i}\mathcal{F}_1(t,\mathcal{L})f_+.$$

Note that these functions satisfy the equation (4.32) with right hand side $\mathrm{e}^{-t\mathcal{L}^{1/2}}f_+$ provided

$$-\partial_t\mathcal{F}_0(t,\mathcal{L}) + \mathcal{L}\mathcal{F}_1(t,\mathcal{L}) = \mathrm{e}^{-t\mathcal{L}^{1/2}}.$$

Moreover, in view of establishing the estimates (4.33) and in order to make use of the fact that $P_{\mathcal{N},+}$ is bounded, it is convenient to rewrite $\mathcal{F}_0(t,\mathcal{L})f_+$ and $\mathcal{F}_1(t,\mathcal{L})f_+$ as the operators on $\mathrm{L}^2(\mu)$

$$F_0(t,\mathcal{L}) = \mathcal{F}_0(t,\mathcal{L})\mathrm{e}^{tL}, \qquad F_1(t,\mathcal{L}) = \mathcal{F}_1(t,\mathcal{L})\mathrm{e}^{tL}, \tag{4.42}$$

applied to the function $(P_{\mathcal{N},+}f)(t) = \mathrm{e}^{-tL}f_+$. This motivates looking for

$$\mathcal{F}_0(t,\mathcal{L}) = P_0\left(\mathrm{e}^{-tL}\right), \qquad \mathcal{F}_1(t,\mathcal{L}) = P_1\left(\mathrm{e}^{-tL}\right), \tag{4.43}$$

where P_0, P_1 are polynomial functions satisfying the operator equation

$$-\frac{d}{dt}\left[P_0\left(\mathrm{e}^{-tL}\right)\right] + \mathrm{L}^2 P_1\left(\mathrm{e}^{-tL}\right) = \mathrm{e}^{-tL}, \tag{4.44}$$

together with the boundary conditions

$$P_0(1) = P_0\left(\mathrm{e}^{-\tau L}\right) = P_1(1) = P_1\left(\mathrm{e}^{-\tau L}\right) = 0. \tag{4.45}$$

In view of the boundary conditions (4.45), and the presence of the operators L and L^2 in (4.44), a natural ansatz is

$$\begin{cases} P_0\left(\mathrm{e}^{-tL}\right) = A_1 L^{-1}\left(1-\mathrm{e}^{-tL}\right)\left(\mathrm{e}^{-tL}-\mathrm{e}^{-\tau L}\right)\left(\mathrm{e}^{-tL}+A_2\right), \\ P_1\left(\mathrm{e}^{-tL}\right) = A_3 L^{-2}\left(1-\mathrm{e}^{-tL}\right)\left(\mathrm{e}^{-tL}-\mathrm{e}^{-\tau L}\right), \end{cases}$$

for operators A_1, A_2, A_3 to be determined.

Note already that the operators $P_0\left(\mathrm{e}^{-tL}\right)$ and $P_1\left(\mathrm{e}^{-tL}\right)$ contain a composition with $\mathrm{e}^{-tL}-\mathrm{e}^{-\tau L}$, which allows to further compose the operators with e^{tL} in order to make sense of (4.42). In order to find the operators A_1, A_2, A_3, which are assumed to be functions of L (and hence commute with all operators at hand), we evaluate (4.44): denoting by $\theta_\star = \mathrm{e}^{-\tau L}$,

$$A_1\frac{d}{d\theta}\left[(1-\theta)(\theta-\theta_\star)(\theta+A_2)\right] + A_3(1-\theta)(\theta-\theta_\star) = \theta,$$

where the various functions are evaluated at $\theta = \mathrm{e}^{-tL}$. The left hand side of this equation is a second order polynomial in θ, so that the operators A_1, A_2, A_3 are obtained by identifying the prefactors of the various terms θ^a for $a = 0,1,2$. More precisely,

$$\begin{cases} -(A_1 + A_3)\theta_\star + A_1 A_2(1+\theta_\star) = 0, \\ (2A_1 + A_3)(1+\theta_\star) - 2A_1 A_2 = 1, \\ -3A_1 - A_3 = 0. \end{cases}$$

The last condition implies that $A_3 = -3A_1$. Plugging this equality in the first condition leads to $A_2 = -2\mathrm{e}^{-\tau L}(1+\mathrm{e}^{-\tau L})^{-1}$. Finally, the second condition gives $A_1(1+\theta_\star + 2A_2) = -1$, so that

$$A_1 = -\left[1+\mathrm{e}^{-\tau L} - 4\mathrm{e}^{-\tau L}\left(1+\mathrm{e}^{-\tau L}\right)^{-1}\right]^{-1} = -\left(1+\mathrm{e}^{-\tau L}\right)\left(1-\mathrm{e}^{-\tau L}\right)^{-2}.$$

Combining (4.42) and (4.43) then leads to the expressions of F_0, F_1 in (4.34).

4.6.4 Proof of the estimates on the divergence equation

The estimate on the first term in (4.34) is a direct consequence of Lemma 67.

Proof of Lemma 67. The existence and uniqueness of the weak solution satisfying the estimate is established in [79, Lemma 2.4]. To make the constants precise, we revisit the steps of [79], just highlighting the differences.

1. There exists a unique solution $u \in \mathrm{H}^1(\mathrm{U}_\tau \otimes \mu)$ to (4.39), which satisfies

$$\|u\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le C_{\mathrm{P}}^2 \|f\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}, \tag{4.46}$$

$$\|\nabla_{t,x} u\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le C_{\mathrm{P}} \|f\|_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}. \tag{4.47}$$

   This is a consequence of Lax-Milgram's Theorem for the tensorised measure $\mathrm{U}_\tau \otimes \mu$ and the Poincaré inequality (4.14).

2. Thanks to (4.21) we have

$$\|\nabla_{t,x}^2 u\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} \le 2\|f\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)} + 2c'_\phi\left(\sqrt{d} + 2\max\left(8c'_\phi, \sqrt{c''_\phi\, d}\right)\right)\|\nabla_{t,x} u\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}.$$

3. Hence, by (4.46),

$$\|u\|^2_{\mathrm{H}^2(\mathrm{U}_\tau\otimes\mu)} \le \left(2 + C_P^2 + C_P^4 + 2C_{\mathrm{P}}^{-2}\, c'_\phi\left(\sqrt{d} + 2\max\left(8c'_\phi, \sqrt{c''_\phi\, d}\right)\right)\right)\|f\|^2_{\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu)}.$$

We conclude by showing that the sought boundary conditions are met, if $f \in \mathcal{N}^\perp$. Consider an element $w \in \mathcal{N}$, written as (4.35) for two smooth functions $w_+, w_- \in \mathrm{L}_0^2(\mu)$ with compact support. Recall that $(-\partial_t^2 + \mathcal{L})w = 0$. Then, the weak formulation (4.39) and integration by parts give

$$\begin{aligned}\int_{\mathcal{X}} \left(u(\tau,\cdot)\partial_t w(\tau,\cdot) - u(0,\cdot)\partial_t w(0,\cdot)\right] d\mu &= \int_0^\tau \int_{\mathcal{X}} [(\partial_t u)(\partial_t w) + \nabla_x u\cdot\nabla_x w]\,\mathrm{U}_\tau(dt)\,\mu(dx) \\ &= \int_0^\tau \int_{\mathcal{X}} f w\,\mathrm{U}_\tau(dt)\,\mu(dx) = 0.\end{aligned}$$

the latter integral being equal to 0 since $f \in \mathcal{N}^\perp$ by assumption. Therefore,

$$\int_{\mathcal{X}} \left[u(\tau,\cdot)\mathcal{L}^{1/2}\left(-\mathrm{e}^{-\tau\mathcal{L}^{1/2}} w_+ + w_-\right) - u(0,\cdot)\mathcal{L}^{1/2}\left(-w_+ + \mathrm{e}^{-\tau\mathcal{L}^{1/2}} w_-\right)\right] d\mu = 0.$$

Fixing w_- and choosing $w_+ = \mathrm{e}^{-\tau\mathcal{L}^{1/2}} w_-$ leads to

$$\int_{\mathcal{X}} u(\tau,\cdot)\mathcal{L}^{1/2}\left(1-\mathrm{e}^{-2\tau\mathcal{L}^{1/2}}\right) w_- \, d\mu = 0.$$

Since this equality holds for any smooth $w_- \in \mathrm{L}^2_0(\mu)$ with compact support, and the operator

$$\mathcal{L}^{1/2}\left(1-\mathrm{e}^{-2\tau\mathcal{L}^{1/2}}\right)$$

is invertible on $\mathrm{L}^2_0(\mu)$, we can conclude that $u(\tau,\cdot) = 0$. A similar reasoning allows to conclude that $u(0) = 0$. □

We next show that the second and third terms in (4.34) satisfy (4.33). Given that the two terms are extremely similar, we only treat the second one. Note first that the operators $T_i = \partial_{x_i} L^{-1}$ are bounded on $\mathrm{L}^2(\mu)$ and have an operator norm smaller than 1 since $T_1^* T_1 + \cdots + T_d^* T_d = 1$. Moreover, the operators $\partial_{x_i}\partial_{x_j} L^{-2}$ for $1 \le i,j \le d$ are also bounded:

$$\|\partial_{x_i}\partial_{x_j} L^{-2}\|_{\mathcal{B}(\mathrm{L}^2(\mathrm{U}_\tau\otimes\mu))} \le \sqrt{2+2\frac{c_\phi'}{C_\mathrm{P}^2}\left(\sqrt{d}+2\max\left(8c_\phi', \sqrt{c_\phi'' d}\right)\right)},$$

by (4.21); see also [127], as well as [79, Lemma 2.3] and [44, Proposition 3.5]. The estimates (4.33) are therefore satisfied provided there exists $C_\tau \in \mathbb{R}_+$ such that

$$\sup_{t\in[0,\tau]} \max\left(\|F_0(t,\mathcal{L})\|_{\mathcal{B}(\mathrm{L}^2(\mu))}, \|\partial_t F_0(t,\mathcal{L})\|_{\mathcal{B}(\mathrm{L}^2(\mu))}, \|LF_0(t,\mathcal{L})\|_{\mathcal{B}(\mathrm{L}^2(\mu))}\right) \le C_\tau,$$
$$\sup_{t\in[0,\tau]} \max\left(\|LF_1(t,\mathcal{L})\|_{\mathcal{B}(\mathrm{L}^2(\mu))}, \|L\partial_t F_1(t,\mathcal{L})\|_{\mathcal{B}(\mathrm{L}^2(\mu))}, \left\|\mathrm{L}^2 F_1(t,\mathcal{L})\right\|_{\mathcal{B}(\mathrm{L}^2(\mu))}\right) \le C_\tau.$$

The explicit expressions of F_0, F_1 show that C_τ can be chosen of order

$$\left\|1-\mathrm{e}^{-\tau L}\right\|^{-2}_{\mathcal{B}(\mathrm{L}^2_0(\mu))} = \left(\frac{1}{1-\mathrm{e}^{-\tau c_\phi}}\right)^2.$$

Note that C_τ scales as τ^{-2} for τ small.

Part III

Functional inequalities

Chapter 5

Logarithmic Sobolev and interpolation inequalities on the sphere: constructive stability results

This chapter corresponds to [P2], a paper in collaboration with Jean Dolbeault and Nikita Simonov.

Abstract

We consider Gagliardo–Nirenberg inequalities on the sphere which interpolate between the Poin-caré inequality and the Sobolev inequality, and include the logarithmic Sobolev inequality as a special case. We establish explicit stability results in the subcritical regime using spectral decomposition techniques, and entropy and *carré du champ* methods applied to nonlinear diffusion flows.

5.1 Introduction and main results

Functional inequalities are essential in many areas of mathematics. The knowledge of optimal constants, or at least good estimates of them, is crucial for various applications. Whether optimality cases are achieved is a standard issue in analysis. The next natural question is to understand how the deficit, say the difference of the two sides of the functional inequality, measures the distance to the set of optimal functions. Such a question has been actively studied in critical Sobolev inequalities, but much less in subcritical interpolation inequalities. In the case of the sphere, a global stability result based on Bianchi-Egnell type methods was recently obtained for a family of Gagliardo-Nirenberg inequalities by R. Frank in [144], with the striking observation that only the power 4 of a natural distance is controlled by the deficit. Here we give a more detailed picture, which includes the logarithmic Sobolev inequality, and provide explicit estimates.

On the sphere $\mathbb{S}^d$ with $d \geq 1$, the *logarithmic Sobolev inequality* can be written as

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \geq \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu) \tag{LS}$$

where $d\mu_d$ denotes the uniform probability measure. The equality case is achieved by constant functions and $d/2$ is the optimal constant as shown by taking the test functions $F_\varepsilon(x) = 1 + \varepsilon\, x \cdot \nu$, for some arbitrary

$\nu \in \mathbb{S}^d$, in the limit as $\varepsilon \to 0$. Our first result is an improved inequality under an orthogonality constraint, which improves upon [122, Proposition 5.4].

Theorem 68. *Let $d \geq 1$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$ such that*

$$\int_{\mathbb{S}^d} x\, F\, d\mu_d = 0\,, \tag{5.1}$$

we have

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d - \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \geq \mathscr{C}_d \int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d \tag{5.2}$$

with $\mathscr{C}_d = \frac{2}{d+2}$.

Since equality in (LS) is achieved if and only if F is a constant function, the right-hand side in (5.2) is an estimate of the distance to the set of optimal functions under the constraint $\int_{\mathbb{S}^d} x\, F\, d\mu_d = 0$. Alternatively, Theorem 68 amounts to the *improved logarithmic Sobolev inequality*

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d \geq \frac{d+2}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu) \text{ s.t. } \int_{\mathbb{S}^d} x\, F\, d\mu_d = 0\,.$$

Without Condition (5.1), there is no such inequality as (5.2). With $F_\varepsilon(x) = 1 + \varepsilon\, x \cdot \nu$ as above, as $\varepsilon \to 0$ one can indeed check that

$$\|\nabla F_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{S}^d)} - \frac{d}{2} \int_{\mathbb{S}^d} F_\varepsilon^2 \log\left(\frac{F_\varepsilon^2}{\|F_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d = O(\varepsilon^4) = O\left(\|\nabla F_\varepsilon\|^4_{\mathrm{L}^2(\mathbb{S}^d)}\right).$$

In absence of an additional constraint, like (5.1), such a behaviour is in fact optimal. The following estimate arises from the *carré du champ* method.

Proposition 69. *Let $d \geq 1$, $\gamma = 1/3$ if $d = 1$ and $\gamma = (4d-1)(d-1)^2/(d+2)^2$ if $d \geq 2$. Then, for any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$ we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d - \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \geq \frac{1}{2} \frac{\gamma\, \|\nabla F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\gamma\, \|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + d\, \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\,. \tag{5.3}$$

With $\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = 1$, notice that the deficit can be estimated from below by

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d - \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(F^2\right) d\mu_d \geq \frac{\gamma}{2\,d} \|\nabla F\|^4_{\mathrm{L}^2(\mathbb{S}^d)} + o\left(\|\nabla F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}\right)$$

if $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ is small enough.

Let $\Pi_1 F$ denote the orthogonal projection of a function $F \in \mathrm{L}^2(\mathbb{S}^d)$ on the spherical harmonics corresponding to the first eigenvalue on $\mathbb{S}^d$, *i.e.*,

$$\Pi_1 F(x) = (d+1)\, x \cdot \int_{\mathbb{S}^d} y\, F(y)\, d\mu_d(y) \quad \forall\, x \in \mathbb{S}^d\,.$$

Our main stability result for the logarithmic Sobolev inequality combines the results of Theorem 68 and Proposition 69 as follows.

Theorem 70. *Let $d \geq 1$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{2} \int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \geq \mathscr{S}_d \left(\frac{\|\nabla \Pi_1 F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \frac{d}{2}\,\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} + \|\nabla(\mathrm{Id} - \Pi_1)\,F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right)$$

for some stability constant $\mathscr{S}_d > 0$.

An explicit estimate of $\mathscr{S}_d$ is given in Section 5.4.

We also consider the subcritical *Gagliardo-Nirenberg inequalities*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \geq \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu) \tag{GN}$$

for any $p \in [1,2) \cup (2, 2^*)$. Here $d\mu_d$ again denotes the uniform probability measure on $\mathbb{S}^d$, the critical Sobolev exponent is $2^* := 2\,d/(d-2)$ if $d \geq 3$ and we adopt the convention that $2^* = +\infty$ if $d = 1$ or $d = 2$. Inequality (GN) with $p = 1$ is the Poincaré inequality. If $d \geq 3$, Inequality (GN) also holds for the critical exponent $p = 2^*$ and it is in fact Sobolev's inequality with optimal constant on $\mathbb{S}^d$, but this is out of the scope of our paper which focuses on the subcritical regime $p < 2^*$. The logarithmic Sobolev inequality (LS) is obtained from (GN) by taking the limit as $p \to 2$ and the counterpart of the above results for $p \neq 2$, in the *subcritical range* $p < 2^*$, goes as follows.

Theorem 71. *Assume that $d \geq 1$ and $p \in (1,2) \cup (2, 2^*)$. For any function $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$ such that the orthogonality condition* (5.1) *holds, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \geq \mathscr{C}_{d,p} \int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \tag{5.4}$$

with $\mathscr{C}_{d,p} = \frac{2\,d - p\,(d-2)}{2\,(d+p)}$.

Taking $F_\varepsilon(x) = 1 + \varepsilon\, x \cdot \nu$ as above shows that (5.1) is needed in Theorem 71. We also have a higher order estimate of the deficit as a consequence of the *carré du champ* method.

Proposition 72. *Let $d \geq 1$ and $p \in (1,2) \cup (2, 2^*)$. There is a convex function ψ on $\mathbb{R}^+$ with $\psi(0) = \psi'(0) = 0$ such that, for any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \geq \|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\, \psi\left(\frac{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\right).$$

An explicit expression of ψ will be given in Section 5.3. The two results of Theorem 71 and Proposition 72 can be combined to prove the analogue of Theorem 70 for $p \neq 2$, with an explicit constant: see Section 5.4.

Theorem 73. *Let $d \geq 1$ and $p \in (1,2) \cup (2, 2^*)$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - \frac{d}{p-2}\left(\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \geq \mathscr{S}_{d,p} \left(\frac{\|\nabla \Pi_1 F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} + \|\nabla(\mathrm{Id} - \Pi_1)\,F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right)$$

for some explicit stability constant $\mathscr{S}_{d,p} > 0$.

Let us give a brief account of the literature. In this paper, we address the distinction between *improved inequalities* (inequalities with improved constants under orthogonality constraints) and *quantitative stability* (as a measure of a distance to the set of optimal functions). There are many adjacent directions of research like, for instance, stability in weaker norms (see for instance [125, 175] for Sobolev's inequality) or notions of stability with no explicit notion of distance. To our knowledge, not so much has been done in subcritical interpolation inequalities (see [63, 144] and some references therein), except for the logarithmic Sobolev inequality, for which we refer to [130, 138] and [173, 140, 172, 182].

The *Gagliardo-Nirenberg inequalities* (GN) on the sphere have been established with optimal constant for any $p \in (2, 2^*)$ in [48, Corollary 6.1] and in [37]. In dimension $d = 2$, Onofri's inequality is obtained from (GN) in the limit as $p \to 2^* = +\infty$: see [37, 80]. With $p \in [1, 2)$ or $p > 2$ but not too large (if $d \geq 2$), Inequality (GN) was known earlier from [24]. A Markovian point of view is presented in [26], with many more references therein on related questions. On the Euclidean space, such inequalities go back to [240, 147, 222]. The *logarithmic Sobolev inequality* (LS) is a well known limit case as $p \to 2$ and can be considered in a common framework with (GN). Whenever possible, we will adopt this point of view. For an overview of early results on the sphere, we refer to [Section 6, (iv)] [156]. The literature on (LS) on the circle and on the sphere can be retraced back at least to [260], [216, Theorem 1, page 268] with computations based on the ultraspherical operator, and [234] for a more variational approach. The inequality with optimal constant is stated in [24, Inequality (13) page 195] as a consequence of the carré du champ method. Also see [25] and [83, page 342] for related results and [113, 121, 122] for a PDE approach based on entropy estimates and the *carré du champ* method. After symmetrization, the problem is reduced to a simpler family of interpolation inequalities involving only the ultraspherical operator.

The interest for stability issues was raised by [73] and the stability result of Bianchi and Egnell in [47], on the Euclidean space. Over the years, various approaches have been developed, based on compactness methods and contradiction arguments as in [47, 97], spectral analysis and orthogonality conditions as in [122, Proposition 5.4] and [159], or entropy methods and improved inequalities as in [15, 129, 120, 122, 118]. For spectral methods, a fruitful strategy relies on the Funk-Hecke formula and the approach of [196, 37], which applies to the stability result for fractional interpolation inequalities of [97] and [131, Corollary 2.3]. This is the method we use in Section 5.2. Stability issues for (GN) have recently been discussed in [144] with methods of Bianchi-Egnell type, with the drawback that no estimate of the stability constant is known. This drawback can be cured by a *carré du champ* method as we shall see in Section 5.3. Without entering into details, let us mention some recent progress on stability in [63, 119, 184, 96] for related critical inequalities.

This paper is organized as follows. Section 5.2 is devoted to the proof by spectral methods of Theorem 74 (see below), which is an extension of Theorems 68 and 71: under orthogonality constraints, these results are reduced to estimates of improved constants in inequalities (LS) and (GN), with various refinements based on a decomposition in spherical harmonics. An explicit stability result without constraints corresponding to Propositions 5.3 and 72 is proved in Section 5.3. The proof of Theorems 70 and 73, in Section 5.4, is based on the spectral decomposition method developed by R. Frank in [144]. We collect the previous estimates (with and without orthogonality constraints) in global results, with explicit constants. Various additional results are stated in two appendices: the extension of the method to interpolation inequalities for the Gaussian measure on the Euclidean space and a discussion of its limitations in Appendix 5.5, the details of the computations of the carré du champ method on the sphere and its application in order to establish improved functional inequalities in Appendix 5.6.

5.2 Improvements under orthogonality constraints

In this section, we prove Theorems 68 and 71 in the slightly more general framework of Theorem 74 below. Let us consider the generalized entropy functionals

$$\mathcal{E}_2[F] := \frac{1}{2}\int_{\mathbb{S}^d} F^2 \log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d \quad \text{and} \quad \mathcal{E}_p[F] := \frac{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{p-2} \quad \text{if} \quad p \neq 2 .$$

With these notations, we can rephrase (LS) and (GN) as

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \geq d\, \mathcal{E}_p[F] \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$$

for any $p \in [1, 2^*)$. The optimality case is achieved by considering the test function $F_\varepsilon = 1 + \varepsilon\,\varphi_1$ in the limit as $\varepsilon \to 0$, where φ_1 is an eigenfunction of the Laplace-Beltrami operator such that $-\Delta\varphi_1 = d\,\varphi_1$, for instance $\varphi_1(x) = x\cdot\nu$ for some $\nu \in \mathbb{S}^d$ as in Section 5.1.

Let us consider the decomposition into spherical harmonics of $\mathrm{L}^2(\mathbb{S}^d, d\mu)$

$$\mathrm{L}^2(\mathbb{S}^d, d\mu) = \bigoplus_{\ell=0}^{\infty} \mathcal{H}_\ell ,$$

where $\mathcal{H}_\ell$ is the subspace of spherical harmonics of degree $\ell \geq 0$. See for instance [218, 242, 43, 213]. For any integer $k \geq 1$, let us define Π_k as the orthogonal projection with respect to $\mathrm{L}^2(\mathbb{S}^d, d\mu)$ onto $\bigoplus_{\ell=1}^{k} \mathcal{H}_\ell$. The following statement extends Theorems 68 and 71.

Theorem 74. *Assume that $d \geq 1$, $p \in (1, 2^*)$ and let $k \geq 1$ be an integer. For any function F in $\mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d - d\, \mathcal{E}_p[F] \geq \mathscr{C}_{d,p,k} \int_{\mathbb{S}^d} \left|\nabla (\mathrm{Id} - \Pi_k)\, F\right|^2 d\mu_d \tag{5.5}$$

for some explicit constant $\mathscr{C}_{d,p,k} \in (0,1)$ such that $\mathscr{C}_{d,p,k} \leq \mathscr{C}_{d,p,1} = \frac{2d - p\,(d-2)}{2\,(d+p)}$.

The expression of $\mathscr{C}_{d,p,k}$ is given below in the proof. Inequality (5.5) can be seen as an improvement of (LS) and (GN), namely

$$(1 - \mathscr{C}_{d,p,k}) \int_{\mathbb{S}^d} |\nabla F|^2 \, d\mu_d \geq d\, \mathcal{E}_p[F]$$

for any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$ such that $\Pi_k F = 0$. With $k = 1$, this establishes (5.2) and (5.4), thus proving Theorem 68 if $p = 2$, and Theorem 71 if $p \neq 2$.

Proof. Let $(F_j)_{j\in\mathbb{N}}$ be the decomposition of F along $\mathcal{H}_j$ for any $j \in \mathbb{N}$. We learn from [37, Ineq. (19)] or [131, Ineq. (1.6)] that the *subcritical interpolation inequalities*

$$\mathcal{E}_p[F] \leq \sum_{j=1}^{\infty} \zeta_j(p) \int_{\mathbb{S}^d} |F_j|^2 \, d\mu_d \quad \forall\, F \in \mathrm{H}^1(\mathbb{S}^d, d\mu) \tag{5.6}$$

hold for any $p \in (1,2) \cup (2, 2^*)$ with

$$\zeta_j(p) := \frac{\gamma_j\left(\frac{d}{p}\right) - 1}{p-2} \quad \text{and} \quad \gamma_j(x) := \frac{\Gamma(x)\,\Gamma(j+d-x)}{\Gamma(d-x)\,\Gamma(x+j)} .$$

Notice that $\zeta_j(p) \geq 0$ for any $p \in (1,2) \cup (2, 2^*)$. According to [131, Lemma 2.2], the function ζ_j is strictly monotone increasing on $(1,\infty)$ for any $j \ge 2$ and the limits

$$\lambda_j = d \lim_{p\to 2^*} \zeta_j(p)$$

are the eigenvalues of the Laplace-Beltrami on the sphere, with $\lambda_j = j\,(j+d-1)$. Hence

$$d\,\mathcal{E}_p[F] \le \sum_{j=1}^{\infty} \lambda_j \int_{\mathbb{S}^d} |F_j|^2\,d\mu_d = \int_{\mathbb{S}^d} |\nabla F|^2\,d\mu_d\,,$$

which is the essence of the proof of (GN) in [37] and also the main idea for the proof of the stability result for fractional interpolation inequalities of [131, Corollary 2.3]. Here we draw some consequences in standard norms for non-fractional operators and identify estimates of the stability constant in the corresponding stability result.

▷ *The case* $p \neq 2$. Let $x = d/p \in ((d-2)/2, d]$ if $d \ge 2$ and $x \in (0,d]$ if $d = 1$. We consider

$$\xi_j(x) := \frac{|\gamma_j(x)-1|}{j\,(j+d-1)} \quad \text{and} \quad h_j(x) = \frac{j\,(j+d-1)\,(j+d-x)}{(j+1)\,(j+d)\,(j+x)}\,,$$

notice that $\gamma_j(x) > 1$ for $x < d/2$, while $\gamma_j(x) < 1$ for $x > d/2$. An elementary computation shows that $0 < h_j(x) < 1$. Since $\gamma_{j+1}(x)\,\lambda_\kappa = h_j(x)\,\lambda_{j+1}\,\gamma_j(x)$, we obtain

$$\xi_{j+1}(x) = h_j(x)\,\xi_j(x) + (1-h_j(x))\,\xi_j^\star(x) \tag{5.7}$$

where

$$\xi_j^\star(x) := \frac{1}{1-h_j(x)} \left|\frac{h_j(x)}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right| = \frac{|d-2x|}{j\,(j+d)\,(2x-d+2)+d\,x}\,.$$

Notice that $(\xi_j^\star(x))_{j\ge2}$ is a monotone decreasing sequence for any fixed, admissible value of x. We start at $j=2$ with the observation that $\xi_2^\star(x) < \xi_2(x)$ if x is admissible. This gives, by using (5.7), the following estimate

$$\xi_3(x) = h_2(x)\,\xi_2(x) + \big(1-h_2(x)\big)\,\xi_2^\star(x) < \xi_2(x)\,.$$

Using $\xi_3^\star(x) < \xi_2^\star(x)$, we can iterate and conclude by induction that $\xi_j(x) < \xi_2(x)$ for all $j \ge 3$. As a consequence, we obtain

$$\sup_{j\ge3} \frac{\zeta_j(p)}{j\,(j+d-1)} < \frac{\zeta_2(p)}{2\,(d+1)} = \frac{p}{2\,(d+p)} < \frac{1}{d} \quad \forall\, p \in (1,2)\cup(2,2^*)\,.$$

We deduce from (5.6) that

$$\mathcal{E}_p[F] \le \int_{\mathbb{S}^d} |F_1|^2\,d\mu_d + \frac{p}{2\,(d+p)} \sum_{j=2}^{\infty} j\,(j+d-1) \int_{\mathbb{S}^d} |F_j|^2\,d\mu_d$$
$$= \frac{1}{d} \int_{\mathbb{S}^d} |\nabla F|^2\,d\mu_d + \frac{2\,d-p\,(d-2)}{2\,d\,(d+p)} \int_{\mathbb{S}^d} |\nabla(\mathrm{Id}-\Pi_1)\,F|^2\,d\mu_d\,,$$

which proves the result with $k=1$ and gives the expression of $\mathscr{C}_{d,p,1}$.

Let us consider the case $k > 1$. We already know that $\xi_2(x) > \xi_2^\star(x)$. For any $j \ge 2$, we deduce from (5.7) that

$$\xi_{j+1}(x) - \xi_{j+1}^\star(x) = h_j(x)\,\big(\xi_j - \xi_j^\star(x)\big) + \xi_j^\star(x) - \xi_{j+1}^\star(x) \ge h_j(x)\,\big(\xi_j - \xi_j^\star(x)\big)$$

because $j\mapsto\xi_j^\star(x)$ is monotone decreasing. By induction, this proves that $\xi_j(x)>\xi_j^\star(x)$ for any $j\ge2$. As a consequence of (5.7), $j\mapsto\xi_j(x)$ is also monotone decreasing and

$$\sup_{j\ge k+2}\frac{\zeta_j(p)}{j\,(j+d-1)}<\frac{\zeta_{k+1}(p)}{(k+1)\,(k+d)}<\frac{1}{d}\quad\forall\,p\in(1,2)\cup(2,2^*)\,.$$

Altogether, for any $k\ge1$, we have

$$d\,\mathcal{E}_p[F]\le\int_{\mathbb{S}^d}|\nabla\Pi_kF|^2\,d\mu_d+\frac{d\,\zeta_{k+1}(p)}{(k+1)\,(k+d)}\int_{\mathbb{S}^d}|\nabla F|^2\,d\mu_d\,,$$

and the stability constant in (5.5) is estimated by

$$\mathscr{C}_{d,p,k}=1-\frac{d\,\zeta_{k+1}(p)}{(k+1)\,(k+d)}\,.$$

In our method, this constant cannot be improved as shown by a test function such that $F_j=0$ for any $j\in\mathbb{N}$ such that $j\neq0$ and $j\neq k+1$, but this does not prove the optimality of $\mathscr{C}_{d,p,k}$.

▷ *The case* $p=2$. By taking the limit as $p\to2_+$ in (5.6), we obtain that

$$\eta_j:=\frac{2}{d}\lim_{p\to2_+}\zeta_j(p)=\psi(j+d/2)-\psi(d/2)$$

where $\psi(z)=\Gamma'(z)/\Gamma(z)$ is the *digamma function*, and

$$\frac12\int_{\mathbb{S}^d}F^2\log\left(\frac{F^2}{\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right)d\mu_d\le\frac{d}{2}\sum_{j=1}^{\infty}\eta_j\int_{\mathbb{S}^d}|F_j|^2\,d\mu_d\quad\forall\,F\in\mathrm{H}^1(\mathbb{S}^d,d\mu)\,.$$

From $\psi(z+1)=\psi(z)+1/z$ obtained by differentiating the identity $\Gamma(z+1)=z\,\Gamma(z)$ with respect to z, we learn that

$$\eta_{j+1}=\eta_j+\frac{2}{d+2\,j}\,.$$

We claim that

$$\eta_2\le\eta_j\le\frac{2\,\lambda_j}{d\,(d+2)}\quad\forall\,j\ge2$$

because there is equality for $j=2$ as $\eta_2=\frac{4\,(d+1)}{d\,(d+2)}$ and $\lambda_2=2\,(d+1)$ on the one hand, and

$$\eta_{j+1}-\eta_j=\frac{2}{d+2\,j}\le\frac{2\,(d+2\,j)}{d\,(d+2)}=\frac{2\,(\lambda_{j+1}-\lambda_j)}{d\,(d+2)}$$

on the other hand, so that the result follows by induction.

Using $\lambda_{j+1}=\lambda_j+(d+2\,j)=\lambda_j+2\,z_j$ where $z_j:=j+d/2$, we also have

$$\frac{\eta_{j+1}}{\lambda_{j+1}}=\frac{\eta_j+\frac{1}{z_j}}{\lambda_j+2\,z_j}<\frac{\eta_j}{\lambda_j}$$

where the inequality follows from

$$z_j^2>\frac{\lambda_j}{2\,\eta_j}\quad\forall\,j\ge1\,.$$

This inequality is indeed true for $j=1$ because $\eta_1=2/d$ and we obtain the result for any $j\ge1$ by induction using

$$\eta_{j+1}-\eta_j=\frac{2}{d+2\,j}\geq\frac{\lambda_{j+1}}{2\,z_{j+1}^2}-\frac{\lambda_j}{2\,z_j^2}=2\,\frac{4\,j^2+2\,(d^2+2)\,j+d^3}{(d+2\,j)^2\,(d+2+2\,j)^2}\,.$$

Altogether, for any $k\geq 1$, we have

$$d\,\mathcal{E}_2[F]\leq\int_{\mathbb{S}^d}|\nabla\Pi_k F|^2\,d\mu_d+\frac{d\,\eta_{k+1}}{(k+1)\,(k+d)}\int_{\mathbb{S}^d}|\nabla F|^2\,d\mu_d$$

and the constant in (5.5) is given by

$$\mathscr{C}_{d,2,k}=1-\frac{d\,\eta_{k+1}}{(k+1)\,(k+d)}\,.$$

In the framework of our method, this estimate of the constant cannot be improved as shown by a test function such that $F_j=0$ for any $j\in\mathbb{N}$ such that $j\neq 0$ and $j\neq k+1$, but again this does not prove the optimality of $\mathscr{C}_{d,p,k}$. □

5.3 Improvements by the carré du champ method

We improve upon Frank's stability result in [144] by giving a constructive estimate based on the *carré du champ* method, without assuming any additional constraint. Various computations that are needed for a complete proof, most of them already known in the literature, are collected in Appendix 5.6.

5.3.1 A simple estimate based on the heat flow, below the Bakry-Emery exponent

Let us consider the constant γ given by

$$\gamma:=\left(\frac{d-1}{d+2}\right)^2(p-1)\,(2^\#-p)\quad\text{if}\quad d\geq 2\,,\quad\gamma:=\frac{p-1}{3}\quad\text{if}\quad d=1\,,\tag{5.8}$$

where $2^\#:=\frac{2\,d^2+1}{(d-1)^2}$ is the *Bakry-Emery exponent*. Notice that $\gamma=2-p$ with $1\leq p\leq 2^\#$ means that

$$\begin{aligned}&d=1\quad\text{and}\quad p=7/4=p_\star(1)\,,\\&d>1\quad\text{and}\quad p=p_\star(d):=\frac{3+d+2\,d^2-2\sqrt{4\,d+4\,d^2+d^3}}{(d-1)^2}\,.\end{aligned}$$

Let us define

$$s_\star:=\frac{1}{p-2}\quad\text{if}\quad p>2\quad\text{and}\quad s_\star:=+\infty\quad\text{if}\quad p\leq 2\,.\tag{5.9}$$

For any $s\in[0,s_\star)$, let

$$\begin{aligned}&\varphi(s)=\tfrac{1-(p-2)\,s-(1-(p-2)\,s)^{-\frac{\gamma}{p-2}}}{2-p-\gamma}&&\text{if}\quad\gamma\neq 2-p\quad\text{and}\quad p\neq 2\,,\\&\varphi(s)=\tfrac{1}{2-p}\,(1+(2-p)\,s)\log\big(1+(2-p)\,s\big)&&\text{if}\quad\gamma=2-p\neq 0\,,\\&\varphi(s)=\tfrac{1}{\gamma}\,(e^{\gamma\,s}-1)&&\text{if}\quad p=2\,.\end{aligned}\tag{5.10}$$

In [118, Theorem 2.1] (also see [120] and earlier related references therein) the *improved Gagliardo-Nirenberg* *inequalities*

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \ge d\,\varphi\left(\frac{\mathcal{E}_p[F]}{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\right)\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} \quad \forall\,F\in\mathrm{H}^1(\mathbb{S}^d) \tag{5.11}$$

are stated with γ given by (5.8) under the conditions

$$d\ge1 \quad\text{and}\quad 1\le p\le 2^\# \quad\text{if}\quad d\ge2\,,\quad p\ge1 \quad\text{if}\quad d=1\,.$$

Why this estimate is based on the heat flow is explained in Appendix 5.6. Additional justifications and the discussion of the case $p=2$ are also given in Appendix 5.6.

Since $\varphi(0)=0$, $\varphi'(0)=1$, and φ is convex increasing, with an asymptote at $s=s_\star$ if $p\in(2,2^\#)$, we know that $\varphi:[0,s_\star)\to\mathbb{R}^+$ is invertible and $\psi:\mathbb{R}^+\to[0,s_\star)$, $s\mapsto\psi(s):=s-\varphi^{-1}(s)$, is convex increasing with $\psi(0)=\psi'(0)=0$, $\lim_{t\to+\infty}\big(t-\psi(t)\big)=s_\star$, and

$$\psi''(0)=\varphi''(0)=\frac{(d-1)^2}{(d+2)^2}\left(2^\#-p\right)(p-1)>0 \quad \forall\,p\in(1,2^\#)\,.$$

Proposition 75. *With the above notations, $d\ge1$ and $p\in(1,2^\#)$, we have*

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}-d\,\mathcal{E}_p[F]\ge d\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\,\psi\left(\frac{1}{d}\,\frac{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\right) \quad \forall\,F\in\mathrm{H}^1(\mathbb{S}^d)\,.$$

If $p=2$, notice that ψ is explicit and given by

$$\psi(t):=t-\frac{1}{\gamma}\log(1+\gamma\,t) \quad \forall\,t\ge0\,.$$

The proof of Proposition 69 follows from the observation that $\psi(t)\ge\frac{\gamma}{2}\,\frac{t^2}{1+\gamma\,t}$ for any $t\ge0$.

5.3.2 An estimate based on the fast diffusion flow, valid up to the critical exponent

The subcritical range $p\in[2^\#,2^*)$ corresponding to exponents between the *Bakry-Emery exponent* and the critical Sobolev exponent is not covered in Section 5.3.1. In that case, we rely on entropy methods based on a fast diffusion or porous medium equation of exponent m which are detailed in Section 5.6 (with corresponding references), to establish that an improved inequality (5.11) holds for any $\varphi=\varphi_{m,p}$ where

$$\varphi_{m,p}(s):=\int_0^s\exp\left[-\zeta\left((1-(p-2)\,z)^{1-\delta}-(1-(p-2)\,s)^{1-\delta}\right)\right]dz \tag{5.12}$$

provided $m\in\mathscr{A}_p:=\mathscr{A}_p:=\left\{m\in\left[m_-(d,p),m_+(d,p)\right]\,:\,\frac{2}{p}\le m<1 \text{ if } p<4\right\}$ where

$$m_\pm(d,p):=\frac{1}{(d+2)\,p}\left(d\,p+2\pm\sqrt{d\,(p-1)\left(2\,d-(d-2)\,p\right)}\right) \tag{5.13}$$

while the parameters δ and ζ are defined by

$$\delta := 1 + \frac{(m-1)\,p^2}{4\,(p-2)}\,,$$
$$\zeta := \frac{(d+2)^2\,p^2\,m^2 - 2\,p\,(d+2)\,(d\,p+2)\,m + d^2\left(5\,p^2 - 12\,p + 8\right) + 4\,d\,(3-2\,p)\,p + 4}{(1-m)\,(d+2)^2\,p^2}\,.$$

Let $s_\star := 1/(p-2)$ as in (5.9) and consider the inverse function $\varphi_{m,p}^{-1} : \mathbb{R}^+ \to [0, s_\star)$ and $\psi_{m,p}(s) := s - \varphi_{m,p}^{-1}(s)$. Exactly as in the case $m = 1$, we have the improved entropy – entropy production inequality

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \ge d\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\,\varphi_{m,p}\left(\frac{\mathcal{E}_p[F]}{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\right) \quad \forall\,F \in \mathrm{H}^1(\mathbb{S}^d)\,,$$

which provides us with the following stability estimate.

Proposition 76. *With above notations, $d \ge 1$, $p \in (2, 2^*)$ and $m \in \mathscr{A}_p$, we have*

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} - d\,\mathcal{E}_p[F] \ge d\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\,\psi_{m,p}\left(\frac{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{d\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\right) \quad \forall\,F \in \mathrm{H}^1(\mathbb{S}^d)\,.$$

The function $\varphi_{m,p}$ can be expressed in terms of the *incomplete* Γ *function*, while $\psi_{m,p}$ is known only implicitly.

