Functional and isoperimetric inequalities for probability measures on H-type groups

by

Vasilis Kontis

A thesis presented for the degree of

Doctor of Philosophy of Imperial College London

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2BZ

2011

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Vasilis Kontis

Abstract

We investigate isoperimetric and functional inequalities for probability measures in the sub-elliptic setting and more specifically, on groups of Heisenberg type. The approach we take is based on U-bounds as well as a Laplacian comparison theorem for H-type groups. We derive different forms of functional inequalities (of Φ-entropy and F-Sobolev type) and show that they can be equivalently stated as isoperimetric inequalities at the level of sets. Furthermore, we study transportation of measure via Talagrand-type inequalities. The methods used allow us to obtain gradient bounds for the heat semigroup. Finally, we examine some properties of more general operators given in Hörmander's sum of squares form and show that the associated semigroup converges to a probability measure as $t \to \infty$.

Acknowledgements

I would like to warmly thank my advisor, Professor Bogusław Zegarliński, for his support and advice through these years. It has been a privilege to study close to him and I have benefited enormously from his insight, without which this thesis would not have been possible.

I consider myself very lucky to have had the opportunity to work with Drs. James Inglis and Federica Dragoni and I would like to thank them for their valuable ideas, many of which are included in this thesis.

Furthermore, I am indebted to Professors Sergey Bobkov, Cyril Roberto, Simon Peszat and Robert Olkiewicz for their hospitality as well as many motivating mathematical discussions.

It is a pleasure to have Professors Franck Barthe, Gordon Blower and Ari Laptev as examiners and I would like to thank them for accepting this task.

Moreover, all the members of the analysis group at Imperial deserve special thanks and in particular Drs. Dan Crisan and Thomas Ø. Sørensen who very patiently read my transfer report and made useful remarks.

Last but not least, I am grateful to my parents for always being there and striving to provide me with the best education possible.

Table of contents

Chapter 1

Introduction

One of the most striking proofs of the isoperimetric inequality in $\mathbb{R}^n$ is given via the well-known Gagliardo-Nirenberg-Sobolev inequality, which states that, for all sufficiently smooth functions $f : \mathbb{R}^n \to \mathbb{R}$,

$$\left(\int_{\mathbb{R}^n} |f(x)|^{\frac{n}{n-1}} dx\right)^{\frac{n-1}{n}} \leq \frac{1}{n\omega_n^{1/n}} \int_{\mathbb{R}^n} \sqrt{\sum_{i=1}^{n} (\partial_i f)^2(x)} dx, \tag{1.1}$$

where ω_n is the volume of the unit ball in $\mathbb{R}^n$. Passing to the limit as f approximates the characteristic function of a set $A \subset \mathbb{R}^n$ in (1.1) yields the isoperimetric inequality

$$vol(A) \leq \frac{1}{n\omega_n^{1/n}} S(\partial A), \tag{1.2}$$

where $vol(A)$ and $S(A)$ stand for the volume of A and the surface measure of its boundary, respectively. Equality is achieved when $A = B_r$ is the Euclidean ball of radius $r > 0$, hence the balls solve the *isoperimetric problem*, which consists of minimising the surface area amongst all sets of fixed volume.

In this thesis, we shall focus on the study of isoperimetry for probability measures on metric spaces with a particular emphasis on a class of sub-Riemannian spaces known as Heisenberg type (for short, H-type) groups. In a metric space, there is a notion of surface area which is given by the surface measure μ^+, known as the Minkowski content of a set, but one cannot hope for an inequality such as (1.1). This is due to the fact that the latter implies a doubling condition on the measure, i.e. that for each $r > 0$, there is a constant $C > 0$ such that $\mu(B_{2r}) \leq C\mu(B_r)$; a condition which fails for probability measures. Nevertheless, one may start from a functional inequality which is weaker than (1.1), such as a Poincaré or a logarithmic Sobolev-type inequality, and hope to deduce some information about the isoperimetric function (or isoperimetric profile) of μ defined by

$$\mathcal{I}_\mu(a) = \inf\{\mu^+(A) : A \text{ Borel with } \mu(A) = a\}, \tag{1.3}$$

for $a \in [0,1]$; it is the best function in the isoperimetric inequality

$$\mathcal{I}_\mu(a) \leq \mu^+(A).$$

The study of such inequalities for probability measures can be traced back to the works of Sudakov and Tsirelson [ST74] and Borell [Bor74] who independently studied the problem in Gauss space $(\mathbb{R}^n, |\cdot|, \gamma_n)$, where $\gamma_n(dx) = (2\pi)^{-n/2}e^{-|x|^2/2}dx =: \phi_n(x)dx$ is the standard Gaussian measure on $\mathbb{R}^n$. By showing that half spaces $\{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \leq \alpha\}$, $\alpha \in \mathbb{R}$, solve the isoperimetric problem, they deduced that the function $\mathcal{I}_{\gamma_n}$ is of the form $\phi_1 \circ \Phi_1^{-1}$, for all n, where $\Phi_1(a) = \int_{-\infty}^a \phi_1(x)dx$ is the one-dimensional Gaussian distribution function.

Identifying the extremal sets for a measure μ can prove to be a very difficult goal, which is why one usually tries to solve the easier problem of estimating the function $\mathcal{I}_\mu$ from below. For instance, such an estimate can be of the form

$$\mathcal{I}_\mu(a) \geq c\hat{a}(-\log \hat{a})^\beta, \tag{1.4}$$

for some $c > 0$ and $\beta \in [0,1]$, where $\hat{a} = \min(a, 1-a)$. The motivation for such an inequality is that the function $a \mapsto \hat{a}(-\log \hat{a})^\beta$ is equivalent to the isoperimetric function of the prototype measure $d\nu_p(x) = Z^{-1}e^{-|x|^p/p}dx$ on the real line, where $p \geq 1$ and $\beta = (p-1)/p$, in the sense that their ratio is bounded from above and below by universal constants. There is a vast number of works in the literature dealing with such estimates for probability measures on $\mathbb{R}^n$ (see e.g. [Bob99, Bar02, BH97b, BH97a, Hue09, Bob97a, Bob96a, Bob96b, Bob02, KLS95, Mil09b] and references therein).

In addition to the study of the isoperimetric and functional inequalities in Euclidean space, there has been a large amount of literature devoted to extending known results to Riemannian manifolds. Here, a very efficient strategy makes use of tools from semigroup theory. The Bakry-Emery semigroup approach to hypercontractivity and logarithmic Sobolev inequalities [BÉ85] inspired many proofs of functional and isoperimetric inequalities (see e.g. [BCR06, BCR07, BL96, Fou00, BL06, Led09, Mil09a, Mil09b] as well as [ABC+00, Bak94, Led00, GZ03] for a more comprehensive account of this theory). A central ingredient in such proofs consists of the so called gradient bounds for the semigroup $P_t = e^{tL}$, which are inequalities of the form

$$\Gamma(P_t f, P_t f)^{\frac{q}{2}} \leq C(t) P_t(\Gamma(f,f)^{\frac{q}{2}}), \tag{1.5}$$

for some $q \geq 1$ and a constant C dependent on t only and decaying to 0 at infinity, where Γ is the *carré du champ* operator, defined as

$$\Gamma(f,g) = \frac{1}{2}(L(fg) - fLg - gLf).$$

The strongest inequality occurs when $q = 1$ and in this case one may derive isoperimetric information for the measure μ as well as the logarithmic Sobolev inequality (although $q = 2$ still suffices for many purposes, such as the Poincaré inequality).

As explained in [BÉ85], the inequality (1.5) can be achieved via a curvature-dimension criterion of the form

$$\Gamma_2(f,f) \geq \lambda \Gamma(f,f), \tag{1.6}$$

where $\Gamma_2(f,f) = \frac{1}{2} L\Gamma(f,f) - \Gamma(f, Lf)$. For example, when $Lf(x) = \Delta f(x) - x \cdot \nabla f(x)$ is the Ornstein-Uhlenbeck operator on $\mathbb{R}^n$, which has as invariant measure the standard Gaussian measure γ_n, we have

$$\Gamma_2(f,f) = |\nabla f|^2 + \|\mathrm{Hess} f\|_2^2 = \Gamma(f,f) + \|\mathrm{Hess} f\|_2^2,$$

where

$$\|\mathrm{Hess} f\|_2^2 = \sum_{i,j=1}^{n} (\partial_i \partial_j f)^2$$

is the Hilbert-Schmidt norm of the matrix of second derivatives of f. Therefore, in this case (1.6) is satisfied with $\lambda = 1$. An interesting situation occurs when an inequality like (1.6) is not available. In the Riemannian setting, the curvature-dimension condition amounts via Bochner's formula to a lower bound for the Ricci tensor (see e.g. [Led00]), which is not available in sub-Riemannian geometries, to which H-type groups belong. In the simplest such space, the Heisenberg group, gradient bounds were recently studied for the heat semigroup [Li06, BBBC08, DM05, HZ10] and then extended to the more general class of H-type groups [Eld10, Eld09]. It is not known how to prove (1.5) for more general semigroups, but an alternative approach to proving logarithmic Sobolev inequalities, which does not rely on the curvature lower bound, was given in [HZ10] and then extended in various directions in [Pap10, IP09, IKZ11]. A central assumption in this strategy is that there exists a finite constant K_d such that

$$\Delta d \leq K_d, \quad \text{for all } x \text{ such that } d(x) := d(x, 0) \geq 1, \tag{1.7}$$

where the notions of d and Δ are understood appropriately depending on the space we are in (e.g. in Euclidean space Δ is the Laplacian and $d(x) = |x|$ the Euclidean distance, while on H-type groups, d is the Carnot-Carathéodory distance and Δ is the sub-Laplacian). Under this assumption, the authors proved certain estimates, which they called U-bounds, of the form

$$\int |f|^q \left(|U| + |\nabla U|^q\right) d\mu \leq C \int |\nabla f|^q d\mu + D \int |f|^q d\mu, \tag{1.8}$$

which then lead to other functional inequalities, including the Poincaré inequality and under additional assumptions on U, the logarithmic Sobolev inequality. The Laplacian comparison principle (1.7) in a sense replaces the assumption on the existence of a lower bound for the Ricci curvature and allows one to recover at the same time the Euclidean, Riemannian and sub-Riemannian cases (it should be noted, however, that the constants are in general dependent on the dimension; one may think, for example, of the Euclidean case, where for $n \geq 2$, $\Delta|x| = (n-1)/|x|$). A natural question therefore arises of what is the weakest condition on the geometry of the space for the results described above to hold.

In the study of functional inequalities and isoperimetry in Euclidean and Riemannian spaces, a large number of papers is devoted to measures which are *log-concave*, i.e. which satisfy

$$\mu(tA + (1-t)B) \geq \mu(A)^t \mu(B)^{1-t}, \tag{1.9}$$

for all measurable sets A and B and all $t \in [0,1]$. By Borell's characterisation [Bor74], this is equivalent to μ being concentrated on some convex set $\Omega \subset \mathbb{R}^n$ where it is absolutely continuous with respect to the Lebesgue measure and where its density $p(x)$ satisfies

$$p(tx + (1-t)y) \geq p(x)^t p(y)^{1-t},$$

for all $t \in [0,1]$ (in other words, $\log p$ is a concave function). In the sub-Riemannian setting, the concept of convexity is not as clearly defined. A notion of weak convexity was introduced in [DGN03] (see also [JLMS07] and [BD] for other approaches to convexity), where it was shown that the Folland-Kaplan gauge N is weakly convex and so are the balls $\{N \leq r\}$. On the other hand, the Carnot-Carathéodory distance d is not convex, despite being the intrinsic metric. This could suggest studying the measures

$$\mu(dx) = Z^{-1} \mathrm{e}^{-N^p(x)} dx, \tag{1.10}$$

as possible analogues of the prototype log-concave measures ν_p defined above in the Euclidean setting. Nevertheless, a remarkable result of [HZ10] states that the measure given by (1.10) does not satisfy a logarithmic Sobolev inequality for *any* $p \geq 1$. On the other hand, as we will see, the measures defined by

$$\mu(dx) = Z^{-1} \mathrm{e}^{-d^p(0,x)/p} dx, \tag{1.11}$$

for $p \geq 1$, share many common isoperimetric as well as functional analytic properties with their Euclidean counterparts, defined by replacing d with the Euclidean norm.

The organisation of the thesis is as follows. In Chapter 2 we describe the basic background which is necessary for Chapters 3-5. This includes the basic notions pertaining to homogeneous Lie groups and H-type groups, as well as the definitions of the inequalities that we study. In Chapter 3 we derive some functional and isoperimetric inequalities for measures of the form (1.11), starting from U-bounds of the type

$$\int |f(x)| d(0,x)^{p-1} \mu(dx) \leq C \int |\nabla f|(x) \mu(dx) + D \int |f| \mu(dx). \tag{1.12}$$

Such an inequality is stable under perturbations, which allows us to extend the results to a large class of measures. We start by showing that (1.12) implies the L^1 Poincaré (or Cheeger) inequality, which already implies an isoperimetric inequality of the form

$$\min(\mu(A), 1 - \mu(A)) \leq C \mu^+(A).$$

This hints the connection between the measure μ and the measure ν_p introduced above, which is known to have a similar property (see e.g. [Bob99]). We then move on to derive a Ledoux inequality starting from the U-bound estimate, which reads

$$\int |f| \log_+^{1/q} \left(\frac{|f|}{\int |f| d\mu} \right) d\mu \leq C \int |\nabla f| d\mu + D \int |f| d\mu.$$

The motivation for such an inequality comes from [Led88], where it was proved for the Gaussian measure on $\mathbb{R}^n$ (with $q = 2$), as well as from subsequent results of [BH97b, Mil09a, Mil09b, RZ07, BG99, BH97a]. As we will see, this inequality combined with the Cheeger inequality imply the following L^1 Φ-Entropy inequality with $\Phi(t) = t \log^{1/q}(1+t)$,

$$\int |f| \log^{1/q}(1 + |f|) d\mu - \int |f| d\mu \log^{1/q} \left(1 + \int |f| d\mu \right) \leq C \int |\nabla f| d\mu,$$

or, equivalently, at the level of sets

$$\min(\mu(A), 1-\mu(A)) \log^{1/q}\left(\frac{1}{\min(\mu(A), 1-\mu(A))}\right) \leq C\mu^+(A).$$

Such Φ-Entropy inequalities (with the L^2 norm of the gradient on the right) were already used the literature as a tool to study isoperimetry, in particular in relation to the behaviour of the isoperimetric profile under tensorisation [BCR07, BCR06]. In the final section of this chapter, we turn to the problem of transportation of measure, which involves looking at inequalities of the form

$$\mathcal{W}_p^p(\mu, \nu) \leq C \int \log\left(\frac{d\nu}{d\mu}\right) d\nu,$$

where ν is a probability measure, absolutely continuous with respect to μ. The quantity appearing on the right here is the Kullback-Leibler information, or relative entropy, defined by

$$H(\nu\|\mu) = \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu, & \text{if } \nu \text{ is absolutely continuous with respect to } \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

The study of such inequalities in relation to functional inequalities as well as isoperimetry and concentration of measure is a topic that attracted a lot of attention in recent years; our work was motivated by various papers, including [Blo03, BB06, Tal96, BGL01, BG99, LV07, BEHM09, CGW10].

In Chapter 4, we engage in the study of gradient bounds for the heat semigroup. Such estimates in the sub-elliptic setting motivated a considerable amount of research in recent years. They were studied in the Heisenberg group [DM05, Li06, BBBC08, Bon09, LP10, Mel04, HZ10, IKZ11], on H-type groups [Eld10], as well as in some other spaces, such as $\mathbf{SU}(2)$ and $\mathbf{SL}(2,\mathbb{R})$ [BB09, Bon09]. Although the results we give in this section are already known, we present a different approach which is based on estimates similar to (1.12) and the Gaussian estimates for the heat kernel of [Eld09].

Finally, Chapter 5 is devoted to proving some gradient bounds and Li-Yau estimates for Hörmander operators. These are second-order differential operators, given in terms of fields satisfying Hörmander's condition [Hör67]. The gradient bounds presented here were used in [DKZ10] as a tool to prove the ergodicity of the semigroup in infinite dimensions. Li-Yau estimates [LY86] are inequalities of the form

$$\frac{\Gamma(P_t f)}{(P_t f)^2} \le C_1 \frac{L P_t f}{P_t f} + \frac{C_2}{t} + C_3,$$

where C_1, C_2, C_3 are nonnegative constants. Such estimates provide useful heat kernel bounds and were used as a tool to reach functional inequalities [BL06] as well as isoperimetric statements [Led94] in the elliptic setting. In sub-Riemannian spaces, they were considered in [BBBQ09, BG11, Bon09] for the heat semigroup.

Some of the results contained in this thesis can also be found in [IKZ11] and [DKZ10].

Chapter 2

Basic definitions

2.1 Homogeneous Lie Groups

Let us start by giving some basic definitions concerning homogeneous Lie groups. More details, as well as proofs to the results that we state below, can be found in [BLU07, FS82].

2.1.1 Lie Groups and Lie algebras on $\mathbb{R}^n$

Let X be a smooth vector field on $\mathbb{R}^n$. Then, for all differentiable functions f, X has the representation

$$(Xf)(x) = \sum_{i=1}^{n} \alpha_i(x)(\partial_i f)(x),$$

where $\alpha_i : \mathbb{R}^n \to \mathbb{R}$ are smooth scalar functions, called the *components* of X, and ∂_i denotes differentiation with respect to the i^{th} coordinate. Following [BLU07], we will denote by

$$(XI)(x) = \begin{pmatrix} \alpha_1(x) \\ \alpha_2(x) \\ \vdots \\ \alpha_n(x) \end{pmatrix}$$

the n-tuple of components of X. We recall that a Lie group over $\mathbb{R}^n$ is a group $\mathbb{G} = (\mathbb{R}^n, \cdot)$ such that for $x, y \in \mathbb{R}^n$, the maps

$$(x, y) \mapsto x \cdot y$$

and

$$x \mapsto x^{-1}$$

are smooth. To a Lie group is associated its Lie algebra, which is a vector space of vector fields and can be thought of as the tangent space at the identity of the group, which can be assumed to be the origin, without loss of generality [1] . More specifically, we have the following characterisation of the Lie algebra.

Definition 2.1. *Let $\mathbb{G} = (\mathbb{R}^n, \cdot)$ be a Lie group and let $\mathfrak{g}$ be its Lie algebra. A vector field X belongs to $\mathfrak{g}$ if and only if there exists a vector $\xi \in \mathbb{R}^n$ such that for every smooth $f : \mathbb{R}^n \to \mathbb{R}$,*

$$(Xf)(x) = \frac{d}{ds}\bigg|_{s=0} f(x \cdot s\xi) = \lim_{s \to 0} \frac{f(x \cdot s\xi) - f(x)}{s}, \tag{2.1}$$

where $s\xi = (s\xi_1, \ldots s\xi_n)$. In this case, $\xi = XI(0)$.

Remark 2.2. *The fact that $\xi = XI(0)$ is a direct consequence of (2.1). Namely, by choosing $f(x) = \pi_i(x) = x_i$ we see that*

$$(X\pi_i)(0) = \xi_i,$$

while the definition of the vector $(XI)(0)$ directly implies that $(X\pi_i)(0) = ((XI)(0))_i$.

It follows by the associativity of the group law and definition (2.1) that the vector fields satisfying the condition (2.1) are left-invariant. In other words, if $y \in \mathbb{R}^n$ and $L_y : \mathbb{R}^n \to \mathbb{R}^n$ is the map of left-translations given by

$$L_y x = y \cdot x,$$

[1] Indeed, if $0 \neq e \in \mathbb{R}^n$ is the identity, we may consider new coordinates on $\mathbb{R}^n$ given by the smooth diffeomorphism $x \mapsto x - e$.

then we have

$$(Xf) \circ L_y = X(f \circ L_y),$$

for all smooth f and all $y \in \mathbb{R}^n$, where $\circ$ denotes composition of functions. To see this, we note that by (2.1) we have, for all $x \in \mathbb{R}^n$ and all smooth f,

$$(Xf)(L_y(x)) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f(L_y(x) \cdot s\xi) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f(L_y(x \cdot (s\xi))) = X(f \circ L_y)(x).$$

The Lie algebra is equipped with a Lie bracket $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, defined by

$$[X, Y] := XY - YX.$$

It may not be immediately obvious from the definition that $[X, Y] \in \mathfrak{g}$, since the bracket of X and Y is defined as the difference of two second order operators. However a direct calculation shows that $[X, Y]$ is indeed of the form $\sum_i \alpha_i \partial_i$, with

$$\alpha_i = ([X, Y]I)_i = X(YI)_i - Y(XI)_i.$$

The Lie bracket is a bilinear operation which satisfies $[X, X] = 0$ for all $X \in \mathfrak{g}$, as well as the *Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0,$$

for all $X, Y, Z \in \mathfrak{g}$.

The Lie algebra $\mathfrak{g}$ has a natural basis, known as the *Jacobian basis*, which can be obtained via Definition 2.1 by choosing $\xi = e_i$, where $\{e_i\}_{i=1}^n$ denotes the standard basis of $\mathbb{R}^n$.

Proposition 2.3. *Let $Z_1, \ldots, Z_n$ be defined as the vector fields satisfying*

$$Z_i f(x) = \frac{d}{ds}\bigg|_{s=0} f(x \cdot se_i), \quad (ZI)(0) = e_i, \tag{2.2}$$

for all $x \in \mathbb{R}^n$ and all differentiable $f : \mathbb{R}^n \to \mathbb{R}$. Then $\{Z_1, \cdots, Z_n\}$ is a basis for the Lie algebra.

Given a set $\mathcal{V}$ of vector fields, we define the Lie algebra generated by $\mathcal{V}$ as the smallest Lie algebra containing $\mathcal{V}$ and denote it by $\mathrm{Lie}(\mathcal{V})$. In other words,

$$\mathrm{Lie}(\mathcal{V}) = \bigcap_{\mathfrak{h}:\mathcal{V}\subset\mathfrak{h}} \mathfrak{h}.$$

A very useful observation is that $\mathrm{Lie}(\mathcal{V})$ consists of nothing more than combinations of nested brackets of elements in $\mathcal{V}$.

Proposition 2.4. *Let $\mathcal{V}$ be a set of vector fields on $\mathbb{R}^n$. Then*

$$\mathrm{Lie}(\mathcal{V}) = \mathrm{span}\{[X_1, [X_2 \cdots [X_{n-1}, X_n]]] : n \geq 1, X_i \in \mathcal{V} \text{ for } i = 1, \dots n\},$$

where, for $n = 1$, the nested bracket on the right is understood as being a single vector field from $\mathcal{V}$.

2.1.2 Homogeneous stratified groups and H-type groups

Groups of Heisenberg type belong to the more general class of Carnot groups, which are defined below. More details on these groups can be found in [BLU07, FS82].

Definition 2.5. *Let $\mathbb{G} = (\mathbb{R}^n, \cdot)$ be a Lie group that admits a decomposition*

$$\mathbb{R}^n = \mathbb{R}^{r_1} \times \cdots \times \mathbb{R}^{r_k}$$

and is equipped with a one-parameter family of automorphisms $\{\delta_\lambda\}_{\lambda\geq 0}$, $\delta_\lambda : \mathbb{R}^n \to \mathbb{R}^n$, defined by

$$\delta_\lambda(x) = (\lambda x^{(1)}, \lambda^2 x^{(2)}, \dots, \lambda^k x^{(k)}),$$

where $x \in \mathbb{R}^n$ and $x^{(i)} \in \mathbb{R}^{r_i}$. Let $X_1, \dots X_{r_1}$ be the first r_1 vector fields in the Jacobian basis, i.e. the left-invariant vector fields given by (2.2), which are such that

$$(X_i f)(0) = (\partial_i f)(0),$$

for $i = 1, \dots, r_1$. Let $\mathcal{Z}(x) = \{ZI(x) : Z \in \mathrm{Lie}\{X_1, \dots X_{r_1}\}\} \subset \mathbb{R}^n$ and suppose, in *addition, that for every* $x \in \mathbb{R}^n$,

$$\dim \mathcal{Z}(x) = n,$$

where $\dim \mathcal{Z}$ *denotes the dimension of* $\mathcal{Z}$ *as a vector space over* $\mathbb{R}$.

Then we say that $(\mathbb{R}^n, \cdot, \delta_\lambda)$ *is a homogeneous stratified group (or a Carnot group) of step* k *and* r_1 *generators.*

The family $\{\delta_\lambda\}_{\lambda \geq 0}$ is known as the family of dilations of the group. A vector field X is said to be homogeneous of degree r with respect to dilations (or δ_λ-homogeneous of degree r), if for any smooth f and any $x \in \mathbb{R}^n$,

$$X(f(\delta_\lambda(x))) = \lambda^r (Xf)(\delta_\lambda(x)).$$

Therefore, the definition says simply that the left-invariant vector fields which are homogeneous of degree 1 together with their commutators must generate the whole Lie algebra at any point of the group.

Any Carnot group $\mathbb{G}$ is a connected and simply connected nilpotent Lie group, equipped with dilations. The Lie algebra of such a group admits the decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k,$$

where $\mathfrak{g}_i$ is the space spanned by the vector fields which are δ_λ-homogeneous of degree i. In other words, if $\{Z_i\}_{i=1}^n$ is the Jacobian basis of $\mathbb{G}$, $\mathfrak{g}_1 = \text{span}\{Z_i : i = 1, \ldots, r_1\}$, $\mathfrak{g}_2 = \text{span}\{Z_i : i = r_1 + 1, \ldots, r_2\}, \ldots, \mathfrak{g}_k = \{Z_i : i = r_{k-1}, \ldots, r_k\}$. The dilations on the group give rise to dilations at the level of the Lie algebra as follows. If $X \in \mathfrak{g}$ has the representation

$$X = \sum_{i=1}^{n} \alpha_i Z_i,$$

then

$$\delta_\lambda(X) = \sum_{i=1}^{r_1} \alpha_i \lambda Z_i + \sum_{i=r_1}^{r_2} \alpha_i \lambda^2 Z_i \; + \cdots + \sum_{i=r_{k-1}}^{r_k} \alpha_i \lambda^k Z_i.$$

The homogeneous dimension of the group is defined as the number

$$Q = \sum_{i=1}^{k} i r_i.$$

We note that this is always greater than the dimension of the underlying space, i.e. $Q \geq n$, while the case $Q = n$ corresponds to the Euclidean situation, i.e. when $\cdot = +$ and $\delta_\lambda(x) = \lambda x$. Moreover, if $\mathbb{G}$ is not the standard Euclidean group, then necessarily $Q \geq 4$. This follows immediately from the definition above, since in this case $r_1 \neq n$ and $r_1 \geq 2$, because there must be at least two linearly independent vector fields in $\mathfrak{g}_1$ whose commutator generates $\mathfrak{g}_2$. The number Q is the degree of homogeneity of the Lebesgue measure with respect to dilations. Indeed, one may compute

$$|\delta_\lambda(A)| = \lambda^Q |A|,$$

for all $A \subset \mathbb{R}^n$, where $|A|$ is the Lebesgue measure of A and $\delta_\lambda(A) = \{\delta_\lambda(a) : a \in A\}$.

We now give the definition of groups of Heisenberg type. A detailed presentation of these groups together with their basic properties can be found in Chapter 18 of [BLU07]. We recall that the centre of a Lie algebra $\mathfrak{g}$ is defined as the vector space $\mathfrak{z}$ such that $[\mathfrak{g}, \mathfrak{z}] = 0$, i.e. $[X, Z] = 0$ for all $X \in \mathfrak{g}$, $Z \in \mathfrak{z}$.

Definition 2.6 (H-type groups)**.** *Let $\mathfrak{g}$ be a Lie algebra whose centre is $\mathfrak{z}$ and let $\mathfrak{v} := \mathfrak{z}^\perp$. We say that $\mathfrak{g}$ is of Heisenberg-type (or simply H-type) if*

$$[\mathfrak{v}, \mathfrak{v}] = \mathfrak{z}$$

and there exists an inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ with $\langle \mathfrak{z}, \mathfrak{v} \rangle = 0$ such that for any $Z \in \mathfrak{z}$, the map $J_Z : \mathfrak{v} \mapsto \mathfrak{v}$ given by

$$\langle J_Z X, Y \rangle = \langle [X, Y], Z \rangle,$$

for $X, Y \in \mathfrak{v}$, is an orthogonal map whenever $\langle Z, Z \rangle = 1$. An H-type group is a connected and simply connected Lie group $\mathbb{G}$ whose Lie algebra is of H-type.

H-type groups were introduced by Kaplan [Kap80] (see also [Mét80]). Such a group is a Carnot group of step 2 whose Lie algebra has the decomposition

$$\mathfrak{g} = \mathfrak{v} \oplus \mathfrak{z}.$$

It can be shown that any H-type Lie algebra is isomorphic to $\mathbb{R}^m \oplus \mathbb{R}^r$ for some integers m, r equipped with a bracket operation $[\cdot, \cdot]$ such that $\mathbb{R}^r$ is the centre of the Lie algebra. Moreover, the corresponding H-type group is isomorphic to $(\mathbb{R}^{m+r}, \cdot)$ where $\cdot$ is given by

$$y \cdot w = y + w + \frac{1}{2}[y, w],$$

for $y, w \in \mathbb{R}^n$, $n = m + r$ (which is the Baker-Campbell-Hausdorff formula for step 2 Lie algebras $\mathfrak{g}$, i.e. those which satisfy $[\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] = 0$). At the level of the group, the bracket $[y, w]$ is defined by identifying a vector field W from the Lie algebra with an element $w \in \mathbb{R}^n$ of the group by

$$w = (w_1, \ldots, w_n) \longleftrightarrow W = \sum_{i=1}^{n} w_i Z_i,$$

where $\{Z_i\}_{i=1}^n$ is the Jacobian basis (in other words, the exponential map is the identity). In what follows, we will always assume that a group of H-type is of this form. By a result of [Kap80], given $m, r \in \mathbb{N}$, the existence of such a group is equivalent to the condition

$$r < \varrho(m),$$

where ϱ is the Hurwitz-Radon function defined as $\varrho(m) = 8\alpha + \beta$, where α and β are the natural numbers in the representation $m = m_o \cdot 2^{4\alpha+\beta}$, $\beta \leq 3$, where m_o is an odd number (so $\varrho(m) = 0$ is m itself is odd). In particular, we always have $r < m$.

