hexagon:2611.00669v1

Gradient structures on metric measure spaces

Sobolev spaces, Heat kernels and Examples

Fabrice Baudoin

Department of Mathematics, Aarhus University
fbaudoin@math.au.dk

Version 1, November 2026

Contents

Preface iv

Introduction viii

I Gradient structures and first-order analysis 1

1 Gradient structures 1
1.1 A model: two measures on an interval . . . 1
1.2 Setting and notation . . . 3
1.3 The definition . . . 4
1.4 Calculus . . . 5
1.5 Fibres and rank . . . 7
1.6 Cores, cutoffs and spanning . . . 8
1.7 Recovery of the metric . . . 10

2 Sobolev spaces, bounded variation and energy forms 11
2.1 Admissible exponents and Sobolev spaces . . . 11
2.2 The trace on the skeleton and the chain rule . . . 12
2.3 Functional-analytic properties . . . 16
2.4 The p-energy form . . . 17
2.5 The Dirichlet form . . . 19
2.6 Relaxation and bounded variation . . . 22
2.7 Divergence duality and coarea . . . 22

3 Doubling, cutoffs and Poincaré inequalities 26
3.1 Scale functions . . . 26
3.2 The three hypotheses . . . 27
3.3 Discrete convolution . . . 29
3.4 Comparison of the scales . . . 32
3.5 The distance cutoff . . . 34
3.6 Overview of the examples . . . 34

4 Korevaar–Schoen theory 36
4.1 The two basic inequalities . . . 36
4.2 Closability . . . 37
4.3 Characterization by difference quotients . . . 38
4.4 Discrete gradients and compatibility of the exponents . . . 39
4.5 The endpoint $p = 1$ . . . 40
4.6 Critical exponents . . . 41
4.7 Gagliardo–Nirenberg inequalities . . . 44

5 The rank theorem 47
5.1 Statements . . . 47
5.2 The metric input: oscillation and Euclidean images . . . 49
5.3 The rank, the volume and the Poincaré scale . . . 52
5.4 Two auxiliary lemmas . . . 53

5.5 Proof of the rank theorem . . . 56

6 Self-improvement of Poincaré inequalities 63
6.1 Setting and the self-improvement theorem . . . 63
6.2 Scale comparison and chains . . . 65
6.3 The cutoff modification . . . 67
6.4 The compact case . . . 68
6.5 Localization by uniform domains . . . 76

II Dirichlet spaces and heat kernels 80

7 Dirichlet spaces and Harnack inequalities 80
7.1 Dirichlet spaces . . . 80
7.2 Capacity and cutoff Sobolev inequalities . . . 82
7.3 The parabolic Harnack inequality . . . 84
7.4 Cutoffs and Sobolev inequalities . . . 86
7.5 Mean value inequalities . . . 88
7.6 The logarithm and the proof of the Harnack inequality . . . 93

8 Heat kernel estimates 96
8.1 The heat kernel under volume doubling . . . 96
8.2 Uniform volume growth and Nash inequalities . . . 100
8.3 Upper bounds at small times . . . 102
8.4 Lower bounds . . . 106
8.5 Global estimates . . . 111

9 Applications to gradient structures 118
9.1 The Dirichlet form of a gradient structure . . . 118
9.2 A trace inequality . . . 119
9.3 Cutoff Sobolev inequalities for distinct scales . . . 120
9.4 Matching scales: the cutoff Sobolev inequality . . . 122
9.5 Matching scales: sub-Gaussian heat kernel estimates . . . 123
9.6 Heat kernel characterization of Sobolev and BV functions . . . 126

III Examples of gradient structures 129

10 Riemannian and sub-Riemannian structures 130
10.1 Riemannian manifolds . . . 130
10.2 Measurable Riemannian metrics . . . 132
10.3 Sub-Riemannian manifolds . . . 134
10.4 Sub-Riemannian manifolds: global estimates . . . 136

11 RCD(K, N) spaces 138
11.1 The gradient structure . . . 138
11.2 The exponents $1 < p < \infty$ . . . 139
11.3 Poincaré inequalities and scales . . . 139
11.4 The exponent one . . . 140

11.5 Energies, heat kernels and volume . . . 141

12 Cubical inverse limits of Laakso type 142
12.1 The inverse system and its measure . . . 142
12.2 The gradient on the sheets . . . 144
12.3 Poincaré inequalities and recovery of the metric . . . 145

13 The harmonic Sierpiński gasket 147
13.1 The harmonic gasket . . . 147
13.2 The gradient structure . . . 148
13.3 Scales and critical exponents . . . 149

14 Liouville Brownian motion 151
14.1 The Liouville measure and its Dirichlet form . . . 151
14.2 Admissibility and relaxation . . . 152
14.3 Proof of Proposition 14.2 . . . 154
14.4 Local Poincaré and cutoff inequalities . . . 156
14.5 Volume, exit times and heat kernels . . . 158

15 Riemannian gluings and the tree of cubes 161
15.1 Finite Riemannian gluings . . . 161
15.2 Local trees . . . 162
15.3 Countable Riemannian gluings . . . 164
15.4 The tree of cubes: construction and volume growth . . . 165
15.5 The tree of cubes: gradient and cutoffs . . . 168
15.6 The tree of cubes: admissibility and Poincaré inequalities . . . 170
15.7 The tree of cubes: density of the algebra $\mathcal{A}$ . . . 174
15.8 The tree of cubes: critical Korevaar–Schoen spaces . . . 176

References 179

Preface

Motivation

The starting point of this book is my work with Li Chen on the Vicsek set [21] and with Li Chen and Meng Yang on local trees [22]. In a *real tree*, any two distinct points x, y are joined by a unique arc $[x, y]$, and this arc is isometric to $[0, d(x, y)]$. A *local tree* is a metric space in which every point has a ball that is a real tree with the restricted metric. The *skeleton* $\mathcal{S}$ is the union of the interiors of the geodesic arcs in such balls, and its length measure is $\nu = \mathcal{H}^1 \restriction_{\mathcal{S}}$.

Fix a root o in a separable real tree and orient the arcs $[o, y]$ away from o. For Lipschitz u, we can differentiate u along each arc parametrized by arclength. The oriented parametrizations agree on common initial segments, so the derivatives define a single function ∂u ν-almost everywhere. The fundamental theorem of calculus gives

$$u(y) - u(o) = \int_{(o,y)} \partial u \, d\nu.$$

For a continuous u, the same identity characterizes the weak derivative whenever there is a density integrable along each rooted arc. On a local tree the construction applies on tree balls.

Trees can still have a complicated geometry and genuinely be fractal sets; For instance, the Vicsek set with its intrinsic metric is a real tree. Its skeleton is, up to endpoints, the union of the increasing cable approximations (Figure 1). The natural reference measure m on X is the Hausdorff measure of dimension $\log 5 / \log 3$ and gives zero mass to $\mathcal{S}$.

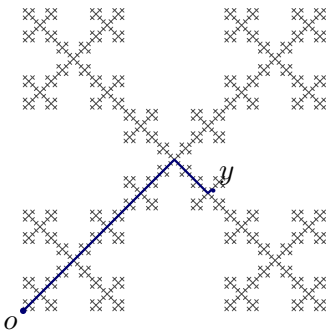


Figure 1: The level-four Vicsek cable approximation and the arc $[o, y]$ to a point y.

For local trees with uniform volume growth, the Sobolev theory uses $u \in L^p(X, m)$ and $\partial u \in L^p(\mathcal{S}, \nu)$. The BV energy is obtained by $L^1(X, m)$-relaxation of $\int_{\mathcal{S}} |\partial u| \, d\nu$ on $W^{1,1}(X, m)$ [22, Section 2.3]. Intuitively, a local tree has *rank one* because its derivative is scalar.

The motivation for this work was to develop a theory of *higher rank* spaces. Obviously, an n-dimensional Riemannian manifold should have *rank* n since the gradient takes values in an n-dimensional space. The notion of a gradient structure on metric spaces we are interested accommodates both cases and allows for fractal examples.

Definition (Definition 1.2)**.** A *gradient structure* on a metric space (X, d) is a triple $(\mathcal{S}, \nu, \nabla)$, where $\mathcal{S} \subset X$ is Borel, ν is a σ-finite Borel measure carried by $\mathcal{S}$, and

$$\nabla : \mathrm{Lip}_b(X) \longrightarrow L^\infty(\mathcal{S}, \mathsf{H}, \nu)$$

is a linear map into the measurable sections of a field $\mathsf{H} = (H_x)_{x \in \mathcal{S}}$ of nonzero subspaces of a fixed $\mathbb{R}^N$. It satisfies the Leibniz rule, the bound $|\nabla u| \le \mathrm{Lip}(u)$ and a weak-$*$ continuity condition. Its pointwise rank is $n(x) = \dim H_x$, and its global rank is $n_* = \operatorname{ess\,sup}_\nu n(x) \le N$.

There will be reference measure on (X, d), often the Hausdorff measure, so that from the beginning we will consider metric measures spaces (X, d, m). The definition of gradient structure, and this the fundamental feature allowing genuinely fractal examples, allows ν to be singular with respect to the *reference measure* m and infinite on every ball.

The Sobolev space $W^{1,p}(X, m)$ is the domain of the closure ∇_p of

$$\nabla : \mathcal{C}_p \subset L^p(X, m) \longrightarrow L^p(\mathcal{S}, \mathsf{H}, \nu)\,, \qquad \mathcal{C}_p = \{u \in \mathrm{Lip}_c(X) : \nabla u \in L^p(\mathcal{S}, \mathsf{H}, \nu)\}\,,$$

when the structure is p-admissible, meaning that the core is uniformly dense in $C_c(X)$ and the operator is closable. The p-energy is then $\mathcal{E}_p(f) = \int_{\mathcal{S}} |\nabla_p f|^p \, d\nu$. For $p = 2$ one obtains a strongly local regular Dirichlet form on $L^2(X, m)$ with energy measures $|\nabla_2 f|^2 \nu$ and pointwise index $n(x)$ in the sense of Hino [85, 86]. Its index and the martingale dimension of the associated diffusion are both n_* (Proposition 2.7). The book develops the theory of these Sobolev spaces and their associated heat kernels.

Use of generative AI

This book was generated almost entirely with generative AI. I used approximately 200 prompts and spent about 40 hours working with Astra 6 (standard speed, max or ultra settings), 20 hours with Sol 6.1 (standard speed, various settings), and 20 hours with Fable 5 and Fable 5.2 (Extra effort). I used Astra 6 to develop proofs, Sol 6.1 for editing, proofreading and bibliographical work, and Fable 5 and Fable 5.2 for writing.[1] The methodology was the following:

1. I had an informal notion of a gradient structure in mind and looked for examples before formalizing it. Trees give examples of rank one. I was particularly interested in examples of rank $n \ge 2$ with $m \perp \nu$. For trees whose length measures are singular with respect to their reference measures, the natural product still has rank one, since the measures carrying the partial derivatives are mutually singular. In my first prompt to Astra 6, I supplied earlier work, including [22], and asked, in substance, whether limits of cubes glued along their interfaces could carry a weak gradient obtained from the ordinary gradients on the cubes, just as the derivative on the Vicsek set is obtained from the derivatives on the cable approximations. Astra returned an initial example of the construction in Section 15.4. Figure 2 shows three stages of this construction in dimension two.

2. I next studied the Sobolev spaces and heat kernel of this example, focusing on the critical L^p Korevaar-Schoen exponents, a longstanding interest of mine [31], and on weak monotonicity. I was surprised that, at the critical exponent, weak monotonicity holds precisely for $p > n$, where n is the rank. Its failure at $p = 2$ suggested an obstruction to two-sided sub-Gaussian heat kernel estimates. This led me to examine rates for Poincaré inequalities and capacity estimates. Both hold, but with different sharp scales.

[1] I used the `math-paper` skill, available at `https://github.com/amelieloher/math-paper-skills`.

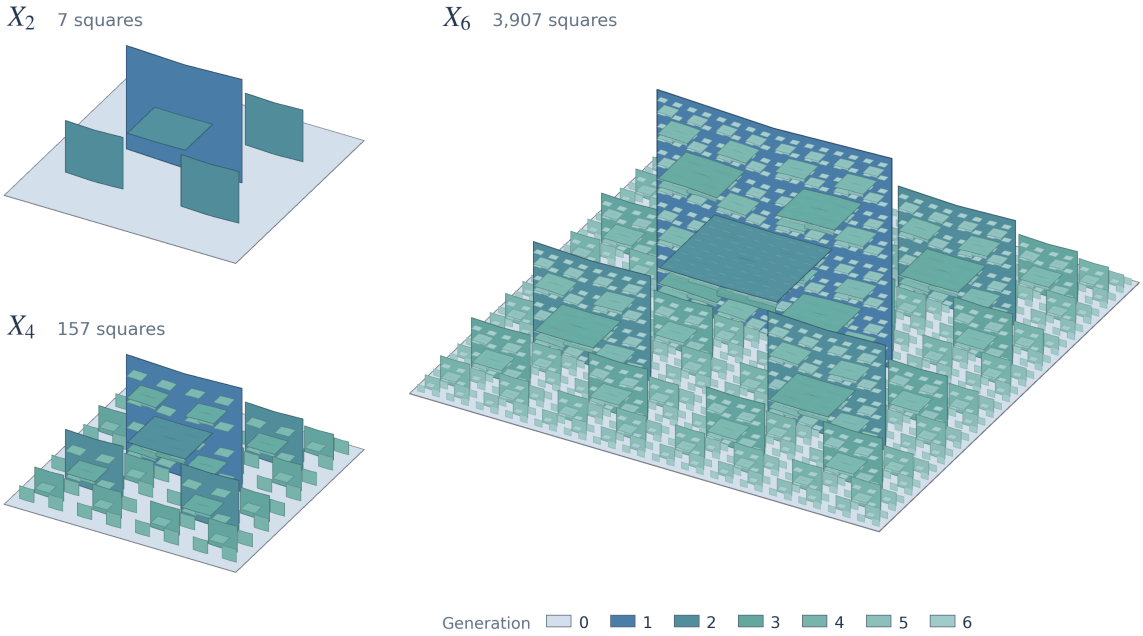


Figure 2: Stages X_2, X_4 and X_6 of the tree of cubes with $n = 2$, Hausdorff dimension $Q = \log_2 5$. Each mesh square has four ordinary dyadic children and one new square attached along a segment. The embedding in $\mathbb{R}^3$ compresses the transverse lengths of the new squares.

3. I then looked for a family of examples in which the scales could be varied. Astra obtained the family of trees of cubes described in Section 15.4. The difficulty of finding examples of rank at least two with genuinely sub-Gaussian heat kernel estimates suggested a structural obstruction. This question led to the rank theorem in Chapter 5: If a gradient structure has sub-Gaussian heat kernel estimates, then the rank has to be one.

4. Since gradient structures must have have different Poincaré and capacity scales, I then investigated how the theory changes when the scales differ. The theory of heat kernels and Sobolev spaces at matching scales is well established [17, 10, 3]. In the two-scale Harnack inequality, the Poincaré scale Ψ_P and the cutoff Sobolev scale Ψ_S enter the Harnack constant through $\Psi_P(r)/T + T/\Psi_S(r)$, where r is the radius and T the time parameter (Theorem 7.5). Under uniform volume growth, Ψ_P governs the prefactor in the heat kernel upper bound and Ψ_S governs the off-diagonal decay (Theorems 8.6 and 8.17).

5. With these results in hand, I returned to the foundations and developed with Astra and Fable the framework of Chapters 1 and 2.

6. To test the theory, I chose to study the problem self-improvement of Poincaré inequalities, a classical phenomenon in metric measure spaces [101]. It is a natural question here because the same gradient is used across different L^p spaces. This leads to Chapter 6.

7. I proposed most of the examples and used the models to check if the axioms were satisfied and the relevant properties. The cubical inverse limit of Laakso type in Chapter 12 came from Astra, which I had asked for a higher-rank version of Laakso's construction.

8. Successive rounds of writing, polishing and editing with Sol and Fable reduced a 350-page draft to the present version. I am entirely responsible for the architecture of the document section by section.

Comments on Version 1

This is a first draft. Some of the results may admit sharper forms (like for the heat kernel estimates). There likely will be revisions. My main aim was to develop a framework in which new questions about first-order analysis on metric spaces could be formulated.

Acknowledgments

This work was supported in part by the Independent Research Fund Denmark (grant 10.46540/4283-00175B), by the Villum Investigator grant *Stochastic Analysis in Aarhus*, and by the European Research Council (ERC) through the European Union's Horizon Europe research and innovation programme (RanGe project, Grant Agreement No. 101199772).

Introduction

Two measures

Let (X, d, m) be a metric measure space. First-order analysis on X starts from an energy of the form

$$\mathcal{E}_p(u) = \int_{\mathcal{S}} |\nabla u|^p \, d\nu \,,$$

where ∇u is a gradient of the function u and ν is a measure. In the classical situations $\nu = m$, and the gradient is defined at m-almost every point. This book studies the situation where ν differs from m. Four examples, all treated in Part III, indicate the range of the phenomena.

- *Manifolds.* On a Riemannian manifold, or on an RCD space, $\nu = m$ is the volume, and the theory reduces to the classical one.
- *Trees.* On a local tree, ∇u is the derivative along the arcs and ν is the length measure, carried by the union $\mathcal{S}$ of the open arcs. The reference measure m is arbitrary within a class of measures with controlled volume growth; it may give no mass to $\mathcal{S}$, and ν may be infinite on every ball.
- *The tree of cubes.* Replacing the arcs by cubes of dimension $n \geq 2$, attached to each other through small windows, gives a space where ν is locally the Lebesgue measure of dimension n, while m has dimension $Q > n$ and is singular with respect to ν.
- *Liouville Brownian motion.* Here X is the plane, ∇ is the Euclidean gradient, ν is the Lebesgue measure, and m is a Gaussian multiplicative chaos.

In all these examples the gradient at a point takes its values in a Euclidean space of finite dimension, which may depend on the point.

Gradient structures and Sobolev spaces

A *gradient structure* on (X, d, m) (Definition 1.2) consists of a Borel set $\mathcal{S}$, the *skeleton*, a σ-finite measure ν carried by $\mathcal{S}$, and a linear map ∇ from the bounded Lipschitz functions to the bounded measurable sections of a measurable field $(H_x)_{x\in\mathcal{S}}$ of Euclidean spaces, which satisfies the Leibniz rule, the bound $|\nabla u| \leq \mathrm{Lip}(u)$, and a condition of weak-$*$ continuity. This is a derivation in the sense of Weaver [152], with values in a field of finite-dimensional spaces. The dimension $n(x)$ of H_x is the *pointwise rank*. The reference measure does not appear in the definition; the measures m and ν may be mutually singular, and ν need not be finite on balls. The bound $|\nabla u| \leq \mathrm{Lip}(u)$ says that the gradient is controlled by the metric. The converse, that the metric is recovered from the gradient, is an additional property, the local Sobolev-to-Lipschitz property of Definition 1.10, which holds in all the examples.

The reference measure enters with the Sobolev spaces (Chapter 2). The structure is *p-admissible* if the gradient, defined on a core $\mathcal{C}_p$ of Lipschitz functions with compact support, is closable from $L^p(X, m)$ to $L^p(\mathcal{S}, \nu)$; its closure ∇_p then defines the Sobolev space $W^{1,p}(X, m)$. Closability is a compatibility between the two measures, and it may fail: for Liouville Brownian motion it holds exactly for $p > \gamma^2/2$, and for the tree of cubes exactly for $p \geq p_*$, with an explicit threshold (Proposition 14.2 and Theorem 15.8). At an admissible exponent, Chapter 2 establishes the following.

- *Trace.* A function of $W^{1,p}(X, m)$ is a class modulo m-null sets, and ν may be singular with respect to m. Nevertheless such a function has a trace on the part of the skeleton which

carries the gradients, defined ν-almost everywhere, and this trace gives the chain rule for compositions with Lipschitz functions (Proposition 2.3).

- *Functional analysis.* $W^{1,p}(X,m)$ is a Banach space, reflexive for $1<p<\infty$, with a divergence defined by duality (Proposition 2.5).
- *Energies.* The energy $\mathcal{E}_p$ is a regular p-energy form with the generalized contraction property of Kajino and Shimizu [98] (Proposition 2.6). At $p=2$ it is a strongly local regular Dirichlet form whose index, in the sense of Hino [85, 86], is the rank of the gradient structure when the core spans the fibres (Proposition 2.7).
- *Bounded variation.* At $p=1$, the relaxation of the energy defines the total variation and the space $BV(X,m)$, with a coarea formula (Theorem 2.11). This definition requires no admissibility.

Three hypotheses and two scales

The quantitative theory (Chapter 3) rests on three hypotheses, each required at small radii or at all radii:

- *volume doubling* VD for the reference measure;
- a *Poincaré inequality* $\mathrm{PI}_p(\Psi_{P,p})$,

$$\int_{B(x,r)} |u-u_{B(x,r)}|^p\,dm \le C\,\Psi_{P,p}(r)\int_{B(x,\Lambda r)\cap\mathcal{S}} |\nabla u|^p\,d\nu\,;$$

- the *cutoff condition* $\mathrm{B}(\Psi_{C,2})$: every ball $B(x,r)$ has a Lipschitz cutoff ψ_x^r, equal to one on $B(x,r)$, vanishing outside $B(x,2r)$, with Lipschitz constant at most C/r, such that

$$\int_{\mathcal{S}} |\nabla\psi_x^r|\,d\nu \le C\,\frac{r}{\Psi_{C,2}(r)}\,m(B(x,r))\,.$$

When $\nu=m$, volume doubling gives the cutoff bound for the distance cutoff at the scale $\Psi_{C,2}(r)=r^2$. The bound (3.11) on the ν-measure of $\{|\nabla\psi_x^r|>0\}$ is a sufficient condition for the integrated estimate. The cutoff estimate and the Lipschitz bound give $\int|\nabla\psi_x^r|^p\,d\nu \le C\,m(B(x,r))/\Psi_{C,p}(r)$, with

$$\Psi_{C,p}(r)=r^{p-2}\,\Psi_{C,2}(r)\,.$$

The functions $\Psi_{P,p}$ and $\Psi_{C,p}$ are the *Poincaré scale* and the *cutoff scale*. Under the three hypotheses, on a connected space, $\Psi_{C,p}\le C\,\Psi_{P,p}$ at small radii (Proposition 3.7). The scales *match* at the exponent p if they are comparable. The distinction between m and ν is essential in the arguments of approximation which follow: a set on which a function is modified may have an infinite measure for ν, even when its measure for m is small, and the cutoff condition supplies the control of the gradients on such sets. An additional hypothesis $\mathrm{V}(\Phi)$ of uniform volume growth gives explicit formulas, and is used only where it is stated.

The results of Part I

Korevaar–Schoen characterizations and closability (Chapter 4). Let $D_p(f,r)$ be the integral of $|f(x)-f(y)|^p$ over the pairs of points at distance less than r, normalized by the volume

of the balls, as in (3.17). Under the three hypotheses, the relaxed energy $\overline{\mathcal{E}}_p$ satisfies, for every $f \in L^p(X,m)$,

$$\sup_r \frac{D_p(f,r)}{\Psi_{P,p}(r)} \le C\,\overline{\mathcal{E}}_p(f)\,, \qquad \overline{\mathcal{E}}_p(f) \le C \liminf_{r\to 0} \frac{D_p(f,r)}{\Psi_{C,p}(r)}$$

(Theorem 4.4). Each inequality involves its own scale. If the ratio $\Psi_{P,p}/\Psi_{C,p}$ is bounded along one sequence of radii tending to zero, the structure is p-admissible (Theorem 4.3), and the two inequalities characterize $W^{1,p}(X,m)$ for $p>1$, and $BV(X,m)$ for $p=1$, by the boundedness of the difference quotients $D_p(f,r)/\Psi_{C,p}(r)$ along this sequence (Theorems 4.4 and 4.8). At matching power scales this determines the critical exponents of the Korevaar–Schoen spaces (Theorem 4.12), and a lower bound for the volume of balls gives Gagliardo–Nirenberg inequalities, by the method of Bakry, Coulhon, Ledoux and Saloff-Coste [16, 31] (Theorem 4.13).

The rank theorem (Chapter 5). The scales restrict the number of independent gradients. If the scales match at an exponent $p>1$ and the common scale has a lower scaling exponent greater than p, then

$$n(x) < p \qquad \text{for } \nu\text{-almost every } x$$

(Theorems 5.1 and 5.2). At $p=2$ this says that a Dirichlet form of a gradient structure whose Poincaré and capacity scales match, at a scale of higher order than r^2, has index one (Theorem 5.4). The tree of cubes shows that the bound cannot be improved: its rank is n and its scales match exactly for $p>n$.

Self-improvement (Chapter 6). A Poincaré inequality at the cutoff scale improves itself, as in the theorem of Keith and Zhong [101]: a $(1,p)$-Poincaré inequality, together with a cutoff condition of order $b<p$, yields an (s,q)-Poincaré inequality with $q<p<s$, in which the gradient is integrated over the ball on which the oscillation is measured (Theorem 6.1). Hence, if $p>1$ is a matching exponent, so is every exponent of an interval $[p-\varepsilon, p]$.

These arguments connect the calculus of derivations [152, 144] with the theory of Korevaar and Schoen [112, 31, 22] and with the self-improvement of Poincaré inequalities [101, 106, 109].

Heat kernels: the results of Part II

Part II starts independently, with a strongly local regular Dirichlet space $(X,d,m,\mathcal{E},\mathcal{F})$ which satisfies volume doubling, a Poincaré inequality $\mathrm{PI}(\Psi_P)$, and a cutoff Sobolev inequality in annuli $\mathrm{CSA}(\Psi_S)$ in the sense of Andres and Barlow [10]. The classical theory assumes $\Psi_P \asymp \Psi_S$. Here the two scales may differ; on a geodesic space, necessarily $\Psi_S \le C\,\Psi_P$ (Proposition 7.4).

Harnack inequality (Chapter 7). A nonnegative solution u of the heat equation in a cylinder of radius r and time length $2T$ satisfies

$$\operatorname*{ess\,sup}_{Q_-} u \le \exp\Big[C\Big(1+\frac{\Psi_P(r)}{T}+\frac{T}{\Psi_S(r)}\Big)\Big]\operatorname*{ess\,inf}_{Q_+} u$$

(Theorem 7.5). The radius and the time are independent parameters, and each scale enters through one term. When the scales are comparable and T is the common scale at the radius r, the constant is uniform and one recovers the parabolic Harnack inequality.

Heat kernel bounds (Chapter 8). The Harnack inequality, the method of Davies, Nash inequalities and chaining arguments give upper and lower bounds for the heat kernel, at small times under local hypotheses and at all times under global ones. In the upper bounds, the exponential factor is governed by the cutoff Sobolev scale, through the rate function $\mathcal{A}_{\Psi_S}$ of (8.1), and, under uniform volume growth, the prefactor is governed by the Poincaré scale. The lower bounds involve both scales. Table 2 lists the statements. At matching scales $\Psi_P \asymp \Psi_S \asymp \Psi$, Corollary 8.12 gives the two-sided sub-Gaussian estimate

$$\frac{c\, e^{-C\mathcal{A}_\Psi(d(x,y),t)}}{m(B(x,\Psi^{-1}(t)))} \le p_t(x,y) \le \frac{C\, e^{-c\mathcal{A}_\Psi(d(x,y),t)}}{m(B(x,\Psi^{-1}(t)))}\,.$$

The hypothesis $\mathrm{V}(\Phi)$ is invoked separately where uniform prefactors are stated.

Gradient structures (Chapter 9). For the Dirichlet form of a 2-admissible gradient structure, the cutoff condition gives an upper bound for capacities, but not a cutoff Sobolev inequality. At matching scales, the capacity bound and the Poincaré inequality imply the cutoff Sobolev inequality, by a theorem of Eriksson-Bique [64]; hence volume doubling, $\mathrm{PI}_2(\Psi)$ and $\mathrm{B}(\Psi)$ imply the two-sided sub-Gaussian estimates (Theorem 9.8), and Sobolev and BV functions are characterized by the heat kernel (Theorem 9.10). For distinct scales, a cutoff Sobolev inequality is derived from a condition on the energy of the cutoffs in small balls, the condition $\mathrm{B}^+(\Psi_{C,2})$, which follows from the cutoff condition and doubling when the cutoff scale is r^2 (Proposition 9.3).

The examples of Part III

Table 1 in Section 3.6 lists the examples with their ranks and scales. Chapters 10 and 11 treat the spaces with $\nu = m$ and Gaussian scales: Riemannian manifolds, measurable Riemannian metrics, sub-Riemannian manifolds and $\mathrm{RCD}(K,N)$ spaces. The remaining chapters contain the following results.

- *Cubical inverse limits* (Chapter 12). For every integer $n \ge 1$ and every real $Q > n$ there is an Ahlfors Q-regular space with $m = \nu$ and a gradient structure of rank n, admissible for every $p \ge 1$, with Gaussian scales (Theorem 12.1).
- *The harmonic Sierpiński gasket* (Chapter 13). The gradient has rank one and is closable for every $p \ge 1$; the kernel of the closed gradient consists of the constants, and the critical Korevaar–Schoen exponents are equal to one for every p (Proposition 13.1). There is no uniform volume function. Whether the Poincaré inequality $\mathrm{PI}_p(r^p)$ holds for all $p \ge 1$ is not known.
- *Liouville Brownian motion* (Chapter 14). For the parameter $0 < \gamma < 2$, almost surely, the Euclidean gradient is closable exactly for $p > \gamma^2/2$, and at and below this threshold the relaxed energy vanishes identically (Proposition 14.2). Locally, the Poincaré inequality and the cutoff Sobolev inequality hold at different power scales.
- *The tree of cubes* (Chapter 15). The structure has constant rank n, the measures m and ν are mutually singular, and the scales match exactly for $p > n$. The critical Korevaar–Schoen exponents are equal to $1 + (Q-n)/p$ and are attained for every $p \ge 1$; for $p > n$ the critical space is $W^{1,p}(X,m)$, and for admissible $1 < p \le n$ it is a proper dense subspace (Theorem 15.11).

Relation to the literature

First-order calculus. The framework combines derivations on metric measure spaces [152, 144] with the measurable differential calculus of Dirichlet forms [54, 85, 86, 90]. Closures of p-energies derived from a quadratic gradient are studied in [89, 119], and definitions of Sobolev and BV spaces by derivations are developed in [61]. In the present framework the metric is prescribed and bounds the derivation, and the quantitative estimates relate the two measures. Laakso's construction with cubes [120, Remark 3.2] and the cubical inverse systems of Cheeger and Kleiner [50, Section 11] provide the geometric setting of Chapter 12; the results of Schioppa on these systems identify the coordinate derivations [143, Theorem 3.20].

Approximation, rank and self-improvement. The arguments of approximation belong to the literature on discrete convolutions and on Korevaar–Schoen energies [107, 31, 22, 146]. The difference with the construction of a p-energy as a limit of metric energies [5, 99] is that here the derivation is prescribed, and its closability must be proved. The proof of the rank theorem uses the methods of image measures and capacities of [87, 88, 52], and the proof of the self-improvement adapts the method of maximal functions of [101, 109]. The hypotheses of these works are specified where their arguments enter.

Heat kernels. The theory with one scale is due to the works on Gaussian and sub-Gaussian Harnack inequalities [77, 140, 151, 17, 10, 75, 121]. The perturbation method of Davies, Nash inequalities and chaining arguments supply the methods for the upper and lower bounds [48, 130, 55, 78]. The estimates with two distinct scales in this book are derived from its stated hypotheses; the citations to the theory with one scale identify the methods, and do not replace those derivations.

How to read this book

Figure 3 shows the dependencies between the chapters. Chapters 1, 2 and 3 are the foundation of everything which follows in Parts I and III. Chapters 5 and 6 are independent of each other; from Chapter 4 they use mainly the admissibility at matching scales, Theorem 4.3. Chapters 7 and 8 can be read without Part I, by a reader interested in Dirichlet spaces. Chapter 9 requires Chapters 2, 3 and 4, together with Chapters 7 and 8. The chapters of Part III are independent of each other, apart from occasional references. They rest on Chapters 1 to 3, and they quote the results of Chapters 4 and 9 when they apply them; a reader may wish to consult an example, for instance the local trees of Section 15.2 or the Riemannian manifolds of Chapter 10, while reading Part I.

Conventions. Statements are numbered consecutively within each chapter. The letters c and C denote positive constants whose value may change from line to line, and $A \asymp B$ means that $c\,B \le A \le C\,B$. The dependence of the constants is specified in each chapter, usually by the phrase *structural constants*. References to the literature are given at the point of use, with the number of the statement which is cited whenever this is possible. Where a proof adapts a known method to the present setting, the source of the method is indicated, and the hypotheses under which the original statement was proved are those of the reference, not those of this book.

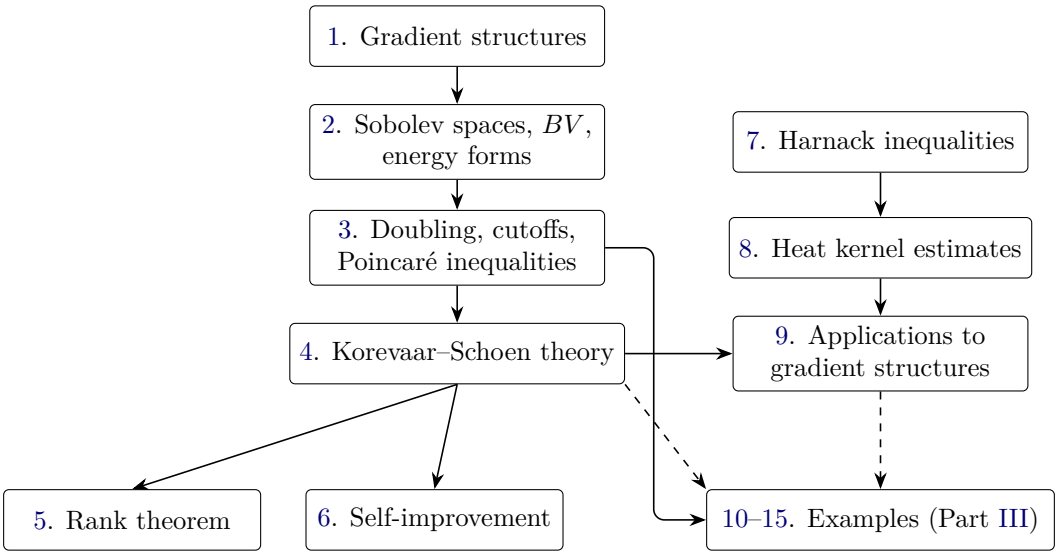


Figure 3: Dependencies between the chapters. A solid arrow means that the chapter at its head relies on the chapter at its tail; a dashed arrow means that results are quoted where they are applied.

Part I
Gradient structures and first-order analysis

1 Gradient structures

On a Riemannian manifold the energy $\int |\nabla u|^2 \, d\mathrm{vol}$ integrates the gradient against the same measure that defines L^2. On singular spaces this coincidence often fails. On a tree whose mass is carried by its boundary, on a self-similar fractal whose energy measures are singular with respect to the Hausdorff measure [118, 84], or after a time change of a diffusion [53, Chapter 5], the measure which carries the energy is singular with respect to the measure which defines the function spaces. This chapter sets up a first-order calculus in which the two measures are independent data.

The data are a metric measure space (X, d, m), a Borel set $\mathcal{S} \subset X$ carrying a second measure ν, and a linear map ∇ which assigns to every bounded Lipschitz function on X a bounded field of vectors over $\mathcal{S}$. Three axioms are imposed: the Leibniz rule, the bound $|\nabla u| \le \mathrm{Lip}(u)$, and a continuity property with respect to bounded pointwise convergence. They are the axioms of a derivation in the sense of Weaver [152], written for a vector-valued map. The measure m does not appear in them.

The chapter is organized as follows. Section 1.1 treats the unit interval with an arbitrary reference measure, where every phenomenon of interest can be computed by hand. Sections 1.2 and 1.3 fix the notation and give the definition. Section 1.4 derives locality and the chain rule from the axioms. Section 1.5 constructs the measurable field of fibres in which the gradients take their values and defines the rank. Section 1.6 introduces the cores $\mathcal{C}_p$ of Lipschitz functions with p-integrable gradient and gives a criterion for these cores to be large. Section 1.7 formulates the condition under which the gradient determines the metric. Closures of the gradient in $L^p(X, m)$ are the subject of Chapter 2.

1.1 A model: two measures on an interval

Let

$$X = [0,1]\,, \quad d(x,y) = |x-y|\,, \quad \mathcal{S} = (0,1)\,, \quad \nu = \mathcal{L}^1 \restriction_{(0,1)}\,, \quad \nabla u = u' \quad (u \in \mathrm{Lip}([0,1]))\,,$$

and let m be a finite Borel measure on $[0,1]$ with $\operatorname{supp} m = [0,1]$. Nothing else is assumed on m: it may be singular with respect to the Lebesgue measure, and it may be purely atomic, with atoms at the rational points. For $1 \le p < \infty$ and $u \in \mathrm{Lip}([0,1])$ consider the norm

$$\|u\|_{W^{1,p}}^p = \int_{[0,1]} |u|^p \, dm + \int_0^1 |u'|^p \, dx\,.$$

The first integral uses m and the second uses ν. Recall that the classical Sobolev space $W^{1,p}(0,1)$ consists of the absolutely continuous functions on $[0,1]$ whose derivative belongs to $L^p((0,1), dx)$. A continuous function on $[0,1]$ is determined by its m-equivalence class, because m has full support; in this sense $W^{1,p}(0,1) \subset L^p([0,1], m)$.

Proposition 1.1. *Let $1 \le p < \infty$. The operator*

$$\mathrm{Lip}([0,1]) \subset L^p([0,1], m) \longrightarrow L^p((0,1), dx)\,, \qquad u \longmapsto u'\,,$$

is closable. The domain of its closure is $W^{1,p}(0,1)$, and on this space $\|\cdot\|_{W^{1,p}}$ is equivalent to the classical norm $\big(\int_0^1 |u|^p \, dx + \int_0^1 |u'|^p \, dx\big)^{1/p}$.

Proof. Step 1: a uniform bound. Let u be absolutely continuous with $u' \in L^p(0,1)$. For all $x, y \in [0,1]$, Hölder's inequality gives

$$|u(y) - u(x)| \le \int_0^1 |u'|\, dt \le \|u'\|_{L^p(0,1)}\,.$$

We average in x with respect to m and then apply Hölder's inequality for m:

$$\Big|u(y) - \fint_{[0,1]} u\, dm\Big| \le \|u'\|_{L^p(0,1)}\,, \qquad \|u\|_\infty \le m([0,1])^{-1/p}\|u\|_{L^p(m)} + \|u'\|_{L^p(0,1)}\,. \tag{1.1}$$

The same argument with the Lebesgue measure in place of m bounds $\|u\|_\infty$ by the classical Sobolev norm. Since $\|u\|_{L^p(m)} \le m([0,1])^{1/p}\|u\|_\infty$ and $\|u\|_{L^p(dx)} \le \|u\|_\infty$, the two Sobolev norms are equivalent.

Step 2: closability. Let (u_k) be a sequence in $\mathrm{Lip}([0,1])$ which is Cauchy for $\|\cdot\|_{W^{1,p}}$, and let g be the limit of (u_k') in $L^p(0,1)$. By (1.1), applied to the differences $u_k - u_l$, the sequence converges uniformly to a continuous function u. Passing to the limit in $u_k(y) - u_k(x) = \int_x^y u_k'\, dt$ gives

$$u(y) - u(x) = \int_x^y g(t)\, dt \qquad (x, y \in [0,1])\,.$$

Thus $u \in W^{1,p}(0,1)$ and $u' = g$. If moreover $u_k \to 0$ in $L^p(m)$, then $u = 0$ m-almost everywhere. Since u is continuous and m has full support, $u = 0$ everywhere, and therefore $g = 0$. This is closability, and the argument shows that the domain of the closure is contained in $W^{1,p}(0,1)$.

Step 3: the domain. Conversely, let $u \in W^{1,p}(0,1)$ and let g_k be continuous functions with $g_k \to u'$ in $L^p(0,1)$. The functions $u_k(x) = u(0) + \int_0^x g_k\, dt$ are Lipschitz, converge uniformly to u, and satisfy $u_k' = g_k \to u'$ in $L^p(0,1)$. Hence u belongs to the domain of the closure. □

Three features of this example recur throughout the book.

The source of closability. Closability came from two facts: the estimate (1.1), which bounds the oscillation of u by the $L^p(\nu)$ norm of its derivative, and the full support of m. A bound of the oscillation on a ball by the energy on a comparable ball is a Poincaré inequality. In Chapter 4, Poincaré inequalities at a scale comparable to the energy of cutoff functions are the general source of closability (Theorem 4.3).

The role of the reference measure. The domain $W^{1,p}(0,1)$ does not depend on the measure m; only the norm does. The operator associated with the quadratic form $\int_0^1 |u'|^2\, dx$ on $L^2([0,1], m)$ is the Krein–Feller operator $\frac{d}{dm}\frac{d}{dx}$ with Neumann boundary conditions, and the associated diffusion is the time change of reflecting Brownian motion with speed measure m; see [53, Chapter 5] for time changes of symmetric Markov processes.

Higher rank. In dimension one a function of finite energy is continuous. For gradients with $n \ge 2$ independent components and exponents $p \le n$ this fails, and closability is no longer automatic. The tree of cubes of Chapter 15 has rank n, and its gradient is closable exactly in the range of exponents given in Theorem 15.8.

1.2 Setting and notation

Throughout the book, (X, d) is a locally compact separable metric space and m is a Radon measure on X with full support. Thus m is finite on compact sets and positive on nonempty open sets, X is σ-compact, and m is σ-finite. Further assumptions on (X, d), such as completeness or the existence of geodesics, are stated where they are used. We use the following notation.

- $B(x, r) = \{y \in X : d(x, y) < r\}$ is the open ball of centre x and radius r. For a ball $B = B(x, r)$ and $\lambda > 0$ we write $\lambda B = B(x, \lambda r)$ and $r_B = r$. A ball is always given together with its centre and its radius.
- $\mathrm{Lip}(u) = \sup_{x \neq y} |u(x) - u(y)|/d(x, y)$ is the Lipschitz constant of $u : X \to \mathbb{R}$. The spaces $\mathrm{Lip}_b(X)$ and $\mathrm{Lip}_c(X)$ consist of the bounded Lipschitz functions and of the Lipschitz functions with compact support.
- For a Borel set A with $0 < m(A) < \infty$ and $f \in L^1(A, m)$,
$$f_A = \fint_A f \, dm = \frac{1}{m(A)} \int_A f \, dm \,.$$
- For a function f and a set B, $\operatorname{osc}_B f = \sup_B f - \inf_B f$.
- $a \vee b = \max\{a, b\}$, $a \wedge b = \min\{a, b\}$, $a_+ = a \vee 0$ and $a_- = (-a) \vee 0$.
- $\mathbf{1}_A$ is the indicator function of a set A, and $\mu \restriction_A$ is the restriction of a measure μ to A.
- For $1 \le p \le \infty$, the conjugate exponent is $p' = p/(p-1)$, with $1' = \infty$ and $\infty' = 1$.

Constants denoted by C or c may change from line to line; their dependence is indicated when it matters.

Measurable fields of Euclidean spaces. The gradient of a function will be a field of vectors $x \mapsto \nabla u(x) \in H_x$, where the spaces H_x may depend on the point and may have different dimensions at different points. We realize all of them inside one Euclidean space. Let $\mathcal{S} \subset X$ be a Borel set, let ν be a σ-finite Borel measure on X with $\nu(X \setminus \mathcal{S}) = 0$, and let $N \ge 1$ be an integer. A *measurable field* of subspaces of $\mathbb{R}^N$ over $\mathcal{S}$ is a family $\mathsf{H} = (H_x)_{x \in \mathcal{S}}$ of linear subspaces $H_x \subset \mathbb{R}^N$ such that the entries of the matrix of the orthogonal projection $\pi_x : \mathbb{R}^N \to H_x$ are Borel functions of x. Each H_x carries the Euclidean norm $|\cdot|$ and the inner product $\langle \cdot, \cdot \rangle$ of $\mathbb{R}^N$. For $1 \le p \le \infty$, the space of *p-integrable sections* is

$$L^p(\mathcal{S}, \mathsf{H}, \nu) = \big\{ V \in L^p(\mathcal{S}, \mathbb{R}^N, \nu) : \ V(x) \in H_x \text{ for } \nu\text{-almost every } x \big\} \,. \tag{1.2}$$

It is a closed subspace of $L^p(\mathcal{S}, \mathbb{R}^N, \nu)$, and $(\pi V)(x) = \pi_x V(x)$ defines a projection π of norm at most one of $L^p(\mathcal{S}, \mathbb{R}^N, \nu)$ onto it. When no confusion is possible we write $L^p(\nu)$ for $L^p(\mathcal{S}, \mathsf{H}, \nu)$. The projection π transfers three properties of the spaces $L^p(\mathcal{S}, \mathbb{R}^N, \nu)$ to the spaces of sections.

- *Duality.* Let $1 \le p < \infty$. The pairing $(V, W) \mapsto \int_{\mathcal{S}} \langle V, W \rangle \, d\nu$ identifies the dual of $L^p(\mathcal{S}, \mathsf{H}, \nu)$ isometrically with $L^{p'}(\mathcal{S}, \mathsf{H}, \nu)$. Indeed, if ℓ is a bounded linear form on $L^p(\mathcal{S}, \mathsf{H}, \nu)$, then $\ell \circ \pi$ is a bounded linear form on $L^p(\mathcal{S}, \mathbb{R}^N, \nu)$. Since ν is σ-finite, it is represented by some $W \in L^{p'}(\mathcal{S}, \mathbb{R}^N, \nu)$ with $\|W\|_{p'} = \|\ell \circ \pi\| \le \|\ell\|$. For $V \in L^p(\mathcal{S}, \mathsf{H}, \nu)$ we get $\ell(V) = \int \langle V, W \rangle \, d\nu = \int \langle V, \pi W \rangle \, d\nu$ and $\|\pi W\|_{p'} \le \|\ell\|$; Hölder's inequality gives the reverse inequality.
- *Interpolation.* The same projection π maps $L^p(\mathcal{S}, \mathbb{R}^N, \nu)$ onto $L^p(\mathcal{S}, \mathsf{H}, \nu)$ for every p. Hence the interpolation theorems for linear operators between L^p spaces apply to operators with values in the spaces $L^p(\mathcal{S}, \mathsf{H}, \nu)$.

- *Weak-$*$ topology.* The space $L^\infty(\mathcal{S}, \mathsf{H}, \nu)$ is the dual of $L^1(\mathcal{S}, \mathsf{H}, \nu)$. A sequence (V_k) converges weak-$*$ to V in $L^\infty(\mathcal{S}, \mathsf{H}, \nu)$ if $\int_{\mathcal{S}} \langle V_k, g\rangle \, d\nu \to \int_{\mathcal{S}} \langle V, g\rangle \, d\nu$ for every $g \in L^1(\mathcal{S}, \mathsf{H}, \nu)$.

1.3 The definition

The axioms below are those of a derivation of the Lipschitz algebra in the sense of Weaver [152, Definition 21], imposed on each component of a vector-valued map. The values are taken in a measurable field of Euclidean spaces, as in the first-order calculus of Dirichlet forms [54, 86, 90]; compare the representation of derivations by measurable bundles in [152, Theorem 10].

Definition 1.2. A *gradient structure* on (X, d, m) is a triple $(\mathcal{S}, \nu, \nabla)$, where $\mathcal{S} \subset X$ is a Borel set, ν is a σ-finite Borel measure on X with $\nu(X \setminus \mathcal{S}) = 0$, and, for some measurable field $\mathsf{H} = (H_x)_{x \in \mathcal{S}}$ of nonzero subspaces of some $\mathbb{R}^N$,

$$\nabla : \mathrm{Lip}_b(X) \longrightarrow L^\infty(\mathcal{S}, \mathsf{H}, \nu)$$

is a linear map satisfying the following condition.

(G) *Leibniz rule and Lipschitz bound.* For $u, v \in \mathrm{Lip}_b(X)$,

$$\nabla(uv) = u\,\nabla v + v\,\nabla u\,, \qquad \|\nabla u\|_{L^\infty(\nu)} \le \mathrm{Lip}(u)\,.$$

Weak-$*$ continuity. If $u_k, u \in \mathrm{Lip}_b(X)$, $u_k \to u$ pointwise on X and $\sup_k(\|u_k\|_\infty + \mathrm{Lip}(u_k)) < \infty$, then

$$\int_{\mathcal{S}} \langle \nabla u_k, g\rangle \, d\nu \longrightarrow \int_{\mathcal{S}} \langle \nabla u, g\rangle \, d\nu \qquad \text{for every } g \in L^1(\mathcal{S}, \mathsf{H}, \nu)\,.$$

The field H is always taken to be the smallest measurable field such that $\nabla u(x) \in H_x$ for ν-almost every x, for each $u \in \mathrm{Lip}_b(X)$; it is constructed in Lemma 1.6. Points with $H_x = \{0\}$ are discarded, by restricting $\mathcal{S}$ and ν to the set where $H_x \neq \{0\}$. The *pointwise rank* and the *rank* of the structure are

$$n(x) = \dim H_x\,, \qquad n_* = \operatorname*{ess\,sup}_{x,\nu} n(x) \le N\,, \tag{1.3}$$

with $n_* = 0$ if $\nu = 0$. The structure has *constant rank* n if $n(x) = n$ for ν-almost every x.

We call $\mathcal{S}$ the *skeleton*, following [22], ν the *gradient measure*, m the *reference measure*, and H_x the *fibre* at x. In the right-hand side of the Leibniz rule, u and v stand for their restrictions to $\mathcal{S}$, which are defined at every point because u and v are continuous.

Comments on the definition.

- *The two measures.* The reference measure does not appear in (G), and no relation between m and ν is assumed. They may be equal, as on a Riemannian manifold. They may be mutually singular, as on a tree whose reference measure is carried by its leaves. The measure ν may give infinite mass to every ball, as in the tree of cubes of Chapter 15. The relation between the two measures enters only through the quantitative hypotheses of Chapter 3.
- *The Lipschitz bound.* The inequality $|\nabla u| \le \mathrm{Lip}(u)$ is the only link between the gradient and the metric in (G). It bounds the gradient by the metric. The converse bound, which recovers the metric from the gradient, is a separate condition and is discussed in Section 1.7.

- *Weak-$*$ continuity.* The hypothesis on (u_k) says that $u_k \to u$ in the weak-$*$ topology of the Lipschitz algebra, restricted to bounded sequences. The conclusion says that $\nabla u_k \to \nabla u$ weak-$*$ in $L^\infty(\mathcal{S}, \mathsf{H}, \nu)$. This axiom replaces integration by parts: on $\mathbb{R}^n$ it holds because $\int \langle \nabla u_k, g \rangle \, dx = -\int u_k \operatorname{div} g \, dx$ for smooth compactly supported g.
- *Derivations.* Each coordinate of ∇ in $\mathbb{R}^N$ is a derivation with respect to ν in the sense of [152, Definition 21]. Schioppa [144] relates derivations to Alberti representations of ν, that is, to decompositions of ν into measures carried by curve fragments; this relation is used in Lemma 5.6. On a space with a doubling measure and a Poincaré inequality, the differentiable structures of Cheeger [49] and Keith [103] provide derivations with $\nu = m$; see also [110, 19].

Example 1.3 (Euclidean space)**.** Let $X = \mathcal{S} = \mathbb{R}^n$ with the Euclidean distance, $m = \nu = \mathcal{L}^n$, and let ∇u be the almost everywhere defined gradient of $u \in \operatorname{Lip}_b(\mathbb{R}^n)$, which exists by Rademacher's theorem. The Leibniz rule and the bound $|\nabla u| \le \operatorname{Lip}(u)$ hold almost everywhere. If $u_k \to u$ pointwise with $\sup_k(\|u_k\|_\infty + \operatorname{Lip}(u_k)) < \infty$ and $g \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)$, then

$$\int \langle \nabla u_k, g \rangle \, dx = -\int u_k \operatorname{div} g \, dx \longrightarrow -\int u \operatorname{div} g \, dx = \int \langle \nabla u, g \rangle \, dx$$

by dominated convergence. Since $C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n, \mathbb{R}^n)$ and the fields ∇u_k are uniformly bounded, the convergence holds for every $g \in L^1$. The gradients of the coordinate functions, multiplied by cutoffs, span $\mathbb{R}^n$ at every point, so $H_x = \mathbb{R}^n$ and the structure has constant rank n. The interval of Section 1.1 is a gradient structure of constant rank one, for every choice of m, by the same argument.

Example 1.4 (A partial gradient)**.** On $X = \mathcal{S} = \mathbb{R}^2$ with $m = \nu = \mathcal{L}^2$, the map $\nabla u = \partial_{x_1} u$ satisfies (G), by the argument of Example 1.3. It is a gradient structure of constant rank one on a space of dimension two. The function $u(x) = x_2$, truncated at large distance, has $\nabla u = 0$ near the origin without being locally constant. Condition (G) therefore does not prevent a gradient structure from ignoring some directions of the metric space. Section 1.7 gives the condition which excludes this.

Further examples are the subject of Part III; Section 3.6 contains a summary.

1.4 Calculus

Locality and the chain rule follow from (G). Locality is the derivation property of [152, Lemma 27]; we give the argument for the present σ-finite gradient measure.

Lemma 1.5. *Let $(\mathcal{S}, \nu, \nabla)$ satisfy (G), and let $u, v \in \operatorname{Lip}_b(X)$.*

(i) Constants. $\nabla c = 0$ *for every constant c.*

(ii) Locality. *If $u = v$ on a Borel set A, then $\nabla u = \nabla v$ ν-almost everywhere on A.*

(iii) Local Lipschitz bound. *If $U \subset X$ is open, then $|\nabla u| \le \operatorname{Lip}(u|_U)$ ν-almost everywhere on U.*

(iv) Chain rule. *If $\varphi \in C^1(\mathbb{R})$, then $\varphi \circ u \in \operatorname{Lip}_b(X)$ and $\nabla(\varphi \circ u) = \varphi'(u) \nabla u$. If $a < b$, then $\nabla((u \vee a) \wedge b) = \mathbf{1}_{\{a<u<b\}} \nabla u$. Moreover,*

$$\nabla(u \vee v) = \mathbf{1}_{\{u \ge v\}} \nabla u + \mathbf{1}_{\{u<v\}} \nabla v \,, \qquad \nabla(u \wedge v) = \mathbf{1}_{\{u \ge v\}} \nabla v + \mathbf{1}_{\{u<v\}} \nabla u \,.$$

Proof. (i) By the Leibniz rule, $\nabla 1 = \nabla(1 \cdot 1) = 2\,\nabla 1$. Hence $\nabla 1 = 0$ and $\nabla c = c\,\nabla 1 = 0$.

(ii) Let $w = u - v$. We must show that $\nabla w = 0$ ν-almost everywhere on A.

Open sets. Suppose first that $w = 0$ on an open set U. If $U = X$, then $\nabla w = 0$ by linearity. Otherwise let $\chi = 1 \wedge d(\cdot, X \setminus U) \in \mathrm{Lip}_b(X)$. Then $w\chi = 0$ on X, so $0 = \chi\,\nabla w + w\,\nabla\chi$ by the Leibniz rule. On U we have $w = 0$ and $\chi > 0$, hence $\nabla w = 0$ ν-almost everywhere on U.

Borel sets. Suppose now that $w = 0$ on a Borel set A. The functions

$$w_k = \bigl(w - \tfrac{1}{k}\bigr)_+ - \bigl(w + \tfrac{1}{k}\bigr)_-\,, \qquad k \geq 1\,,$$

belong to $\mathrm{Lip}_b(X)$ and satisfy $\|w_k\|_\infty \leq \|w\|_\infty$, $\mathrm{Lip}(w_k) \leq \mathrm{Lip}(w)$ and $|w_k - w| \leq 1/k$. Moreover $w_k = 0$ on the open set $\{|w| < 1/k\}$, which contains A. By the first case, $\nabla w_k = 0$ ν-almost everywhere on A. Since ν is σ-finite, there is a Borel function $h : \mathcal{S} \to (0,1]$ with $\int_{\mathcal{S}} h\,d\nu < \infty$. The field $g = h\,\mathbf{1}_A \nabla w$ belongs to $L^1(\mathcal{S}, \mathsf{H}, \nu)$, and the weak-$*$ continuity in (G) gives

$$\int_A h\,|\nabla w|^2\,d\nu = \int_{\mathcal{S}} \langle \nabla w, g\rangle\,d\nu = \lim_{k\to\infty} \int_{\mathcal{S}} \langle \nabla w_k, g\rangle\,d\nu = 0\,.$$

Hence $\nabla w = 0$ ν-almost everywhere on A.

(iii) Let $L = \mathrm{Lip}(u|_U)$ and $M = \|u\|_\infty$; we may assume $U \neq \varnothing$. The truncated McShane extension [125]

$$\tilde u(y) = \Bigl(\inf_{z\in U} \bigl(u(z) + L\,d(y,z)\bigr) \vee (-M)\Bigr) \wedge M\,, \qquad y \in X\,,$$

is L-Lipschitz, bounded by M, and equal to u on U. By (ii) and (G), $|\nabla u| = |\nabla \tilde u| \leq L$ ν-almost everywhere on U.

(iv) For a monomial $\varphi(t) = t^k$, the formula $\nabla(u^k) = k\,u^{k-1}\nabla u$ follows from the Leibniz rule by induction on $k \geq 1$, and from (i) for $k = 0$. By linearity it holds for polynomials. Let $\varphi \in C^1(\mathbb{R})$ and $M = \|u\|_\infty$. The function $\varphi \circ u$ is bounded and Lipschitz because φ is Lipschitz on $[-M, M]$. By the Weierstrass theorem applied to φ', there are polynomials P_k with $P_k \to \varphi$ and $P_k' \to \varphi'$ uniformly on $[-M, M]$. The function $(\varphi - P_k) \circ u$ has Lipschitz constant at most $\|\varphi' - P_k'\|_{L^\infty([-M,M])}\,\mathrm{Lip}(u)$, so (G) gives

$$\bigl|\nabla(\varphi\circ u) - P_k'(u)\,\nabla u\bigr| = \bigl|\nabla\bigl((\varphi - P_k)\circ u\bigr)\bigr| \leq \|\varphi' - P_k'\|_{L^\infty([-M,M])}\,\mathrm{Lip}(u) \longrightarrow 0\,,$$

while $P_k'(u)\nabla u \to \varphi'(u)\nabla u$ uniformly. This proves the chain rule.

For the truncation $w = (u \vee a) \wedge b$, the space X is the union of the three Borel sets $\{u \leq a\}$, $\{u \geq b\}$ and $\{a < u < b\}$, on which w is equal to a, to b and to u respectively. By (ii) and (i), ∇w is equal to 0, to 0 and to ∇u on these sets. The formula for $u \vee v$ follows in the same way from the two sets $\{u \geq v\}$ and $\{u < v\}$, and the formula for $u \wedge v = u + v - u \vee v$ follows by linearity. □

Three consequences are used repeatedly.

- By (ii) and (i), $\nabla u = 0$ ν-almost everywhere on every level set $\{u = c\}$, and in particular outside $\operatorname{supp} u$.
- The gradient of a finite maximum or minimum of functions of $\mathrm{Lip}_b(X)$ coincides, on each set of a finite Borel partition of $\mathcal{S}$, with the gradient of one of them. In particular, for $u_1, \dots, u_k \in \mathrm{Lip}_b(X)$ and $1 \leq p < \infty$,

$$|\nabla(u_1 \vee \cdots \vee u_k)|^p \leq \sum_{i=1}^k |\nabla u_i|^p \qquad \nu\text{-almost everywhere.} \tag{1.4}$$

- For $u, v \in \mathrm{Lip}_b(X)$,

$$|\nabla(u \vee v)| + |\nabla(u \wedge v)| = |\nabla u| + |\nabla v| \qquad \nu\text{-almost everywhere.} \tag{1.5}$$

Functions of several variables. The chain rule extends to several functions. If $u_1, \dots, u_k \in \mathrm{Lip}_b(X)$ and $\varphi \in C^1(\mathbb{R}^k)$, then $\varphi(u_1, \dots, u_k) \in \mathrm{Lip}_b(X)$ and

$$\nabla \varphi(u_1, \dots, u_k) = \sum_{i=1}^{k} \partial_i \varphi(u_1, \dots, u_k)\, \nabla u_i \,. \tag{1.6}$$

The proof is the same: the formula holds for polynomials by the Leibniz rule, and polynomials are dense in $C^1(Q)$ for every compact cube $Q \subset \mathbb{R}^k$.

1.5 Fibres and rank

Definition 1.2 takes H to be the smallest measurable field containing all gradients. We now construct it, and we show that the rank can be read off from countably many functions.

Lemma 1.6. *Let $\mathcal{A} \subset L^\infty(\mathcal{S}, \mathbb{R}^N, \nu)$ be an arbitrary family of bounded fields.*

(i) *There is a measurable field $\mathsf{H}^{\mathcal{A}} = (H_x^{\mathcal{A}})_{x \in \mathcal{S}}$ such that every $V \in \mathcal{A}$ satisfies $V(x) \in H_x^{\mathcal{A}}$ for ν-almost every x, and such that $H_x^{\mathcal{A}} \subset H'_x$ for ν-almost every x whenever (H'_x) is a measurable field with the same property. It is unique up to ν-null sets.*

(ii) *There is a sequence (V_k) in $\mathcal{A}$ with $H_x^{\mathcal{A}} = \mathrm{span}\{V_k(x) : k \geq 1\}$ for ν-almost every x.*

(iii) *For a gradient structure, the families $\{\nabla u : u \in \mathrm{Lip}_b(X)\}$ and $\{\nabla u : u \in \mathrm{Lip}_c(X)\}$ generate the same field H.*

We call $\mathsf{H}^{\mathcal{A}}$ the field generated *by $\mathcal{A}$, and a sequence as in (ii) a* spanning sequence.

Proof. (i), (ii) Since ν is σ-finite, there is a finite measure $\nu' = h\nu$ with $h > 0$. The space $L^1(\mathcal{S}, \mathbb{R}^N, \nu')$ is separable and contains $\mathcal{A}$, so there is a sequence (V_k) in $\mathcal{A}$ which is dense in $\mathcal{A}$ for the norm of $L^1(\nu')$. Fix Borel representatives of the fields V_k and put $H_x = \mathrm{span}\{V_k(x) : k \geq 1\}$. For $j \geq 1$ let $A_j(x)$ be the $N \times j$ matrix with columns $V_1(x), \dots, V_j(x)$. The orthogonal projection onto the range of $A_j(x)$ is

$$P_j(x) = \lim_{\varepsilon \downarrow 0} A_j(x) \big(A_j(x)^{\mathsf{T}} A_j(x) + \varepsilon\big)^{-1} A_j(x)^{\mathsf{T}} \,,$$

as one sees from the singular value decomposition of $A_j(x)$; hence P_j has Borel entries. The ranges of the projections $P_j(x)$ increase with j and their union is H_x. Since H_x has finite dimension, $P_j(x) = \pi_x$ for large j. Thus $\pi_x = \lim_j P_j(x)$ has Borel entries and (H_x) is a measurable field.

Let $V \in \mathcal{A}$. A subsequence of (V_k) converges to V in $L^1(\nu')$, and a further subsequence converges ν-almost everywhere. Since H_x is closed, $V(x) \in H_x$ for ν-almost every x. Conversely, if (H'_x) is a measurable field which contains every field of $\mathcal{A}$ almost everywhere, then for ν-almost every x it contains $V_k(x)$ for all k, and therefore H_x. This proves the minimality, and the uniqueness follows.

(iii) Let H^c be the field generated by the gradients of $\mathrm{Lip}_c(X)$; clearly $H_x^c \subset H_x$. Since X is locally compact and separable, it is covered by countably many open sets U_i for which there are $\chi_i \in \mathrm{Lip}_c(X)$ with $\chi_i = 1$ on U_i. Let $u \in \mathrm{Lip}_b(X)$. Then $u\chi_i \in \mathrm{Lip}_c(X)$, and $\nabla u = \nabla(u\chi_i)$ ν-almost everywhere on U_i by Lemma 1.5(ii). Hence $\nabla u(x) \in H_x^c$ for ν-almost every x, and $H_x \subset H_x^c$ by minimality. □

By (iii), a gradient structure has a spanning sequence (∇u_k) with $u_k \in \mathrm{Lip}_c(X)$. If ν is finite on compact sets, the gradients of the functions of $\mathrm{Lip}_c(X)$ are integrable, because they are bounded and vanish outside a compact set. When ν is infinite on balls this fails, and the existence of spanning sequences with integrable gradients is a genuine condition; Proposition 1.9 gives a criterion.

Lemma 1.7. *Let $(u_k)_{k\ge1}$ be a sequence in $\mathrm{Lip}_b(X)$ whose gradients span H_x for ν-almost every x. Then, for ν-almost every x,*

$$n(x) = \sup_{j\ge1} \operatorname{rank}\big(\langle \nabla u_a(x), \nabla u_b(x)\rangle\big)_{a,b=1}^{j}. \tag{1.7}$$

In particular $x \mapsto n(x)$ is ν-measurable, and for every integer $k \ge 1$, up to a ν-null set,

$$\{x : n(x) \ge k\} = \bigcup_{j_1,\dots,j_k\ge1} \Big\{x : \det\big(\langle \nabla u_{j_a}(x), \nabla u_{j_b}(x)\rangle\big)_{a,b=1}^{k} > 0\Big\}. \tag{1.8}$$

Consequently, if $\nu(\{n \ge k\}) > 0$, then there are indices $j_1, \dots, j_k$ such that $\nabla u_{j_1}, \dots, \nabla u_{j_k}$ are linearly independent on a set of positive ν-measure.

Proof. Fix Borel representatives of the fields ∇u_k, and let x be a point with $\operatorname{span}\{\nabla u_k(x) : k \ge 1\} = H_x$. The rank of the Gram matrix of finitely many vectors of a Euclidean space is the dimension of their span. The subspaces $\operatorname{span}\{\nabla u_1(x), \dots, \nabla u_j(x)\}$ increase with j and their union is H_x; since H_x has finite dimension, they are equal to H_x for large j. This proves (1.7). Moreover $n(x) \ge k$ if and only if k of the vectors $\nabla u_j(x)$ are linearly independent, that is, if and only if their Gram determinant is positive. This is (1.8). The sets on its right-hand side are Borel and countably many, which gives the measurability and the last assertion. □

Thus a set of positive measure on which the rank is at least k always carries k functions whose gradients are independent on a subset of positive measure. In Chapter 5 this observation replaces any assumption that the rank is locally constant. All pointwise assertions on the rank are understood ν-almost everywhere, not m-almost everywhere.

Change of frame. The space $\mathbb{R}^N$ and the embedding of the fibres into it are auxiliary. Let $\mathsf{H}' = (H'_x)$ be another measurable field over $\mathcal{S}$, possibly in another $\mathbb{R}^{N'}$, and let $O_x : H_x \to H'_x$ be surjective linear isometries such that $x \mapsto O_x\pi_x$ has Borel matrix entries. Then $\nabla' u(x) = O_x \nabla u(x)$ satisfies (G): the Leibniz rule and the Lipschitz bound are clear, and the weak-$*$ continuity follows from $\int\langle \nabla' u_k, g\rangle\, d\nu = \int\langle \nabla u_k, O^* g\rangle\, d\nu$ for $g \in L^1(\mathcal{S}, \mathsf{H}', \nu)$. The images of a spanning sequence span the new fibres, so H' is the field generated by ∇'. We identify two structures related by such a change of frame. All quantities considered in this book, such as $|\nabla u|$, $\langle \nabla u, \nabla v\rangle$ and $n(x)$, are invariant. For a structure of constant rank n, a measurable orthonormal frame of H identifies the fibres with $\mathbb{R}^n$. When $n(x) = 1$ almost everywhere, a change of frame is a measurable choice of sign, that is, an orientation of the skeleton.

1.6 Cores, cutoffs and spanning

The gradient of a bounded Lipschitz function is bounded, but it need not be integrable to any power, because ν may be infinite on balls. The functions whose gradient is p-integrable form the natural domain on which closures will be taken in Chapter 2.

Definition 1.8. For $1 \le p < \infty$, the *core at the exponent p* is

$$\mathcal{C}_p = \big\{u \in \mathrm{Lip}_c(X) :\ \nabla u \in L^p(\mathcal{S}, \mathsf{H}, \nu)\big\}.$$

We say that $\mathcal{C}_p$ *spans the fibres* if the measurable field generated by $\{\nabla u : u \in \mathcal{C}_p\}$ is H.

Since ∇u is bounded, $\int_{\mathcal{S}} |\nabla u|^q \, d\nu \le \|\nabla u\|_\infty^{q-p} \int_{\mathcal{S}} |\nabla u|^p \, d\nu$ for $p \le q$, hence

$$\mathcal{C}_1 \subset \mathcal{C}_p \subset \mathcal{C}_q \qquad (1 \le p \le q < \infty) \,. \tag{1.9}$$

If ν is finite on compact sets, then $\mathcal{C}_p = \mathrm{Lip}_c(X)$ for every p. By Lemma 1.5 and (1.4), each $\mathcal{C}_p$ is an algebra, and it is stable under $u \mapsto \varphi \circ u$ for $\varphi \in C^1(\mathbb{R})$ with $\varphi(0) = 0$, under the truncations $u \mapsto (u \vee a) \wedge b$ with $a \le 0 \le b$, and under $(u, v) \mapsto u \vee v$ and $(u, v) \mapsto u \wedge v$.

In general the core may be small, and its gradients may fail to span the fibres; an example is given in Example 2.4. The next proposition shows that both defects are excluded as soon as balls admit cutoff functions of finite energy. The hypothesis is qualitative: the energy of the cutoffs is finite, but no bound on it is required.

Proposition 1.9. *Let $1 \le b < \infty$ and $C \ge 1$. Suppose that for every compact set $K \subset X$ there is $r_K > 0$ such that, for every $x \in K$ and $0 < r < r_K$, there is $\chi \in \mathrm{Lip}_c(X)$ with*

$$\mathbf{1}_{B(x,r)} \le \chi \le \mathbf{1}_{B(x,2r)} \,, \qquad \mathrm{Lip}(\chi) \le C/r \,, \qquad \int_{\mathcal{S}} |\nabla \chi|^b \, d\nu < \infty \,.$$

Then, for every $p \in [b, \infty)$,

(i) *$\mathcal{C}_p$ is dense in $C_c(X)$ for the uniform norm, and*

(ii) *$\mathcal{C}_p$ spans the fibres: there is a sequence in $\mathcal{C}_p$ whose gradients span H_x for ν-almost every x.*

Proof. By (1.9) it suffices to treat $p = b$.

Step 1: approximation of a Lipschitz function. Let $f \in \mathrm{Lip}_c(X)$ be nonnegative and not identically zero, let $L > 0$ be its Lipschitz constant, and let $K = \operatorname{supp} f$. We claim that for every sufficiently small $t > 0$ there is $f_t \in \mathcal{C}_b$ with

$$\|f_t - f\|_\infty \le t \,, \qquad \mathrm{Lip}(f_t) \le 8CL \,. \tag{1.10}$$

Let $r = t/(8L)$ and assume $r < r_K$. For each integer $k \ge 1$ with $\{f \ge kt\} \ne \varnothing$, the set $\{f \ge kt\}$ is a compact subset of K. Cover it by finitely many balls of radius r centred on it, and let χ_k be the maximum of the corresponding cutoffs. By (1.4), $\chi_k \in \mathcal{C}_b$. Moreover $\mathrm{Lip}(\chi_k) \le C/r$, $\chi_k = 1$ on $\{f \ge kt\}$, and χ_k vanishes outside the $2r$-neighbourhood of $\{f \ge kt\}$. Since $2rL = t/4$,

$$\chi_k = 1 \text{ on } \{f \ge kt\} \,, \qquad \chi_k = 0 \text{ on } \{f \le kt - t/4\} \,.$$

Only finitely many k occur, and we put $f_t = t \sum_k \chi_k \in \mathcal{C}_b$.

Let $x \in X$ and let $j \ge 0$ be the integer with $jt \le f(x) < (j+1)t$. Then $\chi_k(x) = 1$ for $k \le j$ and $\chi_k(x) = 0$ for $k \ge j + 2$, so $jt \le f_t(x) \le (j+1)t$ and $|f_t(x) - f(x)| \le t$. Let now $x, y \in X$. If $|f(x) - f(y)| \ge 3t/4$, then

$$|f_t(x) - f_t(y)| \le |f(x) - f(y)| + 2t \le \tfrac{11}{3} |f(x) - f(y)| \le \tfrac{11}{3} L \, d(x, y) \,.$$

If $|f(x) - f(y)| < 3t/4$ and $f(x) \le f(y)$, with $jt \le f(x) < (j+1)t$, then $f(y) < (j+2)t - t/4$. Hence $\chi_k(x) = \chi_k(y) = 1$ for $k \le j$ and $\chi_k(x) = \chi_k(y) = 0$ for $k \ge j + 2$, and only the term $k = j + 1$ can differ:

$$|f_t(x) - f_t(y)| = t \, |\chi_{j+1}(x) - \chi_{j+1}(y)| \le t \, \frac{C}{r} \, d(x, y) = 8CL \, d(x, y) \,.$$

This proves (1.10), since $C \ge 1$.

Step 2: density. By Step 1, applied to the positive and negative parts, every function of $\mathrm{Lip}_c(X)$ is a uniform limit of functions of $\mathcal{C}_b$. Since $\mathrm{Lip}_c(X)$ is uniformly dense in $C_c(X)$, this proves (i).

Step 3: spanning. Let H^b be the field generated by the gradients of $\mathcal{C}_b$, with projections π^b_x. The set of $V \in L^\infty(\mathcal{S}, \mathsf{H}, \nu)$ with $V(x) \in H^b_x$ almost everywhere is the kernel of the weak-$*$ continuous map $V \mapsto V - \pi^b V$, hence it is weak-$*$ closed. Let $f \in \mathrm{Lip}_c(X)$ be nonnegative and let $f_t \in \mathcal{C}_b$ be as in Step 1. As $t \to 0$, $f_t \to f$ uniformly with bounded Lipschitz constants, so $\nabla f_t \to \nabla f$ weak-$*$ by (G). Hence $\nabla f(x) \in H^b_x$ almost everywhere, and the same holds for every $f \in \mathrm{Lip}_c(X)$ by linearity. By Lemma 1.6(iii), $\mathsf{H}^b = \mathsf{H}$. The spanning sequence is given by Lemma 1.6(ii). □

1.7 Recovery of the metric

Condition (G) bounds the gradient by the Lipschitz constant. Example 1.4 shows that the converse may fail. The following condition asserts the converse locally.

Definition 1.10. A gradient structure has the *local Sobolev-to-Lipschitz property* (SL) if for every $x \in X$ there is a radius $\iota(x) > 0$ such that, for all $u \in \mathrm{Lip}_b(X)$ and all $y, z \in B(x, \iota(x))$,

$$|u(y) - u(z)| \le d(y,z)\, \|\nabla u\|_{L^\infty(B(x,3\iota(x))\cap\mathcal{S},\nu)}\,. \tag{1.11}$$

The structure is *uniform* if the radii in (1.11) can be chosen with $\inf_{x\in X} \iota(x) > 0$.

The name refers to the Sobolev-to-Lipschitz property of metric measure spaces, which asserts that a Sobolev function with gradient bounded by one has a 1-Lipschitz representative; see [56, Definition 1.6 and Theorem 1.7] for its formulation with minimal weak upper gradients. Here the property is imposed on Lipschitz functions only, and the gradient is measured by ν. Both properties in Definition 1.10 are invariant under a change of frame. In the model of Section 1.1, (1.11) holds with $\iota \equiv 1$, for every reference measure m.

Proposition 1.11. *If* (1.11) *holds and* $y, z \in B(x, \iota(x))$, *then*

$$d(y,z) = \max\{u(y) - u(z) :\ u \in \mathrm{Lip}_b(X),\ |\nabla u| \le 1\ \nu\text{-almost everywhere on } B(x, 3\iota(x))\}\,.$$

Proof. Let $u \in \mathrm{Lip}_b(X)$ with $|\nabla u| \le 1$ ν-almost everywhere on $B(x, 3\iota(x))$. By (1.11), $u(y) - u(z) \le d(y,z)$. Conversely, the function $u = d(\cdot, z) \wedge 2\iota(x)$ belongs to $\mathrm{Lip}_b(X)$ and has $\mathrm{Lip}(u) \le 1$, hence $|\nabla u| \le 1$ by (G). Moreover $u(y) - u(z) = d(y,z)$, because $d(y,z) < 2\iota(x)$. □

Comparison with the intrinsic metric. The formula of Proposition 1.11 should not be confused with the intrinsic metric of a strongly local Dirichlet form [150], which is defined by the constraint that the energy measure of u is bounded by the reference measure. In the notation of Chapter 2 this constraint reads $|\nabla_2 u|^2 \nu \le m$. When $\nu \perp m$ it forces $\nabla_2 u = 0$, and the intrinsic metric carries no information. In Proposition 1.11 the constraint $|\nabla u| \le 1$ is imposed ν-almost everywhere, and the reference measure plays no role.

Remark 1.12. The axioms of a gradient structure and the Sobolev-to-Lipschitz property concern the metric and the gradient measure; they do not involve the reference measure. In particular they do not imply that the gradient is closable in $L^p(X, m)$. Admissibility, introduced in Definition 2.1, is an additional property, which relates the two measures. The examples of Part III verify it directly, or by the criterion of Theorem 4.3.

2 Sobolev spaces, bounded variation and energy forms

A gradient structure defines the gradient of a Lipschitz function. This chapter extends it to larger classes of functions, in two ways. For an exponent $1 \le p < \infty$ at which the gradient is closable in $L^p(X,m)$, the domain of its closure is the Sobolev space $W^{1,p}(X,m)$. For $p=1$, where closability may fail and the closure is in any case too small, the lower semicontinuous envelope of the energy defines the total variation and the space $BV(X,m)$.

No volume, cutoff or Poincaré estimate is used in this chapter. Closability is a hypothesis; Chapter 4 proves it under quantitative assumptions. The main results are the following.

- *A trace on the skeleton* (Proposition 2.3). A Sobolev function is an m-equivalence class and has, a priori, no values on the skeleton when $m(\mathcal{S})=0$. We show that it has a well-defined ν-measurable trace on the part of the skeleton which carries the gradients of the core, and that the chain rule holds with this trace.
- *Energy forms* (Propositions 2.5, 2.6 and 2.7). For $p>1$ the space $W^{1,p}(X,m)$ is reflexive and $(\mathcal{E}_p, W^{1,p}(X,m))$ is a regular p-energy form with the generalized contraction property of Kajino and Shimizu [98]. For $p=2$ it is a strongly local regular Dirichlet form, whose energy measures are $|\nabla_2 f|^2\nu$ and whose index in the sense of Hino [85] is the rank of the gradient structure.
- *Bounded variation* (Proposition 2.10 and Theorem 2.11). The total variation is the support function of the divergences of bounded vector fields, and it satisfies the coarea formula.

2.1 Admissible exponents and Sobolev spaces

Let $(\mathcal{S},\nu,\nabla)$ be a gradient structure on (X,d,m). Recall from Definition 1.8 the core

$$\mathcal{C}_p = \{u \in \mathrm{Lip}_c(X) :\ \nabla u \in L^p(\mathcal{S},\mathsf{H},\nu)\}\,, \qquad 1 \le p < \infty\,.$$

The Sobolev space at the exponent p is the domain of the closure of the gradient on this core.