5.3.3 Comparison with other estimates

Let us assume that $p \in (2, 2^*)$. In [144], R. Frank proves the existence of a positive constant $\mathsf{c}_\star(d,p)$ such that

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} - d\,\mathcal{E}_p[F] \ge \mathsf{c}_\star(d,p)\,\frac{\left(\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \left\|F - \overline{F}\right\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right)^2}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \frac{d}{p-2}\,\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} \quad \forall\,F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$$

where $\overline{F} := \int_{\mathbb{S}^d} F\,d\mu_d$, which in particular implies the existence of a positive constant $\mathsf{c}(d,p)$ such that

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} - d\,\mathcal{E}_p[F] \ge \mathsf{c}(d,p)\,\frac{\|\nabla F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \frac{d}{p-2}\,\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} \quad \forall\,F \in \mathrm{H}^1(\mathbb{S}^d, d\mu) \tag{5.14}$$

for all $p \in (2, 2^\star)$. The value of the constant $\mathsf{c}_\star(d,p)$ found in [144] is unknown as it follows from a compactness argument, in the spirit of [47], but the exponent 4 in the r.h.s. of (5.14) is optimal. With the test functions $F_\varepsilon(x) = 1 + \varepsilon\,x\cdot\nu$ for some arbitrary $\nu \in \mathbb{S}^d$, we can indeed check that

$$\lim_{\varepsilon\to0}\frac{1}{\varepsilon^4}\left(\|\nabla F_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{S}^d)} - d\,\mathcal{E}_p[F_\varepsilon]\right) = \frac{(d+p)\,(p-1)}{2\,d\,(d+3)}\,,$$

which gives the upper bounds

$$\mathsf{c}(d,p) \le \frac{(p-1)\,(d+p)}{2\,(p-2)\,(d+3)} \quad \text{and} \quad \mathsf{c}_\star(d,p) \le \frac{d^2}{(d+1)^2}\,\frac{(p-1)\,(d+p)}{2\,(p-2)\,(d+3)}\,.$$

Let us notice that $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \ge d\, \|F-\overline{F}\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ by the Poincaré inequality, so that we have

$$\left(\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|F-\overline{F}\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right)^2 \ge \|\nabla F\|^4_{\mathrm{L}^2(\mathbb{S}^d)} \ge \frac{d^2}{(d+1)^2}\left(\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|F-\overline{F}\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right)^2$$

and, at least if $\mathsf{c}_\star(d,p)$ and $\mathsf{c}(d,p)$ are the optimal constants

$$\frac{d^2}{(d+1)^2}\,\mathsf{c}(d,p) \le \mathsf{c}_\star(d,p) \le \mathsf{c}(d,p)\,.$$

We claim that the *carré du champ* method provides us with a constructive estimate of $\mathsf{c}(d,p)$. Let

$$\phi_c(s) := \frac{d}{2\,(1-c)}\left(2\,c\,s - s_\star + \sqrt{s_\star^2 + 4\,c\,s\,(s-s_\star)}\right).$$

Corollary 77. *Let $p \in (2,2^*)$. With the notations of Proposition 76, Inequality (5.14) holds with*

$$\mathsf{c} = \sup\left\{c>0\,:\,\exists\, m \in \mathscr{A}_p \text{ such that } \phi_c(s) \le \varphi_{m,p}(s)\;\forall\, s \in [0,s_\star)\right\}.$$

Proof. With no loss of generality, let us assume that $\|F\|_{\mathrm{L}^p(\mathbb{S}^d)} = 1$ and define

$$\mathsf{i} = \|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \quad \text{and} \quad \mathsf{e} := \frac{1-\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{p-2}$$

so that $\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = 1-(p-2)\,\mathsf{e}$. With $c = \mathsf{c}(d,p)$, we can rewrite (5.14) as

$$\mathsf{i} - d\,\mathsf{e} \ge \frac{c\,\mathsf{i}^2}{\mathsf{i} + \frac{d}{p-2} - d\,\mathsf{e}}$$

which amounts to

$$\mathsf{i} - d\,\mathsf{e} \ge \phi_c(\mathsf{e})\,.$$

Since we know that $\mathsf{i} - d\,\mathsf{e} \ge \varphi_{m,p}(\mathsf{e})$, the conclusion follows for the largest possible $c>0$ such that $\varphi_{m,p} \ge \phi_c$. □

5.4 Global stability results

We collect the statements of Theorems 70 and 73 in a single result. The whole Section is devoted to its proof.

Theorem 78. *Let $d \ge 1$ and $p \in (1,2^*)$. For any $F \in \mathrm{H}^1(\mathbb{S}^d, d\mu)$, we have*

$$\int_{\mathbb{S}^d} |\nabla F|^2\, d\mu_d - d\,\mathcal{E}_p[F] \ge \mathscr{S}_{d,p}\left(\frac{\|\nabla \Pi_1 F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} + \|\nabla(\mathrm{Id}-\Pi_1)\,F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \tag{5.15}$$

for some explicit stability constant $\mathscr{S}_{d,p} > 0$.

The value of $\mathscr{S}_{d,p}$ is elementary and explicit but its expression is lengthy. We explain in the proof how to compute it with all necessary details to obtain a numerical expression of $\mathscr{S}_{d,p}$ for given p and d, if needed.

Proof. By homogeneity of (5.15), we can assume that $\|F\|_{\mathrm{L}^2(\mathbb{S}^d)} = 1$ without loss of generality. For clarity, we subdivide the proof in various steps. Let us start with the case $p > 2$.

• *An estimate based on the carré du champ method.* If $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} / \|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} \ge \vartheta_0 > 0$, we know by the convexity of $\psi_{m,p}$ that

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} - d\,\mathcal{E}_p[F] \ge d\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\,\psi_{m,p}\left(\frac{1}{d}\,\frac{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\right) \ge \frac{d}{\vartheta_0}\,\psi_{m,p}\left(\frac{\vartheta_0}{d}\right)\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\,. \tag{5.16}$$

In that case, we conclude from $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = \|\nabla \Pi_1 F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|\nabla(\mathrm{Id}-\Pi_1)\,F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$ and

$$\|\nabla \Pi_1 F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \ge \frac{\|\nabla \Pi_1 F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}\,.$$

Let us assume now that $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} < \vartheta_0\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}$. By taking into account (GN), we obtain

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} < \vartheta_0\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)} \le \vartheta_0\left(\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \frac{p-2}{d}\,\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right).$$

Using $\|F\|_{\mathrm{L}^2(\mathbb{S}^d)} = 1$, and

$$\vartheta = \frac{d\,\vartheta_0}{d-(p-2)\,\vartheta_0} > 0$$

then we know that

$$\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} < \vartheta\,. \tag{5.17}$$

Notice that so far the parameter ϑ (or equivalently ϑ_0) still has to be chosen.

• *An estimate of the average.* Let us estimate $\Pi_0 F := \int_{\mathbb{S}^d} F\,d\mu_d$. By the Poincaré inequality, we have

$$1 = \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = \left(\int_{\mathbb{S}^d} F\,d\mu_d\right)^2 + \|(\mathrm{Id}-\Pi_0)\,F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \le \left(\int_{\mathbb{S}^d} F\,d\mu_d\right)^2 + \frac{\vartheta}{d}\,,$$

and on the other hand we know that $\left(\int_{\mathbb{S}^d} F\,d\mu_d\right)^2 \le \|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = 1$ by the Cauchy-Schwarz inequality, so that

$$\frac{d-\vartheta}{d} < \left(\int_{\mathbb{S}^d} F\,d\mu_d\right)^2 \le 1\,. \tag{5.18}$$

We assume in the sequel that

$$\vartheta < d\,. \tag{5.19}$$

• *Partial decomposition on spherical harmonics.* With no loss of generality, let us write

$$F = \mathscr{M}\,(1+\varepsilon\,\mathscr{Y}+\eta\,G) \tag{5.20}$$

such that $\mathscr{M} = \Pi_0 F$ and $\Pi_1 F = \varepsilon\,\mathscr{Y}$ where $\mathscr{Y}(x) = \sqrt{\frac{d+1}{d}}\,x\cdot\nu$ for some given $\nu\in\mathbb{S}^d$. Here the functions $\mathscr{Y}$ and G are normalized so that $\|\nabla\mathscr{Y}\|_{\mathrm{L}^2(\mathbb{S}^d)} = \|\nabla G\|_{\mathrm{L}^2(\mathbb{S}^d)} = 1$ and

$$\mathscr{M}^{-2}\,\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = \varepsilon^2+\eta^2 \quad\text{and}\quad \mathscr{M}^{-2}\,\|F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = 1+\frac{1}{d}\,\varepsilon^2+\eta^2\,\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\,.$$

We observe that $\Pi_0(F-\mathscr{M}) = 0$. Using (GN) and the Poincaré inequality, we have

$$\|F-\mathscr{M}\|^2_{\mathrm{L}^p(\mathbb{S}^d)} \le \|F-\mathscr{M}\|^2_{\mathrm{L}^2(\mathbb{S}^d)} + \frac{p-2}{d}\,\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \le \frac{p-1}{d}\,\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\,.$$

Similarly, by (5.5)

$$\frac{d\,p}{2\,(d+p)}\,\|\nabla G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\left(1-\frac{2\,d-p\,(d-2)}{2\,(d+p)}\right)\|\nabla G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\ge d\,\mathcal{E}_p[G]$$

and the improved Poincaré inequality (5.5) written with $p=1$ and $k=1$

$$\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\le\frac{1}{2\,(d+1)}\,\|\nabla G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\frac{1}{2\,(d+1)}\,,$$

we have

$$\|G\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\le\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}+\frac{p\,(p-2)}{2\,(d+p)}\,\|\nabla G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\le C_{p,d}$$

using $\|\nabla G\|_{\mathrm{L}^2(\mathbb{S}^d)}=1$, with $C_{p,d}:=\frac{1}{2\,(d+1)}+\frac{p\,(p-2)}{2\,(d+p)}$. By the Cauchy-Schwarz inequality, we also have

$$\|G\|_{\mathrm{L}^1(\mathbb{R}^d)}\le\frac{1}{\sqrt{2\,(d+1)}}\,,$$

We recall that the eigenvalues of $-\Delta$ on $\mathbb{S}^d$ are $\lambda_k=k\,(k+d-1)$ with $k\in\mathbb{N}$. In preparation for a detailed Taylor expansion as in [144], let us consider the function

$$\mathscr{Y}(x):=\sqrt{\frac{d+1}{d}}\,x\cdot\nu\,,$$

which is such that $-\Delta\mathscr{Y}=\lambda_1\,\mathscr{Y}$ with $\lambda_1=d$ and

$$\|\nabla\mathscr{Y}\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=1\,,\quad\|\mathscr{Y}\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\frac{1}{d}\,,$$
$$\|\mathscr{Y}\|^4_{\mathrm{L}^4(\mathbb{S}^d)}=\frac{3\,(d+1)}{(d+3)\,d^2}\,,\quad\|\mathscr{Y}\|^6_{\mathrm{L}^6(\mathbb{S}^d)}=\frac{15\,(d+1)^2}{(d+3)\,(d+5)\,d^2}\,.$$

The function $\mathscr{Y}_2:=\mathscr{Y}^2-\frac{1}{d}$ is such that $-\Delta\mathscr{Y}_2=\lambda_2\,\mathscr{Y}_2$ with $\lambda_2=2\,(d+1)$ and

$$\|\mathscr{Y}_2\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\frac{2}{d\,(d+3)}\,,\quad\|\nabla\mathscr{Y}_2\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\frac{4\,(d+1)}{d\,(d+3)}\,.$$

The function $\mathscr{Y}_3:=\mathscr{Y}^3-\frac{3\,(d+1)}{d\,(d+3)}\,\mathscr{Y}$ is such that $-\Delta\mathscr{Y}_3=\lambda_3\,\mathscr{Y}_3$ with $\lambda_3=3\,(d+2)$ and

$$\|\mathscr{Y}_3\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\frac{6\,(d+1)^2}{(d+5)\,(d+3)^2\,d^2}\,,\quad\|\nabla\mathscr{Y}_3\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=\frac{18\,(d+2)\,(d+1)^2}{(d+5)\,(d+3)^2\,d^2}\,.$$

As a consequence of (5.20), we know that $\Pi_0G=\Pi_1G=0$ and $\|\nabla G\|_{\mathrm{L}^2(\mathbb{S}^d)}=1$. Let

$$g_2:=\frac{\int_{\mathbb{S}^d}\nabla\mathscr{Y}_2\cdot\nabla G\,d\mu_d}{\|\nabla\mathscr{Y}_2\|_{\mathrm{L}^2(\mathbb{S}^d)}}\quad\text{and}\quad g_3:=\frac{\int_{\mathbb{S}^d}\nabla\mathscr{Y}_3\cdot\nabla G\,d\mu_d}{\|\nabla\mathscr{Y}_3\|_{\mathrm{L}^2(\mathbb{S}^d)}}\,.$$

With $k=1,\,2$, using $-\Delta\mathscr{Y}_k=\lambda_k\,\mathscr{Y}_k$ with $\lambda_k=\|\nabla\mathscr{Y}_k\|^2_{\mathrm{L}^2(\mathbb{S}^d)}/\|\mathscr{Y}_k\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$, we compute

$$\int_{\mathbb{S}^d}\mathscr{Y}^k\,G\,d\mu_d=\int_{\mathbb{S}^d}\mathscr{Y}_k\,G\,d\mu_d=\frac{\|\mathscr{Y}_k\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla\mathscr{Y}_k\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\int_{\mathbb{S}^d}\nabla\mathscr{Y}_k\cdot\nabla G\,d\mu_d=g_k\,\frac{\|\mathscr{Y}_k\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla\mathscr{Y}_k\|_{\mathrm{L}^2(\mathbb{S}^d)}}$$

and obtain

$$\int_{\mathbb{S}^d} \mathscr{Y}^2\, G\, d\mu_d = \int_{\mathbb{S}^d} \mathscr{Y}_2\, G\, d\mu_d = \frac{g_2}{\sqrt{d\,(d+1)\,(d+3)}}\,,$$

$$\int_{\mathbb{S}^d} \mathscr{Y}^3\, G\, d\mu_d = \int_{\mathbb{S}^d} \mathscr{Y}_3\, G\, d\mu_d = c_3\, g_3 \quad \text{with} \quad c_3 := \frac{d+1}{d\,(d+3)} \sqrt{\frac{2}{(d+2)\,(d+5)}}\,.$$

• *Taylor expansions (1).* Let us start with elementary estimates of $\|1+\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}$. If $2 \le p < 3$ and $|s| < 1$, we have

$$\frac{1}{2}\left((1+s)^p + (1-s)^p\right) \le 1 + \frac{p}{2}\,(p-1)\,s^2 \left(1 + \frac{1}{12}\,(p-2)\,(p-3)\,s^2\right).$$

because all other terms in the series expansion of the l.h.s. around $s = 0$ correspond to even powers of s and appear with non-positive coefficients. If either $1 \le p < 2$ or $p > 3$ and $|s| < 1/2$, let

$$f_p(s) := \frac{1}{2}\left((1+s)^p + (1-s)^p\right) - \left(1 + \frac{p}{2}\,(p-1)\,s^2\right)$$

and notice that $f_p''(s) = \frac{p}{2}\,(p-1)\left((1+s)^{p-2} + (1-s)^{p-2} - 2\right) \ge 0$ by convexity of the function $y \mapsto y^{p-2}$ so that $c_p^{(+)}$ defined as the maximum of $s \mapsto f_p(s)/s^6$ on $[-1/2, 1/2] \ni s$ is finite and we have

$$\frac{1}{2}\left((1+s)^p + (1-s)^p\right) \le 1 + \frac{p}{2}\,(p-1)\,s^2 \left(1 + \frac{1}{12}\,(p-2)\,(p-3)\,s^2\right) + c_p^{(+)}\,s^6\,. \tag{5.21}$$

We adapt the convention that $c_p^{(+)} = 0$ if $p \in [2,3)$. Using the fact that $\mathscr{Y}(-x) = -\mathscr{Y}(x)$,

$$\|1+\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}^p = \frac{1}{2}\left(\|1+\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}^p + \|1-\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}^p\right).$$

For any $\varepsilon \in (0, 1/2)$ we use (5.21) to write

$$\|1+\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}^p - \left(1 + \frac{p}{2}\,(p-1)\left(\|\mathscr{Y}\|_{\mathrm{L}^2(\mathbb{S}^d)}^2 + \frac{1}{12}\,(p-2)\,(p-3)\,\|\mathscr{Y}\|_{\mathrm{L}^4(\mathbb{S}^d)}^4\,\varepsilon^2\right)\varepsilon^2\right) \le c_p^{(+)}\,\|\mathscr{Y}\|_{\mathrm{L}^6(\mathbb{S}^d)}^6\,\varepsilon^6\,.$$

For similar reasons, one can prove that there is another constant $c_p^{(-)}$ which provides us with a lower bound $c_p^{(-)}\,\|\mathscr{Y}\|_{\mathrm{L}^6(\mathbb{S}^d)}^6\,\varepsilon^6$. Altogether, this amounts to

$$c_{p,d}^{(-)}\,\varepsilon^6 \le \|1+\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}^p - \left(1 + a_{p,d}\,\varepsilon^2 + b_{p,d}\,\varepsilon^4\right) \le c_{p,d}^{(+)}\,\varepsilon^6 \tag{5.22}$$

with

$$a_{p,d} := \frac{p\,(p-1)}{2\,d}\,,\quad b_{p,d} := \frac{1}{4}\,(p-2)\,(p-3)\,\frac{d+1}{d\,(d+3)}\,a_{p,d}\,,\quad c_{p,d}^{(\pm)} := \frac{15\,(d+1)^2}{(d+3)\,(d+5)\,d^2}\,c_p^{(\pm)}\,.$$

Estimate (5.22) is valid under the condition that $\varepsilon < 1/2$. We shall therefore request that

$$\vartheta < \frac{1}{4}\,, \tag{5.23}$$

which is an obvious sufficient condition according to (5.17). Now we draw two consequences of equation (5.22). First, let us give an upper estimate of $\|1+\varepsilon\,\mathscr{Y}\|_{\mathrm{L}^p(\mathbb{S}^d)}^2$. Using

$$(1+s)^{\frac{2}{p}} \le 1 + 2\,\frac{s}{p} - (p-2)\,\frac{s^2}{p^2} + \frac{2}{3}\,(p-1)\,(p-2)\,\frac{s^3}{p^3}\,,$$

we obtain

$$\|1+\varepsilon\,\mathscr{Y}\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\le 1+\frac{2}{p}\,a_{p,d}\,\varepsilon^2+\frac{1}{p^2}\left(2\,p\,b_{p,d}-(p-2)\,a_{p,d}^2\right)\varepsilon^4+r^{(+)}\,\varepsilon^6 \tag{5.24}$$

where the remainder term $r^{(+)}$ is explicitly estimated by

$$\begin{aligned}96\,p^3\,r^{(+)}=64\,a_{p,d}^3\left(p^2-3\,p+2\right)+48\,a_{p,d}^2\left(p^2-3\,p+2\right)(2\,b_{p,d}+c_{p,d})\\ +12\,a_{p,d}\,(p-2)\,(2\,b_{p,d}+c_{p,d})\,(2\,b_{p,d}\,(p-1)+c_{p,d}\,(p-1)-8\,p)\\ +8\,b_{p,d}^3\left(p^2-3\,p+2\right)+12\,b_{p,d}^2\,(p-2)\,(c_{p,d}\,(p-1)-4\,p)\\ +6\,b_{p,d}\,c_{p,d}\,(p-2)\,(c_{p,d}\,(p-1)-8\,p)\\ +c_{p,d}\left(c_{p,d}^2\left(p^2-3\,p+2\right)-12\,c_{p,d}\,(p-2)\,p+192\,p^2\right).\end{aligned}$$

To do this estimate, we simply write that $\varepsilon^\alpha\le 2^{6-\alpha}\,\varepsilon^6$ for any $\alpha>6$ using the (non-optimal) bound $\varepsilon^2<1/2$. Similarly, using

$$\begin{aligned}(1+s)^{\frac{2}{p}-1}\le 1-(p-2)\,\frac{s}{p}+(p-1)\,(p-2)\,\frac{s^2}{p^2}-\frac{1}{3}\,(p-1)\,(p-2)\,(3\,p-2)\,\frac{s^3}{p^3}\\ +\frac{1}{6}\,(p-1)\,(p-2)\,(3\,p-2)\,(2\,p-1)\,\frac{s^4}{p^4},\end{aligned}$$

we obtain

$$\|1+\varepsilon\,\mathscr{Y}\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\le 1+\frac{p-2}{p}\,a_{p,d}\,\varepsilon^2-\frac{p-2}{p^2}\left(p\,b_{p,d}-(p-1)\,a_{p,d}^2\right)\varepsilon^4+r^{(-)}\,\varepsilon^6 \tag{5.25}$$

where the remainder term $r^{(-)}$ has also an explicit expression in terms of $a_{p,d}$, $b_{p,d}$ and $c_{p,d}^{(-)}$, which is not given here.

• *Taylor expansions (2).* With $u\ge0$, $u+r\ge0$ and $p>2$, we claim that

$$(u+r)^p\le u^p+p\,u^{p-1}\,r+\frac{p}{2}\,(p-1)\,u^{p-2}\,r^2+\sum_{2<k<p}C^p_k\,u^{p-k}|r|^k+K_p\,|r|^p$$

for some constant $K_p>0$, where the coefficients

$$C^p_k:=\frac{\Gamma(p+1)}{\Gamma(k+1)\,\Gamma(p-k+1)}$$

are the binomial coefficients if p is an integer. It is proved in [119] that $K_p=1$ if $p\in(2,4]\cup\{6\}$. The proof is similar to the above analysis and left to the reader. Let us integrate this inequality and raise both sides to the power $2/p$ to get

$$\|u+r\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\le\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\,(1+s)^{\frac{2}{p}}$$

with

$$\begin{aligned}s=\frac{1}{\|u\|^p_{\mathrm{L}^p(\mathbb{S}^d)}}\Bigg(p\int_{\mathbb{S}^d}u^{p-1}\,r\,d\mu_d+\frac{p}{2}\,(p-1)\int_{\mathbb{S}^d}u^{p-2}\,r^2\,d\mu_d\\ +\sum_{2<k<p}C^p_k\int_{\mathbb{S}^d}u^{p-k}|r|^k\,d\mu_d+K_p\int_{\mathbb{S}^d}|r|^p\,d\mu_d\Bigg).\end{aligned}$$

By assumption $2/p<1$ so that we may use the identity $(1+s)^{2/p}\le 1+2\,s/p$ for any $s\ge-1$. Notice that we can assume that $u+r\geq 0$ and deduce from (1) that $s\geq -1$. As a consequence, we have

$$\|u+r\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\leq \|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)} + \frac{2}{p}\,\|u\|^{2-p}_{\mathrm{L}^p(\mathbb{S}^d)}\left(p\int_{\mathbb{S}^d}u^{p-1}\,r\,d\mu_d+\frac{p}{2}\,(p-1)\int_{\mathbb{S}^d}u^{p-2}\,r^2\,d\mu_d + \sum_{2<k<p}C^p_k\int_{\mathbb{S}^d}u^{p-k}|r|^k\,d\mu_d+K_p\int_{\mathbb{S}^d}|r|^p\,d\mu_d\right)$$

We apply these computations to $u=1+\varepsilon\mathscr{Y}$ and $r=\eta\,G$.

$$\mathscr{M}^{-2}\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}-\|1+\varepsilon\mathscr{Y}\|^2_{\mathrm{L}^p(\mathbb{S}^d)} \leq \frac{2}{p}\,\|1+\varepsilon\mathscr{Y}\|^{2-p}_{\mathrm{L}^p(\mathbb{S}^d)}\,\eta\left(p\int_{\mathbb{S}^d}(1+\varepsilon\mathscr{Y})^{p-1}\,G\,d\mu_d + \frac{p}{2}\,(p-1)\,\eta\int_{\mathbb{S}^d}(1+\varepsilon\mathscr{Y})^{p-2}\,|G|^2\,d\mu_d + \sum_{2<k<p}C^p_k\,\eta^k\int_{\mathbb{S}^d}(1+\varepsilon\mathscr{Y})^{p-k}\,|G|^k\,d\mu_d+K_p\,\eta^p\int_{\mathbb{S}^d}|G|^p\,d\mu_d\right).$$

Let us detail the expansion of each of the terms involving G in the right-hand side of this estimates. For any $s\in(-1/2,1/2)$, using the expansion

$$(1+s)^{p-1}\leq 1+(p-1)\,s+\frac{1}{2}\,(p-1)\,(p-2)\,s^2+\frac{1}{6}\,(p-1)\,(p-2)\,(p-3)\,s^3+R_p\,s^4$$

for some constant $R_p>0$ applied with $s=1+\varepsilon\mathscr{Y}$, we obtain

$$\eta\int_{\mathbb{S}^d}(1+\varepsilon\mathscr{Y})^{p-1}\,G\,d\mu_d\leq\frac{1}{2}\,(p-1)\,(p-2)\,\frac{g_2}{\sqrt{d\,(d+1)\,(d+3)}}\,\eta\,\varepsilon^2 + \frac{1}{6}\,(p-1)\,(p-2)\,(p-3)\,c_3\,g_3\,\eta\,\varepsilon^3+\frac{R_p\,\eta\,\varepsilon^4}{\sqrt{2\,(d+1)}}\,.$$

The other terms admit simpler expansions:

$$\eta^2\int_{\mathbb{S}^d}(1+\varepsilon\mathscr{Y})^{p-2}\,|G|^2\,d\mu_d\leq\eta^2\,(1+\varepsilon)^{p-2}\,\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \leq\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\,\eta^2+\frac{1}{2\,(d+1)}\,\eta^2\left((1+\varepsilon)^{p-2}-1\right)$$

and

$$\sum_{2<k<p}C^p_k\,\eta^k\int_{\mathbb{S}^d}(1+\varepsilon\mathscr{Y})^{p-k}\,|G|^k\,d\mu_d+K_p\,\eta^p\int_{\mathbb{S}^d}|G|^p\,d\mu_d \leq\sum_{2<k<p}C^p_k\,\eta^k\,(1+\varepsilon)^{p-k}\,\|G\|^k_{\mathrm{L}^p(\mathbb{S}^d)}+K_p\,\eta^p\,\|G\|^p_{\mathrm{L}^p(\mathbb{S}^d)} \leq\sum_{2<k<p}C^p_k\,\eta^k\,(1+\varepsilon)^{p-k}\,C^{k/p}_{p,d}+K_p\,\eta^p\,C_{p,d}\,.$$

Collecting (5.24), (5.25), and the above estimates, we arrive at

$$\begin{aligned}
&\mathscr{M}^{-2}\,\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\\
&\le 1+\frac{2}{p}\,a_{p,d}\,\varepsilon^2+\frac{1}{p^2}\left(2\,p\,b_{p,d}-(p-2)\,a_{p,d}^2\right)\varepsilon^4+r^{(+)}\,\varepsilon^6\\
&\quad+\left(1+\frac{p-2}{p}\,a_{p,d}\,\varepsilon^2-\frac{p-2}{p^2}\left(p\,b_{p,d}-(p-1)\,a_{p,d}^2\right)\varepsilon^4+r^{(-)}\,\varepsilon^6\right)\\
&\qquad\cdot\Bigg[(p-1)\,(p-2)\,\frac{g_2}{\sqrt{d\,(d+1)\,(d+3)}}\,\eta\,\varepsilon^2+\frac{1}{3}\,(p-1)\,(p-2)\,(p-3)\,c_3\,g_3\,\eta\,\varepsilon^3\\
&\qquad\quad+\frac{2\,R_p\,\eta\,\varepsilon^4}{\sqrt{2\,(d+1)}}+(p-1)\left(\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\,\eta^2+\frac{1}{2\,(d+1)}\,\eta^2\left((1+\varepsilon)^{p-2}-1\right)\right)\\
&\qquad\qquad+\frac{2}{p}\left(\sum_{2<k<p}C^p_k\,\eta^k\,(1+\varepsilon)^{p-k}\,C^{k/p}_{p,d}+K_p\,\eta^p\,C_{p,d}\right)\Bigg].
\end{aligned}$$

Using $|g_2|<1$, $|g_3|<1$, and $2\,(d+1)\,\|G\|^2_{\mathrm{L}^2(\mathbb{S}^d)}<1$, this gives rise to an explicit although lengthy expression for a positive constant $\mathcal{R}_{p,d}$ such that, assuming (5.17), *i.e.*, $\varepsilon^2+\eta^2<\vartheta$, we have

$$\mathscr{M}^{-2}\left(\int_{\mathbb{S}^d}|\nabla F|^2\,d\mu_d-d\,\mathcal{E}_p[F]\right)\ge A\,\varepsilon^4-B\,\varepsilon^2\,\eta+C\,\eta^2-\mathcal{R}_{p,d}\left(\vartheta^p+\vartheta^{5/2}\right)$$

with $A:=\frac{(p-1)\,(d+p)}{2\,d\,(d+3)}$, $B:=\frac{d\,(p-1)}{\sqrt{d\,(d+1)\,(d+3)}}$ and $C:=\frac{d+2}{2\,(d+1)}$. The discriminant

$$B^2-4\,A\,C=-\frac{1}{d\,(d+3)}\,(p-1)\left(2\,d-p\,(d-2)\right)$$

is negative if $p\in(1,2^*)$, so that we can write

$$A\,s^2-B\,s+C=(A-\lambda)\,s^2-B\,s+(C-\lambda)+\lambda\left(s^2+1\right)\ge\lambda\left(s^2+1\right)$$

where

$$\lambda:=\frac{1}{2}\left(A+C+\sqrt{(A-C)^2+B^2}\right)$$

is given by the condition that $B^2-4\,(A-\lambda)\,(C-\lambda)=0$. Altogether, we obtain

$$\mathscr{M}^{-2}\left(\int_{\mathbb{S}^d}|\nabla F|^2\,d\mu_d-d\,\mathcal{E}_p[F]\right)\ge\lambda\left(\varepsilon^4+\eta^2\right)-\mathcal{R}_{p,d}\left(\vartheta^p+\vartheta^{5/2}\right).$$

• *Conclusion if* $p>2$. We choose $\vartheta>0$ such that (5.19) and (5.23) are fulfilled and we further assume that it is taken small enough so that

$$\mathcal{R}_{p,d}\left(\vartheta^p+\vartheta^{5/2}\right)\le\frac{\lambda}{2}\left(\vartheta^2+\vartheta\right).$$

With

$$\vartheta_{p,d}:=\left\{\vartheta>0\,:\,\mathcal{R}_{p,d}\left(\vartheta^p+\vartheta^{5/2}\right)=\frac{\lambda}{2}\left(\vartheta^2+\vartheta\right)\right\}$$

this last condition amounts to

$$\vartheta\le\vartheta_{p,d}\,.$$

It is elementary to check that under the condition $\varepsilon^2+\eta^2<\vartheta$, we have

$$\varepsilon^4+\eta^2\le\frac{1}{2+\vartheta}\left(\frac{\varepsilon^4}{\varepsilon^2+\eta^2+1}+\eta^2\right).$$

For any F such that $\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\le\vartheta$, we obtain

$$\int_{\mathbb{S}^d}|\nabla F|^2\,d\mu_d-d\,\mathcal{E}_p[F]\ge\frac{\lambda\,\mathscr{M}^2}{2\,(2+\vartheta)}\left(\frac{\|\nabla\Pi_1 F\|^4_{\mathrm{L}^2(\mathbb{S}^d)}}{\|\nabla F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}+\|F\|^2_{\mathrm{L}^p(\mathbb{S}^d)}}+\|\nabla(\mathrm{Id}-\Pi_1)\,F\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right).$$

Using (5.16) and (5.18), this completes the proof of Theorem 78 if $p>2$ with

$$\vartheta=\min\left\{\frac12\,d,\,\frac14,\,\vartheta_{p,d}\right\}$$

$$\text{and}\quad\mathscr{S}_{d,p}=\min\left\{\frac{d+(p-2)\,\vartheta}{\vartheta}\,\psi_{m,p}\left(\frac{\vartheta}{d+(p-2)\,\vartheta}\right),\,\frac{\lambda\,(d-\vartheta)}{2\,d\,(2+\vartheta)}\right\}.$$

• *The case* $p\le2$. The strategy is the same, with some simplifications, so we only sketch the proof and emphasize only the changes compared to the case $p>2$. Let us notice that

$$(1+s)^p\le1+p\,s+\frac{p}{2}\,(p-1)\,s^2\quad\text{if}\quad1\le p<2$$

and $(1+s)^2\log\big((1+s)^2\big)\le2\,s+2\,s^2+\frac23\,s^3$ in the limit case $p=2$. The estimates involving $1+\varepsilon\,\mathscr{Y}$ are therefore essentially the same if we assume $\varepsilon<1/2$, while the computation of $\|u+r\|^2_{\mathrm{L}^p(\mathbb{S}^d)}$ is in fact simpler, when applied to $u=1+\varepsilon\,\mathscr{Y}$ and $r=\eta\,G$. The estimate on the average is simplified because $\|F\|_{\mathrm{L}^p(\mathbb{S}^d)}\le\|F\|_{\mathrm{L}^2(\mathbb{S}^d)}$ by Hölder's inequality, since $d\mu$ is a probability measure on $\mathbb{S}^d$. Spectral estimates are exactly the same and the Taylor expansions present no additional difficulty, as we can use (GN) for some exponent $q\in(2,2^*)$ to control the remainder terms if $p=2$, so that $(1+s)^2\log\big((1+s)^2\big)\le2\,s+2\,s^2+\kappa_q\,s^q$ for some $\kappa_q>0$. The conclusion is the same as for $p>2$ except that we have to replace $\psi_{m,p}$ by ψ defined as in Proposition 75. □

5.5 Appendix: Improved Gaussian inequalities, hypercontractivity and stability

Whether the results of Theorems 68, 71 and 74 can be extended to the Euclidean case with the Gaussian measure is a very natural question. Spherical harmonics can indeed be replaced by Hermite polynomials and there is a clear correspondence for spectral estimates. The answer is yes for a whole family of interpolation inequalities, but it is no for the logarithmic Sobolev inequality, which is an endpoint of the family.

Let us consider the *normalized Gaussian measure* on $\mathbb{R}^d$ defined by

$$d\sigma(x)=(2\,\pi)^{-\frac d2}\,e^{-\frac12|x|^2}\,dx\,.$$

For any $p\in[1,2)$, W. Beckner in [35] established the family of interpolation inequalities

$$\frac{\|f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}-\|f\|^2_{\mathrm{L}^p(\mathbb{R}^d,d\gamma)}}{2-p}\le\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}\quad\forall\,f\in\mathrm{H}^1(\mathbb{R}^d,d\sigma)\,.\tag{5.26}$$

With $p=1$, Inequality (5.26) is the *Gaussian Poincaré inequality* while one recovers the Gaussian logarithmic

Sobolev inequality of [155] in the limit as $p \to 2$. For any $p \in [1,2)$, the inequality is optimal: using $f_\varepsilon := 1 + \varepsilon\,\varphi$ as a test function, where φ is such that $\int_{\mathbb{R}^d} \varphi\,d\gamma = 0$, we recover the *Gaussian Poincaré inequality* with optimal constant in the limit as $\varepsilon \to 0$, so that the constant in (5.26) cannot be improved. Based on [187, 14], the improved version of the inequality

$$\frac{\|f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^d,d\gamma)}}{2-p} \le \frac{p}{2}\,\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^d,d\sigma) \tag{5.27}$$

holds under the additional condition

$$\int_{\mathbb{R}^d} x\,f(x)\,d\gamma = 0\,. \tag{5.28}$$

Let us give a short proof of (5.27). Assume that $f = \sum_{k\in\mathbb{N}} f_k$ is a decomposition on Hermite functions such that $\mathcal{L} f_k = k\,f_k$ where $\mathcal{L} = \Delta - x\cdot\nabla$ is the Ornstein-Uhlenbeck operator, and let $a_k := \|f_k\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}$ for any $k \in \mathbb{N}$, so that

$$\|f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} = \sum_{k\in\mathbb{N}} a_k \quad \text{and} \quad \|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} = \sum_{k\in\mathbb{N}} k\,a_k\,.$$

Let us consider the solution of

$$\frac{\partial u}{\partial t} = \mathcal{L} u \tag{5.29}$$

with initial datum $u(t=0,\cdot) = f$ and notice that

$$\|u(t,\cdot)\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} = \sum_{k\in\mathbb{N}} a_k\,e^{-2\,k\,t}\,.$$

Hence, if (5.28) holds, $a_1 = 0$ and

$$\begin{aligned} \|f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \|u(t,\cdot)\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} &= \sum_{k\ge 2} a_k\left(1 - e^{-2\,k\,t}\right) \\ &\le \frac{1}{2}\left(1 - e^{-4\,t}\right)\sum_{k\in\mathbb{N}} k\,a_k = \frac{1}{2}\left(1 - e^{-4\,t}\right)\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} \end{aligned} \tag{5.30}$$

because $k \mapsto \left(1 - e^{-2\,k\,t}\right)/k$ is monotone nonincreasing for any given $t \ge 0$. Next, we use Nelson's hypercontractivity estimate in [220, Theorem 3] to find $t_\star > 0$ such that

$$\|u(t_\star,\cdot)\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} \le \|f\|^2_{\mathrm{L}^p(\mathbb{R}^d,d\gamma)}\,.$$

As noted in [155], this estimate can be seen as a consequence of the *Gaussian logarithmic Sobolev inequality*

$$\int_{\mathbb{R}^d} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}}\right) d\gamma \le 2\int_{\mathbb{R}^d} |\nabla v|^2\,d\gamma \quad \forall\, v \in \mathrm{H}^1(\mathbb{R}^d,d\sigma)\,, \tag{5.31}$$

and the argument goes as follows. With $h(t) := \|u(t,\cdot)\|_{\mathrm{L}^{q(t)}(\mathbb{R}^d,d\gamma)}$ for some exponent q depending on t and u solving (5.29), we have

$$\frac{h'}{h} = \frac{q'}{q^2}\int_{\mathbb{R}^d} \frac{|u|^q}{h^q}\log\left(\frac{|u|^q}{h^q}\right) d\gamma - \frac{4}{h^q}\,\frac{q-1}{q^2}\int_{\mathbb{R}^d} \left|\nabla\left(|u|^{q/2}\right)\right|^2 d\gamma \le 0$$

by (5.31) applied to $v = |u|^{q/2}$, if $t \mapsto q(t)$ solves the ordinary differential equation

$$q' = 2\,(q-1)\,.$$

With $q(0) = p < 2$, we obtain $q(t) = 1+(p-1)\,e^{2t}$ and find that Nelson's time t_* is determined by the condition $q(t_*) = 2$ which means $e^{-2t_*} = p-1$. Replacing $t = t_*$ in (5.30) completes the proof of (5.27), which can be recast in the form of a stability result for (5.26).

Theorem 79. *Let $d \ge 1$ and $p \in [1,2)$. For any $f \in \mathrm{H}^1(\mathbb{R}^d, d\sigma)$ such that (5.28) holds,*

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^d,d\gamma)}\right) \ge \frac{2-p}{2}\,\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}\,.$$

As a byproduct of the proof, with $t = t_*$ in (5.30), we have the mode-by mode interpolation inequality

$$\frac{\|f\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^d,d\gamma)}}{2-p} \le \sum_{k\ge1} \frac{1-(p-1)^k}{k\,(2-p)}\,\|\nabla f_k\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^d, d\sigma)\,,$$

without imposing the condition (5.28), for any $p \in [1,2)$. For any $k \ge 1$,

$$\lim_{p\to2_-} \frac{1-(p-1)^k}{k\,(2-p)} = \lim_{p\to2_-} \frac{1-\big(1-(2-p)\big)^k}{k\,(2-p)} = 1\,,$$

so that no improvement should be expected by this method. This is very similar to the case of the critical exponent on the sphere of dimension $d \ge 3$. In this sense $p = 2$ is the critical case in presence of a Gaussian weight, as *all modes* are equally involved in the estimate of the constant. This is a limitation of the method which does not forbid a stability result for (5.31), to be established by other methods.

Let us conclude this appendix by some bibliographic comments for the literature on inequality (5.26), for the Gaussian measure. The analogue of Proposition 72 in the Gaussian case is known from [15]; also see [126, Section 2.5]). Assuming that not only condition (5.28) is satisfied, but also orthogonality conditions with all modes up to order $k_0 \ge 2$, then an improvement of the order of

$$\frac{1-(p-1)^{k_0}}{k_0\,(2-p)}$$

can be achieved for Inequality (5.26), which is the counterpart of Theorem 71 in the Gaussian case. This has been studied in [187] but we can refer to [14] for a more abstract setting and later papers, *e.g.*, to [259, 257] for results on compact manifolds and generalizations involving weights. For an overview of interpolation between Poincaré and logarithmic Sobolev inequalities from the point of view of Markov processes, and for some spectral considerations, we refer to [258, Chapter 6]. Notice that hypercontractivity appears as one of the main motivation of the founding paper [24] of the *carré du champ* method.

5.6 Appendix: Carré du champ method and improved inequalities

For sake of completeness, we collect various results of [121, 120, 124, 118] and draw some new consequences. Computations similar to those of Section 5.6.1 can be found in [48] for the study of rigidity results in elliptic equations. For nonlinear parabolic flows, also see [112, 113]. Other sections of this appendix collects results which are scattered in the literature, but additional details needed in Section 5.3 are given, for instance a sketch of the proof Proposition 83 or the computations in the case $p = 2$.