For such groups, one can compute the Jacobian fields directly. Let us write the elements of the group as $w = (x, z) \in \mathbb{R}^n$ with $x \in \mathbb{R}^m$ and $z \in \mathbb{R}^r$. We have

$$\begin{aligned} Z_i f(w) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f(w \cdot se_i) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\left(w + se_i + \frac{s}{2}[w, e_i]\right). \end{aligned}$$

Since $[w, e_i] = [(x, z), e_i] = (0, [x, e_i])$ if $i \in \{1, \ldots, m\}$ and is zero otherwise, for $i = 1, \ldots, m$ we have

$$\begin{aligned} Z_i f(w) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\left(\left(x + se_i, z + \frac{s}{2}[x, e_i]\right)\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\left(\left(x + se_i, z + \frac{s}{2}[x, e_i]\right)\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\left(\left(x + se_i, z + \frac{s}{2}\sum_{j=m+1}^{n} \left\langle J_{e_j} x, e_i \right\rangle e_j\right)\right) \\ &= \left(\partial_i + \frac{1}{2}\sum_{j=m+1}^{n} \left\langle J_{e_j} x, e_i \right\rangle \partial_j\right) f(w), \end{aligned}$$

while if $i = m + 1, \ldots, n$,

$$Z_i f(w) = \partial_i f(w).$$

The map J_Z appearing the the definition has the following properties (see e.g. [Kap80]).

Proposition 2.7. *Let $Z, Z' \in \mathfrak{z}$ and $X, X' \in \mathfrak{v}$. Then*

$$\begin{aligned} \langle J_Z(X), X\rangle &= 0, \\ \langle J_Z(X), X'\rangle &= -\langle X, J_Z(X')\rangle, \\ |J_Z(X)| &= |Z||X|, \\ \langle J_Z(X), J_{Z'}(X)\rangle &= \langle Z, Z'\rangle |X|^2, \\ [X, J_Z(X)] &= |X|^2 Z. \end{aligned}$$

In particular, this shows that the transformations J_Z are linear and skew-symmetric, which in turn implies that m (the dimension of $\mathfrak{v}$) must be even, or the determinant of the orthogonal map J_Z, for $\langle Z, Z\rangle = 1$, would be 0.

The simplest example of a Carnot group is given by the (first) Heisenberg group $\mathbb{H}$. This is a group over $\mathbb{R}^3$ with group law

$$(x, y, z) \cdot (x', y', z') = \left(x + x', y + y', z + z' + \frac{1}{2}(xy' - x'y)\right).$$

Alternatively (see e.g. [CG90]), it can be seen as the group of upper triangular matrices

$$M(x, y, z) = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

equipped with the (noncommutative) product

$$M(x, y, z)M(x', y', z') = M(x + x', y + y', z + z' + xy').$$

The Heisenberg group is equipped with dilations $\delta_\lambda(x, y, z) = (\lambda x, \lambda y, \lambda^2 z)$, which induce the decomposition $\mathbb{H} = \mathbb{R}^2 \times \mathbb{R}$.

Let us compute the Jacobian fields for $\mathbb{H}$. Let $p = (x, y, z)$. We have,

$$\begin{aligned} Z_1 f(p) &= \lim_{\varepsilon \to 0} \frac{f\left((x, y, z) \cdot (\varepsilon, 0, 0)\right)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{f\left(x + \varepsilon, y, z - \frac{1}{2} y \varepsilon\right)}{\varepsilon} \\ &= \partial_x f(p) - \frac{1}{2} y \partial_z f(p) \end{aligned}$$

and similarly

$$Z_2 f(p) = \partial_y f(p) + \frac{1}{2} x \partial_z f(p)$$

and

$$Z_3 f(p) = \partial_z f(p).$$

Observe that $[Z_1, Z_2] = Z_3$ and the family $\{Z_1, Z_2, Z_3\}$ indeed forms a basis for the Lie algebra. We have the decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2,$$

where $\mathfrak{g}_1 = \mathrm{span}\{Z_1, Z_2\}$ and $\mathfrak{g}_2 = \mathrm{span}\{Z_3\}$. Moreover, $[\mathfrak{g}, \mathfrak{g}_2] = 0$ so that indeed $\mathfrak{g}_2$ is the centre of the Lie algebra. In particular, the Heisenberg group is an H-type group with a one dimensional centre. One may define higher dimensional Heisenberg groups $\mathbb{H}^n = (\mathbb{R}^{2n+1}, \cdot)$ where, denoting points in $p \in \mathbb{R}^{2n+1}$ by $p = (x, y, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$,

$$(x, y, t) \cdot (x', y', t') = \left(x + x', y + y', t + t' + \frac{1}{2} S(p, p') \right),$$

with $S(p, p') = \sum_{i=1}^n (x_i y_{n+i} - x_{n+i} y_i)$. The Jacobian basis of $\mathbb{H}^n$ is given by

$$\begin{aligned} Z_i &= \partial_{x_i} - \frac{1}{2} y_{n+i} \partial_t, \text{ for } i = 1, \dots, n, \\ Z_i &= \partial_{y_{i-n}} + \frac{1}{2} x_{i-n} \partial_t, \text{ for } i = n, \dots, 2n, \\ Z_{2n+1} &= \partial_t. \end{aligned}$$

The Heisenberg Lie algebra is the only H-type Lie algebra of dimension $2n + 1$ with a one dimensional centre.

2.1.3 The sub-Laplacian operator

Let $\mathbb{G} = \mathbb{R}^{r_1} \times \cdots \times \mathbb{R}^{r_k}$ be a Carnot group and let $Z_1, \dots, Z_n$ denote its Jacobian basis. There is a natural operator on this group, which can be thought of as the analogue of the Laplacian in Euclidean space.

Definition 2.8. *The operator*

$$\Delta = \sum_{i=1}^{r_1} Z_i^2$$

is called the sub-Laplacian of $\mathbb{G}$.

In other words, the sub-Laplacian is given as the sum of the squares of the Jacobian vector fields that belong to the first layer of the Lie algebra and consequently, it is homogeneous of degree 2 with respect to dilations. In an analogous way, the *sub-gradient* (or *horizontal* gradient) is defined as

$$\nabla = (Z_1, \dots, Z_{r_1}).$$

A deep result for operators defined as sums of squares of vector fields is Hörmander's theorem [Hör67].

Theorem 2.9. *Consider the operator*

$$L = \sum_{i=1}^{m} Z_i^2 + Z_0,$$

where $Z_0, Z_1, \dots, Z_m$ are smooth vector fields on $\mathbb{R}^n$ and assume that

$$\mathrm{Lie}\{Z_0, \dots, Z_m\} = \mathbb{R}^n \tag{2.3}$$

for every $x \in \mathbb{R}^n$. Then, for every distribution u and any open set $\Omega \subset \mathbb{R}^n$, if Lu is smooth on Ω then u must be smooth on Ω.

In particular, the assumption of the theorem is satisfied when L is the sub-Laplacian operator of a Carnot group. An operator satisfying the conclusion of the theorem is called *hypoelliptic*.

2.1.4 The distance on a homogeneous group

Suppose that $\mathbb{G} = (\mathbb{R}^n, \cdot, \delta_\lambda)$ is a homogeneous stratified group, of step k and r_1 generators, $Z_1, \dots, Z_{r_1}$. There is an intrinsic metric on this group, called the *Carnot-Carathéodory* distance, which can be introduced in the following way. Consider two points $x, y \in \mathbb{G}$ and let $\gamma : [0,1] \to \mathbb{G}$ be a path joining them, such that $\gamma(0) = x$ and $\gamma(1) = y$. Such a path is said to be *horizontal* if there exist functions $a_1, \dots, a_{r_1}$ such that

$$\dot\gamma(t) = \sum_{i=1}^{r_1} a_i(t) Z_i(\gamma(t))$$

(the space spanned by $Z_1, \ldots, Z_{r_1}$ is called the *horizontal subspace* of the group). We define the length of γ by

$$\ell(\gamma) = \int_0^1 \sqrt{\sum_{i=1}^{r_1} a_i(t)^2}\, dt.$$

Definition 2.10. *The Carnot-Carathéodory distance between x and y is defined as*

$$d(x,y) = \inf\{\ell(\gamma) \mid \gamma : [0,1] \to \mathbb{G} \text{ is a horizontal path joining } x \text{ and } y\}.$$

Hereafter, $d(x,y)$ will always denote the Carnot-Carathéodory metric and we will write $d(x) = d(0,x)$. On H-type groups (and more generally for stratified Lie algebras), it can be shown that the infimum in the definition is actually a minimum. The Carnot-Carathéodory distance is well-defined thanks to the following theorem of Chow (see e.g. [BLU07]).

Theorem 2.11. *Let $\mathfrak{g}$ be a Lie algebra, generated by the vector fields $Z_1, \ldots, Z_m$, for some $m \in \mathbb{N}$. Then, for any two points x and y there exists a path $\gamma : [0,1] \to \mathbb{G}$ such that $\gamma(0) = x$, $\gamma(1) = y$ and*

$$\dot\gamma(t) = \sum_{i=1}^{m} a_i(t) Z_i(\gamma(t)),$$

for some coefficients $a_1, \ldots, a_m$.

Alternatively, one may define the Carnot-Carathéodory distance between two points by

$$d(x,y) = \sup\{|f(x) - f(y)| : |\nabla f| \le 1\},$$

where $\nabla = (Z_1, \ldots, Z_{r_1})$ denotes the sub-gradient. We moreover have that

$$|\nabla f|(x) = \limsup_{d(x,y)\to 0} \frac{|f(x) - f(y)|}{d(x,y)}. \tag{2.4}$$

This is useful when it comes to passing from functional Sobolev inequalities to isoperimetric inequalities on sets, because it enables us to use several results on metric spaces (such as the coarea inequality) which use the above definition for the gradient.

If $\gamma : [0, d(x,y)] \to \mathbb{G}$ is a horizontal path joining x and y such that $|\dot{\gamma}| = 1$ and $\gamma(0) = x$, $\gamma(d(x,y)) = y$, then for any differentiable f,

$$\begin{aligned} |f(x) - f(y)| &= \left| \int_0^{d(x,y)} \frac{\mathrm{d}}{\mathrm{d}s} f(\gamma(s)) ds \right| \\ &\le \int_0^{d(x,y)} |\nabla f|(\gamma(s)) ds \\ &\le d(x,y) \sup_{0 \le s \le d(x,y)} |\nabla f|(\gamma(s)). \end{aligned}$$

Similarly, one can show that

$$|f(x) - f(y)| \ge d(x,y) \inf_{0 \le s \le d(x,y)} |\nabla f|(\gamma(s)).$$

Dividing by $d(x,y)$ and taking the lim sup as $d(x,y) \to 0$, we arrive at (2.4). A second property that is very important to us is that d satisfies the *eikonal equation*

$$|\nabla d|(x) = 1 \tag{2.5}$$

(where $d = d(\cdot, 0)$ is viewed as a function of one variable). We can see this by considering a geodesic joining 0 to x, i.e. a path $\gamma : [0, d(x)] \to \mathbb{G}$ with $\gamma(0) = 0$, $\gamma(d(x)) = x$ and $|\dot{\gamma}| = 1$ such that for any $s \in [0, d(x)]$, $s = d(\gamma(s))$. Differentiating this equality at $s = d(x)$, we obtain the eikonal equation. One of the basic results that we will need in the sequel (especially in Chapter 3) is the following Laplacian comparison theorem for H-type groups, proved in [HZ10].

Theorem 2.12. *If $\mathbb{G}$ is an H-type group, there exists a constant $K_d > 0$ such that for all $x \in \mathbb{G}$ with $d(x) \ge 1$*

$$\Delta d(x) \le K_d, \tag{2.6}$$

in the sense of distributions.

In fact, as noted in [IP09], this implies that

$$\Delta d \leq \frac{K_d}{d} \tag{2.7}$$

on $\mathbb{G}$. Indeed, by dilations and homogeneity of the sub-Laplacian, for $x \neq 0$,

$$\Delta d(x) = \lambda(\Delta d)(\delta_\lambda(x)),$$

so that by letting $\lambda = 1/d(x)$ and considering $\tilde{x} = \delta_\lambda(x)$ we have $d(\tilde{x}) = 1$ and

$$\Delta d(x) = \frac{\Delta d(\tilde{x})}{d(x)}.$$

Therefore, knowing (2.6) only for x with $d(x) = 1$ is sufficient to deduce (2.7).

The above theorem extends previously known results to a large class of sub-Riemannian spaces. Let us recall that in Euclidean space, if Δ represents the Laplace operator on $\mathbb{R}^n$, it can be checked directly that $\Delta|x| = (n-1)|x|^{-1}$, so that $\Delta|x| \leq n-1$ outside the unit ball. On a Riemannian manifold, if Δ is the Laplace-Beltrami operator and d_R the Riemannian distance, it is known (see e.g. [CLN06, Kas82]) that, if $\mathrm{Ric} \geq (n-1)M$ with some $M \in \mathbb{R}$, then $\Delta d_R \leq (n-1)(d^{-1} + \sqrt{-\min\{M, 0\}})$ so that $\Delta d_R \leq (n-1)(1 + \sqrt{-\min\{M, 0\}})$ outside the unit ball.

It is often not convenient to work with the Carnot-Carathéodory distance because of the lack of an explicit formula for it, as well as because of the fact that it is not smooth (it is not differentiable at any point on the vertical axis, the centre of the group). Instead, one may consider the following norm, known as the Folland-Kaplan gauge (or the Korányi gauge in the case of the Heisenberg group), defined by

$$N(x) = \left(\sum_{j=1}^{k} |x^{(j)}|^{\frac{2k!}{j}} \right)^{\frac{1}{2k!}}, \tag{2.8}$$

where $x^{(j)} \in \mathbb{R}^{n_j}$ (and $|\cdot|$ denotes Euclidean distance). This is a homogeneous norm, i.e. it satisfies

- $N(\delta_\lambda(x)) = \lambda N(x)$ for every $\lambda > 0$ and $x \in \mathbb{G}$,

- $N(x) > 0 \iff x \neq 0$,
- $N(x^{-1}) = N(x)$, for all $x \in \mathbb{G}$.

The Carnot-Carathéodory norm $d(x) = d(0, x)$ is also a homogeneous norm since $d(x, y)$ is a homogeneous metric. The gauge N is smooth out of the origin and it induces a pseudo-distance (see e.g. [BLU07]) defined by

$$d_N(x, y) = N(y^{-1}x). \tag{2.9}$$

Kaplan [Kap80] showed that on H-type groups, there is a constant C such that the locally integrable function

$$\Gamma(x, y) = \frac{C}{d_N(x, y)^{Q-2}}$$

is the fundamental solution of the sub-Laplacian.

Given any two homogeneous norms ρ, $\tilde{\rho}$, there exist constants $c_1, c_2 > 0$ such that for all $x \in \mathbb{G}$

$$c_1\rho(x) \leq \tilde{\rho}(x) \leq c_2\rho(x).$$

This follows from the fact that the norms are homogeneous, since we may take

$$c_1 = \min_{\rho(x)=1} \tilde{\rho}(x)$$

and

$$c_2 = \max_{\rho(x)=1} \tilde{\rho}(x),$$

so that the inequality is satisfied for all x with $\rho(x) = 1$ and hence for all $x \in \mathbb{G}$ by dilations. It is also possible to prove that on any compact set K, a homogeneous norm ρ is comparable to the Euclidean norm in the sense that for all $x \in \mathbb{G}$

$$\frac{1}{C_K}|x|^{1/k} \leq \rho(x) \leq C_K|x|^{1/k},$$

with some constant $C_K > 0$, where k is the step of the group. For the Folland-Kaplan gauge N, this can be seen straight from the definition.

2.1.5 Sobolev and local Poincaré inequalities on H-type groups

The study of Sobolev spaces on sub-Riemannian manifolds has a long history (see e.g. [Jer86, FS82, VSCC92, SC02, RS76, GN96, HL08] and references therein). In what follows, we describe two functional inequalities that are available in the spaces that we are considering and will be essential for the methods used in Chapters 3-5. The first one is a Sobolev inequality, while the second one is an L^1 Poincaré (or Cheeger) inequality in the ball.

Let $\mathbb{G} = (\mathbb{R}^{m+r}, \cdot)$ be an H-type group and let dx denote the Haar measure of the group, which can be identified with the Lebesgue measure on $\mathbb{R}^n$, where $n = m + r$ [FS82].

Theorem 2.13. *There exist constants $\alpha_{\text{cs}} > 1, A_{\text{CS}} > 0$, $B_{\text{CS}} \geq 0$ such that for all locally Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$*

$$\left(\int |f|^{\alpha_{\text{cs}}} dx\right)^{\frac{1}{\alpha_{\text{cs}}}} \leq A_{\text{CS}} \int |\nabla f| dx + B_{\text{CS}} \int |f| dx. \tag{2.10}$$

Theorem 2.14. *For any $r > 0$ there exists a constant $m_r > 0$ such that for all locally Lipschitz f*

$$\int_{B_r} |f - \bar{f}| \frac{dx}{|B_r|} \leq m_r \int_{B_r} |\nabla f| \frac{dx}{|B_r|}, \tag{2.11}$$

where $B_r = \{d < r\}$, $|B_r| = \int_{B_r} dx$ and $\bar{f} = \int_{B_r} f \frac{dx}{|B_r|}$ is the average of f over the ball B_r.

A proof can be found, for instance, in [VSCC92, SC02].

2.2 Markov semigroups

In this section we give a brief outline of some basic notions in the theory of Markov semigroups, which we will need for Chapters 4 and 5. For a more detailed introduction to Markov semigroups we refer the reader to [DP06]; in [ABC+00, Wan06, Bak94, Bak06, GZ03] they are studied in relation to functional inequalities, while further connections to isoperimetry can be found for instance in [BL96, BCR06, BCR07].

Consider the space $C_b(\mathbb{R}^n)$ of uniformly continuous and bounded functions on $\mathbb{R}^n$ taking real values.

Definition 2.15. *A Markov semigroup $(P_t)_{t\geq 0}$ is a linear family of operators $C_b(\mathbb{R}^n) \to C_b(\mathbb{R}^n)$, such that for all functions $f \in C_b(\mathbb{R}^n)$*

1. $P_{t=0} f = f$,
2. $P_{t+s} f = P_t P_s f$, *for all* $s, t \geq 0$,
3. $P_t 1 = 1$, *for all* $t \geq 0$,
4. $f \geq 0 \Rightarrow P_t f \geq 0$,
5. *For any* $T > 0$ *and any* $f \in C_b(\mathbb{R}^n)$, *the map*

$$\begin{aligned} [0,T] \times \mathbb{R}^n &\to \mathbb{R} \\ (x,t) &\mapsto P_t f(x) \end{aligned}$$

   is continuous.

The semigroup is said to be strongly continuous if the map $s \mapsto P_s$, s>0, is strongly continuous, i.e. for $f \in C_b(\mathbb{R}^n)$, $\|P_t f - f\|_\infty \to 0$ as $t \to 0$.

A Markov semigroup is given by a family of Markov probability kernels $(p_t)_{t\geq 0}$, such that

$$P_t f(x) = \int_{\mathbb{R}^n} f(y) p_t(x, dy).$$

As a consequence, for every $t > 0$ the operator P_t is a contraction, i.e. $\|P_t f\|_\infty \leq \|f\|_\infty$ for all $f \in C_b(\mathbb{R}^n)$, where as usual, $\|\cdot\|_\infty$ denotes the supremum norm.

The infinitesimal generator L of P_t is defined as the operator

$$Lf = \lim_{t\to 0} \frac{P_t f - f}{t}$$

with domain $\mathcal{D}(L)$ consisting of functions $f \in C_b(\mathbb{R}^n)$ satisfying the following condition

$$\exists g \in C_b(\mathbb{R}^n) \text{ such that } \lim_{t \to 0} \frac{P_t f - f}{t} = g \text{ and } \sup_{t \in (0,1]} \frac{\|P_t f - f\|_\infty}{t} < \infty.$$

The Hille-Yosida theorem [Yos95] gives necessary and sufficient conditions for an operator to be the generator of a strongly continuous semigroup, and can be generalised under the weaker assumption of (pointwise) continuity of the semigroup [DP06].

A probability measure μ is said to be *invariant* under P_t, if for all $f \in C_b(\mathbb{R}^n)$ and all $t \geq 0$,

$$\int P_t f d\mu = \int f d\mu.$$

At the level of the generator, this property reads

$$\int L(f) d\mu = 0.$$

When the semigroup is defined as an operator on $L^2(\mu)$, such that for all $f \in L^2(\mu)$

$$\|P_t f - f\|_{L^2(\mu)} \to 0, \text{ as } t \to 0,$$

a stronger requirement is that the semigroup is *symmetric* (or *reversible*), i.e. that it satisfies

$$\int f P_t g d\mu = \int (P_t f) g d\mu$$

for all f, g, or, equivalently, that

$$\int f(Lg) d\mu = \int (Lf) g d\mu.$$

The *carré du champ* is defined as the nonnegative bilinear form

$$\Gamma(f,g) = \frac{1}{2} L(fg) - gLf - fLg,$$

while the Γ_2 operator is defined by

$$\Gamma_2(f,g) = \frac{1}{2}L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(g,Lf).$$

This gives rise to a Dirichlet form on $L^2(\mu)$ which is defined by

$$\mathcal{E}(f,f) = \int \Gamma(f)d\mu = \int f(-Lg)d\mu.$$

We will write for short $\Gamma(f) = \Gamma(f,f)$ and $\Gamma_2(f,f) = \Gamma_2(f)$.

For example, for a smooth function V such that $\int \mathrm{e}^{-V}dx < \infty$, we consider the generator

$$L = \Delta - (\nabla V)\cdot\nabla,$$

where Δ and ∇ denote the Laplacian and the Euclidean gradient respectively, acting on smooth functions. Then L generates a strongly continuous Markov semigroup $P_t = \mathrm{e}^{tL}$ and the corresponding invariant probability measure is

$$\mu(dx) = \frac{\mathrm{e}^{-V}}{\int \mathrm{e}^{-V}dx}dx.$$

In this case, one can calculate that $\Gamma(f) = |\nabla f|^2$ and $\Gamma_2(f) = \|D^2 f\|_{HS} + \langle D^2 V\nabla f, \nabla f\rangle$, where $(D^2 f)_{ij} = (\partial_i\partial_j f)_{ij}$ denotes the matrix of second derivatives of f and $\|\cdot\|_{HS}$ the Hilbert-Schmidt norm of a matrix .

We will say that L is a diffusion generator, if for all smooth φ,

$$L(\varphi(f)) = \varphi'(f)L(f) + \varphi''(f)\Gamma(f). \tag{2.12}$$

For $\rho > 0$ and $n \in \mathbb{N}$, the *curvature-dimension condition* $CD(\rho,n)$ is satisfied if for all $f \in \mathcal{D}(L)$

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{n}(Lf)^2,$$

while the $CD(\rho,\infty)$ condition is said to hold if

$$\Gamma_2(f) \geq \rho\Gamma(f).$$

The latter inequality has several equivalent characterisations in the context of Riemannian manifolds (see e.g. Proposition 3.3 of [Bak06]) one of which is the following gradient bound for the semigroup

$$\Gamma(P_t f) \leq \mathrm{e}^{-2\rho t} P_t \Gamma(f). \tag{2.13}$$

Such inequalities and their applications to Sobolev inequalities and hypercontractivity were investigated in [BÉ85]. The semigroup P_t is said to be *hypercontractive* if for all $1 \leq p \leq q < \infty$, there exists a constant $T = T(p,q)$ such that

$$\|P_t f\|_q \leq \|f\|_p$$

for all $t > T$.

In the sub-elliptic setting, there is an equivalent formulation of Hörmander's theorem in probabilistic language (see e.g. [Nua06]).

Theorem 2.16. *Let L be as in Theorem 2.9 and suppose that $P_t = \mathrm{e}^{tL}$ is the semigroup generated by L. Then P_t has a smooth density with respect to the Lebesgue measure.*

2.3 Functional inequalities, isoperimetry and transportation

2.3.1 Sobolev-type inequalities

We begin by introducing two inequalities that have been studied extensively in the literature, the Poincaré inequality and the logarithmic Sobolev inequality. For a more detailed account of the theory of functional inequalities, we refer the reader to [GZ03, ABC$^+$00, Bak94, Wan06]. Let d be a metric on $\mathbb{R}^n$ and suppose that the gradient of a function is given by (2.4). A probability measure μ is said to satisfy the *Poincaré inequality* if there exists a constant C_P such that for all locally Lipschitz f

[2]

$$\operatorname{Var}_\mu(f) = \int (f - \mu(f))^2 d\mu \le C_P \int |\nabla f|^2 d\mu, \tag{2.14}$$

where $\mu(f) := \int f d\mu$. The Poincaré inequality is also known as the *spectral gap* inequality, because the optimal constant C_P is the reciprocal of the first (nontrivial) eigenvalue of the generator associated to the measure μ (see e.g. [Wan06]) . More generally, we will say that, for $q \ge 1$, a q-Poincaré inequality holds if there exists a constant C such that

$$\int |f - \mu(f)|^q d\mu \le C \int |\nabla f|^q d\mu,$$

for all locally Lipschitz f. The case $q = 1$ is of special interest in relation to the isoperimetric problem, and is known as *Cheeger's inequality.* A strong result of E. Milman [Mil08] states that under certain convexity conditions on the space if the q-Poincaré inequality holds for some q then it holds for all q. This generalises previous results of Buser [Bus82] (see also [Led94]) and Cheeger [Che70] which state that the Cheeger inequality and the Poincaré inequality (for $q = 2$) are equivalent (always under convexity assumptions). In the absence of convexity assumptions, if $q < q'$, a q-Poincaré inequality always implies a q'-Poincaré inequality. The following result will be of use to us in the future.

Proposition 2.17. *Suppose that a measure μ satisfies*

$$\int |f - \mu(f)| d\mu \le C_{Che} \int |\nabla f| d\mu,$$

for all locally Lipschitz f, with some constant C_{Che} independent of f. Then, for all $q \ge 1$ there exists a constant C_P depending only on C_{Che} and q, such that for all

[2] It suffices for the inequality to hold for bounded smooth f with compact support. An approximation argument then shows that it holds for locally Lipschitz functions, which are almost everywhere differentiable by Rademacher's theorem.

nonnegative locally Lipschitz f

$$\int |f-\mu(f)|^q \leq C_P \int |\nabla f|^q d\mu.$$

A proof can be found in [BZ05].

The measure μ is said to satisfy a *logarithmic Sobolev inequality* if there exists a constant C_{LS} such that for all locally Lipschitz f

$$\operatorname{Ent}_\mu(f^2) = \int f^2 \log\left(\frac{f^2}{\mu(f^2)}\right) d\mu \ \leq C_{LS} \int |\nabla f|^2 d\mu. \tag{2.15}$$

This inequality was studied by Gross [Gro75] for the Gaussian measure, in relation to the study of hypercontractivity. By applying it to the function $f = 1+\varepsilon g$ for some $\varepsilon > 0$ and then letting $\varepsilon \to 0$ we recover the Poincaré inequality (2.14), which is strictly weaker than (2.15). In Euclidean space $(\mathbb{R}^n, |\cdot|)$, for instance, it is known that (2.14) is satisfied for all log-concave measures [Bob99], while (2.15) implies the integrability condition

$$\int e^{\varepsilon|x|^2} d\mu < \infty,$$

for some $\varepsilon > 0$ [Wan97, Bob99]. A property that is shared by both (2.14) and (2.15) is the stability under tensorisation. In other words, if the inequality holds for two measures μ on $(\mathbb{R}^n, d)$ and ν on $(\mathbb{R}^m, \rho)$, then it also holds for the measure $\mu \otimes \nu$ on $(\mathbb{R}^{n+m}, \delta)$, equipped with the distance $\delta = \sqrt{d^2+\rho^2}$.

There are numerous generalisations of the logarithmic Sobolev inequality in the literature, some of which we present below. For $q \in (1,2]$, the q-logarithmic Sobolev inequality (see e.g. [GZ03]) states that there exists a constant C such that for all locally Lipschitz f

$$\operatorname{Ent}_\mu(f^q) \leq C \int |\nabla f|^q d\mu.$$

Let $\Phi : \mathbb{R} \to \mathbb{R}$ be a smooth convex function. We define the Φ-entropy of a function by

$$\operatorname{Ent}_\mu^\Phi(f) = \int \Phi(f) d\mu - \Phi\left(\int f d\mu\right).$$

It is a nonnegative quantity as Jensen's inequality shows. The measure μ is said to satisfy an L^q Φ-entropy inequality if there exists a constant C_Φ such that for all locally Lipschitz f

$$\mathrm{Ent}_\mu^\Phi(f) \le C_\Phi \int |\nabla f|^q d\mu.$$

A comprehensive study of these inequalities can be found in [Cha04]. Typical examples of the function $\Phi(t)$ include t^2, which corresponds to the Poincaré inequality, $t^q \log(t^q)$ which gives the q-logarithmic Sobolev inequality, while for $\beta \in (0,1)$, one may consider the functions $|t|(\log(1+|t|))^\beta$. In Chapter 3 we will study L^1 Φ-entropy inequalities.