Definition 2.1. Let $1 \le p < \infty$. A gradient structure $(\mathcal{S},\nu,\nabla)$ is *admissible at the exponent* p, or *p-admissible*, if

(A1) $\mathcal{C}_p$ is dense in $C_c(X)$ for the uniform norm, and

(A2) the operator $\nabla : \mathcal{C}_p \subset L^p(X,m) \to L^p(\mathcal{S},\mathsf{H},\nu)$ is closable.

In this case ∇_p denotes the closure of this operator, and we set

$$W^{1,p}(X,m) = \mathrm{Dom}(\nabla_p)\,, \qquad \|f\|^p_{W^{1,p}} = \|f\|^p_{L^p(X,m)} + \|\nabla_p f\|^p_{L^p(\mathcal{S},\nu)}\,, \qquad \mathcal{E}_p(f) = \int_{\mathcal{S}} |\nabla_p f|^p \, d\nu\,,$$

with $\mathcal{E}_p(f) = +\infty$ for $f \in L^p(X,m) \setminus W^{1,p}(X,m)$. We call ∇_p the *closed gradient*, $\mathcal{E}_p$ the *p-energy*, and $\|\cdot\|_{W^{1,p}}$ the *graph norm*.

Explicitly, $f \in W^{1,p}(X,m)$ and $\nabla_p f = G$ if and only if there are $u_k \in \mathcal{C}_p$ with $u_k \to f$ in $L^p(X,m)$ and $\nabla u_k \to G$ in $L^p(\mathcal{S},\mathsf{H},\nu)$. Condition (A2) says that G is determined by f:

$$u_k \in \mathcal{C}_p\,, \qquad u_k \to 0 \text{ in } L^p(X,m)\,, \qquad \nabla u_k \to G \text{ in } L^p(\mathcal{S},\mathsf{H},\nu) \qquad \Longrightarrow \qquad G = 0\,. \tag{2.1}$$

A function $u \in \mathcal{C}_p$ is identified with its m-equivalence class; since u is continuous and m has full support, the class determines u, so that $\mathcal{C}_p \subset W^{1,p}(X,m)$ and $\nabla_p u = \nabla u$ for $u \in \mathcal{C}_p$.

Why closability is a condition. On $\mathbb{R}^n$ with $m = \nu = \mathcal{L}^n$, closability follows from integration by parts against smooth vector fields. In general no such fields are given. When $m(\mathcal{S}) = 0$ the difficulty is visible directly: convergence in $L^p(X, m)$ gives no information on the values of u_k on $\mathcal{S}$, so (2.1) asserts that a sequence cannot oscillate on the skeleton with bounded energy while converging to zero m-almost everywhere. Two general criteria for (A2) are available in this book. Proposition 2.5(iv) restates it as the existence of sufficiently many vector fields on $\mathcal{S}$ whose divergence is a function on X. Theorem 4.3 derives (A1) and (A2) from volume, cutoff and Poincaré hypotheses.

Dependence on the exponent. Admissibility depends on p. For the tree of cubes of Chapter 15, the admissible exponents are precisely those at or above the threshold of Theorem 15.8. When several exponents are admissible, the closed gradients ∇_p and ∇_q need not agree a priori on $W^{1,p}(X, m) \cap W^{1,q}(X, m)$. Proposition 4.7 proves that they agree under Poincaré inequalities.

Relation to other definitions. Sobolev spaces defined as the closure of a derivation associated with a Dirichlet form are studied in [90, Section 6] and [89, Section 2 and Appendix B]. Kuwae [119, Assumption 1 and Definition 1.3] distinguishes the reference measure from a minimal energy-dominant measure and assumes closability, as we do. Definition 2.1 is not, without further comparison, a definition by upper gradients as in [145], nor is $W^{1,p}(X, m)$ identified with the spaces defined by relaxed slopes in [49, 9]. For the examples with $\nu = m$ of Part III, the identification is made in the corresponding chapters.

Cutoffs in the core. Condition (A1) provides cutoff functions in $\mathcal{C}_p$ and implies that ∇_p is densely defined. By Proposition 1.9, it holds as soon as small balls admit cutoff functions of finite p-energy.

Lemma 2.2. *Let $1 \le p < \infty$ and assume (A1). Let $K \subset U \subset X$ with K compact and U open.*

(i) *There is $\chi \in \mathcal{C}_p$ with $0 \le \chi \le 1$, $\chi = 1$ on a neighbourhood of K, and $\operatorname{supp}\chi \subset U$.*

(ii) *$\mathcal{C}_p$ is dense in $L^p(X, m)$.*

Proof. (i) Since X is locally compact, there are a compact neighbourhood K' of K and $g \in C_c(X)$ with $0 \le g \le 2$, $g = 2$ on K' and $\operatorname{supp} g \subset U$. Choose $v \in \mathcal{C}_p$ with $\|v - g\|_\infty \le 1/2$, and put

$$\chi = \big((v \vee \tfrac{1}{2}) \wedge \tfrac{3}{2}\big) - \tfrac{1}{2}\,.$$

Then χ is Lipschitz, $0 \le \chi \le 1$, and $\chi = 1$ on $\{v \ge \frac{3}{2}\} \supset \{g = 2\} \supset K'$. Moreover $\{\chi \ne 0\} = \{v > \frac{1}{2}\} \subset \{g > 0\}$, so $\operatorname{supp}\chi \subset \operatorname{supp} g$ is a compact subset of U. By Lemma 1.5(i),(iv), $\nabla\chi = \mathbf{1}_{\{1/2<v<3/2\}}\nabla v \in L^p(\mathcal{S}, \mathsf{H}, \nu)$. Hence $\chi \in \mathcal{C}_p$.

(ii) The space $C_c(X)$ is dense in $L^p(X, m)$. Let $f \in C_c(X)$, let $\chi \in \mathcal{C}_p$ be given by (i) for $K = \operatorname{supp} f$ and $U = X$, and let $u_k \in \mathcal{C}_p$ converge uniformly to f. Then $\chi u_k \in \mathcal{C}_p$ converges uniformly to $\chi f = f$, with supports in the compact set $\operatorname{supp}\chi$, which has finite m-measure. Hence $\chi u_k \to f$ in $L^p(X, m)$. □

2.2 The trace on the skeleton and the chain rule

Let the structure be p-admissible. A function $f \in W^{1,p}(X, m)$ is an m-equivalence class, while its gradient $\nabla_p f$ is a ν-equivalence class of fields over $\mathcal{S}$. To state a chain rule $\nabla_p(\varphi \circ f) = \varphi'(f)\nabla_p f$ for a Lipschitz function φ, one needs the values of f at ν-almost every point where $\nabla_p f \ne 0$.

When $\nu \perp m$ these values are not provided by the class of f. They are provided by the following construction.

The carrier of the core. By Lemma 1.6, there is a sequence (v_k) in $\mathcal{C}_p$ whose gradients span, at ν-almost every point, the fibre of the field generated by $\{\nabla u : u \in \mathcal{C}_p\}$. Put

$$\mathcal{S}_p = \bigcup_{k\geq 1} \{x \in \mathcal{S} : |\nabla v_k(x)| > 0\}\,. \tag{2.2}$$

Up to ν-null sets, $\mathcal{S}_p$ is the smallest Borel subset of $\mathcal{S}$ outside which every gradient of $\mathcal{C}_p$ vanishes. In particular $\mathcal{S}_p = \mathcal{S}$ up to a ν-null set if $\mathcal{C}_p$ spans the fibres, which is the case under the hypothesis of Proposition 1.9 with $b \leq p$. If $u_k \in \mathcal{C}_p$ and $\nabla u_k \to G$ in $L^p(\nu)$, then $G = 0$ ν-almost everywhere outside $\mathcal{S}_p$. Hence

$$\nabla_p f = 0 \quad \nu\text{-almost everywhere on } \mathcal{S} \setminus \mathcal{S}_p \qquad (f \in W^{1,p}(X,m))\,. \tag{2.3}$$

Convergence in measure. Let $L^0(\mathcal{S}_p, \nu)$ be the space of ν-equivalence classes of Borel functions $\mathcal{S}_p \to \mathbb{R}$, with the topology of convergence in measure on sets of finite ν-measure. If μ is a finite measure with the same null sets as $\nu \restriction_{\mathcal{S}_p}$, this topology is defined by the complete translation-invariant metric $\int \min\{1, |f-g|\}\, d\mu$; indeed, a sequence converges for either notion if and only if every subsequence has a further subsequence which converges ν-almost everywhere on $\mathcal{S}_p$.

Proposition 2.3. *Let $1 \leq p < \infty$, and suppose that the gradient structure is p-admissible.*

(i) Trace. *The restriction map $\mathcal{C}_p \to L^0(\mathcal{S}_p, \nu)$, $u \mapsto u|_{\mathcal{S}_p}$, extends uniquely to a continuous linear map*

$$T_p : W^{1,p}(X,m) \longrightarrow L^0(\mathcal{S}_p, \nu)\,.$$

(ii) Image of the energy. *For every $f \in W^{1,p}(X,m)$, the measure $|\nabla_p f|^p \nu$ is carried by $\mathcal{S}_p$, and*

$$(T_p f)_{\#}(|\nabla_p f|^p \nu) \ll \mathcal{L}^1\,.$$

(iii) Chain rule. *If $f \in W^{1,p}(X,m)$ and $\varphi : \mathbb{R} \to \mathbb{R}$ is Lipschitz with $\varphi(0) = 0$, then $\varphi \circ f \in W^{1,p}(X,m)$ and*

$$T_p(\varphi \circ f) = \varphi(T_p f)\,, \qquad \nabla_p(\varphi \circ f) = \varphi'(T_p f)\, \nabla_p f \qquad \nu\text{-almost everywhere on } \mathcal{S}_p\,, \tag{2.4}$$

and $\nabla_p(\varphi \circ f) = 0$ ν-almost everywhere outside $\mathcal{S}_p$.

In particular, the trace is defined on the whole skeleton if $\mathcal{C}_p$ spans the fibres.

In (2.4) the derivative φ' is defined outside a Lebesgue null set $Z \subset \mathbb{R}$. By (ii), $\nabla_p f = 0$ ν-almost everywhere on $\{T_p f \in Z\}$, and the product $\varphi'(T_p f)\nabla_p f$ is set equal to zero there. Assertion (ii) is the analogue, for the closed gradient, of the energy image density property of a single function in a Dirichlet space [53, Theorem 4.3.8]. Applied to a one-point set it gives the locality of the closed gradient:

$$\nabla_p f = 0 \qquad \nu\text{-almost everywhere on } \{T_p f = c\}\,, \qquad c \in \mathbb{R}\,. \tag{2.5}$$

Proof. The measure $|\nabla_p f|^p \nu$ is carried by $\mathcal{S}_p$ by (2.3).

Step 1: restriction is continuous on the core. Let $u_j \in \mathcal{C}_p$ tend to zero in the graph norm; we show that $u_j|_{\mathcal{S}_p} \to 0$ in $L^0(\mathcal{S}_p, \nu)$. Put $\theta(t) = t^2/(1+t^2)$, so that $0 \le \theta \le 1$, $\theta(t) \le |t|$, $|\theta'| \le 1$ and $\theta(0) = 0$. Fix $v \in \mathcal{C}_p$ and let $\mu_v = |\nabla v|^p \nu$, a finite measure. We claim that

$$\int_{\mathcal{S}} \theta(u_j)\, d\mu_v \longrightarrow 0\,. \tag{2.6}$$

The sequence $\theta(u_j)$ is bounded in $L^\infty(\mu_v)$, and $L^1(\mu_v)$ is separable. It therefore suffices to show that if a subsequence, still denoted by $\theta(u_j)$, converges weak-$*$ in $L^\infty(\mu_v)$ to some h, then $h = 0$; the claim then follows by testing against the constant function $1 \in L^1(\mu_v)$.

Let $V \in L^{p'}(\mathcal{S}, \mathsf{H}, \nu)$. The function $\rho = \langle V, \nabla v\rangle |\nabla v|^{-p} \mathbf{1}_{\{\nabla v \neq 0\}}$ satisfies $\langle V, \nabla v\rangle \nu = \rho\, \mu_v$ and $\int |\rho|\, d\mu_v \le \|V\|_{p'} \|\nabla v\|_p < \infty$. Hence

$$\int_{\mathcal{S}} \langle V, \theta(u_j)\nabla v\rangle\, d\nu = \int_{\mathcal{S}} \theta(u_j)\, \rho\, d\mu_v \longrightarrow \int_{\mathcal{S}} h\, \rho\, d\mu_v = \int_{\mathcal{S}} \langle V, h\nabla v\rangle\, d\nu\,,$$

that is, $\theta(u_j)\nabla v \rightharpoonup h\nabla v$ weakly in $L^p(\mathcal{S}, \mathsf{H}, \nu)$. On the other hand $v\,\theta(u_j) \in \mathcal{C}_p$, and by Lemma 1.5,

$$|v\,\theta(u_j)| \le \|v\|_\infty |u_j|\,, \qquad \nabla(v\,\theta(u_j)) = \theta(u_j)\nabla v + v\,\theta'(u_j)\nabla u_j\,.$$

Thus $v\,\theta(u_j) \to 0$ in $L^p(X, m)$, and the last term tends to zero in $L^p(\nu)$, because $|v\,\theta'(u_j)\nabla u_j| \le \|v\|_\infty |\nabla u_j|$. The pairs $(v\,\theta(u_j), \nabla(v\,\theta(u_j)))$ therefore converge weakly to $(0, h\nabla v)$ in $L^p(X, m) \oplus L^p(\mathcal{S}, \mathsf{H}, \nu)$. The graph of ∇_p is a closed linear subspace, hence weakly closed, so $h\nabla v = \nabla_p 0 = 0$. Thus $h = 0$ μ_v-almost everywhere, and (2.6) is proved.

Since $\inf_{|t|\ge\varepsilon} \theta(t) > 0$ for every $\varepsilon > 0$, (2.6) shows that $u_j \to 0$ in μ_v-measure. Apply this to the functions v_k of (2.2). The finite measure

$$\mu = \sum_{k\ge 1} 2^{-k} \frac{\mu_{v_k}}{1 + \mu_{v_k}(\mathcal{S})}$$

has the same null sets as $\nu\restriction_{\mathcal{S}_p}$, and $u_j \to 0$ in μ-measure, because $\mu(\{|u_j| > \varepsilon\}) \le \sum_{k\le k_0} \mu_{v_k}(\{|u_j| > \varepsilon\}) + 2^{-k_0}$ for every k_0. This is the convergence in $L^0(\mathcal{S}_p, \nu)$.

Step 2: extension. By Step 1, restriction is a linear map from $\mathcal{C}_p$, with the graph norm, to the complete metrizable topological vector space $L^0(\mathcal{S}_p, \nu)$, and it is continuous at zero. It is therefore uniformly continuous, and it extends uniquely by continuity to the closure $W^{1,p}(X, m)$ of $\mathcal{C}_p$. This proves (i).

Step 3: the chain rule for $\varphi \in C^1(\mathbb{R})$ with bounded derivative and $\varphi(0) = 0$. Let $u_j \in \mathcal{C}_p$ converge to f in the graph norm. Then $\varphi \circ u_j \in \mathcal{C}_p$, $\varphi \circ u_j \to \varphi \circ f$ in $L^p(X, m)$, and $\nabla(\varphi \circ u_j) = \varphi'(u_j)\nabla u_j$. We have

$$\|\varphi'(u_j)\nabla u_j - \varphi'(T_p f)\nabla_p f\|_{L^p(\nu)} \le \|\varphi'\|_\infty \|\nabla u_j - \nabla_p f\|_{L^p(\nu)} + \|(\varphi'(u_j) - \varphi'(T_p f))\nabla_p f\|_{L^p(\nu)}\,.$$

The first term tends to zero. For the second, $u_j \to T_p f$ in $L^0(\mathcal{S}_p, \nu)$, hence in measure for the finite measure $|\nabla_p f|^p \nu$, which is absolutely continuous with respect to $\nu\restriction_{\mathcal{S}_p}$. Since φ' is continuous and bounded, the second term tends to zero by dominated convergence. As ∇_p is closed, $\varphi \circ f \in W^{1,p}(X, m)$ and $\nabla_p(\varphi \circ f) = \varphi'(T_p f)\nabla_p f$. Moreover $\varphi \circ u_j \to \varphi \circ f$ in the graph norm, so $T_p(\varphi \circ f) = \lim_j \varphi(u_j)|_{\mathcal{S}_p} = \varphi(T_p f)$ by the continuity of T_p and of φ.

Step 4: the image of the energy. Let $K \subset \mathbb{R}$ be a compact set with $\mathcal{L}^1(K) = 0$. Choose open sets $U_j \supset K$ decreasing to K with $\mathcal{L}^1(U_j) \to 0$, and smooth functions $0 \le \eta_j \le 1$ with $\eta_j = 1$

on K and $\operatorname{supp}\eta_j \subset U_j$. The functions $\varphi_j(t) = \int_0^t \eta_j(s)\,ds$ satisfy $|\varphi_j(t)| \le |t| \wedge \mathcal{L}^1(U_j)$. Hence $\varphi_j \circ f \to 0$ in $L^p(X,m)$ by dominated convergence, and by Step 3,

$$\nabla_p(\varphi_j \circ f) = \eta_j(T_p f)\nabla_p f \longrightarrow \mathbf{1}_{\{T_p f \in K\}}\nabla_p f \qquad \text{in } L^p(\mathcal{S},\mathsf{H},\nu)\,.$$

Since ∇_p is closed, the limit is zero, that is, $|\nabla_p f|^p\nu(\{T_p f \in K\}) = 0$. A finite Borel measure on $\mathbb{R}$ is inner regular, so the image measure vanishes on every Lebesgue null Borel set. This proves (ii).

Step 5: the chain rule for Lipschitz φ. Let (ρ_j) be a smooth approximation of the identity on $\mathbb{R}$ and $\varphi_j = \varphi * \rho_j - (\varphi*\rho_j)(0)$. Then $\varphi_j \in C^1(\mathbb{R})$, $\varphi_j(0) = 0$, $|\varphi_j'| \le \operatorname{Lip}(\varphi)$, $\varphi_j \to \varphi$ uniformly, and $\varphi_j' \to \varphi'$ Lebesgue-almost everywhere. By (ii), $\varphi_j'(T_p f) \to \varphi'(T_p f)$ almost everywhere for $|\nabla_p f|^p\nu$, and dominated convergence gives $\varphi_j'(T_p f)\nabla_p f \to \varphi'(T_p f)\nabla_p f$ in $L^p(\nu)$. Since $|\varphi_j \circ f| \le \operatorname{Lip}(\varphi)|f|$, also $\varphi_j \circ f \to \varphi \circ f$ in $L^p(X,m)$. By Step 3 and the closedness of ∇_p, $\varphi \circ f \in W^{1,p}(X,m)$ with the gradient (2.4), and $\varphi_j \circ f \to \varphi \circ f$ in the graph norm. The continuity of T_p gives $T_p(\varphi \circ f) = \lim_j \varphi_j(T_p f) = \varphi(T_p f)$. The last assertion of (iii) is (2.3). □

At $p = 2$ the trace has a classical description: $T_2 f$ is the restriction to $\mathcal{S}_2$ of the quasi-continuous representative of f (Remark 2.8). Section 9.2 gives quantitative L^p bounds for the trace with respect to the energy measures of cutoff functions.

The restriction to $\mathcal{S}_p$ in Proposition 2.3 cannot be removed. The following example shows that, without a spanning hypothesis, a structure can be admissible for a trivial reason.

Example 2.4 (An admissible structure with zero closed gradient)**.** Let $E \subset (0,1)$ be a compact nowhere dense set of positive Lebesgue measure. On $X = [0,1]^2$ take $m = \mathcal{L}^2$, $\mathcal{S} = E \times (\mathbb{Q} \cap [0,1])$ and

$$\nu = \sum_{q \in \mathbb{Q}\cap[0,1]} (\mathcal{L}^1 \restriction_E) \otimes \delta_q\,, \qquad \nabla u(x,q) = \partial_x u(x,q) \quad (x \in E)\,.$$

This is a gradient structure of constant rank one. The Leibniz rule and the bound $|\partial_x u| \le \operatorname{Lip}(u)$ hold on each horizontal line. If $u_k \to u$ pointwise with bounded Lipschitz constants, then on each line $\partial_x u_k(\cdot,q) \to \partial_x u(\cdot,q)$ weak-$*$ in $L^\infty(0,1)$, by integration by parts against smooth functions and density, as in Example 1.3. For $g \in L^1(\nu)$ the norms $\|g(\cdot,q)\|_{L^1(E)}$ are summable in q, and dominated convergence for the sum over q gives the weak-$*$ continuity in (G). The coordinate function $(x,y) \mapsto x$ has gradient one on the skeleton.

Every function of the core has zero gradient. Let $1 \le p < \infty$ and $u \in \mathcal{C}_p$, so that $\sum_q \|\partial_x u(\cdot,q)\|^p_{L^p(E)} < \infty$. Let $q \in \mathbb{Q} \cap [0,1]$ and let $q_j \to q$ be distinct rationals; then $\|\partial_x u(\cdot,q_j)\|_{L^p(E)} \to 0$. The slices $u(\cdot,q_j)$ converge uniformly to $u(\cdot,q)$ with bounded Lipschitz constants, so their derivatives converge weak-$*$ in $L^\infty(0,1)$ to $\partial_x u(\cdot,q)$. Hence $\partial_x u(\cdot,q) = 0$ almost everywhere on E. Thus $\nabla u = 0$ for every $u \in \mathcal{C}_p$, and $\mathcal{S}_p = \varnothing$.

The structure is p-admissible. Since ∇ vanishes on $\mathcal{C}_p$, condition (A2) holds. For (A1), let $\varepsilon > 0$ and partition $[0,1]$ into finitely many intervals I of length at most ε. Since E is nowhere dense, $\mathcal{L}^1(I \setminus E) > 0$. Let $\tau : [0,1] \to [0,1]$ be the Lipschitz map which fixes the endpoints of each I and has derivative $|I|\,\mathbf{1}_{I\setminus E}/\mathcal{L}^1(I \setminus E)$ on I. Then $\|\tau - \mathrm{id}\|_\infty \le \varepsilon$ and $\tau' = 0$ almost everywhere on E. For a Lipschitz function u on X, the function $v(x,y) = u(\tau(x),y)$ is Lipschitz, $\|v-u\|_\infty \le \operatorname{Lip}(u)\,\varepsilon$, and

$$|v(x+h,y) - v(x,y)| \le \operatorname{Lip}(u)\,|\tau(x+h) - \tau(x)| = o(h) \qquad \text{at every point } x \text{ with } \tau'(x) = 0\,.$$

Hence $\partial_x v(\cdot,y) = 0$ almost everywhere on E for every y, and $v \in \mathcal{C}_p$. This proves (A1).

Conclusions. The closure of ∇ on $\mathcal{C}_p$ is the zero operator on $W^{1,p}(X,m) = L^p(X,m)$, at every finite p, although the fibres of the structure are not zero: the core does not span the fibres.

Moreover, restriction to the whole skeleton admits no continuous extension to $W^{1,p}(X,m)$. Indeed, for a fixed rational q, the functions $(x,y)\mapsto (1-j|y-q|)_+$ belong to $\mathcal{C}_p$, tend to zero in the graph norm, and are equal to one on the set $E\times\{q\}$, which has positive finite ν-measure.

2.3 Functional-analytic properties

The next proposition collects the properties of $W^{1,p}(X,m)$ for $p>1$. The closed-graph and measurable-field arguments are those of [89, Section 3 and Appendix A]; compare also the uniform convexity and contraction results of [119, Theorems 1.8–1.9].

Proposition 2.5. *Let $1<p<\infty$, and let $(\mathcal{S},\nu,\nabla)$ be p-admissible.*

(i) *$W^{1,p}(X,m)$ is a separable, uniformly convex Banach space; in particular it is reflexive.*

(ii) *$\mathcal{E}_p$ is lower semicontinuous on $L^p(X,m)$.*

(iii) *If $f\in W^{1,p}(X,m)$ and $\varphi:\mathbb{R}\to\mathbb{R}$ is Lipschitz with $\varphi(0)=0$, then $\varphi\circ f\in W^{1,p}(X,m)$ and $|\nabla_p(\varphi\circ f)|\le \mathrm{Lip}(\varphi)|\nabla_p f|$ ν-almost everywhere.*

(iv) *Let $\mathcal{D}_{p'}$ be the set of $V\in L^{p'}(\mathcal{S},\mathsf{H},\nu)$ for which there is $h\in L^{p'}(X,m)$ with*

$$\int_{\mathcal{S}}\langle V,\nabla u\rangle\,d\nu=-\int_X u\,h\,dm\qquad\text{for all } u\in\mathcal{C}_p\,. \tag{2.7}$$

Under (A1), condition (A2) holds if and only if $\mathcal{D}_{p'}$ is dense in $L^{p'}(\mathcal{S},\mathsf{H},\nu)$.

Proof. (i) The map $f\mapsto(f,\nabla_p f)$ is an isometry of $W^{1,p}(X,m)$ onto the graph of ∇_p, which is a closed subspace of $L^p(X,m)\oplus L^p(\mathcal{S},\mathsf{H},\nu)$ with the norm $(\|f\|_p^p+\|G\|_p^p)^{1/p}$. This direct sum is a closed subspace of the space of p-integrable functions with values in the Euclidean space $\mathbb{R}^{N+1}$ over the disjoint union of (X,m) and $(\mathcal{S},\nu)$. Such a space is uniformly convex for $1<p<\infty$ [58], and it is separable because both measures are σ-finite Borel measures on a separable metric space. Closed subspaces inherit both properties, and uniformly convex Banach spaces are reflexive.

(ii) Let $f_k\to f$ in $L^p(X,m)$ with $\ell=\liminf_k \mathcal{E}_p(f_k)<\infty$. After passing to a subsequence, $\mathcal{E}_p(f_k)\to\ell$ and (f_k) is bounded in $W^{1,p}(X,m)$. By reflexivity, a further subsequence of $(f_k,\nabla_p f_k)$ converges weakly in $L^p(X,m)\oplus L^p(\mathcal{S},\mathsf{H},\nu)$. The graph of ∇_p is closed and convex, hence weakly closed, and the first component of the weak limit is f. Thus $f\in W^{1,p}(X,m)$ and $\nabla_p f_k\rightharpoonup\nabla_p f$ weakly in $L^p(\mathcal{S},\mathsf{H},\nu)$, and $\mathcal{E}_p(f)\le\ell$ by the weak lower semicontinuity of the norm.

(iii) This follows from Proposition 2.3(iii), since $|\varphi'|\le\mathrm{Lip}(\varphi)$.

(iv) By Lemma 2.2, the operator $(\nabla,\mathcal{C}_p)$ is densely defined, and $\mathcal{D}_{p'}$ is the domain of its adjoint. A densely defined operator between reflexive Banach spaces is closable if and only if its adjoint is densely defined [100, Section III.5.5]. We include the argument.

Suppose that $\mathcal{D}_{p'}$ is dense, and let (u_k) and G be as in (2.1). For $V\in\mathcal{D}_{p'}$, with h as in (2.7),

$$\int_{\mathcal{S}}\langle V,G\rangle\,d\nu=\lim_{k\to\infty}\int_{\mathcal{S}}\langle V,\nabla u_k\rangle\,d\nu=-\lim_{k\to\infty}\int_X u_k\,h\,dm=0\,.$$

Since $\mathcal{D}_{p'}$ is dense in the dual $L^{p'}(\mathcal{S},\mathsf{H},\nu)$ of $L^p(\mathcal{S},\mathsf{H},\nu)$, we get $G=0$.

Conversely, suppose that (A2) holds and that $\mathcal{D}_{p'}$ is not dense. By the Hahn–Banach theorem and the duality of Section 1.2, there is $G\in L^p(\mathcal{S},\mathsf{H},\nu)$, $G\neq0$, with $\int_{\mathcal{S}}\langle V,G\rangle\,d\nu=0$ for all $V\in\mathcal{D}_{p'}$. We claim that $(0,G)$ belongs to the closure of the graph of $(\nabla,\mathcal{C}_p)$ in $L^p(X,m)\oplus L^p(\mathcal{S},\mathsf{H},\nu)$,

which contradicts (A2). Otherwise, by the Hahn–Banach theorem again, there is $(h, V) \in L^{p'}(X, m) \oplus L^{p'}(\mathcal{S}, \mathsf{H}, \nu)$ such that

$$\int_X u\, h\, dm + \int_{\mathcal{S}} \langle V, \nabla u \rangle\, d\nu = 0 \quad (u \in \mathcal{C}_p) \qquad \text{and} \qquad \int_{\mathcal{S}} \langle V, G \rangle\, d\nu \neq 0\,.$$

The first property says that $V \in \mathcal{D}_{p'}$, and the second then contradicts the choice of G. □

Divergence. For $V \in \mathcal{D}_{p'}$ the function h in (2.7) is unique, because $\mathcal{C}_p$ is dense in $L^p(X, m)$; we write $h = \operatorname{div} V$. By approximation in the graph norm, (2.7) extends to

$$\int_{\mathcal{S}} \langle V, \nabla_p f \rangle\, d\nu = -\int_X f\, \operatorname{div} V\, dm \qquad (f \in W^{1,p}(X, m),\ V \in \mathcal{D}_{p'})\,, \tag{2.8}$$

so that $-\operatorname{div}$ is the adjoint of ∇_p. A vector field lives on $\mathcal{S}$ and is measured by ν; its divergence is a function on X and is measured by m. Closability is the density of the fields which have a divergence in this sense. Integration by parts also underlies Di Marino's construction of Sobolev spaces on metric measure spaces by derivations [61, Definition 1.5 and Theorem 1.7]. There the derivations and their divergences are formulated relative to one measure; in (2.8) the pairing uses two measures.

The endpoint $p = 1$. If the structure is 1-admissible, the graph norm makes $W^{1,1}(X, m)$ a Banach space, and Proposition 2.3 applies. The space is not reflexive in general, and $\mathcal{E}_1$ need not be lower semicontinuous on $L^1(X, m)$. For this reason functions of bounded variation are defined in Section 2.6 through the lower semicontinuous envelope of $\mathcal{E}_1$.

2.4 The p-energy form

For $1 < p < \infty$, nonlinear analogues of Dirichlet forms have been axiomatized under the name of p-energy forms. We recall the conventions of Kajino and Shimizu [98, Definitions 3.1, 3.24 and 3.25]. A *p-energy form* on (X, m) is a pair $(\mathcal{E}, \mathcal{F})$, where $\mathcal{F}$ is a linear space of m-measurable functions and $\mathcal{E} : \mathcal{F} \to [0, \infty)$ is a functional such that $\mathcal{E}^{1/p}$ is a seminorm. It is *closed* on $L^p(X, m)$ if $\mathcal{F} \subset L^p(X, m)$ is complete for the norm $(\|f\|_p^p + \mathcal{E}(f))^{1/p}$, and *regular* if $\mathcal{F} \cap C_c(X)$ contains a subspace which is dense in $\mathcal{F}$ for this norm and dense in $C_c(X)$ for the uniform norm.

The Markov property of a Dirichlet form is replaced by the *generalized p-contraction property* of [98, Definition 2.1], which is the inequality (2.9) below with $A = X$. It contains, as special cases, the contraction under 1-Lipschitz maps, strong subadditivity and the p-Clarkson inequalities [98, Proposition 2.2]. The next proposition shows that the energy of a p-admissible gradient structure has this property, together with the measures $|\nabla_p f|^p \nu$. No comparison between m and ν is required.

Proposition 2.6. *Let $1 < p < \infty$, and suppose that the gradient structure is p-admissible. Then $(\mathcal{E}_p, W^{1,p}(X, m))$ is a closed regular p-energy form. For $f \in W^{1,p}(X, m)$ define*

$$\Gamma_p(f)(A) = \int_{A \cap \mathcal{S}} |\nabla_p f|^p\, d\nu \qquad (A \subset X \text{ Borel})\,.$$

These are finite Radon measures, $\Gamma_p(f)(X) = \mathcal{E}_p(f)$, and $f \mapsto \Gamma_p(f)(A)^{1/p}$ is a seminorm for every Borel set A.

Let $k, l \geq 1$, $0 < q \leq p \leq r \leq \infty$, and let $T = (T_1, \dots, T_l) : \mathbb{R}^k \to \mathbb{R}^l$ satisfy $T(0) = 0$ and

$$\|T(a) - T(b)\|_{\ell^r} \leq \|a - b\|_{\ell^q} \qquad (a, b \in \mathbb{R}^k)\,.$$

For $f_1, \dots, f_k \in W^{1,p}(X, m)$, every $T_j(f_1, \dots, f_k)$ belongs to $W^{1,p}(X, m)$, and

$$\Big\| \big(\Gamma_p(T_j(f))(A)^{1/p}\big)_{j=1}^{l} \Big\|_{\ell^r} \leq \Big\| \big(\Gamma_p(f_i)(A)^{1/p}\big)_{i=1}^{k} \Big\|_{\ell^q} \qquad (A \subset X \text{ Borel})\,. \tag{2.9}$$

Thus both the form and its measures satisfy the generalized p-contraction property of [98, Definition 2.1 and Section 4.1].

The proof reduces (2.9) to a pointwise inequality between the gradients, which in turn follows from an inequality for linear contractions between finite-dimensional sequence spaces. We adapt the Gaussian argument of [98, Proposition A.16 and Theorem A.19]; the numbering refers to the author revision specified in the bibliography.

Proof. The seminorm property follows from the linearity of ∇_p and Minkowski's inequality in $L^p(A, \nu)$. Closedness and regularity follow from the definition of p-admissibility, with core $\mathcal{C}_p$. Each $\Gamma_p(f)$ is a finite Borel measure on a locally compact separable metric space, hence is Radon.

Step 1: a linear inequality. Let $A = (A_{ji}) : \mathbb{R}^k \to \mathbb{R}^l$ be linear with $\|Aa\|_{\ell^r} \leq \|a\|_{\ell^q}$ for all a, and let $v_1, \dots, v_k$ be vectors of a finite-dimensional Euclidean space. Then

$$\left\| \left(\left| \sum_i A_{ji} v_i \right| \right)_{j=1}^{l} \right\|_{\ell^r} \leq \big\| (|v_i|)_{i=1}^{k} \big\|_{\ell^q}\,. \tag{2.10}$$

Indeed, let Z be a standard Gaussian vector in that space. For every vector v, the random variable $\langle v, Z\rangle$ is centred Gaussian with variance $|v|^2$, so its L^q quasi-norm is $c_q|v|$, where $c_q > 0$ depends only on q. Since $q \leq r$, Minkowski's inequality with exponent $r/q \geq 1$ gives

$$\begin{aligned} c_q \left\| \left(\left| \sum_i A_{ji} v_i \right| \right)_j \right\|_{\ell^r} &\leq \big(\mathbb{E}\, \|A(\langle v_i, Z\rangle)_i\|_{\ell^r}^q\big)^{1/q} \\ &\leq \left(\mathbb{E} \sum_i |\langle v_i, Z\rangle|^q \right)^{1/q} = c_q \|(|v_i|)_i\|_{\ell^q}\,. \end{aligned}$$

For $r = \infty$, the first inequality holds because the maximum of the individual L^q quasi-norms is at most the L^q quasi-norm of the maximum.

Step 2: smooth T and functions of the core. Suppose that T is smooth and $u_1, \dots, u_k \in \mathcal{C}_p$. Differentiating the hypothesis on T shows that $\|DT(a)h\|_{\ell^r} \leq \|h\|_{\ell^q}$ for all $a, h \in \mathbb{R}^k$. The chain rule (1.6) and (2.10) therefore give

$$\|(|\nabla T_j(u)|)_j\|_{\ell^r} \leq \|(|\nabla u_i|)_i\|_{\ell^q} \qquad \nu\text{-almost everywhere}\,. \tag{2.11}$$

The functions $T_j(u)$ are Lipschitz, they have compact support because $T(0) = 0$, and (2.11) shows that their gradients are p-integrable. Thus $T_j(u) \in \mathcal{C}_p$.

Step 3: general T and functions of the core. Let (ρ_ε) be a smooth approximation of the identity on $\mathbb{R}^k$ and $T^\varepsilon = T * \rho_\varepsilon - (T * \rho_\varepsilon)(0)$. Since $\|\cdot\|_{\ell^r}$ is a norm, T^ε satisfies the hypothesis on T, and $T^\varepsilon \to T$ locally uniformly. For $u_1, \dots, u_k \in \mathcal{C}_p$, the functions $T_j^\varepsilon(u)$ vanish outside $\bigcup_i \operatorname{supp} u_i$ and converge uniformly, hence in $L^p(X, m)$, to $T_j(u)$. By Step 2 their gradients are bounded in

$L^p(\nu)$. Since $L^p(\mathcal{S},\mathsf{H},\nu)$ is reflexive and the graph of ∇_p is weakly closed, $T_j(u)\in W^{1,p}(X,m)$, and along a sequence $\varepsilon\to 0$ the gradients $\nabla T_j^{\varepsilon}(u)$ converge weakly to $\nabla_p T_j(u)$ for every j. The set of l-tuples of fields (W_j) with $\|(|W_j|)_j\|_{\ell^r}\le\|(|\nabla u_i|)_i\|_{\ell^q}$ almost everywhere is convex and closed in $L^p(\nu)^l$, hence weakly closed. Therefore (2.11) holds with $\nabla_p T_j(u)$ on the left-hand side.

Step 4: general functions. Let $f_i\in W^{1,p}(X,m)$ and choose $u_{i,n}\in\mathcal{C}_p$ converging to f_i in the graph norm. The hypothesis on T gives $|T_j(a)-T_j(b)|\le C\sum_i|a_i-b_i|$, so $T_j(u_n)\to T_j(f)$ in $L^p(X,m)$. Set

$$h_n=\|(|\nabla u_{i,n}|)_i\|_{\ell^q}\,,\qquad h=\|(|\nabla_p f_i|)_i\|_{\ell^q}\,.$$

After passing to a subsequence, the gradients $\nabla u_{i,n}$ converge ν-almost everywhere and have a common majorant in $L^p(\nu)$; dominated convergence gives $h_n\to h$ in $L^p(\nu)$. By Step 3 the gradients $\nabla_p T_j(u_n)$ are bounded in $L^p(\nu)$ by h_n. Pass to a subsequence along which they converge weakly for every j; since ∇_p is closed, the limits are $\nabla_p T_j(f)$, and in particular $T_j(f)\in W^{1,p}(X,m)$. By Mazur's lemma, applied simultaneously to these gradients and to h_n, suitable convex combinations converge strongly in $L^p(\nu)$ and, along a further subsequence, almost everywhere. Since $\|\cdot\|_{\ell^r}$ is convex, the bound of Step 3 passes to the limit:

$$\|(|\nabla_p T_j(f)|)_j\|_{\ell^r}\le\|(|\nabla_p f_i|)_i\|_{\ell^q}\qquad\nu\text{-almost everywhere}\,. \tag{2.12}$$

Step 5: integration. Let $A\subset X$ be a Borel set. Minkowski's inequality with exponent $r/p\ge 1$, then (2.12), then Minkowski's inequality with exponent $p/q\ge 1$, give

$$\begin{aligned}\big\|\big(\|\nabla_p T_j(f)\|_{L^p(A,\nu)}\big)_j\big\|_{\ell^r}&\le\big\|\,\|(|\nabla_p T_j(f)|)_j\|_{\ell^r}\big\|_{L^p(A,\nu)}\\&\le\big\|\,\|(|\nabla_p f_i|)_i\|_{\ell^q}\big\|_{L^p(A,\nu)}\le\big\|\big(\|\nabla_p f_i\|_{L^p(A,\nu)}\big)_i\big\|_{\ell^q}\,.\end{aligned}$$

This is (2.9). □

Related constructions. Proposition 2.6 starts from a prescribed gradient and obtains the energy measures as $|\nabla_p f|^p\nu$. In the literature on fractals the order is usually reversed. Limits of Korevaar–Schoen energies give p-energy forms and energy measures in [5, 99]; self-similar constructions are developed in [129]. Sasaya [141] constructs canonical energy measures for strongly local p-energy forms under his axioms, and [142] extends the construction beyond homogeneous energies. Identifying the measures $\Gamma_p(f)$ with the canonical measures of those theories requires checking the corresponding locality and calculus assumptions, and is not claimed here.

One measure for all exponents. In a gradient structure, the single measure ν dominates the energy measures $\Gamma_p(f)$ at every admissible exponent. This is additional structure, which a family of p-energy forms need not have: in the self-similar examples of [13], the domains of the energy forms at two different exponents can intersect only in the constants. Such examples are compatible with Proposition 4.7, which assumes one gradient structure and Poincaré inequalities at both exponents.

2.5 The Dirichlet form

At the exponent $p=2$ the closed gradient defines a quadratic form, and the theory of Dirichlet forms [69] applies. We recall the notions needed to state the result. Let $(\mathcal{E},\mathcal{F})$ be a strongly local regular Dirichlet form on $L^2(X,m)$. The *energy measure* $\Gamma(f)$ of $f\in\mathcal{F}$ is normalized by $\Gamma(f)(X)=\mathcal{E}(f,f)$; for bounded f it is characterized by

$$\int_X\varphi\,d\Gamma(f)=\mathcal{E}(\varphi f,f)-\tfrac12\,\mathcal{E}(f^2,\varphi)\qquad(\varphi\in\mathcal{F}\cap C_c(X))\,, \tag{2.13}$$

and $\Gamma(f,g)$ is defined by polarization; see [69, Section 3.2], where the measure $\mu_{\langle f\rangle} = 2\Gamma(f)$ is used. The following two notions are due to Hino [85, 86].

- A σ-finite Borel measure μ is *minimal energy-dominant* if $\Gamma(f) \ll \mu$ for every $f \in \mathcal{F}$, and if $\mu(A) = 0$ for every Borel set A such that $\Gamma(f)(A) = 0$ for all $f \in \mathcal{F}$.
- The *pointwise index* is the smallest measurable function, up to μ-null sets, which bounds the rank of the matrix $(d\Gamma(f_i,f_j)/d\mu)_{i,j\le k}$ μ-almost everywhere, for every finite family $f_1,\dots,f_k \in \mathcal{F}$. Its μ-essential supremum is the *index* of the form.

The index is equal to the martingale dimension of the associated diffusion process [85, Theorem 3.4].

Proposition 2.7. *Let $(\mathcal{S},\nu,\nabla)$ be a 2-admissible gradient structure, and let*

$$\mathcal{F} = W^{1,2}(X,m)\,, \qquad \mathcal{E}(f,g) = \int_{\mathcal{S}} \langle \nabla_2 f, \nabla_2 g\rangle \, d\nu\,.$$

(i) *$(\mathcal{E},\mathcal{F})$ is a regular, strongly local Dirichlet form on $L^2(X,m)$ with core $\mathcal{C}_2$.*

(ii) *The energy measure of $f \in \mathcal{F}$ is $d\Gamma(f) = |\nabla_2 f|^2\, d\nu$. In particular, if $f \in \mathcal{F}$ is m-almost everywhere constant on an open set U, then $\nabla_2 f = 0$ ν-almost everywhere on U.*

(iii) *The measure $\nu\restriction_{\mathcal{S}_2}$ is minimal energy-dominant. On $\mathcal{S}_2$, the pointwise index is the dimension of the field generated by the gradients of $\mathcal{C}_2$. If $\mathcal{C}_2$ spans the fibres, then ν is minimal energy-dominant, the pointwise index is $n(x)$, and the index and the martingale dimension of the associated diffusion are equal to the rank n_*.*

Proof. (i) The form is closed by the definition of ∇_2, and it is Markovian by Proposition 2.5(iii) applied to $\varphi(t) = (t\vee 0)\wedge 1$. By construction $\mathcal{C}_2$ is dense in $\mathcal{F}$, and by (A1) it is dense in $C_c(X)$; hence the form is regular with core $\mathcal{C}_2$. For the strong locality we use the Beurling–Deny decomposition [69, Theorem 3.2.1]: for $u,v \in \mathcal{F}\cap C_c(X)$,

$$\mathcal{E}(u,v) = \mathcal{E}^{(c)}(u,v) + \int_{X\times X\setminus \mathrm{diag}} (u(x)-u(y))(v(x)-v(y))\, J(dx\,dy) + \int_X uv\, dk\,,$$

where $\mathcal{E}^{(c)}$ is strongly local, J is a symmetric positive Radon measure off the diagonal and k is a positive Radon measure. We show that $J = 0$ and $k = 0$.

The jump measure vanishes. Let $K_1, K_2 \subset X$ be disjoint compact sets. By Lemma 2.2(i), applied in disjoint open neighbourhoods of K_1 and K_2, there are nonnegative $u,v \in \mathcal{C}_2$ with disjoint supports, $u = 1$ on K_1 and $v = 1$ on K_2. By Lemma 1.5(ii), $\nabla u = 0$ ν-almost everywhere outside $\operatorname{supp} u$ and $\nabla v = 0$ ν-almost everywhere outside $\operatorname{supp} v$, so $\mathcal{E}(u,v) = 0$. Moreover $\mathcal{E}^{(c)}(u,v) = 0$ and $uv = 0$, and $(u(x)-u(y))(v(x)-v(y)) = -u(x)v(y) - u(y)v(x)$. The decomposition and the symmetry of J give

$$2J(K_1\times K_2) \le 2\int u(x)v(y)\, J(dx\,dy) = -\mathcal{E}(u,v) = 0\,.$$

The complement of the diagonal is a countable union of such products, hence $J = 0$.

The killing measure vanishes. Let $K \subset X$ be compact. By Lemma 2.2(i) there are $v \in \mathcal{C}_2$ with $0 \le v \le 1$ and $v = 1$ on K, and then $u \in \mathcal{C}_2$ with $u = 1$ on a neighbourhood W of $\operatorname{supp} v$. By Lemma 1.5(i),(ii), $\nabla u = 0$ ν-almost everywhere on W and $\nabla v = 0$ ν-almost everywhere outside W; hence $\mathcal{E}(u,v) = 0$. Also $\mathcal{E}^{(c)}(u,v) = 0$, because u is constant on a neighbourhood of $\operatorname{supp} v$. Since $J = 0$, the decomposition gives $k(K) \le \int_X v\, dk = \int_X uv\, dk = 0$. Hence $k = 0$ and $\mathcal{E} = \mathcal{E}^{(c)}$.

(ii) For $f, \varphi \in \mathcal{C}_2$, the Leibniz rule gives $\nabla(\varphi f) = \varphi\,\nabla f + f\,\nabla\varphi$ and $\nabla(f^2) = 2f\,\nabla f$, hence

$$\mathcal{E}(\varphi f, f) - \tfrac{1}{2}\,\mathcal{E}(f^2, \varphi) = \int_{\mathcal{S}} (\varphi\,|\nabla f|^2 + f\langle \nabla\varphi, \nabla f\rangle - f\langle \nabla f, \nabla\varphi\rangle)\,d\nu = \int_{\mathcal{S}} \varphi\,|\nabla f|^2\,d\nu\,.$$

By (2.13), the two finite measures $\Gamma(f)$ and $|\nabla f|^2\nu$ have the same integral against every $\varphi \in \mathcal{C}_2$. Since $\mathcal{C}_2$ is uniformly dense in $C_c(X)$, they are equal. For a general $f \in \mathcal{F}$, let $u_k \in \mathcal{C}_2$ with $u_k \to f$ in $\mathcal{F}$. For every Borel set A,

$$\big|\Gamma(f)(A)^{1/2} - \Gamma(u_k)(A)^{1/2}\big| \le \mathcal{E}(f - u_k, f - u_k)^{1/2}\,,$$

by [69, Section 3.2], and the same inequality holds for the measures $|\nabla_2 f|^2\nu$ and $|\nabla u_k|^2\nu$, by the triangle inequality in $L^2(A,\nu)$. Letting $k \to \infty$ gives $\Gamma(f)(A) = \int_A |\nabla_2 f|^2\,d\nu$. The last assertion holds because the energy measure of a function of the domain of a strongly local regular Dirichlet form does not charge an open set on which the function is constant [69, Section 3.2].

By polarization, the identity in (ii) gives

$$\frac{d\Gamma(f,g)}{d\nu}(x) = \langle \nabla_2 f(x), \nabla_2 g(x)\rangle \qquad (f, g \in \mathcal{F})\,, \tag{2.14}$$

and all these measures are carried by $\mathcal{S}_2$, by (2.3).

(iii) Every energy measure is absolutely continuous with respect to $\nu\restriction_{\mathcal{S}_2}$, by (ii) and (2.3). If a Borel set A has $\Gamma(f)(A) = 0$ for all $f \in \mathcal{F}$, then the functions v_k of (2.2) have zero gradient almost everywhere on A, so $\nu(A \cap \mathcal{S}_2) = 0$. Thus $\nu\restriction_{\mathcal{S}_2}$ is minimal energy-dominant. Every closed gradient takes its values in the measurable field generated by the gradients of $\mathcal{C}_2$, by approximation in the graph norm. By (2.14), the matrices $(d\Gamma(f_i, f_j)/d\nu)$ are Gram matrices of vectors of this field. Their ranks are bounded by its dimension, and the family (v_k) attains that dimension, by (1.7) applied to its span; compare [85, Lemma 2.3 and Proposition 2.10]. When $\mathcal{C}_2$ spans the fibres, $\mathcal{S}_2 = \mathcal{S}$ up to a null set, the dimension is $n(x)$ and the index is n_*. The index is equal to the martingale dimension by [85, Theorem 3.4]; compare [86, Definition 2.2 and Theorem 3.4]. □

Remark 2.8 (The trace at $p = 2$). For a 2-admissible structure, the trace $T_2 f$ of Proposition 2.3 is the restriction to $\mathcal{S}_2$ of a quasi-continuous representative $\tilde{f}$ of f. Indeed, let $u_j \in \mathcal{C}_2$ converge to f in $\mathcal{F}$. A subsequence converges to $\tilde{f}$ quasi-everywhere [69, Theorem 2.1.4]. Energy measures do not charge sets of zero capacity [69, Lemma 3.2.4], and by the proof of Proposition 2.3 the measure $\nu\restriction_{\mathcal{S}_2}$ has the same null sets as a countable combination of energy measures. Hence the subsequence converges to $\tilde{f}$ ν-almost everywhere on $\mathcal{S}_2$, and its limit in $L^0(\mathcal{S}_2, \nu)$ is $T_2 f$.

From Dirichlet forms to gradients. Conversely, every strongly local regular Dirichlet form admits a first-order calculus with values in a measurable field of Hilbert spaces whose fibre dimensions are given by the pointwise index [86]; see also [54, 92, 90]. A gradient structure contains more information: it is defined on all bounded Lipschitz functions of a specified metric, and it is bounded by the Lipschitz constant. Under the additional condition (SL), Proposition 1.11 recovers that metric locally.

Generator and semigroup. The operator $L = \operatorname{div}\nabla_2$, with domain $\{f \in \mathcal{F} : \nabla_2 f \in \mathcal{D}_2\}$, is the generator of $(\mathcal{E}, \mathcal{F})$, by (2.8). We denote by $P_t = e^{tL}$, $t \ge 0$, the associated *heat semigroup* on $L^2(X, m)$. Part II studies P_t and its kernel.

Singular energy measures. When $\nu \perp m$, Proposition 2.7(ii) shows that every energy measure is singular with respect to m. On self-similar fractals this singularity is a theorem [118, 32, 84]. Kajino and Murugan [96] proved it for symmetric diffusions with full off-diagonal sub-Gaussian heat kernel estimates, under a condition on the space-time scale function which is satisfied by $\Psi(r) = r^\beta$ with $\beta > 2$. Gradient structures with $\nu \perp m$ are a framework in which this singularity is built into the data.

2.6 Relaxation and bounded variation

Closure requires closability and produces a closed operator. Relaxation requires nothing and produces a lower semicontinuous functional. For $1 \le p < \infty$, the *relaxed energy* of $f \in L^p(X,m)$ is

$$\overline{\mathcal{E}}_p(f) = \inf\Big\{ \liminf_{k\to\infty} \int_{\mathcal{S}} |\nabla u_k|^p \, d\nu : \ u_k \in \mathcal{C}_p \,, \ u_k \to f \text{ in } L^p(X,m) \Big\} \,, \tag{2.15}$$

with the convention $\inf \varnothing = +\infty$. It is the largest lower semicontinuous functional on $L^p(X,m)$ which is bounded by $\int_{\mathcal{S}} |\nabla u|^p \, d\nu$ on $\mathcal{C}_p$. For a p-admissible structure with $p > 1$, the lower semicontinuity of $\mathcal{E}_p$ (Proposition 2.5(ii)) and the density of $\mathcal{C}_p$ in the graph norm give $\overline{\mathcal{E}}_p = \mathcal{E}_p$. The relaxed energy is defined without any closability assumption, and Chapter 4 compares it with averaged difference quotients.

At $p = 1$ the relaxed energy is the total variation. This is the definition of BV by relaxation of [126, 6] for metric measure spaces, and of [22, Definition 2.17] for local trees.

Definition 2.9. For $f \in L^1(X,m)$, the *total variation* of f is

$$\|Df\|(X) = \overline{\mathcal{E}}_1(f) = \inf\Big\{ \liminf_{k\to\infty} \int_{\mathcal{S}} |\nabla u_k| \, d\nu : \ u_k \in \mathcal{C}_1 \,, \ u_k \to f \text{ in } L^1(X,m) \Big\} \,,$$

and $BV(X,m) = \{ f \in L^1(X,m) : \|Df\|(X) < \infty \}$ is the space of *functions of bounded variation*. The *perimeter* of a Borel set E with $m(E) < \infty$ is $P(E) = \|D\mathbf{1}_E\|(X)$.

The functional $f \mapsto \|Df\|(X)$ is convex, even, positively homogeneous of degree one, and lower semicontinuous on $L^1(X,m)$. Consequently it is a seminorm on $BV(X,m)$, and $BV(X,m)$ is a Banach space for the norm $\|f\|_{L^1} + \|Df\|(X)$. The definition requires neither closability at $p = 1$ nor any quantitative hypothesis. In this book $\|Df\|(X)$ is a number; we do not show that it is the total mass of a measure $\|Df\|$ on X, and the notation is chosen only for its familiarity.

2.7 Divergence duality and coarea

We prove two structural properties of the total variation: it is the support function of the divergences of the bounded vector fields, and it satisfies the coarea formula. The divergence representation is analogous to [61, Theorem 3.4]; relaxation and coarea follow the approach to BV on metric measure spaces of [126, 6]. Both statements follow from the calculus of Lemma 1.5 and from the definition by relaxation. Throughout this section we write

$$J(f) = \|Df\|(X) \,, \qquad f \in L^1(X,m) \,.$$

Duality. A bounded field $V \in L^\infty(\mathcal{S}, \mathsf{H}, \nu)$ has a *bounded divergence* if there is $h \in L^\infty(X,m)$ such that

$$\int_{\mathcal{S}} \langle \nabla u, V \rangle \, d\nu = -\int_X u \, h \, dm \qquad \text{for all } u \in \mathcal{C}_1 \,; \tag{2.16}$$

we then write $h = \operatorname{div} V$. When $\mathcal{C}_1$ is dense in $L^1(X,m)$ the function h is determined by V. In general, the supremum in the next statement is taken over the pairs (V,h) satisfying (2.16).

Proposition 2.10. *Let $(\mathcal{S}, \nu, \nabla)$ be a gradient structure. For every $f \in L^1(X, m)$,*

$$\|Df\|(X) = \sup\Big\{ \Big| \int_X f \, \operatorname{div} V \, dm \Big| : \ V \in L^\infty(\mathcal{S}, \mathsf{H}, \nu) \,, \ \|V\|_\infty \le 1 \,, \ \operatorname{div} V \in L^\infty(X, m) \Big\} \,. \quad (2.17)$$

On $\mathbb{R}^n$, with $m = \nu = \mathcal{L}^n$ and the Euclidean gradient, this is the classical definition of the total variation by duality with vector fields; see [7, Chapter 3]. In the present setting the fields are sections over the skeleton, bounded with respect to ν, and their divergences are functions on X, bounded with respect to m.

Proof. Step 1: the polar set. The functional J is the lower semicontinuous envelope on $L^1(X, m)$ of the functional which is equal to $\|\nabla u\|_{L^1(\nu)}$ on $\mathcal{C}_1$ and to $+\infty$ elsewhere, and the latter is convex and positively homogeneous. Hence J is convex, lower semicontinuous and positively homogeneous. By the Hahn–Banach theorem, applied to its closed convex epigraph, it is the support function of a subset of the dual space $L^\infty(X, m)$:

$$J(f) = \sup_{h \in \mathcal{K}} \int_X f \, h \, dm \,, \qquad \mathcal{K} = \Big\{ h \in L^\infty(X, m) : \ \int_X v \, h \, dm \le J(v) \text{ for all } v \in L^1(X, m) \Big\} \,. \quad (2.18)$$

Moreover $h \in \mathcal{K}$ if and only if

$$\Big| \int_X u \, h \, dm \Big| \le \|\nabla u\|_{L^1(\nu)} \qquad \text{for every } u \in \mathcal{C}_1 \,. \quad (2.19)$$

Indeed, $J(u) \le \|\nabla u\|_{L^1(\nu)}$ on $\mathcal{C}_1$, and $\mathcal{C}_1$ is stable under $u \mapsto -u$. Conversely, if (2.19) holds and $u_k \in \mathcal{C}_1$ converges to v in $L^1(X, m)$, then $\int_X v \, h \, dm = \lim_k \int_X u_k \, h \, dm \le \liminf_k \|\nabla u_k\|_{L^1(\nu)}$.

Step 2: representation by a field. Let $h \in \mathcal{K}$. By (2.19), the map $\nabla u \mapsto \int_X u \, h \, dm$ is well defined on the linear subspace $\{\nabla u : u \in \mathcal{C}_1\}$ of $L^1(\mathcal{S}, \mathsf{H}, \nu)$, and it has norm at most one. By the Hahn–Banach theorem it extends to a linear form of norm at most one on $L^1(\mathcal{S}, \mathsf{H}, \nu)$, which by the duality of Section 1.2 is represented by a field $W \in L^\infty(\mathcal{S}, \mathsf{H}, \nu)$ with $\|W\|_\infty \le 1$. Thus $V = -W$ satisfies (2.16), that is, $h = \operatorname{div} V$.

Step 3: conclusion. Conversely, if $\|V\|_\infty \le 1$ and (V, h) satisfies (2.16), then h satisfies (2.19), so $h \in \mathcal{K}$ and $-h \in \mathcal{K}$. Formula (2.17) is therefore (2.18). □

The coarea formula. The classical coarea formula for functions of bounded variation [7, Theorem 3.40] expresses the total variation as the integral of the perimeters of the superlevel sets.

Theorem 2.11. *For every $f \in L^1(X, m)$,*

$$\|Df\|(X) = \int_0^\infty \big(P(\{f > t\}) + P(\{f < -t\}) \big) \, dt \,, \quad (2.20)$$

where the equality holds in $[0, \infty]$. In particular, for a function of $BV(X, m)$ and almost every $t > 0$, the sets $\{f > t\}$ and $\{f < -t\}$ have finite perimeter. For Borel sets E, F of finite m-measure,

$$P(E \cup F) + P(E \cap F) \le P(E) + P(F) \,. \quad (2.21)$$

The proof adapts the truncation and relaxation argument of [126, Proposition 4.2]. The inequality "$\ge$" is proved by truncating a recovery sequence; the inequality "$\le$" follows from the duality formula (2.18). Only the global functional J is used.

Proof. For $f \in L^1(X,m)$, both maps $t \mapsto \mathbf{1}_{\{f>t\}}$ and $t \mapsto \mathbf{1}_{\{f<-t\}}$ are right-continuous from $(0,\infty)$ to $L^1(X,m)$. Their compositions with the lower semicontinuous functional J are therefore Borel measurable, and the integral in (2.20) is defined.

Step 1: the core. For $u \in \mathcal{C}_1$, $t \geq 0$ and $\varepsilon > 0$, the functions

$$u^{\pm}_{t,\varepsilon} = \min\Big\{1, \frac{(\pm u - t)_+}{\varepsilon}\Big\}$$

belong to $\mathcal{C}_1$ and converge in $L^1(X,m)$ to $\mathbf{1}_{\{\pm u>t\}}$ as $\varepsilon \downarrow 0$. The chain rule of Lemma 1.5(iv) gives

$$|\nabla u^{\pm}_{t,\varepsilon}| = \varepsilon^{-1}\, \mathbf{1}_{\{t<\pm u<t+\varepsilon\}}\, |\nabla u|\,.$$

For each point x, the two sets $\{t>0 : t < \pm u(x) < t+\varepsilon\}$ have total length at most ε. The lower semicontinuity of J, Fatou's lemma and Tonelli's theorem give

$$\int_0^\infty \big(J(\mathbf{1}_{\{u>t\}}) + J(\mathbf{1}_{\{u<-t\}})\big)\, dt \leq \liminf_{\varepsilon\downarrow 0} \int_0^\infty \!\!\int_{\mathcal{S}} \big(|\nabla u^+_{t,\varepsilon}| + |\nabla u^-_{t,\varepsilon}|\big)\, d\nu\, dt \leq \int_{\mathcal{S}} |\nabla u|\, d\nu\,. \tag{2.22}$$

Step 2: passage to the relaxation. If $J(f) < \infty$, choose $u_k \in \mathcal{C}_1$ with $u_k \to f$ in $L^1(X,m)$ and $\|\nabla u_k\|_{L^1(\nu)} \to J(f)$. The identity $|a-b| = \int_0^\infty \big(|\mathbf{1}_{\{a>t\}} - \mathbf{1}_{\{b>t\}}| + |\mathbf{1}_{\{a<-t\}} - \mathbf{1}_{\{b<-t\}}|\big)\, dt$ for real numbers a, b gives

$$\int_0^\infty \Big(\|\mathbf{1}_{\{u_k>t\}} - \mathbf{1}_{\{f>t\}}\|_{L^1(m)} + \|\mathbf{1}_{\{u_k<-t\}} - \mathbf{1}_{\{f<-t\}}\|_{L^1(m)}\Big)\, dt = \|u_k - f\|_{L^1(m)} \longrightarrow 0\,.$$

Along a subsequence both families of indicators converge in $L^1(X,m)$ for almost every t. The lower semicontinuity of J, Fatou's lemma and (2.22) therefore yield

$$\int_0^\infty \big(P(\{f>t\}) + P(\{f<-t\})\big)\, dt \leq \liminf_{k\to\infty} \int_{\mathcal{S}} |\nabla u_k|\, d\nu = J(f)\,.$$

This inequality is trivial if $J(f) = \infty$.

Step 3: the reverse inequality. Let $h \in \mathcal{K}$, with $\mathcal{K}$ as in (2.18); recall that $-h \in \mathcal{K}$, because J is even. The identity $f = \int_0^\infty (\mathbf{1}_{\{f>t\}} - \mathbf{1}_{\{f<-t\}})\, dt$ and Fubini's theorem give

$$\begin{aligned}\int_X f\, h\, dm &= \int_0^\infty \Big(\int_{\{f>t\}} h\, dm - \int_{\{f<-t\}} h\, dm\Big)\, dt \\ &\leq \int_0^\infty \big(P(\{f>t\}) + P(\{f<-t\})\big)\, dt\,.\end{aligned}$$

Taking the supremum over h proves (2.20).

Step 4: the perimeter. Let $P(E) + P(F) < \infty$, and let $u_k, v_k \in \mathcal{C}_1$ with $u_k \to \mathbf{1}_E$, $v_k \to \mathbf{1}_F$ in $L^1(X,m)$, $\|\nabla u_k\|_{L^1(\nu)} \to P(E)$ and $\|\nabla v_k\|_{L^1(\nu)} \to P(F)$. Then $u_k \vee v_k \to \mathbf{1}_{E\cup F}$ and $u_k \wedge v_k \to \mathbf{1}_{E\cap F}$ in $L^1(X,m)$, and by (1.5),

$$|\nabla(u_k \vee v_k)| + |\nabla(u_k \wedge v_k)| = |\nabla u_k| + |\nabla v_k| \qquad \nu\text{-almost everywhere.}$$

The lower semicontinuity of J proves (2.21). □

Applying (2.20) to f_+ and f_- gives the additivity

$$\|Df\|(X) = \|Df_+\|(X) + \|Df_-\|(X) \qquad (f \in L^1(X,m))\,. \tag{2.23}$$

Compact spaces. Formula (2.20) is written with two families of level sets of finite measure because $m(X)$ may be infinite. If X is compact, constants belong to $\mathcal{C}_1$ and have zero gradient; then $\|D(f+c)\|(X) = \|Df\|(X)$ and $P(X \setminus E) = P(E)$. Since $m(\{f = t\}) > 0$ for at most countably many t, (2.20) takes the usual form

$$\|Df\|(X) = \int_{\mathbb{R}} P(\{f > t\})\, dt\,.$$

3 Doubling, cutoffs and Poincaré inequalities

The results of Chapter 2 are qualitative: they assume that the gradient is closable and describe the resulting spaces. The quantitative theory rests on three hypotheses, which are introduced in this chapter:

- volume doubling VD for the reference measure m;
- a cutoff condition $\mathrm{B}(\Psi_{C,2})$, which bounds the gradient of a cutoff function of a ball, integrated against ν, by the m-measure of the ball;
- a Poincaré inequality $\mathrm{PI}_p(\Psi_{P,p})$, which bounds the oscillation of a function, integrated against m, by its energy, integrated against ν.

The last two hypotheses are the two places where the measures m and ν are compared, and they compare them in opposite directions. Each involves a function of the radius, the *cutoff scale* $\Psi_{C,2}$ and the *Poincaré scale* $\Psi_{P,p}$. In the theory of heat kernels on metric measure spaces [17, 75, 10], a single space-time scale function Ψ governs both the Poincaré inequality and the cutoff Sobolev inequality. Here the two scales are kept separate, because there are natural examples in which they differ (Proposition 15.10), and because several results use only one of them.

The technical tool of the chapter is the discrete convolution of Lemma 3.5, which replaces a function by a combination of its averages on balls of radius ε and bounds the energy of the result by averaged difference quotients of the function. It plays the role of mollification, and all later approximation arguments rely on it. Two first applications are given in Section 3.4: the cutoff scale is at most quadratic, with $\nu \ll m$ in the quadratic case, and on a connected space the cutoff scale is dominated by the Poincaré scale.

3.1 Scale functions

Range of radii. Fix $r_0 \in (0, \infty]$. All hypotheses are imposed for radii below a fixed multiple of r_0, with constants which do not depend on the centre of the ball. A statement which holds *at small radii* holds for $r < c\, r_0$, where $c > 0$ depends only on the constants of the hypotheses; reducing r_0 by a fixed factor accommodates the dilations of balls which occur in the proofs. The hypotheses are *global* if $r_0 = \infty$; the conclusions then hold at every radius.

Growth conditions. A positive function S on $(0, r_0)$ has *weak scaling* if there are constants $c, C > 0$ and exponents $0 < a \le b < \infty$ with

$$c\Big(\frac{R}{r}\Big)^a \le \frac{S(R)}{S(r)} \le C\Big(\frac{R}{r}\Big)^b \qquad (0 < r \le R < r_0)\,. \tag{3.1}$$

The upper bound is a doubling condition and the lower bound a reverse doubling condition; the model is $S(r) = r^\beta$, with $a = b = \beta$. Some approximation arguments need only comparability at comparable radii:

$$C^{-1} \le \frac{S(R)}{S(r)} \le C \qquad (0 < r \le R \le 2r < r_0)\,. \tag{3.2}$$

The cutoff scales. We fix a positive Borel function $\Psi_{C,2}$ on $(0, r_0)$ satisfying (3.2), the *cutoff scale*, and define

$$\Psi_{C,p}(r) = r^{p-2}\Psi_{C,2}(r) \qquad (1 \le p < \infty)\,. \tag{3.3}$$

The index 2 is a normalization: $\Psi_{C,2}$ will be the scale of the quadratic energy of cutoffs, and (3.3) is the scale of their p-energy, as (3.10) below shows. On $\mathbb{R}^n$ one has $\Psi_{C,2}(r) = r^2$ and $\Psi_{C,p}(r) = r^p$.

The Poincaré scales. For each exponent p under consideration, $\Psi_{P,p}$ is a positive function on $(0, r_0)$, the *Poincaré scale* at p. In the quadratic theory we write $\Psi_P = \Psi_{P,2}$. Unless a weaker condition is specified, the cutoff and Poincaré scales at the exponents under consideration satisfy (3.1).

Definition 3.1. The cutoff and Poincaré scales *match* at the exponent p if

$$\Psi_{P,p}(r) \le C\,\Psi_{C,p}(r) \qquad (0 < r < r_0)\,. \tag{3.4}$$

On a connected space, Proposition 3.7 gives the reverse inequality at small radii. Matching therefore means that the two scales are comparable, and that the hypotheses reduce to the usual situation of one scale function.

3.2 The three hypotheses

Let $(\mathcal{S}, \nu, \nabla)$ be a gradient structure on (X, d, m).

Volume doubling: VD. All balls of radius less than r_0 have positive finite measure, and

$$m(B(x, 2r)) \le C_D\, m(B(x, r)) \qquad (x \in X,\ 0 < 2r < r_0)\,. \tag{3.5}$$

Iteration gives constants $C, D > 0$ such that

$$\frac{m(B(x, R))}{m(B(x, r))} \le C\Big(\frac{R}{r}\Big)^D \qquad (x \in X,\ 0 < r \le R < r_0)\,. \tag{3.6}$$

In particular, two balls of comparable radii whose centres are at distance at most a fixed multiple of these radii have comparable measures. No comparison is asserted for distant centres: the measure of a ball may depend on its centre.

Poincaré inequality: $\mathrm{PI}_p(\Psi_{P,p})$. There are $\Lambda \ge 1$ and $C_P > 0$ such that every $u \in \mathcal{C}_p$ and every ball $B = B(x, r)$ with $0 < r < r_0/\Lambda$ satisfy

$$\int_B |u - u_B|^p\, dm \le C_P\, \Psi_{P,p}(r) \int_{\Lambda B \cap \mathcal{S}} |\nabla u|^p\, d\nu\,. \tag{3.7}$$

This is a weak (p, p)-Poincaré inequality in the terminology of [79], with the gradient integrated against ν. At a p-admissible exponent, convergence in the graph norm extends (3.7) to every $f \in W^{1,p}(X, m)$, with the closed gradient $\nabla_p f$.

Cutoff condition: $\mathrm{B}(\Psi_{C,2})$. There is $C_B \ge 1$ such that for every $x \in X$ and $0 < r < r_0/4$ there is $\psi_x^r \in \mathrm{Lip}_c(X)$ with

$$\mathbf{1}_{B(x,r)} \le \psi_x^r \le \mathbf{1}_{B(x,2r)}\,, \qquad \mathrm{Lip}(\psi_x^r) \le \frac{C_B}{r}\,, \tag{3.8}$$

and

$$\int_{\mathcal{S}} |\nabla \psi_x^r|\, d\nu \le C_B\, \frac{r}{\Psi_{C,2}(r)}\, m(B(x, r))\,. \tag{3.9}$$

The condition is an $L^1(\nu)$ bound on the gradient of a Lipschitz cutoff. Because $|\nabla\psi_x^r| \le C_B/r$ by (G), it controls every $L^p(\nu)$ norm: $|\nabla\psi_x^r|^p \le (C_B/r)^{p-1}|\nabla\psi_x^r|$, and therefore

$$\int_{\mathcal{S}} |\nabla\psi_x^r|^p \, d\nu \le C_B^p \, \frac{r^{2-p}}{\Psi_{C,2}(r)} \, m(B(x,r)) = C_B^p \, \frac{m(B(x,r))}{\Psi_{C,p}(r)} \qquad (1 \le p < \infty)\,. \tag{3.10}$$

In particular the cutoffs ψ_x^r belong to every core $\mathcal{C}_p$, and the hypothesis of Proposition 1.9 holds with $b = 1$: under $\mathrm{B}(\Psi_{C,2})$, every core $\mathcal{C}_p$ is uniformly dense in $C_c(X)$ and spans the fibres. For $p = 2$, (3.10) is an upper bound for the capacity of the condenser $(B(x,r), B(x,2r))$, of the kind used in [75]. The right-hand side of (3.10) is an upper bound; no optimality of the cutoff scale is claimed.

A sufficient condition for (3.9) is a bound on the ν-measure of the set where the cutoff gradient is nonzero,

$$\nu(\{|\nabla\psi_x^r| > 0\}) \le C \, \frac{r^2}{\Psi_{C,2}(r)} \, m(B(x,r))\,, \tag{3.11}$$

by the bound $|\nabla\psi_x^r| \le C_B/r$. If $\nu \le C_0\, m$ and VD holds, the distance cutoffs of Section 3.5 satisfy (3.11) with $\Psi_{C,2}(r) = r^2$. When $\nu \perp m$, (3.11) says that the cutoff varies only on a part of the skeleton whose ν-measure is controlled; on a tree, for instance, this part consists of boundedly many segments of length r.

Uniform volume growth: $\mathrm{V}(\Phi)$. This is an additional hypothesis, which is not part of the standing assumptions. We write $\mathrm{V}(\Phi)$ if there is a nondecreasing function Φ on $(0, r_0)$ with

$$C_V^{-1}\Phi(r) \le m(B(x,r)) \le C_V\, \Phi(r) \qquad (x \in X,\ 0 < r < r_0)\,, \tag{3.12}$$

where

$$c\Big(\frac{R}{r}\Big)^{d_-} \le \frac{\Phi(R)}{\Phi(r)} \le C_\Phi\Big(\frac{R}{r}\Big)^{d_+}, \qquad 0 < r \le R < r_0\,, \quad 0 < d_- \le d_+ < \infty\,. \tag{3.13}$$

This implies VD. The case $\Phi(r) = r^Q$ is Ahlfors Q-regularity. Under $\mathrm{V}(\Phi)$ it is convenient to write the cutoff scale in terms of a second function of the radius, the *gauge*

$$\Theta(r) = \frac{r^2\,\Phi(r)}{\Psi_{C,2}(r)}\,, \qquad \text{so that} \qquad \Psi_{C,p}(r) = \frac{r^p\,\Phi(r)}{\Theta(r)}\,. \tag{3.14}$$

The gauge is comparable at comparable radii. Under $\mathrm{V}(\Phi)$, the integrated condition (3.9) reads $\int_{\mathcal{S}} |\nabla\psi_x^r|\, d\nu \le C\Theta(r)/r$. Its sufficient condition (3.11) reads $\nu(\{|\nabla\psi_x^r| > 0\}) \le C\Theta(r)$. If $\Phi(r) = r^Q$ and $\Theta(r) = r^s$, then $\Psi_{C,p}(r) = r^{p+Q-s}$. This notation is used in the explicit examples.

Coverings. The following covering lemma is the usual packing consequence of doubling; see [83, Chapter 10] and [31, Proposition 2.1]. A set $\{x_i\} \subset X$ is *ε-separated* if $d(x_i, x_j) \ge \varepsilon$ for $i \ne j$.

Lemma 3.2. *Let $\varepsilon > 0$. A maximal ε-separated set $\{x_i\}$ is countable, and the balls $B(x_i, \varepsilon)$ cover X. Under* VD*, for each $\lambda \ge 1$ every point of X belongs to at most $N(\lambda) = C(1+\lambda)^D$ of the balls $B(x_i, \lambda\varepsilon)$, provided $2(\lambda+1)\varepsilon < r_0$.*

Proof. Separability gives countability, and maximality gives the cover. If $y \in B(x_i, \lambda\varepsilon)$, then $B(x_i, \varepsilon/2) \subset B(y, (\lambda+1)\varepsilon) \subset B(x_i, 2(\lambda+1)\varepsilon)$. By (3.6), $m(B(x_i, \varepsilon/2)) \ge c(1+\lambda)^{-D} m(B(y, (\lambda+1)\varepsilon))$. Since the balls $B(x_i, \varepsilon/2)$ are disjoint, summing over the indices i with $y \in B(x_i, \lambda\varepsilon)$ proves the assertion. □

Reverse doubling on connected spaces. On a connected space, doubling implies a reverse volume comparison; this is the standard reverse doubling argument; see, for instance, [35, Corollary 3.8].

Lemma 3.3. *Assume* VD*, and suppose that X is connected with positive diameter. There is $M > 2$, depending only on C_D, such that*

$$m(B(x, Mr)) \geq 2\, m(B(x, 2r)) \tag{3.15}$$

for every $x \in X$ and every $r > 0$ with $Mr < \min\{r_0, \operatorname{diam} X\}/8$. On an unbounded connected space with global doubling, (3.15) *holds at every radius.*

Proof. Let $8\rho < \min\{r_0, \operatorname{diam} X\}$. Since X is connected and contains a point at distance greater than 3ρ from x, the continuous function $d(x, \cdot)$ takes the value 3ρ at some point y. The ball $B(y, \rho/2)$ lies in $B(x, 4\rho) \setminus B(x, 2\rho)$, and $B(x, 2\rho) \subset B(y, 8\rho)$, so $m(B(y, \rho/2)) \geq C_D^{-4} m(B(x, 2\rho))$ by doubling. Hence $m(B(x, 4\rho)) \geq (1 + C_D^{-4})\, m(B(x, 2\rho))$. Iterating k times, with $(1 + C_D^{-4})^k \geq 2$, gives (3.15) with $M = 2^{k+1}$. □

Lower semicontinuity. The weak-$*$ continuity of the gradient gives the lower semicontinuity of the energy on Borel sets, along bounded sequences of Lipschitz functions.

Lemma 3.4. *Let $u_k, u \in \operatorname{Lip}_b(X)$ with $u_k \to u$ pointwise and $\sup_k(\|u_k\|_\infty + \operatorname{Lip}(u_k)) < \infty$. Then, for every Borel set A and $1 \leq p < \infty$,*

$$\int_A |\nabla u|^p \, d\nu \leq \liminf_{k\to\infty} \int_A |\nabla u_k|^p \, d\nu \,. \tag{3.16}$$

Proof. Let g be a bounded field which vanishes outside a subset of A of finite ν-measure, with $\|g\|_{L^{p'}(\nu)} \leq 1$. Then $g \in L^1(\mathcal{S}, \mathsf{H}, \nu)$, and (G) and Hölder's inequality give

$$\int_{\mathcal{S}} \langle \nabla u, g \rangle \, d\nu = \lim_{k\to\infty} \int_{\mathcal{S}} \langle \nabla u_k, g \rangle \, d\nu \leq \liminf_{k\to\infty} \Big(\int_A |\nabla u_k|^p \, d\nu \Big)^{1/p} .$$

Since ν is σ-finite, the supremum of the left-hand side over all such g is $\|\nabla u\|_{L^p(A,\nu)}$, also for $p = 1$. □

3.3 Discrete convolution

On $\mathbb{R}^n$ a function is regularized by convolution with a mollifier. On a metric measure space the substitute is the *discrete convolution*: one fixes a partition of unity subordinate to a cover by balls of radius ε and replaces f by the combination of its averages on these balls. For Sobolev functions on metric spaces the construction is used in [107, Section 3]. Versions with control of the energy appear in [31, Theorem 5.2], [22, Theorem 3.5 and Proposition 3.6] and [146, Lemmas 3.10–3.11]. The estimates below follow the same argument, with the gradients integrated against ν and the oscillations against m.

Averaged difference quotients. The error and the energy of the discrete convolution are controlled by a quantity which involves only the metric and the reference measure. For $f \in L^p(X, m)$, $r > 0$ and a positive function S, put

$$D_p(f, r) = \int_X \fint_{B(x,r)} |f(x) - f(y)|^p \, dm(y)\, dm(x) \,, \qquad E_{p,S}(f, r) = \frac{D_p(f, r)}{S(r)} \,. \tag{3.17}$$

For $S(r) = r^p$, $E_{p,S}(f,r)$ is the approximate energy of Korevaar and Schoen [112] at the scale r; Chapter 4 studies its behaviour as $r \to 0$. If m is doubling at small radii, then $D_p(\cdot,r)^{1/p}$ is a continuous seminorm on $L^p(X,m)$ and, uniformly at small radii,

$$D_p(f,r) \le C\,\|f\|^p_{L^p(X,m)}\,. \tag{3.18}$$

Indeed, $|f(x)-f(y)|^p \le 2^{p-1}(|f(x)|^p + |f(y)|^p)$, and in the term involving $|f(y)|^p$ one interchanges the order of integration and uses $m(B(x,r)) \ge c\,m(B(y,r))$ for $d(x,y) < r$.

