Coercive Inequalities for Generators of Hörmander Type

A thesis presented for the degree of

Doctor of Philosophy of the University of London

and the

Diploma of Imperial College

by

James Devear Inglis

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2AZ

NOVEMBER 28, 2010

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.

Signed:

Copyright

To my parents, Stephen and Moira.

Abstract

This thesis investigates coercive inequalities, such as the logarithmic Sobolev and spectral gap inequalities, for generators defined as the sum of squares of degenerate and non-commuting vector fields (such generators are said to be of Hörmander type). Situations in which the sum is both finite and infinite are considered. Particular attention is paid to the setting of H-type groups, which are naturally equipped with such generators and an associated sub-Riemannian geometry. The bulk of the monograph consists of three self-contained but strongly related projects. In the first of these projects the spectral properties of some Hörmander-type generators on H-type groups are examined via coercive inequalities. In another direction, it is shown that certain non-trivial Gibbs measures with quadratic interaction potentials on an infinite product of H-type groups satisfy logarithmic Sobolev inequalities. The thesis concludes with a study of the ergodicity of an infinite particle system described by a highly degenerate generator, in which coercive inequalities again play a role.

Acknowledgements

I would like to thank first and foremost my supervisor, Prof. Bogusław Zegarliński for his patience, kindness and the gift of his knowledge. I have learned an incredible amount under his guidance.

I thank all the members of the Imperial analysis group for providing a stimulating atmosphere in which to work during my time here. In particular I thank my colleagues Ioannis Papageorgiou, Vasilis Kontis and Misha Neklyudov for their good ideas, support and co-operation during our collaborations. I also thank Dr. Jens Wirth, Dr. Piotr Ługiewicz, Dr. Thomas Sørensen, Prof. W. Hebisch and Prof. Ari Laptev for the many interesting discussions I have been lucky enough to have had with them.

I am grateful to Prof. Laurent Saloff-Coste, who was kind enough to host me in Cornell, and to point me in intriguing directions.

I would also like to thank Dr. Glenys Luke and Mr. Peter Thomas who inspired me and taught me how to enjoy mathematics.

Finally I thank my parents, family, and friends, for their unerring support and encouragement through the highs as well as the lows. I could not have done this without you.

James D. Inglis

Table of contents

List of Publications

Part of the research presented in this thesis can also be found in the following publications:

[80] J. Inglis and I. Papageorgiou, *Logarithmic Sobolev inequalities for infinite dimensional Hörmander type generators on the Heisenberg group*. Journal of Potential Analysis, **31** (2009) pp. 79-102.
Abstract:
"The Heisenberg group is one of the simplest sub-Riemannian settings in which we can define non-elliptic Hörmander type generators. We can then consider coercive inequalities associated to such generators. We prove that certain non-trivial Gibbs measures with quadratic interaction potentials on an infinite product of Heisenberg groups satisfy logarithmic Sobolev inequalities. "
This can be downloaded from:
`http://www.springerlink.com/content/0926-2601`.

[79] J. Inglis, M. Neklyudov and B. Zegarliński, *Liggett-type inequalities and interacting particle systems: the Gaussian case*. To appear in the proceedings of the 7th International Society for Analysis, its Applications and Computation (ISAAC) Congress, 2010.
Abstract:
"We describe Liggett-type inequalities for certain degenerate infinite dimensional sub-elliptic generators and obtain estimates on the long-time behaviour of the corresponding Markov semigroups."

[76] J. Inglis, *Operators on the Heisenberg group with discrete spectra*. To appear in the proceedings of the 7th ISAAC Congress, 2010.

Abstract:
"We show that a certain class of hypoelliptic operators on the Heisenberg group have discrete spectra, using both a spectral representation of the Heisenberg Laplacian and methods based on functional inequalities."

[77] J. Inglis, V. Kontis and B. Zegarliński, *From U-bounds to isoperimetry with applications to H-type groups*. To appear in the Journal of Functional Analysis.
Abstract:
"In this paper we study U-bounds in relation to L_1-type coercive inequalities and isoperimetric problems for a class of probability measures on a general metric space $(\mathbb{R}^N, d)$. We prove the equivalence of an isoperimetric inequality with several other coercive inequalities in this general framework. The usefulness of our approach is illustrated by an application to the setting of H-type groups, and an extension to infinite dimensions."
A preprint of this can be downloaded from:
`http://arxiv.org/abs/0912.0236`.

The following paper has been submitted:

[78] J. Inglis, M. Neklyudov and B. Zegarliński, *Ergodicity for infinite particle systems with locally conserved quantities*.
Abstract:
"We analyse certain degenerate infinite dimensional sub-elliptic generators, and obtain estimates on the long-time behaviour of the corresponding Markov semigroups that describe a certain model of heat conduction. In particular, we establish ergodicity of the system for a family of invariant measures, and show that the optimal rate of convergence to equilibrium is polynomial. Consequently, there is no spectral gap, but a Liggett-Nash-type inequality is shown to hold."
A preprint of this can be downloaded from:
`http://arxiv.org/abs/1002.0282`.

Chapter 1

Introduction

The central idea of this thesis is to investigate the behaviour of certain classes of Markov generators which take the following form:

$$\mathcal{L} = \sum_{i \in \mathcal{R}} X_i^2 \tag{1.1}$$

where $\{X_i : i \in \mathcal{R}\}$ is a given family of degenerate and non-commuting vector fields, and $\mathcal{R}$ is either a finite or infinite (but countable) index set. Such generators are of Hörmander type. This description is a rather broad one, since in fact several more specific problems are dealt with which fit into this general framework. The behaviours we are interested in will be primarily expressed in the form of coercive functional inequalities. Loosely, these are "forcing" inequalities which, when satisfied, necessitate that the generator and associated semigroup behave in a certain way. We will be particularly interested in the so-called logarithmic Sobolev and spectral gap inequalities, which have been extensively studied over the past 30 years (see Chapter 2 for a brief review of this body of work). For generators given by (1.1), establishing these inequalities pose interesting problems, since the degeneracy severely restricts the methods available.

One of the simplest settings in which families of non-commuting vector fields naturally occur is that of an H-type group. It is for this reason, combined with the fact that such groups have attracted a lot of attention recently (see [9, 51, 52, 53, 69, 91, 99]), that we concentrate, at least to start with, on this setting. On an H-type group, the canonical Laplacian operator (called the sub-Laplacian) takes the form (1.1) where $\mathcal{R}$ is a finite set and $|\mathcal{R}|$ is strictly less than the dimension of the space. The sub-Laplacian is therefore not elliptic, but it is *hypoelliptic* by Hörmander's celebrated result i.e. the associated heat kernel is smooth. This is because the set of fields $\{X_i : i \in \mathcal{R}\}$ together with all the commutators span the tangent space at every point.

In the work that follows, generators on H-type groups are explored in both finite and infinite dimensional set-ups. In finite dimensions, coercive inequalities are used to gain information about the spectra of the associated generators. More specifically it is shown that certain generators of Hörmander-type on H-type groups have a spectral gap, and in some cases have entirely discrete spectra. For an infinite dimensional environment, we consider an infinite product of H-type groups. In this context we again study coercive inequalities for Hörmander-type generators and their associated symmetric measures, which are now defined on an infinite dimensional space.

An alternative infinite dimensional setting is introduced in the final strand of this monograph, where $\mathcal{R}$ in (1.1) is taken to be the lattice $\mathbb{Z}^D$, and the family of vector fields X_i to be even more degenerate than in the case of H-type groups, so that not even Hörmander's condition is satisfied. We analyse in detail a situation when the vector fields are specifically given, and use coercive inequalities to show that the associated interacting particle system is ergodic with an explicit rate of convergence.

The majority of the author's own work is contained in Chapters 4, 5 and 6. Chapter 6 can be thought of as a stand-alone chapter, and it is for this reason that we include very little directly relevant background material in Chapters 2 and 3 for the work presented there, preferring instead to include an expanded introduction at the beginning of Chapter 6 and an appendix.

The outline of this thesis is thus as follows. In Chapter 2 we review the literature surrounding the subject area dealt with in Chapters 3–5, which has provided both the motivation and the inspiration for the work described there. The necessary notation, definitions and basic results are then set out in Chapter 3.

Chapter 4 is concerned with proving results about the spectra of certain operators on H-type groups. Two approaches are taken — the first one uses a unitary representation of the sub-Laplacian in the special case of the Heisenberg group, while the second more general one proceeds through functional inequalities. The chapter finishes with an investigation into an interesting class of generator which does not fit into the preceding framework, but for which we are still able to prove the existence of a spectral gap.

The main result of Chapter 5 is that certain infinite dimensional Gibbs measures with unbounded quadratic interaction potentials on an infinite product of H-type groups satisfy logarithmic Sobolev inequalities. The necessary finite dimensional material is prepared in Section 5.3, before the passage to infinity is described in Section 5.4. An alternative interaction potential is considered in the final section. The ideas of this chapter formed part of a joint work with I. Papageorgiou.

Finally, in Chapter 6 we deal with the specific situation when the generator $\mathcal{L}$ is given by

$$\mathcal{L} = \sum_{\substack{i,j \in \mathbb{Z}^D \\ i \sim j}} X_{i,j}^2$$

where the sum is taken over all nearest neighbours $i \sim j$ in the lattice $\mathbb{Z}^D$, and

$$X_{i,j}^2 = (\partial_i V \partial_j - \partial_j V \partial_i)^2 ,$$

with $\partial_i V$ indicating some linear coefficients. Such generators are interesting, since they appear in physical models of heat conduction and are highly degenerate. Moreover, it can be shown that they do not have a spectral gap. Despite this fact we prove that the system is still ergodic, with polynomial rate of convergence, via some coercive inequalities. This was part of a joint work with M. Neklyudov and B. Zegarliński.

Chapter 2

Background for Chapters 3–5

The purpose of this chapter is to put the contents of Chapters 3–5 into context, by giving an overview of the related work that appears in the literature. It is important to view the results stated in the ensuing chapters in this context for two reasons: firstly because it makes the actual contribution made in this thesis more apparent, and secondly because the background material provides the essential motivation for the investigations we engage in. With this in mind, we offer here a discussion of the literature without any precise definitions or statements of results in the name of clarity and readability, leaving the necessary formalities until Chapter 3.

Due to the volume of work that has been carried out, we cannot hope for completeness in this overview. Instead we aim to describe key results that have had a direct influence on the work that follows.

A central concept will be the logarithmic Sobolev inequality, so we begin in Section 2.1 with a discussion of the origins of this inequality and the development of sufficient criteria for it to hold. In the next section (Section 2.2) we describe generalisations and related inequalities that have been well studied, together with some applications to areas such as isoperimetry and spectral theory. Section 2.3 attends to the recent trend of investigating these inequalities in sub-elliptic settings, while in the final section of this chapter we recount the use of logarithmic Sobolev inequalities in the study of spin systems and statistical mechanics, leading to efforts to prove that they hold in various infinite dimensional settings.

2.1 The logarithmic Sobolev inequality and the curvature condition

In his seminal work [64] of 1975, L. Gross showed that, for the Gaussian measure μ on $\mathbb{R}^n$, it holds that

$$\int_{\mathbb{R}^n} f^2(x)\log f^2(x)d\mu(x) \leq 2\int_{\mathbb{R}^n} |\nabla f(x)|^2 d\mu(x) + \|f\|_2^2 \log \|f\|_2^2 \tag{2.1}$$

where $\|f\|_p$ denotes the $L^p(\mu)$ norm of f, and ∇ is the standard gradient on $\mathbb{R}^n$. Inequality (2.1) asserts that $\int f^2 \log f^2 d\mu$ is finite whenever f and ∇f are in $L^2(\mu)$, and was designated a *logarithmic Sobolev inequality*. Although these inequalities had been formally considered in [54], it was only in [64] that their importance was highlighted with two key observations, which opened the door to further research. The first remarkable fact is that the inequality is uniform with respect to the space dimension n, and therefore extends easily to infinite dimensions. Secondly, Gross proved that if $\mathcal{L}$ is the non-positive self-adjoint operator on $L^2(\mu)$ such that

$$(-\mathcal{L}f, f)_{L^2(\mu)} = \int_{\mathbb{R}^n} |\nabla f(x)|^2 d\mu(x),$$

then the logarithmic Sobolev inequality (2.1) is equivalent to the fact that the semigroup generated by $\mathcal{L}$ is *hypercontractive*, that is for $P_t = e^{t\mathcal{L}}$ and $q(t) \leq 1 + (q-1)e^{2t}$ with $q > 1$ we have

$$\|P_t f\|_{q(t)} \leq \|f\|_q \tag{2.2}$$

for all $f \in L^q(\mu)$. Hypercontractivity thus provides detailed information about the smoothing properties of the semigroup P_t, and has many important applications.

Given the potential usefulness of these inequalities, a natural question was whether or not (2.1) holds in any other situations. The first major breakthrough in this direction, which tremendously enlarged the class of probability measures known to satisfy such inequalities, was due to D. Bakry and M. Emery who in [12] generalised the concept of a logarithmic Sobolev inequality, before giving a sufficient condition for the inequality to hold.

To be more precise, given a non-positive and self-adjoint Markov generator[1] $\mathcal{L}$ acting on $L^2(\mu)$, where (Ω, μ) is some probability space, we say that μ satisfies a logarithmic Sobolev inequality if there exists a constant c such that

$$\mu\left(f^2 \log \frac{f^2}{\mu f^2}\right) \leq c\mu(f(-\mathcal{L}f)) \tag{2.3}$$

for all $f \in \mathcal{D}(\mathcal{L})$, where $\mu(f) \equiv \int f d\mu$. The proof of Gross showing that a logarithmic Sobolev inequality is equivalent to hypercontractivity of the associated semigroup carries over to this more general situation too, as does the observation that such inequalities are uniform in the dimension of the space (see for example Chapters 2 and 3 of [5]).

An important fact is that the logarithmic Sobolev inequality implies the well-known *Poincaré* or *spectral gap inequality*, that is

$$\mu(f - \mu f)^2 \leq c\mu(f(-\mathcal{L}f)). \tag{2.4}$$

To see this implication one can replace f by $1 + \varepsilon f$ in (2.3) and let $\varepsilon \to 0$. These types of inequalities date back to Poincaré, and imply exponential convergence of the associated semigroup to the invariant measure (see for example [66]).

Given this general set-up, Bakry and Emery (following P. A. Meyer) introduced the so-called *carré du champ* operator as the symmetric bilinear form Γ given by

$$\Gamma(f, g) := \frac{1}{2}\left(\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right),$$

and the Γ_2 operator as the symmetric bilinear form given by

$$\Gamma_2(f, g) := \frac{1}{2}\left(\mathcal{L}\Gamma(f, g) - \Gamma(f, \mathcal{L}g) - \Gamma(\mathcal{L}f, g)\right).$$

[1]An operator $\mathcal{L}$ on a Banach space $\mathcal{B}$ is a Markov generator if it generates a Markov semigroup $(P_t)_{t\geq 0}$ i.e. if $\mathcal{L}f = \lim_{t\to 0} \frac{1}{t}(P_t - I)f$ for a Markov semigroup $(P_t)_{t\geq 0}$. Such operators are characterised by the Hille-Yosida Theorem (see for example Theorem 1.7 of [66]).

Their idea was then to study the condition

$$\Gamma_2(f,f) \geq \rho\Gamma(f,f) \tag{2.5}$$

for some constant $\rho \in \mathbb{R}$, which has become known as the $CD(\rho,\infty)$ condition, or the "curvature-dimension" condition (for reasons that will become clear). Their renowned result is that, under the condition that $\mathcal{L}$ is a diffusion[2], the $CD(\rho,\infty)$ condition with $\rho > 0$ is sufficient to ensure that the symmetric measure μ satisfies a logarithmic Sobolev inequality (2.3) with constant $\frac{2}{\rho}$. The $CD(\rho,\infty)$ condition with $\rho > 0$ is sometimes referred to as the *Bakry-Emery condition*.

It is instructive to illustrate the meaning of this condition in a concrete set-up. A fundamental example considered by Bakry and Emery in [11] is the situation of a smooth Riemannian manifold M, equipped with standard gradient ∇ and Laplace-Beltrami operator Δ. It can be noticed that the Markov generator $\mathcal{L} = \Delta - \nabla U \cdot \nabla$ is symmetric and non-positive in $L^2(\mu_U)$, where $\mu_U(dx) = Z^{-1}e^{-U(x)}dx$ is a probability measure on M, with dx the standard Riemannian volume element and $Z = \int e^{-U(x)}dx$ the normalisation constant. In this case it can be calculated that $\Gamma(f,f) = |\nabla f|^2$, and by Bochner's formula

$$\Gamma_2(f,f) = \mathrm{Ric}(\nabla f, \nabla f) + |\mathrm{Hess} f|^2 + \langle \mathrm{Hess}(U)\nabla f, \nabla f\rangle.$$

Thus the Bakry-Emery condition is satisfied if $\inf_{x\in M} k(x) > 0$ where

$$k(x) = \inf\{\mathrm{Ric}(X,X) + \langle \mathrm{Hess}(U)X, X\rangle : X \in T_xM, |X| = 1\}, \quad \forall x \in M,$$

so that for the condition to be satisfied we need some control over the curvature of the space. The $CD(\rho,\infty)$ condition therefore establishes a deep and fundamental link between coercive inequalities (and all their consequences) and the geometry of the underlying space. This relationship has proved extremely useful and has provided the basis for a huge amount of further research.

It is worth making a few remarks at this point:

[2]The operator $\mathcal{L}$ acting on a Banach space $\mathcal{B}$ is a diffusion if for all smooth functions f on $\mathcal{B}$ and Ψ on $\mathbb{R}$ we have that $\mathcal{L}\Psi(f) = \Psi'(f)\mathcal{L}f + \Psi''(f)\Gamma(f,f)$.

- One case of particular interest is when $M = \mathbb{R}^n$. In this setting the curvature is zero, so that by the Bakry-Emery criterion μ_U satisfies a logarithmic Sobolev inequality whenever $\mathrm{Hess}(U)$ is bounded from below by a positive constant, in the sense of quadratic forms. In this case the measure μ_U is often described as being *log-concave*. It is thus clear that the Bakry-Emery condition includes the result of Gross.

- The Bakry-Emery condition is not necessary for a logarithmic Sobolev inequality. Indeed, even if $\rho \leq 0$ in (2.5), in some cases we can still conclude that the invariant measure satisfies a logarithmic Sobolev inequality. For information in this direction we refer to [123, 125] and references therein.

- If we are working in a space where the Ricci curvature is not bounded from below, the $CD(\rho, \infty)$ condition will not hold. This will be important for us, since our focus will be on such settings (see Section 2.3 for details).

- The methods of Bakry and Emery rely heavily on semigroup techniques, and it turns out that that the $CD(\rho, \infty)$ condition is also extremely useful for proving related inequalities involving the associated semigroup. Indeed, under the $CD(\rho, \infty)$ condition (now for any $\rho \in \mathbb{R}$), it can be shown that when $\mathcal{L}$ is a diffusion,

$$\sqrt{\Gamma(P_t f)} \leq e^{-\rho t} P_t(\sqrt{\Gamma(f)})$$

for all $t \geq 0$, where $P_t := e^{t\mathcal{L}}$. In the fundamental example described above, this translates into a commutation relation between P_t and ∇: $|\nabla P_t| \leq e^{-\rho t} P_t |\nabla f|$, which is a well-studied and important relationship.

For a thorough review of the $CD(\rho, \infty)$ condition we refer the reader to [5], which includes all the details and important results, although one can also consult [7] and [114].

2.2 Applications and generalisations

2.2.1 Isoperimetric inequalities

One reason why the logarithmic Sobolev inequality has received so much attention is because it has many applications and connections to other areas. To illustrate this we begin this section by mentioning one of these, namely the role of the logarithmic Sobolev inequality in the study of isoperimetric problems. Although such problems will not be considered below, we include this discussion because in a forthcoming paper co-authored with V. Kontis and B. Zegarliński ([77]) we use methods related to those used in this thesis to prove isoperimetric inequalities in a sub-elliptic setting.

The isoperimetric problem is concerned with controlling the volume of a given set in terms of its surface area. More precisely, given a probability measure μ on a metric space (M, d), we would like to estimate the largest function $\mathcal{I}_\mu : [0,1] \to \mathbb{R}^+$ such that

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

for all measurable sets A, where $\mu^+(A)$ is the μ-surface area of A, defined by

$$\mu^+(A) := \lim_{\varepsilon \to 0} \frac{\mu(A^\varepsilon) - \mu(A)}{\varepsilon},$$

with $A^\varepsilon := \{x \in M : \exists y \in A \text{ such that } d(x,y) < \varepsilon\}$. In pursuing this goal, it turns out that the following two inequalities are of special interest:

$$\mu^+(A) \geq c \min\{\mu(A), 1 - \mu(A)\} \tag{2.6}$$

$$\mu^+(A) \geq c\,\mathcal{U}(\mu(A)). \tag{2.7}$$

Here $c > 0$ is a constant and $\mathcal{U} = \varphi \circ \Phi^{-1}$, where Φ is the distribution function of the normal distribution on $\mathbb{R}$, with $\Phi'(t) = \varphi(t) = (1/\sqrt{2\pi})e^{-\frac{1}{2}t^2}$ for $t \in \mathbb{R}$. Inequalities (2.6) and (2.7) are important because the isoperimetric profile $\mathcal{I}_\mu(t) = \min\{t, 1-t\}$ for the exponential measure on the real line (see [120]) while $\mathcal{I}_\mu(t) = \mathcal{U}(t)$ when μ is the Gaussian measure (see [39, 119]).

These kinds of isoperimetric inequalities were first associated with functional inequalities by J. Cheeger in 1970 ([45]), where it was noted that an inequality of the form (2.6), which is sometimes referred to as a *Cheeger inequality*, implies a Poincaré inequality. Inequality (2.7) is stronger than Cheeger's inequality, and in [88] M. Ledoux connected it with a logarithmic Sobolev inequality, by showing that if (2.7) holds then so does a logarithmic Sobolev inequality. This connection was further strengthened by Bakry and Ledoux in [13], in which they showed that, in a space with Ricci curvature bounded from below, the isoperimetric inequality (2.7) is actually equivalent to the logarithmic Sobolev inequality. Subsequently a large volume of work has been done detailing the role of functional inequalities in isoperimetry, see for example [15, 27, 29, 30, 32, 26, 56, 87, 100] and [113].

2.2.2 q-logarithmic Sobolev inequalities

The first generalisation of the logarithmic Sobolev inequality that will be important for us is the so-called q-logarithmic Sobolev inequality,

$$\mu\left(|f|^q \log \frac{|f|^q}{\mu|f|^q}\right) \leq c\mu|\nabla f|^q \tag{2.8}$$

where $q \in (1, 2]$ and, if we are on a metric space, $|\nabla f|$ comes naturally via the identity

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

For example, given a number $p \in (1, \infty)$, we may equip $\mathbb{R}^n$ with the l^p-metric $d(x, y) = \|x - y\|_p \equiv (\sum_{i=1}^n |x_i - y_i|^p)^{\frac{1}{p}}$, and then obtain

$$|\nabla f|(x) = \left(\sum_{i=1}^n \left|\frac{\partial}{\partial x_i} f(x)\right|^q\right)^{\frac{1}{q}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Inequality (2.8) was introduced in [31], and was shown to hold for probability measures on $\mathbb{R}^n$ with "sufficiently log-concave" densities, for example $\mu(dx) = Z^{-1}e^{-|x|^p}dx$ with $p > 2$. The study was then taken up by S. Bobkov and B. Zegarliński in [32] where it was shown that a q-logarithmic Sobolev inequality with $q \in (1, 2)$ serves as a certain sharpening of the standard inequality. Indeed, they proved that under a q-logarithmic Sobolev inequality one gets a much stronger decay of tails estimate than in the classical "Gaussian" case when $q = 2$ (see Proposition 3.1.5). Moreover, under some weak conditions, it also implies a stronger contractivity property of the associated semigroup P_t than the hypercontractivity one gets when $q = 2$, in that when the dimension of the space is finite, P_t is *ultracontractive* i.e. $\|P_t f\|_\infty \leq \|f\|_p$ for all $t \geq 0$ and $p \in [1, \infty)$ (see Theorem 3.1.13). The q-logarithmic Sobolev inequality is introduced rigorously in Chapter 3.

2.2.3 Related inequalities and their applications

Another generalisation of the logarithmic Sobolev inequality is the so-called Φ-entropy inequality, which, when we are in the general set-up with a non-positive self-adjoint Markov generator $\mathcal{L}$ acting on $L^2(\mu)$ for some measure μ, takes the form

$$\mu(\Phi(f)) - \Phi(\mu(f)) \leq c\mu\left(\Phi''(f)f(-\mathcal{L}f)\right). \tag{2.9}$$

Here Φ is a smooth convex function on an interval I. The left-hand side of this inequality, which is positive by Jensen's inequality, is called the Φ-entropy of f and is often written as $\mathbf{Ent}_\mu^\Phi(f)$.

The Φ-entropy inequality includes many interesting inequalities as particular cases. Indeed, when $\Phi : [0, \infty) \to \mathbb{R}$ is given by $\Phi(x) = x \log x$, (2.9) is nothing but the logarithmic Sobolev inequality, while if $\Phi(x) = x^2$ we recover the Poincaré inequality. When $\Phi(x) = x^p$ for $p \in (1, 2]$ the Φ-entropy describes another important family of inequalities called the *Beckner* inequalities (see [19]), which were later generalised in [86]. Beckner inequalities interpolate between the Poincaré and logarithmic Sobolev inequalities.

A general framework for Φ-entropy inequalities was proposed in [43, 44], though the concept of a Φ-entropy dates back at least to I. Csiszár in the early '70s ([48]). In this framework the additional assumptions that $\Phi'' > 0$ and $\frac{-1}{\Phi''}$ is convex are made, which then allow one to show that Φ-entropy inequalities are tensorisable i.e. that if two measures μ_1, μ_2 satisfy (2.9) with the same constant c, then so does the product measure $\mu_1 \otimes \mu_2$ (extending the result for the standard logarithmic Sobolev inequality). It is further noted that one can perturb the measure by a bounded function, and the inequality remains valid, which generalises an idea of R. Holley and D. Stroock described in [72] (see Proposition 3.1.8 below).

The literature concerning Φ-entropy inequalities is large — see for example [6, 8, 35, 74, 111] and references therein. In particular, in [6] and the recent work [35], they have been used as a tool in studying the convergence to equilibrium of Fokker-Planck-type equations.

It should be noted that in general the Φ-entropy inequality is not homogeneous. For this reason, amongst others, it is useful to introduce a slightly different inequality which is homogeneous. We say that μ satisfies a homogeneous F-Sobolev inequality if there exist constants c_1 and c_2 such that

$$\mu\left(f^2F(f^2)\right) \leq c_1\mu(f(-\mathcal{L}f)) + c_2, \quad \mu(f^2) = 1, \tag{2.10}$$

where $F : [0, \infty) \to [0, \infty)$ with $F(\infty) := \lim_{x\to\infty} F(x) = \infty$. These inequalities appear in the work of F. Y. Wang ([61, 124, 126, 127]), and more recently in [14] where they are studied in relation to contractivity properties, capacity and the Φ-entropy inequality.

In particular in [124], (2.10) was studied with regards to the spectral properties of the generator $-\mathcal{L}$. It is well known that the Poincaré inequality (2.4) is equivalent to the fact that the operator $-\mathcal{L}$ has a gap at the bottom of its spectrum (hence the alternative name for the inequality). The idea was to extend this equivalence, and show that under the stronger inequality (2.10) one can conclude something more about the nature of the spectrum. Wang proved the striking result that, under some conditions, inequality (2.10) holds if and only if the essential spectrum of $-\mathcal{L}$ is empty.

This idea is studied by the introduction of an intermediate inequality designated a *super-Poincaré inequality*:

$$\mu(f^2) \leq r\mu(f(-\mathcal{L}f)) + \beta(r)\left(\mu|f|\right)^2, \quad \forall r > r_0, \tag{2.11}$$

where $\beta : (r_0, \infty) \to (0, \infty)$ is a positive and decreasing function. The author first shows that (2.11) is equivalent to the fact that $\sigma_{ess}(-\mathcal{L}) \subset [r_0^{-1}, \infty)$, before proving that (2.11) with $r_0 = 0$ is in turn equivalent to the homogeneous F-Sobolev inequality (2.10), for a properly chosen F. A similar result was obtained independently by F. Cipriani in [47], where the emphasis was on a specific class of operators and Sobolev embeddings. We summarise their results in more detail in Theorem 4.4.4 below.

2.3 Inequalities for Hörmander-type generators in the sub-Riemannian setting

Recently, a lot of attention has been given to investigating coercive inequalities and their consequences in sub-Riemannian settings, which can be thought of as spaces in which "one can only move in certain directions". These spaces are especially interesting in terms of functional inequalities since the natural Laplacian is no longer elliptic, but is of Hörmander-type and has some degeneracy. Because of this degeneracy the $CD(\rho, \infty)$ condition *cannot hold*, so that the methods of Bakry et al. do not apply.

Following [9], to illustrate this we consider one of the simplest sub-Riemannian settings, namely the Heisenberg group $\mathbb{H}$. This group is introduced rigorously in Section 3.2.2, but can be thought of as $\mathbb{R}^3$, equipped with the vector fields

$$X = \partial_x - \frac{y}{2}\partial_z, \qquad Y = \partial_y + \frac{x}{2}\partial_z.$$

The sub-Riemannian structure is then generated by only considering paths in $\mathbb{R}^3$ which are integral curves of X and Y. The natural sub-Laplacian is given by $\mathcal{L} = X^2 + Y^2$, which is self-adjoint for the Lebesgue measure on $\mathbb{R}^3$. The matrix of second order derivatives associated to $\mathcal{L}$ is degenerate, and thus $\mathcal{L}$ is not elliptic. One should also note that

$$Z := [X, Y] = \partial_z, \quad \text{and} \quad [X, Z] = [Y, Z] = 0,$$

so that X, Y and $[X, Y]$ span the tangent space of $\mathbb{H}$ at every point, or in other words that Hörmander's condition is satisfied. Thus $\mathcal{L}$ is *hypoelliptic* in the sense of Hörmander, which means that, despite being non-elliptic, the heat semigroup $(P_t)_{t\geq 0} = (e^{t\mathcal{L}})_{t\geq 0}$ still admits a smooth density with respect to the Lebesgue measure on $\mathbb{R}^3$ i.e. there exists a smooth function p_t such that

$$P_t f(x) = \int f(xy) p_t(y) dy.$$

For this operator $\mathcal{L}$, one can easily calculate that

$$\Gamma(f, f) = X(f)^2 + Y(f)^2 =: |\nabla_{\mathbb{H}} f|^2$$

and

$$\begin{aligned}\Gamma_2(f, f) = X^2(f)^2 + Y^2(f)^2 + \frac{1}{2}(XY + YX)(f)^2 + \frac{1}{2} Z(f)^2 \\ + 2(XZ(f)Y(f) - YZ(f)X(f)).\end{aligned}$$

The presence of $YZ(f)$ and $XZ(f)$ in the above expression forbids the existence of a constant ρ such that the $CD(\rho, \infty)$ is satisfied. In other words, by the considerations of Section 2.1, the Ricci tensor is everywhere $-\infty$.

Despite this degeneracy, B. Driver and T. Melcher heightened interest in this setting by proving in [51] the existence of a constant C_p such that

$$|\nabla_{\mathbb{H}} P_t f|^p(x) \leq C_p P_t |\nabla_{\mathbb{H}} f|^p(x) \tag{2.12}$$

for all $p > 1, x \in \mathbb{H}$ and smooth functions f. As in the elliptic case, such a gradient bound with $p = 2$ implies that the heat kernel measure satisfies a spectral gap inequality, that is

$$P_t(f^2) - (P_t f)^2 \leq 2tC_2 P_t |\nabla_{\mathbb{H}} f|^2.$$

Driver and Melcher noticed that, due to the group action and homogeneity, it is sufficient to prove (2.12) at the identity and for $t = 1$. The proof then follows using methods from Malliavin calculus. Their result was later extended by Melcher in [99] to include all finite dimensional Lie groups $\mathbb{G}$, where $\{X_i\}_{i=1}^k$ generates the Lie algebra and $\nabla_{\mathbb{G}} = (X_1, \ldots, X_k)$.

Unfortunately the probabilistic approach of Driver and Melcher could not handle the important case $p = 1$. However, using very different methods, H. Q. Li verified in [91] that

(2.12) does indeed hold for $p = 1$, and as a corollary it follows that

$$P_t\left(f^2 \log \frac{f^2}{P_t f^2}\right) \le C_1^2 t P_t |\nabla_{\mathbb{H}} f|^2.$$

The key to proving Li's result is two very precise estimates on the heat kernel p_t, namely that

$$p_1(x) \sim \frac{e^{-\frac{1}{4}d^2(x)}}{(1 + \|x\| d(x))^{\frac{1}{2}}} \tag{2.13}$$

and

$$|\nabla_{\mathbb{H}} p_1(x)| \le C d(x) p_1(x), \tag{2.14}$$

where d is the natural distance function on $\mathbb{H}$ (see Section 3.2). The proofs are given in [91], though they rely on results from [75] and [18]. These precise bounds are strictly necessary; indeed an explicit example was given in [69] showing that the standard exponential bounds on the heat kernel, as described in [49, 122], are not enough for the logarithmic Sobolev inequality to hold. Other simplified proofs of Li's result, which also make use of the estimates (2.13) and (2.14), have since been given in [9], where the symmetry of the group was exploited, and in [77].

An investigation into these types of precise heat kernel bounds on more general H-type groups was undertaken by N. Eldredge in [52], which was then used to obtain gradient estimates with $p = 1$ on such groups in [53].

A slightly different approach to proving coercive inequalities in a sub-Riemannian setting, in particular on H-type groups, was developed by W. Hebisch and B. Zegarliński in [69]. This approach is described in more detail in Section 3.2.5 below and provides some of the main motivation for the work presented here. In this paper an effective technology to study coercive inequalities on very general measure metric spaces was introduced, which does not require a bound on the curvature of the space. Their method is based on so-called U-bounds, which, given that we are working on a general metric space equipped with a collection of possibly non-commuting vector fields $\{X_1, \ldots, X_k\}$, are estimates of the form

$$\int |f|^q U(d)^{\gamma_q} d\mu \le A_q \int |\nabla f|^q d\mu + B_q \int |f|^q d\mu, \tag{2.15}$$

where $d\mu \equiv Z^{-1}e^{-U(d)}d\lambda$ is a probability measure, with $U(d)$ a function having suitable growth at infinity, and $d\lambda$ a natural underlying measure. Here $q \in (1, \infty)$, γ_q is a constant depending on q and U, and d is a metric associated to the gradient $\nabla := (X_1, \dots, X_k)$. The main result of this paper is that, under some weak conditions on the measure $d\lambda$, (2.15) implies that the measure μ satisfies both a spectral gap inequality and a q-logarithmic Sobolev inequality. Moreover, in the case when $U(d) = \alpha d^p$ for $p \geq 2$ and $\alpha > 0$, an inequality of the form (2.15) holds with $\frac{1}{p} + \frac{1}{q} = 1$ and $\gamma_q = 1$ whenever

$$\frac{1}{\sigma} \leq |\nabla d| \leq \sigma \tag{2.16}$$

almost everywhere, for some $\sigma > 0$, and

$$\Delta d \leq K \tag{2.17}$$

outside the unit ball, where $\Delta := \sum_{i=1}^{k} X_i^2$. It happens that conditions (2.16) and (2.17) can be shown to be satisfied in the setting of H-type groups. It thus follows that the measure $d\mu = Z^{-1}e^{-\alpha d^p}d\lambda$ with $p \geq 2$ and $\alpha > 0$ on an H-type group satisfies a q-logarithmic Sobolev inequality, with $\frac{1}{q} + \frac{1}{p} = 1$. Another consequence is that, using the heat kernel bounds (2.13) and (2.14) (and their generalisations), one can use the U-bound to recover the gradient bounds of [9, 53] and [91].

To conclude this section we remark that the question of coercive inequalities and gradient bounds of the type (2.12) for $p = 1$ on groups other than of H-type remains largely open. Some progress has been made on other groups of step 2, including $SU(2)$ and $SL(2)$ (see [10]), but apart from the work of Melcher, almost nothing has been done on groups with step greater than 2.

2.4 Logarithmic Sobolev inequalities in infinite dimensions

In the final section of this chapter we aim to describe a particularly fruitful application of the theory of logarithmic Sobolev inequalities, namely to the infinite dimensional setting of statistical mechanics and spin systems. These considerations provide strong motivation for the work of Chapter 5.

Let μ be a probability measure on a manifold $\mathbf{M}$ which satisfies a logarithmic Sobolev inequality with a constant c. Let $\Omega = \mathbf{M}^{\mathbb{Z}^D}$. By the tensorisation property of the logarithmic Sobolev inequality, it follows that the product measure $\mu_\Lambda := \mu^{\otimes\Lambda}$ also satisfies the inequality with constant c for all $\Lambda \subset \mathbb{Z}^D$. In particular, the inequality makes sense for $\nu = \mu^{\otimes\mathbb{Z}^D}$. We would, however, like to be able to handle more non-trivial situations, when the infinite dimensional measure is not a product measure. Such situations appear in the setting of statistical mechanics and spin systems, where one is often given a family of conditional expectations $\{\mathbb{E}_\Lambda^\omega\}$ indexed by the finite subsets $\Lambda \subset \mathbb{Z}^D$ and $\omega \in \mathbf{M}^{\Lambda^c}$, where $\mathbb{E}_\Lambda^\omega$ is a function of the boundary conditions ω and integrates over the coordinates in Λ. Typically the measures $\mathbb{E}_\Lambda^\omega$ take the following form

$$d\mathbb{E}_\Lambda^\omega = \frac{e^{-U_\Lambda^\omega}}{Z_\Lambda^\omega} d\mu_\Lambda. \tag{2.18}$$

Under some mild conditions, it can be shown that there exists a probability measure ν on Ω, the so-called *Gibbs measure*, which has $\mathbb{E}_\Lambda^\omega$ as its finite volume conditional measures. The Gibbs measure ν is therefore characterised by the condition

$$\nu(\mathbb{E}_\Lambda^\cdot f) = \nu(f)$$

for every finite subset Λ of $\mathbb{Z}^D$ and bounded measurable function f. This is known as the Dobrushin-Lanford-Ruelle (DLR) equation. The conditional measures model the evolution of an interacting particle system whose equilibrium measure is the Gibbs state ν. Originally the purpose of studying such systems was to gain a better understanding of phase transition, though as time has passed it has been noted that very similar mathematical structures can also be naturally formulated in other contexts — neural networks and the spread of infection for example — which illustrates the importance of this type of scheme (see [93] for a comprehensive review of this topic).

A fundamental question within this framework is: when is the system ergodic (i.e. when does it converge to its equilibrium state), and if so how fast and in what sense? This question was addressed by M. Aizenmann and R. Holley in [3] (see also references therein), where it was shown that the associated dynamics is in fact ergodic in the uniform norm with an exponential rate of convergence. The conditions they assumed, however, turned out to be too strong for many meaningful and interesting models, in which their method broke down. Fortunately the theory of logarithmic Sobolev inequalities came to the rescue: a clever new strategy based on the hypercontractivity property was developed in [72, 73] and [116] which overcame these difficulties. The main idea was to deduce the uniform ergodicity from $L^2(\nu)$ ergodicity and hypercontractivity, which we have if and only if the Gibbs measure ν satisfies a logarithmic Sobolev inequality. In view of this work it thus became important to determine in what situations the infinite dimensional measure ν does in fact satisfy a logarithmic Sobolev inequality.

This problem has attracted a lot of attention over the years. The first non-trivial class of examples of non-product Gibbs measures in infinite dimensions which satisfied the logarithmic Sobolev inequality was given in [41]. Later the theory was extended and applied to spin systems when the underlying space is compact by B. Zegarliński [133, 132], D. Stroock and B. Zegarliński [117, 118], S. L. Lu and H. T. Yau [94] and F. Martinelli and E. Olivieri [97, 98]. It has since been reviewed in [66]. The more delicate case of non-compact systems with unbounded interactions was considered by B. Zegarliński [134], N. Yosida [129, 130, 131], B. Helffer [70, 71], B. Helffer and T. Bodineau [33, 34] and others. A self-contained review of this material, which simplifies some of the proofs, was provided by M. Ledoux in [90]. More recently a new criterion in a special setting was given in [104].

It is useful to give an example of a commonly considered unbounded spin system, since we will try to emulate such systems in the new setting examined in Chapter 5 below. Following [90], suppose we are in the situation described at the beginning of this section, with $\mathbf{M} = \mathbb{R}$. We take $\mu(dx) = Z^{-1}e^{-V(x)}dx$ with V strictly convex at infinity (for instance $V(x) = x^4 - \beta x^2$ with $\beta \in \mathbb{R}$), so that μ does indeed satisfy a logarithmic Sobolev inequality by the Bakry-Emery criterion combined with the stability of the inequality under bounded perturbations. Let $\{\mathbb{E}^\omega_\Lambda\}$ be given by (2.18), where, for finite subsets $\Lambda \subset \mathbb{Z}^D$ and $\omega \in \mathbb{R}^{\mathbb{Z}^D}$,

$$U^\omega_\Lambda(x) = J \sum_{i,j\in\Lambda: i\sim j} x_i x_j + J \sum_{i\in\Lambda, j\notin\Lambda: i\sim j} x_i \omega_j, \tag{2.19}$$

with $x = (x_i)_{i\in\Lambda} \in \mathbb{R}^\Lambda$, $J \in \mathbb{R}$ and where the summation is taken over the nearest neighbours $i \sim j$ in the lattice $\mathbb{Z}^D$. Results in the preceding references assert that the logarithmic Sobolev inequality holds for $\mathbb{E}_\Lambda^\omega$ uniformly in both Λ and ω when $|J|$ is small enough, so that it also holds for the Gibbs measure ν by a convergence argument.

To conclude this chapter, we mention that some efforts have been made by B. Zegarliński and P. Ługiewicz in [95] to prove similar infinite dimensional results when we are in a set-up in which there is a given family of *degenerate* vector fields on the underlying space. To be more specific, in this work the authors concentrate on the situation when the underlying spin space $\mathbb{M}$ is a compact manifold without boundary, equipped with a family of degenerate vector fields $\{X_1, \ldots, X_k\}$ satisfying Hörmander's condition. Under some assumptions, they are able to prove that the Gibbs measure corresponding to finite volume conditional measures $\mathbb{E}_\Lambda^\omega$ defined with bounded interactions satisfies a logarithmic Sobolev inequality involving a gradient purely defined in terms of the fields $X_1, \ldots, X_k$. The authors finish by using the proven inequalities to deduce some uniform decay to equilibrium of infinite-dimensional semigroups generated by Hörmander type generators, with exponential rate of convergence.

The work of Chapter 5 below extends this idea of using an underlying spin space naturally equipped with families of Hörmander fields, by trying to combine the techniques for proving coercive inequalities in the sub-Riemannian setting recounted in Section 2.3, with those for proving them for non-trivial Gibbs measures in infinite dimensions. It offers something different to the work of Ługiewicz and Zegarliński in that we consider the more difficult and intriguing case of non-compact spin spaces and unbounded interactions.

Chapter 3

Definitions and Basic Facts

In this chapter we present the definitions, notation and basic facts that will be used in Chapters 4 and 5. At the end of each section we include some notes containing references together with some discussion of the origin of these results. We give proofs where they are sufficiently short and self-contained, and refer the reader to the references given in the notes at the end where they are not.

3.1 Logarithmic Sobolev inequalities

3.1.1 Definitions

In all of what follows, we will be working in measure metric spaces, and therefore we restrict ourselves to this setting.

Indeed, let (Ω, μ) be a probability space equipped with a metric $d : \Omega \times \Omega \to [0, \infty)$. Then for all non-negative measurable functions $f : \Omega \to \mathbb{R}$, we define the entropy functional

$$\mathbf{Ent}_\mu(f) := \int f \log f d\mu - \left(\int f d\mu\right) \log \left(\int f d\mu\right),$$

which is positive by Jensen's inequality.

Furthermore, for a measurable function f we will write

$$\mu(f) \equiv \int f d\mu,$$

and as usual denote by $L^p(\mu)$ the set of all measurable functions such that $\|f\|_p := (\mu|f|^p)^{\frac{1}{p}} < \infty$ for $p \geq 1$, and by $L^\infty(\mu)$ the set of all essentially bounded functions.

Given that we are working on a metric space, we can also introduce the "modulus of the gradient" via the natural identity

$$|\nabla f(x)| \equiv \limsup_{d(x,y)\to 0^+} \frac{|f(x) - f(y)|}{d(x,y)}.$$

Definition 3.1.1. *For $q \in (1, 2]$ we say that μ satisfies a q-*logarithmic Sobolev inequality*, or an LS_q inequality for short, if there exists a constant $c \in (0, \infty)$ such that*

$$\mathbf{Ent}_\mu(|f|^q) \equiv \mu\left(|f|^q \log \frac{|f|^q}{\mu|f|^q}\right) \leq c\mu|\nabla f|^q \qquad (LS_q)$$

for all locally Lipschitz functions f. Moreover, we say that μ satisfies a defective q-logarithmic Sobolev inequality*, or a DLS_q inequality, if there exist constants $c_1, c_2 \in (0, \infty)$ such that*

$$\mathbf{Ent}_\mu(|f|^q) \leq c_1\mu|\nabla f|^q + c_2\mu|f|^q \qquad (DLS_q)$$

for all locally Lipschitz functions f.

The following q-spectral gap inequality will also play an important role.

Definition 3.1.2. *For $q \in (1, 2]$ we say that μ satisfies a q-*spectral gap inequality*, or an SG_q inequality for short, if there exists a constant $c_0 \in (0, \infty)$ such that*

$$\mu|f - \mu f|^q \leq c_0\mu|\nabla f|^q \qquad (SG_q)$$

for all locally Lipschitz functions f.

3.1.2 Basic results

As in the previous section, suppose we are working in a probability space (Ω, μ) equipped with a metric d. The first result gives two equivalent elementary formulations of the entropy functional, together with an associated inequality.

Lemma 3.1.3. *For all positive measurable functions f we have that*

$$\begin{aligned}\mathbf{Ent}_\mu(f) &\equiv \sup\left\{\mu(fg) : \mu\left(e^g\right) = 1\right\} \\ &\equiv \inf_{t>0} \mu\left(f\log f - f\log t - f + t\right).\end{aligned}$$

In particular we have the following relative entropy inequality:

$$\mu(fg) \leq \frac{1}{t}\mu(f)\log\mu\left(e^{tg}\right) + \frac{1}{t}\mathbf{Ent}_\mu(f) \tag{3.1}$$

for all $t > 0$ and measurable functions $f \geq 0$ and g.

Proof. The first identity follows from the elementary inequality $uv \leq u\log u - u + e^v$ for $u \geq 0$ and $v \in \mathbb{R}$. Indeed, for $f \geq 0$ such that $\mu(f) = 1$ we then have that

$$\mu(f\log f) \geq \mu(fg) + 1 - \mu(e^g)$$

so that $\mu(f\log f) \geq \sup\left\{\mu(fg) : \mu(e^g) = 1\right\}$. Furthermore $\mu(e^{\log f}) = 1$ so that

$$\sup\left\{\mu(fg) : \mu(e^g) \geq 1\right\} \geq \mu(f\log f).$$

The assumption that $\mu(f) = 1$ can then be removed by replacing f with $\frac{f}{\mu(f)}$.

The second identity simply follows by calculating the minimum of the functional

$$t \mapsto \mu\left(f\log f - f\log t - f + t\right)$$

and noting that $x \mapsto x\log x$ is twice differentiable and convex.

For the final inequality, note that $\mu\left(e^{\log\frac{e^g}{\mu(e^g)}}\right) = 1$, so that by the first identity

$$\mathbf{Ent}_\mu(f) \geq \mu\left(f\log\frac{e^g}{\mu(e^g)}\right) = \mu(fg) - \mu(f)\log\mu(e^g).$$

Replacing g with tg and rearranging then gives the desired inequality. $\square$

We have that the LS_q inequality is stronger than the SG_q inequality in the following sense:

Proposition 3.1.4. *Suppose the measure μ satisfies an LS_q inequality with a constant c for $q \in (1,2]$. Then μ satisfies an SG_q inequality with constant $\frac{4c}{\log 2}$.*

The next result shows that under the LS_q inequality, one can prove some exponential bounds, and moreover that we have a "decay of tails" estimate.

Proposition 3.1.5. *Suppose the measure μ satisfies an LS_q inequality with a constant c for $q \in (1,2]$. Then for every bounded locally Lipschitz function f such that $|\nabla f| \leq M$ μ-a.e. for $M \in (0,\infty)$, we have*

$$\mu\left(e^{tf}\right) \leq \exp\left\{\frac{cM^q}{q^q(q-1)}t^q + t\mu(f)\right\} \tag{3.2}$$

for all $t > 0$. Moreover

$$\mu\left\{|f-\mu(f)| \geq s\right\} \leq 2\exp\left\{-\frac{(q-1)^p}{M^p c^{p-1}}s^p\right\} \tag{3.3}$$

for all such f and $s > 0$, where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Let f be a bounded locally Lipschitz function such that $|\nabla f| \leq M$ μ-a.e. Applying the LS_q inequality to the function $F = e^{tf/q}, t > 0$ yields

$$\mu\left(tfe^{tf}\right) - \mu(e^{tf})\log\mu(e^{tf}) \leq c\frac{M^q t^q}{q^q}\mu(e^{tf}). \tag{3.4}$$

We can write $\int e^{tf} \equiv e^{tv(t)}$ for some function v which is smooth in $t > 0$ and satisfies $\lim_{t\to 0} v(t) = \mu(f)$. We then note that

$$t^2 v'(t)e^{tv(t)} = \mu\left(tfe^{tf}\right) - \mu(e^{tf})\log\mu(e^{tf}),$$

so that (3.4) yields the following differential inequality

$$v'(t) \leq c\frac{M^q t^{q-2}}{q^q}.$$

Thus

$$v(t) - \mu(f) = \int_0^t v'(s)ds \leq c\frac{M^q}{q^q}\int_0^t s^{q-2}ds = \frac{cM^q}{q^q(q-1)}t^{q-1},$$

from which it follows that

$$\mu\left(e^{tf}\right) \leq \exp\left\{\frac{cM^q}{q^q(q-1)}t^q + t\mu(f)\right\},$$

i.e. (3.2) holds. Applying the same argument to $-f$ we then arrive at

$$\mu\left(e^{t|f-\mu(f)|}\right) \leq 2\exp\left\{\frac{cM^q}{q^q(q-1)}t^q\right\}$$

for all $t > 0$. Now, given $s > 0$, by Chebyshev's inequality[1] and (3.2), we see that

$$\mu\{|f-\mu(f)| \geq s\} \leq e^{-ts}\int e^{t|f-\mu(f)|}d\mu \leq 2\exp\left\{-ts + \frac{cM^q}{q^q(q-1)}t^q\right\}$$

for all $t > 0$. Optimisation over t then yields (3.3). ☐

Corollary 3.1.6. *Suppose the measure μ satisfies an LS_q inequality with a constant c for $q \in (1, 2]$. Then for every bounded locally Lipschitz function f such that $|\nabla f| \leq M$ μ-a.e. for $M \in (0, \infty)$, we have that*

$$\mu\left(e^{t|f-\mu(f)|^p}\right) \leq 1 + \frac{2t}{t_0 - t}$$

for all $t \in (0, t_0)$, where $t_0 = \frac{(q-1)^p}{M^p c^{p-1}}$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Define the function $G(s) = \int_{\{|f-\mu(f)|\geq s\}} d\mu$. If $\sigma(s) := \mu(|f-\mu(f)| \leq s)$ is the distribution of $|f-\mu(f)|$ then $G(s) = 1 - \sigma(s)$, so that

$$\int e^{t|f-\mu(f)|^p}d\mu = \int_0^\infty e^{ts^p}d\sigma(s) = -\int_0^\infty e^{ts^p}dG(s).$$

[1] Chebyshev's inequality states that $\mu\{f(x) \geq s\} \leq \frac{1}{g(s)}\int g \circ f d\mu$ for any non-negative and non-decreasing measurable function g.