Although Φ-entropy inequalities preserve the additive nature of the logarithmic Sobolev inequality, they are not necessarily homogeneous any more. The homogeneity is preserved by another generalisation of (2.15), known as the F-Sobolev inequality. Consider an increasing function $F : [0,\infty) \to \mathbb{R}$ such that $F(0) = 1$. The measure μ satisfies an L^q F-Sobolev inequality with constant C_F, if for all locally Lipschitz f

$$\int |f|^q F\left(\frac{|f|^q}{\mu(|f|^q)}\right) d\mu \le C_F \int |\nabla f|^q d\mu.$$

In relation to the isoperimetric problem, Φ-entropy and F-Sobolev inequalities (with $q = 2$) were used for instance in [BCR06, BCR07], as well as [Led88] and [Mil09a] (under the more general form of Orlicz-Sobolev inequalities).

We conclude this section with the following theorem, which generalises ideas of Rothaus [Rot85].

Theorem 2.18. *Let $q \in (1,2]$ and suppose that a defective q-logarithmic Sobolev inequality holds, i.e. that there exist constants C, D such that for all locally Lipschitz f*

$$\int |f|^q \log\left(\frac{|f|^q}{\mu|f|^q}\right) d\mu \le C \int |\nabla f|^q d\mu + D \int |f|^q d\mu.$$

If the measure μ also satisfies a q-Poincaré inequality, then the q-logarithmic Sobolev inequality holds.

For the proof we need the following results from [BZ05]. We define the Orlicz space $L_{N_q}(\mu)$ generated by the function $N_q(x) = |x|^q \log(1+|x|^q)$ to be the space of measurable functions f such that

$$\|f\|_{N_q}^q := \inf\left\{\lambda > 0 : \int N_q\left(\frac{f}{\lambda}\right) d\mu \leq 1\right\} < \infty.$$

Lemma 2.19. *If $f \geq 0 \in L_{N_q}(\mu)$ then*

$$\|f\|_{N_q}^q \leq \int f^q \log \frac{f^q}{\mu(f^q)} d\mu + \int f^q d\mu.$$

Lemma 2.20. *If $f \in L_{N_q}(\mu)$ and $\mu(f) = 0$ then*

$$\sup_{a\in\mathbb{R}} \int |f+a|^q \log \frac{|f+a|^q}{\mu|f+a|^q} d\mu \leq 16\|f\|_{N_q}^q.$$

With these results in hand, suppose that we have a defective LS_q inequality

$$\int |f|^q \log\left(\frac{|f|^q}{\mu|f|^q}\right) d\mu \leq C\int |\nabla f|^q + D\int |f|^q d\mu, \tag{2.16}$$

as well as a q-Poincaré inequality

$$\int |f - \mu f|^q d\mu \leq C_P \int |\nabla f|^q d\mu. \tag{2.17}$$

By Lemma 2.19 applied to the function $|f - \mu(f)|$ we have

$$\begin{aligned}
\|f-\mu(f)\|_{N_q}^q &\leq \int |f-\mu(f)|^q \log\left(\frac{|f-\mu(f)|^q}{\int |f-\mu(f)|^q d\mu}\right) d\mu \ + \int |f-\mu(f)|^q d\mu \\
&\leq C\int |\nabla f|^q d\mu + (D+1)\int |f-\mu(f)|^q d\mu \\
&\leq (C + C_P(D+1)) \int |\nabla f|^q d\mu,
\end{aligned}$$

where we first used (2.16) and then (2.17). Finally, by applying Lemma 2.20 to

$f - \mu(f)$ we obtain

$$\int |f|^q \log\left(\frac{|f|^q}{\mu|f|^q}\right) d\mu = \int |f - \mu(f) + \mu(f)|^q \log\left(\frac{|f - \mu(f) + \mu(f)|^q}{\int |f - \mu(f) + \mu(f)|^q d\mu}\right) d\mu$$
$$\leq 16\|f - \mu f\|_{N_q}^q$$
$$\leq 16(C + K(D+1)) \int |\nabla f|^q d\mu.$$

2.3.2 Isoperimetric inequalities

Consider a metric d on $\mathbb{R}^n$ and let μ be a probability measure. The surface measure of a Borel set A is defined by

$$\mu^+(A) = \liminf_{h\to 0} \frac{\mu(A^h) - \mu(A)}{h},$$

where $A^h = \{x \in A : d(x, A) < h\}$ is the open h-neighbourhood (or h-enlargement) of A (with respect to d). The *isoperimetric problem* consists of minimising the surface area $\mu^+(A)$ over all sets of equal measure $\mu(A) = t$, for $t \in [0,1]$ (see e.g. [Led01] for an introduction to the subject). In other words, one looks for those sets A that have measure t and among all sets B such that $\mu(A) = \mu(B) = t$, A has minimal surface area $\mu^+(A)$. Such sets are called extremal and identifying them is a very difficult question in most cases, especially in dimension $n \geq 2$. A second question, which we will study in what follows, is to estimate the *isoperimetric function* (or *isoperimetric profile*) of μ, which is defined as the optimal $\mathcal{I} = \mathcal{I}_\mu$ in the isoperimetric inequality

$$\mu^+(A) \geq \mathcal{I}(\mu(A)), \tag{2.18}$$

valid for all measurable sets A. In other words, for $t \in [0,1]$

$$\mathcal{I}(t) = \inf\{\mu^+(A) : A \text{ has measure } \mu(A) = t\}$$

(and so extremal sets are the ones for which (2.18) becomes equality). The isoperimetric function is explicitly known only in a few cases, some of which we present below.

Let μ denote the Lebesgue measure on $\mathbb{R}^n$. Then $\mathcal{I}_\mu(t) = n\ \sqrt[n]{\omega_n}\ t^{\frac{n-1}{n}}$, where ω_n is the volume of the n-dimensional Euclidean ball of radius 1. In this setting, it is known that Euclidean balls solve the isoperimetric problem.

For the two-sided exponential distribution $\nu_1(dx) = \frac{1}{2}e^{-|x|}dx$ on the real line, the isoperimetric profile is $\mathcal{I}_{\nu_1}(t) = \min(t, 1-t)$. A systematic approach to determining the isoperimetric profile for a large class of probability measures on the real line was given in [BH97b].

On $\mathbb{R}$ with Gaussian measure $\nu_2 = \gamma$ of density $\phi(x) = \frac{e^{-|x|^2/2}}{(2\pi)^{n/2}}$ the isoperimetric function is known to be $I_\gamma = \phi \circ \Phi^{-1}$ where $\Phi' = \phi$. This is result of Sudakov-Tsirelson [ST74] and Borell [Bor74] who showed that half-lines $H_s = \{x \in \mathbb{R} : x \leq s\}$ solve the isoperimetric problem for γ. Moreover, they showed that, for any $n \geq 2$, $\mathcal{I}_{\gamma^{\otimes n}} = \mathcal{I}_\gamma$; this property actually characterises Gaussian measures [BH96].

More generally, for the measure ν_p, with density $f_p(x)dx = \frac{e^{-|x|^p/p}}{Z_p}dx$, $x \in \mathbb{R}$, $p \geq 1$, the isoperimetric function is $\mathcal{I}_{\nu_p} = f_p \circ F_p^{-1}$, where $F_p' = f_p$. For $p \in [1, 2]$, we have $\mathcal{I}_{\nu_p^{\otimes n}} \asymp \mathcal{I}_{\nu_p}$ (with equality when $p = 2$ as mentioned above) [BCR06], where $\asymp$ denotes Lipschitz equivalence, i.e. that there exists a constant $\kappa > 0$ such that

$$\frac{1}{\kappa} \leq \frac{\mathcal{I}_{\nu_p^{\otimes n}}}{\mathcal{I}_{\nu_p}} \leq \kappa.$$

It is moreover known (see e.g. [BZ05]) that if $q = p/(p-1)$ denotes the conjugate exponent of p, then

$$\mathcal{I}_{\nu_p}(t) \asymp \hat{t}\left(-\log \hat{t}\right)^{1/q},$$

where $\hat{t} = \min(t, 1-t)$. This recovers the case $p = 1$ mentioned above, as well as the Gaussian case ($p = 2$) where a Taylor expansion at 0 shows that $\phi \circ \Phi^{-1}(t) \approx t\sqrt{-2\log t}$ for small t (and similarly for t near 1).

Definition 2.21. *We will say that a probability measure μ satisfies a q-*isoperimetric inequality*, if there exists a constant $C_{iso} > 0$ such that, for all measurable A with*

$\mu(A) = a$,

$$\mu^+(A) \geq C_{iso}\hat{a}(-\log \hat{a})^{1/q}, \tag{2.19}$$

where $\hat{a} = \min(a, 1-a)$.

Equivalently, there exists a constant $D > 0$ such that

$$\mathcal{I}_\mu \geq D\mathcal{I}_{\nu_p},$$

where ν_p are the measures introduced above. The inequality (2.19) has various equivalent functional forms, amongst which, as we shall see, are L^1 Φ-Entropy inequalities and F-Sobolev inequalities. These equivalent forms were extensively studied in the literature (see e.g. for instance [Bar01, BK08, BL96, Led94, Bob97a, Bob96b, BM00, BH97b, Mil09a, Mil09b, BCR06, BCR07, BZ05]). A full equivalence between Orlicz-Sobolev inequalities and isoperimetric inequalities was given in [Mil09a], under some convexity assumptions on the space, in particular involving a lower bound on the Ricci curvature. These convexity assumptions are usually needed in order to deduce an isoperimetric inequality on sets such as (2.19), starting from a functional inequality involving the q^{th} power of the length of the gradient on the right-hand-side, with $q > 1$, such as the q-logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu(|f|^q) \leq C \int |\nabla f|^q.$$

In the sub-elliptic case, the convexity assumptions made in the aforementioned works are not satisfied. Nevertheless, it is still possible to deduce isoperimetric information from functional inequalities, which include the L^1 norm of the gradient, such as an F-Sobolev inequality of the form

$$\int f(\log_+ f)^{1/q} d\mu \leq C \int |\nabla f| + (\log 2)^{1/q} \int f d\mu.$$

To pass from such an inequality on functions to an inequality on sets, one considers a Lipschitz approximation $(f_n)_{n\in\mathbb{N}}$ of the characteristic function of a set. In the functional inequality for f_n, it is then possible to pass to the limit as $n \to \infty$. This procedure can be carried out in the general setting of a metric measure space, as following result of [BH97a] suggests.

Proposition 2.22. *Suppose (X, d) is a metric space and μ a probability measure on X. Let $A \subset X$ be a Borel set of measure $\mu(A) \in (0, 1)$ and let $\bar{A}$ be its closure. Then, there exists a sequence of Lipschitz functions $(f_n)_{n\in\mathbb{N}}$, $f_n : X \to [0, 1]$, such that for all $x \in X$*

$$f_n(x) \to \chi_{\bar{A}}(x)$$

and

$$\limsup_{n\to\infty} \int |\nabla f_n| d\mu \leq \mu^+(A)$$

This gives us a way to approximate the indicator function of the closure of a set, rather than of the set itself, but it is sufficient to obtain the isoperimetric inequality. To see this, suppose that we start from some functional inequality and given a set A, we apply it to f_n as above to deduce, in the limit, $\mathcal{I}(\mu(\bar{A})) \leq \mu^+(A)$. Then, either $\mu(A) = \mu(\bar{A})$, in which case we arrive at the isoperimetric inequality for μ, or $\mu(\bar{A}) > \mu(A)$, in which case the definition of surface measure gives $\mu^+(A) = \infty$. Either way, the isoperimetric inequality $\mathcal{I}(\mu(A)) \leq \mu^+(A)$ is satisfied.

Conversely, to pass from an isoperimetric inequality on sets to an analytic inequality, we will use the following generalised coarea inequality of [BH97a] (which extends the usual coarea formula, see e.g. [Fed69]).

Proposition 2.23. *Let f be a Lipschitz function on a metric measure space (X, d, μ). Then*

$$\int |\nabla f| d\mu \geq \int_{-\infty}^{\infty} \mu^+(\{f \geq t\}) dt. \tag{2.20}$$

We conclude this section by describing two functional forms of the isoperimetric inequality, which were introduced by S.G. Bobkov in [Bob96b] and [Bob97a] in relation to the isoperimetric problem for Gaussian measures (see also [BL96] where the inequalities were extended to other probability measures on Riemannian manifolds).

The first one states that for all locally Lipschitz $f : \mathbb{R}^n \to [0,1]$,

$$\mathcal{I}(\mu(f)) - \int \mathcal{I}(f) d\mu \leq C \int |\nabla f| d\mu, \tag{2.21}$$

while the second one reads

$$\mathcal{I}(\mu(f)) \leq \int \sqrt{\mathcal{I}(f)^2 + C|\nabla f|^2} d\mu. \tag{2.22}$$

Since $a^2 + b^2 \leq (a+b)^2$ for positive numbers a, b, the second inequality implies the first one. On the other hand, both inequalities imply the isoperimetric inequality

$$\mathcal{I}(\mu(A)) \leq C\mu^+(A)$$

in the limit as f approximates the indicator function of a set A. The advantage of the second form (2.22) is that it has the tensorisation property.

In the case where $\mathcal{I}$ is the Gaussian isoperimetric function, we have the following equivalence theorem of [BM00].

Theorem 2.24. *Let μ be a probability measure on $(\mathbb{R}^n, d)$ and let $\mathcal{U} = \phi \circ \Phi^{-1}$ where ϕ is the density of the standard Gaussian measure on $\mathbb{R}$ and $\Phi' = \phi$. Then the following inequalities are equivalent (with the same constant C):*

(1) *$\mathcal{U}(\mu(A)) \leq C\mu^+(A)$, for all Borel measurable A,*

(2) *$\mathcal{U}(\mu(f)) - \int \mathcal{U}(f) d\mu \leq C \int |\nabla f| d\mu$, for all locally Lipschitz $f : \mathbb{R}^n \to [0,1]$,*

(3) *$\mathcal{U}(\mu(f)) \leq \int \sqrt{\mathcal{U}(f)^2 + C^2|\nabla f|^2} d\mu$.*

The proof given in [BM00] is in the setting of Riemannian manifolds, but it extends to arbitrary metric spaces, provided that we have a coarea inequality, which is guaranteed by Proposition 2.23.

Outside the Gaussian case, the inequality (2.22) is known to hold for the two-sided exponential distribution ν_1 on $\mathbb{R}$ [Bob97b, Bob09]

$$\mathcal{I}_{\nu_1}(\nu_1(f)) \leq \int \sqrt{\mathcal{I}_{\nu_1}(f)^2 + \frac{1}{2\sqrt{6}}|\nabla f|^2} d\nu_1,$$

but the question of whether the measures ν_p introduced above on the real line, with $p \in (1,2)$, satisfy

$$\mathcal{I}_{\nu_p}(\nu_p(f)) \leq \int \sqrt{\mathcal{I}_{\nu_p}(f)^2 + C_p^2 |\nabla f|^2} d\nu_p,$$

for some constant C_p, remains open. Such a conjecture is expected to be true, since it would reflect the fact that the isoperimetric profile of these measures is stable with respect to the dimension [BCR06], in other words that there exists a constant $c > 0$ such that for all $n \in \mathbb{N}$

$$c\mathcal{I}_{\nu_p} \leq \mathcal{I}_{\nu_p^{\otimes n}} \leq \mathcal{I}_{\nu_p}$$

(the inequality on the right is always true; e.g. considering the set $A \times \mathbb{R} \subset \mathbb{R}^2$, for some $A \subset \mathbb{R}$, shows that $\mathcal{I}_{\mu\otimes\mu} \leq \mathcal{I}_\mu$). Although such an inequality cannot be true for $p > 2$ (because the stability with respect to the dimension breaks in this case), it was conjectured in [BZ05, Zeg01] that the following inequality holds with $q = p/(p-1)$

$$\mathcal{I}_{\nu_p}(\nu_p(f)) \leq \int \left(\mathcal{I}_{\nu_p}(f)^q + C_p^2 |\nabla f|^q\right)^{1/q} d\nu_p.$$

2.3.3 Transportation of measure

Let $p \geq 1$ and suppose μ, ν are two probability measures on $\mathbb{R}^n$ with finite p^{th} moments, i.e. for some $y \in \mathbb{R}^n$ (and hence for all y by the triangle inequality),

$$\int_{\mathbb{R}^n} d(x,y)^p d\mu(x), \int_{\mathbb{R}^n} d(x,y)^p d\nu(x) < \infty.$$

We define the p-*Wasserstein distance* between μ and ν by

$$\mathcal{W}_p(\mu,\nu)^p = \inf \int \frac{d(x,y)^p}{p} d\pi(x,y), \tag{2.23}$$

where the infimum is taken over all couplings π of μ, ν, i.e. all probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x) d\pi(x,y) = \int_{\mathbb{R}^n} \phi(x) d\mu(x), \quad \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(y) d\pi(x,y) = \int_{\mathbb{R}^n} \phi(y) d\nu(y),$$

for all bounded and measurable ϕ on $\mathbb{R}^n$. The Wasserstein distance has the following characterisation, known as the Kantorovich-Rubinstein duality.

Theorem 2.25. *For μ and ν as above,*

$$\mathcal{W}_p(\mu,\nu)^p = \sup\left(\int f d\mu - \int g d\nu\right), \tag{2.24}$$

where the supremum is taken over all bounded Lipschitz $f, g : \mathbb{R}^n \to \mathbb{R}$ satisfying

$$f(x) \le g(y) + \frac{d(x,y)^p}{p}$$

for all $x, y \in \mathbb{R}^n$.

This theorem can be extended to arbitrary cost functions in Polish spaces; for a concise introduction to the theory of optimal transportation we refer the reader to [Vil03]. When $p = 1$, we have

$$\mathcal{W}_1(\mu,\nu) = \sup\left(\int f d\mu - \int f d\nu\right), \tag{2.25}$$

where the supremum is taken over all Lipschitz f with $\|f\|_{\mathrm{Lip}} \le 1$. Given a bounded Lipschitz function f, we can define a one-parameter family of operators by

$$(Q_t f)(x) = \inf_{y\in\mathbb{R}^n}\left(f(y) + \frac{d(x,y)^p}{pt^{p-1}}\right), \tag{2.26}$$

for $t > 0$. This family is known to form a semigroup, known as the *Hamilton-Jacobi semigroup*, which solves the equation

$$\partial_t Q_t f = -\frac{|\nabla Q_t f|^q}{q}$$

([Eva10], see also [LV07, BEHM09] for some basic properties of Q_t). By the *infimum convolution* of f, we mean the function $Qf = Q_1 f$. Observe that by the duality theorem,

$$\mathcal{W}_p(\mu,\nu)^p = \sup\left(\int Qf d\mu - f d\nu\right), \tag{2.27}$$

where the supremum is taken over all bounded Lipschitz f.

A measure μ is said to satisfy a $\mathbb{T}_p$ transportation inequality with $p \in [1,2]$, if for all probability measures ν which are absolutely continuous with respect to μ, we have

$$\mathcal{W}_p(\mu,\nu) \leq C_T \sqrt{\mathrm{Ent}_\mu(\rho)} \tag{2.28}$$

for some constant C_T independent of ν, where $\rho = d\nu/d\mu$. For $p \geq 2$, we will say that μ satisfies a p-transportation inequality $\mathbb{T}_p$ if, for probability measures ν as above,

$$\mathcal{W}_p(\mu,\nu) \leq C_T \left(\mathrm{Ent}_\mu(\rho)\right)^{1/p}. \tag{2.29}$$

The particular case $p = 2$ is known as Talagrand's inequality [Tal96]. Following [BEHM09], we remark that if ρ is of the form $1 + \varepsilon\tilde{\rho}$, with some function $\tilde{\rho}$, then as $\varepsilon \to 0$, $\mathrm{Ent}_\mu(\rho)$ is of order ε^2, while $\mathcal{W}_p(\mu,\nu)^p$ is typically of order ε^p . Therefore, we see that for $p < 2$ the second inequality (2.29) cannot hold.

Chapter 3

Sobolev-type inequalities and isoperimetry

This chapter is devoted to the study of functional inequalities for probability measures and their consequences on isoperimetry. The results that we present below apply to the following three models.

1. The n-dimensional Euclidean space $(\mathbb{R}^n, |\cdot|, dx)$, where $|\cdot|$ denotes the standard Euclidean norm and dx the Lebesgue measure.

2. A Riemannian manifold $(\mathbb{R}^n, d_R, d_{vol})$ where d_R is a Riemannian metric and d_{vol} the Riemannian volume element.

3. An H-type group $(\mathbb{R}^n, d_{cc}, dx)$ where d_{cc} denotes the Carnot-Carathéodory metric and dx denotes the Haar measure of the group.

Although the first two cases are of great interest, as mentioned in the introduction, they were well-studied in the literature and we will therefore focus on the third one. We will work in the setting of an H-type group on $\mathbb{R}^n$ and denote by $d = d_{cc}$ the Carnot-Carathéodory distance. We moreover write $d(x) = d(x, 0)$, while ∇ stands for the sub-gradient.

3.1 U-bounds

Consider the probability measure $\mu(dx) = Z^{-1}\mathrm{e}^{-U}dx$, where U is a smooth function, while $Z = \int \mathrm{e}^{-U}dx$ is a normalisation constant. We will say that μ satisfies a *U-bound* if there are constants $\beta \in (0,1]$ and $C_U, D_U \geq 0$ such that

$$\int |f|\left(|U|^\beta + |\nabla U|\right) d\mu \leq C_U \int |\nabla f| d\mu + D_U \int |f| d\mu, \tag{3.1}$$

for all locally Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$. This inequality is an L^1 analogue of the U-bounds introduced in [HZ10] as a tool to prove functional inequalities.

Let us examine the above inequality when $U(x) = d(x)^p/p$ with some $p > 1$. In this case, since d satisfies the eikonal equation, (3.1) reads

$$\int |f|\left(d^{\beta p} + p d^{p-1}\right) d\mu \leq C \int |\nabla f| d\mu + D \int |f| d\mu.$$

We want this to be satisfied for some $\beta \in (0,1]$ and a natural choice would be

$$\beta = \frac{1}{q} := 1 - \frac{1}{p},$$

since then we would have $\beta p = p - 1$. This inequality is indeed satisfied, as the following theorem shows.

Theorem 3.1 ([HZ10]). *Let $\mu(dx) = Z^{-1}\mathrm{e}^{-d(x)^p/p}dx$. For all locally Lipschitz f*

$$\int |f| d^{p-1} d\mu \leq \int |\nabla f| d\mu + \left(2^{p-1} + K_d + 1\right) \int |f| d\mu. \tag{3.2}$$

Proof. We may assume that $f \geq 0$. Suppose first that f is supported outside the unit ball, so that by (2.6) we have $-\Delta d \geq -K_d$ on $\mathrm{supp}(f)$. Taking the inner product with ∇d of the expression

$$\nabla\left(f\mathrm{e}^{-d^p/p}\right) = (\nabla f)\,\mathrm{e}^{-d^p/p} - f d^{p-1}(\nabla d)\mathrm{e}^{-d^p/p}$$

and using that $|\nabla d| = 1$, then integrating over $\mathbb{R}^n$ with respect to Lebesgue measure we obtain

$$\int \left\langle \nabla \left(f \mathrm{e}^{-d^p/p} \right), \nabla d \right\rangle dx = \int \left\langle \nabla f, \nabla d \right\rangle \mathrm{e}^{-d^p/p} dx - \int f d^{p-1} \mathrm{e}^{-d^p/p} dx.$$

For the left-hand-side, an integration by parts and Theorem 2.12 give

$$\int \left\langle \nabla \left(f \mathrm{e}^{-d^p/p} \right), \nabla d \right\rangle dx = \int f(-\Delta d) \mathrm{e}^{-d^p/p} dx \geq -K_d \int f \mathrm{e}^{-d^p/p} dx,$$

and using the Cauchy-Schwarz inequality $\langle \nabla f, \nabla d \rangle \leq |\nabla f||\nabla d| = |\nabla f|$, we arrive at

$$-K_d \int f \mathrm{e}^{-d^p/p} dx \leq \int |\nabla f| \mathrm{e}^{-d^p/p} dx - \int f d^{p-1} \mathrm{e}^{-d^p/p} dx.$$

Rearranging and dividing by the normalization constant Z we conclude that if f is supported outside the unit ball, then

$$\int f d^{p-1} d\mu \leq \int |\nabla f| d\mu + K_d \int f d\mu. \tag{3.3}$$

For general $f \geq 0$, we may decompose f into two parts as $f = \phi f + (1-\phi) f =: f_1 + f_2$, where $\phi : \mathbb{R}^n \to [0,1]$ is defined as

$$\phi(x) = \begin{cases} 1, & \text{for } x \in \{d < 1\}, \\ 2 - d(x), & \text{for } x \in \{1 \leq d < 2\}, \\ 0, & \text{for } x \in \{d \geq 2\}. \end{cases} \tag{3.4}$$

Notice that $\operatorname{supp}(f_2) \cap \{d < 1\} = \emptyset$ and so we may apply (3.3) to f_2. We have

$$\begin{aligned} \int f d^{p-1} d\mu &= \int_{\{d<2\}} f d^{p-1} d\mu \ + \int_{\{d \geq 2\}} f d^{p-1} d\mu \\ &\leq 2^{p-1} \int f d\mu + \int f_2 d^{p-1} d\mu \\ &\leq 2^{p-1} \int f d\mu + \int |\nabla f_2| d\mu + K_d \int f_2 d\mu. \end{aligned}$$

It remains to note that $|\nabla f_2| \le |\nabla f| + f$ which gives

$$\int f d^{p-1} d\mu \le \int |\nabla f| d\mu + \left(2^{p-1} + K_d + 1\right) \int f d\mu,$$

which is what we wanted to show. ☐

Remark 3.2. *The above result is actually a special case of the following stronger statement proved in [HZ10]: For all locally Lipschitz f and all $q \ge 1$,*

$$\int |f|^q d^{q(p-1)} d\mu \le \int |\nabla f|^q d\mu + D_q \int |f|^q d\mu,$$

for some constant $D_q > 0$.

In principle, given a general potential U, proving (3.1) may not be an easy task. Nevertheless, the following perturbation result holds, which, coupled with Theorem 3.1, allows us to extend our results to a large class of probability measures.

Proposition 3.3. *Suppose that a probability measure $\mu(dx) = Z^{-1} \mathrm{e}^{-U(x)} dx$ satisfies*

$$\int |f| \left(|U|^\beta + |\nabla U|\right) d\mu \le C_U \int |\nabla f| d\mu + D_U \int |f| d\mu \tag{3.5}$$

and let $\tilde{\mu}(dx) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{U}} dx$ be another probability measure, where $\tilde{U} = U + W$ and $W : \mathbb{R}^n \to \mathbb{R}$ is a smooth function such that

$$\int \mathrm{e}^{-W} d\mu < \infty$$

and

$$|\nabla W| \le \delta \left(|U|^\beta + |\nabla U|\right) + L_\delta, \quad |W|^\beta \le c_1 \left(|U|^\beta + |\nabla U|\right) + c_2, \tag{3.6}$$

for some constants $0 \le \delta < C_U^{-1}$ and $L_\delta, c_1, c_2 \ge 0$. Then, the measure $\tilde{\mu}$ satisfies

$$\int |f| \left(|\tilde{U}|^\beta + |\nabla \tilde{U}|^q\right) d\tilde{\mu} \le C_{\tilde{U}} \int |\nabla f| d\tilde{\mu} + D_{\tilde{U}} \int |f| d\tilde{\mu},$$

for some constants $C_{\tilde{U}}, D_{\tilde{U}}$.

Proof. Without loss of generality, we may assume that $f \geq 0$. Applying (3.5) to the function $f e^{-W}$ and multiplying both sides by $Z/\tilde{Z}$ we obtain

$$\begin{aligned}\int f\left(|U|^\beta + |\nabla U|\right) d\tilde{\mu} &\leq C_U \int |\nabla f| d\tilde{\mu} + C_U \int f |\nabla W| d\tilde{\mu} + D_U \int f d\tilde{\mu} \\ &\leq C_U \int |\nabla f| d\tilde{\mu} + \delta C_U \int f\left(|U|^\beta + |\nabla U|\right) d\tilde{\mu} \\ &\quad + (L_\delta C_U + D_U) \int f d\tilde{\mu},\end{aligned}$$

by the triangle inequality. Since $\delta C_U < 1$ by assumption, we may rearrange this inequality to get

$$\int f\left(|U|^\beta + |\nabla U|\right) d\tilde{\mu} \leq C_1 \int |\nabla f| d\tilde{\mu} + C_2 \int f d\tilde{\mu},$$

with

$$C_1 = \frac{C_U}{1 - \delta C_U}, \quad C_2 = \frac{L_\delta C_U + D_U}{1 - \delta C_U}.$$

By our assumptions,

$$\begin{aligned}\int f\left(|W|^\beta + |\nabla W|\right) d\tilde{\mu} &\leq (\delta + c_1) \int f\left(|U|^\beta + |\nabla U|\right) d\tilde{\mu} + (L_\delta + c_2) \int f d\tilde{\mu} \\ &\leq (\delta + c_1) C_1 \int |\nabla f| d\tilde{\mu} + C_3 \int f d\tilde{\mu},\end{aligned}$$

with

$$C_3 = L_\delta + c_2 + (\delta + c_1) C_2.$$

Finally, using the inequality $(a+b)^\beta \leq a^\beta + b^\beta$ for $a, b \geq 0$ and $\beta \in (0, 1]$ we have

$$\begin{aligned}\int f\left(|\tilde{U}|^\beta + |\nabla \tilde{U}|\right) d\tilde{\mu} &\leq \int f\left(|U + W|^\beta + |\nabla U + \nabla W|\right) d\tilde{\mu} \\ &\leq \int f(|U|^\beta + |\nabla U|) d\tilde{\mu} + \int f\left(|W|^\beta + |\nabla W|\right) d\tilde{\mu} \\ &\leq C_{\tilde{U}} \int |\nabla f| d\tilde{\mu} + D_{\tilde{U}} \int f d\tilde{\mu},\end{aligned}$$

with

$$C_{\tilde{U}} = \frac{C_U}{1-\delta C_U}(1+\delta+c_1), \quad D_{\tilde{U}} = \frac{L_\delta C_U + D_U}{1-\delta C_U}(1+\delta+c_1) + L_\delta + c_2.$$

□

When $U(x) = d(x)^p/p$ and $\beta = (p-1)/p$, the assumptions (3.6) on W become

$$|\nabla W| \leq 2\delta d^{p-1} + L_\delta, \quad |W|^{(p-1)/p} \leq 2c_1 d^{p-1} + c_2.$$

Examples of functions satisfying these assumptions are bounded functions with bounded derivative (in which case we may take $\delta = c_1 = 0$) as well as polynomials in d of order less than p, whereas if W is a polynomial of order p we need the leading coefficient to be small enough. The U-bound inequality is also stable under perturbation by a function of bounded oscillation. We formulate the above remarks as a corollary, which will be particularly useful in Chapter 4.