The lemma. The approximation requires an energy bound for cutoffs at one exponent only. We state it in this generality; this is the form used in the rank theorem of Chapter 5. In the next lemma, $\Psi_{C,p}$ may be any positive function satisfying (3.2). Under VD and $\mathrm{B}(\Psi_{C,2})$ the hypothesis (3.19) holds with the cutoff scale (3.3) and the cutoffs $\chi_{B(x,r)} = \psi^r_x$, by (3.10).

Lemma 3.5. *Let $1 \le p < \infty$. Assume that m is doubling at small radii and that $\Psi_{C,p}$ satisfies (3.2). Suppose that each ball $B = B(x,r)$ of sufficiently small radius admits $\chi_B \in \mathcal{C}_p$ with*

$$\mathbf{1}_B \le \chi_B \le \mathbf{1}_{2B}\,, \qquad \mathrm{Lip}(\chi_B) \le \frac{C}{r}\,, \qquad \int_S |\nabla\chi_B|^p\,d\nu \le C\,\frac{m(B)}{\Psi_{C,p}(r)}\,. \tag{3.19}$$

Then $\mathcal{C}_p$ is uniformly dense in $C_c(X)$. Given a maximal ε-separated set $\{x_i\}$, put

$$B_i = B(x_i,\varepsilon)\,, \qquad \varphi_i = \frac{\chi_{B_i}}{\sum_j \chi_{B_j}}\,.$$

For sufficiently small ε, the functions φ_i form a locally finite partition of unity in $\mathcal{C}_p$, with $\mathrm{Lip}(\varphi_i) \le C/\varepsilon$ and $\operatorname{supp}\varphi_i \subset \overline{2B_i}$. For $f \in L^p(X,m)$ write

$$f_\varepsilon = \sum_i f_{B_i}\,\varphi_i\,, \qquad G_\varepsilon f = \sum_i f_{B_i}\,\nabla\varphi_i\,, \qquad F_N = \sum_{i\le N} f_{B_i}\,\varphi_i\,. \tag{3.20}$$

For every $f \in L^p(X,m)$ and every sufficiently small ε the following hold.

(i) *$F_N \in \mathcal{C}_p$, and $F_N \to f_\varepsilon$ in $L^p(m)$ and $\nabla F_N \to G_\varepsilon f$ in $L^p(\nu)$ as $N \to \infty$.*

(ii) *One has*

$$\|f - f_\varepsilon\|^p_{L^p(m)} \le C\,D_p(f,5\varepsilon)\,, \qquad \|G_\varepsilon f\|^p_{L^p(\nu)} \le C\,\frac{D_p(f,5\varepsilon)}{\Psi_{C,p}(\varepsilon)}\,. \tag{3.21}$$

Moreover $D_p(f,r) \to 0$ as $r \downarrow 0$.

(iii) *If $f \in \mathrm{Lip}_c(X)$, then $f_\varepsilon \in \mathcal{C}_p$, $\nabla f_\varepsilon = G_\varepsilon f$, and*

$$\|f_\varepsilon\|_\infty \le \|f\|_\infty\,, \qquad \|f - f_\varepsilon\|_\infty \le 3\varepsilon\,\mathrm{Lip}(f)\,, \qquad \mathrm{Lip}(f_\varepsilon) \le C\,\mathrm{Lip}(f)\,.$$

Under global doubling and (3.19) at all radii, the estimates hold for every $\varepsilon > 0$.

We write G_ε for $G_\varepsilon f$ when the function is clear. The function f_ε is the *discrete convolution* of f at the scale ε, and the field $G_\varepsilon f$ is its *discrete gradient*. Both are defined for every $f \in L^p(X,m)$. By (i), the pair $(f_\varepsilon, G_\varepsilon f)$ belongs to the closure of the graph of $(\nabla, \mathcal{C}_p)$; if the structure is p-admissible, this means $f_\varepsilon \in W^{1,p}(X,m)$ and $\nabla_p f_\varepsilon = G_\varepsilon f$. By (ii), the energy of the discrete convolution is bounded by an averaged difference quotient of f at the scale 5ε, normalized by the cutoff scale. This is the estimate from which the closability theorem and the Korevaar–Schoen characterizations of Chapter 4 are derived.

Proof. We denote by $N(\lambda)$ the overlap constant of Lemma 3.2.

Step 1: density. The set $\mathcal{C}_p$ is a subalgebra of the algebra $C_0(X)$ of continuous functions vanishing at infinity. It separates points and vanishes nowhere: for $x \neq y$, the cutoff of a ball $B(x,r)$ with $2r < d(x,y)$ is one at x and zero at y. The Stone–Weierstrass theorem gives the uniform density of $\mathcal{C}_p$ in $C_0(X)$, hence in $C_c(X)$.

Step 2: the partition. By Lemma 3.2, $1 \leq \sum_j \chi_{B_j} \leq N(2)$ on X. At most $2N(2)$ terms of this sum are nonzero at one of two given points, so the sum has Lipschitz constant at most C/ε, and so has each φ_i. The support of φ_i is the support of χ_{B_i}, a compact subset of $\overline{2B_i}$. On $2B_i$ the sum $\sum_j \chi_{B_j}$ coincides with the finite sum $\sigma_i = \sum_{d(x_i,x_j)<4\varepsilon} \chi_{B_j} \in \mathcal{C}_p$. Choose $\rho \in C^1(\mathbb{R})$ with $\rho(t) = 1/t$ for $t \geq 1$. Then $\varphi_i = \chi_{B_i}\, \rho(\sigma_i)$ on X, so $\varphi_i \in \mathcal{C}_p$, and by Lemma 1.5,

$$|\nabla \varphi_i| \leq |\nabla \chi_{B_i}| + \sum_{d(x_i,x_j)<4\varepsilon} |\nabla \chi_{B_j}| \qquad \nu\text{-almost everywhere on } 2B_i\,,$$
$$\nabla \varphi_i = 0 \qquad \nu\text{-almost everywhere outside } 2B_i\,.$$

The number of indices in this sum is at most $N(4)$, and $m(B_j) \leq C\, m(B_i)$ for these indices by doubling. With (3.19) this gives

$$\int_{\mathcal{S}} |\nabla \varphi_i|^p \, d\nu \leq C \sum_{d(x_i,x_j)<4\varepsilon} \int_{\mathcal{S}} |\nabla \chi_{B_j}|^p \, d\nu \leq C\, \frac{m(B_i)}{\Psi_{C,p}(\varepsilon)}\,. \tag{3.22}$$

Finally, on each ball B_j the sum $\sum_i \varphi_i = 1$ reduces to the finite sum over the indices with $d(x_i,x_j) < 3\varepsilon$, and the remaining functions φ_i vanish on B_j. By locality (Lemma 1.5(ii)),

$$\sum_i \nabla \varphi_i = 0 \qquad \nu\text{-almost everywhere.} \tag{3.23}$$

Step 3: proof of (i). At each point at most $N(2)$ of the functions φ_i are nonzero, and at ν-almost every point at most $N(2)$ of the fields $\nabla\varphi_i$ are nonzero. By Jensen's inequality $|f_{B_i}|^p\, m(B_i) \leq \int_{B_i} |f|^p \, dm$. With (3.22) and doubling this gives

$$\Big\| \sum_{i>N} f_{B_i} \varphi_i \Big\|^p_{L^p(m)} \leq C \sum_{i>N} |f_{B_i}|^p\, m(2B_i) \leq C \sum_{i>N} \int_{B_i} |f|^p \, dm\,,$$
$$\Big\| \sum_{i>N} f_{B_i} \nabla \varphi_i \Big\|^p_{L^p(\nu)} \leq C \sum_{i>N} |f_{B_i}|^p \int_{\mathcal{S}} |\nabla \varphi_i|^p \, d\nu \leq \frac{C}{\Psi_{C,p}(\varepsilon)} \sum_{i>N} \int_{B_i} |f|^p \, dm\,.$$

Both tend to zero as $N \to \infty$, because $\sum_i \int_{B_i} |f|^p \, dm \leq N(1) \|f\|^p_{L^p(m)}$.

Step 4: proof of (ii). If $\varphi_i(x) \neq 0$, then $x \in 2B_i$, hence $B_i \subset B(x, 5\varepsilon)$ and $m(B(x,5\varepsilon)) \leq C\, m(B_i)$. Jensen's inequality gives

$$|f(x) - f_{B_i}|^p \leq \fint_{B_i} |f(x) - f(y)|^p \, dm(y) \leq C \fint_{B(x,5\varepsilon)} |f(x) - f(y)|^p \, dm(y)\,.$$

Since $f(x) - f_\varepsilon(x) = \sum_i (f(x) - f_{B_i}) \varphi_i(x)$ is a convex combination, integrating in x gives the first inequality in (3.21).

The second inequality uses the cancellation (3.23), as in [146, proof of Theorem 3.12, equations (3.23)–(3.25)]. Fix j. Since $\nabla \varphi_i = 0$ almost everywhere on B_j unless $d(x_i,x_j) < 3\varepsilon$, (3.23) gives, ν-almost everywhere on B_j,

$$G_\varepsilon f = \sum_{d(x_i,x_j)<3\varepsilon} (f_{B_i} - f_{B_j})\, \nabla \varphi_i\,.$$

For these indices $B_i \subset B(x,5\varepsilon)$ whenever $x \in B_j$, and Jensen's inequality and doubling give

$$|f_{B_i}-f_{B_j}|^p \le \fint_{B_j}\fint_{B_i} |f(x)-f(y)|^p\,dm(y)\,dm(x) \le \frac{C}{m(B_j)}\int_{B_j}\fint_{B(x,5\varepsilon)} |f(x)-f(y)|^p\,dm(y)\,dm(x)\,.$$

We multiply by (3.22), use $m(B_i)\le C\,m(B_j)$ and the bound $N(3)$ on the number of indices:

$$\int_{B_j\cap\mathcal{S}} |G_\varepsilon f|^p\,d\nu \le \frac{C}{\Psi_{C,p}(\varepsilon)}\int_{B_j}\fint_{B(x,5\varepsilon)} |f(x)-f(y)|^p\,dm(y)\,dm(x)\,.$$

Summing over j, with the bounded overlap of the balls B_j, gives the second inequality.

Finally, for $f\in C_c(X)$ the uniform continuity gives $D_p(f,r)\to 0$, since the r-neighbourhood of $\operatorname{supp} f$ lies in a fixed compact set for small r. The general case follows from (3.18) and the density of $C_c(X)$ in $L^p(X,m)$.

Step 5: proof of (iii). Only finitely many balls B_i meet the compact support of f, so f_ε is a finite sum, $f_\varepsilon\in\mathcal{C}_p$ and $\nabla f_\varepsilon = G_\varepsilon f$. The bound $\|f_\varepsilon\|_\infty\le\|f\|_\infty$ is clear. If $x\in\operatorname{supp}\varphi_i$, every point of B_i is at distance at most 3ε from x, so $|f(x)-f_{B_i}|\le 3\varepsilon\operatorname{Lip}(f)$, and the second estimate follows. Let $d(x,y)<\varepsilon$. Since $\sum_i(\varphi_i(x)-\varphi_i(y))=0$,

$$f_\varepsilon(x)-f_\varepsilon(y)=\sum_i (f_{B_i}-f(x))\,(\varphi_i(x)-\varphi_i(y))\,.$$

A term is nonzero only if x or y belongs to $\operatorname{supp}\varphi_i$, in which case $|f_{B_i}-f(x)|\le 4\varepsilon\operatorname{Lip}(f)$. There are at most $2N(2)$ such terms and $|\varphi_i(x)-\varphi_i(y)|\le Cd(x,y)/\varepsilon$, hence $|f_\varepsilon(x)-f_\varepsilon(y)|\le C\operatorname{Lip}(f)\,d(x,y)$. For $d(x,y)\ge\varepsilon$ the same bound follows from $\|f-f_\varepsilon\|_\infty\le 3\varepsilon\operatorname{Lip}(f)$. □

One partition for all exponents. Under VD and $\mathrm{B}(\Psi_{C,2})$ the cutoffs ψ_x^r satisfy (3.19) at every exponent, by (3.10). The resulting partition of unity does not depend on p, and for it Lemma 3.5 gives, for every $1\le p<\infty$,

$$\int_{\mathcal{S}}|\nabla\varphi_i|\,d\nu \le C\frac{m(B_i)}{\Psi_{C,1}(\varepsilon)}\,,\qquad \int_{\mathcal{S}}|\nabla\varphi_i|^p\,d\nu \le C\frac{m(B_i)}{\Psi_{C,p}(\varepsilon)}\,,\qquad \sum_i\nabla\varphi_i=0\quad \nu\text{-almost everywhere.} \tag{3.24}$$

In particular the discrete convolutions of compactly supported Lipschitz functions belong to $\mathcal{C}_1$, and the discrete gradient $G_\varepsilon f$ is the same field at every exponent. This is what allows the comparison of the closed gradients at different exponents in Proposition 4.7.

3.4 Comparison of the scales

The hypotheses restrict the scale functions. First, because the gradient is bounded by the Lipschitz constant, the cutoff scale is at most quadratic, and in the quadratic case the gradient measure is absolutely continuous with respect to the reference measure. This is the counterpart, for a prescribed gradient measure, of the dichotomy for the energy measures of diffusions with sub-Gaussian heat kernel estimates: they are singular when the space-time scale is r^β with $\beta>2$ [96], and absolutely continuous in the Gaussian case [117]. The proof below uses only the discrete convolution; no heat kernel or closed form is involved.

Proposition 3.6. *Assume* VD *and* $\mathrm{B}(\Psi_{C,2})$*, with* $\nu\neq 0$.

(i) $\Psi_{C,2}(r)\le Cr^2$ *at sufficiently small radii.*

(ii) *If* $\ell = \liminf_{r\downarrow 0} r^2/\Psi_{C,2}(r) < \infty$, *then* $\nu \ll m$, *and for every* $u \in \mathrm{Lip}_b(X)$

$$|\nabla u|\,\nu \le C\,\ell\,\mathrm{Lip}(u)\,m \qquad \text{as measures on } X\,. \tag{3.25}$$

(iii) *Under the additional hypothesis* $\mathrm{V}(\Phi)$, $\Theta(r) \ge c\,\Phi(r)$ *at small radii. If* $\Phi(r) \asymp r^Q$ *and* $\Theta(r) \asymp r^s$, *then* $s \le Q$, *with* $s < Q$ *if* $\nu \not\ll m$.

Proof. If $\ell = \infty$, then $\Psi_{C,2}(r) \le r^2$ for small r and (i) holds. Suppose that $\ell < \infty$. Let $u \in \mathrm{Lip}_b(X)$ and let $A \subset X$ be compact. Let $u_\varepsilon = \sum_i u_{B_i}\varphi_i$ be the discrete convolution of u for the partition of unity (3.24). The sum is locally finite. The pointwise argument of Lemma 3.5(iii) does not use the compactness of the support, and shows that $u_\varepsilon \in \mathrm{Lip}_b(X)$, with $\|u_\varepsilon\|_\infty \le \|u\|_\infty$, $\mathrm{Lip}(u_\varepsilon) \le C\,\mathrm{Lip}(u)$ and $\|u - u_\varepsilon\|_\infty \le 3\varepsilon\,\mathrm{Lip}(u)$. By locality and the cancellation in (3.24),

$$\nabla u_\varepsilon(z) = \sum_i (u_{B_i} - u(z))\nabla\varphi_i(z)\,, \qquad |u_{B_i} - u(z)| \le 3\varepsilon\,\mathrm{Lip}(u) \quad (z \in \operatorname{supp}\varphi_i)\,.$$

If $\operatorname{supp}\varphi_i$ meets A, then B_i lies in the neighbourhood $A_{3\varepsilon} = \{d(\cdot, A) < 3\varepsilon\}$. The bounded overlap of the balls B_i and (3.24) therefore imply

$$\int_A |\nabla u_\varepsilon|\,d\nu \le C\,\mathrm{Lip}(u)\frac{\varepsilon}{\Psi_{C,1}(\varepsilon)} \sum_{i:\,\operatorname{supp}\varphi_i \cap A \ne \varnothing} m(B_i) \le C\,\mathrm{Lip}(u)\frac{\varepsilon^2}{\Psi_{C,2}(\varepsilon)}\,m(A_{3\varepsilon})\,.$$

As $\varepsilon \to 0$, $u_\varepsilon \to u$ uniformly with a common Lipschitz bound, and $m(A_{3\varepsilon}) \to m(A)$, since small neighbourhoods of A are relatively compact. Applying Lemma 3.4 along a sequence $\varepsilon_k \to 0$ with $\varepsilon_k^2/\Psi_{C,2}(\varepsilon_k) \to \ell$ gives

$$\int_A |\nabla u|\,d\nu \le C\,\ell\,\mathrm{Lip}(u)\,m(A) \qquad (A \subset X \text{ compact})\,.$$

In particular the measure $|\nabla u|\nu$ is finite on compact sets, hence Radon, and by inner regularity the bound extends to all Borel sets. This is (3.25).

Let (u_k) be a sequence in $\mathrm{Lip}_b(X)$ whose gradients span the fibres (Lemma 1.6). If $m(A) = 0$, then by (3.25) every ∇u_k vanishes ν-almost everywhere on A; since the fibres are nonzero, $\nu(A) = 0$. Thus $\nu \ll m$. If $\ell = 0$, the same argument with $A = X$ shows that $\nu = 0$, which is excluded. Hence $\ell > 0$, and $\Psi_{C,2}(r) \le Cr^2$ at small radii. This proves (i) and (ii), and (iii) follows from (3.14): $\Theta/\Phi = r^2/\Psi_{C,2}$ is bounded below, and it tends to infinity if $\nu \not\ll m$. □

Second, on a connected space a cutoff tested in the Poincaré inequality shows that the cutoff scale is dominated by the Poincaré scale. Together with Definition 3.1, this shows that matching scales are comparable.

Proposition 3.7. *Assume* VD, $\mathrm{B}(\Psi_{C,2})$ *and* $\mathrm{PI}_p(\Psi_{P,p})$, *and suppose that* X *is connected with positive diameter. Then* $\Psi_{C,p}(r) \le C\Psi_{P,p}(r)$ *at sufficiently small radii. Under global hypotheses on an unbounded connected space this holds at every radius.*

Proof. Let $M > 2$ be the constant of Lemma 3.3, let $B = B(x, r)$, and let $\psi = \psi_x^r$. Since $\psi \le \mathbf{1}_{B(x,2r)}$, the mean of ψ on MB is at most $m(B(x,2r))/m(MB) \le 1/2$, while $\psi = 1$ on B. The Poincaré inequality on the ball MB and (3.10) give

$$2^{-p}\,m(B) \le \int_{MB} |\psi - \psi_{MB}|^p\,dm \le C_P\,\Psi_{P,p}(Mr)\int_{\mathcal{S}} |\nabla\psi|^p\,d\nu \le C\,\frac{\Psi_{P,p}(Mr)}{\Psi_{C,p}(r)}\,m(B)\,.$$

We cancel $m(B)$ and compare $\Psi_{P,p}(Mr)$ with $\Psi_{P,p}(r)$ by (3.1). □

3.5 The distance cutoff

When the gradient measure is dominated by the reference measure, the cutoff condition holds at the quadratic scale, with the simplest cutoffs. For $x \in X$ and $r > 0$, the *distance cutoff* is

$$\psi_x^r(y) = \big(3 - 2d(x,y)/r\big)_+ \wedge 1\,. \tag{3.26}$$

It is equal to one on $B(x,r)$, its support lies in $\overline{B}(x,3r/2) \subset B(x,2r)$, and $\mathrm{Lip}(\psi_x^r) \le 2/r$.

Lemma 3.8. *Let $(\mathcal{S},\nu,\nabla)$ be a gradient structure on (X,d,m), and suppose that the closed balls of radius less than r_0 are compact. Then, for $0 < r < r_0/2$, the distance cutoff belongs to $\mathrm{Lip}_c(X)$ and*

$$\int_{B(y,s)\cap\mathcal{S}} |\nabla\psi_x^r|\,d\nu \le \frac{2}{r}\,\nu(B(y,s)) \qquad (y \in X\,,\ s > 0)\,. \tag{3.27}$$

In particular, if m is doubling and $\nu \le C_0\,m$, then the distance cutoffs satisfy $\mathrm{B}(\Psi_{C,2})$ with $\Psi_{C,2}(r) = r^2$, and $\Psi_{C,p}(r) = r^p$ for every $1 \le p < \infty$.

Proof. By (G), $|\nabla\psi_x^r| \le 2/r$ ν-almost everywhere, which gives (3.27). Suppose that $\nu \le C_0\,m$. By locality, $\nabla\psi_x^r = 0$ ν-almost everywhere outside $B(x,2r)$. Taking $y = x$ and $s = 2r$ in (3.27) and applying doubling,

$$\int_{\mathcal{S}} |\nabla\psi_x^r|\,d\nu \le \frac{2C_0}{r}\,m(B(x,2r)) \le \frac{2C_0C_D}{r}\,m(B(x,r))\,,$$

which is (3.9) with $\Psi_{C,2}(r) = r^2$. □

3.6 Overview of the examples

The examples are constructed in Part III. Table 1 lists those which satisfy VD, the cutoff condition $\mathrm{B}(\Psi_{C,2})$ and Poincaré inequalities at matching scales, with estimates which are uniform in the centre at sufficiently small radii. The column "Exponents" gives the range of p for which $\mathrm{PI}_p(\Psi_{P,p})$ is proved at the matching scale $\Psi_{P,p} = \Psi_{C,p}$; all these exponents are admissible. The last column records a uniform volume function when there is one; it is not required for the other columns.

Space	Rank	Exponents	$\Psi_{C,2}(r)$	$\Psi_{P,p}(r)$	$\Phi(r)$
Smooth or measurable Riemannian metric	n	$p \ge 1$	r^2	r^p	r^n under noncollapsing
Compact equiregular sub-Riemannian manifold	k	$p \ge 1$	r^2	r^p	r^Q
RCD(K,N) space, $N < \infty$	k	$p \ge 1$	r^2	r^p	not required
Cubical inverse limit	n	$p \ge 1$	r^2	r^p	r^Q, $Q > n$
Harmonic Sierpiński gasket	1	$p > 2 - \varepsilon$	r^2	r^p	none
Finite Riemannian gluing	n	$p \ge 1$	r^2	r^p	r^n
Uniform local tree	1	$p \ge 1$	$r\,\Phi(r)$	$r^{p-1}\Phi(r)$	assumed
Tree of cubes	n	$p > n$	r^{Q+2-n}	r^{Q+p-n}	r^Q

Table 1: Examples with volume doubling and matching scales.

Comments on the table.

- *Riemannian metrics* (Chapter 10). Completeness and a lower bound on the Ricci curvature give the three hypotheses. The measurable case assumes uniform ellipticity relative to such a background metric. A uniform lower bound for the volume of unit balls is needed only for $\mathrm{V}(r^n)$.
- *Sub-Riemannian manifolds* (Section 10.3). A Carnot group satisfies the estimates globally. Proposition 10.8 gives further global examples under a nonnegative generalized curvature-dimension bound, including the complete Sasakian manifolds and totally geodesic foliations described there, with Poincaré inequalities for $p > 2 - \varepsilon$. The sub-Laplacian comparison criterion of Proposition 10.9 gives all $p \geq 1$ in the Sasakian and H-type classes specified there, through the measure contraction property.
- RCD *spaces* (Chapter 11). Every finite-dimensional RCD space has the stated local properties; compact noncollapsed $\mathrm{RCD}(K, n)$ spaces satisfy in addition $\mathrm{V}(r^n)$.
- *Cubical inverse limits* (Chapter 12). The inverse limits of Theorem 12.1 realize every integer rank $n \geq 1$ and every real volume exponent $Q > n$, with $m = \nu$.
- *The harmonic gasket* (Chapter 13). The structure is admissible for all finite $p \geq 1$ and has no uniform volume function. The table records the range of exponents for which matching Poincaré inequalities are proved in this book.
- *Gluings and trees* (Chapter 15). The finite gluing has the pieces and the interface of Proposition 15.1. The local trees are those of [22], under the uniform volume assumption of Proposition 15.2; here ν is the length measure and $\Theta(r) = r$.
- *The tree of cubes* (Chapter 15). Here $n \geq 2$, $Q > n$, and the construction depends on a parameter $0 < \sigma < 2$. The gauge is $\Theta(r) = r^n$. The structure is p-admissible exactly for

$$p \geq p_* := \frac{n\sigma}{n - 2 + \sigma}\,, \qquad 1 \leq p < \infty\,;$$

  for $n = 2$ this threshold is $p_* = 2$. The scales match precisely for $p > n$, and Proposition 15.10 gives the two distinct scales at the smaller admissible exponents. In particular the quadratic structure is admissible although its scales do not match.

An example without doubling. Liouville Brownian motion (Chapter 14) has a Euclidean gradient structure of rank two with $m = M_\gamma$, the Gaussian multiplicative chaos measure, and $\nu = dx$. For each fixed $0 < \gamma < 2$, almost surely, it is p-admissible precisely when $p > \gamma^2/2$, among finite $p \geq 1$. For $1 \leq p \leq \gamma^2/2$ the relaxed energy is identically zero; in particular the total variation vanishes identically when $\gamma \geq \sqrt{2}$ (Proposition 14.2). On each fixed compact set the structure satisfies $\mathrm{PI}_2(r^a)$, the cutoff condition $\mathrm{B}(r^b)$ and the cutoff Sobolev inequality $\mathrm{CSA}(r^b)$ for $0 < a < (2 - \gamma)^2/2$ and $b > (2 + \gamma)^2/2$. Global volume doubling fails.

4 Korevaar–Schoen theory

Korevaar and Schoen [112] defined the Sobolev energy of a map on a Riemannian domain through the behaviour, as $r \to 0$, of averaged difference quotients of the type

$$\frac{1}{r^p} \int_X \fint_{B(x,r)} |f(x) - f(y)|^p \, dm(y) \, dm(x) \,,$$

which involve only the metric and the measure; see also [40] for a related characterization of Sobolev functions on Euclidean domains. The relation of this approach to Sobolev spaces on metric measure spaces and to Dirichlet forms is developed in [116, 114, 31]. This chapter compares the energy of a gradient structure with the quantity $D_p(f, r)$ of (3.17), normalized by the Poincaré scale or by the cutoff scale. The comparison rests on two inequalities which use the two hypotheses separately.

- *From the energy to difference quotients.* The Poincaré inequality bounds $D_p(f, r)/\Psi_{P,p}(r)$ by the energy, uniformly in r (Lemma 4.1).
- *From difference quotients to the energy.* The discrete convolution bounds the energy by the lower limit of $D_p(f, r)/\Psi_{C,p}(r)$ as $r \to 0$ (Lemma 4.2).

Both inequalities hold for the relaxed energy $\overline{\mathcal{E}}_p$ of (2.15), before any closability is known. When the two scales are comparable along one sequence of radii tending to zero, the two inequalities close up, and they have the following consequences: the gradient is closable (Theorem 4.3); the Sobolev space for $p > 1$ and the space BV for $p = 1$ are characterized by difference quotients (Theorems 4.4 and 4.8); the core $\mathcal{C}_1$ is dense in every Sobolev space (Corollary 4.5); and the closed gradients at different exponents coincide (Proposition 4.7). Sections 4.6 and 4.7 determine the critical exponents of the Korevaar–Schoen spaces at matching power scales and derive Gagliardo–Nirenberg inequalities.

The proofs by controlled cutoffs are close to [31, Theorem 5.2], [22, Theorems 3.7–3.8] and [146, Section 3]. Two features are specific to the present setting: the Poincaré and cutoff scales are kept separate, and closability is a conclusion, whereas in those references a closed form or a Sobolev space is given in advance.

4.1 The two basic inequalities

Lemma 4.1. *If m is doubling at small radii and $\mathrm{PI}_p(\Psi_{P,p})$ holds with $\Psi_{P,p}$ satisfying (3.2), then*

$$D_p(u, r) \le C \, \Psi_{P,p}(r) \int_{\mathcal{S}} |\nabla u|^p \, d\nu \qquad (u \in \mathcal{C}_p) \tag{4.1}$$

for sufficiently small r, or for all $r > 0$ under global hypotheses. For a p-admissible structure the inequality extends to $f \in W^{1,p}(X, m)$, with $\nabla_p f$ in place of ∇u.

Proof. This is the covering argument of [31, proof of Theorem 5.1], also used in [146, Propositions 3.3 and 3.5]. Cover X by the balls $B_i = B(x_i, r)$ of Lemma 3.2. For $x \in B_i$ one has $B(x, r) \subset 2B_i$ and $m(2B_i) \le C \, m(B(x, r))$. Writing $u(x) - u(y) = (u(x) - u_{2B_i}) - (u(y) - u_{2B_i})$ and integrating first in $y \in B(x, r)$ and then in $x \in B_i$, we get

$$\int_{B_i} \fint_{B(x,r)} |u(x) - u(y)|^p \, dm(y) \, dm(x) \le C \int_{2B_i} |u - u_{2B_i}|^p \, dm \le C \, \Psi_{P,p}(r) \int_{2\Lambda B_i \cap \mathcal{S}} |\nabla u|^p \, d\nu \,,$$

by (3.7) on the ball $2B_i$ and (3.2). Summing over i, with the bounded overlap of the balls $2\Lambda B_i$, gives (4.1). The extension follows from the continuity of $D_p(\cdot, r)^{1/p}$ on $L^p(X, m)$. □

Lemma 4.2. *Let $1 \le p < \infty$, and assume the hypotheses of Lemma 3.5.*

(i) *For every $f \in L^p(X,m)$,*
$$\overline{\mathcal{E}}_p(f) \le C \liminf_{r\downarrow 0} E_{p,\Psi_{C,p}}(f,r)\,.$$

(ii) *For every $u \in \mathcal{C}_p$,*
$$\int_{\mathcal{S}} |\nabla u|^p \, d\nu \le C \liminf_{r\downarrow 0} E_{p,\Psi_{C,p}}(u,r)\,.$$

(iii) *If $\mathrm{PI}_p(\Psi_{P,p})$ also holds, with $\Psi_{P,p}$ satisfying (3.2), then for sufficiently small $r_1 > 0$, or $r_1 = \infty$ under global hypotheses,*
$$\sup_{0<r\le r_1} E_{p,\Psi_{P,p}}(f,r) \le C\,\overline{\mathcal{E}}_p(f) \qquad (f \in L^p(X,m))\,.$$

Assertion (ii) is not a special case of (i): on the core, $\overline{\mathcal{E}}_p(u) \le \int_{\mathcal{S}} |\nabla u|^p \, d\nu$, and the inequality may be strict when the gradient is not closable.

Proof. (i) Assume that the lower limit is finite, and choose $\varepsilon_k \downarrow 0$ such that $E_{p,\Psi_{C,p}}(f,5\varepsilon_k)$ converges to it. By Lemma 3.5(ii) and (3.2),
$$f_{\varepsilon_k} \longrightarrow f \ \text{ in } L^p(m)\,, \qquad \|G_{\varepsilon_k} f\|^p_{L^p(\nu)} \le C\, E_{p,\Psi_{C,p}}(f,5\varepsilon_k)\,.$$
By Lemma 3.5(i), each pair $(f_{\varepsilon_k}, G_{\varepsilon_k} f)$ is the limit in $L^p(m)\times L^p(\nu)$ of pairs $(F_N, \nabla F_N)$ with $F_N \in \mathcal{C}_p$. A diagonal sequence $u_k \in \mathcal{C}_p$ satisfies $u_k \to f$ in $L^p(m)$ and $\liminf_k \int_{\mathcal{S}} |\nabla u_k|^p \, d\nu \le C \liminf_{r\downarrow 0} E_{p,\Psi_{C,p}}(f,r)$.

(ii) For $u \in \mathcal{C}_p$, the functions $u_{\varepsilon_k} \in \mathcal{C}_p$ converge uniformly to u with a common Lipschitz bound, by Lemma 3.5(iii), and $\nabla u_{\varepsilon_k} = G_{\varepsilon_k} u$. Lemma 3.4 gives $\int_{\mathcal{S}} |\nabla u|^p \, d\nu \le \liminf_k \|G_{\varepsilon_k} u\|^p_{L^p(\nu)}$, and the right-hand side is bounded as in (i).

(iii) Let $u_k \in \mathcal{C}_p$ with $u_k \to f$ in $L^p(m)$. Lemma 4.1 gives $D_p(u_k,r) \le C\Psi_{P,p}(r) \int_{\mathcal{S}} |\nabla u_k|^p \, d\nu$, and $D_p(u_k,r) \to D_p(f,r)$ for each fixed r. Taking the lower limit in k and then the infimum over the sequences (u_k) gives the assertion. □

4.2 Closability

The relaxed energy of a function of the core is at most its energy. Closability is essentially the converse inequality, up to a constant, and this converse follows from the two basic inequalities when the scales are comparable along a sequence.

Theorem 4.3. *Let $1 \le p < \infty$. Assume the hypotheses of Lemma 3.5 and $\mathrm{PI}_p(\Psi_{P,p})$, with $\Psi_{P,p}$ satisfying (3.2). If*
$$\liminf_{r\downarrow 0} \frac{\Psi_{P,p}(r)}{\Psi_{C,p}(r)} < \infty\,, \tag{4.2}$$
then the structure is p-admissible. In particular, this holds under VD, $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}_p(\Psi_{P,p})$ *and* (4.2).

Proof. Condition (A1) is part of Lemma 3.5. For (A2), choose $r_j \downarrow 0$ such that $\Psi_{P,p}(r_j) \le C_0 \Psi_{C,p}(r_j)$, all within the range of radii of Lemma 4.2. Lemma 4.2(iii) gives
$$\sup_j E_{p,\Psi_{C,p}}(f,r_j) \le C\,\overline{\mathcal{E}}_p(f) \qquad (f \in L^p(X,m))\,, \tag{4.3}$$

and with Lemma 4.2(ii) we obtain

$$\int_{\mathcal{S}} |\nabla u|^p \, d\nu \le C \liminf_{j\to\infty} E_{p,\Psi_{C,p}}(u, r_j) \le C' \, \overline{\mathcal{E}}_p(u) \qquad (u \in \mathcal{C}_p) \,. \tag{4.4}$$

Let $u_k \in \mathcal{C}_p$ tend to zero in $L^p(m)$, with $\nabla u_k \to G$ in $L^p(\nu)$. For fixed j, the sequence $(u_j - u_k)_k$ lies in $\mathcal{C}_p$ and converges to u_j in $L^p(m)$. By the definition of the relaxed energy and (4.4),

$$\|\nabla u_j\|^p_{L^p(\nu)} \le C' \, \overline{\mathcal{E}}_p(u_j) \le C' \lim_{k\to\infty} \|\nabla u_j - \nabla u_k\|^p_{L^p(\nu)} = C' \|\nabla u_j - G\|^p_{L^p(\nu)} \,.$$

Letting $j \to \infty$ gives $\|G\|_{L^p(\nu)} = 0$. The last assertion follows from (3.10). □

Condition (4.2) holds in particular when the scales match in the sense of (3.4). It is weaker: it requires the comparison only along one sequence of radii. In the model of Section 1.1, the Poincaré inequality is the oscillation bound (1.1), and Theorem 4.3 is the general form of the argument given there.

Under the matching condition (3.4), Lemma 4.1 gives, at all sufficiently small radii,

$$D_p(u, r) \le C \, \Psi_{C,p}(r) \int_{\mathcal{S}} |\nabla u|^p \, d\nu \qquad (u \in \mathcal{C}_p) \,. \tag{4.5}$$

4.3 Characterization by difference quotients

For a regular p-energy form satisfying his standing axioms, Shimizu [146, Corollaries 1.14–1.15] proves characterizations closely related to the next theorem, under volume doubling, a Poincaré inequality and an upper capacity estimate at one scale. Here the two inequalities are stated for the relaxed energy at possibly different scales, and the exponent $p = 1$ is included.

Theorem 4.4. *Let $1 \le p < \infty$. Assume the hypotheses of Lemma 3.5 and $\mathrm{PI}_p(\Psi_{P,p})$, with $\Psi_{P,p}$ satisfying (3.2). Let $r_1 > 0$ be sufficiently small, or let $r_1 = \infty$ if all the hypotheses hold at every radius.*

(i) *For every $f \in L^p(X, m)$,*

$$\sup_{0<r\le r_1} \frac{D_p(f, r)}{\Psi_{P,p}(r)} \le C \, \overline{\mathcal{E}}_p(f) \,, \qquad \overline{\mathcal{E}}_p(f) \le C \liminf_{r\to 0} \frac{D_p(f, r)}{\Psi_{C,p}(r)} \,. \tag{4.6}$$

The second inequality does not require $\mathrm{PI}_p(\Psi_{P,p})$. These conclusions hold, in particular, under VD, $\mathrm{B}(\Psi_{C,2})$ and $\mathrm{PI}_p(\Psi_{P,p})$, without any admissibility assumption.

(ii) *If $p > 1$ and the structure is p-admissible, then $\overline{\mathcal{E}}_p = \mathcal{E}_p$ on $L^p(X, m)$.*

(iii) *Assume (4.2), and let $r_j \downarrow 0$ be any sequence in $(0, r_1]$ on which $\Psi_{P,p}(r_j)/\Psi_{C,p}(r_j)$ is bounded. Then, for every $f \in L^p(X, m)$,*

$$\sup_j E_{p,\Psi_{C,p}}(f, r_j) \le C \, \overline{\mathcal{E}}_p(f) \le C' \liminf_{r\downarrow 0} E_{p,\Psi_{C,p}}(f, r) \le C' \liminf_{j\to\infty} E_{p,\Psi_{C,p}}(f, r_j) \,. \tag{4.7}$$

For $p > 1$, the finiteness of any of these quantities characterizes $W^{1,p}(X, m)$. For $p = 1$, it characterizes $BV(X, m)$.

Proof. (i) is Lemma 4.2. For (ii), approximation in the graph norm gives $\overline{\mathcal{E}}_p \le \mathcal{E}_p$, and the lower semicontinuity of $\mathcal{E}_p$ (Proposition 2.5(ii)) gives the reverse inequality. For (iii), (4.6) along the sequence (r_j) gives (4.7). By Theorem 4.3 the structure is p-admissible, so that for $p > 1$ the middle term is $\mathcal{E}_p(f)$, which is finite exactly on $W^{1,p}(X, m)$. At $p = 1$, $\overline{\mathcal{E}}_1(f) = \|Df\|(X)$ by definition. □

Matching scales and weak monotonicity. At matching scales the supremum in (4.7) may be taken over all sufficiently small radii:

$$\sup_{0<r\le r_1} E_{p,\Psi_{C,p}}(f,r) \asymp \overline{\mathcal{E}}_p(f) \asymp \liminf_{r\downarrow 0} E_{p,\Psi_{C,p}}(f,r) \qquad (f\in L^p(X,m))\,. \tag{4.8}$$

In particular the supremum over the radii of the normalized difference quotients is bounded by a constant times their lower limit. This property of a space and of a scale is the *weak monotonicity* property introduced in [31], where it is denoted $\mathcal{P}(p,\alpha)$ for the scale $r^{p\alpha}$. It is the source of the strong Sobolev inequalities of Section 4.7.

Examples. On a Riemannian manifold $\Psi_{C,p}(r)=r^p$, and (4.8) is the classical comparison of the gradient energy with averaged difference quotients [112, 40]. On a local tree $\Psi_{C,p}(r)=r^{p-1}\Phi(r)$, and (4.8) is [22, Theorem 3.7]. Under V(Φ), the normalization is $\Psi_{C,p}(r)=r^p\,\Phi(r)/\Theta(r)$, by (3.14).

Corollary 4.5. *Let $1<p<\infty$ and assume* VD, B($\Psi_{C,2}$), PI$_p$($\Psi_{P,p}$) *and* (4.2). *Then $\mathcal{C}_1$ is dense in $W^{1,p}(X,m)$.*

Proof. Let W_1 be the closure of $\mathcal{C}_1$ in $W^{1,p}(X,m)$, and let $f\in W^{1,p}(X,m)$. We use the partitions of unity (3.24), which belong to $\mathcal{C}_1$. By Lemma 3.5(i), $f_\varepsilon\in W_1$ and $\nabla_p f_\varepsilon = G_\varepsilon f$. Choose r_j as in Theorem 4.4(iii) and put $\varepsilon_j=r_j/5$. By Lemma 3.5(ii) and (4.7), $f_{\varepsilon_j}\to f$ in $L^p(X,m)$ and $\|G_{\varepsilon_j}f\|^p_{L^p(\nu)}\le C\,\mathcal{E}_p(f)$. The sequence (f_{ε_j}) is bounded in the reflexive space $W^{1,p}(X,m)$, so a subsequence converges weakly, and its weak limit is f. Since W_1 is a closed subspace, it is weakly closed, and $f\in W_1$. □

The core $\mathcal{C}_1$ does not depend on p. Corollary 4.5 thus provides one space of test functions which is dense in all the Sobolev spaces at exponents $p>1$ satisfying (4.2).

4.4 Discrete gradients and compatibility of the exponents

When two exponents p and q are admissible, the closed gradients ∇_p and ∇_q are closures of the same operator in different topologies, and it is not obvious that they agree on a function which belongs to both Sobolev spaces. At matching scales they do. The reason is that the discrete gradient $G_\varepsilon f$ of (3.20), for the partition of unity (3.24), is defined without reference to the exponent, and converges to the closed gradient at each exponent.

Lemma 4.6. *Assume* VD, B($\Psi_{C,2}$), PI$_p$($\Psi_{P,p}$) *and* (3.4) *at an exponent $1\le p<\infty$, and let $G_\varepsilon f$ be defined with the partition of unity* (3.24). *For $f\in W^{1,p}(X,m)$ and sufficiently small ε,*

$$\|G_\varepsilon f\|_{L^p(\nu)} \le C\,\|\nabla_p f\|_{L^p(\nu)}\,, \tag{4.9}$$

and, for every bounded section g of H which vanishes outside a set of finite ν-measure,

$$\lim_{\varepsilon\downarrow 0}\int_{\mathcal{S}}\langle G_\varepsilon f, g\rangle\,d\nu = \int_{\mathcal{S}}\langle \nabla_p f, g\rangle\,d\nu\,. \tag{4.10}$$

For $p>1$, $G_\varepsilon f \rightharpoonup \nabla_p f$ weakly in $L^p(\mathcal{S},\mathsf{H},\nu)$.

Proof. The bound (4.9) follows from Lemma 3.5(ii) and from (4.5), extended to $W^{1,p}(X,m)$ by continuity. For $u\in\mathcal{C}_p$, the discrete convolutions u_ε converge uniformly to u with a common

Lipschitz bound, and $\nabla u_\varepsilon = G_\varepsilon u$; since $g \in L^1(\mathcal{S}, \mathsf{H}, \nu)$, (G) gives (4.10). For $f \in W^{1,p}(X,m)$, choose $u_k \in \mathcal{C}_p$ converging to f in the graph norm. Since $g \in L^{p'}(\nu)$ and G_ε is linear,

$$\Big|\int_{\mathcal{S}} \langle G_\varepsilon f - \nabla_p f, g\rangle \, d\nu\Big| \le C\, \|\nabla_p(f-u_k)\|_{L^p(\nu)} \|g\|_{L^{p'}(\nu)} + \Big|\int_{\mathcal{S}} \langle G_\varepsilon u_k - \nabla u_k, g\rangle \, d\nu\Big|,$$

by (4.9) applied to $f - u_k$. Let $\varepsilon \downarrow 0$ and then $k \to \infty$. For $p > 1$ the sections g of the statement are dense in $L^{p'}(\mathcal{S}, \mathsf{H}, \nu)$, and the weak convergence follows from (4.9). □

Proposition 4.7. *Assume* VD *and* $\mathrm{B}(\Psi_{C,2})$*, and assume the Poincaré inequalities with matching scales at two exponents* $p, q \in [1,\infty)$*. For* $f \in W^{1,p}(X,m) \cap W^{1,q}(X,m)$ *one has* $\nabla_p f = \nabla_q f$ *ν-almost everywhere.*

Proof. By Theorem 4.3 the structure is admissible at p and at q. The same partitions of unity serve both exponents, so $G_\varepsilon f$ is the same field at p and at q. By Lemma 4.6, applied at both exponents, $\int_{\mathcal{S}} \langle \nabla_p f, g\rangle \, d\nu = \int_{\mathcal{S}} \langle \nabla_q f, g\rangle \, d\nu$ for every bounded section g which vanishes outside a set of finite ν-measure. Since ν is σ-finite, this implies $\nabla_p f = \nabla_q f$ ν-almost everywhere. □

Proposition 4.7 identifies the gradients on the intersection of two Sobolev spaces; it does not say when a function of $W^{1,p}(X,m)$ belongs to $W^{1,q}(X,m)$.

4.5 The endpoint $p = 1$

At $p = 1$ two spaces must be distinguished: the domain $W^{1,1}(X,m)$ of the closed gradient, and the space $BV(X,m)$ defined by relaxation. Difference quotients characterize the second. The comparison follows the relaxation method of [126, 6] and its version for trees [22, Theorem 3.8]. For BV energies obtained as Γ-limits of metric difference quotients, see [5].

Theorem 4.8. *Assume* VD, $\mathrm{B}(\Psi_{C,2})$ *and* $\mathrm{PI}_1(\Psi_{P,1})$.

(i) *For every* $f \in L^1(X,m)$,

$$\sup_{0<r\le r_1} \frac{D_1(f,r)}{\Psi_{P,1}(r)} \le C\, \|Df\|(X) \qquad \text{and} \qquad \|Df\|(X) \le C \liminf_{r\to 0} \frac{D_1(f,r)}{\Psi_{C,1}(r)}\,.$$

(ii) *Assume in addition* (4.2) *at* $p = 1$*. For every* $u \in \mathcal{C}_1$,

$$C^{-1} \int_{\mathcal{S}} |\nabla u| \, d\nu \le \|Du\|(X) \le \int_{\mathcal{S}} |\nabla u| \, d\nu\,.$$

(iii) *Under the same additional assumption, the structure is* 1*-admissible,* $W^{1,1}(X,m) \subset BV(X,m)$*, and*

$$C^{-1} \int_{\mathcal{S}} |\nabla_1 f| \, d\nu \le \|Df\|(X) \le \int_{\mathcal{S}} |\nabla_1 f| \, d\nu \qquad (f \in W^{1,1}(X,m))\,.$$

Proof. (i) This is Theorem 4.4(i) at $p = 1$, since $\overline{\mathcal{E}}_1(f) = \|Df\|(X)$.

(ii) The constant sequence $u_k = u$ gives the upper bound. The lower bound is (4.4) at $p = 1$.

(iii) The admissibility is Theorem 4.3 at $p = 1$. Let $f \in W^{1,1}(X,m)$ and choose $u_k \in \mathcal{C}_1$ converging to f in the graph norm. The sequence (u_k) is admissible in the definition of $\|Df\|(X)$, and for fixed k the sequence $(u_k - u_j)_j$ is admissible for $u_k - f$; hence

$$\|Df\|(X) \le \int_{\mathcal{S}} |\nabla_1 f| \, d\nu\,, \qquad \|D(u_k - f)\|(X) \le \int_{\mathcal{S}} |\nabla u_k - \nabla_1 f| \, d\nu \longrightarrow 0\,.$$

By the triangle inequality for the total variation, $\|Du_k\|(X) \to \|Df\|(X)$, and (ii) gives

$$C^{-1}\int_{\mathcal{S}} |\nabla_1 f|\, d\nu = C^{-1}\lim_{k\to\infty}\int_{\mathcal{S}} |\nabla u_k|\, d\nu \le \lim_{k\to\infty}\|Du_k\|(X) = \|Df\|(X)\,. \qquad \square$$

On a local tree, [22, Lemma 2.18] gives the equality $\|Df\|(X) = \int_{\mathcal{S}} |\partial f|\, d\nu$ on $W^{1,1}(X,m)$; Theorem 4.8(iii) replaces it by a two-sided bound.

4.6 Critical exponents

Replacing the scale function by a power $r^{p\alpha}$ gives a one-parameter family of spaces. For large α they contain only constants, and for small α they are dense in L^p. The critical exponents separating these regimes were introduced in [31] as invariants of the metric measure space. In this section we compute them from the Poincaré and cutoff scales.

Throughout this section, X is connected and contains at least two points, and m is doubling at small radii. Fix a sufficiently small finite radius $r_1 > 0$. For $1 \le p < \infty$ and $\alpha \ge 0$, the *Korevaar–Schoen space* of order α is

$$KS^{\alpha,p}(X) = \left\{ f \in L^p(X,m) : [f]^p_{KS^{\alpha,p}} := \sup_{0<r\le r_1} \frac{D_p(f,r)}{r^{p\alpha}} < \infty \right\}. \tag{4.11}$$

The space does not depend on the choice of a sufficiently small r_1: the contribution of the radii bounded away from zero is controlled by (3.18). The same bound shows that $f \in KS^{\alpha,p}(X)$ if and only if $\limsup_{r\downarrow 0} r^{-p\alpha} D_p(f,r) < \infty$, as in [31, Lemma 3.2 and Section 4.2]. The finiteness of the lower limit is a different condition; the weak monotonicity (4.8) identifies the two at matching scales.

Definition 4.9. The *nontriviality critical exponent* and the *density critical exponent* are, respectively,

$$\alpha_p = \sup\{\alpha \ge 0 : KS^{\alpha,p}(X) \text{ contains a function which is not } m\text{-almost everywhere constant}\}\,,$$
$$\alpha_p^{\mathrm{dens}} = \sup\{\alpha \ge 0 : KS^{\alpha,p}(X) \text{ is dense in } L^p(X,m)\}\,.$$

The first definition is the formulation at small radii of [31, Definition 4.1]. The second records the stronger requirement of density. Since the spaces decrease as α increases, these definitions specify thresholds; they do not assert either property at the threshold itself. If $m(X) = \infty$, the only constant in $L^p(X,m)$ is zero. In either case the constants form a closed proper subspace of $L^p(X,m)$, so

$$1 \le \alpha_p^{\mathrm{dens}} \le \alpha_p\,. \tag{4.12}$$

Indeed, compactly supported Lipschitz functions are dense in $L^p(X,m)$ and belong to $KS^{1,p}(X)$, because their oscillation on a ball of radius r is at most their Lipschitz constant times $2r$.

The dependence on p is controlled by truncation and by power transforms [31, Theorem 4.1 and Lemma 4.3].

Proposition 4.10. *For $1 \le p \le q < \infty$,*

$$p\,\alpha_p \le q\,\alpha_q\,, \qquad \alpha_q \le \alpha_p\,.$$

Proof. If $f \in KS^{\alpha,p}(X)$ is nonconstant, then $f_N = (-N) \vee (f \wedge N)$ is nonconstant for sufficiently large N, belongs to L^q, and

$$D_q(f_N, r) \le (2N)^{q-p} D_p(f, r)\,.$$

Thus $f_N \in KS^{p\alpha/q,q}(X)$, which proves the first inequality. Conversely, for a nonconstant $f \in KS^{\alpha,q}(X)$ the signed power $g = \operatorname{sgn}(f)|f|^{q/p}$ belongs to L^p and is nonconstant. The elementary inequality $|g(x) - g(y)| \le \frac{q}{p}|f(x) - f(y)|\,(|f(x)|^{q/p-1} + |f(y)|^{q/p-1})$, Hölder's inequality with the exponents q/p and $q/(q-p)$, and doubling give

$$D_p(g, r) \le C\,\|f\|_{L^q}^{q-p} D_q(f, r)^{p/q}\,.$$

Hence $g \in KS^{\alpha,p}(X)$, proving the second inequality. □

Bounds from the two scales. The two scales give opposite bounds on the critical exponents. The Poincaré inequality supplies a dense class of functions with controlled oscillation, which bounds the critical exponents from below. The cutoff approximation shows that a function whose oscillation decays too fast has zero relaxed energy and is constant, which bounds them from above.

Proposition 4.11. *Assume* $\mathrm{B}(\Psi_{C,2})$ *and* $\mathrm{PI}_p(\Psi_{P,p})$*, with the scale hypotheses of Theorem 4.4. Let* S *be a positive function satisfying* (3.2)*.*

(i) *If* $\Psi_{P,p}(r) \le CS(r)$ *at small radii, then*

$$\left\{ f \in L^p(X,m) : \sup_{0<r\le r_1} E_{p,S}(f,r) < \infty \right\}$$

is dense in $L^p(X,m)$*.*

(ii) *If* $S(r)/\Psi_{C,p}(r) \to 0$ *as* $r \downarrow 0$*, then every* $f \in L^p(X,m)$ *with* $\sup_{0<r\le r_1} E_{p,S}(f,r) < \infty$ *is* m*-almost everywhere constant.*

Consequently, if

$$\Psi_{P,p}(r) \asymp r^{\beta_P}\,, \qquad \Psi_{C,p}(r) \asymp r^{\beta_C}$$

at small radii, with $\beta_P, \beta_C > 0$*, then*

$$\max\left\{1, \frac{\beta_P}{p}\right\} \le \alpha_p^{\mathrm{dens}} \le \alpha_p \le \frac{\beta_C}{p}\,. \tag{4.13}$$

Proof. (i) By Lemma 4.1, the class contains $\mathcal{C}_p$, which is dense in $L^p(X,m)$ by Lemmas 3.5 and 2.2.

(ii) The hypothesis gives

$$\liminf_{r\downarrow 0} \frac{D_p(f,r)}{\Psi_{C,p}(r)} = 0\,.$$

By the second inequality in (4.6), $\overline{\mathcal{E}}_p(f) = 0$, and then, by the first, $D_p(f,r) = 0$ for every sufficiently small r. Fix such an r. For m-almost every x, $f = f(x)$ almost everywhere on $B(x,r)$; hence f is almost everywhere equal to a constant c_z on each ball $B(z, r/2)$. Two such balls which intersect have an intersection of positive measure, because m has full support, and therefore the same constant. Thus $z \mapsto c_z$ is locally constant, and constant because X is connected. This is the argument of [31, Lemma 4.7], with the two scales used separately.

Taking $S(r) = r^{\beta_P}$ in (i) and $S(r) = r^{p\alpha}$ with $p\alpha > \beta_C$ in (ii) proves (4.13), together with (4.12). □

Connectedness is essential in (ii). If X has a component of finite measure at positive distance from its complement, the indicator of this component has zero oscillation at every sufficiently small radius without being constant.

Theorem 4.12. *Assume the scale hypotheses of Theorem 4.4, together with* $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}_p(\Psi_{P,p})$ *and matching at* p. *If*

$$\Psi_{C,p}(r) \asymp r^{\beta} \qquad (r \downarrow 0)\,,$$

then

$$\alpha_p^{\mathrm{dens}} = \alpha_p = \frac{\beta}{p}\,. \tag{4.14}$$

Both thresholds are attained: $KS^{\beta/p,p}(X)$ *is dense in* $L^p(X,m)$ *and contains nonconstant functions. Moreover,*

$$KS^{\beta/p,p}(X) = \begin{cases} W^{1,p}(X,m)\,, & p > 1\,, \\ BV(X,m)\,, & p = 1\,, \end{cases} \tag{4.15}$$

and, for every $f \in L^p(X,m)$,

$$\sup_{0<r\le r_1} \frac{D_p(f,r)}{r^{\beta}} \asymp \overline{\mathcal{E}}_p(f) \asymp \liminf_{r\downarrow 0} \frac{D_p(f,r)}{r^{\beta}}\,. \tag{4.16}$$

Here $\overline{\mathcal{E}}_p = \mathcal{E}_p$ *for* $p > 1$ *and* $\overline{\mathcal{E}}_1(f) = \|Df\|(X)$.

Proof. Matching gives $\Psi_{P,p}(r) \le Cr^{\beta}$. By Proposition 4.11, the dense core $\mathcal{C}_p$ lies in $KS^{\beta/p,p}(X)$, and every $KS^{\alpha,p}(X)$ with $\alpha > \beta/p$ consists of constants. This proves (4.14) and the attainment. Theorems 4.3, 4.4 and 4.8 give (4.15) and (4.16). No admissibility assumption is needed beyond matching and the stated hypotheses. □

At $p = 1$, (4.15) identifies the space $BV(X,m)$ defined by relaxation, not the domain of the closed gradient. The latter is $W^{1,1}(X,m)$, which is a subspace of $BV(X,m)$ by Theorem 4.8.

Power volume growth. Suppose in addition that $\mathrm{V}(\Phi)$ holds and that, at small radii,

$$\Phi(r) \asymp r^{Q}\,, \qquad \Theta(r) \asymp r^{s}\,.$$

By (3.14), $\Psi_{C,p}(r) \asymp r^{p+Q-s}$. At every matching exponent p, Theorem 4.12 gives

$$\alpha_p^{\mathrm{dens}} = \alpha_p = 1 + \frac{Q-s}{p}\,. \tag{4.17}$$

By Proposition 3.6, $s \le Q$, so that $\alpha_p \ge 1$, with $\alpha_p > 1$ exactly when $s < Q$. The number s is the exponent of the gauge Θ. It need not describe the ν-measure of balls, which may be infinite.

Examples.

- For the examples of Table 1 with matching scale r^p, one obtains $\alpha_p = 1$.
- For a local tree with $\Phi(r) \asymp r^{Q}$, Proposition 15.2 gives the matching scale $r^{p-1}\Phi(r)$, and hence

  $$\alpha_p^{\mathrm{dens}} = \alpha_p = 1 + \frac{Q-1}{p}\,, \qquad 1 \le p < \infty\,.$$

  The critical spaces are $W^{1,p}$ for $p > 1$ and BV for $p = 1$, as in [22, Theorems 3.7–3.8].

- For the tree of n-dimensional cubes of Chapter 15, Theorem 15.11 gives

$$\alpha_p^{\mathrm{dens}} = \alpha_p = 1 + \frac{Q-n}{p} \qquad (1 \le p < \infty)\,,$$

with both thresholds attained, although the scales match only for $p > n$. Weak monotonicity at the critical exponent holds precisely for $p > n$, and the critical space is then $W^{1,p}$. For admissible $1 < p \le n$ the critical space is a proper subspace of $W^{1,p}$, dense in the graph norm.

4.7 Gagliardo–Nirenberg inequalities

Sobolev-type inequalities follow from two estimates on the averaging operators $M_r f(x) = \fint_{B(x,r)} f\,dm$: a bound on the error $f - M_r f$ in terms of the energy, and a bound on $M_r f$ in L^∞. The first is the Poincaré estimate of Lemma 4.1. The second follows from a uniform lower bound for the volume of balls. Combined with a truncation property of the energy, these two estimates give the Gagliardo–Nirenberg inequalities, by the method of Bakry, Coulhon, Ledoux and Saloff-Coste [16, Theorem 9.1]; its formulation for Korevaar–Schoen spaces is developed in [31, Section 4.4], and the analogous argument with the heat semigroup in [4, Section 4]. No heat kernel is needed here.

We state the inequalities for the relaxed energy. Thus the endpoint $p = 1$ concerns BV, and the conclusions do not require closability. For $p > 1$ at an admissible exponent, $\overline{\mathcal{E}}_p = \mathcal{E}_p$ by Theorem 4.4.

Theorem 4.13. *Let $1 \le p < \infty$, $\alpha, Q > 0$ and $0 < R \le \infty$. Suppose that, for every $x \in X$, $0 < r < R$ and $f \in L^p(X,m)$,*

$$c_V\, r^Q \le m(B(x,r)) < \infty\,, \qquad D_p(f,r) \le C_P\, r^{p\alpha}\, \overline{\mathcal{E}}_p(f)\,. \tag{4.18}$$

Let $a, b \in [1,\infty]$ and $0 < \theta \le 1$ satisfy

$$\frac{1}{a} = \theta\left(\frac{1}{p} - \frac{\alpha}{Q}\right) + \frac{1-\theta}{b}\,. \tag{4.19}$$

If $p\alpha = Q$ and $p > 1$, assume $a < \infty$. Then

$$\|f\|_{L^a(m)} \le C\left(\overline{\mathcal{E}}_p(f) + R^{-p\alpha}\|f\|^p_{L^p(m)}\right)^{\theta/p} \|f\|^{1-\theta}_{L^b(m)} \tag{4.20}$$

for every $f \in L^p(m) \cap L^b(m)$ of finite relaxed energy. Here $R^{-p\alpha} = 0$ when $R = \infty$. If $\theta = 1$, the L^b assumption and factor are omitted. The constant depends on $p, \alpha, Q, a, b, \theta, c_V, C_P$, and not on f or R.

Proof. We verify the hypotheses of [16, Section 2 and Theorem 9.1].

Step 1: the averaging estimates. The first assumption in (4.18) and Jensen's inequality give, for $0 < r < R$,

$$\|M_r f\|_\infty \le c_V^{-1} r^{-Q} \|f\|_1\,, \qquad \|f - M_r f\|_p \le D_p(f,r)^{1/p} \le C_P^{1/p}\, r^\alpha\, \overline{\mathcal{E}}_p(f)^{1/p}\,.$$

These are the estimates of [31, Lemmas 4.12–4.13]. The change of parameter $t = r^\alpha$ puts them in the form required by the cited theorem, with dimension parameter Q/α and upper parameter R^α.

Step 2: the truncation property. For $f \ge 0$, $\rho > 1$ and $j \in \mathbb{Z}$, set $f_j = (f - \rho^j)_+ \wedge (\rho^{j+1} - \rho^j)$. We claim that

$$\sum_{j\in\mathbb{Z}} \overline{\mathcal{E}}_p(f_j) \le \overline{\mathcal{E}}_p(f)\,, \qquad \sum_{j\in\mathbb{Z}} \|f_j\|_p^p \le \|f\|_p^p\,. \tag{4.21}$$

The second inequality follows pointwise from $\sum_j f_j = f$. For the first, let $u \in \mathcal{C}_p$ be nonnegative. By Lemma 1.5(iv), $u_j \in \mathcal{C}_p$ and $|\nabla u_j| = \mathbf{1}_{\{\rho^j < u < \rho^{j+1}\}} |\nabla u|$; since the intervals (ρ^j, ρ^{j+1}) are disjoint, $\sum_j |\nabla u_j|^p \le |\nabla u|^p$. Let now $u_k \in \mathcal{C}_p$ with $u_k \to f$ in $L^p(m)$ and $\int_{\mathcal{S}} |\nabla u_k|^p \, d\nu \to \overline{\mathcal{E}}_p(f)$; replacing u_k by $(u_k)_+$, we may assume $u_k \ge 0$. For each j, $(u_k)_j \to f_j$ in $L^p(m)$. For a finite set J of indices, the definition of the relaxed energy gives

$$\sum_{j \in J} \overline{\mathcal{E}}_p(f_j) \le \liminf_{k \to \infty} \sum_{j \in J} \int_{\mathcal{S}} |\nabla (u_k)_j|^p \, d\nu \le \overline{\mathcal{E}}_p(f) \,,$$

and monotone convergence in J proves (4.21). The same argument shows that the relaxed energy does not increase under the truncations $f \mapsto (f - s)_+ \wedge t$, and Minkowski's inequality gives the triangle inequality for $\overline{\mathcal{E}}_p^{1/p}$. The functionals $\overline{\mathcal{E}}_p^{1/p}$ and $(\overline{\mathcal{E}}_p + R^{-p\alpha} \| \cdot \|_p^p)^{1/p}$ therefore satisfy the seminorm and truncation hypotheses of [16]. This is the step which gives strong inequalities rather than weak ones, as in [31, Theorem 4.3].

Step 3: conclusion. Theorem 9.1 of [16], together with its Theorems 3.1–3.4, which distinguish the three ranges of $1/p - \alpha/Q$, now gives (4.20). Its term involving the upper parameter is $R^{-\alpha} \|f\|_p$, equivalently the term displayed in (4.20). One first applies the theorem to $(|f| - \varepsilon)_+ \wedge M$, which is bounded and vanishes outside a set of finite measure. Letting $\varepsilon \downarrow 0$ and $M \uparrow \infty$, the contraction of the energy under truncation and Fatou's lemma prove the assertion for f. □

The hypotheses. Under VD and $\mathrm{PI}_p(r^{p\alpha})$, the second estimate in (4.18) follows from Lemma 4.2(iii); in the local case, R is chosen within the range of radii of that lemma. The lower bound on the volume is an additional assumption. It follows from $\mathrm{V}(r^Q)$, but no upper bound of that order is needed. On a noncompact space, volume doubling alone does not give a positive uniform lower bound for the volume of balls of a fixed radius. Neither the cutoff condition nor matching scales are needed for (4.20).

Sobolev and Nash inequalities. If $p\alpha < Q$, taking $\theta = 1$ gives the Sobolev inequality

$$\|f\|_{L^{pQ/(Q-p\alpha)}(m)} \le C \left(\overline{\mathcal{E}}_p(f) + R^{-p\alpha} \|f\|_p^p \right)^{1/p} . \tag{4.22}$$

At $p = 1$ this applies to BV functions. For a set E of finite measure and finite perimeter it gives the isoperimetric inequality

$$m(E)^{1-\alpha/Q} \le C\big(P(E) + R^{-\alpha} m(E)\big) \,.$$

For every $p > 1$, irrespective of the sign of $Q - p\alpha$, the choice $a = p$, $b = 1$ gives (4.20) with $\theta = (p-1)Q/(p\alpha + (p-1)Q)$. In the quadratic case this is the Nash inequality

$$\|f\|_2^{2+4\alpha/Q} \le C \left(\overline{\mathcal{E}}_2(f) + R^{-2\alpha} \|f\|_2^2 \right) \|f\|_1^{4\alpha/Q} . \tag{4.23}$$

The term involving $\|f\|_p$ disappears under global hypotheses. On a compact space it is necessary, since nonzero constants have zero energy.

Critical and supercritical ranges. If $p\alpha = Q$ and $p > 1$, every finite $a > b \ge 1$ is allowed, with $\theta = 1 - b/a$. The constants need not remain bounded as $a \to \infty$, and there is no general L^∞ conclusion. The critical case $p = 1$, $\alpha = Q$ is different:

$$\|f\|_\infty \le C\big(\|Df\|(X) + R^{-\alpha} \|f\|_1\big) \,. \tag{4.24}$$

The distinction is the endpoint in [16, Theorem 3.4]. It can also be seen from the coarea formula, as follows. Let $u \in L^p(X,m)$ with $\sigma = m(\{u \neq 0\}) < \infty$. If $m(B(x,r)) \geq 2\sigma$ for every x, then at least half of each ball $B(x,r)$ lies in $\{u = 0\}$, so $D_p(u,r) \geq \frac{1}{2}\|u\|_p^p$. Choosing r with $c_V r^Q = 2\sigma$ when this radius is less than R, and otherwise using the lower-order term, gives

$$\|u\|_p \leq C\,\sigma^{\alpha/Q}\left(\overline{\mathcal{E}}_p(u) + R^{-p\alpha}\|u\|_p^p\right)^{1/p}. \tag{4.25}$$

At $p = 1$ and $\alpha = Q$, applying this to the indicator of a set E of positive finite measure yields $P(E) + R^{-\alpha}m(E) \geq C^{-1}$. For $0 < t < \|f\|_\infty$, at least one of the sets $\{f > t\}$ and $\{f < -t\}$ has positive measure. Integrating the last inequality in t and using Theorem 2.11 proves (4.24).

If $p\alpha > Q$, then for every $1 \leq b < \infty$ the inequality (4.20) holds with $a = \infty$ and

$$\theta = \frac{pQ}{pQ + b(p\alpha - Q)}.$$

Interpolation gives the finite target exponents in (4.19). These are estimates of the essential supremum; the existence of a continuous representative is a separate question.

Relation with the critical exponent. Suppose that X is connected and has at least two points, that the hypotheses of Theorem 4.4 hold, and that $\Psi_{P,p}(r) \asymp \Psi_{C,p}(r) \asymp r^{p\alpha}$ at small radii. Then $\alpha = \alpha_p$ by Theorem 4.12, and the relaxed energy in (4.20) is comparable to

$$\liminf_{r\downarrow 0} r^{-p\alpha_p} D_p(f,r).$$

For $p > 1$ it is also the energy of the closed gradient, and for $p = 1$ it is the total variation. This is the intrinsic Korevaar–Schoen form of the inequalities. The truncation property (4.21) cannot be obtained for the functional $\sup_r r^{-p\alpha} D_p(f,r)$ by interchanging a supremum over the radii and a sum over the truncations. In an abstract formulation with difference quotients, weak monotonicity supplies the passage from the supremum to the lower limit [31, Theorem 4.3]; compare the weak and strong inequalities for heat-kernel Besov spaces in [1, Theorems 6.3 and 6.9].

Under the additional power laws $\Phi(r) \asymp r^Q$ and $\Theta(r) \asymp r^s$, matching at p gives $p\,\alpha_p = p + Q - s$. The Sobolev exponent in the subcritical range is therefore $pQ/(s-p)$; the critical condition is $p = s$, and the supercritical condition is $p > s$. Thus the threshold between the subcritical and the supercritical range is the exponent s of the gauge, not the volume exponent Q. These statements require matching at the exponent under consideration. When the two scales differ, (4.20) still uses the Poincaré exponent, but that exponent is not thereby identified with α_p.

5 The rank theorem

The rank $n(x) = \dim H_x$ counts the gradients which are linearly independent at a point of the skeleton. At the exponent $p = 2$, and when the core spans the fibres, it is the pointwise index of the Dirichlet form, and its essential supremum n_* is the martingale dimension of the associated diffusion (Proposition 2.7). This chapter shows that the quantitative hypotheses of Chapter 3 restrict the rank. The main result is the following: *if the Poincaré inequality and the cutoff energy bound hold at a common scale whose order is strictly higher than r^p, with $p > 1$, then $n_* < p$.* In particular, a common quadratic scale of order higher than r^2 forces rank one.

For self-similar fractals, results of this type go back to Kusuoka [118], who proved that the martingale dimension of the Brownian motion on the Sierpiński gasket is one, and to Hino [87, Section 3], who bounded the martingale dimension by the spectral dimension using harmonic coordinates and their Euclidean image measures. Without self-similarity, Hino [88, Theorem 3.3] gave a capacity criterion for rank one, and Chen and Murugan [52, Theorem 1.7] bounded the martingale dimension by the spectral dimension under heat kernel estimates. The theorem proved here concerns a gradient structure on a metric space and an arbitrary exponent $p > 1$. It uses the bound of the gradient by the Lipschitz constant, and not only a representation of energy measures; the references above supply the methods which we adapt, under different hypotheses.

The chapter is organized as follows. Section 5.1 states the results and explains the idea of the proof. Section 5.2 contains the metric input: a Lipschitz function oscillates linearly near almost every point where its gradient does not vanish, and k independent gradients push the gradient measure forward to a measure on $\mathbb{R}^k$ which is absolutely continuous. Section 5.3 deduces elementary bounds of the rank by the volume growth and the Poincaré scale. Section 5.4 proves two auxiliary lemmas, and Section 5.5 proves the rank theorem.

5.1 Statements

All assertions concern a rank which may vary from point to point. Recall that a positive function Ψ has the lower scaling exponent a if $\Psi(R)/\Psi(r) \geq c(R/r)^a$ for $0 < r \leq R < r_0$, as in (3.1).

Theorem 5.1. *Let $1 < p < \infty$, and let $(\mathcal{S}, \nu, \nabla)$ be a gradient structure satisfying* VD, $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}_p(\Psi_{P,p})$ *and the matching condition* (3.4)*. If $\Psi_{C,p}$ has a lower scaling exponent $a > p$ in* (3.1)*, then*

$$n_* < p\,.$$

Theorem 5.1 is a special case of the following statement, in which neither the cutoff condition $\mathrm{B}(\Psi_{C,2})$ nor a volume function appears, and in which the cutoff functions need not be Lipschitz with controlled constant.

Theorem 5.2. *Let $1 < p < \infty$, and let $(\mathcal{S}, \nu, \nabla)$ be a p-admissible gradient structure with $\nu \neq 0$ whose core $\mathcal{C}_p$ spans the fibres. Suppose that m is doubling at small radii and that $\Psi : (0, r_0) \to (0, \infty)$ satisfies* (3.1) *with a lower scaling exponent $a > p$. Assume that there are $C < \infty$ and $\Lambda \geq 2$ such that, for every sufficiently small ball $B = B(x, r)$,*

$$\int_B |u - u_B|^p \, dm \leq C\,\Psi(r) \int_{\Lambda B \cap \mathcal{S}} |\nabla u|^p \, d\nu \qquad (u \in \mathcal{C}_p)\,, \tag{5.1}$$

and that B admits $\chi \in \mathcal{C}_p$ with

$$\mathbf{1}_B \leq \chi \leq \mathbf{1}_{\Lambda B}\,, \qquad \int_{\mathcal{S}} |\nabla \chi|^p \, d\nu \leq C\,\frac{m(B)}{\Psi(r)}\,. \tag{5.2}$$

Then $n_* < p$; equivalently, $n(x) < p$ for ν-almost every x. For $p = 2$, the same conclusion holds if the functions χ are only assumed to belong to $\mathcal{F} = W^{1,2}(X, m)$, with the pointwise bounds in (5.2) understood m-almost everywhere and $\nabla\chi$ replaced by $\nabla_2\chi$.

Proof of Theorem 5.1, assuming Theorem 5.2. If $\nu = 0$, the rank is zero and the conclusion holds. Otherwise Theorem 4.3 gives p-admissibility, and Proposition 1.9 gives spanning by $\mathcal{C}_p$ from $\mathrm{B}(\Psi_{C,2})$. Hypothesis $\mathrm{PI}_p(\Psi_{P,p})$ and (3.4) give (5.1) with $\Psi = \Psi_{C,p}$. The functions ψ_x^r of $\mathrm{B}(\Psi_{C,2})$ belong to $\mathcal{C}_p$, and by (3.10) they satisfy (5.2) with $\Lambda \geq 2$. Theorem 5.2 applies. □

The case of power scales. Under uniform volume growth, the theorem combines with the elementary bounds of Section 5.3 and with Proposition 3.6 into the following statement.

Corollary 5.3. *Let $1 < p < \infty$, and let $(\mathcal{S}, \nu, \nabla)$ be a gradient structure satisfying $\mathrm{V}(\Phi)$, $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}_p(\Psi_{P,p})$ and (3.4) at small radii, with $\Phi(r) = r^Q$ and $\Theta(r) = r^s$. Then*

$$n_* \leq s \leq Q\,, \qquad s < Q \implies n_* < p\,, \qquad s = Q \implies \Psi_{C,p}(r) = r^p \ \text{ and } \ \nu \ll m\,. \tag{5.3}$$

In particular, if $n_* \geq p$, then $s = Q$ and $\nu \ll m$. If $p = 2$ and $\nu \not\ll m$, then $n(x) = 1$ for ν-almost every x, and $Q > 1$.

The corollary is proved at the end of Section 5.3. It exhibits a dichotomy at matching scales. Either the scale is the Euclidean one, $\Psi_{C,p}(r) = r^p$, and the gradient measure is absolutely continuous; or the scale is of higher order, and the rank is smaller than the exponent. For the tree of cubes of Chapter 15, which has rank n and $s = n < Q$, the scales match exactly for $p > n$ (Proposition 15.10), in accordance with the corollary; this shows that the bound $n_* < p$ cannot be improved. The endpoint $p = 1$ behaves differently: local trees have rank one and the matching scale $\Psi_{P,1} = \Psi_{C,1} = \Phi$, whose lower scaling exponent can be greater than one (Proposition 15.2).

The quadratic case. For $p = 2$ and cutoff functions in the domain of the Dirichlet form, Theorem 5.2 reads as follows. In the language of Section 7.2, its hypotheses are a Poincaré inequality and a capacity upper bound at the same scale Ψ.

Theorem 5.4. *Let $(\mathcal{S}, \nu, \nabla)$ be a 2-admissible gradient structure with $\nu \neq 0$ whose core $\mathcal{C}_2$ spans the fibres, and let $(\mathcal{E}, \mathcal{F})$ be the associated Dirichlet form, with energy measures Γ. Let m be doubling at small radii and let Ψ satisfy (3.1) with a lower scaling exponent $a_2 > 2$, so that*

$$\frac{\Psi(R)}{\Psi(r)} \geq c\Big(\frac{R}{r}\Big)^{a_2}\,, \qquad 0 < r \leq R < r_0\,. \tag{5.4}$$

Assume that, for fixed $\Lambda \geq 2$ and $C < \infty$, each sufficiently small ball $B = B(x, r)$ admits $\chi \in \mathcal{F}$ with $\mathbf{1}_B \leq \chi \leq \mathbf{1}_{\Lambda B}$ m-almost everywhere, and that

$$\int_B |u - u_B|^2\, dm \leq C\,\Psi(r) \int_{\Lambda B} d\Gamma(u) \qquad (u \in \mathcal{F})\,, \tag{5.5}$$

$$\mathcal{E}(\chi, \chi) \leq C\,\frac{m(B)}{\Psi(r)}\,. \tag{5.6}$$

Then $n(x) = 1$ for ν-almost every x. Equivalently, the Dirichlet form has index one, and the associated diffusion has martingale dimension one.

Proof. By Proposition 2.7(ii), (5.5) restricted to $u \in \mathcal{C}_2$ is (5.1), and (5.6) is the energy bound in (5.2). Theorem 5.2, with $p = 2$ and cutoff functions in $\mathcal{F}$, gives $n_* < 2$. Since $1 \leq n(x) \leq n_*$ ν-almost everywhere, $n(x) = 1$ ν-almost everywhere. The last assertion is Proposition 2.7(iii). □

Idea of the proof of Theorem 5.2. In $\mathbb{R}^N$ with the Lebesgue measure, a point has zero p-capacity if and only if $N \geq p$. The proof transports this fact to the skeleton. Suppose that $n_* \geq p$, and let $N = \lceil p \rceil$. One chooses N functions of the core with independent gradients, uses them as coordinates near a suitable point x of the skeleton, and rescales at a sequence of radii $r_j \to 0$. The reference measure and the energy measure of the coordinates are pushed forward to $\mathbb{R}^N$ by the rescaled coordinates, and one passes to weak limits η and κ.

- Because the scale Ψ has order higher than r^p, the energy of the coordinates concentrates: the set where the gradients are independent occupies a small proportion, in m-measure, of each small ball. At every scale there is therefore a ball of comparable size on which the energy is negligible, and on which the coordinates are almost constant, by the Poincaré inequality. In the limit, *the image η of the reference measure has an atom.* The linear oscillation of Lipschitz functions (Lemma 5.6) shows that η is not reduced to this atom.
- The image κ of the energy measure satisfies an integration by parts inequality, inherited from the closability of the gradient in its dual form (Proposition 2.5(iv)). When $N \geq p$ this inequality implies that *every point of $\mathbb{R}^N$ has zero p-capacity with respect to κ* (Lemma 5.11).
- The Poincaré inequality passes to the pair (η, κ). Applied to functions which approximate the indicator of the atom with vanishing energy, it gives a contradiction.

Remark 5.5. (i) The proof of Theorem 5.2 uses the Poincaré inequality and the energy bound for the cutoff functions at the same scale Ψ. It uses neither heat kernel estimates nor a cutoff Sobolev inequality.

(ii) The Lipschitz bound and the weak-$*$ continuity in (G) enter through Lemma 5.6. Together with the Poincaré inequality, this lemma yields the lower bound (5.14) and the uniform Lipschitz bound (5.20) at the rescaled radii. The result therefore concerns the specified gradient structure on the metric space (X, d), rather than an arbitrary first-order calculus representing a Dirichlet form.