Now, by integration by parts, and Proposition 3.1.5, we therefore have

$$\begin{aligned}\int e^{t|f-\mu(f)|^q} d\mu &= 1 + pt\int_0^\infty s^{p-1}e^{ts^p}\mu\{|f-\mu(f)|\geq s\}\,ds \\ &\leq 1+2t\int_0^\infty ps^{p-1}e^{s^p(t-t_0)}ds,\end{aligned}$$

where $t_0 = \frac{(q-1)^p}{M^p c^{p-1}}$. Thus

$$\int e^{t|f-\mu(f)|^q} d\mu \leq 1 + \frac{2t}{t-t_0}\int_0^\infty \frac{d}{ds}\left(e^{s^p(t-t_0)}\right)ds = 1-\frac{2t}{t-t_0},$$

for all $0<t<t_0$. □

An important result that will be used extensively is the following one, which states that a defective q-logarithmic Sobolev inequality can be tightened using a q-spectral gap inequality.

Proposition 3.1.7. *Suppose that the measure μ satisfies a defective q-logarithmic Sobolev inequality for $q\in(1,2]$, i.e. there exist constants c_1, c_2 such that*

$$\mu\left(|f|^q\log\frac{|f|^q}{\mu|f|^q}\right)\leq c_1\mu|\nabla f|^q + c_2\mu|f|^q.$$

Suppose moreover that μ satisfies an SG_q inequality, i.e. there exists a constant c_0 such that

$$\mu|f-\mu f|^q \leq c_0\mu|\nabla f|^q.$$

Then μ satisfies an LS_q inequality.

The next two results show that the LS_q inequality is stable under bounded perturbations and tensorisation.

Proposition 3.1.8. *Suppose that the measure μ satisfies an LS_q inequality for $q\in(1,2]$ with a constant c, and define $d\hat{\mu}=\rho d\mu$, where ρ is some strictly positive and bounded density such that $\int\rho d\mu=1$. Then the measure $\hat{\mu}$ also satisfies an LS_q inequality with constant $\hat{c}=c\exp\{\sup(\log\rho)-\inf(\log\rho)\}$.*

Proof. By Lemma 3.1.3 we can write

$$\begin{aligned}
\mathbf{Ent}_{\hat{\mu}}(|f|^q) &= \inf_{t>0} \hat{\mu}\left(|f|^q \log|f|^q - |f|^q \log t - |f|^q + t\right) \\
&\leq e^{\sup(\log\rho)} \inf_{t>0} \mu\left(|f|^q \log|f|^q - |f|^q \log t - |f|^q + t\right) \\
&\leq c e^{\sup(\log\rho)} \mu|\nabla f|^q \\
&\leq c e^{\sup(\log\rho) - \inf(\log\rho)} \hat{\mu}|\nabla f|^q,
\end{aligned}$$

since we have assumed that μ satisfies an LS_q inequality with a constant c. $\square$

Proposition 3.1.9. *Let* (Ω_i, μ_i, d_i) *for* $i \in \{1,2\}$ *be two metric measure spaces. Suppose* μ_i *satisfies an* LS_q *inequality for* $q \in (1,2]$ *with constant* c_i *for* $i \in \{1,2\}$. *Then the product measure* $\mu_1 \otimes \mu_2$ *also satisfies an* LS_q *inequality with constant* $c = \max\{c_1, c_2\}$, *in the sense that*

$$\mathbf{Ent}_{\mu_1\otimes\mu_2}(|f|^q) \leq c\mu_1 \otimes \mu_2 |\nabla f|_q^q \tag{3.5}$$

where $|\nabla f|_q^q := |\nabla_{\Omega_1} f|^q + |\nabla_{\Omega_2} f|^q$, *and* $|\nabla_{\Omega_i} f|$ *is the length of the gradient of* f *as a function on* Ω_i.

Remark 3.1.10. *When we refer to an* LS_q *inequality on a product space, we will be alluding to an inequality of the form* (3.5). *This can be reconciled with Definition 3.1.1 by equipping the probability space* $(\Omega_1 \otimes \Omega_2, \mu_1 \otimes \mu_2)$ *with the metric* $(d_1^p + d_2^p)^{\frac{1}{p}}$, *where* $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. We first claim that

$$\mathbf{Ent}_{\mu_1\otimes\mu_2}(f) \leq \mu_1(\mathbf{Ent}_{\mu_2}(f)) + \mu_2(\mathbf{Ent}_{\mu_1}(f)). \tag{3.6}$$

Indeed, let g be a measurable function on $\Omega_1 \otimes \Omega_2$ such that $\mu_1 \otimes \mu_2(e^g) = 1$. Then we can write

$$g = g_1 + g_2$$

where $g_1 = g - \log \int e^g d\mu_1$ and $g_2 = \log \int e^g d\mu_1$. Note that $\mu_1(e^{g_1}) = \mu_2(e^{g_2}) = 1$. Thus, by Lemma 3.1.3, we see that

$$\mu_1(fg_1) + \mu_2(fg_2) \leq \mathbf{Ent}_{\mu_1}(f) + \mathbf{Ent}_{\mu_2}(f).$$

Therefore

$$\mu_1 \otimes \mu_2(fg) = \mu_1 \otimes \mu_2(fg_1 + fg_2) \leq \mu_2\left(\mathbf{Ent}_{\mu_1}(f)\right) + \mu_1\left(\mathbf{Ent}_{\mu_2}(f)\right).$$

Taking the supremum over all such g, and using the characterisation of the entropy given in Lemma 3.1.3 once more, proves the claim.

Now, applying the LS_q inequalities for the measures μ_1 and μ_2 in (3.6) yields

$$\mathbf{Ent}_{\mu_1\otimes\mu_2}(|f|^q) \leq c_2\mu_1 \otimes \mu_2|\nabla_{\Omega_2} f|^q + c_1\mu_1 \otimes \mu_2|\nabla_{\Omega_1} f|^q,$$

which proves the result. □

The next two results show that on a finite product space, when $q < q'$, the SG_q inequality is stronger than the $SG_{q'}$ inequality, and similarly LS_q is stronger than $LS_{q'}$.

Proposition 3.1.11. *Let (Ω_i, μ_i, d_i) for $i \in \{1, \dots n\}$ be metric measure spaces and $q, q' > 1$ be such that $q < q'$. Suppose that $\mu \equiv \otimes_{i=1}^n \mu_i$ satisfies an SG_q inequality, in the sense that there exists a constant c_0 such that*

$$\mu|f - \mu f|^q \leq c_0\mu|\nabla f|_q^q$$

where $|\nabla f|_q^q = \sum_{i=1}^n \mu|\nabla_{\Omega_i} f|^q$. Then it also satisfies an $SG_{q'}$ inequality i.e there exists a constant c_0' such that

$$\mu|f - \mu f|^{q'} \leq c_0'\mu|\nabla f|_{q'}^{q'}.$$

Proof. Let f be a locally Lipschitz function such that $\mu(f) = 0$. Then by assumption

$$\int |f|^q d\mu \leq c_0 \int |\nabla f|_q^q d\mu. \tag{3.7}$$

Denote by $m = m(f)$ a median of f, so that $\mu\{f \leq m\} \geq \frac{1}{2}$ and $\mu\{f \geq m\} \geq \frac{1}{2}$. Assume for definiteness that $m > 0$. Since by Chebyshev's inequality $\mu\{f \geq m\} \leq \frac{1}{m^q}\mu|f|^q$, we see that $m^q \leq 2\mu|f|^q$. Thus for general m we have

$$|m|^q \leq 2\mu|f|^q.$$

Hence by (3.7) we see that

$$\int |f - m(f)|^q d\mu \leq 3.2^{q-1} c_0 \int |\nabla f|_q^q d\mu. \tag{3.8}$$

This inequality is invariant under translation, so it also holds for all locally Lipschitz functions with arbitrary mean. Now take such a function and assume that $m(f) = 0$. Consider the locally Lipschitz functions $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. Note that $|\nabla f^+|_q$ and $|\nabla f^-|_q$ respectively vanish on the sets $\{f < 0\}$ and $\{f > 0\}$, and coincide with $|\nabla f|_q$ on $\{f > 0\}$ and $\{f < 0\}$. Since $m(f^+) = m(f^-) = 0$, an application of (3.8) to $(f^+)^{q'/q}$ and $(f^-)^{q'/q}$ gives respectively

$$\int_{\{f\geq 0\}} |f|^{q'} d\mu \leq 3.2^{q-1} c_0 \left(\frac{q'}{q}\right)^q \int_{\{f\geq 0\}} |f|^{q'-q} |\nabla f|_q^q d\mu$$
$$\int_{\{f\leq 0\}} |f|_q^{q'} d\mu \leq 3.2^{q-1} c_0 \left(\frac{q'}{q}\right)^q \int_{\{f\leq 0\}} |f|^{q'-q} |\nabla f|_q^q d\mu.$$

Summing these yields

$$\int |f|^{q'} d\mu \leq 3c_0 \left(\frac{2q'}{q}\right)^q \int |f|^{q'-q} |\nabla f|_q^q d\mu. \tag{3.9}$$

Note that by definition

$$\int |f|^{q'-q} |\nabla f|_q^q d\mu = \mu\left(|f|^{q'-q}\left(\sum_{i=1}^n |\nabla_{\Omega_i} f|^q\right)\right).$$

We can then use Hölder's inequality to see that

$$\begin{aligned}\int |f|^{q'-q}|\nabla f|_q^q d\mu &\leq \left[\mu|f|^{q'}\right]^{\frac{q'-q}{q'}} \left[\mu\left(\sum_{i=1}^n |\nabla_{\Omega_i} f|^q\right)^{\frac{q'}{q}}\right]^{\frac{q}{q'}} \\ &\leq \left[\mu|f|^{q'}\right]^{\frac{q'-q}{q'}} n^{\frac{q'-q}{q'}} \left[\mu\left(\sum_{i=1}^n |\nabla_{\Omega_i} f|^{q'}\right)\right]^{\frac{q}{q'}}. \end{aligned} \tag{3.10}$$

Using this in (3.9) yields

$$\begin{aligned}\left[\mu|f|^{q'}\right]^{\frac{q}{q'}} &\leq 3c_0\left(\frac{2q'}{q}\right)^q n^{\frac{q'-q}{q'}}\left[\mu|\nabla f|_{q'}^{q'}\right]^{\frac{q}{q'}} \\ \Rightarrow \mu|f|^{q'} &\leq (3c_0)^{\frac{q'}{q}}\left(\frac{2q'}{q}\right)^{q'} n^{\frac{q'-q}{q}}\mu|\nabla f|_{q'}^{q'}.\end{aligned}$$

Since $\mu|f-\mu f|^{q'} \leq 2^{q'-1}\mu|f|^{q'}$, we arrive at

$$\mu|f-\mu f|^{q'}d\mu \leq c_0'\mu|\nabla f|_{q'}^{q'},$$

where $c_0' = 2^{q'-1}(3c_0)^{\frac{q'}{q}}\left(\frac{2q'}{q}\right)^{q'} n^{\frac{q'-q}{q}}$. The assumption that f has a zero median may then be omitted due to the translational invariance of the inequality. ☐

Proposition 3.1.12. *Let (Ω_i, μ_i, d_i) for $i \in \{1, \dots n\}$ be metric measure spaces and $q, q' \in (1, 2]$ be such that $q < q'$. Suppose that $\mu \equiv \otimes_{i=1}^n \mu_i$ satisfies an LS_q inequality. Then it also satisfies an $LS_{q'}$ inequality.*

Proof. Suppose μ satisfies an LS_q inequality with constant c. We can apply this inequality to $f^{\frac{q'}{q}}$ to see that

$$\mu\left(|f|^{q'}\log\frac{|f|^{q'}}{\mu|f|^{q'}}\right) \leq c\mu\left|\nabla f^{\frac{q'}{q}}\right|_q^q = c\left(\frac{q'}{q}\right)^q \mu\left(|f|^{q'-q}|\nabla f|_q^q\right). \tag{3.11}$$

We can bound the right-hand side using (3.10) of Proposition 3.1.11, since once again we are supposing that the underlying space is finite dimensional. Indeed, using (3.10) in (3.11) yields

$$\mu\left(|f|^{q'}\log\frac{|f|^{q'}}{\mu|f|^{q'}}\right)\le c\left(\frac{q'}{q}\right)^q\left[\mu|f|^{q'}\right]^{\frac{q'-q}{q'}}n^{\frac{q'-q}{q'}}\left[\mu\left(\sum_{i=1}^n|\nabla_{\Omega_i}f|^{q'}\right)\right]^{\frac{q}{q'}}.$$

We may then apply Young's inequality $ab\le\frac{a^r}{r}+\frac{r-1}{r}b^{\frac{r}{r-1}}$ for all $a,b\ge0$ with $r=\frac{q'}{q'-q}$ to see that

$$\mu\left(|f|^{q'}\log\frac{|f|^{q'}}{\mu|f|^{q'}}\right)\le c\left(\frac{q'}{q}\right)^{q'-1}\mu|\nabla f|_{q'}^{q'}+c\frac{(q'-q)n}{q'}\mu|f|^{q'}.$$

Thus we see that μ satisfies a $DLS_{q'}$ inequality. By Propositions 3.1.4 and 3.1.11 we also have that μ satisfies an $SG_{q'}$ inequality, so that we may conclude with an application of Proposition 3.1.7. □

We finish this section by stating a consequence of the LS_q inequality, to do with the contractivity properties of the associated semigroup.

Theorem 3.1.13. *Let (Ω_i,μ_i,d_i) for $i\in\{1,\dots n\}$ be metric measure spaces and suppose that $\mu\equiv\otimes_{i=1}^n\mu_i$ satisfies an LS_q inequality for $q\in(1,2)$. Suppose also that $\mathcal{L}$ is an operator such that*

$$\mu(f\mathcal{L}f)=-\mu|\nabla f|_2^2.$$

Then the semigroup $P_t=e^{t\mathcal{L}}$ is ultracontractive, that is for any $p\in[1,\infty)$ and $t>0$ the operator $P_t:L^p\to L^\infty$ is bounded.

3.1.3 Notes

As already mentioned, the logarithmic Sobolev inequality for the Gaussian measure on $\mathbb{R}^n$ was introduced by L. Gross in [64]. Although we only give the definition in the context of metric measure spaces, it can also be given in a more general setting in terms of an infinitesimal Markov generator and the so-called *carré du champ* operator, as briefly mentioned in Chapter 2. For further information in this direction we refer the reader to [5, 7, 11, 12, 66] and references therein. The q-logarithmic Sobolev inequality first appeared in [31] and was extensively studied in [32].

The fact that a q-logarithmic Sobolev inequality implies a q-spectral gap inequality (Proposition 3.1.4) in the case when $q = 2$ is well known; indeed one can see this by applying the logarithmic Sobolev inequality to the function $1+\varepsilon f$ and developing the limit as $\varepsilon \to 0$. For general $q \in (1, 2]$ this was shown in [32] from which the proof given above is taken.

Proposition 3.1.5 has its origins in an unpublished letter by I. Herbst. Indeed in the case $q = 2$ the proof presented is known as the Herbst argument. In this case it was also mentioned in [50] and further developed in [1, 2, 28, 65] and [89] in relation to concentration of measure results. The argument in the case of general q was given in [32]. Corollary 3.1.6 is adapted from Lemma 7.3.2 of [71].

In the case $q = 2$, Proposition 3.1.7 was first shown by O. Rothaus in [113], and is indeed sometimes referred to as the Rothaus lemma, with the general case following from results found in [32]. The general case is also given explicitly in Appendix B of [77]. Propositions 3.1.11 and 3.1.12 are both shown in [32].

The tensorisation property of the logarithmic Sobolev inequality (Proposition 3.1.9) was first noticed by Gross in [64]. The bounded perturbation result (Proposition 3.1.8) first appeared in in [72]. Both these results are well known and can be found together with some discussion in, for example, Chapter 3 of [5]. Once again the general case for $q \in (1, 2]$ is shown in [32].

The final result of this section (Theorem 3.1.13) is found in [32]. However, it also has its origins in the work of Gross ([64]), where the equivalence of the standard logarithmic Sobolev inequality and hypercontractivity of the associated semigroup was shown.

3.2 H-type groups

In this section we introduce the concept of an H-type group and the sub-Riemannian geometry in which we will be working in Chapters 4 and 5. We also state some results that will be important for what follows.

3.2.1 Definitions and structure

Definition 3.2.1 (H-type group). *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra equipped with Lie bracket $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Let $\mathfrak{z}$ denote the centre of $\mathfrak{g}$, that is*

$$\mathfrak{z} = \{X \in \mathfrak{g} : [X, Y] = 0 \,\forall\, Y \in \mathfrak{g}\}\,.$$

The Lie algebra $\mathfrak{g}$ is said to be of H-type *if it can be endowed with an inner product $\langle\cdot,\cdot\rangle$ such that*

$$[\mathfrak{z}^\perp, \mathfrak{z}^\perp] = \mathfrak{z},$$

and moreover, for every fixed $Z \in \mathfrak{z}$, the map $J_Z : \mathfrak{z}^\perp \to \mathfrak{z}$ defined by

$$\langle J_Z(X), Y\rangle = \langle Z, [X, Y]\rangle \qquad \forall Y \in \mathfrak{z}^\perp$$

is an orthogonal map[2] *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.*

We remark that the map $J_{(\cdot)} : \mathfrak{z} \to \mathrm{End}(\mathfrak{z}^\perp)$ in the above definition is well-defined and linear. Indeed, for fixed $Z \in \mathfrak{z}$ and $X \in \mathfrak{z}^\perp$ the map

$$\Psi : \mathfrak{z}^\perp \to \mathbb{R}, \qquad Y \mapsto \Psi(Y) := \langle Z, [X, Y]\rangle$$

is linear. Hence there exists exactly one $W \in \mathfrak{z}^\perp$ such that $\Psi(Y) = \langle W, Y\rangle$ for every $Y \in \mathfrak{z}^\perp$, and we set $J_Z(X) = W$. It can then be checked that for fixed $Z \in \mathfrak{z}$, $J_Z(\cdot)$ is linear, and moreover that for fixed $X \in \mathfrak{z}^\perp$, $J_{(\cdot)}(X) : \mathfrak{z} \to \mathfrak{z}^\perp$ is also linear.

The following Theorem provides an explicit characterisation of H-type groups.

Theorem 3.2.2. *$\mathbb{G}$ is an H-type group if and only if $\mathbb{G}$ is (isomorphic to) $\mathbb{R}^{n+m}$ with the*

[2] Recall that J_Z is orthogonal if $\langle J_Z(X), J_Z(Y)\rangle = \langle X, Y\rangle$ for all $X, Y \in \mathfrak{z}^\perp$.

group law

$$(w,z)\circ(\omega,\zeta)=\begin{pmatrix} w_i+\omega_i, \quad i=1,\dots,n \\ z_j+\zeta_j+\frac{1}{2}\langle U^{(j)}w,\omega\rangle, \quad j=1,\dots,m \end{pmatrix}, \tag{3.12}$$

for $w,\omega\in\mathbb{R}^n$, $z,\zeta\in\mathbb{R}^m$ *and where the matrices* $U^{(1)},\dots,U^{(m)}$ *have the following properties:*

(1) $U^{(j)}$ *is an* $n\times n$ *skew-symmetric and orthogonal matrix for every* $j\in\{1,\dots,m\}$*;*

(2) $U^{(k)}U^{(j)}+U^{(j)}U^{(k)}=0$ *for every* $k,j\in\{1,\dots,m\}$ *with* $k\neq j$.

Thus, without any loss of generality, we will henceforth assume that any H-type group $\mathbb{G}$ is of this form. For an H-type group $\mathbb{G}$ and $x\in\mathbb{G}$, we will therefore use the notation

$$x=(w,z)=(w_1,\dots,w_n,z_1,\dots,z_m),$$

for $w\in\mathbb{R}^n$ and $z\in\mathbb{R}^m$. It is clear that the point $(0,0)$ is the identity in $\mathbb{G}$ and the inverse operation is $(w,z)^{-1}=(-w,-z)$.

We can identify $\mathfrak{g}$ with the space spanned by the left-invariant vector fields

$$\{X_1,\dots,X_n,Z_1,\dots,Z_m\}$$

on $\mathbb{G}$, where $X_i(0)=\frac{\partial}{\partial w_i}$ and $Z_j(0)=\frac{\partial}{\partial z_j}$. This is the canonical basis for $\mathfrak{g}$. If we let $\{e^{(1)},\dots,e^{(n)}\}$ and $\{u^{(1)},\dots,u^{(m)}\}$ denote the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, using the group operation (3.12) we can explicitly calculate the left-invariant vector fields. Indeed, for $i\in\{1,\dots,n\}$, $j\in\{1,\dots,m\}$ and a smooth function f we have

$$\begin{aligned}(X_if)(w,z)&=\frac{d}{ds}\bigg|_{s=0} f\left((w,z)\circ(se^{(i)},0)\right)\\ &=\frac{d}{ds}\bigg|_{s=0} f\left(w+se^{(i)},z+\frac{1}{2}s\sum_{k=1}^m\sum_{l=1}^n U^{(k)}_{il}w_lu^{(k)}\right)\\ &=\left(\frac{\partial}{\partial w_i}+\frac{1}{2}\sum_{k=1}^m\sum_{l=1}^n U^{(k)}_{il}w_l\frac{\partial}{\partial z_k}\right)f(w,z)\end{aligned}$$

and

$$\begin{aligned}(Z_j f)(w,z) &= \frac{d}{ds}\bigg|_{s=0} f\left((w,z)\circ(0,su^{(j)})\right)\\ &= \frac{d}{ds}\bigg|_{s=0} f\left(w, z+su^{(j)}\right)\\ &= \frac{\partial}{\partial z_j} f(w,z).\end{aligned}$$

Thus $\mathfrak{g}$ is spanned by the vector fields

$$X_i = \frac{\partial}{\partial w_i} + \frac{1}{2}\sum_{k=1}^{m}\sum_{l=1}^{n} U_{il}^{(k)} w_l \frac{\partial}{\partial z_k}, \qquad Z_j = \frac{\partial}{\partial z_j}, \tag{3.13}$$

for $i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}$. The key point is that the algebra generated by the vector fields $\{X_1,\ldots,X_n\}$ together with their first order commutators is actually the whole of $\mathfrak{g}$. This follows from the observation that

$$[X_i, X_j] = \sum_{k=1}^{m} U_{ji}^{(k)} Z_k \tag{3.14}$$

for every $i, j \in \{1,\ldots,n\}$, and the fact that $U^{(1)},\ldots,U^{(m)}$ are linearly independent, which follows from properties of the matrices given in Theorem 3.2.2 (see Remark 18.2.3 of [38]). Thus by taking linear combinations of $[X_i, X_j]$ for $i, j \in \{1,\ldots n\}$, one can obtain the vector fields Z_k for $k \in \{1,\ldots,m\}$. In other words

$$\operatorname{span}\{X_i, [X_j, X_k] : i,j,k \in \{1,\ldots n\}\} = \mathfrak{g},$$

which is equivalent to saying that the H-type group $\mathbb{G}$ is a Carnot group of step 2. To see how this structure relates to Definition 3.2.1, one can introduce the inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$ to be the standard inner product with respect to the canonical basis (3.13) and show by direct calculation that $\mathfrak{z}^\perp = \operatorname{span}\{X_1,\ldots,X_n\}$, and $\mathfrak{z} = \operatorname{span}\{Z_1,\ldots,Z_m\} = [\mathfrak{z}^\perp,\mathfrak{z}^\perp]$.

In view of this, we make the following definitions of the *sub-gradient* and *sub-Laplacian* on $\mathbb{G}$.

Definition 3.2.3. *The second order differential operator*

$$\Delta_{\mathbb{G}} = \sum_{j=1}^{n} X_i^2$$

is called the canonical sub-Laplacian *on* $\mathbb{G}$. *The vector-valued operator*

$$\nabla_{\mathbb{G}} = (X_1, \ldots, X_n)$$

will be called the canonical sub-gradient *on* $\mathbb{G}$.

Remark 3.2.4. *As mentioned in Section 2.3, one of the reasons that these spaces are of interest to us is that the existing methods of Bakry and Emery to prove logarithmic Sobolev inequalities do not work here. Indeed, their methods rely on the existence of a constant* $\rho \in \mathbb{R}$ *such that*

$$\Gamma_2(f) \geq \rho |\nabla_{\mathbb{G}} f|^2$$

where $\Gamma_2(f) = \frac{1}{2}\left(\Delta_{\mathbb{G}}|\nabla_{\mathbb{G}} f|^2 - 2\nabla_{\mathbb{G}} f \cdot \nabla_{\mathbb{G}}(\Delta_{\mathbb{G}} f)\right)$, *which is equivalent to having a bound from below on the curvature of the space. However, as in the case of the Heisenberg group outlined in Section 2.3, by direct calculation one can see that no such* ρ *exists when we are working on an H-type group. For more details of this calculation see [9].*

It is worth noting here that the vector fields $\{X_1, \ldots, X_n\}$ satisfy *Hörmander's condition* i.e. the set $\{X_1, \ldots, X_n\}$ together with their commutators span the tangent space at each point $x \in \mathbb{G}$. This has two consequences — the first being that we can therefore reach every point of the space just by travelling along integral curves of $X_1, \ldots, X_n$, which in turn allows us to define a sub-Riemannian distance function on $\mathbb{G}$ (see Section 3.2.3). The other consequence is that by Hörmander's theorem the sub-Laplacian is *hypoelliptic*, that is if u is a distribution such that $\Delta_{\mathbb{G}} u \in C^\infty$, then $u \in C^\infty$. This is equivalent to the fact that there exists a *smooth* function $\rho_t : \mathbb{G} \to \mathbb{R}, t > 0$ called the *heat kernel* such that

$$e^{t\Delta_{\mathbb{G}}} f(x) = f * \rho_t(x) = \int_{\mathbb{G}} f(x \circ y)\rho_t(y)dy, \qquad \forall x \in \mathbb{G}.$$

Given the above structure we make some further remarks. The first one is that there is a natural family of dilation operators on an H-type group.

Definition 3.2.5. *Let* $\mathbb{G} = \mathbb{R}^{m+n}$ *be an H-type group. For* $\lambda > 0$ *define the map* $\delta_\lambda : \mathbb{G} \to \mathbb{G}$ *by*

$$\delta_\lambda(w, z) := (\lambda w, \lambda^2 z)$$

for $(w, z) \in \mathbb{G}$*. Then* δ_λ *is a group homomorphism in the sense that*

$$\delta_\lambda(x \circ y) = \delta_\lambda(x) \circ \delta_\lambda(y) \qquad \forall\, x, y \in \mathbb{G}.$$

The family $(\delta_\lambda)_{\lambda>0}$ *is referred to as the family of* dilations*, and the triple* $(\mathbb{G}, \circ, \delta_\lambda)$ *is said to be a homogeneous group.*

The second remark is that the Lebesgue measure on $\mathbb{R}^{n+m}$ is invariant with respect to the group action i.e. it is the Haar measure.

Lemma 3.2.6. *Let* $\mathbb{G} = \mathbb{R}^{n+m}$ *be an H-type group. Then the Lebesgue measure on* $\mathbb{R}^{n+m}$ *is invariant under both left and right translations on* $\mathbb{G}$ *i.e. if we denote by* $|E|$ *the Lebesgue measure of a measurable set* $E \subset \mathbb{R}^{n+m}$ *we have*

$$|x \circ E| = |E| = |E \circ x|, \qquad \forall x \in \mathbb{G}.$$

Proof. To see this consider the maps $y \mapsto x \circ y$ and $y \mapsto y \circ x$. One can calculate the Jacobian matrices of these maps directly using the group product (3.12) to see that they are lower triangular with 1s on the diagonal, so that their determinant is 1. Indeed, the Jacobian of the map $L_x : \mathbb{G} \to \mathbb{G}$ where $L_x(y) = x \circ y$ is given by the matrix $(a_{ij})_{1 \le i,j \le n+m}$ where

$$a_{ij} = \frac{\partial}{\partial x_j}(L_x(y))_i.$$

Moreover, by the definition of the group law, one can then see that $a_{ii} = 1$ for all $i \in \{1, \ldots, n+m\}$ and $a_{ij} = 0$ if $i < j$. □

We also note here, since it will be used later, that the vector fields $X_1, \ldots, X_n$ are divergence-free with respect to the Lebesgue measure on $\mathbb{R}^{n+m}$.

Finally, using the same method as in the above Lemma, we can also see that

$$|\delta_\lambda(E)| = \lambda^Q |E|$$

for all $\lambda > 0$, where $Q = n + 2m$. In view of this we make the following definition:

Definition 3.2.7. *Let* $\mathbb{G} = \mathbb{R}^{n+m}$ *be an H-type group. Then* $Q = n + 2m$ *is called the* homogeneous dimension *of* $\mathbb{G}$.

3.2.2 Example: The Heisenberg group

The main example of an H-type group to keep in mind is the Heisenberg group $\mathbb{H}$. In fact H-type groups were introduced as a generalisation of the Heisenberg group. $\mathbb{H}$ can be realised as $\mathbb{R}^{2+1}$ with the group operation

$$(w_1, w_2, z) \circ (\omega_1, \omega_2, \zeta) = \left(w_1 + \omega_1, w_2 + \omega_2, z + \zeta + \frac{1}{2}(w_1\omega_2 - w_2\omega_1) \right)$$

for $w = (w_1, w_2), \omega = (\omega_1, \omega_2) \in \mathbb{R}^2$ and $z, \zeta \in \mathbb{R}$. We can see that $\mathbb{H}$ is an H-type group in the sense of Theorem 3.2.2, since

$$(w_1, w_2, z) \circ (\omega_1, \omega_2, \zeta) = \left(w_1 + \omega_1, w_2 + \omega_2, z + \zeta + \frac{1}{2}\langle Uw, \omega\rangle \right)$$

where

$$U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The left-invariant vector fields on $\mathbb{H}$ are given by

$$X_1 = \partial_{w_1} - \frac{1}{2} w_2 \partial_z, \quad X_2 = \partial_{w_2} + \frac{1}{2} w_1 \partial_z, \quad Z = \partial_z,$$

and one can easily calculate that $[X_1, X_2] = Z, [X_1, Z] = [X_2, Z] = 0$. It is known as the Heisenberg group because Heisenberg wrote down these bracket relations in his work on quantum mechanics. Higher dimensional Heisenberg groups can similarly be defined, and all have the common characteristic of a one dimensional centre. It should however be noted that H-type groups with centres of arbitrarily high dimension can also be defined.

3.2.3 Natural homogeneous metrics

Throughout this section we will suppose that $\mathbb{G} = \mathbb{R}^{n+m}$ is an H-type group with Lie algebra $\mathfrak{g} = \operatorname{span}\{X_i, [X_j, X_k] : i, j, k \in \{1, \ldots, n\}\}$ as above. We describe two different but natural ways to define a metric on $\mathbb{G}$. The first way is to use the structure we have on $\mathbb{G}$ to define a geometry in which we "only move in certain directions", or more precisely only along the integral curves of the vector fields $X_1, \ldots, X_n$. The second metric appears naturally in the fundamental solution of the sub-Laplacian.

Definition 3.2.8. *Let $\gamma : [0,1] \to \mathbb{G}$ be an absolutely continuous path. We say that γ is* horizontal *if there exist measurable functions $a_1, \ldots, a_n : [0,1] \to \mathbb{R}$ such that*

$$\dot{\gamma}(t) = \sum_{i=1}^{n} a_i(t) X_i(\gamma(t))$$

for almost all $t \in [0,1]$ i.e. $\dot{\gamma}(t) \in \operatorname{span}\{X_1(\gamma(t)), \ldots, X_n(\gamma(t))\}$ almost everywhere. For such a horizontal curve γ, we define the length of γ to be

$$|\gamma| := \int_0^1 \left(\sum_{i=1}^{n} a_i^2(t) \right)^{\frac{1}{2}} dt.$$

We then define the Carnot-Carathéodory distance *$d(x,y)$ between two points $x, y \in \mathbb{G}$ to be*

$$d(x,y) := \inf\left\{ |\gamma| \text{ such that } \gamma : [0,1] \to \mathbb{G} \text{ is horizontal and } \gamma(0) = x, \gamma(1) = y \right\}.$$

We will write $d(x) := d(x,0)$.

It is not immediately clear, and it is non-trivial, that this distance function is well-defined. We therefore need the following result, which relies on the fact that the vector fields $X_1, \ldots, X_n$ satisfy the Hörmander condition:

Theorem 3.2.9 (Chow). *Let $x, y \in \mathbb{G}$ with $x \neq y$. Then there exists a horizontal path $\gamma : [0,1] \to \mathbb{G}$ such that $\gamma(0) = x$ and $\gamma(1) = y$.*

Thus the Carnot-Carathéodory distance is well-defined, and can be shown to be a metric. We also have that the infimum in the definition is achieved by some horizontal path:

Theorem 3.2.10. *For any two points $x, y \in \mathbb{G}$, there exists a horizontal path $\gamma : [0,1] \to \mathbb{G}$ with $\gamma(0) = x$ and $\gamma(1) = y$ such that $d(x,y) = |\gamma|$.*

It is worth remarking that by a scaling argument an equivalent definition of the Carnot-Carathéodory distance is

$$d(x,y) = \inf \left\{ t | \gamma : [0,t] \to \mathbb{G} \text{ is horizontal}, \gamma(0) = x, \gamma(t) = y, |\dot{\gamma}(s)| \leq 1 \ \forall s \in [0,t] \right\}.$$

We use this observation to see that d is associated to the sub-gradient via the identity

$$|\nabla_{\mathbb{G}} f(x)| = \limsup_{d(x,y) \to 0} \frac{|f(x) - f(y)|}{d(x,y)}. \tag{3.15}$$

Indeed

$$\begin{aligned} |f(x) - f(y)| &= \left| \int_0^t \frac{d}{ds} f(\gamma(s)) ds \right| \\ &= \left| \int_0^t \nabla_{\mathbb{G}} f(\gamma(s)) \cdot \dot{\gamma}(s) ds \right| \end{aligned}$$

where $\gamma : [0,t] \to \mathbb{R}$ is a horizontal path from x to y such that $|\dot{\gamma}(s)| \leq 1$ which realises the distance $d(x,y)$, so that $t = d(x,y)$. Then

$$|f(x) - f(y)| \leq \int_0^t |\nabla_{\mathbb{G}} f(\gamma(s))| ds \leq t \sup_{s \in [0,t]} |\nabla_{\mathbb{G}} f(\gamma(s))|.$$

By dividing by t and taking the limit supremum as $t \to 0$ we arrive at (3.15).

The following result describes some important properties of d.

Proposition 3.2.11. *The function $d : \mathbb{G} \to [0, \infty)$ is continuous (with respect to the Euclidean topology) and is such that*

(i) $d(x) > 0$ *if and only if* $x \neq 0$*;*

(ii) $d(x^{-1}) = d(x)$ *for all* $x \in \mathbb{G}$*;*

(iii) $d(\delta_\lambda(x)) = \lambda d(x)$ *for all* $\lambda > 0$ *and* $x \in \mathbb{G}$.

We say that d is a *symmetric homogeneous norm on* $\mathbb{G}$. In fact we have the following result, which asserts the equivalence of all homogeneous norms on $\mathbb{G}$:

Proposition 3.2.12. *Let $\tilde{d}$ be another homogeneous norm on $\mathbb{G}$. Then there exists a constant $C > 0$ such that*

$$C^{-1}d(x) \leq \tilde{d}(x) \leq Cd(x), \quad x \in \mathbb{G}.$$

Despite this fact, as we will see in Chapter 4, homogeneous norms can behave quite differently. We now introduce an alternative homogeneous norm which arises naturally from the fundamental solution of the sub-Laplacian.

Theorem 3.2.13. *Define the function*

$$F(x) := N(x)^{2-Q},$$

where $N(x) = (|w|^4 + 16|z|^2)^{1/4}$ for $x = (w, z) \in \mathbb{G}$, and where $Q = n + 2m$ is the homogeneous dimension of $\mathbb{G}$ as in Definition 3.2.7. Then F is a fundamental solution of $\Delta_{\mathbb{G}}$, in the sense that F is smooth out of the origin and

$$\Delta_{\mathbb{G}} F(x) = 0 \qquad in \quad \mathbb{G} \backslash \{0\}.$$

Definition 3.2.14. *The function $N : \mathbb{G} \to [0, \infty)$ defined by $N(x) = (|w|^4 + 16|z|^2)^{1/4}$ for $x = (w, z) \in \mathbb{G}$ is a symmetric homogeneous norm (one can easily check this), which we will call the* Kaplan *distance.*

Remark 3.2.15. *Perhaps the most important difference between the Carnot-Carathéodory distance d and the Kaplan distance N, as we will see in the next section, is the fact that N is smooth on $\mathbb{G} \backslash \{0\}$ while d is not differentiable on $\{x = (w, z) \in \mathbb{G} : w = 0\}$.*

3.2.4 Preliminary calculations and inequalities

Let $\mathbb{G} = \mathbb{R}^{n+m}$ be an H-type group and $\Delta_{\mathbb{G}}$ and $\nabla_{\mathbb{G}}$ be the sub-Laplacian and sub-gradient respectively. Moreover, let $d : \mathbb{G} \to [0, \infty)$ be the Carnot-Carathéodory distance and $N : \mathbb{G} \to [0, \infty)$ be the Kaplan distance of a point from the origin.

The first useful result describes the behaviour of the sub-gradient of the two distance functions.

Proposition 3.2.16. (i) $d : \mathbb{G} \to [0, \infty)$ *is smooth on the set* $\{x = (w, z) \in \mathbb{G} : w \neq 0\}$, *and*

$$|\nabla_{\mathbb{G}} d(x)| = 1$$

for all $x = (w, z) \in \mathbb{G}$ *such that* $w \neq 0$.

(ii) $N : \mathbb{G} \to [0, \infty)$ *is smooth on* $\mathbb{G}\backslash\{0\}$, *and*

$$|\nabla_{\mathbb{G}} N(x)| = \frac{\|x\|}{N(x)}$$

for all $x = (w, z) \in \mathbb{G}$ *such that* $x \neq 0$, *where* $\|x\| := |w| = (\sum_{i=1}^{n} w_i^2)^{\frac{1}{2}}$.

Proof. The fact that $d : \mathbb{G} \to [0, \infty)$ is smooth on $\{x = (w, z) \in \mathbb{G} : w \neq 0\}$ is shown in Lemma 6.2 of [69]. We also have that

$$|d(x) - d(y)| \leq d(x \circ y^{-1}) = d(x, y) \qquad \forall\, x, y \in \mathbb{G},$$

so that the function d is trivially 1-Lipschitz. We can then apply a generalisation of Rademacher's Theorem (see for example Theorem 3.7 of [102] or [59]) to conclude that $X_i d(x)$ exists for all $x = (w, z) \in \mathbb{G}$ with $w \neq 0, i \in \{1, \ldots, n\}$ and moreover that

$$|\nabla_{\mathbb{G}} d(x)| = \left(\sum_{i=1}^{n} (X_i d(x))^2 \right)^{\frac{1}{2}} \leq 1.$$

For the reverse inequality let $x = (w, z) \in \mathbb{G}$ be a point where this inequality holds. Let

$\gamma : [0,t] \to \mathbb{G}$ be a horizontal geodesic joining 0 to x such that $|\dot{\gamma}(s)| \le 1$. Thus

$$\dot{\gamma}(s) = \sum_{i=1}^{n} a_i(s) X_i(\gamma(s))$$

with $\sum_{i=1}^{n} a_i^2(s) \le 1$. We can then differentiate the identity $s = d(\gamma(s))$ to see that

$$\begin{aligned} 1 &= \frac{d}{ds} d(\gamma(s)) = \nabla_{\mathbb{G}} d(\gamma(s)) \cdot \dot{\gamma}(s) \\ &= \sum_{i=1}^{n} a_i(s) X_i d(\gamma(s)) \le |\nabla_{\mathbb{G}} d(\gamma(s))| \end{aligned}$$

for all $s \in [0,t]$, by the Cauchy-Schwarz inequality. By taking $s = t$ this proves (i).

For (ii) we make a direct calculation. Indeed, for $i \in \{0, \ldots, n\}$ and $x = (w,z) \neq 0$ we have

$$X_i N(x) = \frac{1}{N^3(x)} \left(|w|^2 w_i + 4 \sum_{k=1}^{m} \sum_{l=1}^{n} U_{il}^{(k)} w_l z_k \right).$$

Thus

$$\begin{aligned} |\nabla_{\mathbb{G}} N(x)|^2 &= \sum_{i=1}^{n} (X_i N(x))^2 \\ &= \frac{1}{N^6(x)} \sum_{i=1}^{n} \left(|w|^2 w_i + 4 \sum_{k=1}^{m} \sum_{l=1}^{n} U_{il}^{(k)} w_l z_k \right)^2 \\ &= \frac{1}{N^6(x)} \left[|w|^6 + 8|w|^2 \sum_{k=1}^{m} \left(\sum_{l,i=1}^{n} U_{il}^{(k)} w_i w_l \right) z_k \right. \\ &\qquad \left. + 16 \sum_{i=1}^{n} \left(\sum_{k=1}^{m} \sum_{l=1}^{n} U_{il}^{(k)} w_l z_k \right)^2 \right]. \end{aligned}$$

Since $U^{(k)}$ is skew-symmetric for all $k \in \{1, \ldots, m\}$ we have that $\sum_{l,i=1}^{n} U_{il}^{(k)} w_i w_l = 0$. Moreover

$$\sum_{i=1}^{n} \left(\sum_{k=1}^{m} \sum_{l=1}^{n} U_{il}^{(k)} w_l z_k \right)^2 = \sum_{i=1}^{n} \sum_{k=1}^{m} \left(\sum_{l=1}^{n} U_{il}^{(k)} w_l \right)^2 z_k^2,$$

since the matrices are such that $U^{(i)}U^{(j)} + U^{(j)}U^{(i)} = 0$ for all $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Now for all $k \in \{1, \ldots, m\}$

$$\sum_{i=1}^{n} \left(\sum_{l=1}^{n} U_{il}^{(k)} w_l \right)^2 = |U^{(k)}w|^2 = |w|^2$$

since $U^{(k)}$ is orthogonal, so that by above

$$\sum_{i=1}^{n} \left(\sum_{k=1}^{m} \sum_{l=1}^{n} U_{il}^{(k)} w_l z_k \right)^2 = |w|^2|z|^2.$$

Therefore

$$\begin{aligned} |\nabla_{\mathbb{G}} N(x)|^2 &= \frac{1}{N^6(x)} \left(|w|^6 + 16|w|^2|z|^2 \right) \\ &= \frac{|w|^2}{N^6(x)} \left(|w|^4 + 16|z|^2 \right) \\ &= \frac{|w|^2}{N^2(x)} \end{aligned}$$

as claimed. □

In what follows we will also have to deal with terms involving $\Delta_{\mathbb{G}} d$. Care is needed, since d is not smooth everywhere so that there will be singularities on the set $\{x = (w, z) \in \mathbb{G} : w = 0\}$. However, the next result provides some control of these singularities, as well as an explicit calculation of $\Delta_{\mathbb{G}} N$.

Proposition 3.2.17. (i) *There exists a constant $K \in (0, \infty)$ such that*

$$\Delta_{\mathbb{G}} d \leq \frac{K}{d}$$

where $\Delta_{\mathbb{G}} d$ is understood in the sense of distributions.

(ii) *For all $g \in \mathbb{G} \backslash \{0\}$,*

$$\Delta_{\mathbb{G}} N(x) = (Q-1) \frac{\|x\|^2}{N^3(x)}$$

where as above, for $x = (w, z) \in \mathbb{R}^{n+m}$, $\|x\| = |w|$.

Proof. For part (i) it suffices to show that $\Delta_{\mathbb{G}} d \leq K$ on the set $\{d(x) = 1\}$. Indeed, using dilations and the homogeneity of the sub-Laplacian, we have that

$$\Delta_{\mathbb{G}} d(x) = \lambda(\Delta_{\mathbb{G}} d)(\delta_\lambda(x))$$

for all $x \neq 0$ and $\lambda > 0$, so that for any $x \in \mathbb{G}\backslash\{0\}$

$$\Delta_{\mathbb{G}} d(x) \leq \frac{1}{d(x)} \sup_{\{d(y)=1\}} \Delta_{\mathbb{G}} d(y).$$

The claim that $\Delta_{\mathbb{G}} d \leq K$ on $\{d(x) = 1\}$ is proved in Theorem 6.1 of [69].

For (ii) again we can just make the calculation. Indeed, using the fact that $F = N^{2-Q}$ is a fundamental solution of the sub-Laplacian as in Theorem 3.2.13, we can calculate that for $x \neq 0$

$$\begin{aligned}
\Delta_{\mathbb{G}} N(x) &= \Delta_{\mathbb{G}} \left(F^{\frac{1}{2-Q}} \right)(x) \\
&= \nabla_{\mathbb{G}} \cdot \left(\frac{1}{2-Q} F^{\frac{1}{2-Q}-1} \nabla_{\mathbb{G}} F \right)(x) \\
&= \frac{Q-1}{(2-Q)^2} F^{\frac{1}{2-Q}-2}(x) |\nabla_{\mathbb{G}} F|^2(x) + \frac{1}{2-Q} F^{\frac{Q-1}{2-Q}}(x) \Delta_{\mathbb{G}} F(x) \\
&= \frac{Q-1}{(2-Q)^2} F^{\frac{1}{2-Q}-2}(x) |\nabla_{\mathbb{G}} F|^2(x).
\end{aligned}$$

Moreover, using part (ii) of Proposition 3.2.16 we have that

$$|\nabla_{\mathbb{G}} F|^2(x) = (2-Q)^2 N^{-2Q}(x) \|x\|^2.$$

Using this in the above calculation yields

$$\Delta_{\mathbb{G}} N(x) = (Q-1) \frac{\|x\|^2}{N^3(x)}$$

for all $x \neq 0$, as required. □

The last two results show that both the classical Sobolev and the Poincaré inequality hold in the setting of H-type groups.

Theorem 3.2.18. *For $r > 0$ and $x \in \mathbb{G}$, let $B_r(x) = \{y \in \mathbb{G} : d(x,y) \leq r\}$ be the ball of radius r centred at x. Then, for all $p \in [1,\infty)$ there exists a constant $P_0(r) = P_0(r,p)$ such that for all $f \in C^\infty(\mathbb{G})$*

$$\int_{B_r(x)} \left| f(y) - f_{B_r(x)} \right|^p dy \leq P_0(r) \int_{B_r(x)} |\nabla_{\mathbb{G}} f(y)|^p dy$$

where $f_{B_r(x)} := \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)dy$.

Theorem 3.2.19. *There exist constants $a, b \in [0,\infty)$ such that for $p \geq Q$*

$$\left(\int f^{\frac{p}{p-1}}(x)dx \right)^{\frac{p-1}{p}} \leq a \int |\nabla_{\mathbb{G}} f(x)|dx + b \int |f(x)|dx$$

for all $f \in C_0^\infty(\mathbb{G})$.

3.2.5 U-bounds and their consequences on H-type groups

A major motivator for the work contained within Chapters 4 and 5 is the paper of W. Hebisch and B. Zegarliński [69], in which some useful machinery was introduced to study coercive inequalities that can be applied in the setting of H-type groups. For this reason, together with the fact that we sometimes directly make use of the results, here we briefly summarise the important points from that paper.

Let $\mathbb{G} = \mathbb{R}^{n+m}$ be an H-type group and $\Delta_{\mathbb{G}}$ and $\nabla_{\mathbb{G}}$ be the sub-Laplacian and sub-gradient respectively. For $p \in (1,\infty)$ let μ_p be the probability measure on $\mathbb{G}$ given by

$$\mu_p(dx) := \frac{e^{-\alpha d^p(x)}}{Z} dx, \tag{3.16}$$

where $d : \mathbb{G} \to [0,\infty)$ is the Carnot-Carathéodory distance of a point from the origin, $Z = \int e^{-\alpha d^p(x)} dx$, $\alpha > 0$, and dx is the Lebesgue measure on $\mathbb{G}$.

Theorem 3.2.20 (U-bound)**.** *Let μ_p be given by* (3.16).

(i) *Let $p \geq 2$. Then there exist constants $A, B \in (0,\infty)$ such that*

$$\int |f|^q d^p d\mu_p \leq A \int |\nabla_{\mathbb{G}} f|^q d\mu_p + B \int |f|^q d\mu_p$$

for all locally Lipschitz functions f, and where $\frac{1}{q}+\frac{1}{p}=1$.

(ii) *Let $p\in(1,2]$. Then there exist constants $A, B\in(0,\infty)$ such that*

$$\int f^2 d^{2(p-1)}d\mu_p \leq A\int |\nabla_{\mathbb{G}} f|^2 d\mu_p + B\int f^2 d\mu_p$$

for all locally Lipschitz functions f.

Remark 3.2.21. *The proof of this result is relatively simple, and only relies on integration by parts together with the facts that $|\nabla_{\mathbb{G}} d| = 1$ almost everywhere and $\Delta_{\mathbb{G}} d\leq K$ outside the unit ball (i.e. Propositions 3.2.16 and 3.2.17). In fact, this result is true in a general metric space when these two bounds hold.*

Using Theorem 3.2.20, we can then pass to a q-spectral gap inequality.

Theorem 3.2.22. *Let μ_p be given by* (3.16).

(i) *Let $p\geq 2$. Then there exists a constant c_0 such that*

$$\mu_p|f-\mu_p f|^q \leq c_0\mu_p|\nabla_{\mathbb{G}} f|^q$$

for all locally Lipschitz functions f, and where $\frac{1}{q}+\frac{1}{p}=1$.

(ii) *Let $p\in(1,2]$. Then there exists a constant c_0 such that*

$$\mu_p|f-\mu_p f|^2 \leq c_0\mu_p|\nabla_{\mathbb{G}} f|^2$$

for all locally Lipschitz functions f.

We can finally combine both the above results to arrive at an LS_q inequality.

Theorem 3.2.23. *Let μ_p be given by* (3.16).

(i) *Let $p\geq 2$. Then there exists a constant c such that*

$$\mu_p\left(|f|^q\log\frac{|f|^q}{\mu_p|f|^q}\right) \leq c\mu_p|\nabla_{\mathbb{G}} f|^q$$

for all locally Lipschitz functions f, and where $\frac{1}{q}+\frac{1}{p}=1$.

(ii) *Let* $p \in (1,2]$. *Then there exists a constant* c *such that*

$$\mu_p F(f^2) - F(\mu_p f^2) \leq c\mu_p |\nabla_{\mathbb{G}} f|^q$$

for all locally Lipschitz functions f, *where* $F(t) = t\,(\log(1+t))^{\frac{2(p-1)}{p}}$.

3.2.6 Notes

General H-type groups were first introduced in [82]. The definition we have given is not exactly the original one, but it is the one usually adopted in the more recent literature.

The main reference for this section is the book [38] where most of the results of Sections 3.2.1 and 3.2.3 can be found, including the characterisation result (Theorem 3.2.2) which is proved in Chapter 18 of that book. It also contains a detailed introduction to Carnot groups in general.

The Carnot-Carathéodory distance was introduced in [40]. The fundamental theorem Theorem 3.2.9 (indeed a more general version on Carnot groups) is due to W. L. Chow in [46], though an earlier version in the case of $\mathbb{R}^3$ with two Hörmander vector fields appeared in [40]. Modern proofs can be found in for example [20, 63] and [122]. A proof of Theorem 3.2.10 can be found in Appendix D of [101].

The homogeneous norm N related to the fundamental solution of the sub-Laplacian was discovered by A. Kaplan in [82] on general H-type groups, extending the work of G.B. Folland [55] on the Heisenberg group.

The proof of the fact that the Carnot-Carathéodory distance on the Heisenberg group satisfies the eikonal equation (Proposition 3.2.16) is due to R. Monti [102]. The bound on the sub-Laplacian of the Carnot-Carathéodory distance is due to W. Hebisch and B. Zegarliński and can be found in [32].

The Poincaré inequality (Theorem 3.2.18) is quoted from Theorem 5.6.1 of [115]. The classical Sobolev inequality (Theorem 3.2.19) comes from Chapter IV of the book of N. Varopoulos, L. Saloff-Coste and T. Coulhon [122], which also includes a comprehensive discussion of inequalities on Lie groups.

Chapter 4

Operators on H-type Groups with Discrete Spectra

4.1 Introduction

In the classical setting of $\mathbb{R}^n$ an extensive study has been made of operators of the form

$$\mathcal{L} = -\Delta + V$$

where Δ is the standard Laplacian on $\mathbb{R}^n$ and V is some potential. The initial value problem for the Schrödinger equation may be reduced to the investigation of the spectrum of these operators acting on a Hilbert space, and hence they have become known as *Schrödinger operators*. A classical reference detailing this study is the book of M. Reed and B. Simon [110].