Corollary 3.4 ([HZ10])**.** *Let $\mu(dx) = Z^{-1}\mathrm{e}^{-V-W-d^p/p}dx$ be a probability measure, where V is a function such that*

$$\operatorname{osc}(V) = \sup V - \inf V < \infty$$

and W satisfies

$$|\nabla W| \leq \delta d^{p-1} + L_\delta,$$

for some constants $\delta < 1$ and $L_\delta \geq 0$. Then there exist constants C, D such that

$$\int f d^{p-1} d\mu \leq C \int |\nabla f| d\mu + D \int f d\mu$$

for all nonnegative locally Lipschitz f.

Proof. For the case where $V = 0$ we apply (3.2) to the function $f\mathrm{e}^{-W}$ to get

$$\int f\mathrm{e}^{-W} d^{p-1}\mathrm{e}^{-d^p/p}\frac{dx}{Z} \le \int |\nabla f|\mathrm{e}^{-W-d^p/p}\frac{dx}{Z} + \int f|\nabla W|\mathrm{e}^{-W-d^p/p}\frac{dx}{Z} + \left(2^{p-1}+K_d+1\right)\int f\mathrm{e}^{-W-d^p/p}\frac{dx}{Z},$$

with $Z = \int \mathrm{e}^{-d^p/p}dx$. We can now argue as above in order to conclude that

$$\int f d^{p-1}\mathrm{e}^{-W-d^p/p}\frac{dx}{Z} \le C\int |\nabla f|\mathrm{e}^{-W-d^p/p}\frac{dx}{Z} + D'\int f\mathrm{e}^{-W-d^p/p}\frac{dx}{Z},$$

with

$$C = \frac{1}{1-\delta}, \quad D' = \frac{2^{p-1}+K_d+1+L_\delta}{1-\delta}.$$

Then we can multiply through by $Z/\int \mathrm{e}^{-W-d^p/p}dx$ to get the result. When $V \neq 0$, we have

$$\begin{aligned}\int f d^{p-1}\frac{\mathrm{e}^{-V-W-d^p/p}}{\int \mathrm{e}^{-V-W-d^p/p}dx}dx &\le \mathrm{e}^{\mathrm{osc}(V)}\int f d^{p-1}\frac{\mathrm{e}^{-W-d^p/p}}{\int \mathrm{e}^{-W-d^p/p}}dx \\ &\le \mathrm{e}^{\mathrm{osc}(V)}\left(C\int |\nabla f|\frac{\mathrm{e}^{-W-d^p/p}}{\int \mathrm{e}^{-W-d^p/p}}dx + D'\int f\frac{\mathrm{e}^{-W-d^p/p}}{\int \mathrm{e}^{-W-d^p/p}}dx\right) \\ &\le \mathrm{e}^{2\mathrm{osc}(V)}\left(C\int |\nabla f|\frac{\mathrm{e}^{-V-W-d^p/p}}{\int \mathrm{e}^{-V-W-d^p/p}}dx + D'\int f\frac{\mathrm{e}^{-V-W-d^p/p}}{\int \mathrm{e}^{-V-W-d^p/p}}dx\right).\end{aligned}$$

□

The inequality (3.1) has a number of interesting consequences, including isoperimetric as well as entropy inequalities. As we will see in the following sections, it turns out that the exponent of d on the left-hand-side reflects the correct behaviour of the tails of the measure.

As a concluding remark for this section, we would like to mention a connection between U-bounds and a method which has been recently developed for proving functional inequalities under a Lyapunov condition [BBCG08] (see also [CG10, CGW10, CG06, CGWW09, CGGR10] and the references therein for further developments). More specifically, consider the metric d on $\mathbb{R}^n$ and the measure $\mu(dx) = Z^{-1}\mathrm{e}^{-U}dx$. The corresponding operator is $L = \Delta - (\nabla U)\cdot\nabla$ (where the Δ and ∇ are to be understood appropriately, depending on the metric). A *Lyapunov function* for L is a smooth function $W \geq 1$ for which there exist constants $\theta > 0$ and $b, r \geq 0$ such that

$$LW \leq -\theta W + b\chi_{B_r}.$$

In some cases, the U-bound inequality may be seen as an integrated version of this Lyapunov condition. For example, when $U = d^2/2$, if we know that

$$|\nabla d| = 1 \text{ and } \Delta d \leq \frac{K}{d}$$

(in the sense of distributions), we can then compute

$$L\left(1 + \frac{d^2}{2}\right) = |\nabla d|^2 + d\Delta d - d^2|\nabla d|^2 \leq 3 + K - 2\left(\frac{d^2}{2} + 1\right),$$

which is a Lyapunov-type condition with $W = 1 + d^2/2$. Therefore, if we multiply by f^2, where f is a smooth function, and integrate with respect to $\mu(dx)$, we obtain

$$\int f^2 d^2 d\mu \leq -\int f^2 L\left(1 + \frac{d^2}{2}\right) d\mu + (K+1)\int f^2 d\mu.$$

After an integration by parts and an application of Young's inequality, this implies that for all $\varepsilon > 0$,

$$\int f^2 d^2 d\mu \leq \frac{1}{\varepsilon}\int |\nabla f|^2 d\mu + \varepsilon \int f^2 d^2 d\mu + (K+1)\int f^2 d\mu,$$

so that choosing $\varepsilon < 1$ and rearranging, we arrive at the U-bound

$$\int f^2 d^2 d\mu \leq \frac{1}{\varepsilon(1-\varepsilon)}\int |\nabla f|^2 d\mu + \frac{K+1}{1-\varepsilon}\int f^2 d\mu$$

(which is (3.2) with $q = 2$). It is also worth mentioning that both the U-bound inequality and the Lyapunov condition are related to the following assumption of

[KS85] on the potential of the measure

$$\min\left(U, -\Delta U + \frac{1}{2}|\nabla U|^2\right) \geq c, \tag{3.7}$$

for some constant $c > 0$. Indeed, for a potential U growing to infinity as $d(x) \to \infty$, the above is satisfied if

$$-\Delta U + \frac{1}{2}|\nabla U|^2 \geq cU \geq cU(r), \tag{3.8}$$

outside a ball B_r for some $r > 0$. In this case, as explained in [BBCG08], it is possible to find a Lyapunov function. Moreover, arguing as in the proof of Theorem 3.1, we may start from the expression

$$\nabla(f\mathrm{e}^{-U}) = (\nabla f)\mathrm{e}^{-U} - f(\nabla U)\mathrm{e}^{-U},$$

take the inner product with ∇U and integrate (with respect to dx), to obtain

$$\int f\left(-\Delta U + |\nabla U|^2\right) d\mu \leq \int |\nabla f||\nabla U| d\mu,$$

after an integration by parts and an application of the Cauchy-Schwarz inequality. Applying this to f^2 and using Young's inequality, we obtain that for all $\varepsilon > 0$,

$$\int f^2\left(-\Delta U + (1-\varepsilon)|\nabla U|^2\right) d\mu \leq \frac{1}{\varepsilon}\int |\nabla f|^2 d\mu.$$

Therefore, under the assumption (3.8), choosing $\varepsilon = 1/2$ we can conclude that

$$\int f^2 U d\mu \leq 2\int |\nabla f|^2 d\mu,$$

which is a slightly weaker inequality than (1.8) with $q = 2$. The fact that all the above assumptions are related is not surprising, since the motivation for them comes from the need to control the terms that appear after integrating by parts.

3.2 Cheeger inequality

In what follows, we continue to work with the measure $\mu(dx) = Z^{-1}\mathrm{e}^{-U}dx$. At this point, we need to introduce the following assumption on the function U. Let $\beta \in (0,1]$. We assume that for every $M \geq 0$, there exists a constant r depending only on U, M, β, such that

$$\{|U|^{\beta} \leq M\} \subset \{d < r\}. \tag{3.9}$$

We will eventually make use of this assumption with β being the constant appearing in (3.1). In other words, we need that

$$\lim_{d(x)\to\infty} U(x) = \infty.$$

Lemma 3.5. *Suppose that the measure $\mu(dx) = Z^{-1}\mathrm{e}^{-U}dx$ satisfies the U-bound (3.1) with some $\beta \in (0,1]$, and assume that U satisfies (3.9). Then, μ satisfies the following Cheeger inequality*

$$\int |f - \mu(f)|d\mu \leq C_{Che} \int |\nabla f| d\mu, \tag{3.10}$$

for all nonnegative locally Lipschitz f, with some constant $C_{Che} > 0$ independent of f.

Proof. By using an approximation argument, we may assume that f is smooth and bounded. For all $\alpha \in \mathbb{R}$, by the triangle inequality,

$$\int |f - \mu(f)|d\mu \leq \int |f - \alpha|d\mu + |\alpha - \mu(f)| \leq 2\int |f - \alpha|d\mu. \tag{3.11}$$

Let $M > 0$ to be chosen later. By our assumption, there exists $r > 0$ such that, if $|U(x)|^{\beta} \leq M$ then $x \in B_r$, where B_r is the ball of radius r centred at 0. We split the range of integration as follows

$$\int |f - \alpha|d\mu = \int |f - \alpha|\chi_{\{|U(d)|^{\beta} < M\}}d\mu + \int |f - \alpha|\chi_{\{|U(d)|^{\beta} \geq M\}}$$

and treat each integral separately. If we choose

$$\alpha = \bar{f} = \int_{B_r} f \frac{dx}{|B_r|},$$

then the first integral can be estimated by using the Cheeger inequality in the ball (2.11) as follows

$$\begin{aligned}\int |f-\alpha|\chi_{\{|U|^\beta<M\}} d\mu &\leq \int_{B_r} |f-\bar{f}| d\mu \\ &\leq \frac{|B_r|}{Z} \sup_{B_r} \mathrm{e}^{-U} \int_{B_r} |f-\bar{f}| \frac{dx}{|B_r|} \\ &\leq \frac{|B_r|}{Z} \sup_{B_r} \mathrm{e}^{-U} m_r \int_{B_r} |\nabla f| \frac{dx}{|B_r|} \\ &\leq m_r \frac{\sup_{B_r} \mathrm{e}^{-U}}{\inf_{B_r} \mathrm{e}^{-U}} \int |\nabla f| d\mu.\end{aligned}$$

For the second integral, we have

$$\begin{aligned}\int |f-\alpha|\chi_{\{|U|^\beta\geq M\}} &\leq \frac{1}{M}\int |f-\bar{f}||U|^\beta d\mu \\ &\leq \frac{C_U}{M}\int |\nabla f| d\mu + \frac{D_U}{M}\int |f-\bar{f}| d\mu,\end{aligned}$$

by (3.1). Combining these facts, we conclude that

$$\int |f-\bar{f}| d\mu \leq \left(m_r \frac{\sup_{B_r} \mathrm{e}^{-U}}{\inf_{B_r} \mathrm{e}^{-U}} + \frac{C_U}{M} \right) \int |\nabla f| d\mu + \frac{D_U}{M} \int |f-\bar{f}| d\mu.$$

Finally, choosing $M = 2D_U$ and rearranging we arrive at

$$\int |f-\bar{f}| d\mu \leq \frac{C_{Che}}{2} \int |\nabla f| d\mu,$$

with

$$C_{Che} = 4\left(m_r \frac{\sup_{B_r} \mathrm{e}^{-U}}{\inf_{B_r} \mathrm{e}^{-U}} + \frac{C_U}{2D_U} \right),$$

which, together with (3.11), gives the desired result. □

The assumption (3.9) is evidently true when $U = d^p/p$ and $\beta = (p-1)/p$. Therefore, we have shown the following.

Corollary 3.6. *For $p > 1$, the measure $\mu(dx) = Z^{-1}\mathrm{e}^{-d^p/p}$ satisfies*

$$\int |f - \mu(f)| d\mu \leq C_{Che} \int |\nabla f| d\mu$$

for all locally Lipschitz f with some constant C_{Che} independent of f.

By Proposition 2.17, a further consequence of this is that the measure above satisfies a q- Poincaré inequality for every $q \geq 1$, i.e.

$$\int |f - \mu(f)|^q \leq C_q \int |\nabla f|^q d\mu,$$

for all nonnegative locally Lipschitz f with some constant C_q independent of f.

At the level of sets, it is well-known that the inequality (3.10) is equivalent to the existence of a constant C such that for all measurable A

$$\min(\mu(A), 1 - \mu(A)) \leq C\mu^+(A). \tag{3.12}$$

As we will see in the next sections, the measure $Z^{-1}\mathrm{e}^{-d^p/p}$ actually satisfies a much stronger property, namely the q-isoperimetric inequality.

3.3 Ledoux Inequality

We will say that a measure μ satisfies a β-Ledoux inequality, where $\beta \in (0, 1]$, if there exist constants C_L, D_L such that

$$\int f \log_+^\beta \left(\frac{f}{\mu(f)} \right) d\mu \leq C_L \int |\nabla f| d\mu + D_L \int f d\mu,$$

for all nonnegative locally Lipschitz functions f. This is an L^1 F-Sobolev inequality, with $F(t) = t \log_+^\beta t$. Starting from the Gaussian isoperimetric inequality, the inequality above was derived in [Led88] for the Gaussian measure on $\mathbb{R}^n$ with $\beta = 1/2$ (in an equivalent formulation in the language of Orlicz-spaces).

Theorem 3.7. *Let $\mu(dx) = Z^{-1}e^{-U}dx$. If the U-bound (3.1) holds, then, for all nonnegative locally Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ with $\mu(f) = 1$,*

$$\int f(\log_+ f)^\beta d\mu \leq C_L \int |\nabla f| d\mu + \ D_L, \tag{3.13}$$

with some constants $C_L, D_L \geq 0$ independent of f.

Proof. We may assume that $U \geq 0$; if not, we can replace U by $U - \inf U$. Suppose that f is a nonnegative smooth function with $\mu(f) = 1$. We consider the quantity $\int f \log_+^\beta (f\mathrm{e}^{-U}) d\mu$ and show that

$$\int f \log_+^\beta f d\mu - \int_{\{f \geq 1\}} f U^\beta d\mu \leq \int f \log_+^\beta (f\mathrm{e}^{-U}) d\mu \tag{3.14}$$

and

$$\int f \log_+^\beta (f\mathrm{e}^{-U}) d\mu \leq k_1 \int |\nabla f| d\mu + k_2 \int f |\nabla U| d\mu + k_3, \tag{3.15}$$

with some constants k_1, k_2, k_3. Once this has been established, we will be able to conclude that

$$\int f \log_+^\beta f d\mu \leq k_1 \int |\nabla f| d\mu + \max(k_2, 1) \int f(U^\beta + |\nabla U|) d\mu + k_3$$

which, combined with the U-bound (3.1), will give the result. It is not difficult to see (3.14). Indeed, using the inequality $|a - b|^\beta \geq a^\beta - b^\beta$, which holds for all $a, b \geq 0$

since $\beta \in (0,1]$, we have

$$\begin{aligned}
\int f \log_+^\beta \left(f \mathrm{e}^{-U}\right) d\mu &= \int_{\{f \geq \mathrm{e}^U\}} f(\log(f) - U)^\beta d\mu \\
&\geq \int_{\{f \geq \mathrm{e}^U\}} f \log^\beta f d\mu - \int_{\{f \geq \mathrm{e}^U\}} f U^\beta d\mu \\
&= \int f \log_+^\beta f d\mu - \int_{\{f \geq \mathrm{e}^U\}} f U^\beta d\mu \\
&\quad - \int_{\{1 \leq f \leq \mathrm{e}^U\}} f \log_+^\beta f d\mu \\
&\geq \int f \log_+^\beta f d\mu - \int_{\{f \geq 1\}} f U^\beta d\mu.
\end{aligned}$$

For (3.15), recall that by the classical Sobolev inequality (2.10), there exist constants $\alpha_{\mathrm{cs}} > 1$, $A_{\mathrm{CS}}, B_{\mathrm{CS}}$ such that

$$\left(\int |f|^{\alpha_{\mathrm{cs}}} dx\right)^{\frac{1}{\alpha_{\mathrm{cs}}}} \leq A_{\mathrm{CS}} \int |\nabla f| dx + B_{\mathrm{CS}} \int |f| dx. \tag{3.16}$$

We use Jensen's inequality for the probability measure

$$\nu_f(dx) = \frac{f \chi_{\{f \geq \mathrm{e}^U\}} \mathrm{e}^{-U}}{Z_f} dx = \frac{f \chi_{\{f \geq \mathrm{e}^U\}} \mathrm{e}^{-U}}{\int_{\{f \geq \mathrm{e}^U\}} f \mathrm{e}^{-U} dx} dx$$

and the positive concave function $t \mapsto \log^{\beta} t$, $t \geq 1$, to get

$$\begin{aligned}
\int f \log_+^{\beta}(f\mathrm{e}^{-U})d\mu &= \frac{Z_f}{Z}\int \log^{\beta}(f\mathrm{e}^{-U})d\nu_f \\
&= \frac{Z_f}{(\alpha_{\mathrm{cs}}-1)^{\beta}Z}\int \log^{\beta}\left(f^{(\alpha_{\mathrm{cs}}-1)}\mathrm{e}^{-(\alpha_{\mathrm{cs}}-1)U}\right)d\nu_f \\
&\leq \frac{Z_f}{(\alpha_{\mathrm{cs}}-1)^{\beta}Z}\log^{\beta}\left(\int_{\{f\geq \mathrm{e}^U\}} f^{\alpha_{\mathrm{cs}}}\mathrm{e}^{-\alpha_{\mathrm{cs}}U}\frac{dx}{Z_f}\right) \\
&= \frac{\alpha_{\mathrm{cs}}^{\beta}Z_f}{(\alpha_{\mathrm{cs}}-1)^{\beta}Z}\log^{\beta}\left(\left(\int_{\{f\geq \mathrm{e}^U\}} f^{\alpha_{\mathrm{cs}}}\mathrm{e}^{-\alpha_{\mathrm{cs}}U}\frac{dx}{Z_f}\right)^{\frac{1}{\alpha_{\mathrm{cs}}}}\right) \\
&\leq \frac{\alpha_{\mathrm{cs}}^{\beta}Z_f^{1-\frac{1}{\alpha_{\mathrm{cs}}}}}{(\alpha_{\mathrm{cs}}-1)^{\beta}Z}\left(\int_{\{f\geq \mathrm{e}^U\}} f^{\alpha_{\mathrm{cs}}}\mathrm{e}^{-\alpha_{\mathrm{cs}}U}dx\right)^{\frac{1}{\alpha_{\mathrm{cs}}}} \\
&\leq \frac{\alpha_{\mathrm{cs}}^{\beta}Z_f^{1-\frac{1}{\alpha_{\mathrm{cs}}}}}{(\alpha_{\mathrm{cs}}-1)^{\beta}Z}\left(\int f^{\alpha_{\mathrm{cs}}}\mathrm{e}^{-\alpha_{\mathrm{cs}}U}dx\right)^{\frac{1}{\alpha_{\mathrm{cs}}}},
\end{aligned}$$

where in the last but one step we used the elementary inequality $(\log x)^{\beta} \leq x$ for $x \geq 1$. Applying the classical Sobolev inequality to the function $f\mathrm{e}^{-U}$ and using that $|\nabla(f\mathrm{e}^{-U})| \leq (|\nabla f| + f|\nabla U|)\mathrm{e}^{-U}$, we arrive at

$$\begin{aligned}
\int f(\log_+(f\mathrm{e}^{-U}))^{\beta}d\mu &\leq \frac{\alpha_{\mathrm{cs}}^{\beta}Z_f^{1-\frac{1}{\alpha_{\mathrm{cs}}}}}{(\alpha_{\mathrm{cs}}-1)^{\beta}Z}\left(A_{CS}\int(|\nabla f| + f|\nabla U|)\mathrm{e}^{-U}dx + B_{CS}Z\right) \\
&= \frac{\alpha_{\mathrm{cs}}^{\beta}}{(\alpha_{\mathrm{cs}}-1)^{\beta}}Z_f^{1-\frac{1}{\alpha_{\mathrm{cs}}}}\left(A_{CS}\int(|\nabla f| + f|\nabla U|)d\mu \; + B_{CS}\right).
\end{aligned}$$

Finally, we observe that

$$1 = \int f d\mu \geq \frac{Z_f}{Z} = \frac{\int f\chi_{\{f\geq \mathrm{e}^U\}}\mathrm{e}^{-U}dx}{Z}$$

and thus

$$\int f(\log_+(f\mathrm{e}^{-U}))^{\beta}d\mu \leq \frac{\alpha_{\mathrm{cs}}^{\beta}}{(1-\alpha_{\mathrm{cs}})^{\beta}}Z^{1-\frac{1}{\alpha_{\mathrm{cs}}}}\left(A_{CS}\int(|\nabla f| + f|\nabla U|)d\mu + \; B_{CS}\right).$$

The claim now follows by (3.1) with

$$C_L = k_\alpha A_{\mathrm{CS}} + \max\{1, k_\alpha A_{\mathrm{CS}}\} C_U, \quad D_L = k_\alpha B_{\mathrm{CS}} + \max\{1, k_\alpha A_{\mathrm{CS}}\} D_U,$$

where $k_\alpha = \frac{\alpha_{\mathrm{cs}}^\beta}{(1-\alpha_{\mathrm{cs}})^\beta} Z^{1-\frac{1}{\alpha_{\mathrm{cs}}}}$. □

In particular, the assumptions of the theorem are satisfied for the measure $\mu(dx) = Z^{-1}\mathrm{e}^{-d^p/p}dx$ with $\beta = 1/q = (p-1)/p$.

Remark 3.8. *If the result holds for a measure μ defined with a potential U and $\tilde{\mu}$ is the perturbed measure defined in Proposition 3.3, then the result continues to hold for $\tilde{\mu}$.*

The Ledoux inequality (3.13) already gives us some isoperimetric information. Approximating the indicator function of a set A of measure $a \leq 1/2$, it leads to

$$a\,(-\log a)^\beta \leq C_L \mu^+(A) + D_L a.$$

Therefore, if a is sufficiently small, say

$$a \leq a_0 := \mathrm{e}^{-(2D_L)^{1/\beta}} < \frac{1}{2},$$

then

$$a(-\log a)^\beta \leq 2C_L \mu^+(A), \tag{3.17}$$

i.e. $\mathcal{I}_\mu(a) \geq a(-\log a)^\beta / 2C_L$ for $a \leq a_0$. By an analogous argument, considering now an approximation of $1 - \chi_A$, we can get the result for the range $a \in [1-a_0, 1]$, so that

$$\mathcal{I}_\mu(a) \geq \hat{a}(-\log \hat{a})^\beta / 2C_L, \tag{3.18}$$

with $\hat{a} = \min(a, 1-a)$, for all $a \in [0, a_0] \cup [1-a_0, 1]$. It remains to prove the inequality for sets of measure a near $1/2$. Suppose we know that the isoperimetric profile is concave. In the Euclidean setting as well as for Riemannian manifolds with

Ricci curvature bounded from below, this is indeed the case for measures which are log-concave (see [MJ00, Bay04, Bob96a, Mil08] and references therein). Then, the inequality (3.18) for small a immediately gives us the inequality on the whole range $a \in [0,1]$, since we may extend the lower bound for $\mathcal{I}_\mu$ by letting it be constant over the interval $(a_0, 1-a_0)$. More precisely, for all $a \in [0,1]$,

$$\mathcal{I}_\mu(a) \geq \Psi(a) \geq C\hat{a}(-\log \hat{a})^\beta,$$

with some constant $C > 0$, where

$$\Psi(a) = \begin{cases} \frac{a(-\log a)^\beta}{2C_L}, & \text{for } a \leq a_0, \\ \frac{a_0(-\log a_0)^\beta}{2C_L} = \frac{D_L}{C_L} \mathrm{e}^{-(2D_L)^{1/\beta}}, & \text{for } a \in (a_0, 1-a_0), \\ \frac{(1-a)(-\log(1-a))^\beta}{2C_L}, & \text{for } a \geq 1-a_0. \end{cases}$$

We are not aware of a result about the concavity of the isoperimetric function for the types of measures we are considering on H-type groups. Another approach to reach the isoperimetric inequality, which does not rely on the concavity of $\mathcal{I}_\mu$, would be to use the inequality (3.17) for small a combined with the Cheeger inequality (3.10) for large a. We have already seen that if a measure μ satisfies the U- bound (3.1) together with the assumption (3.9), it then satisfies the Cheeger inequality. Recall that this can be stated equivalently as

$$\min(a, 1-a) \leq C\mu^+(A),$$

for all A with $\mu(A) = a$. Let now $a \in (a_0, 1/2)$. Since $(-\log a)^\beta < 2D_L$ we have

$$a(-\log a)^\beta \leq 2D_L a \leq 2D_L C\mu^+(A),$$

so we conclude that

$$a(-\log a)^\beta \leq \max(2D_L C, 2C_L)\mu^+(A), \tag{3.19}$$

for all $a = \mu(A) \in [0, 1/2]$. A similar argument for $a \in (1/2, 1 - a_0)$ shows that

$$\hat{a}(-\log \hat{a})^{1/\beta} \leq \max(2D_L C, 2C_L)\mu^+(A),$$

for all $a \in [0, 1]$. A more rigorous proof of this result, starting from a Φ-Entropy inequality, will be given in Section 3.4.

Conversely, we may start from (3.19) and use an idea of [Led88] to go back to the functional form (3.13).

Theorem 3.9. *If the measure μ satisfies*

$$\hat{a}(-\log \hat{a})^{\beta} \leq C\mu^+(A)$$

for some constant C and all sets A of measure $a = \mu(A) \in [0, 1/2)$, then there exist constants C_L, D_L depending only on C such that

$$\int f(\log_+ f)^{\beta} d\mu \leq C_L \int |\nabla f| d\mu + D_L$$

for all positive locally Lipschitz functions f with $\mu(f) = 1$.

Proof. Let f be nonnegative, with $\mu(f) = 1$. The coarea inequality (2.20) together with our assumption imply

$$\int |\nabla f| d\mu \geq \int_0^{\infty} \mu^+(\{f > s\}) ds \geq \frac{1}{C} \int_2^{\infty} \mu(\{f > s\})(-\log \mu(\{f > s\}))^{\beta} ds.$$

By Markov's inequality,

$$\mu(\{f > s\}) \leq \mu(\{f > 2\}) \leq 1/2,$$

for all $s \geq 2$. Therefore,

$$\begin{aligned}\int |\nabla f| d\mu \geq \frac{1}{C} \int_0^\infty \mu(\{f > s\}) \left(-\log \mu(\{f > s\})\right)^\beta ds \\ - \frac{1}{C} \int_0^2 \mu(\{f > s\}) \left(-\log \mu(\{f > s\})\right)^\beta ds \\ \geq \frac{1}{C} \int_0^\infty \mu(\{f > s\}) \left(-\log \mu(\{f > s\})\right)^\beta ds - \frac{2}{C} M,\end{aligned}$$

where $M = \sup_{t \in [0,1]} t \log^\beta \frac{1}{t}$. Next, again by Markov's inequality, $\mu(\{f > s\}) \leq 1/s$. Therefore, when $s \geq 1$ we have $-\log \mu(\{f > s\}) \geq \log s$ and we always have $-\log \mu(\{f > s\}) \geq 0$. Therefore, $-\log \mu(\{f > s\}) \geq \log_+ s$, which implies

$$\int |\nabla f| d\mu \ \geq \frac{1}{C} \int_0^\infty \left(\log_+ s\right)^\beta \mu(\{f > s\}) ds - \frac{2M}{C} \geq C_L \int f \left(\log_+ f\right)^\beta d\mu - D_L,$$

with $C_L = 1/C$ and $D_L = (2M/C) + 1$. To see the last inequality, let $F(s) = \int_0^s (\log_+ t)^\beta dt$ and $H(s) = s(\log_+ s)^\beta - s$. Then $F(s) \geq 0 \geq H(s)$ on $[0, e]$ and when $s \geq e$, $F'(s) = (\log s)^\beta$ and $H'(s) = (\log s)^\beta + \beta(\log s)^{\beta - 1} - 1$. Therefore, since $\log s \geq 1$ and $\beta \in (0, 1]$, $F' \geq H'$ for $s \geq e$ from which it follows that $F \geq H$ on $[0, \infty)$. Therefore,

$$\begin{aligned}\int_0^\infty (\log_+ s)^\beta \mu(\{f > s\}) ds &= \int_0^\infty F'(s) \mu(\{f > s\}) ds \\ &= \int F(f) d\mu \\ &\geq \int H(f) d\mu \\ &= \int f (\log_+ f)^\beta d\mu - 1.\end{aligned}$$

□

Finally, let us remark that the inequality

$$\hat{a}(-\log \hat{a})^\beta \leq C \mu^+(A)$$

directly implies

$$\min(a, 1-a) \leq \frac{C}{(-\log 2)^\beta} \mu^+(A),$$

or, equivalently,

$$\int |f - \mu(f)| d\mu \leq C_{Che} \int |\nabla f| d\mu$$

for all locally Lipschitz f, with some constant C_{Che}.