(iii) The restriction on the rank uses the behaviour of the scale at small radii only. No restriction on the exponents at large radii is inferred from it.

(iv) The conclusions concern gradients defined ν-almost everywhere. They do not assert that (X, d, m) is a Lipschitz differentiability space. In the setting of doubling Dirichlet spaces of [66, Theorem 4.1], a common Poincaré and capacity scale of order strictly higher than r^2 excludes this property. This distinction must be kept in mind when comparing the theorem with the identification of the analytic and martingale dimensions under Gaussian heat kernel bounds [131, Theorem 3.2].

5.2 The metric input: oscillation and Euclidean images

The bound $|\nabla u| \leq \mathrm{Lip}(u)$ has a pointwise converse at small scales: near almost every point where its gradient does not vanish, a Lipschitz function oscillates at least linearly at all small radii. This is the property of the metric which the rank theorem uses. Its proof relies on Schioppa's correspondence between derivations and Alberti representations, which we first recall.

Alberti representations. We use the formulation with curve fragments of [19, Section 2] and [144, Section 2]. Let (K, d) be a compact metric space and let μ be a finite Borel measure on K. A *fragment* in K is a bi-Lipschitz map $\gamma : T_\gamma \to K$ defined on a compact set $T_\gamma \subset \mathbb{R}$ of

positive Lebesgue measure. An *Alberti representation* of μ is a probability measure P on the set of fragments, together with Borel measures $\mu_\gamma \ll \mathcal{H}^1 \restriction_{\gamma(T_\gamma)}$, such that

$$\mu(E) = \int \mu_\gamma(E)\, dP(\gamma) \qquad \text{for every Borel set } E \subset K\,.$$

If $f : K \to \mathbb{R}$ is Lipschitz and γ is a fragment, then $f \circ \gamma$ is a Lipschitz function on T_γ, and its derivative $(f\circ\gamma)'(t)$ along T_γ exists at almost every $t \in T_\gamma$. Schioppa's theorem [144, Theorem 3.60] asserts that if D is a derivation of $\mathrm{Lip}(K)$ with respect to μ, in the sense of Weaver, and if $Df \geq \delta > 0$ μ-almost everywhere for some Lipschitz function f, then μ has an Alberti representation along whose fragments f has speed bounded from below.

Lemma 5.6 (Lower oscillation)**.** *Let $(\mathcal{S}, \nu, \nabla)$ satisfy (G), and let $f \in \mathrm{Lip}_b(X)$. For ν-almost every $y \in \{\nabla f \neq 0\}$ there are $c_y > 0$ and $\rho_y > 0$ such that*

$$\mathrm{osc}_{B(y,r)}\, f \geq c_y\, r\,, \qquad 0 < r < \rho_y\,.$$

Proof. Step 1: reduction to a compact set. Fix a Borel representative of ∇f and a Borel function $h : X \to (0,1]$ with $\int_X h\, d\nu < \infty$. The measure $h\nu$ is a finite Borel measure on a locally compact separable metric space, hence it is inner regular with respect to compact sets. Therefore each set $\{|\nabla f| \geq 1/i\}$, $i \geq 1$, is the union of countably many compact sets and of a ν-null set. It suffices to prove the following: if $c > 0$ and $K_0 \subset \{|\nabla f| \geq c\}$ is compact, then the conclusion of the lemma holds for ν-almost every $y \in K_0$.

Step 2: a derivation on K_0. Let $\mu_0 = (h\nu)\restriction_{K_0}$, a finite Borel measure on the compact metric space (K_0, d). For a Lipschitz function v on K_0, with Lipschitz constant L and $\|v\|_\infty = M$, the truncated McShane extension [125]

$$\tilde v(y) = \Big(\inf_{z\in K_0} \big(v(z) + L\, d(y,z)\big) \vee (-M) \Big) \wedge M\,, \qquad y \in X\,,$$

belongs to $\mathrm{Lip}_b(X)$, is equal to v on K_0, and satisfies $\mathrm{Lip}(\tilde v) \leq L$. Define

$$Dv = \frac{\langle \nabla f, \nabla \tilde v\rangle}{|\nabla f|^2} \qquad \text{on } K_0\,.$$

By Lemma 1.5(ii), $\nabla u = \nabla \tilde v$ ν-almost everywhere on K_0 for every $u \in \mathrm{Lip}_b(X)$ which is equal to v on K_0. Hence D is linear, $D(vw) = v\,Dw + w\,Dv$ by the Leibniz rule, $|Dv| \leq c^{-1}\,\mathrm{Lip}(v)$ by (G), and $D(f|_{K_0}) = 1$. For the weak-$*$ continuity, let $v_k \to v$ pointwise on K_0 with $\sup_k(\|v_k\|_\infty + \mathrm{Lip}(v_k)) < \infty$. Since K_0 is compact and the functions are equi-Lipschitz, the convergence is uniform. With common constants L and M in the extension formula, the extensions u_k of v_k converge uniformly on X to the extension u of v, and they are bounded in $\mathrm{Lip}_b(X)$. For $g_0 \in L^1(K_0, \mu_0)$, the field $g = g_0\, h\, \mathbf{1}_{K_0} \nabla f / |\nabla f|^2$ belongs to $L^1(\mathcal{S}, \mathsf{H}, \nu)$, and (G) gives

$$\int_{K_0} g_0\, Dv_k\, d\mu_0 = \int_{\mathcal{S}} \langle \nabla u_k, g\rangle\, d\nu \longrightarrow \int_{\mathcal{S}} \langle \nabla u, g\rangle\, d\nu = \int_{K_0} g_0\, Dv\, d\mu_0\,.$$

Thus D is a derivation of the Lipschitz algebra of K_0 with respect to μ_0.

Step 3: an Alberti representation. By Schioppa's theorem [144, Theorem 3.60], applied on K_0 to the derivation D and to the function $f|_{K_0}$, for which $D(f|_{K_0}) = 1$, there are $\delta_0 > 0$ and an Alberti representation $(P, (\mu_\gamma))$ of μ_0 such that, for P-almost every fragment γ,

$$|t - t'| \leq d(\gamma(t), \gamma(t')) \leq 2\,|t - t'| \quad (t, t' \in T_\gamma)\,, \qquad (f\circ\gamma)'(t) \geq \delta_0 \quad \text{for almost every } t \in T_\gamma\,.$$

Step 4: oscillation along a fragment. Let

$$G = \{y \in K_0 : \text{ there is } \rho > 0 \text{ such that } \operatorname{osc}_{B(y,r)} f \geq \tfrac{\delta_0}{12}\, r \text{ for all } 0 < r < \rho\}\,.$$

For fixed r the function $y \mapsto \operatorname{osc}_{B(y,r)} f$ is lower semicontinuous, and for fixed y the function $r \mapsto \operatorname{osc}_{B(y,r)} f$ is nondecreasing and left-continuous, because the balls are open. Hence the condition defining G may be restricted to rational r and ρ, and G is a Borel set. Let γ be a fragment as in Step 3, and let $t \in T_\gamma$ be a Lebesgue density point of T_γ at which $(f \circ \gamma)'(t)$ exists and is at least δ_0. For every sufficiently small $\ell > 0$, the set $T_\gamma \cap [t + \ell/2, t + \ell]$ is nonempty, and each of its points t' satisfies

$$f(\gamma(t')) - f(\gamma(t)) \geq \frac{\delta_0}{2}(t' - t) \geq \frac{\delta_0 \ell}{4}\,, \qquad d(\gamma(t'), \gamma(t)) \leq 2\ell\,.$$

Hence $\operatorname{osc}_{B(\gamma(t),3\ell)} f \geq \delta_0 \ell/4$ for all small ℓ, that is, $\gamma(t) \in G$. Almost every point of T_γ has these two properties and γ is Lipschitz, so $\mathcal{H}^1(\gamma(T_\gamma) \setminus G) = 0$ and therefore $\mu_\gamma(K_0 \setminus G) = 0$.

Step 5: conclusion. Integrating with respect to P gives $\mu_0(K_0 \setminus G) = \int \mu_\gamma(K_0 \setminus G)\, dP(\gamma) = 0$. Since $h > 0$, we get $\nu(K_0 \setminus G) = 0$, which is the statement of Step 1 with $c_y = \delta_0/12$. □

The proof shows that the sets $\{y : \operatorname{osc}_{B(y,r)} f \geq r/i \text{ for all } 0 < r < 1/i\}$, $i \geq 1$, are Borel and cover $\{\nabla f \neq 0\}$ up to a ν-null set. Hence, for every finite measure $\mu \ll \nu$ with $\mu(\{\nabla f \neq 0\}) > 0$, inner regularity gives a compact set K with $\mu(K) > 0$ and constants $c_0, \rho_0 > 0$ such that

$$\operatorname{osc}_{B(y,r)} f \geq c_0\, r \qquad (y \in K,\ 0 < r < \rho_0)\,. \tag{5.7}$$

A single nonzero gradient thus detects a linear oscillation. Independent gradients give more: the map formed by the corresponding functions pushes the gradient measure forward to an absolutely continuous measure. This follows from Schioppa's theorem and from the theorem of De Philippis, Marchese and Rindler [60], according to which the image under a Lipschitz map $F : X \to \mathbb{R}^k$ of a measure with k F-independent Alberti representations is absolutely continuous with respect to the Lebesgue measure.

Lemma 5.7 (Euclidean images)**.** *Let $(\mathcal{S}, \nu, \nabla)$ satisfy (G). Let $F = (f_1, \dots, f_k)$ have components in $\operatorname{Lip}_b(X)$, let μ be a finite Borel measure absolutely continuous with respect to ν, and let $K \subset X$ be compact. If the Gram matrix $(\langle \nabla f_i, \nabla f_j \rangle)_{i,j \leq k}$ is uniformly positive definite on K, then*

$$F_\#(\mu \restriction_K) \ll \mathcal{L}^k\,.$$

In particular, $\mu(K) > 0$ implies $\mathcal{H}^k(K) > 0$.

Proof. The inverse of the Gram matrix defines bounded measurable sections $V_1, \dots, V_k$ of H over K, which are linear combinations of $\nabla f_1, \dots, \nabla f_k$ and satisfy $\langle V_i, \nabla f_j \rangle = \delta_{ij}$. For a Lipschitz function v on K, let $\tilde{v}$ be a bounded Lipschitz extension of v to X, and set $D_i v = \langle V_i, \nabla \tilde{v} \rangle$ on K. By locality this does not depend on the extension. As in Step 2 of the proof of Lemma 5.6, D_i is a derivation of the Lipschitz algebra of K with respect to $\mu \restriction_K$: its weak-$*$ continuity follows from (G), applied to extensions with common Lipschitz and supremum bounds and to the integrable field $g_0\, V_i\, (d\mu/d\nu)\, \mathbf{1}_K$ for $g_0 \in L^1(\mu \restriction_K)$. Moreover $D_i(f_j|_K) = \delta_{ij}$. Schioppa's theorem [144, Theorem 3.60], applied to small cones about the coordinate directions of $\mathbb{R}^k$, gives k F-independent Alberti representations of $\mu \restriction_K$. By the theorem of De Philippis, Marchese and Rindler [60, Lemmas 3.2–3.3], in the form [20, Theorem 2.16], $F_\#(\mu \restriction_K) \ll \mathcal{L}^k$. Finally, if $\mathcal{H}^k(K) = 0$, then $\mathcal{L}^k(F(K)) = 0$ because F is Lipschitz, which is incompatible with $F_\#(\mu \restriction_K)(F(K)) = \mu(K) > 0$. □

For strongly local regular Dirichlet forms, the corresponding absolute continuity on the set where the energy matrix is nondegenerate is the energy image density theorem of Eriksson-Bique and Murugan [65, Theorem 2.65]. Lemma 5.7 uses the axioms of a derivation on the metric space and does not assume closability.

5.3 The rank, the volume and the Poincaré scale

Lemma 5.7 bounds the rank by the Hausdorff dimension of the sets which carry the gradient measure. A lower bound for the volume of balls and a Poincaré inequality control this dimension. The comparison of independent directions with the Hausdorff measure follows [20, Theorem 2.17 and Corollary 2.20].

Proposition 5.8. *Let $(\mathcal{S}, \nu, \nabla)$ be a gradient structure with $\nu \neq 0$. Suppose that $\mathcal{C}_p$ spans the fibres for some $1 \le p < \infty$ and that, for some $Q > 0$, $a > 0$, $c, C > 0$ and $\Lambda \ge 1$, all sufficiently small balls satisfy*

$$m(B(x,r)) \ge c\, r^Q\,, \qquad \int_{B(x,r)} |u - u_{B(x,r)}|^p \, dm \le C\, r^a \int_{B(x,\Lambda r)\cap \mathcal{S}} |\nabla u|^p \, d\nu \quad (u \in \mathcal{C}_p)\,.$$

Then

$$n_* \le Q + p - a\,. \tag{5.8}$$

In particular, under $\mathrm{V}(\Phi)$, $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}_p(\Psi_{P,p})$ *and matching power scales $\Phi(r) = r^Q$, $\Theta(r) = r^s$, one has $n_* \le s$.*

Proof. Put $N = n_*$. Since $\nu \neq 0$ and $n(x)$ takes positive integer values, $\nu(\{n = N\}) > 0$.

Step 1: independent gradients. By Lemma 1.7, applied to a spanning sequence in $\mathcal{C}_p$, there are $f_1, \dots, f_N \in \mathcal{C}_p$ whose gradients are linearly independent on a set of positive ν-measure. Put $F = (f_1, \dots, f_N)$ and

$$w = \sum_{i=1}^{N} |\nabla f_i|^2\,, \qquad \mu = w^{p/2}\,\nu\,.$$

The measure μ is finite because $\nabla f_i \in L^p(\nu)$. We fix Borel representatives of the gradients. On the set of independence $\nabla f_1 \neq 0$. By (5.7) and the inner regularity of μ, there are a compact set K with $\mu(K) > 0$ and constants $c_0, r_K > 0$ such that the Gram matrix $(\langle \nabla f_i, \nabla f_j \rangle)_{i,j}$ is uniformly positive definite on K and

$$\operatorname{osc}_{B(y,r)} f_1 \ge c_0\, r \qquad (y \in K,\ 0 < r < r_K)\,.$$

Step 2: a lower bound for μ. Let $y \in K$ and let r be small. Choose two points of $B(y,r)$ at which the values of f_1 differ by at least $c_0 r/2$, and let $\varepsilon = \min(1, c_0/(8\,\mathrm{Lip}(f_1)))$. The balls of radius εr about these two points are contained in $B(y, 2r)$, and the values of f_1 on these two balls are separated by at least $c_0 r/4$. Hence, for every constant b, $|f_1 - b| \ge c_0 r/8$ on at least one of the two balls. The lower bound for the volume and the Poincaré inequality on $B(y,2r)$, applied to $f_1 \in \mathcal{C}_p$ with $|\nabla f_1|^p \le w^{p/2}$, give

$$c\, r^{Q+p} \le \int_{B(y,2r)} |f_1 - (f_1)_{B(y,2r)}|^p \, dm \le C\, r^a\, \mu(B(y, 2\Lambda r))\,.$$

After changing the radius by a fixed factor, $\mu(B(y,r)) \ge c\, r^{Q+p-a}$ for all $y \in K$ and all small r.

Step 3: the Hausdorff measure of K. Let $\{y_i\}$ be a maximal r-separated subset of K. The balls $B(y_i, r/2)$ are pairwise disjoint, so by Step 2 the family has at most $C\mu(X)\, r^{-(Q+p-a)}$ members, and the balls $B(y_i, r)$ cover K. Hence, if $N > Q + p - a$, then $\mathcal{H}^N(K) \le \liminf_{r\to 0} C\, r^{N-(Q+p-a)} = 0$.

Step 4: conclusion. By Lemma 5.7, applied to F, μ and K, one has $\mathcal{H}^N(K) > 0$. By Step 3 this excludes $N > Q + p - a$, which proves (5.8). Under $\mathrm{B}(\Psi_{C,2})$, Proposition 1.9 gives spanning by $\mathcal{C}_p$. At matching power scales, $\mathrm{PI}_p(\Psi_{P,p})$ holds with the scale $\Psi_{C,p}(r) = r^{Q+p-s}$, so $a = Q + p - s$ and $n_* \le s$. □

Remark 5.9 (The rank and the volume). The lower bound $m(B(x,r)) \ge c\, r^Q$ alone gives $n_* \le Q$. Indeed, if $N = n_* > 0$, choose N functions of $\mathrm{Lip}_c(X)$ with independent gradients on a set of positive ν-measure. With a finite measure equivalent to ν, extract a compact set K of positive measure on which their Gram matrix is uniformly positive definite. Lemma 5.7 gives $\mathcal{H}^N(K) > 0$. The disjoint balls of radius $r/2$ centred at a maximal r-separated subset of K lie in a neighbourhood of K of finite m-measure, and each has measure at least $c\, r^Q$, so there are at most $C\, r^{-Q}$ centres. The balls of radius r cover K, which gives $\mathcal{H}^Q(K) < \infty$ and $N \le Q$.

For the related bound of the martingale dimension by the Hausdorff dimension under Gaussian heat kernel estimates, see [131, Theorem 3.2 and Corollary 3.3]. The rank can be strictly smaller than the volume exponent even when $m = \nu$: Theorem 12.1 realizes $n_* = n < Q$ for every integer $n \ge 1$ and every real $Q > n$, and Kleiner and Schioppa construct spaces with a Poincaré inequality, of arbitrary topological dimension, whose measurable tangent bundle is one-dimensional [111, Theorem 1.1].

Proof of Corollary 5.3. Hypothesis $\mathrm{V}(\Phi)$ excludes isolated points, since $\Phi(r) \to 0$ as $r \downarrow 0$. We first show that $\nu \ne 0$. If $\nu = 0$, the Poincaré inequality and the full support of m make every function of $\mathcal{C}_p = \mathrm{Lip}_c(X)$ constant on each sufficiently small ball. Applied to $u = (1 - d(x,\cdot)/\varepsilon)_+$, for small ε, this shows that every point x is isolated, a contradiction.

Proposition 5.8 gives $n_* \le s$. Proposition 3.6 gives $s \le Q$, and $\nu \ll m$ when $s = Q$; in that case $\Psi_{C,p}(r) = r^{p+Q-s} = r^p$. For $s < Q$, the scale $\Psi_{C,p}(r) = r^{Q+p-s}$ has a lower scaling exponent greater than p, so Theorem 5.1 gives $n_* < p$. If $p = 2$ and $\nu \not\ll m$, the case $s = Q$ is excluded, and $n_* < 2$ gives $n(x) = 1$ almost everywhere; then $1 = n_* \le s < Q$. □

5.4 Two auxiliary lemmas

The proof of Theorem 5.2 uses two further lemmas, which are independent of gradient structures. The first differentiates a finite measure, which need not be doubling, along suitable radii. The second shows that a point of $\mathbb{R}^N$ has zero p-capacity with respect to a measure which satisfies an integration by parts inequality, when $N \ge p$.

Differentiation along doubling radii. The energy measure of the coordinates is a finite measure λ which is not doubling in general, so the Lebesgue differentiation theorem is not available for it. Following Hytönen [91, Lemma 3.3 and Corollary 3.6], one differentiates along the radii at which λ is doubling at the given point; almost every point has arbitrarily small radii of this kind, because the ambient space is doubling. The same device is used to choose scales for energy measures in [52, Lemma 2.2]. In the following lemma the reference measure m supplies the doubling property of the space, and the measure which is differentiated is λ.

Lemma 5.10. *Assume that m is doubling at small radii, let λ be a finite Borel measure on X, and let $R \ge 1$. There is a constant C_R, depending only on R and on the doubling constant of m, with the following properties.*

(i) *For λ-almost every $x \in X$ there is a sequence $r_j \downarrow 0$ such that*

$$0 < \lambda(B(x, 5Rr_j)) \le C_R\, \lambda(B(x, r_j)) \qquad (j \ge 1)\,. \tag{5.9}$$

(ii) *Let $g \in L^1(X, \lambda)$. For λ-almost every $x \in X$, every sequence $r_j \downarrow 0$ satisfying (5.9) satisfies*

$$\lim_{j\to\infty} \frac{1}{\lambda(B(x, r_j))} \int_{B(x,Rr_j)} |g - g(x)|\, d\lambda = 0\,.$$

Proof. By doubling there are $D \ge 1$ and $r_0 > 0$ such that $m(B(y, 5Rr)) \le D\, m(B(y, r))$ for $y \in X$ and $0 < 5Rr < r_0$. Let $C_R = 2D$. We use the covering theorem [83, Theorem 1.2]: every family of balls of bounded radii contains a subfamily of pairwise disjoint balls $B(x_i, r_i)$ such that the balls $B(x_i, 5r_i)$ cover the union of the family. We write λ^* for the outer measure defined by λ. Since X is separable, λ-almost every point belongs to the support of λ.

(i) *Step 1: a point without doubling radii has zero density.* Let x be a point of the support of λ for which no sequence as in (i) exists. Then there is $\rho \in (0, r_0)$ such that $\lambda(B(x, 5Rr)) > C_R\, \lambda(B(x, r))$ for all $0 < r < \rho$. Let $\rho_i = (5R)^{-i}\rho$. By induction, $\lambda(B(x, \rho_i)) \le C_R^{-i}\lambda(X)$ and $m(B(x, \rho_i)) \ge D^{-i} m(B(x, \rho))$, and for $\rho_{i+1} \le r \le \rho_i$,

$$\frac{\lambda(B(x, r))}{m(B(x, r))} \le \frac{\lambda(B(x, \rho_i))}{m(B(x, \rho_{i+1}))} \le \frac{D\, \lambda(X)}{m(B(x, \rho))} \Big(\frac{D}{C_R}\Big)^i = \frac{D\, \lambda(X)}{m(B(x, \rho))}\, 2^{-i}\,.$$

Hence x belongs to the set Z of the points at which $\lambda(B(x, r))/m(B(x, r)) \to 0$ as $r \to 0$.

Step 2: $\lambda^*(Z) = 0$. Fix $x_0 \in X$, $0 < \rho_0 < r_0/10$ and $\varepsilon > 0$. For $x \in Z \cap B(x_0, \rho_0)$ choose $r_x \in (0, \rho_0)$ with $\lambda(B(x, 5r_x)) \le \varepsilon\, m(B(x, r_x))$; this is possible because $m(B(x, 5r)) \le D\, m(B(x, r))$. The covering theorem gives pairwise disjoint balls $B(x_i, r_i)$, $r_i = r_{x_i}$, such that the balls $B(x_i, 5r_i)$ cover $Z \cap B(x_0, \rho_0)$. These balls are contained in $B(x_0, 2\rho_0)$ and have positive m-measure, so the family is countable, and

$$\lambda^*(Z \cap B(x_0, \rho_0)) \le \sum_i \lambda(B(x_i, 5r_i)) \le \varepsilon \sum_i m(B(x_i, r_i)) \le \varepsilon\, m(B(x_0, 2\rho_0))\,.$$

Letting $\varepsilon \to 0$ and covering X by countably many balls $B(x_0, \rho_0)$ gives $\lambda^*(Z) = 0$.

(ii) *Step 3: a maximal inequality.* We use the maximal operator with enlarged denominator of [91, Proposition 3.5]. For $g \in L^1(X, \lambda)$ and x in the support of λ, let

$$\mathcal{M}g(x) = \sup_{0<r<1} \frac{1}{\lambda(B(x, 5r))} \int_{B(x,r)} |g|\, d\lambda\,.$$

Let $t > 0$. Every point x with $\mathcal{M}g(x) > t$ is the centre of a ball $B(x, r_x)$, $r_x < 1$, with $\int_{B(x,r_x)} |g|\, d\lambda > t\, \lambda(B(x, 5r_x))$. The covering theorem gives pairwise disjoint balls $B(x_i, r_i)$ of this type such that the balls $B(x_i, 5r_i)$ cover $\{\mathcal{M}g > t\}$; the family is countable because $\int_{B(x_i,r_i)} |g|\, d\lambda > 0$. Hence

$$\lambda^*(\{\mathcal{M}g > t\}) \le \sum_i \lambda(B(x_i, 5r_i)) \le \frac{1}{t} \sum_i \int_{B(x_i,r_i)} |g|\, d\lambda \le \frac{1}{t}\, \|g\|_{L^1(\lambda)}\,. \tag{5.10}$$

Step 4: differentiation. Let x be a point of the support of λ, let (r_j) satisfy (5.9), and let $u \in C_c(X)$. For $Rr_j < 1$, the inequality $\lambda(B(x, r_j)) \ge C_R^{-1}\lambda(B(x, 5Rr_j))$ gives

$$\begin{aligned}\frac{1}{\lambda(B(x, r_j))} \int_{B(x,Rr_j)} |g - g(x)|\, d\lambda \le{}& \frac{C_R}{\lambda(B(x, 5Rr_j))} \int_{B(x,Rr_j)} |g - u|\, d\lambda \\ &+ C_R\Big(\sup_{B(x,Rr_j)} |u - u(x)| + |u(x) - g(x)|\Big)\,,\end{aligned}$$

so that the upper limit of the left-hand side is at most $C_R(\mathcal{M}(g-u)(x)+|g-u|(x))$. By (5.10) and Chebyshev's inequality, for every $t>0$ the set of the points x at which this upper limit exceeds $2C_Rt$, for some sequence satisfying (5.9), has outer measure at most $2t^{-1}\|g-u\|_{L^1(\lambda)}$. Since λ is a finite Borel measure on a locally compact separable metric space, $C_c(X)$ is dense in $L^1(X,\lambda)$. Hence this set is λ-null for every t, which proves (ii). □

Points of zero capacity. In the Euclidean space $\mathbb{R}^N$ with the Lebesgue measure, a point has zero p-capacity if and only if $N \geq p$. The next lemma extends the sufficiency to any finite measure κ which admits an integration by parts with an error bounded in $L^{p'}(\kappa)$; the Lebesgue measure on a ball is an example. The proof first derives, from the integration by parts, a Hardy inequality for κ, by testing with the radial vector fields of the classical Hardy inequalities [124]; the capacity estimate then follows with the usual cutoffs, which are logarithmic in the critical case $N=p$. The use of logarithmic cutoffs to exclude higher rank is present in [87, Section 3] and [88, proof of Theorem 3.3].

Lemma 5.11. *Let $1<p<\infty$, let N be an integer with $N\geq p$, and let κ be a finite Borel measure on $\mathbb{R}^N$. Suppose that, for some $C_0<\infty$,*

$$\Big|\int_{\mathbb{R}^N} \operatorname{div}\varphi\, d\kappa\Big| \leq C_0\Big(\int_{\mathbb{R}^N} |\varphi|^{p'}\, d\kappa\Big)^{1/p'} \qquad (\varphi\in C_c^\infty(\mathbb{R}^N,\mathbb{R}^N))\,. \tag{5.11}$$

Then, for every $z\in\mathbb{R}^N$, there are $\phi_k\in C_c^\infty(\mathbb{R}^N)$ such that

$$0\leq\phi_k\leq 1\,, \qquad \phi_k\longrightarrow \mathbf{1}_{\{z\}} \text{ pointwise}\,, \qquad \int_{\mathbb{R}^N}|\nabla\phi_k|^p\, d\kappa\longrightarrow 0\,. \tag{5.12}$$

Proof. By translation we may take $z=0$. We write $\ell(w)=\log(1/|w|)$ for $0<|w|<1$.

Step 1: a radial field. If $N>p$, set

$$V(w)=\frac{w}{|w|^p}\,, \qquad \text{so that} \qquad \operatorname{div}V(w)=\frac{N-p}{|w|^p}\,, \qquad |V(w)|^{p'}=\frac{1}{|w|^p}\,.$$

If $N=p$, set instead

$$V(w)=\frac{w}{|w|^p\,\ell(w)^{p-1}}\,, \qquad \text{so that} \qquad \operatorname{div}V(w)=\frac{p-1}{|w|^p\,\ell(w)^p}\,, \qquad |V(w)|^{p'}=\frac{1}{|w|^p\,\ell(w)^p}\,.$$

In both cases $\operatorname{div}V = c\,|V|^{p'}$ on $\{0<|w|<1\}$, where $c=N-p>0$ in the first case and $c=p-1>0$ in the second, and $\langle\nabla\ell, V\rangle<0$.

Step 2: a Hardy inequality for κ. Choose $\chi\in C_c^\infty(B(0,3/4))$ with $0\leq\chi\leq 1$ and $\chi=1$ on $B(0,1/2)$, and a nonincreasing function $\theta\in C^\infty(\mathbb{R})$ with $0\leq\theta\leq 1$, $\theta=1$ on $(-\infty,1]$ and $\theta=0$ on $[2,\infty)$. For $\varepsilon>0$, the field

$$\varphi_\varepsilon(w)=\chi(w)\,\theta(\varepsilon\ell(w))\,V(w)$$

vanishes for $|w|\leq e^{-2/\varepsilon}$, and it extends by zero to a smooth compactly supported field on $\mathbb{R}^N$. Its divergence is

$$\operatorname{div}\varphi_\varepsilon=\chi\,\theta(\varepsilon\ell)\,\operatorname{div}V+\theta(\varepsilon\ell)\,\langle\nabla\chi,V\rangle+\varepsilon\,\chi\,\theta'(\varepsilon\ell)\,\langle\nabla\ell,V\rangle\,.$$

The last term is nonnegative, and the second is bounded by a constant C independent of ε, because $\nabla\chi$ is supported in $\{1/2\leq|w|\leq 3/4\}$. With

$$I_\varepsilon=\int\chi\,\theta(\varepsilon\ell)\,|V|^{p'}\,d\kappa<\infty\,,$$

the hypothesis (5.11) and the inequality $|\varphi_\varepsilon|^{p'} \le \chi\,\theta(\varepsilon\ell)\,|V|^{p'}$ give

$$c\,I_\varepsilon - C\,\kappa(\mathbb{R}^N) \le \int \operatorname{div}\varphi_\varepsilon\,d\kappa \le C_0\,I_\varepsilon^{1/p'}\,.$$

By Young's inequality, I_ε is bounded independently of ε. As $\varepsilon \downarrow 0$, $\theta(\varepsilon\ell)$ increases to 1 on $\{0<|w|<1\}$, and the monotone convergence theorem gives

$$\int_{\{0<|w|<1/2\}} \frac{d\kappa(w)}{|w|^p} < \infty \quad (N>p)\,, \qquad \int_{\{0<|w|<1/2\}} \frac{d\kappa(w)}{|w|^p \log^p(1/|w|)} < \infty \quad (N=p)\,. \tag{5.13}$$

Step 3: the case $N > p$. Let ϕ_k be a smooth radial function which is equal to one on $B(0,2^{-k-1})$, vanishes outside $B(0,2^{-k})$, and satisfies $0 \le \phi_k \le 1$ and $|\nabla\phi_k| \le C2^k$. Then $\phi_k \to \mathbf{1}_{\{0\}}$ pointwise, and by (5.13)

$$\int |\nabla\phi_k|^p\,d\kappa \le C \int_{\{2^{-k-1}\le|w|\le 2^{-k}\}} \frac{d\kappa(w)}{|w|^p} \longrightarrow 0\,.$$

Step 4: the case $N = p$. Choose a nondecreasing function $\tau \in C^\infty(\mathbb{R})$ with $0 \le \tau \le 1$, $\tau = 0$ on $(-\infty,1]$ and $\tau = 1$ on $[2,\infty)$, and put $\phi_k(w) = \tau(\ell(w)/k)$ for $0 < |w| < 1$, $\phi_k(0) = 1$, and $\phi_k = 0$ outside the unit ball. Then $\phi_k \in C_c^\infty(\mathbb{R}^N)$, $\phi_k \to \mathbf{1}_{\{0\}}$ pointwise, and $\nabla\phi_k$ is supported in $\{e^{-2k} \le |w| \le e^{-k}\}$, where $k \le \ell \le 2k$ and

$$|\nabla\phi_k(w)|^p \le \frac{C}{k^p\,|w|^p} \le \frac{C}{|w|^p\,\ell(w)^p}\,.$$

By the second part of (5.13) and dominated convergence, $\int |\nabla\phi_k|^p\,d\kappa$ tends to zero. □

5.5 Proof of the rank theorem

We prove Theorem 5.2 by contradiction. Throughout this section the hypotheses of the theorem are in force, and we suppose that $n_* \ge p$. Let $N = \lceil p \rceil$, so that $N \ge p$ and $\nu(\{n \ge N\}) > 0$.

The argument adapts the approach by Euclidean images of [87, Section 3] and [88, Proposition 4.3]; compare also [52, Proposition 2.6 and Lemmas 3.4–3.5]. Two points differ from these references. In place of harmonic coordinates, whose energy flux is divergence-free, we use arbitrary functions of the core and approximate their flux by fields which have a divergence; such fields are dense because the gradient is closable. Second, the atom which produces the final contradiction comes from the bound of the gradient by the Lipschitz constant, through Lemma 5.6.

Plan. The proof has eight steps.

- Steps 1 to 3 choose N coordinate functions, a base point x and radii $r_j \to 0$, and normalize the coordinates, the reference measure and the energy measure at the scale r_j.
- Steps 4 and 5 use the dual form of closability to show that the normalized coordinates have oscillation of order one on the ball $B(x, r_j)$.
- Step 6, the only one which uses $a > p$, produces a ball of radius comparable to r_j on which the normalized coordinates become constant.
- Steps 7 and 8 pass to the image measures on $\mathbb{R}^N$ and conclude.

Constants denoted by C and c do not depend on the index j, and the estimates hold for j large enough.

Proof of Theorem 5.2. Step 1: coordinates, and a lower bound for their energy measure. By Lemma 1.7, applied to a spanning sequence in $\mathcal{C}_p$, there are $f_1, \dots, f_N \in \mathcal{C}_p$ whose gradients are linearly independent on a Borel subset of $\{n \geq N\}$ of positive ν-measure. Let

$$F = (f_1, \dots, f_N)\,, \qquad w = \sum_{i=1}^{N} |\nabla f_i|^2\,, \qquad \mu = w^{p/2}\,\nu\,, \qquad M_{il} = \frac{\langle \nabla f_i, \nabla f_l \rangle}{w}\,\mathbf{1}_{\{w>0\}}\,.$$

The measure μ is finite because $\nabla f_i \in L^p(\nu)$; it plays the role of the p-energy measure of the map F. The matrix M is symmetric and nonnegative, with trace one on $\{w > 0\}$, so $|M| \leq 1$, and it is positive definite on the set where the gradients are independent. As in Step 1 of the proof of Proposition 5.8, (5.7) and inner regularity give a compact subset K of this set, with $\mu(K) > 0$, and constants $c_0, \rho_0 > 0$ such that

$$\operatorname{osc}_{B(y,r)} f_1 \geq c_0\, r \qquad (y \in K,\ 0 < r < \rho_0)\,.$$

We claim that, for every $y \in K$, every small r and every $b \in \mathbb{R}$,

$$\int_{B(y,2r)} |f_1 - b|^p\, dm \geq c\, r^p\, m(B(y,r))\,.$$

Indeed, choose two points of $B(y,r)$ at which the values of f_1 differ by at least $c_0 r/2$, and let $\varepsilon = \min(1, c_0/(8\operatorname{Lip}(f_1)))$. On the balls of radius εr about these points, which are contained in $B(y,2r)$, the values of f_1 remain separated by at least $c_0 r/4$. Hence $|f_1 - b| \geq c_0 r/8$ on one of these balls, and by doubling each of them has measure at least $c\, m(B(y,r))$. We apply (5.1) to f_1 on $B(y,2r)$ and use $|\nabla f_1|^p \leq w^{p/2}$. After changing the radius by a fixed factor, and using doubling and the comparability of Ψ at comparable radii, this yields $\rho_1 > 0$ such that

$$\mu(B(y,r)) \geq c\, \frac{r^p\, m(B(y,r))}{\Psi(r)} \qquad (y \in K,\ 0 < r < \rho_1)\,. \tag{5.14}$$

Since $\Psi(r)/r^p \to 0$, this says that the energy measure concentrates near K: the μ-measure of a small ball centred on K is much larger than its m-measure.

Step 2: approximation by fields with a divergence, and choice of the base point. For $1 \leq i \leq N$, define

$$T_i = w^{(p-2)/2}\, \nabla f_i \quad \text{on } \{w > 0\}\,, \qquad T_i = 0 \quad \text{on } \{w = 0\}\,.$$

These fields are the components of the flux of the p-energy of F: they satisfy $\langle T_i, \nabla f_l \rangle = M_{il}\, w^{p/2}$. Since $|T_i|^{p'} \leq w^{(p-1)p'/2} = w^{p/2}$, they belong to $L^{p'}(\mathcal{S}, \mathsf{H}, \nu)$. The structure is p-admissible, so by Proposition 2.5(iv), for all i and $k \geq 1$ there is $V_{ik} \in \mathcal{D}_{p'}$, with $h_{ik} = \operatorname{div} V_{ik} \in L^{p'}(X, m)$, such that

$$\int_{\mathcal{S}} |T_i - V_{ik}|^{p'}\, d\nu \leq 4^{-k}\,.$$

Choose $a_{ik} > 0$ such that $\sum_{i,k} a_{ik} \|h_{ik}\|^{p'}_{L^{p'}(m)} < \infty$, and put

$$\lambda = \mu + \sum_{i=1}^{N} \sum_{k \geq 1} 2^k\, |T_i - V_{ik}|^{p'}\, \nu + \sum_{i=1}^{N} \sum_{k \geq 1} a_{ik}\, |h_{ik}|^{p'}\, m\,.$$

This is a finite measure, and for every Borel set E,

$$\int_E |T_i - V_{ik}|^{p'}\, d\nu \leq 2^{-k} \lambda(E)\,, \qquad \int_E |h_{ik}|^{p'}\, dm \leq a_{ik}^{-1} \lambda(E)\,. \tag{5.15}$$

Thus the single measure λ controls μ and the quality of all the approximations on every set.

Let $R = 3\Lambda^3$ and $\varrho = d\mu/d\lambda \in [0,1]$. We apply Lemma 5.10 to λ and to the bounded functions ϱ, $\varrho\,\mathbf{1}_K$ and $\varrho\,M_{il}$. Since $\varrho > 0$ μ-almost everywhere and $\mu(K) > 0$, we may choose $x \in K$, with $\varrho(x) > 0$, and radii $r_j \downarrow 0$ satisfying (5.9), for which the conclusion of Lemma 5.10(ii) holds for these finitely many functions. We write

$$B_j(b) = B(x, b\,r_j)\,, \qquad B_j = B_j(1)\,.$$

The differentiation gives

$$\begin{aligned} \mu(B_j) &= \varrho(x)\,\lambda(B_j) + o(\lambda(B_j))\,, \\ \mu(B_j(R)\setminus K) &= \int_{B_j(R)} (\varrho - \varrho\,\mathbf{1}_K)\,d\lambda = o(\lambda(B_j))\,, \\ \int_{B_j(R)} |M - M(x)|\,d\mu &\le \int_{B_j(R)} \big(|\varrho M - \varrho(x)M(x)| + |\varrho - \varrho(x)|\big)\,d\lambda = o(\lambda(B_j))\,. \end{aligned}$$

Together with $\lambda(B_j(R)) \le C_R\lambda(B_j)$, this implies

$$\lambda(B_j(R)) \le C\,\mu(B_j)\,, \qquad \mu(B_j(R)\setminus K) + \int_{B_j(R)} |M - M(x)|\,d\mu = o(\mu(B_j))\,. \tag{5.16}$$

In words: at the scales r_j, the measure μ is doubling at x, it is almost entirely carried by K, and the matrix M is almost constant. In particular, the two bounds in (5.15) hold on $E = B_j(R)$ with $\lambda(E)$ replaced by $C\mu(B_j)$.

Step 3: normalization. We normalize the Gram matrix as in the quadratic construction of [52, Proposition 2.6]. Let A be the symmetric positive definite matrix with $A\,M(x)\,A = I$, and set

$$\vartheta_j = \Big(\frac{\Psi(r_j)\,\mu(B_j)}{m(B_j)}\Big)^{1/p}, \qquad H_j = \frac{A\,(F - F(x))}{\vartheta_j}\,, \qquad m_j = \frac{m}{m(B_j)}\,, \qquad \mu_j = \frac{\mu}{\mu(B_j)}\,.$$

For $v \in \mathcal{C}_p + \mathbb{R}$ define

$$\mathcal{E}_j(v) = \frac{\Psi(r_j)}{m(B_j)} \int_{\mathcal{S}} |\nabla v|^p\,d\nu\,, \qquad T_{j,i} = \vartheta_j^{1-p} \sum_{l=1}^{N} A_{il}\,T_l\,.$$

The map $H_j : X \to \mathbb{R}^N$ is the rescaled coordinate map, m_j and μ_j are the measures normalized to have mass one on B_j, $\mathcal{E}_j$ is the energy in the units in which a cutoff of B_j has energy of order one, and $T_{j,i}$ is the rescaled flux. The number ϑ_j is chosen so that the rescaled energy measure of H_j is μ_j. Indeed, $\nabla H_j^l = \vartheta_j^{-1}\sum_k A_{lk}\nabla f_k$, and the definitions of T_l and M give

$$\frac{\Psi(r_j)}{m(B_j)}\,\langle T_{j,i}, \nabla H_j^l\rangle\,\nu = \frac{\Psi(r_j)}{m(B_j)\,\vartheta_j^p}\,(AMA)_{il}\,w^{p/2}\,\nu = (AMA)_{il}\,\mu_j\,. \tag{5.17}$$

In the same way, using $|T_l|^{p'} \le w^{p/2}$ and $(p-1)p' = p$,

$$\frac{\Psi(r_j)}{m(B_j)}\,|T_{j,i}|^{p'}\,\nu \le C\,\mu_j\,, \qquad \frac{\Psi(r_j)}{m(B_j)}\,|\nabla H_j^i|^p\,\nu \le C\,\mu_j\,. \tag{5.18}$$

We record five further properties of the normalized objects.

- *Mass and Gram matrix.* By (5.16),

$$\mu_j(B_j(R)) \le C\,, \qquad \int_{B_j(R)} |AMA - I|\, d\mu_j \longrightarrow 0\,. \tag{5.19}$$

- *Lipschitz bound.* By (5.14), applied at $y = x \in K$, $\vartheta_j \ge c\, r_j$. Since F is Lipschitz,

$$H_j(x) = 0\,, \qquad |H_j(y) - H_j(y')| \le C\,\frac{d(y,y')}{r_j}\,, \qquad \sup_{B_j(b)} |H_j| \le C\, b\,. \tag{5.20}$$

- *Volume.* Doubling gives, for every fixed $b \in (0, R]$,

$$0 < c(b) \le m_j(B(y, b\, r_j)) \le C(b) < \infty \qquad (y \in B_j(R))\,. \tag{5.21}$$

- *Cutoffs.* Choose cutoff functions $\chi_j, \zeta_j \in \mathcal{C}_p$ as in (5.2) for the balls B_j and $B_j(2\Lambda^2)$. The comparability of Ψ at comparable radii and doubling give

$$\mathbf{1}_{B_j} \le \chi_j \le \mathbf{1}_{B_j(\Lambda)}\,, \qquad \mathbf{1}_{B_j(2\Lambda^2)} \le \zeta_j \le \mathbf{1}_{B_j(2\Lambda^3)}\,, \qquad \mathcal{E}_j(\chi_j) + \mathcal{E}_j(\zeta_j) \le C\,. \tag{5.22}$$

- *Supports.* All these functions vanish outside $B_j(2\Lambda^3) \subset B_j(R)$.

Step 4: the flux is almost divergence-free. We claim that there is $\varepsilon_j \to 0$ such that, for $1 \le i \le N$ and every $v \in \mathcal{C}_p$ with support in $B_j(R)$,

$$\Big|\frac{\Psi(r_j)}{m(B_j)} \int_{\mathcal{S}} \langle T_{j,i}, \nabla v\rangle\, d\nu\Big| \le \varepsilon_j \big(\mathcal{E}_j(v)^{1/p} + \|v\|_{L^p(m_j)}\big)\,. \tag{5.23}$$

For a harmonic coordinate the left-hand side would vanish. Here the estimate comes from the approximation of Step 2: for each fixed k we write $T_l = (T_l - V_{lk}) + V_{lk}$. By locality $\nabla v = 0$ ν-almost everywhere outside $B_j(R)$. By (2.7), Hölder's inequality and (5.15) on $B_j(R)$,

$$\begin{aligned}\Big|\int_{\mathcal{S}} \langle T_l, \nabla v\rangle\, d\nu\Big| &\le \Big(\int_{B_j(R)} |T_l - V_{lk}|^{p'}\, d\nu\Big)^{1/p'} \Big(\int_{\mathcal{S}} |\nabla v|^p\, d\nu\Big)^{1/p} + \Big|\int_X v\, h_{lk}\, dm\Big| \\ &\le C\, 2^{-k/p'} \mu(B_j)^{1/p'} \Big(\int_{\mathcal{S}} |\nabla v|^p\, d\nu\Big)^{1/p} + C_k\, \mu(B_j)^{1/p'} m(B_j)^{1/p} \|v\|_{L^p(m_j)}\,.\end{aligned}$$

We multiply by $\Psi(r_j)/(m(B_j)\vartheta_j^{p-1})$ and sum with the coefficients A_{il}. Since

$$\vartheta_j^{p-1} = \Big(\frac{\Psi(r_j)\,\mu(B_j)}{m(B_j)}\Big)^{1/p'}\,,$$

the result is

$$\Big|\frac{\Psi(r_j)}{m(B_j)} \int_{\mathcal{S}} \langle T_{j,i}, \nabla v\rangle\, d\nu\Big| \le C\, 2^{-k/p'} \mathcal{E}_j(v)^{1/p} + C_k\, \Psi(r_j)^{1/p} \|v\|_{L^p(m_j)}\,.$$

The infimum over k of $C 2^{-k/p'} + C_k \Psi(r_j)^{1/p}$ tends to zero as $j \to \infty$, because $\Psi(r_j) \to 0$. This proves (5.23).

Step 5: the normalized oscillation is of order one. Step 3 gave $\vartheta_j \ge c\, r_j$. We now prove the reverse inequality, by a test with a cutoff which is the analogue of the nonconcentration

argument in [52, Lemma 3.5(1)]. Put $S = \sup_{B_j(\Lambda)} |H_j^1|$, which is uniformly bounded by (5.20). We test (5.23) with $i = 1$ and $v = \chi_j^p H_j^1$, for which $\nabla v = \chi_j^p \nabla H_j^1 + p\,\chi_j^{p-1} H_j^1 \nabla \chi_j$. By (5.17), the contribution of $\chi_j^p \nabla H_j^1$ is

$$I_j = \int \chi_j^p \, (AMA)_{11} \, d\mu_j \,, \qquad 1 + o(1) \le I_j \le C \,,$$

by (5.19), $\chi_j = 1$ on B_j and $\mu_j(B_j) = 1$. By Hölder's inequality, (5.18) and (5.22), the remaining contribution has absolute value at most

$$pS\Big(\frac{\Psi(r_j)}{m(B_j)} \int \chi_j^p |T_{j,1}|^{p'} \, d\nu\Big)^{1/p'} \mathcal{E}_j(\chi_j)^{1/p} \le CS \,.$$

The same bounds and (5.21) give $\mathcal{E}_j(v) \le C(1 + S^p)$ and $\|v\|_{L^p(m_j)} \le CS$. Thus (5.23) yields

$$1 \le C \sup_{B_j(\Lambda)} |H_j^1| + o(1) \le C \, \frac{r_j}{\vartheta_j} + o(1) \,. \tag{5.24}$$

It follows that $\vartheta_j \le C r_j$, that is, $\mu(B_j) \le C r_j^p m(B_j)/\Psi(r_j)$. Together with (5.16), this gives

$$\mu(B_j(R)) \le C \, \frac{r_j^p \, m(B_j)}{\Psi(r_j)} \,. \tag{5.25}$$

Since (5.14) gave the reverse bound, $\vartheta_j \asymp r_j$: the normalization by ϑ_j is a rescaling of the metric by the factor r_j.

Step 6: a ball on which the coordinates become constant. This is the step which uses $a > p$. The lower bound (5.14) holds at every point of K and at every small scale, whereas the upper bound (5.25) holds at the scale r_j. Comparing them at the scales δr_j and r_j shows that K occupies a small part of B_j. Fix $0 < \delta < 1/(4\Lambda)$, and choose a maximal δr_j-separated family (y_i) in $K \cap B_j$. The balls $B(y_i, \delta r_j/2)$ are pairwise disjoint and lie in $B_j(2)$. Let $W_j = \bigcup_i B(y_i, (\Lambda+2)\delta r_j)$. Doubling, (5.14) at the points $y_i \in K$, (5.25) and the lower scaling of Ψ give

$$\begin{aligned} m(W_j) &\le C \sum_i m(B(y_i, \delta r_j/2)) \le C \, \frac{\Psi(\delta r_j)}{\delta^p \, r_j^p} \, \mu(B_j(2)) \\ &\le C \, \frac{\Psi(\delta r_j)}{\delta^p \, \Psi(r_j)} \, m(B_j) \le C \, \delta^{a-p} \, m(B_j) \,, \end{aligned} \tag{5.26}$$

with a constant independent of δ. Since $a > p$, we may fix δ so small that $m(W_j) < m(B_j(1/2))$. Then there is $z_j \in B_j(1/2) \setminus W_j$. The ball $U_j = B(z_j, \delta r_j)$ satisfies $\Lambda U_j \subset B_j$, and $\Lambda U_j \cap K = \varnothing$ by the maximality of the separated family. By (5.1) on U_j, the comparability of Ψ at comparable radii, (5.18) and (5.16),

$$\int_{U_j} |H_j^i - (H_j^i)_{U_j}|^p \, dm_j \le C_\delta \, \mu_j(\Lambda U_j) \le C_\delta \, \mu_j(B_j \setminus K) \longrightarrow 0 \,.$$

The vectors $(H_j)_{U_j}$ are bounded by (5.20); after passing to a subsequence, they converge to some $z \in \mathbb{R}^N$. By (5.21),

$$\int_{U_j} |H_j - z|^p \, dm_j \longrightarrow 0 \,, \qquad m_j(U_j) \ge c_\delta > 0 \,. \tag{5.27}$$

Step 7: image measures. Define finite measures on $\mathbb{R}^N$ by

$$\eta_j = (H_j)_{\#}\big(m_j \restriction_{B_j(2\Lambda)}\big)\,, \qquad \kappa_j = (H_j)_{\#}(\zeta_j^p\,\mu_j)\,.$$

By (5.20), (5.21) and (5.19), their supports lie in a fixed compact ball and their masses are bounded. Along a subsequence they converge weakly to finite measures η and κ. We prove three properties of the limits.

(a) The measure η has an atom at z and is not concentrated at z:

$$0 < \eta(\{z\}) < \eta(\mathbb{R}^N)\,. \tag{5.28}$$

Indeed, (5.27) and Chebyshev's inequality give, for every $\varepsilon > 0$,

$$\eta_j(\overline{B}(z,\varepsilon)) \geq m_j(U_j) - \varepsilon^{-p}\int_{U_j} |H_j - z|^p\,dm_j \geq c_\delta - o(1)\,.$$

Since $\limsup_j \eta_j(E) \leq \eta(E)$ for every closed set E, letting $\varepsilon \downarrow 0$ gives $\eta(\{z\}) \geq c_\delta$. To show that η is not a point mass, put $\ell(w) = (A^{-1}w)_1$, so that $\ell(H_j) = (f_1 - f_1(x))/\vartheta_j$. The lower oscillation of f_1 on B_j and its Lipschitz bound give, as in Step 1, two balls of radius εr_j in $B_j(2)$ on which the values of f_1 are separated by cr_j, with fixed $c, \varepsilon > 0$. Since $\vartheta_j \asymp r_j$ and each ball has m_j-measure bounded below by (5.21),

$$\inf_{b\in\mathbb{R}} \int |\ell(w) - b|^p\,d\eta_j(w) \geq c' > 0\,.$$

The supports of η_j lie in a fixed compact set, so the infimum can be restricted to a fixed compact interval of b and passes to the weak limit. The same positive lower bound holds for η, which proves (5.28).

(b) The measure κ satisfies the integration by parts inequality (5.11). We adapt the integration by parts for image measures of [87, Section 3] and [52, Lemma 3.4(2)]. Fix $\varphi = (\varphi_1, \dots, \varphi_N) \in C_c^\infty(\mathbb{R}^N, \mathbb{R}^N)$ and test (5.23) with $v_i = \zeta_j^p\,\varphi_i(H_j)$. By (5.18) and (5.22), $\mathcal{E}_j(v_i)$ and $\|v_i\|_{L^p(m_j)}$ are bounded uniformly in j, so the pairings tend to zero. Expanding the gradients by the chain rule (1.6) and using (5.17),

$$o(1) = \sum_{i,l}\int \zeta_j^p\,\partial_l\varphi_i(H_j)\,(AMA)_{il}\,d\mu_j + \frac{p\,\Psi(r_j)}{m(B_j)}\sum_i \int \zeta_j^{p-1}\varphi_i(H_j)\,\langle T_{j,i}, \nabla\zeta_j\rangle\,d\nu\,.$$

By (5.19), the first term is $\int \operatorname{div}\varphi\,d\kappa_j + o(1)$. By Hölder's inequality, (5.18) and (5.22), the absolute value of the second term is at most

$$C\sum_i \Big(\int \zeta_j^p\,|\varphi_i(H_j)|^{p'}\,d\mu_j\Big)^{1/p'} \mathcal{E}_j(\zeta_j)^{1/p} \leq C\Big(\int |\varphi|^{p'}\,d\kappa_j\Big)^{1/p'}\,.$$

Passing to the limit proves (5.11) for κ, with a constant independent of φ.

(c) The pair (η, κ) satisfies a Poincaré inequality:

$$\inf_{b\in\mathbb{R}} \int_{\mathbb{R}^N} |\phi - b|^p\,d\eta \leq C\int_{\mathbb{R}^N} |\nabla\phi|^p\,d\kappa \qquad (\phi \in C_c^\infty(\mathbb{R}^N))\,. \tag{5.29}$$

The passage to the image measures follows [52, proof of Lemma 3.5], with the exponent p. We apply (5.1) to $u = \phi(H_j) \in \mathcal{C}_p + \mathbb{R}$ on the ball $B_j(2\Lambda)$. Its gradient satisfies the identity

$$\frac{\Psi(r_j)}{m(B_j)}\,|\nabla u|^p\,\nu = \big\langle \nabla\phi(H_j), AMA\,\nabla\phi(H_j)\big\rangle^{p/2}\mu_j\,,$$

which is proved as (5.17). The matrices AMA are uniformly bounded and converge to I in $L^1(\mu_j)$ on $B_j(R)$, by (5.19), and $t \mapsto t^{p/2}$ is uniformly continuous on bounded intervals. Since $\zeta_j = 1$ on the ball $B_j(2\Lambda^2)$, over which the energy is integrated in (5.1), we get

$$\int |\phi - b_j|^p \, d\eta_j \le C \int |\nabla\phi|^p \, d\kappa_j + o(1) \,, \qquad b_j = (\phi(H_j))_{B_j(2\Lambda)} \,.$$

Since $|b_j| \le \|\phi\|_\infty$, we may pass to a further subsequence along which b_j converges, and the weak convergence proves (5.29).

Step 8: the contradiction. Since $N \ge p$ and κ satisfies (5.11), Lemma 5.11 provides functions ϕ_k satisfying (5.12) for the point z and the measure κ. Since $0 \le \phi_k \le 1$, the infimum over b in (5.29) may be restricted to $[0,1]$. By dominated convergence, uniformly for $b \in [0,1]$,

$$\int |\phi_k - b|^p \, d\eta \longrightarrow (1-b)^p \, \eta(\{z\}) + b^p (\eta(\mathbb{R}^N) - \eta(\{z\})) \,.$$

By (5.28), the minimum of the right-hand side over $b \in [0,1]$ is positive. On the other hand, the right-hand side of (5.29) for $\phi = \phi_k$ tends to zero. This contradiction proves that $n_* < p$. □

Cutoff functions in the domain of the Dirichlet form. Let $p = 2$, and suppose that the cutoff functions of (5.2) are only assumed to belong to $\mathcal{F} = W^{1,2}(X,m)$. The proof is the same, with the following modifications.

- The inequality (5.1) and the identity (2.7) extend to $\mathcal{F}$ by approximation in the graph norm. The test functions belong to $\mathcal{F}$, and ∇v stands for $\nabla_2 v$; the claim (5.23) is stated for $v \in \mathcal{F}$ with m-essential support in $B_j(R)$.
- In products and in integrals against ν we use quasi-continuous representatives, denoted by a tilde. By Proposition 2.7(iii), ν is minimal energy-dominant, and it charges no set of zero capacity [69, Section 3.2]. The bounds $0 \le \chi \le 1$ and the constant values on open sets pass to the quasi-continuous representative [69, Section 2.1].
- For bounded $u, v \in \mathcal{F}$, the product belongs to $\mathcal{F}$ and

  $$\nabla_2(uv) = \tilde{u} \, \nabla_2 v + \tilde{v} \, \nabla_2 u \qquad \nu\text{-almost everywhere} \,.$$

  Indeed, the pairings of both sides with the gradients of a spanning sequence in $\mathcal{C}_2$ are equal, by the product rule for mutual energy measures [69, Section 3.2] and by (2.14). The formula extends to the case where one bounded factor belongs to $\mathcal{F} + \mathbb{R}$. Thus the products $\chi_j^2 (H_j^1 - b)$ and $\zeta_j^2 \varphi_i(H_j)$ belong to $\mathcal{F}$.
- By strong locality, $\nabla_2 v = 0$ ν-almost everywhere outside the m-essential support of v. The bounds (5.22) hold m-almost everywhere; the quasi-continuous representatives of χ_j and ζ_j are equal to one ν-almost everywhere on the inner balls and vanish ν-almost everywhere outside the closures of the outer balls, which lie in $B_j(R)$. The bounds for H_j extend to these closures by continuity.

With these facts, the computations of Steps 4 to 8 apply to cutoff functions in $\mathcal{F}$.

6 Self-improvement of Poincaré inequalities

This chapter studies how the Poincaré inequality at the cutoff scale depends on the exponent. Its main result is a self-improvement theorem. Keith and Zhong [101, Theorem 1.0.1] proved that, on a complete metric space with a doubling measure, a $(1,p)$-Poincaré inequality with $p>1$ implies a $(1,q)$-Poincaré inequality for some $q<p$. Theorem 6.1 is the analogue for gradient structures: a $(1,p)$-Poincaré inequality at the cutoff scale, together with a cutoff bound of some order $b<p$, implies an (s,q)-Poincaré inequality with $q<p<s$, in which the oscillation and the gradient are integrated over the same ball. Consequently the set of exponents at which the scales match is open from below (Corollary 6.2).

The difficulty and the method. The theorem is proved by modifying a function on the set where a maximal function of its oscillation is large, as in the proof of Keith and Zhong, in the form given by Kinnunen, Lehrbäck, Vähäkangas and Zhong [109, Theorems 4.3 and 5.8]. In the classical arguments, the energy of the modified function on the bad set is bounded by a power of its Lipschitz constant times the measure of the bad set. Here the gradient is integrated against ν, and the ν-measure of the bad set may be infinite, even when its m-measure is small. The modification of Lemma 6.5 uses cutoffs satisfying, for $B=B(x,r)$,

$$r^b \int_{\mathcal{S}} |\nabla \psi_x^r|^b \, d\nu \le C\,\Theta(B)\,, \qquad \Theta(B)=\frac{r^2\,m(B)}{\Psi_{C,2}(r)}\,.$$

The Poincaré inequality controls the sum of $\Theta(B)$ over the disjoint stopping balls whose fixed dilations cover the bad set.

Related results. A quantitative proof of the theorem of Keith and Zhong through connectivity, including a localization, is given in [63, Theorems 1.4–1.6 and Section 3.1]; the axiomatic approach to Sobolev spaces relevant here is developed in [68]. A self-improvement theorem for two measures is proved in [106, Theorem 6.1], for two doubling measures which are finite on balls, under balance, independence and absolute continuity conditions. It does not cover the case where m and ν are mutually singular or ν is infinite on balls, which is the case of interest here.

6.1 Setting and the self-improvement theorem

Standing assumptions. Throughout this chapter, $(\mathcal{S},\nu,\nabla)$ is a gradient structure on (X,d,m), where (X,d) is *proper*, that is, its closed balls are compact, and *geodesic*, and $0<r_0\le \operatorname{diam} X$, with $r_0=\infty$ allowed on an unbounded space. We assume VD and the comparability (3.2) of the cutoff scale $\Psi_{C,2}$ for radii below r_0.

The gauge of a ball. For a ball $B=B(x,r)$ put

$$\Theta(B)=\frac{r^2\,m(B)}{\Psi_{C,2}(r)}=\frac{r^p\,m(B)}{\Psi_{C,p}(r)} \qquad (1\le p<\infty)\,. \tag{6.1}$$

Under $\mathrm{V}(\Phi)$ this is comparable to the gauge $\Theta(r)$ of (3.14); in general it depends on the ball. By doubling and (3.2), $\Theta(\lambda B)\asymp\Theta(B)$ for a fixed dilation λ.

Cutoffs of order b. For $1 \le b < \infty$, the *cutoff condition of order b* requires, for every ball $B = B(x,r)$ with $0 < r < r_0/4$, a function $\psi_x^r \in \mathrm{Lip}_c(X)$ with

$$\mathbf{1}_B \le \psi_x^r \le \mathbf{1}_{2B}\,, \qquad \mathrm{Lip}(\psi_x^r) \le \frac{C_B}{r}\,, \qquad \int_{\mathcal{S}} |\nabla \psi_x^r|^b \, d\nu \le C_B \, \frac{\Theta(B)}{r^b}\,. \tag{6.2}$$

Since $|\nabla \psi_x^r| \le C_B/r$, the condition of order b implies the condition of every order $e \ge b$, with the constant C_B^{1+e-b}. In particular Lemma 3.5 applies at those exponents. The condition of order one is $\mathrm{B}(\Psi_{C,2})$:

$$\int_{\mathcal{S}} |\nabla \psi_x^r| \, d\nu \le C_B \, \frac{\Theta(B)}{r}\,. \tag{6.3}$$

Matching exponents. An exponent $p \in [1,\infty)$ is called *matching* if there are C and Λ such that, for all sufficiently small balls $B = B(x,r)$,

$$\Big(\fint_B |v - v_B|^p \, dm \Big)^{1/p} \le C\, r \Big(\frac{1}{\Theta(B)} \int_{\Lambda B \cap \mathcal{S}} |\nabla v|^p \, d\nu \Big)^{1/p} \qquad (v \in \mathcal{C}_p)\,. \tag{6.4}$$

We write M for the set of matching exponents. Since $r^p/\Theta(B) = \Psi_{C,p}(r)/m(B)$, the inequality (6.4) is $\mathrm{PI}_p(\Psi_{P,p})$ with $\Psi_{P,p} = \Psi_{C,p}$: an exponent is matching if the Poincaré inequality holds at the cutoff scale. Under (6.2) with $b \le p$, the core $\mathcal{C}_p$ spans the fibres, by Proposition 1.9, and a matching exponent p is admissible, by Theorem 4.3. The inequality (6.4) then extends to $W^{1,p}(X,m)$, with the closed gradient; we call this extension the *closed Poincaré inequality*.

Theorem 6.1 (Self-improvement)**.** *Let $1 < p < \infty$ and $1 \le b < p$. Assume (6.2) and the $(1,p)$-Poincaré inequality at the cutoff scale*

$$\fint_B |v - v_B| \, dm \le C\, r \Big(\frac{1}{\Theta(B)} \int_{\Lambda B \cap \mathcal{S}} |\nabla v|^p \, d\nu \Big)^{1/p} \qquad (v \in \mathcal{C}_p) \tag{6.5}$$

at radii $r < r_0/\Lambda$. There exist $s > p$, $0 < \varepsilon_0 < p - b$, $A \ge 2$ and $C < \infty$ such that for every $q \in [p - \varepsilon_0, p]$, every $u \in \mathcal{C}_q$ and every ball $B = B(x,r)$ with $r < r_0/A$,

$$\Big(\fint_B |u - u_B|^s \, dm \Big)^{1/s} \le C\, r \Big(\frac{1}{\Theta(B)} \int_{B \cap \mathcal{S}} |\nabla u|^q \, d\nu \Big)^{1/q}\,. \tag{6.6}$$

The constants depend only on p, b, the doubling constant, the constants in (3.2), and the constants in (6.2) and (6.5).

We call constants with this dependence *structural*. The theorem improves the hypothesis in three ways: the exponent of the gradient decreases from p to q, the exponent of the oscillation increases from 1 to s, and the dilation Λ of the ball on the right-hand side is removed.

By Jensen's inequality, (6.6) implies $[p - \varepsilon_0, p] \subset M$. Each such exponent q is greater than b, so the structure is q-admissible, and applying (6.6) to differences of approximating sequences extends it to $W^{1,q}(X,m)$, with the closed gradient ∇_q.

Corollary 6.2. *Under $\mathrm{B}(\Psi_{C,2})$, if $M \ne \varnothing$ and $\inf M > 1$, then $\inf M \notin M$.*

Proof. If $p = \inf M > 1$ belonged to M, the theorem, with $b = 1$, would give a matching exponent strictly below p. □

For the tree of cubes of Chapter 15, $M = (n,\infty)$, and every ball has infinite gradient measure (Propositions 15.5 and 15.10). Thus the set of matching exponents can be an open half-line with an endpoint larger than one, and the theorem does not extend to the endpoint.

Strategy of the proof. The proof has four parts, which occupy the next four sections.

- *Scales and chains* (Section 6.2). The hypotheses imply that $\Psi_{C,2}(r)/r^2$ is almost nondecreasing. With this information, a chaining argument along geodesics raises the exponent of the oscillation and removes the dilation of the ball.
- *Cutoff modification* (Section 6.3). A Lipschitz function is modified on a family of balls, with a gradient on these balls whose $L^b(\nu)$ norm is bounded by the sum of the gauges of the balls.
- *The compact case* (Section 6.4). On a compact space, the exponent of the gradient is lowered in a Poincaré inequality on the whole space. This is the core of the proof: the stopping-time argument of Keith and Zhong, with the gauge in place of the measure.
- *Localization* (Section 6.5). Each ball is enclosed in a uniform domain, whose completion is a compact space to which the compact case applies.

6.2 Scale comparison and chains

On a Riemannian manifold $\Psi_{C,2}(r)/r^2$ is constant, and on a local tree it is $\Phi(r)/r$. The first lemma shows that this ratio is always almost nondecreasing when a Poincaré inequality holds at the cutoff scale. The proof tests the Poincaré inequality at the radius R with the discrete convolution, at a smaller radius r, of a cutoff of radius R.

Lemma 6.3. *Let $1 \le b \le p < \infty$ and assume (6.2) and (6.5) at p. There are structural constants $c_0 > 0$ and $C < \infty$ such that*

$$\frac{\Psi_{C,2}(r)}{r^2} \le C\,\frac{\Psi_{C,2}(R)}{R^2} \qquad (0 < r \le R < c_0 r_0)\,. \tag{6.7}$$

Consequently, for every $1 \le s < \infty$,

$$\Psi_{C,s}(r) \le C\Big(\frac{r}{R}\Big)^s \Psi_{C,s}(R) \qquad (0 < r \le R < c_0 r_0)\,. \tag{6.8}$$

In particular $\Psi_{C,s}$ satisfies (3.1), with the lower exponent s and a finite upper exponent.

Proof. Fix $x \in X$, take $R < c_0 r_0$ and $0 < r \le R/6$, and let $f = \psi_x^R$. Let f_r be the discrete convolution of f at the radius r (Lemma 3.5, at the exponent p). Since f is C/R-Lipschitz and vanishes outside $B(x, 2R)$, doubling gives

$$D_p(f, 5r) \le C\, m(B(x,R))\Big(\frac{r}{R}\Big)^p .$$

Lemma 3.5(ii) and the identity $r^p/\Psi_{C,p}(r) = r^2/\Psi_{C,2}(r)$ therefore yield

$$\int_{\mathcal{S}} |\nabla f_r|^p \, d\nu \le C\,\frac{m(B(x,R))}{R^p}\,\frac{r^2}{\Psi_{C,2}(r)}\,. \tag{6.9}$$

Since $\varphi_i(y) \ne 0$ implies $B_i \subset B(y, 3r)$, the function f_r is equal to one on $B(x, R/2)$ and vanishes outside $B(x, 3R)$.

Choose c_0 so that $8R < \operatorname{diam} X$ and all the dilated balls below lie in the range of the hypotheses. Since X is geodesic, there is a point z with $d(x,z) = 4R$. The ball $B(z, R/2)$ lies in $B(x, 5R) \setminus B(x, 3R)$, and by doubling

$$m(B(x, R/2)) \ge c\, m(B(x, 6R))\,, \qquad m(B(z, R/2)) \ge c\, m(B(x, 6R))\,.$$

Since $f_r = 1$ on the first ball and $f_r = 0$ on the second, for every constant a,

$$\fint_{B(x,6R)} |f_r - a|\, dm \geq c(|1-a| + |a|) \geq c\,.$$

We apply (6.5) to f_r on $B(x,6R)$, and use (6.9), doubling and (3.2):

$$c \leq C\,R\Big(\frac{\Psi_{C,2}(R)}{R^2\, m(B(x,R))} \cdot \frac{m(B(x,R))}{R^p}\,\frac{r^2}{\Psi_{C,2}(r)}\Big)^{1/p} = C\,\Big(\frac{\Psi_{C,2}(R)}{R^2}\,\frac{r^2}{\Psi_{C,2}(r)}\Big)^{1/p}.$$

This is (6.7) for $r \leq R/6$. For $R/6 < r \leq R$ it follows from (3.2). Multiplying (6.7) by r^s gives (6.8), and iterating the upper bound in (3.2) gives a finite upper scaling exponent. □

In the rest of the proof of Theorem 6.1 we replace r_0 by $c_0 r_0$, so that (6.7) and (6.8) hold for $0 < r \leq R < r_0$.

Chains. A Poincaré inequality with a dilated ball on the right-hand side implies, in a geodesic space, the same inequality without dilation and with a higher exponent for the oscillation. This is the chaining argument of Hajłasz and Koskela [79, Theorems 9.3 and 9.10]. In the next lemma the energy measure and the reference measure enter the chain estimate separately, and the scale decay (6.8) provides the summability.

Lemma 6.4. *Let $1 \leq e \leq s < \infty$. Assume the scale decay (6.8) and the inequality (6.5) with the exponent e in place of p, for $v \in \mathcal{C}_e$. If D_m is a doubling exponent in (3.6) and*

$$D_m\left(\frac{1}{e} - \frac{1}{s}\right) < 1\,, \tag{6.10}$$

then for every $u \in \mathcal{C}_e$ and all sufficiently small balls $B = B(o,R)$,

$$\Big(\fint_B |u - u_B|^s\, dm\Big)^{1/s} \leq C\,R\Big(\frac{1}{\Theta(B)}\int_{B\cap\mathcal{S}} |\nabla u|^e\, d\nu\Big)^{1/e}. \tag{6.11}$$

Proof. Step 1: a chain of balls. Fix $0 < a < 1/(2\Lambda)$. For $y \in B$ let γ_y be a geodesic from y to o, parametrized by arc length. Put $t_j = 2^{-j}R$ and

$$D_j(y) = B(\gamma_y(\min\{t_j, d(y,o)\}), a\,t_j)\,, \qquad j \geq 0\,.$$

Then $D_0(y) = B(o, aR) =: D_0$ does not depend on y, and the centres of the balls $D_j(y)$ converge to y. Between $D_j(y)$ and $D_{j+1}(y)$ insert at most C/a balls of radius $a\,t_{j+1}$ whose consecutive centres are at distance at most $a\,t_{j+1}/2$ on γ_y. Two consecutive balls of the resulting chain contain a common ball of comparable volume. If β is a ball of the chain between the indices j and $j+1$, then

$$\Lambda\beta \subset B \cap B(y, 2t_j)\,,$$

because the distance from $\gamma_y(t)$ to $X \setminus B$ is at least t and $\Lambda a < 1/2$.

Step 2: the chain estimate. Apply (6.5), with the exponent e, to each ball of the chain, in order to bound the difference of the averages of u on two consecutive balls. By doubling, the measure of each ball of the chain between the indices j and $j+1$ is comparable to $m(B(y,t_j))$. Since u is continuous, the averages along the chain converge to $u(y)$. With the measure $\eta = \mathbf{1}_{B\cap\mathcal{S}}|\nabla u|^e\,\nu$, we obtain

$$|u(y) - u_{D_0}| \leq C\sum_{j\geq 0} \Psi_{C,e}(t_j)^{1/e}\left(\frac{\eta(B(y,2t_j))}{m(B(y,t_j))}\right)^{1/e} \qquad (y \in B)\,. \tag{6.12}$$

Step 3: integration. For $0 < t \le R$, Fubini's theorem and doubling give

$$\fint_B \frac{\eta(B(y,2t))}{m(B(y,t))}\,dm(y) \le \frac{C}{m(B)}\int_B \frac{m(B(z,2t))}{m(B(z,t))}\,d\eta(z) \le C\,\frac{\eta(B)}{m(B)}\,.$$

Also $B \subset B(y,2R)$ for $y \in B$, so (3.6) gives

$$\sup_{y\in B}\frac{\eta(B(y,2t))}{m(B(y,t))} \le C\Big(\frac{R}{t}\Big)^{D_m}\frac{\eta(B)}{m(B)}\,.$$

Interpolating between these two bounds, the norm in $L^s(B, m/m(B))$ of the function $y \mapsto (\eta(B(y,2t))/m(B(y,t)))^{1/e}$ is at most $C(R/t)^{D_m(1/e-1/s)}(\eta(B)/m(B))^{1/e}$. By (6.8), $\Psi_{C,e}(t_j)^{1/e} \le C\,2^{-j}\Psi_{C,e}(R)^{1/e}$. By Minkowski's inequality and (6.10), the norm of the right-hand side of (6.12) is therefore at most

$$C\,\Psi_{C,e}(R)^{1/e}\Big(\frac{\eta(B)}{m(B)}\Big)^{1/e}\sum_{j\ge 0}2^{-j(1-D_m(1/e-1/s))} \le C\,R\,\Big(\frac{\eta(B)}{\Theta(B)}\Big)^{1/e},$$

by (6.1). Replacing u_{D_0} by u_B costs at most a factor two. □

In particular, with $e = s = p$, the hypothesis (6.5) implies (6.4) with $\Lambda = 1$, after decreasing the range of radii.

6.3 The cutoff modification

The next lemma is the main technical device of the chapter. Given a Lipschitz function v and a family of balls B_i, it produces a Lipschitz function $w \le v$ which is constant on each B_i near the lower bound of v, and which coincides with v away from the balls. On the union of the balls the gradient of w is bounded by a scalar function F, built from the gradients of cutoff functions of the balls, and the b-th moment of F is controlled by the sum of the numbers Θ_i which bound the energies of these cutoffs. No estimate of the ν-measure of the union of the balls is involved. The lemma is used with the stopping balls of a sharp maximal function, in Section 6.4.

Lemma 6.5 (Cutoff modification)**.** *Let $b \ge 1$, $v \in \mathrm{Lip}_b(X)$, $L = \mathrm{Lip}(v)$, and let (B_i) be a nonempty finite or countable family of balls $B_i = B(x_i, r_i)$ with $\sup_i r_i < \infty$. Suppose that $\psi_i \in \mathrm{Lip}_b(X)$ satisfy*

$$\mathbf{1}_{B_i} \le \psi_i \le \mathbf{1}_{2B_i}\,, \qquad \mathrm{Lip}(\psi_i) \le \frac{C_B}{r_i}\,, \qquad \int_{\mathcal{S}} |\nabla\psi_i|^b\,d\nu \le C_B\,\frac{\Theta_i}{r_i^b} \tag{6.13}$$

for numbers $\Theta_i \ge 0$. Put $a_i = \inf_{B_i} v$ and

$$T_i = \min\{v\,,\ \max\{v, a_i + 4Lr_i\} - 4Lr_i\psi_i\}\,, \qquad w = \inf_i T_i\,. \tag{6.14}$$

Then the following assertions hold.

(i) *The function w is $(1+4C_B)L$-Lipschitz, $v - 3L\sup_i r_i \le w \le v$, and $w = v$ outside $\bigcup_i 2B_i$.*

(ii) *The function $F = 4L\sup_i r_i|\nabla\psi_i|$ satisfies $0 \le F \le 4C_BL$, vanishes ν-almost everywhere outside $\bigcup_i 2B_i$, and*

$$|\nabla w| \le \mathbf{1}_{X\setminus\bigcup_i B_i}|\nabla v| + F\,, \qquad \int_{A\cap\mathcal{S}} F^b\,d\nu \le 4^bC_BL^b\sum_{\{i:\,2B_i\cap A\ne\varnothing\}}\Theta_i \tag{6.15}$$

for every Borel set $A \subset X$.

(iii) $\displaystyle\int_{\bigcup_i B_i} |\nabla w|^b \, d\nu \le 4^b C_B L^b \sum_i \Theta_i$.

For the empty family we set $w = v$ and $F = 0$.

The function T_i is the modification of v on the single ball B_i: on $2B_i$ it is the minimum of v and of the function $a_i + 4Lr_i(1-\psi_i)$, which is equal to a_i on B_i and exceeds v where $\psi_i = 0$.

Proof. Step 1: the functions T_i. If $z \in 2B_i$, every point of B_i is at distance less than $3r_i$ from z, so $v(z) \le a_i + 3Lr_i$. Hence

$$T_i = \min\{v, a_i + 4Lr_i(1-\psi_i)\} \quad \text{on } 2B_i\,, \qquad T_i = v \quad \text{outside } 2B_i\,. \tag{6.16}$$

Thus $T_i = a_i$ on B_i and $v - 3Lr_i \le T_i \le v$. The formula (6.14) shows that each T_i is $(1 + 4C_B)L$-Lipschitz, and this bound passes to the infimum. This proves (i).

Step 2: finite families. Put $w_n = \min_{i\le n} T_i$. On $\{w_n < v\}$, which is contained in $\bigcup_{i\le n} 2B_i$, the function w_n coincides, on each set of a finite Borel partition, with one of the functions $a_i + 4Lr_i(1-\psi_i)$, so locality identifies ∇w_n with one of the fields $-4Lr_i\nabla\psi_i$. On $\{w_n = v\} \cap B_i$ one has $v = w_n \le T_i = a_i \le v$, so that v is constant on this set and $\nabla w_n = \nabla v = 0$ there. Consequently

$$|\nabla w_n| \le \mathbf{1}_{X\setminus\bigcup_{i\le n} B_i} |\nabla v| + F\,.$$

Step 3: countable families. The functions w_n decrease to w, with uniform bounds on the supremum norm and the Lipschitz constant, so (G) gives the weak-$*$ convergence of the gradients. For every $\xi \in L^1(\mathcal{S}, \mathsf{H}, \nu)$, the inequality of Step 2 and dominated convergence yield

$$\Big|\int_{\mathcal{S}} \langle \nabla w, \xi\rangle \, d\nu\Big| \le \int_{\mathcal{S}} (\mathbf{1}_{X\setminus\bigcup_i B_i} |\nabla v| + F)|\xi| \, d\nu\,.$$

Testing with fields ξ which are supported on sets of finite measure and point in the direction of ∇w proves the pointwise inequality in (6.15).