In this chapter we consider a direct analogue of this type of operator, but now defined in the sub-Riemannian setting of H-type groups, and where we replace the full Laplacian with the more natural sub-Laplacian. Given an H-type group $\mathbb{G}$, we will be particularly interested in the sub-elliptic operators

$$\mathcal{L} = -\Delta_{\mathbb{G}} + \nabla_{\mathbb{G}} U \cdot \nabla_{\mathbb{G}} \tag{4.1}$$

where $\Delta_{\mathbb{G}}$ and $\nabla_{\mathbb{G}}$ are the sub-Laplacian and sub-gradient respectively. When considered as an operator acting on $L^2(\mu_U)$ with $\mu_U \equiv Z^{-1}e^{-U}dx$, such operators are positive and self-adjoint. Moreover, when U is given as a power of the Carnot-Carathéodory distance d, our investigation will tie in nicely with the results of the recent work of Hebisch and Zegarliński described in [69], where such measures are thoroughly studied (see also Section 3.2.5). Our principal aim is to show that when $U(x) = \alpha d^p(x)$ for $\alpha \in (0, \infty)$ and $p > 1$, the operator (4.1) acting on $L^2(\mu_U)$ has empty essential spectrum, or in other words that it has a purely discrete set of eigenvalues.

We begin our pursuit of this goal by working in the Heisenberg group, and in the first section below prove a generalisation of a classical result in $\mathbb{R}^n$. This generalisation is of interest because in the classical case the corresponding theorem provides us with information about the spectrum of operators corresponding to those we wish to study. However, in the Heisenberg group things are more complicated, since the Carnot-Carathéodory distance function is not smooth out of the origin. It turns out that the potentials we are interested in are not smooth enough to be easily handled by the generalised result (see Remark 4.3.5). To avoid these problems in Section 4.4 we take a different approach. We instead achieve our objective by exploiting some results of F. Y. Wang and F. Cipriani ([47, 124]) about the relationship between functional inequalities and the spectrum.

In the final section we deal with the situation when we replace the Carnot-Carathéodory distance with the Kaplan distance, and show that subtle differences in the behaviour of these distance functions result in notable differences in the properties of the corresponding generators. Indeed, for $p \in (1, 2)$, the operators defined with the Kaplan distance do not even have a spectral gap, let alone an empty essential spectrum. However, we show that for $p \geq 2$ they do at least have a spectral gap, and conjecture that when $p > 2$ they will also have a discrete spectrum.

The work of this section came about as a result of some discussions with Prof. L. Saloff-Coste, to whom the author is very grateful, and the results are partially published in [76].

4.2 Notation

Let $(T, \mathcal{D}(T))$ be a closed[1] operator on a Hilbert space $\mathcal{H}$. A complex number λ is in the *resolvent set* $\rho(T)$ if $\lambda I - T$ is a bijection of $\mathcal{D}(T)$ onto $\mathcal{H}$ with bounded inverse. The *spectrum* of the operator T is defined to be $\sigma(T) := \mathbb{C} \backslash \rho(T)$. The *discrete spectrum* of T, $\sigma_{disc}(T)$, consists of all isolated eigenvalues of T with finite multiplicity. The *essential spectrum* of T is defined by $\sigma_{ess}(T) := \sigma(T) \backslash \sigma_{disc}(T)$.

4.3 Generalisation of a classical result

There are some well-known criteria that ensure classical Schrödinger operators defined on $\mathbb{R}^n$ have empty essential spectra. For example, we can consider the Schrödinger operator $\mathcal{L} = -\Delta + V$ on $\mathbb{R}^n$, with $V \in L^\infty_{loc}(\mathbb{R}^n)$. If $V(x) \to \infty$ as $|x| \to \infty$, we can then conclude that $\mathcal{L}$ has a purely discrete spectrum (see Theorem XIII.67 of [110]). In this section we prove a generalisation of this result in the Heisenberg group (Theorem 4.3.3 below), and apply it to the situation when $\mathcal{L}$ is given by (4.1).

Let $\mathbb{H} = \mathbb{R}^3$ be the Heisenberg group, as described in Section 3.2.2, with $\Delta_{\mathbb{H}} = X_1^2 + X_2^2$ and $\nabla_{\mathbb{H}} = (X_1, X_2)$ the sub-Laplacian and sub-gradient respectively. Recall that

$$X_1 f(x) = \left(\partial_{w_1} - \frac{1}{2} w_2 \partial_z \right) f(x), \qquad X_2 f(x) = \left(\partial_{w_2} + \frac{1}{2} w_1 \partial_z \right) f(x),$$

for $x = (w, z) \in \mathbb{H}$, where $w \in \mathbb{R}^2$ and $z \in \mathbb{R}$. As usual, we let $d : \mathbb{H} \to [0, \infty)$ denote the Carnot-Carathéodory distance of a point from the origin.

We will make use of the following neat observation from [68] (see also [57] and [67]). Denote by $\mathcal{F}_3$ the partial Fourier transform with respect to the third variable:

$$\mathcal{F}_3 f(w_1, w_2, \zeta) := (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-iz\zeta} f(w_1, w_2, z) dz.$$

[1] $(T, \mathcal{D}(T))$ is closed if $\{(\varphi, T\varphi) \in \mathcal{H} \times \mathcal{H} : \varphi \in \mathcal{D}(T)\}$ is a closed subset of $\mathcal{H} \times \mathcal{H}$

Note that, by integration by parts,

$$\begin{aligned}\mathcal{F}_3(iX_1f)(w_1,w_2,\zeta) &= (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-iz\zeta}\left(i\frac{\partial}{\partial w_1}-\frac{1}{2}iw_2\frac{\partial}{\partial z}\right)f(w_1,w_2,z)dz\\ &= (2\pi)^{-1/2}\left(i\frac{\partial}{\partial w_1}+\frac{1}{2}w_2\zeta\right)\int_{-\infty}^{\infty} e^{-iz\zeta}f(w_1,w_2,z)dz\\ &= \left(i\frac{\partial}{\partial w_1}+\frac{1}{2}w_2\zeta\right)\mathcal{F}_3f(w_1,w_2,\zeta).\end{aligned}$$

Hence

$$\mathcal{F}_3(-X_1^2f)(w_1,w_2,\zeta) = \left(i\frac{\partial}{\partial w_1}+\frac{1}{2}w_2\zeta\right)^2\mathcal{F}_3f(w_1,w_2,\zeta),$$

and similarly

$$\mathcal{F}_3(-X_2^2f)(w_1,w_2,\zeta) = \left(i\frac{\partial}{\partial w_2}-\frac{1}{2}w_1\zeta\right)^2\mathcal{F}_3f(w_1,w_2,\zeta).$$

Thus

$$\begin{aligned}\mathcal{F}_3(-\Delta_{\mathbb{H}}f)(w,\zeta) &= \left[\left(i\frac{\partial}{\partial w_1}+\frac{1}{2}w_2\zeta\right)^2+\left(i\frac{\partial}{\partial w_2}-\frac{1}{2}w_1\zeta\right)^2\right]\mathcal{F}_3f(w,\zeta)\\ &= \left(i\nabla_w+\zeta\mathbf{A}(w)\right)^2\mathcal{F}_3f(w,\zeta), \qquad (4.2)\end{aligned}$$

where $w=(w_1,w_2)\in\mathbb{R}^2$, $\nabla_w=(\partial_{w_1},\partial_{w_2})$ and $\mathbf{A}(w)=\frac{1}{2}(-w_2,w_1)$.

The key observation is that, for fixed $\zeta\in\mathbb{R}$, the operator $\left(i\nabla_w+\zeta\mathbf{A}(w)\right)^2$ is well-known and corresponds to the Hamiltonian of a particle moving in a uniform magnetic field (see [84] and the references therein). The spectral analysis of these operators goes back to Landau and Fock and the birth of quantum mechanics. In particular, as described in [84], the spectrum is discrete, and the eigenvalues, or energy levels, are given by

$$\lambda_k(\zeta) := |\zeta|(2k+1), \quad k\in\{0,1,\dots\}.$$

The eigenvalue $|\zeta|(2k+1)$ is sometimes called the k-th *Landau level*. Moreover, the eigenspace corresponding to each eigenvalue is infinite dimensional, and the corresponding orthogonal eigenprojections $\mathcal{P}_k$ are explicit, and given by

$$\mathcal{P}_k f(w) = \int_{\mathbb{R}^2} f(w')\pi_k(w,w')dw',$$

for $w \in \mathbb{R}^2$, where

$$\pi_k(w,w') = \frac{|\zeta|}{2\pi} e^{-\frac{|\zeta|}{2} i(w_1 w_2' - w_2 w_1') - \frac{|\zeta|}{4}|w-w'|^2} L_k\left(\frac{|\zeta|}{2}|w-w'|^2\right)$$

and L_k is the k-th Laguerre polynomial, given by

$$L_k(t) = \frac{1}{k!} e^t \frac{d^k}{dt^k}\left(t^k e^{-t}\right), \quad t \geq 0.$$

Note that $\pi_k(w,w')$ is constant on the diagonal:

$$\pi_k(w,w) = \frac{|\zeta|}{2\pi}. \tag{4.3}$$

Using these facts in (4.2), we arrive at the following spectral decomposition

$$\mathcal{F}_3(-\Delta_{\mathbb{H}} f)(w,\zeta) = \sum_{k=0}^{\infty} \lambda_k(\zeta)\mathcal{P}_k\mathcal{F}_3 f(w,\zeta), \qquad w \in \mathbb{R}^2, \zeta \in \mathbb{R}.$$

Moreover, note that

$$\begin{aligned}
&\sum_{k=0}^{\infty} \int_{\mathbb{R}} |\zeta|(2k+1)\|\mathcal{P}_k\mathcal{F}_3 f\|_{L^2(dw)}^2 d\zeta \\
&\quad = \sum_{k=0}^{\infty} \int_{\mathbb{R}} |\zeta|(2k+1)\left(\int_{\mathbb{R}^2} \mathcal{P}_k^2\mathcal{F}_3 f(w,\zeta)\mathcal{F}_3 f(w,\zeta)dw\right)d\zeta \\
&\quad = \sum_{k=0}^{\infty} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \mathcal{F}_3^*\left(|\zeta|(2k+1)\mathcal{P}_k\mathcal{F}_3 f\right)(w,z) f(w,z)dwdz \\
&\quad = \int_{\mathbb{R}^3} \mathcal{F}_3^*\left(\sum_{k=0}^{\infty} |\zeta|(2k+1)\mathcal{P}_k\mathcal{F}_3 f\right)(x) f(x)dx \\
&\quad = \int_{\mathbb{H}} f(x)(-\Delta_{\mathbb{H}} f)(x)dx,
\end{aligned}$$

for $x = (w, z) \in \mathbb{R}^3$. In view of this we make the following definition:

Definition 4.3.1. *For a function $f \in L^2(\mathbb{H})$, define*

$$\hat{f}(z, k) := \|\mathcal{P}_k \mathcal{F}_3 f(w, z)\|_{L^2(dw)}$$

for $z \in \mathbb{R}, k \in \{0, 1, \dots\}$. Then by the above calculation

$$\int_{\mathbb{H}} f(x)(-\Delta_{\mathbb{H}} f)(x) dx = \sum_{k=0}^{\infty} \int_{\mathbb{R}} |\zeta|(2k+1) \left| \hat{f}(\zeta, k) \right|^2 d\zeta. \tag{4.4}$$

Thus, by the spectral theorem, we can define a functional calculus for the operator $-\Delta_{\mathbb{H}}$. Indeed, for any Borel function $\varphi : [0, \infty) \to \mathbb{R}$, we define

$$\varphi(-\Delta_{\mathbb{H}}) := \varphi\left(\lambda_k(\zeta)\right), \tag{4.5}$$

where the right hand side represents the operator $\mathcal{F}_3^* \sum_k \varphi(\lambda_k(\zeta)) \mathcal{P}_k \mathcal{F}_3$ with domain

$$\mathcal{D}(\varphi(-\Delta_{\mathbb{H}})) = \left\{ f \in L^2(\mathbb{H}) : \sum_{k=0}^{\infty} \int_{\mathbb{R}} \varphi\left(|\zeta|(2k+1)\right) \left| \hat{f}(\zeta, k) \right|^2 d\zeta < \infty \right\}.$$

For $f \in \mathcal{D}(\varphi(-\Delta_{\mathbb{H}}))$, we have

$$\int_{\mathbb{H}} f(x) \varphi(-\Delta_{\mathbb{H}}) f(x) dx = \sum_{k=0}^{\infty} \int_{\mathbb{R}} \varphi\left(|\zeta|(2k+1)\right) \left| \hat{f}(\zeta, k) \right|^2 d\zeta. \tag{4.6}$$

To prove the main result of this section, we will also make use of the Min-Max principle for general self-adjoint operators, which we briefly recall now (see for example [110]).

Theorem 4.3.2 (Min-Max Principle)**.** *Let $\mathbf{L}$ be a self-adjoint operator on a Hilbert space that is bounded from below, i.e. $\mathbf{L} \geq \kappa I$ for some $\kappa \in \mathbb{R}$. Define, for $k \in \mathbb{N}$,*

$$\mu_k(\mathbf{L}) = \sup_{\varphi_1, \dots, \varphi_{k-1}} U_{\mathbf{L}}(\varphi_1, \dots, \varphi_{k-1})$$

where

$$U_{\mathbf{L}}(\varphi_1,\ldots,\varphi_m) = \inf_{\substack{\psi\in\mathcal{D}(\mathbf{L}):\|\psi\|=1\\ \psi\in[\varphi_1,\ldots,\varphi_m]^\perp}} (\psi,\mathbf{L}\psi).$$

Then exactly one of the following holds:

(a) $\mu_k(\mathbf{L})$ *is the* k*-th eigenvalue below the bottom of the essential spectrum, counting multiplicity;*

(b) $\mu_k(\mathbf{L})$ *is the bottom of the essential spectrum,* $\mu_k(\mathbf{L}) = \mu_{k+1}(\mathbf{L}) = \mu_{k+2}(\mathbf{L}) = \ldots,$ *and there are at most* $k-1$ *eigenvalues (counting multiplicity) below* $\mu_k(\mathbf{L})$.

We are now in a position to state and prove the main result of this section:

Theorem 4.3.3. *Suppose* V *is in* $L^\infty_{loc}(\mathbb{H})$ *and is bounded from below. Suppose also that for every* $L>0$ *there exists* $R_L>0$ *such that*

$$V(x)\geq L \quad whenever \quad d(x)\geq R_L.$$

Then the operator $\mathcal{L} = -\Delta_{\mathbb{H}} + V$ *on* $L^2(\mathbb{H})$ *has empty essential spectrum. In particular it has a purely discrete set of eigenvalues and a complete set of eigenfunctions.*

Proof. Let $\mu_m(\mathcal{L})$ be as in the Min-Max principle (i.e. Theorem 4.3.2). To prove that $\mathcal{L}$ has discrete spectrum, by the Min-Max principle, it is sufficient to show that $\mu_m(\mathcal{L})\to\infty$ as $m\to\infty$.

Suppose W is a bounded function, supported in a compact set $\Omega\subset\mathbb{R}^3$, so that

$$\sup_{x\in\Omega}|W(x)|\leq M,$$

for some $M\in\mathbb{R}$. For $\varepsilon>0$ consider the operator

$$W\phi_\varepsilon(-\Delta_{\mathbb{H}}),$$

where $\phi_\varepsilon(t) = (\varepsilon t^2+t+1)^{-1}$ for $t\in\mathbb{R}$. Using (4.5) and (4.6), and following [68], we have that

$$
\begin{aligned}
\mathbf{Tr}(W^2\phi_\varepsilon^2(-\Delta_{\mathbb{H}})) &\leq M^2\frac{1}{2\pi}\int_\Omega\int_{-\infty}^{\infty}\sum_{k=0}^{\infty}\phi_\varepsilon^2(\lambda_k(\zeta))\pi_k(w,w)d\zeta dw\\
&= M^2\frac{|\Omega|}{4\pi^2}\sum_{k=0}^{\infty}\int_{-\infty}^{\infty}\phi_\varepsilon^2(\lambda_k(\zeta))|\zeta|d\zeta\\
&= M^2\frac{|\Omega|}{2\pi^2}\sum_{k=0}^{\infty}\int_0^{\infty}\frac{|\zeta|}{\left(\varepsilon\lambda_k(\zeta)^2+\lambda_k(\zeta)+1\right)^2}d\zeta\\
&= M^2\frac{|\Omega|}{2\pi^2}\sum_{k=0}^{\infty}\int_0^{\infty}\frac{|\zeta|}{\left(\varepsilon|\zeta|^2(2k+1)^2+|\zeta|(2k+1)+1\right)^2}d\zeta\\
&< \infty.
\end{aligned}
$$

Since $W\phi_\varepsilon(-\Delta_{\mathbb{H}})$ is positive and self-adjoint on $L^2(\mathbb{H})$, we thus have that $W\phi_\varepsilon(-\Delta_{\mathbb{H}})$ is Hilbert-Schmidt for all $\varepsilon > 0$. Moreover,

$$
\left(\varepsilon\lambda_k(\zeta)^2+\lambda_k(\zeta)+1\right)^{-1} \to \left(\lambda_k(\zeta)+1\right)^{-1} \equiv \phi_0(\lambda_k(\zeta))
$$

in $L^\infty(\mathbb{R})\times l^\infty(\mathbb{N}\cup\{0\})$ as $\varepsilon \to 0$. Indeed

$$
\begin{aligned}
|\left(\phi_0-\phi_\varepsilon\right)\left(\lambda_k(\zeta)\right)| &= \left|\frac{1}{\lambda_k(\zeta)+1}-\frac{1}{\varepsilon\lambda_k(\zeta)^2+\lambda_k(\zeta)+1}\right|\\
&= \frac{\varepsilon\lambda_k(\zeta)^2}{(\lambda_k(\zeta)+1)(\varepsilon\lambda_k(\zeta)^2+\lambda_k(\zeta)+1)}\\
&\leq \varepsilon\frac{\lambda_k(\zeta)^2}{(\lambda_k(\zeta)+1)^2} \leq \varepsilon.
\end{aligned}
$$

Therefore $W\phi_0(-\Delta_{\mathbb{H}})$ is a norm-limit of Hilbert-Schmidt operators:

$$
\begin{aligned}
\|W\left(\phi_0-\phi_\varepsilon\right)(-\Delta_{\mathbb{H}})\psi\|_2^2 &\leq M^2\sum_{k=0}^{\infty}\int_{\mathbb{R}}\left(\phi_0-\phi_\varepsilon\right)^2\left(\lambda_k(\zeta)\right)\left|\hat{\psi}(\zeta,k)\right|^2 d\zeta\\
&\leq \varepsilon^2M^2\sum_{k=0}^{\infty}\int_{\mathbb{R}}\left|\hat{\psi}(\zeta,k)\right|^2 d\zeta\\
&= \varepsilon^2M^2\|\psi\|_2^2,
\end{aligned}
$$

using (4.6). We can thus conclude that $W\phi_0(-\Delta_{\mathbb{H}})$ is a compact operator, or in other words that W is relatively compact with respect to $-\Delta_{\mathbb{H}}$.

Since $W\phi_0(-\Delta_{\mathbb{H}})$ is compact, by Weyl's Theorem (see Corollary 2 of Theorem XIII.14 of [110]),

$$\sigma_{ess}(-\Delta_{\mathbb{H}} + W) = \sigma_{ess}(-\Delta_{\mathbb{H}}) = [0, \infty).$$

Therefore by the Min-Max principle $\mu_m(-\Delta_{\mathbb{H}} + W) \geq -1$ for m sufficiently large.

Now, given $a > 0$, define V_a by

$$V_a(x) = \min\{V(x), a+1\} - a - 1.$$

Then V_a has compact support, since $V(x) \to \infty$ as $d(x) \to \infty$, and V_a is bounded since V is locally bounded. Thus, by the above considerations, $\mu_m(-\Delta_{\mathbb{H}} + V_a) \geq -1$ for large m. Finally, since

$$\mu_m(\mathcal{L}) \geq \mu_m\left(-\Delta_{\mathbb{H}} + V_a\right) + a + 1,$$

we see that $\mu_m(\mathcal{L}) \geq a$ for large m. Since a is arbitrary we reach the desired conclusion. $\square$

As mentioned in the introduction to this chapter, we are actually interested in probability measures on $\mathbb{H}$ of the form

$$\mu_U(dx) := \frac{e^{-U(x)}}{Z} dx, \tag{4.7}$$

where $Z = \int_{\mathbb{H}} e^{-U(x)} dx < \infty$, with which we can associate a positive and self-adjoint operator $\mathcal{L} = -\Delta_{\mathbb{H}} + \nabla_{\mathbb{H}} U \cdot \nabla_{\mathbb{H}}$ on $L^2(d\mu_U)$. We will pay particular attention to the case when U is a power of the distance function (see Remark 4.3.5 and Section 4.4.2).

In the corollary below, we use the above theorem to obtain some conditions on U that ensure the operator $\mathcal{L}$ acting on $L^2(d\mu_U)$ has empty essential spectrum.

Corollary 4.3.4. *Let μ_U be a probability measure on $\mathbb{H}$, and suppose that U is twice differentiable almost everywhere. Suppose also that*

$$V = \frac{1}{4}|\nabla_{\mathbb{H}} U|^2 - \frac{1}{2}\Delta_{\mathbb{H}} U$$

is in $L^\infty_{loc}(\mathbb{H})$*, is bounded from below, and is such that* $V(x) \to \infty$ *as* $d(x) \to \infty$*. Let* $\mathcal{L} = -\Delta_{\mathbb{H}} + \nabla_{\mathbb{H}} U \cdot \nabla_{\mathbb{H}}$*, so that* $\mathcal{L}$ *is a positive self-adjoint operator on* $L^2(d\mu_U)$*. Then* $\sigma_{ess}(\mathcal{L}) = \emptyset$*.*

Proof. This follows from the observation that for $g = fe^{-\frac{1}{2}U}$,

$$\begin{aligned}\int_{\mathbb{H}} f(-\Delta_{\mathbb{H}} + \nabla_{\mathbb{H}} U \cdot \nabla_{\mathbb{H}}) f d\mu_U &= \int_{\mathbb{H}} |\nabla_{\mathbb{H}} f|^2 d\mu_U \\ &= \int_{\mathbb{H}} g\left(-\Delta_{\mathbb{H}} + \frac{1}{4}|\nabla_{\mathbb{H}} U|^2 - \frac{1}{2}\Delta_{\mathbb{H}} U\right) g dx \\ &= \int_{\mathbb{H}} g\left(-\Delta_{\mathbb{H}} + V\right) g dx.\end{aligned}$$

Hence the spectrum of the operator $\mathcal{L}$ in $L^2(d\mu_U)$ is contained within the spectrum of the operator $-\Delta_{\mathbb{H}} + V$ on $L^2(dx)$. Since we have assumed $V(x) \in L^\infty_{loc}$ and $V \to \infty$ as $d \to \infty$, the result follows by Theorem 4.3.3. □

Remark 4.3.5. *Suppose* $U(x) = \alpha d^p(x)$*, with* $p \in (1, \infty)$ *and* $\alpha > 0$*. In this case we can formally calculate that*

$$\begin{aligned}V = \frac{1}{4}|\nabla_{\mathbb{H}} U|^2 - \frac{1}{2}\Delta_{\mathbb{H}} U &= \frac{\alpha^2 p^2}{4} d^{2(p-1)} |\nabla_{\mathbb{H}} d|^2 \\ &\quad - \frac{\alpha p(p-1)}{2} d^{p-2} |\nabla_{\mathbb{H}} d|^2 - \frac{\alpha p}{2} d^{p-1} \Delta_{\mathbb{H}} d \\ &= \frac{\alpha^2 p^2}{4} d^{2(p-1)} - \frac{\alpha p(p-1)}{2} d^{p-2} - \frac{\alpha p}{2} d^{p-1} \Delta_{\mathbb{H}} d\end{aligned}$$

almost everywhere, where we have used Proposition 3.2.16. As noted in Section 3.2.4, we must understand this expression in the sense of distributions, since $\Delta_{\mathbb{H}} d$ *is not defined on the centre of the group. Hence it is not straight forward to apply Corollary 4.3.4 to conclude that* $\mathcal{L} = -\Delta_{\mathbb{H}} + \nabla_{\mathbb{H}} U \cdot \nabla_{\mathbb{H}}$ *has purely discrete spectrum.*

We note that this is in contrast to the Euclidean setting, where such a strategy would yield the desired result, at least for $p \geq 2$*. This is because in* $\mathbb{R}^n$ *we can explicitly write* $\Delta d = \frac{n-1}{d}$*, where* d *is the Euclidean distance and* Δ *the standard Laplacian.*

To get around this problem, in the next section we take an alternative approach, which is also more general in that it can easily include the case of general H-type groups. The above techniques and ideas are not so easily extended to general H-type groups since we do not have such a representation of the sub-Laplacian as the one used above in the general case.

4.4 Empty essential spectrum via functional inequalities

The relationship between functional inequalities and the spectrum of operators is a very interesting and much studied one. Indeed, if (Ω, μ) is a probability space and $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is a positive self-adjoint operator on $L^2(\mu)$, then it is well-known that $\mathcal{L}$ has a gap at the bottom of its spectrum if and only if there exists a constant $c_0 > 0$ such that

$$\mu \left(f - \mu(f)\right)^2 \leq c_0 \mathcal{E}(f, f),$$

where $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is the Dirichlet form[2] associated to $\mathcal{L}$ i.e. the closure of the form

$$\mathcal{E}(f, g) = \mu(f \mathcal{L} g), \quad f, g \in \mathcal{D}(\mathcal{L}).$$

More recently this relationship has been further illustrated by the work of F. Cipriani ([47]) and F. Y. Wang ([124]) in which functional inequalities are introduced that characterise the essential spectra of operators under very general conditions. In this section we aim to use functional inequalities to overcome the problems encountered in Remark 4.3.5.

4.4.1 Super-Poincaré inequalities

To state the results of Wang and Cipriani in full generality, we first need the following two technical definitions.

Definition 4.4.1. *A topological space Ω is a* Lusin *space if Ω is homeomorphic to a Borel subset of a compact metric space.*

Remark 4.4.2. *It should be noted that, as shown in Theorem 82.5 of [112], every complete metric space is a Lusin space. In particular, any H-type group $\mathbb{G}$ is a Lusin space.*

[2] Recall that a Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a densely defined, closed quadratic form on $L^2(\mu)$ such that $\mathcal{E}(f \wedge 1, f \wedge 1) \leq \mathcal{E}(f, f)$ for all $f \in \mathcal{D}(\mathcal{E})$.

Definition 4.4.3. *Let Ω be a Lusin space, and μ a positive Radon measure* [3] *on Ω having full topological support. A positive, self-adjoint operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ on $L^2(\mu)$, with associated closed Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ defined by*

$$\mathcal{E}(f, g) = \mu(f\mathcal{L}g), \qquad f, g \in \mathcal{D}(\mathcal{L}),$$

is called a Persson operator *if*

$$\inf \sigma_{ess}(\mathcal{L}) = \sup \{\Sigma(K) : K \subset \Omega \text{ is compact}\}$$

where

$$\Sigma(K) := \inf \left\{ \frac{\mathcal{E}(f, f)}{\|f\|_2^2} : f \in \mathcal{D}(\mathcal{E}),\ \operatorname{supp}(f) \subset K^c \right\}.$$

This class of operator was introduced by A. Persson in [108]. The result below is a combination of the independent work of Wang and Cipriani, and is explicitly stated in [126].

Theorem 4.4.4 (Wang/Cipriani)**.** *Let Ω be a Lusin space, μ a positive Radon measure on Ω having full topological support, and $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ a Persson operator on $L^2(\mu)$. Then the inequality*

$$\mu(f^2) \leq r\mu(f\mathcal{L}f) + \beta(r)(\mu|f|)^2, \qquad \forall r > r_0, \quad f \in \mathcal{D}(\mathcal{L}), \tag{4.8}$$

for some decreasing function $\beta : (r_0, \infty) \to (0, \infty)$ and $r_0 \geq 0$ holds if and only if $\sigma_{ess}(\mathcal{L}) \subset [r_0^{-1}, \infty)$. In particular, (4.8) *is satisfied with $r_0 = 0$ if and only if $\sigma_{ess}(\mathcal{L}) = \emptyset$.*

Inequality (4.8) is known as a *super-Poincaré* inequality. In a similar way to the generalisation of the standard logarithmic Sobolev inequality to the LS_q inequality, we can generalise the super-Poincaré inequality to a *q-super-Poincaré inequality*:

Definition 4.4.5. *Let (Ω, μ) be a probability space, equipped with a metric $d : \Omega \times \Omega \to [0, \infty)$. For $q \in (1, 2]$, we say that μ satisfies a q*-super-Poincaré inequality*, or SP_q for*

[3] μ is a Radon measure if $\mu(A) = \sup\{\mu(K) : K \subset A,\ K \text{ compact}\}$ for all Borel sets A and every point of Ω has a neighbourhood of finite measure.

short, with constant r_0, if

$$\mu|f|^q \leq r\mu\,|\nabla f|^q + \beta(r)\left(\mu|f|^{\frac{q}{2}}\right)^2, \qquad \forall r > r_0, \qquad (SP_q) \tag{4.9}$$

for all locally Lipschitz functions f and some $\beta : (r_0, \infty) \to (0, \infty)$, where $|\nabla f|(x) \equiv \limsup_{d(x,y)\to 0} |f(x) - f(y)|/d(x, y)$.

Remark 4.4.6. *For the remainder of this chapter we will be working in an H-type group $\mathbb{G}$ equipped with the Carnot-Carathéodory distance d and a probability measure $\mu_U(dx) := Z^{-1}e^{-U(x)}dx$. In this case $|\nabla_{\mathbb{G}} f|(x) = \limsup_{d(x,y)\to 0} |f(x) - f(y)|/d(x, y)$ and for $\mathcal{L} = -\Delta_{\mathbb{G}} + \nabla_{\mathbb{G}} U \cdot \nabla_{\mathbb{G}}$ we have $\mu_U(f\mathcal{L}f) = \mu_{\mathbb{G}}|\nabla_{\mathbb{G}} f|^2$.*

4.4.2 Applications to H-type groups

Let $\mathbb{G}$ be an H-type group as usual, equipped with the Carnot-Carathéodory distance d. Let μ_p be the probability measure on $\mathbb{G}$ defined by

$$\mu_p(dx) := \frac{e^{-\alpha d^p(x)}}{Z}dx \tag{4.10}$$

where $Z = \int e^{-\alpha d^p(x)}dx$ is the normalisation constant, and $p \in (1, \infty), \alpha > 0$. Define

$$\mathcal{L}_p := -\Delta_{\mathbb{G}} + \nabla_{\mathbb{G}}(\alpha d^p) \cdot \nabla_{\mathbb{G}} = -\Delta_{\mathbb{G}} + \alpha p d^{p-1}\nabla_{\mathbb{G}} d \cdot \nabla_{\mathbb{G}} \tag{4.11}$$

as a positive self-adjoint operator acting on $L^2(\mu_p)$. The associated Dirichlet form $\mathcal{E}_p(f, g)$ is then given by

$$\mathcal{E}_p(f, g) = \mu_p(f\mathcal{L}_p g) = \int_{\mathbb{G}} \nabla_{\mathbb{G}} f \cdot \nabla_{\mathbb{G}} g d\mu_p. \tag{4.12}$$

We are thus in the situation of Remark 4.4.6 with $U = \alpha d^p$.

We aim to prove the following:

Theorem 4.4.7. *For any $p > 1$ the positive self-adjoint operator $\mathcal{L}_p$ on $L^2(\mu_p)$ given by* (4.11) *has purely discrete spectrum i.e. $\sigma_{ess}(\mathcal{L}_p) = \emptyset$.*

The idea is to use Theorem 4.4.4. It is clear that we first need to show that $\mathcal{L}_p$ is a

Persson operator. We make use of the following very general result stated in the setting of Dirichlet forms, proved by G. Grillo in [62] (and also stated explicitly in [47]).

Theorem 4.4.8 (Grillo). *Let (Ω, μ) be a locally compact, separable metric space, and $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ a regular*[4]*, strongly local*[5] *Dirichlet form on $L^2(\Omega)$, with associated positive self-adjoint operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$.*

Define the associated intrinsic pseudo-metric ρ on Ω by

$$\rho(x, y) := \sup \{|f(x) - f(y)| : f \in \mathcal{D}(\mathcal{E}) \cap C_0(\Omega), \Gamma(f, f) \leq 1\}$$

where for $f \in \mathcal{D}(\mathcal{E}), \Gamma(f, f)$ is such that

$$\int_\Omega g\Gamma(f, f)d\mu = \mathcal{E}(gf, f) - \frac{1}{2}\mathcal{E}(f^2, g), \qquad f, g \in \mathcal{D}(\mathcal{E}) \cap C_0(\Omega).$$

Suppose ρ is a true metric generating the original topology of Ω. Then the operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is a Persson operator.

Corollary 4.4.9. *The operator $\mathcal{L}_p$ given by* (4.11) *acting on $L^2(\mu_p)$ is a Persson operator.*

Proof. First of all it is clear that $\mathbb{G}$ is a locally compact separable metric space. Moreover, $(\mathcal{E}_p, \mathcal{D}(\mathcal{E}_p))$ is a regular Dirichlet form. Indeed, $C_0^2(\mathbb{G})$ is dense in $\mathcal{D}(\mathcal{E}_p)$, with respect to the norm induced by $\mathcal{E}_p$, and in $C_0(\mathbb{G})$ with respect to the uniform norm. It is also clear that it is strongly local by (4.12). Finally we have that

$$\int_\mathbb{G} g|\nabla_\mathbb{G} f|^2 d\mu_p = \int_\mathbb{G} gf\mathcal{L}_p f d\mu_p - \frac{1}{2}\int_\mathbb{G} g\mathcal{L}_p f^2 d\mu_p.$$

Thus the associated intrinsic pseudo-metric is given by

$$\rho(x, y) = \sup \left\{|f(x) - f(y)| : |\nabla_\mathbb{G} f|^2 \leq 1\right\}.$$

This is nothing more than the Carnot-Carathéodory distance (by definition), so that ρ is

[4] $\mathcal{E}$ is regular if $\mathcal{D}(\mathcal{E}) \cap C_0(\Omega)$ is dense in $C_0(\Omega)$ with respect to the uniform norm, and in $\mathcal{D}(\mathcal{E})$ with respect to the norm induced by $\mathcal{E}_1(f, g) = \mathcal{E}(f, g) + (f, g)$ — $\mathcal{D}(\mathcal{E}) \cap C_0(\Omega)$ is said to be a *core* of $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$.

[5] $\mathcal{E}$ is strongly local if $\mathcal{E}(f, g) = 0$ whenever f is constant on $\mathrm{supp}(g)$.

indeed a true metric generating the original topology of $\mathbb{G}$. Hence the result follows from Theorem 4.4.8. □

The next result we prove on route to Theorem 4.4.7 is that the measures μ_p satisfy certain super-Poincaré inequalities.

Theorem 4.4.10. *Let μ_p be the probability measure on $\mathbb{G}$ given by* (4.10).

(i) *Suppose $p \geq 2$. Then μ_p satisfies an SP_q inequality with constant $r_0 = 0$ i.e.*

$$\mu_p|f|^q \leq r\mu|\nabla_{\mathbb{G}} f|^q + \beta(r)\left(\mu|f|^{\frac{q}{2}}\right)^2, \qquad \forall r > 0,$$

where $\frac{1}{p}+\frac{1}{q} = 1$, for some function $\beta : (0,\infty) \to (0,\infty)$ and for all locally Lipschitz functions f.

(ii) *Suppose $p \in (1,2]$. Then μ_p satisfies an SP_2 inequality with constant $r_0 = 0$ i.e.*

$$\mu_p(f^2) \leq r\mu|\nabla_{\mathbb{G}} f|^2 + \beta(r)\left(\mu|f|\right)^2, \qquad \forall r > 0,$$

for some function $\beta : (0,\infty) \to (0,\infty)$ and for all locally Lipschitz functions f.

Proof. The idea is to pass from a logarithmic Sobolev inequality for the measure μ_p, which is true by Theorem 3.2.23, to a super-Poincaré inequality by adapting the methods of F. Y. Wang described in [124].

We first deal with the case $p \geq 2$. Without loss of generality we may assume that $f \geq 0$. By part (i) of Theorem 3.2.23, we have that μ_p satisfies an LS_q inequality i.e. there exists a constant c such that

$$\mu_p\left(f^q \log \frac{f^q}{\mu_p f^q}\right) \leq c\mu_p|\nabla_{\mathbb{G}} f|^q \tag{4.13}$$

where $\frac{1}{p}+\frac{1}{q} = 1$.

Let $g : (0,\infty) \to \mathbb{R}$ be given by $g(\xi) = t\xi - \xi\log\left(\frac{\xi^2}{a}\right)$ for any $t, a > 0$. By simple differentiation, it can be shown that

$$\max_{\{\xi>0\}} g(\xi) = 2\sqrt{ae^{t-2}}. \tag{4.14}$$

Indeed $g'(\xi) = t - \log\left(\frac{\xi^2}{a}\right) - 2$ so that g is maximum at $\xi_0 = \sqrt{ae^{t-2}}$.

Suppose that $\mu_p(f^{\frac{q}{2}}) = 1$, and set $a = \mu_p(f^q)$. Then by (4.14), for all $t > 0$,

$$\begin{aligned} tf^{\frac{q}{2}} - f^{\frac{q}{2}} \log\left(\frac{f^q}{a}\right) &\leq 2\sqrt{ae^{t-2}} \\ \Rightarrow tf^q - f^q \log\left(\frac{f^q}{a}\right) &\leq 2\sqrt{ae^{t-2}} f^{\frac{q}{2}} \\ \Rightarrow \mu_p\left(f^q \log\frac{f^q}{a}\right) &\geq ta - 2\sqrt{ae^{t-2}}, \end{aligned} \tag{4.15}$$

using the fact that $f \geq 0$ and $\mu_p(f^{\frac{q}{2}}) = 1$. Setting $b = \mu_p|\nabla_{\mathbb{G}} f|^q$, by (4.13), we then have

$$ta - 2\sqrt{ae^{t-2}} - cb \leq 0.$$

Solving this quadratic inequality gives

$$\sqrt{a} \leq \frac{2\sqrt{e^{t-2}}}{2t} + \frac{\sqrt{4e^{t-2} + 4tcb}}{2t}$$

for $t > 0$, so that

$$a \leq \frac{2c}{t} b + 4\frac{e^{t-2}}{t^2}$$

for $t > 0$. In other words

$$\mu_p(f^q) \leq \frac{2c}{t} \mu_p|\nabla_{\mathbb{G}} f|^q + 4\frac{e^{t-2}}{t^2}$$

for all $t > 0$ and f such that $\mu_p(f^{\frac{q}{2}}) = 1$. Replacing f by $\frac{f}{\mu_p(f^{q/2})^{2/q}}$ yields

$$\mu_p(f^q) \leq \frac{2c}{t} \mu_p|\nabla_{\mathbb{G}} f|^q + 4\frac{e^{t-2}}{t^2} \left(\mu_p(f^{\frac{q}{2}})\right)^2$$

for all $t > 0$. Taking $r = \frac{2c}{t}$ we see that SP_q holds, so that we have proved part (i).

In the case where $p \in (1, 2)$, we no longer have an inequality of the type (4.13). However, by part (ii) of Theorem 3.2.23 there exists a constant $c \in (0, \infty)$ such that

$$\mu_p\left(f^2 \left[\log(1 + f^2)\right]^\theta\right) \leq c\mu_p|\nabla_{\mathbb{G}} f|^2 + (\log 2)^\theta, \qquad \mu_p(f^2) = 1, \tag{4.16}$$

where $\theta = \frac{2(p-1)}{p}$. In this case we instead let $g : (0, \infty) \to \mathbb{R}$ be given by $g(\xi) = t\xi - \xi \left[\log\left(1 + \frac{\xi^2}{a}\right)\right]^\theta$ for $t, a > 0$, and claim that

$$\sup_{\{\xi > 0\}} g(\xi) \leq t\sqrt{a(e^{t^{1/\theta}} - 1)}. \tag{4.17}$$

Indeed, since g is smooth we may differentiate to get

$$g'(\xi) = t - \left[\log\left(1 + \frac{\xi^2}{a}\right)\right]^\theta - \frac{2\theta\xi^2}{a + \xi^2}\left[\log\left(1 + \frac{\xi^2}{a}\right)\right]^{\theta - 1}.$$

If we then let $\xi_0 > 0$ be such that

$$\left[\log\left(1 + \frac{\xi_0^2}{a}\right)\right]^\theta + \frac{2\theta\xi_0^2}{a + \xi_0^2}\left[\log\left(1 + \frac{\xi_0^2}{a}\right)\right]^{\theta - 1} = t, \tag{4.18}$$

we see $g(\xi) \leq g(\xi_0)$ for all $\xi > 0$. Now, using (4.18),

$$\begin{aligned} g(\xi_0) &= \xi_0 \frac{2\theta\xi_0^2}{a + \xi_0^2}\left[\log\left(1 + \frac{\xi_0^2}{a}\right)\right]^{\theta - 1} \\ &\leq t\xi_0. \end{aligned}$$

Moreover, again by (4.18), we have

$$\begin{aligned} \left[\log\left(1 + \frac{\xi_0^2}{a}\right)\right]^\theta &\leq t \\ \Rightarrow \xi_0 &\leq \sqrt{a\left(e^{t^{1/\theta}} - 1\right)}, \end{aligned}$$

which proves the claim (4.17). Proceeding now in a very similar way as in the proof of part (i), we arrive at an SP_2 inequality. □

The final result we need is that SP_q inequalities are stronger than SP_2 inequalities (at least when the dimension of the underlying space is finite).

Lemma 4.4.11. *Suppose an arbitrary probability measure μ on $\mathbb{G}$ satisfies an SP_q inequal-*

ity with $q \in (1, 2]$ *and constant* $r_0 = 0$ *i.e.*

$$\mu|f|^q \leq r\mu|\nabla_{\mathbb{G}} f|^q + \beta(r)\left(\mu|f|^{\frac{q}{2}}\right)^2, \qquad \forall r > 0,$$

for some $\beta : (0, \infty) \to (0, \infty)$ *and all locally Lipschitz functions* f*. Then* μ *also satisfies an* SP_2 *inequality with constant* $r_0 = 0$.

Proof. As usual, without loss of generality we may suppose $f \geq 0$. Let $q < 2$ (there is nothing to prove if $q = 2$). Applying the SP_q inequality to $f^{\frac{2}{q}}$ yields,

$$\mu(f^2) \leq r\mu\left|\nabla_{\mathbb{G}} f^{\frac{2}{q}}\right|^q + \beta(r)\left(\mu f\right)^2, \qquad \forall r > 0.$$

Therefore for all $r > 0$, we have by Hölder's inequality followed by Young's inequality,

$$\begin{aligned}\mu(f^2) &\leq \frac{2^q r}{q^q}\mu\left(f^{2-q}|\nabla_{\mathbb{G}} f|^q\right) + \beta(r)\left(\mu f\right)^2 \\ &\leq \frac{2^{q-1} r}{q^{q-1}}\tau^{\frac{2-q}{q}}\mu|\nabla_{\mathbb{G}} f|^2 + \frac{2^{q-1} r(2-q)}{q^q}\tau^{-1}\mu(f^2) + \beta(r)\left(\mu f\right)^2\end{aligned}$$

for all $\tau > 0$. Taking $\tau = \frac{2^q r(2-q)}{q^q}$ we see that

$$\begin{aligned}\frac{1}{2}\mu(f^2) &\leq \frac{2^{q-1} r}{q^{q-1}}\tau^{\frac{2-q}{q}}\mu|\nabla_{\mathbb{G}} f|^2 + \beta(r)\left(\mu f\right)^2 \\ &= \frac{2^{q-1}}{q^{q-1}}\left(\frac{2}{q}\right)^{2-q}(2-q)^{\frac{2-q}{q}} r^{\frac{2}{q}}\mu|\nabla_{\mathbb{G}} f|^2 + \beta(r)\left(\mu f\right)^2 \\ \Rightarrow \mu(f^2) &\leq \frac{4}{q}(2-q)^{\frac{2-q}{q}} r^{\frac{2}{q}}\mu|\nabla_{\mathbb{G}} f|^2 + 2\beta(r)\left(\mu f\right)^2.\end{aligned}$$

Taking $s = \frac{4}{q}(2-q)^{\frac{2-q}{q}} r^{\frac{2}{q}}$ we see that

$$\mu(f^2) \leq s\mu|\nabla_{\mathbb{G}} f|^2 + \tilde{\beta}(s)\left(\mu f\right)^2, \qquad s > 0,$$

where $\tilde{\beta}(s) = 2\beta\left((2-q)^{\frac{q-2}{2}}\left(\frac{qs}{4}\right)^{\frac{q}{2}}\right)$. □

Proof of Theorem 4.4.7. We can now combine all of the above results to arrive at Theorem 4.4.7. Indeed by Theorem 4.4.10 and Lemma 4.4.11, we have that the measures μ_p satisfy a super-Poincaré inequality with constant $r_0 = 0$ for all $p > 1$. Moreover, by Corollary 4.4.9, $\mathcal{L}_p$ is a Persson operator, so that we may conclude by applying Theorem 4.4.4. $\square$

Corollary 4.4.12. *Let $\rho_t(x, y)$ be the heat kernel at time t on an H-type group $\mathbb{G}$ i.e. $\rho_t(x, y)$ is the function (smooth by Hörmander's theorem) such that*

$$e^{t\Delta_{\mathbb{H}}} f(x) = \int_{\mathbb{G}} \rho_t(x, y) f(y) dy.$$

Let $\rho(x) := \rho_1(x, e)$ and define

$$\mathcal{L}_H := -\Delta_{\mathbb{G}} + \nabla_{\mathbb{G}} \log \rho \cdot \nabla_{\mathbb{G}}.$$

Then $\mathcal{L}_H$ is a positive self-adjoint operator on $L^2(\mu_H)$, where $\mu_H(dx) = \rho(x)dx$, and $\sigma_{ess}(\mathcal{L}_H) = \emptyset$, so that $\mathcal{L}_H$ has a purely discrete spectrum.

Remark 4.4.13. *$\mathcal{L}_H$ can be regarded as the natural Ornstein-Uhlenbeck generator on $\mathbb{G}$, as suggested in [17], and the resulting Markov process is the natural OU-process associated to the hypoelliptic diffusion on $\mathbb{G}$.*

Proof. It follows exactly as above, once we have recalled that μ_H satisfies a logarithmic Sobolev inequality i.e. there exists a constant c such that

$$\mu_H \left(f^2 \log \frac{f^2}{\mu_H f^2} \right) \leq c \mu_H |\nabla_{\mathbb{G}} f|^2$$

(see [9, 69] and [91]). $\square$

4.5 Spectral information for measures defined with the Kaplan distance

In the previous section we have focused on probability measures on an H-type group $\mathbb{G} = \mathbb{R}^{n+m}$ given by $\mu_p(dx) = Z^{-1} e^{-\alpha d^p(x)} dx$ for $p > 1$ and $\alpha > 0$, together with their associated generators, where d is the Carnot-Carathéodory distance. However, as we have mentioned in Section 3.2, there is another natural homogeneous distance function on an H-type group, namely the Kaplan distance $N : \mathbb{G} \to [0, \infty)$ given by

$$N(x) = \left(|w|^4 + 16|z|^2\right)^{\frac{1}{4}}, \quad \text{for } x = (w, z) \in \mathbb{R}^n \times \mathbb{R}^m, \tag{4.19}$$

which appears in the fundamental solution of the sub-Laplacian. Therefore, an obvious question to ask is whether one can replace the Carnot-Carathéodory distance with the Kaplan distance in the above work. At first glance such a question might seem simple, since all homogeneous metrics on $\mathbb{G}$ are equivalent (see Proposition 3.2.12). However, as we will see, this is not the case, and there are some fundamental differences between the two situations.

To make things precise, suppose now that we are working in an H-type group $\mathbb{G} = \mathbb{R}^{n+m}$ equipped with a probability measure

$$\nu_p(dx) := \frac{e^{-\alpha N^p(x)}}{Z} dx, \tag{4.20}$$

where $p \in (1, \infty)$, $\alpha > 0$ and $Z = \int e^{-\alpha N^p(x)} dx$ is the normalisation constant as usual. The associated positive self-adjoint operator on $L^2(\nu_p)$ is then given by

$$\mathcal{T}_p = -\Delta_{\mathbb{G}} + \alpha p N^{p-1} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}}. \tag{4.21}$$

The aim now is to gain some spectral information about these operators. We first try to apply the functional inequality approach of Section 4.4. The key idea there was to pass from a logarithmic Sobolev inequality to a super-Poincaré inequality. However, we immediately come up against a problem in the form of Theorem 6.3 from [69]:

Theorem 4.5.1 (Hebisch-Zegarliński)**.** *The measure ν_p on $\mathbb{G}$ given by* (4.20) *with $p > 1$ satisfies no LS_q inequality with $q \in (1, 2]$.*

Thus we cannot simply follow the proof of Theorem 4.4.7 to conclude that the operator $\mathcal{T}_p$ given by (4.21) has empty essential spectrum. Theorem 4.5.1 illustrates a major difference in the behaviour of the measures defined with the Carnot-Carathéodory distance and those defined with the Kaplan distance.

Given that it is not simple to apply the functional inequalities method, we may instead try to apply the results of Section 4.3 in the setting of the Heisenberg group $\mathbb{H}$ (in particular Corollary 4.3.4). However, there is a problem here too. Indeed, for $U = \alpha N^p$, by using Propositions 3.2.16 and 3.2.17, we can directly calculate that

$$\begin{aligned} V(x) &:= \left(\frac{1}{4}|\nabla_{\mathbb{H}}U|^2 - \frac{1}{2}\Delta_{\mathbb{H}}U\right)(x) \\ &= \left(\frac{\alpha^2 p^2}{4}N^{2(p-1)}|\nabla_{\mathbb{H}}N|^2 - \frac{\alpha p(p-1)}{2}N^{p-2}|\nabla_{\mathbb{H}}N|^2 - \frac{\alpha p}{2}N^{p-1}\Delta_{\mathbb{H}}N\right)(x) \\ &= \frac{\alpha^2 p^2}{4}N^{2p-4}(x)\|x\|^2 - \left(\frac{\alpha p(p-1)}{2} + \frac{\alpha p}{2}(Q-1)\right)N^{p-4}(x)\|x\|^2, \end{aligned}$$

for $x = (w, z) \in \mathbb{H}\backslash\{0\} = \mathbb{R}^2 \times \mathbb{R}$, and where $\|x\| = |w|$. It is then clear that $V(x) = 0$ for all $x = (0, z) \in \mathbb{H}$, so that it is certainly not true that $V(x) \to \infty$ as $x \to \infty$.

In view of these two observations, it seems that the problem of gaining spectral information about the operator $\mathcal{T}_p$ given by (4.21) is an interesting one. We therefore start by asking whether such operators have a spectral gap. This question is completely answered by Theorems 4.5.2 and 4.5.5 below.

Theorem 4.5.2. *If $p < 2$, then the measure ν_p given by* (4.20) *does not satisfy a spectral gap inequality. In particular the operator $\mathcal{T}_p$ given by* (4.21) *does not have a spectral gap, and hence it does not have empty essential spectrum.*

To prove this, we make use of the following lemma, quoted from [69].

Lemma 4.5.3. *Let f be a smooth function on $\mathbb{G}$ and d the Carnot-Carathéodory distance as usual. Then at points $x_0 \in \mathbb{G}$ such that $(\nabla_{\mathbb{G}} f)(x_0) = 0$ we have*

$$|f(x) - f(x_0)| \leq O\left(d^2(x, x_0)\right)$$

for all $x \in \mathbb{G}$.