5.6.1 Algebraic preliminaries

Let us denote the *Hessian* by $\mathrm{H}v$ and define the *trace free Hessian* by

$$\mathrm{L}v := \mathrm{H}v - \frac{1}{d}\,(\Delta v)\,g_d\,.$$

We also consider the following trace free tensor

$$\mathrm{M}v := \frac{\nabla v \otimes \nabla v}{v} - \frac{1}{d}\,\frac{|\nabla v|^2}{v}\,g_d\,,$$

where $(\nabla v \otimes \nabla v)_{ij} := \partial_i v\,\partial_j v$ and $\|\nabla v \otimes \nabla v\|^2 = |\nabla v|^4 = (g_d^{ij}\,\partial_i v\,\partial_j v)^2$ using Einstein's convention. Using

$$\mathrm{L} : g_d = 0\,, \quad \mathrm{M} : g_d = 0$$

where $a : b$ denotes $a^{ij}\,b_{ij}$ and $\|a\|^2 := a : a$, and

$$\|\mathrm{L}v\|^2 = \|\mathrm{H}v\|^2 - \frac{1}{d}\,(\Delta v)^2\,,$$
$$\|\mathrm{M}v\|^2 = \left\|\frac{\nabla v \otimes \nabla v}{v}\right\|^2 - \frac{1}{d}\,\frac{|\nabla v|^4}{v^2} = \frac{d-1}{d}\,\frac{|\nabla v|^4}{v^2}\,,$$

we deduce from

$$\int_{\mathbb{S}^d} \Delta v\,\frac{|\nabla v|^2}{v}\,d\mu_d = \int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\,d\mu_d - 2\int_{\mathbb{S}^d} \mathrm{H}v : \frac{\nabla v \otimes \nabla v}{v}\,d\mu_d$$
$$= \frac{d}{d-1}\int_{\mathbb{S}^d} \|\mathrm{M}v\|^2\,d\mu_d - 2\int_{\mathbb{S}^d} \mathrm{L}v : \frac{\nabla v \otimes \nabla v}{v}\,d\mu_d - \frac{2}{d}\int_{\mathbb{S}^d} \Delta v\,\frac{|\nabla v|^2}{v}\,d\mu_d$$

a first identity that reads

$$\int_{\mathbb{S}^d} \Delta v\,\frac{|\nabla v|^2}{v}\,d\mu_d = \frac{d}{d+2}\left(\frac{d}{d-1}\int_{\mathbb{S}^d} \|\mathrm{M}v\|^2\,d\mu_d - 2\int_{\mathbb{S}^d} \mathrm{L}v : \frac{\nabla v \otimes \nabla v}{v}\,d\mu_d\right). \tag{5.32}$$

The Bochner-Lichnerowicz-Weitzenböck formula on $\mathbb{S}^d$ takes the simple form

$$\frac{1}{2}\,\Delta\,(|\nabla v|^2) = \|\mathrm{H}v\|^2 + \nabla(\Delta v)\cdot\nabla v + (d-1)\,|\nabla v|^2$$

where the last term, *i.e.*, $\mathrm{Ric}(\nabla v, \nabla v) = (d-1)\,|\nabla v|^2$, accounts for the Ricci curvature tensor contracted with $\nabla v \otimes \nabla v$. An integration of this formula on $\mathbb{S}^d$ shows a second identity,

$$\int_{\mathbb{S}^d} (\Delta v)^2\,d\mu_d = \frac{d}{d-1}\int_{\mathbb{S}^d} \|\mathrm{L}v\|^2\,d\mu_d + d\int_{\mathbb{S}^d} |\nabla v|^2\,d\mu_d\,. \tag{5.33}$$

Hence

$$\mathscr{K}[v] := \int_{\mathbb{S}^d}\left(\Delta v + \kappa\,\frac{|\nabla v|^2}{v}\right)\left(\Delta v + (\beta-1)\,\frac{|\nabla v|^2}{v}\right)d\mu_d$$
$$= \int_{\mathbb{S}^d} (\Delta v)^2\,d\mu_d + (\kappa+\beta-1)\int_{\mathbb{S}^d} \Delta v\,\frac{|\nabla v|^2}{v}\,d\mu_d + \kappa\,(\beta-1)\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\,d\mu_d$$

can be rewritten using (5.32) and (5.33) as

$$
\begin{aligned}
\mathscr{K}[v] = \frac{d}{d-1}\int_{\mathbb{S}^d}\|\mathrm{L}v\|^2\,d\mu_d + d\int_{\mathbb{S}^d}|\nabla v|^2\,d\mu_d \\
+(\kappa+\beta-1)\,\frac{d}{d+2}\left(\frac{d}{d-1}\int_{\mathbb{S}^d}\|\mathrm{M}v\|^2\,d\mu_d - 2\int_{\mathbb{S}^d}\mathrm{L}v:\mathrm{M}v\,d\mu_d\right) \\
+\kappa\,(\beta-1)\,\frac{d}{d-1}\int_{\mathbb{S}^d}\|\mathrm{M}v\|^2\,d\mu_d \\
=\frac{d}{d-1}\left(\|\mathrm{L}v\|^2 - 2\,b\,\mathrm{L}v:\mathrm{M}v + c\,\|\mathrm{M}v\|^2\right) + d\int_{\mathbb{S}^d}|\nabla v|^2\,d\mu_d \\
=\frac{d}{d-1}\left(\|\mathrm{L}v - b\,\mathrm{M}v\|^2 + \left(c-b^2\right)\|\mathrm{M}v|^2\right) + d\int_{\mathbb{S}^d}|\nabla v|^2\,d\mu_d \\
=\frac{d}{d-1}\,\|\mathrm{L}v - b\,\mathrm{M}v\|^2 + \left(c-b^2\right)\int_{\mathbb{S}^d}\frac{|\nabla v|^4}{v^2}\,d\mu_d + d\int_{\mathbb{S}^d}|\nabla v|^2\,d\mu_d
\end{aligned}
$$

where

$$
b = (\kappa+\beta-1)\,\frac{d-1}{d+2}\quad\text{and}\quad c = \frac{d}{d+2}\,(\kappa+\beta-1)+\kappa\,(\beta-1)\,.
$$

Let $\kappa = \beta\,(p-2)+1$. The condition $\gamma := c - b^2 \ge 0$ amounts to

$$
\gamma = \frac{d}{d+2}\,\beta\,(p-1) + \big(1+\beta\,(p-2)\big)\,(\beta-1) - \left(\frac{d-1}{d+2}\,\beta\,(p-1)\right)^2 \tag{5.34}
$$

where $\gamma = -\left(A\,\beta^2 - 2\,B\,\beta + C\right)$ with

$$
A = \left(\frac{d-1}{d+2}\,(p-1)\right)^2 + 2 - p\,,\quad B = \frac{d+3-p}{d+2}\quad\text{and}\quad C = 1\,.
$$

A necessary and sufficient condition for the existence of a β such that $\gamma \ge 0$ is that the reduced discriminant is nonnegative, which amounts to

$$
B^2 - A\,C = \frac{4\,d\,(d-2)}{(d+2)^2}\,(p-1)\,(2^*-p) \ge 0\,.
$$

Summarizing, we have the following result, which can be found in [121] for a general manifold.

Lemma 80. *With the above notations, for any smooth function v on $\mathbb{S}^d$, we have*

$$
\mathscr{K}[v] \ge \gamma\int_{\mathbb{S}^d}\frac{|\nabla v|^4}{v^2}\,d\mu_d + d\int_{\mathbb{S}^d}|\nabla v|^2\,d\mu_d
$$

for some $\gamma > 0$ given in terms of β by (5.34) if $p \in (1, 2^*)$.

Notice that we recover the expression of γ in (5.8) if we take $\beta = 1$. The case $p = 2$ does not add any difficulty compared to $p \neq 2$.

5.6.2 Diffusion flow and monotonicity

Assume that u is a positive solution of

$$
\frac{\partial u}{\partial t} = u^{-p\,(1-m)}\left(\Delta u + (m\,p-1)\,\frac{|\nabla u|^2}{u}\right). \tag{5.35}
$$

In the linear case $m=1$, u^p solves the heat equation. Otherwise we deal with the nonlinear case either of a fast diffusion flow with $m<1$ or of a solution of the porous media equation with $m>1$. We claim that

$$\frac{d}{dt}\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)}=0 \quad\text{and}\quad \frac{d}{dt}\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=2\,(p-2)\int_{\mathbb{S}^d} u^{-p\,(1-m)}\,|\nabla u|^2\,d\mu_d\,.$$

Let us assume that the parameters β and m are related by

$$m=1+\frac{2}{p}\left(\frac{1}{\beta}-1\right). \tag{5.36}$$

If v is a function such that $u=v^\beta$, then v solves

$$\frac{\partial v}{\partial t}=v^{2-2\beta}\left(\Delta v+\kappa\,\frac{|\nabla v|^2}{v}\right)$$

with $\kappa=\beta\,(p-2)+1$ and as a consequence we find that

$$\frac{d}{dt}\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=2\,(p-2)\,\beta^2\int_{\mathbb{S}^d}|\nabla v|^2\,d\mu_d\,.$$

Similarly, we find that

$$\frac{d}{dt}\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}=-2\int_{\mathbb{S}^d}\left(\beta\,v^{\beta-1}\,\frac{\partial v}{\partial t}\right)\left(\Delta v^\beta\right)d\mu_d=-2\,\beta^2\,\mathscr{K}[v] \tag{5.37}$$

By eliminating β in (5.34) using (5.36), we obtain

$$\gamma=\frac{(d+2)^2\,p^2\,m^2-2\,p\,(d+2)\,(d\,p+2)\,m+d^2\left(5\,p^2-12\,p+8\right)+4\,d\,(3-2\,p)\,p+4}{(d+2)^2\left(2-p\,(1-m)\right)^2}\,. \tag{5.38}$$

The condition $\gamma\ge0$ determines the range $m_-(d,p)\le m\le m_+(d,p)$ of *admissible* parameters m, where $m_\pm(d,p)$ is given by (5.13). Summarizing, we have the following result (also see [121]).

Lemma 81. *Assume that* $p\in(1,2^*)$ *and* $m\in[m_-(d,p),m_+(d,p)]$. *If* u *solves* (5.35), *then we have*

$$\frac{1}{2\,\beta^2}\,\frac{d}{dt}\left(\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}-d\,\mathcal{E}_p[u]\right)\le-\gamma\int_{\mathbb{S}^d}\frac{|\nabla v|^4}{v^2}\,d\mu_d \tag{5.39}$$

where $v=u^{1/\beta}$ *with* β *and* γ *given in terms of* m *respectively by* (5.34) *and* (5.38).

Notice that the case of the linear flow corresponds to the case $m=\beta=1$ and $v=u$.

Proof. For a smooth solution, the result follows from (5.37) and Lemma 80. The result for a general solution is obtained by standard regularization procedures. □

5.6.3 Interpolation

Depending on the value of p, we shall consider various interpolation inequalities. Let us define

$$\delta:=\frac{2-(4-p)\,\beta}{2\,\beta\,(p-2)}\quad\text{if}\quad p>2\,,\quad \delta:=1\quad\text{if}\quad p\in[1,2]\,. \tag{5.40}$$

Lemma 82. *If one of the following conditions is satisfied:*

(i) $p \in (1, 2^\#)$ *and* $\beta = 1$ *(so that* $\delta = 1$*),*

(ii) $p \in (4, 2^*)$ *if* $d \geq 5$,

(iii) $p \in \big(2, \min\{4, 2^*\}\big)$ *and* $1 < \beta < 2/(4-p)$,

then $u = v^\beta$ *is such that*

$$\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{|v|^2}\, d\mu_d \geq \frac{1}{\beta^2} \frac{\int_{\mathbb{S}^d} |\nabla u|^2\, d\mu_d \int_{\mathbb{S}^d} |\nabla v|^2\, d\mu_d}{\left(\int_{\mathbb{S}^d} |u|^2\, d\mu_d\right)^\delta \left(\int_{\mathbb{S}^d} |u|^p\, d\mu_d\right)^{\frac{\beta-1}{\beta(p-2)}}}. \tag{5.41}$$

Case (iii) was originally proved in [112, 113] and we refer to [120] for a proof in the case of the ultraspherical operator.

Proof. In case (i), $v = u$ and Inequality (5.41) is a consequence of the Cauchy-Schwarz inequality

$$\int_{\mathbb{S}^d} |\nabla v|^2\, d\mu_d = \int_{\mathbb{S}^d} \frac{|\nabla v|^2}{v} \cdot v\, d\mu_d \leq \left(\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\, d\mu_d\right)^{\frac12} \left(\int_{\mathbb{S}^d} |u|^2\, d\mu_d\right)^{\frac12},$$

Cases (i) and (ii) follow from two Hölder inequalities.
1) With $\frac12 + \frac{\beta-1}{2\beta} + \frac{1}{2\beta} = 1$, we deduce from

$$\int_{\mathbb{S}^d} |\nabla v|^2\, d\mu_d = \int_{\mathbb{S}^d} \frac{|\nabla v|^2}{v} \cdot 1 \cdot v\, d\mu_d \leq \left(\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\, d\mu_d\right)^{\frac12} \left(\int_{\mathbb{S}^d} 1\, d\mu_d\right)^{\frac{\beta-1}{2\beta}} \left(\int_{\mathbb{S}^d} |u|^2\, d\mu_d\right)^{\frac{1}{2\beta}},$$

and the assumption that $d\mu$ is a probability measure, the first estimate

$$\left(\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\, d\mu_d\right)^{\frac12} \geq \frac{\int_{\mathbb{S}^d} |\nabla v|^2\, d\mu_d}{\left(\int_{\mathbb{S}^d} |u|^2\, d\mu_d\right)^{\frac{1}{2\beta}}}.$$

2) With $\frac12 + \frac{\beta-1}{\beta(p-2)} + \frac{2-(4-p)\beta}{2\beta(p-2)} = 1$, Hölder's inequality shows that

$$\begin{aligned}\frac{1}{\beta^2} \int_{\mathbb{S}^d} |\nabla u|^2\, d\mu_d = \int_{\mathbb{S}^d} v^{2(\beta-1)} |\nabla v|^2\, d\mu_d &= \int_{\mathbb{S}^d} \frac{|\nabla v|^2}{v} \cdot v^{\frac{p(\beta-1)}{p-2}} \cdot v^{2\beta\delta}\, d\mu_d \\ &\leq \left(\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\, d\mu_d\right)^{\frac12} \left(\int_{\mathbb{S}^d} |u|^p\, d\mu_d\right)^{\frac{\beta-1}{\beta(p-2)}} \left(\int_{\mathbb{S}^d} |u|^2\, d\mu_d\right)^{\delta}\end{aligned}$$

from which we deduce the second estimate

$$\left(\int_{\mathbb{S}^d} \frac{|\nabla v|^4}{v^2}\, d\mu_d\right)^{\frac12} \geq \frac{1}{\beta^2} \frac{\int_{\mathbb{S}^d} |\nabla u|^2\, d\mu_d}{\left(\int_{\mathbb{S}^d} |u|^2\, d\mu_d\right)^\delta \left(\int_{\mathbb{S}^d} |u|^p\, d\mu_d\right)^{\frac{\beta-1}{\beta(p-2)}}}.$$

The combination of our two estimates proves (5.41). □

Using (5.36), Condition (iii) in Lemma 82 is changed into the condition that $\frac{2}{p} \leq m < 1$ if $2 < p < 4$ and we may notice as in [112, 113] that it is always satisfied if we choose $\beta = 4/(6-p)$ corresponding to an *admissible* fast diffusion exponent $m = (p+2)/(2p)$, for any $p \in (2, 2^*)$. By the *admissible*, one has to understand $m_-(d,p) \leq m \leq m_+(d,p)$, so that γ is nonnegative. With the choice of $m = (p+2)/(2p)$, we find $\delta = 1/4$.

5.6.4 Improved functional inequalities

Let us denote the *entropy* and the *Fisher information* respectively by

$$\mathsf{e} := \frac{1}{p-2}\left(\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\right) \quad \text{and} \quad \mathsf{i} := \|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}$$

and let γ and δ be given respectively by (5.34) and (5.40). Up to the replacement of u by $u/\|u\|_{\mathrm{L}^p(\mathbb{S}^d)}$, with no loss of generality, we shall assume that

$$\|u\|_{\mathrm{L}^p(\mathbb{S}^d)} = 1\,.$$

We learn from (5.39) and (5.41) that

$$(\mathsf{i} - d\,\mathsf{e})' \le \frac{\gamma\,\mathsf{i}\,\mathsf{e}'}{\big(1-(p-2)\,\mathsf{e}\big)^\delta}\,. \tag{5.42}$$

Solving the ordinary differential equation in the equality case of (5.42) with $\mathsf{e}' = -2\,\beta^2\,\mathsf{i}$, that is,

$$\left(\mathsf{e}' + 2\,d\,\beta^2\,\mathsf{e}\right)' = \frac{\gamma\,(\mathsf{e}')^2}{\big(1-(p-2)\,\mathsf{e}\big)^\delta}\,,$$

is equivalent to solve

$$\frac{d}{dt}\big(\mathsf{i} - d\,\varphi(\mathsf{e})\big) = \frac{\gamma}{\beta^2}\,\frac{\mathsf{e}'}{\big(1-(p-2)\,\mathsf{e}\big)^\delta}\,\big(\mathsf{i} - d\,\varphi(\mathsf{e})\big)$$

where φ solves

$$\varphi'(s) = 1 + \frac{\gamma}{\beta^2}\,\frac{\varphi(s)}{\big(1-(p-2)\,s\big)^\delta}\,. \tag{5.43}$$

The reader is invited to check that the solution of (5.43) with initial datum $\varphi(0) = 0$ is given by (5.10) if $m=1$ and by (5.12) with $\zeta = 2\,\gamma/\big(\beta\,(1-\beta)\big)$ in the nonlinear case. We learn from (5.42) that

$$\big(\mathsf{i} - d\,\varphi(\mathsf{e})\big)' \le \frac{\gamma}{\beta^2}\,\frac{\mathsf{e}'}{\big(1-(p-2)\,\mathsf{e}\big)^\delta}\,\big(\mathsf{i} - d\,\varphi(\mathsf{e})\big)\,.$$

This is enough to prove the following result.

Proposition 83. *With the above notations, we claim that*

$$\mathsf{i} \ge d\,\varphi(\mathsf{e})\,.$$

Proof. Let us give the scheme of a proof. Let $\tilde\gamma := \gamma/\beta^2$, in order to simplify notations. We can argue as follows:

1. $\mathsf{i}' + 2\,d\,\mathsf{i} = (\mathsf{i} - d\,\mathsf{e})' \le 0$ shows that
$$0 \le \mathsf{i}(t) \le \mathsf{i}(0)\,e^{-2\,d\,t}$$
and in particular $\lim_{t\to+\infty} \mathsf{i}(t) = 0$.
2. As $t \to +\infty$, e converges to a constant, hence $\lim_{t\to+\infty} \mathsf{e}(t) = 0$.
3. From (5.42), we learn that
$$(\mathsf{i} - d\,\mathsf{e})' \le d\,\tilde\gamma\,\mathsf{e}\,\mathsf{e}' = \frac{1}{2}\,d\,\tilde\gamma\,(\mathsf{e}^2)'$$
where the inequality follows from $1-(p-2)\,\mathsf{e} \le 1$ and $\mathsf{i} \ge d\,\mathsf{e}$.
4. It follows from $(\mathsf{i} - d\,\mathsf{e})' \le 0$ that $\mathsf{i} \ge d\,\mathsf{e}$ using an integration from any $t \ge 0$ to $+\infty$.

5. Unless u is a constant, we read from $(\mathsf{i}-d\,\mathsf{e})' \le \frac{1}{2}\,\tilde\gamma\,d\,(\mathsf{e}^2)'$ that $\mathsf{i}-d\,\mathsf{e} > \frac{1}{2}\,\tilde\gamma\,d\,\mathsf{e}^2$ using again an integration from any $t\ge 0$ to $+\infty$.

6. Take some $\vartheta\in(0,1)$ and consider the solution of
$$\overline{\varphi}'(s) = 1 + \frac{\vartheta\,\tilde\gamma\,\overline{\varphi}(s)}{(1-(p-2)\,s)^\delta}\,,\quad \overline{\varphi}(0)=0\,. \tag{5.44}$$
In the spirit of (5.42), we have a following chain of elementary estimates
$$(\mathsf{i}-d\,\vartheta\,\overline{\varphi}(\mathsf{e}))' \le (\mathsf{i}-d\,\overline{\varphi}(\mathsf{e}))' + d\,(1-\vartheta)\,(\overline{\varphi}(\mathsf{e}))' \le (\mathsf{i}-d\,\overline{\varphi}(\mathsf{e}))'$$
and obtain
$$(\mathsf{i}-d\,\vartheta\,\overline{\varphi}(\mathsf{e}))' \le \frac{\tilde\gamma\,\mathsf{e}'}{\big(1-(p-2)\,\mathsf{e}\big)^\delta}\,(\mathsf{i}-d\,\vartheta\,\overline{\varphi}(\mathsf{e}))\,. \tag{5.45}$$
We know that $\overline{\varphi}(0)=0$ and read from (5.44) that $\overline{\varphi}'(0)=1$ and $\overline{\varphi}''(0)=\vartheta\,\tilde\gamma\,\overline{\varphi}'(0)=\vartheta\,\tilde\gamma$ so that $\overline{\varphi}(\mathsf{e})-\mathsf{e}\sim \frac{1}{2}\,\vartheta\,\tilde\gamma\,\mathsf{e}^2$ as $\mathsf{e}\to 0$. Using $\mathsf{i}-d\,\mathsf{e} > \frac{1}{2}\,\tilde\gamma\,d\,\mathsf{e}^2$, we learn that
$$\mathsf{i}-d\,\overline{\varphi}(\mathsf{e}) \ge \frac{1}{2}\,\tilde\gamma\,d\,(1-\vartheta)\,\mathsf{e}^2\,\big(1+O(\mathsf{e})\big)$$
for $\mathsf{e}=\mathsf{e}(t)$ small enough, *i.e*, for $t>0$ large enough.

7. It is simple to check from (5.45) that $\mathsf{i}-d\,\vartheta\,\overline{\varphi}(\mathsf{e})$ cannot change sign.

8. We conclude as above that $\mathsf{i}-d\,\vartheta\,\overline{\varphi}(\mathsf{e})\ge 0$ using an integration from any $t\ge 0$ to $+\infty$.

9. Finally, we consider the limit as $\vartheta\to 1_-$.

Altogether, we conclude that $\mathsf{i}\ge d\,\varphi(\mathsf{e})$ where φ solves (5.43). This completes the scheme of the proof of Proposition 83. □

Chapter 6

On Gaussian interpolation inequalities

This chapter corresponds to the paper [P3], accepted in Comptes Rendues Mathématique, in collaboration with Jean Dolbeault and Nikita Simonov.

Abstract

This paper is devoted to Gaussian interpolation inequalities with endpoint cases corresponding to the Gaussian Poincaré and the logarithmic Sobolev inequalities, seen as limits in large dimensions of Gagliardo-Nirenberg-Sobolev inequalities on spheres. Entropy methods are investigated using not only heat flow techniques but also nonlinear diffusion equations as on spheres. A new stability result is established for the Gaussian measure, which is directly inspired by recent results for spheres.

Résumé

Cet article est consacré à des inégalités d'interpolation Gaussiennes, avec comme cas extrêmes l'inégalité de Poincaré Gaussienne et l'inégalité de Sobolev logarithmique, vues comme limites en grandes dimensions des inégalités de Gagliardo-Nirenberg-Sobolev sur les sphères.

Les méthodes d'entropie sont abordées en utilisant non seulement des techniques basée sur l'équation de la chaleur mais aussi sur des équations de diffusion non-linéaires, comme pour les sphères. Un nouveau résultat de stabilité est établi pour les mesures Gaussiennes, qui s'inspire directement de résultats récents sur les sphères.

6.1 Introduction and main results

Let us consider the *Gagliardo-Nirenberg-Sobolev inequalities* on the unit d-dimensional sphere

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)} \ge \frac{d}{p-2}\left(\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d,d\mu_d)} - \|u\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)}\right) \quad \forall\, u \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d) \tag{6.1}$$

for any $p \in [1,2) \cup (2,+\infty)$ if $d = 1$, 2, and for any $p \in [1,2) \cup (2,2^*]$ if $d \ge 3$. Here $d\mu_d$ denotes the uniform probability measure on $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ and, if $d \ge 3$, $2^* = 2\,d/(d-2)$ is the critical Sobolev exponent. By convention, we take $2^* = +\infty$ if $d = 1$ or 2. The purpose of this paper is to clarify the links of these interpolation inequalities with the family of *Gaussian interpolation inequalities*

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \ge \frac{1}{2-p}\left(\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right) \quad \forall\, v \in \mathrm{H}^1(\mathbb{R}^n,d\gamma) \tag{6.2}$$

where the exponent p is taken in the range $1 \le p < 2$. Inequality (6.2) is intermediate between the Poincaré inequality corresponding to $p = 1$ and the *Gaussian logarithmic Sobolev inequality*

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \ge \frac{1}{2}\int_{\mathbb{R}^n}|v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right)d\gamma \quad \forall\, v \in \mathrm{H}^1(\mathbb{R}^n,d\gamma)$$

obtained as a limit case of (6.2) as $p \to 2_-$. Here $d\gamma(y) := (2\pi)^{-n/2}\,e^{-\frac{1}{2}|y|^2}\,dy$ denotes the centred normalized Gaussian probability measure and the dimension n is any positive integer.

It is somewhat classical that, if we consider the Sobolev inequality on the sphere, *i.e.*, Inequality (6.1) with $p = 2^*$ and $d \ge 3$, rescale it to the sphere of radius $\sqrt{d}$ and fix a function depending on n variables only on this sphere, since the curvature tends to 0 at the right order as one takes the limit as $d \to +\infty$, then the sphere tends to the flat space with Gaussian measure. For instance, we read (with adapted notations) from [36, p. 4818] that *if we rescale this inequality so as to be on a sphere of radius* $\sqrt{d}$ *and take the limit* $d \to \infty$ *or* $p \to 2$ *we obtain in the Poincaré limit the Gross logarithmic inequality for the Gaussian measure since* $-(1/d)\,\Delta$ *on* $\mathbb{S}^d$ *goes to* $-\Delta + x\cdot\nabla$ *on the infinite-dimensional limit.* The last part of the sentence refers to a result known as the MaxwellPoincaré lemma : see [208] and [252, Remark 4, p. 254] for some historical comments. The statement of [36] has been made more precise later in [39, 38] using a slightly different limit. However, to our knowledge, the infinite-dimensional limit has not been considered in the subcritical range $p < 2^*$.

On the sphere, Inequality (6.1) follows from [24, 25] for any $p \ge 1$ if $d = 1$ and any $p \in [1, 2^\#]$ with $2^\# := (2\,d^2+1)/(d-1)^2$ if $d \ge 2$. The proof in the range $p \in (2^\#, +\infty)$ if $d = 2$ and $p \in (2^\#, 2^*]$ if $d \ge 3$ can be found in [48, Corollary 6.1], [49] and [37]. Also see [196] in the case $p = 2^*$. In the case of the Gaussian measure, we refer to [35] (also see [187]) for a first proof of Inequality (6.2). The formal analogy of (6.1) and (6.2) is striking. Although computations are somewhat standard, our first purpose is to make this point rigorous and recover (6.2) as a special limit of (6.1) as $d \to +\infty$.

Theorem 84. *Let n be a positive integer, $p \in [1,2)$ and consider a function $v \in \mathrm{H}^1(\mathbb{R}^n,d\gamma)$ with compact support. For any $d \ge n$, large enough, if $u_d(\omega) = v\big(\omega_1/\sqrt{d},\omega_2/\sqrt{d},\ldots,\omega_n/\sqrt{d}\big)$ where $\omega = (\omega_1,\omega_2,\ldots,\omega_{d+1}) \in \mathbb{R}^{d+1} \supset \mathbb{S}^d$ is such that $|\omega| = 1$, then*

$$\lim_{d\to+\infty} d\left(\|\nabla u_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)} - \frac{d}{2-p}\left(\|u_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)} - \|u_d\|^2_{\mathrm{L}^p(\mathbb{S}^d,d\mu_d)}\right)\right)$$
$$= \|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2-p}\left(\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right).$$

The *carré du champ* method has frequently been applied to prove Gaussian interpolation inequalities ranging between the logarithmic Sobolev inequality and the Poincaré inequality like (6.2) using the linear flow associated with the Ornstein-Uhlenbeck operator; see [24], and [26] for an overview. Still in the case of the Gaussian measure, we adopt here a new point of view by using *nonlinear diffusion equations* in order to prove the same inequalities, but with different remainder terms. This is a very natural point of view when dealing with inequalities like (6.1) on the sphere, as shown for instance in [121] (see earlier references therein). In that case, linear flows are indeed limited to exponents $p \le 2^\#$ if $d \ge 2$. To overcome this difficulty if either $d = 2$ and $p > 2^\#$ or $d \ge 3$ and $p \in (2^\#, 2^*]$, one has to consider fast diffusion flows. Before explaining the results for the Gaussian measure, let us summarize the main known results on the sphere.

On $\mathbb{S}^d$, let us consider a positive solution u of

$$\frac{\partial u}{\partial t} = u^{-p(1-m)} \left(\Delta u + (m\,p-1)\,\frac{|\nabla u|^2}{u} \right) \tag{6.3}$$

where Δ denotes the Laplace-Beltrami operator on $\mathbb{S}^d$. In the case $m = 1$, u^p solves the heat equation and for this reason, we shall call it the *linear case*. Otherwise u^p solves a nonlinear diffusion equation corresponding either to a fast diffusion flow with $m < 1$ or to a porous media equation with $m > 1$. In any case, we claim that

$$\frac{d}{dt}\,\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)} = 0 \quad \text{and} \quad \frac{d}{dt}\,\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)} = 2\,(p-2) \int_{\mathbb{S}^d} u^{-p(1-m)}\,|\nabla u|^2\,d\mu_d\,.$$

Let us define

$$m_\pm(d,p) := \frac{1}{(d+2)\,p} \left(d\,p + 2 \pm \sqrt{d\,(p-1)\,\big(2\,d-(d-2)\,p\big)} \right). \tag{6.4}$$

The following result can be found in [113, 121] with additional details in [122, 118] and [P2].

Proposition 85 ([113, 121]). *Assume that* $d \ge 1$*, with either* $p \in [1,2) \cup (2,+\infty)$ *if* $d = 2$ *or* $p \in [1,2) \cup (2, 2^*]$ *if* $d \ge 3$*, and let* $m \in [m_-(d,p), m_+(d,p)]$*. If* $u > 0$ *solves* (6.3) *with an initial datum in* $\mathrm{L}^2 \cap \mathrm{L}^p(\mathbb{S}^d, d\mu_d)$*, then*

$$\frac{d}{dt} \left(\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d, d\mu_d)} - \frac{d}{p-2} \left(\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d, d\mu_d)} - \|u\|^2_{\mathrm{L}^2(\mathbb{S}^d, d\mu_d)} \right) \right) \le 0 \quad \forall\, t > 0\,.$$

The limit, as $t \to +\infty$, of any solution of (6.3) is a constant. This means that the *deficit*, that is, the difference of the two sides in Inequality (6.1), converges to 0. Then, by Proposition 85, it follows that the deficit is non-negative, which directly proves (6.1). The same monotonicity property applies to the deficit of the *logarithmic Sobolev inequality on the sphere*

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d, d\mu_d)} \ge \frac{d}{2} \int_{\mathbb{S}^d} |u|^2 \log \left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} \right) d\mu_d \quad \forall\, u \in \mathrm{H}^1(\mathbb{S}^d, d\mu_d)$$

in the limit case corresponding to $p = 2$. The admissible values of the parameters are limited to $m_-(d,p) \le m \le m_+(d,p)$ and $1 \le p \le 2^*$ if $d \ge 3$. Moreover, at the endpoints, we have $m_\pm(d,1) = 1$ and $m_\pm(d,2^*) = (d-1)/d$ if $d \ge 3$, while $m_+(d,2^\#) = 1$ so that $m = 1$ is admissible if and only if $1 \le p \le 2^\#$ when $d \ge 2$. See Fig. 6.1. For appropriate initial data, it is shown in [122] that the monotonicity property of the deficit along the flow of (6.3) is violated for any $p \in [2, 2^*)$ or $p = 2^*$ if $d \ge 3$ as soon as either $m < m_-(d,p)$ or $m > m_+(d,p)$.

In view of the results of Theorem 84, it is natural to ask whether there is also a monotonicity property of the deficit associated to the Gaussian interpolation inequalities (6.2) when we rely on a nonlinear diffusion flow on $\mathbb{R}^n$. Let $m_\pm(p) := \lim_{d\to+\infty} m_\pm(d,p)$ and notice that

$$m_\pm(p) = 1 \pm \frac{1}{p} \sqrt{(p-1)\,(2-p)}\,. \tag{6.5}$$

The diffusion operator associated to the Gaussian measure is the *Ornstein-Uhlenbeck operator* $\mathcal{L} = \Delta - x \cdot \nabla$ and we consider now the nonlinear parabolic equation

$$\frac{\partial v}{\partial t} = v^{-p(1-m)} \left(\mathcal{L} v + (m\,p-1)\,\frac{|\nabla v|^2}{v} \right). \tag{6.6}$$

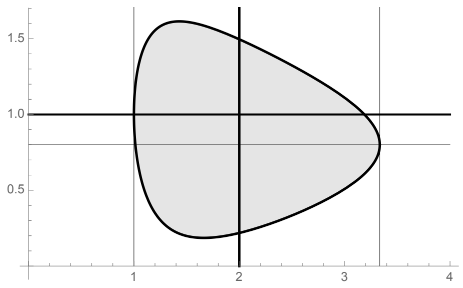

Figure 6.1: Case $d = 5$. The admissible parameters p and m correspond to the grey area. The boundary of the admissible set is tangent to the vertical lines $p = 1$ at $(m, p) = (1,1)$ and $p = 2^* = 10/3$ at $(m, p) = (4/5, 10/3)$. Qualitatively, this figure does not change as d increases but gets squeezed in the interval $1 \le p \le 2$ as $d \to +\infty$.

In the definition of $\mathcal{L}$, Δ denotes the standard Laplacian on $\mathbb{R}^n$. The following result is new for $m \neq 1$ while the case $m = 1$ follows from the method of the *carré du champ* developed by Bakry and Emery in [24].

Theorem 86. *Assume that $n \ge 1$, $p \in [1,2)$. If $v > 0$ solves (6.6) with $m \in [m_-(p), m_+(p)]$ for an initial datum in $\mathrm{L}^2 \cap \mathrm{L}^p(\mathbb{R}^n, d\gamma)$, then*

$$\frac{d}{dt}\left(\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{p-2}\left(\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} - \|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\right)\right) \le 0 \quad \forall\, t > 0\,.$$

The limiting case $p = 2$ corresponding to the Gaussian logarithmic Sobolev inequality is also covered but it is obtained as a standard application of the linear *carré du champ* method known from [24] because $m_\pm(2) = 1$. See Fig. 6.2.

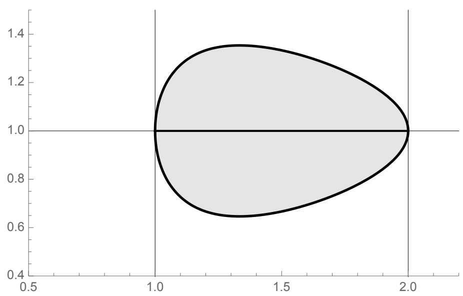

Figure 6.2: The admissible parameters p and m correspond to the grey area and are independent of the dimension n. The boundary of the admissible set is tangent to the vertical lines $p = 1$ at $(m, p) = (1,1)$ and $p = 2$ at $(m, p) = (1,2)$. It is the limit set of the admissible parameters for Proposition 85 as $d \to +\infty$.

Said in simple words, the result of Theorem 86 is that the admissible range of exponents of the nonlinear flow, for which the deficit associated to (6.2) is monotone non-increasing, is obtained as the limit of the range of the corresponding exponents on the sphere, in the large dimensions limit. Moreover, $p = 2$ appears as a *critical exponent* for the Gaussian measure.

Let us now focus on stability results. The main result of [P2] for the sphere is a constructive stability estimate for Inequality (6.1), limited to the subcritical range $p \in (1, 2^*)$, which measures the distance to optimal functions, and distinguishes the subspaces generated by constant functions, spherical harmonic functions associated to the first positive eigenvalue of the Laplace-Beltrami operator, and the orthogonal directions. Optimal exponents in the stability estimate measuring the distance to the set of optimal functions differ, depending on the directions. Here we have the exact counterpart in the Gaussian case. Let Π_1 denote the orthogonal projection of $\mathrm{L}^2(\mathbb{R}^n,d\gamma)$ onto the $(n+1)$-dimensional function space generated by 1 and x_i with $i=1,2,\ldots,n$.

Theorem 87. *For all $n\ge1$, and all $p\in(1,2)$, there is an explicit constant $c_{n,p}>0$ such that, for all $v\in\mathrm{H}^1(d\gamma)$, it holds*

$$\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\frac{1}{p-2}\left(\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}-\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\right)$$
$$\ge c_{n,p}\left(\|\nabla(\mathrm{Id}-\Pi_1)\,v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}+\frac{\|\nabla\Pi_1\,v\|^4_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}+\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right).$$

Exponents 2 and 4 which appear in the right-hand side of the inequality are sharp and the constant $c_{n,p}$ has an explicit although complicated expression given in the proof. If $p=1$, the last term drops and the distance to optimal functions is measured only by

$$\|\nabla(\mathrm{Id}-\Pi_1)\,v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\,,$$

and a decomposition on Hermite polynomials shows that the estimate cannot be improved. If $p=2^*$ and $n\ge3$, the recent stability result on Sobolev's inequality on $\mathbb{R}^n$ of [119], which quantifies the estimate of Bianchi and Egnell in [47] can be translated into a stability result for (6.1) on $\mathbb{S}^d$, that can be recast in the form of [P2, Theorem 6]. By a large dimension argument, a stability on the Gaussian logarithmic Sobolev inequality is also shown in [119], although the distance is measured only by an $\mathrm{L}^2(\mathbb{R}^n,d\gamma)$ norm. Whether a stronger estimate can be obtained in the limiting case $p=2$, eventually under some restriction, is therefore so far an open question.

This paper is organized as follows. In Section 6.2, we give a new proof of Inequality (6.2) as a consequence of Inequality (6.1) by taking a large dimensions limit, applied to the inequality written for a function depending only on a fixed number n of real variables. To our knowledge, this is new except for the limit case $p=2$ of the logarithmic Sobolev inequality. Section 6.3 is devoted to the proof of Theorem 86: we characterize the nonlinear diffusion flows of porous medium or fast diffusion type such that the deficit is monotone nonincreasing and recover the picture known on the sphere in the large dimensions limit. Moreover, by the *carré du champ* method, we establish improved inequalities that provide us with first stability results. The stability result of Theorem 87 for the Gaussian measure is proved in Section 6.4 using a detailed Taylor expansion and the improved inequalities of Section 6.3.

6.2 From subcritical interpolation inequalities on the sphere to Gaussian interpolation

In this section we explain how Inequality (6.2) can be seen as the limit of Inequality (6.1) in the large dimensions limit, that is, as $d\to+\infty$, and prove Theorem 84. Comments on the limit case $p=2$ can be found at the end of this section.

The unit sphere $\mathbb{S}^d$ is parametrized in terms of the *stereographic coordinates* by

$$\omega_j=\frac{2\,x_j}{1+|x|^2}\quad\text{if}\quad 1\le j\le d\quad\text{and}\quad \omega_{d+1}=\frac{1-|x|^2}{1+|x|^2}$$

where $\omega=(\omega_1,\omega_2,\dots,\omega_{d+1})$ denote the coordinates in $\mathbb{R}^{d+1}\supset\mathbb{S}^d$ and $x=(x_1,x_2,\dots,x_d)$ are Cartesian coordinates in $\mathbb{R}^d$. To a function u on $\mathbb{S}^d$, we associate a function w on $\mathbb{R}^d$ using the stereographic projection such that

$$\left(\frac{2}{1+|x|^2}\right)^{\frac{d-2}{2}}u(\omega)=w(x)\quad\forall\,x\in\mathbb{R}^d\,.$$

It is a standard result that

$$\int_{\mathbb{S}^d}|u|^p\,d\mu_d=2^{\frac{\delta(p)}{2}}\left|\mathbb{S}^d\right|^{-1}\int_{\mathbb{R}^d}\langle x\rangle^{-\delta(p)}\,|w|^p\,dx$$

and that

$$\int_{\mathbb{S}^d}|\nabla u|^2\,d\mu_d+\tfrac14\,d\,(d-2)\int_{\mathbb{S}^d}|u|^2\,d\mu_d=\left|\mathbb{S}^d\right|^{-1}\int_{\mathbb{R}^d}|\nabla w|^2\,dx\,.$$

where $\langle x\rangle:=\sqrt{1+|x|^2}$ and $\delta(p):=2\,d-p\,(d-2)$. Using the stereographic projection, Inequality (6.1) can be written on the Euclidean space $\mathbb{R}^d$ as the weighted interpolation inequality

$$\int_{\mathbb{R}^d}|\nabla w|^2\,dx+\frac{d\,\delta(p)}{p-2}\int_{\mathbb{R}^d}\frac{|w|^2}{\langle x\rangle^4}\,dx\ge\frac{\mathcal{C}_{d,p}}{p-2}\left(\int_{\mathbb{R}^d}\frac{|w|^p}{\langle x\rangle^{\delta(p)}}\,dx\right)^{\frac2p}\quad\text{with}\quad\mathcal{C}_{d,p}=2^{\frac{\delta(p)}{p}}\,d\left|\mathbb{S}^d\right|^{1-\frac2p}.\tag{6.7}$$

See [118] for details. Equality is achieved by the Aubin-Talenti function $w_\star(x):=\langle x\rangle^{2-d}$. Assume that $d\ge4$. Let us consider $f=w/w_\star$ and notice that the inequality rewritten in terms of f is

$$\int_{\mathbb{R}^d}|\nabla f|^2\,w_\star^2\,dx+\frac{4\,d}{p-2}\int_{\mathbb{R}^d}|f|^2\,w_\star^{2^*}\,dx\ge\frac{\mathcal{C}_{d,p}}{p-2}\left(\int_{\mathbb{R}^d}|f|^p\,w_\star^{2^*}\,dx\right)^{\frac2p}.$$

At this point, we may notice that for any $d\ge3$, we can choose freely any $p\in[1,2)\cup(2,2^*]$ if $d\ge3$. Since we are interested in the limit as $d\to+\infty$, in order to consider a fixed, given value of p, we have no choice but to restrict p to the case $1\le p<2$.

With $g(x)=f\big(x/\sqrt d\big)$, we obtain after changing variables that

$$\frac14\int_{\mathbb{R}^d}|\nabla g|^2\,\frac{dx}{\left(1+\frac1d\,|x|^2\right)^{d-2}}+\frac{1}{p-2}\int_{\mathbb{R}^d}|g|^2\,\frac{dx}{\left(1+\frac1d\,|x|^2\right)^{d}}\ge\frac{\mathcal{C}_{d,p}\,d^{\,d\frac{p-2}{2p}}}{4\,d\,(p-2)}\left(\int_{\mathbb{R}^d}|g|^p\,\frac{dx}{\left(1+\frac1d\,|x|^2\right)^{d}}\right)^{\frac2p}.$$

Let us assume that $n\ge1$ is a given integer and take $d>\max\{n,3\}$. With $x=(y,z)\in\mathbb{R}^n\times\mathbb{R}^{d-n}\approx\mathbb{R}^d$, we also assume that the function g depends only on y. In other words, we write $g=g_d$ where $g_d(y,z)=v(y)$ for some function v defined on $\mathbb{R}^n$, that is,

$$g_d(y,z)=v(y)\quad\forall\,(y,z)\in\mathbb{R}^n\times\mathbb{R}^{d-n}\tag{6.8}$$

Here we use an index d in order to emphasize that g_d has to be considered as a function on $\mathbb{R}^d$. Let us define

$$c_d:=(d\,\pi)^{\frac d2}\,\frac{\Gamma(d/2)}{\Gamma(d)}\,.$$

Lemma 88. *Let n be a positive integer, $p\in[1,2)$, consider a function $v\in\mathrm{H}^1(\mathbb{R}^n,d\gamma)$ with compact support and* *define g_d according to (6.8). Then we have*

$$\lim_{d\to+\infty}\frac{1}{4\,c_d}\int_{\mathbb{R}^d}|\nabla g_d|^2\,\frac{dx}{\left(1+\frac{1}{d}|x|^2\right)^{d-2}} + \frac{1}{2-p}\lim_{d\to+\infty}\frac{1}{c_d}\left(\frac{\mathcal{C}_{d,p}\,d^{d\frac{p-2}{2p}}}{4\,d}\left(\int_{\mathbb{R}^d}|g_d|^p\,\frac{dx}{\left(1+\frac{1}{d}|x|^2\right)^d}\right)^{\frac{2}{p}}-\int_{\mathbb{R}^d}|g_d|^2\,\frac{dx}{\left(1+\frac{1}{d}|x|^2\right)^d}\right)$$
$$=\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\frac{1}{2-p}\left(\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right).$$

In other words, we prove that the infinite dimensional limit of (6.7), for functions depending only on a finite number n of real variables, is (6.2). The assumption of compact support can be removed if g_d is square integrable with respect to the measure $\left(1+\frac{1}{d}|x|^2\right)^{-d}dx$ and ∇g_d is square integrable with respect to the measure $\left(1+\frac{1}{d}|x|^2\right)^{2-d}dx$, at least for some $d\in\mathbb{N}$ large enough.