Step 4: the bounds on F. By (6.13) and (G), $r_i|\nabla\psi_i| \le C_B$, and by locality $\nabla\psi_i = 0$ ν-almost everywhere outside $2B_i$. A supremum of nonnegative functions is bounded by their sum, so

$$\int_{A\cap\mathcal{S}} F^b \, d\nu \le (4L)^b \sum_{\{i:\, 2B_i\cap A\ne\varnothing\}} r_i^b \int_{\mathcal{S}} |\nabla\psi_i|^b \, d\nu\,,$$

which gives the integral bound in (6.15). Assertion (iii) follows from (ii) with $A = X$, since $|\nabla w| \le F$ on $\bigcup_i B_i$. □

The proof uses only the Leibniz rule, the Lipschitz bound and the weak-$*$ continuity in (G), through the locality of Lemma 1.5. It therefore applies to any linear map $\nabla : \mathrm{Lip}(X) \to L^\infty(\mathcal{S}, \mathsf{H}, \nu)$ with these properties; this remark is used on the auxiliary compact spaces of the next section.

6.4 The compact case

The core of the proof of Theorem 6.1 is an estimate on a compact space, in which the oscillation and the gradient are integrated over the whole space. We state its hypotheses abstractly, as conditions (C1)–(C4) on an auxiliary compact space, because in the application this space is the completion of a domain of X, with the restricted measures and gradient.

Setting of this section.

- (X,d) is a compact geodesic space of diameter $D>0$, m is a finite Borel measure on X with full support, ν is a σ-finite Borel measure carried by a Borel set $\mathcal{S}$, H is a measurable field over $\mathcal{S}$, and $\nabla : \mathrm{Lip}(X) \to L^\infty(\mathcal{S},\mathsf{H},\nu)$ is a linear map which satisfies (G).
- $1 \le b < p < \infty$, and $\Psi_{C,2}$ is positive and comparable at comparable radii in $(0,D]$. We set $\Psi_{C,s}(r) = r^{s-2}\Psi_{C,2}(r)$, define $\Theta(B)$ by (6.1), and put $\Theta(X) = D^2 m(X)/\Psi_{C,2}(D)$ and $r_X = D$.

A ball $B = B(x,r_B)$ is given with its centre and its radius. We assume that, for a constant $C_* < \infty$:

(C1) *Doubling.* $m(B(x,2r)) \le C_* m(B(x,r))$ for all $x \in X$ and $0 < r \le D$.

(C2) *Cutoffs.* Each ball $B = B(x,r)$ with $0 < r \le D$ admits $\psi_x^r \in \mathrm{Lip}(X)$ such that

$$\mathbf{1}_B \le \psi_x^r \le \mathbf{1}_{2B}\,, \qquad \mathrm{Lip}(\psi_x^r) \le \frac{C_*}{r}\,, \qquad \int_{\mathcal{S}} |\nabla \psi_x^r|^b \, d\nu \le C_* \, \frac{\Theta(B)}{r^b}\,.$$

(C3) *Scale comparison.* $\Psi_{C,2}(r)/r^2 \le C_* \Psi_{C,2}(R)/R^2$ for $0 < r \le R \le D$.

(C4) *Poincaré inequality.* For every $u \in \mathrm{Lip}(X)$ with $\nabla u \in L^p(\mathcal{S},\mathsf{H},\nu)$, for every ball B of radius at most D, and also for $B = X$ with $r_X = D$,

$$\int_B |u - u_B|^p \, dm \le C_* \, \Psi_{C,p}(r_B) \int_{B\cap\mathcal{S}} |\nabla u|^p \, d\nu\,.$$

In this section, C denotes a constant which depends only on p, b, C_* and the constants of the comparability of $\Psi_{C,2}$ at comparable radii.

Proposition 6.6. *In this setting, there are $\varepsilon_0 \in (0, p-b)$ and $C < \infty$ such that for every $q \in [p-\varepsilon_0, p)$ and every $u \in \mathrm{Lip}(X)$ with $\nabla u \in L^q(\mathcal{S},\mathsf{H},\nu)$,*

$$\Big(\fint_X |u - u_X|^p \, dm \Big)^{1/p} \le C\, D \Big(\frac{1}{\Theta(X)} \int_{\mathcal{S}} |\nabla u|^q \, d\nu \Big)^{1/q}\,. \tag{6.17}$$

The proof adapts the stopping-time argument of [109, Sections 4.4–4.6]. Throughout, $u \in \mathrm{Lip}(X)$ with $\nabla u \in L^p(\mathcal{S},\mathsf{H},\nu)$ is fixed, and

$$g = |\nabla u|\,.$$

Outline. Let $\omega(B)$ be the normalized L^p oscillation of u on a ball B, and $u^\sharp$ the corresponding maximal function. On the set where $u^\sharp \le t$ the function u is Ct-Lipschitz. The set $U_t = \{u^\sharp > t\}$ is covered by stopping balls, and its size is measured by the number $\Theta_u(t)$, the largest sum of the gauges of disjoint stopping balls. The Poincaré inequality gives the weak-type bound

$$t^p \, \Theta_u(t) \le C \int_{\mathcal{S}} g^p \, d\nu\,.$$

The goal is to improve the exponent p in this bound to some $q < p$, with $\int g^q \, d\nu$ on the right-hand side. For this, one modifies u on the stopping balls at k consecutive dyadic levels by Lemma 6.5, and averages the k modified functions. The averaged function coincides with u on the good set, and the p-th power of its gradient on the bad set carries a factor k^{-p} in front of a sum of k terms. Applying the Poincaré inequality to the averaged function gives a recurrence between the numbers $\lambda^p \Theta_u(\lambda)$ at different levels (Lemma 6.14), which can be absorbed after multiplication by λ^{q-p}, when $p-q$ is small.

The sharp maximal function

For a ball B of radius at most D, and for $B = X$ with $r_X = D$, let

$$\omega(B) = \frac{1}{r_B}\Big(\fint_B |u - u_B|^p \, dm\Big)^{1/p}, \qquad \lambda_0 = \omega(X)\,, \tag{6.18}$$

and

$$u^\sharp(x) = \max\Big\{\lambda_0\,,\ \sup_{B \ni x} \omega(B)\Big\}\,, \qquad U_t = \{u^\sharp > t\} \quad (t > 0)\,. \tag{6.19}$$

The supremum is over the balls of radius at most D which contain x. Whenever a ball of radius larger than D occurs in an argument, it is replaced by X, with $r_X = D$. The function $u^\sharp$ is lower semicontinuous, because the balls are open, so the sets U_t are open, and

$$\lambda_0 \le u^\sharp \le 2\,\mathrm{Lip}(u)\,.$$

If $\lambda_0 = 0$, then u is constant, because it is continuous and m has full support, and (6.17) holds. We assume $\lambda_0 > 0$. We shall use repeatedly that, by (C1), if $B \subset B'$ and $r_{B'} \le C r_B$, where B' may be X, then

$$\omega(B) \le C\,\omega(B')\,. \tag{6.20}$$

The first lemma says that u is Lipschitz on the set where $u^\sharp$ is bounded. It is the Lipschitz case of the telescoping argument in [109, Lemma 4.12].

Lemma 6.7. *For every $t \ge \lambda_0$,*

$$|u(x) - u(y)| \le C\,t\,d(x,y) \qquad (x, y \in X \setminus U_t)\,, \tag{6.21}$$

and there is $v_t \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(v_t) \le Ct$, $\inf u \le v_t \le \sup u$ and $v_t = u$ on $X \setminus U_t$. Moreover

$$g \le C\,u^\sharp \qquad \nu\text{-almost everywhere.} \tag{6.22}$$

Proof. Let $x, y \in X \setminus U_t$ and $\rho = d(x,y)$. The balls $B(x, 2^{-j}\rho)$, $j \ge 0$, contain x, so $\omega \le t$ on them, and

$$|u_{B(x,2^{-j-1}\rho)} - u_{B(x,2^{-j}\rho)}| \le C \fint_{B(x,2^{-j}\rho)} |u - u_{B(x,2^{-j}\rho)}|\, dm \le C\,t\,2^{-j}\rho\,.$$

Summing over j gives $|u(x) - u_{B(x,\rho)}| \le Ct\rho$, and similarly at y. The balls $B(x,\rho)$ and $B(y,\rho)$ are contained in $B(x, 2\rho)$, on which $\omega \le t$. Hence $|u_{B(x,\rho)} - u_{B(y,\rho)}| \le Ct\rho$. This proves (6.21). If $X \setminus U_t \ne \varnothing$, let v_t be the McShane extension [83, pp. 43–44] of $u|_{X\setminus U_t}$ with the constant Ct, truncated at $\inf u$ and $\sup u$. When $X \setminus U_t$ is empty, take v_t constant. By Lemma 1.5(ii) and (G), $g = |\nabla v_t| \le Ct$ ν-almost everywhere on $\{u^\sharp \le t\}$. Applying this to the rational numbers $t \ge \lambda_0$ gives (6.22). □

Stopping balls

The set U_t is covered by balls on which the oscillation exceeds t and which are maximal for this property. The stopping rule follows [109, Section 4.4]. The difference with the classical argument is the way in which the size of U_t is measured: by the gauges of the stopping balls, and not by the measure of U_t.

Definition 6.8. Let $t \geq \lambda_0$. A ball B is a *stopping ball at the level* t if $\omega(B) > t$ and if there is a point $x \in B$ such that $\omega(B') \leq t$ for every ball $B' \ni x$ with $r_{B'} \geq 2r_B$. We put

$$\Theta_u(t) = \sup\Big\{ \sum_{B\in\mathcal{B}} \Theta(B) : \ \mathcal{B} \text{ is a finite family of disjoint stopping balls at the level } t \Big\} \quad (t \geq \lambda_0)\,, \tag{6.23}$$

and $\Theta_u(t) = \Theta(X)$ for $0 < t < \lambda_0$.

For $t < \lambda_0$ one has $U_t = X$, and the space X itself plays the role of a stopping ball.

Lemma 6.9. *Let $t \geq \lambda_0$.*

(i) *Every point of U_t belongs to a stopping ball at the level t, and every stopping ball B at the level t satisfies $B \subset U_t$ and $t < \omega(B) \leq C\,t$.*

(ii) *There is $L_0 < \infty$ such that every stopping ball at a level $t \geq L_0\lambda_0$ has radius less than $D/12$.*

(iii) $\Theta_u(t) \leq \dfrac{C}{t^p} \displaystyle\int_{U_t\cap\mathcal{S}} g^p \, d\nu$.

(iv) *There is a countable family (b_i) of disjoint stopping balls at the level t such that*

$$U_t \subset \bigcup_i 5b_i\,, \qquad \sum_i \Theta(b_i) \leq \Theta_u(t)\,. \tag{6.24}$$

Proof. (i) Let $x \in U_t$, and let ρ_x be the supremum of the radii of the balls $B \ni x$ with $\omega(B) > t$. Since $x \in U_t$ and $t \geq \lambda_0$, we have $0 < \rho_x \leq D$. Choose such a ball with $r_B > \rho_x/2$. Every ball $B' \ni x$ with $r_{B'} \geq 2r_B$ has $\omega(B') \leq t$, and so has X. Thus B is a stopping ball. If B is a stopping ball with the point x, then $B \subset U_t$ by (6.19), and the concentric ball $2B$ contains x, so that $\omega(B) \leq C\omega(2B) \leq Ct$ by (6.20).

(ii) If $r_B \geq D/12$, then $\omega(B) \leq C\omega(X) = C\lambda_0$ by (6.20). Taking $L_0 > C$ excludes such a stopping ball.

(iii) By (C4) and (6.1), a stopping ball satisfies $t^p < \omega(B)^p \leq C\,\Theta(B)^{-1} \int_{B\cap\mathcal{S}} g^p \, d\nu$. Sum over a disjoint family, and use $B \subset U_t$.

(iv) Apply the $5r$-covering theorem [83, Theorem 1.2] to the family of all stopping balls at the level t, whose radii are at most D and which cover U_t by (i). The family (b_i) is countable because X is separable, and the bound on the sum follows from (6.23) applied to its finite subfamilies. □

The next lemma compares the size of U_t inside stopping balls of a much lower level λ with the number $\Theta_u(t)$. It is the only place where the scale comparison (C3) is used: it converts the reference measure of $B \cap U_t$, weighted by the gauge of the larger ball B, into gauges of the smaller balls which cover U_t.

Lemma 6.10. *There is $H_0 < \infty$ such that the following holds for $\lambda > \lambda_0/2$ and $t \geq H_0\lambda$. If $\lambda \geq \lambda_0$ and $\mathcal{B}$ is a finite family of disjoint stopping balls at the level λ, then*

$$\sum_{B\in\mathcal{B}} \frac{\Theta(B)}{m(B)}\, m(B \cap U_t) \leq C\,\Theta_u(t)\,. \tag{6.25}$$

If $\lambda_0/2 < \lambda < \lambda_0$, the same inequality holds for $\mathcal{B} = \{X\}$.

Proof. Let (b_i) be the family of Lemma 6.9(iv) at the level t. By Lemma 6.9(ii), $r_{b_i} < D/8$ if H_0 is large enough.

Step 1: the balls b_i are smaller. Let $B \in \mathcal{B}$ be a stopping ball at the level λ, with the point x, and suppose that $5b_i \cap B \neq \varnothing$ and $r_{b_i} \geq r_B$. Then $B \subset 8b_i$, and $8b_i$ is a ball of radius at least $2r_B$ which contains x. Hence $\omega(8b_i) \leq \lambda$. On the other hand $\omega(8b_i) \geq c\,\omega(b_i) > c\,t \geq c\,H_0\lambda$ by (6.20). This is impossible if H_0 is large. Thus $r_{b_i} < r_B$ whenever $5b_i$ meets B. For $B = X$ this holds because $r_{b_i} < D$.

Step 2: summation. By Step 1 and (C3), whenever $5b_i$ meets B,

$$\frac{\Theta(B)}{m(B)} = \frac{r_B^2}{\Psi_{C,2}(r_B)} \leq C\,\frac{r_{b_i}^2}{\Psi_{C,2}(r_{b_i})} = C\,\frac{\Theta(b_i)}{m(b_i)}\,.$$

The same relation holds for $B = X$, with $r_X = D$. Since the balls of $\mathcal{B}$ are disjoint and $U_t \subset \bigcup_i 5b_i$, doubling gives

$$\begin{aligned}\sum_{B\in\mathcal{B}} \frac{\Theta(B)}{m(B)}\, m(B\cap U_t) &\leq C \sum_i \frac{\Theta(b_i)}{m(b_i)} \sum_{B\in\mathcal{B}} m(B\cap 5b_i)\\ &\leq C\sum_i \frac{\Theta(b_i)}{m(b_i)}\, m(5b_i) \leq C\sum_i \Theta(b_i) \leq C\,\Theta_u(t)\,.\end{aligned}$$

□

Modification at one level

We now modify u at a level t: first by the Lipschitz extension v_t from the good set, then by the cutoff modification on the stopping balls.

Lemma 6.11. *There are $L_0 < \infty$ and $c \in (0,1)$ such that, for every $t \geq L_0\lambda_0$, there are $w_t \in \mathrm{Lip}(X)$ and a nonnegative Borel function F_t on $\mathcal{S}$ satisfying $\mathrm{Lip}(w_t) \leq Ct$ and*

$$w_t = u \quad \text{on } X \setminus U_{ct}\,, \qquad |\nabla w_t| \leq g\,\mathbf{1}_{\{u^\sharp \leq t\}} + F_t\,, \tag{6.26}$$

with

$$0 \leq F_t \leq Ct\,, \qquad \int_{\mathcal{S}} F_t^b\,d\nu \leq Ct^b\,\Theta_u(t)\,. \tag{6.27}$$

In particular $\int_{\mathcal{S}} F_t^p\,d\nu \leq Ct^p\,\Theta_u(t)$ and $\nabla w_t \in L^p(\nu)$.

Proof. If $U_t = \varnothing$, take $w_t = u$ and $F_t = 0$. Otherwise let v_t be the Lipschitz extension of Lemma 6.7 and let (b_i) be the disjoint stopping balls of (6.24). Choose L_0 so that $11r_{b_i} \leq D$, by Lemma 6.9(ii). Apply Lemma 6.5 to v_t, to the balls $5b_i$ and to their cutoff functions from (C2), with $\Theta_i = \Theta(5b_i)$. The resulting functions w_t and F_t satisfy (6.27), because $\Theta(5b_i) \leq C\,\Theta(b_i)$ and $\sum_i \Theta(b_i) \leq \Theta_u(t)$. Since $U_t \subset \bigcup_i 5b_i$ and $\nabla v_t = \nabla u$ ν-almost everywhere outside U_t by locality, (6.15) gives the gradient bound in (6.26).

If $z \in 10b_i$, then $z \in 11b_i$ and $\omega(11b_i) \geq c\,\omega(b_i) > ct$ by (6.20). Thus $\bigcup_i 10b_i \subset U_{ct}$. Outside U_{ct} both the cutoff modification and the Lipschitz extension agree with u, which proves the equality in (6.26). The last assertion follows from $F_t^p \leq (Ct)^{p-b}F_t^b$ and from $g \in L^p(\nu)$. □

The oscillation on the bad part of a ball

The next lemma bounds the oscillation of u on the part E of a ball which lies in U_t by the oscillation on the complementary part G, up to a term which is small when the level t is much larger than the oscillation on the whole ball. It adapts the argument with balanced balls of [109, Lemma 4.15], whose predecessor is [101, Lemma 3.1.2]. Its estimates use only the reference measure m and the oscillation of u.

Lemma 6.12. *Let $C_0 < \infty$. There are $\alpha \in (0,p)$ and $C < \infty$, depending also on C_0, with the following property. Let $\lambda > 0$, let $k \geq 1$ be an integer with $t = 2^k\lambda \geq \lambda_0$, and let B be a ball, or X, with $\omega(B) \leq C_0\lambda$. Put $E = B \cap U_t$ and $G = B \setminus U_t$. If $m(E) < m(B)/2$, then*

$$\frac{1}{r_B^p} \int_E |u - u_G|^p \, dm \leq C\, 2^{-k\alpha}\, t^p\, m(E) + \frac{C}{r_B^p} \int_G |u - u_G|^p \, dm\,. \tag{6.28}$$

One may take $\alpha = p^2/(p + d_D)$, where d_D is a doubling exponent of m.

Proof. Write $R = r_B$.

Step 1: balanced balls. Every sphere has zero m-measure, by geodesicity and doubling; see [109, Lemma 2.5]. Indeed, if y is on a sphere of centre x and radius s, moving a distance $r/2$ from y towards x, for $0 < r < s$, locates a ball of radius $r/4$ inside $B(y,r)$ and disjoint from the sphere. Its measure is a fixed fraction of $m(B(y,r))$ by doubling, so no point of the sphere is a density point of the sphere, and the Lebesgue differentiation theorem gives the assertion. For $B = B(o,R)$, moving from $x \in B$ towards o by a distance $\min\{r/2, d(x,o)\}$ locates a ball of radius $r/2$ in $B(x,r) \cap B$ when $r \leq 2R$. Consequently

$$m(B(x,r)) \leq C\, m(B(x,r) \cap B) \qquad (x \in B\,,\ 0 < r \leq 2R)\,, \tag{6.29}$$

which also holds for $B = X$ with $R = D$. For $x \in E$ the ratio $m(B(x,r) \cap E)/m(B(x,r) \cap B)$ is a continuous function of r. It is equal to one at small radii, since E is open, and it is smaller than $1/2$ at $r = \min\{2R, D\}$, where $B(x,r)$ contains B up to a null set. Thus there is a ball $B'_x = B(x, r'_x)$ with $0 < r'_x \leq \min\{2R, D\}$ and

$$m(B'_x \cap E) = m(B'_x \cap G)\,.$$

By the $5r$-covering theorem there are disjoint balls $B'_j = B(x_j, r'_j)$ of this family with $E \subset \bigcup_j 5B'_j$. Put

$$Q_j = 5B'_j \cap B\,, \qquad O_j = B'_j \cap G\,, \qquad \rho_j = r'_j/R\,.$$

By (6.29) and the doubling property, $m(Q_j) \asymp m(O_j) \asymp m(B'_j) \asymp m(B'_j \cap E)$.

Step 2: two bounds on Q_j. The set O_j contains a point at which $u^\sharp \leq t$, so $\omega(5B'_j) \leq t$; if $5r'_j > D$, use X with $r_X = D$ and $m(X) \leq C\, m(B'_j)$. Comparing the means on Q_j and O_j with the mean on $5B'_j$, or on X, gives

$$\fint_{Q_j} |u - u_{O_j}|^p \, dm \leq C\, t^p\, (r'_j)^p = C\, t^p\, R^p\, \rho_j^p\,.$$

On the other hand $Q_j \cup O_j \subset B$ and $\omega(B) \leq C_0\lambda = C_0 2^{-k} t$, so that, by doubling,

$$\fint_{Q_j} |u - u_{O_j}|^p \, dm \leq C\, \frac{m(B)}{m(B'_j)}\, \lambda^p\, R^p \leq C\, 2^{-kp}\, t^p\, R^p\, \rho_j^{-d_D}\,.$$

The first bound is good for small balls and the second for large ones. Since $\min\{\rho^p, 2^{-kp}\rho^{-d_D}\} \leq 2^{-kp^2/(p+d_D)}$ for every $\rho > 0$, we obtain, with $\alpha = p^2/(p+d_D) < p$,

$$\int_{Q_j} |u - u_{O_j}|^p \, dm \leq C\, 2^{-k\alpha}\, t^p\, R^p\, m(B'_j \cap E)\,.$$

Step 3: summation. On Q_j we write $u - u_G = (u - u_{O_j}) + (u_{O_j} - u_G)$. By Jensen's inequality, $m(Q_j)|u_{O_j} - u_G|^p \leq C \int_{O_j} |u - u_G|^p \, dm$. Since $E \subset \bigcup_j Q_j$, the balls B'_j are disjoint and $O_j \subset G$,

$$\int_E |u - u_G|^p \, dm \leq C\, 2^{-k\alpha}\, t^p\, R^p\, m(E) + C \int_G |u - u_G|^p \, dm\,. \qquad \square$$

Averaging over the levels

A single modification w_t has a gradient of size t on the bad set, and its energy there is of the same order $t^p \Theta_u(t)$ as the weak-type bound: nothing is gained. Following [101, proof of Proposition 3.1.1] and [109, proof of Lemma 4.23], one averages the modifications at k consecutive dyadic levels.

Let $A_1 < \infty$ be such that $cA_1 \geq 1$ and $A_1 \geq 2L_0$, where c and L_0 are the constants of Lemma 6.11. For $\lambda > \lambda_0/2$ and an integer $k \geq 1$ with $2^k \geq H_0$, where H_0 is the constant of Lemma 6.10, we put

$$t_i = A_1\, 2^i\, \lambda \quad (k \leq i \leq 2k-1)\,, \qquad \bar{w} = \frac{1}{k} \sum_{i=k}^{2k-1} w_{t_i}\,. \tag{6.30}$$

Since $U_{ct_i} \subset U_{2^k\lambda}$ and $t_i \leq A_1 4^k \lambda$, Lemma 6.11 gives

$$\bar{w} = u \quad \text{on } X \setminus U_{2^k\lambda}\,, \qquad |\nabla \bar{w}| \leq g\, \mathbf{1}_{\{u^\sharp \leq A_1 4^k \lambda\}} + \frac{1}{k} \sum_{i=k}^{2k-1} F_{t_i}\,. \tag{6.31}$$

The gain comes from the fact that the functions F_{t_i} have geometrically increasing sizes: the p-th power of their sum is comparable to the sum of their p-th powers, so that the factor k^{-1} in the average becomes a factor k^{-p} in front of a sum of only k terms.

Lemma 6.13. *With the notation (6.30) and the functions F_t of Lemma 6.11,*

$$\int_{\mathcal{S}} \Big(\frac{1}{k} \sum_{i=k}^{2k-1} F_{t_i} \Big)^p \, d\nu \leq \frac{C}{k^p} \sum_{i=k}^{2k-1} t_i^p\, \Theta_u(t_i)\,.$$

Proof. We first prove a numerical inequality. Let β_i be numbers with $0 \leq \beta_i \leq C' t_i$, and put $S = \sum_i \beta_i$. We claim that

$$S^p \leq C \sum_i t_i^{p-b} \beta_i^b\,. \tag{6.32}$$

Let $S > 0$. Since the numbers t_i are geometrically spaced, the indices with $t_i < S/(4C')$ contribute at most $S/2$ to the sum. On the remaining indices, if $b > 1$, Hölder's inequality gives

$$\frac{S}{2} \leq \Big(\sum_i t_i^{p-b} \beta_i^b \Big)^{1/b} \Big(\sum_{t_i \geq S/(4C')} t_i^{-(p-b)/(b-1)} \Big)^{1-1/b} \leq C \Big(\sum_i t_i^{p-b} \beta_i^b \Big)^{1/b} S^{-(p-b)/b}\,,$$

which is (6.32). For $b = 1$, use $t_i^{p-1} \geq (S/(4C'))^{p-1}$ on the remaining indices. Apply (6.32) to $\beta_i = F_{t_i}$ at each point, integrate using (6.27), and divide by k^p. □

The following recurrence corresponds to [109, equation (4.24)], with the gauge of the stopping balls in place of the measure of the level sets.

Lemma 6.14. *Let $\lambda > \lambda_0/2$ and let $k \geq 1$ be an integer with $2^k \geq H_0$. Then, with the constant α of Lemma 6.12 for a suitable C_0,*

$$\lambda^p\, \Theta_u(\lambda) \leq C\, 2^{-k\alpha}\, (2^k\lambda)^p\, \Theta_u(2^k\lambda) + \frac{C}{k^p} \sum_{i=k}^{2k-1} t_i^p\, \Theta_u(t_i) + C \int_{\{\lambda < u^\sharp \leq A_1 4^k \lambda\}} g^p \, d\nu\,. \tag{6.33}$$

Proof. If $\lambda \geq \lambda_0$, let $\mathcal{B}$ be a finite family of disjoint stopping balls at the level λ. If $\lambda_0/2 < \lambda < \lambda_0$, let $\mathcal{B} = \{X\}$. In both cases every $B \in \mathcal{B}$ satisfies $\lambda < \omega(B) \leq C_0\lambda$ and $B \subset U_\lambda$, by Lemma 6.9(i). Let $B \in \mathcal{B}$, $t = 2^k\lambda$, $E = B \cap U_t$ and $G = B \setminus U_t$.

Case 1: $m(E) \geq m(B)/2$. Then, since $\alpha < p$,

$$\lambda^p\,\Theta(B) \leq 2\,\lambda^p\,\frac{\Theta(B)}{m(B)}\,m(E) \leq 2\cdot 2^{-k\alpha}\,t^p\,\frac{\Theta(B)}{m(B)}\,m(B\cap U_t)\,.$$

Case 2: $m(E) < m(B)/2$. Since $\omega(B) > \lambda$,

$$\lambda^p\,r_B^p\,m(B) < \int_B |u-u_B|^p\,dm \leq 2^p\int_B |u-u_G|^p\,dm = 2^p\int_E |u-u_G|^p\,dm + 2^p\int_G |u-u_G|^p\,dm\,.$$

By (6.31), $\bar{w} = u$ on G. Applying (C4) to $\bar{w}$ gives

$$\int_G |u-u_G|^p\,dm = \int_G |\bar{w}-\bar{w}_G|^p\,dm \leq 2^p\int_B |\bar{w}-\bar{w}_B|^p\,dm \leq C\,\frac{r_B^p\,m(B)}{\Theta(B)}\int_{B\cap\mathcal{S}} |\nabla\bar{w}|^p\,d\nu\,.$$

With Lemma 6.12 this gives

$$\lambda^p\,\Theta(B) \leq C\,2^{-k\alpha}\,t^p\,\frac{\Theta(B)}{m(B)}\,m(B\cap U_t) + C\int_{B\cap\mathcal{S}} |\nabla\bar{w}|^p\,d\nu\,.$$

Summation. We sum over $B \in \mathcal{B}$. The first terms are bounded by Lemma 6.10. The balls are disjoint and contained in U_λ, so by (6.31) and Lemma 6.13,

$$\sum_{B\in\mathcal{B}}\int_{B\cap\mathcal{S}} |\nabla\bar{w}|^p\,d\nu \leq C\int_{\{\lambda<u^\sharp\leq A_1 4^k\lambda\}} g^p\,d\nu + \frac{C}{k^p}\sum_{i=k}^{2k-1} t_i^p\,\Theta_u(t_i)\,.$$

Taking the supremum over the families $\mathcal{B}$ gives (6.33). For $\lambda < \lambda_0$, the left-hand side is $\lambda^p\,\Theta(X)$. □

Absorption

Proof of Proposition 6.6. We use the absorption argument of [109, Section 4.6], applied here to a supremum over the levels. Let $q \in [p-\varepsilon_0, p)$, where ε_0 will be chosen below, and let $u \in \mathrm{Lip}(X)$ with $g = |\nabla u| \in L^q(\nu)$. Since g is bounded, $g \in L^p(\nu)$. We may assume that $\lambda_0 > 0$. Put

$$J = \sup_{\lambda>\lambda_0/2} \lambda^q\,\Theta_u(\lambda)\,.$$

For $\lambda \geq \lambda_0$, Lemma 6.9(iii) gives

$$\lambda^q\,\Theta_u(\lambda) \leq C\lambda_0^{q-p}\int_{\mathcal{S}} g^p\,d\nu\,.$$

Since $\Theta_u(\lambda) = \Theta(X)$ when $\lambda_0/2 < \lambda < \lambda_0$, it follows that

$$\Theta(X)\,\lambda_0^q \leq J < \infty\,. \tag{6.34}$$

Multiply (6.33) by λ^{q-p}. Each of the levels $2^k\lambda$ and t_i belongs to the range of the supremum defining J, so the first two terms are bounded by

$$\Big(C\,2^{k(p-q-\alpha)} + C\,k^{1-p}A_1^{p-q}4^{k(p-q)}\Big)J\,.$$

On the set of integration in the last term, (6.22) gives $g \le Cu^\sharp \le CA_1 4^k \lambda$, and hence

$$\lambda^{q-p} g^p \le (CA_1 4^k)^{p-q} g^q \,.$$

Taking the supremum over $\lambda > \lambda_0/2$, we obtain

$$J \le \Bigl(C\, 2^{k(p-q-\alpha)} + C\, k^{1-p} A_1^{p-q} 4^{k(p-q)} \Bigr) J + C_k \int_S g^q \, d\nu \,. \tag{6.35}$$

Since $p > 1$ and $\alpha > 0$, first choose k so that $2^k \ge H_0$ and $C2^{-k\alpha} + Ck^{1-p} \le 1/4$. Then choose $\varepsilon_0 \in (0, p-b)$ so small that the coefficient of J in (6.35) is at most $1/2$ whenever $0 < p - q \le \varepsilon_0$. The constant C_k is bounded uniformly over this interval. Since $J < \infty$, absorption and (6.34) give

$$\Theta(X)\, \lambda_0^q \le C \int_S g^q \, d\nu \,.$$

By the definition (6.18) of λ_0, this is (6.17). □

6.5 Localization by uniform domains

We return to the setting of Theorem 6.1. To apply Proposition 6.6 to a ball, we need a compact geodesic space which contains the ball, is contained in a fixed dilate of it, and inherits the hypotheses (C1)–(C4). A ball with its induced metric is not geodesic in general, and the Poincaré inequality on a ball need not hold up to its boundary. Uniform domains provide the substitute.

A domain $\Omega \subset X$ is *A_2-uniform* if any two points $x, y \in \Omega$ are joined by a curve $\gamma : [0, \ell] \to \Omega$, parametrized by arc length, such that

$$\ell \le A_2\, d(x,y) \,, \qquad \operatorname{dist}\bigl(\gamma(t), X \setminus \Omega\bigr) \ge A_2^{-1} \min\{t, \ell - t\} \quad (0 \le t \le \ell) \,. \tag{6.36}$$

Rajala [136, Theorem 1.1] proved that a bounded domain of a doubling quasiconvex metric space can be approximated from inside and from outside by uniform domains. The use of such domains to replace global hypotheses by local ones is taken from [37, Section 3, especially Theorem 3.2]. We need a quantitative form for balls.

Lemma 6.15. *There are structural constants $A_2 < \infty$ and $c_2 > 0$ such that, for every ball $B(o, R)$ with $R < c_2 r_0$, there is an A_2-uniform domain Ω with*

$$B(o,R) \subset \Omega \subset B(o, 2R) \,. \tag{6.37}$$

Proof. Rescale to $R = 1$, with r_0/R larger than a fixed constant, and set $E_1 = B(o,1)$. We use Rajala's covering construction [136, Lemma 2.1 and proof of Theorem 1.1]: the domain is obtained from the ball by adding, at the scales δ^k, the balls of a covering which meet a neighbourhood of the previous stage.

Step 1: coverings with separated radii. At each scale $s \le 1$, a maximal s-separated net has at most N points in any ball of radius $5s$, by the doubling property VD. Enumerate the net and assign successively a radius $r_i \in [s, 2s]$ to each point x_i, in such a way that, with $c = 1/(2N)$,

$$d(x_i, x_j) - r_i - r_j \notin (0, cs) \qquad (i \ne j) \,.$$

This is possible, because the balls to which a radius has already been assigned exclude a union of at most N intervals of length cs from $[s, 2s]$. The maximality of the net shows that the balls

$B(x_i, r_i)$ cover X. By construction, two balls of the covering are either at distance at least cs, or their centres are at distance at most the sum of their radii.

Step 2: the domain. Take $\delta = \min\{c/(20+c), 1/11\}$ and let E_{k+1} be the union of the balls of the covering of Step 1 at the scale $s = \delta^k$ which meet $B(E_k, \delta^k) = \{x : \mathrm{dist}(x, E_k) < \delta^k\}$. The construction gives

$$B(E_k, \delta^k) \subset E_{k+1} \subset B(E_k, 5\delta^k)\,,$$

and the domain $\Omega = \bigcup_k E_k$ satisfies

$$B(o,1) \subset \Omega \subset B\Big(o, 1 + \frac{5\delta}{1-\delta}\Big) \subset B(o, 3/2)\,, \qquad \mathrm{dist}(E_1, X \setminus \Omega) \geq \delta\,.$$

This gives (6.37).

Step 3: curves between the stages. A point of E_j, $j \geq 2$, can be joined to E_{j-1} by a curve of length at most $5\delta^{j-1}$ in E_j: from the point, go to the centre of a covering ball containing it, then to a point of that ball within δ^{j-1} of E_{j-1}, and finally to E_{j-1} by a geodesic of length less than δ^{j-1}. The last segment lies in $B(E_{j-1}, \delta^{j-1})$. Along this curve the distance to $X \setminus \Omega$ is at least δ^j. Concatenating these curves from the stage j down to the stage 1, the length from the initial point to any point of the part in E_i, $i \leq j$, is at most $5\delta^{i-1}/(1-\delta)$.

Step 4: uniformity. Let $x, y \in \Omega$. Suppose first that $d(x,y) \geq c\delta^2/4$. Follow the curves of Step 3 from x and from y to E_1, and join their endpoints by radial geodesics through o. The middle part has length at most 2 and distance at least δ from $X \setminus \Omega$. The total length is bounded, so (6.36) holds with a structural constant.

Suppose now that $0 < d(x,y) < c\delta^2/4$, and choose $n \geq 2$ with $c\delta^{n+1}/4 \leq d(x,y) < c\delta^n/4$. If one of the points lies in E_{n-1}, a geodesic joining the points has distance at least $\delta^{n-1} - d(x,y)$ from $X \setminus \Omega$, and it satisfies (6.36). Otherwise follow the curves of Step 3 to points $x', y' \in E_n$. The two curves each have length at most $5\delta^n/(1-\delta) \leq c\delta^{n-1}/4$, and hence $d(x', y') < c\delta^{n-1}$. By the separation property of Step 1, the covering balls at the scale δ^{n-1} which contain x' and y' have centres at distance at most the sum of their radii. Join x' to y' through those centres by geodesics. This path lies in the union of the two closed balls, has length at most $8\delta^{n-1}$ and stays at distance at least δ^n from $X \setminus \Omega$. The length of the concatenated curve from x to y is at most $9\delta^{n-1} \leq 36\, d(x,y)/(c\delta^2)$. The distance estimates on the middle part and on the curves of Step 3 give the second inequality of (6.36). □

The next lemma transfers the hypotheses to the completion of the uniform domain. The part of the argument which concerns the reference measure parallels [37, Theorem 3.2]. The restriction of the gradient and the cutoff estimates are verified separately. Note that the construction uses the restricted measure $\nu\lceil_\Omega$, whether or not $\nu(\partial\Omega) > 0$, and that it does not require $\nu(\Omega)$ to be finite.

Lemma 6.16. *Assume the hypotheses of Theorem 6.1 and let Ω be the domain of Lemma 6.15 for a ball $B(o,R)$ with $R < c_3 r_0$, where $c_3 > 0$ is a small structural constant, and let (Y, d_Y) be the completion of Ω for its inner length metric d_Ω. Let*

$$m_Y = m\lceil_\Omega\,, \qquad \nu_Y = \nu\lceil_\Omega\,, \qquad \mathcal{S}_Y = \Omega \cap \mathcal{S}\,, \qquad \nabla_Y v = (\nabla \widetilde{v})|_{\Omega \cap \mathcal{S}} \quad (v \in \mathrm{Lip}(Y))\,,$$

where m_Y and ν_Y give no mass to $Y \setminus \Omega$, the field over $\mathcal{S}_Y$ is the restriction of H, and $\widetilde{v} \in \mathrm{Lip}_b(X)$ is an extension of v. Then (Y, d_Y) is a compact geodesic space with $\operatorname{diam} Y \asymp R$, the map ∇_Y is well defined and satisfies (G), and $(Y, d_Y, m_Y, \nu_Y, \nabla_Y)$ satisfies (C1)–(C4) with the restricted scale $\Psi_{C,2}$, the gauge $\Theta_Y(B_Y(z,r)) = r^2 m_Y(B_Y(z,r))/\Psi_{C,2}(r)$, and a structural constant C_*.

Proof. Step 1: the space. By (6.36), $d \le d_\Omega \le A_2\, d$ on Ω. Hence Y is identified with the closure of Ω in X, which is compact because X is proper, and $d \le d_Y \le A_2\, d$ on Y. The completion of a length space is a length space, and a compact length space is geodesic [43, Part I, Chapter 3]. Since $R < r_0 \le \operatorname{diam} X$ and X is geodesic, $B(o,R)$ contains a geodesic of length $R/2$ issued from o. Together with $\Omega \subset B(o,2R)$, this gives $R/2 \le \operatorname{diam} Y \le 4A_2 R$.

Step 2: the gradient. A function $v \in \operatorname{Lip}(Y)$ is Lipschitz for d, and it has an extension $\widetilde v \in \operatorname{Lip}_b(X)$. By locality, $\nabla \widetilde v$ on Ω does not depend on the extension, so ∇_Y is well defined, linear, and satisfies the Leibniz rule. If β is a ball with $3\beta \subset \Omega$, a geodesic of X between two points of β stays in 3β. Thus $d_\Omega = d$ on β. Extend $v|_\beta$ to X with the constant $\operatorname{Lip}_{d_Y}(v)$, and use locality on a countable cover of Ω by such balls to obtain $|\nabla_Y v| \le \operatorname{Lip}_{d_Y}(v)$ ν_Y-almost everywhere. Let $v_k \to v$ pointwise on Y with uniformly bounded supremum norms and Lipschitz constants. The convergence is uniform, because Y is compact, and the McShane extensions with a common Lipschitz constant, truncated at a common level, converge uniformly on X. Testing the weak-$*$ convergence on X against fields of $L^1(\mathcal{S}, \mathsf{H}, \nu)$ which vanish outside Ω gives the required weak-$*$ convergence on Y.

Step 3: interior balls. We claim that there is a structural $c > 0$ such that for every $z \in Y$ and $0 < r \le \operatorname{diam} Y$ there is $z_r \in \Omega$ with

$$d_Y(z, z_r) \le r/2\,, \qquad B(z_r, c\,r) \subset \Omega \cap B_Y(z,r)\,, \qquad \operatorname{dist}(z_r, X \setminus \Omega) \ge c\,r\,. \tag{6.38}$$

For $z \in \Omega$, choose $y \in \Omega$ with $d_Y(z,y) \ge r/3$, which exists because $r \le \operatorname{diam} Y$, and let γ be a curve from z to y as in (6.36). Its length is at least $r/3$. The point $z_r = \gamma(r/6)$ satisfies $\operatorname{dist}(z_r, X \setminus \Omega) \ge r/(6A_2)$ and $d_Y(z, z_r) \le r/6$, and $d_Y \le A_2 d$ gives the inclusion for $c = 1/(12A_2^2)$. For $z \in Y \setminus \Omega$, apply this to a point of Ω close to z.

Step 4: volume, cutoffs and scales. By (6.38) and the doubling property of m,

$$c\, m(B(z,r)) \le m_Y(B_Y(z,r)) \le m(B(z,r)) \qquad (z \in Y,\ 0 < r \le \operatorname{diam} Y)\,. \tag{6.39}$$

Indeed, the inner ball in (6.38) has radius cr, and its centre is within distance $r/2$ of z for the metric d, so its measure is comparable from below to $m(B(z,r))$. This also proves that m_Y has full support. It implies the doubling property of m_Y when $2r \le \operatorname{diam} Y$. If $r > \operatorname{diam} Y/2$, use $Y \subset B(z, 2\operatorname{diam} Y)$ and the doubling property of m to obtain the same conclusion. This is (C1). Moreover $m_Y(Y) \asymp m(B(o,R))$ and $\operatorname{diam} Y \asymp R$.

For the cutoffs, set $\delta_1 = 1/(8A_2)$ and choose a maximal family $z_1, \dots, z_N$ in $\Omega \cap B_Y(z,r)$ which is $\delta_1 r$-separated for d_Y. The inherited doubling gives $N \le C$. The restriction to Y of

$$\psi = \max_{i \le N} \psi_{z_i}^{2\delta_1 r}$$

is equal to one on $B_Y(z,r)$, vanishes outside $B_Y(z,2r)$, and has Lipschitz constant at most C/r for d_Y. By (1.4), (6.2) and (3.2),

$$\int |\nabla_Y \psi|^b \, d\nu_Y \le C \sum_{i \le N} \frac{(2\delta_1 r)^{2-b}\, m(B(z_i, 2\delta_1 r))}{\Psi_{C,2}(2\delta_1 r)} \le C\, \frac{\Theta_Y(B_Y(z,r))}{r^b}\,.$$

This proves (C2). Condition (C3) is (6.7), which does not involve the measure or the centre.

Step 5: a Poincaré inequality on Y. Following the chain method for John domains [79, Theorems 9.3 and 9.10], we prove that for $v \in \operatorname{Lip}(Y)$ with $\nabla_Y v \in L^p(\nu_Y)$ and $D' = B_Y(z,r)$, $0 < r \le \operatorname{diam} Y$,

$$\int_{D'} |v - v_{D'}|^p \, dm_Y \le C\, \Psi_{C,p}(r) \int_{B_Y(z, C_2 r)} |\nabla_Y v|^p \, d\nu_Y\,. \tag{6.40}$$

Let z_r be as in (6.38). For $y \in D' \cap \Omega$ let $\gamma_y : [0, \ell_y] \to \Omega$ be a curve from y to z_r as in (6.36). Then $\ell_y \le C_1 r$ and

$$\operatorname{dist}(\gamma_y(t), X \setminus \Omega) \ge c_1\, t \qquad (0 \le t \le \ell_y)\,. \tag{6.41}$$

Indeed, if $\ell_y - t \ge cr/2$, then $\min\{t, \ell_y - t\} \ge \min\{1, c/(2C_1)\}\, t$ and (6.36) applies. If $\ell_y - t < cr/2$, then $d(\gamma_y(t), z_r) < cr/2$, and (6.38) gives a distance at least $cr/2 \ge c\, t/(2C_1)$. Let $t_j = 2^{-j} C_1 r$, and let $D_j(y) = B(\gamma_y(\min\{t_j, \ell_y\}), a\, t_j)$ for a small structural $a > 0$. As in the proof of Lemma 6.4, $D_0(y) = B(z_r, aC_1 r)$ does not depend on y, consecutive balls are joined by boundedly many balls of radius at_{j+1} centred on γ_y, and by (6.41) and (6.38) the closed Λ-dilates of all these balls are compact subsets of Ω, contained in $B_Y(z, C_2 r) \cap B(y, C_3 t_j)$.

The inequality (6.5) of X can be applied to v on each of these balls β, although an extension $\widetilde{v}$ may have infinite energy outside Ω. Indeed, by compactness there is a finite maximum ζ of cutoff functions of (6.2), with support in Ω, equal to one on the closure of $\Lambda\beta$. Then $\zeta\widetilde{v} \in \mathcal{C}_p$: in $\nabla(\zeta\widetilde{v}) = \zeta\nabla\widetilde{v} + \widetilde{v}\,\nabla\zeta$ the first term vanishes outside Ω and is p-integrable on Ω, and the second is p-integrable by (6.2) and the Lipschitz bound. Moreover $\nabla(\zeta\widetilde{v}) = \nabla_Y v$ on $\Lambda\beta$.

The telescoping argument gives, for $y \in D' \cap \Omega$,

$$|v(y) - v_{D_0}| \le C \sum_{j \ge 0} \left(\frac{\Psi_{C,p}(t_j)}{m(B(y,t_j))}\, \eta(B(y, C_3 t_j)) \right)^{1/p}, \qquad \eta = \mathbf{1}_{B_Y(z,C_2 r)} |\nabla_Y v|^p\, \nu_Y\,.$$

The denominator is the volume for m, since the balls of the chain lie in Ω. Using $m_Y \le m$, Fubini's theorem and the doubling property of m give

$$\int_{D'} \frac{\eta(B(y, C_3 t))}{m(B(y,t))}\, dm_Y(y) \le C \int \frac{m(B(w, C_3 t))}{m(B(w,t))}\, d\eta(w) \le C\, \eta(Y)\,.$$

Minkowski's inequality and (6.8) imply

$$\|v - v_{D_0}\|_{L^p(D', m_Y)} \le C\, \Psi_{C,p}(r)^{1/p}\, \eta(Y)^{1/p}\,.$$

This proves (6.40).

Step 6: removal of the dilation. Apply the proof of Lemma 6.4, with $e = s = p$, in the compact geodesic space Y, using the inherited doubling, (6.40), and the scale comparison (6.8). The Fubini estimate now has $m_Y(B_Y(y,t))$ in its denominator. The comparison of the volumes of neighbouring balls follows from the inherited doubling, including radii comparable to $\operatorname{diam} Y$. This gives (C4) on every ball. For $B = Y$, join all points to a fixed centre and use the same chains with $R = \operatorname{diam} Y$; there is no restriction on their dilates. □

Proof of Theorem 6.1. For a sufficiently small ball $B = B(o, R)$, let Ω and Y be as in Lemma 6.16. With $D = \operatorname{diam} Y$, we have $D \asymp R$ and $\Theta_Y(Y) \asymp \Theta(B)$, by doubling, (3.2) and $B \subset \Omega \subset 2B$. Let $q \in [p - \varepsilon_0, p)$ and $u \in \mathcal{C}_q$, where ε_0 is given by Proposition 6.6 for the constant C_* of Lemma 6.16. We apply Proposition 6.6 to the restriction of u to Y. The norms on Y are computed with $m\restriction_\Omega$ and $\nu\restriction_\Omega$, so

$$\Big(\fint_B |u - u_B|^p\, dm \Big)^{1/p} \le C \Big(\fint_\Omega |u - u_\Omega|^p\, dm \Big)^{1/p} \le C\, R \left(\frac{1}{\Theta(B)} \int_{2B \cap \mathcal{S}} |\nabla u|^q\, d\nu \right)^{1/q}.$$

For $q = p$ the same inequality holds by Lemma 6.4 with $e = s = p$. Bounding the L^1 oscillation on the left-hand side by the L^p oscillation, we obtain the hypothesis of Lemma 6.4 at the exponent $e = q$, for every $q \in [p - \varepsilon_0, p]$. Choose $s > p$ and decrease ε_0 so that $D_m(1/(p - \varepsilon_0) - 1/s) < 1$. Then (6.11) gives (6.6) for every such q. The chain estimate is uniform in q on this interval, by (6.8) and the strict inequality (6.10). □

Part II
Dirichlet spaces and heat kernels

7 Dirichlet spaces and Harnack inequalities

This part of the book studies the heat equation. Its first two chapters are independent of Part I: they concern a strongly local regular Dirichlet form on a metric measure space, without assuming that it comes from a gradient structure. Chapter 9 returns to gradient structures.

Background. On a complete Riemannian manifold, Grigor'yan [77] and Saloff-Coste [140, Theorem 3.1] proved that the parabolic Harnack inequality is equivalent to the conjunction of volume doubling and the Poincaré inequality at the scale r^2, and to two-sided Gaussian bounds for the heat kernel; Sturm [151] extended this to strongly local Dirichlet spaces. On fractals the natural space-time scale is $\Psi(r) = r^\beta$ with $\beta > 2$, and volume doubling and the Poincaré inequality $\mathrm{PI}(\Psi)$ no longer suffice. Barlow, Bass and Kumagai [17] showed that the parabolic Harnack inequality at the scale Ψ is equivalent to volume doubling, $\mathrm{PI}(\Psi)$ and a *cutoff Sobolev inequality* at the same scale, which asserts the existence of cutoff functions with controlled energy measure. Simpler forms of this third condition were given by Andres and Barlow [10] and by Grigor'yan, Hu and Lau [75], and Lierl [121] carried out Moser's iteration [127, 128] in this setting.

Two scales. In all these results one scale function governs both the Poincaré inequality and the cutoff Sobolev inequality. For the Dirichlet form of a gradient structure this need not be the case: the Poincaré scale and the cutoff scale may differ (Proposition 15.10). In this part the Poincaré inequality holds at a scale Ψ_P and the cutoff Sobolev inequality at a scale Ψ_S, with $\Psi_S \lesssim \Psi_P$ (Proposition 7.4), and the dependence of the estimates on the two scales is made explicit. The main result of this chapter, Theorem 7.5, is a parabolic Harnack inequality on a cylinder of radius r and time length T, whose constant is

$$\exp\Big[C\Big(1 + \frac{\Psi_P(r)}{T} + \frac{T}{\Psi_S(r)}\Big)\Big]\,.$$

When the two scales are comparable and $T \asymp \Psi(r)$, this is the usual parabolic Harnack inequality. The proof follows Moser's method in the form of [121], keeping track of the two coefficients $\Psi_P(r)/T$ and $T/\Psi_S(r)$ at each step.

7.1 Dirichlet spaces

The form. Throughout Part II, (X, d) is a proper metric space, m is a Radon measure with full support, and $(\mathcal{E}, \mathcal{F})$ is a symmetric, strongly local, regular Dirichlet form on $L^2(X, m)$, with generator L and semigroup $P_t = e^{tL}$; we refer to [69] for the theory of Dirichlet forms. The energy measure of $f \in \mathcal{F}$ is denoted by $\Gamma(f)$, and $\Gamma(f, g)$ is its polarization, with the normalization

$$\mathcal{E}(f, f) = \int_X d\Gamma(f)\,.$$

For the form of a 2-admissible gradient structure, $d\Gamma(f) = |\nabla_2 f|^2\, d\nu$ by Proposition 2.7. We use the following notation and conventions.

- For an open set U, $\mathcal{F}_{\mathrm{loc}}(U)$ is the space of the functions which coincide on every relatively compact open subset of U with a function of $\mathcal{F}$. The energy measure of a function of $\mathcal{F}_{\mathrm{loc}}(U)$ is defined on U by locality.
- $\mathcal{F}_0(U)$ is the closure in $\mathcal{F}$ of the functions of $\mathcal{F}$ with compact support in U, that is, the domain of the part of the form on U.
- In integrals with respect to energy measures, functions of $\mathcal{F}$ are represented by their quasi-continuous versions, and conditions on functions of $\mathcal{F}$ are understood quasi-everywhere.
- We use the Leibniz and chain rules for energy measures [69, Section 3.2].

Scale functions. A positive function S on a range of radii has *two-sided weak scaling* if there are $0 < s_- \le s_+ < \infty$ such that

$$c_S\Big(\frac{R}{r}\Big)^{s_-} \le \frac{S(R)}{S(r)} \le C_S\Big(\frac{R}{r}\Big)^{s_+} \qquad (0 < r \le R) \tag{7.1}$$

on this range. This is condition (3.1). The Poincaré scale Ψ_P and the cutoff Sobolev scale Ψ_S below satisfy (7.1), and they are taken continuous and strictly increasing whenever their inverses are used. We fix the following exponents:

- $Q_D > 0$ is a doubling exponent of m, so that $m(B(x,R))/m(B(x,r)) \le C(R/r)^{Q_D}$;
- $a_- > 0$ is a lower scaling exponent of Ψ_P;
- $\beta_+ \ge 2$ is an upper scaling exponent of Ψ_S.

Volume doubling and Poincaré inequality. For the balls of admissible radius,

$$m(B(x,2r)) \le C_D\, m(B(x,r))\,, \tag{7.2}$$

$$\int_{B(x,r)} |f - f_{B(x,r)}|^2\, dm \le C\,\Psi_P(r) \int_{B(x,\Lambda r)} d\Gamma(f)\,, \qquad f \in \mathcal{F}_{\mathrm{loc}}(X)\,, \tag{7.3}$$

for a fixed $\Lambda \ge 1$. These are VD and PI(Ψ_P). For the form of a gradient structure, the Poincaré inequality $\mathrm{PI}_2(\Psi_{P,2})$ of Chapter 3, stated on the core, extends to $\mathcal{F}$ by approximation in the form norm, and to $\mathcal{F}_{\mathrm{loc}}(X)$ by locality. The proof of the Harnack inequality uses (7.3) with $\Lambda = 1$. On a geodesic space, (7.2) and (7.3) imply the inequality with $\Lambda = 1$ by a Whitney covering and a chain argument, after reducing the range of radii by a fixed factor; see [121, Remark 3.6], and Lemma 6.4 for the same argument for gradient structures.

Solutions. Let $I \subset \mathbb{R}$ be an open interval and $\Omega \subset X$ an open set. A *local weak subsolution* of the heat equation in $I \times \Omega$ is a function $u \in L^2_{\mathrm{loc}}(I; \mathcal{F}_{\mathrm{loc}}(\Omega))$ which satisfies $\partial_t u \le Lu$ in the variational sense:

$$-\int_I \int_X u\, \partial_t \phi \, dm\, dt + \int_I \mathcal{E}(u(t), \phi(t))\, dt \le 0$$

for every nonnegative $\phi \in L^2(I; \mathcal{F})$ with $\partial_t \phi \in L^2(I; L^2(X,m))$ and with compact support in $I \times \Omega$. It is a *local weak solution*, and we also say that u is *caloric* in $I \times \Omega$, if equality holds for all such ϕ, without sign condition. A function $u \in \mathcal{F}_{\mathrm{loc}}(\Omega)$ is *weakly harmonic* in Ω if $\mathcal{E}(u,\phi) = 0$ for every $\phi \in \mathcal{F}$ with compact support in Ω.

Energy computations for weak solutions are performed with Steklov averages in time and truncated test functions, and then passed to the limit. This justification is carried out in [121, Lemmas 4.4 and 5.1], and we do not repeat it. The computations below specify the coefficients and the boundary terms in time which are needed here.

7.2 Capacity and cutoff Sobolev inequalities

In Part I cutoff functions were Lipschitz functions with an integral bound on their gradient. For a Dirichlet form the corresponding notions are formulated in the domain of the form.

Cutoff functions and capacity. For $R < R'$, a *cutoff function* for $B(x,R) \subset B(x,R')$ is a function $\chi \in \mathcal{F}$, with $0 \le \chi \le 1$, equal to one on a neighbourhood of $\overline{B(x,R)}$ and with compact support in $B(x,R')$. The *capacity* of the condenser is

$$\operatorname{Cap}(B(x,R), B(x,R')) = \inf_{\chi} \mathcal{E}(\chi,\chi)\,, \tag{7.4}$$

where the infimum is over these cutoff functions.

Definition 7.1. A positive function Ψ_C, comparable at comparable radii, is a *capacity scale* if

$$\operatorname{Cap}(B(x,r), B(x,2r)) \le C\,\frac{m(B(x,r))}{\Psi_C(r)} \tag{7.5}$$

uniformly over the specified radii and centres. We refer to this bound as (Cap).

For the form of a gradient structure satisfying $\mathrm{B}(\Psi_{C,2})$, the cutoffs ψ_x^r show that the cutoff scale $\Psi_{C,2}$ of Chapter 3 is a capacity scale, by (3.10). The capacity in (7.4) is the usual capacity of a condenser. It should be distinguished from the generalized capacity condition of [75, Section 1.3], which tests the cutoff function against an additional function and allows the cutoff function to depend on that function. Consequently the equivalence theorems of that paper cannot be applied under (Cap) alone.

Cutoff Sobolev inequalities. An upper bound on the capacity controls the total energy of a cutoff function. The heat kernel estimates at scales of order higher than r^2 need more: a control of the energy measure of the cutoff function, integrated against the square of an arbitrary function.

Definition 7.2. Let Ψ_S be a positive function, comparable at comparable radii. The *cutoff Sobolev inequality in annuli* $\mathrm{CSA}(\Psi_S)$ of Andres and Barlow [10, Definitions 1.9–1.10] holds if each annulus $A = B(x,R+r) \setminus B(x,R)$ with $0 < r \le R$ admits a cutoff function χ for $B(x,R) \subset B(x,R+r)$ such that

$$\int_A f^2\,d\Gamma(\chi) \le \frac{1}{8}\int_A \chi^2\,d\Gamma(f) + \frac{C}{\Psi_S(r)}\int_A f^2\,dm \qquad (f \in \mathcal{F})\,. \tag{7.6}$$

The cutoff function is chosen before f, and the constant is uniform in x, R, r. The inequality is *local in scale* if it holds for all sufficiently small widths r, with arbitrary inner radii, and *global* if it holds at every width.

The left-hand side of (7.6) is the term which appears when the energy of a product $f\chi$ is expanded by the Leibniz rule. The inequality says that this term is controlled by the energy of f in the annulus, with a small constant, and by the mass of f in the annulus at the scale Ψ_S of the width. Three remarks on the definition:

- By locality, the inequality extends to test functions which belong to $\mathcal{F}$ only locally, on a neighbourhood of the closed annulus.
- The restriction $r \le R$ can be lifted by reducing the width of the transition to $(R+r)/2$ and increasing the inner radius to $(R+r)/2$.
- The constant $1/8$ is a normalization. The *unweighted* form of the inequality is

$$\int_A f^2 \, d\Gamma(\chi) \le C_0 \, \Gamma(f)(A) + \frac{C_1}{\Psi_S(r)} \int_A f^2 \, dm \,. \tag{7.7}$$

  If m is doubling and Ψ_S has a finite upper scaling exponent, the construction of Andres and Barlow [10, Lemma 5.1] converts (7.7) into (7.6); see Lemma 7.8. Only smaller transition widths enter that construction. Proposition 9.3 gives a version which preserves Lipschitz continuity, for gradient structures, under the condition $\mathrm{B}^+(\Psi_{C,2})$.

The third hypothesis. In this part we assume the cutoff Sobolev inequality in its unweighted form: for every annulus $A = B(z, R+s) \setminus \overline{B(z,R)}$ with admissible width $0 < s \le R$ there is a cutoff function χ for $B(z,R) \subset B(z,R+s)$ such that

$$\int_X f^2 \, d\Gamma(\chi) \le C_0 \int_A d\Gamma(f) + \frac{C_0}{\Psi_S(s)} \int_A f^2 \, dm \,, \qquad f \in \mathcal{F}_{\mathrm{loc}}(X) \,. \tag{7.8}$$

The same function χ serves all f. Applying (7.8) to a function equal to one on the annulus shows that Ψ_S is a capacity scale:

$$\mathrm{Cap}(B(z,r), B(z,2r)) \le C \, \frac{m(B(z,r))}{\Psi_S(r)} \,. \tag{7.9}$$

Ranges of radii and constants. The hypotheses (7.2), (7.3) and (7.8) are *global* if they hold for all radii on an unbounded space, and *local* if they hold for the radii and the widths below some $r_0 < \infty$. In the local case every fixed dilation of a ball which occurs in a proof must have a radius below r_0. For bounds at all distances, the cutoff inequality (7.8) must be uniform in the inner radius R even when the width s is restricted to small values; the proof of the Harnack inequality uses only annuli contained in the ball under consideration. A constant is called *structural* if it depends only on the constants in (7.1)–(7.3) and (7.8).

From balls to annuli. In applications one often has cutoff functions for balls, with a cutoff Sobolev inequality in which a dilated ball appears on the right-hand side. The next proposition converts such an inequality into the annular form.

Proposition 7.3. *Suppose that m is doubling at sufficiently small radii, and let S be a positive function which is comparable at comparable radii. Assume that for some fixed $\Lambda \ge 2$, every sufficiently small ball $B = B(y,h)$ admits a cutoff function χ_B with $\chi_B = 1$ on B, $\chi_B = 0$ outside $2B$, and*

$$\int_X f^2 \, d\Gamma(\chi_B) \le C \int_{\Lambda B} d\Gamma(f) + \frac{C}{S(h)} \int_{\Lambda B} f^2 \, dm \qquad (f \in \mathcal{F}) \,. \tag{7.10}$$

Then (7.7) holds with $\Psi_S = S$ at every sufficiently small width, uniformly in the inner radius, for $f \in \mathcal{F}_{\mathrm{loc}}(X)$. If the assumptions hold at all radii, the conclusion holds at every width.

Proof. Let $A = B(x, R+r) \setminus \overline{B(x,R)}$, choose $h = r/(8(2+\Lambda))$ and a finite maximal h-separated set of centres y_i in $\overline{B(x, R+r/2)}$. Put $\chi = \max_i \chi_{B(y_i,h)}$. The balls $B(y_i,h)$ cover the closed ball $\overline{B(x,R+r/2)}$. Thus $\chi = 1$ there and $\chi = 0$ outside $\overline{B(x,R+r/2+2h)}$, quasi-everywhere. By strong locality, the energy measure of χ is carried by

$$K = \{y : R + r/2 \le d(x,y) \le R + r/2 + 2h\} \subset A\,,$$

and, since $\Gamma(\max\{u,v\}) \le \Gamma(u) + \Gamma(v)$, it is bounded by the sum of the measures $\Gamma(\chi_{B(y_i,h)})$ restricted to K. If $\overline{B(y_i,2h)}$ meets K, then $d(x,y_i) \ge R+r/2-2h$, and consequently $B(y_i,\Lambda h) \subset A$. These balls have uniformly bounded overlap, by doubling and separation. Summing (7.10) over these indices and using $S(h) \asymp S(r)$ proves (7.7). Since the balls which occur on the right-hand side have compact closures in A, locality allows $f \in \mathcal{F}_{\mathrm{loc}}(X)$. □

The order of the scales. The Poincaré inequality bounds the capacity of a condenser from below, and a cutoff Sobolev inequality bounds it from above. Hence the cutoff Sobolev scale is dominated by the Poincaré scale. This is the analogue, for Dirichlet forms, of Proposition 3.7.

Proposition 7.4. *Let X be geodesic with positive diameter, and let $(\mathcal{E},\mathcal{F})$ be a strongly local regular Dirichlet form satisfying* VD *and* (7.3).

(i) *Every scale Ψ_S for which* CSA(Ψ_S) *holds is a capacity scale.*

(ii) *For every capacity scale Ψ_C, one has $\Psi_C(r) \lesssim \Psi_P(r)$ at sufficiently small radii, and at all radii under global hypotheses.*

In particular,

$$\Psi_S \lesssim \Psi_P\,, \qquad \Psi_C \lesssim \Psi_P\,. \tag{7.11}$$

Proof. (i) Let χ be the cutoff function of CSA(Ψ_S) for $B(x,r) \subset B(x,2r)$, and test (7.6) with a function $f \in \mathcal{F}$ which is equal to one on a neighbourhood of $\operatorname{supp}\chi$. By strong locality $\Gamma(f)$ does not charge $\{\chi \ne 0\}$, and $\Gamma(\chi)$ is carried by the annulus A. Hence $\mathcal{E}(\chi,\chi) = \int_A f^2\,d\Gamma(\chi) \le C\,m(A)/\Psi_S(r)$, and doubling gives (7.5).

(ii) We use the Poincaré inequality for a function which vanishes on a set of comparable measure; compare [121, Lemma 3.7]. Restrict first to $8r < \operatorname{diam} X$. Since X is geodesic, there is a point z at distance $3r$ from x. The ball $B(z,r/4)$ lies in $B(x,4r) \setminus B(x,2r)$, and doubling gives $m(B(z,r/4)) \ge c\,m(B(x,r))$. A cutoff function χ for $B(x,r) \subset B(x,2r)$ is equal to one on $B(x,r)$ and to zero on $B(z,r/4)$. Hence $\int_{B(x,4r)} |\chi - a|^2\,dm \ge c\,m(B(x,r))$ for every constant a. The Poincaré inequality on $B(x,4r)$ gives $c\,m(B(x,r)) \le C\,\Psi_P(r)\,\mathcal{E}(\chi,\chi)$, that is,

$$\operatorname{Cap}(B(x,r),B(x,2r)) \ge c\,\frac{m(B(x,r))}{\Psi_P(r)}\,. \tag{7.12}$$

Comparison with (7.5) proves the assertion. On an unbounded space the restriction on the diameter is absent. □

7.3 The parabolic Harnack inequality

In the rest of this chapter X is geodesic, and the hypotheses (7.2), (7.3) with $\Lambda = 1$, and (7.8) hold on the balls which are used. In the local case the radius r of a cylinder is restricted by $32\,r < r_0$, and on a bounded space also by $32\,r < \operatorname{diam} X$; we call such x and r *admissible*.

Theorem 7.5 (Parabolic Harnack inequality with two scales)**.** *There is a structural constant $C < \infty$ such that, for all admissible x and r, all $t_0 \in \mathbb{R}$ and $T > 0$, every nonnegative local weak solution u of the heat equation in $(t_0 - T, t_0 + T) \times B(x, r)$ satisfies*

$$\operatorname*{ess\,sup}_{Q_-} u \le \exp\Big[C\Big(1 + \frac{\Psi_P(r)}{T} + \frac{T}{\Psi_S(r)}\Big)\Big] \operatorname*{ess\,inf}_{Q_+} u\,, \tag{7.13}$$

where

$$Q_- = (t_0 - \tfrac{3}{4}T,\, t_0 - \tfrac{1}{2}T) \times B(x, r/2)\,, \qquad Q_+ = (t_0 + \tfrac{1}{2}T,\, t_0 + \tfrac{3}{4}T) \times B(x, r/2)\,.$$

When the scales coincide, $\Psi_P = \Psi_S = \Psi$, with a lower scaling exponent at least two, and $T \asymp \Psi(r)$, this is [121, Theorem 5.3]. The two terms in the exponent have different origins. Heuristically, the term $\Psi_P(r)/T$ is large when the time is short compared with the Poincaré scale, so that the solution has not had the time to equilibrate over the ball; the term $T/\Psi_S(r)$ is large when the time is long compared with the cutoff scale, so that the values of u near the boundary of the ball, which are not controlled, have had the time to influence the interior.

Corollary 7.6 (Elliptic Harnack inequality)**.** *Every nonnegative weakly harmonic function u in $B(x, r)$ satisfies*

$$\operatorname*{ess\,sup}_{B(x,r/2)} u \le \exp\Big[C\Big(1 + \sqrt{\Psi_P(r)/\Psi_S(r)}\Big)\Big] \operatorname*{ess\,inf}_{B(x,r/2)} u\,.$$

Proof. Apply Theorem 7.5 to the stationary solution $u(t, \cdot) = u$, with $T = \sqrt{\Psi_P(r)\,\Psi_S(r)}$, which minimizes $\Psi_P(r)/T + T/\Psi_S(r)$. □

Remark 7.7. When the two scales are comparable, the corollary is the elliptic Harnack inequality with a constant independent of r. In general the constant depends on the ratio $\Psi_P(r)/\Psi_S(r)$, which may be unbounded as $r \to 0$. The inequalities are stated with essential suprema and infima. A pointwise formulation requires a suitable choice of representative, whose continuity is not asserted here.

Outline of the proof. After normalization the cylinder has radius one and time length one, the Poincaré inequality has the coefficient $A = \Psi_P(r)/T$ and the cutoff Sobolev inequality the coefficient $B = T/\Psi_S(r)$. The proof has four steps, which occupy the next three sections.

(1) *Cutoffs and Sobolev inequalities* (Section 7.4). The cutoff Sobolev inequality is put in a weighted form with a small coefficient in front of the energy, and the Poincaré inequality gives a Sobolev inequality with a gain of integrability.

(2) *Mean value inequalities* (Section 7.5). Moser's iteration bounds the supremum of a subsolution by its L^2 norm (Lemma 7.11), and the suprema of u^p and u^{-p} by their integrals for small $p > 0$ (Lemma 7.12).

(3) *The logarithm* (Section 7.6). The mean of $\log u$ on a ball is almost monotone in time, and the measure of the set where $\log u$ deviates from its mean at the time t_0 is controlled (Lemma 7.13).

(4) *Conclusion.* A lemma of Bombieri and Giusti (Lemma 7.14) combines the mean value inequality for small powers with the logarithmic estimate, and bounds u from above before t_0 and from below after t_0 in terms of the same number.

The linear dependence of the exponent in (7.13) on $A + B$ comes from two facts: the mean value inequality for the powers $u^{\pm p}$ has a constant independent of A and B when $p \le c/(1 + A + B)$, and the logarithmic estimate has a constant proportional to A and a drift proportional to B.

7.4 Cutoffs and Sobolev inequalities

The first lemma converts the cutoff Sobolev inequality (7.8), which has a fixed coefficient in front of the energy, into a weighted inequality with an arbitrarily small coefficient. The price is a power of this coefficient in front of the zeroth-order term, determined by the upper scaling exponent β_+ of Ψ_S. The construction, which sums cutoff functions of a sequence of thin annuli with geometrically decreasing widths and weights, is due to Andres and Barlow [10, Lemma 5.1]; the quantitative weighted form is [121, Lemma 2.3].

Lemma 7.8 (Weighted cutoff functions)**.** *For every admissible pair $B(z,R) \subset B(z,R+s)$ and every $\varepsilon \in (0,1)$, there is a cutoff function ψ for these balls such that*

$$\int_A f^2 \, d\Gamma(\psi) \le \varepsilon \int_A \psi^2 \, d\Gamma(f) + \frac{C\varepsilon^{1-\beta_+/2}}{\Psi_S(s)} \int_A \psi f^2 \, dm \tag{7.14}$$

for every $f \in \mathcal{F}_{\mathrm{loc}}(X)$, where $A = B(z,R+s) \setminus \overline{B(z,R)}$. If $s \asymp R$ and ε is fixed, the cutoff function also satisfies the weighted Poincaré inequality

$$\int |f - f_\psi|^2 \psi^2 \, dm \le C \, \Psi_P(R) \int \psi^2 \, d\Gamma(f) \,, \qquad f_\psi = \frac{\int f\psi^2 \, dm}{\int \psi^2 \, dm} \,, \tag{7.15}$$

with a constant which depends on the fixed ε and on the comparison between s and R.

Proof. We follow [121, Lemma 2.3 and Theorem 3.4], keeping the two scales separate. Write c_1 for the coefficient of the energy in (7.8) and c_2 for the other constant, and let $\beta = \beta_+$.

Step 1: construction. Put $p = e^{-\lambda}$, where

$$\lambda = \log(1 + \sqrt{\varepsilon/c_1}) \,, \qquad s_n = \frac{s}{2} \, (p^{-1/\beta} - 1) \, p^{n/\beta} \,, \qquad B_n = B\Big(z, R + \sum_{j=1}^{n} s_j\Big) \,.$$

Then $\sum_{n \ge 1} s_n = s/2$. Let χ_n be the cutoff function of (7.8) for $B_{n-1} \subset B_n$, and set

$$\alpha_n = (1-p) \, p^{n-1} \,, \qquad \psi = \sum_{n \ge 1} \alpha_n \chi_n \,.$$

Since $\sum_n \alpha_n = 1$, the series converges uniformly outside a set of zero capacity. Its support lies in $\overline{B(z, R + s/2)}$, and it is equal to one on a neighbourhood of $\overline{B(z,R)}$ quasi-everywhere. On the annulus $B_n \setminus B_{n-1}$ the functions χ_j with $j > n$ are equal to one and those with $j < n$ vanish, so that

$$\psi = \alpha_n \chi_n + p^n \ge p^n \qquad \text{on } B_n \setminus B_{n-1} \,.$$

Step 2: the cutoff Sobolev inequality. By strong locality,

$$\int f^2 \, d\Gamma(\psi) = \sum_{n \ge 1} \alpha_n^2 \int_{B_n \setminus B_{n-1}} f^2 \, d\Gamma(\chi_n) \,.$$

By the choice of λ, $c_1\alpha_n^2 = \varepsilon\, p^{2n}$, and by the upper scaling of Ψ_S, $\Psi_S(s)/\Psi_S(s_n) \le C(s/s_n)^\beta$, where $p^{-1/\beta} - 1 \asymp \lambda \asymp \sqrt{\varepsilon}$. Hence, on $B_n \setminus B_{n-1}$,

$$c_1\alpha_n^2 \le \varepsilon\,\psi^2\,, \qquad \frac{c_2\,\alpha_n^2}{\Psi_S(s_n)} \le \frac{C\varepsilon^{1-\beta/2}}{\Psi_S(s)}\,p^n \le \frac{C\varepsilon^{1-\beta/2}}{\Psi_S(s)}\,\psi\,.$$

Summing the inequalities (7.8) for the functions χ_n proves (7.14). The same estimates, with a function equal to one on the annulus, show that the partial sums of the series are Cauchy in $\mathcal{F}$; this justifies the identity for the energy measure of the infinite sum.

Step 3: the weighted Poincaré inequality. Define $\Xi = \sum_{n\ge1}\alpha_n \mathbf{1}_{B_n}$. Since $\chi_n \le \mathbf{1}_{B_n}$ and $\chi_{n+1} = 1$ on B_n,

$$\psi \le \Xi \le p^{-1}\psi\,.$$

When $s \asymp R$, all the balls B_n have comparable radii and volumes. Since $B_0 \subset B_n$, the Poincaré inequality with $\Lambda = 1$ gives

$$\int_{B_n} |f - f_{B_0}|^2\,dm \le C\int_{B_n} |f - f_{B_n}|^2\,dm \le C\,\Psi_P(R)\int_{B_n} d\Gamma(f)\,.$$

The weighted mean f_ψ minimizes the weighted square error. Expanding $\Xi^2 = \sum_{j,k\ge1}\alpha_j\alpha_k\mathbf{1}_{B_{\min\{j,k\}}}$ and applying the preceding estimate to each term gives

$$\int |f - f_\psi|^2\psi^2\,dm \le \int |f - f_{B_0}|^2\,\Xi^2\,dm \le C\,\Psi_P(R)\int \Xi^2\,d\Gamma(f) \le Cp^{-2}\,\Psi_P(R)\int \psi^2\,d\Gamma(f)\,.$$

Since ε is fixed, p^{-2} is a structural constant. □

The next two lemmas are Sobolev inequalities. They follow from the volume estimate and the Poincaré inequality alone, by the telescoping argument of Hajłasz and Koskela [79]; compare the Sobolev inequalities for Dirichlet forms in [121, Section 3.2, especially Theorem 3.9] and the Gaussian case in [140, Theorem 2.1]. The gain of integrability is determined by the lower scaling exponent a_- of the Poincaré scale. No cutoff inequality is used.

Lemma 7.9 (Interior Sobolev inequality)**.** *Fix $D > \max\{Q_D, a_-\}$ and $1 < \kappa < D/(D - a_-)$. For $1/4 \le R' < R \le 1$, put $h = R - R'$. Then*

$$\left(\frac{1}{m(B(x,r))}\int_{B(x,R'r)} |v|^{2\kappa}\,dm\right)^{1/\kappa} \le \frac{C\,h^{-D}}{m(B(x,r))}\left(\Psi_P(r)\int_{B(x,Rr)} d\Gamma(v) + \int_{B(x,Rr)} v^2\,dm\right)$$

for every $v \in \mathcal{F}_{\mathrm{loc}}$ on a neighbourhood of $\overline{B(x,Rr)}$. The constant depends only on the constants of the volume and Poincaré hypotheses, on D and on κ.

Proof. Step 1: normalization. Divide the distance by r, the measure by $m(B(x,r))$, and the energy by $m(B(x,r))/\Psi_P(r)$. In these units $m(B(x,1)) = 1$; for $y \in B(x,1)$ and $0 < s \le 1$, doubling gives $m(B(y,s)) \ge c\,s^{Q_D} \ge c\,s^D$; and the coefficient of the Poincaré inequality at a radius $s \le 1$ is at most Cs^{a_-}. Put $e = \mathbf{1}_{B(x,R)}\Gamma(v)$ and $E = e(X)$.

Step 2: a pointwise bound. The Lebesgue differentiation theorem and the Poincaré inequality, applied to the balls $B(y, 2^{-j}h/8)$, which lie in $B(x,R)$, give

$$|v(y)| \le C\,h^{-D/2}\|v\|_{L^2(B(x,R))} + C\sum_{j\ge0}(2^{-j}h)^{a_-/2}\left(\frac{e(B(y,2^{-j}h/8))}{m(B(y,2^{-j}h/8))}\right)^{1/2} \tag{7.16}$$

for almost every $y \in B(x,R')$. The first term bounds the average on the largest ball. Let $\mathcal{M}e(y)$ be the supremum of the ratios $e(B(y,s))/m(B(y,s))$ over the radii $s \le h/8$. Each term of the sum in (7.16) has the two upper bounds

$$C(2^{-j}h)^{a_-/2}\,\mathcal{M}e(y)^{1/2} \qquad \text{and} \qquad C(2^{-j}h)^{(a_- - D)/2}E^{1/2}\,.$$

Using the first bound for the radii below $(E/\mathcal{M}e(y))^{1/D}$ and the second for the others gives

$$|v(y)| \le C\,h^{-D/2}\|v\|_{L^2(B(x,R))} + C\,E^{a_-/(2D)}\,\mathcal{M}e(y)^{(D-a_-)/(2D)}\,.$$

Step 3: integration. The $5r$-covering argument for the maximal function [83, Chapter 2], applied to the finite measure e, and doubling imply the weak-type bound

$$m(\{y \in B(x,R') : \mathcal{M}e(y) > \lambda\}) \le \frac{C\,E}{\lambda}\,. \tag{7.17}$$

The distribution function of the second term of the pointwise bound is therefore at most $C\min\{1,(E/t^2)^{D/(D-a_-)}\}$. Since $\kappa < D/(D-a_-)$, integration against $2\kappa\,t^{2\kappa-1}\,dt$ gives

$$\|v\|^2_{L^{2\kappa}(B(x,R'))} \le C\,h^{-D}\|v\|^2_{L^2(B(x,R))} + C\,E\,.$$

Returning to the original units proves the assertion. □

Lemma 7.10 (Sobolev inequality)**.** *There are $\kappa > 1$ and $C < \infty$, depending only on the constants of the volume and Poincaré hypotheses and on a_-, such that*

$$\left(\frac{1}{m(B(z,r))}\int_{B(z,r)} |f|^{2\kappa}\,dm\right)^{1/\kappa} \le \frac{C\,\Psi_P(r)}{m(B(z,r))}\,\mathcal{E}(f,f) \tag{7.18}$$

for every $f \in \mathcal{F}$ with compact support in $B(z,r)$. One may choose any $D > \max\{Q_D, a_-\}$ and any $1 < \kappa < D/(D-a_-)$.