Proof. Let x, x_0 be arbitrary points in $\mathbb{G}$ and $\gamma : [0, 1] \to \mathbb{G}$ a horizontal curve joining x

and x_0 which realises $d(x, x_0)$. Then

$$\begin{aligned} |f(x) - f(x_0)| &\leq \int_0^1 \left| \frac{d}{ds} f(\gamma(s)) \right| ds \\ &\leq \int_0^1 |\nabla_{\mathbb{G}} f(\gamma(s))| \, |\gamma'| ds \\ &\leq |\gamma| \sup_{s \in [0,1]} |\nabla_{\mathbb{G}} f(\gamma(s))|. \end{aligned}$$

Put $r = d(x, x_0)$, and $B_r(x) = \{y \in \mathbb{G} : d(y, x) \leq r\}$. Since $|\gamma| = r$, $\gamma(s) \in B_r(x)$ for $s \in [0, 1]$, and thus

$$|f(x) - f(x_0)| \leq r \sup_{y \in B_r(x)} |\nabla_{\mathbb{G}} f(y)|.$$

Since f is smooth, the supremum is finite. Now suppose in addition that x_0 is such that $(\nabla_{\mathbb{G}} f)(x_0) = 0$. By applying the above argument to the components of $\nabla_{\mathbb{G}} f$, it follows that

$$\sup_{y \in B_r(x)} |\nabla_{\mathbb{G}} f(y)| \leq r \sup_{y \in B_r(x)} |\nabla_{\mathbb{G}} \nabla_{\mathbb{G}} f(y)|,$$

so that

$$|f(x) - f(x_0)| \leq r^2 \sup_{y \in B_r(x)} |\nabla_{\mathbb{G}} \nabla_{\mathbb{G}} f(y)|.$$

□

Proof of Theorem 4.5.2. Let $p < 2$ and suppose for a contradiction that there exists a constant c_0 such that

$$\nu_p(f^2) - (\nu_p f)^2 \leq c_0 \nu_p |\nabla_{\mathbb{G}} f|^2 \tag{4.22}$$

for all locally Lipschitz functions f.

Fix $x_0 = (0, z) \in \mathbb{G}$ for $z \in \mathbb{R}^m \setminus \{0\}$. Then $|\nabla_{\mathbb{G}} N(x_0)| = \frac{\|x_0\|}{N(x_0)} = 0$ by Proposition 3.2.16, so that $\nabla_{\mathbb{G}} N(x_0) = 0$. Similarly $\nabla_{\mathbb{G}} N(-x_0) = 0$. Let $r_0 > 0$ be small enough so that $0 \notin B_{r_0}(x_0) = \{y' \in \mathbb{G} : d(y', x_0) \leq r_0\}$. Then N is smooth on $B_{r_0}(x_0)$, and by Lemma 4.5.3 there exists a constant C_1 such that

$$|N(y) - N(x_0)| \leq C_1 r_0^2, \tag{4.23}$$

for all $y \in B_{r_0}(x_0)$. The same holds for $y \in B_{r_0}(-x_0)$. We now dilate by a factor of $t > 0$. Since N is homogeneous, we have that

$$|N(y) - N(\delta_t(x_0))| = t|N(\delta_{t^{-1}}(y)) - N(x_0)| \le C_1 t r_0^2$$

for $\delta_{t^{-1}}(y) \in B_{r_0}(x_0) \Leftrightarrow y \in B_{tr_0}(\delta_t(x_0))$, where the family of dilations $(\delta_t)_{t>0}$ is given by Definition 3.2.5. The same holds for $y \in B_{tr_0}(\delta_t(-x_0))$.

Let $r = tr_0$. We have for $y \in B_r(\delta_t(x_0))$ or $y \in B_r(\delta_t(-x_0))$

$$\begin{aligned} |N^p(y) - N^p(\delta_t(x_0))| &\le C_2 N^{p-1}(\delta_t(x_0))\, |N(y) - N(\delta_t(x_0))| \\ &\le C_3 t^{p-1} t r_0^2 = C_3 t^p r_0^2 \end{aligned}$$

for some constants C_2, C_3, using the mean value theorem. Thus if we take t large enough so that $r_0 = t^{-\frac{p}{2}}$, we have

$$|N^p(y) - N^p(\delta_t(x_0))| \le C_3, \quad \forall y \in B_r(\delta_t(x_0)) \cup B_r(\delta_t(-x_0)),$$

so that

$$\left| \frac{e^{-\beta N^p(y)}}{e^{-\beta N^p(\delta_t(x_0))}} \right| \approx 1 \tag{4.24}$$

for all $y \in B_r(\delta_t(x_0)) \cup B_r(\delta_t(-x_0))$. Now define

$$\begin{aligned} \varphi(y) = \max&\left\{ \min\left\{ 2 - \frac{N(y, \delta_t(x_0))}{r}, 1 \right\}, 0 \right\} \\ &- \max\left\{ \min\left\{ 2 - \frac{N(y, \delta_t(-x_0))}{r}, 1 \right\}, 0 \right\}. \end{aligned} \tag{4.25}$$

Then φ is a Lipschitz function supported on balls of radius r centred at $\delta_t(x_0)$ and $\delta_t(-x_0)$, which is equal to 1 on balls of radius $r/2$ around these two points and decays to zero linearly in between $r/2$ and r. We can note that by construction, and since the measure ν_p is symmetric about the origin,

$$\int_{\mathbb{G}} \varphi(y) d\nu_p(y) = 0.$$

Applying the spectral gap inequality (4.22) to the function φ, then yields

$$\int_{B_r(\delta_t(x_0))\cup B_r(\delta_t(-x_0))} \varphi^2(y)d\nu_p(y) \leq c_0 \int_{B_r(\delta_t(x_0))\cup B_r(\delta_t(-x_0))} |\nabla_{\mathbb{G}}\varphi(y)|^2 d\nu_p(y). \quad (4.26)$$

Now, using (4.24), there exist positive constants C_4 and C_5 such that

$$\begin{aligned}\int_{B_r(\delta_t(x_0))\cup B_r(\delta_t(-x_0))} \varphi^2(y)d\nu_p(y) &\geq 2\int_{B_{\frac{r}{2}}(\delta_t(x_0))} d\nu_p(y)\\ &\geq C_4 r^Q e^{-\beta t^p N^p(x_0)},\end{aligned}$$

and

$$\begin{aligned}\int_{B_r(\delta_t(x_0))\cup B_r(\delta_t(-x_0))} |\nabla_{\mathbb{G}}\varphi(y)|^2 d\nu_p(y) &\leq 2r^{-2}\int_{B_r(\delta_t(x_0))} d\nu_p(y)\\ &\leq C_5 r^{-2+Q} e^{-\beta t^p N^p(x_0)},\end{aligned}$$

where $Q = n + 2m$ is the homogeneous dimension of the group. Using these estimates in (4.26) yields

$$C_4 \leq c_0 C_5 r^{-2},$$

where $r = tr_0 = t^{1-\frac{p}{2}}$, so that the above equation reads

$$C_4 \leq c_0 C_5 t^{p-2}.$$

Since $p < 2$ and t can be taken arbitrarily large, this is a contradiction.

$\square$

Remark 4.5.4. *Theorem 4.5.2 provides another illustration of a fundamental difference between the operators $\mathcal{L}_p$ defined by* (4.11) *with the Carnot-Carathéodory distance and the operators $\mathcal{T}_p$ defined by* (4.21) *with the Kaplan distance. Indeed, with $p \in (1, 2)$, by Theorem 4.4.7 $\mathcal{L}_p$ has empty essential spectrum, while $\mathcal{T}_p$ does not even have a spectral gap.*

Theorem 4.5.5. *If $p \geq 2$, the measure ν_p given by* (4.20) *satisfies a q-spectral gap inequality, i.e. there exists a constant c_0 such that*

$$\nu_p|f - \nu_p f|^q \leq c_0 \nu_p |\nabla_{\mathbb{G}} f|^q$$

for all locally Lipschitz functions f, where $\frac{1}{q} + \frac{1}{p} = 1$. In particular, for $p \geq 2$ the operator $\mathcal{T}_p$ associated to ν_p given by (4.21) *has a spectral gap.*

To prove this we adapt the methods of Hebisch and Zegarliński in [69], and proceed through an intermediate inequality which is similar to the U-bound studied there.

Lemma 4.5.6. *For $p \geq 2$ there exist constants A, B such that*

$$\nu_p(f^q N^{p-2} \| \cdot \|^2) \leq A \nu_p |\nabla_{\mathbb{G}} f|^q + B \nu_p |f|^q \tag{4.27}$$

for all locally Lipschitz functions f, where $\frac{1}{p} + \frac{1}{q} = 1$, and $\|x\| = |w|$ for $x = (w, z) \in \mathbb{G} = \mathbb{R}^n \times \mathbb{R}^m$.

Proof. We can suppose as usual that $f \geq 0$ and moreover that $f \in C_0^\infty(\mathbb{G})$ (since the result will then follow by an approximation argument). By the Leibniz rule, we can write

$$(\nabla_{\mathbb{G}} f) e^{-\alpha N^p} = \nabla_{\mathbb{G}} \left(f e^{-\alpha N^p} \right) + \alpha p f N^{p-1} (\nabla_{\mathbb{G}} N) e^{-\alpha N^p}.$$

We now take the inner product of both sides of this equation with $\frac{N}{\|x\|} \nabla_{\mathbb{G}} N$ and integrate over $\mathbb{G}$ to arrive at

$$\begin{aligned} \int \frac{N}{\|x\|} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} f e^{-\alpha N^p} dx = & \int \frac{N}{\|x\|} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} \left(f e^{-\alpha N^p} \right) dx \\ & + \alpha p \int f \frac{N^p}{\|x\|} |\nabla_{\mathbb{G}} N|^2 e^{-\alpha N^p} dx. \end{aligned}$$

By the Cauchy-Schwarz inequality, we then have

$$\begin{aligned} & \int \frac{N}{\|x\|} |\nabla_{\mathbb{G}} N| |\nabla_{\mathbb{G}} f| e^{-\alpha N^p} dx \\ & \geq \int \frac{N}{\|x\|} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} \left(f e^{-\alpha N^p} \right) dx + \alpha p \int f \frac{N^p}{\|x\|} |\nabla_{\mathbb{G}} N|^2 e^{-\alpha N^p} dx, \end{aligned}$$

so that by Proposition 3.2.16 and integration by parts,

$$\begin{aligned}\int |\nabla_{\mathbb{G}} f| e^{-\alpha N^p} dx &\geq \int \frac{N}{\|x\|} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} \left(f e^{-\alpha N^p} \right) dx + \alpha p \int f N^{p-2} \|x\| e^{-\alpha N^p} dx \\ &= -\int f \nabla_{\mathbb{G}} \cdot \left(\frac{N}{\|x\|} \nabla_{\mathbb{G}} N \right) e^{-\alpha N^p} dx + \alpha p \int f N^{p-2} \|x\| e^{-\alpha N^p} dx.\end{aligned} \tag{4.28}$$

Note that

$$\begin{aligned}\nabla_{\mathbb{G}} \cdot \left(\frac{N}{\|x\|} \nabla_{\mathbb{G}} N \right) &= \frac{|\nabla_{\mathbb{G}} N|^2}{\|x\|} + \frac{N}{\|x\|} \Delta_{\mathbb{G}} N - \frac{N}{\|x\|^2} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} \|x\| \\ &= \frac{\|x\|}{N^2} + (Q-1) \frac{\|x\|}{N^2} - \frac{N}{\|x\|^2} \nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} \|x\|.\end{aligned} \tag{4.29}$$

Moreover, denoting $x = (w, z) \in \mathbb{G}$ and recalling the definitions from Section 3.2,

$$\begin{aligned}\nabla_{\mathbb{G}} N \cdot \nabla_{\mathbb{G}} \|x\| &= \sum_{i=1}^{n} X_i N X_i \|x\| \\ &= \sum_{i=1}^{n} \frac{1}{N^3} \left(|w|^2 w_i + 4 \sum_{k=1}^{m} \sum_{j=1}^{n} U_{ij}^{(k)} w_j z_k \right) \times \frac{w_i}{|w|} \\ &= \frac{1}{N^3 |w|} \left(|w|^4 + 4 \sum_{k=1}^{m} \left(\sum_{i,j=1}^{n} U_{ij}^{(k)} w_i w_j \right) z_k \right) \\ &= \frac{|w|^3}{N^3} = \frac{\|x\|^3}{N^3},\end{aligned}$$

where we have used the fact that $U^{(k)}$ is skew-symmetric for all $k \in \{1, \dots, m\}$, so that $\sum_{i,j=1}^{n} U_{ij}^{(k)} w_i w_j = 0$. Using this in (4.29) yields

$$\nabla_{\mathbb{G}} \cdot \left(\frac{N}{\|x\|} \nabla_{\mathbb{G}} N \right) = (Q-1) \frac{\|x\|}{N^2}. \tag{4.30}$$

Putting (4.30) in (4.28) and using the definition of ν_p then gives

$$\alpha p \nu_p (f N^{p-2} \| \cdot \|) \leq \nu_p |\nabla_{\mathbb{G}} f| + (Q-1) \nu_p \left(f \frac{\| \cdot \|}{N^2} \right).$$

Replacing f by $f\|\cdot\|$, we see that

$$\begin{aligned}\alpha p\nu_p(fN^{p-2}\|\cdot\|^2) &\leq \nu_p\left(\|\cdot\||\nabla_{\mathbb{G}}f|\right) + \nu_p\left(f|\nabla_{\mathbb{G}}\|\cdot\||\right) + (Q-1)\nu_p\left(f\frac{\|\cdot\|^2}{N^2}\right) \\ &\leq \nu_p\left(\|\cdot\||\nabla_{\mathbb{G}}f|\right) + Q\nu_p\left(f\right), \end{aligned} \tag{4.31}$$

using the fact that $\|x\| \leq N(x)$ and $|\nabla_{\mathbb{G}}\|x\|| = 1$. Now, by replacing f by f^q with $\frac{1}{q}+\frac{1}{p}=1$ in (4.31), we then arrive at

$$\begin{aligned}\alpha p\nu_p(f^qN^{p-2}\|\cdot\|^2) &\leq q\nu_p\left(\|\cdot\|f^{q-1}|\nabla_{\mathbb{G}}f|\right) + Q\nu_p\left(f^q\right) \\ &\leq \frac{1}{\varepsilon^{q-1}}\nu_p|\nabla_{\mathbb{G}}f|^q + \frac{q}{p}\varepsilon\nu_p\left(\|\cdot\|^pf^q\right) + Q\nu_p\left(f^q\right),\end{aligned}$$

for all $\varepsilon > 0$, using Young's inequality. Thus

$$\alpha p\nu_p(f^qN^{p-2}\|\cdot\|^2) \leq \frac{1}{\varepsilon^{q-1}}\nu_p|\nabla_{\mathbb{G}}f|^q + \frac{q}{p}\varepsilon\nu_p\left(N^{p-2}\|\cdot\|^2f^q\right) + Q\nu_p\left(f^q\right),$$

so that, by taking $\varepsilon < \frac{p^2}{q}\alpha$, we see that

$$\nu_p(f^qN^{p-2}\|\cdot\|^2) \leq A\nu_p|\nabla_{\mathbb{G}}f|^q + B\nu_p\left(f^q\right),$$

with

$$A = \frac{1}{\varepsilon^{q-1}(\alpha p - \frac{q}{p}\varepsilon)}, \quad B = \frac{Q}{\alpha p - \frac{q}{p}\varepsilon}.$$

$\square$

We are now in a position to prove Theorem 4.5.5.

Proof of Theorem 4.5.5. First note that

$$\nu_p|f-\nu_pf|^q \leq 2^q\nu_p|f-m|^q, \tag{4.32}$$

for all $m \in \mathbb{R}$. Now, for $R > 0$ and $L > 1$,

$$\nu_p|f-m|^q = \nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2} \geq R\}}\right) + \nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2} \leq R\}} \mathbf{1}_{\{N \leq L\}}\right)$$
$$+ \nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2} \leq R\}} \mathbf{1}_{\{N \geq L\}}\right). \tag{4.33}$$

We treat each of the three terms of (4.33) separately.

First term of (4.33)**:** This can be estimated using Lemma 4.5.6. Indeed

$$\begin{aligned} \nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2} \geq R\}}\right) &\leq \frac{1}{R}\nu_p\left(|f-m|^q N^{p-2}\|\cdot\|^2\right) \\ &\leq \frac{A}{R}\nu_p\left|\nabla_{\mathbb{G}} f\right|^q + \frac{B}{R}\nu_p|f-m|^q. \end{aligned} \tag{4.34}$$

Second term of (4.33)**:** We have

$$\begin{aligned} \nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2} \leq R\}} \mathbf{1}_{\{N \leq L\}}\right) &\leq \nu_p\left(|f-m|^q \mathbf{1}_{\{N \leq L\}}\right) \\ &= \frac{1}{Z}\int_{\{N \leq L\}} |f(x)-m|^q e^{-\alpha N^p(x)} dx \\ &\leq \frac{1}{Z}\int_{\{N \leq L\}} |f(x)-m|^q dx. \end{aligned}$$

Since all homogeneous norms on $\mathbb{G}$ are equivalent (see Proposition 3.2.12), we know that there exist L_1, L_2 such that

$$\{N \leq L\} \subset B_{L_1} := \{x \in \mathbb{G} : d(x) \leq L_1\} \subset \{N \leq L_2\},$$

where d is the Carnot-Carathéodory distance as usual. Choosing

$$m = \frac{1}{|B_{L_1}|}\int_{B_{L_1}} f(x) dx,$$

we then see that by the Poincaré inequality in balls (Theorem 3.2.18),

$$\begin{aligned}
\nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2}\le R\}}\mathbf{1}_{\{N\le L\}}\right) &\le \frac{1}{Z}\int_{\{N\le L\}} |f(x)-m|^q dx \\
&\le \frac{1}{Z}\int_{\{d\le L_1\}} |f(x)-m|^q dx \\
&\le \frac{P_0(L_1)}{Z}\int_{\{d\le L_1\}} |\nabla_{\mathbb{G}} f(x)|^q dx \\
&\le \frac{P_0(L_1)}{Z}\int_{\{N\le L_2\}} |\nabla_{\mathbb{G}} f(x)|^q dx \\
&\le \frac{P_0(L_1)}{Z} e^{\alpha L_2^p}\int_{\{N\le L_2\}} |\nabla_{\mathbb{G}} f(x)|^q e^{-\alpha N^p(x)} dx \\
&\le P_0(L_1) e^{\alpha L_2^p}\nu_p|\nabla_{\mathbb{G}} f|^q. \qquad (4.35)
\end{aligned}$$

Third term of (4.33)**:** Set $\bar{f} = f - m$ and $A_{L,R} := \{x \in \mathbb{G} : \|x\|^2 \le R, N(x) \ge L\}$. Note that since $L > 1$ we have

$$\left\{x \in \mathbb{G} : \|x\|^2 N^{p-2}(x) \le R, N(x) \ge L\right\} \subset A_{L,R}.$$

Thus

$$\nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2}\le R\}}\mathbf{1}_{\{N\ge L\}}\right) \le \int_{A_{L,R}} |\bar{f}(x)|^q d\nu_p(x).$$

Recall that we can write $x = (w, z) \in \mathbb{G}$ for $w \in \mathbb{R}^n$ and $z \in \mathbb{R}^m$. For $e \in \{0,1\}^m$, set

$$\mathcal{S}_e := \{x = (w,z) \in \mathbb{G} : (-1)^{e_1} z_1 \ge 0, \dots, (-1)^{e_m} z_m \ge 0\},$$

so that $\mathbb{G} = \cup_{e\in\{0,1\}^m} \mathcal{S}_e$. The reason for introducing these sets, as we will see, is so that in a particular $\mathcal{S}_e$, the signs of z_j for $j \in \{1,\dots,m\}$ are known. By above, we then have

$$\nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2}\le R\}}\mathbf{1}_{\{N\ge L\}}\right) \le \sum_{e\in\{0,1\}^m} \int_{\mathcal{S}_e\cap A_{L,R}} |\bar{f}(x)|^q d\nu_p(x). \qquad (4.36)$$

We consider $\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(x)|^q d\nu_p(x)$ with $e=(0,\dots,0)$ (the other cases are similar). Let $h\in\mathbb{G}$ be such that $\|h\|=2\sqrt{R}$. Then we may write

$$\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(x)|^q d\nu_p(x)\le 2^{q-1}\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(x)-\bar{f}(xh)|^q d\nu_p(x)$$
$$+\,2^{q-1}\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(xh)|^q d\nu_p(x). \tag{4.37}$$

Let $\gamma:[0,t]\to\mathbb{G}$ be a horizontal geodesic from 0 to h such that $|\dot{\gamma}(s)|\le 1$ for $s\in[0,t]$. Then, by Hölder's inequality,

$$\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(x)-\bar{f}(xh)|^q d\nu_p(x)=\int_{\mathcal{S}_e\cap A_{L,R}}\left|\int_0^t\frac{d}{ds}\bar{f}(x\gamma(s))ds\right|^q d\nu_p(x)$$
$$\le t^{\frac{q}{p}}\int_0^t\int_{\mathcal{S}_e\cap A_{L,R}}|\nabla_{\mathbb{G}}f(x\gamma(s))|^q d\nu_p(x)ds. \tag{4.38}$$

Using this estimate in (4.37), we arrive at

$$\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(x)|^q d\nu_p(x)\le 2^{q-1}d^{\frac{q}{p}}(h)\int_0^t\int_{\mathcal{S}_e\cap A_{L,R}}|\nabla_{\mathbb{G}}f(x\gamma(s))|^q d\nu_p(x)ds$$
$$+\,2^{q-1}\int_{\mathcal{S}_e\cap A_{L,R}}|\bar{f}(xh)|^q d\nu_p(x). \tag{4.39}$$

Since we have chosen h such that $\|h\|=2\sqrt{R}$, we have for $x\in A_{L,R}$

$$\|xh\|\ge\|h\|-\|x\|\ge 2\sqrt{R}-\sqrt{R}=\sqrt{R}. \tag{4.40}$$

We now claim that, for fixed R, we can choose h depending only on R, with $\|h\|=2\sqrt{R}$ and such that for large enough L

$$N^p(xh)\le N^p(x),\qquad\forall\, x\in\mathcal{S}_e\cap A_{L,R} \tag{4.41}$$

i.e. translation by h shifts points of $\mathcal{S}_e\cap A_{L,R}$ closer to the origin (with respect to the distance N).

Proof of claim (4.41)**:** For $x=(w,z)\in\mathcal{S}_e\cap A_{L,R}$, we have

$$\|x\|=|w|\leq\sqrt{R},\qquad N(x)\geq L,\qquad\text{and}\qquad z_1\geq 0,\dots z_m\geq 0.$$

Let $h=(2\sqrt{R},0,\dots,0,h_1,\dots,h_m)\in\mathbb{G}=\mathbb{R}^{n+m}$, for $h_1,\dots,h_m$ only depending on R to be chosen later. Then, by the definition of the group law (see Theorem 3.2.2),

$$\begin{aligned}xh=\Bigg(&w_1+2\sqrt{R},\dots,w_n,\\&z_1+h_1+\sqrt{R}\left(\sum_{j=1}^n U_{1j}^{(1)}w_j\right),\dots,z_m+h_m+\sqrt{R}\left(\sum_{j=1}^n U_{1j}^{(m)}w_j\right)\Bigg),\end{aligned}$$

so that

$$\begin{aligned}&N^4(xh)-N^4(x)\\&\quad=\left((w_1+2\sqrt{R})^2+w_2^2+\dots w_n^2\right)^2+16\left(z_1+h_1+\sqrt{R}\left(\sum_{j=1}^n U_{1j}^{(1)}w_j\right)\right)^2\\&\qquad+\dots+16\left(z_m+h_m+\sqrt{R}\left(\sum_{j=1}^n U_{1j}^{(m)}w_j\right)\right)^2\\&\qquad-\left(w_1^2+\dots+w_n^2\right)^2-16(z_1^2+\dots+z_m^2).\end{aligned}$$

After expansion and cancellation, since we are taking x such that $\|x\|=|w|\leq\sqrt{R}$, we can bound all the remaining terms in the above expression that only involve $w_1,\dots,w_n$ from above in terms of R. This will leave us with

$$\begin{aligned}N^4(xh)-N^4(x)\leq K(R)&+32z_1\left(h_1+\sqrt{R}\left(\sum_{j=1}^n U_{1j}^{(1)}w_j\right)\right)\\&+\dots+32z_m\left(h_m+\sqrt{R}\left(\sum_{j=1}^n U_{1j}^{(m)}w_j\right)\right)\end{aligned}$$

for some constant K depending on R and the matrices $U^{(i)}$ for $i \in \{1, \ldots, m\}$. Now, for $i \in \{1, \ldots, m\}$ let $K_i(R)$ be the constant such that

$$\sqrt{R} \left| \sum_{j=1}^{n} U_{1j}^{(i)} w_j \right| \leq K_i(R)$$

for all $w \in \mathbb{R}^n$ such that $|w| \leq \sqrt{R}$ (so that K_i also depends on the matrix $U^{(i)}$). Then, since $z_i \geq 0$ for $i \in \{1, \ldots, m\}$ by assumption, we have

$$N^4(xh) - N^4(x) \leq K(R) + 32z_1 \left(h_1 + K_1(R)\right) + \cdots + 32z_m \left(h_m + K_m(R)\right).$$

Let $\varepsilon > 0$, and take $h_i = -K_i(R) - \varepsilon$ for $i \in \{1, \ldots, m\}$. Then

$$N^4(xh) - N^4(x) \leq K(R) - 32\varepsilon z_1 - \cdots - 32\varepsilon z_m. \tag{4.42}$$

Now, since we are assuming that $N(x) \geq L$ and $|w| \leq \sqrt{R}$, it follows that

$$|z|^2 \geq \frac{1}{16}(L^4 - R^2).$$

Thus $z_j \geq \frac{1}{4\sqrt{m}}(L^4 - R^2)^{\frac{1}{2}}$ for at least one $j \in \{1, \ldots, m\}$, so that by (4.42) we have

$$N^4(xh) - N^4(x) \leq K(R) - 8\varepsilon \frac{1}{\sqrt{m}}(L^4 - R^2)^{\frac{1}{2}}. \tag{4.43}$$

For big enough L the right-hand side of (4.43) is negative, which proves the claim (4.41).

We now use (4.40) and (4.41) to estimate the terms of (4.39). Indeed, using (4.40) we have that

$$\begin{aligned} 2^{q-1} \int_{\mathcal{S}_e \cap A_{L,R}} |\bar{f}(xh)|^q d\nu_p(x) &\leq \frac{2^{q-1}}{R} \int_{\mathcal{S}_e \cap A_{L,R}} |\bar{f}(xh)|^q \|xh\|^2 d\nu_p(x) \\ &\leq \frac{2^{q-1}}{R(L - N(h))^{p-2}} \int_{\mathcal{S}_e \cap A_{L,R}} |\bar{f}(xh)|^q \|xh\|^2 N^{p-2}(xh) d\nu_p(x) \end{aligned} \tag{4.44}$$

for L large in comparison with $N(h)$. By (4.41) we also have

$$d\nu_p(x) = Z^{-1}e^{-\alpha N^p(x)}dx \le Z^{-1}e^{-\alpha N^p(xh)}dx = d\nu_p(xh)$$

on $\mathcal{S}_e \cap A_{L,R}$, so that we can continue (4.44) to see that

$$\begin{aligned} & 2^{q-1}\int_{\mathcal{S}_e\cap A_{L,R}} |\bar f(xh)|^q d\nu_p(x) && (4.45)\\ &\le \frac{2^{q-1}}{R(L-N(h))^{p-2}}\int_{\mathcal{S}_e\cap A_{L,R}} |\bar f(xh)|^q \|xh\|^2 N^{p-2}(xh)d\nu_p(xh) \\ &\le \frac{2^{q-1}}{R(L-N(h))^{p-2}}\nu_p(|\bar f|^q\|\cdot\|^2N^{p-2}) \\ &\le \frac{2^{q-1}A}{R(L-N(h))^{p-2}}\nu_p|\nabla_{\mathbb{G}} f|^q + \frac{2^{q-1}B}{R(L-N(h))^{p-2}}\nu_p|f-m|^q && (4.46)\end{aligned}$$

where we have used the translational invariance of the Lebesgue measure, and Lemma 4.5.6 again.

For the first term of (4.39), note that there exists a constant $\tilde K = \tilde K(h)$ depending only on h (and hence only on R) such that

$$N^p(x\gamma(s)) - N^p(x) \le \tilde K(h), \qquad \forall x \in \mathcal{S}_e \cap A_{L,R},\ s\in[0,t].$$

This is because $N^p(x\gamma(s)) - N^p(x) \to 0$ as $N(x)\to\infty$ by the mean value theorem. Then

$$\begin{aligned}\int_0^t\int_{\mathcal{S}_e\cap A_{L,R}} |\nabla_{\mathbb{G}} f(x\gamma(s))|^q d\nu_p(x)ds &\le e^{\tilde K(h)}\int_0^t\int_{\mathcal{S}_e\cap A_{L,R}} |\nabla_{\mathbb{G}} f(x\gamma(s))|^q d\nu_p(x\gamma(s))ds \\ &\le d(h)e^{\tilde K(h)}\nu_p|\nabla_{\mathbb{G}} f|^q. && (4.47)\end{aligned}$$

Using (4.47) together with (4.45) in (4.39) yields

$$\begin{aligned}\int_{\mathcal{S}_e\cap A_{L,R}} |\bar f(x)|^q d\nu_p(x) &\le 2^{q-1}\left(d^{\frac{q}{p}+1}(h)e^{\tilde K(h)} + \frac{A}{R(L-N(h))^{p-2}}\right)\nu_p|\nabla_{\mathbb{G}} f|^q \\ &\quad + \frac{2^{q-1}B}{R(L-N(h))^{p-2}}\nu_p|f-m|^q. && (4.48)\end{aligned}$$

The key point is that the coefficient $\frac{2^{q-1}B}{R(L-N(h))^{p-2}}$ can be made as small as we wish by taking R large enough, provided L remains large in comparison. Although we have done the calculations for a specific $e \in \{0,1\}^m$, the same may be done for arbitrary e (with a different choice of h). Thus, by (4.36), we see that there exist constants $C(R,L)$ and $\delta(R,L)$ such that

$$\nu_p\left(|f-m|^q \mathbf{1}_{\{\|\cdot\|^2 N^{p-2}\leq R\}}\mathbf{1}_{\{N\geq L\}}\right) \leq C(R,L)\nu_p|\nabla_{\mathbb{G}} f|^q + \delta(L,R)\nu_p|f-m|^q, \quad (4.49)$$

where $\delta(L,R)$ may be made as small as we wish by taking L and R large enough. This completes the estimate of the third term of (4.33).

It remains to insert the estimates (4.34), (4.35) and (4.49) into (4.33). Doing this we arrive at

$$\nu_p|f-m|^q \leq \left(\frac{A}{R} + P_0(L_1)e^{\alpha L_2^p} + C(R,L)\right)\nu_p|\nabla_{\mathbb{G}} f|^q + \left(\frac{B}{R} + \delta(R,L)\right)\nu_p|f-m|^q,$$

where R and L may be taken large enough so that $\frac{B}{R} + \delta(R,L) < 1$. Upon rearrangement, this inequality, combined with the observation (4.32), proves Theorem 4.5.5. □

Remark 4.5.7. *Although our current techniques do not allow us to conclude that $\mathcal{T}_p$ given by* (4.21) *has empty essential spectrum, we conjecture that this will be true for $p > 2$. This is a clear direction for further investigation.*

Remark 4.5.8. *It has recently come to the author's attention that some similar ideas to those contained in this chapter have been discussed in [127], where conditions for empty essential spectrum for hypoelliptic generators are put forward. In particular it is proved that when $\mu_U(dx) = Z^{-1}e^{-U}dx$ is a probability measure on an H-type group $\mathbb{G}$ with $U \in C^\infty(\mathbb{G})$, and $\mathcal{L} = -\Delta_{\mathbb{G}} + \nabla_{\mathbb{G}} U\cdot\nabla_{\mathbb{G}}$ (so that $\mathcal{L}$ is hypoelliptic, positive and symmetric in $L^2(\mu_U)$), then the condition that*

$$\liminf_{\rho\to\infty} |\mathcal{L}\rho| = \infty, \quad (4.50)$$

for some smooth compact function ρ with $|\nabla_{\mathbb{G}}\rho|^2 \leq 1$, implies that $\sigma_{ess}(\mathcal{L}) = \emptyset$. However, *we note that this result is not easily applicable in either of the situations dealt with above: in the case when $U = \alpha d^p$ it is clear that U is not smooth, and in the case when $U = \alpha N^p$, the obvious choice for ρ is N (cf. Corollary 2.3 of [124]), for which* (4.50) *does not hold, since $\mathcal{L}N(w, z) = 0$ for $w = 0, z \neq 0$.*

Chapter 5

Logarithmic Sobolev Inequalities on an Infinite Product of H-type groups

5.1 Introduction

The aim of this chapter is to prove that certain non-trivial Gibbs measures with unbounded interaction potentials on an infinite product of H-type groups satisfy q-logarithmic Sobolev inequalities. We consider a D-dimensional lattice, and impose interactions between points in the lattice described by a potential. Our approach is similar to those described in the literature where the underlying space is Euclidean (cf. [34, 33, 66, 70, 90, 129, 131, 133, 134]), in that we first prove that each of the single site measures satisfies a q-logarithmic Sobolev inequality with a constant independent of the boundary conditions, before passing to infinity using a telescopic expansion argument. However, the methods we use here to prove that the single site measures satisfy LS_q inequalities are necessarily very different different to those described in the references, since the Γ_2 calculus of Bakry and Emery is not applicable in the setting of H-type groups (see Remark 3.2.4). The alternative methods we use are strongly motivated by those of Hebisch and Zegarliński in [69], and we similarly pass through an intermediate inequality of the type studied there. Moreover, our passage to infinity is also non-standard, since we are interested in general LS_q inequalities rather than just LS_2 inequalities. Although this was considered in [32], the case of unbounded interactions was only hinted at.

Throughout this chapter we consider interaction potentials that grow at most quadratically. However, it may also be asked if similar results can hold when we have potentials that grow faster than quadratically, and some results in this direction have been recently obtained by I. Papageorgiou (see [105, 106] and [107]).

The chapter is organised as follows. We first introduce the infinite dimensional setting with the necessary notation, and then state the main result of the chapter. The proof of the result can be split into two parts: firstly we state and prove the results for the single site measures in Section 5.3, before describing the passage to infinity in Section 5.4. We finish with a similar result for an alternative interaction potential.

Some of the results of this chapter formed part of a joint project with I. Papageorgiou, and have been published in [80].

5.2 Infinite dimensional setting and main result

The Lattice: Let $\mathbb{Z}^D$ be the D-dimensional square lattice, for some fixed $D \in \mathbb{N}$. We equip $\mathbb{Z}^D$ with the l^1 lattice metric $dist(\cdot, \cdot)$, defined by

$$dist(i, j) := \sum_{l=1}^{D} |i_l - j_l|$$

for $i = (i_1, \ldots, i_D), j = (j_1, \ldots, j_D) \in \mathbb{Z}^D$. For $i, j \in \mathbb{Z}^D$ we will also write

$$i \sim j \qquad \Leftrightarrow \qquad dist(i, j) = 1$$

i.e. $i \sim j$ when i and j are nearest neighbours in the lattice.

For $\Lambda \subset \mathbb{Z}^D$, we will write $|\Lambda|$ for the cardinality of Λ, and $\Lambda \subset\subset \mathbb{Z}^D$ when $|\Lambda| < \infty$.

The Configuration Space: Let $\mathbb{G} = \mathbb{R}^{n+m}$ be an H-type group (as defined in Section 3.2) and let $\Omega = (\mathbb{G})^{\mathbb{Z}^D}$ be the *configuration space*. We will say that $\mathbb{G}$ is the *spin space*. We introduce the following notation. Given $\Lambda \subset \mathbb{Z}^D$ and $\omega = (\omega_i)_{i \in \mathbb{Z}^D} \in \Omega$, let $\omega_\Lambda := (\omega_i)_{i \in \Lambda} \in \mathbb{G}^\Lambda$ (so that $\omega \mapsto \omega_\Lambda$ is the natural projection of Ω onto $\mathbb{G}^\Lambda$).

Let $f\colon \Omega \to \mathbb{R}$. Then for $i \in \mathbb{Z}^D$ and $\omega \in \Omega$ define $f_i(\cdot|\omega)\colon \mathbb{G} \to \mathbb{R}$ by

$$f_i(x|\omega) := f(x \bullet_i \omega)$$

where the configuration $x \bullet_i \omega \in \Omega$ is defined by declaring its i-th coordinate to be equal to $x \in \mathbb{G}$ and all the other coordinates coinciding with those of $\omega \in \Omega$. Let $C^{(l)}(\Omega), l \in \mathbb{N}$ denote the set of all functions f for which we have $f_i(\cdot|\omega) \in C^{(l)}(\mathbb{G})$ for all $i \in \mathbb{Z}^D$. For $i \in \mathbb{Z}^D, k \in \{1, \dots, n\}$ and $f \in C^{(1)}(\Omega)$, define

$$X_{i,k} f(\omega) := X_k f_i(x|\omega)|_{x=\omega_i}$$

where $X_1, \dots X_n$ are the left-invariant vector fields on $\mathbb{G}$ as in (3.13).

Define similarly $\nabla_i f(\omega) := \nabla_{\mathbb{G}} f_i(x|\omega)|_{x=\omega_i}$ and $\Delta_i f(\omega) := \Delta_{\mathbb{G}} f_i(x|\omega)|_{x=\omega_i}$ for suitable f, where $\nabla_{\mathbb{G}}$ and $\Delta_{\mathbb{G}}$ are the sub-gradient and the sub-Laplacian on $\mathbb{G}$ respectively. For $\Lambda \subset \mathbb{Z}^D$, set $\nabla_\Lambda f = (\nabla_i f)_{i\in\Lambda}$ and

$$|\nabla_\Lambda f|^q := \sum_{i\in\Lambda} |\nabla_i f|^q.$$

We will write $\nabla_{\mathbb{Z}^D} = \nabla$, since it will not cause any confusion.

Finally, a function f on Ω is said to be *localised* in a set $\Lambda \subset \mathbb{Z}^D$ if f is only a function of those coordinates in Λ.

Local Specification and Gibbs Measure: Let $\Phi = (\phi_{\{i,j\}})_{\{i,j\}\subset\mathbb{Z}^D, i\sim j}$ be a family of C^2 functions such that $\phi_{\{i,j\}}$ is localised in $\{i,j\}$. Assume that there exists $M \in (0,\infty)$ such that $\|\phi_{\{i,j\}}\|_\infty \le M$ and $\|\nabla_i\nabla_j\phi_{\{i,j\}}\|_\infty \le M$ for all $i,j \in \mathbb{Z}^D$ such that $i \sim j$. We say Φ is a bounded potential of range 1. For $\omega \in \Omega$, define

$$H^\omega_\Lambda(x_\Lambda) = \sum_{\substack{\{i,j\}\cap\Lambda\neq\emptyset \\ i\sim j}} \phi_{\{i,j\}}(x_i, x_j)$$

for $x_\Lambda = (x_i)_{i\in\Lambda} \in \mathbb{G}^\Lambda$, where the summation is taken over couples of nearest neighbours $i \sim j$ in the lattice with at least one point in Λ, and where $x_i = \omega_i$ for $i \notin \Lambda$.

Now let $(\mathbb{E}^\omega_\Lambda)_{\Lambda\subset\subset\mathbb{Z}^D,\ \omega\in\Omega}$ be the local specification defined by

$$\mathbb{E}^\omega_\Lambda(dx_\Lambda) = \frac{e^{-U^\omega_\Lambda(x_\Lambda)}}{\int e^{-U^\omega_\Lambda(x_\Lambda)}dx_\Lambda}dx_\Lambda \equiv \frac{e^{-U^\omega_\Lambda(x_\Lambda)}}{Z^\omega_\Lambda}dx_\Lambda \tag{5.1}$$

where dx_Λ is the Lebesgue product measure on $\mathbb{G}^\Lambda$ and

$$U^\omega_\Lambda(x_\Lambda) = \alpha\sum_{i\in\Lambda} d^p(x_i) + \varepsilon \sum_{\substack{\{i,j\}\cap\Lambda\neq\emptyset \\ i\sim j}} (d_\sigma(x_i)+\rho d_\sigma(x_j))^2 + \theta H^\omega_\Lambda(x_\Lambda), \tag{5.2}$$

for $\alpha,\sigma > 0$, $\varepsilon,\rho,\theta\in\mathbb{R}$, and $p\geq 2$, where as above $x_i=\omega_i$ for $i\notin\Lambda$. Here $d:\mathbb{G}\to[0,\infty)$ is the Carnot-Carathéodory distance on $\mathbb{G}$ and

$$d_\sigma(x) := \chi_\sigma(x)d(x), \quad \forall x\in\mathbb{G},$$

where χ_σ is a Lipschitz function given by

$$\chi_\sigma(x) := \begin{cases} 1 & \text{if } d(x)\geq\sigma, \\ \frac{2}{\sigma}d(x)-1 & \text{if } \frac{\sigma}{2}\leq d(x)\leq\sigma, \\ 0 & \text{if } d(x)\leq\frac{\sigma}{2}. \end{cases}$$

We say that the *product part* of the measure $\mathbb{E}^\omega_\Lambda$ is $e^{-\alpha\sum_{i\in\Lambda}d^p(x_i)}$, whilst the *interaction potential* is given by

$$\varepsilon \sum_{\substack{\{i,j\}\cap\Lambda\neq\emptyset \\ i\sim j}} (d_\sigma(x_i)+\rho d_\sigma(x_j))^2 + \theta H^\omega_\Lambda(x_\Lambda). \tag{5.3}$$

Remark 5.2.1. *In the case when $p=2$, we must assume that $\varepsilon > -\frac{\alpha}{2D}$ to ensure that $\int e^{-U_\Lambda}dx_\Lambda<\infty$.*

We define an infinite volume *Gibbs measure* ν on Ω to be a solution of the so-called Dobrushin-Lanford-Ruelle (DLR) equation:

$$\nu\mathbb{E}^\cdot_\Lambda f = \nu f$$

for all bounded measurable functions f on Ω and $\Lambda \subset \mathbb{Z}^D$. The measure ν on Ω has $(\mathbb{E}^{\omega}_{\Lambda})_{\omega\in\Omega,\Lambda\subset\subset\mathbb{Z}^D}$ as its finite volume conditional measures.

The main result of this chapter is the following:

Theorem 5.2.2. *Let ν be a Gibbs measure corresponding to the local specification defined by* (5.1) *and* (5.2)*. Let q be dual to p i.e. $\frac{1}{p}+\frac{1}{q}=1$ and suppose $\varepsilon\rho>0$, with the additional condition that $\varepsilon>-\frac{\alpha}{2D}$ when $p=2$. Then there exist $\varepsilon_0,\theta_0>0$ such that for $|\varepsilon|<\varepsilon_0$ and $|\theta|<\theta_0$, ν is unique and satisfies an LS_q inequality i.e. there exists a constant C such that*

$$\nu\left(|f|^q\log\frac{|f|^q}{\nu|f|^q}\right)\leq C\nu\left(\sum_{i\in\mathbb{Z}^D}|\nabla_i f|^q\right)$$

for all f for which the right-hand side is well defined.

Remark 5.2.3. *One might ask why we consider an interaction potential* (5.3) *involving a cut-off version of the distance function, d_σ. Indeed, in the situation when the underlying spin space is Euclidean (and where the distance function is now the natural Euclidean one), the corresponding interaction potential with d_σ replaced by d is convex at infinity. By the Bakry-Emery criterion, one therefore has that the associated single site measures all satisfy a logarithmic Sobolev inequality with a constant independent of the boundary conditions, allowing passage to infinity in the same way as in Section 5.4. However, in our setting, where the spin space is an H-type group, things are more complicated, in that we cannot use the Bakry-Emery condition. The reason that we take d_σ in the interaction potential is thus a technical one — it will remove the singularity at the origin that will allow our methods to proceed. While not completely satisfactory, the given interaction potential still fulfils the main criteria of being unbounded and quadratic.*

We briefly mention some consequences of Theorem 5.2.2. The first follows directly from Proposition 3.1.4.

Corollary 5.2.4. *Let ν be as in Theorem 5.2.2. Then ν satisfies the q-spectral gap inequality. Indeed*

$$\nu\,|f-\nu f|^q\leq\frac{4C}{\log 2}\nu\left(\sum_{i\in\mathbb{Z}^D}|\nabla_i f|^q\right)$$

where C is as in Theorem 5.2.2.

The proofs of the next two results follow from Propositions 3.1.5 and 3.1.13 respectively.

Corollary 5.2.5. *Let ν be as in Theorem 5.2.2 and suppose $f : \Omega \to \mathbb{R}$ is such that $\||\nabla f|^q\|_\infty < 1$. Then*

$$\nu\left(e^{\lambda f}\right) \leq \exp\left\{\lambda\nu(f) + \frac{C}{q^q(q-1)}\lambda^q\right\}$$

for all $\lambda > 0$, where C is as in Theorem 5.2.2. Moreover, by applying Chebyshev's inequality, and optimising over λ, we arrive at the following 'decay of tails' estimate

$$\nu\left\{\left|f - \int f d\nu\right| \geq h\right\} \leq 2\exp\left\{-\frac{(q-1)^p}{C^{p-1}}h^p\right\}$$

for all $h > 0$, where $\frac{1}{p} + \frac{1}{q} = 1$.

Corollary 5.2.6. *Suppose that our configuration space is actually finite dimensional, so that we replace $\mathbb{Z}^D$ by some finite graph $\mathcal{G}$, and $\Omega = (\mathbb{G})^{\mathcal{G}}$. Then Theorem 5.2.2 still holds, and implies that the semigroup $P_t = e^{t\mathcal{L}}$ is ultracontractive, where $\mathcal{L}$ is a Dirichlet operator satisfying*

$$\nu\left(f\mathcal{L}f\right) = -\nu|\nabla f|^2.$$

Remark 5.2.7. *In the above set-up we are only considering interactions of range 1, but our methods could be generalised to handle interactions of range R.*

5.3 Results for the single site measure

The aim of this section is to show that the single site measures

$$\mathbb{E}^\omega_{\{i\}}(dx_i) =: \mathbb{E}^\omega_i(dx_i) = \frac{e^{-U^\omega_i(x_i)}}{Z^\omega_i}dx_i, \qquad i \in \mathbb{Z}^D,$$

each satisfy an LS_q inequality uniformly on the boundary conditions $\omega \in \Omega$ i.e. with a constant independent of ω. We will often drop the ω in the notation for convenience.

As mentioned, the work is strongly motivated by the methods of Hebisch and Zegarliński described in [69].

Theorem 5.3.1. *Suppose* $(\mathbb{E}^{\omega}_{\Lambda})_{\Lambda \subset\subset \mathbb{Z}^D, \omega\in\Omega}$ *is the local specification defined by* (5.1) *and* (5.2). *Let* $\frac{1}{q}+\frac{1}{p}=1$, *and* $\varepsilon\rho>0$ *with the additional condition that* $\varepsilon > -\frac{\alpha}{2D}$ *when* $p=2$. *Then there exists a constant* c, *independent of the boundary conditions* ω, *such that*

$$\mathbb{E}^{\omega}_{i}\left(|f|^q \log\frac{|f|^q}{\mathbb{E}^{\omega}_{i}|f|^q}\right) \leq c\mathbb{E}^{\omega}_{i}|\nabla_i f|^q$$

for all locally Lipschitz f, $i\in\mathbb{Z}^D$ *and* $\omega\in\Omega$.

We first note that it is sufficient to prove Theorem 5.3.1 in the case when $\theta=0$ (so that we no longer have the bounded interaction term in (5.2)), since LS_q inequalities are stable under bounded perturbations (see Proposition 3.1.8). Moreover, it is clear that

$$\mathbb{E}^{\omega}_{i}(dx_i) = \frac{e^{-\alpha d^p(x_i) - \varepsilon\sum_{j:j\sim i}(d_\sigma(x_i)+\rho d_\sigma(\omega_j))^2}}{\int e^{-\alpha d^p(x_i) - \varepsilon\sum_{j:j\sim i}(d_\sigma(x_i)+\rho d_\sigma(\omega_j))^2}dx_i} = \frac{e^{-\tilde{U}^{\omega}_{i}}}{\tilde{Z}^{\omega}_{i}},$$

where

$$\tilde{U}^{\omega}_{i}(x_i) := \alpha d^p(x_i) + 2D\varepsilon d^2_\sigma(x_i) + 2\varepsilon\rho d_\sigma(x_i)\sum_{j:j\sim i} d_\sigma(\omega_j)$$

and $\tilde{Z}^{\omega}_{i} := \int e^{-\tilde{U}^{\omega}_{i}}dx_i$.

The proof of the theorem will be in several steps. We first concentrate on proving some inequalities of 'U-bound' type, which were introduced in [69].

Lemma 5.3.2. *Let* $\frac{1}{p}+\frac{1}{q}=1$ *and suppose* $\varepsilon\rho>0$, *with the additional condition that* $\varepsilon > -\frac{\alpha}{2D}$ *when* $p=2$. *Then there exist constants* $A_1, B_1\in(0,\infty)$, *independent of* ω, *such that*

$$\mathbb{E}^{\omega}_{i}\left(|f|^q|\nabla_i\tilde{U}^{\omega}_{i}|\right) \leq A_1\mathbb{E}^{\omega}_{i}|\nabla_i f|^q + B_1\mathbb{E}^{\omega}_{i}|f|^q$$

for all locally Lipschitz f, $i\in\mathbb{Z}^D$ *and* $\omega\in\Omega$.

Proof. Without loss of generality assume $f\geq 0$ (otherwise we can apply the inequality to the positive and negative parts of f separately). We can also treat f as a function of the i-th coordinate only, and assume that $f\in C^\infty_0$, since the result will then follow by an approximation argument. By the Leibniz rule, we have

$$(\nabla_i f)e^{-\tilde{U}_i} = \nabla_i(fe^{-\tilde{U}_i}) + f\nabla_i\tilde{U}_i e^{-\tilde{U}_i} \tag{5.4}$$

almost everywhere. Taking the inner product of both sides of (5.4) with $\chi_{\frac{\sigma}{2}}\nabla_i d$ yields

$$\begin{aligned}\int_{\mathbb{G}} f\chi_{\frac{\sigma}{2}}\nabla_i d\cdot\nabla_i\tilde{U}_i e^{-\tilde{U}_i}dx_i &\le \int_{\mathbb{G}}\chi_{\frac{\sigma}{2}}|\nabla_i d||\nabla_i f|e^{-\tilde{U}_i}dx_i - \int_{\mathbb{G}}\chi_{\frac{\sigma}{2}}\nabla_i d\cdot\nabla_i\left(fe^{-\tilde{U}_i}\right)dx_i\\ &\le \int_{\mathbb{G}}|\nabla_i f|e^{-\tilde{U}_i}dx_i + \int_{\mathbb{G}} f\nabla_i\cdot(\chi_{\frac{\sigma}{2}}\nabla_i d)e^{-\tilde{U}_i}dx_i,\end{aligned}$$

where we have used Proposition 3.2.16 and integration by parts. Now

$$\begin{aligned}\nabla_i\cdot(\chi_{\frac{\sigma}{2}}\nabla_i d) &= \nabla_i\chi_{\frac{\sigma}{2}}\cdot\nabla_i d + \chi_{\frac{\sigma}{2}}\Delta_i d\\ &= \frac{4}{\sigma}|\nabla_i d|^2\mathbf{1}_{\{\frac{\sigma}{4}\le d\le\frac{\sigma}{2}\}} + \chi_{\frac{\sigma}{2}}\Delta_i d\\ &\le \frac{4}{\sigma}\mathbf{1}_{\{\frac{\sigma}{4}\le d\le\frac{\sigma}{2}\}} + K\chi_{\frac{\sigma}{2}}\frac{1}{d}\end{aligned}$$

in the sense of distributions, by Proposition 3.2.17. Thus

$$\begin{aligned}\int_{\mathbb{G}} f\chi_{\frac{\sigma}{2}}\nabla_i d\cdot\nabla_i\tilde{U}_i e^{-\tilde{U}_i}dx_i &\le \int_{\mathbb{G}}|\nabla_i f|e^{-\tilde{U}_i}dx_i + \frac{4}{\sigma}\int_{\mathbb{G}} fe^{-\tilde{U}_i}dx_i + K\int_{\mathbb{G}} f\chi_{\frac{\sigma}{2}}\frac{1}{d}e^{-\tilde{U}_i}dx_i\\ &\le \int_{\mathbb{G}}|\nabla_i f|e^{-\tilde{U}_i}dx_i + \frac{4}{\sigma}(1+K)\int_{\mathbb{G}} fe^{-\tilde{U}_i}dx_i. \qquad (5.5)\end{aligned}$$

We now claim that there exist constants $a > 0$ and $b \ge 0$ independent of ω such that

$$\chi_{\frac{\sigma}{2}}\nabla_i d\cdot\nabla_i\tilde{U}_i \ge a|\nabla_i\tilde{U}| - b \tag{5.6}$$

almost everywhere. To see this, first note that

$$\chi_{\frac{\sigma}{2}}\nabla_i d\cdot\nabla_i\tilde{U}_i = p\alpha d^{p-1}\chi_{\frac{\sigma}{2}} + \chi_{\frac{\sigma}{2}}\left(4D\varepsilon d_\sigma + 2\varepsilon\rho\sum_{j:j\sim i} d_\sigma(\omega_j)\right)\nabla_i d\cdot\nabla_i d_\sigma \tag{5.7}$$

almost everywhere. By the definition of d_σ, we have

$$\nabla_i d \cdot \nabla_i d_\sigma = |\nabla_i d|^2 \chi_\sigma + \frac{2}{\sigma} d |\nabla_i d|^2 \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} = \chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}}$$

almost everywhere.