Proof. Using

$$\left(1+\tfrac{1}{d}|x|^2\right)^{2-d}=\left(1+\tfrac{1}{d}\left(|y|^2+|z|^2\right)\right)^{2-d}=\left(1+\tfrac{1}{d}|y|^2\right)^{2-d}\left(1+\tfrac{1}{d}|\zeta|^2\right)^{2-d}\quad\text{with}\quad \zeta=\frac{z}{\sqrt{1+\frac{1}{d}|y|^2}},$$

we can integrate with respect to z and obtain

$$\int_{\mathbb{R}^{d-n}}\left(1+\tfrac{1}{d}|x|^2\right)^{2-d}dz=\left(1+\tfrac{1}{d}|y|^2\right)^{2-\frac{d+n}{2}}\int_{\mathbb{R}^{d-n}}\left(1+\tfrac{1}{d}|\zeta|^2\right)^{2-d}d\zeta\,.$$

Let $d\gamma(y):=(2\pi)^{-n/2}\,e^{-\frac{1}{2}|y|^2}\,dy$ be the centred normalized Gaussian probability measure. We recall that $|\mathbb{S}^{k-1}|=2\pi^{k/2}/\Gamma(k/2)$ for any $k\in\mathbb{N}\setminus\{0\}$ and

$$\int_0^{+\infty}r^{a-1}\left(1+\tfrac{1}{d}r^2\right)^{-b}dr=d^{\frac{a}{2}}\,\frac{\Gamma\left(\frac{a}{2}\right)\Gamma\left(b-\frac{a}{2}\right)}{2\,\Gamma(b)}$$

if $0<a<2b$. Applying these formulas with $a=k=d-n$ and $b=d-2>2-n$, we find that

$$\int_{\mathbb{R}^{d-n}}\left(1+\tfrac{1}{d}|\zeta|^2\right)^{2-d}d\zeta=(d\,\pi)^{\frac{d-n}{2}}\,\frac{\Gamma\left(\frac{d+n}{2}-2\right)}{\Gamma(d-2)}\,.$$

Applying these formulas with $a=b=k=d\ge 2$, we find that

$$\int_{\mathbb{R}^d}\left(1+\tfrac{1}{d}|x|^2\right)^{-d}dx=c_d$$

and, as a consequence

$$\lim_{d\to+\infty}\frac{1}{c_d}\int_{\mathbb{R}^{d-n}}\left(1+\tfrac{1}{d}|\zeta|^2\right)^{2-d}d\zeta=\frac{4}{(2\pi)^{n/2}}$$

using Stirling's formula. Since

$$\lim_{d\to+\infty}\left(1+\tfrac{1}{d}|y|^2\right)^{2-\frac{d+n}{2}}=e^{-\frac{1}{2}|y|^2},$$

we obtain

$$\lim_{d\to+\infty}\frac{1}{c_d}\int_{\mathbb{R}^d}|\nabla g_d(y)|^2\left(1+\tfrac{1}{d}|x|^2\right)^{2-d}dx=4\int_{\mathbb{R}^n}|\nabla v|^2\,d\gamma\,.$$

Similar computations show that

$$\int_{\mathbb{R}^d} |g_d(y)|^2 \left(1+\tfrac{1}{d}|x|^2\right)^{-d} dx = \int_{\mathbb{R}^n} |g_d(y)|^2 \left(1+\tfrac{1}{d}|y|^2\right)^{-\frac{d+n}{2}} dy \int_{\mathbb{R}^{d-n}} \left(1+\tfrac{1}{d}|\zeta|^2\right)^{-d} d\zeta\,,$$

$$\lim_{d\to+\infty} \frac{1}{c_d} \int_{\mathbb{R}^d} |g_d(y)|^2 \left(1+\tfrac{1}{d}|x|^2\right)^{-d} dx = \int_{\mathbb{R}^n} |v|^2\, d\gamma\,,$$

and

$$\int_{\mathbb{R}^d} |g_d(y)|^p \left(1+\tfrac{1}{d}|x|^2\right)^{-d} dx = \int_{\mathbb{R}^n} |g_d(y)|^p \left(1+\tfrac{1}{d}|y|^2\right)^{-\frac{d+n}{2}} dy \int_{\mathbb{R}^{d-n}} \left(1+\tfrac{1}{d}|\zeta|^2\right)^{-d} d\zeta\,,$$

$$\lim_{d\to+\infty} \frac{\mathcal{C}_{d,p}\, d^{d\,\frac{p-2}{2p}}}{4\,d\,c_d} \left(\int_{\mathbb{R}^d} |g_d(y)|^p \left(1+\tfrac{1}{d}|x|^2\right)^{-d} dx\right)^{\frac{2}{p}} = \left(\int_{\mathbb{R}^n} |v|^p\, d\gamma\right)^{\frac{2}{p}}.$$

This completes the proof of Lemma 88. □

Proof of Theorem 84. Applied to the function u_d, Inequality (6.1) is transformed into Inequality (6.7) applied to

$$g_d(x) = u_d\left(\frac{2\,y}{1+\frac{1}{d}|x|^2}\right) \quad \forall\, x=(y,z)\in\mathbb{R}^n\times\mathbb{R}^{d-n}\,.$$

Since the right-hand side uniformly converges to $v(y)$ for any smooth and compactly supported function v, the same conclusion holds for Theorem 84 as for Lemma 88. □

It is a natural question to ask what happens in (6.1) to the marginals depending only on a finite number n of variables if $p=2$ or in the case $2<p\le 2^*=2\,d/(d-2)$. We may notice that $\lim_{d\to+\infty} 2\,d/(d-2)=2$ and it is known, for instance from [36], that one recovers the *Gaussian logarithmic Sobolev inequality* as a limit case of Sobolev's inequality on $\mathbb{S}^d$ corresponding to $p=2^*$ when $d\to+\infty$. This is also true if we consider a sequence $(p_d)_{d\in\mathbb{N}}$ with $1<p_d<2^*$, depending on d, if its limit is also 2, as shown next. By convention, when $p_d=2$, we consider the Gaussian logarithmic Sobolev inequality instead of (6.2).

Proposition 89. *Let n be a positive integer and consider a function $v\in \mathrm{H}^1(\mathbb{R}^n,d\gamma)$ with compact support. For any $d\ge n$, large enough, let*

$$u_d(\omega) = v\big(\omega_1/\sqrt{d},\omega_2/\sqrt{d},\ldots,\omega_n/\sqrt{d}\big),$$

where $\omega=(\omega_1,\omega_2,\ldots,\omega_d)\in\mathbb{R}^{d+1}\supset\mathbb{S}^d$ is such that $|\omega|=1$, as in Theorem 84. Then we have

$$\lim_{d\to+\infty} d\left(\|\nabla u_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)} - \frac{d}{2}\int_{\mathbb{S}^d} |u_d|^2 \log\left(\frac{|u_d|^2}{\|u_d\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu_d\right)$$
$$= \|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^n} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right) d\gamma\,.$$

If $(p_d)_{d\ge3}$ is a sequence of real numbers such that $p_d\in(1,2)\cup(2,2^*)$ and $\lim_{d\to+\infty} p_d=2$, then

$$\lim_{d\to+\infty} d\left(\|\nabla u_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)} - \frac{d}{2-p_d}\left(\|u_d\|^2_{\mathrm{L}^2(\mathbb{S}^d,d\mu_d)} - \|u_d\|^2_{\mathrm{L}^{p_d}(\mathbb{S}^d,d\mu_d)}\right)\right)$$
$$= \|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^n} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right) d\gamma\,.$$

Proof. The proof is an adaptation of the proof of Theorem 84 and, in the case $p_d \neq 2$, relies on the standard observation that

$$\lim_{p\to 2} \frac{\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} - \|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{p-2} = \frac{1}{2}\int_{\mathbb{R}^n} |v|^2 \log\left(\frac{|v|^2}{\|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right) d\gamma\,.$$

As this computation raises no special difficulty, details are omitted. □

6.3 Entropy methods and nonlinear flows for Gaussian measures

In this section, we prove the result of Theorem 86 for the Gaussian measure $d\gamma$ and extend it to the slightly more general framework of a uniformly strictly log-concave measure $d\mu$, before drawing some consequences. Most of the results are similar to computations usually done on the sphere, but we are not aware of the use of nonlinear flows ($m \neq 1$) in the context of Gaussian measures. This approach is very natural in the perspective of spheres in the large-dimensional limit.

6.3.1 Gaussian interpolation inequalities: a proof by *carré du champ*

On $\mathbb{R}^n$, let us consider the probability measure

$$d\mu = Z^{-1}\, e^{-\phi}\, dy \quad \text{with} \quad Z = \int_{\mathbb{R}^n} e^{-\phi}\, dy \tag{6.9}$$

and redefine the *Ornstein-Uhlenbeck operator* by

$$\mathcal{L} := \Delta - \nabla\phi\cdot\nabla \tag{6.10}$$

on $\mathrm{L}^2(\mathbb{R}^n, d\mu)$. This generalizes the case of the harmonic potential $\phi(y) = \frac{1}{2}|y|^2$ considered in the introduction. We assume that ϕ satisfies the *Bakry-Emery condition*

$$\mathrm{Hess}\,\phi \geq \lambda_\star\,\mathrm{Id} \quad \text{a.e.} \tag{6.11}$$

for some $\lambda_\star > 0$. The harmonic potential corresponds to the equality case with $\lambda_\star = 1$. Under Assumption (6.11), it is well known (see for instance [26, Section 7.6.2]) that, with $\lambda = \lambda_\star$ and for any $p \in [1,2)$,

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} \geq \frac{\lambda}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right) \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^n, d\mu) \tag{6.12}$$

and also, by taking the limit as $p \to 2_-$, that

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} \geq \frac{\lambda}{2}\int_{\mathbb{R}^n} |f|^2 \log\left(\frac{|f|^2}{\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}}\right) d\mu \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^n, d\mu)\,.$$

The classical proof by the carré du champ, as in [23, 24, 19], relies on the *Ornstein-Uhlenbeck flow* $\partial_t\rho = \mathcal{L}\rho$ applied to the solution with initial datum $\rho(t=0,\cdot) = |f|^p$. Here we consider a more general strategy and compute as in the carré du champ method using the *nonlinear diffusion flow*

$$\frac{\partial\rho}{\partial t} = \frac{1}{m}\,\mathcal{L}\rho^m \quad \forall\, t \geq 0\,, \quad \rho(t=0,\cdot) = |f|^p\,. \tag{6.13}$$

Our goal is to understand the range of m for which we have a monotonicity property of the deficit as in the case $m=1$. Let $m_\pm(p)$ be defined as in (6.5).

Theorem 90. *Assume that $n \geq 1$, $p \in [1,2)$ and $m \in [m_-(p), m_+(p)]$. We consider the measure $d\mu$ as in (6.9) such that (6.11) holds for some $\lambda_\star > 0$. If $\rho > 0$ solves (6.13) with $\mathcal{L}$ defined by (6.10) for an initial datum $\rho(t=0,\cdot) = |f|^p$ in $\mathrm{L}^{2/p} \cap \mathrm{L}^1(\mathbb{R}^n, d\mu)$, then*

$$\frac{d}{dt}\left(\|\nabla \rho^{1/p}\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \frac{\lambda_\star}{2-p}\left(\|\rho^{1/p}\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \|\rho\|^{2/p}_{\mathrm{L}^1(\mathbb{R}^n,d\mu)}\right)\right) \leq 0 \quad \forall\, t > 0\,.$$

The result of Theorem 90 is new for $m \neq 1$. Theorem 86 corresponds to the special case of the harmonic potential $\phi(y) = \frac{1}{2}|y|^2$ with $v = \rho^{1/p}$ in Theorem 90. The range of the admissible parameters (m,p) is the same in Theorem 90 and shown in Fig. 6.2. With the additional observation that $\rho(t,\cdot)$ converges to a constant as $t \to +\infty$ so that the limit of the deficit is 0, the monotonicity of the deficit of Theorem 90 provides us with a proof of (6.12).

Proof of Theorem 90. In order to do computations, a very convenient reformulation is obtained with the flow

$$\frac{\partial w}{\partial t} = w^{2-2\beta}\left(\mathcal{L} w + \kappa\,\frac{|\nabla w|^2}{w}\right) \tag{6.14}$$

for any $t \geq 0$, with initial datum $w(t=0,\cdot) = |f|^{1/\beta}$, where

$$\kappa := \beta\,(p-2)+1 \quad \text{and} \quad \beta = \frac{2}{2-p\,(1-m)}\,.$$

A first computation shows that $\int_{\mathbb{R}^n} w^{\beta p}\,d\mu = \int_{\mathbb{R}^n} |f|^p\,d\mu$ is independent of t because

$$\frac{d}{dt}\int_{\mathbb{R}^n} w^{\beta p}\,d\mu = \beta\,p\int_{\mathbb{R}^n} w^\kappa\left(\mathcal{L} w + \kappa\,\frac{|\nabla w|^2}{w}\right)d\mu = 0\,.$$

A second useful computation goes as follows:

$$\begin{aligned}
&-\frac{1}{2\,\beta^2}\,\frac{d}{dt}\int_{\mathbb{R}^n}\left(\left|\nabla w^\beta\right|^2 + \frac{\lambda_\star}{p-2}\,w^{2\beta}\right)d\mu\\
&= \int_{\mathbb{R}^n}\left(\mathcal{L} w + (\beta-1)\,\frac{|\nabla w|^2}{w} - \frac{\lambda_\star\,w}{\beta\,(p-2)}\right)\left(\mathcal{L} w + \kappa\,\frac{|\nabla w|^2}{w}\right)d\mu\\
&= \int_{\mathbb{R}^n}(\mathcal{L} w)^2\,d\mu + (\kappa+\beta-1)\int_{\mathbb{R}^n}(\mathcal{L} w)\,\frac{|\nabla w|^2}{w}\,d\mu + \kappa\,(\beta-1)\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu - \lambda_\star\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu\,.
\end{aligned}$$

The purely technical purpose of introducing the exponent β is to make the last line of the above computation 2-homogeneous in w, which makes the discussion easier to read. Inserting the two following estimates,

$$\begin{aligned}
\int_{\mathbb{R}^n}(\mathcal{L} w)^2\,d\mu = -\int_{\mathbb{R}^n}\nabla w\cdot\nabla(\mathcal{L} w)\,d\mu &= -\int_{\mathbb{R}^n}\nabla w\cdot(\mathcal{L}\nabla w)\,d\mu + \int_{\mathbb{R}^n}\nabla w\cdot[\mathcal{L},\nabla]\,w\,d\mu\\
&= \int_{\mathbb{R}^n}\|\mathrm{Hess}\,w\|^2\,d\mu + \int_{\mathbb{R}^n}\nabla w\cdot[\mathcal{L},\nabla]\,w\,d\mu\\
&= \int_{\mathbb{R}^n}\|\mathrm{Hess}\,w\|^2\,d\mu + \int_{\mathbb{R}^n}\mathrm{Hess}\,\phi : \nabla w\otimes\nabla w\,d\mu\\
&\geq \int_{\mathbb{R}^n}\|\mathrm{Hess}\,w\|^2\,d\mu + \lambda_\star\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu
\end{aligned}$$

and
$$\int_{\mathbb{R}^n} (\mathcal{L}w)\,\frac{|\nabla w|^2}{w}\,d\mu = -2\int_{\mathbb{R}^n} \mathrm{Hess}\,w : \frac{\nabla w\otimes\nabla w}{w}\,d\mu + \int_{\mathbb{R}^n} \frac{|\nabla w|^4}{w^2}\,d\mu\,,$$
we obtain that
$$\frac{d}{dt}\int_{\mathbb{R}^n}\left(\left|\nabla w^\beta\right|^2 + \frac{\lambda_\star}{p-2}\,w^{2\beta}\right)d\mu \le 0$$
if, for any function w, we have
$$\mathcal{Q}_\beta[w] := \int_{\mathbb{R}^n} \|\mathrm{Hess}\,w\|^2\,d\mu - 2\,(\kappa+\beta-1)\int_{\mathbb{R}^n}\mathrm{Hess}\,w : \frac{\nabla w\otimes\nabla w}{w}\,d\mu \\ + \big(\kappa\,(\beta-1)+\kappa+\beta-1\big)\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu \ge 0\,.$$
A sufficient condition is obtained if the reduced discriminant is negative, that is, if
$$(\kappa+\beta-1)^2 - \big(\kappa\,(\beta-1)+\kappa+\beta-1\big) \le 0\,.$$
Altogether, this gives the condition
$$\beta_-(p) \le \beta \le \beta_+(p) \quad \text{with} \quad \beta_\pm(p) := \frac{1\pm\sqrt{(p-1)\,(2-p)}}{1-(p-1)\,(2-p)}\,. \tag{6.15}$$
See Fig. 6.3.

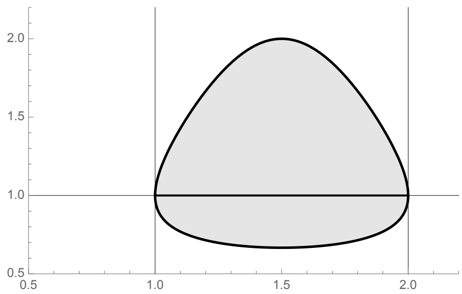

Figure 6.3: The admissible parameters p and β correspond to the grey area and are independent of the dimension n. The boundary of the admissible set is tangent to the vertical lines $p=1$ at $(\beta,p)=(1,1)$ and $p=2$ at $(\beta,p)=(1,2)$. This figure corresponds to Fig. 6.2 up to the transformation of $m\to\beta=2/\big(2-p\,(1-m)\big)$.

Equivalently, written in terms of m, the condition is $m_-(p)\le m\le m_+(p)$ with $m_\pm(p)$ defined by (6.5). Summarizing our computations, we learn that
$$\frac{d}{dt}\left(\|\nabla w\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \frac{\lambda_\star}{p-2}\left(\|w\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)} - \|w\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}\right)\right) = -2\,\beta^2\,\mathcal{Q}_\beta[w] \le 0$$
because $\mathcal{Q}_\beta[w]$ is non-negative under Condition (6.15).

To make these computations rigorous, one has to justify all integrations by parts. This is by now rather standard and can be done using the following scheme.

1. Let $h>0$ be large enough and consider $\Omega_h := \{x\in\mathbb{R}^n : \phi(x) < h\}$. We can consider the evolution equation restricted to Ω_h, with no flux boundary conditions. Then apply the *carré du champ* method and keep track of the boundary terms. Since Ω_h is a bounded convex domain, these terms have a sign due to Grisvard's lemma: see for instance [148, Lemma 5.2], [197, Proposition 4.2], [154, 128] and [178,

Lemma A.3].

2. Extend Inequality (6.12) written on Ω_h to $\mathbb{R}^d$ by taking the limit as $h \to +\infty$ and then argue by density.

This completes the proof of Theorem 90. □

Remark 91. *For any $p \in [1,2)$, if the condition $m_-(p) \le m \le m_+(p)$ is not satisfied, then one can find a positive initial datum such that the solution v of (6.6) is such that*

$$\frac{d}{dt}\left(\|\nabla v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \frac{d}{p-2}\left(\|v\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)} - \|v\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}\right)\right)_{|t=0} > 0\,.$$

See [122] for a similar statement on the sphere and its proof.

6.3.2 Improved inequalities based on the *carré du champ* method

In the proof of Theorem 90, using only $\mathcal{Q}_\beta[w] \ge 0$ is a crude estimate. Let us explain how one can obtain improved estimates by making a better use of $\mathcal{Q}_\beta[w] \ge 0$. Under Condition (6.15), we can indeed rewrite $\mathcal{Q}_\beta[w]$ as an integral of a sum of squares,

$$\mathcal{Q}_\beta[w] = \int_{\mathbb{R}^n}\left\|\mathrm{Hess}\, w - (\kappa+\beta-1)\,\frac{\nabla w\otimes\nabla w}{w}\right\|^2 d\mu + \delta\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu \tag{6.16}$$

with

$$\delta := \kappa\,(\beta-1)+\kappa+\beta-1-(\kappa+\beta-1)^2 = \big(\beta-\beta_-(p)\big)\big(\beta_+(p)-\beta\big) > 0\,. \tag{6.17}$$

As in [15, Theorem 2], let us consider the special case $m = \beta = 1$ of the linear flow. Let us define the *entropy* and the *Fisher information* by

$$\mathcal{E}[w] := \frac{1}{2-p}\left(\|w\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \|w\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right) \quad\text{and}\quad \mathcal{I}[w] := \|\nabla w\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}\,.$$

Inequality (6.12) amounts simply to

$$\mathcal{I}[w] - \lambda_\star\,\mathcal{E}[w] \ge 0\,.$$

We can now state a first *improved entropy–entropy production inequality.*

Proposition 92. *Let n be any positive integer. We consider the measure $d\mu$ as in (6.9) such that (6.11) holds for some $\lambda_\star > 0$. For any $p \in (1,2)$, let $\varphi(s) := 1+s-(1+s)^{p-1}$. With the above notations, we have*

$$\mathcal{I}[f] \ge \frac{\lambda_\star}{(2-p)^2}\,\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\,\varphi\!\left(\frac{(2-p)\,\mathcal{E}[f]}{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}}\right) \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^n,d\mu)\,. \tag{6.18}$$

Since $p \in (1,2)$, the function φ is convex with $\varphi(0) = 0$ and $\varphi'(0) = 2-p$, which implies in particular that $\varphi(s) \ge (2-p)\,s$ for any $s \ge 0$, so that Inequality (6.18) is stronger than (6.12). Inequality (6.18) amounts to

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} \ge \frac{\lambda_\star}{(2-p)^2}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \|f\|^{2(p-1)}_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}\,\|f\|^{2(2-p)}_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right)$$

and can also be rewritten as

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \frac{\lambda_\star}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right)$$
$$\geq \frac{\lambda_\star}{(2-p)^2}\left((p-1)\,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)} + (2-p)\,\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)} - \|f\|^{2(p-1)}_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}\,\|f\|^{2(2-p)}_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right). \tag{6.19}$$

We claim no originality in either Proposition 92 or Inequality (6.19) and refer to [15, Theorem 2] and [14, Ineq. (3.3)] for earlier results. Also see [118, Theorem 2.1] for an improved inequality like (6.19) in the case of the sphere. However, let us give a proof for completeness.

Proof. With $\beta = 1$, notice that (6.16) holds with $\delta = (2-p)\,(p-1)$. Using the Cauchy-Schwarz inequality, we obtain

$$\big(\mathcal{I}[w]\big)^2 = \left(\int_{\mathbb{R}^n} |\nabla w|^2\,d\mu\right)^2 \leq \int_{\mathbb{R}^n} |w|^2\,d\mu \int_{\mathbb{R}^n} \frac{|\nabla w|^4}{w^2}\,d\mu$$

and, with $M := \left(\int_{\mathbb{R}^n} |w|^p\,d\mu\right)^{2/p}$, we can also write that

$$\int_{\mathbb{R}^n} |w|^2\,d\mu = (2-p)\,\mathcal{E}[w] + M\,.$$

Altogether, with $\mathsf{e}(t) := (2-p)\,M^{-1}\,\mathcal{E}[w(t,\cdot)]$, if w solves (6.14) with $\beta = 1$, then we have the differential inequality

$$\mathsf{e}'' + 2\,\lambda_\star\,\mathsf{e}' - (p-1)\,\frac{(\mathsf{e}')^2}{1+\mathsf{e}} \geq 0\,.$$

We claim that $(2-p)\,\mathsf{e}' + 2\,\lambda_\star\left(1+\mathsf{e}-(1+\mathsf{e})^{p-1}\right) \leq 0$, which follows from the observation that the equation

$$y'' + \mathsf{a}\,y' - \mathsf{b}\,\frac{(y')^2}{1+y} = 0$$

can be solved using the ansatz $y' = \frac{\mathsf{a}}{\mathsf{b}-1}\,\varphi(y)$ with $\varphi(0) = 0$ if

$$\varphi' - \mathsf{b}\,\frac{\varphi}{1+s} = 1-\mathsf{b}\,. \tag{6.20}$$

It is straightforward to check that the unique solution is $\varphi(s) = 1+s-(1+s)^{\mathsf{b}}$. With $\mathsf{a} = 2\,\lambda_\star$ and $\mathsf{b} = p-1$, we obtain

$$\left(\mathsf{e}' + \frac{2\,\lambda_\star}{2-p}\,\varphi(\mathsf{e})\right)' \geq (p-1)\,\frac{\mathsf{e}'}{1+\mathsf{e}}\left(\mathsf{e}' + \frac{2\,\lambda_\star}{2-p}\,\varphi(\mathsf{e})\right).$$

By integrating from $t = 0$ to $+\infty$ using the facts that $\mathsf{e}' \leq 0$, $\mathsf{e}' + \frac{2\,\lambda_\star}{2-p}\,\mathsf{e} \leq 0$ and

$$\lim_{t\to+\infty} \mathsf{e}'(t) = \lim_{t\to+\infty} \mathsf{e}(t) = 0,$$

we conclude with $M = \|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}$ that

$$\frac{2-p}{M}\left(\mathcal{I}[w] - \frac{\lambda_\star\,M}{(2-p)^2}\,\varphi\left(\frac{(2-p)\,\mathcal{E}[w]}{M}\right)\right) = -\frac{1}{2}\left(\mathsf{e}' + \frac{2\,\lambda_\star}{2-p}\,\varphi(\mathsf{e})\right) \geq 0 \quad \forall\, t \geq 0$$

See [P2, Appendix B.4] for a more detailed proof in the similar case of the sphere. □

Inspired once more by the results on the sphere (see [118] and earlier references therein), it is a very natural question to wonder if *improved entropy–entropy production inequalities* can be achieved with $\beta \neq 1$.

The answer goes as follows. Let us consider the function φ_β given by

$$\chi_\beta(s)=\big(1-\mathsf{b}(\beta)\big)\big(1+s-(1+s)^{\mathsf{b}(\beta)}\big)\quad\text{where}\quad \mathsf{b}(\beta)=\frac{\delta(\beta)}{\beta^2}\,\frac{2-p}{\lambda_\star}$$

and $\delta(\beta)$ is defined by (6.17).

Proposition 93. *Let n be any positive integer. We consider the measure $d\mu$ as in (6.9) such that (6.11) holds for some $\lambda_\star>0$. For any $p\in(1,2)$, take $\beta\in\big[1,\beta_+(p)\big]$. With the same notations as in Proposition 92 and χ_β as above, we have*

$$\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\,\chi_\beta\left(\frac{(2-p)\,\mathcal{E}[f]}{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}}\right)\le\mathcal{I}[f]\quad\forall\,f\in\mathrm{H}^1(\mathbb{R}^n,d\mu)\,. \tag{6.21}$$

Proof. With the notations

$$\mathsf{e}:=\frac{\|w^\beta\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}}{\|w^\beta\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}}\quad\text{and}\quad \mathsf{i}:=\frac{\|\nabla(w^\beta)\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}}{\|w^\beta\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}}\,,$$

we have now to consider the differential inequality

$$\left(\mathsf{i}'-\frac{\lambda_\star}{2-p}\,\mathsf{e}\right)'\le-\delta\,\beta^2\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\,.$$

The key ingredient is to replace an estimate due to Demange in the case of the sphere for $p>2$ (see [112, 113] and also [Lemma 15, (iii)]) by its counterpart for log-concave measures and $p\in(1,2)$. Compared to the linear case, the Cauchy-Schwarz inequality in the proof of Proposition 92 has to be replaced by the Hölder inequalities

$$\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu=\int_{\mathbb{R}^n}\left(\frac{|\nabla w|^2}{w}\,w\,1\right)d\mu\le\left(\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\right)^{\frac12}\left(\int_{\mathbb{R}^n}|w|^{2\beta}\,d\mu\right)^{\frac{1}{2\beta}}1^{\frac{\beta-1}{2\beta}}\,,$$
$$\frac{1}{\beta^2}\int_{\mathbb{R}^n}|\nabla w^\beta|^2\,d\mu=\int_{\mathbb{R}^n}\left(\frac{|\nabla w|^2}{w}\,w^{2\beta-1}\,1\right)d\mu\le\left(\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\right)^{\frac12}\left(\int_{\mathbb{R}^n}|w|^{2\beta}\,d\mu\right)^{\frac{2\beta-1}{2\beta}}1^{\frac{1}{2\beta}}\,,$$

after observing that $\beta\ge\beta_-(p)\ge\beta_-(3/2)=2/3>1/2$, so that

$$\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\ge\frac{\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu\int_{\mathbb{R}^n}|\nabla w^\beta|^2\,d\mu}{\beta^2\int_{\mathbb{R}^n}w^{2\beta}\,d\mu}=-\frac{\mathsf{i}\,\mathsf{e}'}{\beta^4\,(1+\mathsf{e})}\,.$$

Hence

$$\left(\mathsf{i}-\frac{\lambda_\star}{2-p}\,\mathsf{e}\right)'\le\frac{\delta}{\beta^2}\,\frac{\mathsf{i}\,\mathsf{e}'}{1+\mathsf{e}}\,.$$

Let us compute

$$\left(\mathsf{i}-\frac{\lambda_\star}{2-p}\,\chi(\mathsf{e})\right)'=\left(\mathsf{i}-\frac{\lambda_\star}{2-p}\,\mathsf{e}\right)'+\frac{\lambda_\star}{2-p}\,(\mathsf{e}-\chi(\mathsf{e}))'\le\frac{\delta}{\beta^2}\,\frac{\mathsf{e}'}{1+\mathsf{e}}\left(\mathsf{i}-\frac{\lambda_\star}{2-p}\,\chi(\mathsf{e})\right)$$

on the condition that χ solves

$$\chi'(s)=1+\frac{\delta}{\beta^2}\,\frac{2-p}{\lambda_\star}\,\frac{\chi(s)}{1+s}\,,\quad\chi(0)=0\,.$$

The solution $\chi(s)=(1-\mathsf{b})\,\varphi(s)$ is such that φ solves (6.20) with $\mathsf{b}=\delta\,\beta^{-2}\,(2-p)/\lambda_\star$, which shows that $\chi=\chi_\beta$. The proof then follows from the same considerations as for Proposition 92 (also see [P2, Appendix B.4] for details). □

If $\phi(y)=|y|^2/2$ is the harmonic potential so that $d\mu=d\gamma$, by testing (6.2) with $f_\varepsilon(y)=1+\varepsilon\,y_1$ where y_1 denotes the first coordinate of $y=(y_1,y_2,\ldots,y_n)\in\mathbb{R}^n$, we find that

$$\|\nabla f_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\frac{1}{2-p}\left(\|f_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\|f_\varepsilon\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right)=\frac{1}{2}\,(p-1)\,\varepsilon^4+O(\varepsilon^5)\quad\text{as}\quad\varepsilon\to0\,. \tag{6.22}$$

This is the standard computation for checking that $\lambda=\lambda_\star=1$ is the optimal constant in (6.2). Since

$$\mathcal{I}[f_\varepsilon]-\frac{1}{(2-p)^2}\,\|f_\varepsilon\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\,\varphi\left(\frac{(2-p)\,\mathcal{E}[f_\varepsilon]}{\|f_\varepsilon\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}}\right)=\frac{1}{2}\,(p-1)^2\,\varepsilon^4+O(\varepsilon^5)\quad\text{as}\quad\varepsilon\to0\,,$$

we also learn that (6.18) involves the optimal exponent at least in the limit as $\varepsilon\to0$. After observing that $\|\nabla f_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}=\varepsilon^2$, we may wonder whether the *deficit*

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\frac{1}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right)$$

measures the distance in terms of $\|\nabla f\|^4_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}$. The detailed answer is not limited to the case $\mu=\gamma$ and goes as follows. For simplicity, we take $\beta=1$ and consider $\varphi(s):=1+s-(1+s)^{p-1}$ as in Proposition 92. We recall that φ is monotone increasing and convex on $\mathbb{R}^+$, such that $\varphi'(0)=2-p$, hence invertible of inverse φ^{-1} such that $\psi(t):=t-(2-p)\,\varphi^{-1}(t)$ is also a convex, non-negative, monotone increasing function.

Corollary 94. *Let $p\in(1,2)$ and n be a positive integer. We consider the measure $d\mu$ as in (6.9) such that (6.11) holds for some $\lambda_\star>0$. With ψ as above, for any $f\in\mathrm{H}^1(d\gamma)$ we have*

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}-\frac{\lambda_\star}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}-\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right)\ge\frac{\lambda_\star}{2-p}\,\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\,\psi\left(\frac{2-p}{\lambda_\star}\,\frac{\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}}{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}}\right).$$

Moreover, there is some $\kappa>0$ such that

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}-\frac{\lambda_\star}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}-\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}\right)\ge\frac{\kappa\,\|\nabla f\|^4_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}}{\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}+\frac{\lambda_\star}{2-p}\,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}}\,.$$

The constant κ depends only on p and its value is estimated in the proof below.

Proof. Let $M=\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}$. We deduce from Proposition 92 that

$$\mathsf{i}:=\frac{(2-p)^2}{\lambda_\star\,M}\,\mathcal{I}[f]\ge\varphi(\mathsf{e})\quad\text{where}\quad\mathsf{e}:=\frac{(2-p)\,\mathcal{E}[f]}{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\mu)}}\,,$$

which is equivalent to $-\mathsf{e}\ge-\varphi^{-1}(\mathsf{i})$, so that

$$\mathsf{i}-(2-p)\,\mathsf{e}\ge\mathsf{i}-(2-p)\,\varphi^{-1}(\mathsf{i})=\psi(\mathsf{i})$$

and, as a consequence,

$$\mathcal{I}[f]-\lambda_\star\,\mathcal{E}[f]=\frac{\lambda_\star\,M}{(2-p)^2}\left(\mathsf{i}-(2-p)\,\mathsf{e}\right)\ge\frac{\lambda_\star\,M}{(2-p)^2}\,\psi(\mathsf{i})=\frac{\lambda_\star\,M}{(2-p)^2}\,\psi\left(\frac{2-p}{\lambda_\star\,M}\,\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}\right).$$

Since $\varphi(s) \sim s$ as $s \to +\infty$, we deduce that $\psi(t) \sim (p-1)\,t$ as $t \to +\infty$. On the other hand, since

$$\psi''(t) = (2-p)\,\frac{\varphi''(s)}{\big(\varphi'(s)\big)^3} \quad \text{with} \quad s = \varphi^{-1}(t)\,,$$

we learn that $\psi''(0) = (p-1)/(2-p) > 0$ and $t \mapsto \psi''(t)$ is non-increasing because $\varphi'(s)^5\,\psi'''(t) = \varphi'(s)\,\varphi'''(s) - \varphi''(s)^2 < 0$. This allows us to define

$$\kappa := \inf_{t>0} t^{-2}\,(1+t)\,\psi(t)\,.$$

Using $\psi(0) = \psi'(0) = 0$, we know that $\kappa > 0$ and $\psi(t) \ge \kappa\,t^2/(1+t)$ concludes the proof using $\|f\|_{\mathrm{L}^p(\mathbb{R}^n,d\mu)} \le \|f\|_{\mathrm{L}^2(\mathbb{R}^n,d\mu)}$. □

Remark 95. *If $\phi(y) = |y|^2/2$ is the harmonic potential so that $d\mu = d\gamma$, then $\lambda = \lambda_\star = 1$ is the optimal constant in (6.2) and the results of Proposition 92 and Corollary 94 are both stability estimates. Even in the general case of a measure $d\mu$ as in (6.9) such that (6.11) holds for some $\lambda_\star > 0$, Proposition 92 and Corollary 94 provide improvements to the basic inequality with a (generically non-optimal) constant $\lambda_\star$.*

6.3.3 From the *carré du champ* method to Obata's theorem

As a side result, we consider an improvement of the *carré du champ* method as in [195, Theorem 2.1] or [121], which goes as follows. Let us consider the optimal constant $\lambda_1 > 0$ in the Poincaré inequality

$$\int_{\mathbb{R}^n} |\nabla w|^2\,d\mu \ge \lambda_1 \int_{\mathbb{R}^n} |w - \bar{w}|^2\,d\mu \quad \forall\, w \in \mathrm{H}^1(\mathbb{R}^n, d\mu)\,, \tag{6.23}$$

where $\bar{w} := \int_{\mathbb{R}^n} w\,d\mu$. By expanding $\int_{\mathbb{R}^n} \big(\mathcal{L}w + \lambda_1\,(w - \bar{w})\big)^2\,d\mu \ge 0$, we obtain

$$\int_{\mathbb{R}^n} (\mathcal{L}w)^2\,d\mu \ge \lambda_1 \int_{\mathbb{R}^n} |\nabla w|^2\,d\mu\,.$$

On the other hand, by the computation of Section 6.3.1 with $\beta = p = 1$, we know that

$$-\frac{1}{2}\,\frac{d}{dt} \int_{\mathbb{R}^n} \left(|\nabla w|^2 + \frac{\lambda_\star}{p-2}\,w^2\right) d\mu = \int_{\mathbb{R}^n} (\mathcal{L}w)^2\,d\mu - \lambda_\star \int_{\mathbb{R}^n} |\nabla w|^2\,d\mu = \mathcal{Q}_1[w] \ge 0\,,$$

which proves that

$$\lambda_1 \ge \lambda_\star\,.$$

Lemma 96. *Assume that $n \ge 1$, $p \in [1,2)$ and consider the measure $d\mu$ as in (6.9) such that (6.11) holds for some $\lambda_\star > 0$. Then (6.12) holds with*

$$\lambda = (2-p)\,\lambda_1 + (p-1)\,\lambda_\star\,. \tag{6.24}$$

As a consequence, we have $\lambda \ge \lambda_\star$ with equality for the optimal value of λ in (6.12) if and only if $\lambda_\star = \lambda_1$ and $\phi(y) = \lambda_\star\,|y - y_0|^2/2$ for some $y_0 \in \mathbb{R}^n$.

Proof. The *carré du champ* method applied as in Section 6.3.1 with $p = 1$ shows that $\lambda_1 \ge \lambda_\star$. Coming back to the computations of Section 6.3.1, we can rearrange the integrals in the expression $\mathcal{Q}_\beta[w]$ differently and get

$$\begin{aligned}
&-\frac{1}{2\beta^2}\frac{d}{dt}\int_{\mathbb{R}^n}\left(\left|\nabla w^\beta\right|^2+\frac{\lambda}{p-2}\,w^{2\beta}\right)d\mu\\
&\quad=(1-\theta)\int_{\mathbb{R}^n}(\mathcal{L}w)^2\,d\mu-\lambda\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu\\
&\qquad+\theta\int_{\mathbb{R}^n}(\mathcal{L}w)^2\,d\mu+(\kappa+\beta-1)\int_{\mathbb{R}^n}(\mathcal{L}w)\,\frac{|\nabla w|^2}{w}\,d\mu+\kappa\,(\beta-1)\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\\
&\quad=\big((1-\theta)\,\lambda_1+\theta\,\lambda_\star-\lambda\big)\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu\\
&\qquad+\theta\left(\int_{\mathbb{R}^n}\|\mathrm{Hess}\,w\|^2\,d\mu+\lambda_\star\int_{\mathbb{R}^n}|\nabla w|^2\,d\mu\right)\\
&\qquad-2\,(\kappa+\beta-1)\int_{\mathbb{R}^n}\mathrm{Hess}\,w:\frac{\nabla w\otimes\nabla w}{w}\,d\mu+\big(\kappa\,(\beta-1)+\kappa+\beta-1\big)\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\,.
\end{aligned}$$

With the choice of θ such that

$$(\kappa+\beta-1)^2-\theta\,\big(\kappa\,(\beta-1)+\kappa+\beta-1\big)=0\,,$$

which means $\theta=\theta(\beta)$ with

$$\theta(\beta):=\frac{(p-1)^2\,\beta^2}{(p-2)\,\beta^2+2\,\beta-1}\,,$$

we can write that

$$\begin{aligned}
\theta\int_{\mathbb{R}^n}\|\mathrm{Hess}\,w\|^2\,d\mu-2\,(\kappa+\beta-1)\int_{\mathbb{R}^n}\mathrm{Hess}\,w:\frac{\nabla w\otimes\nabla w}{w}\,d\mu+\big(\kappa\,(\beta-1)+\kappa+\beta-1\big)\int_{\mathbb{R}^n}\frac{|\nabla w|^4}{w^2}\,d\mu\\
=\theta\int_{\mathbb{R}^n}\left\|\mathrm{Hess}\,w-\frac{\beta\,(p-1)}{\theta}\,\frac{\nabla w\otimes\nabla w}{w}\right\|^2\,d\mu\ge0\,.
\end{aligned}$$

Altogether, we have shown that

$$\frac{d}{dt}\int_{\mathbb{R}^n}\left(\left|\nabla w^\beta\right|^2+\frac{\lambda}{p-2}\,w^{2\beta}\right)d\mu\le0$$

if $\lambda=(1-\theta)\,\lambda_1+\theta\,\lambda_\star$. Notice that $\theta\big(\beta_\pm(p)\big)=1$ for any $p\in[1,2)$. The observation that

$$\min_{\beta\in[\beta_-(p),\beta_+(p)]}\theta(\beta)=\theta(1)=p-1$$

completes the proof of (6.24).