We conclude this section with a theorem which essentially summarises the above results for our model measures.

Theorem 3.10. *Let $\mu(dx) = Z^{-1} e^{-d^p/p} dx$. Then, there exists constants C_L, D_L and C_{Che} such that*

$$\int f (\log_+ f)^{1/q} d\mu \leq C_L \int |\nabla f| d\mu + D_L$$

and

$$\int |f - \mu(f)| d\mu \leq C_{Che} \int |\nabla f| d\mu,$$

for all nonnegative locally Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$, where $q = p/(p-1)$. Equivalently, there exists a constant C_{iso} such that for all measurable sets A of measure a

$$\hat{a}(-\log \hat{a})^{1/q} \leq C_{iso} \mu^+(A),$$

where $\hat{a} = \min(a, 1-a)$, i.e. $\mathcal{I}_\mu(a) \geq \hat{a}(-\log \hat{a})^{1/q}$.

3.4 Φ-Entropy inequality

As we will see, the inequality (3.13) can be formally strengthened by controlling the defective term D_L. To do this, we pass to an additive form of the inequality (Φ-entropy) and then use a Rothaus-type argument together with the Cheeger inequality (3.10). Throughout this section, we denote $\Phi(t) = t \log^\beta(1+t)$. Recall that the Φ-Entropy of a function is defined by

$$\mathrm{Ent}_\mu^\Phi(|f|) := \int \Phi(|f|) d\mu - \Phi(\mu|f|).$$

It is a nonnegative quantity, since Φ is a convex function. The following theorem is the main result of this section.

Theorem 3.11. *Suppose that $\mu(dx) = Z^{-1}\mathrm{e}^{-U}dx$ satisfies the U-bound (3.1) with some $\beta \in (0,1]$ and suppose that U satisfies (3.9). Then for all locally Lipschitz f*

$$\mathrm{Ent}_\mu^\Phi(|f|) \leq C_\Phi \int |\nabla f| d\mu, \tag{3.20}$$

with some constant C_Φ independent of f.

For the proof, we will need a monotonicity property of the entropy, which is stated in the following lemma.

Lemma 3.12. *For any probability measure μ and for any functions f, g with $0 \leq g \leq f$, $\mu(f) < \infty$, one has*

$$\mathrm{Ent}_\mu^\Phi(g) \leq \int f \log_+ \left(\frac{f}{\mu f}\right)^\beta d\mu \ + D_{\mathrm{re}}\mu(f),$$

for some constant D_{re} independent of f and g.

Proof. The proof follows an idea of [FRZ07], where a generalised relative entropy inequality was proved. We start by noticing that, since $x^\beta - y^\beta \leq |x-y|^\beta$,

$$\begin{aligned}
\mathrm{Ent}_\mu^\Phi(g) &= \int \left(g\left(\log^\beta(1+g) - \log^\beta(1+\mu(g))\right)\right) d\mu \\
&\leq \int g \log^\beta\left(1 + \frac{g}{\mu(g)}\right) d\mu \\
&\leq \int f \log^\beta\left(1 + \frac{g}{\mu(g)}\right) d\mu,
\end{aligned} \tag{3.21}$$

where in the last line we used our assumption that $g \leq f$. The function $F(x) = \log^\beta(1+x)$, defined for $x \in [0,\infty)$, is increasing and concave, with $F(0) = 0$. Let θ be a constant such that $xF'(x) \leq \theta$ for all $x \geq 0$. We claim that $xF(y) \leq xF(x) + \theta y$ for all $x, y \geq 0$. This is clear when $y \leq x$, while when $x \leq y$ we have

$$x(F(y) - F(x)) = x\left(\frac{F(y)-F(x)}{y-x}\right)(y-x) \leq xF'(x)y \leq \theta y.$$

Choosing $x = f/\mu(f)$ and $y = g/\mu(g)$ and integrating both sides with respect to the measure μ we arrive at

$$\int f \log^\beta \left(1 + \frac{g}{\mu(g)}\right) d\mu \leq \int f \log^\beta \left(1 + \frac{f}{\mu(f)}\right) d\mu + \theta\mu(f).$$

Thus, by (3.21)

$$\mathrm{Ent}_\mu^\Phi(g) \leq \int f \log^\beta \left(1 + \frac{f}{\mu(f)}\right) d\mu + \theta\mu(f). \tag{3.22}$$

Now

$$\begin{aligned}\int f \log^\beta \left(1 + \frac{f}{\mu(f)}\right) d\mu &= \int_{\{f \leq \mu(f)\}} f \log^\beta \left(1 + \frac{f}{\mu(f)}\right) d\mu \\ &\quad + \int_{\{f \geq \mu(f)\}} f \log^\beta \left(1 + \frac{f}{\mu(f)}\right) d\mu \\ &\leq (\log 2)^\beta \mu(f) + \int_{\{f \geq \mu(f)\}} f \log^\beta \left(\frac{2f}{\mu(f)}\right) d\mu \\ &\leq 2(\log 2)^\beta \mu(f) + \int f \log_+^\beta \left(\frac{f}{\mu(f)}\right) d\mu\end{aligned}$$

using once again the inequality $(x+y)^\beta \leq x^\beta + y^\beta$ for $x, y \geq 0$ and $\beta \in (0, 1]$. The inequality is proved with $D_{re} = 2(\log 2)^\beta + \theta$. □

We are now in position to prove the theorem.

Proof of Theorem 3.11. By Lemma A.1 of the appendix of [ŁZ07], there exist constants $\tilde{a}$ and $\tilde{b}$ such that

$$\mathrm{Ent}_\mu^\Phi(f^2) \leq \tilde{a}\mathrm{Ent}_\mu^\Phi\left((f - \mu f)^2\right) + \tilde{b}\mathrm{Var}_\mu(f).$$

Thus, for any $t \in \mathbb{R}$, we have that

$$\begin{aligned}\mathrm{Ent}_\mu^\Phi|f+t| &= \mathrm{Ent}_\mu^\Phi\left[\left(|f+t|^{\frac{1}{2}}\right)^2\right] \\ &\leq \tilde{a}\mathrm{Ent}_\mu^\Phi\left[\left(|f+t|^{\frac{1}{2}} - \mu|f+t|^{\frac{1}{2}}\right)^2\right] + \tilde{b}\mathrm{Var}_\mu\left(|f+t|^{\frac{1}{2}}\right).\end{aligned} \tag{3.23}$$

Let $G = \left(|f+t|^{\frac{1}{2}} - \int |f+t|^{\frac{1}{2}} d\mu\right)^2$. Note that we can write

$$\begin{aligned} G &= \left(\int |f(\omega)+t|^{\frac{1}{2}} - |f(\tilde{\omega})+t|^{\frac{1}{2}} d\mu(\tilde{\omega})\right)^2 \\ &\leq \int \left(|f(\omega)+t|^{\frac{1}{2}} - |f(\tilde{\omega})+t|^{\frac{1}{2}}\right)^2 d\mu(\tilde{\omega}) \\ &\leq \int |f(\omega) - f(\tilde{\omega})| d\mu(\tilde{\omega}) \\ &\leq |f| + \int |f| d\mu, \end{aligned}$$

using the elementary inequality $\left||x+t|^{\frac{1}{2}} - |y+t|^{\frac{1}{2}}\right| \leq |x-y|^{\frac{1}{2}}$ in the last but one step. Hence, since $\mu(G) = \mathrm{Var}_\mu\left(|f+t|^{\frac{1}{2}}\right)$, we have by (3.23) that

$$\mathrm{Ent}_\mu^\Phi |f+t| \leq \tilde{a}\mathrm{Ent}_\mu^\Phi(G) + 2\tilde{b}\int |f| d\mu. \tag{3.24}$$

Since $0 \leq G \leq |f| + \int |f| d\mu$, by Lemma 3.12 and Theorem 3.7, we have

$$\begin{aligned} \mathrm{Ent}_\mu^\Phi(G) &\leq \int \left(|f| + \int |f| d\mu\right) \log_+^\beta \frac{|f| + \int |f| d\mu}{\mu\left(|f| + \int |f| d\mu\right)} d\mu \ + 2D_{re}\int |f| d\mu \\ &\leq C_L \int |\nabla f| d\mu \ + 2(D_L + D_{re}) \int |f| d\mu. \end{aligned} \tag{3.25}$$

Combining (3.24) and (3.25) yields

$$\sup_{t\in\mathbb{R}} \mathrm{Ent}_\mu^\Phi |f+t| \leq \tilde{a} C_L \int |\nabla f| d\mu \ + 2(\tilde{a}(D_L + D_{re}) + \tilde{b}) \int |f| d\mu \ . \tag{3.26}$$

This implies the following bound

$$\mathrm{Ent}_\mu^\Phi |f| \leq \tilde{a} C_L \int |\nabla f| d\mu + 2(\tilde{a}(D_L + D_{re}) + \tilde{b}) \int |f - \mu(f)| d\mu \ . \tag{3.27}$$

Finally we can apply the Cheeger inequality of Lemma 3.5 to the last term on the right-hand side of (3.27) to arrive at

$$\mathrm{Ent}_\mu^\Phi\,(|f|) \leq C_\Phi \int |\nabla f| d\mu \ ,$$

with $C_\Phi = \tilde{a} C_L + 2C_{Che}\left(\tilde{a}(D_L + D_{re}) + \tilde{b}\right)$. ☐

Corollary 3.13. *Consider a probability measure*

$$\mu(dx) = Z^{-1} \mathrm{e}^{-V-W-d^p/p} dx \tag{3.28}$$

on an H- type group, where $p > 1$, $V$is a function such that

$$\mathrm{osc}(V) = \sup V - \inf V < \infty$$

and W satisfies

$$|\nabla W| \leq \delta d^{p-1} + L_\delta,$$

for some constants $\delta < 1$ and $L_\delta \geq 0$. There exists a constant C_Φ such that for all locally Lipschitz f

$$\mathrm{Ent}_\mu^\Phi(|f|) \leq C_\Phi \int |\nabla f| d\mu.$$

Proof. The proof follows from Corollary 3.4 and Theorem 3.11. ☐

The conclusion of the theorem allows us to 'tighten' inequality (3.13), in a sense made precise below. Let us apply the Φ-Entropy inequality to a function f with mean 1. We then obtain the following homogeneous inequality

$$\int f(\log_+(1+f))^\beta d\mu \leq C_\Phi \int |\nabla f| d\mu + (\log 2)^\beta. \tag{3.29}$$

In other words, by enlarging the constant in front of $\int |\nabla f| d\mu$ if necessary, we may assume that $D_L = (\log 2)^\beta$ in (3.13).

This now gives an alternative proof of the isoperimetric inequality (3.19) obtained in the previous section.

Corollary 3.14. *The measure $\mu(dx) = Z^{-1}e^{-d^p/p}dx$ satisfies the isoperimetric inequality*

$$\hat{a}(-\log \hat{a})^{1/q} \leq C\mu^+(A)$$

for some constant C and all sets A of measure $\mu(A) = a$ where $q = p/(p-1)$ and $\hat{a} = \min(a, 1-a)$. In other words, $\mathcal{I}_\mu(a) \geq \hat{a}(-\log \hat{a})^{1/q}/C$ for all $a \in [0,1]$.

Proof. The assumptions of Theorem 3.11 are satisfied for μ with $\beta = 1/q = (p-1)/p$. Therefore, for all nonnegative locally Lipschitz f,

$$\int f \log\left(1 + \frac{f}{\mu(f)}\right)^{1/q} d\mu \leq C \int |\nabla f| d\mu + (\log 2)^{1/q} \int f d\mu. \tag{3.30}$$

Let A be a Borel set with measure $a = \mu(A) \leq 1/2$. By Proposition 2.22 (and the remark following it) we may assume that $\mu(A) = \mu(\bar{A})$ and approximate the indicator function of $\bar{A}$ by a sequence of Lipschitz functions $(f_n)_{n\in\mathbb{N}}$ satisfying $0 \leq f_n \leq 1$ and

$$\limsup_{n\to\infty} \int |\nabla f_n| d\mu \ \leq \mu^+(A).$$

Taking f_n in (3.30), in the limit as $n \to \infty$ we obtain

$$a\left(\log\left(1 + \frac{1}{a}\right)^{1/q} - (\log 2)^{1/q}\right) \leq C\mu^+(A). \tag{3.31}$$

We now observe that for $a \in \left[0, \frac{1}{2}\right]$ we have

$$\eta\left(\log\left(\frac{1}{a}\right)\right)^{1/q} \leq \left(\log\left(1 + \frac{1}{a}\right)\right)^{1/q} - (\log 2)^{1/q}, \tag{3.32}$$

with $\eta = \left(\frac{\log 3}{\log 2}\right)^{1/q} - 1 > 0$. This implies

$$a\left(\log\left(\frac{1}{a}\right)\right)^{1/q} \leq \frac{C}{\eta}\,\mu^+(A), \tag{3.33}$$

for all $a \in \left[0, \frac{1}{2}\right]$.

Now suppose that $a = \mu(A) \in \left(\frac{1}{2}, 1\right]$. For functions $f \in [0,1]$, we can apply (3.30) to $1-f$, which yields

$$\int (1-f)\left(\left(\log\left(1+\frac{1-f}{1-\mu(f)}\right)\right)^{1/q} - (\log 2)^{\beta}\right) d\mu \le C \int |\nabla f| d\mu.$$

If we now take f_n in this inequality (where $(f_n)_{n\in\mathbb{N}}$ is again the Lipschitz approximation of the characteristic function of $\bar{A}$) and pass to the limit as $n \to \infty$, we see that

$$(1-a)\left(\left(\log\left(1+\frac{1}{1-a}\right)\right)^{1/q} - (\log 2)^{1/q}\right) \le C\mu^+(A).$$

Arguing as before, this implies that

$$s\left(\log\left(\frac{1}{s}\right)\right)^{1/q} \le \frac{C}{\eta}\,\mu^+(A)$$

for $s = 1 - a \le 1/2$. □

Corollary 3.15. *The* Φ*-entropy inequality* (3.20) *implies the Cheeger inequality.*

Proof. This can be directly seen at the level of sets. Since (3.20) implies

$$\hat{a}(-\log \hat{a})^{\beta} \le C\mu^+(A)$$

with some constant C, for all sets A with $\mu(A) = a$, we immediately deduce that

$$\hat{a} \le \frac{C}{(-\log 2)^{\beta}}\mu^+(A),$$

which is the Cheeger inequality. □

We have seen that under the compactness assumption (3.9) on U, the U-bound (3.1) implies the Φ-Entropy inequality (3.20). We now show that under a convexity assumption on U, the converse also holds.

Theorem 3.16. *Suppose $\mu(dx) = Z^{-1}e^{-U}dx$ satisfies the Φ-entropy inequality (3.20) with some $\beta \in [1/2, 1]$ and assume that there exist constants a, b such that*

$$|\nabla U| \leq a|U|^{\beta} + b. \tag{3.34}$$

Then there exist constants C_U, D_U such that

$$\int f(|U|^{\beta} + |\nabla U|)d\mu \leq C_U \int |\nabla f| d\mu + D_U \int f d\mu$$

for all nonnegative locally Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$.

Let us observe that, typically (e.g. when $U(d) = d^p$), we have $U'(d(x))/U^{\beta}(d(x)) \to 1$ as $d(x) \to \infty$ so that assumption (3.34) is satisfied. For the proof we will need the following result.

Lemma 3.17. *Let μ be a probability measure. Then*

$$\int fh d\mu \leq s^{-1}\mathrm{Ent}_{\mu}^{\Phi}(f) + s^{-1}\Theta(sh) \tag{3.35}$$

for all $s > 0$ and suitable functions $f, h \geq 0$ such that $\mu(f) = 1$, where

$$\Theta(h) \equiv \left(\theta + (\log 2)^{\beta} + \left(\log \int e^{h^q} d\mu\right)^{\beta}\right)$$

with $\theta = \sup_{x\geq 0} \beta x(\log(1+x))^{\beta-1}/(1+x)$.

Moreover, if μ satisfies (3.20) for some $\beta \in [1/2, 1]$ and $g \geq 0$ is a locally Lipschitz function such that

$$|\nabla g|^{1/\beta} \leq ag + b \tag{3.36}$$

for some constants $a, b \in (0, \infty)$, then $\Theta(s^{\beta}g^{\beta}) < \infty$ for sufficiently small $s > 0$ and

$$\int fg^{\beta}d\mu \leq \frac{c}{s^{\beta}} \int |\nabla f| d\mu + \frac{c}{s^{\beta}}\Theta(s^{\beta}g^{\beta})\mu(f), \tag{3.37}$$

for all locally Lipschitz functions $f \geq 0$.

Proof. Let $f, h \geq 0$, $\mu(f) = 1$, with $s \in (0, \infty)$ and $\beta \equiv \frac{1}{q} \in (0,1)$. Setting $H = \mathrm{e}^{s^q h^q}$, we have

$$\begin{aligned}\int f h d\mu &= s^{-1} \int f \log^\beta H d\mu \\ &\leq s^{-1} \int_{\{H \geq \mu(H)\}} f \log^\beta \left(1 + \frac{H}{\mu(H)}\right) \\ &\quad + s^{-1} \log^\beta \mu(H).\end{aligned}$$

By the generalised relative entropy inequality of [FRZ07], we have

$$\begin{aligned}\int f \log^\beta \left(1 + \frac{H}{\mu(H)}\right) d\mu &\leq \int f \log^\beta (1+f)\, d\mu + \theta \\ &\leq \mathrm{Ent}_\mu^\Phi(f) + \theta + (\log 2)^\beta,\end{aligned}$$

since $\mu(f) = 1$. We therefore get the following bound

$$\int f h d\mu \leq s^{-1} \mathrm{Ent}_\mu^\Phi(f) + s^{-1} \left(\theta + (\log 2)^\beta + \log^\beta \mu(H)\right). \tag{3.38}$$

This ends the proof of the first part of the lemma.
Replacing h by g^β and s by s^β in (3.38), we see that the second part is a consequence of the fact that, for g satisfying (3.36), $\int e^{tg} d\mu < \infty$ (see [HZ10]). □

Proof of Theorem 3.16. We may assume that $f \geq 0$ and $U \geq 0$ (otherwise we can shift it by a constant). We ote that from (3.34), it follows that

$$|\nabla U|^{1/\beta} \leq \tilde{a} U + \tilde{b}.$$

Hence we may apply Lemma 3.17, to see that

$$\int f U^\beta d\mu \leq \frac{c}{s^\beta} \int |\nabla f| d\mu + \frac{c}{s^\beta} \Theta(s^\beta U^\beta) \int f d\mu,$$

with $\Theta(s^\beta U^\beta) < \infty$ for sufficiently small s. □

3.5 Transportation inequalities

In this section we discuss some consequences of the inequalities obtained above in terms of measure transportation. As a preliminary observation, we show that the inequality

$$\int fd\, d\mu \leq C \int |\nabla f| d\mu + D \int f \tag{3.39}$$

directly implies a $\mathbb{T}_1$ transportation inequality (defined in (2.28)). As we already saw in Section 3.1, the measures $\mu(dx) = Z^{-1} \mathrm{e}^{-d^p/p} dx$, as well as the perturbed measures defined in Corollary 3.4, satisfy

$$\int fd^{p-1} d\mu \leq C \int |\nabla f| d\mu + D \int f d\mu, \tag{3.40}$$

for all nonnegative locally Lipschitz f. Therefore, if $p \geq 2$, using the inequality $d^{p-1} \geq d - 1$ and enlarging the constant D, we see they also satisfy (3.39). Note that for such a measure we have $\int d(x) d\mu(x) < \infty$. Suppose that the compactness assumption (3.9) also holds for the potential of the measure. By the results of Section 3.2, the inequality (3.40) implies a Cheeger inequality, which when combined with (3.39), applied to the function $|f - \mu(f)|$, gives

$$\int |f - \mu(f)| d\, d\mu \leq C \int |\nabla f| d\mu,$$

with some constant C independent of f. Let g be a bounded continuous Lipschitz function with $\mu(g) = 0$ and $\|g\|_{\mathrm{Lip}} \leq 1$. Since, for all $x \in \mathbb{R}^n$,

$$|g(x) - g(0)| \leq d(x),$$

we arrive at

$$\int (f(x) - \mu f)(g(x) - g(0)) d\mu \leq C \int |\nabla f| d\mu,$$

which in turn implies

$$\int fg d\mu \leq C \int |\nabla f| d\mu,$$

using the assumption $\mu(g) = 0$. Now, choosing $f(x) = e^{tg(x)}$ with some $t \in \mathbb{R}$ and using the fact that $|\nabla g| \leq \|g\|_{\text{Lip}} \leq 1$ almost everywhere, we arrive at

$$\int g e^{tg} d\mu \leq C \int t e^{tg} d\mu,$$

which can be written as a differential inequality for the function $H(t) = \int e^{tg} d\mu$, as

$$H'(t) \leq CtH(t).$$

After integration, this yields

$$\int e^{tg} d\mu \leq e^{t^2 C/2}, \tag{3.41}$$

for all bounded continuous g with $\|g\|_{\text{Lip}} \leq 1$ and $\mu(g) = 0$. As observed in [BG99], (3.41) is equivalent to the $\mathbb{T}_1$ transportation inequality. To see this, recall that the duality formula for the Wasserstein distance states that

$$\mathcal{W}_1(\mu, \nu) = \sup \left(\nu(g) - \mu(g) \right), \tag{3.42}$$

where the supremum is taken over all continuous and bounded g with $\|g\|_{\text{Lip}} \leq 1$. Let $\varphi = d\nu/d\mu$ and g be such that $\|g\|_{\text{Lip}} \leq 1$ (we now drop the assumption that $\mu(g) = 0$). Applying (3.41) to the function $g - \mu(g)$ we obtain

$$\int e^{tg - \frac{Ct^2}{2} - t\mu g} d\mu \leq 1.$$

Using Sanov's variational characterisation of the entropy, which reads

$$\text{Ent}_\mu(\varphi) = \sup \left\{ \int \varphi \psi d\mu : \int e^\psi d\mu \leq 1 \right\},$$

we conclude that

$$\begin{aligned} \text{Ent}_\mu(\varphi) &\geq \int \left(tg - \frac{Ct^2}{2} - t\mu(g) \right) \varphi d\mu \\ &= t(\nu(g) - \mu(g)) - \frac{C}{2} t^2. \end{aligned}$$

Rearranging this, we obtain

$$\frac{C}{2}t + \frac{1}{t}\mathrm{Ent}_\mu(\varphi) \geq \nu(g) - \mu(g),$$

and optimising in t we arrive at

$$\sqrt{2C\mathrm{Ent}_\mu(\varphi)} \geq \nu(g) - \mu(g).$$

Finally, taking supremum over g we conclude by (3.42) that

$$\sqrt{2C\mathrm{Ent}_\mu(\varphi)} \geq \mathcal{W}_1(\mu, \nu).$$

As we will see below, the stronger inequality $\mathbb{T}_2$ actually holds for the measures satisfying (3.39). The proof uses a result of [OV00] (see also [BGL01]), which says that the logarithmic Sobolev inequality implies the $\mathbb{T}_2$ inequality. It would be interesting to find a direct proof of the $\mathbb{T}_p$ inequality starting from (3.40).

Theorem 3.18. *Let $p \geq 2$ and consider a probability measure*

$$\mu(dx) = Z^{-1}\mathrm{e}^{-V-W-d^p/p}dx \tag{3.43}$$

on an H- type group, where V is a function such that

$$\mathrm{osc}(V) = \sup V - \inf V < \infty$$

and W satisfies

$$|\nabla W| \leq \delta d^{p-1} + L_\delta,$$

for some constants $\delta < 1$ and $L_\delta \geq 0$. Then, for all probability measures which are absolutely continuous with respect to μ, we have

$$\mathcal{W}_p(\mu, \nu)^p \leq C\mathrm{Ent}_\mu\left(\frac{d\nu}{d\mu}\right),$$

with some constant C independent of ν.

Let us now state the Otto-Villani theorem.

Theorem 3.19. *Suppose that a measure μ satisfies the logarithmic Sobolev inequality*

$$\mathrm{Ent}(f^2) \le C_{LS} \int |\nabla f|^2 d\mu.$$

Then the measure μ satisfies a Talagrand inequality

$$\mathcal{W}_2(\mu,\nu)^2 \le C_T \mathrm{Ent}_\mu\left(\frac{d\nu}{d\mu}\right)$$

with constant $C_T = C_{LS}/2$.

The result actually holds for arbitrary $q \in (1,2]$: the q-logarithmic Sobolev inequality implies $\mathbb{T}_p$ (where $p \ge 2$ is the conjugate exponent to q). This can be seen by following the proof of [BGL01], which uses the Hamilton-Jacobi semigroup and a Herbst-type argument. We apply the q-logarithmic Sobolev inequality

$$\int f^q \log \frac{f^q}{\mu(f^q)} d\mu \le C_{\mathrm{LS}} \int |\nabla f|^q d\mu$$

to the function $f = \mathrm{e}^{at^\alpha Q_t g/q}$, where g is a bounded Lipschitz function and a, α are constants to be determined. We obtain

$$\int \mathrm{e}^{at^\alpha Q_t g} at Q_t g d\mu - G(t)\log G(t) \le C_{\mathrm{LS}} \int \frac{a^q t^{\alpha q}}{q^q} |\nabla Q_t g|^q \mathrm{e}^{at^\alpha Q_t g} d\mu,$$

where

$$G(t) = \int \mathrm{e}^{at^\alpha Q_t g} d\mu.$$

Recall that $Q_t g$ solves

$$\partial_t Q_t f = -\frac{|\nabla Q_t f|^q}{q}$$

and therefore the inequality above can be rewritten as

$$\int \mathrm{e}^{at^\alpha Q_t g}\left(\alpha a t^{\alpha-1} Q_t g + \frac{\alpha C_{\mathrm{LS}}}{q^{q-1}} a^q t^{\alpha q-1} \partial_t Q_t f\right) d\mu \le \alpha \, \frac{G(t)\log G(t)}{t}.$$

If we now choose

$$\alpha = p - 1 = \frac{p}{q}, \quad a = \left(\frac{q^{q-1}}{\alpha C_{\mathrm{LS}}} \right)^{\frac{1}{q-1}},$$

this can be rewritten as

$$G'(t) = \int \mathrm{e}^{at^{p-1} Q_t g} \partial_t \left(a t^{p-1} Q_t g \right) d\mu \leq (p-1) \frac{G(t) \log G(t)}{t}.$$

In other words,

$$\left(\frac{\log G(t)}{t^{p-1}} \right)' \leq 0.$$

In particular,

$$\log G(1) \leq \frac{\log G(\varepsilon)}{\varepsilon^{p-1}}$$

for any $\varepsilon \leq 1$ and letting $\varepsilon \to 0$ we conclude that, since $G'(0) = a\mu(g)$ [1],

$$\int \mathrm{e}^{aQ(g)} d\mu \ \leq \mathrm{e}^{a\mu(g)}$$

(recall that we denote $Q = Q_1$). This is known as an infimum-convolution inequality and is equivalent to the $\mathbb{T}_q$ transportation inequality, as observed in [BGL01]. Let ν be a probability measure which is absolutely continuous with respect to μ and let $\varphi = d\nu/d\mu$. Arguing as in the proof of $\mathbb{T}_1$, the infimum–convolution inequality implies that

$$\mathrm{Ent}_\mu(\varphi) = \sup \left\{ \int \varphi \psi d\mu : \int \mathrm{e}^{\psi} \leq 1 \right\} \geq a \int \varphi \left(Q(g) - \mu(g) \right) d\mu.$$

Taking supremum over all bounded Lipschitz g, by the dual description of the Wasserstein distance (2.27), we obtain

$$\mathcal{W}_p(\mu, \nu)^p = \sup \left(\int Qf d\mu - f d\nu \right) \leq \frac{1}{a} \mathrm{Ent}_\mu(\varphi), \tag{3.44}$$

which is the $\mathbb{T}_p$ inequality (2.29). Therefore, to establish Theorem 3.18 it remains to

[1]For a proof of all the regularity properties of Q_t used in this discussion we refer the reader to [LV07] ([BEHM09]).

prove that the measure defined in (3.43) satisfies the q- logarithmic Sobolev inequality. Recall that we have already seen in Corollary 3.13 that the measure satisfies a Φ-Entropy inequality with $\Phi(x) = t\log^{1/q}(1+t)$. We now show that this implies the q-logarithmic Sobolev inequality.

Theorem 3.20. *Suppose that a measure μ satisfies the Φ-entropy inequality (3.20) for some $\beta \in [1/2, 1]$ and set $q = \frac{1}{\beta} \in [1, 2]$. Then there exists a constant C_q such that the q-logarithmic Sobolev inequality holds*

$$\mathrm{Ent}_\mu(f^q)d\mu \leq c_q \int |\nabla f|^q d\mu, \tag{3.45}$$

for all nonnegative locally Lipschitz functions f.