Therefore,

$$\begin{aligned}
\chi_{\frac{\sigma}{2}} \nabla_i d \cdot \nabla_i \tilde{U}_i &= p\alpha d^{p-1} \chi_{\frac{\sigma}{2}} + \chi_{\frac{\sigma}{2}} \left(4D\varepsilon d_\sigma + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j) \right) \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} \right) \\
&= p\alpha d^{p-1} \chi_{\frac{\sigma}{2}} + \left(4D\varepsilon d_\sigma + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j) \right) \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} \right) \\
&\ge p\alpha d^{p-1} \chi_{\frac{\sigma}{2}} - 4D|\varepsilon| d \chi_{\frac{\sigma}{2}} - 8D|\varepsilon|\sigma \\
&\quad + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j) \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} \right), \qquad (5.8)
\end{aligned}$$

using the fact that $4D\varepsilon d_\sigma \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} \right) \le 4D|\varepsilon| d \chi_{\frac{\sigma}{2}} + 8D|\varepsilon|\sigma$. Now, for $p > 2$, we have that for all $\delta > 0$ there exists a constant $C(\delta)$ such that $d \le \delta d^{p-1} + C(\delta)$. We can thus continue (5.8) in this case to see that

$$\begin{aligned}
\chi_{\frac{\sigma}{2}} \nabla_i d \cdot \nabla_i \tilde{U}_i &\ge (p\alpha - 4\delta D|\varepsilon|)\, d^{p-1} \chi_{\frac{\sigma}{2}} - 4D|\varepsilon|(C(\delta) + 2\sigma) \\
&\quad + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j) \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} \right).
\end{aligned}$$

Taking δ small enough so that $p\alpha - 4\delta D|\varepsilon| > 0$, and since we are assuming $\varepsilon\rho > 0$, we have that

$$\chi_{\frac{\sigma}{2}} \nabla_i d \cdot \nabla_i \tilde{U}_i \ge a_1 \left(d^{p-1} \chi_{\frac{\sigma}{2}} + \sum_{j:j\sim i} d_\sigma(\omega_j) \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \le d \le \sigma\}} \right) \right) - b_1 \qquad (5.9)$$

where

$$a_1 = \min\{p\alpha - 4\delta D|\varepsilon|, 2\rho\varepsilon\} > 0, \qquad b_1 = 4D|\varepsilon|(C(\delta) + 2\sigma).$$

When $p = 2$ we assume $\varepsilon > -\frac{\alpha}{2D}$, so that from (5.8) we can see that (5.9) is also valid in this case, albeit with adjusted constants.

On the other hand,

$$
\begin{aligned}
|\nabla_i \tilde{U}_i| &= \left| p\alpha d^{p-1}\nabla_i d + \left(4D\varepsilon d_\sigma + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j)\right)\nabla_i d_\sigma \right| \\
&= \left| p\alpha d^{p-1} + \left(4D\varepsilon d_\sigma + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j)\right)\left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2}\le d\le \sigma\}}\right) \right| \\
&\le p\alpha d^{p-1} + 12D|\varepsilon| d + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j)\left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2}\le d\le \sigma\}}\right) \\
&\le \left(p\alpha + 12D|\varepsilon|\right) d^{p-1} + 12D|\varepsilon| + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j)\left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2}\le d\le \sigma\}}\right),
\end{aligned}
$$

using the fact that $d_\sigma \le d$. Now, since $d^{p-1} \le d^{p-1}\chi_{\frac{\sigma}{2}} + \left(\frac{\sigma}{2}\right)^{p-1}$, we then have that

$$
\begin{aligned}
|\nabla_i \tilde{U}_i| &\le \left(p\alpha + 12D|\varepsilon|\right) d^{p-1}\chi_{\frac{\sigma}{2}} + \left(p\alpha + 12D|\varepsilon|\right)\left(\frac{\sigma}{2}\right)^{p-1} + 12D|\varepsilon| \\
&\quad + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j)\left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2}\le d\le \sigma\}}\right) \\
&\le a_2\left(d^{p-1}\chi_{\frac{\sigma}{2}} + \sum_{j:j\sim i} d_\sigma(\omega_j)\left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2}\le d\le \sigma\}}\right)\right) + b_2 \qquad (5.10)
\end{aligned}
$$

where

$$
a_2 = \max\{p\alpha + 12D|\varepsilon|, 2\varepsilon\rho\} > 0, \quad b_2 = \left(p\alpha + 12D|\varepsilon|\right)\left(\frac{\sigma}{2}\right)^{p-1} + 12D|\varepsilon|.
$$

Combining (5.9) and (5.10) proves the claim (5.6). Hence, by (5.5) we have

$$
\int_{\mathbb{G}} f|\nabla_i \tilde{U}_i| e^{-\tilde{U}_i} dx_i \le \frac{1}{a}\int_{\mathbb{G}} |\nabla_i f| e^{-\tilde{U}_i} dx_i + \frac{4}{a\sigma}(1+K+b)\int_{\mathbb{G}} f e^{-\tilde{U}_i} dx_i.
$$

We can finally replace f by f^q in the above, and since $|\nabla_i f^q| = q f^{q-1}|\nabla_i f| \le |\nabla_i f|^q + \frac{q}{p} f^q$, the result follows. $\square$

Corollary 5.3.3. *Let $\frac{1}{p} + \frac{1}{q} = 1$ and suppose $\varepsilon\rho > 0$, with the additional condition that $\varepsilon > -\frac{\alpha}{2D}$ when $p = 2$. Then there exist constants $A_2, B_2 \in (0, \infty)$, independent of ω, such* *that*

$$\mathbb{E}_i^\omega \left(|f|^q \mathcal{W}_i^\omega\right) \leq A_2 \mathbb{E}_i^\omega |\nabla_i f|^q + B_2 \mathbb{E}_i^\omega |f|^q$$

for all locally Lipschitz f, $i \in \mathbb{Z}^D$ and $\omega \in \Omega$, and where $\mathcal{W}_i^\omega$ is defined by

$$\mathcal{W}_i^\omega(x_i) := d^{p-1}(x_i) + \mathbf{1}_{\{d(x_i) \geq \frac{\sigma}{2}\}}(x_i) \sum_{j \in \mathbb{Z}^D : j \sim i} d_\sigma(\omega_j) \tag{5.11}$$

for $x_i \in \mathbb{G}$.

Proof. This follows simply by directly inserting estimate (5.9) into (5.5) in the proof of the above Lemma, before noting that

$$\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\left\{\frac{\sigma}{2} \leq d \leq \sigma\right\}} \geq \mathbf{1}_{\left\{d \geq \frac{\sigma}{2}\right\}}$$

and

$$d^{p-1} \chi_{\frac{\sigma}{2}} \geq d^{p-1} - \left(\frac{\sigma}{2}\right)^{p-1}.$$

□

Lemma 5.3.4. *Let $\frac{1}{p} + \frac{1}{q} = 1$ and suppose $\varepsilon\rho > 0$, with the additional condition that $\varepsilon > -\frac{\alpha}{2D}$ when $p = 2$. Then there exist constants $A_3, B_3 \in (0, \infty)$, independent of ω, such that*

$$\mathbb{E}_i^\omega \left(|f|^q \tilde{U}_i^\omega\right) \leq A_3 \mathbb{E}_i^\omega |\nabla_i f|^q + B_3 \mathbb{E}_i^\omega |f|^q$$

for all locally Lipschitz f, $i \in \mathbb{Z}^D$ and $\omega \in \Omega$.

Proof. The proof of this result is similar to that of Lemma 5.3.2. Once again our starting point is

$$(\nabla_i f) e^{-\tilde{U}_i} = \nabla_i (f e^{-\tilde{U}_i}) + f \nabla_i \tilde{U}_i e^{-\tilde{U}_i},$$

so that by taking the inner product of both sides with $d\nabla_i d$ and integrating yields

$$\begin{aligned}\int_{\mathbb{G}} f d\nabla_i d \cdot \nabla_i \tilde{U}_i e^{-\tilde{U}_i} dx_i &\leq \int_{\mathbb{G}} d|\nabla_i d||\nabla_i f| e^{-\tilde{U}_i} dx_i - \int_{\mathbb{G}} d\nabla_i d \cdot \nabla_i \left(f e^{-\tilde{U}_i} \right) dx_i \\ &= \int_{\mathbb{G}} d|\nabla_i f| e^{-\tilde{U}_i} dx_i + \int_{\mathbb{G}} f \nabla_i \cdot (d\nabla_i d) e^{-\tilde{U}_i} dx_i \\ &= \int_{\mathbb{G}} d|\nabla_i f| e^{-\tilde{U}_i} dx_i + \int_{\mathbb{G}} f(|\nabla_i d|^2 + d\Delta_i d) e^{-\tilde{U}_i} dx_i \\ &\leq \int_{\mathbb{G}} d|\nabla_i f| e^{-\tilde{U}_i} dx_i + (1+K) \int_{\mathbb{G}} f e^{-\tilde{U}_i} dx_i\end{aligned}$$

again using Propositions 3.2.16 and 3.2.17. Replacing f by f^q in this inequality, yields

$$\int_{\mathbb{G}} f^q d\nabla_i d \cdot \nabla_i \tilde{U}_i e^{-\tilde{U}_i} dx_i \leq q \int_{\mathbb{G}} d f^{q-1} |\nabla_i f| e^{-\tilde{U}_i} dx_i + (1+K) \int_{\mathbb{G}} f^q e^{-\tilde{U}_i} dx_i.$$

Now, by Young's inequality, we have that

$$f^{q-1}|\nabla_i f| \leq \frac{1}{q\tau d} |\nabla_i f|^q + \frac{1}{p} \tau^{p-1} d^{p-1} f^q$$

for all $\tau > 0$, so that we then arrive at

$$\begin{aligned}\int_{\mathbb{G}} f^q d\nabla_i d \cdot \nabla_i \tilde{U}_i e^{-\tilde{U}_i} dx_i \leq \frac{1}{\tau} \int_{\mathbb{G}} |\nabla_i f|^q e^{-\tilde{U}_i} dx_i + \frac{q}{p} \tau^{p-1} \int_{\mathbb{G}} f^q d^p e^{-\tilde{U}_i} dx_i \\ + (1+K) \int_{\mathbb{G}} f^q e^{-\tilde{U}_i} dx_i \qquad (5.12)\end{aligned}$$

for all $\tau > 0$.

We can now calculate that

$$\begin{aligned}d\nabla_i d \cdot \nabla_i \tilde{U}_i = p\alpha d^p + d \left(4D\varepsilon d_\sigma + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j) \right) \left(\chi_\sigma + \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \leq d \leq \sigma\}} \right) \\ = p\alpha d^p + \left(4D\varepsilon d_\sigma^2 + 2\varepsilon\rho d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) \right) \\ + \left(4D\varepsilon d_\sigma + 2\varepsilon\rho \sum_{j:j\sim i} d_\sigma(\omega_j) \right) \frac{2}{\sigma} d \mathbf{1}_{\{\frac{\sigma}{2} \leq d \leq \sigma\}}. \qquad (5.13)\end{aligned}$$

For $\varepsilon\rho > 0$, we therefore have that there exist constants $\tilde{a}_1, \tilde{b}_1 \in (0, \infty)$ such that

$$d\nabla_i d \cdot \nabla_i \tilde{U}_i \geq \tilde{a}_1 \left(d^p + d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) \right) - \tilde{b}_1 \tag{5.14}$$

almost everywhere. This is clear if $\varepsilon > 0$, since we can neglect the last term of (5.13) (as it is positive) and take $\tilde{a}_1 = \min\{p\alpha, 2\varepsilon\rho\}$ and $\tilde{b}_1 = 0$.

If $\varepsilon < 0$ and $p > 2$ then, since $d \geq d_\sigma$, and using again the fact that for any $\delta \in (0, 1)$ there exists a positive constant $C(\delta)$ such that $d^2 \leq \delta d^p + C(\delta)$, we have

$$\begin{aligned} d\nabla_i d \cdot \nabla_i \tilde{U}_i &\geq p\alpha d^p - 4D|\varepsilon| d^2 + 2\varepsilon\rho d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) - 8D|\varepsilon|\sigma \\ &\geq (p\alpha - 4D|\varepsilon|\delta) d^p + 2\varepsilon\rho d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) - 4D|\varepsilon|(C(\delta) + 2\sigma). \end{aligned}$$

Thus, taking δ small enough to ensure that $p\alpha - 4D|\varepsilon|\delta > 0$, in (5.14) we can take

$$\tilde{a}_1 = \min\{p\alpha - 4D|\varepsilon|\delta, 2\varepsilon\rho\} > 0, \qquad \tilde{b}_1 = 4D|\varepsilon|(C(\delta) + 2\sigma) > 0.$$

In the case $p = 2$, recall that we must assume $\varepsilon > -\frac{\alpha}{2D}$, and then assertion (5.14) similarly follows.

Using (5.14) in (5.12) we see that

$$\begin{aligned} &\left(\tilde{a}_1 - \frac{q}{p}\tau^{p-1} \right) \int_{\mathbb{G}} f^q \left(d^p + d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) \right) e^{-\tilde{U}_i} dx_i \\ &\qquad \leq \frac{1}{\tau} \int_{\mathbb{G}} |\nabla_i f|^q e^{-\tilde{U}_i} dx_i + (1 + K + \tilde{b}_1) \int_{\mathbb{G}} f^q e^{-\tilde{U}_i} dx_i, \end{aligned} \tag{5.15}$$

where we may choose τ small enough to ensure that $\tilde{a}_1 - \frac{q}{p}\tau^{p-1} > 0$. Finally, we also have that

$$\begin{aligned}\tilde{U}_i &= \alpha d^p + 2D\varepsilon d_\sigma^2 + 2\varepsilon\rho d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) \\ &\leq \tilde{a}_2 \left(d^p + d_\sigma \sum_{j:j\sim i} d_\sigma(\omega_j) \right) + \tilde{b}_2 \end{aligned} \tag{5.16}$$

where

$$\tilde{a}_2 = \max\{\alpha + 2D|\varepsilon|, 2\varepsilon\rho\} > 0, \quad \tilde{b}_2 = 2D|\varepsilon|.$$

Using (5.16) in (5.15) then yields

$$\int_{\mathbb{G}} f^q \tilde{U}_i e^{-\tilde{U}_i} dx_i \leq A_3 \int_{\mathbb{G}} |\nabla_i f|^q e^{-\tilde{U}_i} dx_i + B_3 \int_{\mathbb{G}} f^q e^{-\tilde{U}_i} dx_i,$$

where

$$A_3 = \frac{\tilde{a}_2}{\tau\left(\tilde{a}_1 - \frac{q}{p}\tau^{p-1}\right)}, \quad B_3 = \frac{\tilde{a}_2}{\tilde{a}_1 - \frac{q}{p}\tau^{p-1}}\left(1 + K + \tilde{b}_1\right) + \tilde{b}_2,$$

as required. □

We are now in a position to prove that the single site measures each satisfy a q-spectral gap inequality, with a constant independent of the boundary conditions ω.

Proposition 5.3.5. *Let $\frac{1}{p} + \frac{1}{q} = 1$ and suppose $\varepsilon\rho > 0$, with the additional condition that $\varepsilon > -\frac{\alpha}{2D}$ when $p = 2$. Then $\mathbb{E}_i^\omega$ satisfies a q-spectral gap inequality uniformly on the boundary conditions i.e. there exists a constant $c_0 \in (0, \infty)$, independent of ω, such that*

$$\mathbb{E}_i^\omega |f - \mathbb{E}_i^\omega f|^q \leq c_0 \mathbb{E}_i^\omega |\nabla_i f|^q$$

for locally Lipschitz f, $i \in \mathbb{Z}^D$ and $\omega \in \Omega$.

Proof. First note that, for all $L \geq 0$, we may write

$$\begin{aligned}\mathbb{E}_i|f - \mathbb{E}_i f|^q &= \mathbb{E}_i\left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \leq \frac{\sigma}{2}\}}\right) \\ &\quad + \mathbb{E}_i\left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \geq \frac{\sigma}{2}\}} \mathbf{1}_{\{\mathcal{W}_i \geq L\}}\right) + \mathbb{E}_i\left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \geq \frac{\sigma}{2}\}} \mathbf{1}_{\{\mathcal{W}_i \leq L\}}\right),\end{aligned} \tag{5.17}$$

where $\mathcal{W}_i^\omega = d^{p-1} + \mathbf{1}_{\{d \geq \frac{\sigma}{2}\}} \sum_{j:j\sim i} d_\sigma(\omega_j)$ is as in Corollary 5.3.3. We will estimate each term of (5.17) separately, treating f as a function of x_i only, by fixing all other coordinates.

Estimate of first term: We have

$$\mathbb{E}_i \left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \leq \frac{\sigma}{2}\}} \right) \leq 2^q \mathbb{E}_i \left(|f - m_1|^q \mathbf{1}_{\{d \leq \frac{\sigma}{2}\}} \right), \tag{5.18}$$

where $m_1 := \frac{1}{|\{d \leq \frac{\sigma}{2}\}|} \int_{d \leq \frac{\sigma}{2}} f dx$. Then, using Theorem 3.2.18,

$$\begin{aligned}
\mathbb{E}_i \left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \leq \frac{\sigma}{2}\}} \right) &\leq \frac{2^q}{\tilde{Z}_i^\omega} \int_{\{d \leq \frac{\sigma}{2}\}} |f(x_i) - m_1|^q dx_i \\
&\leq P_0 \left(\frac{\sigma}{2} \right) \frac{2^q}{\tilde{Z}_i^\omega} \int_{\{d \leq \frac{\sigma}{2}\}} |\nabla_i f|^q (x_i) dx_i \\
&\leq c_1 \mathbb{E}_i |\nabla_i f|^q
\end{aligned} \tag{5.19}$$

for

$$c_1 = 2^q P_0 \left(\frac{\sigma}{2} \right) e^{\alpha \frac{\sigma^p}{2^p}}.$$

Estimate of second term: By Corollary 5.3.3, we have

$$\begin{aligned}
\mathbb{E}_i \left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \geq \frac{\sigma}{2}\}} \mathbf{1}_{\{\mathcal{W}_i \geq L\}} \right) &\leq \mathbb{E}_i \left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{\mathcal{W}_i \geq L\}} \right) \\
&\leq \frac{1}{L} \mathbb{E}_i \left(|f - \mathbb{E}_i f|^q \mathcal{W}_i \right) \\
&\leq \frac{A_2}{L} \mathbb{E}_i |\nabla_i f|^q + \frac{B_2}{L} \mathbb{E}_i |f - \mathbb{E}_i f|^q .
\end{aligned} \tag{5.20}$$

Estimate of third term: Set $R = L^{1/(p-1)}$, and recall that $B_R := \{x \in \mathbb{G} : d(x) \leq R\}$. We have

$$\mathbb{E}_i \left(|f - \mathbb{E}_i f|^q \mathbf{1}_{\{d \geq \frac{\sigma}{2}\}} \mathbf{1}_{\{\mathcal{W}_i \leq L\}} \right) \leq 2^q \mathbb{E}_i \left(|f - m_2|^q \mathbf{1}_{\{d \geq \frac{\sigma}{2}\}} \mathbf{1}_{\{\mathcal{W}_i \leq L\}} \right),$$

for $m_2 := \frac{1}{|B_R|} \int_{B_R} f(x_i) dx_i$. Note that, by definition of $\mathcal{W}_i^\omega$,

$$\sum_{j:j\sim i} d_\sigma(\omega_j) \leq L \qquad \text{and} \qquad d(x_i) \leq R$$

whenever x_i is such that $d(x_i) \geq \frac{\sigma}{2}$ and $\mathcal{W}_i^\omega(x_i) \leq L$. Thus, by making use of Theorem 3.2.18 again, we see that

$$\begin{aligned}
\mathbb{E}_i\left(|f-\mathbb{E}_i f|^q \mathbf{1}_{\{d\geq\frac{\sigma}{2}\}}\mathbf{1}_{\{\mathcal{W}_i\leq L\}}\right) &\leq 2^q \mathbf{1}_{\{\sum_{j:j\sim i} d_\sigma(\omega_j)\leq L\}}\mathbb{E}_i\left(|f-m_2|^q \mathbf{1}_{B_R}\right) \\
&\leq 2^q \mathbf{1}_{\{\sum_{j:j\sim i} d_\sigma(\omega_j)\leq L\}}\frac{e^{2D|\varepsilon|R^2}}{\tilde{Z}_i^\omega}\int_{B_R}|f(x_i)-m_2|^q dx_i \\
&\leq 2^q \mathbf{1}_{\{\sum_{j:j\sim i} d_\sigma(\omega_j)\leq L\}}P_0(R)\frac{e^{2D|\varepsilon|R^2}}{\tilde{Z}_i^\omega}\int_{B_R}|\nabla_i f(x_i)|^q dx_i \\
&\leq 2^q \mathbf{1}_{\{\sum_{j:j\sim i} d_\sigma(\omega_j)\leq L\}}P_0(R)\frac{e^{4D|\varepsilon|R^2+\alpha R^p}}{\tilde{Z}_i^\omega}\int_{B_R}|\nabla_i f(x_i)|^q e^{-\alpha d^p(x_i)-2D\varepsilon d_\sigma^2(x_i)} dx_i \\
&\leq c_2 \mathbb{E}_i|\nabla_i f|^q
\end{aligned} \tag{5.21}$$

for

$$c_2 = 2^q P_0(R) e^{4D|\varepsilon|R^2+\alpha R^p+2\varepsilon\rho RL}.$$

To finish we use the estimates (5.19), (5.20) and (5.21) in (5.17), which yields

$$\mathbb{E}_i|f-\mathbb{E}_i f|^q \leq \left(c_1+c_2+\frac{A_2}{L}\right)\mathbb{E}_i|\nabla_i f|^q + \frac{B_2}{L}\mathbb{E}_i|f-\mathbb{E}_i f|^q$$

for all $L>0$. If we then take L large enough to ensure that $\frac{B_2}{L}<1$, a rearrangement of this inequality gives the result. □

We can now prove Theorem 5.3.1 :

Proof of Theorem 5.3.1 . Our starting point is the classical Sobolev inequality on H-type groups for the Lebesgue measure: there exists a $t>0$ such that

$$\left(\int_{\mathbb{G}}|f|^{1+t}dx_i\right)^{\frac{1}{1+t}} \leq a\int_{\mathbb{G}}|\nabla_i f|dx_i + b\int_{\mathbb{G}}|f|dx_i \tag{5.22}$$

for some constants $a, b \in (0,\infty)$. Indeed, by Theorem 3.2.19 we may take t such that $1+t=\frac{Q}{Q-1}$ where Q is the homogeneous dimension of $\mathbb{G}$. Once again, without loss of generality we may assume that $f \geq 0$. Suppose also that $\mathbb{E}_i(f^q) = 1$. Now, if we set

$$g \equiv \frac{f^q e^{-\tilde{U}_i}}{\tilde{Z}_i},$$

then

$$\mathbb{E}_i(f^q \log f^q) = \int_{\mathbb{G}} g \log g dx_i + \mathbb{E}_i(f^q \tilde{U}_i) + \log \tilde{Z}_i. \tag{5.23}$$

Now by Jensen's inequality

$$\begin{aligned}
\int_{\mathbb{G}} g \log g dx_i &= \frac{1}{t} \int_{\mathbb{G}} g \log g^t dx_i \\
&\leq \frac{1+t}{t} \log \left(\int_{\mathbb{G}} g^{1+t} dx_i \right)^{\frac{1}{1+t}} \\
&\leq \frac{1+t}{t} \left(\int_{\mathbb{G}} g^{1+t} dx_i \right)^{\frac{1}{1+t}} \\
&\leq \frac{a(1+t)}{t} \int_{\mathbb{G}} |\nabla_i g| dx_i + \frac{1+t}{t} b,
\end{aligned}$$

where we have used the classical Sobolev inequality (5.22) and the elementary inequality $\log x \leq x$. Hence by (5.23)

$$\begin{aligned}
\mathbb{E}_i(f^q \log f^q) \leq{}& \frac{a(1+t)}{t} \int \left| \nabla_i \left(\frac{f^q e^{-\tilde{U}_i}}{\tilde{Z}_i} \right) \right| dx_i + \mathbb{E}_i(f^q \tilde{U}_i) + \frac{1+t}{t} b + \log \tilde{Z}_i \\
\leq{}& \frac{a(1+t)}{t} \mathbb{E}_i(q f^{q-1} |\nabla_i f|) + \frac{a(1+t)}{t} \mathbb{E}_i(f^q |\nabla_i \tilde{U}_i|) + \mathbb{E}_i(f^q \tilde{U}_i) \\
&+ \frac{1+t}{t} b + \log \tilde{Z}_i \\
\leq{}& \frac{a(1+t)}{t} \mathbb{E}_i |\nabla_i f|^q + \frac{a(1+t)}{t} \mathbb{E}_i(f^q |\nabla_i \tilde{U}_i|) + \mathbb{E}_i(f^q \tilde{U}_i) \\
&+ \frac{1+t}{t} b + \frac{aq(1+t)}{pt} + \log \tilde{Z}_i,
\end{aligned} \tag{5.24}$$

where we have used Young's inequality i.e. $q f^{q-1} |\nabla_i f| \leq |\nabla_i f|^q + (q/p) f^q$. Note that, since $\varepsilon\rho > 0$, we have that $\tilde{Z}_i^\omega \leq C_1$ for some constant $C_1 \in (0, \infty)$ independent of ω. We also recognise that the second and third terms in (5.24) can be bounded by Lemmas 5.3.2

and 5.3.4 respectively. Using these bounds allows us to conclude that

$$\mathbb{E}_i(f^q \log f^q) \leq C_2 \mathbb{E}_i |\nabla_i f|^q + C_3 \tag{5.25}$$

where

$$C_2 = \frac{a(1+t)}{t}(1+A_1) + A_3, \quad C_3 = \frac{a(1+t)}{t}B_1 + B_3 + \frac{1+t}{t}b + \frac{aq(1+t)}{pt} + C_1.$$

Replacing f^q by $\frac{f^q}{\mathbb{E}_i f^q}$ in (5.25) gives

$$\mathbb{E}_i \left(f^q \log \frac{f^q}{\mathbb{E}_i f^q} \right) \leq C_2 \mathbb{E}_i |\nabla_i f|^q + C_3 \mathbb{E}_i(f^q), \tag{5.26}$$

so that $\mathbb{E}_i^\omega$ satisfies the defective q-logarithmic Sobolev inequality, DLS_q, with constants independent of the boundary conditions.

Since we also have that $\mathbb{E}_i$ satisfies an SG_q inequality with constant independent of the boundary conditions (Proposition 5.3.5), we can finally apply the Rothaus argument (Proposition 3.1.7) to conclude that there exists a constant c, independent of ω, such that

$$\mathbb{E}_i \left(f^q \log \frac{f^q}{\mathbb{E}_i f^q} \right) \leq c \mathbb{E}_i |\nabla_i f|^q,$$

which proves Theorem 5.3.1 . □

5.4 Passage to infinite dimensions

In this section we show how to pass from the uniform LS_q inequality for the single site measures $\mathbb{E}_i^\omega$ to the LS_q inequality for the corresponding Gibbs measure ν on the entire configuration space $\Omega = (\mathbb{G})^{\mathbb{Z}^D}$. As mentioned in the introduction to this chapter, in the more standard case when $q = 2$ this problem has been thoroughly investigated, whilst the procedure for the case $q < 2$ with unbounded interactions has only been hinted at (see [32]). It is for this reason that we describe the argument here in detail, which is primarily based on ideas introduced in [133] and [134].

We work in greater generality than is required for Theorem 5.2.2 , though the results of

Section 5.3 show that in the specific case where the local specification is defined by (5.1) and (5.2), the hypotheses **(H0)** and **(H1)** below are satisfied. Theorem 5.2.2 then follows from Theorem 5.4.1 (see Corollary 5.4.2 below).

Consider a local specification $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\omega\in\Omega}$ defined by

$$\mathbb{E}_\Lambda^\omega(dx_\Lambda) = \frac{e^{-\sum_{i\in\Lambda}\varphi(x_i)-\sum_{\{i,j\}\cap\Lambda\neq\emptyset,i\sim j}J_{ij}V(x_i,x_j)}dx_\Lambda}{Z_\Lambda^\omega}, \tag{5.27}$$

where Z_Λ^ω is the normalisation factor and the summation is taken over couples of nearest neighbours $i \sim j$ in the lattice with at least one point in Λ and where $x_i = \omega_i$ for $i \notin \Lambda$, as before. Thus the product part of the measure $\mathbb{E}_\Lambda^\omega$ is $e^{-\sum_{i\in\Lambda}\varphi(x_i)}$, while the interaction potential is given by $\sum_{\{i,j\}\cap\Lambda\neq\emptyset,i\sim j} J_{ij}V(x_i,x_j)$.

We suppose that $|J_{ij}| \leq J_0$ for all i,j and some $J_0 > 0$. Moreover, as above, we suppose that ν is a Gibbs measure corresponding to this local specification i.e. ν is a solution to the DLR equation

$$\nu\mathbb{E}_\Lambda^{\cdot} f = \nu f \tag{5.28}$$

for all bounded measurable functions f on Ω and $\Lambda \subset \mathbb{Z}^D$.

We will work with the following hypotheses:

(H0): The one-dimensional single site measure $\mathbb{E}_i^\omega$ satisfies LS_q with a constant c which is independent of the boundary conditions ω, for all $i \in \mathbb{Z}^D$ and $\omega \in \Omega$.

(H1): There exists a constant $\tilde{M} \in (0,\infty)$ such that

$$\|\nabla_i\nabla_j V(x_i,x_j)\|_\infty \leq \tilde{M}$$

uniformly in i and j.

Theorem 5.4.1. *Suppose the local specification* $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\omega\in\Omega}$ *defined by* (5.27) *satisfies* **(H0)** *and* **(H1)**. *Then, for sufficiently small J_0, the corresponding infinite dimensional*

Gibbs measure ν is unique and satisfies the LS_q inequality

$$\nu\left(|f|^q \log \frac{|f|^q}{\nu|f|^q}\right) \leq C\nu\left(\sum_{i\in\mathbb{Z}^D} |\nabla_i f|^q\right)$$

for some positive constant C and all f for which the right-hand side is well-defined.

Corollary 5.4.2. *Theorem 5.2.2 holds*

Proof. In the setting of Theorem 5.2.2, we have $\varphi(x_i) = \alpha d^p(x_i)$ and

$$V(x_i, x_j) = (d_\sigma(x_i) + \rho d_\sigma(x_j))^2 + \phi_{\{i,j\}}(x_i, x_j)$$

for $\alpha, \sigma > 0$, $\rho \in \mathbb{R}$ and $p \geq 2$. By Theorem 5.3.1 **(H0)** holds. It thus remains to check **(H1)**:

$$\begin{aligned}|\nabla_i \nabla_j V(x_i, x_j)| &\leq 2|\rho|\,|\nabla_i d_\sigma(x_i) \cdot \nabla_j d_\sigma(x_j)| + \left|\nabla_i \nabla_j \phi_{\{i,j\}}(x_i, x_j)\right| \\ &\leq 18|\rho| + M,\end{aligned}$$

by our assumptions on the potential Φ and since $|\nabla_i d_\sigma| \leq 3$. Hence Theorem 5.2.2 follows from an application of Theorem 5.4.1. □

The proof of Theorem 5.4.1 will rely on several lemmata, which we prove in the following subsection.

5.4.1 Lemmata

Define the following sets

$$\begin{aligned}&\Gamma_0 = (0,0) \cup \{j \in \mathbb{Z}^D : dist(j,(0,0)) = 2m \ \text{ for some } \ m \in \mathbb{N}\}, \\ &\Gamma_1 = \mathbb{Z}^D \smallsetminus \Gamma_0.\end{aligned}$$

where $dist(\cdot,\cdot)$ is as in Section 5.2 . Note that $dist(i,j) > 1$ for all $i \neq j$ in Γ_k, and $\Gamma_0 \cap \Gamma_1 = \emptyset$. Moreover $\mathbb{Z}^D = \Gamma_0 \cup \Gamma_1$. As above, for the sake of notation, we will write

$\mathbb{E}_{\Gamma_k} = \mathbb{E}^{\omega}_{\Gamma_k}$ for $k = 0, 1$. We will also define

$$\mathcal{P} := \mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}.$$

Lemma 5.4.3. *Suppose the local specification* $(\mathbb{E}^{\omega}_{\Lambda})_{\Lambda\subset\subset\mathbb{Z}^D,\omega\in\Omega}$ *defined by* (5.27) *satisfies* **(H0)** *and* **(H1)**. *Then, for sufficiently small J_0, there exist constants $K_1 > 0$ and $\eta_1 \in (0, 1)$ such that*

$$\nu\,|\nabla_{\Gamma_k}(\mathbb{E}_{\Gamma_l} f)|^q \le K_1\nu\,|\nabla_{\Gamma_k} f|^q + \eta_1\nu\,|\nabla_{\Gamma_l} f|^q$$

for $k, l \in \{0, 1\}$ such that $k \neq l$.

Proof. For convenience, suppose $k = 1$ and $l = 0$. The case $k = 0, l = 1$ follows similarly. Define $\{\sim i\} := \{j : j \sim i\}$. By construction, the measure $\mathbb{E}_{\Gamma_0}$ is actually a product measure. This is because interactions only occur between nearest neighbours of the lattice, and all points in Γ_0 are at least a distance 2 apart. We can then write $\mathbb{E}_{\Gamma_0} = \mathbb{E}_{\Gamma_0\setminus\{\sim i\}}\mathbb{E}_{\{\sim i\}}$ for any $i \in \Gamma_1$, so that

$$\begin{aligned}
\mathcal{I} := \nu\,|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0} f)|^q &= \nu\sum_{i\in\Gamma_1}|\nabla_i(\mathbb{E}_{\Gamma_0} f)|^q \\
&= \nu\sum_{i\in\Gamma_1}\left|\nabla_i(\mathbb{E}_{\Gamma_0\setminus\{\sim i\}}\mathbb{E}_{\{\sim i\}} f)\right|^q \\
&= \nu\sum_{i\in\Gamma_1}\left|\mathbb{E}_{\Gamma_0\setminus\{\sim i\}}\nabla_i(\mathbb{E}_{\{\sim i\}} f)\right|^q \le \nu\sum_{i\in\Gamma_1}\left|\nabla_i(\mathbb{E}_{\{\sim i\}} f)\right|^q,
\end{aligned}$$

where we have used Jensen's inequality and the DLR equation (5.28). Note that in the third line we can bring the sub-gradient inside the first expectation because $\mathbb{E}_{\Gamma_0\setminus\{\sim i\}}$ does not depend on the ith coordinate. Continuing the above, and using the explicit form of $\mathbb{E}_{\{\sim i\}}$ given by (5.27), we can then calculate that

$$\begin{aligned}
\mathcal{I} &\le \nu\sum_{i\in\Gamma_1}\left|\nabla_i(\mathbb{E}_{\{\sim i\}} f)\right|^q \\
&\le 2^{q-1}\nu\sum_{i\in\Gamma_1}\left|\mathbb{E}_{\{\sim i\}}\nabla_i f\right|^q + 2^{q-1}J_0^q(2D)^{\frac{q}{p}}\nu\sum_{i\in\Gamma_1}\sum_{j\in\{\sim i\}}\left|\mathbb{E}_{\{\sim i\}}\left(f\bar{\mathcal{V}}_{ij}\right)\right|^q
\end{aligned} \tag{5.29}$$

where $\frac{1}{p}+\frac{1}{q}=1$ as usual, and we have denoted $\mathcal{V}_{ij}(x_i,x_j):=\nabla_i V(x_i,x_j)$ and

$$\bar{\mathcal{V}}_{ij}=\mathcal{V}_{ij}-\mathbb{E}_{\{\sim i\}}\mathcal{V}_{ij}.$$

Then

$$\begin{aligned}\mathcal{I} &\leq 2^{q-1}\nu\sum_{i\in\Gamma_1}\mathbb{E}_{\{\sim i\}}\left|\nabla_i f\right|^q+2^{q-1}J_0^q(2D)^{\frac{q}{p}}\nu\sum_{i\in\Gamma_1}\sum_{j\in\{\sim i\}}\left|\mathbb{E}_{\{\sim i\}}\left((f-\mathbb{E}_{\{\sim i\}}f)\bar{\mathcal{V}}_{ij}\right)\right|^q\\ &\leq 2^{q-1}\nu\sum_{i\in\Gamma_1}\mathbb{E}_{\{\sim i\}}\left|\nabla_i f\right|^q\\ &\quad+2^{q-1}J_0^q(2D)^{\frac{q}{p}}\nu\sum_{i\in\Gamma_1}\mathbb{E}_{\{\sim i\}}\left|f-\mathbb{E}_{\{\sim i\}}f\right|^q\sum_{j\in\{\sim i\}}\left(\mathbb{E}_{\{\sim i\}}\left|\bar{\mathcal{V}}_{ij}\right|^p\right)^{q/p}\end{aligned}\tag{5.30}$$

using Hölder's inequality and the fact that $\mathbb{E}_{\{\sim i\}}\bar{\mathcal{V}}_{ij}=0$ for $j\in\{\sim i\}$. As already noted, no interactions occur between points of the set $\{\sim i\}$, so that the measure $\mathbb{E}^{\omega}_{\{\sim i\}}$ is a product measure i.e. $\mathbb{E}^{\omega}_{\{\sim i\}}=\otimes_{j\in\{\sim i\}}\mathbb{E}^{\omega}_j$. Moreover, by **(H0)**, all measures $\mathbb{E}^{\omega}_j, j\in\{\sim i\}$ satisfy the LS_q inequality with a constant c uniformly on the boundary conditions. Therefore, since the LS_q inequality is stable under tensorisation (see Proposition 3.1.9), we have that the product measure $\mathbb{E}^{\omega}_{\{\sim i\}}$ also satisfies the LS_q inequality with the same constant c. By Proposition 3.1.4, it follows that $\mathbb{E}^{\omega}_{\{\sim i\}}$ satisfies a q-spectral gap inequality with constant $c_0=\frac{4c}{\log 2}$ i.e.

$$\mathbb{E}_{\{\sim i\}}\left|f-\mathbb{E}_{\{\sim i\}}f\right|^q\leq c_0\mathbb{E}_{\{\sim i\}}\left|\nabla_{\{\sim i\}}f\right|^q.\tag{5.31}$$

By Proposition 3.1.11, since $q<p$, we also have that there exists a constant $\tilde{c}_0$ such that for any $j\in\{\sim i\}$

$$\begin{aligned}\mathbb{E}_{\{\sim i\}}\left|\bar{\mathcal{V}}_{ij}\right|^p=\mathbb{E}_j\left|\bar{\mathcal{V}}_{ij}\right|^p&=\mathbb{E}_j\left|\mathcal{V}_{ij}-\mathbb{E}_j\mathcal{V}_{ij}\right|^p\\ &\leq\tilde{c}_0\mathbb{E}_j\left|\nabla_j\mathcal{V}_{ij}\right|^p\\ &\leq\tilde{c}_0\mathbb{E}_j\left|\nabla_j\nabla_i V(x_i,x_j)\right|^p\leq\tilde{c}_0\tilde{M}^p\end{aligned}\tag{5.32}$$

by **(H1)**.

If we combine (5.30), (5.31) and (5.32) we obtain

$$\begin{aligned}
\nu \left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0} f)\right|^q \leq & \, 2^{q-1} \nu \left(\sum_{i \in \Gamma_1} \mathbb{E}_{\{\sim i\}} \left|\nabla_i f\right|^q \right) \\
& + 2^{q-1} c_0 (\tilde{c}_0)^{q/p} (2D)^{1+q/p} \tilde{M}^q J_0^q \nu \left(\sum_{i \in \Gamma_1} \mathbb{E}_{\{\sim i\}} \left|\nabla_{\{\sim i\}} f\right|^q \right) \\
\leq & \, 2^{q-1} \nu \left(\sum_{i \in \Gamma_1} \left|\nabla_i f\right|^q \right) \\
& + 2^{q-1} c_0 (\tilde{c}_0)^{q/p} (2D)^{2+q/p} \tilde{M}^q J_0^q \nu \left(\sum_{i \in \Gamma_0} \left|\nabla_i f\right|^q \right).
\end{aligned}$$

Therefore, choosing J_0 sufficiently small so that $2^{q-1} c_0 (\tilde{c}_0)^{q/p} (2D)^{2+q/p} \tilde{M}^q J_0^q < 1$, we see that

$$\nu \left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0} f)\right|^q \leq K_1 \nu \left|\nabla_{\Gamma_1} f\right|^q + \eta_1 \nu \left|\nabla_{\Gamma_0} f\right|^q$$

with $K_1 = 2^{q-1}$ and $\eta_1 = 2^{q-1} c_0 (\tilde{c}_0)^{q/p} (2D)^{2+q/p} \tilde{M}^q J_0^q < 1$, as required. □

Lemma 5.4.4. *Suppose the local specification* $(\mathbb{E}^\omega_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^D, \omega \in \Omega}$ *defined by* (5.27) *satisfies* **(H0)** *and* **(H1)**. *Define* $\mathcal{V}_{ij}(x_i, x_j) := \nabla_i V(x_i, x_j)$*, as in the proof of Lemma 5.4.3. Then there exists a constant* κ*, independent of the boundary conditions, such that*

$$\left|\mathbb{E}_{\{\sim i\}} \left(|f|^q ; \mathcal{V}_{ij}\right)\right| \leq \left(\mathbb{E}_{\{\sim i\}} |f|^q\right)^{\frac{1}{p}} \left(\kappa \mathbb{E}_{\{\sim i\}} \left|\nabla_{\{\sim i\}} f\right|^q\right)^{\frac{1}{q}}$$

for all $i \in \mathbb{Z}^D$ *and* $j \in \{\sim i\}$*, where* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $\mathbb{E}_{\{\sim i\}}(g; h) := \mathbb{E}_{\{\sim i\}}(gh) - \mathbb{E}_{\{\sim i\}}(g)\mathbb{E}_{\{\sim i\}}(h)$ *for any functions* g, h.

Proof. Without loss of generality, we may suppose that $f \geq 0$. Let $\hat{\mathbb{E}}_{\{\sim i\}}$ be an isomorphic copy of $\mathbb{E}_{\{\sim i\}}$. Then for $i \in \mathbb{Z}^D$ and $j \in \{\sim i\}$ we have

$$\begin{aligned}
\mathbb{E}_{\{\sim i\}}(f^q; \mathcal{V}_{ij}) &= \frac{1}{2} \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}} \left(\left(f^q - \hat{f}^q \right) \left(\mathcal{V}_{ij} - \hat{\mathcal{V}}_{ij} \right) \right) \\
&= \frac{1}{2} \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}} \left[\left(\int_0^1 \frac{d}{ds} F_s^q ds \right) \left(\mathcal{V}_{ij} - \hat{\mathcal{V}}_{ij} \right) \right]
\end{aligned}$$

where $F_s = sf + (1-s)\hat{f}$ for $s \in [0,1]$. Then

$$\begin{aligned}
|\mathbb{E}_{\{\sim i\}}(f^q;\mathcal{V}_{ij})| &= \frac{q}{2}\left|\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left[\left(\int_0^1 F_s^{q-1}ds\right)\left(f-\hat{f}\right)\left(\mathcal{V}_i - \hat{\mathcal{V}}_{ij}\right)\right]\right| \\
&\leq \frac{q}{2}\left\{\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left(\int_0^1 F_s^{q-1}ds\right)^p\right\}^{\frac{1}{p}} \\
&\quad \times \left\{\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left|f-\hat{f}\right|^q\left|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}\right|^q\right\}^{\frac{1}{q}}. \qquad (5.33)
\end{aligned}$$

Now by Jensen's inequality and convexity of the function $y \mapsto y^q$ we have

$$\begin{aligned}
\left\{\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left(\int_0^1 F_s^{q-1}ds\right)^p\right\}^{\frac{1}{p}} &\leq \left\{\int_0^1 \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}} F_s^q ds\right\}^{\frac{1}{p}} \\
&\leq \left\{\int_0^1 \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left(sf^q + (1-s)\hat{f}^q\right)ds\right\}^{\frac{1}{p}} \\
&= \left(\mathbb{E}_{\{\sim i\}} f^q\right)^{\frac{1}{p}}. \qquad (5.34)
\end{aligned}$$

Moreover,

$$\begin{aligned}
&\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left|f-\hat{f}\right|^q\left|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}\right|^q \qquad (5.35) \\
&\quad \leq 2^q \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left|f - \mathbb{E}_{\{\sim i\}}f\right|^q\left|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}\right|^q.
\end{aligned}$$

Recalling the relative entropy inequality from Lemma 3.1.3, we have that $\forall \tau > 0$

$$\begin{aligned}
&\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left|f-\hat{f}\right|^q\left|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}\right|^q \\
&\quad \leq \frac{2^q}{\tau}\mathbb{E}_{\{\sim i\}}\left|f-\mathbb{E}_{\{\sim i\}}f\right|^q \log \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left(e^{\tau|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}|^q}\right) \\
&\quad + \frac{2^q}{\tau}\mathbb{E}_{\{\sim i\}}\left(\left|f-\mathbb{E}_{\{\sim i\}}f\right|^q \log \frac{\left|f-\mathbb{E}_{\{\sim i\}}f\right|^q}{\mathbb{E}_{\{\sim i\}}\left|f-\mathbb{E}_{\{\sim i\}}f\right|^q}\right). \qquad (5.36)
\end{aligned}$$

Now, since both **(H0)** and **(H1)** are satisfied, we can apply Corollary 3.1.6 to see that there exists a constant $\Theta > 0$ independent of ω such that

$$\mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}}\left(e^{\tau|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}|^q}\right) = \mathbb{E}_j \otimes \hat{\mathbb{E}}_j\left(e^{\tau|\mathcal{V}_{ij}-\hat{\mathcal{V}}_{ij}|^q}\right) \leq \Theta$$

for sufficiently small τ. Indeed, let $G = \mathcal{V}_{ij} - \hat{\mathcal{V}}_{ij}$, so that $\mathbb{E}_j \otimes \hat{\mathbb{E}}_j(G) = 0$. Then

$$|\nabla_j G| + |\hat{\nabla}_j G| \leq 2 \left\| \nabla_j \mathcal{V}_{ij} \right\|_\infty = 2 \|\nabla_j \nabla_i V(x_i, x_j)\|_\infty \leq 2\tilde{M}.$$

Thus by Corollary 3.1.6, we have that

$$\mathbb{E}_j \otimes \hat{\mathbb{E}}_j \left(e^{\tau |\mathcal{V}_{ij} - \hat{\mathcal{V}}_{ij}|^q} \right) \leq e^\tau \mathbb{E}_j \otimes \hat{\mathbb{E}}_j \left(e^{\tau |\mathcal{V}_{ij} - \hat{\mathcal{V}}_{ij}|^p} \right) \leq \Theta$$

for τ sufficiently small and where Θ depends only on $\tilde{M}$ and c. We can also use **(H0)** to bound the second term of (5.36). Altogether this gives

$$\begin{aligned} \mathbb{E}_{\{\sim i\}} \otimes \hat{\mathbb{E}}_{\{\sim i\}} \left| f - \hat{f} \right|^q \left| \mathcal{V}_{ij} - \hat{\mathcal{V}}_{ij} \right|^q &\leq \frac{2^q \log \Theta}{\tau} \mathbb{E}_{\{\sim i\}} \left| f - \mathbb{E}_{\{\sim i\}} f \right|^q \\ &\quad + \frac{2^q c}{\tau} \mathbb{E}_{\{\sim i\}} \left| \nabla_{\{\sim i\}} f \right|^q \\ &\leq \frac{2^q}{\tau} \left(c_0 \log \Theta + c \right) \mathbb{E}_{\{\sim i\}} \left| \nabla_{\{\sim i\}} f \right|^q, \end{aligned} \tag{5.37}$$

where $c_0 = \frac{4c}{\log 2}$, by Proposition 3.1.4 once again.

Putting estimates (5.34) and (5.37) into (5.33) we see that

$$\mathbb{E}_{\{\sim i\}}(f^q; \mathcal{V}_{ij}) \leq \left(\mathbb{E}_{\{\sim i\}} f^q \right)^{\frac{1}{p}} \left(\frac{q^q}{\tau} (c_0 \log \Theta + c) \mathbb{E}_{\{\sim i\}} \left| \nabla_{\{\sim i\}} f \right|^q \right)^{\frac{1}{q}},$$

which gives the desired result. $\square$

Lemma 5.4.5. *Suppose the local specification $(\mathbb{E}^\omega_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^D, \omega \in \Omega}$ defined by (5.27) satisfies* **(H0)** *and* **(H1)**. *Then, for sufficiently small J_0, there exist constants $K_2 > 0$ and $\eta_2 \in (0, 1)$ such that*

$$\nu \left| \nabla_{\Gamma_k} (\mathbb{E}_{\Gamma_l} |f|^q)^{\frac{1}{q}} \right|^q \leq K_2 \nu \left| \nabla_{\Gamma_k} f \right|^q + \eta\nu_2 \left| \nabla_{\Gamma_l} f \right|^q$$

for $k, l \in \{0, 1\}, k \neq l$.