With $\beta=1$ and $p=1$, the computation in the proof of Theorem 90 shows that, if the initial datum $w(t=0,\cdot)$ is an optimal function for the Poincaré inequality (6.23), then

$$0=\frac{d}{dt}\int_{\mathbb{R}^n}\big(|\nabla w|^2-\lambda_\star\,w^2\big)\,d\mu=-2\left(\mathcal{Q}_1[w]+\int_{\mathbb{R}^n}(\mathrm{Hess}\,\phi-\lambda_\star\mathrm{Id}):\nabla w\otimes\nabla w\,d\mu\right)$$

at $t=0$. Here we keep all terms and in particular do not use the fact that $\mathrm{Hess}\,\phi-\lambda_\star\mathrm{Id}\ge0$ a.e. in the sense of positive matrices. Since $\nabla w\neq0$ a.e. and $\mathcal{Q}_1[w]\ge0$, we find that $\mathrm{Hess}\,\phi-\lambda_\star\mathrm{Id}=0$ a.e. This completes the proof in the equality case $\lambda_\star=\lambda_1$. □

Remark 97. *The proof of Lemma 96 is reminiscent of [195, 121]. The result when $\lambda_\star=\lambda_1$ points in the direction of* Obata's theorem *(also known as the* ObataLichnerowicz theorem*) and in some sense, it is the analogue for Gaussian measures of the result of [194, p. 135] (also see for instance [43, p. 179]) on the sphere. The case $\lambda_\star=\lambda_1$ is easy to understand in dimension $d=1$: with $\beta=1$ and $p=1$, we apply the computation of the proof of Lemma 96 to a function u in the eigenspace associated with λ_1 and obtain that $u''=0$ almost everywhere.*

This means that $u(y) = a\,y + b$ for some real constants $a \neq 0$ and b, and there is no loss of generality if we take $a = 1$. Using now the eigenvalue equation $\mathcal{L}u + \lambda_1\,u = 0$, we read that $\phi'(y) = \lambda_1\,(y - b)$, which means that ϕ is an harmonic potential. In higher dimensions, one has to remember that Inequality (6.12) can be tensorized on product spaces: see for instance [187, 91, 126]. This is however responsible for some technicalities, which are dealt with in greatest generality, e.g., *in [150].*

6.3.4 Improved inequalities under orthogonality conditions

Let Π_1 be the $\mathrm{L}^2(\mathbb{R}^n, d\gamma)$orthogonal projection onto the space generated by the constants and the coordinate functions, corresponding to the Hermite polynomials of order less or equal than 1. The following result was recently proved in [P2, Appendix A] on the basis of Nelson's hypercontractivity estimate in [220, Theorem 3] and its relation with Gross' logarithmic Sobolev inequality in [155] (also see [187, 14] for earlier results).

Proposition 98. *Let $n \geq 1$ and $p \in [1,2)$. For any $f \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$, we have*

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{p-2}\left(\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} - \|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\right) \geq \frac{1}{2}\,(2-p)\,\|\nabla(\mathrm{Id} - \Pi_1)\,f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\,.$$

Compared to (6.2), this result provides us with an *improved entropy–entropy production inequalities* under orthogonality conditions. As noted in [P2], such an improvement is not optimal. There are other possible approaches. For instance, a finer analysis of entropy methods has been used in [122] on the sphere, that could probably be adapted to the case of the Gaussian measure. Alternatively, the convex interpolation of [187], with the possible advantage that the result would not degenerate in the limit as $p \to 2$ using the recent stability result of [119, Theorem 2].

6.4 Stability results for the Gaussian measure in the subcritical range

The whole Section is devoted to the proof of Theorem 87. We split it in four lemmas. The key estimate is obtained in Lemma 102.

① Let us start with the easy case, far away from the optimizers of (6.2) in the sense that for some $\theta > 0$, we assume

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \geq \theta\,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\,. \tag{6.25}$$

By homogenity of the inequalities, we can fix $\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = 1$ without loss of generality.

Lemma 99. *Let $n \geq 1$ and $\theta \in (0,1)$. For any function $f \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$ such that $\|f\|_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = 1$ and $\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \geq \theta$, we have the estimate*

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right) \geq \kappa_\star(\theta)\,\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}$$

In case (6.25), this already proves the result of Theorem 87 with $c_{n,p} \leq c^{(1)}_{n,p} := \kappa_\star(\theta)$.

Proof. From Corollary 94 and its proof applied with $\lambda_\star = 1$, we obtain

$$\begin{aligned}\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right)\\ \ge \frac{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}}{2-p}\,\frac{1}{2}\,\psi\big((2-p)\,\theta\big)\left((2-p)\,\frac{\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}}\right)^2\\ \ge \frac{1}{2}\,(2-p)\,\psi\big((2-p)\,\theta\big)\,\frac{\|\nabla f\|^4_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}} \ge \kappa_\star(\theta)\,\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\end{aligned}$$

with $\kappa_\star(\theta) = \frac{1}{2}\,(2-p)\,\psi\big((2-p)\,\theta\big)\,\theta$ because $t \mapsto \psi''(t)$ is non-increasing, $\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \ge \theta$ and $\|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} \le \|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}$. □

② From now on we work in a neighbourhood of the constants which, by homogeneity of the inequalities, is defined as

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \le \theta\,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\,. \tag{6.26}$$

With $\theta > 0$ small, we claim that $\int_{\mathbb{R}^n} f\,d\gamma$ is close to 1 if $\|f\|_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = 1$.

Lemma 100. *Let $n \ge 1$ and $\theta \in (0,1)$. For any function*

$$f \in \mathrm{H}^1(\mathbb{R}^n,d\gamma) \quad \textit{such that} \quad \|f\|_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = 1 \quad \textit{and} \quad \|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \le \theta\,, \tag{6.27}$$

we have the estimate

$$\sqrt{1-\theta} \le \int_{\mathbb{R}^n} f\,d\gamma \le 1\,.$$

Proof. With $\bar f := \int_{\mathbb{R}^n} f\,d\gamma$, the result follows from the Gaussian Poincaré inequality according to

$$1 = \|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = \int_{\mathbb{R}^n} |f-\bar f|^2\,d\gamma + \bar f^2 \le \|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} + \bar f^2 \le \theta + \bar f^2\,.$$

□

③ Assume that f is as in (6.27) and let us decompose $u_f(x) := f(x)/\int_{\mathbb{R}^n} f\,d\gamma$ as

$$u_f(x) = 1 + \varepsilon\,x\cdot\nu + \eta\,r(x)$$

where $\nu \in \mathbb{S}^{n-1}$ is such that $\varepsilon\,\nu = \int_{\mathbb{R}^n} x\,u_f(x)\,d\gamma$ with $\varepsilon > 0$, η is a positive number and r is a function in $\mathrm{H}^1(\mathbb{R}^n,d\gamma) \cap (\mathrm{Id}-\Pi_1)\mathrm{L}^2(\mathbb{R}^n,d\gamma)$ such that $\|\nabla r\|_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = 1$ and $\|r\|_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \le 1/2$ by the Gaussian Poincaré inequality after taking into account the additional orthogonality condition $\int_{\mathbb{R}^n} r\,x_i\,d\gamma = 0$ for any $i = 1, 2, \ldots, n$.

Lemma 101. *Let $n \ge 1$ and $\theta \in (0,1)$. Let $f \in \mathrm{H}^1(\mathbb{R}^n,d\gamma)$ be such that (6.27) holds. With the above notations, we have*

$$\|u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = 1 + \varepsilon^2 + \eta^2\,\|r\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \le 1+\theta \quad \textit{and} \quad \|\nabla u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} = \varepsilon^2 + \eta^2 \le \frac{\theta}{1-\theta}$$

and, if $\eta > t\,\varepsilon^2$ for some $t > 0$, then

$$\|\nabla u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2-p}\left(\|u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right) \ge \frac{1}{4}\,(2-p)\left(\eta^2 + \frac{t^2\,\varepsilon^4}{1+\varepsilon^2+\eta^2}\right).$$

By homogeneity, if (6.27) holds and $\eta > t\,\varepsilon^2$ for some given $t>0$, using Lemma 100, we obtain

$$\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \frac{1}{2-p}\left(\|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \|f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right) \\ \geq \frac{1}{4}\,(2-p)\,(1-\theta)\left(\|\nabla(\mathrm{Id}-\Pi_1)\,f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} + \frac{t^2\,\|\nabla(\Pi_1 f)\|^4_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}{\|\nabla f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} + \|f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}}\right).$$

This case already covers the result of Theorem 87 with $c_{n,p} \leq c^{(2)}_{n,p} := \frac{1}{4}\,(2-p)\,(1-\theta)\min\{t,1\}$.

Proof. The result follows from Proposition 98 and from the chain of elementary inequalities

$$\eta^2 \geq \frac{1}{2}\left(\eta^2 + t^2\,\varepsilon^4\right) \geq \frac{1}{2}\left(\eta^2 + \frac{t^2\,\varepsilon^4}{1+\varepsilon^2+\eta^2}\right).$$

□

④ The next part of the proof relies on a Taylor expansion of $\|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}$. With no loss of generality, by rotational invariance, we can assume that $\nu = (1,0,\ldots,0)$ so that with Cartesian coordinates $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we write $u_f(x) = 1 + \varepsilon\,x_1 + \eta\,r(x)$. The following result is at the core of our strategy. It heavily relies on the Gaussian logarithmic Sobolev inequality and new estimates for the remainder terms based on the boundedness of $\int_{\mathbb{R}^n} r^2 \log r^2\,d\gamma$.

Lemma 102. *Let $n \geq 1$ and $f \in \mathrm{H}^1(\mathbb{R}^n, d\gamma)$ be a non-negative function such that (6.27) holds. We keep the same notations as in Lemma 101 and further assume that $\eta \leq t\,\varepsilon^2$ for some $t>0$. Then there is a constant $\mathcal{C}>0$, depending only on n, p and t, such that*

$$\|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)} \geq 1 + (p-1)\left(\varepsilon^2 + \eta^2\,\|r\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\right) \\ + (p-1)\,(2-p)\left(\frac{1}{2}\,\varepsilon^4 + \varepsilon^2\,\eta\int_{\mathbb{R}^n} x_1^2\,r(x)\,d\gamma\right) - \frac{\mathcal{C}\,\varepsilon^4}{\log\left(\frac{1}{\varepsilon}\right)} \quad \text{as} \quad \varepsilon \to 0_+\,.$$

Proof. This proof is elementary although a little bit lengthy. Let us split it into three steps.

Step 1. Let us start with a list of preliminary remarks.

• Let B_ε be the centred ball of radius $1/(2\,\varepsilon)$, that is,

$$B_\varepsilon := \left\{x \in \mathbb{R}^n : 2\,\varepsilon\,|x| < 1\right\}$$

and let $B_\varepsilon^c = \mathbb{R}^n \setminus B_\varepsilon$. We observe that

$$\gamma(B_\varepsilon^c) = \left|\mathbb{S}^{n-1}\right| \int_{1/(2\,\varepsilon)}^{+\infty} r^{n-1}\,e^{-\frac{r^2}{2}}\,dr = c_n\,\varepsilon^{2-n}\,e^{-\frac{1}{8\varepsilon^2}}\left(1 + O\left(\varepsilon^2\right)\right) \quad \text{as} \quad \varepsilon \to 0_+$$

with $c_n = 2^{3\,(2-n)/2}/\Gamma(n/2)$. Let $\xi_p := \sup_{\varepsilon\in(0,1/2)} \varepsilon^{-5}\left(\gamma(B_\varepsilon^c)\right)^{(2-p)/2}$. Hence we have

$$\int_{B_\varepsilon^c} |g|^p\,d\gamma \leq \|g\|^p_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\left(\gamma(B_\varepsilon^c)\right)^{(2-p)/2} \leq \xi_p\,\|g\|^p_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\,\varepsilon^5$$

for any $g \in \mathrm{L}^2(\mathbb{R}^n, d\gamma)$, by Hölder's inequality and, as a consequence,

$$\int_{B_\varepsilon} |g|^p\,d\gamma \leq \int_{\mathbb{R}^n} |g|^p\,d\gamma \leq \int_{B_\varepsilon} |g|^p\,d\gamma + \xi_p\,\|g\|^p_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}\,\varepsilon^5 \quad \forall\,\varepsilon \in (0,1/2)\,. \tag{6.28}$$

From now on, we assume without further notice that $x \in B_\varepsilon$ unless it is specified.

• An expansion in Taylor series of $(1+s)^p$ for $s \le 0$ shows that all terms are non-negative:

$$(1+s)^p \ge 1 + p\,s + \frac{1}{2}\,p\,(p-1)\,s^2 \quad \forall\, s \in (-1,0]\,.$$

Applied to $u_f = 1 + \varepsilon\,x_1 + \eta\,r$ whenever $1 + \varepsilon\,x_1 > 0$ and $r \le 0$, we obtain

$$\begin{aligned}
|u_f|^p = (1+\varepsilon\,x_1+\eta\,r)^p &= (1+\varepsilon\,x_1)^p \left(1 + \frac{\eta\,r}{1+\varepsilon\,x_1}\right)^p \\
&\ge (1+\varepsilon\,x_1)^p \left(1 + p\,\frac{\eta\,r}{1+\varepsilon\,x_1} + \frac{1}{2}\,p\,(p-1)\left(\frac{\eta\,r}{1+\varepsilon\,x_1}\right)^2\right) \\
&= (1+\varepsilon\,x_1)^p + p\,(1+\varepsilon\,x_1)^{p-1}\,\eta\,r + \frac{1}{2}\,p\,(p-1)\,(1+\varepsilon\,x_1)^{p-2}\,\eta^2\,r^2\,.
\end{aligned}$$

• Let us consider the case $1 + \varepsilon\,x_1 > 0$ and $r > 0$. The function

$$\rho(s) := \frac{1}{s^2}\left((1+s)^p - 1 - p\,s - \frac{1}{2}\,p\,(p-1)\,s^2\right) \quad \forall\, s \ge 0$$

is bounded. Let us extend ρ by 0 on $(-1,0)$ and define

$$\psi_{\varepsilon,\eta,r}(x) := (1+\varepsilon\,x_1)^{p-2}\,\rho\left(\frac{\eta\,r(x)}{1+\varepsilon\,x_1}\right).$$

With this definition, using $u_f \ge 0$ by hypothesis, we obtain

$$\begin{aligned}
|u_f|^p &\ge (1+\varepsilon\,x_1+\eta\,r)^p \\
&= (1+\varepsilon\,x_1)^p + p\,(1+\varepsilon\,x_1)^{p-1}\,\eta\,r + \frac{1}{2}\,p\,(p-1)\,(1+\varepsilon\,x_1)^{p-2}\,\eta^2\,r^2 + \frac{1}{2}\,p\,(p-1)\,\eta^2\,r(x)^2\,\psi_{\varepsilon,\eta,r}(x)
\end{aligned}$$

with equality whenever $r \ge 0$.

As $\eta \to 0_+$, $\psi_{\varepsilon,\eta,r}$ converges a.e. to 0 on B_ε uniformly with respect to $\varepsilon \in (0,1/2)$. The dominated convergence theorem is enough to conclude that

$$\lim_{\eta\to 0_+} \int_{B_\varepsilon} r(x)^2\,\psi_{\varepsilon,\eta,r}(x)\,d\gamma = 0$$

for a given function r, but this is not enough to conclude uniformly with respect to r. To do this, we need more detailed estimates. Notice however that

$$\|\psi_{\varepsilon,\eta,r}\|_{\mathrm{L}^\infty(B_\varepsilon)} \le M := 2^{2-p}\,\|\rho\|_{\mathrm{L}^\infty(\mathbb{R}^+)}$$

where M is independent of ε, η and r.

• Since $\int_{B_\varepsilon} x_1^{2\,k}\,d\gamma \ge \int_{B_\varepsilon} x_1^{2\,k+1}\,d\gamma = 0$ and $\int_{\mathbb{R}^n} x_1^{2\,k}\,d\gamma \ge \int_{\mathbb{R}^n} x_1^{2\,k+1}\,d\gamma = 0$ for any $k \in \mathbb{N}$, an expansion in Taylor series of $(1+\varepsilon\,x_1)^p$ gives

$$\begin{aligned}
\int_{\mathbb{R}^n} (1+\varepsilon\,x_1)^p\,d\gamma &\ge \int_{\mathbb{R}^n} \left(1 + \frac{1}{2}\,p\,(p-1)\,\varepsilon^2\,x_1^2 + \frac{1}{24}\,p\,(p-1)\,(p-2)\,(p-3)\,x_1^4\,\varepsilon^4\right) d\gamma \\
&= 1 + \frac{1}{2}\,p\,(p-1)\,\varepsilon^2 + \frac{1}{8}\,p\,(p-1)\,(p-2)\,(p-3)\,\varepsilon^4\,.
\end{aligned}$$

By applying (6.28) with $g = 1 + \varepsilon\, x_1$, we obtain

$$\int_{\mathbb{R}^n} |1+\varepsilon\, x_1|^p \, d\gamma \le \int_{B_\varepsilon} |1+\varepsilon\, x_1|^p \, d\gamma + \xi_p \left(1+\varepsilon^2\right)^{p/2} \varepsilon^5 \,.$$

Summing up with $\varepsilon^2 \le \theta$, we have

$$\int_{B_\varepsilon} |1+\varepsilon\, x_1|^p \, d\gamma \ge 1 + \frac{1}{2}\, p\,(p-1)\,\varepsilon^2 + \frac{1}{8}\, p\,(p-1)\,(p-2)\,(p-3)\,\varepsilon^4 - \xi_p\,(1+\theta)^{p/2}\,\varepsilon^5 \,.$$

• Let us estimate $\|u_f\|^p_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}$ using

$$\int_{\mathbb{R}^n} |u_f|^p \, d\gamma \ge \int_{B_\varepsilon} |u_f|^p \, d\gamma$$

and

$$\begin{aligned}\int_{B_\varepsilon} |u_f|^p \, d\gamma \ge \int_{B_\varepsilon} |1+\varepsilon\, x_1|^p \, d\gamma &\\ + p\,\eta \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-1}\, r \, d\gamma + \frac{1}{2}\, p\,(p-1)\,\eta^2 \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-2}\, r^2 \, d\gamma &\\ + \frac{1}{2}\, p\,(p-1)\,\eta^2 \int_{B_\varepsilon} r(x)^2\, \psi_{\varepsilon,\eta,r}(x)\, d\gamma \,.&\end{aligned}$$

We obtain

$$\begin{aligned}\int_{\mathbb{R}^n} |u_f|^p \, d\gamma \ge 1 + \frac{1}{2}\, p\,(p-1)\,\varepsilon^2 + \frac{1}{8}\, p\,(p-1)\,(p-2)\,(p-3)\,\varepsilon^4 - \xi_p\,(1+\theta)^{p/2}\,\varepsilon^5 &\\ + p\,\eta \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-1}\, r \, d\gamma + \frac{1}{2}\, p\,(p-1)\,\eta^2 \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-2}\, r^2 \, d\gamma &\\ + \frac{1}{2}\, p\,(p-1)\,\eta^2 \int_{B_\varepsilon} r(x)^2\, \psi_{\varepsilon,\eta,r}(x)\, d\gamma \,.&\end{aligned}$$

Step 2. We prove that $\eta^2 \int_{B_\varepsilon} r(x)^2\, \psi_{\varepsilon,\eta,r}(x)\, d\gamma$ is of order $o\big(\varepsilon^4\big)$ as $\eta \le t\,\varepsilon^2 \to 0$ for a given $t > 0$.

• By the logarithmic Sobolev inequality,

$$\int_{\mathbb{R}^n} h^2 \log h^2 \, d\gamma \le 2 \int_{\mathbb{R}^n} |\nabla h|^2 \, d\gamma + \int_{\mathbb{R}^n} h^2 \, d\gamma \log\left(\int_{\mathbb{R}^n} h^2 \, d\gamma\right),$$

applied to $h = 1 + (r-1)_+$, we learn that

$$\int_{\mathbb{R}^n} h^2 \log h^2 \, d\gamma \le 2 \int_{\mathbb{R}^n} |\nabla r|^2 \, d\gamma + \left(1 + \int_{\mathbb{R}^n} r^2 \, d\gamma\right) \log\left(1 + \int_{\mathbb{R}^n} r^2 \, d\gamma\right) \le 2\log(2\,e)\,.$$

Let $\chi := \mathbb{1}_{\{h>s_0\}}$ for any $s_0 > 1$ and consider $A_{\varepsilon,\eta,r,s_0} := \{x \in B_\varepsilon : \eta\, r(x) \le s_0\}$. Then we have

$$\int_{\mathbb{R}^n} h^2 \log h^2 \, d\gamma \ge \int_{B_\varepsilon \setminus A_{\varepsilon,\eta,r,s_0}} h^2 \log h^2 \, d\gamma = \int_{B_\varepsilon} \chi\, h^2 \log h^2 \, d\gamma \ge \log\left(\frac{s_0}{\eta}\right)^2 \int_{B_\varepsilon} \chi\, r^2 \, d\gamma$$

because $h \ge 1$ a.e. and $h > s_0/\eta$ on $B_\varepsilon \setminus A_{\varepsilon,\eta,r,s_0}$ and, as a consequence,

$$\eta^2 \int_{B_\varepsilon \setminus A_{\varepsilon,\eta,r,s_0}} r(x)^2\, \psi_{\varepsilon,\eta,r}(x)\, d\gamma \ge -\frac{\log(2\,e)\, M}{\log s_0 + \log\left(\frac{1}{\eta}\right)}\, \eta^2 \,.$$

• Let us notice that

$$\mathcal{M} = \sup_{s>0} \frac{|\rho(s)|}{\log(1+s)}$$

is finite. This allows us to write that

$$\eta^2 \int_{A_{\varepsilon,\eta,r,s_0}} r(x)^2 \, \psi_{\varepsilon,\eta,r}(x) \, d\gamma \geq -2^{2-p} \mathcal{M} \int_{A_{\varepsilon,\eta,r,s_0}} \eta^2 \, r^2 \log(1+2\eta\, r) \, \mathrm{Id}_{\{r>0\}} \, d\gamma \,,$$

where the restriction to the set $\{r > 0\}$ comes from the fact that $\psi_{\varepsilon,\eta,r}(x) = 0$ whenever $r(x) \leq 0$. Now we estimate $\log(1+2\eta\, r)$ by

$$\begin{aligned} &\log(1+2\eta\, r) \leq \log\big(1+2\sqrt{\eta}\big) && \text{if} \quad 0 \leq r \leq \frac{1}{\sqrt{\eta}}\,, \\ &\log(1+2\eta\, r) \leq \frac{\log(1+2\, s_0)}{\log\left(\frac{1}{\eta}\right)} \log r^2 && \text{if} \quad \frac{1}{\sqrt{\eta}} \leq r \leq \frac{s_0}{\eta}\,, \end{aligned}$$

and conclude using $\int_{\mathbb{R}^n} r^2 \, d\gamma \leq 1$ and $\int_{\mathbb{R}^n} r^2 \log r^2 \, d\gamma \leq 2 \int_{\mathbb{R}^n} |\nabla r|^2 \, d\gamma = 2$ by the logarithmic Sobolev inequality that

$$\eta^2 \int_{A_{\varepsilon,\eta,r,s_0}} r(x)^2 \, \psi_{\varepsilon,\eta,r}(x) \, d\gamma \geq -2^{2-p} \mathcal{M}\, \eta^2 \left(\theta \log\big(1+2\sqrt{\eta}\big) + \frac{2 \log(1+2\, s_0)}{\log\left(\frac{1}{\eta}\right)} \right),$$

Step 3. We compute the contribution of

$$p\,\eta \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-1} \, r \, d\gamma \quad \text{and} \quad \frac{1}{2}\, p\,(p-1)\,\eta^2 \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-2} \, r^2 \, d\gamma$$

to the expansion of $\int_{\mathbb{R}^n} |u_f|^p \, d\gamma$.

• Using (6.28) applied to $g = |r|^{1/p}$ and the orthogonality constraints on r, we obtain

$$\begin{aligned} p\,\eta \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-1} \, r \, d\gamma &\geq p\,\eta \left(\int_{\mathbb{R}^n} (1+\varepsilon\, x_1)^{p-1} \, r \, d\gamma - 2^{1-p} \, \xi_p \, \varepsilon^5 \right) \\ &\geq p\,\eta\,\varepsilon^2 \left(\frac{1}{2}\,(p-1)\,(p-2) \int_{\mathbb{R}^n} x_1^2 \, r \, d\gamma - \mathsf{c}_1\, \varepsilon^3 - 2^{1-p}\, \xi_p \, \varepsilon^5 \right) \end{aligned}$$

where $\mathsf{c}_1 := \sqrt{15}\, \sup_{s \in (-1,0)\cup(0,+\infty)} \left| (1+s)^{p-1} - 1 - (p-1)\, s - \frac{1}{2}\,(p-1)\,(p-2)\, s^2 \right| / s^3$

• A similar computation shows that

$$\frac{1}{2}\, p\,(p-1)\,\eta^2 \int_{B_\varepsilon} (1+\varepsilon\, x_1)^{p-2} \, r^2 \, d\gamma = \frac{1}{2}\, p\,(p-1)\,\eta^2 \left(\|r\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} - \mathsf{c}_2\, \varepsilon \right)$$

where $\mathsf{c}_2 := \sup_{s \in (-1,0)\cup(0,+\infty)} \left| (1+s)^{p-2} - 1 \right| / s$

Step 4. Collecting all terms, we have

$$\begin{aligned} \int_{\mathbb{R}^n} |u_f|^p \, d\gamma \geq 1 + \frac{1}{2}\, p\,(p-1)\,\varepsilon^2 + \frac{1}{8}\, p\,(p-1)\,(p-2)\,(p-3)\,\varepsilon^4 + \frac{1}{2}\, p\,(p-1)\,\eta^2 \, \|r\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)} \\ + \frac{1}{2}\, p\,(p-1)\,(p-2)\,\eta\,\varepsilon^2 \int_{\mathbb{R}^n} x_1^2 \, r \, d\gamma - C \frac{\varepsilon^4}{\log\left(\frac{1}{\varepsilon}\right)} \end{aligned}$$

for some constant $C > 0$ that is explicitly given in terms of t, ξ_p, M, $\mathcal{M}$, c_1 and c_2. In order to conclude, we notice that for any $s>-1$, $\frac{d^4}{ds^4}(1+s)^{\frac{2}{p}}>0$ implies that

$$(1+s)^{\frac{2}{p}}\geq 1+\frac{2}{p}\,s+\frac{1}{p^2}\,(2-p)\,s^2-\frac{1}{3\,p^3}\,(2-p)\,(4-p)\,s^3\,.$$

Applied to

$$\|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}=\left(\int_{\mathbb{R}^n}(1+\varepsilon\,x_1+\eta\,r)^p\,d\gamma\right)^{\frac{2}{p}}\,,$$

this completes the proof of Lemma 102. □

Proof of Theorem 87. Up to the replacement of f by $|f|$, there is no restriction in assuming that f is non-negative: we can rely on Lemma 102, which is the main ingredient of the proof. The strategy is now very similar to the proof of [P2, Theorem 7]. Let is consider the Hermite polynomial $h_1(x):=x_1^2-1$ and decompose r according to

$$r(x):=\alpha\,h_1(x)+\beta\,\tilde r(x)$$

with $\|\tilde r\|_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}=1$ so that $\|r\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}=\alpha^2+\beta^2\leq 1/2$ and $\int_{\mathbb{R}^n}x_1^2\,r(x)\,d\gamma=\alpha$. With these notations we have

$$\begin{aligned}
&\|\nabla u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}=\varepsilon^2+\eta^2\,,\\
&\|u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}=1+\varepsilon^2+\eta^2\left(\alpha^2+\beta^2\right),\\
&\|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\geq 1+(p-1)\left(\varepsilon^2+\eta^2\left(\alpha^2+\beta^2\right)\right)+(p-1)\,(2-p)\left(\frac{1}{2}\,\varepsilon^4+\alpha\,\varepsilon^2\,\eta\right)-\frac{\mathcal{C}\,\varepsilon^4}{\log\left(\frac{1}{\varepsilon}\right)}\,,
\end{aligned}$$

where the estimate of $\|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}$ comes from Lemma 102. Hence, for some $\lambda>0$ to be fixed,

$$\begin{aligned}
&\|\nabla u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\frac{1}{2-p}\left(\|u_f\|^2_{\mathrm{L}^2(\mathbb{R}^n,d\gamma)}-\|u_f\|^2_{\mathrm{L}^p(\mathbb{R}^n,d\gamma)}\right)-\lambda\left(\eta^2+\frac{t^2\,\varepsilon^4}{1+\varepsilon^2+\eta^2}\right)\\
&\geq\eta^2\left(1-\alpha^2-\beta^2-\lambda\right)+\left(\frac{p-1}{2}-\frac{\lambda\,t^2}{1+\varepsilon^2+\eta^2}\right)\varepsilon^4+(p-1)\,\alpha\,\varepsilon^2\,\eta-\frac{1}{2-p}\,\frac{\mathcal{C}\,\varepsilon^4}{\log\left(\frac{1}{\varepsilon}\right)}\\
&\geq\left(t^2\left(1-\alpha^2-\beta^2-2\,\lambda\right)+(p-1)\,\alpha\,t+\frac{1}{2}\,(p-1)-\frac{1}{2-p}\,\frac{\mathcal{C}}{\log\left(\frac{1}{\varepsilon}\right)}\right)\varepsilon^4\,.
\end{aligned}$$

By writing

$$\begin{aligned}
t^2&\left(1-\alpha^2-\beta^2-2\,\lambda\right)+(p-1)\,\alpha\,t+\frac{1}{2}\,(p-1)\\
&=t^2\left(1-\frac{p+1}{2}\,\alpha^2-\beta^2-2\,\lambda\right)+\frac{1}{2}\,(p-1)(1+\alpha\,t)^2\\
&\geq\frac{1}{4}\,t^2\,(3-p-8\,\lambda)+\frac{1}{2}\,(p-1)(1+\alpha\,t)^2\,,
\end{aligned}$$

for any given $t>0$ and $\lambda<(3-p)/8$, if $\varepsilon>0$ is small enough, we obtain the result with $c_{n,p}=(1-\theta)\,\lambda$. □

Chapter 7

Stability for the logarithmic Sobolev inequality

This chapter corresponds to [P4], in collaboration with Jean Dolbeault and Nikita Simonov.

Abstract

This paper is devoted to stability results for the Gaussian logarithmic Sobolev inequality, where stability is measured in strong norms. The strategy of our proofs relies on entropy methods and the use of diffusion flows based on the Ornstein-Uhlenbeck operator. Under appropriate constraints, explicit constants are obtained. Our approach covers several cases involving the strongest possible norm which is natural for the logarithmic Sobolev inequality.

7.1 Introduction and main results

Let us consider the *Gaussian logarithmic Sobolev inequality*

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} \ge \frac{1}{2}\int_{\mathbb{R}^d} |u|^2 \log\left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{R}^d)}}\right) d\gamma \quad \forall\, u \in \mathrm{H}^1(\mathbb{R}^d, d\gamma) \tag{7.1}$$

where $d\gamma = \gamma(x)\,dx$ is the normalized Gaussian probability measure with density

$$\gamma(x) = (2\pi)^{-\frac{d}{2}}\, e^{-\frac{1}{2}|x|^2} \quad \forall\, x \in \mathbb{R}^d .$$

In this paper we are interested in stability results, that is, in estimating the difference of the two terms in (7.1) from below, by a distance to the set of optimal functions. According to [81, 24], equality in (7.1) is achieved by functions in the manifold

$$\mathscr{M} := \left\{ w_{a,c} : (a,c) \in \mathbb{R}^d \times \mathbb{R} \right\}$$

where

$$w_{a,c}(x) = c\, e^{-a\cdot x} \quad \forall\, x \in \mathbb{R}^d$$

and only by these functions. The ultimate goal of *stability estimates* is to find a notion of distance d, an explicit constant $\beta > 0$ and an explicit exponent $\alpha > 0$, which may depend on d, such that

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^d} |u|^2 \log\left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{R}^d)}}\right) d\gamma \ge \beta \inf_{w\in\mathscr{M}} \mathsf{d}(u,w)^\alpha \tag{$\mathscr{S}$}$$

for any given $u \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$. In this paper we consider the slightly simpler question of finding a specific $w_u \in \mathscr{M}$ such that

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^d} |u|^2 \log\left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{R}^d)}}\right) d\gamma \ge \beta\,\mathsf{d}(u,w_u)^\alpha\,, \tag{$\star$}$$

which provides us with no more than an estimate for ($\mathscr{S}$): any estimate of α and β for ($\star$) is also an estimate for ($\mathscr{S}$). In order to illustrate the difference between the two questions, let us consider the following elementary example. Assume that $d = 1$ and consider the functions $u_\varepsilon(x) = 1 + \varepsilon\,x$ in the limit as $\varepsilon \to 0$. With $\mathsf{d}(u,w) = \|u' - w'\|_{\mathrm{L}^2(\mathbb{R},d\gamma)}$, which is the strongest possible notion of distance that we can expect to control in ($\star$), elementary computations show that the *deficit* of the logarithmic Sobolev inequality, *i.e.*, the left hand-side in ($\star$), is

$$\|\nabla u_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^d} |u_\varepsilon|^2 \log\left(\frac{|u_\varepsilon|^2}{\|u_\varepsilon\|^2_{\mathrm{L}^2(\mathbb{R}^d)}}\right) d\gamma = \frac{1}{2}\,\varepsilon^4 + O\big(\varepsilon^6\big)\,,$$

while, using the test function $w_{a_\varepsilon,c_\varepsilon} \in \mathscr{M}$ where $a_\varepsilon = 2\,\varepsilon$ and $c_\varepsilon = e^{-a_\varepsilon^2/4}$, we obtain

$$d(u_\varepsilon,1)^2 = \|u_\varepsilon'\|^2_{\mathrm{L}^2(\mathbb{R},d\gamma)} = \varepsilon^2 \quad \text{and} \quad \inf_{w\in\mathscr{M}} \mathsf{d}(u_\varepsilon,w)^\alpha \le d\left(u_\varepsilon, w_{a_\varepsilon,c_\varepsilon}\right)^2 = \frac{1}{2}\,\varepsilon^4 + O\big(\varepsilon^6\big)\,.$$

In practice we will consider only the case

$$w_u = 1$$

in ($\star$) and the above example shows that the best we can hope for without additional restriction is $\alpha \ge 4$. Similar examples in higher dimensions can be obtained by considering for an arbitrary given $\nu \in \mathbb{S}^{d-1}$ the functions $u_\varepsilon(x) = 1 + \varepsilon\,x\cdot\nu$ in the limit as $\varepsilon \to 0$. This is not a surprise in view of [138, 130, 171], and also of the detailed Taylor expansions of [144, 74, 75]. Still with $w_u = 1$, we can expect to have $\alpha = 2$ in ($\star$) *under additional conditions*, including for $\mathsf{d}(u,w) = \|\nabla u - \nabla w\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}$, while it is otherwise banned as shown for instance from [171, Theorem 1.2 (2)], or simply from considering the above example. Before entering the details, let us mention a recent stability result for ($\mathscr{S}$) with $\alpha = 2$ involving a constructive although very delicate expression for $\beta > 0$ and $\mathsf{d}(u,w) = \|u - w\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}$ that appeared in [119]. Here we aim at stronger estimates under additional constraints, with $w_u = 1$, which is a different point of view. Let us start by a first stability result.

Proposition 103. *For all $u \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ such that $\|u\|_{\mathrm{L}^2(\mathbb{R}^d)} = 1$ and $\|x\,u\|^2_{\mathrm{L}^2(\mathbb{R}^d)} \le d$, we have*

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^d} |u|^2 \log|u|^2\,d\gamma \ge \frac{1}{2\,d}\left(\int_{\mathbb{R}^d} |u|^2 \log|u|^2\,d\gamma\right)^2 \tag{7.2}$$

and, with $\psi(s) := s - \frac{d}{4}\log\big(1 + \frac{4}{d}\,s\big)$, we also have the stronger estimate

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2}\int_{\mathbb{R}^d} |u|^2 \log|u|^2\,d\gamma \ge \psi\left(\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}\right). \tag{7.3}$$

Similar results are already known in the literature (see for instance [58, 138, 130, 171]) and we claim no originality for the the results. Also see references to earlier proofs at the end of the introduction. Our method is based on the *carré du champ* method. Even if some ideas go back to [15], it is elementary, new as far as we know, and of some use for our other results.

Coming back to ($\star$), we may notice that there is no loss of generality in imposing the condition $\|u\|_{\mathrm{L}^2(\mathbb{R}^d)} = 1$, as we can always replace u by $u/\|u\|_{\mathrm{L}^2(\mathbb{R}^d)}$. Because of the Csiszár-Kullback-Pinsker inequality

$$\int_{\mathbb{R}^d} |u|^2 \log|u|^2 \, d\gamma \ge \frac{1}{4} \left\| |u|^2 - 1 \right\|^2_{\mathrm{L}^1(\mathbb{R}^d, d\gamma)} \tag{7.4}$$

and $\big||u|-1\big| = \big||u|^2-1\big| / \big||u|+1\big| \le \big||u|^2-1\big|$, we find that (7.2) implies ($\star$) type with

$$\mathsf{d}(u,w) = \|u-w\|_{\mathrm{L}^1(\mathbb{R}^d, d\gamma)}$$

for nonnegative functions u, $\alpha = 4$, and $\beta = 1/(32\,d)$. For functions far away from the optimal functions, say such that $\|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)} \ge A$ under the conditions of Proposition 103, Inequality (7.3) provides us with an even stronger stability result of ($\star$) type with $\alpha = 2$ and $\mathsf{d}(u,w) = \|\nabla u - \nabla w\|_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)}$, but with a positive constant β which depends on $A > 0$. Again, notice that ($\star$) with such a distance cannot hold without constraints.

Next we aim at explicit results with $\alpha = 2$, under other constraints. Let

$$\mathscr{C}_\star = 1 + \frac{1}{1728} \approx 1.0005787\,.$$

Theorem 104. *For all $u \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ such that $u^2\gamma$ is log-concave and such that*

$$\int_{\mathbb{R}^d} (1,x)\,|u|^2\, d\gamma = (1,0) \quad \textit{and} \quad \int_{\mathbb{R}^d} |x|^2\,|u|^2\, d\gamma \le d\,, \tag{7.5}$$

we have

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)} - \frac{\mathscr{C}_\star}{2} \int_{\mathbb{R}^d} |u|^2 \log|u|^2\, d\gamma \ge 0\,. \tag{7.6}$$

The condition $\int_{\mathbb{R}^d} |x|^2\,|u|^2\, d\gamma \le d$ in (7.5) is a simplifying assumption. A result like (7.6) also holds if $\int_{\mathbb{R}^d} |x|^2\,|u|^2\, d\gamma > d$, but with a constant that differs from $\mathscr{C}_\star$ and actually depends on $\int_{\mathbb{R}^d} |x|^2\,|u|^2\, d\gamma$. We refer to Section 7.3.5: see Proposition 109 for an extension of Theorem 104, and also for further comments on the extension of Proposition 103. The constant $\mathscr{C}_\star$ in (7.8) relies on an estimate of [55].

Inequality (7.6) with improved constant $\mathscr{C}_\star > 1$ compared to (7.1) can be recast in the form of a stability inequality of type ($\star$) around the normalised Gaussian as

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)} - \frac{1}{2} \int_{\mathbb{R}^d} |u|^2 \log|u|^2\, d\gamma \ge \frac{1}{2}\,(\mathscr{C}_\star - 1) \int_{\mathbb{R}^d} |u|^2 \log|u|^2\, d\gamma$$

for all functions $u \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ such that $\|u\|_{\mathrm{L}^2(\mathbb{R}^d)} = 1$, which covers the case $\alpha = 2$, $\beta = (\mathscr{C}_\star - 1)/8$ and $\mathsf{d}(u,w) = \|u-w\|_{\mathrm{L}^1(\mathbb{R}^d, d\gamma)}$ in ($\star$) for nonnegative functions by (7.4), or even in the stronger $\dot{\mathrm{H}}^1(\mathbb{R}^d, d\gamma)$ semi-norm, as

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)} - \frac{1}{2} \int_{\mathbb{R}^d} |u|^2 \log|u|^2\, d\gamma \ge \frac{\mathscr{C}_\star - 1}{\mathscr{C}_\star}\, \|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)}$$

for all functions $u \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ such that $\|u\|_{\mathrm{L}^2(\mathbb{R}^d)} = 1$, which corresponds to $\alpha = 2$, $\beta = (\mathscr{C}_\star - 1)/\mathscr{C}_\star$ and $\mathsf{d}(u,w) = \|\nabla(u-w)\|_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)}$ in ($\star$). By the Gaussian Poincaré inequality, notice that the case of ($\star$) with $\alpha = 2$, $\beta = (\mathscr{C}_\star - 1)/\mathscr{C}_\star$ and the standard distance $\mathsf{d}(u,w) = \|u-w\|_{\mathrm{L}^2(\mathbb{R}^d, d\gamma)}$ is also covered.

Log-concavity might appear as a rather restrictive assumption, but this is useful because a function which is compactly supported at time $t=0$ evolves through the diffusion flow into a logarithmically concave function after some finite time that can be estimated by the heat flow estimates of [192]. This is enough to produce a stability result with an explicit constant. Compact support is in fact a too restrictive condition and we have the following result.

Theorem 105. *Let $d\ge 1$. For any $\varepsilon>0$, there is some explicit $\mathscr{C}>1$ depending only on ε such that, for any $u\in \mathrm{H}^1(\mathbb{R}^d,d\gamma)$ satisfying (7.5) and*

$$\int_{\mathbb{R}^d}|u|^2\,e^{\varepsilon\,|x|^2}\,d\gamma<\infty\,, \tag{7.7}$$

then we have

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}\ge\frac{\mathscr{C}}{2}\int_{\mathbb{R}^d}|u|^2\,\log|u|^2\,d\gamma\,. \tag{7.8}$$

Additionally, if u is compactly supported in a ball of radius $R>0$, then (7.8) holds with

$$\mathscr{C}=1+\frac{\mathscr{C}_\star-1}{1+\mathscr{C}_\star\,R^2}\,.$$

This expression of the constant $\mathscr{C}$ in (7.8) is given in the proof, in Section 7.3.4. The simpler estimate in terms of R relies on Theorem 104.