Proof. We apply (3.20) to the function $g = f\left(1 + \log(1+f)\right)^{1-\beta} \geq f \geq 0$, where f is such that $\mu(f) = 1$. Note that $\mu(g) \geq 1$. We have

$$\begin{aligned}
\int g \log\left(1 + \frac{g}{\mu(g)}\right)^\beta d\mu &= \int f\left(1 + \log(1+f)\right)^{1-\beta} \log\left(1 + \frac{g}{\mu(g)}\right)^\beta d\mu \\
&\geq \int f\left(1 + \log(1+f)\right)^{1-\beta} \log\left(1 + \frac{f}{\mu(g)}\right)^\beta d\mu \\
&\geq \int f\left(1 + \log\left(1 + \frac{f}{\mu(g)}\right)\right)^{1-\beta} \log\left(1 + \frac{f}{\mu(g)}\right)^\beta d\mu \\
&\geq \int f \log\left(1 + \frac{f}{\mu(g)}\right) d\mu \\
&= \int f(\log(\mu(g) + f) - \log \mu(g)) d\mu \\
&\geq \int f \log(1+f) d\mu - \mu(g).
\end{aligned}$$

Thus for all $f \geq 0$ with $\mu(f) = 1$,

$$\begin{aligned}
\int f \log(1+f) d\mu \leq{}& C_\Phi \int \left| \nabla \left(f \left(1 + \log\left(1+f\right)\right)^{1-\beta} \right) \right| d\mu \\
&+ \left((\log 2)^\beta + 1 \right) \int f \left(1 + \log(1+f)\right)^{1-\beta} d\mu \\
\leq{}& C_\Phi \int \left(1 + \log(1+f)\right)^{1-\beta} |\nabla f| d\mu \\
&+ C_\Phi (1-\beta) \int \frac{f}{\left(1 + \log(1+f)\right)^\beta} \frac{1}{1+f} |\nabla f| d\mu \\
&+ \left((\log 2)^\beta + 1 \right) \int f \left(1 + \log(1+f)\right)^{1-\beta} d\mu \\
\leq{}& C_\Phi \int \left(1 + \log(1+f)\right)^{1-\beta} |\nabla f| d\mu \ + C_\Phi (1-\beta) \int |\nabla f| d\mu \\
&+ \left((\log 2)^\beta + 1 \right) \int f \left(1 + \log(1+f)\right)^{1-\beta} d\mu. \qquad (3.46)
\end{aligned}$$

Since we have assumed $\beta \geq \frac{1}{2}$, we have $1 - \beta \leq \beta$ and hence

$$\begin{aligned}
\int f \left(1 + \log(1+f)\right)^{1-\beta} d\mu &= 1 + \int f \log(1+f)^{1-\beta} d\mu \\
&\leq \int f \log(1+f)^\beta d\mu + 2 \\
&\leq C_\Phi \int |\nabla f| d\mu + (\log 2)^\beta + 2
\end{aligned}$$

by another application of (3.20) the last step. Summarising, we have shown that for $f \geq 0$,

$$\begin{aligned}
\int f \log\left(1 + \frac{f}{\mu f}\right) d\mu \leq{}& C_\Phi \int \left(1 + \log\left(1 + \frac{f}{\mu f}\right)\right)^{1-\beta} |\nabla f| d\mu \\
&+ C_\Phi (2 - \beta + (\log 2)^\beta) \int |\nabla f| d\mu \\
&+ \left((\log 2)^\beta + 1 \right) \left((\log 2)^\beta + 2 \right) \int f d\mu.
\end{aligned}$$

Replacing f by f^q with $q = \frac{1}{\beta}$ in the above and using Young's inequality we arrive at

$$\begin{aligned}\int f^q \log\left(1+\frac{f^q}{\mu f^q}\right) d\mu \leq qC_\Phi \int \left(1+\log\left(1+\frac{f^q}{\mu f^q}\right)\right)^{1-\beta} f^{q-1}|\nabla f| d\mu \\ + qC_\Phi(2-\beta+(\log 2)^\beta) \int f^{q-1}|\nabla f| d\mu \\ + \left((\log 2)^\beta+1\right)\left((\log 2)^\beta+2\right) \int f^q d\mu \\ \leq \frac{qC_\Phi \varepsilon^{p-1}}{p} \int f^q \left(1+\log\left(1+\frac{f^q}{\mu f^q}\right)\right) d\mu \\ + \left(\frac{C_\Phi}{\varepsilon}+c(2-\beta+(\log 2)^\beta)\right)\int |\nabla f|^q d\mu \\ + \tilde{c}\mu(f^q),\end{aligned}$$

for all $\varepsilon > 0$, with

$$\tilde{c} = \frac{qC_\Phi}{p}(2-\beta+(\log 2)^\beta) + \left((\log 2)^\beta+1\right)\left((\log 2)^\beta+2\right).$$

Choosing $qc\varepsilon^{p-1}/p < 1$, we can simplify this bound as follows

$$\int f^q \log\left(1+\frac{f^q}{\mu f^q}\right) d\mu \leq C\int |\nabla f|^q d\mu + D\int f^q d\mu,$$

where

$$C' = \frac{\frac{C_\Phi}{\varepsilon}+C_\Phi(2-\beta+(\log 2)^\beta)}{1-\frac{qC_\Phi\varepsilon^{p-1}}{p}}, \qquad D' = \frac{p\tilde{c}}{p-qC_\Phi\varepsilon^{p-1}}.$$

From this one obtains the defective q-log Sobolev inequality, which for all $f \geq 0$ such that $\mu(f^q) = 1$ can be equivalently represented as

$$\int f^q \log f^q d\mu \ \leq C' \int |\nabla f|^q d\mu \ + D'. \tag{3.47}$$

Let us now recall that by Corollary 3.15 and Proposition 2.17, there exists a constant C_q such that

$$\int |f-\mu f|^q d\mu \ \leq C_q \int |\nabla f|^q d\mu.$$

Finally, by Theorem 2.18, we can remove the defective term in (3.47) to arrive at the result. □

Chapter 4

Gradient bounds for the heat semigroup

Let $\mathbb{G} = (\mathbb{R}^{m+r}, \cdot)$ be an H-type group and consider the sub-Laplacian

$$\Delta = \Delta_{\mathbb{G}} = \sum_{i=1}^{m} X_i^2.$$

Let $P_t = \mathrm{e}^{t\Delta}$ be the heat semigroup. It is known that P_t is given by a convolution kernel, i.e.

$$P_t f(q) = \int_{\mathbb{G}} f(q \cdot w) h_t(w) dw$$

and moreover if $w = (x, z)$ with $x \in \mathbb{R}^m$, $z \in \mathbb{R}^r$, then $h_t(w)$ has the representation [Ran96]

$$h_t(w) = \int_{\mathbb{R}^r} \mathrm{e}^{i\langle q,z\rangle - \frac{1}{4}|q| \coth(t|q|)|x|^2} \left(\frac{|q|}{\sinh(t|q|)^n} \right) dq. \tag{4.1}$$

Note that this depends only on the norms of x and z. For simplicity, in the sequel we will write $h = h_1$ while we denote $\nu(dw) = h(w)dw$.

Theorem 4.1. *Let $q \geq 1$. For all smooth $f : \mathbb{G} \to \mathbb{R}$ and all $x \in \mathbb{G}$, $t \geq 0$*

$$|\nabla P_t f|^q \leq \kappa_q P_t(|\nabla f|^q), \tag{4.2}$$

for some constant $\kappa_q > 0$.

The above gradient bound was naturally first studied in the case of the Heisenberg group. For $q > 1$ the result first appeared in [DM05]. Eventually, the whole range $q \geq 1$ was treated in [Li06]. Another proof for $q > 1$ was later given in [LP10]. In [BBBC08], the authors gave various proofs of (4.2) for $q \geq 1$ (still restricted to the Heisenberg group case). Finally, their methods were extended in [Eld10], where the result for arbitrary H-type groups was obtained. Let us point out that the case $q = 1$ is of particular interest, because it implies the logarithmic Sobolev inequality as well as Bobkov's isoperimetric inequality (2.21) (in a local form for the semigroup P_t).

By Jensen's inequality it is plain that the gradient bound for $q = 1$ implies the others (up to a universal constant) since

$$|\nabla P_t f| \leq \kappa_1(t) P_t |\nabla f| \leq \kappa_1(t) \left(P_t |\nabla f|^q\right)^{1/q}$$

and therefore (4.2) is satisfied with $\kappa_q(t) = \kappa_1(t)^q$.

We also note that, by the structure of the group, it suffices to prove (4.2) at $t = 1$ and $x = 0$. Following [BBBC08, DM05], given $y \in \mathbb{G}$, consider the left-translation $L_y(x) = y \cdot x,\ x \in \mathbb{G}$. Since Δ is given as a sum of squares of left-invariant vector fields, it follows that Δ is left-invariant and thus P_t commutes with left-translations, i.e.

$$L_y P_t(f)(x) = (P_t f)(L_y(x)) = P_t(L_y f)(x).$$

Therefore, choosing $x = 0$ we see that, for all y,

$$\nabla(P_t f)(y) = \nabla(P_t f \circ L_y)(0) = \nabla(P_t(f \circ L_y))(0),$$

since ∇ is invariant under translations. Moreover, for any $\lambda, t > 0$, we can see directly from formula (4.1) that

$$P_t(f \circ \delta_\lambda) = (P_{\lambda^2 t} f) \circ \delta_\lambda,$$

so that P_1 gives the whole of P_t.

We begin by stating the following result of [Eld09] (see also [Li06]), which provides a Gaussian estimate for the heat kernel. Recall that we use the notation $(x, z) = w$

and $d(0,w) = d(w)$.

Proposition 4.2. *Let $\mathbb{G} = (\mathbb{R}^{m+r}, \cdot)$ be an H-type group. There exists $R > 0$ such that for all points $w = (x,z) = (x_1, \ldots, x_m, z_1, \ldots, z_r)$ with $d(w) = d(0,(x,z)) > R$,*

$$h(x,z) \asymp \frac{d(w)^{m-r-1}}{1 + (|x| d(w))^{(m-1)/2}} \mathrm{e}^{-d(w)^2/4} \tag{4.3}$$

and

$$|\nabla \log h(w)| \leq C_h (1 + d(w)), \tag{4.4}$$

for some constant C_h independent of w.

As a consequence, when combined with Theorem 3.1 and Proposition 3.3, the above bounds imply that the heat kernel measure satisfies a U-bound as well as a Cheeger inequality.

Lemma 4.3. *For all nonnegative smooth f,*

$$\int f d \, d\nu \leq C \int |\nabla f| d\nu + D \int f d\nu$$

and

$$\int |f - \nu(f)| d\nu \leq C' \int |\nabla f| d\nu$$

where $C, D, C' \geq 0$ are constants independent of f.

Before proving the lemma, let us see how it can be used to reach (4.2). Given a smooth f, we observe that

$$\left| \int f(\nabla h) dx \right| = \left| \int (f - \nu(f)) (\nabla h) dx \right| \leq \int |f - \nu(f)| \, |\nabla \log h| \, d\nu, \tag{4.5}$$

which, combined with (4.4), implies

$$\left| \int f(\nabla h) dx \right| \leq C_h \int |f - \nu(f)| d \, d\nu + \; C_h \int |f - \nu(f)| d\nu.$$

Now Lemma 4.3 applied to the function $|f-\nu(f)|$ allows us to conclude that

$$\left|\int f(\nabla h)dx\right| \le C_hC\int |\nabla f|d\nu + C_h(1+D)\int |f-\nu(f)|d\nu \le \kappa\int |\nabla f|d\nu,$$

with

$$\kappa = C_hC + C'C_h(1+D).$$

We have thus established (4.2) for $q=1$ at $t=1$ and $x=0$.

Proof of Lemma 4.3. The proof is based on perturbative arguments, similar to the ones used in [HZ10] (Theorem 7.1). We note that if f,g are functions such that $f\asymp g$, then we may write $f=\mathrm{e}^{\psi}g$, where ψ is a function of bounded oscillation. Indeed, we may define ψ pointwise by

$$\psi(x) = \log\frac{f(x)}{g(x)}.$$

Then ψ is well defined since the ratio f/g is bounded from above and below by absolute constants. Now, consider the function

$$W(w) = \log\left(\frac{d(w)^{m-r-1}}{(1+\varepsilon|x|d(w))^{(m-1)/2}}\right),$$

on $\{d>R\}$, where R is as in Proposition 4.2, with some $\varepsilon\in(0,1)$ to be determined later. Since

$$(1+|x|d(w))^{(m-1)/2} \asymp 1+(|x|d(w))^{(m-1)/2}$$

and

$$\varepsilon^{(m-1)/2}(1+|x|d(w))^{(m-1)/2} \le (1+\varepsilon|x|d(w))^{(m-1)/2} \le (1+|x|d(w))^{(m-1)/2},$$

we see that

$$\mathrm{e}^{W(w)} \asymp \frac{d(w)^{m-r-1}}{1+(|x|d(w))^{(m-1)/2}}.$$

Therefore, by (4.3), on the set $\{d > R\}$, we may write

$$h(w) = \mathrm{e}^{\psi(w)+W(w)}\mathrm{e}^{-d(w)^2/4},$$

with a function ψ satisfying $\mathrm{osc}(\psi) < \infty$. Let f be a nonnegative smooth function. By Corollary 3.4, the proof of

$$\int_{\{d>R\}} fd\, d\nu \leq C \int_{\{d>R\}} |\nabla f| d\nu + D \int_{\{d>R\}} f d\nu$$

amounts to showing that

$$|\nabla W| \leq \delta d + L_\delta,$$

for some $\delta < 1$ and $L_\delta \geq 0$. Indeed, using the triangle inequality we compute

$$\begin{aligned} |\nabla W| &\leq (m-r-1)\frac{|\nabla d(w)|}{d(w)} + \varepsilon \frac{m-1}{2} \frac{|\nabla|x||d(w) + |x||\nabla d(w)|}{1+\varepsilon|x|d(w)} \\ &\leq \frac{m-r-1}{R} + \varepsilon\,(m-1)\,d(w) \end{aligned}$$

where we used that $|\nabla|x|| = |\nabla d| = 1$ and $|x| \leq d(w)$. To conclude, we choose $\varepsilon < (m-1)^{-1}$. Finally, we have

$$\int fd\, d\nu \leq \int_{\{d>R\}} fd\, d\nu + R\int f d\nu \leq C\int |\nabla f| d\nu + (D+R)\int f d\nu.$$

The proof of the Cheeger inequality is essentially the same as the proof of Lemma 3.5, so we sketch it briefly. Let $r > 0$ to be determined and set $\bar{f} = \int_{B_r} f dx/|B_r|$. We have

$$\int |f - \nu(f)| d\nu \leq 2\int |f - \bar{f}| d\nu = 2\int_{\{d<r\}} |f-\bar{f}| d\nu + 2\int_{\{d\geq r\}} |f - \bar{f}| d\nu.$$

For the first term we use the Cheeger inequality in the ball to conclude that

$$\int_{\{d<r\}} |f - \bar{f}| d\nu \leq C_r \int |\nabla f| d\nu,$$

with some $C_r > 0$, while for the second one, we use the U-bound for the function $|f - \bar{f}|$ to obtain

$$\int_{\{d \geq r\}} |f - \bar{f}| d\nu \leq \frac{1}{r} \int |f - \bar{f}| d d\nu \leq \frac{C_1}{r} \int |\nabla f| d\nu + \frac{D}{r} \int |f - \bar{f}| d\nu.$$

Choosing r large enough and rearranging we obtain the desired result. □

In conclusion, let us mention that the gradient bound

$$|\nabla P_t f| \leq \kappa P_t |\nabla f|$$

has many interesting consequences in terms of functional and isoperimetric inequalities for the semigroup, which are outlined in [BBBC08, Bon09] (see also [BÉ85, BL96, BL06, Fou00]).

Chapter 5

Markov Semigroups with Hörmander Generators on H-type groups

In this chapter, we consider second order differential operators which are given in terms of vector fields satisfying Hörmander's condition. We will prove gradient bounds for the associated semigroup, with the full gradient (which includes all the vector fields needed to span the Lie algebra, instead of only the fields appearing in the sub-gradient), as well as some Li-Yau estimates.

Consider an H-type group on $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^r$. As before, we write the elements of $\mathbb{R}^n$ as $w = (x, z)$ with $x \in \mathbb{R}^m$, $z \in \mathbb{R}^r$. We denote by $X_1, \ldots, X_m$ the vector fields that belong to the first layer of the stratification, while we denote by $\{Z_k\}_{k=1}^n$, a basis for the Lie algebra. In other words, we take $X_k = Z_k$, for $k = 1, \ldots, m$, while the remaining $Z_{m+1}, \ldots, Z_n$ are ordered commutators of length 2. The Lie algebra is naturally equipped with a first order operator D which generates dilations and satisfies

$$e^{sD} Z_k e^{-sD} = e^{sl_k} Z_k \text{ and } [Z_k, D] = l_k Z_k, \tag{5.1}$$

for all $k = 1, \ldots, n$ and $s > 0$, where $l_k = 1$ for $k = 1, \ldots m$ and $l_k = 2$ otherwise (the constants l_k reflect the layer of the Lie algebra that Z_k belongs to). More specifically,

D is given as the generator of the dilations $\delta_\lambda(w) = (\lambda x, \lambda^2 z)$, by

$$D = \partial_\lambda \mid_{\lambda=1} \delta_\lambda(w) = x \cdot \nabla_m + 2z \cdot \nabla_r, \tag{5.2}$$

where ∇_m and ∇_r denote the Euclidean gradients on $\mathbb{R}^m$ and $\mathbb{R}^r$, respectively.

We consider the operator

$$\mathrm{L} = \sum_{i=1}^{m} (\delta_{ij} + G_{ij}) X_i X_j + \sum_{k=1}^{n} \alpha_k Z_k - \beta D, \tag{5.3}$$

acting on smooth bounded functions on $\mathbb{R}^n$, where $\beta > 0$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ are constants, while $G = (G_{ij})_{i,j=1}^m$ is a constant matrix satisfying

$$G^* \geq 0,$$

where $G^*_{ij} = (G_{ij} + G_{ji})/2$ is the symmetrised matrix of G. Let $(P_t)_{t \geq 0}$ denote the semigroup generated by L. By Theorem 2.16 , for every $x \in \mathbb{R}^n$, $P_t f(x)$ is given by a smooth Markov kernel. The sub-gradient of a smooth function f, will be denoted by

$$\overline{\Gamma}(f) = \sum_{i=1}^{m} (X_i f)^2, \tag{5.4}$$

while the full gradient of f will be denoted by

$$\Gamma(f) = \sum_{k=1}^{n} (Z_k f)^2. \tag{5.5}$$

The corresponding quadratic forms are given by

$$\overline{\Gamma}(f, g) = \sum_{i=1}^{m} (X_i f)(X_i g), \quad \Gamma(f, g) = \sum_{k=1}^{n} \big(Z_k f\big)\big(Z_k g\big),$$

respectively. Finally, we define $\hat{\Gamma}(f, g) := \Gamma(f, g) - \overline{\Gamma}(f, g)$. The operator Γ_2 is defined as

$$\Gamma_2(f, g) = \frac{1}{2}(\mathrm{L}\Gamma(f, g) - \Gamma(f, \mathrm{L}g) - \Gamma(g, \mathrm{L}f)).$$

We define similarly

$$\overline{\Gamma}_2(f,g) = \frac{1}{2}(\mathrm{L}\overline{\Gamma}(f,g) - \overline{\Gamma}(f,\mathrm{L}g) - \overline{\Gamma}(g,\mathrm{L}f))$$

and

$$\hat{\Gamma}_2(f,g) = \frac{1}{2}(\mathrm{L}\hat{\Gamma}(f,g) - \hat{\Gamma}(f,\mathrm{L}g) - \hat{\Gamma}(g,\mathrm{L}f))$$

and set $\Gamma_2(f) = \Gamma_2(f,f)$ (and similarly for $\overline{\Gamma}_2$ and $\hat{\Gamma}_2$). Our definition of Γ is convenient for the analysis below, but it is not the standard one, i.e. it doesn't denote the carré du champ of L, defined as $\Gamma_{\mathrm{L}}(f) := \frac{1}{2}\mathrm{L}f^2 - f\mathrm{L}f$. However, we have

$$\overline{\Gamma}(f) \leq \Gamma_{\mathrm{L}}(f) = \sum_{i,j=1}^{m}(\delta_{ij} + G^*_{ij})(X_i f)(X_j f) \leq \left(1 + \max_i \sum_{j=1}^{m} G^*_{ij}\right)\overline{\Gamma}(f),$$

where the estimate on the left follows from the assumption that $G^* \geq 0$, while the one on the right by Young's inequality. Therefore, $\overline{\Gamma}(f)$ can be thought of as the carré du champ of L, up to a constant.

For $i, j = 1, \ldots, m$, we set

$$[X_i, X_j] =: Y_{ij}. \tag{5.6}$$

For example, in the Heisenberg group we have $X_1 = (1, 0, -y/2)^T$ and $X_2 = (0, 1, x/2)^T$ on $(x, y, z) \in \mathbb{R}^3$. In this case $Z_1 = X_1$, $Z_2 = X_2$ and $Z_3 = Y_{12} = [X_1, X_2] = (0, 0, 1)^T$.

We conclude this introductory section with a few remarks about the derivatives of the Folland-Kaplan gauge N. Recall that this is defined as

$$N(w) = \left(|x|^4 + 16|z|^2\right)^{1/4}. \tag{5.7}$$

A computation shows [DGN03] that the sub-gradient and the sub-Laplacian of N read

$$\sum_{i=1}^{m} |X_i N|^2(w) = \frac{|x|^2}{N^2(w)} \tag{5.8}$$

and

$$\sum_{i=1}^{m} X_i^2 N(w) = 3\frac{|x|^2}{N^3(w)}, \tag{5.9}$$

respectively, where $w = (x, z)$ and $|\cdot|$ is Euclidean norm on $\mathbb{R}^m$. In particular, $\overline{\Gamma}(N) \leq 1$, since $|x| \leq N$. Moreover, recalling that for $j > m$, $Z_j = \partial_j$, for all $i = 1, \ldots, r$ we have

$$Z_{m+i}N = \frac{8z_i}{N^3},$$

and therefore

$$\sum_{i=1}^{n} |Z_i N|^2(w) = \frac{|x|^2}{N^2(w)} + \frac{64|z|^2}{N^6(w)} \leq 1 + \frac{4}{N^2(w)},$$

using that $|x| \leq N(w)$ and $16|z|^2 \leq N^4(w)$. Let $(\mathrm{Hess}N)^*$ denote the symmetrised Hessian of N, i.e. the matrix with elements

$$(\mathrm{Hess}N)^*_{ij} = \frac{1}{2}(X_i X_j N + X_j X_i N),$$

for $i, j = 1, \ldots, m$. It was shown in [GT10] that

$$(\mathrm{Hess}N)^*_{ij} = \frac{1}{N^7}\left(N^4|x|^2\delta_{ij} + 2N^4\left(x_i x_j + \sum_{s=1}^{r} B_{is}B_{js}\right) - 3\langle A, e_i\rangle\langle A, e_j\rangle\right),$$

where, for $s = 1, \ldots, r$,

$$B_{is} = B_{is}(w) = \left\langle J_{e_{m+s}} x, e_i \right\rangle$$

and $A = |x|^2 x + 4J_z x$. Using the identities $|J_z x| = |x||z|$ and $\langle J_z x, x\rangle = 0$ (see Proposition 2.7), we see that $B_{is} \leq |J_{e_{m+s}} x| = |x|$ and $|A|^2 = |x|^4|x|^2 + 16|x|^2|z|^2 = N^4|x|^2$. Using Young's inequality, we thus arrive at the estimate

$$\begin{aligned}
\left|(\mathrm{Hess}N)^*_{ij}\right| &\leq \frac{1}{N^7}\left(N^4|x|^2\delta_{ij} + 2N^4\left(\frac{x_i^2 + x_j^2}{2} + r|x|^2\right) + 3|A|^2\right) \\
&\leq \frac{1}{N^7}\left(N^4|x|^2\delta_{ij} + 2(1+r)N^4|x|^2 + 3N^4|x|^2\right) \\
&\leq \frac{\delta_{ij} + 2r + 5}{N},
\end{aligned}$$

where we used once again that $|x| \leq N$. Summing over i, j, we obtain

$$\sum_{i,j=1}^{m} \left|(\mathrm{Hess}N)^*_{ij}\right| \leq \frac{m + 2rm^2 + 5m^2}{N}.$$

Let us also observe that

$$DN = \sum_{i=1}^{m} x_i \partial_i N + 2\sum_{i=1}^{r} z_i \partial_{m+i} N = N^{-3}\left(|x|^4 + 16|z|^2\right) = N.$$

Combining the above estimates and using the Cauchy-Schwarz inequality, we conclude that

$$\begin{aligned}
\mathrm{L}N &= \sum_{i,j=1}^{m} (\delta_{ij} + G_{ij})(X_i X_j N) + \sum_{i=1}^{n} \alpha_i Z_i N - \beta DN \\
&\leq \max_{i,j}(\delta_{ij} + G_{ij}) \sum_{i,j=1}^{m} \left|(\mathrm{Hess} N)^*_{ij}\right| + |\alpha| \sqrt{\sum_{i=1}^{n} |Z_i N|^2} - \beta DN \\
&\leq \frac{c_1}{N} + |\alpha| \sqrt{1 + \frac{4}{N^2}} - \beta N,
\end{aligned}$$

where $c_1 = \max_{i,j}(\delta_{ij} + G_{ij})(m + 2rm^2 + 5m^2)$ and $|\alpha|^2 = \sum_{i=1}^{n} \alpha_i^2$. Let

$$W(w) = 1 + N(w)^2. \tag{5.10}$$

Then $W \geq 1$ is a smooth function. Since L is a diffusion generator, we have

$$\begin{aligned}
\mathrm{L}W = 2N\mathrm{L}N + 2\overline{\Gamma}(N) &\leq 2N\left(\frac{c_1}{N} + |\alpha|\sqrt{1 + \frac{4}{N^2}} - \beta N\right) + 2 \\
&= -2\beta N^2 + 2|\alpha|\sqrt{N^2 + 4} + 2(c_1 + 1).
\end{aligned}$$

Recall that, if d denotes the Carnot-Carathéodory distance and $d_N(w, y) = N(y^{-1}w)$ is the distance induced by N, then there is a universal constant R such that

$$\frac{1}{R} d(w, y) \leq d_N(w, y) \leq R d(w, y),$$

for all $w, y \in \mathbb{R}^n$. There exists therefore a number $\tilde{r} > 0$ such that if $N(w) \geq \tilde{r}$, then

$$\mathrm{L}W(w) \leq -\beta W(w).$$

Setting $r = \tilde{r}/R$, we conclude that W satisfies the Lyapunov condition

$$\mathrm{L}W(w) \leq -\beta W(w) + B\chi_{B_r}(w), \tag{5.11}$$

for some constant $B > 0$, where χ_{B_r} is the indicator function of the ball of radius r, $B_r = \{d(x) < r\}$. As we already mentioned in Section 3.1, if there is a reversible invariant measure μ for this generator, such an estimate can be used to prove functional inequalities for μ.

5.1 Gradient Bounds

We start by proving a gradient bound for the semigroup $P_t = \mathrm{e}^{t\mathrm{L}}$ and the gradient Γ. We then use this bound to prove that the semigroup converges weakly to a probability measure on $\mathbb{R}^n$ as $t \to \infty$.

Theorem 5.1. *For any $q > 1$ there exists a constant $\kappa_q \in \mathbb{R}$ such that*

$$\Gamma(P_t f)^{\frac{q}{2}} \leq \mathrm{e}^{-\kappa_q t} P_t \left(\Gamma(f)^{\frac{q}{2}}\right), \tag{5.12}$$

for all smooth f and all $t > 0$. Moreover, there exists $\beta_q > 0$ such that if $\beta > \beta_q$, then $\kappa_q > 0$.

Let us begin with a preliminary calculation.