Proof. Again we may suppose $f \geq 0$. For $k = 1, l = 0$ (the other case is similar), we can write

$$\nu \left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0} f^q)^{\frac{1}{q}}\right|^q \leq \nu \sum_{i\in\Gamma_1} \left|\nabla_i(\mathbb{E}_{\{\sim i\}} f^q)^{\frac{1}{q}}\right|^q$$
$$= \nu \sum_{i\in\Gamma_1} \frac{1}{q^q}(\mathbb{E}_{\{\sim i\}} f^q)^{-\frac{q}{p}} \left|\nabla_i(\mathbb{E}_{\{\sim i\}} f^q)\right|^q. \tag{5.38}$$

We will compute the terms in the sum on the right-hand side of (5.38). For $i \in \Gamma_1$, we have

$$\nabla_i(\mathbb{E}_{\{\sim i\}} f^q) = q(\mathbb{E}_{\{\sim i\}} f^{q-1}\nabla_i f) - \sum_{j\in\{\sim i\}} J_{i,j}\mathbb{E}_{\{\sim i\}}\left(f^q; \nabla_i V(x_i, x_j)\right)$$
$$\Rightarrow \left|\nabla_i(\mathbb{E}_{\{\sim i\}} f^q)\right| \leq q\left(\mathbb{E}_{\{\sim i\}} f^q\right)^{1/p}\left(\mathbb{E}_{\{\sim i\}}|\nabla_i f|^q\right)^{1/q} + J_0 \sum_{j\in\{\sim i\}} \left|\mathbb{E}_{\{\sim i\}}\left(f^q; \mathcal{V}_{ij}\right)\right|$$

where $\mathcal{V}_{ij}(x_i, x_j) = \nabla_i V(x_i, x_j)$ as above, so that

$$\left|\nabla_i(\mathbb{E}_{\{\sim i\}} f^q)\right|^q \leq 2^{q-1}q^q\left(\mathbb{E}_{\{\sim i\}} f^q\right)^{\frac{q}{p}}\left(\mathbb{E}_{\{\sim i\}}|\nabla_i f|^q\right)$$
$$+ 2^{q-1}(2D)^{\frac{q}{p}} J_0^q \sum_{j\in\{\sim i\}} \left|\mathbb{E}_{\{\sim i\}}\left(f^q; \mathcal{V}_{ij}\right)\right|^q.$$

We can use Lemma 5.4.4 to bound the correlation in the second term. Indeed, this gives

$$\left|\nabla_i(\mathbb{E}_{\{\sim i\}} f^q)\right|^q \leq 2^{q-1}\left(\mathbb{E}_{\{\sim i\}} f^q\right)^{\frac{q}{p}}\left(q^q\mathbb{E}_{\{\sim i\}}|\nabla_i f|^q + \kappa(2D)^{\frac{q}{p}+1}J_0^q\mathbb{E}_{\{\sim i\}}\left|\nabla_{\{\sim i\}} f\right|^q\right).$$

Using this in (5.38) yields

$$\nu \left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0} f^q)^{\frac{1}{q}}\right|^q \leq 2^{q-1}\nu \sum_{i\in\Gamma_1}\left(\mathbb{E}_{\{\sim i\}}|\nabla_i f|^q + \frac{\kappa}{q^q}(2D)^{\frac{q}{p}+1}J_0^q\mathbb{E}_{\{\sim i\}}\left|\nabla_{\{\sim i\}} f\right|^q\right)$$
$$= 2^{q-1}\nu\left|\nabla_{\Gamma_1} f\right|^q + \frac{2^{q-1}\kappa}{q^q}(2D)^{\frac{q}{p}+1}J_0^q\nu \sum_{i\in\Gamma_1}\left|\nabla_{\{\sim i\}} f\right|^q$$
$$= 2^{q-1}\nu\left|\nabla_{\Gamma_1} f\right|^q + \frac{2^{q-1}\kappa}{q^q}(2D)^{\frac{q}{p}+2}J_0^q\nu\left|\nabla_{\Gamma_0} f\right|^q.$$

Finally, taking J_0 such that $\frac{2^{q-1}\kappa}{q^q}(2D)^{\frac{q}{p}+2}J_0^q < 1$ we see that

$$\nu \left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}f^q)^{\frac{1}{q}}\right|^q \leq K_2\nu\left|\nabla_{\Gamma_1}f\right|^q + \eta_2\nu\left|\nabla_{\Gamma_0}f\right|^q,$$

where $K_2 = 2^{q-1}$ and $\eta_2 = \frac{2^{q-1}\kappa}{q^q}(2D)^{\frac{q}{p}+2}J_0^q < 1$, as required. □

Lemma 5.4.6. *Suppose the local specification* $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\omega\in\Omega}$ *defined by* (5.27) *satisfies* **(H0)** *and* **(H1)**. *Then, for sufficiently small* J_0, $\mathcal{P}^n f$ *converges* ν*-almost everywhere to* νf, *where we recall that* $\mathcal{P} = \mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}$. *In particular,* ν *is unique.*

Proof. We will follow the argument given in Chapter 5 of [66]. We have

$$\begin{aligned}\nu\left|f - \mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}f\right|^q &\leq 2^{q-1}\nu\mathbb{E}_{\Gamma_0}\left|f - \mathbb{E}_{\Gamma_0}f\right|^q + 2^{q-1}\nu\mathbb{E}_{\Gamma_1}\left|\mathbb{E}_{\Gamma_0}f - \mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}f\right|^q \\ &\leq 2^{q-1}c_0\nu\left|\nabla_{\Gamma_0}f\right|^q + 2^{q-1}c_0\nu\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}f)\right|^q,\end{aligned}$$

since by **(H0)** and Proposition 3.1.9 both the measures $\mathbb{E}_{\Gamma_0}$ and $\mathbb{E}_{\Gamma_1}$ satisfy the SG_q inequality with constant $c_0 = \frac{4c}{\log 2}$ independent of the boundary conditions. For sufficiently small J_0, we may use Lemma 5.4.3, which yields

$$\nu\left|f - \mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}f\right|^q \leq 2^{q-1}c_0\nu\left|\nabla_{\Gamma_0}f\right|^q + 2^{q-1}c_0(K_1\nu\left|\nabla_{\Gamma_1}f\right|^q + \eta_1\nu\left|\nabla_{\Gamma_0}f\right|^q).$$

From the last inequality we obtain that for any $n \in \mathbb{N}$,

$$\begin{aligned}\nu\left|\mathcal{P}^n f - \mathcal{P}^{n+1}f\right|^q &\leq 2^{q-1}c_0\nu\left|\nabla_{\Gamma_0}\mathcal{P}^n f\right|^q + 2^{q-1}c_0\eta_1\nu\left|\nabla_{\Gamma_0}\mathcal{P}^n f\right|^q \\ &= 2^{q-1}c_0(1+\eta_1)\nu\left|\nabla_{\Gamma_0}\mathcal{P}^n f\right|^q,\end{aligned}$$

using the fact that $\mathcal{P}^n f$ does not depend on coordinates in Γ_1 by definition, so that $\nabla_{\Gamma_1}\mathcal{P}^n f = 0$. By repeated applications of Lemma 5.4.3 we see that,

$$\begin{aligned}\nu\left|\mathcal{P}^n f - \mathcal{P}^{n+1}f\right|^q &\leq 2^{q-1}c_0(1+\eta_1)\eta_1^{2n-1}\nu\left|\nabla_{\Gamma_1}\mathbb{E}_{\Gamma_0}f\right|^q \\ &\leq 2^{q-1}c_0(1+\eta_1)\eta_1^{2n-1}\left(K_1\nu\left|\nabla_{\Gamma_1}f\right|^q + \eta_1\nu\left|\nabla_{\Gamma_0}f\right|^q\right).\end{aligned}$$

Since $\eta_1 < 1$, this clearly tends to zero as $n \to \infty$, so that the sequence $\{\mathcal{P}^n f\}_{n\in\mathbb{N}}$ is

Cauchy in $L^q(\nu)$. Moreover, for all $\varepsilon > 0$, we have

$$\begin{aligned}\sum_{n=1}^{\infty} \nu \left\{ \left| \mathcal{P}^n - \mathcal{P}^{n+1} f \right| \geq \varepsilon \right\} &\leq \frac{1}{\varepsilon^q} \sum_{n=1}^{\infty} \nu \left| \mathcal{P}^n f - \mathcal{P}^{n+1} f \right|^q \\ &\leq \frac{2^{q-1} c_0}{\varepsilon^q} (1+\eta_1) \left(K_1 \nu |\nabla_{\Gamma_1} f|^q + \eta_1 \nu |\nabla_{\Gamma_0} f|^q \right) \sum_{n=1}^{\infty} \eta_1^{2n-1} \\ &< \infty,\end{aligned}$$

again since $\eta_1 < 1$. Thus by the Borel-Cantelli lemma, the sequence $\{\mathcal{P}^n f\}_{n \in \mathbb{N}}$ is convergent ν-almost surely. We can similarly show that $\{|\nabla \mathcal{P}^n f|\}_{n \in \mathbb{N}}$ converges to zero almost surely. Thus

$$\{\mathcal{P}^n f\}_{n \in \mathbb{N}}$$

converges ν-a.s. to a constant, so that the limit of $\mathcal{P}^n f - \nu \mathcal{P}^n f = \mathcal{P}^n f - \nu f$ is identical to zero. □

5.4.2 Proof of Theorem 5.4.1

Recall that we want to extend the LS_q inequality from the single-site measures to the Gibbs measure corresponding to the local specification $(\mathbb{E}_\Lambda^\omega)_{\Lambda \subset\subset \mathbb{Z}^D, \omega \in \Omega}$.

Again without loss of generality, suppose $f \geq 0$. We can write

$$\begin{aligned}\nu \left(f^q \log \frac{f^q}{\nu f^q} \right) =& \nu \mathbb{E}_{\Gamma_0} \left(f^q \log \frac{f^q}{\mathbb{E}_{\Gamma_0} f^q} \right) + \nu \mathbb{E}_{\Gamma_1} \left(\mathbb{E}_{\Gamma_0} f^q \log \frac{\mathbb{E}_{\Gamma_0} f^q}{\mathbb{E}_{\Gamma_1} \mathbb{E}_{\Gamma_0} f^q} \right) \\ &+ \nu \left(\mathbb{E}_{\Gamma_1} \mathbb{E}_{\Gamma_0} f^q \log \mathbb{E}_{\Gamma_1} \mathbb{E}_{\Gamma_0} f^q \right) - \nu \left(f^q \log \nu f^q \right). \qquad (5.39)\end{aligned}$$

As already noted, since the measures $\mathbb{E}_{\Gamma_0}$ and $\mathbb{E}_{\Gamma_1}$ are product measures, by **(H0)** we know that they both satisfy an LS_q inequality with constant c independent of the boundary conditions. Using this fact in (5.39) yields

$$\nu\left(f^q\log\frac{f^q}{\nu f^q}\right)\leq c\nu(\mathbb{E}_{\Gamma_0}\left|\nabla_{\Gamma_0}f\right|^q)+c\nu\mathbb{E}_{\Gamma_1}\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}f^q)^{\frac{1}{q}}\right|^q$$
$$+\nu\left(\mathcal{P}f^q\log\mathcal{P}f^q\right)-\nu\left(f^q\log\nu f^q\right). \qquad (5.40)$$

For the third term of (5.40) we can similarly write

$$\nu\left(\mathcal{P}f^q\log\mathcal{P}f^q\right)=\nu\mathbb{E}_{\Gamma_0}\left(\mathcal{P}f^q\log\frac{\mathcal{P}f^q}{\mathbb{E}_{\Gamma_0}\mathcal{P}f^q}\right)+\nu\mathbb{E}_{\Gamma_1}\left(\mathbb{E}_{\Gamma_0}\mathcal{P}f^q\log\frac{\mathbb{E}_{\Gamma_0}\mathcal{P}f^q}{\mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}\mathcal{P}f^q}\right)$$
$$+\nu\left(\mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}\mathcal{P}f^q\log\mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}\mathcal{P}f^q\right).$$

If we use again the LS_q inequality for the measures $\mathbb{E}_{\Gamma_k}$ $(k=0,1)$ we get

$$\nu\left(\mathcal{P}f^q\log\mathcal{P}f^q\right)\leq c\nu\left|\nabla_{\Gamma_0}(\mathcal{P}f^q)^{\frac{1}{q}}\right|^q+c\nu\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}\mathcal{P}f^q)^{\frac{1}{q}}\right|^q+\nu\left(\mathcal{P}^2f^q\log\mathcal{P}^2f^q\right). \quad (5.41)$$

Working similarly for the last term $\nu\left(\mathcal{P}^2f^q\log\mathcal{P}^2f^q\right)$ of (5.41) and inductively for any term $\nu(\mathcal{P}^kf^q\log\mathcal{P}^kf^q)$, then, by combining this observation with (5.40), after n steps we see that

$$\nu\left(f^q\log\frac{f^q}{\nu f^q}\right)\leq c\sum_{k=0}^{n-1}\nu\left|\nabla_{\Gamma_0}(\mathcal{P}^kf^q)^{\frac{1}{q}}\right|^q+c\sum_{k=0}^{n-1}\nu\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}\mathcal{P}^kf^q)^{\frac{1}{q}}\right|^q$$
$$+\nu\left(\mathcal{P}^nf^q\log\mathcal{P}^nf^q\right)-\nu\left(f^q\log\nu f^q\right). \qquad (5.42)$$

In order to deal with the first and second term on the right-hand side of (5.42) we will use Lemma 5.4.5. Indeed, using the bound given there we have, for any $k\in\mathbb{N}$,

$$\begin{aligned}\nu\left|\nabla_{\Gamma_0}(\mathcal{P}^kf^q)^{\frac{1}{q}}\right|^q&=\nu\left|\nabla_{\Gamma_0}(\mathbb{E}_{\Gamma_1}\mathbb{E}_{\Gamma_0}\mathcal{P}^{k-1}f^q)^{\frac{1}{q}}\right|^q\\&\leq K_2\nu\left|\nabla_{\Gamma_0}(\mathbb{E}_{\Gamma_0}\mathcal{P}^{k-1}f^q)^{\frac{1}{q}}\right|^q+\eta_2\nu\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}\mathcal{P}^{k-1}f^q)^{\frac{1}{q}}\right|^q\\&=\eta_2\nu\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}\mathcal{P}^{k-1}f^q)^{\frac{1}{q}}\right|^q\\&\leq\eta_2^2\nu\left|\nabla_{\Gamma_0}(\mathcal{P}^{k-1}f^q)^{\frac{1}{q}}\right|^q.\end{aligned}$$

We can iterate this inequality to see that

$$\begin{aligned}\nu\left|\nabla_{\Gamma_0}(\mathcal{P}^k f^q)^{\frac{1}{q}}\right|^q &\leq \eta_2^2\nu\left|\nabla_{\Gamma_0}(\mathcal{P}^{k-1} f^q)^{\frac{1}{q}}\right|^q \\ &\leq \eta_2^{2k-1}\nu\left|\nabla_{\Gamma_0}(\mathbb{E}_{\Gamma_0} f^q)^{\frac{1}{q}}\right|^q \\ &\leq \eta_2^{2k-1}K_2\nu\left|\nabla_{\Gamma_1} f\right|^q + \eta_2^{2k}\nu\left|\nabla_{\Gamma_0} f\right|^q, \end{aligned} \tag{5.43}$$

where the last line follows from a final application of Lemma 5.4.5. Similarly,

$$\nu\left|\nabla_{\Gamma_1}(\mathbb{E}_{\Gamma_0}\mathcal{P}^k f^q)^{\frac{1}{q}}\right|^q \leq \eta_2^{2k}K_2\nu\left|\nabla_{\Gamma_1} f\right|^q + \eta_2^{2k+1}\nu\left|\nabla_{\Gamma_0} f\right|^q. \tag{5.44}$$

Using (5.43) and (5.44) in (5.42) yields

$$\begin{aligned}\nu\left(f^q \log\frac{f^q}{\nu f^q}\right) \leq\; & cK_2\left(\eta_2^{-1}+1\right)\left(\sum_{k=0}^{n-1}\eta_2^{2k}\right)\nu\left|\nabla_{\Gamma_1} f\right|^q \\ &+ c\,(1+\eta_2)\left(\sum_{k=0}^{n-1}\eta_2^{2k}\right)\nu\left|\nabla_{\Gamma_0} f\right|^q \\ &+ \nu\left(\mathcal{P}^n f^q \log\mathcal{P}^n f^q\right) - \nu(f^q\log\nu f^q). \end{aligned} \tag{5.45}$$

By Lemma 5.4.6 we have that $\lim_{n\to\infty}\mathcal{P}^n f^q = \nu f^q$, $\nu - a.s.$ Therefore, taking the limit as $n \to \infty$ in (5.45) yields

$$\nu\left(f^q \log\frac{f^q}{\nu f^q}\right) \leq cK_2\left(\frac{1}{\eta_2}+1\right)K_3\nu\left|\nabla_{\Gamma_1} f\right|^q + c(1+\eta_2)K_3\nu\left|\nabla_{\Gamma_0} f\right|^q,$$

where $K_3 = \sum_{k=0}^{\infty}\eta^{2k} = \frac{1}{1-\eta^2}$ for $\eta_2 < 1$. Hence

$$\nu\left(f^q \log\frac{f^q}{\nu f^q}\right) \leq C\nu|\nabla f|^q$$

for $C = \max\left\{cK_2\left(\frac{1}{\eta_2}+1\right)K_3, c(1+\eta_2)K_3\right\}$, which completes the proof.

5.5 An alternative interaction potential

The purpose of this section is to show that we may work with a local specification defined with an alternative interaction potential to the one considered in Section 5.2, and prove a similar result. To be specific, suppose we are again in the situation of Section 5.2, with $\Omega = (\mathbb{G})^{\mathbb{Z}^D}$ for some H-type group $\mathbb{G}$, but now consider a local specification $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\ \omega\in\Omega}$ given by

$$\mathbb{E}_\Lambda^\omega(dx_\Lambda) = \frac{e^{-U_\Lambda^\omega(x_\Lambda)}}{Z_\Lambda^\omega} dx_\Lambda \tag{5.46}$$

with

$$U_\Lambda^\omega(x_\Lambda) = \alpha \sum_{i\in\Lambda} d^p(x_i) + \varepsilon \sum_{\substack{\{i,j\}\cap\Lambda\neq\emptyset \\ i\sim j}} d(x_i x_j^{-1}), \tag{5.47}$$

for $\alpha > 0$, $p \geq 2$, $\varepsilon \in \mathbb{R}$, and where $x_i = \omega_i$ for $i \notin \Lambda$ and $d : \mathbb{G} \to [0, \infty)$ is the Carnot-Carathéodory distance as usual.

Remark 5.5.1. *Given the results of the preceding sections, we would actually like to be able to include quadratic interactions in* (5.47)*, of the form*

$$\sum_{\substack{\{i,j\}\cap\Lambda\neq\emptyset \\ i\sim j}} d^2(x_i x_j^{-1}).$$

However, it seems that this case is more delicate, and the methods below cannot easily handle it. We therefore restrict ourselves to linear interactions of this form, but keep the quadratic case in mind as interesting avenue of further study.

Remark 5.5.2. *It should be noted that, in exactly the same way as above, we may also include a small bounded interaction term θH_Λ^ω in* (5.47) *without affecting the validity of any of the following results.*

The result we prove is the following:

Theorem 5.5.3. *Let ν be a Gibbs measure corresponding to the local specification defined by* (5.46) *and* (5.47)*. Let q be dual to p i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Then there exists an $\varepsilon_0 > 0$ such that for $|\varepsilon| < \varepsilon_0$, ν is unique and satisfies an LS_q inequality i.e. there exists a constant C*

such that

$$\nu\left(|f|^q \log \frac{|f|^q}{\nu|f|^q}\right) \le C\nu\left(\sum_{i\in\mathbb{Z}^D} |\nabla_i f|^q\right)$$

for all f for which the right-hand side is well defined.

The work to be done here involves showing that the single site measures $\mathbb{E}_i^\omega$ satisfy an LS_q inequality with a constant independent of the boundary conditions i.e. that **(H0)** of Section 5.4 is satisfied. This is because the passage to infinity can be achieved in very similar way to the one described in Section 5.4. Indeed, the proofs in that section can actually be simplified somewhat, because now we have $V(x_i, x_j) = d(x_i x_j^{-1})$ so that

$$|\nabla_i V(x_i, x_j)| \le 1$$

for all $i, j \in \mathbb{Z}^D$ i.e. the *first* derivative of the interaction potential is uniformly bounded. Then, wherever the condition **(H1)** is needed in the arguments of Section 5.4, we may use instead this observation.

Thus it is sufficient to prove the following.

Theorem 5.5.4. *Suppose $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\ \omega\in\Omega}$ is given by* (5.46) *and* (5.47)*, and let $\frac{1}{q} + \frac{1}{p} = 1$. Then there exists a constant c, independent of the boundary conditions $\omega \in \Omega$, such that*

$$\mathbb{E}_i^\omega\left(|f|^q \log \frac{|f|^q}{\mathbb{E}_i^\omega |f|^q}\right) \le c\mathbb{E}_i^\omega |\nabla_i f|^q$$

for all locally Lipschitz f, $i \in \mathbb{Z}^D$ and $\omega \in \Omega$.

The route to proving this result will be similar to that of Theorem 5.3.1, in that it will be in three steps: 'U-bound $+$ $SG_q \Rightarrow LS_q$'. To prove the necessary intermediate results we will explicitly make use of the results of Section 3.2.5, together with some perturbation techniques.

Lemma 5.5.5. *Suppose $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\ \omega\in\Omega}$ is given by* (5.46) *and* (5.47)*, and let $\frac{1}{q} + \frac{1}{p} = 1$. Then there exist constants $A, B \in (0, \infty)$, independent of ω, such that*

$$\mathbb{E}_i^\omega(|f|^q d^p) \le A\mathbb{E}_i^\omega |\nabla_i f|^q + B\mathbb{E}_i^\omega |f|^q$$

for all locally Lipschitz f, $i \in \mathbb{Z}^D$ and $\omega \in \Omega$.

Proof. As usual we can suppose $f \geq 0$. By part (i) of Theorem 3.2.20 there exist constants $\tilde{A}$ and $\tilde{B}$ such that

$$\int_{\mathbb{G}} f^q(x_i) d^p(x_i) e^{-\alpha d^p(x_i)} dx_i \leq \tilde{A} \int_{\mathbb{G}} |\nabla_i f|^q(x_i) e^{-\alpha d^p(x_i)} dx_i + \tilde{B} \int_{\mathbb{G}} f^q(x_i) e^{-\alpha d^p(x_i)} dx_i$$

for all locally Lipschitz functions. Replacing f in the above with

$$f(x_i) e^{-\frac{\varepsilon}{q} \sum_{j:j\sim i} d(x_i \omega_j^{-1})}$$

yields

$$\mathbb{E}_i^\omega (f^q d^p) \leq 2^{q-1} \tilde{A} \mathbb{E}_i^\omega |\nabla_i f|^q + \left(\frac{2^{2q-1} |\varepsilon|^q D^q}{q^q} + \tilde{B} \right) \mathbb{E}_i^\omega |f|^q,$$

where we have used the fact that $\sum_{j:j\sim i} |\nabla_i d(x_i \omega_j^{-1})| \leq 2D$ almost everywhere, by Proposition 3.2.16. Thus we can take $A = 2^{q-1}\tilde{A}$ and $B = \left(\frac{2^{2q-1}|\varepsilon|^q D^q}{q^q} + \tilde{B} \right)$, and the lemma is proved. □

Lemma 5.5.6. *There exist constants $a_1, a_2 \in (0, \infty)$, independent of ω, such that*

$$a_1 e^{-2D|\varepsilon| d(x_i)} \leq \frac{e^{-\varepsilon \sum_{j:j\sim i} d(x_i \omega_j^{-1})}}{Z_i^\omega} \leq a_2 e^{2D|\varepsilon| d(x_i)}$$

for all $x_i \in \mathbb{G}$, $i \in \mathbb{Z}^D$ and $\omega \in \Omega$.

Proof. First suppose that $\varepsilon < 0$. Then by definition,

$$\begin{aligned} \frac{e^{-\varepsilon \sum_{j:j\sim i} d(x_i \omega_j^{-1})}}{Z_i^\omega} &= \frac{e^{-\varepsilon \sum_{j:j\sim i} d(x_i \omega_j^{-1})}}{\int e^{-\alpha d^p(x_i) - \varepsilon \sum_{j:j\sim i} d(x_i \omega_j^{-1})} dx_i} \\ &\geq \frac{e^{-\varepsilon \sum_{j:j\sim i} (d(\omega_j) - d(x_i))}}{\int e^{-\alpha d^p(x_i) - \varepsilon \sum_{j:j\sim i} (d(x_i) + d(\omega_j))} dx_i} \\ &= \frac{e^{2D\varepsilon d(x_i)}}{\int e^{-\alpha d^p(x_i) - 2D\varepsilon d(x_i)} dx_i}, \end{aligned}$$

so that the lower bound is proved with $a_1^{-1} = \int e^{-\alpha d^p(x_i)+2D|\varepsilon|d(x_i)}dx_i < \infty$. Similarly

$$\frac{e^{-\varepsilon\sum_{j:j\sim i} d(x_i\omega_j^{-1})}}{Z_i^\omega} \leq \frac{e^{-\varepsilon\sum_{j:j\sim i}(d(\omega_j)+d(x_i))}}{\int e^{-\alpha d^p(x_i)-\varepsilon\sum_{j:j\sim i}(d(\omega_j)-d(x_i))}dx_i}$$
$$= \frac{e^{-2D\varepsilon d(x_i)}}{\int e^{-\alpha d^p(x_i)+2D\varepsilon d(x_i)}dx_i},$$

so that $a_2^{-1} = \int e^{-\alpha d^p(x_i)-2D|\varepsilon|d(x_i)}dx_i$. The case when $\varepsilon > 0$ is similar. □

Lemma 5.5.7. *Suppose $(\mathbb{E}_\Lambda^\omega)_{\Lambda\subset\subset\mathbb{Z}^D,\ \omega\in\Omega}$ is given by* (5.46) *and* (5.47)*, and let $\frac{1}{q}+\frac{1}{p}=1$. Then $\mathbb{E}_i^\omega$ satisfies an SG_q inequality uniformly on the boundary conditions i.e. there exists a constant c_0, independent of ω, such that*

$$\mathbb{E}_i^\omega\left|f-\mathbb{E}_i^\omega f\right|^q \leq c_0\mathbb{E}_i^\omega|\nabla_i f|^q$$

for all locally Lipschitz f, $i\in\mathbb{Z}^D$ and $\omega\in\Omega$.

Proof. Again we follow [69]. We have

$$\mathbb{E}_i^\omega\left|f-\mathbb{E}_i^\omega f\right|^q \leq 2^q\mathbb{E}_i^\omega\left|f-m\right|^q \tag{5.48}$$

for all $m\in\mathbb{R}$. Now for all $L>0$,

$$\mathbb{E}_i^\omega\left|f-m\right|^q = \mathbb{E}_i^\omega\left|f-m\right|^q\mathbf{1}_{\{d\leq L\}} + \mathbb{E}_i^\omega\left|f-m\right|^q\mathbf{1}_{\{d\geq L\}}. \tag{5.49}$$

By Lemma 5.5.6, for the first term we have

$$\mathbb{E}_i^\omega\left|f-m\right|^q\mathbf{1}_{\{d\leq L\}} = \int_{\{d\leq L\}}|f(x_i)-m|^q\frac{e^{-\alpha d^p(x_i)-\varepsilon\sum_{j:j\sim i}d(x_i\omega_j^{-1})}}{Z_i^\omega}dx_i$$
$$\leq a_2e^{2D|\varepsilon|L}\int_{B_L}|f(x_i)-m|^q dx_i,$$

where $B_L=\{x_i\in\mathbb{G}: d(x_i)\leq L\}$. Taking $m=|B_L|^{-1}\int_{B_L}f(x_i)dx_i$, we can continue this using Theorem 3.2.18, which yields

$$\begin{aligned}\mathbb{E}_i^\omega |f-m|^q \mathbf{1}_{\{d\le L\}} &= a_2 e^{2D|\varepsilon|L} P_0(L) \int_{B_L} |\nabla_i f(x_i)|^q dx_i \\ &\le a_2 a_1^{-1} e^{4D|\varepsilon|L+\alpha L^p} P_0(L) \mathbb{E}_i^\omega |\nabla_i f|^q, \qquad (5.50)\end{aligned}$$

where again we have used Lemma 5.5.6. For the second term of (5.49), we can write

$$\begin{aligned}\mathbb{E}_i^\omega |f-m|^q \mathbf{1}_{\{d\ge L\}} &\le \frac{1}{L^p}\mathbb{E}_i^\omega \left(d^p |f-m|^q\right) \qquad (5.51)\\ &\le \frac{A}{L^p}\mathbb{E}_i^\omega|\nabla_i f|^q + \frac{B}{L^p}\mathbb{E}_i^\omega|f-m|^q, \qquad (5.52)\end{aligned}$$

using Lemma 5.5.5. Putting estimates (5.50) and (5.51) in (5.49) yields

$$\mathbb{E}_i^\omega |f-m|^q \le \left(a_2 a_1^{-1} e^{4D|\varepsilon|L+\alpha L^p} P_0(L) + \frac{A}{L^p}\right) \mathbb{E}_i^\omega|\nabla_i f|^q + \frac{B}{L^p}\mathbb{E}_i^\omega|f-m|^q.$$

Taking L large enough so that $B/L^p < 1$, rearranging and combining with (5.48) then yields the result. □

Proof of Theorem 5.5.4. Again we can suppose $f \ge 0$. By part (i) of Theorem 3.2.23, we have that there exists a constant $\tilde{c}$ such that

$$\int f^q \log\left(\frac{f^q}{\int f^q Z^{-1} e^{-\alpha d^p} dx_i}\right) Z^{-1} e^{-\alpha d^p} dx_i \le \tilde{c} \int |\nabla_i f|^q Z^{-1} e^{-\alpha d^p} dx_i$$

for all suitable locally Lipschitz functions f, and where $Z = \int e^{-\alpha d^p} dx_i$. We can then replace f in the above by

$$f(x_i) e^{-\frac{\varepsilon}{q}\sum_{j:j\sim i} d(x_i\omega_j^{-1})} Z^{\frac{1}{q}} (Z_i^\omega)^{-\frac{1}{q}},$$

which yields

$$\begin{aligned}&\mathbb{E}_i^\omega\left(f^q \log\frac{f^q}{\mathbb{E}_i^\omega f^q}\right) + \mathbb{E}_i^\omega\left(f^q \log\left(e^{-\varepsilon\sum_{j:j\sim i} d(x_i\omega_j^{-1})} Z (Z_i^\omega)^{-1}\right)\right)\\ &\le 2^{q-1}\tilde{c}\mathbb{E}_i^\omega|\nabla_i f|^q + 2^{2q-1}\tilde{c}\frac{|\varepsilon|^q D^q}{q^q}\mathbb{E}_i^\omega(f^q)\end{aligned}$$

so that

$$\mathbb{E}_i^\omega\left(f^q \log\frac{f^q}{\mathbb{E}_i^\omega f^q}\right) \leq \tilde{a}\mathbb{E}_i^\omega|\nabla_i f|^q + \tilde{b}\mathbb{E}_i^\omega(f^q) + \mathbb{E}_i^\omega\left(f^q \log\frac{Z_i^\omega}{e^{-\varepsilon\sum_{j:j\sim i} d(x_i\omega_j^{-1})}}\right), \quad (5.53)$$

where $\tilde{a} = 2^{q-1}$ and $\tilde{b} = 2^{2q-1}\tilde{c}\frac{|\varepsilon|^q D^q}{q^q} - \log Z$. We now note that by Lemma 5.5.6,

$$\frac{Z_i^\omega}{e^{-\varepsilon\sum_{j:j\sim i} d(x_i\omega_j^{-1})}} \leq a_1^{-1}e^{2D|\varepsilon|d(x_i)}.$$

Using this in (5.53) then yields

$$\begin{aligned}
\mathbb{E}_i^\omega\left(f^q \log\frac{f^q}{\mathbb{E}_i^\omega f^q}\right) &\leq \tilde{a}\mathbb{E}_i^\omega|\nabla_i f|^q + \left(\tilde{b} - \log a_1\right)\mathbb{E}_i^\omega(f^q) + 2D|\varepsilon|\mathbb{E}_i^\omega(df^q) \\
&\leq \tilde{a}\mathbb{E}_i^\omega|\nabla_i f|^q + \left(\tilde{b} - \log a_1 + 2D|\varepsilon|\right)\mathbb{E}_i^\omega(f^q) + 2D|\varepsilon|\mathbb{E}_i^\omega(d^p f^q) \\
&\leq (\tilde{a} + 2D|\varepsilon|A)\,\mathbb{E}_i^\omega|\nabla_i f|^q + \left(\tilde{b} - \log a_1 + 2D|\varepsilon|(B+1)\right)\mathbb{E}_i^\omega(f^q),
\end{aligned}$$

using Lemma 5.5.5. Thus $\mathbb{E}_i^\omega$ satisfies a defective DLS_q inequality, with constants independent of the boundary conditions. To complete the proof we can tighten the inequality using Proposition 3.1.7 in conjunction with Lemma 5.5.7. □

Chapter 6

Ergodicity for Infinite Particle Systems with Locally Conserved Quantities

6.1 Introduction

In this chapter we present some results obtained as part of a joint research project with M. Neklyudov and B. Zegarliński. We treat it as a stand-alone chapter, since the setting and methods used here are quite different to those of the preceding chapters. Having said that, we do still maintain the central theme of this monograph, in that we continue to study generators given by sums of non-commuting degenerate vector fields. Indeed, here we will be concerned with the long-time behaviour of a class of Markov semigroups $(P_t)_{t\geq 0}$ whose generators are defined in Hörmander form as follows:

$$\mathcal{L} = \sum_i \mathbf{X}_i^2,$$

where the $\mathbf{X}_i$'s form the aforementioned family of degenerate non-commuting vector fields. In the above chapters we have concentrated on the case when $\mathcal{L}$ has been defined in terms of the natural fields on H-type groups, which satisfy Hörmander's condition; here the scene is different, in that we investigate a situation when the family of fields is infinite and a commutator of any order does not remove degeneration. Functional inequalities will again play a major role, as we aim to determine the behaviour of these generators and their associated semigroups.

In particular, we will be interested in the situation when we have "locally conserved quantities", that is when any operator given by

$$\mathcal{L}_\Lambda \equiv \sum_{i\in\Lambda} \mathbf{X}_i^2,$$

where the sum is over a finite set of indices, has a non-trivial set of harmonic functions, while for the full generator this is not the case (at least formally). To model such a situation we consider an infinite product space and fields of the following form

$$\mathbf{X}_{i,j} = \partial_i V(x)\partial_j - \partial_j V(x)\partial_i,$$

with ∂_i denoting the partial derivative with respect to the i-th coordinate and $\partial_i V(x)$ indicating some polynomial coefficients.

Generators of a similar type appear in the study of dissipative dynamics in which certain quantities are preserved — see for example [21, 23] and [58], where systems of harmonic oscillators perturbed by noise are considered. A further example of a physical model very closely related to our setup is the heat conduction model discussed in [22] and [60]. For more information in this direction, in particular in connection with an effort to explain the so-called Fourier law of heat conduction, we refer to a comprehensive review [37], as well as [36] and the references therein.

The classical approach to studying the asymptotic behaviour of conservative reversible interacting particle systems employs either functional inequalities together with some special norm-bound of the semigroup (see for instance [24, 25] and [83]), or some kind of approximation of the dynamics by finite dimensional ones, together with sharp estimates of their spectral gaps ([81, 85]). The approach we take is quite different, in the sense that we do not use any approximation techniques, but rather exploit the structure of the Lie algebra generated by the corresponding vector fields to derive the necessary estimates directly.

One other motivation to study the semigroup $(P_t)_{t\geq 0}$ associated to this particular generator comes from the fact that, since V is formally conserved under the action of P_t, we can see that there is a family of invariant measures formally given by "$e^{-\frac{V}{r}}dx$" for all $r > 0$. On the one hand, the semigroup $(P_t)_{t\geq 0}$ is quite simple, since we can calculate many quantities we are interested in directly. On the other hand, standard methods from interacting particle theory [93, 95] do not help because they require some type of strong non-degeneracy condition such as Hörmander's condition. Another difficulty stems from the intrinsic difference between the infinite dimensional case we consider, and the finite dimensional case i.e. the case when V depends on only a finite number of variables, and instead of the lattice we use its truncation with a periodic boundary condition. Indeed, in the finite dimensional case we can notice that V is a non-trivial fixed point for P_t, and therefore the semigroup is strictly not ergodic[1]. This reasoning turns out to be incorrect in the infinite dimensional case. The situation here is more subtle because the expression V is only formal (and would be equal to infinity on the support set of the invariant measure).

Our goal is to give a detailed study of the case when the coefficients of the fields are linear, and to show that the system is ergodic with polynomial rate of convergence.

The chapter is organised as follows. In Section 6.2 we introduce the basic notation and state an infinite system of stochastic differential equations of interest to us. In Section 6.3 we show the existence of a mild solution to this system, and continue in Section 6.4 by identifying a family of non-trivial invariant measures. Because of the special non-commutative features of the fields and the form of the generator, this is slightly more cumbersome than otherwise. Section 6.5 provides a certain characterisation of invariant Sobolev-type subspaces, while Section 6.6 is devoted to the demonstration of the ergodicity of the system with polynomial rate of convergence. We conclude with a section in which we use previously obtained information to derive Liggett-Nash-type inequalities.

Throughout this chapter we will make use of the theory of stochastic differential equations in Hilbert spaces, as outlined in [109]. For the sake of completeness we include a very brief description of some of the basic ideas in Appendix A, though [109] should be referred to for the details.

---

[1]Recall that the semigroup $P_t = e^{t\mathcal{L}}$ is ergodic in $L^2(\mu)$, where μ is an invariant measure, if and only if $\mathcal{L}u = 0$ for $u \in \mathcal{D}(\mathcal{L}) \Rightarrow u$ is constant — see for example Proposition 2.3 of [4]

6.2 Setting

As in Chapter 5, let $\mathbb{Z}^D$ be the D-dimensional square lattice for some fixed $D \in \mathbb{N}$, equipped with the l_1 lattice metric $dist(\cdot, \cdot)$ defined by

$$dist(i,j) := |i-j|_1 \equiv \sum_{l=1}^{D} |i_l - j_l|$$

for $i = (i_1, \ldots, i_D), j = (j_1, \ldots, j_D) \in \mathbb{Z}^D$. As before, for $i, j \in \mathbb{Z}^D$ we will write $i \sim j$ whenever $dist(i,j) = 1$, i.e. when i and j are neighbours in the lattice.

Let $\Omega = (\mathbb{R})^{\mathbb{Z}^D}$ and define the Hilbert spaces

$$E_\alpha := \left\{ x \in \Omega : |x|^2_{E_\alpha} := \sum_{i \in \mathbb{Z}^D} x_i^2 e^{-\alpha |i|_1} < \infty \right\}$$

for $\alpha > 0$, and

$$H := \left\{ (h^{(1)}, \ldots, h^{(D)}) \in \Omega^D : |(h^{(1)}, \ldots, h^{(D)})|^2_H := \sum_{i \in \mathbb{Z}^D} \sum_{k=1}^{D} \left(h_i^{(k)} \right)^2 < \infty \right\},$$

with inner products given by

$$\langle x, y \rangle_{E_\alpha} := \sum_{i \in \mathbb{Z}^D} x_i y_i e^{-\alpha |i|_1}$$

for $x, y \in E_\alpha$ and

$$\langle (g^{(1)}, \ldots, g^{(D)}), (h^{(1)}, \ldots, h^{(D)}) \rangle_H := \sum_{i \in \mathbb{Z}^D} \sum_{k=1}^{D} g_i^{(k)} h_i^{(k)}$$

for $(g^{(1)}, \ldots, g^{(D)}), (h^{(1)}, \ldots, h^{(D)}) \in H$ respectively.

Let $\mu_{\mathbf{G}}$ be a Gaussian probability measure on $(E_\alpha, \mathcal{B}(E_\alpha))$ with mean zero and covariance $\mathbf{G}$ (see Appendix A). We assume that the inverse $\mathbf{G}^{-1}$ of the covariance is of finite range i.e.

$$\mathbf{M}_{i,j} := \mathbf{G}^{-1}_{i,j} = 0 \quad \text{if } dist(i,j) > R,$$

for some $R > 0$, and that $|\mathbf{M}_{i,j}| \leq M$ for all $i, j \in \mathbb{Z}^D$.

We are now in a position to describe the system we are going to consider. Indeed, let

$$W = \left\{ \left(W^{(1)}, \ldots, W^{(D)} \right) \right\}$$

be a cylindrical Wiener process in H (see Appendix A).

We introduce the following notation: for $i = (i_1, \ldots, i_D) \in \mathbb{Z}^D$ and $k \in \{1, \ldots, D\}$ define

$$i^{\pm}(k) := (i_1, \ldots, i_{k-1}, i_k \pm 1, i_{k+1}, \ldots, i_D).$$

We also define, for $x \in E_\alpha$, $i \in \mathbb{Z}^D$,

$$V_i(x) := \sum_{j \in \mathbb{Z}^D} x_i \mathbf{M}_{i,j} x_j,$$

which is a finite sum since $\mathbf{M}_{i,j} = 0$ if $dist(i,j) > R$, and for all finite subsets $\Lambda \subset \mathbb{Z}^D$ set

$$V_\Lambda(x) := \sum_{i \in \Lambda} V_i(x).$$

Using the formal expression

$$V(x) := \frac{1}{2} \sum_{i \in \mathbb{Z}^D} V_i(x),$$

it will be convenient to simplify the notation for $\partial_i V_i$ as follows

$$\partial_i V(x) = \frac{1}{2} \partial_i \left(\sum_{j,l \in \mathbb{Z}^D} x_j \mathbf{M}_{j,l} x_l \right) \equiv \sum_{j \in \mathbb{Z}^D} \mathbf{M}_{i,j} x_j = \partial_i V_i.$$

We consider the following system of Stratonovich SDEs:

$$dY_i(t) = \sum_{k=1}^{D} \left(\partial_{i^-(k)} V(Y(t)) \circ dW^{(k)}_{i^-(k)}(t) - \partial_{i^+(k)} V(Y(t)) \circ dW^{(k)}_i(t) \right) \tag{6.1}$$

for $i \in \mathbb{Z}^D$ and $t \geq 0$.

Remark 6.2.1. *We consider a system of Stratonovich rather than Itô SDEs, since we are* *trying to write down a system that has a particular generator. As outlined in [16], it is more concise to do this in terms of Stratonovich SDEs. However, in the next section we confirm that this system does indeed give rise to the desired generator, by converting to Itô integrals, and rigorously describe what we mean by a solution.*

6.3 Existence of a mild solution

In this section we show that the system (6.1) has a mild solution $Y(t)$ taking values in the Hilbert space E_α. The first step is to write the system in Itô form. To this end, we have

$$\begin{aligned} dY_i(t) = &\sum_{k=1}^{D} \Big(\partial_{i^-(k)} V(Y(t)) dW^{(k)}_{i^-(k)}(t) - \partial_{i^+(k)} V(Y(t)) dW^{(k)}_{i}(t) \Big) \\ &+ \frac{1}{2} \sum_{k=1}^{D} \Big(d \Big[\partial_{i^-(k)} V(Y(\cdot)), W^{(k)}_{i^-(k)}(\cdot) \Big]_t - d \Big[\partial_{i^+(k)} V(Y(\cdot)), W^{(k)}_{i}(\cdot) \Big]_t \Big) \end{aligned} \tag{6.2}$$

for all $i \in \mathbb{Z}^D$ and $t \geq 0$, where $[\cdot,\cdot]_t$ is the quadratic covariation process, as introduced in Appendix A.

Hence, by Itô's formula,

$$\begin{aligned} \Big[\partial_{i^-(k)} V(Y(\cdot)), W^{(k)}_{i^-(k)}(\cdot) \Big]_t &= \Bigg[\sum_{j\in\mathbb{Z}^D} \int_0^{\cdot} \partial_j \partial_{i^-(k)} V(Y(s)) dY_j(s), \int_0^{\cdot} dW^{(k)}_{i^-(k)}(s) \Bigg]_t \\ &= \sum_{j\in\mathbb{Z}^D} \Big[\int_0^{\cdot} \partial_j \partial_{i^-(k)} V(Y(s)) \partial_{j^-(k)} V(Y(s)) dW^{(k)}_{j^-(k)}(s), \int_0^{\cdot} dW^{(k)}_{i^-(k)}(s) \Big]_t \\ &\quad - \sum_{j\in\mathbb{Z}^D} \Big[\int_0^{\cdot} \partial_j \partial_{i^-(k)} V(Y(s)) \partial_{j^+(k)} V(Y(s)) dW^{(k)}_{j}(s), \int_0^{\cdot} dW^{(k)}_{i^-(k)}(s) \Big]_t \\ &= \int_0^t \partial^2_{i,i^-(k)} V(Y(s)) \partial_{i^-(k)} V(Y(s)) ds - \int_0^t \partial^2_{i^-(k)} V(Y(s)) \partial_i V(Y(s)) ds \end{aligned}$$

for all $i \in \mathbb{Z}^D$. By a similar calculation, and using this in (6.2), we see that

$$\begin{aligned} dY_i(t) = &\sum_{k=1}^{D} \left(\partial_{i^-(k)} V(Y(t)) dW^{(k)}_{i^-(k)}(t) - \partial_{i^+(k)} V(Y(t)) dW^{(k)}_i(t) \right) \\ &- \frac{1}{2} \sum_{k=1}^{D} \Big\{ \left(\partial^2_{i^-(k)} V(Y(t)) + \partial^2_{i^+(k)} V(Y(t)) \right) \partial_i V(Y(t)) \\ &\quad - \partial^2_{i,i^-(k)} V(Y(t)) \partial_{i^-(k)} V(Y(t)) - \partial^2_{i,i^+(k)} V(Y(t)) \partial_{i^+(k)} V(Y(t)) \Big\} \, dt \end{aligned} \tag{6.3}$$

for all $i \in \mathbb{Z}^D$.

Recall now that $\partial_j V(x) = \sum_{l \in \mathbb{Z}^D} \mathbf{M}_{j,l} x_l$ for all $j \in \mathbb{Z}^D$ so that $\partial^2_{i,j} V(x) = \mathbf{M}_{i,j}$, $\forall i, j \in \mathbb{Z}^D$. Thus the system (6.3) can be written as

$$\begin{aligned} dY_i(t) = &\sum_{k=1}^{D} \left(\partial_{i^-(k)} V(Y(t)) dW^{(k)}_{i^-(k)}(t) - \partial_{i^+(k)} V(Y(t)) dW^{(k)}_i(t) \right) \\ &- \frac{1}{2} \sum_{k=1}^{D} \Big\{ \left(\mathbf{M}_{i^-(k), i^-(k)} + \mathbf{M}_{i^+(k), i^+(k)} \right) \partial_i V(Y(t)) \\ &\quad - \mathbf{M}_{i,i^-(k)} \partial_{i^-(k)} V(Y(t)) - \mathbf{M}_{i,i^+(k)} \partial_{i^+(k)} V(Y(t)) \Big\} dt \end{aligned} \tag{6.4}$$

for all $i \in \mathbb{Z}^D$ and $t \geq 0$.

We now claim that we can write this system in operator form:

$$dY(t) = \mathbf{A} Y(t) dt + \mathbf{B}(Y(t)) dW(t), \tag{6.5}$$

where $\mathbf{A}$ is a bounded linear mapping from E_α to E_α given by

$$(\mathbf{A}x)_i := \sum_{k=1}^{D} a^{(k)}_i(x), \quad i \in \mathbb{Z}^D, \tag{6.6}$$

with

$$a_i^{(k)}(x) = -\frac{1}{2}\Bigg\{\big(\mathbf{M}_{i^-(k),i^-(k)} + \mathbf{M}_{i^+(k),i^+(k)}\big)\sum_{l\in\mathbb{Z}^D}\mathbf{M}_{l,i}x_l \\ - \mathbf{M}_{i,i^-(k)}\sum_{l\in\mathbb{Z}^D}\mathbf{M}_{l,i^-(k)}x_l - \mathbf{M}_{i,i^+(k)}\sum_{l\in\mathbb{Z}^D}\mathbf{M}_{l,i^+(k)}x_l\Bigg\}, \tag{6.7}$$

and where $\mathbf{B} : E_\alpha \to L_{HS}(H, E_\alpha)$[2] is a bounded linear operator given by

$$\big(\mathbf{B}(x)(h^{(1)}, \dots, h^{(D)})\big)_i := \sum_{k=1}^{D}\Big(\partial_{i^-(k)}V(x)h^{(k)}_{i^-(k)} - \partial_{i^+(k)}V(x)h^{(k)}_i\Big) \tag{6.8}$$

for $x \in E_\alpha$, $(h^{(1)}, \dots, h^{(D)}) \in H$ and $i \in \mathbb{Z}^D$.

Indeed, the fact that $\mathbf{A} : E_\alpha \to E_\alpha$ is a bounded linear operator follows from the assumption that the constants $\mathbf{M}_{i,j}$ are uniformly bounded by a constant M. To show that $\mathbf{B}$ is bounded from E_α to $L_{HS}(H, E_\alpha)$, first define, for $i \in \mathbb{Z}^D$, $e(i) \in \Omega$ by

$$(e(i))_j := \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise,} \end{cases}$$

and for $i \in \mathbb{Z}^D, k \in \{1, \dots, D\}$, let f_i^k be the element in H given by

$$f_i^k := (0, \dots, 0, e(i), 0, \dots, 0),$$

where the $e(i)$ occurs in the k-th coordinate. Then

$$\big\{f_i^k : i \in \mathbb{Z}^D, k \in \{1, \dots, D\}\big\}$$

[2]We denote by $L_{HS}(H, E_\alpha)$ the space of all Hilbert-Schmidt operators from H to E_α, where we recall that an operator $\mathbf{L} : H \to E_\alpha$ is Hilbert-Schmidt if $\sum_{i\in I}\|\mathbf{L}g_i\|^2_{E_\alpha} < \infty$ for an orthonormal basis $\{g_i : i \in I\}$ of H.

is an orthonormal basis for H. Let $x \in E_\alpha$. Then

$$\|\mathbf{B}(x)\|_{HS}^2 = \sum_{i\in\mathbb{Z}^D}\sum_{k=1}^{D} \left|\mathbf{B}(x)(f_i^k)\right|_{E_\alpha}^2 .$$

Now by definition

$$\left(\mathbf{B}(x)(f_i^k)\right)_j = \partial_{j^-(k)}V(x)\,(e(i))_{j^-(k)} - \partial_{j^+(k)}V(x)(e(i))_j,$$

so that

$$\begin{aligned}
\left|\mathbf{B}(x)(f_i^k)\right|_{E_\alpha}^2 &= \sum_{j\in\mathbb{Z}^D}\left(\partial_{j^-(k)}V(x)\,(e(i))_{j^-(k)} - \partial_{j^+(k)}V(x)(e(i))_j\right)^2 e^{-\alpha|j|_1} \\
&= \left(\partial_i V(x)\right)^2 e^{-\alpha|i^+(k)|_1} + \left(\partial_{i^+(k)}V(x)\right)^2 e^{-\alpha|i|_1} \\
&= \left(\sum_{l:|l-i|_1\le R}\mathbf{M}_{i,l}x_l\right)^2 e^{-\alpha|i^+(k)|_1} + \left(\sum_{l:|l-i^+(k)|_1\le R}\mathbf{M}_{i^+(k),l}x_l\right)^2 e^{-\alpha|i|_1} \\
&\le C\left[\left(\sum_{l:|l-i|_1\le R}x_l^2\right)e^{-\alpha|i^+(k)|_1} + \left(\sum_{l:|l-i^+(k)|_1\le R}x_l^2\right)e^{-\alpha|i|_1}\right] \\
&\le Ce^{\alpha}\left[\left(\sum_{l:|l-i|_1\le R}x_l^2\right)e^{-\alpha|i|_1} + \left(\sum_{l:|l-i^+(k)|_1\le R}x_l^2\right)e^{-\alpha|i^+(k)|_1}\right] \\
&\le Ce^{(R+1)\alpha}\left[\left(\sum_{l:|l-i|_1\le R}x_l^2 e^{-\alpha|l|_1}\right) + \left(\sum_{l:|l-i^+(k)|_1\le R}x_l^2 e^{-\alpha|l|_1}\right)\right]
\end{aligned}$$

where $C = (2R+1)^D M^2$. Thus

$$\begin{aligned}
\|\mathbf{B}(x)\|_{HS} &= \sum_{i\in\mathbb{Z}^D}\sum_{k=1}^{D}\left|\mathbf{B}(x)(f_i^k)\right|_{E_\alpha}^2 \\
&\le Ce^{(R+1)\alpha}\sum_{k=1}^{D}\sum_{i\in\mathbb{Z}^D}\left[\left(\sum_{l:|l-i|_1\le R}x_l^2 e^{-\alpha|l|_1}\right) + \left(\sum_{l:|l-i^+(k)|_1\le R}x_l^2 e^{-\alpha|l|_1}\right)\right] \\
&= 2D(2R+1)^D Ce^{(R+1)\alpha}|x|_{E_\alpha}^2,
\end{aligned}$$

which proves our claim that $\mathbf{B}$ is bounded from E_α to $L_{HS}(H, E_\alpha)$.

We then have the following existence theorem.

Proposition 6.3.1. *Consider the stochastic evolution equation*

$$dY(t) = \mathbf{A}Y(t)dt + \mathbf{B}(Y(t))dW(t), \qquad Y_0 = x \in E_\alpha,\ t \geq 0, \tag{6.9}$$

where $\mathbf{A}$ *and* $\mathbf{B}$ *are given by* (6.6) *and* (6.8) *respectively, and* $(W(t))_{t\geq 0}$ *is a cylindrical Wiener process in* H*. This equation has a mild solution* $(Y(t))_{t\geq 0}$ *(see Appendix A) taking values in* E_α*, which is unique up to processes satisfying*

$$\mathbb{P}\left(\int_0^T |Y(s)|_{E_\alpha} ds < \infty\right) = 1, \tag{6.10}$$

for all $T > 0$*. Moreover,* $(Y(t))_{t\geq 0}$ *has a continuous modification, and is a strong Markov process.*

Proof. We have shown above that $\mathbf{A} : E_\alpha \to E_\alpha$ is a bounded linear operator, so that it is the infinitesimal generator of a strongly continuous semigroup[3] $(S_t)_{t\geq 0}$ on E_α. Indeed, $\mathbf{A}$ can be thought of as a bounded linear perturbation of 0, which is trivially the generator of a strongly continuous semigroup. We have also shown that $\mathbf{B} : E_\alpha \to L_{HS}(H, E_\alpha)$ is bounded. Hence the result follows immediately from Theorem A.4.1 of Appendix A. □

Lemma 6.3.2. *The mild solution* $(Y(t))_{t\geq 0}$ *to* (6.9) *solves the martingale problem for the operator*

$$\mathcal{L} = \frac{1}{4}\sum_{i\in\mathbb{Z}^D}\sum_{j\in\mathbb{Z}^D : j\sim i} (\partial_i V(x)\partial_j - \partial_j V(x)\partial_i)^2.$$

Proof. By Itô's formula, we have for any suitable function f,

$$\begin{aligned} f(Y(t)) = f(Y(0)) &+ \sum_{i\in\mathbb{Z}^D}\int_0^t \partial_i f(Y(s))dY_i(s) \\ &+ \frac{1}{2}\sum_{i,j\in\mathbb{Z}^D}\int_0^t \partial^2_{i,j} f(Y(s))d\left[Y_i, Y_j\right]_s . \end{aligned} \tag{6.11}$$

[3] A semigroup $(P_t)_{t\geq 0}$ on a Banach space $\mathcal{B}$ is strongly continuous if $\lim_{t\to 0} |P_t f - f|_{\mathcal{B}} = 0$ for all $f \in \mathcal{B}$.