Let us conclude this introduction with a review of the literature. The *logarithmic Sobolev inequality* historically appeared in [241, 53], in a form that was later rediscovered as an equivalent scale-invariant form of the Euclidean version of the inequality in [261]. We refer to [155] for the Gaussian version (7.1) of the inequality, and also to [139] for an equivalent result. The optimality case in the inequality has been characterized in [81] but can also be deduced from [24]. Also see [253] for a short introductory review with an emphasis on *information theoretical* aspects. The logarithmic Sobolev inequality can be viewed as a limit case of a family of the Gagliardo-Nirenberg-Sobolev (GNS) as observed in [110], in the Euclidean setting, or as a large dimension limit of the Sobolev inequality according to [36]. See [119, 75] for recent developments and further references. We refer to [12, 160, 235, 26] to reference books on the topic. In a classical result on stability in functional inequalities, Bianchi and Egnell proved in [47] that the deficit in the Sobolev inequality measures the $\dot{\mathrm{H}}^1(\mathbb{R}^d,dx)$ distance to the manifold of the Aubin-Talenti functions. The estimate has been made constructive in [119] where a new $\mathrm{L}^2(\mathbb{R}^d,dx)$ stability result for the logarithmic Sobolev inequality is also established (also see [171] for furrther results in strong norms). Still in the Euclidean setting a first stability result in strong norms for the logarithmic Sobolev inequality appears in [138], where the authors give deficit estimates in various distances for functions inducing a Poincaré inequality. Under the condition $\|x\,u\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}=\sqrt{d}$, a stability result measured by an entropy is given in [130]. For sequential stability results in strong norms, we refer to [172] when assuming a bound on u in $\mathrm{L}^4(\mathbb{R}^d,d\gamma)$ and to [171] when assuming a bound on $|x|^2\,u$ in $\mathrm{L}^2(\mathbb{R}^d,d\gamma)$. Stability according to other notions of distance has been studied in [191, 190, 140].

To our knowledge, the first result of stability for the logarithmic Sobolev inequality is a reinforcement of the inequality due to E. Carlen in [81] where he introduces an additional term involving the Wiener transform. Stability in logarithmic Sobolev inequality is related to deficit in Gaussian isoperimetry and we refer to [58] for an introduction to early results in this direction, [29] for a sharp, dimension-free quantitative Gaussian isoperimetric inequality, and [135] for recent results and further references. Results of Proposition 103 are known from [58, Theorem 1.1] where it is deduced from the HWI inequality due to F. Otto and C. Villani [226]. Such estimates have even been refined in [135]. There are several other proofs. In [138], M. Fathi,

E. Indrei and M. Ledoux use a Mehler formula for the Ornstein-Uhlenbeck semigroup and Poincaré inequalities. The proof in [130] is based on simple scaling properties of the Euclidean form of the logarithmic Sobolev inequality, which also apply to Gagliardo-Nirenberg inequalities. Various stability results have been proved in Wasserstein's distance: we refer to [173, 58, 138, 172, 182, 61, 135, 171]. A key argument for Theorem 104 is the fact that the heat flow preserves log-concavity according, *e.g.*, [236], which is a pretty natural property in this framework: see for instance [102].

In this paper, we carefully distinguish stability results of type ($\mathscr{S}$) where stability is measured w.r.t. $\mathscr{M}$, and of type ($\star$) where the distance to a given function is estimated. Even if this function is normalized and centered, this is not enough as shown in [171]. Many counter-examples to stability are known, involving Wasserstein's distance for instance in [173, 58, 138, 172, 182], weaker distances like p-Wasserstein, or stronger norms like L^p or H^1: see for instance [172, 171]. The main counter-examples which we might try to apply to our setting are [182, Theorem 1.3] and [135, Theorem 4] but, as already noted in [172], they are based on the fact that the second moment diverges along a sequence of test functions, which is forbidden in our assumptions.

This paper is organized as follows. Section 7.2 is devoted to the standard *carré du champ* method and a proof of Proposition 103. Theorem 104 is proved in Section 7.3.3, under a log-concavity assumption. Using properties of the heat flow, the method is extended to the larger class of functions of Theorem 105 in Section 7.3.4.

7.2 Entropy methods and entropy – entropy production stability estimates

This section is devoted to the proof of Proposition 103.

7.2.1 Definitions and preliminary results

Consider the *Ornstein-Uhlenbeck equation* on $\mathbb{R}^d$

$$\frac{\partial h}{\partial t} = \mathscr{L} h\,, \quad h(t=0,\cdot) = h_0\,, \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}^d\,, \tag{7.9}$$

where $\mathscr{L} h := \Delta h - x \cdot \nabla h$ denotes the *Ornstein-Uhlenbeck* operator.

Let us recall some classical results. If $h_0 \in \mathrm{L}^1(\mathbb{R}^d, d\gamma)$ is nonnegative, then there exists a unique nonnegative weak solution to (7.9) (see for instance [137]). The two key properties of the Ornstein-Uhlenbeck operator are

$$\int_{\mathbb{R}^d} v_1\,(\mathscr{L} v_2)\,d\gamma = -\int_{\mathbb{R}^d} \nabla v_1 \cdot \nabla v_2\,d\gamma \quad \text{and} \quad [\nabla, \mathscr{L}]\,v = -\nabla v\,.$$

As a consequence, we obtain the two identities

$$\int_{\mathbb{R}^d} (\mathscr{L} v)^2\,d\gamma = \int_{\mathbb{R}^d} \|\mathrm{Hess}\,v\|^2\,d\gamma + \int_{\mathbb{R}^d} |\nabla v|^2\,d\gamma \tag{7.10}$$

and

$$\int_{\mathbb{R}^d} \mathscr{L} v\,\frac{|\nabla v|^2}{v}\,d\gamma = -2\int_{\mathbb{R}^d} \mathrm{Hess}\,v : \frac{\nabla v \otimes \nabla v}{v}\,d\gamma + \int_{\mathbb{R}^d} \frac{|\nabla v|^4}{v^2}\,d\gamma\,, \tag{7.11}$$

where $\mathrm{Hess}\,v = \nabla \otimes \nabla v$ is the *Hessian matrix* of v. Here we use the following notations. If a and b take values in $\mathbb{R}^d$, $a\otimes b$ denotes the matrix $(a_i\,b_j)_{1\le i,j\le d}$. With matrix valued m and n, we define $m:n=\sum_{i,j=1}^d m_{i,j}\,n_{i,j}$ and $\|m\|^2=m:m$. If h is a nonnegative solution of (7.9), notice that $v=\sqrt{h}$ solves

$$\frac{\partial v}{\partial t}=\mathscr{L}v+\frac{|\nabla v|^2}{v}\,. \tag{7.12}$$

Let us fix $\|v\|_{\mathrm{L}^2(\mathbb{R}^d)}=1$, then the *entropy* and the *Fisher information*, respectively defined by

$$\mathcal{E}[v]:=\int_{\mathbb{R}^d}|v|^2\log|v|^2\,d\gamma\quad\text{and}\quad\mathcal{I}[v]:=\int_{\mathbb{R}^d}|\nabla v|^2\,d\gamma\,,$$

evolve along the flow according to

$$\frac{d}{dt}\mathcal{E}[v]=-4\,\mathcal{I}[v]\quad\text{and}\quad\frac{d}{dt}\mathcal{I}[v]=-2\int_{\mathbb{R}^d}\left((\mathscr{L}v)^2+\mathscr{L}v\,\frac{|\nabla v|^2}{v}\right)d\gamma$$

if v solves (7.12). Using (7.10) and (7.11), we obtain the classical expression of the *carré du champ* method

$$\frac{d}{dt}\mathcal{I}[v]+2\,\mathcal{I}[v]=-2\int_{\mathbb{R}^d}\left\|\mathrm{Hess}\,v-\frac{\nabla v\otimes\nabla v}{v}\right\|^2d\gamma \tag{7.13}$$

as for instance in [19, 128, 26]. By writing that

$$\frac{d}{dt}\left(\mathcal{I}[v]-\frac12\,\mathcal{E}[v]\right)\le0\quad\text{and}\quad\lim_{t\to+\infty}\left(\mathcal{I}[v(t,\cdot)]-\frac12\,\mathcal{E}[v(t,\cdot)]\right)=0\,,$$

we recover the standard proof of the *entropy – entropy production* inequality

$$\mathcal{I}[v]-\frac12\,\mathcal{E}[v]\ge0\,, \tag{7.14}$$

i.e., of (7.1) by the method of [24].

Several of the above expression can be rephrased in terms of the *pressure variable*

$$P:=-\log h=-2\,\log v$$

using the following elementary identities

$$\nabla v=-\frac12\,e^{-P/2}\,\nabla P\,,\quad\frac{\nabla v\otimes\nabla v}{v}=\frac14\,e^{-P/2}\,\nabla P\otimes\nabla P\,,$$
$$\mathrm{Hess}\,v=-\frac12\,e^{-P/2}\,\mathrm{Hess}\,P+\frac14\,e^{-P/2}\,\nabla P\otimes\nabla P\,,$$

so that, by taking into account $v\,\nabla P=-2\,\nabla v$ and $h=v^2$, we have

$$\mathcal{I}[v]=\frac14\int_{\mathbb{R}^d}|\nabla P|^2\,h\,d\gamma\quad\text{and}\quad\int_{\mathbb{R}^d}\left\|\mathrm{Hess}\,v-\frac{\nabla v\otimes\nabla v}{v}\right\|^2d\gamma=\frac14\int_{\mathbb{R}^d}\|\mathrm{Hess}\,P\|^2\,h\,d\gamma\,.$$

7.2.2 Improvements under moment constraints

In standard computations based on the *carré du champ* method, one usually drops the right-hand side in (7.13) which results in the standard exponential decay of $\mathcal{I}[v(t,\cdot)]$ if v solves (7.12) and, after integration on $t\in\mathbb{R}^+$, proves (7.1). Keeping track of the right-hand side in (7.13) provides us with improvements as shown in [15, 112, 129] in various interpolation inequalities but generically fails in the case of the logarithmic

Sobolev inequality. We remedy to this issue by introducing moment constraints. This is not a very difficult result but, as far as we know, it is new in the framework of the *carré du champ* method.

Lemma 106. *With the notations of Section 7.2.1, if* $v \in \mathrm{H}^2(\mathbb{R}^d, d\gamma)$ *is a positive function such that* $\int_{\mathbb{R}^d} |x|^2\,|v|^2\,d\gamma \le d$*, then*

$$4\,\mathcal{I}[v] \le \int_{\mathbb{R}^d} (\Delta P)\,h\,d\gamma \le \sqrt{d \int_{\mathbb{R}^d} \|\mathrm{Hess}\,P\|^2\,h\,d\gamma}\,.$$

Proof. Using $h\,\nabla P = -\nabla h$, we obtain

$$4\,\mathcal{I}[v] = \int_{\mathbb{R}^d} |\nabla P|^2\,h\,d\gamma = -\int_{\mathbb{R}^d} \nabla P \cdot \nabla h\,d\gamma = \int_{\mathbb{R}^d} h\,(\mathscr{L} P)\,d\gamma\,.$$

After recalling that $\mathscr{L} P = \Delta P - x \cdot \nabla P$, using an integration by parts we deduce that

$$-\int_{\mathbb{R}^d} h\,x \cdot \nabla P\,d\gamma = \int_{\mathbb{R}^d} x \cdot \nabla h\,d\gamma = \int_{\mathbb{R}^d} h\left(|x|^2 - d\right)d\gamma = \int_{\mathbb{R}^d} |v|^2\left(|x|^2 - d\right)d\gamma \le 0$$

which proves the first inequality. The second inequality follows from a Cauchy-Schwarz inequality and the arithmetic-geometric inequality

$$(\Delta P)^2 \le d\,\|\mathrm{Hess}\,P\|^2\,.$$

□

Proof of Proposition 103. Let $h = v^2$ be the solution of (7.9) with initial datum $h_0 = u^2$. Since $x \mapsto \left(|x|^2 - d\right)$ is an eigenfunction of $\mathscr{L}$ with corresponding eigenvalue -2 and $\mathscr{L}$ is self-adjoint on $\mathrm{L}^2(\mathbb{R}^d, d\gamma)$, we have

$$\frac{d}{dt} \int_{\mathbb{R}^d} \left(|x|^2 - d\right) h\,d\gamma = \int_{\mathbb{R}^d} \left(|x|^2 - d\right)(\mathscr{L} h)\,d\gamma$$
$$= \int_{\mathbb{R}^d} h\,\mathscr{L}\left(|x|^2 - d\right)d\gamma = -2 \int_{\mathbb{R}^d} \left(|x|^2 - d\right) h\,d\gamma\,. \tag{7.15}$$

The sign of $t \mapsto \int_{\mathbb{R}^d} \left(|x|^2 - d\right) h(t,x)\,d\gamma$ is conserved and in particular we have that $\int_{\mathbb{R}^d} |x|^2\,|v|^2\,d\gamma \le d$ for any $t \ge 0$. For any $i = 1, 2 \ldots d$, we also notice that $x \mapsto x_i$ is also an eigenfunction of $\mathscr{L}$ with corresponding eigenvalue -1 so that

$$\frac{d}{dt} \int_{\mathbb{R}^d} x\,h\,d\gamma = -\int_{\mathbb{R}^d} x\,h\,d\gamma$$

and, as a consequence $\int_{\mathbb{R}^d} x\,h(t,\cdot)\,d\gamma = 0$ for all $t \ge 0$ because $\int_{\mathbb{R}^d} x\,h_0\,d\gamma = 0$.

For smooth enough solutions, we deduce from Lemma 106, (7.13) and (7.14) that

$$\frac{d}{dt}\mathcal{I}[v] + 2\,\mathcal{I}[v] \le -\frac{8}{d}\,\mathcal{I}^2[v] \le \frac{1}{2d}\,\frac{d}{dt}\left(\mathcal{E}[v]\right)^2$$

and obtain by considering the limit as $t \to +\infty$ that

$$\mathcal{I}[v] \ge \frac{1}{2}\,\mathcal{E}[v] + \frac{1}{2d}\left(\mathcal{E}[v]\right)^2.$$

This provides us with (7.2). In the general case, one can get rid of the $\mathrm{H}^2(\mathbb{R}^d, d\gamma)$ regularity of Lemma 106 by a standard approximation scheme, which is classical and will not be detailed here.

As in [74], a better estimate is achieved as follows. Let

$$\phi(s) := \frac{d}{4}\left(e^{\frac{2}{d}\,s} - 1\right) \quad \forall\, s \ge 0\,.$$

Using $\frac{d}{dt}\mathcal{E}[v] = -4\mathcal{I}[v]$, we notice that

$$\frac{d}{dt}\Big(\mathcal{I}[v] - \phi\big(\mathcal{E}[v]\big)\Big) = -\frac{8}{d}\Big(\mathcal{I}[v] - \phi\big(\mathcal{E}[v]\big)\Big).$$

Since $\lim_{t\to+\infty}\mathcal{I}[v(t,\cdot)] = 0$ as can be deduced from a Gronwall estimate relying on $\frac{d}{dt}\mathcal{I}[v] \le -2\mathcal{I}[v]$ and $\lim_{t\to+\infty}\mathcal{E}[v(t,\cdot)] = 0$ as a consequence of (7.1), one knows that

$$\lim_{t\to+\infty}\Big(\mathcal{I}[v(t,\cdot)] - \phi\big(\mathcal{E}[v(t,\cdot)]\big)\Big) = 0.$$

Moreover, Gronwall estimates show that $\mathcal{I}[v(t,\cdot)] - \phi\big(\mathcal{E}[v(t,\cdot)]$ cannot change sign and an asymptotic expansion as $t\to+\infty$ as in [74, Appendix B.4] is enough to obtain that $\mathcal{I}[v(t,\cdot)] - \phi\big(\mathcal{E}[v(t,\cdot)]$ takes nonnegative values for $t>0$ large enough. Altogether, we conclude that

$$\mathcal{I}[v(t,\cdot)] - \phi\Big(\mathcal{E}[v(t,\cdot)]\Big) \ge 0$$

for any $t\ge 0$ and, as a particular case, at $t=0$ for $v(0,\cdot)=u$. The function ϕ is convex increasing and, as such, invertible, so that we can also write

$$\varphi\big(\mathcal{I}[u]\big)^{-1} - \phi\Big(\mathcal{E}[u]\Big) \ge 0.$$

his completes the proof of (7.3) with the convex monotone increasing function

$$\psi(s) := s - \frac{1}{2}\phi^{-1}(s).$$

□

7.3 Stability results

7.3.1 Log-concave measures and Poincaré inequality

According to [27], given a Borel probability measure μ on $\mathbb{R}^d$, its *isoperimetric constant* is defined as

$$h(\mu) := \inf_A \frac{\mathrm{P}_\mu(A)}{\min\big\{\mu(A), 1-\mu(A)\big\}}$$

where the infimum is taken on the set of arbitrary Borel subset $\mathbb{R}^d$ with μ-perimeter or surface measure $\mathrm{P}_\mu(A) := \lim_{\varepsilon\to 0_+}\big(\mu(A_\varepsilon) - \mu(A)\big)/\varepsilon$ and $A_\varepsilon := \{x\in\mathbb{R}^d : |x-a| < \varepsilon$ for some $a\in A\}$. Here and in what follows, we shall say that a measure μ with density $e^{-\psi}$ with respect to Lebesgue's measure is a *log-concave probability measure* if ψ is a convex function, and denote by $\lambda_1(\mu)$ the first positive eigenvalue of $-\mathscr{L}_\psi$ where $\mathscr{L}_\psi$ is the *Ornstein-Uhlenbeck operator* $\mathscr{L}_\psi := \Delta - \nabla\psi\cdot\nabla$. In that case, we learn from [189, Ineq. (5.8)] that

$$\frac{1}{4}h(\mu)^2 \le \lambda_1(\mu) \le 36\,h(\mu)^2$$

where the lower bound is J. Cheeger's inequality that goes back to [93] for Riemannian manifolds and also to earlier works by V.G. Maz'ya [206, 207]. This bound was later improved in [76, 188]. The characterization of $h(\mu)$ has been actively studied, but it is out of the scope of the present paper. We learn from [55, Theorem 1.2] and [55, Ineq. (3.4)] that

$$h(\mu) \geq \frac{1}{6\sqrt{3\int_{\mathbb{R}^d}|x-x_\mu|^2\,d\mu}} \quad \text{where} \quad x_\mu = \int_{\mathbb{R}^d} x\,d\mu$$

for any log-concave probability measure μ. This estimate is closely related with the results by R. Kannan, L. Lovász and M. Simonovits in [179] and their conjecture, which again lies out of the scope of the present paper (see for instance [57] for a recent work on the topic).

Altogether, if μ is a log-concave probability measure with $d\mu = e^{-\psi}\,dx$ such that $\int_{\mathbb{R}^d}|x|^2\,d\mu \leq d$, then we have the *Poincaré inequality*

$$\int_{\mathbb{R}^d}|\nabla f|^2\,d\mu \geq \frac{1}{432}\int_{\mathbb{R}^d}|f|^2\,d\mu \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^d, d\mu) \text{ such that } \int_{\mathbb{R}^d} f\,d\mu = 0\,. \tag{7.16}$$

We refer to [88] and references therein for further estimates on $\lambda_1(\mu)$.

7.3.2 Time evolution, log-concave densities and Poincaré inequality

Lemma 107. *Let us consider consider a solution h of (7.9) with initial datum $h_0 = v^2$ and assume that $\mu_0 := h_0\,\gamma$ is log-concave. Then $\mu_t := h(t,\cdot)\,\gamma$ is log-concave for all $t \geq 0$.*

Proof. The function $g := h\,\gamma$ solves the Fokker-Planck equation

$$\frac{\partial g}{\partial t} = \Delta g + \nabla\cdot(x\,g)\,.$$

The function f such that

$$f(t,x) := g\left(\frac{1}{2}\log(1+2\,t), \frac{x}{\sqrt{1+2\,t}}\right) \quad \forall\,(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$$

solves the heat equation and can be represented using the heat kernel. According for instance to [236, 30], log-concavity is preserved under convolution, which completes the proof. □

The log-concavity property becomes true under the action of the flow of (7.9) after some delay $t_\star$ for large classes of initial data. With the notation of Lemma 107, for any $R > 0$, we read from [192, Theorem 5.1] by K. Lee and J-L. Vázquez that μ_t is log-concave for any

$$t \geq t_\star(R) := \log\left(\sqrt{R^2+1}\right), \tag{7.17}$$

if v is compactly supported in a ball of radius $R > 0$, by reducing the problem to the heat flow as in the above proof. As a consequence, we know that (7.16) holds for any $t \geq t_\star(R)$.

Alternatively, under Assumption (7.7), we learn from a recent paper [95, Theorem 2] by H.-B. Chen, S. Chewi, and J. Niles-Weed that the *Poincaré inequality*

$$\int_{\mathbb{R}^d}|\nabla f|^2\,d\mu_t \geq \lambda_1(\mu_t)\int_{\mathbb{R}^d}|f|^2\,d\mu_t \quad \forall\, f \in \mathrm{H}^1(\mathbb{R}^d, d\mu_t) \text{ such that } \int_{\mathbb{R}^d} f\,d\mu_t = 0 \tag{7.18}$$

holds for all $t \geq t_\star^\varepsilon$ with

$$t_\star^\varepsilon := \log\left(\sqrt{1+\varepsilon^{-1}}\right), \quad \frac{1}{\lambda_1(\mu_t)} \leq \tau\left(\frac{\varepsilon\,\tau}{\varepsilon\,\tau-1} + A^{\frac{1}{\varepsilon\tau-1}}\right) \quad \text{and} \quad \tau = \frac{1}{2}\left(e^{2t}-1\right).$$

7.3.3 Explicit stability results for log-concave densities

Let us start by an elementary observation.

Lemma 108. *If* $h \in \mathrm{H}^1(\mathbb{R}^d, d\gamma)$ *is such that* $\int_{\mathbb{R}^d} x\, h\, d\gamma = 0$ *and* $P = -\log h$ *is the* pressure variable*, then*

$$\int_{\mathbb{R}^d} \nabla P\, h\, d\gamma = 0\,.$$

Proof. The result follows from $\int_{\mathbb{R}^d} \nabla P\, h\, d\gamma = -\int_{\mathbb{R}^d} \nabla h\, d\gamma = \int_{\mathbb{R}^d} x\, h\, d\gamma = 0$. □

With this result in hand, we can now prove our first main result.

Proof of Theorem 104. The function $h = v^2$ is such that $\int_{\mathbb{R}^d} x\, h\, d\gamma = 0$ and Lemma 108 applies. Since $h\gamma$ is log-concave, we can apply (7.16) with $f = \partial P/\partial x_i$ for any $i = 1, 2, \ldots d$ and obtain

$$\int_{\mathbb{R}^d} \|\mathrm{Hess}\, P\|^2\, h\, d\gamma \ge \frac{1}{432} \int_{\mathbb{R}^d} |\nabla P|^2\, h\, d\gamma\,.$$

It follows from (7.13) that

$$\frac{d}{dt} \int_{\mathbb{R}^d} |\nabla v|^2\, d\gamma + 2 \int_{\mathbb{R}^d} |\nabla v|^2\, d\gamma \le -\frac{1}{864} \int_{\mathbb{R}^d} |\nabla v|^2\, d\gamma\,,$$

and the stability result is obtained as in the proof of Proposition 103. □

7.3.4 Extension by entropy methods and flows

This section is devoted to the proof of Theorem 105. The key idea is to evolve the function by the Ornstein-Uhlenbeck equation, so that the solution after an *initial time layer* has the log-concavity property of Theorem 104, at least if the initial datum has compact support. To some extent, the strategy is similar to the one used in [63]. During the initial time layer, we use an improved version of the entropy – entropy production inequality which arises as a consequence of the *carré du champ* method.

Proof of Theorem 105. Let h be the solution to (7.9) with initial datum $h_0 = u^2$ and define

$$\mathcal{Q}(t) := \frac{\mathcal{I}\big[\sqrt{h(t,\cdot)}\big]}{\mathcal{E}\big[\sqrt{h(t,\cdot)}\big]} \quad \forall\, t \ge 0\,.$$

Let us assume first that v has compact support. With no loss of generality, we can assume that v is supported in $B(0,R)$ for some $R > 0$. With $t_\star = t_\star(R)$ given by (7.17), we know from [192, Theorem 5.1] that μ_t is log-concave at $t = t_\star$ and Theorem 104 applies:

$$\mathcal{Q}(t_\star) \ge \frac{\mathscr{C}_\star}{2}\,.$$

With an estimate similar to [63, Lemma 2.9], we learn from Section 7.2 that

$$\frac{d\mathcal{Q}}{dt} \le 2\,\mathcal{Q}\,(2\,\mathcal{Q} - 1)\,. \tag{7.19}$$

An integration on $(0, t_\star)$ shows that

$$\mathcal{Q}(0) \ge \frac{1}{2}\left(1 + \frac{2\,\mathcal{Q}(t_\star) - 1}{1 + 2\,\mathcal{Q}(t_\star)\,(e^{2\,t_\star} - 1)}\right) \ge \frac{1}{2}\left(1 + \frac{\mathscr{C}_\star - 1}{1 + \mathscr{C}_\star\, R^2}\right) = \frac{\mathscr{C}}{2}\,.$$

Under the more general assumption (7.7), we rely on (7.18) and obtain with same notations as above and $t_\star = t_\star^\varepsilon$ that

$$\int_{\mathbb{R}^d} \|\mathrm{Hess}\, P\|^2\, h(t,\cdot)\, d\gamma \ge \lambda_1(\mu_t) \int_{\mathbb{R}^d} |\nabla P|^2\, h(t,\cdot)\, d\gamma \quad \forall\, t \ge t_\star\,.$$

Moreover, for some explicit $t_0 = t_0(\varepsilon) > t_\star$, we notice that $t \mapsto \lambda_1(\mu_t)$ is nonincreasing on $(t_0, +\infty)$. Hence we deduce from

$$\mathcal{I}[v(t_0,\cdot)] - \frac{1}{8} \int_{t_0}^{+\infty} \big(4 + \lambda_1(\mu_s)\big)\, \mathcal{E}'[v(s,\cdot)]\, ds \ge 0$$

after an integration by parts that

$$\mathcal{Q}(t_0) \ge \frac{1}{2}\left(1 + \frac{1}{4}\,\lambda_1(\mu_{t_0})\right) =: \mathscr{C}_0\,.$$

Using (7.19), we obtain

$$\mathscr{C} = 1 + \frac{\mathscr{C}_0 - 1}{1 + \mathscr{C}_0\,(e^{2\,t_0} - 1)}\,.$$

This concludes the proof. □

7.3.5 Normalization issues

If we do not assume that $\|u\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} = 1$ and $\|x\,u\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} \le d$, it is still possible to state the analogue of Theorem 104, but the price to be paid is a dependence on

$$\kappa[u] := \frac{\|u\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}}{\max\left\{\sqrt{d}, \|(x - x_0)\,u\|_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}\right\}} \quad \text{where} \quad x_0 = \int_{\mathbb{R}^d} x\, h_0\, d\gamma\,,$$

which goes as follows.

Proposition 109. *For all $u \in \mathrm{H}^1\big(\mathbb{R}^d, (1+|x|^2)\, d\gamma\big)$ such that $\int_{\mathbb{R}^d} x\, u^2\, d\gamma = 0$, and $u^2\,\gamma$ is log-concave, we have*

$$\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} - \frac{1}{2}\big(1 + (\mathscr{C}_\star - 1)\,\kappa[u]\big) \int_{\mathbb{R}^d} |u|^2 \log\left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)}}\right) d\gamma \ge 0\,.$$

Proof. We learn from (7.15) that

$$\int_{\mathbb{R}^d} |x|^2\, h(t,x)\, d\gamma = d\, \|u\|^2_{\mathrm{L}^2(\mathbb{R}^d,d\gamma)} + e^{-2t} \int_{\mathbb{R}^d} \big(|x|^2 - d\big)\, h_0\, d\gamma \quad \forall\, t \ge 0\,.$$

Hence [55, Theorem 1.2] and [55, Ineq. (3.4)] apply with

$$h(\mu) \ge \frac{\kappa[u]}{6\sqrt{3}}\,.$$

and the remainder of the proof of Theorem 104 is unchanged. □

A similar extension of Theorem 105 can be done on the same basis. Details are left to the reader. As for Proposition 103, we can make the following observations. The case $\int_{\mathbb{R}^d} |x|^2\, |v|^2\, d\gamma \le d$ is already covered in Lemma 106. If

$$A := \int_{\mathbb{R}^d} |u|^2 \big(|x|^2 - d\big)\, d\gamma$$

is positive, let us consider the solution h of (7.9) with initial datum $h_0 = u^2$. We know from (7.15) that

$$\int_{\mathbb{R}^d} h(t,x)\big(|x|^2 - d\big)\, d\gamma = A\, e^{-2t}$$

and deduce as in the proof of Lemma 106 that

$$4\mathcal{I}[v] = \int_{\mathbb{R}^d} |\nabla P|^2\, h\, d\gamma = \int_{\mathbb{R}^d} h\,(\mathscr{L} P)\, d\gamma + A\, e^{-2t} \le \sqrt{d \int_{\mathbb{R}^d} \|\mathrm{Hess}\, P\|^2\, h\, d\gamma} + A\, e^{-2t}$$

with $P = -\log h$. Hence by (7.13), we learn that

$$\frac{d}{dt}\mathcal{I}[v] + 2\mathcal{I}[v] = -\frac{1}{2}\int_{\mathbb{R}^d} \|\mathrm{Hess}\, P\|^2\, h\, d\gamma \le -\frac{1}{2d}\left(4\mathcal{I}[v] - A\, e^{-2t}\right)^2$$

and this estimate can be rephrased with $z(t) := e^{2t}\mathcal{I}[v(t,\cdot)]$ as

$$z' \le -\frac{e^{-2t}}{2d}\,(4z - A)^2\,.$$

Knowing that $z' < 0$ is an improvement on the decay rate $\mathcal{I}[v(t,\cdot)] \le \mathcal{I}[u]\, e^{-2t}$ can be rephrased as an improved entropy – entropy production inequality for $A > 0$.

Part IV

Nonlinear Dirichlet forms

Chapter 8

The normal contraction property for non-bilinear Dirichlet forms

This chapter corresponds to [P5], published in Potential Analysis. The work is a collaboration with Ivailo Hartarsky.

Abstract

We analyse the class of convex functionals $\mathcal{E}$ over $\mathrm{L}^2(X,m)$ for a measure space (X,m) introduced by Cipriani and Grillo [99] and generalising the classical bilinear Dirichlet forms. We investigate whether such non-bilinear forms verify the normal contraction property, i.e., if $\mathcal{E}(\phi\circ f) \leqslant \mathcal{E}(f)$ for all $f \in \mathrm{L}^2(X,m)$, and all 1-Lipschitz functions $\phi : \mathbb{R} \to \mathbb{R}$ with $\phi(0) = 0$. We prove that normal contraction holds if and only if $\mathcal{E}$ is symmetric in the sense $\mathcal{E}(-f) = \mathcal{E}(f)$, for all $f \in \mathrm{L}^2(X,m)$. An auxiliary result, which may be of independent interest, states that it suffices to establish the normal contraction property only for a simple two-parameter family of functions ϕ.

8.1 Introduction

8.1.1 Setting

Bilinear Dirichlet forms

Bilinear Dirichlet forms are a well-established topic, related to the theory of Markov processes and semigroups, see [70, 146, 203]. Let X be a nonempty set, let $\mathcal{F}$ be a $\sigma-$algebra over X, and take a $\sigma-$finite measure $m : \mathcal{F} \to [0,\infty]$. Let $\Lambda : D(\Lambda) \times D(\Lambda) \to \mathbb{R}$, be a symmetric, bilinear, and positive semi-definite form, such that $D(\Lambda) \subset \mathrm{L}^2(X,m)$ is dense. If the form is *closed*, there exists a unique self-adjoint, positive operator $A : D(A) \to \mathrm{L}^2(X,m)$, such that $D(A) \subset D(\Lambda)$, and

$$\langle Af, g\rangle = \Lambda(f,g), \quad \forall f \in D(A), g \in D(\Lambda).$$

Adopting the notation of functional calculus, we also have the formulae $D(\Lambda) = D(A^{1/2})$, and

$$\Lambda(f,g) = \langle A^{1/2} f, A^{1/2} g\rangle, \quad \forall f,g \in D(\Lambda).$$

The bilinear form Λ is called a (bilinear) Dirichlet form if

$$\Lambda(1\wedge f \vee 0, 1\wedge f\vee 0) \leqslant \Lambda(f,f), \quad \forall f \in D(\Lambda).$$

This conditions intuitively means that the form contracts if its argument is truncated. By extension, the term Dirichlet form also refers to the quadratic form

$$\mathcal{E}(f) = \begin{cases} \frac{1}{2}\Lambda(f,f), & iff \in D(\Lambda); \\ +\infty, & \text{otherwise;} \end{cases}$$

associated with a bilinear Dirichlet form Λ. This functional turns out to be always non-negative, convex (since it is quadratic), and lower semicontinuous. Moreover, the subdifferential satisfies $\partial\mathcal{E} = A$.

Bilinear Dirichlet forms correspond to self-adjoint, Markovian linear semigroups, via $T_t = \mathrm{e}^{At}$, for any $t \geqslant 0$. This was one idea behind the definition given in the section below.

Non-bilinear Dirichlet forms

We next turn to defining non-bilinear Dirichlet forms as they will be studied in the present work. Let $\mathcal{E} : \mathrm{L}^2(X,m) \to [0,\infty]$ be a convex and l.s.c. functional. In all the paper we assume that $\mathcal{E}$ is not the constant $+\infty$. Let $(T_t)_{t\geqslant 0}$ be the semigroup of nonlinear operators generated by $-\partial\mathcal{E}$, where ∂ denotes the subdifferential operator, via the differential equation

$$\begin{cases} \partial_t T_t f \in -\partial\mathcal{E}(T_t f), & \forall t \in (0,\infty), \quad \forall u \in \mathrm{L}^2(X,m), \\ T_0 f = f, & \forall f \in \mathrm{L}^2(X,m). \end{cases} \tag{8.1}$$

Equation (8.1) is well-posed for all $f \in \mathrm{L}^2(X,m)$. Its solution is usually called the gradient flow of $\mathcal{E}$ starting at f. See [6, 71] and refer to 8.1.2 for more background.

We say that a non-negative l.s.c. functional $\mathcal{E}$ is a *non-bilinear Dirichlet form* if $\mathcal{E}$ is convex and, for all $t \geqslant 0$, the operator $T_t : \mathrm{L}^2(X,m) \to \mathrm{L}^2(X,m)$ verifies

1. *order preservation*: $T_t f \leqslant T_t g$ for all $f,g \in \mathrm{L}^2(X,m)$ such that $f \leqslant g$ (for the pointwise order up to a negligible set);

2. L^∞*-contraction*: $\|T_t f - T_t g\|_\infty \leqslant \|f-g\|_\infty$ for all $f,g \in \mathrm{L}^2(X,m)$.

This class of forms was introduced by Cipriani and Grillo [99] and we will provide an equivalent "static" definition in Theorem 112 without reference to the underlying semigroup (also see Theorem 114).

Our main goal is to verify the normal contraction property for non-bilinear Dirichlet forms. A *normal contraction* is a 1−Lipschitz function $\phi : \mathbb{R} \to \mathbb{R}$, such that $\phi(0) = 0$. We denote by Φ the set of all normal contractions. We say that a functional $\mathcal{E}$ over $\mathrm{L}^2(X,m)$ has the *normal contraction property* if

$$\mathcal{E}(\phi(f)) \leqslant \mathcal{E}(f), \quad \forall \phi \in \Phi, \quad \forall f \in \mathrm{L}^2(X,m). \tag{8.2}$$

In the literature this property goes also under the name of *Second Beurling-Deny Criterion* since [233].

8.1.2 Background

Bilinear setting

Aside their interest in probability, for which we refer to the bibliography of [70, 146, 203], bilinear Dirichlet forms are also well-linked with linear diffusion equations and semigroups, see [24, 128]. This link gave fruitful results in the theory of metric measure spaces, allowing for an intrinsic/Eulerian approach towards Ricci curvature bounds, [10]. Under mild hypotheses, the authors of [10] could represent any bilinear Dirichlet form $\mathcal{E}$ as a quadratic Cheeger's energy on the base space X. One important point is that Ambrosio, Gigli, and Savaré were able to create an appropriate notion of distance $d_{\mathcal{E}}$ directly from the Dirichlet form $\mathcal{E}$. Then, via a condition *à la Bakry-Emery,* on the *carré du champ* associated with the quadratic form $\mathcal{E}$, the authors give a sense to notions such as Bochner's inequality or a lower bound on the Ricci curvature. Their approach is equivalent to that of Lott and Villani [201] and Sturm [245, 246], based on optimal transport. The creation of a distance from a bilinear form is a technique present also in [51]. Bilinear Dirichlet forms also play a role in potential and capacity theory, see [146, 237].

Historically, bilinear Dirichlet forms have been introduced by Beurling and Deny in [46]. One motivation behind their definition was the fact that being a bilinear Dirichlet form was sufficient to have the normal contraction property (see (8.2)). The fact that controlling one normal contraction is necessary and sufficient to control all of them is nowadays known as the *Beurling-Deny criterion.* To prove such a property, one usually approximates the function f with weighted sums of characteristic functions. The normal contraction property is a cornerstone for many purposes. For instance, for the development of a differential calculus [10] and the classification of linear Markov semigroups [146], both based on bilinear Dirichlet forms.

Non-bilinear setting

Generalising the concept of Dirichlet form to a non-bilinear setting is a more recent problem, started with the two works [52, 99]. A different kind of generalisation is that of [177], but we will not focus on it, since its purpose is different. Using instruments from [31, 42, 71], Cipriani and Grillo [99] provided two equivalent definitions of a non-bilinear Dirichlet form relevant to us, which will be discussed in further detail in Section 8.2. In [99], a number of properties of the class of non-bilinear Dirichlet forms are given, in particular with respect to $\Gamma-$convergence (see [105]).

Two recent works on the topic are [100, 101], where Claus recovers many structural properties for non-bilinear Dirichlet forms, among which we find a nonlinear Beurling–Deny principle, see [100, Theorem 2.39]. In the following sections, he develops a nonlinear theory of capacity. Furthermore, in [100, Corollary 2.40] (also see [101, Theorem 3.22]), the normal contraction property is proved for non-bilinear Dirichlet forms, but only for *non-decreasing* normal contractions and additionally assuming that the form is 0 at 0 (we avert the reader that in [100, Definition 2.31] non-decreasing normal contractions are named simply normal contractions).

Examples Let us mention two classes of basic examples, which generalise corresponding families of local and nonlocal bilinear Dirichlet forms. These lie at the core of the functionals analysed in the references quoted at the end of the section. Let Ω be an open subset of $\mathbb{R}^d$ and $f : \Omega \times \mathbb{R}^d \to \mathbb{R}$ be a Borel-measurable function. Let

$$\mathcal{E}(u) = \begin{cases} \int_\Omega f(x, Du)\,\mathrm{d}x & u \in \mathrm{W}^{1,2}_{\mathrm{loc}}(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{8.3}$$

We have that $\mathcal{E}$ is a non-bilinear Dirichlet form if f is non-negative, measurable in the first argument and convex and lower-semicontinuous in the second one. See [107] for the lower semicontinuity of the functional, while the property of being a non-bilinear Dirichlet form can be inferred as in [99, Theorem 4.1]. In addition, $\mathcal{E}$ is symmetric if $f(\cdot,-v) = f(\cdot,v)$, for all $v \in \mathbb{R}^d$. Finally, $\mathcal{E}$ is always local, due to the locality of Du and the fact that $\mathcal{E}$ is an integral functional. Among local forms, we can consider the following.

Example 110. *Let $\Omega = \mathbb{R}$. Let $f(x,v) = \max(v,0)$. Then, the integral functional $\mathcal{E}$ associated to f by (8.3) is a non-symmetric non-bilinear Dirichlet form, which does not satisfy the normal contraction property (8.2) for the function $\phi = -\mathrm{id}$.*

In this class of local functionals we also have the distinguished subclass of Finsler metrics, where

$$f(x,\cdot) = \|\cdot\|_x, \quad \forall x \in \Omega.$$

The form is bilinear if and only if, for all $x \in \Omega$, the norm $\|\cdot\|_x$ satisfies the parallelogram identity, see [72, Chapter 5].

Some non-local non-bilinear Dirichlet forms appear in [103], for example. In general we can say that any functional $\mathcal{E}$ of the form

$$\mathcal{E}(u) = \int_{\Omega^2} \psi(u(x)-u(y))\,\mathrm{d}x\,\mathrm{d}y, \quad \forall u \in \mathrm{L}^2(\Omega,\mathrm{d}x).$$

is a non-bilinear Dirichlet form for non-negative, l.s.c., convex ψ such that $\psi(0) = 0$. Lower semicontinuity of the functional comes from Fatou's Lemma, its convexity from the convexity of ψ. Finally, one can repeat the computations in [170, Theorem 2] to prove order-preservation and contraction in L^∞ for the semigroup associated with $\mathcal{E}$.

In [99], some interesting examples are developed in detail, ranging from functionals from the calculus of variations to Sobolev seminorms in the context of $C^\star$−monomodules. The theory of [99] can be applied to nonlinear diffusion equations (see [103, 141] and the references therein), analysis on graphs [168, 217], and analysis on spaces with a very irregular geometry [167, 210]. Furthermore, Cheeger's energies on extended metric spaces are known to be non-bilinear Dirichlet forms [8]. We refer to [7, 9, 10] for this theory, which originates from [94, 238]. See also [181, 202] for more estimates and contraction properties of Cheeger's energies.

8.1.3 Main results

Our main result is the following.

Theorem 111. *Let $\mathcal{E}$ be a non-bilinear Dirichlet form. Then $\mathcal{E}$ has the normal contraction property (8.2) if and only if*

$$\mathcal{E}(-f) \leqslant \mathcal{E}(f) \quad \forall f \in \mathrm{L}^2(X,m). \tag{8.4}$$

This theorem goes in the same direction as the well-established one for the bilinear case [70, 146, 203]. We merely prove that a form will operate on all normal contractions, once it operates on the simplest one. Henceforth, we say that $\mathcal{E}$ is *symmetric* if (8.4) holds and, equivalently, $\mathcal{E}(-f) = \mathcal{E}(f)$ for all $f \in \mathrm{L}^2(X,m)$. As witnessed by Example 110, the necessary symmetry assumption (8.4) needs to be made, since this non-bilinear Dirichlet form does not have the normal contraction property.