Lemma 5.2. *There exists a constant $c > 0$ such that for all $\lambda > 0$ and all smooth nonnegative functions f,*

$$\overline{\Gamma}_2(f) + \lambda\hat{\Gamma}_2(f) \geq \frac{1}{2}\sum_{i,k=1}^{m}(X_iX_kf)^2 + \frac{\lambda}{2}\sum_{i=1}^{m}\sum_{k=m+1}^{n}(X_iZ_kf)^2 + \left(\beta - \frac{c}{\lambda}\right)\overline{\Gamma}(f) + \frac{1}{16}\hat{\Gamma}(f). \tag{5.13}$$

Therefore, there exist constants $A, B > 0$ such that

$$\Gamma_2(f) \geq A\sum_{i=1}^{m}\sum_{k=1}^{n}(X_iZ_kf)^2 + B\Gamma(f). \tag{5.14}$$

Proof. We first estimate $\overline{\Gamma}_2(f) = \frac{1}{2}\mathrm{L}\overline{\Gamma}(f) - \overline{\Gamma}(f, \mathrm{L}f)$. On the one hand,

$$\begin{aligned}
\mathrm{L}\overline{\Gamma}(f) &= \sum_{i,j=1}^{m}\sum_{k=1}^{n}(\delta_{ij} + G_{ij})X_iX_j(X_kf)^2 + \sum_{i=1}^{n}\sum_{k=1}^{m}\alpha_i Z_i(X_kf)^2 - \beta\sum_{k=1}^{m} D(X_kf)^2 \\
&= 2\sum_{i,j,k=1}^{m}(\delta_{ij} + G_{ij})(X_iX_kf)(X_jX_kf) + 2\sum_{k=1}^{m}(X_kf)(\mathrm{L}X_kf),
\end{aligned}$$

while on the other hand

$$\overline{\Gamma}(f, \mathrm{L}f) = \sum_{k=1}^{m}(X_kf)(X_k\mathrm{L}f).$$

Therefore,

$$\begin{aligned}
\overline{\Gamma}_2(f) &= \sum_{i,j,k=1}^{m}(\delta_{ij} + G_{ij})(X_iX_kf)(X_jX_kf) \ + \sum_{k=1}^{m}(X_kf)([\mathrm{L}, X_k]f) \\
&= \sum_{i,j,k=1}^{m}(\delta_{ij} + G^*_{ij})(X_iX_kf)(X_jX_kf) \ + \sum_{k=1}^{m}(X_kf)([\mathrm{L}, X_k]f) \\
&\geq \sum_{i,k=1}^{m}(X_iX_kf)^2 + \sum_{k=1}^{m}(X_kf)([\mathrm{L}, X_k]f),
\end{aligned}$$

since $G^* \geq 0$. Next, we estimate the commutator $[\mathrm{L}, X_k]$. For every $k = 1, \dots, m$, we have

$$\begin{aligned}
[\mathrm{L}, X_k] &= \sum_{i,j=1}^{m} (\delta_{ij} + G_{ij})[X_i X_j, X_k] + \sum_{i=1}^{n} \alpha_i [Z_i, X_k] - \beta[D, X_k] \\
&= \sum_{i,j=1}^{m} (\delta_{ij} + G_{ij})(X_i[X_j, X_k] + [X_i, X_k]X_j) + \sum_{i=1}^{m} \alpha_i [X_i, X_k] + \beta X_k \\
&= \sum_{i,j=1}^{m} (\delta_{ij} + G_{ij}) X_i Y^{jk} + \sum_{i,j=1}^{m} (\delta_{ij} + G_{ij}) Y^{ik} X_j + \sum_{i=1}^{m} \alpha_i Y^{ik} + \beta X_k \\
&= 2 \sum_{i,j=1}^{m} (\delta_{ij} + G^*_{ij})(X_i Y^{jk} + Y^{jk} X_i) + \sum_{i=1}^{m} \alpha_i Y^{ik} + \beta X_k \\
&= 4 \sum_{i,j=1}^{m} (\delta_{ij} + G^*_{ij}) X_i Y^{jk} + \sum_{i=1}^{m} \alpha_i Y^{ik} + \beta X_k,
\end{aligned}$$

where we used that $-\beta[D, X_k] = \beta X_k$ and the fact that X_i and Y^{jk} commute for all i, j, k, since Y^{jk} is in the centre of the Lie algebra. Summing over $k \in \{1, \dots, m\}$ and using Young's inequality, we obtain

$$\begin{aligned}
\sum_{k=1}^{m} (X_k f)([\mathrm{L}, X_k] f) =& 4 \sum_{i,j,k=1}^{m} (\delta_{ij} + G^*_{ij})(X_k f)(X_i Y^{jk} f) \\
& + \sum_{i,k=1}^{m} \alpha_i (X_k f)(Y^{ik} f) + \beta \sum_{k=1}^{m} (X_k f)^2 \\
\geq & - 4 \sum_{i,j,k=1}^{m} \left(\frac{(\delta_{ij} + G^*_{ij})^2 (X_k f)^2}{2\varepsilon} + \varepsilon \frac{(X_i Y^{jk} f)^2}{2} \right) \\
& - \sum_{i,k=1}^{m} \left(\frac{\alpha_i^2 (X_k f)^2}{2\varepsilon} + \varepsilon \ \frac{(Y^{ik} f)^2}{2} \right) + \beta \sum_{k=1}^{m} (X_k f)^2 \\
\geq & - \frac{c_1}{\varepsilon} \overline{\Gamma}(f) - 2\varepsilon \sum_{i,j,k=1}^{m} (X_i Y^{jk} f)^2 - \frac{\varepsilon}{2} \sum_{i,k=1}^{m} (Y^{ik} f)^2 + \beta \overline{\Gamma}(f),
\end{aligned} \tag{5.15}$$

for all $\varepsilon > 0$ with

$$c_1 = 2\sum_{i,j=1}^{m}(\delta_{ij} + G^*_{ij})^2 + \frac{1}{2}\sum_{i=1}^{m}\alpha_i^2.$$

Now, for every $j,k = 1,\dots,m$ there exist constants $\sigma_r^{jk} \in \mathbb{R}$ such that

$$Y^{jk} = \sum_{r=m+1}^{n}\sigma_r^{jk}Z_r,$$

since the vector field Y^{ik} is obtained as a commutator of fields from the first layer of the Lie algebra and hence belongs to the second layer, which is spanned by $Z_{m+1},\dots,Z_n$. Therefore,

$$\sum_{i,k,j=1}^{m}(X_iY^{jk}f)^2 \leq \sum_{i,k,j=1}^{m}\sum_{r=m+1}^{n}\left(\sigma_r^{jk}\right)^2(X_iZ_rf)^2 \leq c_2\sum_{i=1}^{m}\sum_{r=m+1}^{n}(X_iZ_rf)^2$$

and

$$\sum_{i,k=1}^{m}(Y^{ik}f)^2 \leq \sum_{i,k=1}^{m}\sum_{r=m+1}^{n}\left(\sigma_r^{ik}\right)^2(Z_rf)^2 \leq c_2\sum_{r=m+1}^{n}(Z_rf)^2,$$

with

$$c_2 = \max_{r=m+1,\dots n}\sum_{k,j=1}^{m}\left(\sigma_r^{jk}\right)^2.$$

Inserting these estimates in (5.15) we arrive at

$$\sum_{k=1}^{m}(X_kf)([\mathrm{L},X_k]f) \geq -\frac{c_1}{\varepsilon}\overline{\Gamma}(f) - 2\varepsilon c_2\sum_{i,j,k=1}^{m}\sum_{r=m+1}^{n}(X_iZ_rf)^2 - c_2\frac{\varepsilon}{2}\hat{\Gamma}(f) + \beta\overline{\Gamma}(f),$$

for any $\varepsilon > 0$. We conclude that

$$\overline{\Gamma}_2(f) \geq \sum_{i,k=1}^{m}(X_iX_kf)^2 + \left(\beta - \frac{c_1}{\varepsilon}\right)\overline{\Gamma}(f) - 2\varepsilon c_2\sum_{i=1}^{m}\sum_{k=m+1}^{n}(X_iZ_kf)^2 - \frac{c_2}{2}\varepsilon\hat{\Gamma}(f). \tag{5.16}$$

A similar computation shows that

$$\hat{\Gamma}_2(f) \geq \sum_{i=1}^{m} \sum_{k=m+1}^{n} (X_i Z_k f)^2 + \sum_{k=m+1}^{n} (Z_k f)([\mathrm{L}, Z_k] f).$$

We note that, for $k \in \{m+1, \ldots, n\}$,

$$[\mathrm{L}, Z_k] = -\beta [D, Z_k] = 2\beta Z_k,$$

since Z_k commutes with Z_j for all $j = 1, \ldots, n$. Therefore,

$$\hat{\Gamma}_2(f) \geq \sum_{i=1}^{m} \sum_{k=m+1}^{n} (X_i Z_k f)^2 + 2\beta \hat{\Gamma}(f). \tag{5.17}$$

Adding up (5.16) and (5.17) we obtain that for any $\varepsilon > 0$ and $\lambda > 0$,

$$\begin{aligned} \overline{\Gamma}_2(f) + \lambda \hat{\Gamma}_2(f) \geq \sum_{i,k=1}^{m} (X_i X_k f)^2 + (\lambda - 2\varepsilon c_2) \sum_{i=1}^{m} \sum_{k=m+1}^{n} (X_i Z_k f)^2 \\ + \left(\beta - \frac{c_1}{\varepsilon}\right) \overline{\Gamma}(f) + \left(2\beta\lambda - \frac{c_2}{2}\varepsilon\right) \hat{\Gamma}(f). \end{aligned} \tag{5.18}$$

Moreover, by Young's inequality, for any $\gamma > 0$,

$$\begin{aligned} \sum_{i,k=1}^{m} (X_i X_k f)^2 &= \frac{1}{2} \sum_{i,k=1}^{m} \big((X_i X_k f)^2 + (X_k X_i f)^2 \big) \\ &= \frac{1}{2} \sum_{i,k=1}^{m} \big((X_i X_k f + X_k X_i f)^2 - 2(X_i X_k f)(X_k X_i f) \big) \\ &= \frac{1}{2} \sum_{i,k=1}^{m} \big((2X_i X_k f + [X_k, X_i] f)^2 - 2(X_i X_k f)(X_k X_i f) \big) \\ &\geq \frac{1}{2} \sum_{i,k=1}^{m} \big((2X_i X_k f + Y^{ki} f)^2 - 2(X_i X_k f)^2 \big) \\ &\geq \sum_{i,k=1}^{m} \left((X_i X_k f)^2 + \frac{1}{2}(Y^{ki} f)^2 - \gamma (X_i X_k f)^2 - \frac{1}{\gamma}(Y^{ki} f)^2 \right). \end{aligned}$$

By choosing $\gamma = 4$, say, and rearranging we obtain

$$\sum_{i,k=1}^{m}(X_iX_kf)^2 \geq \frac{1}{16}\sum_{i,k=1}^{m}(Y^{ki}f)^2 \geq \frac{1}{8}\hat{\Gamma}(f), \tag{5.19}$$

where the inequality on the right follows from the facts that $\{Y^{ik}\}_{i,k=1}^{m} \supset \{Z_r\}_{r=m+1}^{n}$ and that for every $r \in \{m+1,\dots,n\}$, the term $(Z_rf)^2$ appears twice in the sum, since, $Z_r = [X_i, X_k] = Y^{ik} = -Y^{ki}$, for some i and k. Therefore, (5.18) implies that for all $\lambda, \varepsilon > 0$,

$$\begin{aligned}\overline{\Gamma}_2(f) + \lambda\hat{\Gamma}_2(f) \geq \frac{1}{2}\sum_{i,k=1}^{m}(X_iX_kf)^2 &+ (\lambda - 2\varepsilon c_2)\sum_{i=1}^{m}\sum_{k=m+1}^{n}(X_iZ_kf)^2 \\ &+ \left(\beta - \frac{c_1}{\varepsilon}\right)\overline{\Gamma}(f) + \left(\frac{1}{16} + 2\beta\lambda - \frac{c_2}{2}\varepsilon\right)\hat{\Gamma}(f).\end{aligned}$$

Finally, we choose $\varepsilon = \lambda\min(1/4, 4\beta)/c_2$ to obtain

$$\overline{\Gamma}_2(f) + \lambda\hat{\Gamma}_2(f) \geq \frac{1}{2}\sum_{i,k=1}^{m}(X_iX_kf)^2 + \frac{\lambda}{2}\sum_{i=1}^{m}\sum_{k=m+1}^{n}(X_iZ_kf)^2 + \left(\beta - \frac{c}{\lambda}\right)\overline{\Gamma}(f) + \frac{1}{16}\hat{\Gamma}(f),$$

with $c = c_1c_2/\min(1/4, 4\beta)$, which is (5.13). This in turn implies

$$\Gamma_2(f) \geq \frac{\min(1,\lambda)}{2\max(1,\lambda)}\sum_{i=1}^{m}\sum_{k=1}^{n}(X_iZ_kf)^2 + \frac{\min\left(\frac{1}{16}, \beta - \frac{c}{\lambda}\right)}{\max(1,\lambda)}\Gamma(f).$$

Finally, by choosing $\lambda > c/\beta$ we can ensure that the coefficient in front of $\Gamma(f)$ is positive. $\square$

The first inequality in the lemma is close to the generalised curvature dimension inequality of [BBBC08, BG11]. We are now in position to prove the main theorem.

Proof of Theorem 5.1. We follow the classical strategy, as outlined for example in [ABC+00]. We aim to show that

$$\partial_s P_{t-s}\Gamma(f_s)^{\frac{q}{2}} \leq -\kappa P_{t-s}\Gamma(f)^{\frac{q}{2}},$$

with $f_s = P_s f$, from which the result will follow by integration. To this end, we note that, since by the diffusion property L satisfies (2.12),

$$\begin{aligned}\partial_s P_{t-s}\Gamma(f_s)^{\frac{q}{2}} &= P_{t-s}\left(-\mathrm{L}\left(\Gamma(f_s)^{\frac{q}{2}}\right) + q\Gamma(f_s)^{\frac{q}{2}-1}\Gamma(f_s, \mathrm{L}f_s)\right) \\ &= P_{t-s}\left(-\frac{q}{2}\frac{\mathrm{L}\Gamma(f_s)}{\Gamma(f_s)^{1-\frac{q}{2}}} + \frac{q}{2}\left(1-\frac{q}{2}\right)\frac{\overline{\Gamma}(\Gamma(f_s))}{\Gamma(f_s)^{2-\frac{q}{2}}} + q\frac{\Gamma(f_s, \mathrm{L}f_s)}{\Gamma(f_s)^{1-\frac{q}{2}}}\right) \\ &= P_{t-s}\left(-q\frac{\Gamma_2(f_s)}{\Gamma(f_s)^{1-\frac{q}{2}}} + \frac{q}{2}\left(1-\frac{q}{2}\right)\frac{\overline{\Gamma}(\Gamma(f_s))}{\Gamma(f_s)^{2-\frac{q}{2}}}\right)\end{aligned}$$

where we made use of the diffusion property of L. To estimate the first term, we use (5.18) with $\lambda = 1$, to obtain

$$\Gamma_2(f_s) \geq \sum_{k=1}^{n}\overline{\Gamma}(Z_k f_s) - 2\varepsilon c_2 \sum_{k=m+1}^{n}\overline{\Gamma}(Z_k f_s) + \left(\beta - \frac{c_1}{\varepsilon}\right)\overline{\Gamma}(f_s) + \left(2\beta - \varepsilon\frac{c_2}{2}\right)\hat{\Gamma}(f).$$

Therefore,

$$-\Gamma_2(f_s) \leq -\sum_{k=1}^{n}\overline{\Gamma}(Z_k f_s) + 2\varepsilon c_2 \sum_{k=m+1}^{n}\overline{\Gamma}(Z_k f_s) - C_\varepsilon\Gamma(f),$$

with

$$C_\varepsilon = \min\left(\beta - \frac{c_1}{\varepsilon}, 2\beta - \varepsilon\frac{c_2}{2}\right).$$

Moreover, by the Cauchy-Schwarz inequality we have

$$\begin{aligned}\overline{\Gamma}(\Gamma(f_s)) &= \sum_{i=1}^{m}\left(X_i\left(\sum_{k=1}^{n}(Z_k f_s)^2\right)\right)^2 = 4\sum_{i=1}^{m}\left(\sum_{k=1}^{n}(Z_k f_s)(X_i Z_k f_s)\right)^2 \\ &\leq 4\left(\sum_{k=1}^{n}(Z_k f_s)^2\right)\left(\sum_{i=1}^{m}\sum_{k=1}^{n}(X_i Z_k f_s)^2\right) = 4\Gamma(f_s)\left(\sum_{i=1}^{m}\sum_{k=1}^{n}(X_i Z_k f_s)^2\right).\end{aligned}$$

Combining the above, we obtain

$$
\begin{aligned}
&\partial_s P_{t-s}\Gamma(f_s)^{\frac{q}{2}} \\
&\leq P_{t-s}\left(q\Gamma(f_s)^{\frac{q}{2}-1}\left((1-q)\sum_{k=1}^{n}\overline{\Gamma}(Z_k f_s) + 2\varepsilon c_2 \sum_{k=m+1}^{n}\overline{\Gamma}(Z_k f_s) - C_\varepsilon \Gamma(f)\right)\right) \\
&\leq P_{t-s}\left(q\Gamma(f_s)^{\frac{q}{2}-1}\left((1-q+2\varepsilon c_2)\sum_{k=1}^{n}\overline{\Gamma}(Z_k f_s) - C_\varepsilon \Gamma(f)\right)\right).
\end{aligned}
$$

Since $q > 1$, we can choose $\varepsilon = \varepsilon_q$ small enough so that

$$1 - q + 2c_2\varepsilon_q \leq 0, \tag{5.20}$$

which implies

$$\partial_s P_{t-s}\Gamma(f_s)^{\frac{q}{2}} \leq -\kappa_q P_{t-s}\Gamma(f_s)^{\frac{q}{2}},$$

with

$$\kappa_q = qC_{\varepsilon_q} = q\min\left(\beta - \frac{c_1}{\varepsilon_q}, 2\beta - \varepsilon_q \frac{c_2}{2}\right).$$

Therefore, if $\beta > \beta_q := \max(c_1/\varepsilon_q, c_2\varepsilon_q/4)$, we have $\kappa_q > 0$. In the particular case where $q = 2$, we can choose $\varepsilon = \varepsilon_2 := 1/2c_2$ so that (5.20) is satisfied, and therefore κ_2 is positive as soon as $\beta > \max(2c_1c_2, 1/8)$. □

We conclude this section by showing that one can extract a subsequence $(P_{t_k})_{k=1}^{\infty}$ which converges weakly to a probability measure ν on $\mathbb{R}^n$, i.e. that for all bounded Lipschitz f

$$P_{t_k} f \to \int f d\nu$$

as $k \to \infty$. We first use a compactness argument (c.f. [DP06]) to prove that the function $t \mapsto P_t W$ is bounded, where $W = 1 + N^2$. Recall that by (5.11) there exist constants $B, r > 0$ such that

$$\mathrm{L}W \leq -\beta W + B\chi_{B_r}.$$

Differentiating P_tW, we obtain

$$\partial_t P_t W = P_t \mathrm{L} W \leq -\beta P_t W + B,$$

and integrating this inequality we arrive at

$$P_t W \leq W \mathrm{e}^{-\beta t} + \frac{B}{\beta}(1 - \mathrm{e}^{-\beta t}).$$

This shows that P_tW is bounded as a function of t, for all $t > 0$. With the help of this observation, we are in position to prove the following result.

Theorem 5.3. *Suppose that $\beta > \beta_2$, where β_2 is as in Theorem 5.12. Then, there exists a sequence $\{t_k\}_{k=1}^{\infty} \subset \mathbb{R}$ and a probability measure ν on $\mathbb{R}^n$ such that for all bounded and Lipschitz f and all $x \in \mathbb{R}^n$*

$$P_{t_k} f(x) \to \int f d\nu$$

as $k \to \infty$. Moreover, ν is invariant for P_t, i.e. for all $t \geq 0$ and f as above,

$$\int P_t f d\nu = \int f d\nu.$$

Proof. Let $w \in \mathbb{R}^n$. To keep track of the dependence on w, we write $P_t f(w) = P_t^w f$. For $\ell > 0$, the sets $\Upsilon_\ell = \{W \leq \ell\}$ are compact. Markov's inequality together with the fact that P_tW is bounded in time, imply

$$P_t^w(\Upsilon_\ell) \geq 1 - \frac{K}{\ell},$$

for some constant $K > 0$ independent of t. Therefore $(P_t^w)_{t>0}$ represents a tight family of measures on $\mathbb{R}^n$ and we deduce from Prokhorov's theorem (see e.g. [DP06]) that there exists a weakly convergent subsequence $P_{t_k}^w \to \lim_{k\to\infty} P_{t_k}^w =: \nu_w$. Furthermore, the family $(\mu_s)_{s>0}$ defined by

$$\mu_s(E) = \frac{1}{s}\int_0^s P_t^w(E)dt$$

for Borel sets E, is tight, since

$$\mu_s(\Upsilon_\ell) \geq 1 - \frac{K}{\ell},$$

so by the Krylov-Bogoliubov theorem (see e.g. [DP06]) we conclude that the measure ν is invariant under P_t^w. To see that ν_w does in fact not depend on w, let $y \in \mathbb{R}^n$, $T = d(w, y)$, and suppose $\gamma_s : [0, T] \to \mathbb{R}^n$ is a path joining w to y such that $|\dot{\gamma}| = 1$. We will show that if f is a bounded and smooth function with compact support then

$$P_{t_k}^y f \to \nu_w f =: \nu f.$$

Let $\varepsilon > 0$. By the triangle inequality,

$$|P_{t_k}^y f - \nu f| \leq |P_{t_k}^w f - \nu f| + |P_{t_k}^w f - P_{t_k}^y f|.$$

There exists k_0 such that for $k > k_0$ we have

$$|P_{t_k}^w f - \nu f| \leq \frac{\varepsilon}{2}.$$

Moreover, since $\overline{\Gamma}(P_{t_k} f) \leq \Gamma(P_{t_k} f)$, we have

$$\begin{aligned} |P_{t_k} f(w) - P_{t_k} f(y)| &\leq \int_0^T \left| \sqrt{\Gamma(P_{t_k} f)(\gamma_s)} \cdot \dot{\gamma}_s \right| ds \\ &\leq T \mathrm{e}^{-\kappa_2 t_k / 2} \sqrt{\|\Gamma(f)\|_\infty}, \end{aligned}$$

with a constant $\kappa_2 > 0$, by using the gradient bound (5.12) with $q = 2$ together with the fact that P_t is contractive. Therefore, there exists k_1 such that for $k > k_1$,

$$t_k \geq -\frac{2}{\kappa_2} \log \frac{\varepsilon}{2T\sqrt{\|\Gamma(f)\|_\infty}}$$

and thus

$$|P_{t_k} f(w) - P_{t_k} f(y)| \leq \frac{\varepsilon}{2}.$$

Combining the above and choosing we conclude that for all $k > \max(k_0, k_1)$,

$$|P^y_{t_k} f - \nu f| \leq \varepsilon.$$

Since ε was arbitrary, the proof is complete. □

5.2 Li-Yau estimates

In this section, we prove a Li-Yau estimate for the operator

$$\mathrm{L} = \Delta - \frac{1}{2}D = \sum_{i=1}^{m} X_i^2 - \frac{1}{2}D.$$

This is a special case of the operator considered in the previous section (defined in (5.3)), when $G_{ij} = 0$, $\alpha_i = 0$ for all i, j and $\beta = 1/2$. The main idea behind the strategy that we use comes from [BL06], where such estimates were proved using semigroup tools in the elliptic setting, as well as [BBBQ09], where these techniques were extended to the sub-Riemannian setting, for L being the sub-Laplacian of the group. Recently, Li-Yau estimates for the sub-Laplacian were proved for a large class of nilpotent groups [BG11].

Let f be a smooth function and suppose that the functions ζ_i are the coefficients in the representation

$$Df(w) = \sum_{i=1}^{n} \zeta_i(w) Z_i f(w). \tag{5.21}$$

From (5.2) we see that D is homogeneous of degree 0, i.e. $D(f(\delta_\lambda(w))) = (Df)(\delta_\lambda(w))$, and therefore its components ζ_i must be homogeneous of degree 1 for $1 \leq i \leq m$ and homogeneous of degree 2 otherwise, i.e

$$\zeta_i(\delta_\lambda(w)) = \begin{cases} \lambda \zeta_i(w), & \text{if } i \in \{1, \dots, m\} \\ \lambda^2 \zeta_i(w), & \text{if } i \in \{m+1, \dots, n\}. \end{cases}$$

It follows by Proposition 1.3.4 of [BLU07] that the ζ_i are either zero or polynomials of the form

$$\zeta_i(w) = \begin{cases} \sum_{l(\alpha)=1} c_\alpha w^\alpha, & \text{if } i \in \{1, \dots, m\} \\ \sum_{l(\alpha)=2} c_\alpha w^\alpha, & \text{if } i \in \{m+1, \dots, n\}, \end{cases}$$

for some constants c_α, where the length of a multi-index $\alpha \in \mathbb{R}^n$ is defined as

$$l(\alpha) = \sum_{i=1}^{m} \alpha_i + 2 \sum_{i=m+1}^{n} \alpha_i.$$

Therefore, for some constants C and p, we have the bound

$$|\zeta|^2 := \sum_i \zeta_i^2 \le C W^p,$$

where we recall that $W = 1 + N^2$. Hence

$$P_t |\zeta|^2 \le C P_t W^p \le C_\zeta, \tag{5.22}$$

for some constant C_ζ and all $t > 0$, where the inequality on the right follows by considerations similar to the ones in the previous section, noting that $\overline{\Gamma}(W) = 4N^2\overline{\Gamma}(N) \le 4W$ and that, by the diffusion property of L,

$$\mathrm{L}(W^p) = pW^{p-1}\mathrm{L}W + p(p-1)W^{p-2}\overline{\Gamma}(W) \le -AW^p + BW^{p-1},$$

for some constants $A, B > 0$.

In order to justify why the choice of the constants $G_{ij} = \alpha_i = 0$ and $\beta = 1/2$ is of particular interest, let us consider the heat semigroup $S_t = \mathrm{e}^{t\Delta}$. Let $t > 0$ and f be a smooth function. For $s \in [0, t]$ consider the function $F(s) = S_s D S_{t-s} f$. We have

$$F'(s) = S_s(\Delta D S_{t-s} f - D \Delta S_{t-s} f) = S_s[\Delta, D] S_{t-s} f.$$

Since $[\Delta, D] = 2\Delta$ and Δ commutes with S_t, we conclude that

$$F'(s) = 2\Delta S_t f.$$

Integrating this over $s \in [0, t]$, we arrive at the commutation relation

$$S_t D f = D S_t f + 2t\Delta S_t f. \tag{5.23}$$

Since $\zeta_i(0) = 0$, evaluating the above at 0 we obtain

$$S_1 \left(\Delta - \frac{1}{2} D\right) f(0) = 0.$$

In other words, if h is the density of S_1 at 0 and $\nu(dw) = h(w)dw$ (as in Chapter 4) the measure ν is invariant for the operator $L = \Delta - D/2$. This observation was already made in [BBBC08] in the Heisenberg group based on the commutation relation (5.23), which remains true in our setup. For a smooth function f, the carré du champ of L is given by

$$\Gamma_{\mathrm{L}}(f) = \sum_{i=1}^{m} (X_i f)^2 = \overline{\Gamma}(f).$$

Let us also note that $\hat{\Gamma}(f, \overline{\Gamma}(f)) = \overline{\Gamma}(f, \hat{\Gamma}(f))$. Indeed, we compute

$$\begin{aligned}
\overline{\Gamma}(f, \hat{\Gamma}(f)) - \hat{\Gamma}(f, \overline{\Gamma}(f)) &= \sum_{i=1}^{m} \sum_{j=m+1}^{n} ((X_i f)(X_i (Z_j f)^2) - (Z_j f)(Z_j (X_i f)^2)) \\
&= 2 \sum_{i=1}^{m} \sum_{j=m+1}^{n} (X_i f)(Z_j f)[X_i, Z_j] f = 0,
\end{aligned}$$

since Z_j belongs to the centre of the Lie algebra.

Our goal is to prove the following estimate.

Proposition 5.4. *For all nonnegative smooth f and all $x \in \mathbb{R}^n$, $t > 0$,*

$$\frac{\overline{\Gamma}(P_t f)}{P_t f} + \frac{t}{24} \frac{\hat{\Gamma}(P_t f)}{P_t f} \leq c_1 \frac{e^t - t - 1}{t^2} \mathrm{L} P_t f + c_2 \frac{e^t - t - 1}{t^2} D P_t f + c_3 \frac{P_t f}{t}, \tag{5.24}$$

where $c_1, c_2, c_3 > 0$ *are absolute constants.*

For the proof we will need the following commutation relation of P_t and D.

Lemma 5.5. *For all smooth functions* f, *for all* $x \in \mathbb{R}^n$ *and all* $t > 0$,

$$P_t D f = \mathrm{e}^t D P_t f + 2(\mathrm{e}^t - 1)\mathrm{L} P_t f. \tag{5.25}$$

Proof. As before, we define $F(s) = P_s D P_{t-s} f$ and we compute

$$F'(s) = P_s[\mathrm{L}, D]P_{t-s} f = P_s(2\mathrm{L} + D)P_{t-s} f = 2\mathrm{L} P_t f + F(s).$$

Integrating this inequality we obtain the result. □

We are now ready to prove (5.24).