We can then calculate from (6.3) that

$$d[Y_i, Y_j]_t := \begin{cases} -\partial_i V(Y(t))\partial_{i^-(k)} V(Y(t))dt & \text{if } j = i^-(k), \\ \sum_{k=1}^D \left\{ \left(\partial_{i^-(k)} V(Y(t))\right)^2 + \left(\partial_{i^+(k)} V(Y(t))\right)^2 \right\} dt & \text{if } j = i, \\ -\partial_i V(Y(t))\partial_{i^+(k)} V(Y(t))dt & \text{if } j = i^-(k), \end{cases}$$

so that

$$\begin{aligned} &\sum_{i,j \in \mathbb{Z}^D} \int_0^t \partial_{i,j}^2 f(Y(s)) d\left[Y_i, Y_j\right]_s \\ &= \sum_{i \in \mathbb{Z}^D} \int_0^t \partial_i^2 f(Y(s)) \sum_{k=1}^D \left\{ \left(\partial_{i^-(k)} V(Y(t))\right)^2 + \left(\partial_{i^+(k)} V(Y(t))\right)^2 \right\} dt \\ &\quad - 2 \sum_{k=1}^D \sum_{i \in \mathbb{Z}^D} \int_0^t \partial_{i,i^-(k)}^2 f(Y(s)) \partial_i V(Y(t)) \partial_{i^-(k)} V(Y(t)) dt. \end{aligned}$$

Thus, if we set

$$\begin{aligned} \mathcal{L} = &\frac{1}{2} \sum_{i \in \mathbb{Z}^D} \sum_{k=1}^D \left\{ \left(\partial_{i^-(k)} V(x)\right)^2 + \left(\partial_{i^+(k)} V(x)\right)^2 \right\} \partial_i^2 \\ &- \sum_{i \in \mathbb{Z}^D} \sum_{k=1}^D \partial_i V(x) \partial_{i^-(k)} V(x) \partial_{i,i^-(k)}^2 \\ &- \frac{1}{2} \sum_{i \in \mathbb{Z}^D} \sum_{k=1}^D \left\{ \left(\partial_{i^-(k)}^2 V(x) + \partial_{i^+(k)}^2 V(x) \right) \partial_i V(x) \right. \\ &\qquad \left. - \partial_{i,i^-(k)}^2 V(x) \partial_{i^-(k)} V(x) - \partial_{i,i^+(k)}^2 V(x) \partial_{i^+(k)} V(x) \right\} \partial_i, \end{aligned}$$

by combining (6.3) with (6.11) we see that

$$\mathbb{E}\left[f(Y(t)) - f(Y(0)) - \mathcal{L} f(Y(t))\right] = 0$$

i.e. $f(Y(t)) - f(Y(0)) - \mathcal{L} f(Y(t))$ is a martingale, or equivalently, that $\mathcal{L}$ is the generator of our system. One can then check by direct calculation that we have

$$\mathcal{L} = \frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j \in \mathbb{Z}^D : j \sim i} \left(\partial_i V(x) \partial_j - \partial_j V(x) \partial_i \right)^2 .$$

□

For $n \in \{0, 1, \dots\}$, let $\mathcal{U}C_b^n \equiv \mathcal{U}C_b^n(E_\alpha), \alpha > 0$ denote the set of all functions which are uniformly continuous and bounded, together with their Fréchet derivatives up to order n.

Corollary 6.3.3. *The semigroup $(P_t)_{t \geq 0}$ acting on $\mathcal{U}C_b(E_\alpha)$ corresponding to the system* (6.9) *is Feller* [4] *and can be represented by the formula*

$$P_t f(\cdot) = \mathbb{E} f\left(Y(t, \cdot) \right), \quad t \geq 0,$$

where $Y(t, x)$ is a mild solution to the system (6.9) *with initial condition $x \in E_\alpha$.*

Proof. This result is standard and follows immediately from Theorems 9.14 and 9.16 of [109]. □

Example 6.3.4. *Suppose that, for all $i \in \mathbb{Z}^D$,*

$$\mathbf{M}_{i,i} = 1, \quad \mathbf{M}_{i,j} = 0 \quad \text{if} \quad i \neq j.$$

Then $\partial_i V(x) = x_i$, and the system (6.9) *becomes*

$$dY_i(t) = -\sum_{k=1}^{D} Y_i(t) dt + \sum_{k=1}^{D} \left(Y_{i^-(k)}(t) dW_{i^-(k)}^{(k)}(t) - Y_{i^+(k)}(t) dW_i^{(k)}(t) \right)$$

for all $i \in \mathbb{Z}^D$, which has generator

$$\mathcal{L} = \frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j \in \mathbb{Z}^D : j \sim i} \left(x_i \partial_j - x_j \partial_i \right)^2 . \tag{6.12}$$

[4] Recall that a semigroup $(P_t)_{t \geq 0}$ on $\mathcal{U}C_b(E_\alpha)$ is Feller if $(P_t)_{t \geq 0}$ is strongly continuous and such that $0 \leq f \leq 1 \Rightarrow 0 \leq P_t f \leq 1$.

In this case the Gaussian measure $\mu_{\mathbf{G}}$ on E_α is the product Gaussian measure.

As mentioned, very closely related generators are considered in the physical models for heat conduction described in [21, 22, 23, 58] and [60]. A related model is also considered in [42]. However, there are some major differences between the system considered there and the one we investigate. Indeed, in [42] Hörmander's condition is assumed to be satisfied, and the system is finite dimensional. Moreover, it is shown that there is a unique invariant measure for the system they investigate, which as we will see, is not the case in our set-up.

Remark 6.3.5. *Let $(r_{i,j}, \theta_{i,j})$ be polar coordinates in the plane (x_i, x_j). Then*

$$\frac{\partial}{\partial \theta_{i,j}} = x_i \partial_j - x_j \partial_i.$$

Therefore in Example 6.3.4

$$\mathcal{L} = \frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j \in \mathbb{Z}^D : j \sim i} \frac{\partial^2}{\partial \theta_{i,j}^2}.$$

The operator $-\frac{\partial^2}{\partial \theta_{i,j}^2}$ is the Hamiltonian for the rigid rotor on the plane. Thus, the operator $-\mathcal{L}$ is the Hamiltonian of a chain of coupled rigid rotors.

6.4 Invariant measure

Suppose $(Y(t))_{t \geq 0}$ is the unique mild solution to the evolution equation (6.9) in the Hilbert space E_α i.e.

$$dY(t) = \mathbf{A}Y(t)dt + \mathbf{B}(Y(t))dW(t)$$

where $\mathbf{A}$ and $\mathbf{B}$ are given by (6.6) and (6.8) respectively, and $(W(t))_{t \geq 0}$ is a cylindrical Wiener process in H. Let $(P_t)_{t \geq 0}$ be the corresponding semigroup, described in Corollary 6.3.3.

For $i, j \in \mathbb{Z}^D$, define

$$\mathbf{X}_{i,j} := \partial_i V(x) \partial_j - \partial_j V(x) \partial_i,$$

so that by Lemma 6.3.2,

$$\mathcal{L} = \frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j \in \mathbb{Z}^D : j \sim i} \mathbf{X}_{i,j}^2$$

is the generator of our system. We will need the following Lemma:

Lemma 6.4.1.

$$\mu_{r\mathbf{G}}\left(f\mathbf{X}_{i,j}g\right) = -\mu_{r\mathbf{G}}\left(g\mathbf{X}_{i,j}f\right)$$

for all $f, g \in \mathcal{U}C_b^2(E_\alpha)$, $i, j \in \mathbb{Z}^D$ and $r > 0$, where we recall that the measure $\mu_{r\mathbf{G}}$ is the Gaussian measure on E_α with covariance matrix $r\mathbf{G}$.

Proof. For finite subsets $\Lambda \subset \mathbb{Z}^D$ and $\omega \in \Omega$, denote by $\mathbb{E}_\Lambda^\omega$ the conditional measure of $\mu_{r\mathbf{G}}$, given the coordinates outside Λ coincide with those of ω. Then we have that

$$\mathbb{E}_\Lambda^\omega(f) = \int_{\mathbb{R}^\Lambda} f(x_\Lambda \cdot \omega_{\Lambda^c}) \frac{e^{-\frac{1}{2r}\sum_{k\in\Lambda} V_k(x_\Lambda\cdot\omega_{\Lambda^c})}}{Z_\Lambda^\omega} dx_\Lambda,$$

where $x_\Lambda \cdot \omega_{\Lambda^c}$ is the element of Ω given by

$$(x_\Lambda \cdot \omega_{\Lambda^c})_i = \begin{cases} x_i & \text{if } i \in \Lambda, \\ \omega_i & \text{if } i \in \Lambda^c, \end{cases}$$

and Z_Λ^ω is the normalisation constant. Now fix $i, j \in \mathbb{Z}^D$ and suppose that Λ is such that $\{i, j\} \subset \Lambda$. Then for $f, g \in \mathcal{U}C_b^2(E_\alpha)$

$$\begin{aligned}
\mathbb{E}_\Lambda^\omega\left(f\mathbf{X}_{i,j}g\right) &= \int_{\mathbb{R}^\Lambda} f(x_\Lambda \cdot \omega_{\Lambda^c})\mathbf{X}_{i,j}g(x_\Lambda \cdot \omega_{\Lambda^c}) \frac{e^{-\frac{1}{2r}\sum_{k\in\Lambda} V_k(x_\Lambda\cdot\omega_{\Lambda^c})}}{Z_\Lambda^\omega} dx_\Lambda \\
&= -\mathbb{E}_\Lambda^\omega\left(g\mathbf{X}_{i,j}f\right) \\
&\quad + \mathbb{E}_\Lambda^\omega\left(fg\left[\partial_i\partial_j V(x) - \partial_j\partial_i V(x)\right]\right) \\
&\quad + r^{-1}\mathbb{E}_\Lambda^\omega\left(fg\left[\partial_i V(x)\partial_j V(x) - \partial_j V(x)\partial_i V(x)\right]\right) = -\mathbb{E}_\Lambda^\omega\left(g\mathbf{X}_{i,j}f\right)
\end{aligned}$$

by integration by parts. Thus

$$\mu_{r\mathbf{G}}\left(f\mathbf{X}_{i,j}g\right) = \mu_{r\mathbf{G}}\mathbb{E}_\Lambda^\cdot\left(f\mathbf{X}_{i,j}g\right) = -\mu_{r\mathbf{G}}\mathbb{E}_\Lambda^\cdot\left(g\mathbf{X}_{i,j}f\right) = -\mu_{r\mathbf{G}}\left(g\mathbf{X}_{i,j}f\right).$$

□

The following result shows that for all $r > 0$, $\mu_{r\mathbf{G}}$ is reversible for the system (6.9).

Theorem 6.4.2. *For all* $f, g \in \mathcal{U}C_b^2(E_\alpha)$ *and* $r > 0$*, we have*

$$\mu_{r\mathbf{G}}\left(fP_t g\right) = \mu_{r\mathbf{G}}\left(gP_t f\right). \tag{6.13}$$

Proof. It is enough to show that (6.13) holds for $f, g \in \mathcal{U}C_b^2(E_\alpha)$ depending only on a finite number of coordinates. Indeed, by an approximation argument, if (6.13) is true for $f, g \in \mathcal{U}C_b^2(E_\alpha)$ depending only on a finite number of coordinates, it follows that it is also true for general $f, g \in \mathcal{U}C_b^2(E_\alpha)$, using the contractivity of $(P_t)_{t\geq 0}$.

In view of this, suppose $f(x) = f\left(\{x_i\}_{|i|_1 \leq n}\right)$ and $g(x) = g\left(\{x_i\}_{|i|_1 \leq n}\right)$ for some n. Note that the generator $\mathcal{L}$ can be rewritten as

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^{D}\sum_{i\in\mathbb{Z}^D}\mathbf{X}^2_{i,i^+(k)}.$$

We decompose $\mathcal{L}$ further, by writing

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^{D}\sum_{m\in\{0,\ldots,R+1\}^D}\left(\sum_{i\in\otimes_{\sigma=1}^{D}((R+2)\mathbb{Z}+m_\sigma)}\mathbf{X}^2_{i,i^+(k)}\right),$$

and define for $m = (m_1, \ldots, m_D) \in \{0, ..., R+1\}^D$, $k \in \{1, \ldots, D\}$,

$$\mathcal{L}_m^{(k)} := \sum_{i\in\otimes_{\sigma=1}^{D}((R+2)\mathbb{Z}+m_\sigma)}\mathbf{X}^2_{i,i^+(k)}$$

so that

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^{D}\sum_{m\in\{0,\ldots,R+1\}^D}\mathcal{L}_m^{(k)}.$$

By construction, for fixed $k \in \{1, \ldots, D\}$ and $m \in \{0, ..., R+1\}^D$, we claim that for any $i, j \in \otimes_{\sigma=1}^{D}\left((R+2)\mathbb{Z}+m_\sigma\right)$

$$\left[\mathbf{X}_{i,i^+(k)}, \mathbf{X}_{j,j^+(k)}\right] = 0. \tag{6.14}$$

For $i = j$ this is clear. If $i \neq j$, we have

$$\left[\mathbf{X}_{i,i^+(k)}, \mathbf{X}_{j,j^+(k)}\right] = \left[\partial_i V(x)\partial_{i^+(k)} - \partial_{i^+(k)} V(x)\partial_i, \partial_j V(x)\partial_{j^+(k)} - \partial_{j^+(k)} V(x)\partial_j\right]. \tag{6.15}$$

Now, $\partial_j V(x)$ depends only on coordinates l such that $|j - l|_1 \leq R$, and for all such l

$$\begin{aligned} |i^+(k) - l|_1 &\geq |i^+(k) - j|_1 - |j - l|_1 \\ &\geq R + 1 - R \\ &= 1, \end{aligned}$$

so that $\partial_j V(x)$ does not depend on coordinate $i^+(k)$ for any k. Thus

$$\partial_{i^+(k)}\partial_j V(x) = 0.$$

Similarly

$$\partial_{i^+(k)}\partial_{j^+(k)} V(x) = \partial_i\partial_j V(x) = \partial_i\partial_{j^+(k)} V(x) = 0,$$

which, when used in (6.15) proves the claim (6.14). Thus for any $k \in \{1, \ldots, D\}$ and $m \in \{0, ..., R + 1\}^D$,

$$S_t^{(k,m)} := e^{t\mathcal{L}_m^{(k)}} = \prod_{i \in \otimes_{\sigma=1}^D ((R+2)\mathbb{Z} + m_\sigma)} e^{t\mathbf{X}^2_{i,i^+(k)}}$$

i.e. $S_t^{(k,m)}$ is a product semigroup.

The next step is to show that

$$\mu_{r\mathbf{G}}\left(f S_t^{(k,m)} g\right) = \mu_{r\mathbf{G}}\left(g S_t^{(k,m)} f\right), \tag{6.16}$$

for all $r > 0$, $k \in \{1, \ldots, D\}$ and $m \in \{0, ..., R + 1\}^D$. Fix $r > 0$ and let $k = 1$ and $m = (0, \ldots, 0)$ (the other cases are similar). Since g only depends on coordinates i such that $|i|_1 \leq n$, we have

$$S_t^{(1,0)} g(x) = \prod_{\substack{i \in \otimes_{\sigma=1}^D (R+2)\mathbb{Z} \\ |i|_1 \leq n+R+2}} e^{t\mathbf{X}^2_{i,i+(k)}} g(x),$$

which is a finite product. By Lemma 6.4.1, we also have that for any $i, j \in \mathbb{Z}^D$

$$\mu_{r\mathbf{G}} \left(f \mathbf{X}^2_{i,j} g \right) = \mu_{r\mathbf{G}} \left(g \mathbf{X}^2_{i,j} f \right)$$

and hence

$$\begin{aligned} \mu_{r\mathbf{G}} \left(f S_t^{(1,0)} g \right) &= \mu_{r\mathbf{G}} \left(f \prod_{\substack{i \in \otimes_{\sigma=1}^D (R+2)\mathbb{Z} \\ |i|_1 \leq n+R+2}} e^{t\mathbf{X}^2_{i,i+(k)}} g \right) \\ &= \mu_{r\mathbf{G}} \left(g \prod_{\substack{i \in \otimes_{\sigma=1}^D (R+2)\mathbb{Z} \\ |i|_1 \leq n+R+2}} e^{t\mathbf{X}^2_{i,i+(k)}} f \right) \\ &= \mu_{r\mathbf{G}} \left(g S_t^{(1,0)} f \right), \end{aligned}$$

as claimed.

To finish the proof, the idea is to use a version of the Trotter product formula, so that the semigroup we are interested in can be thought of as the limit of compositions of the product semigroups $S_t^{(k,m)}$. We will use the following version of the Trotter product formula, given in [121]:

Theorem 6.4.3. *Let $\mathcal{H}$ and $\mathcal{H}_0$ be two Hilbert spaces, and let $F_i \in Lip(\mathcal{H}, \mathcal{H})$, $G_i \in Lip(\mathcal{H}, L_{HS}(\mathcal{H}_0, \mathcal{H}))$ for $i = 1, 2, 3$. Let $(W(t))_{t \geq 0}$ be a cylindrical Wiener process in $\mathcal{H}_0$. Consider the SDEs, indexed by $i = 1, 2, 3$, given by*

$$dY_i(t) = F_i(Y_i(t))dt + G_i(Y_i(t))dW(t), \qquad Y_i(0) = x \in \mathcal{H},$$

and let $(\mathcal{P}_t^i)_{t\geq 0}$ be the corresponding semigroups on $\mathcal{UC}_b(\mathcal{H})$. Assume that

$$F_3 = F_1 + F_2, \qquad G_3G_3^* = G_1G_1^* + G_2G_2^*,$$

and that the first and second Fréchet derivatives of F_i and G_i are uniformly continuous and bounded on bounded subsets of $\mathcal{H}$. Then

$$\lim_{n\to\infty} \left(\mathcal{P}^1_{\frac{t}{n}}\mathcal{P}^2_{\frac{t}{n}}\right)^n f(x) = \mathcal{P}_t^3 f(x)$$

for all $f \in \mathbb{K}$, where $\mathbb{K}$ is the closure of $\mathcal{UC}_b^2(\mathcal{H})$ in $\mathcal{UC}_b(\mathcal{H})$, and the convergence is uniform in x on any bounded subset of $\mathcal{H}$.

To make use of this result, recall that by above the generator of our system can be decomposed as

$$\mathcal{L} = \frac{1}{2}\sum_{k=1}^{D} \sum_{m\in\{0,\ldots,R+1\}^D} \mathcal{L}_m^{(k)}$$

where, for $k \in \{1, \ldots, D\}$ and $m \in \{0, ..., R+1\}^D$, $\mathcal{L}_m^{(k)}$ is the generator of the semigroup $S_t^{(k,m)}$. By the one-to-one correspondence between SDEs and Markov generators, we see that the SDE associated with $\mathcal{L}_m^{(k)}$ is given by

$$dY(t) = \mathbf{A}_m^{(k)}Y(t)dt + \mathbf{B}_m^{(k)}(Y(t))dW(t),$$

where $\mathbf{A}_m^{(k)} : E_\alpha \to E_\alpha$ and $\mathbf{B}_m^{(k)} : E_\alpha \to L_{HS}(E_\alpha, H)$ are such that

$$\mathbf{A} = \sum_{k=1}^{D} \sum_{m\in\{0,\ldots,R+1\}^D} \mathbf{A}_m^{(k)}$$

and

$$\mathbf{B}\mathbf{B}^* = \sum_{k=1}^{D} \sum_{m\in\{0,\ldots,R+1\}^D} \mathbf{B}_m^{(k)}\left(\mathbf{B}_m^{(k)}\right)^*.$$

We can then apply Theorem 6.4.3 iteratively to get the result. Indeed, order the set

$$\{1, \ldots, D\} \times \{0, \ldots, R+1\}^D = \{\iota_1, \ldots, \iota_I\}$$

where $I = D(R+2)^D$. If $\iota_l = (k,m) \in \{1,\dots,D\} \times \{0,\dots,R+1\}^D$, write

$$\mathbf{A}_m^{(k)} \equiv \mathbf{A}_{\iota_l}, \quad \mathbf{B}_m^{(k)} \equiv \mathbf{B}_{\iota_l}, \quad \mathcal{L}_m^{(k)} \equiv \mathcal{L}_{\iota_l}, \quad S_t^{(k,m)} = S_t^{\iota_l}.$$

Then define, for $1 \le l \le I$,

$$\tilde{\mathbf{A}}_l := \sum_{j=1}^{l} \mathbf{A}_{\iota_j}$$

and $\tilde{\mathbf{B}}_l : E_\alpha \to L_{HS}(E_\alpha, H)$ to be such that

$$\tilde{\mathbf{B}}_l \tilde{\mathbf{B}}_l^* := \sum_{j=1}^{l} \mathbf{B}_{\iota_j} \mathbf{B}_{\iota_j}^*.$$

Consider the SDE

$$d\tilde{Y}_l(t) = \tilde{\mathbf{A}}_l \tilde{Y}_l(t) dt + \tilde{\mathbf{B}}_l \left(\tilde{Y}_l(t)\right) dW(t),$$

which has generator $\tilde{\mathcal{L}}_l = \sum_{j=1}^{l} \mathcal{L}_{\iota_j}$. Let $(\tilde{P}_t^l)_{t\ge0}$ be the semigroup on $\mathcal{U}C_b(E_\alpha)$ associated with $\tilde{\mathcal{L}}_l$. By a first application of Theorem 6.4.3, for all $f \in \mathbb{K}$, we have

$$\lim_{n\to\infty} \left(S_{\frac{t}{n}}^{\iota_1} S_{\frac{t}{n}}^{\iota_2}\right)^n f(x) = \tilde{P}_t^2 f(x),$$

where the convergence is uniform on bounded subsets. Moreover, by claim (6.16) above and the dominated convergence theorem, we have

$$\mu_{r\mathbf{G}}\left(f\tilde{P}_t^2 g\right) = \lim_{n\to\infty} \mu_{r\mathbf{G}}\left(f\left(S_{\frac{t}{n}}^{\iota_1} S_{\frac{t}{n}}^{\iota_2}\right)^n g\right) = \lim_{n\to\infty} \mu_{r\mathbf{G}}\left(g\left(S_{\frac{t}{n}}^{\iota_1} S_{\frac{t}{n}}^{\iota_2}\right)^n f\right) = \mu_{r\mathbf{G}}\left(g\tilde{P}_t^2 f\right) \tag{6.17}$$

for all $f, g \in \mathcal{U}C_b^2(E_\alpha)$. Similarly, for all $f \in \mathbb{K}$, we have

$$\lim_{n\to\infty} \left(\tilde{P}_{\frac{t}{n}}^2 S_{\frac{t}{n}}^{\iota_3}\right)^n f(x) = \tilde{P}_t^3 f(x),$$

where again the convergence is uniform on bounded sets, so that

$$\mu_{r\mathbf{G}}\left(f\tilde{P}_t^3 g\right) = \lim_{n\to\infty} \mu_{r\mathbf{G}}\left(f\left(\tilde{P}_{\frac{t}{n}}^2 S_{\frac{t}{n}}^{\iota_3}\right)^n g\right) = \lim_{n\to\infty} \mu_{r\mathbf{G}}\left(g\left(\tilde{P}_{\frac{t}{n}}^2 S_{\frac{t}{n}}^{\iota_3}\right)^n f\right) = \mu_{r\mathbf{G}}\left(g\tilde{P}_t^3 f\right)$$

using identities (6.16) and (6.17). Continuing in this manner, we see that $P_t = \tilde{P}_t^I$, the semigroup corresponding to the generator $\mathcal{L} = \sum_{j=1}^I \mathcal{L}_{\iota_j}$, is such that

$$\mu_{r\mathbf{G}}\left(fP_t g\right) = \mu_{r\mathbf{G}}\left(gP_t f\right)$$

for all $f, g \in \mathcal{U}C_b^2(E_\alpha)$, as required. □

Finally we can extend the above result to functions in $L^p(\mu_{r\mathbf{G}})$.

Corollary 6.4.4. *The semigroup $(P_t)_{t\geq 0}$ acting on $\mathcal{U}C_b(E_\alpha)$ can be extended to $L^p(\mu_{r\mathbf{G}})$ for any $p \geq 1$ and $r > 0$. Moreover we have*

$$\mu_{r\mathbf{G}}(fP_t g) = \mu_{r\mathbf{G}}(gP_t f)$$

for any $f, g \in L^2(\mu_{r\mathbf{G}})$ and $r > 0$.

Proof. Although the proof is standard, we recall the idea for the sake of completeness. By Corollary 6.3.3 , we have $P_t f(\cdot) = \mathbb{E}f(Y_t(\cdot))$, for $f \in \mathcal{U}C_b(E_\alpha)$ and $t \geq 0$, so that by Jensen's inequality for any $p \geq 1$ and $f \in \mathcal{U}C_b(E_\alpha)$,

$$|P_t f|^p \leq P_t|f|^p.$$

By Theorem 6.4.2 this implies that

$$\mu_{r\mathbf{G}}|P_t f|^p \leq \mu_{r\mathbf{G}}|f|^p, \tag{6.18}$$

for all $f \in \mathcal{U}C_b^2(E_\alpha), r > 0$.

The fact that P_t can be extended to $L^p(\mu_{r\mathbf{G}})$, then follows from an application of the Hahn-Banach theorem. Suppose now that $\{g_n\}_{n\geq 1}$ is a sequence of functions in $\mathcal{U}C_b^2(E_\alpha)$ converging to g in $L^2(\mu_{r\mathbf{G}})$, and let $f \in \mathcal{U}C_b^2(E)$. Then

$$|\mu_{r\mathbf{G}}(fP_t g) - \mu_{r\mathbf{G}}(gP_t f)| \leq |\mu_{r\mathbf{G}}(fP_t(g - g_n))| + |\mu_{r\mathbf{G}}(fP_t g_n - gP_t f)|,$$

so that by Hölder's inequality

$$\begin{aligned}|\mu_{r\mathbf{G}}(fP_tg) - \mu_{r\mathbf{G}}(gP_tf)| &\leq \left(\mu_{r\mathbf{G}}f^2\right)^{\frac{1}{2}} \left(\mu_{r\mathbf{G}}|P_t(g-g_n)|^2\right)^{\frac{1}{2}} \\ &\quad + \left(\mu_{r\mathbf{G}}(g-g_n)^2\right)^{\frac{1}{2}} \left(\mu_{r\mathbf{G}}|P_tf|^2\right)^{\frac{1}{2}} \\ &\leq 2\left(\mu_{r\mathbf{G}}f^2\right)^{\frac{1}{2}} \left(\mu_{r\mathbf{G}}(g-g_n)^2\right)^{\frac{1}{2}},\end{aligned}$$

using inequality (6.18) and Theorem 6.4.2. It is clear that this converges to 0 as $n \to \infty$, so that

$$\mu_{r\mathbf{G}}(fP_tg) = \mu_{r\mathbf{G}}(gP_tf),$$

for all $f \in \mathcal{U}C_b^2(E), g \in L^2(\mu_{r\mathbf{G}})$ and $r > 0$. The result follows by a similar argument, this time by taking approximations $f_n \in \mathcal{U}C_b^2(E)$ of $f \in L^2(\mu_{r\mathbf{G}})$. □

6.5 Symmetry in Sobolev spaces

In this section we show that the generator $\mathcal{L}$ is symmetric in some family of infinite dimensional Sobolev spaces. In the next section this result will be useful in the proof of the ergodicity of the semigroup generated by $\mathcal{L}$. For $r > 0$, we start by introducing the following Dirichlet operator:

$$(f, \mathbf{L}_r g)_{L^2(\mu_{r\mathbf{G}})} = -\sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}(\partial_k f, \partial_l g)_{L^2(\mu_{r\mathbf{G}})},$$

where $\mathbf{G} \equiv \mathbf{M}^{-1}$ is the covariance matrix associated to the measure $\mu_{\mathbf{G}}$, as above. By integration by parts,

$$\begin{aligned}&-\sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}(\partial_k f, \partial_l g)_{L^2(\mu_{r\mathbf{G}})} \\ &= \left(f, \sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\partial_k\partial_l g\right)_{L^2(\mu_{r\mathbf{G}})} - \left(f, r^{-1}\sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\partial_k V \partial_l g\right)_{L^2(\mu_{r\mathbf{G}})}\end{aligned}$$

so that

$$
\begin{aligned}
&-\sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}(\partial_k f, \partial_l g)_{L^2(\mu_r\mathbf{G})} \\
&= \left(f, \sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\partial_k\partial_l g\right)_{L^2(\mu_r\mathbf{G})} - \left(f, r^{-1}\sum_{l,j\in\mathbb{Z}^D}\left(\sum_{k\in\mathbb{Z}^D}\mathbf{G}_{l,k}\mathbf{M}_{k,j}x_j\right)\partial_l g\right)_{L^2(\mu_r\mathbf{G})} \\
&= \left(f, \sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\partial_k\partial_l g\right)_{L^2(\mu_r\mathbf{G})} - \left(f, r^{-1}\sum_{l\in\mathbb{Z}^D} x_l\partial_l g\right)_{L^2(\mu_r\mathbf{G})},
\end{aligned}
$$

since by definition, $\sum_{k\in\mathbb{Z}^D}\mathbf{G}_{l,k}\mathbf{M}_{k,j}x_j = x_l$ if $j = l$, and 0 otherwise. Thus, on a dense domain including $\mathcal{U}C_b^2$, we have

$$
\mathbf{L}_r g = \sum_{k,l\in\mathbb{Z}^D}\mathbf{G}_{k,l}\partial_k\partial_l g - r^{-1}\mathbf{D}g \tag{6.19}
$$

where

$$
\mathbf{D}g \equiv \sum_{l\in\mathbb{Z}^D} x_l\partial_l g. \tag{6.20}
$$

$\mathbf{D}$ will play the role of the dilation generator in our setup. We now make two important observations.

Lemma 6.5.1. *For all $i, j \in \mathbb{Z}^D$,*

$$
[\mathbf{D}, \mathbf{X}_{i,j}] = 0, \tag{6.21}
$$

and

$$
\left[\sum_{k,l\in\mathbb{Z}^D}\mathbf{G}_{k,l}\partial_k\partial_l, \mathbf{X}_{i,j}\right] = 0. \tag{6.22}
$$

Proof. We have

$$
[\mathbf{D}, \mathbf{X}_{i,j}] = \left[\sum_{k\in\mathbb{Z}^D} x_k\partial_k, \sum_{l\in\mathbb{Z}^D}\mathbf{M}_{i,l}x_l\partial_j - \sum_{l\in\mathbb{Z}^D}\mathbf{M}_{j,l}x_l\partial_i\right]
$$

so that

$$\begin{aligned}
[\mathbf{D}, \mathbf{X}_{i,j}] &= \sum_{k,l\in\mathbb{Z}^D} \mathbf{M}_{i,l}[x_k\partial_k, x_l\partial_j] - \sum_{k,l\in\mathbb{Z}^D} \mathbf{M}_{j,l}[x_k\partial_k, x_l\partial_i] \\
&= \sum_{k,l\in\mathbb{Z}^D} \mathbf{M}_{i,l}\left(x_k\delta_{k,l}\partial_j - x_l\delta_{k,j}\partial_k\right) \\
&\quad - \sum_{k,l\in\mathbb{Z}^D} \mathbf{M}_{j,l}\left(x_k\delta_{k,l}\partial_i - x_l\delta_{k,i}\partial_k\right) \\
&= \sum_{l\in\mathbb{Z}^D} \mathbf{M}_{i,l}x_l\partial_j - \sum_{l\in\mathbb{Z}^D} \mathbf{M}_{i,l}x_l\partial_j \\
&\quad - \sum_{l\in\mathbb{Z}^D} \mathbf{M}_{j,l}x_l\partial_i + \sum_{l\in\mathbb{Z}^D} \mathbf{M}_{j,l}x_l\partial_i \\
&= 0,
\end{aligned}$$

so that (6.21) holds.

For (6.22), we calculate that

$$\left[\sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\partial_k\partial_l, \mathbf{X}_{i,j}\right] = \sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\left(\partial_k[\partial_l, \mathbf{X}_{i,j}] + [\partial_k, \mathbf{X}_{i,j}]\partial_l\right),$$

where

$$\begin{aligned}
[\partial_k, \mathbf{X}_{i,j}] &= \left[\partial_k, \sum_{l\in\mathbb{Z}^D} \mathbf{M}_{i,l}x_l\partial_j - \sum_{l\in\mathbb{Z}^D} \mathbf{M}_{j,l}x_l\partial_i\right] \\
&= \mathbf{M}_{i,k}\partial_j - \mathbf{M}_{j,k}\partial_i.
\end{aligned}$$

Thus

$$\left[\sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\partial_k\partial_l, \mathbf{X}_{i,j}\right] = \sum_{k,l\in\mathbb{Z}^D} \mathbf{G}_{k,l}\left(\mathbf{M}_{i,l}\partial_k\partial_j - \mathbf{M}_{j,l}\partial_k\partial_i + \mathbf{M}_{i,k}\partial_j\partial_l - \mathbf{M}_{j,k}\partial_i\partial_l\right)$$

so that

$$
\begin{aligned}
\left[\sum_{k,l\in\mathbb{Z}^D}\mathbf{G}_{k,l}\partial_k\partial_l,\mathbf{X}_{i,j}\right] &= \sum_{k\in\mathbb{Z}^D}\left(\sum_{l\in\mathbb{Z}^D}\mathbf{G}_{k,l}\mathbf{M}_{l,i}\right)\partial_k\partial_j-\sum_{k\in\mathbb{Z}^D}\left(\sum_{l\in\mathbb{Z}^D}\mathbf{G}_{k,l}\mathbf{M}_{l,j}\right)\partial_k\partial_i\\
&\quad+\sum_{l\in\mathbb{Z}^D}\left(\sum_{k\in\mathbb{Z}^D}\mathbf{G}_{l,k}\mathbf{M}_{k,i}\right)\partial_j\partial_l-\sum_{l\in\mathbb{Z}^D}\left(\sum_{k\in\mathbb{Z}^D}\mathbf{G}_{l,k}\mathbf{M}_{k,j}\right)\partial_i\partial_l\\
&=\partial_i\partial_j-\partial_j\partial_i+\partial_j\partial_i-\partial_i\partial_j=0,
\end{aligned}
$$

again using the fact that $\mathbf{G}=\mathbf{M}^{-1}$. □

We thus arrive at the following result.

Proposition 6.5.2. *On $\mathcal{U}C_b^4$, we have*

$$[\mathbf{L}_r,\mathbf{X}_{i,j}]=0$$

for all $i,j\in\mathbb{Z}^D$ and $r>0$, so that

$$[\mathbf{L}_r,\mathcal{L}]=0$$

for all $r>0$.

Proof. For $i,j\in\mathbb{Z}^D$,

$$
\begin{aligned}
[\mathbf{L}_r,\mathbf{X}_{i,j}] &= \left[\sum_{k,l\in\mathbb{Z}^D}\mathbf{G}_{k,l}\partial_k\partial_l-r^{-1}\mathbf{D},\mathbf{X}_{i,j}\right]\\
&=0
\end{aligned}
$$

by (6.21) and (6.22) of Lemma 6.5.1. Hence

$$[\mathbf{L}_r,\mathcal{L}]=\frac{1}{4}\sum_{i\in\mathbb{Z}^D}\sum_{j:j\sim i}[\mathbf{L}_r,\mathbf{X}_{i,j}^2]=\frac{1}{4}\sum_{i\in\mathbb{Z}^D}\sum_{j:j\sim i}\left([\mathbf{L}_r,\mathbf{X}_{i,j}]\mathbf{X}_{i,j}+\mathbf{X}_{i,j}[\mathbf{L}_r,\mathbf{X}_{i,j}]\right)=0.$$

□

With this result in mind, we introduce the following family of Hilbert spaces: for $n \in \mathbb{N} \cup \{0\}$ and $r > 0$, define

$$\mathbb{X}_r^n := \left\{ f \in L^2(\mu_{r\mathbf{G}}) \cap \mathcal{D}(\mathbf{L}_r^n) : |f|^2_{\mathbb{X}_r^n} := |f|^2_{L^2(d\mu_{r\mathbf{G}})} + (f, (-\mathbf{L}_r)^n f)_{L^2(d\mu_{r\mathbf{G}})} < \infty \right\},$$

equipped with the corresponding inner product

$$(f, g)_{\mathbb{X}_r^n} = (f, g)_{L^2(\mu_{r\mathbf{G}})} + (f, (-\mathbf{L}_r)^n f)_{L^2(d\mu_{r\mathbf{G}})},$$

for $f, g \in \mathbb{X}_r^n$. Then we obtain the following fact:

Proposition 6.5.3. *For all $n \in \mathbb{N} \cup \{0\}$ and $r > 0$, on a dense set $\mathcal{D}_r^n \subset \mathbb{X}_r^n$, we have*

$$(f, \mathcal{L}g)_{\mathbb{X}_r^n} = (\mathcal{L}f, g)_{\mathbb{X}_r^n} = -\frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} (\mathbf{X}_{i,j} f, \mathbf{X}_{i,j} g)_{\mathbb{X}_r^n}.$$

Proof. Using the antisymmetry of $\mathbf{X}_{i,j}$ in $L^2(\mu_{r\mathbf{G}})$ (Lemma 6.4.1) for all $i, j \in \mathbb{Z}^D$ and the fact that $\mathbf{L}_r$ commutes with $\mathcal{L}$ by Proposition 6.5.2, we have

$$\begin{aligned}(f, \mathcal{L}g)_{\mathbb{X}_r^n} &= (f, \mathcal{L}g)_{L^2(\mu_{r\mathbf{G}})} + (-1)^n \left(f, \mathbf{L}_r^n \mathcal{L} g\right)_{L^2(\mu_{r\mathbf{G}})} \\ &= -\frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} (\mathbf{X}_{i,j} f, \mathbf{X}_{i,j} g)_{L^2(\mu_{r\mathbf{G}})} + (-1)^n \left(f, \mathcal{L} \mathbf{L}_r^n g\right)_{L^2(\mu_{r\mathbf{G}})}.\end{aligned}$$

Thus

$$\begin{aligned}(f, \mathcal{L}g)_{\mathbb{X}_r^n} &= -\frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} (\mathbf{X}_{i,j} f, \mathbf{X}_{i,j} g)_{L^2(\mu_{r\mathbf{G}})} - \frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} \left(\mathbf{X}_{i,j} f, \mathbf{X}_{i,j} (-\mathbf{L}_r)^n g\right)_{L^2(\mu_{r\mathbf{G}})} \\ &= -\frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} (\mathbf{X}_{i,j} f, \mathbf{X}_{i,j} g)_{L^2(\mu_{r\mathbf{G}})} - \frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} \left(\mathbf{X}_{i,j} f, (-\mathbf{L}_r)^n \mathbf{X}_{i,j} g\right)_{L^2(\mu_{r\mathbf{G}})} \\ &= -\frac{1}{4} \sum_{i \in \mathbb{Z}^D} \sum_{j : j \sim i} (\mathbf{X}_{i,j} f, \mathbf{X}_{i,j} g)_{\mathbb{X}_r^n}.\end{aligned}$$

$\square$

In the case $n = 1$, we have

$$(f, g)_{\mathbb{X}_r} = (f, g)_{L^2(\mu_{r\mathbf{G}})} + \sum_{i,j \in \mathbb{Z}^D} \mu_{r\mathbf{G}}(\mathbf{G}_{i,j} \partial_i f \partial_j g) = (f, g)_{L^2(\mu_{r\mathbf{G}})} + \mu_{r\mathbf{G}}(\mathbf{G}^{\frac{1}{2}} \nabla f \cdot \mathbf{G}^{\frac{1}{2}} \nabla g),$$

where for simplicity here and later we set $\mathbb{X}_r \equiv \mathbb{X}_r^1$ for all $r > 0$.

Remark 6.5.4. *By Proposition 6.5.3, the operator $-\mathcal{L}$ is closable in $\mathbb{X}_r$ for all $r > 0$ (by standard arguments — see for example Proposition 3.3 of [96]) and can be extended to a non-negative self-adjoint operator on $\mathbb{X}_r$ by taking the Friedrichs extension. We continue to denote this extension by the same symbol $\mathcal{L}$. Moreover, $\mathcal{L}$ generates a strongly continuous semigroup $e^{t\mathcal{L}} : \mathbb{X}_r \to \mathbb{X}_r$ such that $e^{t\mathcal{L}} = P_t|_{\mathbb{X}_r}$. Indeed, by the spectral theorem, the strongly continuous contraction semigroup $e^{t\mathcal{L}} : \mathbb{X}_r \to \mathbb{X}_r$ is well defined, and can be extended to $L^2(\mu_{r\mathbf{G}})$. This extension coincides with P_t on a dense set of $L^2(\mu_{r\mathbf{G}})$ (namely $\mathcal{U}C_b^2(E_\alpha)$), so that the extension must coincide with P_t on the whole of $L^2(\mu_{r\mathbf{G}})$. In view of these observations, we can think of $\{\mathbb{X}_r : r > 0\}$ as the natural family of spaces on which P_t acts.*

6.6 Ergodicity

Before we start investigating the ergodicity of the semigroup $(P_t)_{t \geq 0}$, it is useful to think about what kind of convergence to expect. One might initially hope for exponential convergence in $L^2(\mu_{r\mathbf{G}})$, i.e. the existence of a constant $\theta > 0$ such that

$$\mu_{r\mathbf{G}} \left(P_t f - \mu_{r\mathbf{G}} f\right)^2 \leq e^{-2\theta t} \mu_{r\mathbf{G}}(f - \mu_{r\mathbf{G}} f)^2 \tag{6.23}$$

for all $t \geq 0$ and $f \in L^2(\mu_{r\mathbf{G}})$. It is well known (see for example Property 2.4 of [66]), that inequality (6.23) is equivalent to the spectral gap inequality:

$$\theta \mu_{r\mathbf{G}} \left(f - \mu_{r\mathbf{G}} f\right)^2 \leq \mu_{r\mathbf{G}} \left(f(-\mathcal{L} f)\right). \tag{6.24}$$

We claim, however, that (6.24) cannot hold. To this end, suppose that we are in the situation when $\mathrm{M} = \mathbf{Id}$, so that $\mu_{r\mathbf{G}}$ is a product measure, and consider a sequence of functions of the following form:

$$f_\Lambda(x) \equiv \sum_{i\in\Lambda} x_i^2,$$

for a finite set $\Lambda \subset \mathbb{Z}^D$. Then, if we denote by μ_r the centred Gaussian measure on $\mathbb{R}$ with variance $r > 0$, we have

$$\begin{aligned}
\mu_{r\mathbf{G}}(f_\Lambda - \mu_{r\mathbf{G}} f_\Lambda)^2 &= \mu_{r\mathbf{G}} \left(\sum_{i\in\Lambda} (x_i^2 - \mu_{r\mathbf{G}} x_i^2) \right)^2 \\
&= \sum_{i\in\Lambda} \mu_{r\mathbf{G}}(x_i^2 - \mu_{r\mathbf{G}} x_i^2)^2 + 2 \sum_{\substack{i,j\in\Lambda \\ i\neq j}} \mu_{r\mathbf{G}} \left((x_i^2 - \mu_{r\mathbf{G}} x_i^2)(x_j^2 - \mu_{r\mathbf{G}} x_j^2) \right) \\
&= \sum_{i\in\Lambda} \mu_{r\mathbf{G}}(x_i^2 - \mu_{r\mathbf{G}} x_i^2)^2 + 2 \sum_{\substack{i,j\in\Lambda \\ i\neq j}} \mu_r(x_i^2 - \mu_r x_i^2)\mu_r(x_j^2 - \mu_r x_j^2) \\
&= \sum_{i\in\Lambda} \mu_r(x_i^2 - \mu_r x_i^2)^2 \\
&= |\Lambda| \mu_r (x_0^2 - \mu_0 x_0^2)^2 \equiv const \cdot |\Lambda|,
\end{aligned}$$

with $|\Lambda|$ denoting cardinality of Λ. Moreover,

$$\mu_{r\mathbf{G}} \left(f_\Lambda(-\mathcal{L} f_\Lambda) \right) = \frac{1}{4} \sum_{i\in\Lambda} \sum_{j:j\sim i} \mu_{r\mathbf{G}}(\mathbf{X}_{i,j} f_\Lambda)^2,$$

where

$$(\mathbf{X}_{i,j} f_\Lambda)^2 = \begin{cases} 0 & \text{if } \{i,j\} \subset \Lambda \text{ or } \{i,j\} \subset \Lambda^c, \\ 4x_i^2 x_j^2 & \text{otherwise.} \end{cases}$$

Therefore

$$\mu_{r\mathbf{G}} \left(f_\Lambda(-\mathcal{L} f_\Lambda) \right) = 2 \sum_{\substack{\{i,j\}\subset\Lambda \\ i\in\Lambda, j\in\Lambda^c, j\sim i}} \mu_r(x_0^2)^2 = const \cdot |\partial\Lambda|$$

i.e. $\mu_{r\mathbf{G}} \left(f_\Lambda(-\mathcal{L} f_\Lambda) \right)$ depends only on the size of the boundary of the set Λ. Hence

$$\frac{\mu_{r\mathbf{G}} \left(f_\Lambda(-\mathcal{L} f_\Lambda) \right)}{\mu_{r\mathbf{G}} |f_\Lambda - \mu_{r\mathbf{G}} f_\Lambda|^2} \sim \frac{|\partial\Lambda|}{|\Lambda|},$$

which converges to 0 for a suitable sequence of sets Λ invading the lattice. This clearly prohibits the existence of a constant $\theta > 0$ such that (6.24) holds.

The above considerations show that we cannot hope for exponential decay to equilibrium of our semigroup acting on the natural space $\mathbb{X}_r$. However, in the remainder of this section we develop a strategy to show that our semigroup acting on $\mathbb{X}_r$ is still ergodic, for simplicity working in the set-up when the matrix $\mathbf{M}$ is given by $\mathbf{M} = b\mathbf{Id}$ with $b \in (0, \infty)$. Our estimates are optimal in the sense that the rate of decay we give is polynomial.