Let us highlight that Theorem 111 may be viewed as a strengthening of the result of Claus [100, Corollary 2.40], whose proof follows the far more conventional approach of [41, 42]. The class of normal contractions

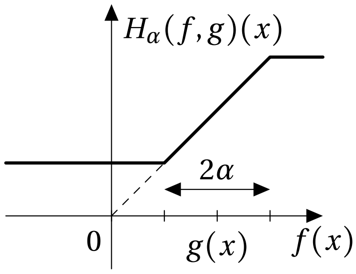

Figure 8.1: Graph of the function $H_\alpha(f,g)(x)$ for fixed $g(x)$.

we consider is richer and it controls, for example, the absolute value of the argument of the non-bilinear Dirichlet form, which can be very useful (see e.g. [146]), as well as more complicated contractions.

In order to prove Theorem 111, we establish two results, both of which may be of independent interest. Firstly, we provide an equivalent characterisation of non-bilinear Dirichlet forms, which turns out to be more widely for our purposes than the other equivalent static characterisation of [99, Theorem 3.8], recalled in Theorem114. To do so, we require a bit of notation. For all $f,g \in \mathrm{L}^2(X,m)$, and $\alpha \in [0,\infty)$ we denote by $f \vee g$ and $f \wedge g$ denote the pointwise maximum and minimum and set $H_\alpha(f,g) = (g-\alpha) \vee f \wedge (g+\alpha)$ (see Figure 8.1), that is,

$$H_\alpha(f,g)(x) = \begin{cases} g(x)-\alpha & f(x)-g(x) < -\alpha, \\ f(x) & f(x)-g(x) \in [-\alpha,\alpha], \\ g(x)+\alpha & f(x)-g(x) > \alpha. \end{cases} \tag{8.5}$$

Theorem 112. *Let $\mathcal{E} : \mathrm{L}^2(X,m) \to [0,\infty]$ be a l.s.c. functional. Then, $\mathcal{E}$ is a non-bilinear Dirichlet form if and only if, for all $f,g \in \mathrm{L}^2(X,m)$, and $\alpha \in [0,\infty)$, $\mathcal{E}$ verifies*

$$\mathcal{E}(f \vee g) + \mathcal{E}(f \wedge g) \leqslant \mathcal{E}(f) + \mathcal{E}(g), \tag{8.6}$$

$$\mathcal{E}(H_\alpha(f,g)) + \mathcal{E}(H_\alpha(g,f)) \leqslant \mathcal{E}(f) + \mathcal{E}(g). \tag{8.7}$$

The advantage of Theorem 112 as compared to Theorem 114 is that conditions (8.7)-(8.6) are easier to verify and useful to develop other functional inequalities such as the normal contraction property (8.2).

The second important step towards Theorem 111 is a reduction.

Lemma 113. *Let G be the set of all normal contractions $\phi \in \Phi$ such that $|\phi'| = 1$ and ϕ' has at most two points of discontinuity. Let $\langle G \rangle$ be the collection of all finite compositions of elements in G. Then, $\langle G \rangle$ is dense in Φ for the pointwise convergence on $\mathbb{R}$.*

We observe that the elements of G are irreducible with respect to composition, so that G is minimal in this sense. While the space G is quite simple, proving the normal contraction property (8.2) for $\phi \in G$ by hand from symmetry and (8.6)-(8.7) is still delicate, albeit elementary.

8.1.4 Plan of the paper

The remainder of the paper is structured as follows. In Section 8.2, we establish Theorem 112. In Section 8.3, we prove Theorem 111, relying on Theorem 112. This is the heart of our work. Finally, we discuss future directions of research in Section 8.4.

8.2 Efficient equivalent characterisation of non-bilinear Dirichlet forms

The goal of the present section is to prove Theorem 112.

8.2.1 Preliminaries

We introduce the subsets C_1 and $C_{2,\alpha}$, for $\alpha \in [0,\infty)$, of $\mathrm{L}^2(X,m;\mathbb{R}^2)$:

$$C_1 = \left\{(f,g) \in \mathrm{L}^2(X,m;\mathbb{R}^2) : f \leqslant g\right\}, \tag{8.8}$$

$$C_{2,\alpha} = \left\{(f,g) \in \mathrm{L}^2(X,m;\mathbb{R}^2) : |f-g| \leqslant \alpha\right\}. \tag{8.9}$$

We notice that for all α, the sets C_1 and $C_{2,\alpha}$ are convex and closed in the L^2−topology. For any closed and convex subset C, the 1−Lipschitz projection operator $P_C : \mathrm{L}^2(X,m;\mathbb{R}^2) \to C$ is defined by

$$P_C(f,g) = \operatorname*{arg\,min}_{(w,z)\in C} \|f-w\|_2^2 + \|g-z\|_2^2.$$

The projection map sends any point (f,g) to the closest point $P_C(f,g)$ in C. We denote by P_C^1 and P_C^2 the two components of the projection operator in $\mathrm{L}^2(X,m)$. More properties of projection maps are studied in [72]. If one considers the sets C_1 and $C_{2,\alpha}$, we have an explicit expression for the projections, thanks to [99, Lemma 3.3]:

$$P_1(f,g) = \left(f - \frac{1}{2}((f-g)\vee 0), g + \frac{1}{2}((f-g)\vee 0)\right), \tag{8.10}$$

$$P_{2,\alpha}(f,g) = \left(g + \frac{1}{2}\varphi_\alpha \circ (f-g), f - \frac{1}{2}\varphi_\alpha \circ (f-g)\right), \tag{8.11}$$

where $\varphi_\alpha : \mathbb{R} \to \mathbb{R}$ is given by

$$\varphi_\alpha(z) = ((z+\alpha)\vee 0) + ((z-\alpha)\wedge 0). \tag{8.12}$$

We further recall [99, Definition 3.1, Remark 3.2, Theorem 3.6].

Theorem 114. *Let $\mathcal{E} : \mathrm{L}^2(X,m) \to [0,\infty]$ be a l.s.c. functional. Then $\mathcal{E}$ is a non-bilinear Dirichlet form if and only if, for all $f,g \in \mathrm{L}^2(X,m)$ and $\alpha \in [0,\infty)$, $\mathcal{E}$ verifies*

$$\mathcal{E}\left(P_1^1(f,g)\right) + \mathcal{E}\left(P_1^2(f,g)\right) \leqslant \mathcal{E}(f) + \mathcal{E}(g), \tag{8.13}$$

$$\mathcal{E}\left(P_{2,\alpha}^1(f,g)\right) + \mathcal{E}\left(P_{2,\alpha}^2(f,g)\right) \leqslant \mathcal{E}(f) + \mathcal{E}(g). \tag{8.14}$$

The key argument is the well-known fact from [31, 71] stating that

$$\mathcal{E}\left(P_C^1(f,g)\right) + \mathcal{E}\left(P_C^2(f,g)\right) \leqslant \mathcal{E}(f) + \mathcal{E}(g)$$

for all $f,g \in \mathrm{L}^2(X,m)$ if and only if the semigroup T_t from (8.1) preserves C:

$$T_t C \subset C, \quad \forall t \geqslant 0,$$

where C can be any convex and closed set. Thus, (8.13)-(8.14) correspond to the order-preservation and the L^∞−contraction properties for $(T_t)_t$, respectively. In [99, Theorem 3.8] one more step is made.

Theorem 115. *Let $\mathcal{E} : \mathrm{L}^2(X,m) \to [0,\infty]$ be a l.s.c. functional. Then, $\mathcal{E}$ satisfies* (8.6) *if and only if $\mathcal{E}$ is convex and satisfies* (8.13).

Indeed, the last statement is a consequence of the more general [31, Proposition 2.5], which we will also use.

Theorem 116. *Let C be a closed convex subset of $\mathrm{L}^2(X,m;\mathbb{R}^2)$, let $P_C = (P_C^1, P_C^2)$ be the associated orthogonal projection. Let $\mathcal{E} : \mathrm{L}^2(X,m) \to [0,\infty]$ be a l.s.c. functional. Let $h, k : \mathrm{L}^2(X,m;\mathbb{R}^2) \to \mathrm{L}^2(X,m)$ be two continuous mappings such that, for all $u, v \in \mathrm{L}^2(X,m)$ and $t, s \in [0,1]$ it holds that*

$$h(u_t, v_s) = u_{1-s}, \qquad\qquad k(u_t, v_s) = v_{1-t}, \tag{8.15}$$

where

$$u_t = (1-t)u + th(u,v), \qquad\qquad v_s = (1-s)v + sk(u,v).$$

Moreover, assume

$$P_C(u,v) = (u_{1/2}, v_{1/2}). \tag{8.16}$$

Then, we have that for all $u, v \in \mathrm{L}^2(X,m)$

$$\mathcal{E}\left(P_C^1(u,v)\right) + \mathcal{E}\left(P_C^2(u,v)\right) \leqslant \mathcal{E}(u) + \mathcal{E}(v),$$

if and only if $\mathcal{E}$ is convex and for all $u, v \in \mathrm{L}^2(X,m)$

$$\mathcal{E}(h(u,v)) + \mathcal{E}(k(u,v)) \leqslant \mathcal{E}(u) + \mathcal{E}(v).$$

Remark 117. *Note that, given Theorems 114 and 115, it is easy to deduce that every non-bilinear Dirichlet form satisfies* (8.6)-(8.7)*, which is the direction of Theorem 112 we will use for proving Theorem 111. Indeed,*

$$H_\alpha(f,g) = \frac{1}{2}P^1_{2,\alpha}(f,g) + \frac{1}{2}P^2_{2,\alpha}(g,f)$$

for all $\alpha \geqslant 0$ and $f, g \in \mathrm{L}^2(X,m)$, so that convexity and (8.14) *give*

$$\begin{aligned}\mathcal{E}(H_\alpha(f,g)) + \mathcal{E}(H_\alpha(g,f))& \\ &\leqslant \frac{1}{2}\left(\mathcal{E}(P^1_{2,\alpha}(f,g)) + \mathcal{E}(P^2_{2,\alpha}(g,f)) + \mathcal{E}(P^1_{2,\alpha}(g,f)) + \mathcal{E}(P^2_{2,\alpha}(f,g))\right)\\ &\leqslant \mathcal{E}(f) + \mathcal{E}(g).\end{aligned}$$

8.2.2 Equivalence of the definitions

To conclude the section, we show that the convex sets $C_{2,\alpha}$ verify the hypotheses of Theorem 116.

Proof of Theorem 112. Fix $\alpha > 0$. Recalling the explicit expression of φ_α from (8.12), for any $u, v \in \mathrm{L}^2(X,m)$ we have

$$\varphi_\alpha \circ (u-v)(x) = \begin{cases} u(x) - v(x) - \alpha & u(x) - v(x) \leqslant -\alpha, \\ 2u(x) - 2v(x) & |u(x) - v(x)| \leqslant \alpha, \\ u(x) - v(x) + \alpha & u(x) - v(x) \geqslant \alpha. \end{cases}$$

Further recalling the expression of $P_{2,\alpha}$ from (8.11), in order to satisfy (8.16), we now choose $h,k : \mathrm{L}^2(X,m;\mathbb{R}^2) \to \mathrm{L}^2(X,m)$ such that

$$v + \frac{1}{2}\varphi_\alpha \circ (u-v) = \frac{u+h(u,v)}{2}, \qquad u - \frac{1}{2}\varphi_\alpha \circ (u-v) = \frac{v+k(u,v)}{2}.$$

Therefore, the expressions for h,k are the following

$$h(u,v)(x) = \begin{cases} v(x)-\alpha & u(x)-v(x) \leqslant -\alpha, \\ u(x) & |u(x)-v(x)| \leqslant \alpha, \\ v(x)+\alpha & u(x)-v(x) \geqslant \alpha, \end{cases}$$
$$k(u,v)(x) = \begin{cases} u(x)+\alpha & u(x)-v(x) \leqslant -\alpha, \\ v(x) & |u(x)-v(x)| \leqslant \alpha, \\ u(x)-\alpha & u(x)-v(x) \geqslant \alpha, \end{cases}$$

and we notice that $h(u,v) = H_\alpha(u,v)$ and $k(u,v) = H_\alpha(v,u)$.

It remains to verify the twist condition (8.15). Fix s,t,u,v as in the hypothesis. Since the values of H_α is defined pointwise, we also fix $x \in X$ and drop this parameter for compactness of notation. Suppose that $|u-v| \leqslant \alpha$, then $H(u,v) = u, H(v,u) = v$, so $u_t = u_{1-s} = u$, $v_s = v$. The case $u-v < -\alpha$ is analogous to that with $u-v > \alpha$, since the role of u and v is symmetric. Hence, we will discuss only the former. Here we have

$$u_t = (1-t)u + t(v-\alpha), \qquad v_s = (1-s)v + s(u+\alpha).$$

We need not discuss more subcases for the expression of $H_\alpha(u_t,v_s)$, since

$$\begin{aligned} u_t - v_s &= (1-t)u + t(v-\alpha) - (1-s)v - s(u+\alpha) \\ &= (1-t-s)(u-v) - (t+s)\alpha < -\alpha. \end{aligned}$$

Hence,

$$H_\alpha(u_t,v_s) = v_s - \alpha = (1-s)v + su + (s-1)\alpha = u_{1-s},$$

The second condition in (8.15) follows similarly, so we omit it. Thus, applying Theorem 116, (8.14) is equivalent to $\mathcal{E}$ being convex and (8.7). Yet, Theorem 115 gives that the convexity and (8.13) are equivalent to (8.6), so Theorem 112 reduces to Theorem 114. □

8.3 The normal contraction property

Throughout this section we fix a measure space (X,m) and a functional on $\mathrm{L}^2(X,m)$ satisfying symmetry and (8.6)-(8.7) for all $f,g \in \mathrm{L}^2(X,m)$ and $\alpha \in [0,\infty)$.

We will prove the normal contraction property (8.2) progressively, starting from simple functions ϕ. More specifically, for $k \in \{0,1,\dots\}$, $x_1,\dots,x_k \in \mathbb{R}$ such that $-\infty = x_0 < x_1 < \cdots < x_k < x_{k+1} = \infty$, we consider the continuous function $\phi_{x_1,\dots,x_k} : \mathbb{R} \to \mathbb{R}$ (see Figure 8.2) defined by $\phi_{x_1,\dots,x_k}(0) = 0$ and

$$\phi'_{x_1,\dots,x_k}(x) = (-1)^i \tag{8.17}$$

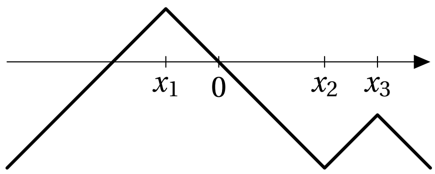

Figure 8.2: Graph of the function ϕ_{x_1,x_2,x_3}.

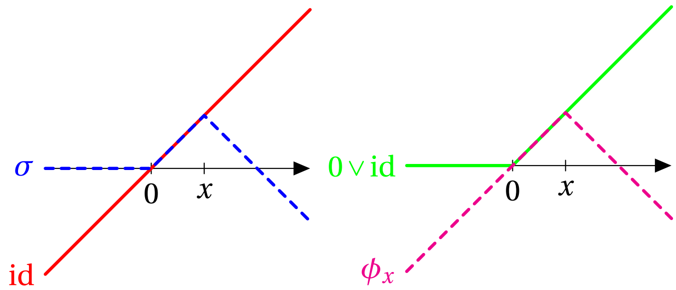

Figure 8.3: Illustration of (8.18).

for $x \in (x_i, x_{i+1})$. Let us denote $F_k = \{\phi_{x_1,\dots,x_k} : x_1 < \dots < x_k \in \mathbb{R}\}$, so that $F_0 = \{\text{id}\}$. We further set $\Phi_{x_1,\dots,x_k} = \mathcal{E} \circ \phi_{x_1,\dots,x_k}$.

8.3.1 Basic contractions

Proposition 118. *For any $x \in \mathbb{R}$ and $f \in L^2(X, m)$ we have $\Phi_x(f) \leqslant \mathcal{E}(f)$.*

Proof. Fix $x \geqslant 0$ (the case $x < 0$ is treated identically) and f. By (8.6)

$$\Phi_x(f) + \mathcal{E}(0 \vee f) \leqslant \mathcal{E}(f) + \mathcal{E}(\sigma \circ f) \tag{8.18}$$

(see Figure 8.3), where

$$\sigma(y) = \begin{cases} 0 & y \leqslant 0, \\ y & y \in (0, x), \\ 2x - y & y \geqslant x. \end{cases}$$

Thus, it suffices to show that $\mathcal{E}(0 \vee f) \geqslant \mathcal{E}(\sigma \circ f)$.

But symmetry and (8.7) with $\alpha = 2x$ (see Figure 8.4) give

$$\begin{aligned} 2\mathcal{E}(\sigma \circ f) &\leqslant \mathcal{E}(\sigma \circ f) + \mathcal{E}(-\sigma \circ f) \\ &\leqslant \mathcal{E}(0 \vee f) + \mathcal{E}(-(0 \vee f)) \leqslant 2\mathcal{E}(0 \vee f), \end{aligned} \tag{8.19}$$

concluding the proof. □

Proposition 119. *For any $0 \leqslant x_1 < x_2$ or $x_1 < x_2 \leqslant 0$ and $f \in L^2(X, m)$ it holds that $\Phi_{x_1,x_2}(f) \leqslant \mathcal{E}(f)$.*

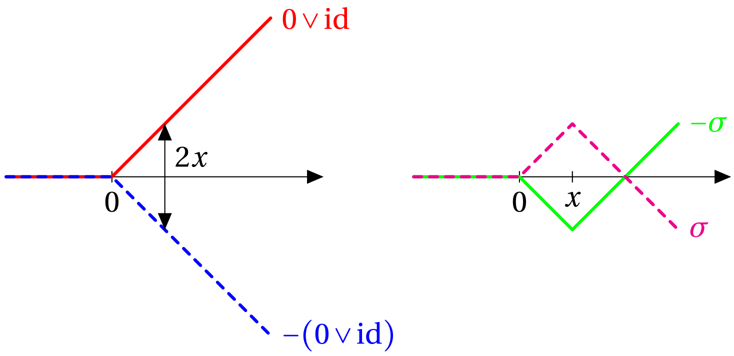

Figure 8.4: Illustration of (8.19).

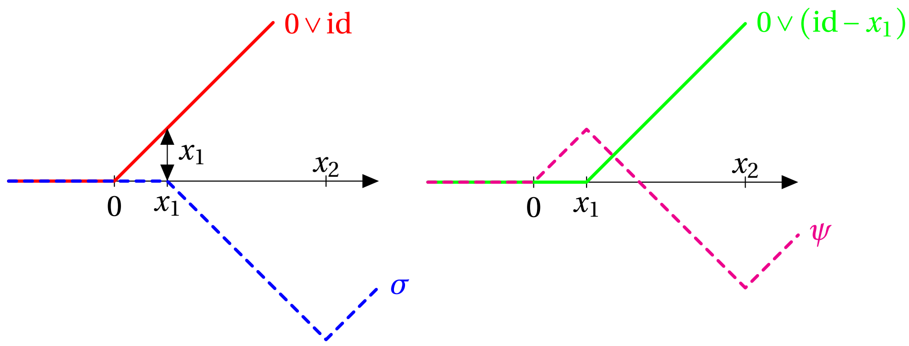

Figure 8.5: Illustration of (8.20).

Proof. Fix $0 \leqslant x_1 < x_2$ and f, the case $x_1 < x_2 \leqslant 0$ being analogous. Let

$$\sigma(x) = \begin{cases} 0 & x \leqslant x_1, \\ x_1 - x & x \in (x_1, x_2), \\ x + x_1 - 2x_2 & x \geqslant x_2, \end{cases} \qquad \psi(x) = \begin{cases} 0 & x \leqslant 0, \\ \phi_{x_1,x_2}(x) & x > 0. \end{cases}$$

Then (8.7) with $\alpha = x_1$ (see Figure 8.5) gives

$$\mathcal{E}(\psi \circ f) + \mathcal{E}(0 \vee (f - x_1)) \leqslant \mathcal{E}(0 \vee f) + \mathcal{E}(\sigma \circ f). \tag{8.20}$$

Moreover, by symmetry and (8.7) for $\alpha = 2(x_2 - x_1)$ (see Figure 8.6) we get

$$\begin{aligned} 2\mathcal{E}(0 \vee (f - x_1)) &\geqslant \mathcal{E}(0 \vee (f - x_1)) + \mathcal{E}(0 \wedge (x_1 - f)) \\ &\geqslant \mathcal{E}(\sigma \circ f) + \mathcal{E}(-\sigma \circ f) \geqslant 2\mathcal{E}(\sigma \circ f), \end{aligned} \tag{8.21}$$

so that $\mathcal{E}(\psi \circ f) \leqslant \mathcal{E}(0 \vee f)$. Furthermore, (8.6) gives

$$\Phi_{x_1,x_2}(f) + \mathcal{E}(0 \vee f) \leqslant \mathcal{E}(\psi \circ f) + \mathcal{E}(f) \tag{8.22}$$

(see Figure 8.7), yielding the desired conclusion. □

Proposition 120. *For any $x_1 < 0 < x_2$ and $f \in L^2(X, m)$ it holds that $\Phi_{x_1,x_2}(f) \leqslant \mathcal{E}(f)$.*

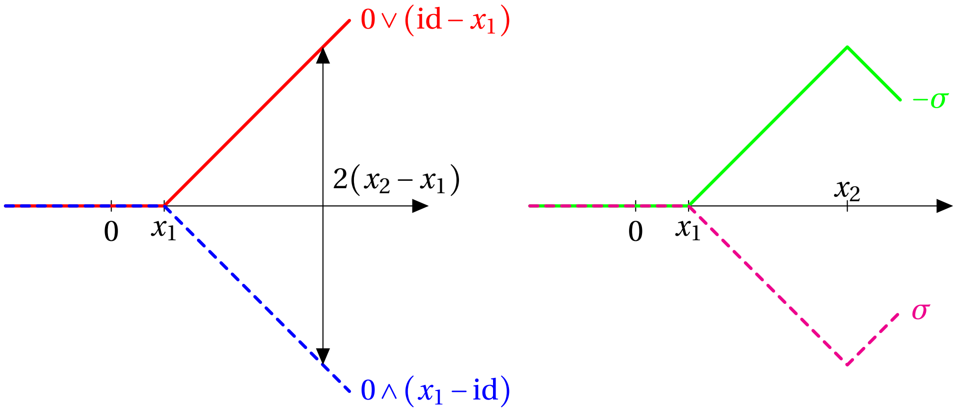

Figure 8.6: Illustration of (8.21).

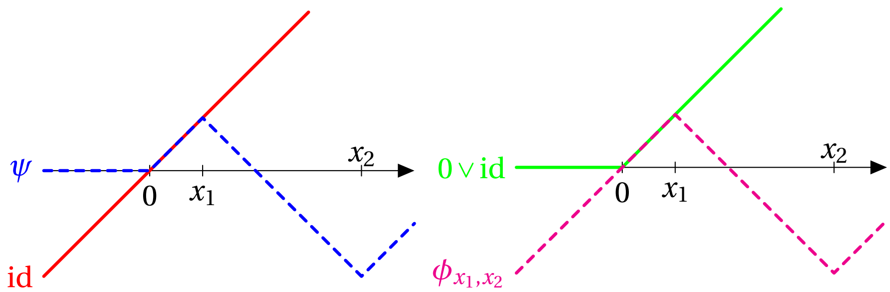

Figure 8.7: Illustration of (8.22).

Proof. Without loss of generality assume that $x_2 > -x_1$ and fix f. Consider

$$\psi(x) = \begin{cases} x - 2x_1 & x < x_1, \\ -x & x_1 \leqslant x \leqslant x_2, \\ -x_2 & x > x_2. \end{cases}$$

Then (8.7) with $\alpha = 2x_2$ (see Figure 8.8) gives

$$\Phi_{x_1,x_2}(f) + \mathcal{E}(f \wedge x_2) \leqslant \mathcal{E}(f) + \mathcal{E}(\psi \circ f). \tag{8.23}$$

Yet, $\psi = \phi_{x_1} \circ (\mathrm{id} \wedge x_2)$, so by Proposition 118 we have $\mathcal{E}(\psi \circ f) \leqslant \mathcal{E}(f \wedge x_2)$. Combining this with (8.23) yields the desired conclusion. □

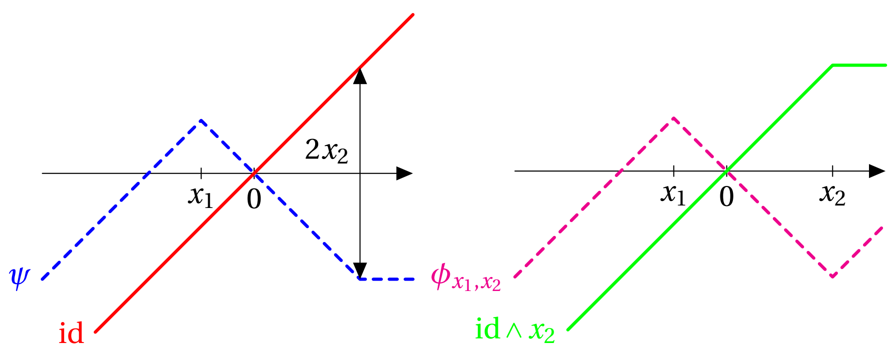

Figure 8.8: Illustration of (8.23).

8.3.2 Reduction to basic contractions

As we will see, the next proposition is essentially Lemma 113.

Proposition 121. *Any $\phi \in F_k$ with $k \geqslant 0$ can be written as $\phi^1 \circ \cdots \circ \phi^{\lfloor k/2 \rfloor} \circ \psi$ with $\phi^i \in F_2$ for all $i \in \{1, \dots, \lfloor k/2 \rfloor\}$ and $\psi \in F_{k-2\lfloor k/2 \rfloor}$.*

Proof. We proceed by induction on k. The statement is trivial for $k \in \{0,1,2\}$. Assume that $\phi = \phi_{x_1,\dots,x_k} \in F_k$ for $k \geqslant 3$, with $-\infty = x_0 < x_1 < \cdots < x_k < x_{k+1} = \infty$. Consider $i \in \{1, \dots, k-1\}$ such that $x_{i+1} - x_i < x_{j+1} - x_j$ for all $j \neq i$ (we may assume that the inequality is strict by perturbing the x_i and taking a limit if necessary). We consider the following cases.

- If $x_{i+1} \leqslant 0$, then set
$$x'_j = \begin{cases} x_j + 2(x_{i+1} - x_i) & 1 \leqslant j < i, \\ x_{j+2} & i \leqslant j \leqslant k-2. \end{cases}$$
- If $x_i \geqslant 0$, then set
$$x'_j = \begin{cases} x_j & 1 \leqslant j < i, \\ x_{j+2} - 2(x_{i+1} - x_i) & i \leqslant j \leqslant k-2. \end{cases}$$
- If $x_i < 0 < x_{i+1}$, then set
$$x'_j = \begin{cases} x_j - x_i & 1 \leqslant j < i, \\ x_{j+2} - x_{i+1} & i \leqslant j \leqslant k-2. \end{cases}$$

Then it suffices to prove that
$$\phi = \phi_{x'_1,\dots,x'_{k-2}} \circ \phi_{x_i,x_{i+1}}.$$
To do this, we verify (8.17) in each case. We will only treat the case $x_i \geqslant 0$, the others two being analogous. We have that
$$\phi'_{x'_1,\dots,x'_{k-2}}(\phi_{x_i,x_{i+1}}(x)) \times \phi'_{x_i,x_{i+1}}(x) \tag{8.24}$$
changes sign at x_i and x_{i+1} due to the second factor. Moreover, $\phi_{x_i,x_{i+1}}$ takes the values in $I = \mathbb{R} \setminus [2x_i - x_{i+1}, x_i]$ exactly once and
$$\phi_{x_i,x_{i+1}}(x_j) = \begin{cases} x'_j & 1 \leqslant j < i, \\ x'_{j-2} & i+2 \leqslant j \leqslant k. \end{cases}$$
But our choice of i implies $I \supset \{x'_1, \dots, x'_{k-2}\}$, so the first factor in (8.24) changes sign precisely at $x_1, \dots, x_{i-1}, x_{i+2}, \dots, x_k$, concluding the proof. □

With Proposition 121 it is immediate to deduce Lemma 113.

Proof of Lemma 113. Observe that $G = \{\mathrm{id}, -\mathrm{id}\} \circ (F_0 \cup F_1 \cup F_2)$. Thus,
$$\langle G \rangle \supset \{\mathrm{id}, -\mathrm{id}\} \circ \langle F_2 \rangle \circ (F_0 \cup F_1) \supset \{\mathrm{id}, -\mathrm{id}\} \circ \bigcup_{k=0}^{\infty} F_k \supset \langle G \rangle, \tag{8.25}$$
where the first and third inclusions follow by definition, while the second one is Proposition 121. Thus, $\langle G \rangle = \{\mathrm{id}, -\mathrm{id}\} \circ \bigcup_{k=0}^{\infty} F_k$. It therefore remains to show that $\langle G \rangle$ is dense in Φ, in order to conclude the proof.

To this extent, note that any $\phi \in \Phi$ coincides with its 1–Lipschitz envelope, i.e.,

$$\phi(x) = \inf_{y \in \mathbb{R}} \phi(y) + |x - y|, \quad \forall x \in \mathbb{R}.$$

By continuity,

$$\phi(x) = \inf_{y \in \mathbb{Q}} \phi(y) + |x - y|, \quad \forall x \in \mathbb{R}.$$

Taking a sequence of finite sets $(Q_n)_n \uparrow \mathbb{Q}$ with $Q_0 = \{0\}$, we can approximate ϕ with $\phi_n \in -\mathrm{id} \circ F_{2k_n - 1}$ for some $k_n \in \{1, \dots, |Q_n|\}$ given by

$$\phi_n(x) := \inf_{y \in Q_n} \phi(y) + |x - y|, \quad \forall x \in \mathbb{R}.$$

The limit $\phi_n \to \phi$ is in uniform convergence on compact sets thanks to equi-continuity, so the proof is complete. □

We are ready to assemble the proof of Theorem 111.

Proof of Theorem 111. By Theorem 112, any non-bilinear Dirichlet form $\mathcal{E}$ satisfies (8.7)-(8.6) and is l.s.c. Since symmetry is a hypothesis of Theorem 111, together with Propositions 120, 119, and 118 it yields that for any $\phi \in G$ (recall Lemma 113) and $f \in \mathrm{L}^2(X, m)$ it holds that $\mathcal{E}(\phi \circ f) \leqslant \mathcal{E}(f)$. Indeed, F_0 is trivial, Proposition 118 deals with F_1, Propositions 119 and 120 give F_2 and then symmetry allows us to take opposites. Therefore, the normal contraction property (8.2) also holds for all $\phi \in \langle G \rangle$.

Fix $f \in \mathrm{L}^2(X, m)$ and an arbitrary normal contraction $\phi \in \Phi$. By Lemma 113, there exists a sequence $\phi_n \in \langle G \rangle$ such that $\phi_n(x) \to \phi(x)$ for all $x \in \mathbb{R}$, as $n \to \infty$, and

$$\mathcal{E}(\phi_n \circ f) \leqslant \mathcal{E}(f)$$

for all n. We have that $\phi_n(f) \to \phi(f)$ pointwise in X, but

$$|\phi_n \circ f|^2 \leqslant |f|^2 \in \mathrm{L}^1(X, m),$$

as all functions ϕ_n are normal contractions. Then, by Lebesgue's dominated convergence theorem

$$\phi_n \circ f \to \phi \circ f$$

in $\mathrm{L}^2(X, m)$. Thus, we obtain the desired inequality via the l.s.c. of $\mathcal{E}$. □

8.3.3 Locality

Let us conclude this section with a concept of locality allowing a much more direct proof of Theorem 111 under this hypothesis. We say that a non-bilinear Dirichlet form $\mathcal{E}$ is *local* if for all $c \in \mathbb{R}$ and $u, v \in \mathrm{L}^2(X, m)$ such that $u(x)(v(x) - c) = 0$ for all $x \in X$, we have

$$\mathcal{E}(u + v) = \mathcal{E}(u) + \mathcal{E}(v).$$

Proof of Theorem 111 in the local case. Fix a symmetric local non-bilinear Dirichlet form $\mathcal{E}$. As in the proof of Theorem 111 it suffices to establish the normal contraction property (8.2) for all $\phi \in \bigcup_{k=1}^{\infty} F_k$ (this part of the proof does not rely on Theorem 112 and Propositions 118, 119, 120,and 121). Fix $\phi = \phi_{x_1, \dots, x_k}$ for some

$x_1 < \cdots < x_k$. Observe that

$$\phi(x) = ((x - x_1) \wedge 0) + \sum_{i=1}^{k} (-1)^i ((0 \vee (x - x_i)) \wedge (x_{i+1} - x_i)).$$

Since all summands satisfy the locality condition, we get

$$\begin{aligned}\mathcal{E}(\phi \circ u) &= \mathcal{E}((u - x_1) \wedge 0)) + \sum_{i=1}^{k} \mathcal{E}\big((-1)^i ((0 \vee (u - x_i)) \wedge (x_{i+1} - x_i))\big) \\ &= \mathcal{E}((u - x_1) \wedge 0)) + \sum_{i=1}^{k} \mathcal{E}((0 \vee (u - x_i)) \wedge (x_{i+1} - x_i)) = \mathcal{E}(u),\end{aligned}$$

using symmetry and locality for the second and third equalities. □

8.4 Future directions

Two challenges which are still open are the following. Firstly, we are not aware of any attempt to obtain a structural decomposition analogous to the one of [146] in the non-bilinear setting. Secondly, the theory of [201, 245, 246] covers even the case where Cheeger's energy of the metric measure space is a non-bilinear form, while an analogue of [10] for the non-bilinear case is missing. It is our opinion that the subject of metric measure spaces would profit from a study in this direction.

These two problems are strong motivations behind our paper, as we foresee that the normal contraction property would be crucial in developing such theories. One difficulty we anticipate is the generalisation of the computations in [24], which looks complicated even in the case of Finsler manifolds. Finally, establishing the normal contraction property adds one structural argument in favour of the choice made by Cipriani and Grillo of the generalisation of bilinear Dirichlet forms to the non-bilinear setting.

Part V

Bibliography

References

[1] F. ACHLEITNER, *A numerical computation of the rates for kinetic Fokker-Planck equation.* Personal communication.

[2] F. ACHLEITNER, A. ARNOLD, AND E. A. CARLEN, *On linear hypocoercive BGK models*, in From Particle Systems to Partial Differential Equations III, Springer, 2016, pp. 1–37.

[3] F. ACHLEITNER, A. ARNOLD, AND E. A. CARLEN, *On multi-dimensional hypocoercive BGK models*, Kinet. Relat. Models, 11 (2018), pp. 953–1009.

[4] F. ACHLEITNER, A. ARNOLD, AND B. SIGNORELLO, *On optimal decay estimates for ODEs and PDEs with modal decomposition*, in International workshop on Stochastic Dynamics out of Equilibrium, Springer, 2017, pp. 241–264.

[5] D. ALBRITTON, S. ARMSTRONG, J.-C. MOURRAT, AND M. NOVACK, *Variational methods for the kinetic Fokker-Planck equation*, Preprint arXiv:1902.04037, (2019).

[6] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[7] ——, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces*, Rev. Mat. Iberoam., 29 (2013), pp. 969–996.

[8] ——, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math., 195 (2014), pp. 289–391.

[9] ——, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J., 163 (2014), pp. 1405–1490.

[10] ——, *Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds*, Ann. Probab., 43 (2015), pp. 339–404.

[11] C. AMROUCHE, P. G. CIARLET, AND C. MARDARE, *On a lemma of Jacques-Louis Lions and its relation to other fundamental results*, Journal de Mathématiques Pures et Appliquées, 104 (2015), pp. 207–226.

[12] C. ANÉ, S. BLACHÈRE, D. CHAFAÏ, P. FOUGÈRES, I. GENTIL, F. MALRIEU, C. ROBERTO, AND G. SCHEFFER, *Sur les inégalités de Sobolev logarithmiques*, vol. 10, Société mathématique de France Paris, 2000.

[13] S. ARMSTRONG AND J.-C. MOURRAT, *Variational methods for the kinetic Fokker-Planck equation*, preprint arXiv:1902.04037, (2019).

[14] A. ARNOLD, J.-P. BARTIER, AND J. DOLBEAULT, *Interpolation between logarithmic Sobolev and Poincaré inequalities*, Commun. Math. Sci., 5 (2007), pp. 971–979.

[15] A. ARNOLD AND J. DOLBEAULT, *Refined convex Sobolev inequalities*, J. Funct. Anal., 225 (2005), pp. 337–351.

[16] A. ARNOLD, J. DOLBEAULT, C. SCHMEISER, AND T. WÖHRER, *Sharpening of Decay Rates in Fourier Based Hypocoercivity methods*, in Recent Advances in Kinetic Equations and Applications, Springer International Publishing, 2021, pp. 1–50.

[17] A. ARNOLD, A. EINAV, B. SIGNORELLO, AND T. WÖHRER, *Large time convergence of the non-homogeneous Goldstein-Taylor Equation*, Journal of statistical physics, 182 (2021), pp. 1–35.

[18] A. ARNOLD AND J. ERB, *Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift*, preprint arXiv:1409.5425, (2014).

[19] A. ARNOLD, P. MARKOWICH, G. TOSCANI, AND A. UNTERREITER, *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations*, Comm. Partial Differential Equations, 26 (2001), pp. 43–100.

[20] V. I. ARNOLD, *Mathematical methods of classical mechanics*, vol. 60 of Graduate Texts in Mathematics, Springer-Verlag, New York, [1989]. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[21] T. AUBIN, *Problemes isopérimétriques et espaces de Sobolev*, Journal of differential geometry, 11 (1976), pp. 573–598.

[22] D. BAKRY, F. BARTHE, P. CATTIAUX, AND A. GUILLIN, *A simple proof of the Poincaré inequality for a large class of probability measures*, Electronic Communications in Probability, 13 (2008), pp. 60 – 66.

[23] D. BAKRY AND M. ÉMERY, *Hypercontractivité de semi-groupes de diffusion*, C. R. Acad. Sci. Paris Sér. I Math., 299 (1984), pp. 775–778.

[24] D. BAKRY AND M. ÉMERY, *Diffusions hypercontractives*, in Séminaire de probabilités, XIX, 1983/84, vol. 1123 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 177–206.

[25] D. BAKRY AND M. ÉMERY, *Inégalités de Sobolev pour un semi-groupe symétrique*, Comptes rendus de l'Académie des sciences. Série 1, Mathématique, 301 (1985), pp. 411–413.

[26] D. BAKRY, I. GENTIL, AND M. LEDOUX, *Analysis and geometry of Markov diffusion operators*, vol. 348 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer, Cham, 2014.

[27] D. BAKRY AND M. LEDOUX, *Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator*, Invent. Math., 123 (1996), pp. 259–281.

[28] R. BALIAN, *From Microphysics to Macrophysics*, Methods and Applications of Statistical Physics, Springer Berlin Heidelberg, 1991.

[29] M. BARCHIESI, A. BRANCOLINI, AND V. JULIN, *Sharp dimension free quantitative estimates for the Gaussian isoperimetric inequality*, Ann. Probab., 45 (2017), pp. 668–697.

[30] J.-B. BARDET, N. GOZLAN, F. MALRIEU, AND P.-A. ZITT, *Functional inequalities for gaussian convolutions of compactly supported measures: Explicit bounds and dimension dependence*, Bernoulli, 24 (2018).

[31] L. BARTHÉLEMY, *Invariance d'un convexe fermé par un semi-groupe associé à une forme non-linéaire*, Abstr. Appl. Anal., 1 (1996), pp. 237–262.

[32] F. BAUDOIN, *Wasserstein contraction properties for hypoelliptic diffusions*, preprint arXiv:1602.04177, (2016).

[33] ——, *Bakry-Émery meet Villani*, J. Funct. Anal., 273 (2017), pp. 2275–2291.

[34] F. BAUDOIN, M. GORDINA, AND D. P. HERZOG, *Gamma calculus beyond Villani and explicit convergence estimates for Langevin dynamics with singular potentials*, Arch. Ration. Mech. Anal., 241 (2021), pp. 765–804.

[35] W. BECKNER, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc., 105 (1989), pp. 397–400.

[36] ——, *Sobolev inequalities, the Poisson semigroup, and analysis on the sphere* $\mathbb{S}^n$, Proc. Nat. Acad. Sci. U.S.A., 89 (1992), pp. 4816–4819.

[37] ——, *Sharp Sobolev inequalities on the sphere and the Moser-Trudinger inequality*, Ann. of Math. (2), 138 (1993), pp. 213–242.

[38] ——, *Geometric asymptotics and the logarithmic Sobolev inequality*, Forum Math., 11 (1999), pp. 105–137.

[39] W. BECKNER AND M. PEARSON, *On sharp Sobolev embedding and the logarithmic Sobolev inequality*, Bull. London Math. Soc., 30 (1998), pp. 80–84.

[40] J.-D. BENAMOU, T. O. GALLOUËT, AND F.-X. VIALARD, *Second-order models for optimal transport and cubic splines on the Wasserstein space*, Foundations of Computational Mathematics, 19 (2019), pp. 1113–1143.

[41] P. Bénilan and M. G. Crandall, *Completely accretive operators*, in Semigroup theory and evolution equations (Delft, 1989), vol. 135 of Lecture Notes in Pure and Appl. Math., Dekker, New York, 1991, pp. 41–75.

[42] P. Bénilan and C. Picard, *Quelques aspects non linéaires du principe du maximum*, in Séminaire de Théorie du Potentiel, No. 4 (Paris, 1977/1978), vol. 713 of Lecture Notes in Math., Springer, Berlin, 1979, pp. 1–37.

[43] M. Berger, P. Gauduchon, and E. Mazet, *Le spectre d'une variété Riemannienne*, Lecture Notes in Mathematics, Vol. 194, Springer-Verlag, Berlin-New York, 1971.

[44] E. Bernard, M. Fathi, A. Levitt, and G. Stoltz, *Hypocoercivity with schur complements*, Annales Henri Lebesgue, 5 (2022), pp. 523–557.

[45] A. Bernou, K. Carrapatoso, S. Mischler, and I. Tristani, *Hypocoercivity for kinetic linear equations in bounded domains with general Maxwell boundary condition*, Annales de l'Institut Henri Poincaré C, (2022).

[46] A. Beurling and J. Deny, *Dirichlet spaces*, Proc. Nat. Acad. Sci. U.S.A., 45 (1959), pp. 208–215.