Proof. Apply Lemma 7.9 between $B(z,r)$ and $B(z,2r)$. Its right-hand side is at most

$$\frac{C}{m(B(z,r))}\bigl(\Psi_P(r)\,\mathcal{E}(f,f) + \|f\|_2^2\bigr)\,.$$

Since X is geodesic, $B(z,4r)\setminus B(z,r)$ contains a ball of radius $r/4$, on which $f = 0$. The Poincaré inequality on $B(z,4r)$ and doubling give

$$\|f\|_2^2 \le C\int_{B(z,4r)} |f - f_{B(z,4r)}|^2\,dm \le C\,\Psi_P(r)\,\mathcal{E}(f,f)\,.$$

Substitution proves (7.18). □

7.5 Mean value inequalities

Normalization. Fix x, r and T, and normalize the distance, the measure and the form by

$$d_r = r^{-1}d\,, \qquad m_r = \frac{m}{m(B(x,r))}\,, \qquad \mathcal{E}_r = \frac{T}{m(B(x,r))}\,\mathcal{E}\,, \tag{7.19}$$

and divide the time by T. With these units, a solution of the heat equation on a cylinder of radius r and time length T becomes a solution, for the normalized form, on a cylinder of radius one and time length one. We write

$$A = \frac{\Psi_P(r)}{T}\,, \qquad B = \frac{T}{\Psi_S(r)}\,, \qquad \beta = \beta_+\,,$$

and drop the subscripts r. In the normalized units:

- the balls of radii between $1/4$ and 1 have uniformly comparable measures;
- the Sobolev inequality (7.18) and the weighted Poincaré inequality (7.15) have the coefficient CA;
- the weighted cutoff Sobolev inequality (7.14), for a cutoff function of width $h \le 1$ and a fixed ε, has the zeroth-order coefficient $CBh^{-\beta}$.

Put

$$\sigma = 2 - \frac{1}{\kappa} > 1\,, \qquad \theta = \frac{\sigma}{\sigma - 1} = \frac{2\kappa - 1}{\kappa - 1}\,. \tag{7.20}$$

In this normalization, Lemma 7.9 gives, for $1/4 \le R' < R \le 1$, $h = R - R'$ and $d_0 = D$,

$$\left(\int_{B(x,R')} |v|^{2\kappa}\, dm \right)^{1/\kappa} \le C\, h^{-d_0} \left(A \int_{B(x,R)} d\Gamma(v) + \int_{B(x,R)} v^2\, dm \right)\,. \tag{7.21}$$

The L^2 mean value inequality. The first estimate is Moser's mean value inequality [127] for subsolutions; compare [121, Theorem 4.7 and Corollary 4.8]. An energy estimate for the powers of u is combined with the Sobolev inequality, and the resulting gain of integrability is iterated.

Lemma 7.11. *Let $u \ge 0$ be a local weak subsolution in $(-1,0) \times B(x,1)$. Then*

$$\operatorname*{ess\,sup}_{(-1/2,0)\times B(x,1/2)} u^2 \le C(1+A)^{\theta-1}(1+B)^{\theta} \int_{-1}^{0} \int_{B(x,1)} u^2\, dm\, dt\,. \tag{7.22}$$

The estimate holds between nested cylinders with positive spatial and initial time gaps, with an additional fixed negative power of the smaller gap. If q belongs to a fixed compact subinterval of $(0,2]$, u^2 can be replaced throughout by u^q, with a constant uniform in q.

Proof. Step 1: the energy estimate. For $p \ge 2$, put $v = u^{p/2}$ and let ψ be a spatial cutoff function. For bounded u, testing the equation with $u^{p-1}\psi^2$ gives

$$\begin{aligned} \frac{d}{dt} \int \psi^2 u^p\, dm &\le -\frac{4(p-1)}{p} \int \psi^2\, d\Gamma(v) - 4 \int \psi v\, d\Gamma(v,\psi) \\ &\le -\int \psi^2\, d\Gamma(v) + C \int v^2\, d\Gamma(\psi)\,, \end{aligned}$$

with a constant uniform in $p \ge 2$. Apply (7.14) with a fixed small parameter and absorb its energy term. Multiplying by a nondecreasing cutoff in time, which vanishes at the initial time and is equal to one after a gap k, and integrating, gives

$$\sup_{I'} \int \psi^2 u^p\, dm + \int_{I'} \int d\Gamma(\psi u^{p/2})\, dt \le C(Bh^{-\beta} + k^{-1}) \int_I \int_{B_R} u^p\, dm\, dt\,. \tag{7.23}$$

Here h is the width of the spatial cutoff function, B_R contains its support, the interval I' begins k after I, and the two intervals have the same final endpoint. For the energy of the product we used $\Gamma(\psi v) \le 2\psi^2\Gamma(v) + 2v^2\Gamma(\psi)$ and the cutoff Sobolev inequality again.

Step 2: removal of the boundedness. The estimate (7.23) does not presuppose that u is locally bounded. As in [121, proof of Lemma 4.4], replace s^p by the function $F_n(s)$ which is equal to s^p for $s \le n$ and to its quadratic Taylor polynomial at n for $s \ge n$. Its derivative is globally Lipschitz and

$$F_n'(s)^2 \le 2F_n(s)F_n''(s)\,, \qquad 2\left|(\sqrt{F_n})'(s)\right|^2 \le F_n''(s)\,.$$

Use $F_n'(u)\psi^2$ as the test function. The calculation of Step 1 with $v = \sqrt{F_n(u)}$ proves (7.23) with constants independent of n and p; the chain rule in time is justified by Steklov averaging, since F_n' is Lipschitz. Since $F_n(s) \uparrow s^p$, monotone convergence and lower semicontinuity give (7.23) whenever its right-hand side is finite. This allows the iteration to start from $p = 2$.

Step 3: gain of integrability. For a function w of space and time with compact support in space, the Sobolev inequality (7.18) and Hölder's inequality give

$$\int_I \int |w|^{2\sigma}\, dm\, dt \le CA \left(\sup_I \int w^2\, dm\right)^{\sigma-1} \int_I \int d\Gamma(w)\, dt\,. \tag{7.24}$$

Step 4: iteration. Choose nested cylinders Q_j decreasing to the cylinder on the left-hand side of (7.22), with spatial and initial time gaps comparable to 2^{-j}, and put $p_j = 2\sigma^j$. Combining (7.23) and (7.24), with $w = \psi u^{p_j/2}$, gives

$$\int_{Q_{j+1}} u^{p_j\sigma}\, dm\, dt \le CA\left[C(1+B)\,2^{j\beta}\int_{Q_j} u^{p_j}\, dm\, dt\right]^{\sigma}\,.$$

Raise this inequality to the power σ^{-j-1} and multiply over j. The sums of the exponents of A and of $1+B$ are $\sum_j \sigma^{-j-1} = \theta - 1$ and $\sum_j \sigma^{-j} = \theta$, respectively, and $\sum_{j\ge 0} j\sigma^{-j} < \infty$. Letting $j \to \infty$ proves (7.22). Smaller geometric gaps introduce only a negative power of those gaps.

Step 5: other exponents. Let q lie in a fixed compact subinterval of $(0, 2]$, and let $S(\delta)$ be the essential supremum of u on the cylinder of a nested family with parameter δ. The L^2 estimate between nested cylinders implies

$$S(\delta') \le \left[M(\delta-\delta')^{-\gamma}\int u^q\, dm\, dt\right]^{1/2} S(\delta)^{1-q/2}\,, \qquad M = C(1+A)^{\theta-1}(1+B)^{\theta}\,,$$

where the integral is over the larger cylinder. By Young's inequality the right-hand side is at most $S(\delta)/2 + C[M(\delta-\delta')^{-\gamma}\int u^q]^{1/q}$. Iterate over a sequence of gaps which decreases slowly enough for the resulting series to converge; this choice is uniform because q stays away from zero. The local boundedness, already proved, removes the terminal supremum. Raising to the power q gives the assertion. □

Small powers. The next lemma is the mean value inequality for small positive and negative powers of a positive solution. Its constant does not depend on A and B, provided that the power is at most a fixed multiple of $(1 + A + B)^{-1}$. This is the step which gives a Harnack constant whose logarithm is linear in $A + B$. The iterations for positive and negative powers adapt [121, Lemmas 4.5–4.6 and Theorem 4.10]; the dependence on the parameters is tracked below.

Lemma 7.12. *Let $u > 0$ be a local weak solution in $(-1,1) \times B(x,1)$. For $1/2 \le \delta \le 1$, set*

$$U_\delta^\pm = \left(\pm\frac{5}{8} - \frac{\delta}{4},\ \pm\frac{5}{8} + \frac{\delta}{4} \right) \times B\left(x, \frac{1}{4} + \frac{\delta}{2} \right) . \tag{7.25}$$

There are structural constants $c_*, C, \gamma > 0$ such that, whenever $0 < p < c_*/(1 + A + B)$ and $1/2 \le \delta' < \delta \le 1$,

$$\begin{aligned} \operatorname*{ess\,sup}_{U_{\delta'}^-} u^p &\le C(\delta - \delta')^{-\gamma} \int_{U_\delta^-} u^p \, dm \, dt \,, \\ \operatorname*{ess\,sup}_{U_{\delta'}^+} u^{-p} &\le C(\delta - \delta')^{-\gamma} \int_{U_\delta^+} u^{-p} \, dm \, dt \,. \end{aligned} \tag{7.26}$$

The estimates hold for $u + \varepsilon$ whenever u is a nonnegative solution and $\varepsilon > 0$, with constants independent of ε.

The cylinders U_δ^- lie in the past, U_δ^+ in the future, and $U_{1/2}^\pm$ are the cylinders $Q_\pm$ of Theorem 7.5 in the normalized units, with $t_0 = 0$.

Proof. We first work with a solution which is bounded above and bounded away from zero on the interior cylinders.

Step 1: energy estimates for positive powers. For $0 < q \le 1/2$, write $v = u^{q/2}$ and $F(t) = \int \psi^2 u^q \, dm$. The heat equation gives

$$F'(t) = \frac{4(1-q)}{q} \int \psi^2 \, d\Gamma(v) - 4 \int \psi v \, d\Gamma(v, \psi) \,.$$

For a power $q < 1$ the energy term has a favourable sign in the direction of increasing time. Young's inequality and (7.14), used with a fixed small parameter, give

$$F'(t) \ge \frac{1}{2q} \int \psi^2 \, d\Gamma(v) - C \, q \, B \, h^{-\beta} \int u^q \, dm \,, \tag{7.27}$$

where h is the width of the cutoff function. The factor q in the error term is essential and is retained.

Suppose that $\psi = 1$ on an intermediate ball B_m and is supported in an outer ball B_R. Choose a nonincreasing cutoff in time, equal to one before b' and to zero at $b > b'$, with derivative bounded by C/k, where $k = b - b'$. Integrating the derivative of the product of this cutoff and of F from t to b gives, for $s < t < b'$, a nonnegative term $F(t)$ and a nonnegative energy term on the left-hand side. Dropping either term, and then letting $t \downarrow s$ for the energy estimate, proves

$$\begin{aligned} \sup_{s<t<b'} \int_{B_m} v^2 \, dm &\le C(k^{-1} + qBh^{-\beta}) \int_s^b \int_{B_R} u^q \, dm \, dt \,, \\ \int_s^{b'} \int_{B_m} d\Gamma(v) \, dt &\le Cq(k^{-1} + qBh^{-\beta}) \int_s^b \int_{B_R} u^q \, dm \, dt \,. \end{aligned} \tag{7.28}$$

No value at the initial time s is prescribed. Estimating $\Gamma(v)$ on the set where $\psi = 1$ preserves the factor q in the second line.

Step 2: one step of the iteration. Fix $\delta' < \delta$ and write $\Delta = \delta - \delta'$. Denote the endpoints in time and the radius of U_δ^- by s, b, R, and those of $U_{\delta'}^-$ by s', b', R'. Put

$$\begin{aligned} b_* &= (b + b')/2 \,, & R_* &= (R + R')/2 \,, \\ b_j &= b_* + 2^{-j}(b - b_*) \,, & R_j &= R_* + 2^{-j}(R - R_*) \,, \end{aligned}$$

and let $V_j = (s, b_j) \times B(x, R_j)$. Choose the cutoff function of the energy estimate between $B(x, (R_j + R_{j+1})/2)$ and $B(x, R_j)$, and apply (7.21) between $B(x, R_{j+1})$ and this intermediate ball. Both spatial gaps are $h_j = (R_j - R_{j+1})/2 \asymp \Delta 2^{-j}$, and the gap at the final time is $k_j = b_j - b_{j+1} \asymp \Delta 2^{-j}$.

Let $J_j(q) = \int_{V_j} u^q \, dm \, dt$. The interpolation inequality

$$\int |v|^{2\sigma} \, dm \le \left(\int v^2 \, dm \right)^{\sigma - 1} \left(\int |v|^{2\kappa} \, dm \right)^{1/\kappa} ,$$

the Sobolev inequality (7.21) and the energy estimates (7.28) give

$$J_{j+1}(q\sigma) \le C \, h_j^{-d_0} \, (k_j^{-1} + qBh_j^{-\beta})^{\sigma - 1} \left(1 + Aqk_j^{-1} + Aq^2 Bh_j^{-\beta} \right) J_j(q)^{\sigma} . \tag{7.29}$$

The term 1 in the last factor comes from the zeroth-order term of (7.21), whose time integral is at most $J_j(q)$.

Step 3: finitely many steps with bounded constants. Choose $c_* < 1/(4\sigma)$ and set $q_j = p \, \sigma^j$. Stop at the first N such that $q_N \ge 1/(2\sigma)$; then $q_N < 1/2$. If $p < c_*/(1 + A + B)$, then

$$Aq_j \le c_* \sigma^j , \qquad Bq_j \le c_* \sigma^j , \qquad ABq_j^2 \le c_*^2 \sigma^{2j} .$$

Thus the coefficient in (7.29), evaluated at $q = q_j$, is at most $C\Delta^{-\gamma_0} \lambda_0^j$, for structural constants $\gamma_0 > 0$ and $\lambda_0 > 1$: the coefficients A and B have disappeared. Raising the successive estimates to the powers σ^{-j-1} and multiplying gives

$$J_N(q_N)^{p/q_N} \le C \Delta^{-\gamma_1} J_0(p) , \tag{7.30}$$

since $\sum_{j \ge 0} \sigma^{-j-1}$ and $\sum_{j \ge 0} j \sigma^{-j-1}$ are finite.

Step 4: from the exponent q_N to the supremum. The cylinder V_N contains $U_{\delta'}^-$ with a spatial gap at least $(R - R')/2$ and an initial time gap $s' - s = \Delta/4$. Apply Lemma 7.11 with the exponent $q_N \in [1/(2\sigma), 1/2]$. Its constant is a structural negative power of Δ times a fixed power of $(1 + A)(1 + B)$. After raising it to the power p/q_N, its dependence on A and B is bounded, because

$$\frac{p}{q_N} \log\bigl((1 + A)(1 + B)\bigr) \le C \, p \log(1 + A + B) \le C .$$

Combining with (7.30) proves the first estimate in (7.26).

Step 5: negative powers. Put $v = u^{-q/2}$ and $F = \int \psi^2 u^{-q} \, dm$. Differentiation and the same cutoff Sobolev inequality give

$$F' \le -\frac{1}{q} \int \psi^2 \, d\Gamma(v) + C \, q \, B \, h^{-\beta} \int u^{-q} \, dm .$$

Now the energy term has a favourable sign in the direction of decreasing time. Use a nondecreasing cutoff in time which vanishes at the initial time. This gives the two estimates in (7.28), with u^{-q} in place of u^q and with a gap at the initial time. Keep the final time of U_δ^+ fixed, increase its initial time towards the midpoint between the initial times of U_δ^+ and $U_{\delta'}^+$, and decrease its radius towards the midpoint of the radii. The recurrence (7.29) and the bounds on its coefficients are unchanged. At the final exponent, apply Lemma 7.11 to u^{-1}, which is a subsolution by the chain rule; a positive gap at the initial time remains. This proves the second estimate.

Step 6: general solutions. For a nonnegative solution u, apply the calculation to $u + \varepsilon$. By strong locality this is a solution, Lemma 7.11 gives its local boundedness from above, and $u + \varepsilon \ge \varepsilon$.

The test functions built from powers have bounded derivatives on the relevant compact cylinders; Steklov averaging justifies the identities as in [121, Lemmas 4.4 and 5.1]. Every constant is independent of ε. For positive u, let $\varepsilon \downarrow 0$, using monotone convergence for the negative powers; the statement is trivial if the integral on the right-hand side is infinite. □

7.6 The logarithm and the proof of the Harnack inequality

The mean value inequalities of the preceding section bound the suprema of u^p and u^{-p} by their integrals. To compare these integrals, Moser's method [127, 128] uses the logarithm of the solution: its weighted mean is almost nondecreasing in time, and the time derivative of the mean controls the variance, by the weighted Poincaré inequality.

Lemma 7.13 (Logarithmic estimate)**.** *Let $\Omega \subset X$ be open, and let $0 \le \psi \le 1$ belong to $\mathcal{F}$, with compact support in Ω. Put $M = \int \psi^2 \, dm > 0$ and $d\mu = M^{-1}\psi^2 \, dm$. Suppose that, for some $A > 0$ and $B \ge 0$,*

$$\int |f - f_\psi|^2 \psi^2 \, dm \le A \int \psi^2 \, d\Gamma(f)\,, \qquad \mathcal{E}(\psi,\psi) \le BM\,, \qquad f_\psi = \int f \, d\mu$$

for every $f \in \mathcal{F}_{\mathrm{loc}}(\Omega)$. If $u \ge 0$ is a local weak solution in $(t_0 - T, t_0 + T) \times \Omega$, then, for $u_\varepsilon = u + \varepsilon$, the function $Y_\varepsilon(t) = \int \log u_\varepsilon(t) \, d\mu$ has a locally absolutely continuous representative satisfying

$$Y_\varepsilon'(t) \ge \frac{1}{2A} \operatorname{Var}_\mu(\log u_\varepsilon(t)) - 2B\,. \tag{7.31}$$

Moreover, for every $\ell > 0$,

$$\begin{aligned} \int_{t_0 - T}^{t_0} \mu\bigl(\{\log u_\varepsilon(t) > Y_\varepsilon(t_0) + 2BT + \ell\}\bigr)\, dt &\le \frac{2A}{\ell}\,, \\ \int_{t_0}^{t_0 + T} \mu\bigl(\{\log u_\varepsilon(t) < Y_\varepsilon(t_0) - 2BT - \ell\}\bigr)\, dt &\le \frac{2A}{\ell}\,. \end{aligned} \tag{7.32}$$

The estimates (7.32) say that, before the time t_0, the solution is rarely much larger than $\exp(Y_\varepsilon(t_0) + 2BT)$, and that after t_0 it is rarely much smaller than $\exp(Y_\varepsilon(t_0) - 2BT)$. The same number $Y_\varepsilon(t_0)$ appears in both; this is what links the past to the future in the Harnack inequality.

Proof. Testing the heat equation with ψ^2/u_ε gives

$$\begin{aligned} M\, Y_\varepsilon'(t) &= \int \psi^2 \, d\Gamma(\log u_\varepsilon) - 2 \int \psi \, d\Gamma(\log u_\varepsilon, \psi) \\ &\ge \frac{1}{2} \int \psi^2 \, d\Gamma(\log u_\varepsilon) - 2\, \mathcal{E}(\psi,\psi)\,. \end{aligned} \tag{7.33}$$

For weak solutions, one first replaces u by its forward Steklov average on a compact time interval. The reciprocal of u_ε and its derivative are bounded, so the test function belongs to $\mathcal{F}$ and has compact support in Ω. The chain rule in L^2 applies, and convergence in the local energy space gives (7.33) and the local absolute continuity. This is the regularization in [121, proof of Lemma 5.1]. The weighted Poincaré inequality now gives (7.31).

Put $Z(t,z) = \log u_\varepsilon(t,z) + 2B(t - t_0)$ and $W(t) = \int Z(t) \, d\mu$. Then

$$W'(t) \ge \frac{1}{2A} \int |Z(t) - W(t)|^2 \, d\mu \ge 0\,. \tag{7.34}$$

For $t < t_0$, set $\rho(t) = W(t_0) - W(t) \ge 0$. On the set $\{Z(t) > W(t_0) + \ell\}$ one has $Z(t) - W(t) > \ell + \rho(t)$; hence, by Chebyshev's inequality and (7.34),

$$\mu\big(\{Z(t) > W(t_0) + \ell\}\big) \le \frac{2A\,W'(t)}{(\ell + \rho(t))^2} = 2A\,\frac{d}{dt}\,\frac{1}{\ell + \rho(t)}\,.$$

Integration gives

$$\int_{t_0 - T}^{t_0} \mu\big(\{Z(t) > W(t_0) + \ell\}\big)\,dt \le \frac{2A}{\ell}\,. \tag{7.35}$$

For $t > t_0$, use $\rho(t) = W(t) - W(t_0)$ and the set $\{Z(t) < W(t_0) - \ell\}$. The same calculation has $-2A\,\frac{d}{dt}(\ell + \rho(t))^{-1}$ on the right-hand side and gives the same bound for the integral. Monotone convergence extends both integrations to the endpoints in time. Since $W(t_0) = Y_\varepsilon(t_0)$ and $|Z - \log u_\varepsilon| \le 2BT$, these estimates imply (7.32). □

The last ingredient is the lemma of Bombieri and Giusti [39], in its parabolic form [128] and in the version of [121, Lemma 5.2]: a mean value inequality for all small powers and a weak-type bound for the logarithm give a bound for the supremum.

Lemma 7.14 (Bombieri–Giusti)**.** *Let U_δ, $1/2 \le \delta \le 1$, be nested measurable sets in a finite measure space, with $|U_1| \le 1$, where $|\cdot|$ denotes the measure. Suppose that $f > 0$ is bounded on U_1, that $M, L \ge 1$, and that*

$$\operatorname*{ess\,sup}_{U_{\delta'}} f^p \le M(\delta - \delta')^{-\gamma} \int_{U_\delta} f^p\,, \qquad |\{\log f > \lambda\} \cap U_1| \le \frac{L}{\lambda} \tag{7.36}$$

for $1/2 \le \delta' < \delta \le 1$, $0 < p < p_0$, and $\lambda > 0$. Then

$$\log\big(\max\{1, \operatorname*{ess\,sup}_{U_{1/2}} f\}\big) \le C\,L\,M^2\,, \tag{7.37}$$

where C depends only on p_0 and γ.

Proof. This is a quantitative form of [121, Lemma 5.2]. Write $S(\delta) = \log\max\{e, \operatorname{ess\,sup}_{U_\delta} f\}$. Splitting the integral at $\log f = S(\delta)/2$ and using the second hypothesis gives

$$\int_{U_\delta} f^p \le \frac{2L}{S(\delta)}\,e^{pS(\delta)} + e^{pS(\delta)/2}\,.$$

If $S = S(\delta) > 2L$, choose $p = 2S^{-1}\log(S/(2L))$, for which the two terms are equal. For S above a constant depending only on p_0, this number lies in $(0, p_0)$, and the first hypothesis in (7.36) implies

$$\log \operatorname*{ess\,sup}_{U_{\delta'}} f \le \frac{S}{2}\left[1 + \frac{\log(2M(\delta - \delta')^{-\gamma})}{\log(S/(2L))}\right].$$

If $S \ge 2L[2M(\delta - \delta')^{-\gamma}]^2$, the right-hand side is at most $3S/4$. If this condition or the preceding threshold fails, then $S \le CLM^2(\delta - \delta')^{-2\gamma}$. In all cases,

$$S(\delta') \le \frac{3}{4}\,S(\delta) + C\,L\,M^2(\delta - \delta')^{-2\gamma}\,.$$

Iterate with $\delta_j = 1 - [2(j+1)]^{-1}$. The resulting series converges, because $\delta_{j+1} - \delta_j \asymp (j+1)^{-2}$ and $\sum_j (3/4)^j (j+1)^{4\gamma} < \infty$, and the terminal term tends to zero since f is bounded on U_1. □

Proof of Theorem 7.5. Normalize as in (7.19), with $t_0 = 0$, and put $u_\varepsilon = u + \varepsilon$ and $K = 1 + A + B$.

Step 1: the logarithmic estimate. Let ψ be the cutoff function of Lemma 7.8 for $B(x, 3/4) \subset B(x, 1)$, with a fixed value of its parameter. It satisfies $\int \psi^2 \, dm \asymp 1$, the weighted Poincaré inequality with the coefficient CA, and $\mathcal{E}(\psi, \psi) \le CB$, by (7.14) applied to a function equal to one on the annulus. Lemma 7.13 applies, with $T = 1$. Let $c_\varepsilon = Y_\varepsilon(0)$, and choose C_0 so that $2\,\mathcal{E}(\psi, \psi) / \int \psi^2 \, dm \le C_0 B$.

Step 2: the two functions. On the families U_δ^- and U_δ^+ of (7.25), respectively, set

$$f_- = (e^{-c_\varepsilon - C_0 B} \, u_\varepsilon)^{1/K} \,, \qquad f_+ = (e^{c_\varepsilon - C_0 B} \, u_\varepsilon^{-1})^{1/K} \,.$$

Since $\psi = 1$ on the spatial projections of these cylinders, (7.32) with $\ell = K\lambda$ gives

$$|\{\log f_\pm > \lambda\} \cap U_1^\pm| \le \frac{CA}{K\lambda} \le \frac{C}{\lambda} \,.$$

For $0 < p < c_*$, Lemma 7.12 applies to the powers $\pm p/K$ of u_ε, and gives the first hypothesis of Lemma 7.14 for $f_\pm$, with structural constants. After multiplication of the space-time measure by a fixed structural factor, the sets $U_1^\pm$ have measure at most one. Lemma 7.11 shows that f_- is bounded on the interior cylinders, and $u_\varepsilon \ge \varepsilon$ shows that f_+ is bounded.

Step 3: conclusion. Lemma 7.14 yields $\operatorname{ess\,sup}_{U_{1/2}^-} f_- \le e^C$ and $\operatorname{ess\,sup}_{U_{1/2}^+} f_+ \le e^C$, that is,

$$\operatorname*{ess\,sup}_{U_{1/2}^-} u_\varepsilon \le e^{c_\varepsilon + C_0 B + CK} \,, \qquad \operatorname*{ess\,inf}_{U_{1/2}^+} u_\varepsilon \ge e^{c_\varepsilon - C_0 B - CK} \,.$$

The number c_ε cancels in the quotient, which is at most $\exp(2C_0 B + 2CK) \le \exp(C'K)$. Since $U_{1/2}^\pm = Q_\pm$, letting $\varepsilon \downarrow 0$ gives (7.13). □

8 Heat kernel estimates

This chapter derives upper and lower bounds for the heat kernel of a strongly local regular Dirichlet form from volume doubling, the Poincaré inequality $\mathrm{PI}(\Psi_P)$ and the cutoff Sobolev inequality at the scale Ψ_S. The guiding principle, which the estimates make precise, is the following: *the Poincaré scale governs the on-diagonal upper bound, and the cutoff Sobolev scale governs the exponential decay away from the diagonal.*

The rate function. Off-diagonal bounds are expressed with the rate function of the sub-Gaussian estimates [10, equation (1.13) and Lemma 1.5]; see also [78, Section 3.3]. For an increasing scale function Ψ with a lower scaling exponent greater than one, set

$$\mathcal{A}_\Psi(R,t) = \sup_{s>0} \left(\frac{R}{s} - \frac{t}{\Psi(s)} \right) . \tag{8.1}$$

For a power $\Psi(s) = s^\beta$, this is a constant multiple of $(R^\beta/t)^{1/(\beta-1)}$, so that $\exp[-c\,\mathcal{A}_\Psi(d(x,y),t)]$ is the sub-Gaussian factor $\exp[-c\,(d(x,y)^\beta/t)^{1/(\beta-1)}]$, and the Gaussian factor for $\beta = 2$.

Conventions. The assertions on kernels hold for $m \otimes m$-almost every pair of points at each fixed time. On the diagonal we use the value given by the semigroup property,

$$p_t(x,x) = \int_X p_{t/2}(x,z)^2 \, dm(z) . \tag{8.2}$$

Results. Table 2 summarizes the main estimates. Each is of the form

$$p_t(x,y) \;\lesssim\; (\text{prefactor}) \times \exp\big[\mp(\text{decay})\big] .$$

The estimates with the volumes of balls in the prefactor require only volume doubling. Those with a function of the time alone require the uniform volume growth $\mathrm{V}(\Phi)$. When the two scales coincide, the upper and lower bounds match and are the classical two-sided sub-Gaussian estimates (Corollaries 8.12 and 8.22). When they differ, the upper and lower bounds are of different orders, and no claim of optimality is made.

Methods. The upper bounds are obtained by Davies' method [57]: the semigroup is conjugated by the exponential of a multiple of a cutoff function, and the cost of the conjugation is controlled by the cutoff Sobolev inequality. The lower bounds are obtained by chaining the parabolic Harnack inequality of Theorem 7.5 along geodesics, starting from a lower bound on the diagonal which comes either from a cutoff function, by the spectral theorem, or from the confinement of the heat in a ball. These methods are classical; what is specific here is that the two scales are kept separate throughout.

8.1 The heat kernel under volume doubling

Existence. The L^2 mean value inequality gives a heat kernel without any uniform lower bound for the volumes of balls.

Proposition 8.1. *Under the local hypotheses of Chapter 7, the semigroup has a symmetric measurable heat kernel. For every $t > 0$ and every compact set $K \subset X$, the map $P_t : L^2(X,m) \to L^\infty(K,m)$ is bounded. The kernel of the semigroup killed outside an open set has the same properties locally in that set.*

Statement	Hypotheses	Prefactor	Decay
Theorem 8.3 (upper)	global VD, $\mathrm{PI}(\Psi_P)$, $\mathrm{CSA}(\Psi_S)$	volumes of balls of an arbitrary radius r, polynomial in $\Psi_P(r)/t$, $t/\Psi_S(r)$	$\mathcal{A}_{\Psi_S}$
Corollary 8.12 (two-sided)	the same, $\Psi_P \asymp \Psi_S = \Psi$	$m(B(x, \Psi^{-1}(t)))^{-1}$	$\mathcal{A}_{\Psi}$
Theorem 8.6 (upper, small t)	local $\mathrm{V}(r^Q)$, $\mathrm{PI}(r^a)$, $\mathrm{CSA}(r^\beta)$	$t^{-Q/a}$	$(d^\beta/t)^{1/(\beta-1)}$
Theorem 8.14 (lower, small t)	the same	$t^{-2Q/(a+\beta)} \exp(-Ct^{-q})$	$(d^{b_-}/t)^{1/(b_--1)}$
Theorem 8.17 (upper)	global $\mathrm{V}(\Phi)$, $\mathrm{PI}(\Psi_P)$, $\mathrm{CSA}(\Psi_S)$	$\Phi(\Psi_P^{-1}(t))^{-1}$	$\mathcal{A}_{\Psi_S}$
Theorem 8.20 (lower)	the same	$\Phi(R_t)^{-1}$	variational, (8.55)
Corollary 8.22 (two-sided)	the same, $\Psi_P \asymp \Psi_S = \Psi$	$\Phi(\Psi^{-1}(t))^{-1}$	$\mathcal{A}_{\Psi}$

Table 2: Heat kernel estimates of Chapter 8. Here $d = d(x, y)$, and q, b_- and R_t are defined in (8.39), (8.42) and (8.52).

Proof. Fix a relatively compact ball $B = B(x, r)$ on which the hypotheses and their required dilations hold, and fix $t > 0$. Apply Lemma 7.11 to $s \mapsto P_s f$ on a cylinder with the spatial ball B and a time interval contained in $(t/2, 3t/2)$, whose smaller cylinder contains the time t. Since P_s is a contraction of L^2,

$$\|P_t f\|^2_{L^\infty(B(x,r/2))} \le \frac{C_{r,t}}{m(B(x,r))} \|f\|_2^2 .$$

For signed f, use $|P_s f| \le P_s|f|$. The constant is finite for each fixed r and t, and the continuity of $s \mapsto P_s f$ in L^2 makes the estimate valid at the time t. A finite cover gives the assertion on K.

The construction of the kernel is the argument of [121, Section 6, proof of Proposition 6.3]. Since $L^2(X, m)$ is separable, the local boundedness from L^2 to L^∞ and the Riesz representation theorem give a measurable kernel for $P_{t/2}$ whose rows belong to L^2. Composing $P_{t/2}$ with itself gives a nonnegative symmetric measurable kernel for P_t, with the semigroup property and (8.2). The same argument applies to the semigroup killed outside an open set: its orbits are caloric in the interior, and it is a contraction of L^2. □

In particular the rows $p_t(x, \cdot)$ of the kernel belong to $L^2(X, m)$ at positive times, and for $\varepsilon > 0$ the function $(t, y) \mapsto p_t(x, y)$ is, for almost every x, a weak solution of the heat equation on $(\varepsilon, \infty) \times X$. The arguments below which apply the Harnack inequality to the kernel use only this fact; no global ultracontractivity is needed.

Davies' perturbations. Let ψ be a bounded function. The kernel of the conjugated operator $e^{\psi} P_t e^{-\psi}$ is $e^{\psi(x)-\psi(y)} p_t(x, y)$. A bound on the norm of this operator therefore gives a bound on $p_t(x, y)$ with the factor $e^{\psi(y)-\psi(x)}$, which is small when ψ is a large multiple of a cutoff function equal to one near x and vanishing near y. The norm of the conjugated semigroup grows at a rate determined by $\int f^2 \, d\Gamma(\psi)$, and this is the quantity which the cutoff Sobolev inequality controls. On a Riemannian manifold one takes ψ Lipschitz, with $d\Gamma(\psi) \le |\nabla \psi|^2 \, dm$; when the energy measures

are singular this is not possible, and the functions ψ must be constructed from the cutoff functions of the hypothesis. This construction is due to Andres and Barlow [10, Lemma 5.1] and to Murugan and Saloff-Coste [130, Lemmas 2.1–2.2 and Proposition 2.3].

We state the construction for a scale function S satisfying (7.1), with an upper exponent $\beta_+ \geq 2$, and for a range of widths bounded by $r_0 \in (0, \infty]$. For every annulus

$$A = B(z, R+s) \setminus \overline{B(z,R)}\,, \qquad z \in X\,, \quad R > 0\,, \quad 0 < s \leq \min\{R, r_0\}\,,$$

assume that there is a cutoff function θ for $B(z,R) \subset B(z,R+s)$ such that

$$\int_X f^2 \, d\Gamma(\theta) \leq C_0 \int_A d\Gamma(f) + \frac{C_0}{S(s)} \int_A f^2 \, dm\,, \qquad f \in \mathcal{F}\,. \tag{8.3}$$

The constant is independent of the centre, of the inner radius and of the width. This is the hypothesis (7.8) with $S = \Psi_S$. The uniformity in the inner radius is what allows upper bounds at arbitrary distances; only the width of the annuli is restricted by r_0.

Lemma 8.2 (Perturbations)**.** *Assume* (8.3)*. Fix $z \in X$, $r > 0$, and $\lambda \geq \max\{1, r/r_0\}$, with $r/r_0 = 0$ if $r_0 = \infty$. Let*

$$\rho(y) = \left(2 - \frac{d(z,y)}{r}\right)_+ \wedge 1\,.$$

For every $p \geq 1$ there is a cutoff function $\varphi_p \in \mathcal{F}$ for $B(z,r) \subset B(z,2r)$ such that $\psi_p = \lambda\varphi_p$ satisfies

$$\|\psi_p - \lambda\rho\|_\infty \leq \frac{C}{p}\,, \tag{8.4}$$

$$\mathcal{E}(e^{\pm\psi_p} f^{2p-1}, e^{\mp\psi_p} f) \geq \frac{1}{2p}\,\mathcal{E}(f^p, f^p) - \frac{C p^{\beta_+ - 1}}{S(r/\lambda)}\,\|f\|_{2p}^{2p} \tag{8.5}$$

for every bounded nonnegative $f \in \mathcal{F}$. The constant depends only on C_0 and on the upper scaling constants of S.

The function $\lambda\rho$ is the Lipschitz profile which one would use on a manifold. The functions ψ_p approximate it uniformly, with an error which improves as the exponent p increases. The inequality (8.5) is the lower bound for the perturbed form which drives the differential inequalities for the L^{2p} norms of the conjugated semigroup.

Proof. We use the averaging construction of [130, Lemma 2.1]. Choose an integer $N \asymp p\lambda$, with a sufficiently large fixed constant of comparison from below. Subdivide the annulus between the radii r and $2r$ into N annuli of width $r/N \leq r_0$, let θ_i be the cutoff functions given by (8.3), and set $\varphi_p = N^{-1} \sum_{i=1}^N \theta_i$. By strong locality, $\Gamma(\varphi_p) = N^{-2} \sum_i \Gamma(\theta_i)$: the energy measures are carried by disjoint annuli. Summing the inequalities (8.3) gives

$$\int_X g^2 \, d\Gamma(\varphi_p) \leq \frac{C_0}{N^2}\,\mathcal{E}(g,g) + \frac{C_0}{N^2 S(r/N)}\,\|g\|_2^2\,, \qquad g \in \mathcal{F}\,. \tag{8.6}$$

On each of the N annuli, all but one of the cutoff functions are equal to zero or to one. Thus $\|\varphi_p - \rho\|_\infty \leq N^{-1}$, which proves (8.4).

Let $\psi = \psi_p$ and $u = f^p$. The calculation of [130, Lemma 2.2], which uses the chain and Leibniz rules, gives

$$\begin{aligned}\mathcal{E}(e^{\psi} f^{2p-1}, e^{-\psi} f) &= \frac{2p-1}{p^2}\,\mathcal{E}(u,u) - \frac{2(p-1)}{p}\int_X u\,d\Gamma(u,\psi) - \int_X u^2\,d\Gamma(\psi) \\ &\geq \frac{1}{p}\,\mathcal{E}(u,u) - p\int_X u^2\,d\Gamma(\psi)\,.\end{aligned}$$

Apply (8.6) with $g = u$ and $\psi = \lambda\varphi_p$. The choice of N makes the coefficient $p\lambda^2 C_0 N^{-2}$ of the energy at most $1/(2p)$. Since $r/N \leq r/\lambda \leq r_0$, the upper scaling of S gives

$$\frac{p\lambda^2}{N^2 S(r/N)} \leq \frac{C}{pS(r/N)} \leq \frac{Cp^{\beta_+ - 1}}{S(r/\lambda)}\,.$$

This proves (8.5), for either sign of ψ. The products belong to $\mathcal{F}$ because f is bounded and by (8.6). □

For $p = 1$, (8.5) gives $\frac{d}{ds}\|e^{\psi} P_s e^{-\psi} f\|_2^2 \leq \frac{C}{S(r/\lambda)}\|e^{\psi} P_s e^{-\psi} f\|_2^2$, with $\psi = \psi_1$, and therefore

$$\|e^{\psi_1} P_s\, e^{-\psi_1}\|_{2\to 2} \leq \exp\Big[\frac{Cs}{S(r/\lambda)}\Big] \qquad (s > 0)\,. \tag{8.7}$$

This is the quadratic perturbation estimate of [10, Proposition 2.3].

The upper bound. Applying the mean value inequality in each of the two variables of the kernel bounds its value by an average in space and time, with a polynomial dependence on the two scale coefficients. The estimate (8.7) for this average gives the decay away from the diagonal. The prefactor involves the volumes of the averaging balls, whose radius r can be chosen independently of the time.

Theorem 8.3. *Suppose that X is unbounded and geodesic, that* VD, PI(Ψ_P) *and the cutoff Sobolev inequality* (7.8) *hold globally, and that Ψ_S has a lower scaling exponent greater than one. There are structural constants $c, C, \kappa > 0$ such that, for every $r, t > 0$ and almost every $x, y \in X$,*

$$p_t(x,y) \leq \frac{C}{\sqrt{m(B(x,r))\,m(B(y,r))}}\left(1 + \frac{\Psi_P(r)}{t}\right)^{\kappa}\left(1 + \frac{t}{\Psi_S(r)}\right)^{\kappa} \times \exp\left[-c\,\mathcal{A}_{\Psi_S}\big((d(x,y) - 8r)_+, t\big)\right]\,. \tag{8.8}$$

Proof. Step 1: the kernel is bounded by an average. By Lemma 7.11 with $q = 1$, after rescaling, a nonnegative caloric function u on $(a - T, a) \times B(z, 2r)$ satisfies

$$\operatorname*{ess\,sup}_{B(z,r)} u(a,\cdot) \leq \frac{C\,(1 + \Psi_P(r)/T)^{\theta-1}(1 + T/\Psi_S(r))^{\theta}}{T\,m(B(z,r))}\int_{a-T}^{a}\int_{B(z,2r)} u(s,v)\,dm(v)\,ds\,,$$

where θ is as in (7.20); the value at the terminal time is obtained from times tending to a and from the continuity in L^2. Apply this estimate successively in the two variables of the symmetric kernel: first at the time t with $T = t/2$, and then at each time $s \in (t/2, t)$ with $T = t/4$. The times which occur lie in $(t/4, t)$, where the rows of the kernel are orbits of the semigroup with initial data in L^2, by Proposition 8.1. Doubling and the scaling of Ψ_P and Ψ_S yield

$$p_t(x,y) \leq \frac{C\,(1 + \Psi_P(r)/t)^{2(\theta-1)}(1 + t/\Psi_S(r))^{2\theta}}{m(B(x,r))\,m(B(y,r))}\sup_{t/4\leq s\leq t}\langle \mathbf{1}_{B(x,2r)}, P_s \mathbf{1}_{B(y,2r)}\rangle\,. \tag{8.9}$$

To obtain this outside one null set for each fixed time, first take the two outer balls in a countable family with dense centres and rational radii, and evaluate at almost every pair of points of the corresponding inner balls. For given x, y and r one can choose such inner balls containing x and y whose outer balls are contained in $B(x,2r)$ and $B(y,2r)$, with radii comparable to r; the volumes and the scale coefficients change by structural factors.

Step 2: the average. Write $d = d(x,y)$. For $d \le 8r$, the contraction property in L^2 bounds the inner product in (8.9) by $C\sqrt{m(B(x,r))\,m(B(y,r))}$. For $d > 8r$ set $R = d/3$ and apply Lemma 8.2 with $p = 1$, the centre x, the radius R, the scale $S = \Psi_S$ and $\lambda \ge 1$. The function $\psi = \psi_1$ is equal to λ on $B(x,2r) \subset B(x,R)$ and to zero on $B(y,2r)$, which lies outside $B(x,2R)$. By (8.7), for all $s \le t$,

$$\begin{aligned}\langle \mathbf{1}_{B(x,2r)}, P_s \mathbf{1}_{B(y,2r)}\rangle &= e^{-\lambda}\langle \mathbf{1}_{B(x,2r)}, e^{\psi} P_s e^{-\psi}\mathbf{1}_{B(y,2r)}\rangle \\ &\le C\sqrt{m(B(x,r))\,m(B(y,r))}\,\exp\left[-\lambda + \frac{Ct}{\Psi_S(R/\lambda)}\right].\end{aligned}$$

Step 3: optimization. Taking the infimum over $\lambda \ge 1$, with $s = R/\lambda$, gives the factor $\exp[-\sup_{s\le R}(R/s - Ct/\Psi_S(s))]$. The scales $s > R$ contribute at most one to the supremum in (8.1), so this factor is at most $C\exp[-\mathcal{A}_{\Psi_S}(R,Ct)]$. By the lower scaling of Ψ_S, $\mathcal{A}_{\Psi_S}(R,Ct) \ge c\,\mathcal{A}_{\Psi_S}(d,t)$; this comparison is proved at the end of the proof of Theorem 8.17. Inserting the result in (8.9) gives (8.8), with $\kappa = 2\theta$. □

Matching scales. Suppose that $\Psi_P \asymp \Psi_S = \Psi$, and choose $r = \Psi^{-1}(t)$ in (8.8). The polynomial factor is bounded. If $d(x,y) \le 16r$, the rate $\mathcal{A}_\Psi(d(x,y),t)$ is bounded by a structural constant; if $d(x,y) > 16r$, then $d(x,y) - 8r \ge d(x,y)/2$ and the scaling of Ψ compares the two rates. Hence

$$p_t(x,y) \le \frac{C\exp[-c\,\mathcal{A}_\Psi(d(x,y),t)]}{\sqrt{m(B(x,\Psi^{-1}(t)))\,m(B(y,\Psi^{-1}(t)))}}\,. \tag{8.10}$$

This is the upper estimate of the theory with one scale; compare [10, Theorem 1.12] and [121, Proposition 5.8]. The matching lower estimate is proved in Corollary 8.12.

8.2 Uniform volume growth and Nash inequalities

This section assumes the uniform volume growth $\mathrm{V}(\Phi)$. Summing the Poincaré inequality over a covering by balls of radius r gives an inequality of Nash type, and hence a bound for the norm of the semigroup from L^1 to L^∞. This is the method of [140, Lemmas 2.3–2.4 and proof of Theorem 2.1], with r^2 replaced by $\Psi_P(r)$. The Poincaré inequality may have a dilation $\Lambda \ge 1$ here, and neither conservativeness nor irreducibility is needed.

Proposition 8.4 (Nash inequality)**.** *Suppose that, uniformly in x and for $0 < r \le r_0$,*

$$m(B(x,r)) \asymp \Phi(r)\,, \qquad \int_{B(x,r)} |f - f_{B(x,r)}|^2\,dm \le C\,\Psi_P(r)\int_{B(x,\Lambda r)} d\Gamma(f)\,,$$

where Φ and Ψ_P are increasing continuous functions satisfying (7.1). There exist $r_* > 0$ and $C < \infty$ such that

$$\|f\|_2^2 \le C\,\Psi_P(r)\,\mathcal{E}(f,f) + \frac{C}{\Phi(r)}\,\|f\|_1^2\,, \qquad 0 < r \le r_*\,, \quad f \in \mathcal{F}\cap L^1(X,m)\,. \tag{8.11}$$

For $\Phi(r) = r^Q$ and $\Psi_P(r) = r^a$, where $Q, a > 0$, this implies

$$\|f\|_2^{2+2a/Q} \le C(\mathcal{E}(f,f) + r_*^{-a}\|f\|_2^2)\|f\|_1^{2a/Q}\,. \tag{8.12}$$

If the hypotheses hold at every radius on an unbounded space, then (8.11) holds for every $r > 0$, and the term $r_*^{-a}\|f\|_2^2$ in (8.12) can be omitted.

Proof. Choose a maximal r-separated family (x_i). The balls $B_i = B(x_i, r)$ cover X, and the balls $B(x_i, \Lambda r)$ have uniformly bounded overlap, by Lemma 3.2. Here r_* is chosen so that all the required dilations have radius at most r_0. Write $\|f\|_2^2 \le 2\sum_i \int_{B_i} |f - f_{B_i}|^2\,dm + 2\sum_i m(B_i)|f_{B_i}|^2$. By the Poincaré inequality and the bounded overlap, the first sum is at most $C\Psi_P(r)\mathcal{E}(f,f)$. For the second,

$$\sum_i m(B_i)|f_{B_i}|^2 \le \frac{C}{\Phi(r)}\sum_i \left(\int_{B_i} |f|\,dm\right)^2 \le \frac{C}{\Phi(r)}\,\|f\|_1^2\,.$$

This proves (8.11).

In the power case, choose r so that $Cr^{-Q}\|f\|_1^2 = \frac{1}{2}\|f\|_2^2$, whenever this radius is at most r_*. Absorbing the second term gives (8.12) without the term $r_*^{-a}\|f\|_2^2$. If the selected radius exceeds r_*, then $\|f\|_2^2 \le Cr_*^{-Q}\|f\|_1^2$, and this term alone proves (8.12). The same choice is always available when the assumptions hold at every radius. □

The passage from a Nash inequality to a bound for the semigroup is Nash's argument [133]; for general profiles, see [51, Proposition 3.1 and Theorem 3.4]. We include it because the volume and Poincaré functions need not be powers.

Proposition 8.5 (Ultracontractivity)**.** *Under the hypotheses of Proposition 8.4, the semigroup admits a symmetric measurable heat kernel, and for some $t_* > 0$,*

$$\|P_t\|_{1\to\infty} \le \frac{C}{\Phi(\Psi_P^{-1}(t))}\,, \qquad 0 < t \le t_*\,. \tag{8.13}$$

Consequently, for $1 \le p \le q \le \infty$,

$$\|P_t\|_{p\to q} \le C\,\Phi(\Psi_P^{-1}(t))^{-(1/p-1/q)}\,. \tag{8.14}$$

If the hypotheses hold at every radius, these conclusions hold for every $t > 0$. In the power case the bound is $Ct^{-Q/a}$.

Proof. Let $f \in L^1 \cap L^2$ with $\|f\|_1 \le 1$, and set $y(t) = \|P_t f\|_2^2$. Then $\|P_t f\|_1 \le 1$, because the semigroup is sub-Markovian and symmetric, and $y'(t) = -2\mathcal{E}(P_t f, P_t f)$. For $y(t)$ above a fixed threshold, insert in (8.11) the radius determined by $\Phi(r) = 2C/y(t)$. Then

$$-y'(t) \ge \frac{c\,y(t)}{H(y(t))}\,, \qquad H(v) = \Psi_P\big(\Phi^{-1}(2C/v)\big)\,.$$

The scaling of Φ and Ψ_P gives $H(2^j v) \le C2^{-j\epsilon}H(v)$ for some $\epsilon > 0$, and hence

$$\int_v^\infty \frac{H(s)}{s}\,ds \le C\,H(v)\,.$$

Integrating the differential inequality gives $ct \le CH(y(t))$, that is, $y(t) \le C/\Phi(\Psi_P^{-1}(t))$ for $0 < t \le t_*$. This is a bound for $\|P_t\|_{1\to 2}^2$. By duality and the semigroup property, $\|P_t\|_{1\to\infty} \le \|P_{t/2}\|_{1\to 2}^2$, which gives (8.13). A bounded positive operator from L^1 to L^∞ on a separable measure space has a measurable kernel; see [51, Section 2.1]. Symmetry and the semigroup property give the symmetric heat kernel. Interpolation with the contraction property in L^p proves (8.14). □

8.3 Upper bounds at small times

Under uniform volume growth, the prefactor of the upper bound can be taken to be the ultracontractive bound of Proposition 8.5, in place of the volumes of two balls. The proof is Davies' method [57] in the formulation with Nash inequalities of Carlen, Kusuoka and Stroock [48, Section 3], adapted to singular energy measures by Murugan and Saloff-Coste [130]. The norm from L^1 to L^∞ of the conjugated semigroup $e^{\psi} P_t e^{-\psi}$ is estimated by iterating differential inequalities for its L^q norms, with $q = 2^{k+1}$, as in the proof of ultracontractivity from the Nash inequality. The Nash inequality involves the Poincaré scale, and the cost of the conjugation is estimated with the perturbations of Lemma 8.2, which involve the cutoff Sobolev scale. A difficulty specific to singular energy measures is that the function ψ_p of Lemma 8.2 depends on the exponent p. It is resolved by comparing the semigroups conjugated by two successive functions [130, equation (16) and Section 2.3]; this is the comparison (8.19) below.

In this section we treat power scales at small times. Under the local hypotheses $m(B(x,r)) \asymp r^Q$ and $\mathrm{PI}(r^a)$, Propositions 8.4 and 8.5, with $D = Q/a$, give

$$\|f\|_2^{2+2/D} \le C_N(\mathcal{E}(f,f) + \|f\|_2^2)\|f\|_1^{2/D}\,, \qquad f \in \mathcal{F} \cap L^1\,, \tag{8.15}$$

and a symmetric measurable heat kernel with

$$\|P_t\|_{1\to\infty} \le C t^{-D}\,, \qquad 0 < t \le 1\,. \tag{8.16}$$

The constant r_*^{-a} of (8.12) is absorbed in C_N, and the semigroup property extends the bound at small times to any fixed bounded interval of times, with a changed constant.

Theorem 8.6. *Assume the Nash inequality (8.15) with $D > 0$ and the cutoff Sobolev inequality (8.3) with $S(s) = s^\beta$, where $\beta \ge 2$. There are $C, c, t_0 > 0$ such that*

$$p_t(x,y) \le C\, t^{-D} \exp\left[-c\left(\frac{d(x,y)^\beta}{t}\right)^{1/(\beta-1)}\right], \qquad 0 < t \le t_0\,. \tag{8.17}$$

In particular, under the local hypotheses $m(B(x,r)) \asymp r^Q$ and $\mathrm{PI}(r^a)$, the prefactor is $t^{-Q/a}$.

Proof. We adapt the iteration of [130, Section 2.3], allowing a Nash exponent which is independent of β and a zeroth-order term in the Nash inequality.

Step 1: the conjugated semigroups. Fix $z \in X$, $r > 0$ and $\lambda \ge \max\{1, r/r_0\}$, and apply Lemma 8.2 with $S(s) = s^\beta$ and $\beta_+ = \beta$. Put $K = (\lambda/r)^\beta$, $p_k = 2^k$, $q_k = 2p_k$, and $\psi_k = \psi_{p_k}$. By (8.4),

$$\|\psi_k - \psi_{k-1}\|_\infty \le \frac{C}{q_k} \quad (k \ge 1)\,, \qquad \psi_k \longrightarrow \lambda\rho \quad \text{in } L^\infty\,. \tag{8.18}$$

For $f \ge 0$ in $L^2 \cap L^\infty$, write

$$v_k(s) = e^{\psi_k} P_s(e^{-\psi_k} f)\,, \qquad U_k(s) = \|v_k(s)\|_{q_k}\,.$$

Since P_s preserves positivity, (8.18) implies

$$e^{-C/q_k}\, v_{k-1}(s) \le v_k(s) \le e^{C/q_k}\, v_{k-1}(s)\,. \tag{8.19}$$

This comparison is what allows the function ψ_k to depend on k.

Step 2: differential inequalities. For $k = 0$, (8.5) gives, as in (8.7),

$$U_0(s) \le e^{CKs} \|f\|_2 \,. \tag{8.20}$$

For $k \ge 1$, differentiate $U_k^{q_k}$. The inequality (8.5) and the Nash inequality (8.15), applied to $v_k^{q_k/2}$, give

$$\frac{d}{ds} U_k^{q_k} \le -c\, U_k^{q_k(1+1/D)} \, \|v_k\|_{q_k/2}^{-q_k/D} + C(1 + q_k^\beta K)\, U_k^{q_k} \,,$$

where the term 1 comes from the zeroth-order term of the Nash inequality. If $U_{k-1}(s) \le M$ on an interval, then $\|v_k\|_{q_k/2} \le e^{C/q_k} M$ by (8.19), and therefore

$$U_k'(s) \le -\frac{c}{q_k}\, M^{-q_k/D}\, U_k(s)^{1+q_k/D} + C\, q_k^{\beta-1}(1+K)\, U_k(s) \,. \tag{8.21}$$

The constants are independent of k: the factor e^{C/q_k} is raised to the power q_k/D and remains bounded. If $M = 0$, then $v_k = 0$ on the interval and there is nothing to prove.

Step 3: integration of the differential inequality. Suppose that

$$U' \le -\frac{c}{q}\, M^{-q/D}\, U^{1+q/D} + H\, U$$

on an interval of length δ. Then, at its right endpoint,

$$U \le \Big(\frac{C}{\delta}\Big)^{D/q} M\, e^{H\delta} \,. \tag{8.22}$$

Indeed, with s measured from the left endpoint, the function $W = (e^{-Hs}U)^{-q/D}$ satisfies $W' \ge \frac{c}{D}\, M^{-q/D}$ on the interval, and $W \ge 0$ at the left endpoint; compare [130, Lemma 2.4]. No bound for U at the left endpoint is needed.

Step 4: iteration. Fix $t > 0$. Choose

$$\delta_k = c_\beta\, t\, 2^{-k\beta} \,, \qquad \sum_{k\ge 1} \delta_k = \frac{t}{2} \,, \qquad t_k = \sum_{i=1}^{k} \delta_i \,.$$

Let $M_k = \sup_{t_k \le s \le t} U_k(s)$, and let $M_0 = \sup_{0\le s\le t} U_0(s)$. For $s \in [t_k, t]$, apply (8.22) on the interval $[s - \delta_k, s]$, which lies in $[t_{k-1}, t]$. We obtain

$$M_k \le \Big(\frac{C}{\delta_k}\Big)^{D/q_k} \exp\big[C\, q_k^{\beta-1}(1+K)\, \delta_k\big]\, M_{k-1} \,.$$

Since

$$\sum_{k\ge 1} \frac{1}{q_k} = \frac{1}{2} \,, \qquad \sum_{k\ge 1} \frac{k}{q_k} < \infty \,, \qquad \sum_{k\ge 1} q_k^{\beta-1}\delta_k \le C_\beta\, t \,,$$

iteration and (8.20) give

$$\limsup_{k\to\infty} U_k(t) \le C\, t^{-D/2}\, e^{C(1+K)t}\, \|f\|_2 \,.$$

By the uniform convergence of the functions ψ_k in (8.18), and since a function whose L^{q_k} norms are bounded as $q_k \to \infty$ is bounded by the same constant, this implies

$$\|e^{\lambda\rho} P_t\, e^{-\lambda\rho}\|_{2\to\infty} \le C\, t^{-D/2}\, e^{C(1+K)t} \,. \tag{8.23}$$

Approximation removes the boundedness of f. The differentiations above are first made for bounded initial data in the domain of the form and for regularized times; multiplication by $e^{\pm\psi_k}$ preserves the domain of the form by (8.6), and the chain rule gives the differential inequalities. Truncation and approximation by the semigroup extend them to the data used in the iteration, as in [130, Section 2.3].

Step 5: the kernel. Replacing all the functions ψ_k by their negatives gives (8.23) with $-\lambda\rho$. By duality, the corresponding estimate from L^1 to L^2 holds with $\lambda\rho$. Composition at the time $t/2$ yields

$$\|e^{\lambda\rho} P_t\, e^{-\lambda\rho}\|_{1\to\infty} \le C\, t^{-D} \exp\bigl[Ct + Ct(\lambda/r)^\beta\bigr]\,.$$

In particular, for $x \in B(z,r)$ and $y \notin B(z,2r)$,

$$p_t(x,y) \le C\, t^{-D} \exp\bigl[Ct + Ct(\lambda/r)^\beta - \lambda\bigr]\,. \tag{8.24}$$

Step 6: optimization. Take r comparable to $d(x,y)$, with x in the inner ball and y outside the outer ball. When $d(x,y)^\beta/t$ is larger than a fixed constant, choose in (8.24)

$$\lambda = c_1 \left(\frac{r^\beta}{t}\right)^{1/(\beta-1)}$$

with a sufficiently small fixed c_1, so that $Ct(\lambda/r)^\beta \le \lambda/2$. The constraint $\lambda \ge 1$ is satisfied after adjusting the threshold. If $r \le r_0$, this is the only constraint. If $r > r_0$, the additional condition $\lambda \ge r/r_0$ holds once $t \le t_0$, where $t_0 > 0$ is a sufficiently small fixed multiple of r_0^β. This proves the exponential bound. In the remaining region $d(x,y)^\beta/t \le C$, use (8.16).

Step 7: exceptional sets. Make the construction for a countable dense set of centres, rational radii, and rational admissible values of λ. For each fixed t, discard the union of the corresponding null sets. Given distinct x and y, one can choose such a centre near x and a rational radius between fixed multiples of $d(x,y)$ so that x lies in the inner ball and y lies outside the outer ball, and rational values of λ approximate the optimal one. Thus (8.17) holds almost everywhere for each t. □

Remark 8.7. If cutoff functions are only known for annuli contained in a prescribed open set, the argument gives the corresponding localized estimate, not an upper bound at arbitrary distances in X. Cutoff functions of small width at arbitrary inner radii, with uniform constants, are what allows the subdivision argument of Lemma 8.2 to handle all distances on an unbounded space. None of the arguments for upper bounds uses conservativeness.

Conservativeness. The semigroup is *conservative* if $P_t 1 = 1$ for every $t > 0$. The following consequence of a criterion of Andres and Barlow [10, Theorem 1.16(a)] requires a cutoff Sobolev inequality at one fixed width and a polynomial bound for the volume growth at infinity. The Poincaré scale plays no role.

Proposition 8.8. *Assume (7.8) at a fixed positive width and at all smaller widths, with constants uniform in the inner radius. If, for some $o \in X$ and $M < \infty$,*

$$m(B(o,R)) \le C(1+R)^M\,, \qquad R \ge 1\,, \tag{8.25}$$

then the semigroup is conservative. On a compact space every strongly local regular Dirichlet form is conservative.

Proof. Fix $0 < \rho \le r_0$ and let $D_j = B(o, j\rho)$. Lemma 7.8, with the coefficient $1/8$, gives cutoff functions χ_j for $D_j \subset D_{j+1}$ satisfying

$$\int f^2 \, d\Gamma(\chi_j) \le \frac{1}{8} \int \chi_j^2 \, d\Gamma(f) + \frac{C}{\Psi_S(\rho)} \int_{D_{j+1} \setminus D_j} f^2 \, dm \,.$$

Thus the condition $\mathrm{CSD}(D_j, D_{j+1}, \theta_j)$ of [10, Definition 1.9] holds with $\theta_j = C\Psi_S(\rho)^{-1}$ independent of j. Since X is proper, the sets D_j are relatively compact, and (8.25) gives

$$4^{-j} \, \theta_j \, m(D_{j+1} \setminus D_j) \longrightarrow 0 \,.$$

The criterion [10, Theorem 1.16(a)] applies. Its hypotheses involve this cutoff condition and this exhaustion; no comparison with the intrinsic metric is required.

If X is compact, the regularity of the form and a normal contraction of a continuous function of $\mathcal{F}$ which approximates 1 uniformly show that $1 \in \mathcal{F}$. Strong locality gives $\mathcal{E}(1,1) = 0$, hence $P_t 1 = 1$. □

Confinement and diagonal lower bounds. An upper bound for the heat kernel away from the diagonal implies that, for a conservative semigroup, most of the mass of $p_t(x, \cdot)$ lies in a ball. By the Cauchy–Schwarz inequality this gives a lower bound on the diagonal; this is the argument of Coulhon and Grigor'yan [55, Lemma 6.2]. When the prefactor and the exponential decay of the upper bound are governed by different scales, the radius of the ball must be enlarged by a logarithmic factor; compare the logarithmic radius in [55, Theorem 6.1].

Proposition 8.9. *Suppose that (8.17) holds with $D = Q/a$, $0 < a \le \beta$, and that $m(B(x,R)) \le CR^Q$ at all radii; on a compact space it suffices to assume this up to the diameter. For $R \ge t^{1/\beta}$,*

$$\int_{X \setminus B(x,R)} p_t(x,y) \, dm(y) \le \min \left\{ 1, \, C \, t^{Q/\beta - Q/a} \exp \left[-c \left(\frac{R^\beta}{t} \right)^{1/(\beta-1)} \right] \right\} . \tag{8.26}$$

If the semigroup is conservative, then

$$p_t(x,x) \ge c \, t^{-Q/\beta} \, [\log(e/t)]^{-Q(\beta-1)/\beta} \,. \tag{8.27}$$

When $a = \beta$, the logarithm can be omitted. These assertions hold for almost every x, with the diagonal values (8.2).

Proof. Put $R_t = t^{1/\beta}$ and $\eta = \beta/(\beta-1)$. Sum (8.17) over the annuli between the radii $2^k R$ and $2^{k+1} R$. The bound on the volume gives

$$\int_{X \setminus B(x,R)} p_t(x,y) \, dm(y) \le C \, t^{-Q/a} \, R_t^Q \sum_{k \ge 0} \Big(\frac{2^k R}{R_t} \Big)^Q e^{-c(2^k R/R_t)^\eta} \,.$$

For $R/R_t \ge 1$, the sum is at most $C \exp[-c'(R/R_t)^\eta]$. The bound by one holds because the semigroup is sub-Markovian. This proves (8.26).

Choose

$$R = K \, t^{1/\beta} \, [\log(e/t)]^{(\beta-1)/\beta} \,. \tag{8.28}$$

For a sufficiently large fixed K, the second quantity in (8.26) is at most $1/2$. Conservativeness gives $\int_{B(x,R)} p_t(x,y) \, dm(y) \ge 1/2$, and hence, by the Cauchy–Schwarz inequality,

$$p_{2t}(x,x) = \int_X p_t(x,y)^2 \, dm(y) \ge \frac{1}{4 \, m(B(x,R))} \ge c \, R^{-Q} \,. \tag{8.29}$$

Substitute (8.28) and change the time by a fixed factor. If $a = \beta$, the polynomial factor in (8.26) is equal to one, and $R = Kt^{1/\beta}$ suffices. □

8.4 Lower bounds

Lower bounds for the heat kernel are obtained in two steps. First, a lower bound on the diagonal, or near the diagonal, at a suitable radius. Second, the propagation of this bound to arbitrary pairs of points by a chain of applications of the Harnack inequality. We begin with the second step, which is common to all the lower bounds of this chapter.

Harnack chains. The following lemma adapts the chaining of [18, Proposition 5.2]; for the corresponding chaining of the semigroup property and the optimization for general scale functions, see [78, Theorem 6.5 and Corollary 6.11]. The cost of a chain is the sum of the Harnack constants of its links, and the two scales contribute to it separately.

Lemma 8.10 (Chain lemma)**.** *Suppose that X is geodesic and that the Harnack inequality (7.13) holds uniformly for every $T > 0$ and all radii in a specified range. Assume that Ψ_P and Ψ_S satisfy (7.1) in that range and that $\Psi_S \le C\Psi_P$. Let $s < v$, $\delta = v - s$, and let $u \ge 0$ be caloric on $(s-\delta, v+\delta) \times X$. For every radius r whose fixed dilations remain in the range,*

$$u(s,x) \le \exp\left[C\left\{\frac{\Psi_P(r)}{\delta}\left(1+\frac{d(x,y)}{r}\right)^2 + \frac{\delta}{\Psi_S(r)}\right\}\right] u(v,y) \tag{8.30}$$

for almost every pair (x,y), where $u(s,\cdot)$ and $u(v,\cdot)$ are the values of the representative which is continuous in time with values in L^2_{loc}. If the semigroup has a measurable kernel whose rows are locally bounded in L^2, as in Proposition 8.1, then, for every $t > 0$,

$$p_t(x,y) \ge \exp\left[-C\left\{\frac{\Psi_P(r)}{t}\left(1+\frac{d(x,y)}{r}\right)^2 + \frac{t}{\Psi_S(r)}\right\}\right] p_{2t/3}(y,y) \tag{8.31}$$

almost everywhere, and the radius can be optimized on a common set of full measure. The same argument applies to the semigroup killed outside an open set, if a geodesic from x to y and all the enlarged balls of the chain lie in this open set.

Proof. Step 1: the chain. Choose an integer $N \asymp 1 + d(x,y)/r$, with a sufficiently large fixed constant of comparison. Subdivide a geodesic from x to y into N links of length at most a fixed small multiple of r, and divide δ into N equal increments. Apply (7.13) on N cylinders of radius comparable to r and time length comparable to δ/N, centred at the points of the subdivision. These cylinders lie in $(s-\delta, v+\delta) \times X$, and the inner balls of consecutive cylinders intersect in balls of radius comparable to r. Multiplying the N inequalities gives a constant whose logarithm is at most

$$C\left(N + \frac{N^2\,\Psi_P(r)}{\delta} + \frac{\delta}{\Psi_S(r)}\right) \le C\left(\frac{N^2\,\Psi_P(r)}{\delta} + \frac{\delta}{\Psi_S(r)}\right),$$

since the sum of the last two terms is at least $2N\sqrt{\Psi_P(r)/\Psi_S(r)} \ge cN$. This proves (8.30).

Step 2: fixed times and exceptional sets. The Harnack inequality compares essential suprema and infima over cylinders in space and time. For fixed times, approximate by times at which the comparison holds, use the continuity in L^2_{loc}, and extract subsequences which converge almost everywhere. Choose the balls in a countable family with dense centres and dyadic radii. Consecutive balls intersect in sets of positive measure, so the essential comparisons propagate along the chain without evaluating the solution at prescribed points. Countable unions of null sets, and Fubini's theorem for the rows of the kernel, justify all these comparisons simultaneously. By (7.1), an

arbitrary radius can be replaced by a dyadic one at the cost of a fixed factor, which allows the optimization in r.

Step 3: the kernel. The rows of the kernel are weak solutions away from the time zero. Apply (8.30) to $u(\tau,\cdot) = p_\tau(\cdot, w)$ between the times $t/3$ and $2t/3$, with the two points interchanged. If A denotes the exponent in (8.30), this gives

$$p_{2t/3}(x,w) \geq e^{-A}\, p_{t/3}(y,w)\,. \tag{8.32}$$

Multiply by $p_{t/3}(y,w)$ and integrate in w. The symmetry of the kernel and the semigroup property give (8.31), with the diagonal value (8.2). □

A near-diagonal lower bound from a cutoff function. The next lemma gives a lower bound for the heat kernel in a ball, without assuming conservativeness. A cutoff function χ of small energy gives a lower bound for the average $\langle P_s\chi, \chi\rangle$, by the spectral theorem [55, equation (2.1) and Proposition 2.1]; the Harnack inequality, applied in the two variables of the kernel, converts this average into a pointwise bound.

Lemma 8.11. *Assume that X is proper and geodesic and that the local hypotheses of Chapter 7 hold. Then there are structural constants $c, C, K > 0$ such that*

$$p_t(y,z) \geq \frac{c}{m(B(x,r))} \exp\left[-C\left(1 + \frac{\Psi_P(r)}{t} + \frac{t}{\Psi_S(r)}\right)\right], \qquad y,z \in B(x,r)\,, \quad t>0\,, \tag{8.33}$$

for $m\otimes m$-almost every pair, whenever all the fixed dilations of the balls which occur are in the range of the hypotheses. The same estimate holds for the kernel of the semigroup killed outside $B(x,Kr)$. With the diagonal values (8.2), the bound holds for $y = z$, for almost every $y \in B(x,r)$, for both kernels.

Proof. Step 1: the average. By (7.9), there is a cutoff function $\chi \in \mathcal{F}_0(B(x,2r))$, with $0 \leq \chi \leq 1$, equal to one on $B(x,r)$, such that

$$\|\chi\|_2^2 \asymp m(B(x,r))\,, \qquad \|\chi\|_1 \asymp m(B(x,r))\,, \qquad \mathcal{E}(\chi,\chi) \leq C\,\frac{m(B(x,r))}{\Psi_S(r)}\,.$$

Let $D = B(x,Kr)$, where K is a sufficiently large structural constant, and let P_s^D be the semigroup killed outside D, with kernel p_s^D. Jensen's inequality for the spectral measure of χ, as in [55, equation (2.1)], gives

$$\langle P_s^D\chi,\chi\rangle \geq \|\chi\|_2^2 \exp\left(-s\,\frac{\mathcal{E}(\chi,\chi)}{\|\chi\|_2^2}\right) \geq c\,m(B(x,r))\exp\Big[-\frac{Cs}{\Psi_S(r)}\Big]\,. \tag{8.34}$$

Step 2: the Harnack inequality in each variable. By Proposition 8.1, the rows of the kernels p_s and p_s^D are weak solutions on (ε,∞) for every $\varepsilon > 0$, with representatives which are continuous in time with values in L^2_{loc}. Apply the Harnack inequality successively in the two variables of the symmetric kernel p^D, comparing the times $t/4$, $t/2$ and t. One may use the ball $B(x,4r)$ in both comparisons, with $(t_0,T) = (3t/8, t/5)$ and $(3t/4, 2t/5)$ respectively. The inner ball $B(x,2r)$ contains $\operatorname{supp}\chi$ and $B(x,r)$, and both cylinders lie in $(t/8,\infty)$. Choose $K > 4$ and restrict r so that these balls and their fixed dilations are in the range of the hypotheses. With

$$H = \exp\left[C\left(1 + \frac{\Psi_P(r)}{t} + \frac{t}{\Psi_S(r)}\right)\right],$$

(7.13) and the scaling of Ψ_P and Ψ_S give

$$p^D_{t/4}(u,v) \le H\, p^D_{t/2}(y,v) \le H^2\, p^D_t(y,z)$$

for almost every $u, v \in \operatorname{supp}\chi$ and $y, z \in B(x,r)$; the comparisons at fixed times are justified as in Lemma 8.10.

Step 3: conclusion. Integrating against $\chi(u)\chi(v)$ and using (8.34) at $s = t/4$,

$$p^D_t(y,z) \ge \frac{\langle P^D_{t/4}\chi, \chi\rangle}{H^2\,\|\chi\|_1^2} \ge \frac{c}{m(B(x,r))} \exp\left[-C\left(1 + \frac{\Psi_P(r)}{t} + \frac{t}{\Psi_S(r)}\right)\right].$$

Since $p_t \ge p^D_t$, the same bound holds for the full semigroup. For the diagonal, apply the bound at the time $t/2$ and integrate its square over $B(x,r)$ in (8.2). This argument applies separately to the two kernels, and it does not evaluate on the diagonal an estimate which holds almost everywhere in the product. □

Two-sided estimates at matching scales. When the two scales coincide, the near-diagonal bound and the semigroup property give the lower bound which matches (8.10). The result is the classical implication from volume doubling, the Poincaré inequality and the cutoff Sobolev inequality at one scale to two-sided sub-Gaussian estimates [17, 10, 75]; it is recovered here from the estimates with two scales.

Corollary 8.12. *Let X be unbounded, proper and geodesic. Suppose that* VD, PI(Ψ) *and* CSA(Ψ) *hold globally, where Ψ is a continuous increasing bijection of $(0,\infty)$ satisfying (7.1) with a lower exponent greater than one. Then the semigroup is conservative, and its heat kernel satisfies*

$$\frac{c}{m(B(x,\Psi^{-1}(t)))}\, e^{-C\mathcal{A}_\Psi(d(x,y),t)} \le p_t(x,y) \le \frac{C}{m(B(x,\Psi^{-1}(t)))}\, e^{-c\mathcal{A}_\Psi(d(x,y),t)} \qquad (t>0) \quad (8.35)$$

for almost every pair (x,y), with structural constants. The same statement holds when the Poincaré scale and the cutoff Sobolev scale are both comparable to Ψ.

Proof. Conservativeness. Global doubling gives a polynomial bound for the volume growth at any base point, and the cutoff Sobolev inequality is uniform in the inner radius at any fixed width. Proposition 8.8 applies.

Upper bound. Put $r_t = \Psi^{-1}(t)$. By doubling,

$$\frac{m(B(x,r_t))}{m(B(y,r_t))} \le C\left(1 + \frac{d(x,y)}{r_t}\right)^{Q_D}.$$

Inserting this in (8.10), the polynomial factor is absorbed by decreasing the constant in the exponential: the upper scaling of Ψ gives $\mathcal{A}_\Psi(d,t) \ge c\,(d/r_t)^{\beta_+/(\beta_+-1)}$ for $d \ge r_t$, while the factor is bounded for $d \le r_t$.

Near-diagonal lower bound. Lemma 8.11 at the radius r_t gives

$$p_t(x,y) \ge \frac{c}{m(B(x,r_t))} \qquad \text{if } d(x,y) \le \varepsilon\, r_t\,,$$

for a fixed $\varepsilon > 0$.

Chaining. We extend this bound to arbitrary distances, following [78, proof of Theorem 6.5 and Corollary 6.11]. Fix a sufficiently small $\delta > 0$, depending on ε. Let N be the least positive integer such that

$$\frac{d(x,y)}{N} \le \delta\, \Psi^{-1}\Big(\frac{t}{N}\Big)\,.$$

Such an integer exists because the lower scaling exponent of Ψ is greater than one. Write $\tau = t/N$ and $\rho = \delta\Psi^{-1}(\tau)$, and divide a geodesic from x to y into N links of length at most ρ. In the semigroup property

$$p_t(x,y) = \int\cdots\int p_\tau(x,z_1)\, p_\tau(z_1,z_2)\cdots p_\tau(z_{N-1},y)\, dm(z_1)\cdots dm(z_{N-1})\,,$$

restrict each variable z_i to the ball of radius ρ about the i-th point of the subdivision. With δ small enough, each factor is bounded below by the near-diagonal estimate at the time τ. By doubling, the volume in the denominator of each factor is comparable to the volume of the ball over which the preceding variable is integrated. All the intermediate volumes cancel, leaving

$$p_t(x,y) \ge \frac{c^N}{m(B(x,\rho))} \ge \frac{c^N}{m(B(x,r_t))}\,, \qquad 0 < c < 1\,.$$

These comparisons hold almost everywhere; countable coverings and Fubini's theorem give the usual exceptional sets.

The number of links. It remains to bound N. If $N > 1$, put $n = N - 1$. By the minimality of N, $t < n\,\Psi(d(x,y)/(\delta n))$. For a sufficiently large structural constant M, the lower scaling of Ψ therefore yields, with $s = Md(x,y)/(\delta n)$ in (8.1),

$$\mathcal{A}_\Psi(d(x,y),t) \ge \frac{\delta n}{M} - \frac{t}{\Psi(Md(x,y)/(\delta n))} \ge \frac{\delta n}{2M}\,.$$

Thus $N \le C(1 + \mathcal{A}_\Psi(d(x,y),t))$, which proves the lower estimate. If the two scales are only comparable to Ψ, the structural constants change. □

Power scales with distinct exponents. We now specialize to power scales which may differ: $\Psi_P(r) = r^a$ and $\Psi_S(r) = r^\beta$, with $a \le \beta$ by Proposition 7.4. The estimates below are derived from the Harnack inequality with two scales. They are not applications of a heat kernel theorem with one scale, and no optimality is claimed. To make the dependence on the Harnack constant explicit, we state them under the following assumptions. The space X is proper and geodesic, the semigroup is ultracontractive at sufficiently small positive times, and, uniformly for $0 < r \le r_0$,

$$m(B(x,r)) \asymp r^Q\,, \qquad \mathrm{Cap}(B(x,r),B(x,2r)) \le C\, r^{Q-\beta}\,, \tag{8.36}$$

where $0 < a \le \beta$ and $a + \beta > 2$. Moreover, nonnegative local weak solutions satisfy a Harnack inequality of the form (7.13) whose constant, on cylinders of radius r and time length T, has a logarithm at most

$$C\left(1 + \frac{r^a}{T} + \frac{T}{r^\beta}\right)\,. \tag{8.37}$$

Theorem 7.5 gives this under $m(B(x,r)) \asymp r^Q$, PI(r^a) and the cutoff Sobolev inequality at the scale r^β. All radii below are restricted so that the fixed dilations of the balls remain in the range of the assumptions. The conclusions hold for $m \otimes m$-almost every pair.

Proposition 8.13. *Under (8.36) and (8.37), there are constants $c, C, K > 0$ such that*

$$p_t(y,z) \geq c\, r^{-Q} \exp\left[-C\left(1 + \frac{r^a}{t} + \frac{t}{r^\beta}\right)\right], \qquad y, z \in B(x,r), \quad t > 0. \tag{8.38}$$

The estimate also holds for the heat kernel killed outside $B(x, Kr)$. In particular, with

$$\alpha = \frac{a+\beta}{2}, \qquad q = \frac{\beta - a}{a+\beta}, \tag{8.39}$$

there is $t_* > 0$ such that

$$p_t(y,z) \geq c\, t^{-Q/\alpha} \exp(-Ct^{-q}), \qquad d(y,z) \leq c\, t^{1/\alpha}, \quad 0 < t \leq t_*. \tag{8.40}$$

When $a = \beta$, the exponential in (8.40) is a constant.

The exponent α is the arithmetic mean of the two exponents: the radius $r = t^{1/\alpha}$ is the one for which the two terms r^a/t and t/r^β of the Harnack constant are equal, and their common value is t^{-q}.