Proof of Proposition 5.4. We follow [BBBQ09]. Let $\Phi_1(s) = P_s(u_s \overline{\Gamma}(\log u_s))$ and

$\Phi_2(s) = P_s(u_s\hat{\Gamma}(\log u_s))$, with $u_s = P_{t-s}f$. We start by computing $\Phi_1'(s)$. We have

$$\begin{aligned}
\Phi_1'(s) &= P_s\left(\mathrm{L}\left(u\sum_{k=1}^m (X_k \log u_s)^2\right) - (\mathrm{L}u_s)\sum_{k=1}^m (X_k \log u_s)^2 + u_s\partial_s \sum_{k=1}^m (X_k \log u_s)^2\right) \\
&= P_s\Bigg((\mathrm{L}u_s)\sum_{k=1}^m (X_k \log u_s)^2 + 2\sum_{i=1}^m (X_i u_s)\left(X_i\sum_{k=1}^m (X_k \log u_s)^2\right) \\
&\qquad + u_s\mathrm{L}\sum_{k=1}^m (X_k \log u_s)^2 - (\mathrm{L}u_s)\sum_{k=1}^m (X_k \log u_s)^2 \\
&\qquad - 2u_s\sum_{k=1}^m (X_k \log u_s) X_k \frac{\mathrm{L}u_s}{u_s}\Bigg) \\
&= P_s\Bigg(u_s\mathrm{L}\overline{\Gamma}(\log u_s) + 2u_s\sum_{i=1}^m (X_i \log u_s)\left(X_i\overline{\Gamma}(\log u_s)\right) \\
&\qquad - 2u_s\sum_{k=1}^m (X_k \log u_s)(X_k \mathrm{L}\log u_s) - 2u_s\sum_{k=1}^m (X_k \log u_s)(X_k\overline{\Gamma}(\log u_s))\Bigg) \\
&= P_s\Bigg(u_s\mathrm{L}\overline{\Gamma}(\log u_s) + 2u_s\overline{\Gamma}(\log u_s, \overline{\Gamma}(\log u_s)) \\
&\qquad - 2u_s\overline{\Gamma}(\log u_s, \mathrm{L}\log u_s) - 2u_s\overline{\Gamma}(\log u_s, \overline{\Gamma}(\log u_s))\Bigg) \\
&= 2P_s\left(u_s\overline{\Gamma}_2(\log u_s)\right),
\end{aligned}$$

where we used that, for smooth functions f,

$$\mathrm{L}\log f = \frac{\mathrm{L}f}{f} - \overline{\Gamma}(\log f).$$

Similarly, using that $\hat{\Gamma}(f, \overline{\Gamma}(f)) = \overline{\Gamma}(f, \hat{\Gamma}(f))$, one can show that

$$\Phi_2'(s) = 2P_s(u_s\hat{\Gamma}_2(\log u_s)).$$

Therefore, by (5.13), we have

$$\Phi_1'(s) + \lambda\Phi_2'(s) \geq P_s\left(u_s\sum_{i=1}^m (X_i^2 \log u_s)^2\right) + \left(1 - \frac{2c}{\lambda}\right)\Phi_1(s) + \frac{1}{8}\Phi_2(s),$$

for all $\lambda > 0$. The first term can be further estimated using the inequality $a^2 \geq 2\gamma a - \gamma^2$, which is valid for all $a, \gamma \in \mathbb{R}$:

$$\begin{aligned}
P_s\left(u_s \sum_{i=1}^m (X_i^2 \log u_s)^2\right) &\geq P_s\left(2\gamma u_s \sum_{i=1}^m X_i^2 \log u_s - \gamma^2 u_s\right) \\
&= P_s\left(2\gamma u_s \mathrm{L} \log u_s + \gamma u_s D \log u_s - \gamma^2 u_s\right) \\
&= P_s\left(2\gamma \mathrm{L} u_s - 2\gamma u_s \overline{\Gamma}(\log u_s) + \gamma D u_s - \gamma^2 u_s\right),
\end{aligned}$$

where we used once again that $u_s \mathrm{L} \log u_s = \mathrm{L} u_s - u_s \overline{\Gamma}(\log u_s)$. The commutation relation (5.25) at time s applied to the function u_s becomes

$$P_s D u_s = \mathrm{e}^s D P_t f + 2(\mathrm{e}^s - 1)\mathrm{L} P_t f.$$

Therefore,

$$P_s\left(u_s \sum_{i=1}^m (X_i^2 \log u_s)^2\right) \geq 2\gamma \mathrm{e}^s \mathrm{L} P_t f - 2\gamma \Phi_1(s) + \gamma \mathrm{e}^s D P_t f - \gamma^2 P_t f.$$

We thus arrive at the estimate

$$\Phi_1'(s) \geq 2\gamma \mathrm{e}^s \mathrm{L} P_t f + \gamma \mathrm{e}^s D P_t f - \gamma^2 P_t f - \left(\frac{2c}{\lambda} + 2\gamma\right)\Phi_1(s) + \frac{1}{8}\Phi_2(s) - \lambda \Phi_2'(s).$$

Therefore, if $b : [0, t] \to [0, \infty)$ is a positive twice differentiable decreasing function, we have

$$\begin{aligned}
(-8b'\Phi_1 + b\Phi_2)' &= -8b'\Phi_1' - 8b''\Phi_1 + b'\Phi_2 + b\Phi_2' \\
&\geq -8b'(2\gamma \mathrm{e}^s \mathrm{L} P_t f + \gamma \mathrm{e}^s D P_t f - \gamma^2 P_t f) + 8\left(b'\left(\frac{2c}{\lambda} + 2\gamma\right) - b''\right)\Phi_1 \\
&\quad + (-b' + b')\,\Phi_2 + (8b'\lambda + b)\Phi_2'
\end{aligned}$$

Choosing

$$\lambda = \frac{-b}{8b'} \geq 0, \quad \gamma = \frac{b''}{2b'} + 8c\frac{b'}{b},$$

we finally arrive at

$$
\begin{aligned}
(-8b'\Phi_1 + b\Phi_2)' &\geq -16b'\gamma \mathrm{e}^s \mathrm{L}P_tf - 8b'\gamma \mathrm{e}^s DP_tf + 8b'\gamma^2 P_tf \\
&= -16b'\left(\frac{b''}{2b'} + 8c\frac{b'}{b}\right)\mathrm{e}^s \mathrm{L}P_tf - 8b'\left(\frac{b''}{2b'} + 8c\frac{b'}{b}\right)\mathrm{e}^s DP_tf + 8b'\left(\frac{b''}{2b'} + 8c\frac{b'}{b}\right)^2 P_tf.
\end{aligned}
$$

In the particular case where $b(s) = (t-s)^3$, this reads

$$
\big(24(t-s)^2\Phi_1(s) + (t-s)^3\Phi_2(s)\big)' \geq -k_1(t-s)\mathrm{e}^s \mathrm{L}P_tf - k_2(t-s)\mathrm{e}^s DP_tf - k_3 P_tf,
$$

with $k_1 = 48(2+48c)$, $k_2 = 48(1+24c)$ and $k_3 = -24(2+48c)^2$. Integrating this inequality over $s \in [0,t]$, we obtain

$$
-24t^2\Phi_1(0) - t^3\Phi_2(0) \geq -k_1(\mathrm{e}^t - t - 1)\mathrm{L}P_tf - k_2(\mathrm{e}^t - t - 1)DP_tf - k_3 t P_tf.
$$

Rearranging, we conclude that

$$
\frac{\overline{\Gamma}(P_tf)}{P_tf} + \frac{t}{24}\frac{\hat{\Gamma}(P_tf)}{P_tf} \leq \frac{k_1}{24}\frac{\mathrm{e}^t - t - 1}{t^2}\mathrm{L}P_tf + \frac{k_2}{24}\frac{\mathrm{e}^t - t - 1}{t^2}DP_tf + \frac{k_3}{24}\frac{P_tf}{t},
$$

which is what we wanted to show. $\square$

As a concluding remark, we note that DP_tf can be controlled using Young's inequality as follows

$$
\begin{aligned}
|DP_tf| &= \left|\sum_{j=1}^n \zeta_j Z_j P_tf\right| \\
&\leq \frac{1}{2\delta}\sum_{j=1}^n \zeta_j(x)^2 P_tf + \frac{\delta}{2}\frac{\Gamma(P_tf)}{P_tf}
\end{aligned}
$$

for all $\delta > 0$. Therefore, (5.24) implies that for all $\delta, t > 0$,

$$\left(1 - c_2\frac{\mathrm{e}^t - t - 1}{t^2}\frac{\delta}{2}\right)\frac{\overline{\Gamma}(P_t f)}{P_t f} + \left(\frac{t}{24} - c_2\frac{\mathrm{e}^t - t - 1}{t^2}\frac{\delta}{2}\right)\frac{\hat{\Gamma}(P_t f)}{P_t f}$$
$$\leq c_1\frac{\mathrm{e}^t - t - 1}{t^2}\mathrm{L}P_t f + \left(\frac{c_2|\zeta|^2}{2\delta}\frac{\mathrm{e}^t - t - 1}{t^2} + \frac{c_3}{t}\right)P_t f.$$

If $t < 24$, we may choose
$$\delta = \frac{1}{c_2}\frac{t}{24}\frac{t^2}{\mathrm{e}^t - t - 1}$$
so that the left-hand-side is nonnegative and the above inequality yields

$$0 \leq c_1\frac{\mathrm{e}^t - t - 1}{t^2}\mathrm{L}P_t f + \left(\frac{24c_2^2|\zeta|^2}{2t}\left(\frac{\mathrm{e}^t - t - 1}{t^2}\right)^2 + \frac{c_3}{t}\right)P_t f$$

If $t > 24$, we can choose
$$\delta = \frac{1}{c_2}\frac{t^2}{\mathrm{e}^t - t - 1}$$
to obtain

$$0 \leq c_1\frac{\mathrm{e}^t - t - 1}{t^2}\mathrm{L}P_t f + \left(\frac{c_2^2|\zeta|^2}{2}\left(\frac{\mathrm{e}^t - t - 1}{t^2}\right)^2 + \frac{c_3}{t}\right)P_t f.$$

Combining the above, we arrive at

$$0 \leq c_1\frac{\mathrm{e}^t - t - 1}{t^2}\mathrm{L}P_t f + \left(\frac{c_2^2|\zeta|^2}{2}\left(1 + \frac{24}{t}\right)\left(\frac{\mathrm{e}^t - t - 1}{t^2}\right)^2 + \frac{c_3}{t}\right)P_t f,$$

for $t > 0$. Since $\mathrm{e}^t - t - 1 = O(t^2)$ as $t \to 0$, for any $T > 0$ there exists a constant C_T such that for all $t \leq T$,
$$1 \leq \frac{\mathrm{e}^t - t - 1}{t^2} \leq C_T.$$
We conclude that for some $K > 0$, for all $T > 0$ and all $t \leq T$

$$(\mathrm{L}P_t f)^- \leq \left(K + \frac{\tilde{C}_T}{t}\right)P_t f \leq \frac{KT + \tilde{C}_T}{t}P_t f,$$

where $\tilde{C}_T$ is a constant depending on T and as usual, we denote the positive and negative parts of a function f by $f^+ = \max(f,0)$ and $f^- = -\min(f,0)$. Following [Led94], since $0 = \int \mathrm{L}P_t f d\nu = \int (\mathrm{L}P_t f)^+ d\nu - \int (\mathrm{L}P_t f)^- d\nu$ and $\int P_t f d\nu = \int f d\nu$, this implies

$$\int |\mathrm{L}P_t f| d\nu = 2\int (\mathrm{L}P_t f)^- d\nu \leq 2\frac{KT + \tilde{C}_T}{t}\int f d\nu.$$

If, in addition, we know that P_t is symmetric in $L^2(\nu)$, such an estimate can be used to derive isoperimetric results for ν [Led94].

References

[ABC$^+$00] 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, With a preface by D. Bakry and M. Ledoux. MR MR1845806 (2002g:46132)

[Bak94] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, Lectures on probability theory (Saint-Flour, 1992), Lecture Notes in Math., vol. 1581, Springer, Berlin, 1994, pp. 1–114. MR MR1307413 (95m:47075)

[Bak06] ______, *Functional inequalities for Markov semigroups*, Probability measures on groups: recent directions and trends, Tata Inst. Fund. Res., Mumbai, 2006, pp. 91–147. MR 2213477 (2007g:60086)

[Bar01] F. Barthe, *Levels of concentration between exponential and Gaussian*, Ann. Fac. Sci. Toulouse Math. (6) **10** (2001), no. 3, 393–404. MR 1923685 (2003f:60038)

[Bar02] ______, *Log-concave and spherical models in isoperimetry*, Geom. Funct. Anal. **12** (2002), no. 1, 32–55. MR 1904555 (2003d:28017)

[Bay04] V. Bayle, *Propriétés de concavité du profil isopérimétrique et applications*, PhD Thesis, Institut Joseph Fourier, Grenoble, 2004.

[BB06] G. Blower and F. Bolley, *Concentration of measure on product spaces with applications to Markov processes*, Studia Math. **175** (2006), no. 1, 47–72. MR 2261699 (2007k:60055)

[BB09] F. Baudoin and M. Bonnefont, *The subelliptic heat kernel on* SU(2)*: representations, asymptotics and gradient bounds*, Math. Z. **263** (2009), no. 3, 647–672. MR 2545862 (2011d:58060)

[BBBC08] D. Bakry, F. Baudoin, M. Bonnefont, and D. Chafaï, *On gradient bounds for the heat kernel on the Heisenberg group*, J. Funct. Anal. **255** (2008), no. 8, 1905–1938. MR MR2462581

[BBBQ09] D. Bakry, F. Baudoin, M. Bonnefont, and B. Qian, *Subelliptic Li-Yau estimates on three dimensional model spaces*, Potential theory and stochastics in Albac, Theta Ser. Adv. Math., vol. 11, Theta, Bucharest, 2009, pp. 1–10. MR 2681833

[BBCG08] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin, *A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case*, Electron. Commun. Probab. **13** (2008), 60–66. MR 2386063 (2009d:60039)

[BCR06] F. Barthe, P. Cattiaux, and C. Roberto, *Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry*, Rev. Mat. Iberoam. **22** (2006), no. 3, 993–1067. MR 2320410 (2008k:47097)

[BCR07] ———, *Isoperimetry between exponential and Gaussian*, Electron. J. Probab. **12** (2007), no. 44, 1212–1237 (electronic). MR MR2346509 (2008j:60046)

[BD] M. Bardi and F. Dragoni, *Convexity and semiconvexity along vector fields*, preprint.

[BÉ85] D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206. MR MR889476 (88j:60131)

[BEHM09] Z. M. Balogh, A. Engoulatov, L. Hunziker, and O. E. Maasalo, *Functional inequalities and Hamilton-Jacobi equations in geodesic spaces*, arXiv:0906.0476v1 [math.FA] (2009).

[BG99] S. G. Bobkov and F. Götze, *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities*, J. Funct. Anal. **163** (1999), no. 1, 1–28. MR 1682772 (2000b:46059)

[BG11] F. Baudoin and N. Garofalo, *Curvature-dimension inequalities and ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, arXiv:1101.3590v1 [math.DG] (2011).

[BGL01] S. G. Bobkov, I. Gentil, and M. Ledoux, *Hypercontractivity of Hamilton-Jacobi equations*, J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696. MR MR1846020 (2003b:47073)

[BH96] S. G. Bobkov and C. Houdré, *Characterization of Gaussian measures in terms of the isoperimetric property of half-spaces*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **228** (1996), no. Veroyatn. i Stat. 1, 31–38, 356. MR 1449845 (98e:60056)

[BH97a] ———, *Isoperimetric constants for product probability measures*, Ann. Probab. **25** (1997), no. 1, 184–205. MR MR1428505 (98g:60032)

[BH97b] ———, *Some connections between isoperimetric and Sobolev-type inequalities*, Mem. Amer. Math. Soc. **129** (1997), no. 616, viii+111. MR MR1396954 (98b:46038)

[BK08] F. Barthe and A. V. Kolesnikov, *Mass transport and variants of the logarithmic Sobolev inequality*, J. Geom. Anal. **18** (2008), no. 4, 921–979. MR 2438906 (2009i:58048)

[BL96] D. Bakry and M. Ledoux, *Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator*, Invent. Math. **123** (1996), no. 2, 259–281. MR MR1374200 (97c:58162)

[BL06] ———, *A logarithmic Sobolev form of the Li-Yau parabolic inequality*, Rev. Mat. Iberoam. **22** (2006), no. 2, 683–702. MR 2294794 (2008m:58051)

[Blo03] G. Blower, *The Gaussian isoperimetric inequality and transportation*, Positivity **7** (2003), no. 3, 203–224. MR 2018596 (2005e:60042)

[BLU07] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007. MR MR2363343

[BM00] F. Barthe and B. Maurey, *Some remarks on isoperimetry of Gaussian type*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 4, 419–434. MR MR1785389 (2001k:60055)

[Bob96a] S.G. Bobkov, *Extremal properties of half-spaces for log-concave distributions*, Ann. Probab. **24** (1996), no. 1, 35–48. MR MR1387625 (97e:60027)

[Bob96b] ———, *A functional form of the isoperimetric inequality for the Gaussian measure*, J. Funct. Anal. **135** (1996), no. 1, 39–49. MR MR1367623 (96m:60091)

[Bob97a] S. G. Bobkov, *An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space*, Ann. Probab. **25** (1997), no. 1, 206–214. MR MR1428506 (98g:60033)

[Bob97b] S.G. Bobkov, *Isoperimetric problems in the theory of infinite dimensional probability distributions*, PhD Thesis, St-Petersburg University, 1997.

[Bob99] S. G. Bobkov, *Isoperimetric and analytic inequalities for log-concave probability measures*, Ann. Probab. **27** (1999), no. 4, 1903–1921. MR 1742893 (2001h:60026)

[Bob02] ———, *A localized proof of the isoperimetric Bakry-Ledoux inequality and some applications*, Teor. Veroyatnost. i Primenen. **47** (2002), no. 2, 340–346. MR MR2001838 (2004j:60038)

[Bob09] ———, *On the isoperimetric constants for product measures*, J. Math. Sci. (N. Y.) **159** (2009), no. 1, 47–53, Problems in mathematical analysis. No. 40. MR 2544042 (2011b:60012)

[Bon09] M. Bonnefont, *Inégalités fonctionnelles pour des noyaux de la chaleur sous-elliptiques*, PhD Thesis, Université Paul-Sabatier, Toulouse III, 2009.

[Bor74] C. Borell, *Convex measures on locally convex spaces*, Ark. Mat. **12** (1974), 239–252.

[Bus82] P. Buser, *A note on the isoperimetric constant*, Ann. Sci. École Norm. Sup. (4) **15** (1982), no. 2, 213–230. MR 683635 (84e:58076)

[BZ05] S. G. Bobkov and B. Zegarliński, *Entropy bounds and isoperimetry*, Mem. Amer. Math. Soc. **176** (2005), no. 829, x+69. MR MR2146071 (2006c:46027)

[CG90] L. Corwin and F. P. Greenleaf, *Representation of nilpotent lie groups and their applications. part i: Basic theory and examples*, Cambridge studies in advanced Mathematics, vol. 18, Cambridge University Press, Cambridge, 1990.

[CG06] P. Cattiaux and A. Guillin, *On quadratic transportation cost inequalities*, J. Math. Pures Appl. (9) **86** (2006), no. 4, 341–361. MR 2257848 (2007g:47066)

[CG10] ———, *Functional inequalities via Lyapunov conditions*, arXiv:1001.1822v2 [math.PR] (2010).

[CGGR10] P. Cattiaux, N. Gozlan, A. Guillin, and C. Roberto, *Functional inequalities for heavy tailed distributions and application to isoperimetry*, Electron. J. Probab. **15** (2010), no. 13, 346–385. MR 2609591 (2011f:60036)

[CGW10] P. Cattiaux, A. Guillin, and L. Wu, *A note on Talagrand's transportation inequality and logarithmic Sobolev inequality*, Probab. Theory Related Fields **148** (2010), no. 1-2, 285–304. MR 2653230

[CGWW09] P. Cattiaux, A. Guillin, F.-Y. Wang, and L. Wu, *Lyapunov conditions for super Poincaré inequalities*, J. Funct. Anal. **256** (2009), no. 6, 1821–1841. MR 2498560

[Cha04] D. Chafaï, *Entropies, convexity, and functional inequalities: on Φ-entropies and Φ-Sobolev inequalities*, J. Math. Kyoto Univ. **44** (2004), no. 2, 325–363. MR 2081075 (2005e:60170)

[Che70] J. Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian*, Problems in analysis (Papers dedicated to Salomon Bochner, 1969), Princeton Univ. Press, Princeton, N. J., 1970, pp. 195–199. MR 0402831 (53 #6645)

[CLN06] B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci flow*, Graduate Studies in Mathematics, vol. 77, American Mathematical Society, Providence, RI, 2006. MR 2274812 (2008a:53068)

[DGN03] D. Danielli, N. Garofalo, and D.-M. Nhieu, *Notions of convexity in Carnot groups*, Comm. Anal. Geom. **11** (2003), no. 2, 263–341. MR 2014879 (2004m:22014)

[DKZ10] F. Dragoni, V. Kontis, and B. Zegarliński, *Ergodicity of Markov semigroups with Hörmander type generators in infinite dimensions*, preprint (2010).

[DM05] B. K. Driver and T. Melcher, *Hypoelliptic heat kernel inequalities on the Heisenberg group*, J. Funct. Anal. **221** (2005), no. 2, 340–365. MR MR2124868 (2005k:58044)

[DP06] G. Da Prato, *An introduction to infinite-dimensional analysis*, Universitext, Springer-Verlag, Berlin, 2006. MR 2244975 (2009a:46001)

[Eld09] N. Eldredge, *Precise estimates for the subelliptic heat kernel on H-type groups*, J. Math. Pures Appl. (9) **92** (2009), no. 1, 52–85. MR 2541147 (2010h:35399)

[Eld10] ______, *Gradient estimates for the subelliptic heat kernel on H-type groups*, J. Funct. Anal. **258** (2010), no. 2, 504–533. MR 2557945 (2011d:35217)

[Eva10] L. C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR 2597943 (2011c:35002)

[Fed69] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[Fou00] P. Fougères, *Hypercontractivité et isopérimétrie gaussienne. Applications aux systèmes de spins*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 5, 647–689. MR MR1792659 (2001m:82004)

[FRZ07] P. Fougères, C. Roberto, and B. Zegarliński, *Sub-Gaussian measures and associated semilinear problems*, preprint (2007).

[FS82] G. B. Folland and E. M. Stein, *Hardy spaces on homogeneous groups*, Mathematical Notes, vol. 28, Princeton University Press, Princeton, N.J., 1982. MR 657581 (84h:43027)

[GN96] N. Garofalo and D.-M. Nhieu, *Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces*, Comm. Pure Appl. Math. **49** (1996), no. 10, 1081–1144. MR 1404326 (97i:58032)

[Gro75] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR MR0420249 (54 #8263)

[GT10] N. Garofalo and J. Tyson, *Riesz potentials and p-superharmonic functions in Lie groups of Heisenberg type*, arXiv:1005.4090v1 [math.AP] (2010).

[GZ03] A. Guionnet and B. Zegarliński, *Lectures on logarithmic Sobolev inequalities*, Séminaire de Probabilités, XXXVI, Lecture Notes in Math., vol. 1801, Springer, Berlin, 2003, pp. 1–134. MR 1971582 (2004b:60226)

[HL08] A. M. Hansson and A. Laptev, *Sharp spectral inequalities for the Heisenberg Laplacian*, Groups and analysis, London Math. Soc. Lecture Note Ser., vol. 354, Cambridge Univ. Press, Cambridge, 2008, pp. 100–115. MR 2528463 (2010k:35075)

[Hör67] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR 0222474 (36 #5526)

[Hue09] N. Huet, *Isoperimetry for spherically symmetric log-concave probability measures*, preprint (2009).

[HZ10] W. Hebisch and B. Zegarliński, *Coercive inequalities on metric measure spaces*, J. Funct. Anal. **258** (2010), no. 3, 814–851. MR 2558178 (2010m:46047)

[IKZ11] J. Inglis, V. Kontis, and B. Zegarliński, *From U-bounds to isoperimetry with applications to H-type groups*, J. Funct. Anal. **260** (2011), no. 1, 76–116. MR 2733571

[IP09] J. Inglis and I. Papageorgiou, *Logarithmic Sobolev inequalities for infinite dimensional Hörmander type generators on the Heisenberg group*, Potential Anal. **31** (2009), no. 1, 79–102. MR 2507447 (2010d:58035)

[Jer86] D. Jerison, *The Poincaré inequality for vector fields satisfying Hörmander's condition*, Duke Math. J. **53** (1986), no. 2, 503–523. MR 850547 (87i:35027)

[JLMS07] P. Juutinen, G. Lu, J.J. Manfredi, and B. Stroffolini, *Convex functions on Carnot groups*, Rev. Mat. Iberoam. **23** (2007), no. 1, 191–200. MR 2351130 (2008m:49165)

[Kap80] A. Kaplan, *Fundamental solutions for a class of hypoelliptic PDE generated by composition of quadratic forms*, Trans. Amer. Math. Soc. **258** (1980), no. 1, 147–153. MR 554324 (81c:58059)

[Kas82] A. Kasue, *A Laplacian comparison theorem and function theoretic properties of a complete Riemannian manifold*, Japan. J. Math. (N.S.) **8** (1982), no. 2, 309–341. MR 722530 (85h:53031)

[KLS95] R. Kannan, L. Lovász, and M. Simonovits, *Isoperimetric problems for convex bodies and a localization lemma*, Discrete Comput. Geom. **13** (1995), no. 3-4, 541–559. MR MR1318794 (96e:52018)

[KS85] S. Kusuoka and D. Stroock, *Some boundedness properties of certain stationary diffusion semigroups*, J. Funct. Anal. **60** (1985), no. 2, 243–264. MR 777238 (87a:60083)

[Led88] M. Ledoux, *Isopérimétrie et inégalités de Sobolev logarithmiques gaussiennes*, C. R. Acad. Sci. Paris Sér. I Math. **306** (1988), no. 2, 79–82. MR 929114 (88m:60013)

[Led94] ———, *A simple analytic proof of an inequality by P. Buser*, Proc. Amer. Math. Soc. **121** (1994), no. 3, 951–959. MR 1186991 (94i:53041)

[Led00] ———, *The geometry of Markov diffusion generators*, Ann. Fac. Sci. Toulouse Math. (6) **9** (2000), no. 2, 305–366, Probability theory. MR MR1813804 (2002a:58045)

[Led01] ———, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001. MR 1849347 (2003k:28019)

[Led09] ———, *From concentration to isoperimetry: semigroup proofs*, preprint (2009).

[Li06] H.Q. Li, *Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg*, J. Funct. Anal. **236** (2006), no. 2, 369–394. MR MR2240167 (2007d:58045)

[LP10] F. Lust-Piquard, *Ornstein-Uhlenbeck semi-groups on stratified groups*, J. Funct. Anal. **258** (2010), no. 6, 1883–1908. MR 2578458

[LV07] J. Lott and C. Villani, *Hamilton-Jacobi semigroup on length spaces and applications*, J. Math. Pures Appl. (9) **88** (2007), no. 3, 219–229. MR 2355455 (2008i:53053)

[LY86] P. Li and S. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** (1986), no. 3-4, 153–201. MR 834612 (87f:58156)

[ŁZ07] P. Ługiewicz and B. Zegarliński, *Coercive inequalities for Hörmander type generators in infinite dimensions*, J. Funct. Anal. **247** (2007), no. 2, 438–476. MR 2323442 (2008e:35025)

[Mel04] T. Melcher, *Hypoelliptic heat kernel inequalities on lie groups*, PhD thesis, University of California, San Diego, 2004.

[Mét80] G. Métivier, *Hypoellipticité analytique sur des groupes nilpotents de rang* 2, Duke Math. J. **47** (1980), no. 1, 195–221. MR 563376 (81i:35037)

[Mil08] E. Milman, *Uniform tail-decay of Lipschitz functions implies Cheeger's isoperimetric inquality under convexity assumptions*, C. R. Math. Acad. Sci. Paris **346** (2008), no. 17-18, 989–994. MR 2449642 (2009h:58069)

[Mil09a] ———, *On the role of convexity in functional and isoperimetric inequalities*, Proc. Lond. Math. Soc. (3) **99** (2009), no. 1, 32–66. MR 2520350 (2011f:32024)

[Mil09b] ______, *On the role of convexity in isoperimetry, spectral gap and concentration*, Invent. Math. **177** (2009), no. 1, 1–43. MR 2507637 (2010j:28004)

[MJ00] F. Morgan and D. L. Johnson, *Some sharp isoperimetric theorems for Riemannian manifolds*, Indiana Univ. Math. J. **49** (2000), no. 3, 1017–1041. MR 1803220 (2002e:53043)

[Nua06] D. Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR 2200233 (2006j:60004)

[OV00] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR 1760620 (2001k:58076)

[Pap10] I. Papageorgiou, *A note on the modified log-Sobolev inequality*, arXiv:1005.4946v2 [math.FA] (2010).

[Ran96] J. Randall, *The heat kernel for generalized heisenberg groups*, J. Geom. Anal. **6(2)** (1996), 287–316.

[Rot85] O. S. Rothaus, *Analytic inequalities, isoperimetric inequalities and logarithmic Sobolev inequalities*, J. Funct. Anal. **64** (1985), no. 2, 296–313. MR 812396 (87f:58181)

[RS76] L. P. Rothschild and E. M. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math. **137** (1976), no. 3-4, 247–320. MR 0436223 (55 #9171)

[RZ07] C. Roberto and B. Zegarliński, *Orlicz-Sobolev inequalities for sub-Gaussian measures and ergodicity of Markov semi-groups*, J. Funct. Anal. **243** (2007), no. 1, 28–66. MR 2289793 (2008g:52015)

[SC02] L. Saloff-Coste, *Aspects of Sobolev-type inequalities*, London Mathematical Society Lecture Note Series, vol. 289, Cambridge University Press, Cambridge, 2002. MR 1872526 (2003c:46048)

[ST74] V. N. Sudakov and B. S. Tsirelson, *Extremal properties of half-spaces for spherically invariant measures*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **41** (1974), 14–24, 165, Problems in the theory of probability distributions, II. MR MR0365680 (51 #1932)

[Tal96] M. Talagrand, *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), no. 3, 587–600. MR 1392331 (97d:60029)

[Vil03] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR 1964483 (2004e:90003)

[VSCC92] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992. MR 1218884 (95f:43008)

[Wan97] F.-Y. Wang, *Logarithmic Sobolev inequalities on noncompact Riemannian manifolds*, Probab. Theory Related Fields **109** (1997), no. 3, 417–424. MR 1481127 (98i:58253)

[Wan06] ———, *Functional inequalities, markov semigroups and spectral theory*, Elsevier Science, 2006.

[Wan09] ———, *Nash and log-Sobolev inequalities for hypoelliptic operators*, Manuscripta Math. **128** (2009), no. 3, 343–358. MR 2481049 (2010b:58045)

[Yos95] K. Yosida, *Functional analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the sixth (1980) edition. MR 1336382 (96a:46001)

[Zeg01] B. Zegarliński, *Isoperimetry for Gibbs measures*, Ann. Probab. **29** (2001), no. 2, 802–819. MR MR1849178 (2002j:26012)