[47] G. Bianchi and H. Egnell, *A note on the Sobolev inequality*, Journal of Functional Analysis, 100 (1991), pp. 18–24.

[48] M.-F. Bidaut-Véron and L. Véron, *Nonlinear elliptic equations on compact Riemannian manifolds and asymptotics of Emden equations*, Invent. Math., 106 (1991), pp. 489–539.

[49] ——, *Erratum: "Nonlinear elliptic equations on compact Riemannian manifolds and asymptotics of Emden equations" [Invent. Math.* **106** *(1991), no. 3, 489–539; MR1134481 (93a:35045)]*, Invent. Math., 112 (1993), p. 445.

[50] I. Bihari, *A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations*, Acta Mathematica Academiae Scientiarum Hungarica, 7 (1956), pp. 81–94.

[51] M. Biroli and U. Mosco, *Sobolev and isoperimetric inequalities for Dirichlet forms on homogeneous spaces*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., 6 (1995), pp. 37–44.

[52] M. Biroli and P. G. Vernole, *Strongly local nonlinear Dirichlet functionals and forms*, Adv. Math. Sci. Appl., 15 (2005), pp. 655–682.

[53] N. Blachman, *The convolution inequality for entropy powers*, IEEE Transactions on Information theory, 11 (1965), pp. 267–271.

[54] A. Blanchet, M. Bonforte, J. Dolbeault, G. Grillo, and J.-L. Vázquez, *Hardy–Poincaré inequalities and applications to nonlinear diffusions*, Comptes Rendus Mathématique, 344 (2007), pp. 431–436.

[55] S. G. Bobkov, *Isoperimetric and analytic inequalities for log-concave probability measures*, Ann. Probab., 27 (1999), pp. 1903–1921.

[56] S. G. Bobkov, *On isoperimetric constants for log-concave probability distributions*, in Geometric aspects of functional analysis, Springer, 2007, pp. 81–88.

[57] S. G. Bobkov and D. Cordero-Erausquin, *KLS-type isoperimetric bounds for log-concave probability measures*, Ann. Mat. Pura Appl. (4), 195 (2016), pp. 681–695.

[58] S. G. Bobkov, N. Gozlan, C. Roberto, and P.-M. Samson, *Bounds on the deficit in the logarithmic Sobolev inequality*, J. Funct. Anal., 267 (2014), pp. 4110–4138.

[59] T. Bodineau, I. Gallagher, and L. Saint-Raymond, *From hard sphere dynamics to the Stokes–Fourier equations: An analysis of the Boltzmann–Grad limit*, Annals of PDE, 3 (2017), p. 2.

[60] M. E. Bogovskii, *Solution of the first boundary value problem for the equation of continuity of an incompressible medium*, in Soviet Math. Dokl., vol. 20, 1979, pp. 1094–1098.

[61] F. Bolley, I. Gentil, and A. Guillin, *Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities*, Ann. Probab., 46 (2018), pp. 261–301.

[62] L. Boltzmann, *Weitere studien über das wärmegleichgewicht unter gasmolekülen*, Aus der kk Hot-und Staatsdruckerei, 1872.

[63] M. Bonforte, J. Dolbeault, B. Nazaret, and N. Simonov, *Stability in Gagliardo-Nirenberg-Sobolev inequalities: Flows, regularity and the entropy method*, Preprint arXiv: 2007.03674 and hal-03160022, to appear in *Memoirs of the AMS*, (2022).

[64] C. Borell, *Convex set functions in d-space*, Periodica Mathematica Hungarica, 6 (1975), pp. 111–136.

[65] E. Bouin, J. Dolbeault, and L. Lafleche, *Fractional hypocoercivity*, Communications in Mathematical Physics, 390 (2022), pp. 1369–1411.

[66] E. Bouin, J. Dolbeault, L. Lafleche, and C. Schmeiser, *Hypocoercivity and sub-exponential local equilibria*, Monatsh. Math., 194 (2021), pp. 41–65.

[67] E. Bouin, J. Dolbeault, S. Mischler, C. Mouhot, and C. Schmeiser, *Hypocoercivity without confinement*, Pure Appl. Anal., 2 (2020), pp. 203–232.

[68] E. Bouin, J. Dolbeault, and L. Ziviani, l^2 *Hypocoercivity methods for kinetic Fokker-Planck equations with factorised Gibbs states*, arXiv preprint arXiv:2304.12040, (2023).

[69] E. Bouin, J. Dolbeaut, and L. Ziviani, *Decay estimates for kinetic FokkerPlanck equations with heavy tailed stationary solutions*, 2023.

[70] N. Bouleau and F. Hirsch, *Dirichlet forms and analysis on Wiener space*, vol. 14 of De Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1991.

[71] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies, No. 5, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.

[72] ———, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.

[73] H. Brezis and E. H. Lieb, *Sobolev inequalities with remainder terms*, J. Funct. Anal., 62 (1985), pp. 73–86.

[74] G. Brigati, J. Dolbeault, and N. Simonov, *Logarithmic Sobolev and interpolation inequalities on the sphere: constructive stability results*, Preprint arXiv:2211.13180 and hal-03868496, (2022).

[75] ———, *On Gaussian interpolation inequalities*, Preprint arXiv: 2302.03926 and hal-03978122, to appear in C.R. Mathématique, (2023).

[76] P. Buser, *A note on the isoperimetric constant*, Ann. Sci. École Norm. Sup. (4), 15 (1982), pp. 213–230.

[77] A. P. Calderón and A. Zygmund, *On singular integrals*, American Journal of Mathematics, 78 (1956), pp. 289–309.

[78] E. Camrud, D. P. Herzog, G. Stoltz, and M. Gordina, *Weighted L^2-contractivity of Langevin dynamics with singular potentials*, Nonlinearity, 35 (2021), p. 998.

[79] Y. Cao, J. Lu, and L. Wang, *On explicit L^2-convergence rate estimate for underdamped Langevin dynamics*, preprint arXiv:1908.04746, (2019).

[80] E. Carlen and M. Loss, *Competing symmetries, the logarithmic HLS inequality and Onofri's inequality on S^n*, Geom. Funct. Anal., 2 (1992), pp. 90–104.

[81] E. A. Carlen, *Superadditivity of Fisher's information and logarithmic Sobolev inequalities*, J. Funct. Anal., 101 (1991), pp. 194–211.

[82] E. A. Carlen, M. C. Carvalho, J. L. Roux, M. Loss, and C. Villani, *Entropy and chaos in the Kac model*, arXiv preprint arXiv:0808.3192, (2008).

[83] E. A. Carlen and D. W. Stroock, *An application of the Bakry-Émery criterion to infinite dimensional diffusions*, Séminaire de probabilités de Strasbourg, 20 (1986), pp. 341–348.

[84] K. Carrapatoso, J. Dolbeault, F. Hérau, S. Mischler, and C. Mouhot, *Weighted Korn and Poincaré-Korn inequalities in the euclidean space and associated operators*, Archive for Rational Mechanics and Analysis, 243 (2022), pp. 1565–1596.

[85] K. Carrapatoso and S. Mischler, *Landau equation for very soft and Coulomb potentials near Maxwellians*, Ann. PDE, 3 (2017), pp. Paper No. 1, 65.

[86] J. A. CARRILLO, M. G. DELGADINO, L. DESVILLETTES, AND J. WU, *The Landau equation as a gradient flow*, Preprint arXiv:2007.08591, (2020).

[87] J. A. CARRILLO, L. WANG, W. XU, AND M. YAN, *Variational Asymptotic Preserving Scheme for the VlasovPoissonFokkerPlanck System*, Multiscale Modeling & Simulation, 19 (2021), pp. 478–505.

[88] P. CATTIAUX AND A. GUILLIN, *Functional inequalities for perturbed measures with applications to log-concave measures and to some bayesian problems*, arXiv preprint arXiv:2101.11257, (2021).

[89] P. CATTIAUX, A. GUILLIN, P. MONMARCHÉ, AND C. ZHANG, *Entropic multipliers method for Langevin diffusion and weighted log-Sobolev inequalities*, Journal of Functional Analysis, 277 (2019), p. 108288.

[90] D. CHAFAI, *Sur les inégalités de Sobolev logarithmiques en théorie de l'information et pour des systèmes de spins conservatifs en mécanique statistique*, PhD thesis, Université Paul Sabatier-Toulouse III, 2002.

[91] D. CHAFAÏ, *Entropies, convexity, and functional inequalities: on* Φ*-entropies and* Φ*-Sobolev inequalities*, J. Math. Kyoto Univ., 44 (2004), pp. 325–363.

[92] S. CHANDRASEKHAR, *Stochastic problems in physics and astronomy*, Rev. Mod. Phys., 15 (1943), pp. 1–89.

[93] J. 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), Princeton Univ. Press, Princeton, N.J., 1970, pp. 195–199.

[94] J. CHEEGER, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal., 9 (1999), pp. 428–517.

[95] H.-B. CHEN, S. CHEWI, AND J. NILES-WEED, *Dimension-free log-sobolev inequalities for mixture distributions*, arXiv preprint arXiv:2102.11476, (2021).

[96] L. CHEN, G. LU, AND H. TANG, *Sharp stability of Log-Sobolev and Moser-Onofri inequalities on the sphere*, 2022.

[97] S. CHEN, R. L. FRANK, AND T. WETH, *Remainder terms in the fractional Sobolev inequality*, Indiana Univ. Math. J., 62 (2013), pp. 1381–1397.

[98] M. CICALESE AND G. P. LEONARDI, *A selection principle for the sharp quantitative isoperimetric inequality*, Archive for Rational Mechanics and Analysis, 206 (2012), pp. 617–643.

[99] F. CIPRIANI AND G. GRILLO, *Nonlinear Markov semigroups, nonlinear Dirichlet forms and applications to minimal surfaces*, Journal für die reine und angewandte Mathematik, (2003).

[100] B. CLAUS, *Nonlinear Dirichlet Forms*, PhD thesis, Technischen Universität Dresden, 2021.

[101] ——, *Energy spaces, Dirichlet forms and capacities in a nonlinear setting*, Potential Anal., (To appear).

[102] T. A. COURTADE, M. FATHI, AND A. PANANJADY, *Quantitative stability of the entropy power inequality*, IEEE Transactions on Information Theory, 64 (2018), pp. 5691–5703.

[103] S. CREO AND M. R. LANCIA, *Fractional* (s,p)*-Robin-Venttsel' problems on extension domains*, NoDEA Nonlinear Differential Equations Appl., 28 (2021), pp. Paper No. 31, 33.

[104] G. CSATÓ, B. DACOROGNA, AND O. KNEUSS, *The pullback equation for differential forms*, vol. 83, Springer Science & Business Media, 2011.

[105] G. DAL MASO, *An introduction to* Γ*-convergence*, vol. 8 of Progress in nonlinear differential equations and their applications, Birkhäuser Boston, Inc., Boston, MA, 1993.

[106] N. G. DE BRUIJN, *A combinatorial problem*, Proceedings of the Section of Sciences of the Koninklijke Nederlandse Akademie van Wetenschappen te Amsterdam, 49 (1946), pp. 758–764.

[107] E. DE GIORGI, G. BUTTAZZO, AND G. DAL MASO, *On the lower semicontinuity of certain integral functionals*, Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, 74 (1983), pp. 274–282.

[108] P. DEGOND, *Global existence of smooth solutions for the vlasov-fokker-planck equation in* 1 *and* 2 *space dimensions*, in Annales scientifiques de l'École Normale Supérieure, vol. 19, 1986, pp. 519–542.

[109] P. DEGOND AND B. LUCQUIN-DESREUX, *The Fokker-Planck asymptotics of the Boltzmann collision operator in the Coulomb case*, Mathematical Models and Methods in Applied Sciences, 2 (1992), pp. 167–182.

[110] M. Del Pino and J. Dolbeault, *Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions*, J. Math. Pures Appl. (9), 81 (2002), pp. 847–875.

[111] M. del Pino and M. Sáez, *On the extinction profile for solutions of* $u_t = \Delta u^{(N-2)/(N+2)}$, Indiana Univ. Math. J., 50 (2001), pp. 611–628.

[112] J. Demange, *Des équations à diffusion rapide aux inégalités de Sobolev sur les modèles de la géométrie*, PhD thesis, Université Paul Sabatier Toulouse 3, 2005.

[113] ——, *Improved Gagliardo-Nirenberg-Sobolev inequalities on manifolds with positive curvature*, J. Funct. Anal., 254 (2008), pp. 593–611.

[114] L. Desvillettes and C. Villani, *On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation*, Comm. Pure Appl. Math., 54 (2001), pp. 1–42.

[115] L. Desvillettes and C. Villani, *On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the boltzmann equation*, Inventiones mathematicae, 159 (2005), pp. 245–316.

[116] H. Dietert, J. Evans, and T. Holding, *Contraction in the wasserstein metric for the kinetic Fokker-Planck equation on the torus*, Kinetic & Related Models, 11 (2018).

[117] H. Dietert, F. Hérau, H. Hutridurga, and C. Mouhot, *Quantitative geometric control in linear kinetic theory*, preprint arXiv:2209.09340, (2022).

[118] J. Dolbeault and M. J. Esteban, *Improved interpolation inequalities and stability*, Adv. Nonlinear Stud., 20 (2020), pp. 277–291.

[119] J. Dolbeault, M. J. Esteban, A. Figalli, R. L. Frank, and M. Loss, *Sharp stability for Sobolev and log-Sobolev inequalities, with optimal dimensional dependence*, 2022.

[120] J. Dolbeault, M. J. Esteban, M. Kowalczyk, and M. Loss, *Improved interpolation inequalities on the sphere*, Discrete and Continuous Dynamical Systems Series S (DCDS-S), 7 (2014), pp. 695–724.

[121] J. Dolbeault, M. J. Esteban, and M. Loss, *Nonlinear flows and rigidity results on compact manifolds*, J. Funct. Anal., 267 (2014), pp. 1338–1363.

[122] ——, *Interpolation inequalities on the sphere: linear vs. nonlinear flows*, Ann. Fac. Sci. Toulouse Math. (6), 26 (2017), pp. 351–379.

[123] J. Dolbeault, M. J. Esteban, and M. Loss, *Symmetry and symmetry breaking: rigidity and flows in elliptic PDEs*, in Proceedings of the International Congress of Mathematicians (ICM 2018) (In 4 Volumes) Proceedings of the International Congress of Mathematicians 2018, World Scientific, 2018, pp. 2261–2285.

[124] J. Dolbeault, M. J. Esteban, M. Loss, and M. Muratori, *Symmetry for extremal functions in subcritical Caffarelli–Kohn–Nirenberg inequalities*, Comptes Rendus Mathématique, 355 (2017), pp. 133 – 154.

[125] J. Dolbeault and G. Jankowiak, *Sobolev and Hardy–Littlewood–Sobolev inequalities*, J. Differential Equations, 257 (2014), pp. 1689–1720.

[126] J. Dolbeault and X. Li, *φ-entropies: convexity, coercivity and hypocoercivity for Fokker-Planck and kinetic Fokker-Planck equations*, Math. Models Methods Appl. Sci., 28 (2018), pp. 2637–2666.

[127] J. Dolbeault, C. Mouhot, and C. Schmeiser, *Hypocoercivity for linear kinetic equations conserving mass*, Transactions of the American Mathematical Society, 367 (2015), pp. 3807–3828.

[128] J. Dolbeault, B. Nazaret, and G. Savaré, *On the Bakry-Emery criterion for linear diffusions and weighted porous media equations*, Commun. Math. Sci., 6 (2008), pp. 477–494.

[129] J. Dolbeault and G. Toscani, *Improved interpolation inequalities, relative entropy and fast diffusion equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 30 (2013), pp. 917–934.

[130] ——, *Stability results for logarithmic Sobolev and Gagliardo-Nirenberg inequalities*, Int. Math. Res. Not. IMRN, (2016), pp. 473–498.

[131] J. Dolbeault and A. Zhang, *Optimal functional inequalities for fractional operators on the sphere and applications*, Adv. Nonlinear Stud., 16 (2016), pp. 863–880.

[132] R. DUAN, *Hypocoercivity of linear degenerately dissipative kinetic equations*, Nonlinearity, 24 (2011), p. 2165.

[133] M. H. DUONG, M. A. PELETIER, AND J. ZIMMER, *Conservative-dissipative approximation schemes for a generalized Kramers equation*, Mathematical Methods in the Applied Sciences, 37 (2014), pp. 2517–2540.

[134] A. EBERLE, A. GUILLIN, AND R. ZIMMER, *Couplings and quantitative contraction rates for Langevin dynamics*, The Annals of Probability, 47 (2019), pp. 1982–2010.

[135] R. ELDAN, J. LEHEC, AND Y. SHENFELD, *Stability of the logarithmic Sobolev inequality via the Föllmer process*, Ann. Inst. Henri Poincaré Probab. Stat., 56 (2020), pp. 2253–2269.

[136] M. ERBAR, C. RIGONI, K.-T. STURM, AND L. TAMANINI, *Tamed spaces–Dirichlet spaces with distribution-valued Ricci bounds*, Journal de Mathématiques Pures et Appliquées, 161 (2022), pp. 1–69.

[137] L. C. EVANS, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010.

[138] M. FATHI, E. INDREI, AND M. LEDOUX, *Quantitative logarithmic Sobolev inequalities and stability estimates*, Discrete Contin. Dyn. Syst., 36 (2016), pp. 6835–6853.

[139] P. FEDERBUSH, *Partially alternate derivation of a result of Nelson*, J. Mathematical Phys., 10 (1969), pp. 50–52.

[140] F. FEO, E. INDREI, M. R. POSTERARO, AND C. ROBERTO, *Some remarks on the stability of the log-Sobolev inequality for the Gaussian measure*, Potential Anal., 47 (2017), pp. 37–52.

[141] F. FEO, J. L. VÁZQUEZ, AND B. VOLZONE, *Anisotropic p-Laplacian evolution of fast diffusion type*, Adv. Nonlinear Stud., 21 (2021), pp. 523–555.

[142] A. FIGALLI AND Y. R.-Y. ZHANG, *Sharp gradient stability for the Sobolev inequality*, Duke Mathematical Journal, 171 (2022), pp. 2407–2459.

[143] M. FRADELIZI, N. GOZLAN, AND S. ZUGMEYER, *Transport proofs of some functional inverse Santaló inequalities*, arXiv preprint arXiv:2109.00871, (2021).

[144] R. L. FRANK, *Degenerate stability of some Sobolev inequalities*, Ann. Inst. H. Poincaré Anal. Non Linéaire (Online first). Doi: 10.4171/AIHPC/35, (2022).

[145] R. L. FRANK AND E. H. LIEB, *A New, Rearrangement-free Proof of the Sharp Hardy–Littlewood–Sobolev Inequality*, in Spectral Theory, Function Spaces and Inequalities, B. M. Brown, J. Lang, and I. G. Wood, eds., Basel, 2012, Springer Basel, pp. 55–67.

[146] M. FUKUSHIMA, Y. OSHIMA, AND M. TAKEDA, *Dirichlet forms and symmetric Markov processes*, vol. 19 of De Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, extended ed., 2011.

[147] E. GAGLIARDO, *Proprietà di alcune classi di funzioni in più variabili*, Ricerche Mat., 7 (1958), pp. 102–137.

[148] U. GIANAZZA, G. SAVARÉ, AND G. TOSCANI, *The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation*, Arch. Ration. Mech. Anal., 194 (2009), pp. 133–220.

[149] N. GIGLI, *Nonsmooth differential geometry–An approach tailored for spaces with Ricci curvature bounded from below*, vol. 251, American Mathematical Society, 2018.

[150] N. GIGLI, C. KETTERER, K. KUWADA, AND S.-I. OHTA, *Rigidity for the spectral gap on* $\mathrm{RCD}(K,\infty)$*-spaces*, Amer. J. Math., 142 (2020), pp. 1559–1594.

[151] A. N. GORBAN, *Hilberts sixth problem: the endless road to rigour*, 2018.

[152] T. GÖTZ, A. KLAR, N. MARHEINEKE, AND R. WEGENER, *A stochastic model and associated Fokker–Planck equation for the fiber lay-down process in nonwoven production processes*, SIAM Journal on Applied Mathematics, 67 (2007), pp. 1704–1717.

[153] H. GRAD, *On the kinetic theory of rarefied gases*, Communications on pure and applied mathematics, 2 (1949), pp. 331–407.

[154] P. GRISVARD, *Elliptic problems in nonsmooth domains*, vol. 69 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Reprint of the 1985 original edition, With a foreword by Susanne C. Brenner.

[155] L. GROSS, *Logarithmic Sobolev inequalities*, Amer. J. Math., 97 (1975), pp. 1061–1083.

[156] ——, *Logarithmic Sobolev inequalities and contractivity properties of semigroups*, in Dirichlet forms (Varenna, 1992), vol. 1563 of Lecture Notes in Math., Springer, Berlin, 1993, pp. 54–88.

[157] M. GROTHAUS AND P. STILGENBAUER, *Hypocoercivity for Kolmogorov backward evolution equations and applications*, Journal of Functional Analysis, 267 (2014), pp. 3515–3556.

[158] M. GROTHAUS AND F.-Y. WANG, *Weak Poincaré inequalities for convergence rate of degenerate diffusion processes*, The Annals of Probability, 47 (2019), pp. 2930–2952.

[159] C. GUI AND A. MORADIFAM, *The sphere covering inequality and its applications*, Invent. Math., 214 (2018), pp. 1169–1204.

[160] A. GUIONNET AND B. ZEGARLINSKI, *Lectures on Logarithmic Sobolev Inequalities*, Séminaire de probabilités de Strasbourg, 36 (2002), pp. 1–134.

[161] Y. GUO, *The Landau equation in a periodic box*, Communications in mathematical physics, 231 (2002), pp. 391–434.

[162] M. HAIRER AND J. C. MATTINGLY, *Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing*, Annals of Mathematics, (2006), pp. 993–1032.

[163] D. HAN-KWAN AND M. IACOBELLI, *From Newtons second law to Eulers equations of perfect fluids*, Proceedings of the American Mathematical Society, 149 (2021), pp. 3045–3061.

[164] F. HÉRAU, *Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation*, Asymptot. Anal., 46 (2006), pp. 349–359.

[165] F. HÉRAU AND F. NIER, *Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential*, Arch. Ration. Mech. Anal., 171 (2004), pp. 151–218.

[166] D. HILBERT, *Mathematische Probleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 1900 (1900), pp. 253–297.

[167] M. HINZ, D. KOCH, AND M. MEINERT, *Sobolev spaces and calculus of variations on fractals*, in Analysis, Probability and Mathematical Physics on Fractals, World Scientific, 2020, pp. 419–450.

[168] M. HOFMANN, *Spectral theory, clustering problems and differential equations on metric graphs*, PhD thesis, Universidade de Lisboa, 2021.

[169] L. HÖRMANDER, *Hypoelliptic second order differential equations*, Acta Mathematica, 119 (1967), pp. 147–171.

[170] E. J. HURTADO, *Non-local Diffusion Equations Involving the Fractional $p(\cdot)$-Laplacian*, Journal of Dynamics and Differential Equations, 32 (2020), pp. 557–587.

[171] E. INDREI, *Sharp stability for LSI*, Preprint arXiv: 2303.05604, (2023).

[172] E. INDREI AND D. KIM, *Deficit estimates for the logarithmic Sobolev inequality*, Differential Integral Equations, 34 (2021), pp. 437–466.

[173] E. INDREI AND D. MARCON, *A quantitative log-Sobolev inequality for a two parameter family of functions*, Int. Math. Res. Not. IMRN, (2014), pp. 5563–5580.

[174] P.-E. JABIN, *A review of the mean field limits for Vlasov equations*, Kinetic & Related Models, 7 (2014), p. 661.

[175] G. JANKOWIAK AND V. H. NGUYEN, *Fractional Sobolev and Hardy-Littlewood-Sobolev inequalities*, 2014.

[176] R. JORDAN, D. KINDERLEHRER, AND F. OTTO, *The variational formulation of the FokkerPlanck equation*, SIAM journal on mathematical analysis, 29 (1998), pp. 1–17.

[177] J. JOST, *Nonlinear Dirichlet forms*, in New directions in Dirichlet forms, vol. 8 of AMS/IP Stud. Adv. Math., Amer. Math. Soc., Providence, RI, 1998, pp. 1–47.

[178] A. JÜNGEL, *Entropy methods for diffusive partial differential equations*, SpringerBriefs in Mathematics, Springer, [Cham], 2016.

[179] R. KANNAN, L. LOVÁSZ, AND M. SIMONOVITS, *Isoperimetric problems for convex bodies and a localization lemma*, Discrete Comput. Geom., 13 (1995), pp. 541–559.

[180] O. KAVIAN, S. MISCHLER, AND M. NDAO, *The Fokker-Planck equation with subcritical confinement force*, J. Math. Pures Appl. (9), 151 (2021), pp. 171–211.

[181] M. KELL, *q-heat flow and the gradient flow of the Renyi entropy in the p-Wasserstein space*, J. Funct. Anal., 271 (2016), pp. 2045–2089.

[182] D. KIM, *Instability results for the logarithmic Sobolev inequality and its application to related inequalities*, Discrete Contin. Dyn. Syst., 42 (2022), pp. 4297–4320.

[183] A. KOLMOGOROFF, *Zufallige Bewegungen (zur Theorie der Brownschen Bewegung)*, Annals of Mathematics, (1934), pp. 116–117.

[184] T. KÖNIG, *On the sharp constant in the Bianchi-Egnell stability inequality*, ArXiv preprint 2210.08482, (2022).

[185] ——, *Stability for the sobolev inequality: existence of a minimizer*, preprint arXiv:2211.14185, (2022).

[186] J. LASALLE, *Uniqueness theorems and successive approximations*, Annals of Mathematics, (1949), pp. 722–730.

[187] R. LATAŁA AND K. OLESZKIEWICZ, *Between Sobolev and Poincaré*, in Geometric aspects of functional analysis, vol. 1745 of Lecture Notes in Math., Springer, Berlin, 2000, pp. 147–168.

[188] M. LEDOUX, *A simple analytic proof of an inequality by P. Buser*, Proc. Amer. Math. Soc., 121 (1994), pp. 951–959.

[189] M. LEDOUX, *Spectral gap, logarithmic Sobolev constant, and geometric bounds*, in Surveys in differential geometry. Vol. IX, vol. 9 of Surv. Differ. Geom., Int. Press, Somerville, MA, 2004, pp. 219–240.

[190] M. LEDOUX, I. NOURDIN, AND G. PECCATI, *Stein's method, logarithmic Sobolev and transport inequalities*, Geom. Funct. Anal., 25 (2015), pp. 256–306.

[191] ——, *A Stein deficit for the logarithmic Sobolev inequality*, Sci. China Math., 60 (2017), pp. 1163–1180.

[192] K.-A. LEE AND J. L. VÁZQUEZ, *Geometrical properties of solutions of the porous medium equation for large times*, Indiana University Mathematics Journal, (2003), pp. 991–1016.

[193] B. LEIMKUHLER, M. SACHS, AND G. STOLTZ, *Hypocoercivity properties of adaptive Langevin dynamics*, SIAM J. Appl. Math., 80 (2020), pp. 1197–1222.

[194] A. LICHNEROWICZ, *Géométrie des groupes de transformations*, Travaux et Recherches Mathématiques, III. Dunod, Paris, 1958.

[195] J. R. LICOIS AND L. VÉRON, *A class of nonlinear conservative elliptic equations in cylinders*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 26 (1998), pp. 249–283.

[196] E. H. LIEB, *Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities*, Ann. of Math. (2), 118 (1983), pp. 349–374.

[197] M. LIERO AND A. MIELKE, *Gradient structures and geodesic convexity for reaction-diffusion systems*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 371 (2013), pp. 20120346, 1–28.

[198] T. M. LIGGETT, L^2 *rates of convergence for attractive reversible nearest particle systems: the critical case*, The Annals of Probability, (1991), pp. 935–959.

[199] P.-L. LIONS, *The concentration-compactness principle in the Calculus of Variations. the locally compact case, part 1.*, in Annales de l'Institut Henri Poincaré C, Analyse non linéaire, vol. 1, Elsevier, 1984, pp. 109–145.

[200] S. LIVINGSTONE, M. F. FAULKNER, AND G. O. ROBERTS, *Kinetic energy choice in Hamiltonian/hybrid Monte Carlo*, Biometrika, 106 (2019), pp. 303–319.

[201] J. LOTT AND C. VILLANI, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2), 169 (2009), pp. 903–991.

[202] G. LUISE AND G. SAVARÉ, *Contraction and regularizing properties of heat flows in metric measure spaces*, Discrete Contin. Dyn. Syst. Ser. S, 14 (2021), pp. 273–297.

[203] Z.-M. MA AND M. RÖCKNER, *Introduction to the theory of (non-symmetric) Dirichlet forms*, Universitext, Springer, Berlin, Heidelberg, 2012.

[204] J. C. MATTINGLY, A. M. STUART, AND D. J. HIGHAM, *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise*, Stochastic processes and their applications, 101 (2002), pp. 185–232.

[205] J. C. MAXWELL, *IV. on the dynamical theory of gases*, Philosophical transactions of the Royal Society of London, (1867), pp. 49–88.

[206] V. G. MAZ'YA, *The negative spectrum of the n-dimensional Schrödinger operator*, in Doklady Akademii Nauk, vol. 144, Russian Academy of Sciences, 1962, pp. 721–722.

[207] ———, *On the solvability of the Neumann problem*, in Doklady Akademii Nauk, vol. 147, Russian Academy of Sciences, 1962, pp. 294–296.

[208] H. P. MCKEAN, *Geometry of differential space*, Ann. Probability, 1 (1973), pp. 197–206.

[209] F. MEHLER, *Ueber die entwicklung einer function von beliebig vielen variablen nach laplaceschen functionen höherer ordnung.*, Journal für die reine und angewandte Mathematik, 65 (1866), pp. 161–176.

[210] M. MEINERT, *Partial differential equations on fractals. Existence, uniqueness and approximation results*, PhD thesis, Universität Bielefeld, 2020.

[211] A. MENEGAKI, *Quantitative rates of convergence to non-equilibrium steady state for a weakly anharmonic chain of oscillators*, Journal of Statistical Physics, 181 (2020), pp. 53–94.

[212] S. MISCHLER AND C. MOUHOT, *Exponential stability of slowly decaying solutions to the kinetic-Fokker-Planck equation*, Archive for rational mechanics and analysis, 221 (2016), pp. 677–723.

[213] M. MORIMOTO, *Analytic functionals on the sphere*, vol. 178 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1998.

[214] E. MOSSEL AND J. NEEMAN, *Robust dimension free isoperimetry in gaussian space*, The Annals of Probability, 43 (2015), pp. 971–991.

[215] C. MOUHOT AND L. NEUMANN, *Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus*, Nonlinearity, 19 (2006), pp. 969–998.

[216] C. E. MUELLER AND F. B. WEISSLER, *Hypercontractivity for the heat semigroup for ultraspherical polynomials and on the n-sphere*, J. Funct. Anal., 48 (1982), pp. 252–283.

[217] D. MUGNOLO, *Semigroup methods for evolution equations on networks*, Understanding Complex Systems, Springer, Cham, 2014.

[218] C. MÜLLER, *Spherical harmonics*, vol. 17 of Lecture Notes in Mathematics, Springer-Verlag, Berlin-New York, 1966.

[219] J. NASH, *Continuity of solutions of parabolic and elliptic equations*, American Journal of Mathematics, 80 (1958), pp. 931–954.

[220] E. NELSON, *The free Markoff field*, J. Funct. Anal., 12 (1973), pp. 211–227.

[221] F. NIER AND B. HELFFER, *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, Springer Science & Business Media, 2005.

[222] L. NIRENBERG, *On elliptic partial differential equations*, Ann. Scuola Norm. Sup. Pisa (3), 13 (1959), pp. 115–162.

[223] S.-I. OHTA, *Nonlinear geometric analysis on Finsler manifolds*, European Journal of Mathematics, 3 (2017), pp. 916–952.

[224] S.-I. OHTA AND K.-T. STURM, *Heat flow on Finsler manifolds*, Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, 62 (2009), pp. 1386–1433.

[225] F. OTTO, *The geometry of dissipative evolution equations: The porous medium equation*, Communications in Partial Differential Equations, 26 (2001), pp. 101–174.

[226] F. OTTO AND C. VILLANI, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173 (2000), pp. 361–400.

[227] M. OTTOBRE AND G. A. PAVLIOTIS, *Asymptotic analysis for the generalized Langevin equation*, Nonlinearity, 24 (2011), pp. 1629–1653.

[228] L. PARESCHI AND G. TOSCANI, *Interacting multiagent systems: kinetic equations and Monte Carlo methods*, OUP Oxford, 2013.

[229] G. A. PAVLIOTIS, G. STOLTZ, AND U. VAES, *Scaling limits for the generalized Langevin equation*, Journal of Nonlinear Science, 31 (2021), p. 8.

[230] A. PERSSON, *Bounds for the discrete part of the spectrum of a semi-bounded Schrödinger operator*, Mathematica Scandinavica, 8 (1960), pp. 143–153.

[231] B. PERTHAME ET AL., *Mathematical tools for kinetic equations*, Bulletin of the American Mathematical Society, 41 (2004), pp. 205–244.

[232] M. RACHID, *Hypoelliptic and spectral estimates for the linearized Landau operator*, arXiv preprint, (2020).

[233] M. REED AND B. SIMON, *Methods of modern mathematical physics, vol. IV, Acad*, Press, New York, 617 (1978).

[234] O. S. ROTHAUS, *Logarithmic Sobolev inequalities and the spectrum of Schrödinger operators*, J. Funct. Anal., 42 (1981), pp. 110–120.

[235] G. ROYER, *An initiation to logarithmic Sobolev inequalities*, vol. 14 of SMF/AMS Texts and Monographs, American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2007. Translated from the 1999 French original by Donald Babbitt.

[236] A. SAUMARD AND J. A. WELLNER, *Log-concavity and strong log-concavity: a review*, Statistics surveys, 8 (2014), p. 45.

[237] M. SCHMIDT, *Energy forms*, PhD thesis, Friedrich-Schiller-Universität Jena, 2016.

[238] N. SHANMUGALINGAM, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Rev. Mat. Iberoamericana, 16 (2000), pp. 243–279.

[239] C. E. SHANNON, *A mathematical theory of communication*, The Bell system technical journal, 27 (1948), pp. 379–423.

[240] S. L. SOBOLEFF, *Sur un théorème d'analyse fonctionnelle*, Rec. Math. Moscou, n. Ser., 4 (1938), pp. 471–497.

[241] A. J. STAM, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Information and Control, 2 (1959), pp. 101–112.

[242] E. M. STEIN AND G. WEISS, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series, No. 32, Princeton University Press, Princeton, N.J., 1971.

[243] G. STOLTZ AND Z. TRSTANOVA, *Langevin dynamics with general kinetic energies*, Multiscale Modeling & Simulation, 16 (2018), pp. 777–806.

[244] R. M. STRAIN AND Y. GUO, *Stability of the relativistic Maxwellian in a collisional plasma*, Communications in mathematical physics, 251 (2004), pp. 263–320.

[245] K.-T. STURM, *On the geometry of metric measure spaces. I*, Acta Math., 196 (2006), pp. 65–131.

[246] ——, *On the geometry of metric measure spaces. II*, Acta Math., 196 (2006), pp. 133–177.

[247] K. TAIRA, *A strong maximum principle for globally hypoelliptic operators*, Rendiconti del Circolo Matematico di Palermo Series 2, 68 (2019), pp. 193–217.

[248] G. TALENTI, *Best constant in Sobolev inequality*, Annali di Matematica pura ed Applicata, 110 (1976), pp. 353–372.

[249] G. TOSCANI, *A strenghtened entropy power inequality for log-concave densities*, IEEE Transactions on Information Theory, 61 (2015), pp. 6550–6559.

[250] M.-B. TRAN, *Convergence to equilibrium of some kinetic models*, Journal of Differential Equations, 255 (2013), pp. 405–440.

[251] J. L. VÁZQUEZ, *Smoothing and decay estimates for nonlinear diffusion equations*, vol. 33 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2006. Equations of porous medium type.

[252] A. M. VERSHIK, *Does there exist a Lebesgue measure in the infinite-dimensional space?*, Proceedings of the Steklov Institute of Mathematics, 259 (2007), pp. 248–272.

[253] C. VILLANI, *A short proof of the "concavity of entropy power"*, IEEE Trans. Inf. Theory, 46 (2000), pp. 1695–1696.

[254] C. VILLANI, *A review of mathematical topics in collisional kinetic theory*, Handbook of mathematical fluid dynamics, 1 (2002), pp. 3–8.

[255] ———, *Hypocoercive diffusion operators*, in International Congress of Mathematicians, vol. 3, 2006, pp. 473–498.

[256] ———, *Hypocoercivity*, Memoirs of the American Mathematical Society, 202 (2009).

[257] F.-Y. WANG, *A generalization of Poincaré and log-Sobolev inequalities*, Potential Anal., 22 (2005), pp. 1–15.

[258] ———, *Functional inequalities, Markov semigroups and spectral theory*, Mathematics monograph series, Elsevier, Burlington, MA, 2006.

[259] ———, *From super Poincaré to weighted log-Sobolev and entropy-cost inequalities*, J. Math. Pures Appl. (9), 90 (2008), pp. 270–285.

[260] F. B. . WEISSLER, *Logarithmic Sobolev inequalities and hypercontractive estimates on the circle*, J. Funct. Anal., 37 (1980), pp. 218–234.

[261] F. B. WEISSLER, *Logarithmic Sobolev inequalities for the heat-diffusion semigroup*, Trans. Amer. Math. Soc., 237 (1978), pp. 255–269.

Publications and preprints

[P1] G. Brigati, *Time averages for kinetic Fokker–Planck equations*, Kinetic and Related Models, https:///doi.org/10.3934/krm.2022037, (2023).

[P2] G. Brigati, J. Dolbeault, and N. Simonov, *Logarithmic Sobolev and interpolation inequalities on the sphere: constructive stability results*, Preprint arXiv: 2211.13180 and hal-03868496, (2022).

[P3] ——, *On Gaussian interpolation inequalities*, Preprint arXiv: 2302.03926 and hal-03978122, to appear in C.R. Mathématique, (2023).

[P4] ——, *Stability for the logarithmic Sobolev inequality*, Preprint arXiv 2303.12926 and hal-04042046, (2023).

[P5] G. Brigati and I. Hartarsky, *The normal contraction property for non-bilinear Dirichlet forms*, Potential Analysis, https://doi.org/10.1007/s11118-022-10057-2, (2022).

[P6] G. Brigati and G. Stoltz, *How to compute decay rates for kinetic Fokker–Planck equations?*, Preprint arXiv: 2302.145506, (2023).

RÉSUMÉ

Cette thèse est consacrée aux équations cinétiques de Fokker-Planck, à la stabilité des inégalités fonctionnelles et aux formes de Dirichlet non linéaires. Des taux de convergence vers l'équilibre sont estimés via un cadre analytique fonctionnel basé sur les normes faibles des solutions. L'équation de Vlasov-Fokker-Planck, avec variable de position confinée dans un tore, est analysée comme modèle de référence. La même stratégie est ensuite etendue à une large classe de modèles cinétiques. Nous considérons également des inégalités de Gagliardo--Nirenberg sur la sphère qui interpolent entre les inégalités de Poincaré et les inégalités de Sobolev. Nous prouvons des résultats constructifs de stabilité, dans la norme la plus forte possible, avec des exposants optimaux. L'estimation de stabilité dégénère sur un sous-espace de dimension finie, ce qui nécessite des précautions supplémentaires. Notre technique combine des développements de Taylor, l'analyse harmonique et des methodes parabòliques. Nous prouvons rigoureusement la convergence de la famille de Gagliardo-Nirenberg sur la sphère vers les inégalités de Beckner gaussiennes dans la limite de grandes dimensions. Ensuite, nous donnons des résultats constructifs de stabilité, en utilisant des diffusions non linéaires sur l'espace gaussien. Enfin, nous traitons l'inégalité de Sobolev logarithmique gaussienne comme cas limite. Nous trouvons des estimations explicites de stabilité pour des densités log-concaves ou à support compact par un argument de log-concavité deduit du flot d'Ornstein--Uhlenbeck et de la méthode du *carré du champ*. Nous contribuons à la théorie des formes de Dirichlet non-linéaires en etendant la propriété de contraction normale. La preuve adopte une nouvelle stratégie, basée sur l'approximation des fonctions Lipschitz réelles par des compositions répétées de fonctions linéaires par morceaux simples.

MOTS CLÉS

Equations cinétiques de Fokker-Planck, inégalités de Gagliardo-Nirenberg, inégalité de Sobolev, inégalité logarithmique de Sobolev, dynamique de Langevin, stabilité, flot de chaleur, formes de Dirichlet, hypocoercivité, fonctions log-concaves.

ABSTRACT

This thesis concerns kinetic Fokker-Planck equations, stability of functionals inequalities, and nonlinear Dirichlet forms. Constructive convergence rates to equilibrium for kinetic equations are computed via a functional analytic framework based on weak norms of solutions. The Vlasov-Fokker-Planck equation, with the space variable confined in a torus, is analysed as a benchmark. Then, the same strategy is generalised to a wide class of kinetic Fokker-Planck models. We also consider Gagliardo-Nirenberg inequalities on the sphere, interpolating between the Poincaré and the Sobolev inequalities. We prove constructive stability results, in the strongest possible norm, with sharp exponents in the distance from optimisers. This term degenerates on a finite-dimensional subspace requiring additional care. Our technique combines Taylor expansions, harmonic analysis, and (non)linear diffusion flows. We rigorously prove convergence of the Gagliardo-Nirenberg family of inequalities, for the dimension of the sphere approaching infinity, to the Gaussian Beckner inequalities. Then, we give constructive stability results for those, using nonlinear diffusion flows on the Gaussian space. Finally, we treat the Gaussian logarithmic Sobolev inequality as a limit case. We find explicit stability estimates for log-concave or compactly-supported densities, thanks to the interplay between log-concavity and the Ornstein-Uhlenbeck flow, using the *carré du champ* method. We contribute to the theory of nonlinear Dirichlet forms. We extend the normal contraction property to the nonlinear setting. The proof adopts a new strategy, based on the approximation of real Lipschitz functions with repeated compositions of elementary piecewise linear functions. 55

KEYWORDS

Kinetic Fokker-Planck equations, Gagliardo-Nirenberg inequalities, Sobolev inequality, logarithmic Sobolev inequality, Langevin dynamics, stability, heat flow, Dirichlet forms, hypocoercivity, log-concave functions.