Proof. The proof of Lemma 8.11 uses the hypotheses of Chapter 7 only through the capacity bound, the Harnack inequality, and the existence of a heat kernel whose rows belong to L^2, which here follows from the ultracontractivity. With $\Phi(r) = r^Q$, $\Psi_P(r) = r^a$ and $\Psi_S(r) = r^\beta$, it gives (8.38) for the full semigroup and for the killed semigroup, and the diagonal bound

$$p_t^D(y,y) = \int_D p_{t/2}^D(y,z)^2\, dm(z) \geq c\, r^{-Q} \exp\left[-C\left(1 + \frac{r^a}{t} + \frac{t}{r^\beta}\right)\right] \tag{8.41}$$

for almost every $y \in B(x,r)$, and the same bound for the full semigroup. Finally take r comparable to $t^{1/\alpha}$. Then $r^a/t \asymp t/r^\beta \asymp t^{-q}$, and a countable covering by such balls gives (8.40). □

Theorem 8.14. *Under the assumptions of Proposition 8.13, suppose in addition that the Harnack inequality with the constant (8.37) holds uniformly along all the chains of Lemma 8.10. Put*

$$b_- = \frac{2\beta}{2 + \beta - a} = \frac{\alpha(1+q)}{1+\alpha q} > 1. \tag{8.42}$$

There are $c, C, t_* > 0$ such that, for $0 < t \leq t_*$,

$$p_t(x,y) \geq c\, t^{-Q/\alpha} \exp\left[-C\left\{t^{-q} + \left(\frac{d(x,y)^{b_-}}{t}\right)^{1/(b_- - 1)}\right\}\right] \tag{8.43}$$

for almost every $(x,y) \in X \times X$. The estimate holds on compact and noncompact spaces, without conservativeness. Neither b_- nor q is asserted to be optimal. If $a = \beta$, then $\alpha = b_- = a$ and $q = 0$, and (8.43) is the usual sub-Gaussian lower bound with the exponent a.

Proof. The bound (8.41) for the full semigroup, with $r \asymp t^{1/\alpha}$ and at the time $2t/3$, gives

$$p_{2t/3}(y,y) \geq c\, t^{-Q/\alpha} \exp(-Ct^{-q}) \tag{8.44}$$

for almost every y. Apply Lemma 8.10 with $\Psi_P(r) = r^a$ and $\Psi_S(r) = r^\beta$. Put $d = d(x,y)$ and choose

$$r = c \min\left\{t^{1/\alpha}, \left(\frac{t}{d}\right)^{1/(\alpha-1)}\right\},$$

where the second term is $+\infty$ when $d = 0$. For sufficiently small t, all the enlarged radii are in the range of the assumptions. If $d \le t^{1/\alpha}$, the exponent in (8.31) is at most Ct^{-q}. Otherwise $r \asymp (t/d)^{1/(\alpha-1)}$ and $d/r \ge c$, so that this exponent is at most

$$C\left(\frac{d^2\, r^{a-2}}{t} + \frac{t}{r^\beta}\right) \le C\, d^{2\beta/(a+\beta-2)}\, t^{-(2+\beta-a)/(a+\beta-2)} = C\left(\frac{d^{b_-}}{t}\right)^{1/(b_- -1)}.$$

Combining with (8.44) proves (8.43). □

When $a < \beta$, Theorem 8.6 gives an upper bound with prefactor $t^{-Q/a}$ and exponent β, while Theorem 8.14 gives a lower bound with prefactor $t^{-2Q/(a+\beta)} \exp(-Ct^{-q})$ and exponent $b_- < \beta$. When $a = \beta$, both bounds have the sub-Gaussian form with walk dimension β.

8.5 Global estimates

In this section the hypotheses (7.2), (7.3) and (7.8) hold globally, X is unbounded and geodesic, and Ψ_S has a lower scaling exponent $\beta_- > 1$. All radii and all positive times are allowed. The estimates are expressed with the scale functions themselves, so that the radii selected in a proof may lie on either side of a change of behaviour of these functions. General scale functions, including different powers below and above the radius one, already occur in [10, Section 1 and Lemma 1.5] and [78, Section 3.3 and Section 6]. The distinction made here is between the two scales Ψ_P and Ψ_S at the same radius.

Order of the scales and Harnack inequality.

Proposition 8.15. *The global hypotheses imply*

$$\Psi_S(r) \le C\, \Psi_P(r)\,, \qquad r > 0\,. \tag{8.45}$$

Proof. This is Proposition 7.4, applied to the capacity scale Ψ_S of (7.9). □

Corollary 8.16. *For every $r, T > 0$, every $x \in X$, and every nonnegative local weak solution on $(t_0 - T, t_0 + T) \times B(x, r)$,*

$$\operatorname*{ess\,sup}_{Q_-} u \le \exp\left[C\left(\frac{\Psi_P(r)}{T} + \frac{T}{\Psi_S(r)}\right)\right] \operatorname*{ess\,inf}_{Q_+} u\,, \tag{8.46}$$

where

$$Q_- = (t_0 - 3T/4, t_0 - T/2) \times B(x, r/2)\,, \qquad Q_+ = (t_0 + T/2, t_0 + 3T/4) \times B(x, r/2)\,.$$

Consequently, nonnegative weakly harmonic functions on $B(x, r)$ satisfy

$$\operatorname*{ess\,sup}_{B(x,r/2)} h \le \exp\left[C\sqrt{\frac{\Psi_P(r)}{\Psi_S(r)}}\right] \operatorname*{ess\,inf}_{B(x,r/2)} h\,. \tag{8.47}$$

Moreover, let $s < v$, put $\delta = v - s$, and let $u \ge 0$ be caloric on $(s - \delta, v + \delta) \times X$. For every $r > 0$ and almost every pair (x, y),

$$u(s, x) \le \exp\left[C\left\{\frac{\Psi_P(r)}{\delta}\left(1 + \frac{d(x, y)}{r}\right)^2 + \frac{\delta}{\Psi_S(r)}\right\}\right] u(v, y)\,, \tag{8.48}$$

with the conventions of Lemma 8.10 for the values at fixed times. The exponent may be minimized over $r > 0$.

Proof. Theorem 7.5 applies at every radius, with structural constants. The constant 1 in its exponent is absorbed, because $\Psi_P(r)/T + T/\Psi_S(r) \geq 2\sqrt{\Psi_P(r)/\Psi_S(r)} \geq c$ by (8.45). This gives (8.46). For (8.47), apply it to the stationary solution with $T = \sqrt{\Psi_P(r)\Psi_S(r)}$. By (8.45), Lemma 8.10 applies and gives (8.48). □

Two remarks on (8.48). First, the hypothesis requires the solution on the interval $(s-\delta, v+\delta)$. If a solution is only given on $(0,\infty)\times X$, one may instead choose a chain with $N \geq C\max\{1 + d(x,y)/r, (v-s)/s\}$ links, which ensures that the first cylinder remains at positive times; its cost is $C(N^2\Psi_P(r)/(v-s) + (v-s)/\Psi_S(r))$. The arguments below use times proportional to the final time and need no such restriction. Second, if Ψ_P/Ψ_S is bounded for $r \geq 1$, the Harnack constant at $T = \sqrt{\Psi_P(r)\Psi_S(r)}$ is uniform for $r \geq 1$.

Upper bound with a uniform prefactor. From here to the end of the chapter we assume in addition $\mathrm{V}(\Phi)$ at every radius, with Φ continuous, strictly increasing and satisfying (7.1). Recall the rate function (8.1) for the cutoff Sobolev scale,

$$\mathcal{A}_{\Psi_S}(d,t) = \sup_{s>0}\left\{\frac{d}{s} - \frac{t}{\Psi_S(s)}\right\}, \qquad d \geq 0\,, \quad t > 0\,. \tag{8.49}$$

It is nonnegative and finite, and positive for $d > 0$: the lower scaling of Ψ_S with an exponent greater than one controls the expression as $s \downarrow 0$ and makes it positive for sufficiently large s.

Theorem 8.17. *Under the global hypotheses and* $\mathrm{V}(\Phi)$*, the heat kernel satisfies, for every* $t > 0$ *and almost every pair* (x,y)*,*

$$p_t(x,y) \leq \frac{C}{\Phi(\Psi_P^{-1}(t))}\,\exp\bigl[-c\,\mathcal{A}_{\Psi_S}(d(x,y),t)\bigr]\,. \tag{8.50}$$

Conservativeness is not used.

Proof. We reduce to the iteration of Theorem 8.6 by a normalization which depends on the time.

Step 1: a normalized Nash inequality. Fix $t > 0$, set $r_t = \Psi_P^{-1}(t)$, and normalize the measure and the form by

$$m_t = \Phi(r_t)^{-1} m\,, \qquad \mathcal{E}_t = t\,\Phi(r_t)^{-1}\mathcal{E}\,.$$

The semigroup of the normalized form at the time s is P_{ts}. Let Q_+ be an upper scaling exponent of Φ and $a_- > 0$ a lower scaling exponent of Ψ_P. The inequality (8.11) at the radius $h\,r_t$ gives, for $0 < h \leq 1$,

$$\|f\|_{2,m_t}^2 \leq C\,h^{a_-}\,\mathcal{E}_t(f,f) + C\,h^{-Q_+}\,\|f\|_{1,m_t}^2\,.$$

Optimizing over $h \leq 1$ yields a Nash inequality with constants independent of t:

$$\|f\|_{2,m_t}^{2+2/D} \leq C(\mathcal{E}_t(f,f) + \|f\|_{2,m_t}^2)\|f\|_{1,m_t}^{2/D}\,, \qquad D = \frac{Q_+}{a_-}\,. \tag{8.51}$$

Step 2: the iteration. For the normalized form, (7.8) is (8.3) with $S(s) = \Psi_S(s)/t$ and $r_0 = \infty$. Fix $z \in X$, $r > 0$ and $\lambda \geq 1$. Lemma 8.2 supplies the perturbations for this scale, with the coefficient $Cp^{\beta_+ - 1}t/\Psi_S(r/\lambda)$ in (8.5). Apply Steps 1 to 5 of the proof of Theorem 8.6 to (8.51) at the normalized time one, with $\beta = \beta_+$ and $K = t/\Psi_S(r/\lambda)$. The zeroth-order term of the Nash inequality contributes a fixed factor. Returning to the original measure gives

$$\|e^{\lambda\rho}P_t\,e^{-\lambda\rho}\|_{1\to\infty} \leq \frac{C}{\Phi(r_t)}\,\exp\left[\frac{Ct}{\Psi_S(r/\lambda)}\right]\,.$$

For x in the inner ball $B(z,r)$ and y outside the outer ball $B(z,2r)$, the kernel is therefore at most

$$\frac{C}{\Phi(r_t)}\,\exp\left[-\lambda+\frac{Ct}{\Psi_S(r/\lambda)}\right].$$

Step 3: optimization. Choose r comparable to $d(x,y)$ and optimize over $\lambda\ge 1$. With $s=r/\lambda$ this gives the supremum over $s\le r$ in (8.49). The values $s>r$ contribute at most one to the supremum, and the bound (8.13) covers the region where the restricted supremum is bounded. Thus

$$p_t(x,y)\le\frac{C}{\Phi(r_t)}\,\exp\bigl[-\mathcal{A}_{\Psi_S}(c_1 d(x,y),C_1 t)\bigr].$$

Step 4: comparison of the rates. We claim that $\mathcal{A}_{\Psi_S}(c_1 d,C_1 t)\ge c\,\mathcal{A}_{\Psi_S}(d,t)$. Choose a fixed $M\ge 1$ so large that $C_1 c_S^{-1}M^{-\beta_-}\le c_1/(2M)$, where c_S is the constant of the lower scaling of Ψ_S. For each s with $d/s-t/\Psi_S(s)>0$, the lower scaling gives $C_1 t/\Psi_S(Ms)\le C_1 c_S^{-1}M^{-\beta_-}\,t/\Psi_S(s)\le c_1 d/(2Ms)$, and therefore

$$\mathcal{A}_{\Psi_S}(c_1 d,C_1 t)\ge\frac{c_1 d}{Ms}-\frac{C_1 t}{\Psi_S(Ms)}\ge\frac{c_1 d}{2Ms}\ge\frac{c_1}{2M}\left(\frac{d}{s}-\frac{t}{\Psi_S(s)}\right).$$

Taking the supremum over s proves the claim. The formulation almost everywhere is obtained with countable dense choices of centres and rational radii and parameters, as in Theorem 8.6. □

Conservativeness and confinement.

Proposition 8.18. *Under the global hypotheses and* $\mathrm{V}(\Phi)$*, the semigroup is conservative.*

Proof. The upper scaling of Φ gives a polynomial bound for the volume growth, and the cutoff Sobolev inequality is uniform in the inner radius, in particular for the widths $0<s\le 1$. Proposition 8.8 applies. □

The diagonal lower bound follows from confinement, as in Proposition 8.9, with a radius chosen by means of the scale functions. For the rest of this section set

$$\rho_t=\Psi_S^{-1}(t),\qquad L_t=1+\log^+\frac{\Phi(\rho_t)}{\Phi(\Psi_P^{-1}(t))},\qquad R_t=K\rho_t\,L_t^{(\beta_+-1)/\beta_+},\tag{8.52}$$

where $\log^+ v=\max\{0,\log v\}$ and K is a sufficiently large structural constant. By (8.45), $\rho_t\ge c\,\Psi_P^{-1}(t)$; the number L_t measures the discrepancy between the two scales at the time t, and $L_t=1$, $R_t=K\rho_t$ when they coincide. For each fixed $c>0$, $R_{ct}\asymp R_t$ uniformly in t, because the inverses of the scale functions change by bounded factors and the logarithm changes by a bounded additive quantity.

Proposition 8.19. *One may choose K in (8.52) so that, for every $t>0$ and almost every x,*

$$\int_{B(x,R_t)}p_t(x,y)\,dm(y)\ge\frac12,\qquad p_t(x,x)\ge\frac{c}{\Phi(R_t)}.\tag{8.53}$$

Moreover, for $R\ge\rho_t$,

$$\int_{X\setminus B(x,R)}p_t(x,y)\,dm(y)\le\min\left\{1,\ C\,\frac{\Phi(\rho_t)}{\Phi(\Psi_P^{-1}(t))}\,\exp\left[-c\Bigl(\frac{R}{\rho_t}\Bigr)^{\beta_+/(\beta_+-1)}\right]\right\}.\tag{8.54}$$

The diagonal value is (8.2).

Proof. Put $\gamma = \beta_+/(\beta_+ - 1)$.

Step 1: a lower bound for the rate. We show that

$$\mathcal{A}_{\Psi_S}(R,t) \geq c\Big(\frac{R}{\rho_t}\Big)^{\gamma}, \qquad R \geq \rho_t\,.$$

Let $k = R/\rho_t$. For k sufficiently large, insert $s = M\rho_t\, k^{-1/(\beta_+-1)}$ in (8.49), with a fixed large M. Then $R/s = M^{-1}k^{\gamma}$, and the upper scaling of Ψ_S, while $s \leq \rho_t$, gives $t/\Psi_S(s) \leq CM^{-\beta_+}k^{\gamma}$, which is at most half of R/s. For bounded $k \geq 1$, insert $s = M'\rho_t$ and use the lower scaling with an exponent greater than one; a sufficiently large fixed M' gives a positive lower bound.

Step 2: the tail. For $w = R/\rho_t \geq 1$, the decomposition into dyadic annuli, Theorem 8.17 and Step 1 give

$$\begin{aligned}\int_{X\setminus B(x,R)} p_t(x,y)\,dm(y) &\leq C\,\frac{\Phi(\rho_t)}{\Phi(\Psi_P^{-1}(t))}\sum_{j\geq 0}(2^j w)^{Q_+}e^{-c(2^j w)^{\gamma}}\\ &\leq C\,\frac{\Phi(\rho_t)}{\Phi(\Psi_P^{-1}(t))}\,e^{-c'w^{\gamma}}\,.\end{aligned}$$

The bound by one holds because the semigroup is sub-Markovian.

Step 3: confinement. At $R = R_t$, a sufficiently large structural K makes the last expression at most $1/2$, by the definition of L_t. Conservativeness gives the first assertion of (8.53). Apply the Cauchy–Schwarz estimate (8.29) at the radius R_t, with $m(B(x,R_t)) \leq C\Phi(R_t)$. Replacing $2t$ by t and using $R_{t/2} \asymp R_t$ proves the diagonal bound. □

Lower bounds.

Theorem 8.20. *Assume the global hypotheses and* $\mathrm{V}(\Phi)$.

(i) *For every $t > 0$ and almost every (x,y),*

$$p_t(x,y) \geq \frac{c}{\Phi(R_t)}\,\exp\left[-C\inf_{r>0}\left\{\frac{\Psi_P(r)}{t}\left(1+\frac{d(x,y)}{r}\right)^2 + \frac{t}{\Psi_S(r)}\right\}\right]. \tag{8.55}$$

(ii) *For every $R, r > 0$,*

$$p_t(x,y) \geq \frac{c}{\Phi(R)}\exp\left[-C\left\{\frac{\Psi_P(R)}{t} + \frac{t}{\Psi_S(R)} + \frac{\Psi_P(r)}{t}\left(1+\frac{d(x,y)}{r}\right)^2 + \frac{t}{\Psi_S(r)}\right\}\right]. \tag{8.56}$$

The proof of this bound uses neither conservativeness nor the upper estimate (8.50).

(iii) *For points in a common ball,*

$$p_t(y,z) \geq \frac{c}{\Phi(R)}\,\exp\left[-C\left(\frac{\Psi_P(R)}{t} + \frac{t}{\Psi_S(R)}\right)\right], \qquad y,z \in B(x,R)\,. \tag{8.57}$$

This estimate holds also for the kernel killed outside $B(x,KR)$, with a fixed structural K.

The bound (i) starts from the diagonal bound given by confinement, and (ii) from the diagonal bound given by a cutoff function at an arbitrary radius R. Neither implies the other in general.

Proof. (i) By Lemma 8.10, $p_t(x,y)$ is at least $p_{2t/3}(y,y)$ times the exponential in (8.31), and Proposition 8.19 gives $p_{2t/3}(y,y) \geq c/\Phi(R_t)$. Optimizing the radius proves (8.55).

(iii) Lemma 8.11 applies at every radius and gives (8.57), for the full kernel and for the kernel killed outside $B(x,KR)$. The constant term in the exponential is absorbed because $\Psi_S \leq C\Psi_P$.

(ii) The diagonal conclusion of Lemma 8.11, at the time $2t/3$, gives

$$p_{2t/3}(y,y) \geq \frac{c}{\Phi(R)} \exp\left[-C\left(\frac{\Psi_P(R)}{t} + \frac{t}{\Psi_S(R)}\right)\right]$$

for almost every $y \in B(x,R)$. A countable covering gives this bound almost everywhere on X; taking dyadic values of R first makes the exceptional set independent of R. Insert it in (8.31) to obtain (8.56). For a killed semigroup the same propagation applies when the ball of radius R and all the enlarged balls of the chain lie in the domain. □

The balanced radius. Let $\Xi(r) = \sqrt{\Psi_P(r)\Psi_S(r)}$. At the radius $R = \Xi^{-1}(t)$ the two terms in the exponent of (8.57) are equal, and

$$p_t(x,y) \geq \frac{c}{\Phi(\Xi^{-1}(t))} \exp\left[-C\sqrt{\frac{\Psi_P(\Xi^{-1}(t))}{\Psi_S(\Xi^{-1}(t))}}\right] \qquad \text{when } d(x,y) \leq c\,\Xi^{-1}(t)\,. \tag{8.58}$$

Different powers at small and large radii. To make the general estimates explicit, assume that the three functions are

$$\begin{aligned} \Phi(r) &= \begin{cases} r^{Q_0}\,, & 0 < r \leq 1\,, \\ r^{Q_\infty}\,, & r \geq 1\,, \end{cases} \qquad \Psi_P(r) = \begin{cases} r^{a_0}\,, & 0 < r \leq 1\,, \\ r^{a_\infty}\,, & r \geq 1\,, \end{cases} \\ \Psi_S(r) &= \begin{cases} r^{\beta_0}\,, & 0 < r \leq 1\,, \\ r^{\beta_\infty}\,, & r \geq 1\,, \end{cases} \end{aligned} \tag{8.59}$$

where $Q_0, Q_\infty, a_0, a_\infty > 0$, $\beta_0 \geq 2$ and $\beta_\infty > 1$. Multiplicative constants and different transition radii for the three functions are covered by the general theorems without change. Proposition 8.15 gives the necessary conditions

$$a_0 \leq \beta_0\,, \qquad \beta_\infty \leq a_\infty\,. \tag{8.60}$$

In particular, two distinct powers r^a and r^β cannot be the Poincaré and cutoff Sobolev scales at every radius: at small radii the Poincaré exponent is the smaller one, and at large radii it is the larger one. The conditions (8.60) are necessary; we do not assert that every such tuple of exponents is realized.

Put $\alpha_i = (a_i + \beta_i)/2$ for $i \in \{0,\infty\}$. With the time length which minimizes the exponent in (8.46), the logarithm of the Harnack constant is bounded as follows:

Radius	Time length	Bound for the logarithm of the Harnack constant
$0 < r \leq 1$	r^{α_0}	$C\,r^{-(\beta_0 - a_0)/2}$
$r \geq 1$	r^{α_∞}	$C\,r^{(a_\infty - \beta_\infty)/2}$

If $a_i = \beta_i$, the constant is uniform in the corresponding range of radii.

Corollary 8.21. *Under* (8.59)*, for every* $t > 0$*,*

$$p_t(x,y) \leq C \begin{cases} t^{-Q_0/a_0}\,, & t \leq 1\,, \\ t^{-Q_\infty/a_\infty}\,, & t \geq 1\,, \end{cases} \times \exp\left[-c\,\mathcal{A}_{\Psi_S}(d(x,y),t)\right], \tag{8.61}$$

where, for $d > 0$,

$$\mathcal{A}_{\Psi_S}(d,t) \asymp \begin{cases} (d^{\beta_0}/t)^{1/(\beta_0-1)}\,, & t \le d\,, \\ (d^{\beta_\infty}/t)^{1/(\beta_\infty-1)}\,, & t \ge d\,. \end{cases} \tag{8.62}$$

The diagonal satisfies

$$p_t(x,x) \ge c \begin{cases} t^{-Q_0/\beta_0}\,[\log(e/t)]^{-Q_0(\beta_0-1)/\beta_0}\,, & 0 < t \le 1\,, \\ t^{-Q_\infty/\beta_\infty}\,[\log(e+t)]^{-Q_\infty(\beta_\infty-1)/\beta_\infty}\,, & t \ge 1\,. \end{cases} \tag{8.63}$$

The logarithm in either line can be omitted when $a_i = \beta_i$ in the corresponding regime.

Proof. The prefactor in (8.61) is obtained by inverting Ψ_P in (8.50).

The rate. For a power s^β, the maximum of $d/s - t/s^\beta$ is attained at $s \asymp (t/d)^{1/(\beta-1)}$ and is comparable to $(d^\beta/t)^{1/(\beta-1)}$. In the present case split the supremum at $s = 1$. If t/d is sufficiently small, the maximizing value of s lies below one and involves β_0; if it is sufficiently large, it lies above one and involves β_∞. When t/d remains between two fixed positive constants, both expressions in (8.62) are comparable to d, and a fixed sufficiently large s gives the lower bound $c\,d$; the same splitting gives the upper bound. This is the computation of [10, Lemma 1.5].

The diagonal at small times. Take $R = K\,t^{1/\beta_0}[\log(e/t)]^{(\beta_0-1)/\beta_0}$. Then $t \le R \le 1$ for sufficiently small t, and every annulus outside $B(x,R)$ is in the first regime of (8.62). The proof of Proposition 8.19, with the exponent $\beta_0/(\beta_0-1)$, shows that the mass outside $B(x,R)$ is at most $1/2$ for a fixed sufficiently large K: the polynomial factor $t^{Q_0/\beta_0 - Q_0/a_0}$ is absorbed by the logarithm. The diagonal bound follows from the Cauchy–Schwarz inequality and $\Phi(R) = R^{Q_0}$.

The diagonal at large times. Take $R = K\,t^{1/\beta_\infty}[\log(e+t)]^{(\beta_\infty-1)/\beta_\infty}$. Then $1 \le R \le t$ for sufficiently large t. The annuli up to the radius t are in the second regime of (8.62), and their contribution is controlled by the same argument. Beyond the radius t, the first regime gives the bound $C\,t^{Q_\infty - Q_\infty/a_\infty}e^{-ct}$ for the mass, which tends to zero. Conservativeness and $\Phi(R) = R^{Q_\infty}$ give the second diagonal bound. The remaining compact interval of times is covered by Proposition 8.19. If $a_i = \beta_i$, the polynomial factor in the relevant estimate of the mass is bounded, and $R = Kt^{1/\beta_i}$ suffices. □

Thus the time alone does not determine the exponent of the off-diagonal decay. For example, even when $t \gg 1$, the distances $d \gg t$ involve β_0 in the exponential of the upper bound, while the prefactor involves Q_∞/a_∞.

Lower bounds for power scales. The near-diagonal bound (8.58) becomes

$$p_t(x,y) \ge \begin{cases} c\,t^{-Q_0/\alpha_0}\,\exp\bigl[-C\,t^{-(\beta_0-a_0)/(a_0+\beta_0)}\bigr]\,, & 0 < t \le 1\,, \\ c\,t^{-Q_\infty/\alpha_\infty}\,\exp\bigl[-C\,t^{(a_\infty-\beta_\infty)/(a_\infty+\beta_\infty)}\bigr]\,, & t \ge 1\,, \end{cases} \tag{8.64}$$

provided that $d(x,y) \le c\,t^{1/\alpha_i}$ in the corresponding line. To make the cost of the chains in (8.55) explicit, take any integer $N \ge 1$ such that

$$d(x,y) \le c\,N\,\Xi^{-1}(t/N)\,, \qquad \Xi = \sqrt{\Psi_P\Psi_S}\,, \tag{8.65}$$

and substitute $r = \Xi^{-1}(t/N)$ in (8.55). Then

$$p_t(x,y) \ge \frac{c}{\Phi(R_t)}\,\exp\left[-C\,N\sqrt{\frac{\Psi_P(\Xi^{-1}(t/N))}{\Psi_S(\Xi^{-1}(t/N))}}\right]. \tag{8.66}$$

In the power case the radius and the exponent are as follows.

Time of a link	Radius	Exponent, up to a constant
$t/N \le 1$	$(t/N)^{1/\alpha_0}$	$N\,(N/t)^{(\beta_0-a_0)/(a_0+\beta_0)}$
$t/N \ge 1$	$(t/N)^{1/\alpha_\infty}$	$N\,(t/N)^{(a_\infty-\beta_\infty)/(a_\infty+\beta_\infty)}$

Since $\alpha_0 > 1$, the condition (8.65) holds for all sufficiently large N, so these formulas reduce the lower bound to an optimization over one integer, for all times and distances.

For $t \ge 1$, choose N comparable to $1 + (d^{\alpha_\infty}/t)^{1/(\alpha_\infty-1)}$ when $d \le t$, and to $1 + (d^{\alpha_0}/t)^{1/(\alpha_0-1)}$ when $d \ge t$, with a sufficiently large fixed multiplier; the first choice has $N \lesssim t$ and the second $N \gtrsim t$. Substituting in the table, the exponent in (8.66) is at most a constant times

$$\begin{cases} t^{(a_\infty-\beta_\infty)/(a_\infty+\beta_\infty)} + d^{2\beta_\infty/(a_\infty+\beta_\infty-2)}\, t^{(a_\infty-\beta_\infty-2)/(a_\infty+\beta_\infty-2)}\,, & d \le t\,, \\ 1 + d^{2\beta_0/(a_0+\beta_0-2)}\, t^{-(2+\beta_0-a_0)/(a_0+\beta_0-2)}\,, & d \ge t\,. \end{cases} \tag{8.67}$$

The denominators $a_i + \beta_i - 2$ are positive. In contrast, $2 + \beta_\infty - a_\infty$ may be zero or negative, so the exponent $2\beta/(2+\beta-a)$ of Theorem 8.14 has no analogue at infinity.

Matching scales.

Corollary 8.22. *Assume the global hypotheses and* $\mathrm{V}(\Phi)$*. If* $\Psi_P \asymp \Psi_S = \Psi$ *at every radius, then the Harnack constant in* (8.46) *is uniform for the time length* $T = \Psi(r)$*, and*

$$\frac{c}{\Phi(\Psi^{-1}(t))}\, e^{-C\mathcal{A}_\Psi(d(x,y),t)} \le p_t(x,y) \le \frac{C}{\Phi(\Psi^{-1}(t))}\, e^{-c\mathcal{A}_\Psi(d(x,y),t)}\,, \qquad t > 0\,. \tag{8.68}$$

Both Φ *and* Ψ *may have different powers at small and large scales.*

Proof. The assertion on the Harnack constant follows from (8.46). The estimates are those of Corollary 8.12, with $m(B(x, \Psi^{-1}(t))) \asymp \Phi(\Psi^{-1}(t))$; the upper bound is also Theorem 8.17. □

On a bounded space the hypotheses concern sufficiently small radii and yield estimates at small times: a volume function with a positive lower scaling exponent cannot describe the volume of balls at every radius on a space of finite measure.

9 Applications to gradient structures

We now apply the results of Chapters 7 and 8 to the Dirichlet form of a 2-admissible gradient structure. Two of the three hypotheses of Part II are available at once: volume doubling is a hypothesis on m, and the Poincaré inequality $\mathrm{PI}_2(\Psi_{P,2})$ on the core extends to the domain of the form. The third hypothesis, the cutoff Sobolev inequality, is not among the hypotheses of Part I. The cutoff condition $\mathrm{B}(\Psi_{C,2})$ gives only an upper bound for the capacity of condensers. This chapter derives cutoff Sobolev inequalities in two situations.

- *Distinct scales* (Section 9.3). If the energy measures of the cutoffs satisfy a growth condition on small balls, denoted $\mathrm{B}^+(\Psi_{C,2})$, a direct argument gives a cutoff Sobolev inequality at a scale Ψ_S computed from Ψ_P and $\Psi_{C,2}$ (Proposition 9.3). The estimates of Chapter 8 with two scales then apply. The argument rests on a trace inequality for measures which satisfy a growth condition (Section 9.2); its proof is the telescoping argument of Anttila [14, Proposition 3.4], with the two scales kept separate.
- *Matching scales* (Sections 9.4 and 9.5). If the Poincaré and cutoff scales are comparable, the capacity bound and the Poincaré inequality imply the cutoff Sobolev inequality at the common scale, by a theorem of Eriksson-Bique [64]. The heat kernel then satisfies two-sided sub-Gaussian estimates (Theorem 9.8).

Section 9.6 characterizes Sobolev and BV functions by the heat kernel.

9.1 The Dirichlet form of a gradient structure

Let $(\mathcal{S}, \nu, \nabla)$ be a 2-admissible gradient structure, and let $(\mathcal{E}, \mathcal{F})$ be its Dirichlet form (Proposition 2.7), with energy measures $d\Gamma(f) = |\nabla_2 f|^2 \, d\nu$. The quantities of Parts I and II correspond as follows.

Poincaré inequality. Under $\mathrm{PI}_2(\Psi_{P,2})$, with $\Psi_P = \Psi_{P,2}$, approximation in the norm of the form extends the inequality from the core to

$$\int_{B(x,r)} |f - f_{B(x,r)}|^2 \, dm \le C \, \Psi_P(r) \, \Gamma(f)(B(x, \Lambda r)) \,, \qquad f \in \mathcal{F} \,, \tag{9.1}$$

and to $\mathcal{F}_{\mathrm{loc}}(X)$ by locality. This is $\mathrm{PI}(\Psi_P)$ in the sense of (7.3).

Capacity. Under VD and $\mathrm{B}(\Psi_{C,2})$, the cutoff scale is a capacity scale in the sense of Definition 7.1:

$$\mathrm{Cap}(B(x,r), B(x,2r)) \le C \, \frac{m(B(x,r))}{\Psi_{C,2}(r)} \,. \tag{9.2}$$

The cutoff ψ_x^r of $\mathrm{B}(\Psi_{C,2})$ is equal to one on $B(x,r)$ only, whereas a cutoff function for $B(x,r) \subset B(x,2r)$ must be equal to one on a neighbourhood of the closed ball, so a buffer has to be inserted. The closed ball $\overline{B(x,5r/4)}$ is compact for small r, because it lies in the support of a cutoff of $\mathrm{B}(\Psi_{C,2})$. Cover it by the balls of radius $r/8$ centred at the points x_i of a maximal $r/8$-separated subset; their number is bounded, by Lemma 3.2. The maximum of the cutoffs $\psi_{x_i}^{r/8}$ is equal to one on $B(x,5r/4)$, and its support lies in $B(x,2r)$. By (1.4) its energy is at most the sum of their energies, and (3.10) gives (9.2). The same argument applies to any fixed ratio of the radii in place of two.

Order of the scales. If X is connected with positive diameter, then $\Psi_{C,2} \le C\,\Psi_P$ at small radii, by Proposition 3.7.

9.2 A trace inequality

Let μ be a measure on X which may be singular with respect to m; the case of interest is the energy measure $\mu = |\nabla\chi|^2\nu$ of a cutoff. The next proposition bounds $\int |u|^p\,d\mu$ by the energy and the $L^p(m)$ norm of u, under a growth condition on the μ-measure of balls. The proof writes u as the limit of its averages on balls of radii $2^{-j}t$: the difference of two successive averages is bounded by the Poincaré inequality, and its integral against μ by the growth condition.

Proposition 9.1. *Assume* VD *and* $\mathrm{PI}_p(\Psi_{P,p})$. *Let μ be a finite Borel measure carried by a compact set K, and suppose that*

$$\mu(B(y,s)) \le M\,\frac{s^2}{\Psi_{C,2}(s)}\,m(B(y,s)) \qquad (y\in X,\ 0<s\le s_0)\,.$$

For all sufficiently small $t \le c\,s_0$ and all $u \in \mathcal{C}_p$,

$$\int_K |u|^p\,d\mu \le C\,M\left(\sum_{j\ge0}\Big(\frac{(2^{-j}t)^2\,\Psi_{P,p}(2^{-j}t)}{\Psi_{C,2}(2^{-j}t)}\Big)^{1/p}\right)^p \int_{K_{Ct}\cap\mathcal{S}} |\nabla u|^p\,d\nu + C\,M\,\frac{t^2}{\Psi_{C,2}(t)}\int_{K_{Ct}}|u|^p\,dm\,, \tag{9.3}$$

where $K_v = \{y : d(y,K) < v\}$. The estimate is asserted when the series converges.

Proof. We use the argument of [14, proof of Proposition 3.4]. Put $u_s(y) = u_{B(y,s)}$. Doubling, Jensen's inequality and the Poincaré inequality give

$$|u_s(y) - u_{s/2}(y)|^p \le \frac{C\,\Psi_{P,p}(s)}{m(B(y,s))}\int_{B(y,\Lambda s)\cap\mathcal{S}} |\nabla u|^p\,d\nu\,.$$

We integrate with respect to μ and exchange the order of integration. For fixed z, doubling compares $m(B(y,s))$ with $m(B(z,s))$ when $d(y,z)<\Lambda s$, and the growth condition on μ and (3.2) imply

$$\int_{B(z,\Lambda s)}\frac{d\mu(y)}{m(B(y,s))} \le C\,\frac{\mu(B(z,\Lambda s))}{m(B(z,s))} \le C\,M\,\frac{s^2}{\Psi_{C,2}(s)}\,.$$

All the radii here are below s_0 after the reduction of t by a fixed factor. Hence

$$\|u_s - u_{s/2}\|^p_{L^p(\mu)} \le C\,M\,\frac{s^2\,\Psi_{P,p}(s)}{\Psi_{C,2}(s)}\int_{K_{\Lambda s}\cap\mathcal{S}}|\nabla u|^p\,d\nu\,.$$

The same calculation gives $\|u_t\|^p_{L^p(\mu)} \le C\,M\,t^2\,\Psi_{C,2}(t)^{-1}\int_{K_t}|u|^p\,dm$. Since u is Lipschitz, $u_{2^{-j}t}\to u$ uniformly. Writing $u = u_t + \sum_{j\ge0}(u_{2^{-j-1}t} - u_{2^{-j}t})$ and applying Minkowski's inequality in $L^p(\mu)$ proves (9.3). □

The series in (9.3) compares the Poincaré scale with the cutoff scale, through the function

$$\frac{s^2\,\Psi_{P,p}(s)}{\Psi_{C,2}(s)} = \frac{s^p\,\Psi_{P,p}(s)}{\Psi_{C,p}(s)}\,,$$

which is comparable to s^p at matching scales.

9.3 Cutoff Sobolev inequalities for distinct scales

To apply the trace inequality to the energy measure of a cutoff, this measure must satisfy the growth condition of Proposition 9.1 at all small scales. The condition $\mathrm{B}(\Psi_{C,2})$ controls only its total mass. We therefore introduce a stronger condition.

Definition 9.2. A gradient structure satisfies $\mathrm{B}^+(\Psi_{C,2})$ if, for every centre x and every radius r in the range of $\mathrm{B}(\Psi_{C,2})$, there is a cutoff ψ_x^r satisfying (3.8) and (3.9) such that

$$\int_{B(y,s)\cap\mathcal{S}} |\nabla\psi_x^r|^2 \, d\nu \le C \left(\frac{s}{r}\right)^2 \frac{m(B(y,s))}{\Psi_{C,2}(s)} \qquad (y \in X\,,\ 0 < s \le r)\,. \tag{$\mathrm{B}^+(\Psi_{C,2})$}$$

The same cutoff serves every y and s.

For $s = r$ this is the bound (3.10) with $p = 2$, localized to balls of radius r. For $s < r$ it says that the energy of the cutoff of radius r in a ball of radius s is at most $(s/r)^2$ times the energy of a cutoff of radius s: the cutoff of the larger ball is, at smaller scales, as regular as a rescaled cutoff of a smaller ball. The condition is the version for the exponent two of the energy inequality for cutoff functions of Anttila [14, Definition 4.8].

At the Gaussian scale $\Psi_{C,2}(r) \asymp r^2$, VD and $\mathrm{B}(\Psi_{C,2})$ imply $\mathrm{B}^+(\Psi_{C,2})$. Indeed, Proposition 3.6 gives $|\nabla\psi_x^r|\,\nu \le C\,\mathrm{Lip}(\psi_x^r)\,m$, and $|\nabla\psi_x^r| \le \mathrm{Lip}(\psi_x^r)$ then gives $|\nabla\psi_x^r|^2\,\nu \le Cr^{-2}\,m$.

Proposition 9.3. *Let X be connected with positive diameter and let a 2-admissible gradient structure satisfy* VD, $\mathrm{PI}(\Psi_P)$ *and* $\mathrm{B}^+(\Psi_{C,2})$. *Set*

$$H(s) = s^2\,\frac{\Psi_P(s)}{\Psi_{C,2}(s)}\,.$$

(i) *For every annulus $A = B(x,R+r)\setminus\overline{B(x,R)}$ with relatively compact outer ball and sufficiently small width r, there is a Lipschitz cutoff function χ for $B(x,R) \subset B(x,R+r)$ such that, for each $0 < t \le c\,r$ for which the sum below is finite and for every $f \in \mathcal{F}$,*

$$\int_A f^2\,d\Gamma(\chi) \le \frac{C}{r^2}\Big(\sum_{j\ge0} H(2^{-j}t)^{1/2}\Big)^2\,\Gamma(f)(A) + \frac{C\,t^2}{r^2\,\Psi_{C,2}(t)}\int_A f^2\,dm\,. \tag{9.4}$$

(ii) *If H has a positive lower scaling exponent, then* $\mathrm{CSA}(\Psi_S)$ *holds with*

$$\Psi_S(r) \asymp \Psi_P\big(H^{-1}(r^2)\big)\,, \tag{9.5}$$

where the inverse is that of a continuous strictly increasing function comparable to H, and the cutoff functions are Lipschitz. The constants are uniform in the inner radius.

(iii) *If the scales match at $p = 2$, then $\Psi_S \asymp \Psi_P \asymp \Psi_{C,2}$.*

The condition of compactness holds for arbitrary annuli when X is proper, and for all sufficiently small outer radii under $\mathrm{B}(\Psi_{C,2})$. *If the hypotheses hold globally, then* $\mathrm{B}(\Psi_{C,2})$ *implies that X is proper, and the conclusions hold at all widths.*

The function H measures the discrepancy between the two scales: $H(s) \asymp s^2$ when they match, and $H(s) \ge c\,s^2$ in general. The number $h = H^{-1}(r^2) \le Cr$ is the radius at which the cutoff of width r is tested in the trace inequality, and the resulting cutoff Sobolev scale is the Poincaré scale at this smaller radius.

Proof. Step 1: the cutoff function and its energy measure. Let $\varepsilon = r/16$. Take a finite maximal ε-separated subset $\{x_i\}$ of the compact set $\overline{B(x, R + r/2)}$, and let $\chi = \max_i \psi^\varepsilon_{x_i}$, with the cutoffs of $\mathrm{B}^+(\Psi_{C,2})$. Then $\chi = 1$ on $B(x, R + r/2)$ and $\operatorname{supp}\chi \subset \overline{B(x, R + r/2 + 2\varepsilon)}$, so that χ is a cutoff function for the annulus, and by locality $\nabla\chi$ vanishes ν-almost everywhere outside the compact set

$$K = \{y : R + r/2 \le d(x, y) \le R + r/2 + 2\varepsilon\}\,.$$

By (1.4), $|\nabla\chi|^2 \le \sum_i |\nabla\psi^\varepsilon_{x_i}|^2$, and by Lemma 3.2 only a bounded number of terms contribute on a ball $B(y, s)$ with $s \le \varepsilon$. By $\mathrm{B}^+(\Psi_{C,2})$,

$$\Gamma(\chi)(B(y, s)) \le \frac{C}{r^2}\,\frac{s^2}{\Psi_{C,2}(s)}\,m(B(y, s)) \qquad (y \in X\,,\ 0 < s \le \varepsilon)\,.$$

Step 2: the trace inequality. Apply Proposition 9.1 with $p = 2$ to $\mu = \Gamma(\chi)$, with $M = C/r^2$, $s_0 = \varepsilon$ and $t \le cr$ so small that $K_{Ct} \subset A$. This is (9.4) for $f \in \mathcal{C}_2$. When the sum is finite, the right-hand side is bounded by a constant times the squared norm of f in $\mathcal{F}$. For $f \in \mathcal{F}$, approximate f in this norm by functions of $\mathcal{C}_2$ and pass to a subsequence which converges quasi-everywhere to the quasi-continuous representative of f [69, Theorem 2.1.4]. Since $\Gamma(\chi)$ charges no set of zero capacity, Fatou's lemma gives (9.4) on $\mathcal{F}$. This proves (i).

Step 3: the unweighted inequality. Suppose now that H has a positive lower scaling exponent. It has a finite upper scaling exponent because it is comparable at comparable radii. The function $t \mapsto \int_0^t H(s)\,ds/s$ is continuous, strictly increasing and comparable to H; we use it to define $h = H^{-1}(r^2)$. Proposition 3.7 gives $H(s) \ge c\,s^2$, hence $h \le Cr$. By the lower scaling, the sum in (9.4) is at most $C\,H(t)^{1/2}$. Given $\eta > 0$, choose $t = c_\eta h$ with $c_\eta > 0$ so small that the coefficient $C\,H(t)/r^2$ of the energy is at most η. Since $t^2/\Psi_{C,2}(t) = H(t)/\Psi_P(t)$, the coefficient of the zeroth-order term is then at most $C_\eta/\Psi_P(h)$. Thus the cutoff function χ of Step 1 satisfies, for every $\eta > 0$,

$$\int_A f^2\,d\Gamma(\chi) \le \eta\,\Gamma(f)(A) + \frac{C_\eta}{\Psi_S(r)}\int_A f^2\,dm\,, \qquad \Psi_S(r) = \Psi_P(H^{-1}(r^2))\,. \tag{9.6}$$

Step 4: the weighted inequality. We pass from (9.6) to (7.6) by the summation over annuli of [10, Lemma 5.1], as in Lemma 7.8, with weights chosen so that the sum remains Lipschitz. The function Ψ_S has a positive lower and a finite upper scaling exponent; let β be an upper exponent. Inside the given annulus choose successive concentric annuli A_j, $j \ge 1$, of widths $r_j = 2^{-j-1}r$, starting at the radius R, and let χ_j be the cutoff function of Step 1 for A_j. Fix $0 < q < \min\{1/2, 2^{-\beta}\}$, put $w_j = (1 - q)q^{j-1}$ and

$$\widetilde{\chi} = \sum_{j\ge1} w_j\,\chi_j\,.$$

The series converges uniformly and $\widetilde{\chi}$ is Lipschitz, because $\mathrm{Lip}(\chi_j) \le C/r_j$ and $q < 1/2$. Its support lies in $\overline{B(x, R + r/2)}$, and it is equal to one on a neighbourhood of $\overline{B(x, R)}$, as is each χ_j. The gradient of χ_j is carried by A_j, and on A_j one has $\chi_i = 1$ for $i > j$, so that $\widetilde{\chi} \ge \sum_{i>j} w_i = q^j$ there. Consequently

$$\eta\,w_j^2 \le \tfrac{1}{8}\,\widetilde{\chi}^2 \ \text{ on } A_j \quad \text{if } \ \eta \le \frac{q^2}{8(1-q)^2}\,, \qquad \text{and} \qquad \frac{w_j^2}{\Psi_S(r_j)} \le \frac{C}{\Psi_S(r)}\,,$$

the second bound because $\Psi_S(r)/\Psi_S(r_j) \le C2^{j\beta}$ and $q < 2^{-\beta}$. Applying (9.6) with this η on each A_j, multiplying by w_j^2 and summing gives (7.6) for $\widetilde{\chi}$. The series also converges in the norm of

the form: testing (9.6) for χ_j against a cutoff function equal to one on a neighbourhood of $\overline{A}$ gives $\mathcal{E}(\chi_j,\chi_j) \le C\, m(A)/\Psi_S(r_j)$, and $\sum_j w_j\, 2^{j\beta/2} < \infty$. Hence $\widetilde{\chi} \in \mathcal{F}$. This proves (ii). At matching scales $H(s) \asymp s^2$, which gives (iii). □

Corollary 9.4 (Power scales)**.** *In the setting of Proposition 9.3, suppose that* $\Psi_P(r) \asymp r^a$ *and* $\Psi_{C,2}(r) \asymp r^b$ *at small radii. If* $b - a < 2$*, then*

$$\mathrm{CSA}(r^\beta) \quad \textit{holds at small widths, with} \qquad \beta = \frac{2a}{2+a-b}\,.$$

Moreover, for each annulus as in Proposition 9.3, the cutoff function of Step 1 of its proof satisfies, for all $0 < \eta \le 1$*,*

$$\int_A f^2\, d\Gamma(\chi) \le \eta\, \Gamma(f)(A) + C\, \eta^{1-\beta/2}\, r^{-\beta} \int_A f^2\, dm\,.$$

One has $a \le b$*,* $b \ge 2$ *and* $\beta \ge b$*, with* $\beta = a$ *when* $a = b$*. If, in addition,* $\mathrm{V}(\Phi)$ *holds with* $\Phi(r) \asymp r^Q$ *and* $\Theta(r) \asymp r^s$*, then* $b = 2 + Q - s$*.*

Proof. Here $H(s) \asymp s^{2+a-b}$, and (9.5) gives the exponent β. Propositions 3.7 and 3.6 give $a \le b$ and $b \ge 2$, and $\beta \ge b$ is equivalent to $(b-a)(b-2) \ge 0$. In (9.4) take $t = c\,(\eta r^2)^{1/(2+a-b)}$; since $a \le b$, this satisfies $t \le cr$ after reducing the range of widths. The coefficient of the energy is then at most η, and the coefficient of the zeroth-order term is $C\, r^{-2} t^{2-b} = C\, \eta^{1-\beta/2} r^{-\beta}$. The last assertion is (3.14). □

Thus a Poincaré scale r^a and a cutoff scale r^b with $a \le b < a + 2$ produce a cutoff Sobolev inequality at the scale $r^{2a/(2+a-b)}$, whose exponent is in general larger than both. With this scale, Theorem 7.5 and the estimates of Chapter 8 for the exponents a and β apply to the gradient structure. The condition $b - a < 2$ is needed for the summation in this argument.

The tree of cubes. The tree of cubes is compact, geodesic, doubling and 2-admissible by Proposition 15.5 and Theorem 15.8. Proposition 15.7 gives $\mathrm{B}^+(\Psi_{C,2})$ with $\Psi_{C,2}(r) = r^{Q+2-n}$, and Proposition 15.10 gives $\Psi_P(r) \asymp r^{Q+2-n-\sigma}$. Since $0 < \sigma < 2$, Corollary 9.4 yields

$$\mathrm{CSA}(r^\beta) \quad \text{at small widths, with} \qquad \beta = \frac{2(Q+2-n-\sigma)}{2-\sigma}\,.$$

9.4 Matching scales: the cutoff Sobolev inequality

When the Poincaré scale and the cutoff scale are comparable, no condition beyond $\mathrm{B}(\Psi_{C,2})$ is needed: the upper bound (9.2) for the capacity and the Poincaré inequality imply the cutoff Sobolev inequality at the common scale. This implication is a theorem of Eriksson-Bique [64]; we only have to check that its proof is local in scale.

Proposition 9.5. *Let* X *be proper and let the gradient structure be* 2*-admissible. Suppose that* VD*,* $\mathrm{PI}(\Psi_{P,2})$ *and* $\mathrm{B}(\Psi_{C,2})$ *hold, with* $\Psi_{P,2} \asymp \Psi_{C,2}$*. For every* $R > 0$ *and every sufficiently small* $0 < r \le R$*, there is a cutoff function* χ *for* $B(x,R) \subset B(x,R+r)$ *such that, with* $A = B(x,R+r) \setminus \overline{B(x,R)}$*,*

$$\int_X f^2\, d\Gamma(\chi) \le C \int_A d\Gamma(f) + \frac{C}{\Psi_{C,2}(r)} \int_A f^2\, dm\,, \qquad f \in \mathcal{F}_{\mathrm{loc}}(X)\,. \tag{9.7}$$

The constant is uniform in x*,* R *and* r*. Consequently* $\mathrm{CSA}(\Psi_{C,2})$ *holds at small widths. Under global hypotheses the estimate holds for every* $0 < r \le R$*. If* X *is connected and* $\mathrm{B}^+(\Psi_{C,2})$ *holds, the cutoff functions may be chosen Lipschitz.*

Proof. Step 1: balls. We apply the implication from the upper bound for the capacity and the Poincaré inequality to the cutoff Sobolev inequality in balls [64, Theorem 3.3 and Lemma 2.13], with $p = 2$ and the common scale $\Psi_{C,2}$. The capacity bound is (9.2), and the Poincaré inequality is (9.1). The resulting estimate (7.10) for a ball of radius h is local in scale: the capacity bound and the Poincaré inequality enter the variational argument and the Whitney construction of [64, Lemmas 3.1, 4.5 and 4.9] only at radii at most a fixed multiple of h. Indeed, the Whitney balls used in the two interpolations lie in a fixed dilation of the given ball and have radii bounded by its radius; the discrete approximations use arbitrarily small meshes; and the packing arguments use only doubling at small radii. Thus (7.10) holds with $S = \Psi_{C,2}$, uniformly for sufficiently small h.

Step 2: annuli. Proposition 7.3 gives (9.7), which is the unweighted inequality (7.7), and Lemma 7.8 gives the weighted inequality (7.6).

Step 3: Lipschitz cutoff functions. Under $\mathrm{B}^+(\Psi_{C,2})$, Proposition 9.3 applies with $H(s) \asymp s^2$ and gives Lipschitz cutoff functions. □

9.5 Matching scales: sub-Gaussian heat kernel estimates

In this section and the next, except in Proposition 9.9, the hypotheses are global: $r_0 = \infty$. We assume that $\Psi_{C,2}$ is a continuous increasing bijection of $(0,\infty)$ onto itself and that, for some $1 < \beta_- \le \beta_+ < \infty$,

$$c\Big(\frac{R}{r}\Big)^{\beta_-} \le \frac{\Psi_{C,2}(R)}{\Psi_{C,2}(r)} \le C\Big(\frac{R}{r}\Big)^{\beta_+}, \qquad 0 < r \le R. \tag{9.8}$$

Since $\Psi_{C,p}(r) = r^{p-2}\Psi_{C,2}(r)$, this implies (3.1) for every $p \ge 1$.

The rate function. The heat kernel estimates are expressed with the radius r_t associated with the time t and with the rate function (8.1) of the scale $\Psi_{C,2}$:

$$r_t = \Psi_{C,2}^{-1}(t), \qquad \mathcal{H}(R,t) = \mathcal{A}_{\Psi_{C,2}}(R,t) = \sup_{\rho>0}\Big(\frac{R}{\rho} - \frac{t}{\Psi_{C,2}(\rho)}\Big). \tag{9.9}$$

The supremum is finite by (9.8), and $\mathcal{H}(0,t) = 0$. For a power $\Psi_{C,2}(r) = r^\beta$,

$$r_t = t^{1/\beta}, \qquad \mathcal{H}(R,t) = c_\beta\Big(\frac{R^\beta}{t}\Big)^{1/(\beta-1)}, \qquad c_\beta = (\beta-1)\,\beta^{-\beta/(\beta-1)},$$

which is $R^2/(4t)$ for $\beta = 2$. We record three elementary properties of $\mathcal{H}$, which follow the estimates for general scales of [78, Section 3.3]; compare also [10, Lemma 1.5]. The constants depend only on those of (9.8).

Lemma 9.6. *Let $R \ge 0$, $t > 0$ and $\gamma = \beta_+/(\beta_+ - 1)$.*

(i) $\mathcal{H}(R,t) \le C$ *if* $R \le r_t$.

(ii) $\mathcal{H}(R,t) \ge c\,(R/r_t)^\gamma$ *if* $R \ge r_t$.

(iii) *For every $c_0 > 0$ there is $c_1 > 0$ such that $\mathcal{H}(c_0 R, t) \ge c_1\,\mathcal{H}(R,t)$.*

Proof. Write $\Psi = \Psi_{C,2}$ and $r = r_t$, so that $t = \Psi(r)$.

(i) Let $R \le r$. If $\rho \ge r$, then $R/\rho \le 1$. If $\rho < r$, then, with $s = r/\rho > 1$, the lower bound in (9.8) gives $R/\rho - t/\Psi(\rho) \le s - c\,s^{\beta_-}$, which is bounded from above because $\beta_- > 1$.

(ii) Let $R \ge r$. For $\rho \le r$ and $s = r/\rho$, the upper bound in (9.8) gives $R/\rho - t/\Psi(\rho) \ge (R/r)\,s - C\,s^{\beta_+}$. If R/r is larger than a suitable constant, the choice $s^{\beta_+ - 1} = R/(2Cr)$ gives

$\mathcal{H}(R,t) \ge c\,(R/r)^{\beta_+/(\beta_+-1)}$. For the remaining values of R/r, we use that $\mathcal{H}(\cdot,t)$ is nondecreasing and that $\mathcal{H}(r,t) \ge 1/A - \Psi(r)/\Psi(Ar) \ge 1/(2A)$ for a suitable fixed $A > 1$.

(iii) If $c_0 \ge 1$, this follows from the monotonicity in R. Let $0 < c_0 < 1$. By (9.8) there is $A > 1$ such that $\Psi(A\rho) \ge (A/c_0)\Psi(\rho)$ for all $\rho > 0$. Restricting the supremum to the radii $A\rho$,

$$\mathcal{H}(c_0 R,t) \ge \sup_{\rho>0}\Big(\frac{c_0 R}{A\rho} - \frac{t}{\Psi(A\rho)}\Big) \ge \frac{c_0}{A}\,\mathcal{H}(R,t)\,. \qquad \square$$

Two-sided estimates.

Definition 9.7. We say that HKE($\Psi_{C,2}$) holds if $(\mathcal{E},\mathcal{F})$ is conservative and has a heat kernel $p_t(x,y)$ such that, for all $t > 0$ and m-almost all $x, y \in X$,

$$\frac{c_1}{m(B(x,r_t))}\,e^{-c_2\mathcal{H}(d(x,y),t)} \le p_t(x,y) \le \frac{c_3}{m(B(x,r_t))}\,e^{-c_4\mathcal{H}(d(x,y),t)}\,. \tag{9.10}$$

For $\Psi_{C,2}(r) = r^\beta$ with $\beta > 2$ these are the sub-Gaussian estimates with the walk dimension β, and for $\beta = 2$ they are the Gaussian estimates.

Theorem 9.8. *Let X be unbounded, proper and geodesic. Assume* VD, PI($\Psi_{P,2}$) *and* B($\Psi_{C,2}$) *at all radii, and $\Psi_{P,2} \asymp \Psi_{C,2}$. If $\Psi_{C,2}$ is a continuous increasing bijection satisfying* (9.8)*, then* HKE($\Psi_{C,2}$) *holds.*

Proof. Theorem 4.3 gives 2-admissibility, and Proposition 2.7 gives a strongly local regular Dirichlet form with the energy measures $d\Gamma(f) = |\nabla_2 f|^2\,d\nu$. The common scale $\Psi_{C,2}$ is the scale of the Poincaré inequality (9.1) on the domain of the form, and the scale of the cutoff Sobolev inequality of Proposition 9.5. Corollary 8.12 gives the conservativeness and the estimates (9.10). $\square$

The implication from volume doubling, the Poincaré inequality and the cutoff Sobolev inequality at one scale to HKE($\Psi_{C,2}$) is the characterization developed in [17, 10, 75]; the equivalence between parabolic Harnack inequalities and heat kernel estimates is treated in [18, Theorems 3.1–3.2]. What Theorem 9.8 adds, for gradient structures, is that the cutoff Sobolev inequality need not be assumed: by Proposition 9.5 it follows from the cutoff condition B($\Psi_{C,2}$). Under the hypotheses of the theorem, the parabolic Harnack inequality (8.46) also holds, with a uniform constant on the cylinders of radius r and time length $\Psi_{C,2}(r)$.

Power scales. Under the additional hypothesis V(Φ), if $\Phi(r) = r^Q$ and $\Theta(r) = r^s$, the common scale is r^β with

$$\beta = Q + 2 - s\,, \tag{9.11}$$

by (3.14). Here $s \le Q$ by Proposition 3.6, so that $\beta \ge 2$, and the estimates (9.10) read

$$c\,t^{-Q/\beta}\exp\Big[-C\Big(\frac{d(x,y)^\beta}{t}\Big)^{1/(\beta-1)}\Big] \le p_t(x,y) \le C\,t^{-Q/\beta}\exp\Big[-c\Big(\frac{d(x,y)^\beta}{t}\Big)^{1/(\beta-1)}\Big]\,.$$

The exponent β is equal to two exactly when $s = Q$, that is, when the gauge is comparable to the volume function. On a local tree, $s = 1$ and the corresponding time scale is $r\,\Phi(r)$; see [22, Theorem 5.1].

The compact case. On a compact space the hypotheses concern small radii, and the estimates hold at small times.

Proposition 9.9. *Suppose that (X,d) is compact and geodesic, and that* VD, $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}(\Psi_{P,2})$ *and* (3.4) *at $p=2$ hold for $0<r<r_0$. Assume that*

$$\Psi_{C,2}(r) \asymp r^{\beta}\,, \qquad 0<r<r_0\,, \qquad \beta \geq 2\,.$$

Then the form is conservative and admits a heat kernel. There are $t_0>0$ and positive constants c_1,c_2,c_3,c_4 such that, for $0<t\leq t_0$ and m-almost all x,y,

$$\frac{c_1}{m(B(x,t^{1/\beta}))}\exp\Big[-c_2\Big(\frac{d(x,y)^{\beta}}{t}\Big)^{1/(\beta-1)}\Big] \leq p_t(x,y) \leq \frac{c_3}{m(B(x,t^{1/\beta}))}\exp\Big[-c_4\Big(\frac{d(x,y)^{\beta}}{t}\Big)^{1/(\beta-1)}\Big]\,. \tag{9.12}$$

The constants may depend on r_0 and on X.

Proof. The arguments are those of Theorem 8.3 and Corollary 8.12; we check that they involve only small radii. Write $d=d(x,y)$.

Step 1: the form. Theorem 4.3 gives 2-admissibility, and Proposition 2.7 gives a strongly local regular Dirichlet form. Since X is compact, $1\in\mathcal{F}$ and $\mathcal{E}(1,1)=0$, so the semigroup is conservative. Proposition 9.5 gives (8.3) with $S(s)=s^{\beta}$, for all widths $s\leq s_0$ and uniformly in the inner radius, for some $s_0>0$. The local hypotheses of Chapter 7 hold, and Proposition 8.1 gives a heat kernel.

Step 2: doubling. Doubling at small radii extends, on a compact space, to doubling at all radii, with a constant depending on X: finitely many balls of a fixed small radius cover X, which gives a positive lower bound for the volume of all the balls of this radius. Consequently, for all $x,y\in X$ and $r>0$,

$$\frac{m(B(x,r))}{m(B(y,r))} \leq C\Big(1+\frac{d}{r}\Big)^{Q_D}\,.$$

Step 3: the upper bound. Set $r=t^{1/\beta}$ and choose $t_0>0$ so that all the fixed dilations of this radius are in the range of the hypotheses for $0<t\leq t_0$. The derivation of (8.9) involves only the balls $B(x,2r)$ and $B(y,2r)$, and its polynomial factor is bounded because $\Psi_{P,2}(r)\asymp\Psi_{C,2}(r)\asymp t$. Hence

$$p_t(x,y) \leq \frac{C}{m(B(x,r))\,m(B(y,r))}\sup_{t/4\leq s\leq t}\langle \mathbf{1}_{B(x,2r)}, P_s\mathbf{1}_{B(y,2r)}\rangle\,.$$

If $d\leq 8r$, the contraction property in L^2 bounds the inner product by $C\sqrt{m(B(x,r))\,m(B(y,r))}$. If $d>8r$, apply Lemma 8.2 with $p=1$, the centre x and the radius $d/3$; it is available for the parameters $\lambda\geq\max\{1,d/(3s_0)\}$, and (8.7) gives, as in [10, Proposition 2.3],

$$\langle \mathbf{1}_{B(x,2r)}, P_s\mathbf{1}_{B(y,2r)}\rangle \leq C\sqrt{m(B(x,r))\,m(B(y,r))}\,\exp\left[-\lambda + C\,t\,\lambda^{\beta}d^{-\beta}\right]\,, \qquad t/4\leq s\leq t\,.$$

Choose $\lambda = c\,(d^{\beta}/t)^{1/(\beta-1)}$ with a sufficiently small fixed $c>0$, so that the exponent is at most $-\lambda/2$. When this value of λ is less than one, the exponential in (9.12) is bounded from below and the contraction property suffices. Otherwise the value is admissible after reducing t_0: since $d>8t^{1/\beta}$, the width $d/(3\lambda)$ is comparable to $(t/d)^{1/(\beta-1)}\leq C\,t^{1/\beta}$, hence at most s_0. This gives the upper bound with the prefactor $(m(B(x,r))\,m(B(y,r)))^{-1/2}$. By Step 2 the ratio of the two volumes is at most $C(1+d/r)^{Q_D}$, and this factor is absorbed by the exponential, because $(d/r)^{\beta/(\beta-1)} = (d^{\beta}/t)^{1/(\beta-1)}$.

Step 4: the lower bound. Lemma 8.11 at the radius $t^{1/\beta}$ gives, for $0 < t \le t_0$,

$$p_t(x,y) \ge \frac{c}{m(B(x,t^{1/\beta}))} \qquad \text{if } d \le t^{1/\beta}\,.$$

Let $\delta > 0$ be small and let N be the least positive integer such that $d/N \le \delta\,(t/N)^{1/\beta}$; then $N \le 1 + C(d^\beta/t)^{1/(\beta-1)}$. Chain along a geodesic from x to y with the time step t/N, as in the proof of Corollary 8.12. All the balls which occur have radii at most a fixed multiple of $t^{1/\beta}$, and are in the range of the hypotheses. The intermediate volumes cancel as in that proof, and

$$p_t(x,y) \ge \frac{c^N}{m(B(x,t^{1/\beta}))}\,, \qquad 0 < c < 1\,,$$

which is the lower bound in (9.12). This is the local version, for a power scale, of the argument of [78, Section 6]. □

9.6 Heat kernel characterization of Sobolev and BV functions

The Korevaar–Schoen characterizations of Chapter 4 have a counterpart in terms of the heat kernel: the mean oscillation $D_p(f,r)$ at the scale r is replaced by the oscillation averaged by the heat kernel at the time t, with $r = r_t$. We keep the global hypotheses of Section 9.5 on doubling and on the scale $\Psi_{C,2}$, and we assume that the structure is 2-admissible and that its Dirichlet form has a measurable heat kernel. For $f \in L^p(X,m)$ and $t > 0$, let

$$I_p(f,t) = \int_X \int_X |f(x) - f(y)|^p\, p_t(x,y)\, dm(y)\, dm(x)\,, \qquad N_p(f,t) = \frac{I_p(f,t)}{\Psi_{C,p}(r_t)}\,.$$

For $\Psi_{C,2}(r) = r^\beta$ one has $\Psi_{C,p}(r_t) = t^{p\alpha_p}$ with $\alpha_p = (\beta + p - 2)/(p\beta)$, and $\sup_{t>0} N_p(f,t)^{1/p}$ is the seminorm of the Besov class $\mathbf{B}^{p,\alpha_p}(X)$ defined by the heat semigroup in [1]. In the Euclidean space, $\beta = 2$ and $N_p(f,t) = t^{-p/2} I_p(f,t)$.

Related results. For general p and power scales, the comparison between the seminorms defined by the metric and by the heat kernel was established by Pietruska-Pałuba [134, Theorem 3.2]; see also [2, Proposition 4.2] and [3, Theorem 2.4]. The identifications with Sobolev and BV spaces are developed in [1, 2, 3, 122]. Related comparisons with Besov spaces defined by the semigroup, under different restrictions on the smoothness, are proved in [76, Theorems 1.5–1.6]. Here the identification with the energy of the closed gradient rests on Chapter 4.

Theorem 9.10. *Let $1 < p < \infty$. Assume that the structure is p-admissible and that* VD *and* $\mathrm{B}(\Psi_{C,2})$ *hold at all scales.*

(i) *If* $\mathrm{PI}(\Psi_{P,p})$ *holds at all scales and, for all $t > 0$ and m-almost all x, y,*

$$p_t(x,y) \le \frac{C}{m(B(x,r_t))}\, e^{-c\,\mathcal{H}(d(x,y),t)}\,, \tag{9.13}$$

then

$$\sup_{t>0} \frac{I_p(f,t)}{\Psi_{P,p}(r_t)} \le C\,\mathcal{E}_p(f)\,, \qquad f \in L^p(X,m)\,. \tag{9.14}$$

(ii) *If there are* $c > 0$ *and* $0 < \varepsilon \le 1$ *such that, for all* $t > 0$ *and* m*-almost all* x, y,

$$p_t(x,y) \ge \frac{c}{m(B(x,r_t))} \qquad \text{when } d(x,y) < \varepsilon\, r_t\,, \tag{9.15}$$

then

$$\mathcal{E}_p(f) \le C \liminf_{t\to 0} N_p(f,t)\,, \qquad f \in L^p(X,m)\,. \tag{9.16}$$

At $p = 1$*, the same implications hold with* $\|Df\|(X)$ *in place of* $\mathcal{E}_p(f)$*, without any assumption of admissibility at the exponent one. Neither implication requires the semigroup to be conservative.*

As in Theorem 4.4, the upper bound involves the Poincaré scale and the lower bound the cutoff scale; the upper bound uses only an upper estimate of the heat kernel, and the lower bound only a near-diagonal lower estimate.

Proof. (i) We use the decomposition into annuli of [2, proof of Proposition 4.2] and [3, proof of Theorem 2.4], with scale functions in place of powers. Put $r = r_t$. On the annulus $2^{k-1}r \le d(x,y) < 2^k r$, $k \ge 1$, Lemma 9.6(ii) and (9.13) bound the kernel by $C\,m(B(x,r))^{-1}e^{-c2^{k\gamma}}$, where $\gamma = \beta_+/(\beta_+ - 1) > 1$. The same bound with $k = 0$ holds in $B(x,r)$. Doubling gives $m(B(x,2^k r))/m(B(x,r)) \le C\,2^{kD}$ for a fixed exponent D. Consequently, if b is an upper scaling exponent of $\Psi_{P,p}$,

$$\begin{aligned} I_p(f,t) &\le C \sum_{k\ge 0} 2^{kD} e^{-c2^{k\gamma}} D_p(f,2^k r) \\ &\le C\,\Psi_{P,p}(r) \sup_{\rho>0} \frac{D_p(f,\rho)}{\Psi_{P,p}(\rho)} \sum_{k\ge 0} 2^{k(D+b)} e^{-c2^{k\gamma}}\,, \end{aligned}$$

and the series converges. The first inequality in (4.6), with $\overline{\mathcal{E}}_p = \mathcal{E}_p$ for $p > 1$ by Theorem 4.4(ii) and $\overline{\mathcal{E}}_1(f) = \|Df\|(X)$ for $p = 1$, proves (9.14).

(ii) Restricting the integral to the pairs with $d(x,y) < \varepsilon r_t$, as in [134, proof of Theorem 3.2], the bound (9.15) and doubling give

$$\begin{aligned} I_p(f,t) &\ge c \int_X \frac{1}{m(B(x,r_t))} \int_{B(x,\varepsilon r_t)} |f(x) - f(y)|^p \, dm(y)\, dm(x) \ge c'\, D_p(f,\varepsilon r_t)\,, \\ N_p(f,t) &\ge c'' \frac{D_p(f,\varepsilon r_t)}{\Psi_{C,p}(\varepsilon r_t)}\,. \end{aligned}$$

The second line also uses (3.2) for $\Psi_{C,p}$. Since $t \mapsto \varepsilon r_t$ is an increasing bijection of $(0,\infty)$ onto itself, the second inequality in (4.6) proves (9.16). That inequality uses neither $\mathrm{PI}(\Psi_{P,p})$ nor (3.4), and at $p = 1$ it gives the statement for $\|Df\|(X)$. □

Corollary 9.11. *Let* $1 < p < \infty$*. Assume* $\mathrm{HKE}(\Psi_{C,2})$*, and assume that* VD, $\mathrm{B}(\Psi_{C,2})$, $\mathrm{PI}(\Psi_{P,p})$ *and* (3.4) *hold at all scales. Then*

$$\sup_{t>0} N_p(f,t) \le C\,\mathcal{E}_p(f)\,, \qquad \mathcal{E}_p(f) \le C \liminf_{t\to 0} N_p(f,t)\,, \qquad f \in L^p(X,m)\,. \tag{9.17}$$

In particular, $f \in L^p(X,m)$ *belongs to* $W^{1,p}(X,m)$ *if and only if* $\sup_{t>0} N_p(f,t) < \infty$*. If* $\mathrm{PI}(\Psi_{P,1})$ *and* (3.4) *at* $p = 1$ *hold at all scales, the same inequalities hold with* $p = 1$ *and* $\|Df\|(X)$ *in place of* $\mathcal{E}_p(f)$*, and* $f \in L^1(X,m)$ *belongs to* $BV(X,m)$ *if and only if* $\sup_{t>0} N_1(f,t) < \infty$*.*

Proof. For $p > 1$, the structure is p-admissible by Theorem 4.3. The upper bound in (9.10) is (9.13). By Lemma 9.6(i), $\mathcal{H}(R,t)$ is bounded when $R \le r_t$, so the lower bound in (9.10) gives (9.15) with $\varepsilon = 1$. Apply Theorem 9.10 and (3.4). The same argument applies at $p = 1$. The characterizations of $W^{1,p}(X,m)$ and $BV(X,m)$ follow, because $\mathcal{E}_p(f) < \infty$ exactly when $f \in W^{1,p}(X,m)$, and $\|Df\|(X) < \infty$ exactly when $f \in BV(X,m)$. □

Part III
Examples of gradient structures

10 Riemannian and sub-Riemannian structures

This part of the book describes the gradient structures summarized in Table 1. For each example, the following points are verified, in this order.

(a) *The structure.* The Leibniz rule, the bound $|\nabla u| \le \mathrm{Lip}(u)$ and the weak-$*$ continuity of condition (G); the fibres and the rank.

(b) *Recovery of the metric.* The local Sobolev-to-Lipschitz property (1.11), and its uniformity.

(c) *Admissibility.* Conditions (A1) and (A2) of Definition 2.1, and the identification of $W^{1,p}(X,m)$ with a known Sobolev space.

(d) *Quantitative hypotheses.* Volume doubling, the cutoff condition, the Poincaré inequalities, and their scales.

In the present chapter the gradient is the classical differential, or its horizontal part, the gradient measure is the reference measure, and all the scales are Gaussian: $\Psi_{C,p}(r) = \Psi_{P,p}(r) = r^p$. The geometric assumptions needed for the uniform quantitative estimates are stated separately from those needed for the structure itself.

10.1 Riemannian manifolds

On a complete Riemannian manifold the gradient measure is the Riemannian volume, and the Sobolev spaces are the classical ones; see [81, Chapter 3]. We begin with the fact which gives the Sobolev-to-Lipschitz property: an almost everywhere bound on the gradient of a Lipschitz function controls its increments along geodesics. Its proof uses absolute continuity on almost every member of a family of curves, as in [79, Proposition 10.1], and then continuity to recover the prescribed geodesic.

Lemma 10.1. *Let (M,g) be a Riemannian manifold with volume measure m, let $\gamma : [0,\ell] \to M$ be an injective unit-speed geodesic, and let u be a locally Lipschitz function on a neighbourhood U of $\gamma([0,\ell])$, with $|\operatorname{grad} u| \le L$ m-almost everywhere on U. Then*

$$|u(\gamma(\ell)) - u(\gamma(0))| \le L\,\ell\,.$$

Proof. Let $E_1, \dots, E_{n-1}$ be a parallel orthonormal frame of the normal bundle of γ, and let

$$F(t,s) = \exp_{\gamma(t)}\Big(\sum_{i=1}^{n-1} s_i E_i(t)\Big)\,, \qquad t \in [0,\ell]\,,\ s \in \mathbb{R}^{n-1}\,,\ |s| < \varepsilon\,.$$

Since γ is injective and $[0,\ell]$ is compact, for ε small enough F is a diffeomorphism onto its image, which is contained in U; these are Fermi coordinates along γ. Let $Z \subset U$ be the set of points at which u is not differentiable or $|\operatorname{grad} u| > L$. It is m-null, by Rademacher's theorem and by the hypothesis. Then $F^{-1}(Z)$ is a Lebesgue null set and, by Fubini's theorem, for almost every s the set $\{t : F(t,s) \in Z\}$ has zero length. For such s, the function $t \mapsto u(F(t,s))$ is Lipschitz on $[0,\ell]$ and has the derivative $\langle \operatorname{grad} u, \partial_t F\rangle$ at almost every t, hence

$$|u(F(\ell,s)) - u(F(0,s))| \le L \int_0^\ell |\partial_t F(t,s)|\,dt\,.$$

Letting $s \to 0$ along such values, the left-hand side converges to $|u(\gamma(\ell)) - u(\gamma(0))|$ and the integral converges to ℓ. □

Proposition 10.2. *Let (M, g) be a complete Riemannian manifold of dimension n, with distance d and volume measure m. Let $(e_1, \dots, e_n)$ be a measurable orthonormal frame of TM, and for $u \in \mathrm{Lip}_b(M)$ let*

$$\nabla u = (\langle \operatorname{grad} u, e_i \rangle)_{1 \le i \le n} .$$

Then (M, m, ∇) is a uniform gradient structure of constant rank n on (M, d, m). It is p-admissible for every $p \in [1, \infty)$, $\mathcal{C}_p = \mathrm{Lip}_c(M)$, and $W^{1,p}(M, m)$ is the classical Sobolev space.

Proof. By Rademacher's theorem, $\operatorname{grad} u$ exists m-almost everywhere and $|\nabla u| = |\operatorname{grad} u| \le \mathrm{Lip}(u)$.

Step 1: the derivation and its fibres. The Riemannian gradient satisfies the Leibniz rule. To verify the weak-$*$ continuity, let (u_k) be as in (G) and let V be a smooth compactly supported vector field. Integration by parts and dominated convergence give

$$\int_M \langle \operatorname{grad} u_k, V \rangle \, dm = - \int_M u_k \operatorname{div} V \, dm \longrightarrow - \int_M u \operatorname{div} V \, dm = \int_M \langle \operatorname{grad} u, V \rangle \, dm .$$

Smooth compactly supported fields are dense in the integrable vector fields, and the uniform bound on $\operatorname{grad} u_k$ extends the convergence to every integrable field. For the fibres, cover M by countably many relatively compact coordinate balls B_i, with coordinates $(x_i^1, \dots, x_i^n)$ defined on a neighbourhood of $\overline{B_i}$, and let $u_{i,j}$ be a smooth compactly supported function equal to x_i^j on B_i. The gradients $\operatorname{grad} u_{i,1}, \dots, \operatorname{grad} u_{i,n}$ are bounded with compact support, and they span $T_x M$ at every $x \in B_i$. Hence $H_x = \mathbb{R}^n$ for almost every x.

Step 2: recovery of the metric. We take $\iota \equiv 1$. Let $y, z \in B(x, 1)$. By completeness there is a minimizing geodesic γ from y to z. It is injective, its length is $d(y, z) < 2$, and it lies in $B(x, 3)$. Lemma 10.1, applied with $U = B(x, 3)$ and $L = \|\nabla u\|_{L^\infty(B(x,3),m)}$, gives (1.11).

Step 3: admissibility. The gradient of a compactly supported Lipschitz function is bounded with compact support, so $\mathcal{C}_p = \mathrm{Lip}_c(M)$, which is uniformly dense in $C_c(M)$; this is (A1). Let $u_k \in \mathrm{Lip}_c(M)$ with $u_k \to 0$ in $L^p(M, m)$ and $\operatorname{grad} u_k \to G$ in L^p. For every smooth compactly supported vector field V,

$$\int_M \langle G, V \rangle \, dm = \lim_{k \to \infty} \int_M \langle \operatorname{grad} u_k, V \rangle \, dm = - \lim_{k \to \infty} \int_M u_k \operatorname{div} V \, dm = 0 ,$$

hence $G = 0$. This proves (A2) for every $1 \le p < \infty$.

Step 4: identification. The classical space $W^{1,p}(M)$ of the functions of L^p with weak gradient in L^p is a Banach space which contains $\mathrm{Lip}_c(M)$, and on which the weak gradient is a closed operator extending ∇. On a complete manifold, smooth compactly supported functions are dense in $W^{1,p}(M)$ [81, Theorem 3.1]. Therefore the closure of ∇ on $\mathcal{C}_p$ is the weak gradient on $W^{1,p}(M)$. □

Quantitative estimates. A lower bound on the Ricci curvature gives the three hypotheses of Chapter 3 at small radii.

Proposition 10.3. *Let M be a complete n-dimensional Riemannian manifold with $\mathrm{Ric} \ge -(n-1)K$, $K \ge 0$, and let m be the Riemannian volume. At small radii, the structure of Proposition 10.2 satisfies*

$$\mathrm{VD} , \qquad \mathrm{B}(r^2) , \qquad \mathrm{PI}_p(r^p) \quad (1 \le p < \infty) .$$

If $\inf_x m(B(x, 1)) > 0$, then $\mathrm{V}(r^n)$ also holds. For $K = 0$, the doubling property, the cutoff condition and the Poincaré inequalities hold at every radius.