For $r > 0$, define

$$\mathcal{A}_r(f) \equiv \left(\sum_{i \in \mathbb{Z}^D} \mu_{r\mathbf{G}} |\partial_i f|^2 \right)^{1/2} \tag{6.25}$$

and

$$\mathcal{B}_r(f) \equiv \left(\sum_{i \in \mathbb{Z}^D} \left(\mu_{r\mathbf{G}} |\partial_i f|^2 \right)^{\frac{1}{2}} \right). \tag{6.26}$$

Lemma 6.6.1. *There exists a constant κ, independent of the dimension D, such that for any $r > 0$, $f \in \mathbb{X}_r$, $i \in \mathbb{Z}^D$ and $t > 0$,*

$$\mu_{r\mathbf{G}} |\partial_i (P_t f)|^2 \leq \frac{\kappa^D}{t^{\frac{D}{2}}} \mathcal{A}_r^2(f). \tag{6.27}$$

Proof. Fix $r > 0$. It is enough to show (6.27) for $f \in \mathcal{U}C_b^4(E_\alpha)$. Indeed, $\mathcal{U}C_b^4(E_\alpha)$ is dense in $\mathbb{X}_r$ and $(P_t)_{t \geq 0}$ is a contraction on $\mathbb{X}_r$. Denote $f_t = P_t f$ for $t \geq 0$. For $i \in \mathbb{Z}^D$, we have

$$|\partial_i f_t|^2 - P_t |\partial_i f|^2 = \int_0^t \frac{d}{ds} P_{t-s} |\partial_i f_s|^2 ds$$

so that

$$\begin{aligned} |\partial_i f_t|^2 - P_t |\partial_i f|^2 &= \int_0^t P_{t-s} (-\mathcal{L}(|\partial_i f_s|^2) + 2\partial_i f_s \mathcal{L} \partial_i f_s + 2\partial_i f_s [\partial_i, \mathcal{L}] f_s) ds \\ &= \int_0^t P_{t-s} \left(-\sum_{k=1}^D \sum_{j \in \mathbb{Z}^D} |\mathbf{X}_{j,j+(k)} (\partial_i f_s)|^2 + 2\partial_i f_s [\partial_i, \mathcal{L}] f_s \right) ds. \end{aligned} \tag{6.28}$$

For $i, j \in \mathbb{Z}^D$ and $k \in \{1, \ldots, D\}$,

$$\begin{aligned}[\partial_i, \mathbf{X}_{j,j^+(k)}] &= b\left[\partial_i, x_j\partial_{j^+(k)} - x_{j^+(k)}\partial_j\right] \\ &= b\left(\delta_{i,j}\partial_{i^+(k)} - \delta_{i^-(k),j}\partial_{i^-(k)}\right),\end{aligned}$$

so that

$$\begin{aligned}[\partial_i, \mathcal{L}] &= \frac{1}{2}\sum_{k=1}^{D}\sum_{j\in\mathbb{Z}^D}[\partial_i, \mathbf{X}^2_{j,j^+(k)}] \\ &= \frac{1}{2}\sum_{k=1}^{D}\sum_{j\in\mathbb{Z}^D}\left([\partial_i, \mathbf{X}_{j,j^+(k)}]\mathbf{X}_{j,j^+(k)} + \mathbf{X}_{j,j^+(k)}[\partial_i, \mathbf{X}_{j,j^+(k)}]\right) \\ &= \sum_{k=1}^{D}\sum_{j\in\mathbb{Z}^D}\left(\mathbf{X}_{j,j^+(k)}[\partial_i, \mathbf{X}_{j,j^+(k)}] + \frac{1}{2}\left[[\partial_i, \mathbf{X}_{j,j^+(k)}], \mathbf{X}_{j,j^+(k)}\right]\right) \\ &= \sum_{k=1}^{D}\left(b\mathbf{X}_{i,i^+(k)}\partial_{i^+(k)} - b\mathbf{X}_{i^-(k),i}\partial_{i^-(k)}\right) \\ &\quad + \frac{b}{2}\sum_{k=1}^{D}\left([\partial_{i^+(k)}, \mathbf{X}_{i,i^+(k)}] + [\partial_{i^-(k)}, \mathbf{X}_{i^-(k),i}]\right) \\ &= b\sum_{k=1}^{D}\left(\mathbf{X}_{i,i^+(k)}\partial_{i^+(k)} + \mathbf{X}_{i,i^-(k)}\partial_{i^-(k)} - b\partial_i\right).\end{aligned}$$

Using this in (6.28), yields

$$\begin{aligned}|\partial_i f_t|^2 - P_t|\partial_i f|^2 = \int_0^t P_{t-s}\Big(&-\sum_{k=1}^{D}\sum_{j\in\mathbb{Z}^D}|\mathbf{X}_{j,j^+(k)}(\partial_i f_s)|^2 \\ &+ 2b\partial_i f_s\sum_{k=1}^{D}(-b\partial_i f_s + \mathbf{X}_{i,i^-(k)}\partial_{i^-(k)}f_s + \mathbf{X}_{i,i^+(k)}\partial_{i^+(k)}f_s)\Big)ds.\end{aligned} \tag{6.29}$$

Integrating (6.29) with respect to the invariant measure $\mu_{r\mathbf{G}}$ then gives

$$\begin{aligned}\mu_{r\mathbf{G}}|\partial_i f_t|^2 - \mu_{r\mathbf{G}}|\partial_i f|^2 = \int_0^t \Big(&- \sum_{k=1}^{D} \sum_{j\in\mathbb{Z}^D} \mu_{r\mathbf{G}}|\mathbf{X}_{j,j^+(k)}(\partial_i f_s)|^2 \\ &- 2Db^2\mu_{r\mathbf{G}}|\partial_i f_s|^2 + 2b\sum_{k=1}^{D} \mu_{r\mathbf{G}}(\partial_i f_s \mathbf{X}_{i,i^-(k)}\partial_{i^-(k)} f_s) \\ &+ 2b\sum_{k=1}^{D} \mu_{r\mathbf{G}}(\partial_i f_s \mathbf{X}_{i,i^+(k)}\partial_{i^+(k)} f_s)\Big) ds. \end{aligned} \tag{6.30}$$

Recall that the fields $\mathbf{X}_{i,j}$, $i,j\in\mathbb{Z}^D$, are anti-symmetric in $L^2(\mu_{r\mathbf{G}})$ (Lemma 6.4.1). Therefore

$$\begin{aligned}\mu_{r\mathbf{G}}|\partial_i f_t|^2 \;&-\; \mu_{r\mathbf{G}}|\partial_i f|^2 = \int_0^t \Big(- \sum_{k=1}^{D} \sum_{j\in\mathbb{Z}^D} \mu_{r\mathbf{G}}|\mathbf{X}_{j,j^+(k)}(\partial_i f_s)|^2 \\ &-\; 2Db^2\mu_{r\mathbf{G}}|\partial_i f_s|^2 - 2b\sum_{k=1}^{D} \mu_{r\mathbf{G}}(\partial_{i^-(k)} f_s \mathbf{X}_{i,i^-(k)}\partial_i f_s) \\ &-\; 2b\sum_{k=1}^{D} \mu_{r\mathbf{G}}(\partial_{i^+(k)} f_s \mathbf{X}_{i,i^+(k)}\partial_i f_s)\Big) ds. \end{aligned} \tag{6.31}$$

Hence, using the elementary fact that $xy \le \frac{b}{2}x^2 + \frac{1}{2b}y^2$ for all $x,y\in\mathbb{R}$, we see that

$$\begin{aligned}\mu_{r\mathbf{G}}|\partial_i f_t|^2 \;&-\; \mu_{r\mathbf{G}}|\partial_i f|^2 \le \int_0^t \Big(- \sum_{k=1}^{D} \sum_{j\in\mathbb{Z}^D} \mu_{r\mathbf{G}}|\mathbf{X}_{j,j^+(k)}(\partial_i f_s)|^2 \\ &-\; 2Db^2\mu_{r\mathbf{G}}|\partial_i f_s|^2 + \sum_{k=1}^{D} b^2\mu_{r\mathbf{G}}|\partial_{i^-(k)} f_s|^2 + \sum_{k=1}^{D} \mu_{r\mathbf{G}}|\mathbf{X}_{i,i^-(k)}\partial_i f_s|^2 \\ &+\; \sum_{k=1}^{D} b^2\mu_{r\mathbf{G}}|\partial_{i^+(k)} f_s|^2 + \sum_{k=1}^{D} \mu_{r\mathbf{G}}|\mathbf{X}_{i,i^+(k)}\partial_i f_s|^2\Big) ds \\ &\le\; \int_0^t b^2 \sum_{k=1}^{D} \Big(\mu_{r\mathbf{G}}|\partial_{i^-(k)} f_s|^2 + \mu_{r\mathbf{G}}|\partial_{i^+(k)} f_s|^2 - 2\mu_{r\mathbf{G}}|\partial_i f_s|^2\Big) ds. \end{aligned} \tag{6.32}$$

Let $\triangle$ denote the Laplacian on the lattice $\mathbb{Z}^D$, that is for functions g on $\mathbb{Z}^D$,

$$\triangle g(i) := \sum_{k=1}^{D} (g(i^+(k)) + g(i^-(k)) - 2g(i)).$$

We recognise that the right-hand side of (6.32) has exactly this form. Indeed, if we set $F(i,t) = \mu_{r\mathbf{G}} |\partial_i (P_t f)|^2$ for $t \geq 0$ and $i \in \mathbb{Z}^D$, then (6.32) yields

$$\partial_t F(i,t) \leq b^2 \triangle F(i,t), \qquad t \in [0,\infty), i \in \mathbb{Z}^D. \tag{6.33}$$

Set $W(i,t) = b^2 \triangle F(i,t) - \partial_t F(i,t)$, so that $W \geq 0$ by (6.33). Then

$$\partial_t F(i,t) = b^2 \triangle F(i,t) - W(i,t), \qquad t \in [0,\infty), i \in \mathbb{Z}^D. \tag{6.34}$$

To solve this equation, we first solve the homogeneous heat equation on the lattice:

$$\begin{aligned} \partial_t u(i,t) &= b^2 \triangle u(i,t), \qquad t \in [0,\infty), i \in \mathbb{Z}^D, \\ u(i,0) &= u_0. \end{aligned} \tag{6.35}$$

This can be done using Fourier transform. Indeed, it is easily seen that the solution to (6.35) is given by

$$u(i,t) = p_t * u_0(i) := \sum_{l \in \mathbb{Z}^D} p_t(i-l) u_0(l),$$

for $t \geq 0, i \in \mathbb{Z}^D$, where

$$p_t(l) \equiv \frac{1}{(2\pi)^D} \int_{[-\pi,\pi]^D} e^{-2tb^2(1-\cos x)} \cos(l \cdot x) dx, \tag{6.36}$$

for $l \in \mathbb{Z}^D$, where $l \cdot x = \sum_{j=1}^{D} l_j x_j$ for $x \in \mathbb{R}^D$. The heat kernel p_t can be recognised as a multidimensional modified Bessel function of the first kind, so that p_t is positive (see [128]). We can now solve (6.34) using Duhamel's principle, to see that

$$F(i,t) = p_t * F(i,0) - \int_0^t p_{t-s} * W(i,s) ds, \tag{6.37}$$

for $t \geq 0$ and $i \in \mathbb{Z}^D$. Since the both the heat kernel p_t and W are positive, this then yields

$$F(i,t) \leq p_t * F(i,0) = \sum_{l \in \mathbb{Z}^D} p_t(i-l)F(l,0),$$

or equivalently

$$\mu_{r\mathbf{G}}|\partial_i(P_t f)|^2 \leq \sum_{l \in \mathbb{Z}^D} p_t(i-l)\mu_{r\mathbf{G}}|\partial_l f|^2$$

for all $t \geq 0, i \in \mathbb{Z}^D$.

To complete the proof, it remains to show that there exists a constant $\kappa \in (0, \infty)$ such that

$$p_t(l) \leq \frac{\kappa^D}{t^{\frac{D}{2}}}, \tag{6.38}$$

for all $l \in \mathbb{Z}^D$. To see this, note that

$$\begin{aligned} p_t(l) &= \frac{1}{(2\pi)^D} \int_{[-\pi,\pi]^D} e^{-2tb^2(1-\cos x)} \cos(l \cdot x)dx \\ &\leq \frac{1}{(2\pi)^D} \left(\int_{-\pi}^{\pi} e^{-2tb^2(1-\cos x)} dx \right)^D. \end{aligned} \tag{6.39}$$

Now, for small $\delta > 0$,

$$\begin{aligned} \int_{-\pi}^{\pi} e^{-2t(1-\cos x)}dx &= 2\int_0^\delta e^{-2t(1-\cos x)}dx + 2\int_\delta^\pi e^{-2t(1-\cos x)}dx \\ &\leq 2\int_0^\delta e^{-2t(1-\cos x)}dx + 2(\pi-\delta)e^{-2t(1-\cos\delta)}. \end{aligned}$$

Moreover, for $x \in (0, \delta)$, we have $\cos x \leq 1 - \frac{x^2}{2}$, so that

$$\int_{-\pi}^{\pi} e^{-2t(1-\cos x)}dx \leq 2\int_0^\delta e^{-tx^2}dx + 2(\pi-\delta)e^{-2t(1-\cos\delta)}.$$

The the right-hand side can be seen to be bounded above by $\kappa t^{-\frac{1}{2}}$ for a constant κ. Using this in (6.39) yields (6.38). $\square$

Remark 6.6.2. *We remark that the convergence in Lemma 6.6.1 cannot be improved, in the sense that the stated rate of convergence is attained. Indeed, we can calculate that for any*

$i \in \mathbb{Z}^D$,

$$\mathcal{L}x_i^2 = b^2 \sum_{k=1}^{D} (x_{i+(k)}^2 + x_{i-(k)}^2 - 2x_i^2)$$

i.e. for $f(x) = x_i^2$, $\mathcal{L}f = b^2 \triangle f$, where as above $\triangle$ is the discrete Laplacian on $\mathbb{Z}^D$. Thus, $P_t f = e^{tb^2\triangle} f$ and using the Fourier representation of the kernel of $e^{tb^2\triangle}$ given in (6.36)

$$e^{tb^2\triangle} f \sim \frac{1}{t^{\frac{D}{2}}},$$

for large t.

Corollary 6.6.3. *For all $r > 0$ and $f \in \mathbb{X}_r$ such that $\mathcal{B}_r(f) < \infty$, we have*

$$\sum_{i \in \mathbb{Z}^D} \mu_{r\mathbf{G}} |\partial_i (P_t f)|^2 \leq \frac{\kappa^{\frac{D}{2}}}{t^{\frac{D}{4}}} \mathcal{A}_r(f) \mathcal{B}_r(f), \tag{6.40}$$

where $\mathcal{A}_r(f)$ and $\mathcal{B}_r(f)$ are given by (6.25) *and* (6.26) *respectively, and κ is the constant that appears in Lemma 6.6.1. Furthermore, there exists a constant $c \in (0, \infty)$ such that*

$$\mu_{r\mathbf{G}} \left((P_t f)^2 \log \frac{(P_t f)^2}{\mu_{r\mathbf{G}}(P_t f)^2} \right) \leq c \frac{\kappa^{\frac{D}{2}}}{t^{\frac{D}{4}}} \mathcal{A}_r(f) \mathcal{B}_r(f), \tag{6.41}$$

i.e. we have convergence of our semigroup in entropy with polynomial rate of convergence. In particular

$$\mu_{r\mathbf{G}} (P_t f - \mu_{r\mathbf{G}}(f))^2 \leq c \frac{\kappa^{\frac{D}{2}}}{t^{\frac{D}{4}}} \mathcal{A}_r(f) \mathcal{B}_r(f). \tag{6.42}$$

Proof. By Proposition 6.5.3, P_t is symmetric in $\mathbb{X}_r$. Therefore we can write

$$\begin{aligned}
\sum_{i \in \mathbb{Z}^D} \mu_{r\mathbf{G}} |\partial_i (P_t f)|^2 &= \sum_{i \in \mathbb{Z}^D} \mu_{r\mathbf{G}} (\partial_i f \partial_i P_{2t} f) \\
&\leq \sum_{i \in \mathbb{Z}^D} \left(\mu_{r\mathbf{G}} |\partial_i f|^2 \right)^{\frac{1}{2}} \left(\mu_{r\mathbf{G}} |\partial_i P_{2t} f|^2 \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{i \in \mathbb{Z}^D} \left(\mu_{r\mathbf{G}} |\partial_i f|^2 \right)^{\frac{1}{2}} \right) \sup_{j \in \mathbb{Z}^D} \left(\mu_{r\mathbf{G}} |\partial_j P_{2t} f|^2 \right)^{1/2}. \qquad (6.43)
\end{aligned}$$

Combining (6.43) with Lemma 6.6.1 we immediately arrive at (6.40). Now inequalities

(6.41) and (6.42) follow from the logarithmic Sobolev and Poincaré inequalities for the product Gaussian measure $\mu_{r\mathbf{G}}$. $\square$

The following result shows that the class of functions for which the system is ergodic is larger than the one considered in Corollary 6.6.3.

Proposition 6.6.4. *The semigroup $(P_t)_{t\geq 0}$ is ergodic in the Orlicz space $L_\Psi(\mu_{r\mathbf{G}})$ [5], with $\Psi(s) \equiv s^2\log(1+s^2)$, in the sense that*

$$\|P_t f - \mu_{r\mathbf{G}} f\|_{L_\Psi(\mu_{r\mathbf{G}})} \to 0$$

as $t \to \infty$, for any $f \in L_\Psi(\mu_{r\mathbf{G}})$ and $r > 0$.

Proof. For $f \in \mathbb{X}_r \cap \left\{ f \in L_\Psi(\mu_{r\mathbf{G}}) : \sum_{i\in\mathbb{Z}^D} \left(\mu_{r\mathbf{G}}|\partial_i f|^2\right)^{\frac{1}{2}} < \infty \right\}$ the result follows from Corollary 6.6.3 and the fact that the logarithmic Sobolev inequality is equivalent to the existence of a constant c such that

$$\|f - \mu_{r\mathbf{G}} f\|^2_{L_\Psi(\mu_{r\mathbf{G}})} \leq c\mu_{r\mathbf{G}}(f(-\mathcal{L}f)),$$

by Proposition 3.1 of [32]. Now it is enough to notice that such a set of functions is dense in $L_\Psi(\mu_{r\mathbf{G}})$. $\square$

6.7 Liggett-Nash-type inequalities

In this final section we will show how to deduce Liggett-Nash type inequalities from the results of the previous section. For $r > 0$, let $\mathcal{A}_r$ and $\mathcal{B}_r$ be defined by (6.25) and (6.26) respectively.

Theorem 6.7.1. *There exist constants $k_1, k_2 \in (0,\infty)$ such that for all $r > 0$ and $f \in \mathbb{X}_r \cap \mathcal{D}(\mathcal{L})$ with $\mathcal{B}_r(f) < \infty$,*

$$\mu_{r\mathbf{G}}(f - \mu_{r\mathbf{G}}(f))^2 \leq k_1 \left(-\mathcal{L}f, f\right)^{\frac{D}{D+4}}_{L^2(\mu_{r\mathbf{G}})} \left(\mathcal{A}_r(f)\mathcal{B}_r(f)\right)^{\frac{4}{D+4}}, \tag{6.44}$$

[5] Recall that the Orlicz space $L_\Psi(\mu_{r\mathbf{G}})$ is defined to be the space of measurable functions f such that $\mu_{r\mathbf{G}}(\Psi(f)) < \infty$, equipped with the norm $\|f\|_{L_\Psi(\mu_{r\mathbf{G}})} := \sup\{\mu_{r\mathbf{G}}|fg| : \mu_{r\mathbf{G}}(\Psi(g)) \leq 1\}$.

and

$$[\mathcal{A}_r(f)]^{2+\frac{4}{D}} \leq k_2 \left[\sum_{i\in\mathbb{Z}^D} \int \partial_i f \partial_i(-\mathcal{L}f) d\mu_{r\mathbf{G}} \right] \mathcal{B}_r(f)^{\frac{4}{D}}. \tag{6.45}$$

Remark 6.7.2. *Note that inequality* (6.45) *can be considered as an analogue of the Nash inequality in* $\mathbb{R}^n$ *(which first appeared in [103]). Indeed, on* $\mathbb{R}^n$*, the Nash inequality states that*

$$|u|_{L^2(\mathbb{R}^n)}^{2+\frac{4}{n}} \leq k(-\Delta u, u)_{L^2(\mathbb{R}^n)} |u|_{L^1(\mathbb{R}^n)}^{\frac{4}{n}}, \quad u \in L^1(\mathbb{R}^n) \cap W^{1,2}(\mathbb{R}^n),$$

for some constant $k > 0$*, and where* Δ *is the standard Laplacian on* $\mathbb{R}^n$*. The main difference with our situation is that the natural space for our operator* $\mathcal{L}$ *is* $\mathbb{X}_r$ *instead of* L^2.

Proof. We follow the method of T. Liggett, described in [92]. As usual, set $f_t = P_t f$. For f such that $\mu_{r\mathbf{G}} f = 0$ we have

$$\begin{aligned} \int f f_t d\mu_{r\mathbf{G}} &\leq \left(\mu_{r\mathbf{G}}(f^2)\right)^{\frac{1}{2}} \left(\mu_{r\mathbf{G}}(f_t^2)\right)^{\frac{1}{2}} \\ &\leq \frac{1}{t^{\frac{D}{8}}} \left(c\kappa^{\frac{D}{2}} \mu_{r\mathbf{G}}(f^2) \mathcal{A}_r(f) \mathcal{B}_r(f) \right)^{\frac{1}{2}}, \end{aligned} \tag{6.46}$$

where we have used inequality (6.42) of Corollary 6.6.3 . Moreover, since $\mathcal{L}$ is symmetric in $L^2(\mu_{r\mathbf{G}})$,

$$\begin{aligned} \frac{d}{ds} \int f \mathcal{L} f_s d\mu_{r\mathbf{G}} &= \frac{d}{ds} \int f_s \mathcal{L} f d\mu_{r\mathbf{G}} \\ &= \int P_s \mathcal{L} f \mathcal{L} f d\mu_{r\mathbf{G}} = \int \left(P_{\frac{s}{2}} \mathcal{L} f\right)^2 d\mu_{r\mathbf{G}} \geq 0. \end{aligned}$$

Thus

$$\begin{aligned} \int f f_t d\mu_{r\mathbf{G}} &= \int f^2 d\mu_{r\mathbf{G}} + \int_0^t \int f \mathcal{L} f_s d\mu_{r\mathbf{G}} ds \\ &\geq \mu_{r\mathbf{G}}(f^2) + t \int f \mathcal{L} f d\mu_{r\mathbf{G}}. \end{aligned}$$

Using this in (6.46), we see that

$$\int f^2 d\mu_{r\mathbf{G}} \leq t \int f(-\mathcal{L}f) d\mu_{r\mathbf{G}} + \frac{1}{t^{\frac{D}{8}}} \left(c\kappa^{\frac{D}{2}} \mu_{r\mathbf{G}}(f^2) \mathcal{A}_r(f) \mathcal{B}_r(f) \right)^{\frac{1}{2}} \tag{6.47}$$

for all $t \geq 0$. We can then optimise the right-hand side over t. Indeed, taking t such that

$$t^{\frac{D+8}{D}} = \frac{D}{8} \times \frac{\left(c\kappa^{\frac{D}{2}} \mu_{r\mathbf{G}}(f^2) \mathcal{A}_r(f) \mathcal{B}_r(f) \right)^{\frac{1}{2}}}{\int f(-\mathcal{L}f) d\mu_{r\mathbf{G}}}$$

yields

$$\begin{aligned} &\mu_{r\mathbf{G}}(f^2)^{1-\frac{4}{D+8}} \\ &\quad \leq c^{\frac{4}{D+8}} \kappa^{\frac{2D}{D+8}} \left(1 + \frac{8}{D}\right) \left(\frac{D}{8}\right)^{\frac{8}{D+8}} \left(\int f(-\mathcal{L}f) d\mu_{r\mathbf{G}} \right)^{\frac{D}{D+8}} \mathcal{A}_r^{\frac{4}{D+8}}(f) \mathcal{B}_r^{\frac{4}{D+8}}(f). \end{aligned}$$

Raising both sides to the power $\frac{D+8}{D+4}$ then gives

$$\mu_{r\mathbf{G}}(f^2) \leq k_1 \left(\int f(-\mathcal{L}f) d\mu_{r\mathbf{G}} \right)^{\frac{D}{D+4}} \mathcal{A}_r^{\frac{4}{D+4}}(f) \mathcal{B}_r^{\frac{4}{D+4}}(f),$$

where $k_1 = c^{\frac{4}{D+4}} \kappa^{\frac{2D}{D+4}} \left(1 + \frac{8}{D}\right)^{\frac{D+8}{D+4}} \left(\frac{D}{8}\right)^{\frac{8}{D+4}}$. Replacing f by $f - \mu_{r\mathbf{G}} f$ then yields (6.44).

The proof of (6.45) is very similar. Indeed, note that by inequality (6.40) of Corollary 6.6.3,

$$\begin{aligned} \sum_{i \in \mathbb{Z}^D} \int \partial_i f \partial_i f_t d\mu_{r\mathbf{G}} &\leq \left(\sum_{i \in \mathbb{Z}^D} \mu_{r\mathbf{G}} |\partial_i f|^2 \right)^{\frac{1}{2}} \left(\sum_{i \in \mathbb{Z}^D} \mu_{r\mathbf{G}} |\partial_i f_t|^2 \right)^{\frac{1}{2}} \\ &\leq \frac{\kappa^{\frac{D}{4}}}{t^{\frac{D}{8}}} \mathcal{A}_r^{\frac{3}{2}}(f) \mathcal{B}_r^{\frac{1}{2}}(f) \end{aligned} \tag{6.48}$$

for all $t \geq 0$. Then, in a similar way to the above, but using the fact that $\mathcal{L}$ is symmetric in $\mathbb{X}_r$ this time, we have

$$\frac{d}{ds} \sum_{i \in \mathbb{Z}^D} \int \partial_i f \partial_i (\mathcal{L} f_s) d\mu_{r\mathbf{G}} \geq 0,$$

so that

$$\begin{aligned}\sum_{i\in\mathbb{Z}^D}\int \partial_i f\partial_i f_t d\mu_{r\mathbf{G}} &= \sum_{i\in\mathbb{Z}^D}\mu_{r\mathbf{G}}|\partial_i f|^2+\int_0^t\sum_{i\in\mathbb{Z}^D}\int \partial_i f\partial_i(\mathcal{L}f_s)d\mu_{r\mathbf{G}}ds\\ &\geq \sum_{i\in\mathbb{Z}^D}\mu_{r\mathbf{G}}|\partial_i f|^2 - t\sum_{i\in\mathbb{Z}^D}\int \partial_i(-\mathcal{L}f)\partial_i f d\mu_{r\mathbf{G}} \qquad (6.49)\end{aligned}$$

for all $t\geq 0$. Using this in (6.48), we obtain

$$\mathcal{A}_r^2(f)\leq t\sum_{i\in\mathbb{Z}^D}\int \partial_i(-\mathcal{L}f)\partial_i f d\mu_{r\mathbf{G}}+\frac{\kappa^{\frac{D}{4}}}{t^{\frac{D}{8}}}\mathcal{A}_r^{\frac{3}{2}}(f)\mathcal{B}_r^{\frac{1}{2}}(f) \qquad (6.50)$$

for all $t\geq 0$. The right-hand side of (6.50) is minimized when

$$t^{\frac{D+8}{8}}=\frac{D}{8}\times\frac{\kappa^{\frac{D}{4}}\mathcal{A}_r^{\frac{3}{2}}(f)\mathcal{B}_r^{\frac{1}{2}}(f)}{\sum_{i\in\mathbb{Z}^D}\int\partial_i(-\mathcal{L}f)\partial_i f d\mu_{r\mathbf{G}}}.$$

For this particular t, we have

$$\mathcal{A}_r^{\frac{2D+4}{D+8}}(f)\leq\left(1+\frac{8}{D}\right)\left(\frac{D}{8}\right)^{\frac{8}{D+8}}\kappa^{\frac{2D}{D+8}}\left(\sum_{i\in\mathbb{Z}^D}\int\partial_i(-\mathcal{L}f)\partial_i f d\mu_{r\mathbf{G}}\right)^{\frac{D}{D+8}}\mathcal{B}_r^{\frac{4}{D+8}}(f).$$

Raising both sides to the power $\frac{D+8}{D}$ then yields

$$\mathcal{A}_r^{2+\frac{4}{D}}(f)\leq k_2\left(\sum_{i\in\mathbb{Z}^D}\int\partial_i(-\mathcal{L}f)\partial_i f d\mu_{r\mathbf{G}}\right)\mathcal{B}_r^{\frac{4}{D}}(f),$$

with $k_2=\left(1+\frac{8}{D}\right)^{\frac{D+8}{D}}\left(\frac{D}{8}\right)^{\frac{8}{D}}\kappa^2$, as claimed.

$\square$

Appendix A

Stochastic Equations in Infinite Dimensions

This appendix very briefly summarises some of the results and ideas from the theory of stochastic equations in infinite dimensions used in Chapter 6. All the material is contained in [109], which should be referred to for the details.

A.1 Gaussian measures on Hilbert spaces

Let U be a separable Hilbert space, with inner product $\langle \cdot, \cdot \rangle$. A probability measure μ on $(U, \mathcal{B}(U))$ is called *Gaussian* if for arbitrary $h \in U$ there exist $m \in \mathbb{R}$ and $q \geq 0$ such that

$$\mu\{x \in U : \langle h, x \rangle \in A\} = \mathcal{N}(m, q)(A)$$

for all $A \in \mathcal{B}(\mathbb{R})$, where $\mathcal{N}(m, q)$ is the standard Gaussian measure on $\mathbb{R}$ with mean m and variance q. If μ is Gaussian, there exist $m \in U$ and a symmetric non-negative continuous linear operator Q such that:

$$\int \langle h, x \rangle \mu(dx) = \langle m, h \rangle, \qquad \forall h \in U,$$

$$\int \langle h_1, x \rangle \langle h_2, x \rangle \mu(dx) - \langle h_1, m \rangle \langle h_2, m \rangle = \langle Q h_1, h_2 \rangle, \qquad \forall h_1, h_2 \in U.$$

The element m is called the *mean* of μ, and Q the *covariance operator*.

A.2 Stochastic processes on Hilbert spaces

Let $(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, \mathbb{P})$ be a filtered probability space and U a separable Hilbert space. A family $(X(t))_{t\geq 0}$ of U-valued random variables such that $X(t)$ is $\mathcal{F}_t$-measurable is called an adapted stochastic process on U. $(X(t))_{t\geq 0}$ is square-integrable if $\mathbb{E}|X(t)|^2 < \infty$, and is a *martingale* if $\mathbb{E}(X(t)|\mathcal{F}_s) = X(s)$ $\mathbb{P}$-a.s. for arbitrary $t \geq s$.

For a U-valued square-integrable martingale $(X(t))_{t\geq 0}$, the *quadratic variation process* of $(X(t))_{t\geq 0}$ is the unique increasing, adapted, continuous process $([X(\cdot)]_t)_{t\geq 0}$ taking values in the space of trace-class[1] operators on U, such that

$$X(t) \otimes X(t) - [X(\cdot)]_t$$

is an $\mathcal{F}_t$-martingale and $[X(\cdot)]_0 = 0$. The *cross quadratic variation* $([X_1(\cdot), X_2(\cdot)]_t)_{t\geq 0}$ of two such processes is then given by

$$[X_1(\cdot), X_2(\cdot)]_t = \frac{1}{4}\left([(X_1 + X_2)(\cdot)]_t - [(X_1 - X_2)(\cdot)]_t\right).$$

A stochastic process $(X(t))_{t\geq 0}$ taking values in U is Gaussian if for all $t_1, \ldots, t_n$, $(X(t_1), \ldots, X(t_n))$ is a Gaussian random variable in U^n.

A.3 Wiener processes in Hilbert spaces

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and U a separable Hilbert space. Let $Q : U \to U$ be a bounded linear operator which is non-negative and such that $\mathbf{Tr}\, Q < \infty$.

Definition A.3.1. *A U-valued stochastic process $W = (W(t))_{t\geq 0}$ is called a Q-Wiener process if*

(i) $W(0) = 0$*;*

[1] An operator T on U is of trace class if $\mathbf{Tr}\, T = \sum_{k=1}^{\infty} \langle Te_k, e_k\rangle < \infty$ for some orthonormal basis $\{e_k\}$ of U.

(ii) *W has continuous trajectories;*

(iii) *W has independent increments;*

(iv) $\mathcal{L}(W(t)-W(s)) = \mathcal{N}(0,(t-s)Q)$ *for all* $t \geq s \geq 0$.

Let W be a Q-Wiener process with $\mathbf{Tr}\, Q < \infty$. Then W is a Gaussian process on U, $\mathbb{E}(W(t)) = 0$ and $[W(\cdot)]_t = tQ$. Moreover, if $\{e_k\}$ is a complete orthonormal system in U and $\{\gamma_k\}$ is a sequence of non-negative numbers such that $Qe_k = \gamma_k e_k$ for $k = 1, 2, \ldots,$ then

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\gamma_k}\beta_k(t)e_k$$

where

$$\beta_k(t) = \frac{1}{\sqrt{\gamma_k}}\langle W(t), e_k\rangle$$

are real-valued mutually independent Brownian motions, and the series is convergent in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. A square-integrable martingale $(X(t))_{t\geq 0}$ such that $X(0) = 0$ is a Q-Wiener process if and only if $[X(\cdot)]_t = tQ$ for all $t \geq 0$.

Now suppose that $Q : U \to U$ is still a non-negative bounded linear operator, but not necessarily of trace-class. Let $U_0 = Q^{\frac{1}{2}}(U)$ with the induced norm $\|\cdot\|_0 = \|Q^{-1/2}(\cdot)\|$. Let U_1 be an arbitrary Hilbert space and $J : U_0 \to U_1$ a Hilbert-Schmidt embedding. Let $\{g_k\}$ be an orthonormal basis for U_0, and $\{\beta_k\}$ a family of independent real-valued standard Wiener processes. Then the formula

$$W(t) = \sum_{k=1}^{\infty} g_k\beta_k(t), \qquad t \geq 0,$$

defines a Q_1-Wiener process W on U_1, where $Q_1 = JJ^*$ is a non-negative bounded linear operator on U_1 such that $\mathbf{Tr}\, Q_1 < \infty$. For arbitrary $h \in U$, the process

$$\langle h, W(t)\rangle = \sum_{k=1}^{\infty}\langle h, g_k\rangle\beta_k(t)$$

is a real-valued Wiener process and

$$\mathbb{E}\langle h_1, W(t)\rangle\langle h_2, W(s)\rangle = \min\{t, s\}\langle Qh_1, h_2\rangle$$

for all $h_1, h_2 \in U$ and $t, s \geq 0$. In the case when Q is of trace class, $Q^{\frac{1}{2}}$ is Hilbert-Schmidt, so we can take $U_1 = U$ to arrive at a Q-Wiener process as defined above. If $\mathbf{Tr}\, Q = \infty$, we will call the constructed process W a *cylindrical* Q-Wiener process on U. When $Q = Id$, we just say that W is a cylindrical Wiener process.

Let H be another separable Hilbert space. Following Chapter 4 of [109], the stochastic integral

$$\int_0^t \Phi(s)dW(s) \tag{A.1}$$

with respect to a (cylindrical) Q-Wiener process W may be constructed for any predictable process $\Phi = (\Phi(t))_{t\geq 0}$ taking values in the space of Hilbert-Schmidt operators from U_0 into H such that

$$\mathbb{P}\left\{\int_0^t \|\Phi(s)\|^2_{L_{HS}(U_0,H)}ds < \infty\right\} = 1.$$

The stochastic integral (A.1) is independent of the choice of U_1 and J. For such Φ, (A.1) is a continuous square-integrable martingale, and its quadratic variation is of the form

$$\left[\int_0^{\cdot} \Phi(s)dW(s)\right]_t = \int_0^t Q_\Phi(s)ds,$$

where $Q_\Phi(s) = (\Phi(s)Q^{\frac{1}{2}})(\Phi(s)Q^{\frac{1}{2}})^*$.

A.4 Solutions to evolution equations

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and suppose U and H are separable Hilbert spaces as above. Let W be a (cylindrical) Q-Wiener process on U. Consider the stochastic evolution equation

$$dX(t) = \mathbf{A}X(t)dt + \mathbf{B}(X(t))dW(t), \qquad X(0) = x \in H, \tag{A.2}$$

where $\mathbf{A}$ is the infinitesimal generator of a strongly continuous semigroup $(S_t)_{t\geq 0} = \left(e^{t\mathbf{A}}\right)_{t\geq 0}$ on H and $\mathbf{B} : H \to L_{HS}(U_0, H)$ is measurable. A predictable H-valued process $X = (X(t))_{t\geq 0}$ is said to be a mild solution of (A.2) if for arbitrary $t \geq 0$

$$\mathbb{P}\left(\int_0^t |X(s)|^2 ds < \infty\right) = 1 \tag{A.3}$$

and

$$X(t) = S_t x + \int_0^t S_{t-s}\mathbf{B}(X(s))dW(s)$$

$\mathbb{P}$-a.s. We have the following existence and uniqueness result (Theorem 7.4 from [109]):

Theorem A.4.1. *Assume that x is an $\mathcal{F}_0$-measurable H-valued random variable. Suppose also that there exists a constant $C > 0$ such that $\|\mathbf{B}(y) - \mathbf{B}(z)\|_{L_{HS}(U_0,H)} \leq C\|y - z\|$ for all $y, z \in H$ and $\|\mathbf{B}(y)\|^2_{L_{HS}(U_0,H)} \leq C^2(1 + \|y\|^2)$ for all $y \in H$. Then there exists a mild solution X to* (A.2)*, unique up to processes satisfying* (A.3).

References

[1] S. Aida, T. Masuda, and I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.*, 126(1):83–101, 1994.

[2] S. Aida and D. Stroock. Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.*, 1:75–86, 1994.

[3] M. Aizenmann and R. Holley. Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin-Shlosman regime. In H. Kesten, editor, *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, pages 1–11. Springer, 1987.

[4] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Ergodicity of L^2-semigroups and extremality of Gibbs states. *J. Funct. Anal.*, 144:394–423, 1997.

[5] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*. Number 10 in Panoramas et Synthèses. Soc. Math. France, Paris, 2000.

[6] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equillibrium for Fokker-Planck type equations. *Comm. Partial Differential Equations*, 26(1-2):43–100, 2001.

[7] D. Bakry. *L'hypercontractivité et son utilisation en théorie des semigroups*, pages 1–114. Number 1581 in Lecture Notes in Math. Springer, 1994.

[8] D. Bakry. Functional inequalities for Markov semigroups. In S. G. Dani and P. Graczyk, editors, *Probability measures on groups: recent directions and trends*, pages 91–147. Tata Inst. Fund. Res., Mumbai, 2006.

[9] D. Bakry, F. Baudoin, M. Bonnefont, and D. Chafaï. On gradient bounds for the heat kernel on the Heisenberg group. *J. Funct. Anal.*, 255:1905–1938, 2008.

[10] D. Bakry, F. Baudoin, M. Bonnefont, and B. Qian. Subelliptic Li-Yau estimates on three dimensional model spaces. In D. Bakry, L. Beznea, N. Boboc, and M. Röckner, editors, *Potential Theory and Stochastics in Albac: Aurel Cornea Memorial Volume*, 2009.

[11] D. Bakry and M. Émery. Hypercontractivité de semi-groupes de diffusion. *C. R. Acad. Sci. Paris Sér. I Math.*, 299(15):775 – 778, 1984.

[12] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités, XIX, 1983/84*, number 1123 in Lecture Notes in Math., pages 177–206. Springer, Berlin, 1985.

[13] D. Bakry and M. Ledoux. Lévy-Gromov's isoperimetric inequality for an infinite dimensional diffusion generator. *Invent. Math.*, 123:259–281, 1996.

[14] F. Barthe, P. Cattiaux, and C. Roberto. Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry. *Rev. Mat. Iberoamericana*, 22(3):993–1067, 2006.

[15] F. Barthe and B. Maurey. Some remarks on isoperimetry of Gaussian type. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(4):419–434, 2000.

[16] F. Baudoin. *An Introduction to the Geometry of Stochastic Flows*. Imperial College Press, 2004.

[17] F. Baudoin, M. Hairer, and J. Teichmann. Ornstein-Uhlenbeck processes on Lie groups. *J. Funct. Anal.*, 255:877–890, 2008.

[18] R. Beals, B. Gaveau, and P. C. Greiner. Hamilton-Jacobi theory and the heat kernel on Heisenberg groups. *J. Math. Pures Appl.*, 79(9):633–689, 2000.

[19] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.*, 105(2):397–400, 1989.

[20] A. Bellaïche. The tangent space in sub-Riemannian geometry. In A. Bellaïche and J. J. Risler, editors, *Sub-Riemannian Geometry*, number 144 in Progr. Math., pages 1–78. Birkhäuser, 1996.

[21] C. Bernardin. Hydrodynamics for a system of harmonic oscillators perturbed by a conservative noise. *Stochastic Process. Appl.*, 117(4):487–513, 2007.

[22] C. Bernardin. Superdiffusivity of asymmetric energy model in dimensions 1 and 2. *J. Math. Phys.*, 49(10):103301–103321, 2008.

[23] C. Bernardin and S. Olla. Fourier's law for a microscopic model of heat conduction. *J. Stat. Phys.*, 121(3-4):271–289, 2005.

[24] L. Bertini and B. Zegarliński. Coercive inequalities for Gibbs measures. *J. Funct. Anal.*, 162(2):257–286, 1999.

[25] L. Bertini and B. Zegarliński. Coercive inequalities for Kawasaki dynamics: the product case. *Markov Process. and Rel. Fields*, 5(2):125–162, 1999.

[26] S. Bobkov. A functional form of the isoperimetric inequality for the Gaussian measure. *J. Funct. Anal.*, 135(1):39–49, 1996.

[27] S. Bobkov. An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. *Ann. Probab.*, 25(1):206–214, 1997.

[28] S. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163:1–28, 1999.

[29] S. Bobkov and C. Houdré. Isoperimetric constants for product probability measures. *Ann. Probab.*, 25(1):184–205, 1997.

[30] S. Bobkov and C. Houdré. Some connections between isoperimetric and Sobolev-type inequalities. *Mem. Amer. Math. Soc.*, 129(616), 1997.

[31] S. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.*, 10(5):1028–1052, 2000.

[32] S. Bobkov and B. Zegarliński. Entropy bounds and isoperimetry. *Mem. Amer. Math. Soc.*, 176(829), 2005.

[33] Th. Bodineau and B. Helffer. The log-Sobolev inequality for unbounded spin systems. *J. Funct. Anal.*, 166:168–178, 1999.

[34] Th. Bodineau and B. Helffer. Correlations, spectral gap and log-Sobolev inequalities for unbounded spin systems. In *Differential Equations and Mathematical Physics (Birmingham, AL, 1999)*, number 16 in AMS/IP Stud. Adv. Math., pages 51–66. Amer. Math. Soc., 2000.

[35] F. Bolley and I. Gentil. Phi-entropy inequalities and Fokker-Planck equations. Preprint: arXiv:1002.4478v1, 2010.

[36] F. Bonetto, J. L. Lebowitz, J. Lukkarinen, and S. Olla. Heat conduction and entropy production in anharmonic crystals with self-consistent stochastic reservoirs. *J. Stat. Phys.*, 134(5-6):1097–1119, 2009.

[37] F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet. Fourier's law: A challenge for theorists. In A. Fokas et al., editor, *Mathematical Physics 2000*, pages 128–150. Imperial College Press, 2000.

[38] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. *Stratified Lie Groups and Potential Theory for their Sub-Laplacians*. Springer Monographs in Mathematics. Springer, 2007.

[39] C. Borell. The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30:207–216, 1975.

[40] C. Carathéodory. Untersuchungen über die Grundlagen der Thermodynamik. *Math. Ann.*, 67:355–386, 1909.

[41] E. Carlen and D. Stroock. An application of the Bakry-Emery criterion to infinite dimensional diffusions. In *Séminaire de Probabilités, XX*, number 1204 in Lecture Notes in Math., pages 341–348. Springer, 1986.

[42] P. Carmona. Existence and uniqueness of an invariant measure for a chain of oscillators with two heat baths. *Stochastic Process. Appl.*, 117(8):1076–1092, 2007.

[43] D. Chafaï. Entropies, convexity and functional inequalities: on Phi-entropies and Phi-Sobolev inequalities. *J. Math. Kyoto Univ.*, 44(2):325–363, 2004.

[44] D. Chafaï. Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue. *ESAIM Probab. Stat. 10*, 10:317–339, 2006.

[45] J. A. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In R. C. Gunning, editor, *Problems in Analysis: A Symposium in Honour of S. Bochner*, pages 195–199, 1970.

[46] W. L. Chow. Über systeme non linearen partiellen Differentialgleichungen erster Odung. *Math. Ann.*, 117:98–105, 1939.

[47] F. Cipriani. Sobolev-Orlicz imbeddings, weak compactness, and spectrum. *J. Funct. Anal.*, 177:89–106, 2000.

[48] I. Csiszár. A class of measures of informativity of observation channels. *I. Period. Math. Hungar.*, 2(191-213), 1972.

[49] E. B. Davies. *Heat Kernels and Spectral Theory*. Cambridge University Press, 1989.

[50] E. B. Davies and B. Simon. Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *J. Funct. Anal.*, 59:335–395, 1984.

[51] B. Driver and T. Melcher. Hypoelliptic heat kernel inequalities on the Heisenberg group. *J. Funct. Anal.*, 221:340–365, 2005.

[52] N. Eldredge. Precise estimates for the subelliptic heat kernel on H-type groups. *J. Math. Pures Appl.*, 92:52–85, 2009.

[53] N. Eldredge. Gradient estimates for the subelliptic heat kernel on H-type groups. *J. Funct. Anal.*, 258(2):504–533, 2010.

[54] P. Federbush. Partially alternative derivation of a result of Nelson. *J. Math. Phys.*, 10:50–52, 1969.

[55] G. B. Folland. Subelliptic estimates and function spaces on nilpotent Lie groups. *Ark. Mat.*, 13:161–207, 1975.

[56] P. Fougères. Hypercontractivité et isopérimétrie Gaussienne. Applications aux systémes de spins. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(5):647–689, 2000.

[57] R. L. Frank and A. Laptev. Inequalities between Dirichlet and Neumann eigenvalues on the Heisenberg group. Preprint: arXiv:0906.1402v1, 2010.

[58] J. Fritz, K. Nagy, and S. Olla. Equilibrium fluctuations for a system of harmonic oscillators with conservative noise. *J. Stat. Phys.*, 122(3):399–415, 2006.

[59] N. Garofalo and D-M. Nhieu. Lipschitz continuity, global smooth approximations and extension theorems for Sobolev functions in Carnot-Carathéodory spaces. *J. Anal. Math.*, 74:67–97, 1998.

[60] C. Giardinà, J. Kurchan, and F. Redig. Duality and exact correlations for a model of heat conduction. *J. Math. Phys.*, 48(3):103301, 2007.

[61] F. Z. Gong and F. Y. Wang. Functional inequalities for uniformly integrable semigroups and application to essential spectrums. *Forum Mathematicum*, 14:293–313, 2002.

[62] G. Grillo. On Persson's theorem in local Dirichlet spaces. *Z. Anal. Anwendungen*, 17:329–338, 1998.

[63] M. Gromov. Carnot-Carathéodory spaces seen from within. In A. Bellaïche and J. J. Risler, editors, *Sub-Riemannian Geometry*, Progress In Mathematics, pages 85–324. Birkhäuser, 1996.

[64] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97:1061–1083, 1975.

[65] L. Gross and O. Rothaus. Herbst inequalities for supercontractive semigroups. *J. Math. Kyoto Univ.*, 38(2):295–318, 1998.

[66] A. Guionnet and B. Zegarliński. Lectures on logarithmic Sobolev inequalites. In *Séminaire de Probabilités, XXXVI*, number 1801 in Lecture Notes in Math., pages 1–134. Springer-Verlag, 2003.

[67] A. M. Hansson. An inequality between Dirichlet and Neumann eigenvalues of the Heisenberg Laplcian. *Comm. Partial Differential Equations*, 33(10-12):2157–2163, 2008.

[68] A. M. Hansson and A. Laptev. Sharp spectral inequalities for the Heisenberg Laplacian. In K. Trent, editor, *Groups and Analysis*, number 354 in London Mathematical Society Lecture Note Series. CUP, 2008.

[69] W. Hebisch and B. Zegarliński. Coercive inequalities on metric measure spaces. *J. Funct. Anal.*, 258:814–851, 2010.

[70] B. Helffer. Remarks on decay of correlations and Witten Laplacians III. Application to logarithmic Sobolev inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 35:483–508, 1999.

[71] B. Helffer. *Semiclassical Analysis, Witten Laplacians, and Statistical Mechanics*. Number 1 in Partial Differential Equations and Applications. World Scientific, 2002.

[72] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46(5-6):1159–1194, 1987.

[73] R. Holley and D. Stroock. Uniform and L^2 convergence in one dimensional stochastic Ising models. *Comm. Math. Phys.*, 123:85–93, 1989.

[74] Y. Z. Hu. A unified approach to several inequalities for Gaussian and diffusion measures. In *Séminaire de Probabilités, XXXIV*, volume 1729 of *Lecture Notes in Math.*, pages 329–335. Springer, 200.

[75] H. Hueber and D. Müller. Asymptotics for some Green kernels on the Heisenberg group and the Martin boundary. *Math. Ann.*, 283:97–119, 1989.

[76] J. Inglis. Operators on the Heisenberg group with discrete spectra. In *Proceedings of the 7th ISAAC congress*. World Scientific, 2010.

[77] J. Inglis, V. Kontis, and B. Zegarliński. From U-bounds to isoperimetry with applications to H-type groups. *To appear in J. Funct. Anal.*

[78] J. Inglis, M. Neklyudov, and B. Zegarliński. Ergodicity for infinite particle systems with locally conserved quantities. Preprint: arXiv:1002.0282v1, 2010.

[79] J. Inglis, M. Neklyudov, and B. Zegarliński. Liggett-type inequalities and interacting particle systems: the Gaussian case. In *Proceedings of the 7th ISAAC congress*. World Scientific, 2010.

[80] J. Inglis and I. Papageorgiou. Logarithmic Sobolev inequalities for infinite dimensional Hörmander type generators on the Heisenberg group. *J. Pot. Anal.*, 31(1):79–102, 2009.

[81] E. Janvresse, C. Landim, J. Quastel, and H. T. Yau. Relaxation to equilibrium of conservative dynamics. I. Zero-range processes. *Ann. Probab.*, 27(1):325–360, 1999.

[82] A. Kaplan. Fundamental solutions for a class of hypoelliptic PDE generated by composition of quadratic forms. *Trans. Amer. Math. Soc.*, 258(1):147–153, 1980.

[83] G. Kondrat, S. Peszat, and B. Zegarliński. Ergodicity for generalized Kawasaki dynamics. *J. Phys. A*, 33:5901–5912, 2000.

[84] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics: Non-relativistic Theory*. Pergamon Press, 1958.

[85] C. Landim and H. T. Yau. Convergence to equilibrium of conservative particle systems on $\mathbb{Z}^d$. *Ann. Probab.*, 31(1):115–147, 2003.

[86] R. Latała and K. Oleszkiewicz. Between Sobolev and Poincaré. In *Geometric Aspects of Functional Analysis*, volume 1745 of *Lecture Notes in Math.*, pages 147–168. Springer, Berlin, 2000.

[87] M. Ledoux. Isopérimétrie et inégalités de Sobolev logarithmiques Gaussiennes. *C. R. Acad. Sci. Paris Sér. I Math.*, 306(2):79–82, 1988.

[88] M. Ledoux. A simple analytic proof of an inequality by P. Buser. *Proc. Amer. Math. Soc.*, 121(3):951–959, 1994.

[89] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, number 1709 in Lecture Notes in Math., pages 120–216. Springer, Berlin, 1999.

[90] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités, XXXV*, number 1755 in Lecture Notes in Math., pages 167–194. Springer, Berlin, 2001.

[91] H. Q. Li. Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg. *J. Funct. Anal.*, 236:369–394, 2006.

[92] T. Liggett. L_2 rates of convergence for attractive reversible nearest particle systems: the critical case. *Ann. Probab.*, 19(3):935–959, 1991.

[93] T. Liggett. *Interacting Particle Systems*. Springer, 2004.

[94] S. L. Lu and H. T. Yau. Spectral gap and logarithmic Sobolev inequalities for Kawasaki and Glauber dynamics. *Comm. Math. Phys.*, 156:399–433, 1993.

[95] P. Ługiewicz and B. Zegarliński. Coercive inequalities for Hörmander type generators in infinite dimensions. *J. Funct. Anal.*, 247(2):438–476, 2007.

[96] Z. M. Ma and M. R. Röckner. *Introduction to the Theory of (non-symmetric) Dirichlet forms*. Universitext. Springer-Verlag, 1992.

[97] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region I. The attactive case. *Comm. Math. Phys.*, 161:447–486, 1994.

[98] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region II. The general case. *Comm. Math. Phys.*, 161:487–514, 1994.

[99] T. Melcher. *Hypoelliptic heat kernel inequalities on Lie groups*. PhD thesis, University of California, San Diego, 2004.

[100] E. Milman. On the role of convexity in functional and isoperimetric inequalities. *Proc. London Math. Soc.*, 99(3):32–66, 2009.

[101] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs*. Amer. Math. Soc., 2002.

[102] R. Monti. Some properties of Carnot-Carathéodory balls in the Heisenberg group. *Rend. Mat. Acc. Lincei*, 11:155–167, 2000.

[103] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.

[104] F. Otto and M. G. Reznikoff. A new criterion for the logarithmic Sobolev inequality and two applications. *J. Funct. Anal.*, 243:121–157, 2007.

[105] I. Papageorgiou. The logarithmic Sobolev inequality in infinite dimensions for unbounded spin systems on the lattice with non-quadratic interactions. Preprint: arXiv:0901.1403v1, 2009.

[106] I. Papageorgiou. Perturbing the logarithmic Sobolev inequality for unbounded spin systems on the lattice with non-quadratic interactions. Preprint: arXiv:0901.1482v1, 2009.

[107] I. Papageorgiou. *The logarithmic Sobolev inequality in infinite dimensions*. PhD thesis, Imperial College London, 2010.

[108] A. Persson. Bounds for the discrete part of the spectrum of a semibounded Schrödinger operator. *Math. Scand.*, 8:143–153, 1960.

[109] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Number 44 in Encyclopedia of Mathematics and its Applications. CUP, 1992.

[110] M. Reed and B. Simon. *IV: Analysis of Operators*. Methods of Modern Mathematical Physics. Academic Press, 1978.

[111] C. Roberto and B. Zegarliński. Orlicz-Sobolev inequalities for sub-Gaussian measures and ergodicity of Markov semigroups. *J. Funct. Anal.*, 243(1):28–66, 2007.

[112] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*. Cambridge Mathematical Library. Cambridge University Press, 2000.

[113] O. Rothaus. Analytic inequalities, isoperimetric inequalities and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 64:296–313, 1985.

[114] O. Rothaus. Hypercontractivity and the Bakry-Emery criterion for compact Lie groups. *J. Funct. Anal.*, 65(3):358–367, 1986.

[115] L. Saloff-Coste. *Aspects of Sobolev-Type Inequalities*. Number 289 in London Mathematical Society Lecture Note Series. Cambridge University Press, 2002.

[116] D. Stroock and B. Zegarliński. The logarithmic Sobolev inequality for continuous spin systems on a lattice. *J. Funct. Anal.*, 104:299–326, 1992.

[117] D. Stroock and B. Zegarliński. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Comm. Math. Phys.*, 149:175–193, 1992.

[118] D. Stroock and B. Zegarliński. On the ergodic properties of Glauber dynamics. *J. Stat. Phys.*, 81:1007–1019, 1995.

[119] V. N. Sudakov and B. S. Tsirel'son. Extremal properties of half-spaces for spherically invariant measures. *J. Sov. Math.*, 9:9–18, 1978.

[120] M. Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric Aspects of Functional Analysis*, number 1469 in Lecture Notes in Math., pages 94–124. Springer-Verlag, 1991.

[121] G. Tessitore and J. Zabczyk. Trotter's formula for transition semigroups. *Semigroup Forum*, 63:114–126, 2001.

[122] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and Geometry on Groups*. Number 100 in Cambridge Tracts in Mathematics. Cambridge University Press, 1992.

[123] F. Y. Wang. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields*, 109:417–424, 1997.

[124] F. Y. Wang. Functional inequalities for empty essential spectrum. *J. Funct. Anal.*, 170:219–245, 2000.

[125] F. Y. Wang. Logarithmic Sobolev inequalities: conditions and counterexamples. *J. Operator Theory*, 46:183–197, 2001.

[126] F. Y. Wang. Functional inequalities and spectrum estimates: the infinite measure case. *J. Funct. Anal.*, 194:288–310, 2002.

[127] F. Y. Wang. Super and weak Poincaré inequalities for hypoelliptic operators. *Acta Math. Appl. Sin. Engl. Ser.*, 25(4):617–630, 2009.

[128] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge Mathematical Library. Cambridge University Press, 1927.

[129] N. Yosida. The log-Sobolev inequality for weakly coupled lattice fields. *Probab. Theory Related Fields*, 115:1–40, 1999.

[130] N. Yosida. Application of log-Sobolev inequality to the stochastic dynamics of unbounded spin systems on the lattice. *J. Funct. Anal.*, 173:74–102, 2000.

[131] N. Yosida. The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on a lattice. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2):223–243, 2001.

[132] B. Zegarliński. Log-Sobolev inequalities for infinite one-dimensional lattice systems. *Comm. Math. Phys.*, 133(147-162), 1990.

[133] B. Zegarliński. On log-Sobolev inequalities for infinite lattice systems. *Lett. Math. Phys.*, 20(3):173–182, 1990.

[134] B. Zegarliński. The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice. *Comm. Math. Phys.*, 175:401–432, 1996.