Proof. The Bishop–Gromov comparison theorem gives doubling uniformly on bounded ranges of radii, and at every radius when $K = 0$; see [79, Section 10.1]. Since $\nu = m$, Lemma 3.8 gives $\mathrm{B}(r^2)$; no lower bound on the volume of unit balls is needed. The 1-Poincaré inequality of Buser under a lower Ricci bound [46], and its consequence in L^p on doubling balls [139, Section 5.6], give $\mathrm{PI}_p(r^p)$. For $K = 0$ the constants are independent of the radius. Finally the Bishop–Gromov theorem also gives

$$c(n,K)\, r^n\, m(B(x,1)) \le m(B(x,r)) \le C(n,K)\, r^n \qquad (0 < r \le 1)\,,$$

so that a lower bound on the volume of unit balls implies $\mathrm{V}(r^n)$. □

The scales match at every exponent: $\Psi_{C,p}(r) = \Psi_{P,p}(r) = r^p$. Under $\mathrm{V}(r^n)$ the gauge is $\Theta(r) = r^n$. Consequently the results of Part I at matching scales apply for every $1 \le p < \infty$, in particular the Korevaar–Schoen characterizations of $W^{1,p}$ and BV (Theorems 4.4 and 4.8). When $K = 0$ and M is noncompact, Theorem 9.8 gives the two-sided Gaussian bounds for the heat kernel.

10.2 Measurable Riemannian metrics

The metric tensor need not be continuous. For a measurable tensor, the length of curves is not a meaningful starting point, since a curve is a null set. The distance is defined instead by duality with functions, so that changing the tensor on a set of measure zero does not change the distance. We use the construction of the intrinsic distance of a strongly local Dirichlet form [147, Section 5]; for measurable elliptic tensors and the distinction between the structures defined by the distance and by the differential, see [115, Sections 2–4].

Proposition 10.4. *Let (M, g_0) be a complete connected smooth Riemannian manifold of dimension n, with distance d_0. Let g be a measurable symmetric tensor satisfying $C^{-1} g_0 \le g \le C g_0$ almost everywhere, for some $C \ge 1$, and set $m = \nu = \mathrm{vol}_g$. Define*

$$d_g(x,y) = \sup\bigl\{u(x) - u(y) :\ u \in W^{1,\infty}_{\mathrm{loc}}(M,g_0) \cap C(M),\ |du|_{g^{-1}} \le 1 \ \text{a.e.}\bigr\}\,. \tag{10.1}$$

Choose a measurable g-orthonormal frame $(e_1, \dots, e_n)$ and put $\nabla u = (du(e_i))_{1 \le i \le n}$. Then (M, ν, ∇) is a uniform gradient structure of constant rank n on (M, d_g, m). It is p-admissible for every $1 \le p < \infty$, with $\mathcal{C}_p = \mathrm{Lip}_c(M, d_g)$ and

$$W^{1,p}(M,m) = W^{1,p}(M,g_0)\,, \qquad \mathcal{E}_p(u) = \int_M |du|^p_{g^{-1}}\, d\mathrm{vol}_g\,.$$

The two Sobolev norms are equivalent, and (M, d_g) is proper and geodesic.

Proof. Step 1: comparison with g_0. We claim that

$$C^{-1/2} d_0 \le d_g \le C^{1/2} d_0\,, \qquad C^{-n/2}\, \mathrm{vol}_{g_0} \le \mathrm{vol}_g \le C^{n/2}\, \mathrm{vol}_{g_0}\,. \tag{10.2}$$

The comparison of the volumes follows from $C^{-n} \det g_0 \le \det g \le C^n \det g_0$ in coordinates. The hypothesis on g is equivalent to $C^{-1} g_0^{-1} \le g^{-1} \le C g_0^{-1}$ on covectors. If u is admissible in (10.1), then $|du|_{g_0^{-1}} \le C^{1/2} |du|_{g^{-1}} \le C^{1/2}$ almost everywhere, so u is $C^{1/2}$-Lipschitz for d_0 by Lemma 10.1, applied along minimizing geodesics of g_0. Hence $d_g \le C^{1/2} d_0$. Conversely $u = C^{-1/2} d_0(\cdot, y)$ satisfies $|du|_{g^{-1}} \le C^{1/2} |du|_{g_0^{-1}} \le 1$, hence $d_g(x,y) \ge C^{-1/2} d_0(x,y)$. It follows that d_g is a distance which induces the topology of M, that (M, d_g) is complete, and that its closed balls are compact.

Step 2: the intrinsic distance. The form $\int_M |du|^2_{g^{-1}}\,dm$, with domain $W^{1,2}(M,g_0)$, is a regular strongly local Dirichlet form on $L^2(M,m)$ with energy measures $|du|^2_{g^{-1}}\,m$. By ellipticity, a continuous function which belongs locally to the domain and satisfies $|du|_{g^{-1}} \le 1$ has bounded weak derivatives in every relatively compact chart; hence the intrinsic distance of this form is d_g. Since d_g induces the original topology, [147, Theorems 5.1 and 5.2] apply: every function u which is Lipschitz for d_g satisfies

$$|\nabla u| = |du|_{g^{-1}} \le \mathrm{Lip}_{d_g}(u) \qquad \text{almost everywhere,} \tag{10.3}$$

and d_g is a length metric. A complete locally compact length space is proper and geodesic, by the Hopf–Rinow theorem.

Step 3: the derivation and its fibres. The differential satisfies the Leibniz rule, and (10.3) gives the Lipschitz bound. Let (u_k) be as in (G). By (10.2) the sequence is equi-Lipschitz for d_0, so in each relatively compact chart the derivatives $\partial_j u_k$ are bounded in L^∞ and converge to $\partial_j u$ in the sense of distributions, hence weak-$*$ in L^∞. Let $G \in L^1(M,\mathbb{R}^n,\nu)$ be supported in such a chart. Writing $e_i = \sum_j e_i^j \partial_j$ and $d\nu = \rho\,dx$,

$$\int_M \langle \nabla u_k, G\rangle\,d\nu = \sum_{i,j} \int \partial_j u_k\; e_i^j\, G_i\,\rho\,dx \longrightarrow \int_M \langle \nabla u, G\rangle\,d\nu\,,$$

because the functions $e_i^j G_i \rho$ are integrable. For a general $G \in L^1(M,\mathbb{R}^n,\nu)$, decompose M into countably many disjoint Borel sets A_l, each contained in a relatively compact chart. The contribution of $\bigcup_{l>l_0} A_l$ to both integrals is at most $\sup_k \mathrm{Lip}_{d_g}(u_k) \int_{\bigcup_{l>l_0} A_l} |G|\,d\nu$, which tends to zero as $l_0 \to \infty$, and the contribution of each A_l converges. This proves the weak-$*$ continuity. As in the proof of Proposition 10.2, the compactly supported smooth functions which agree with the coordinates on the balls of a countable cover have differentials spanning $T^*_x M$, so their gradients span $\mathbb{R}^n$ in the chosen frame.

Step 4: a local Lipschitz bound. Let $U \subset M$ be open, and let $u \in \mathrm{Lip}_b(M,d_g)$ and $L>0$ with $|\nabla u| \le L$ almost everywhere on U. We claim that every $x \in U$ has a neighbourhood on which u is L-Lipschitz for d_g. The truncation below is a direct form of the passage from local to global bounds for intrinsic Lipschitz constants; compare [115, Lemma 4.4] and [147, Lemma 5.4]. Choose a relatively compact open set V with $x \in V \subset \overline{V} \subset U$ and $V \ne M$, let $\rho = d_g(\cdot, M\setminus V)$, and define on M

$$v = \big((u-u(x)) \wedge L\rho\big) \vee (-L\rho)\,.$$

The function v is Lipschitz and vanishes outside V. At almost every point, dv is equal to du, to $L\,d\rho$ or to $-L\,d\rho$, according to which of the three functions $u-u(x)$, $L\rho$, $-L\rho$ is equal to v, and $dv = 0$ almost everywhere on $M\setminus V$. Since $|du|_{g^{-1}} \le L$ almost everywhere on V, and $|d\rho|_{g^{-1}} \le 1$ by (10.3), we get $|dv|_{g^{-1}} \le L$ almost everywhere on M. Thus v/L is admissible in (10.1), and v is L-Lipschitz for d_g on M. Since $\rho(x) > 0$ and u is continuous, $v = u - u(x)$ on a neighbourhood of x, which proves the claim.

Step 5: recovery of the metric. We take $\iota \equiv 1$. Let $y,z \in B_{d_g}(x,1)$ and $L > \|\nabla u\|_{L^\infty(B_{d_g}(x,3),\nu)}$. A minimizing geodesic for d_g from y to z exists by Step 2, and it lies in $U = B_{d_g}(x,3)$. By Step 4 and compactness, it can be divided into finitely many consecutive arcs on each of which u is L-Lipschitz. Adding the corresponding inequalities gives $|u(y)-u(z)| \le L\,d_g(y,z)$, and letting L decrease to $\|\nabla u\|_{L^\infty(B_{d_g}(x,3),\nu)}$ proves (1.11).

Step 6: admissibility and Sobolev spaces. Since ν is locally finite, $\mathcal{C}_p = \mathrm{Lip}_c(M,d_g)$, which is uniformly dense in $C_c(M)$. By (10.2), the norm $\big(\|u\|^p_{L^p(m)} + \int_M |du|^p_{g^{-1}}\,dm\big)^{1/p}$ is equivalent

to the norm of $W^{1,p}(M, g_0)$. The weak differential is a closed operator on $W^{1,p}(M, g_0)$ which extends ∇, so ∇ is closable on $\mathcal{C}_p$. Finally $C_c^\infty(M) \subset \mathcal{C}_p$ is dense in $W^{1,p}(M, g_0)$ by completeness [81, Theorem 3.1], so the closure of ∇ on $\mathcal{C}_p$ is the weak differential on $W^{1,p}(M, g_0)$. □

Remark 10.5. In coordinates, the estimate (10.3) and its local converse in L^∞ are also proved in [115, Lemmas 2.3 and 4.4]. They do not assert the pointwise equality between $|du|_{g^{-1}}$ and the metric slope of u; that equality can fail for measurable tensors [115, Remark 4.5]. Nor should (10.1) be replaced by the infimum of the lengths of curves computed with an arbitrary representative of g: assigning a smaller tensor on a curve, which has zero volume, can change that infimum without changing any weak differential or any energy.

Quantitative estimates.

Proposition 10.6. *Under the hypotheses of Proposition 10.4, assume that g_0 satisfies the lower Ricci bound of Proposition 10.3. Then* VD, $\mathrm{B}(r^2)$ *and* $\mathrm{PI}_p(r^p)$ *hold at small radii for every $1 \le p < \infty$. If the g_0-volumes of unit balls are uniformly bounded from below, then* $\mathrm{V}(r^n)$ *also holds.*

Proof. By (10.2), a ball for d_g lies between two concentric balls for d_0 with radii comparable to its own, and the two volume measures and the two norms of covectors are uniformly comparable. Doubling, and the uniform bounds for the volume when assumed, therefore pass from g_0 to g. For the Poincaré inequality, bound the oscillation on a ball for d_g about the g_0-mean on a ball for d_0 which contains it, and enlarge the ball for d_g by a fixed factor so that it contains the region on which the g_0-energy is integrated. Since $\nu = m$ and (M, d_g) is proper, Lemma 3.8 applies to the distance cutoff for d_g, whose gradient is bounded by $2/r$ by (10.3). □

If g is only locally comparable to g_0, the same statements hold with constants which are uniform on compact sets. Uniform ellipticity relative to the Euclidean metric gives the three estimates at every radius on $\mathbb{R}^n$.

The harmonic Sierpiński gasket is a further measurable Riemannian example, of rank one, which is not a local tree. It is treated in Chapter 13.

10.3 Sub-Riemannian manifolds

Let M be a connected smooth manifold, H a smooth bracket-generating subbundle of TM of rank k with a smooth inner product, and m a positive smooth measure. Equip M with the Carnot–Carathéodory distance d, obtained by minimizing the lengths of the curves tangent to H. This distance induces the manifold topology [132, 79]. In a local orthonormal frame $X_1, \dots, X_k$ of H, the horizontal gradient of a smooth function u is $\nabla_H u = \sum_i (X_i u) X_i$; it is represented by $(X_1 u, \dots, X_k u)$, and $|\nabla_H u|^2 = \sum_i (X_i u)^2$. For a locally integrable u, the horizontal gradient is defined in the sense of distributions, by the formula (10.5) below. To obtain a gradient with values in $\mathbb{R}^k$, we express it in a measurable orthonormal frame of H, as in Proposition 10.2.

Proposition 10.7. *The distributional horizontal gradient ∇_H defines a uniform gradient structure on (M, d, m) with $\nu = m$ and constant rank k. It is p-admissible for every $1 \le p < \infty$, and*

$$W^{1,p}(M, m) = \overline{C_c^\infty(M)}^{\,\|\cdot\|_{L^p(m)} + \|\nabla_H \cdot\|_{L^p(m)}}. \tag{10.4}$$

If (M, d) is complete, this is the space of all the functions of $L^p(M, m)$ whose distributional horizontal gradient belongs to L^p.

Proof. Step 1: the derivation and its fibres. For a Lipschitz function u, the derivative along the flow of each smooth unit horizontal field agrees with the distributional derivative and is bounded by $\mathrm{Lip}(u)$, by [79, Theorem 11.7 and Corollary 11.8]. Taking a countable dense family of unit linear combinations of a local frame gives $|\nabla_H u| \le \mathrm{Lip}(u)$. The product rule along each flow gives the Leibniz rule for ∇_H. For a smooth horizontal field W of compact support,

$$\int_M \langle \nabla_H u, W \rangle \, dm = -\int_M u \, \mathrm{div}_m W \, dm \,. \tag{10.5}$$

The density of these fields in $L^1(H, m)$ and the bound on the gradient give the weak-$*$ continuity. The smooth compactly supported functions which are equal to coordinates on smaller charts have bounded horizontal gradients, are Lipschitz for d by integration along horizontal curves, and their gradients span the horizontal fibres.

Step 2: recovery of the metric. Suppose that $|\nabla_H u| \le L$ almost everywhere on an open set U. In a chart, mollify u. For each smooth field X_i of the frame, the Friedrichs commutator between X_i and the convolution is an integral of $u(y) - u(x)$ against a kernel with uniformly bounded L^1 norm, supported at a distance at most the radius of the mollifier; since u is continuous, it tends to zero uniformly on compact subsets of the chart. Thus the smooth approximations converge uniformly and have horizontal gradients bounded by $L + o(1)$ there; see [79, Theorem 11.9 and Proposition 11.10]. Integrating along the segments of a horizontal curve in U, with a finite subdivision, bounds the variation of u along the curve by L times its length. Let now $y, z \in B(x, 1)$. A horizontal curve from y to z of length less than $d(y, z) + \varepsilon$, with ε small, lies in $B(x, 3)$. Applying the bound with $U = B(x, 3)$ and letting $\varepsilon \to 0$ proves (1.11) with $\iota \equiv 1$.

Step 3: admissibility. The gradient is closable in every L^p by (10.5). Since $\nu = m$ is a Radon measure, $\mathcal{C}_p = \mathrm{Lip}_c(M)$, and the smooth functions in this core are uniformly dense in $C_c(M)$. A partition of unity and the local Friedrichs approximation in L^p [79, Theorem 11.9] give (10.4). If M is complete, the distance cutoffs χ_R at the scale R are compactly supported and satisfy $|\nabla_H \chi_R| \le C/R$. Multiplying a function of the distributional Sobolev space by χ_R gives convergence in the graph norm as $R \to \infty$, by the product rule, and the compactly supported approximations can be smoothed by the same local argument. □

At $p = 2$, $d\Gamma(u) = |\nabla_H u|^2 \, dm$, and the generator of the Dirichlet form on $C_c^\infty(M)$ is the sub-Laplacian $L = \mathrm{div}_m \nabla_H$.

Equiregular structures. Let H^j be the bundle spanned by the brackets of length at most j of sections of H. Suppose that the H^j have constant dimensions n_j, with $n_0 = 0$; the structure is then *equiregular*, and its *homogeneous dimension* is

$$Q = \sum_{j \ge 1} j \, (n_j - n_{j-1}) \,. \tag{10.6}$$

The hypotheses of Chapter 3 hold locally, for the following reasons.

- *Volume.* The theorem of Nagel, Stein and Wainger on the volume of balls gives $m(B(x, r)) \asymp r^Q$ for x in each fixed compact set and sufficiently small radii, with constants uniform on that set [132, Theorems 1 and 4].
- *Poincaré inequalities.* The local Poincaré inequality of Jerison for the vector fields of the frame [93], [79, Theorem 11.20] yields $\mathrm{PI}_p(r^p)$ for all $1 \le p < \infty$, by Hölder's inequality and the Sobolev–Poincaré estimate in L^p.

- *Cutoffs.* The distance cutoff gives $\mathrm{B}(r^2)$, by Lemma 3.8.

Thus on a compact equiregular manifold

$$\Phi(r) = \Theta(r) = r^Q\,, \qquad \Psi_{P,p}(r) = \Psi_{C,p}(r) = r^p\,. \tag{10.7}$$

For the cutoff condition on a compact subset of a noncompact manifold, one uses the bounds for the volume on a relatively compact neighbourhood of the subset: if a small ball meets the set where a cutoff centred in the subset varies, its centre belongs to this neighbourhood when r is small enough.

Carnot groups. On a Carnot group, the left translations and the homogeneous dilations make (10.7) uniform at every radius; these are the homogeneous structures of the horizontal Sobolev theory of [67, Sections 1 and 4] and [79, Section 11.3]. For example, the Heisenberg group $\mathbb{H}^j$ has the horizontal fields

$$X_i = \partial_{x_i} - \tfrac{1}{2} y_i\,\partial_t\,, \qquad Y_i = \partial_{y_i} + \tfrac{1}{2} x_i\,\partial_t\,, \qquad 1 \le i \le j\,.$$

With these fields orthonormal, its rank is $2j$, its homogeneous dimension is $Q = 2j + 2$, and its reference and gradient measures are both the Haar measure. The rank $2j$ is smaller than the exponent $s = Q = 2j + 2$ of the gauge: the bound $n_* \le s$ of Proposition 5.8 is not an equality in general.

10.4 Sub-Riemannian manifolds: global estimates

On a complete sub-Riemannian manifold with $\nu = m$, the distance cutoffs give $\mathrm{B}(r^2)$ at every radius, by Lemma 3.8. Global doubling and global Poincaré inequalities require geometric conditions. We describe two: a generalized curvature-dimension inequality, and a comparison theorem for the sub-Laplacian of the distance.

Generalized curvature-dimension inequalities.

Proposition 10.8. *Suppose that M is complete and satisfies the generalized curvature-dimension inequality $\mathrm{CD}(\rho_1, \rho_2, \kappa, N)$ of [23, Definition 1.3] with $\rho_1 \ge 0$, $\rho_2 > 0$, $\kappa \ge 0$ and $0 < N < \infty$, under the analytic hypotheses* (H.1)–(H.3) *of [24, Section 1]. Then:*

(i) *VD, $\mathrm{B}(r^2)$ and $\mathrm{PI}_2(r^2)$ hold at every radius;*

(ii) *the heat kernel has two-sided Gaussian bounds with the prefactor $m(B(x, \sqrt{t}))^{-1}$, for all $t > 0$;*

(iii) *for some $\varepsilon \in (0, 1)$, $\mathrm{PI}_p(r^p)$ holds at every radius for every $2 - \varepsilon < p < \infty$.*

Proof. The global doubling property, the quadratic Poincaré inequality with the same ball on both sides, and the Gaussian bounds are due to Baudoin, Bonnefont and Garofalo [24, Theorems 1.5 and 4.1]. The quadratic inequality, applied to powers of u centred at a median, gives $\mathrm{PI}_p(r^p)$ for $p \ge 2$. Theorem 6.1 gives the range below 2, as in the theorem of Keith and Zhong [101]. □

Examples include the complete Sasakian manifolds with nonnegative Tanaka–Webster Ricci curvature, the Carnot groups of step two, and the complete totally geodesic Yang–Mills foliations satisfying the curvature and torsion bounds of Baudoin, Kim and Wang [27, Theorems 4.2 and 5.1]. The Sasakian and Carnot cases are treated in [23, Theorem 2.24 and Proposition 2.21]. For principal bundles with a compact semisimple structure group, see the criterion on the curvature of the connection in [23, Proposition 2.28].

Sub-Laplacian comparison and measure contraction.

Proposition 10.9. *Let M be a complete sub-Riemannian manifold with a smooth positive measure m and the sub-Laplacian $L = \operatorname{div}_m \nabla_H$. Suppose that, for each $x \in M$, the cut locus of x is m-null, the distance $r_x = d(x, \cdot)$ is smooth away from x and its cut locus, and the minimizing radial maps are smooth and injective there. If*

$$Lr_x \le \frac{N-1}{r_x}$$

on that set, for some $N \ge 1$ independent of x, then the measure contraction property $\mathrm{MCP}(0, N)$, VD, $\mathrm{B}(r^2)$ *and* $\mathrm{PI}_p(r^p)$ *hold at every radius, for every $1 \le p < \infty$.*

Proof. The comparison gives the measure contraction property, by the estimate of the radial Jacobian of [28, Section 3.6], and $\mathrm{MCP}(0, N)$ implies doubling. The resulting segment inequality gives $\mathrm{PI}_1(r)$ and, by Jensen's inequality along the same geodesics, $\mathrm{PI}_p(r^p)$ for $p > 1$ [30, Section 3]. The cutoff condition is given by the distance cutoffs. □

Two classes of examples satisfy the hypotheses of Proposition 10.9.

- *Sasakian manifolds.* On a complete Sasakian manifold of horizontal rank $2j$ with its pseudo-Hermitian volume, assume that, for each unit horizontal vector X, the Tanaka–Webster sectional curvature of $\operatorname{span}\{X, JX\}$ and the Ricci trace over its horizontal orthogonal complement are nonnegative. Then $\mathrm{MCP}(0, 2j+3)$ holds, and the Riemannian approximations satisfy $\mathrm{MCP}(0, 2j+4)$ [28, Theorem 3.1 and Corollary 3.15].
- *H-type foliations.* On a complete H-type foliation of horizontal and vertical ranks k and ℓ, completely parallel Bott torsion and nonnegative horizontal Bott sectional curvature give $\mathrm{MCP}(0, k+3\ell)$ for the Riemannian volume, by the same argument on the Jacobian and the comparison of [29, Proposition 2.10 and Theorem 3.20].

Consequences. In the two classes above, the theory at matching scales gives the critical exponent $\alpha_p = 1$ for all finite $p \ge 1$, and the characterization of BV at $p = 1$ (Theorems 4.12 and 4.8). Under Proposition 10.8, $\alpha_p = 1$ for $p > 2 - \varepsilon$.

At nonnegative curvature, Baudoin and Kim obtain the pseudo-Poincaré inequality $\|f - P_t f\|_p \le C_p \sqrt{t}\, \|\nabla_H f\|_p$ for $f \in C_c^\infty(M)$ and $1 \le p < \infty$, and relate a global lower bound $m(B(x,r)) \ge c\, r^D$, $D > 1$, to Sobolev and isoperimetric inequalities [26, Proposition 2.3 and Theorem 2.6]. The same lower bound for the volume gives the global Gagliardo–Nirenberg inequalities of Theorem 4.13 with $\alpha = 1$ at each matching exponent. On a Carnot group one takes $D = Q$ and any finite $p \ge 1$.

Compact manifolds and negative lower bounds. Connected compact equiregular manifolds have global doubling and $\mathrm{PI}_p(r^p)$ for every finite $p \ge 1$, by patching the local estimates over a finite connected cover. The volume of their balls is comparable to $\min\{r^Q, 1\}$, and the Gagliardo–Nirenberg inequalities at bounded radii retain the term of lower order. A positive curvature parameter ρ_1 gives compactness, by the sub-Riemannian Bonnet–Myers theorem [23, Theorem 10.1]. A lower bound $\rho_1 = -K < 0$ gives doubling and $\mathrm{PI}_2(r^2)$ uniformly for $0 < r \le R$, with constants depending on KR^2 [25, Theorem 12], [108, Theorem 1.1].

11 RCD(K, N) spaces

Metric measure spaces with synthetic lower bounds on the Ricci curvature carry a first-order calculus, due to Gigli, which fits the framework of this book: the differential of a Sobolev function is a section of a module with fibres of finite constant dimension, and the reference measure is the gradient measure. In this chapter we verify the axioms, identify the Sobolev spaces $W^{1,p}(X, m)$ of Chapter 2 with the intrinsic ones for every $1 \le p < \infty$, and check the quantitative hypotheses. All the scales are Gaussian.

Setting. We use the convention that an RCD(K, N) space is a metric measure space (X, d, m) which satisfies the curvature-dimension condition CD(K, N) and whose Cheeger energy Ch is quadratic, with

$$2\,\mathrm{Ch}(u) = \int_X |Du|^2 \, dm\,, \tag{11.1}$$

where $|Du|$ is the minimal 2-weak upper gradient. Throughout the chapter the space has more than one point, $1 \le N < \infty$, and m has full support. The space is then proper and geodesic, and it is doubling, uniformly on bounded ranges of radii [148, 9].

11.1 The gradient structure

Proposition 11.1. *Gigli's differential, expressed in a measurable orthonormal frame of the cotangent module, defines a 2-admissible gradient structure with $\nu = m$. Its rank is the almost everywhere constant dimension k of the fibres of the module, and its quadratic energy is* $2\,\mathrm{Ch}$.

Proof. Step 1: the module. The representation of the cotangent module by a measurable field of Hilbert spaces is [74, Theorem 1.4.11], and the dimension of its fibres is almost everywhere constant by the theorem of Brué and Semola [45, Corollary 3.10]. In a measurable orthonormal frame, the differential du of a function $u \in W^{1,2}(X, d, m)$ is a measurable field with values in $\mathbb{R}^k$, and $|du| = |Du|$.

Step 2: Lipschitz functions with compact support. Bounded Lipschitz Sobolev functions are dense in $W^{1,2}(X, d, m)$, by the theorem on density in energy [8, Section 7], and the differential is closed by its construction [74, Section 2.2]. Multiplication by distance cutoffs of growing support gives approximation in the graph norm by functions of $\mathrm{Lip}_c(X)$, since the gradients of the cutoffs are $O(R^{-1})$ and the cutoffs tend to one.

Step 3: the fibres. The differentials of a countable family which is dense in the graph norm generate the module. If their pointwise span were a proper subspace on a set of positive measure, a square-integrable section orthogonal to the span, on a subset of finite measure, would contradict the generation. By Step 2 the family can be chosen in $\mathrm{Lip}_c(X)$; its gradients are bounded with compact support, hence integrable, so $\mathcal{C}_1$ spans the fibres.

Step 4: condition (G). The differential satisfies the Leibniz rule, and its norm is bounded by the Lipschitz constant [74, Section 2.2]. For the weak-$*$ continuity, let (u_j) be as in (G), and multiply the sequence by a fixed compactly supported Lipschitz cutoff equal to one on a given ball. The products converge in $L^2(m)$ and their differentials are bounded in L^2, so the weak closedness of the differential identifies their weak limit. Locality yields the convergence of the original gradients on the ball, tested against square-integrable fields, and their uniform bound extends the convergence to all integrable fields, by truncation.

Step 5: recovery of the metric. Let $U \subset X$ be open, let $u \in \mathrm{Lip}_b(X)$ with $|du| \le L$ almost everywhere on U, and let $w \in U$. Apply the truncation of Step 4 of the proof of Proposition 10.4:

with a relatively compact open set V such that $w \in V \subset \overline{V} \subset U$ and with the compactly supported function $\rho = d(\cdot, X \setminus V)$, the function $v = ((u - u(w)) \wedge L\rho) \vee (-L\rho)$ is a compactly supported Lipschitz function with $|dv| \le L$ almost everywhere on X. By the Sobolev-to-Lipschitz property of RCD spaces [74, (3.1.3)], v is L-Lipschitz, and $v = u - u(w)$ near w. Subdividing a geodesic, as in Step 5 of that proof, gives (1.11) with $\iota \equiv 1$.

Step 6: admissibility. The differential is closed on the domain of the Cheeger energy, and $\mathrm{Lip}_c(X) = \mathcal{C}_2$ is dense in this domain for the graph norm, by Step 2. Hence (A2) holds, $W^{1,2}(X, m) = W^{1,2}(X, d, m)$, and $\mathcal{E}_2 = 2\,\mathrm{Ch}$. Condition (A1) holds because $\mathrm{Lip}_c(X)$ is uniformly dense in $C_c(X)$. □

Step 5 shows that the structure is uniform in the sense of Definition 1.10.

11.2 The exponents $1 < p < \infty$

For $1 < p < \infty$, write $|Du|_p$ for the minimal p-weak upper gradient, and $W^{1,p}(X, d, m)$ for the corresponding intrinsic Sobolev space. On a general metric measure space the minimal weak upper gradient may depend on p. On RCD(K, N) spaces it does not: the theorem of Gigli and Han [72, Theorem 3.4] identifies these norms on the common Sobolev domains, and Gigli and Nobili [73, Theorem 4.1 and Proposition 5.9] identify the differentials themselves and prove the density of Lipschitz functions which is needed below.

Proposition 11.2. *The gradient structure of Proposition 11.1 is p-admissible for every $1 < p < \infty$. The domain of its closed gradient is the intrinsic Sobolev space $W^{1,p}(X, d, m)$, and*

$$|\nabla_p u| = |Du|_p \quad m\text{-almost everywhere}, \qquad \mathcal{E}_p(u) = \int_X |Du|_p^p \, dm\,. \tag{11.2}$$

For $1 < p, q < \infty$ and $u \in W^{1,p}(X, m) \cap W^{1,q}(X, m)$, one has $\nabla_p u = \nabla_q u$ m-almost everywhere. In particular the field of fibres and its rank k do not depend on p.

Proof. By [73, Theorem 4.1 and Proposition 5.9], the intrinsic Sobolev spaces admit a common linear differential d with $|du| = |Du|_p$, whose restriction at $p = 2$ is the differential already used. In the same measurable orthonormal frame, du is a section of H.

The differential is closed. Let $u_j \to u$ in $L^p(m)$ and $du_j \to G$ in $L^p(\mathsf{H}, m)$. The identity of the norms makes (u_j) a Cauchy sequence for the intrinsic norm of $W^{1,p}$, so completeness gives $u \in W^{1,p}(X, d, m)$ and $du = G$.

The core is dense. The proof of [73, Proposition 5.9] gives the density, for the norm, of the Lipschitz functions with bounded support. Since X is proper, these functions belong to $\mathrm{Lip}_c(X) = \mathcal{C}_p$.

Thus the closure of the gradient on $\mathcal{C}_p$ is d, in the chosen frame. The common differential also gives the compatibility on intersections. □

11.3 Poincaré inequalities and scales

Proposition 11.3. *For every $1 \le p < \infty$,*

$$\int_{B(x,r)} |u - u_{B(x,r)}|^p \, dm \le C_p \, r^p \int_{B(x,2r)} |du|^p \, dm \qquad (u \in \mathrm{Lip}_c(X))\,. \tag{11.3}$$

The constants are uniform in the centre on each bounded range of radii, and at all radii when $K \ge 0$.

This is the version in L^p of the proof of the Poincaré inequality by medians and optimal transport of Rajala [137, Theorem 4.1]; the input is the bound for the densities of the interpolating measures of [137, Theorem 1.3]. For $N = 1$, apply that bound with the upper dimension parameter 2.

Proof. Let $B = B(x,r)$, let a be a median of u on B, and put $B^{\pm} = B \cap \{\pm(u-a) \geq 0\}$, so that $m(B^{\pm}) \geq m(B)/2$. By [137, Theorem 1.3] there is a transport plan π between $m(B^+)^{-1} m|_{B^+}$ and $m(B^-)^{-1} m|_{B^-}$ which is carried by geodesics of length at most $2r$, contained in $2B$, and whose marginals at the times $t \in [0,1]$ have densities at most $C/m(B)$. This is a test plan, with bounded speed and bounded compression, so the inequality defining weak upper gradients holds along π-almost every curve γ. The values of u at the endpoints of γ lie on opposite sides of a, so that $|u(\gamma_0) - a|$ and $|u(\gamma_1) - a|$ are at most $|u(\gamma_1) - u(\gamma_0)|$. Consequently, by Hölder's inequality,

$$\begin{aligned} \int_B |u-a|^p \, dm &\leq 2\, m(B) \int |u(\gamma_1) - u(\gamma_0)|^p \, d\pi(\gamma) \\ &\leq 2\,(2r)^p \, m(B) \int \int_0^1 |du|^p(\gamma_t) \, dt \, d\pi(\gamma) \leq C\, r^p \int_{2B} |du|^p \, dm \,. \end{aligned}$$

Finally $\int_B |u - u_B|^p \, dm \leq 2^p \int_B |u-a|^p \, dm$, by Jensen's inequality. At $p = 1$ the argument uses the same inequality along curves, for the quadratic differential of the Lipschitz function u; no identification of the Sobolev space at the exponent one is needed for this step. □

Since $\nu = m$, Lemma 3.8 gives $\mathrm{B}(r^2)$, without assuming $\mathrm{V}(\Phi)$. Thus, for every finite $p \geq 1$, the space satisfies

$$\mathrm{VD}\,, \quad \mathrm{B}(r^2)\,, \quad \mathrm{PI}_p(r^p)\,, \qquad \Psi_{C,p}(r) = \Psi_{P,p}(r) = r^p\,, \tag{11.4}$$

on bounded ranges of radii, and at every radius when $K \geq 0$. This has three consequences.

- Theorem 4.3 gives the admissibility at $p = 1$, which is not covered by Proposition 11.2.
- Approximation in the graph norm extends (11.3) to the Sobolev spaces $W^{1,p}(X,m)$.
- The Korevaar–Schoen characterization of Theorem 4.4 applies with the normalization r^p for $p > 1$, and Theorem 4.8 applies with the normalization r at $p = 1$.

11.4 The exponent one

We compare the space $W^{1,1}(X,m)$ and the total variation of Definition 2.9 with the functions of bounded variation of metric measure spaces. For $u \in BV(X,m)$ in the metric sense, denote by $\|Du\|$ its variation measure, which is a finite Radon measure.

Proposition 11.4. *The total variation of Definition 2.9 is equal to the total mass of the metric variation measure, and the space $BV(X,m)$ of Definition 2.9 is the metric space of functions of bounded variation. Moreover*

$$W^{1,1}(X,m) = \{u \in BV(X,m) : \ \|Du\| \ll m\}\,, \qquad \|Du\| = |\nabla_1 u|\, m \quad (u \in W^{1,1}(X,m))\,. \tag{11.5}$$

Thus $W^{1,1}(X,m)$ is the part of BV with absolutely continuous variation; a general function of BV may also have a singular variation. This is the identification at the exponent one of [72, Remark 3.5], together with [6]; the proof below explains its relation to the definition by closure used in this book.

Proof. Step 1: the two relaxations. For $v \in \mathrm{Lip}_c(X)$, the metric variation satisfies $\|Dv\| = |dv|\,m$, by [72, Remark 3.5, equation (3.3)]: its density is bounded by the local Lipschitz constant, and is therefore locally square integrable. The total variation of Definition 2.9 is the relaxation in L^1 of $v \mapsto \int_X |dv|\,dm$ on $\mathrm{Lip}_c(X)$, and the metric total variation is the relaxation of $v \mapsto \int_X \mathrm{lip}\, v\,dm$ on locally Lipschitz functions, where $\mathrm{lip}\, v$ is the local Lipschitz constant. The identity above and the lower semicontinuity of the metric variation show that the first is at least the second. Conversely, an approximating sequence in the definition of the metric variation can be made compactly supported by distance cutoffs, and $|dv| \le \mathrm{lip}\, v$, so the first is at most the second.

Step 2: a function of $W^{1,1}$ has absolutely continuous variation. Let $u_j \in \mathrm{Lip}_c(X)$ with $u_j \to u$ and $du_j \to G$ in L^1. The lower semicontinuity of the variation on open sets gives $\|Du\| \le |G|\,m$.

Step 3: the converse. Let $u \in BV(X,m)$ with $\|Du\| = g\,m$. By [72, Remark 3.5] there are Lipschitz functions v_j such that $v_j \to u$ and $\mathrm{lip}\, v_j \to g$ in L^1. Distance cutoffs of increasing support make $v_j \in \mathrm{Lip}_c(X)$ without changing these convergences. Since $|dv_j| \le \mathrm{lip}\, v_j$, the fields dv_j are uniformly integrable and tight. In the measurable frame, whose fibres have finite dimension, the Dunford–Pettis theorem gives a subsequence which converges weakly in L^1 to a field G, with $|G| \le g$ almost everywhere. By Mazur's lemma, convex combinations converge in the graph norm, so $u \in W^{1,1}(X,m)$ and $G = \nabla_1 u$. Step 2 gives $g \le |G|$, hence $g = |\nabla_1 u|$. This proves (11.5). □

Proposition 4.7 extends the agreement of the closed gradients to the intersections $W^{1,1}(X,m) \cap W^{1,p}(X,m)$.

11.5 Energies, heat kernels and volume

The p-energies. For $1 < p < \infty$ and $u, v \in W^{1,p}(X,m)$, the energy measure and the first variation of the p-energy are

$$d\Gamma_p(u) = |\nabla_p u|^p\,dm\,,$$
$$\frac{d}{dt}\,\frac{1}{p}\,\mathcal{E}_p(u+tv)\Big|_{t=0} = \int_X |\nabla_p u|^{p-2}\langle \nabla_p u, \nabla_p v\rangle\,dm\,,$$

with the convention $|z|^{p-2}z = 0$ at $z = 0$. These formulas follow from Proposition 2.6 and from the differentiation of the Euclidean norm of the fibres. They give the usual weak p-Laplacian $\mathrm{div}(|\nabla_p u|^{p-2}\nabla_p u)$ whenever this divergence is defined. The differential and its rank do not depend on p; the energy is quadratic only at $p = 2$.

Heat kernels. A compact space has the Gaussian bounds of Proposition 9.9, with the prefactor $m(B(x,\sqrt{t}))^{-1}$; on an unbounded space with $K \ge 0$, Theorem 9.8 gives the bounds at all times. These Gaussian estimates are known directly for RCD spaces, by the work of Jiang, Li and Zhang [94, Theorems 1.1–1.2].

Volume. The uniform volume growth $\mathrm{V}(\Phi)$ is an additional condition.

- For a compact noncollapsed RCD(K,n) space, $\mathrm{V}(r^n)$ holds, by the Bishop–Gromov comparison and the bound for the density of noncollapsed spaces [59, Corollary 2.14]. On a noncompact noncollapsed space the same conclusion holds at small radii if the volumes of the balls of one fixed positive radius are bounded from below uniformly.
- The space $([0,\pi], |\cdot|, \sin^{N-1}(t)\,dt)$, $1 < N < 2$, of rank one, is RCD$(N-1,N)$ [59, after Corollary 1.5]. The volume of its balls has the order r^N at the endpoints and the order r at a fixed interior point, so there is no uniform volume function.

12 Cubical inverse limits of Laakso type

In the examples of the two preceding chapters the rank of the structure and the growth of the volume are tied together by the geometry. The spaces of this chapter separate them: the rank is any integer $n \geq 1$, the volume exponent is any real number $Q > n$, and all the scales remain Gaussian.

Laakso constructed Ahlfors Q-regular spaces which support a $(1,1)$-Poincaré inequality, for every $Q > 1$, and proposed to replace the interval of his construction by an n-cube [120, Theorem 2.6 and Remark 3.2]. Cheeger and Kleiner proved the Poincaré inequality for admissible inverse systems of graphs, and formulated its extension to cubical complexes [50, Theorem 1.1 and Section 11]. The spaces below are inverse limits of this kind. They are made of Euclidean n-cubes, the *sheets*, glued along hyperplanes; the dimension of the sheets is the rank, and the number of sheets at each scale determines the volume exponent.

Theorem 12.1. *For each integer $n \geq 1$ and each real $Q > n$, there is a compact geodesic inverse limit (X, d, m) of cubical complexes, in the sense of Cheeger and Kleiner, with the following properties.*

(i) *The measure m is an Ahlfors Q-regular probability measure. With $\mathcal{S} = X$ and $\nu = m$, the Euclidean gradient on the sheets defines a gradient structure of constant rank n, which is p-admissible for every $1 \leq p < \infty$.*

(ii) *The space supports a $(1,1)$-Poincaré inequality. For the gradient on the sheets, $\mathrm{PI}_p(r^p)$ holds for every $1 \leq p < \infty$, and the cutoff conditions $\mathrm{B}(r^2)$ and $\mathrm{B}^+(r^2)$ hold.*

(iii) *For $u \in \mathrm{Lip}(X)$, $x \in X$, $r > 0$ and $y, z \in B(x, r)$,*

$$|u(y) - u(z)| \leq d(y,z)\, \|\nabla u\|_{L^\infty(B(x,3r),m)}\,. \tag{12.1}$$

In particular the structure is uniform, and

$$d(y,z) = \max\{u(y) - u(z) : u \in \mathrm{Lip}(X),\ \|\nabla u\|_{L^\infty(m)} \leq 1\}\,.$$

Thus $\Phi(r) = \Theta(r) = r^Q$ and $\Psi_{C,p}(r) = \Psi_{P,p}(r) = r^p$ for every $1 \leq p < \infty$, while $n_* = n < Q$. The construction occupies Section 12.1; assertion (i) is proved in Sections 12.1 and 12.2, and assertions (ii) and (iii) in Section 12.3.

12.1 The inverse system and its measure

Construction. Put $I = [0,1]$, $h_j = 2^{-j}$ and

$$N_j = 2^{\lfloor j(Q-n) \rfloor}\,, \qquad b_j = \frac{N_j}{N_{j-1}} \quad (j \geq 1)\,.$$

Thus $N_0 = 1$, and the numbers b_j are integers with $1 \leq b_j \leq 2^{\lceil Q-n \rceil}$. Start with $X_0 = I^n$. To pass from X_{j-1} to X_j, subdivide each cube into 2^n equal cubes, of side h_j, and take b_j copies of the subdivided complex. Write $t = (t_1, \dots, t_n)$ for the projection to I^n, and use the coordinates in the cyclic order $1, \dots, n, 1, \dots$. If i is the coordinate used at the level j, identify the corresponding points of the b_j copies on the hyperplanes

$$t_i = (2k+1)\, h_j\,, \qquad 0 \leq k < 2^{j-1}\,.$$

These are the hyperplanes through the centres of the cubes of the level $j-1$, perpendicular to the ith axis. Equip X_j with the Euclidean length metric d_j. The *bonding map* $X_j \to X_{j-1}$ forgets the index of the copy, and the measure m_j has the density N_j^{-1} with respect to the Euclidean volume on each n-cell. The complex X_j consists of N_j copies of the unit cube, glued along sets of zero volume, so m_j is a probability measure, and the bonding maps preserve these measures. Let $X = \varprojlim X_j$ be the inverse limit, and let $\pi_j : X \to X_j$ be the projections.

A point of X is thus described by a point $t \in I^n$ and a sequence of labels $\omega = (\omega_j)_{j\geq 1}$, $\omega_j \in \{1, \dots, b_j\}$, the label ω_j being irrelevant when t lies on a gluing hyperplane of the level j. A path in X can change its label at the level j only by crossing one of these hyperplanes, which are $2h_j$ apart.

The hypotheses of the cubical inverse limit theorem. We verify the hypotheses of [50, Section 11] and [143, Definition 3.1].

- Each bonding map is open and isometric on the cells, and each copy gives an isometric section.
- *Bounded local geometry.* At a vertex, the reduced denominator of each dyadic coordinate determines at most one level at which a gluing occurs. Hence at most $2^n(\sup_j b_j)^n$ cells of dimension n meet there.
- *Connectivity by galleries.* A gallery is a chain of n-cells in which consecutive cells share a face of dimension $n-1$. Every cube of the level j has a face on a gluing hyperplane of this level, and its b_j copies share this face. Therefore two cells suffice to join any pair of cells in a fibre of the bonding map, and the connectivity by galleries follows by induction.
- *Weights.* Adjacent cells have equal weights. At a lift of a face of dimension $n-1$, the fraction of the weight of an incident cell of the base which meets that lift is 1 on a new gluing face and $1/b_j$ elsewhere, independently of the incident cell.

Since I^n supports a $(1,1)$-Poincaré inequality, the cubical inverse limit theorem gives that (X, d, m), with the distance and the measure described below, is doubling and supports a $(1,1)$-Poincaré inequality [143, Theorem 3.8].

The distance. A path of the level $k-1$ lifts to any fixed copy of the level k. A detour from its endpoint to the nearest gluing hyperplane of the level k and back changes the last label, at a cost of at most $2h_k$. Since the projections are 1-Lipschitz, summing over $k > j$ gives

$$d_j(\pi_j y, \pi_j z) \;\leq\; d(y,z) := \lim_{k\to\infty} d_k(\pi_k y, \pi_k z) \;\leq\; d_j(\pi_j y, \pi_j z) + 2h_j\,. \tag{12.2}$$

The compact geodesic complexes X_j thus converge to X in the Gromov–Hausdorff distance, so X is compact and geodesic.

The measure and the sheets. Give $\Omega = \prod_{j\geq 1}\{1, \dots, b_j\}$ the uniform product probability $\mathbb{P}$. The surjection

$$q : I^n \times \Omega \to X$$

assigns to (t, ω) the point with the coordinate t and the labels ω. It is injective away from the gluing hyperplanes. For fixed ω, the image $q(I^n \times \{\omega\})$ is an isometric copy of the Euclidean cube, called a *sheet*: straight segments lift with their lengths, and the projection to I^n is 1-Lipschitz. The measure on X is

$$m = q_*(dt \otimes \mathbb{P})\,;$$

it satisfies $(\pi_j)_* m = m_j$ for every j, as in the construction of Barlow and Evans [149, Proposition 4.1].

Ahlfors regularity. Let C be a closed n-cell of X_j containing $\pi_j x$, and let y be a lift of its centre. By (12.2),

$$B(y, h_j/3) \subset \pi_j^{-1} C \subset B(x, (\sqrt{n}+3)h_j)\,, \qquad m(\pi_j^{-1}C) = N_j^{-1} h_j^n \asymp h_j^Q\,.$$

The first inclusion holds because the centre is at the distance $h_j/2$ from the boundary of C. Doubling compares the two balls with $B(x, h_j)$, since $d(x,y) \le (\sqrt{n}+2)h_j$. It follows that $m(B(x,h_j)) \asymp h_j^Q$, uniformly in x and j. By monotonicity, $m(B(x,r)) \asymp r^Q$ for every $0 < r \le \operatorname{diam} X$; for $r \ge 1$ one uses $m(X) = 1$ and $\operatorname{diam} X \le \sqrt{n}+2$.

12.2 The gradient on the sheets

For $u \in \operatorname{Lip}(X)$ set

$$(\nabla u)(q(t,\omega)) = \nabla_t \big[u(q(t,\omega))\big]\,. \tag{12.3}$$

On each sheet the restriction of u is Lipschitz. Difference quotients, Rademacher's theorem and Fubini's theorem give a measurable gradient, defined m-almost everywhere, since the gluing hyperplanes are null. This extends the longitudinal derivative used in the ternary Laakso space [47, Definition 2.5 and Theorem 2.9]. For cubical inverse systems, compare the coordinate derivations of [143, Theorem 3.20(Wea)].

Condition (G) and the fibres. The Leibniz rule holds on the sheets, and $|\nabla u| \le \operatorname{Lip}(u)$ almost everywhere, because the sheets are isometric copies of the cube. For a bounded and measurable on Ω and $\varphi \in C_c^\infty((0,1)^n)$, integration by parts on the sheets gives

$$\int (\nabla u)_i \circ q \;\varphi\, a \, dt \, d\mathbb{P} = -\int (u \circ q)\, \partial_i \varphi \, a \, dt \, d\mathbb{P}\,. \tag{12.4}$$

If $u_k \to u$ pointwise, with uniformly bounded Lipschitz constants and supremum norms, dominated convergence applies on the right-hand side. The finite sums of the functions $\varphi\, a$ are dense in $L^1(dt \otimes \mathbb{P})$, so the uniform bound on the gradients gives the weak-$*$ continuity in (G). The n coordinate functions of the projection $\pi_0 : X \to I^n$ have the gradients $e_1, \dots, e_n$. Thus the minimal field is $H_x = \mathbb{R}^n$ almost everywhere, and the rank is n.

Admissibility. Let $1 \le p < \infty$, and suppose that $u_k \to 0$ in $L^p(m)$ and $\nabla u_k \to V$ in $L^p(X, \mathbb{R}^n, m)$. Both convergences also hold in $L^1(m)$, since m is finite. Passing to the limit in (12.4) gives $\int (V_i \circ q)\, \varphi\, a \, dt \, d\mathbb{P} = 0$ for all φ and a. These products separate the integrable functions, so $V = 0$. This proves (A2), including at $p = 1$. Since m is finite and X is compact, $\mathcal{C}_p = \operatorname{Lip}(X)$, which is uniformly dense in $C(X)$; this is (A1). At $p = 2$, the product structure also underlies the proof of closability of [149, Section 3] for projective limits of Dirichlet forms.

For $1 \le s \le p < \infty$, convergence in the graph norm at the exponent p implies convergence in the graph norm at the exponent s. Thus

$$W^{1,p}(X,m) \subset W^{1,s}(X,m) \qquad \text{and} \qquad \nabla_s u = \nabla_p u \quad \text{on } W^{1,p}(X,m)\,.$$

12.3 Poincaré inequalities and recovery of the metric

The Poincaré inequality of [143] is stated for the pointwise Lipschitz constant. To obtain it for the gradient on the sheets, we approximate u by functions which depend on finitely many labels, as in [50, Sections 4–5].

Averaging over the fibres. For $u \in \mathrm{Lip}(X)$, let u_j be the function on X_j obtained by averaging u over the labels after the level j, with the point of the base and the first j labels fixed. Fixing the later labels gives an isometric section $X_j \to X$, so u_j is Lipschitz, with $\mathrm{Lip}(u_j) \le \mathrm{Lip}(u)$, and it is well defined on the gluing faces. By (12.2), $\|u_j \circ \pi_j - u\|_\infty \le 2h_j \,\mathrm{Lip}(u)$. In the interior of each n-cell, differentiating under the average, in the sense of distributions, gives

$$(\nabla u_j) \circ \pi_j = \mathbb{E}_m[\nabla u \mid \sigma(\pi_j)] \qquad m\text{-almost everywhere}\,, \tag{12.5}$$

where the gradient on the left-hand side is the Euclidean gradient on that cell. The increasing σ-fields $\sigma(\pi_j)$ generate the Borel σ-field of X. Hence, by the convergence theorem for martingales, these gradients converge to ∇u in $L^p(m)$ for every finite $p \ge 1$.

The Poincaré inequalities. At a point of differentiability of u_j in an open n-cell, the pointwise Lipschitz constant of $u_j \circ \pi_j$ at any lift is equal to $|\nabla u_j|$ at that point: the projection is 1-Lipschitz, and the section through the lift is isometric. These points have full m_j-measure. Apply the $(1,1)$-Poincaré inequality of X to $u_j \circ \pi_j$, with its pointwise Lipschitz constant; this formulation is equivalent to the inequality for upper gradients on complete doubling spaces [102, Theorem 2]. Passing to the limit by (12.5) gives, for $B = B(x,r)$,

$$\fint_B |u - u_B|\, dm \le C\, r \fint_{\Lambda B} |\nabla u|\, dm\,. \tag{12.6}$$

For $p > 1$, compare the average on B with the averages on $B(z, 2^{1-k}r)$, $k \ge 0$, for $z \in B$. Doubling and (12.6) bound the difference from the first average, and the sum of the successive differences, by $C\, r\, M(\mathbf{1}_{(2\Lambda+1)B}|\nabla u|)(z)$, where M is the centred maximal operator for m. The averages tend to $u(z)$. The boundedness of M in L^p gives $\mathrm{PI}_p(r^p)$, as in [79, proof of Theorem 3.2 and Theorem 14.13].

The cutoff conditions. The distance cutoffs give $\mathrm{B}(r^2)$, by Lemma 3.8, and the bound $|\nabla \psi_x^r|^2\, m \le C r^{-2}\, m$ gives $\mathrm{B}^+(r^2)$.

Recovery of the metric. Fix $y, z \in B(x,r)$ and write $L = \|\nabla u\|_{L^\infty(B(x,3r),m)}$. Let j be so large that $2h_j < r$. By (12.2), every lift of a point of $B_{X_j}(\pi_j x, 2r)$ belongs to $B(x,3r)$. Thus (12.5) bounds $|\nabla u_j|$ by L, almost everywhere on the part of each cell which lies in this ball of X_j. On each cell, the Euclidean estimate for functions of $W^{1,\infty}$ gives the local Lipschitz bound L inside this open set, and up to the faces by continuity. Join $\pi_j y$ to $\pi_j z$ by piecewise linear paths whose lengths tend to their distance in X_j. The paths whose length is sufficiently close to the infimum lie in $B_{X_j}(\pi_j x, 2r)$, since each point of a path is within half its length of one of the endpoints. Adding the estimates on the cells and taking the infimum yields

$$|u_j(\pi_j y) - u_j(\pi_j z)| \le L\, d_j(\pi_j y, \pi_j z) \le L\, d(y,z)\,.$$

The uniform convergence of $u_j \circ \pi_j$ to u proves (12.1). Taking $r > \operatorname{diam} X$ gives the global inequality, and the function $u = d(\cdot, z)$ attains the maximum in Theorem 12.1(iii). This completes the proof of Theorem 12.1.

Consequences. Since the scales match, with $\Psi_{C,p}(r) = \Psi_{P,p}(r) = r^p$, Theorem 4.12 gives the critical exponent $\alpha_p = 1$, and it is attained. On the common core $\mathrm{Lip}(X)$ the energies are

$$\mathcal{E}_p(u) = \int_\Omega \int_{I^n} |\nabla_t (u \circ q)(t,\omega)|^p \, dt \, d\mathbb{P}(\omega) \,, \qquad 1 \le p < \infty \,.$$

The Sobolev spaces of these energies are nested. This is in contrast with the self-similar energies on the Laakso diamond studied by Anttila, Eriksson-Bique and Shimizu, whose domains at distinct exponents in $(1,\infty)$ intersect only in the constants [13, Theorem 1.8 and Section 1.4]. Since $n < Q$, Theorem 12.1 also shows that the bound $n_* \le s$ of Proposition 5.8 can be strict for every rank, with $s = Q$ arbitrarily close to n or arbitrarily large.

13 The harmonic Sierpiński gasket

The definition of a gradient structure does not require a uniform volume function, and the hypotheses VD, $\mathrm{B}(\Psi_{C,2})$ and $\mathrm{PI}_p(\Psi_{P,p})$ are stated with the volumes of the balls themselves. The Sierpiński gasket with its harmonic metric and the Kusuoka measure is an example where this generality is needed: it satisfies the three hypotheses at Gaussian scales, but the volume of a ball of radius r is not comparable to a function of r alone. It is also an example of a measurable Riemannian structure, in the sense of Kigami [104], which has rank one and is not a tree.

13.1 The harmonic gasket

The resistance form. Let $K \subset \mathbb{R}^2$ be the Sierpiński gasket, the invariant set of the three contractions $F_i(x) = (x + p_i)/2$, where $V_0 = \{p_1, p_2, p_3\}$ is the set of vertices of an equilateral triangle. For a word $w = w_1 \cdots w_k$, put $|w| = k$, $F_w = F_{w_1} \circ \cdots \circ F_{w_k}$, $K_w = F_w(K)$ and $V_w = F_w(V_0)$; the set K_w is a *cell* of the level k, and V_w is the set of its three boundary vertices. Let $(\mathcal{E}, \mathcal{F})$ be the standard resistance form on K. It is self-similar with the factor $5/3$,

$$\mathcal{E}(u, v) = \frac{5}{3} \sum_{i=1}^{3} \mathcal{E}(u \circ F_i, v \circ F_i)\,,$$

and a function h is *harmonic* if it minimizes the energy among the functions with the same values on V_0. The harmonic functions form a space of dimension three, and for a harmonic h and any $v \in \mathcal{F}$,

$$\mathcal{E}(v, h) = c_0 \sum_{q \in V_0} v(q) \sum_{q' \in V_0} \big(h(q) - h(q')\big)\,, \tag{13.1}$$

where $c_0 > 0$ is fixed by the normalization of the form; see [104, Sections 2 and 4].

The harmonic metric and the Kusuoka measure. Let h_1, h_2 be a basis of the harmonic functions modulo constants which is orthonormal for the energy, and put $H = (h_1, h_2) : K \to \mathbb{R}^2$. The map H is a homeomorphism onto its image, the *harmonic gasket*. Equip K with the *harmonic metric* d, the shortest-path metric of $H(K)$; it is geodesic. Let $\Gamma(u, v)$ denote the energy measures, normalized by $\Gamma(u)(K) = \mathcal{E}(u, u)$. The *Kusuoka measure* and its matrix of densities are

$$\nu = \Gamma(h_1) + \Gamma(h_2)\,, \qquad Z_{ij} = \frac{d\Gamma(h_i, h_j)}{d\nu}\,.$$

By a theorem of Kusuoka [118], the matrix Z is an orthogonal projection of rank one, ν-almost everywhere. There is a measurable differential $\widetilde{\nabla}$, defined on $\mathcal{F}$ with values in the field of lines $\operatorname{Im} Z$, such that

$$\mathcal{E}(u, v) = \int_K \langle \widetilde{\nabla} u, \widetilde{\nabla} v \rangle \, d\nu\,, \qquad \widetilde{\nabla}(F \circ H) = Z\big((DF) \circ H\big) \quad (F \in C^1(\mathbb{R}^2))\,, \tag{13.2}$$

and the energy measures have the same representation on Borel subsets: $d\Gamma(u, v) = \langle \widetilde{\nabla} u, \widetilde{\nabla} v \rangle \, d\nu$. Moreover d is the intrinsic distance of the Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(K, \nu)$. For this differential and for the identification of d, see [118], [104, Sections 2, 4 and 5] and [97, Theorem 4.2]. We take $m = \nu$.

Thus the harmonic gasket is, in a measurable sense, a Riemannian manifold of dimension one: at ν-almost every point there is one tangent direction $\operatorname{Im} Z \subset \mathbb{R}^2$, and the energy is the integral of the square of the derivative in this direction.

13.2 The gradient structure

Proposition 13.1. *Choose a measurable unit vector e in $\operatorname{Im} Z$ and set $\nabla u = \langle e, \widetilde{\nabla} u\rangle$ for $u \in \operatorname{Lip}(K,d)$. This defines a gradient structure (K,ν,∇) of rank one on (K,d,ν), which is p-admissible for every $1 \le p < \infty$. Its Dirichlet form is $(\mathcal{E},\mathcal{F})$, and $\ker \nabla_p$ consists of the constants for every $1 \le p < \infty$. For $u \in \operatorname{Lip}(K,d)$,*

$$|\nabla u|(x) = \limsup_{y\to x,\ y\neq x} \frac{|u(y)-u(x)|}{d(x,y)} \qquad \text{for } \nu\text{-almost every } x\,. \tag{13.3}$$

The proof of the closability for $p \neq 2$ and the identification of the kernel rest on a Gauss–Green formula on cells.

Lemma 13.2. *Let $h = a\cdot H$, $a \in \mathbb{R}^2$, and let K_w be a cell. For every $v \in \operatorname{Lip}(K,d)$,*

$$\int_{K_w} \nabla v\, \nabla h\, d\nu = \sum_{q\in V_w} v(q)\, \partial_n^w h(q)\,, \qquad \partial_n^w h(q) = c_0\Big(\frac{5}{3}\Big)^{|w|} \sum_{q'\in V_w} \big(h(q)-h(q')\big)\,. \tag{13.4}$$

Proof. The integral is the energy $\Gamma(v,h)(K_w)$ of the cell, which by self-similarity is equal to $(5/3)^{|w|}\,\mathcal{E}(v\circ F_w, h\circ F_w)$. Since $h\circ F_w$ is harmonic, (13.1) applies. □

Proof of Proposition 13.1. Step 1: the Lipschitz bound. The Rademacher theorem for the intrinsic distance and the identity (13.3) are [117, Theorems 2.1 and 4.1]: a function $u \in \operatorname{Lip}(K,d)$ belongs to $\mathcal{F}$, and $|\nabla u| = |\widetilde{\nabla} u|$ is ν-almost everywhere equal to the pointwise Lipschitz constant of u. In particular $|\nabla u| \le \operatorname{Lip}(u)$.

Step 2: the Leibniz rule and the weak-$*$ continuity. The C^1 functions of H are dense in $\mathcal{F}$, for the energy and for the uniform norm [104, Theorem 4.8]. For these functions the Leibniz rule follows from (13.2), and it extends by this density and the closedness of $\widetilde{\nabla}$. Let (u_k) be as in (G). A pointwise convergent sequence with bounded Lipschitz constants converges uniformly on the compact space K, and it is bounded in $\mathcal{F}$ by Step 1. The closedness of the form identifies the weak limit in $L^2(\nu)$ of the differentials, and their uniform bound gives the weak-$*$ convergence in $L^\infty(\nu)$.

Step 3: the fibres, the core and the Dirichlet form. Since $Ze = e$, (13.2) gives $\nabla h_i = \langle e, Ze_i\rangle = e_i$, the ith component of e, so that $|\nabla h_1|^2 + |\nabla h_2|^2 = 1$. The functions h_i are 1-Lipschitz for d. Hence the gradients of h_1, h_2 span the fibres, which are equal to $\mathbb{R}$, and the rank is one. Since K is compact and ν is finite, $\mathcal{C}_p = \operatorname{Lip}(K,d)$ for every finite p; it is uniformly dense in $C(K)$, which is (A1). The core contains the C^1 functions of H, which are dense in $\mathcal{F}$, and $\int_K |\nabla u|^2\, d\nu = \mathcal{E}(u,u)$ on the core. Therefore the closure of ∇ in $L^2(\nu)$ has the domain $\mathcal{F}$, and the Dirichlet form of the structure is $(\mathcal{E},\mathcal{F})$.

Step 4: closability. Let $1 \le p < \infty$, and suppose that $u_j \to 0$ and $\nabla u_j \to G$ in $L^p(\nu)$. Both convergences hold in $L^1(\nu)$. Apply (13.4) with $K_w = K$ to $v = u_j\phi$, where ϕ is Lipschitz and vanishes near V_0. The right-hand side vanishes, so

$$\int_K \phi\, \nabla u_j\, \nabla h_i\, d\nu = -\int_K u_j\, \nabla\phi\, \nabla h_i\, d\nu \longrightarrow 0\,.$$

Thus $\int_K \phi\, G\, \nabla h_i\, d\nu = 0$ for all such ϕ, and $G\,\nabla h_i = 0$ almost everywhere outside V_0. The measure ν has no atoms [104, Proposition 2.10], so Step 3 gives $G = 0$. This proves (A2).

Step 5: convergence at the vertices. It remains to identify the kernel. Let $u_j \in \operatorname{Lip}(K,d)$ with $u_j \to u$ and $\nabla u_j \to 0$ in $L^p(\nu)$, hence in $L^1(\nu)$. Fix a cell K_w and $q \in V_w$. The restrictions of $1, h_1, h_2$ to K_w span the harmonic functions on K_w, since the three points of $H(V_w)$ are affinely

independent [104, Proposition 4.4]. Choose $h = a \cdot H$ with $\partial_n^w h(q) \neq 0$, and a Lipschitz function ϕ whose restriction to V_w is $\mathbf{1}_{\{q\}}$. Applying (13.4) to $u_j\phi$ shows that $u_j(q)$ converges, and that

$$\partial_n^w h(q) \lim_{j\to\infty} u_j(q) = \int_{K_w} u \, \nabla\phi \, \nabla h \, d\nu \,. \tag{13.5}$$

Taking $v = u_j$ in (13.4), for every h, shows that the limits at the three vertices of each cell are equal, because the kernel of the graph energy on the three vertices consists of the constants. The graph formed by the vertices of each level is connected and contains V_0. It follows that the limit is the same number c at every vertex of every level.

Step 6: from the vertices to u. Let A_w be the linear part of the affine map induced by F_w in the harmonic coordinates, so that $H \circ F_w = A_w H + b_w$. The self-similarity of the energy gives

$$\nu(K_w) = \Big(\frac{5}{3}\Big)^{|w|} \|A_w\|_{\mathrm{HS}}^2 \,;$$

see [104, Propositions 2.10 and 4.4]. Choose $h = a\cdot H$ with $|a| = 1$ such that $R = \max_{V_w} h - \min_{V_w} h$ satisfies $R \asymp \|A_w\|_{\mathrm{HS}}$; this is possible because the differences of the three points of $H(V_0)$ span $\mathbb{R}^2$. One of the vertices at which h is extreme on V_w, say q, has the property that $h(q)$ differs from the values of h at the two other vertices by at least $R/2$. Take $\phi = (1 - 2|h - h(q)|/R)_+$. Then

$$\phi|_{V_w} = \mathbf{1}_{\{q\}} \,, \qquad |\nabla\phi| \le \frac{2}{R} \,, \qquad |\partial_n^w h(q)| \ge c_0 \Big(\frac{5}{3}\Big)^{|w|} R \,.$$

Subtract from (13.5) the same identity for the constant sequence $u_j = b$, $b \in \mathbb{R}$. This gives

$$|c - b| \le \frac{2}{c_0\,(5/3)^{|w|} R^2} \int_{K_w} |u - b| \, d\nu \le C \fint_{K_w} |u - b| \, d\nu \qquad (b \in \mathbb{R}) \,.$$

The partitions of K into the cells of a given level, outside the null set of the vertices, are nested, and they generate the Borel σ-field because the diameters of the cells tend uniformly to zero. The convergence theorem for martingales, applied to $|u - b|$ for rational b, gives $|c - b| \le C|u(x) - b|$ for every rational b, at almost every x. Letting $b \to u(x)$ proves that $u = c$ almost everywhere. □

Recovery of the metric. The structure is uniform: (1.11) holds with any constant ι. Indeed, suppose that $u \in \mathrm{Lip}(K, d)$ satisfies $|\nabla u| \le L$ almost everywhere on an open set U. Choose $a \in U$ and an open neighbourhood $V \neq K$ of a with $\overline{V} \subset U$, and put $\rho = d(\cdot, K \setminus V)$. The function $v = ((u - u(a)) \wedge L\rho) \vee (-L\rho)$ agrees with $u - u(a)$ near a, and its gradient is bounded by L on K, by locality. Since d is the intrinsic distance and $\Gamma(v) = |\nabla v|^2\nu \le L^2\nu$, the function v is L-Lipschitz on K. Apply this local estimate along a geodesic from y to z, which lies in $B(x, 3r)$ when $y, z \in B(x, r)$, and let L decrease to the essential supremum of $|\nabla u|$ on $B(x, 3r)$.

13.3 Scales and critical exponents

The hypotheses. The quantitative properties of the harmonic gasket are the following.

- *Doubling.* The Kusuoka measure is doubling for d, and the heat kernel satisfies Gaussian bounds with the volume $\nu(B(x, \sqrt{t}))$ [104, Theorems 6.2–6.3].
- *Poincaré inequalities.* The quadratic Poincaré inequality holds at the scale r^2 [117, Proposition 4.1]. By (13.3), this is a Poincaré inequality for the pointwise Lipschitz constant, so the self-improvement theorem of Keith and Zhong and the Sobolev–Poincaré inequality give $\mathrm{PI}_p(r^p)$ for all $p > 2 - \varepsilon$, for some $\varepsilon > 0$ [101, 79]. Whether $\mathrm{PI}_p(r^p)$ holds for all $p \ge 1$ is not known.

- *Cutoffs.* For all $p \geq 1$, the distance cutoffs of Lemma 3.8 satisfy

$$\int_K |\nabla \psi_x^r|^p \, d\nu \leq \Big(\frac{2}{r}\Big)^p \nu(B(x,2r)) \leq C_p \, \frac{\nu(B(x,r))}{r^p} \,. \tag{13.6}$$

  This gives $\mathrm{B}(r^2)$, and (3.19) with $\Psi_{C,p}(r) = r^p$.

Thus the scales match for $p > 2 - \varepsilon$, and the results of Part I at matching scales apply in this range. At $p = 2$, Proposition 9.9 gives the Gaussian bounds for the heat kernel at small times, in agreement with [104].

The critical exponents. The Korevaar–Schoen critical exponents of Definition 4.9 can be computed for every $p \geq 1$, although the Poincaré inequality is only known for $p > 2 - \varepsilon$. Let $\alpha > 1$ and $f \in KS^{\alpha,p}(K)$. Lemma 3.5 gives Lipschitz approximations $f_\delta \to f$ in $L^p(\nu)$ with

$$\|\nabla f_\delta\|_{L^p(\nu)}^p \leq C \, \delta^{p(\alpha-1)} \, [f]_{KS^{\alpha,p}}^p \longrightarrow 0 \,.$$

Hence $f \in W^{1,p}(K,\nu)$ and $\nabla_p f = 0$, and Proposition 13.1 shows that f is constant. Since the Lipschitz functions are dense in $L^p(\nu)$ and belong to $KS^{1,p}(K)$, both thresholds are attained, and

$$\alpha_p^{\mathrm{dens}} = \alpha_p = 1 \qquad (1 \leq p < \infty) \,.$$

The Poincaré inequality is replaced in this argument by the triviality of the kernel of the closed gradient.

A Morrey estimate. A quantitative estimate is available in the range of the Poincaré inequalities, after decreasing ε. By [97, Theorem 3.19],

$$\frac{\nu(B(x,R))}{\nu(B(x,r))} \leq C \Big(\frac{R}{r}\Big)^q \quad (0 < r \leq R) \,, \qquad q = \frac{\log 15}{\log 5} < 2 \,.$$

Choose ε with $2 - \varepsilon > q$, and let $p > 2 - \varepsilon$. In the proof of the Morrey estimate by a telescoping sum of averages [79, Theorem 5.1(3)], the differences of the averages of u on $B(y, 2^{-j}r)$ and on $B(y, 2^{-j-1}r)$ are bounded by

$$C \, \frac{r}{\nu(B(y,r))^{1/p}} \, 2^{-j(1-q/p)} \Big(\int_{B(y,\Lambda r)} |\nabla u|^p \, d\nu \Big)^{1/p} .$$

Summing over j, and comparing the averages centred at y and at z on a common larger ball, gives, with $r = d(y,z) > 0$,

$$|u(y) - u(z)|^p \leq \frac{C \, r^p}{\nu(B(y,r))} \int_{B(y,\Lambda' r)} |\nabla u|^p \, d\nu \qquad (u \in \mathrm{Lip}(K,d)) \,.$$

No uniform volume function. At a boundary vertex $x \in V_0$, $\nu(B(x,r)) \asymp r$ [97, Theorem 5.3]. At ν-almost every point, the local dimension of ν is strictly greater than one, by the spectral asymptotics and the Gaussian estimates with the volume [97, Corollary 4.3 and Theorem 6.1]. Hence the volume of the balls is not comparable to a function of the radius which is uniform in the centre, and $\mathrm{V}(\Phi)$ fails. Finally, the structure has rank one although the space is not a tree: every nonempty open set contains the boundary of a cell, which is a closed curve.

14 Liouville Brownian motion

In the examples of the preceding chapters the gradient measure ν is the reference measure m. In the present example the two measures are mutually singular: the space is the Euclidean plane, the gradient is the Euclidean gradient, ν is the Lebesgue measure, and m is the Liouville measure M_γ, a random multifractal measure. The Dirichlet form of this structure is that of the Liouville Brownian motion of Garban, Rhodes and Vargas [71]. The example shows the following.

- *Admissibility depends on the exponent.* The structure is p-admissible exactly for $p > \gamma^2/2$, and for $p \le \gamma^2/2$ the relaxed energy vanishes identically (Proposition 14.2). In particular, for $\gamma \ge \sqrt{2}$ every integrable function has zero total variation.
- *The scales differ.* On compact sets, the Poincaré inequality holds at the scales r^a with $a < (2-\gamma)^2/2$, while the cutoff condition and the cutoff Sobolev inequality hold at the scales r^b with $b > (2+\gamma)^2/2$ (Propositions 14.6 and 14.7).
- *The volume is not uniform.* There is no volume function as in $\mathrm{V}(\Phi)$, and global doubling fails (Propositions 14.1 and 14.8).

All the statements of this chapter hold almost surely, and the constants may depend on the realization of the measure.

14.1 The Liouville measure and its Dirichlet form

The measure. Let H be the massive planar Gaussian free field of mass one, with covariance

$$G(x,y) = \frac{1}{2}\int_0^\infty e^{-s/2-|x-y|^2/(2s)}\,\frac{ds}{s} = \log\frac{1}{|x-y|} + O(1) \qquad (|x-y| \downarrow 0)\,. \tag{14.1}$$

For $0 < \gamma < 2$, the subcritical Gaussian multiplicative chaos M_γ is the limit in probability, for the vague topology, of

$$M_{\gamma,\varepsilon}(dx) = \exp\Bigl(\gamma H_\varepsilon(x) - \tfrac{1}{2}\gamma^2\,\mathbb{E}[H_\varepsilon(x)^2]\Bigr)\,dx\,, \tag{14.2}$$

where H_ε is a smooth convolution of H. The limit is independent of the mollifier and is almost surely a nonatomic Radon measure with full support, singular with respect to the Lebesgue measure [95, 34, 71].

The process and its form. Let B be the planar Brownian motion with generator Δ, and let A be its positive continuous additive functional with Revuz measure M_γ. The Liouville Brownian motion is the time change

$$Z_t = B_{\tau_t}\,, \qquad \tau_t = \inf\{s \ge 0 : A_s > t\}\,. \tag{14.3}$$

The clock is strictly increasing and the process has infinite lifetime [71, Theorem 2.7]. Its Dirichlet form on $L^2(M_\gamma)$ is

$$\mathcal{F} = \bigl\{\widetilde{u} \in L^2(M_\gamma) :\ u \in H^1_{\mathrm{loc}}(\mathbb{R}^2, dx)\,,\ \nabla u \in L^2(dx)\bigr\}\,, \tag{14.4}$$

$$\mathcal{E}(\widetilde{u},\widetilde{v}) = \int_{\mathbb{R}^2} \nabla u \cdot \nabla v\,dx\,, \tag{14.5}$$

where $\widetilde{u}$ is the quasi-continuous representative of u for the Euclidean capacity [70, Theorem 1.7], [11, Section 2.3]. With this normalization the process runs at twice the speed of the process of [11]. Thus the energy is the Euclidean one, and the measure M_γ enters only through the L^2 space on which the form is considered.

Proposition 14.1. *Almost surely, the Euclidean gradient with $\mathcal{S} = \mathbb{R}^2$ and $\nu = dx$ defines a uniform 2-admissible gradient structure of rank two on $(\mathbb{R}^2, |\cdot|, M_\gamma)$. Its Dirichlet form is (14.4)–(14.5), and $C_c^\infty(\mathbb{R}^2)$ is a core. The distance cutoffs satisfy $\int_{\mathbb{R}^2} |\nabla \psi_x^r|\, dx \le C\, r$. There is no uniform volume function as in* $\mathrm{V}(\Phi)$.

Proof. Condition (G) and the property (1.11) involve only the gradient measure $\nu = dx$; they hold as in Proposition 10.2, with $\iota \equiv 1$. Compactly supported Lipschitz functions belong to $\mathcal{F}$. The closedness of $\mathcal{E}$ on $L^2(M_\gamma)$ and the fact that the smooth compactly supported functions form a core [71, Theorem 2.18] give the 2-admissibility and the identification of the form. The bound for the cutoff follows from its support and its Lipschitz constant.

Suppose that $M_\gamma(B(x,r)) \asymp \Phi(r)$ uniformly for $x \in B(0,1)$ and small r. By Fubini's theorem,

$$\pi r^2\, M_\gamma(B(0,1/2)) \le \int_{B(0,1)} M_\gamma(B(x,r))\, dx \le \pi r^2\, M_\gamma(B(0,2)) \qquad (r < 1/2)\,.$$

Thus $\Phi(r) \asymp r^2$. The resulting lower bound $M_\gamma(B(x,r)) \ge c\, r^2$ contradicts the Lebesgue differentiation theorem for the singular measure M_γ. □

Volume and capacity. The local dimension of M_γ is $2 - \gamma^2/2$ at M_γ-almost every point, and the measure is carried by a Borel set of this Hausdorff dimension [138, Theorems 4.1–4.2]. Uniformly on each compact set of centres, for every $\varepsilon > 0$,

$$c\, r^{(2+\gamma)^2/2+\varepsilon} \le M_\gamma(B(x,r)) \le C\, r^{(2-\gamma)^2/2-\varepsilon} \qquad (0 < r \le 1)\,, \tag{14.6}$$

with random constants [11, Lemma 3.1]. The capacity of an annulus, on the other hand, is the Euclidean one:

$$\mathrm{Cap}(B(x,r), B(x,2r)) = \frac{2\pi}{\log 2}\,. \tag{14.7}$$

The radial logarithmic cutoff gives the upper bound. For the lower bound, let u be a cutoff function for the pair of balls; the Cauchy–Schwarz inequality along almost every radial segment gives $1 \le \log 2 \int_r^{2r} |\partial_\rho u(x+\rho\omega)|^2 \rho\, d\rho$, and one integrates over $\omega \in S^1$. Thus the three exponents $(2-\gamma)^2/2 < 2 - \gamma^2/2 < (2+\gamma)^2/2$ describe the smallest, the typical and the largest rate of decay of the volume of small balls, and the ratio $M_\gamma(B(x,r))/\mathrm{Cap}(B(x,r),B(x,2r))$ is not comparable to a function of r.

14.2 Admissibility and relaxation

Proposition 14.2. *Fix $0 < \gamma < 2$. Almost surely, the Euclidean gradient on $(\mathbb{R}^2, |\cdot|, M_\gamma)$ has the following properties.*

(i) *For $1 \le p < \infty$, the structure is p-admissible if and only if $p > \gamma^2/2$.*

(ii) *For $p > 2$,*

$$W^{1,p}(\mathbb{R}^2, M_\gamma) = \{u \in C(\mathbb{R}^2) \cap W^{1,p}_{\mathrm{loc}}(\mathbb{R}^2, dx) :\ u \in L^p(M_\gamma)\,,\ \nabla u \in L^p(dx)\}\,. \tag{14.8}$$

(iii) *For $1 \le p \le \gamma^2/2$,*

$$\overline{\mathcal{E}}_p(f) = 0 \qquad (f \in L^p(M_\gamma))\,. \tag{14.9}$$

In particular, for $\gamma \ge \sqrt{2}$, $BV(\mathbb{R}^2, M_\gamma) = L^1(M_\gamma)$ and every function has zero total variation.

The threshold has the following meaning. A set of zero $(2-p)$-dimensional Hausdorff measure has zero p-capacity for the Euclidean Sobolev space, and the measure M_γ has dimension $2-\gamma^2/2$. If $p > \gamma^2/2$, pieces of the measure are spread out enough for the functions of the Euclidean Sobolev space $W^{1,p}$ to have a trace on them, and this makes the gradient closable. If $p \le \gamma^2/2$, the measure is carried by a set of zero p-capacity; a function can then be modified on a neighbourhood of this set, at an arbitrarily small cost in energy, and the relaxed energy vanishes. The proof occupies the rest of this section and the next one. It uses three lemmas: the existence of pieces of M_γ with a uniform growth at many scales, a trace inequality for such pieces, and the existence of a carrier of M_γ with zero Hausdorff measure in the critical dimension.

Decomposition into scales. We use the representation of (14.1) by independent scales, as in [71, Section 2.1]. Each interval of integration in (14.1) gives a Gaussian field, and disjoint intervals give independent fields, because the kernel $e^{-|x-y|^2/(2s)}$ is positive definite. The chaos of the truncations to the scales $s \ge \varepsilon$ is a uniformly integrable martingale on bounded sets, by [138, Theorem 2.5], and its limit agrees with the limit of (14.2), by the uniqueness of the subcritical chaos [153, Corollary 18 and Theorems 25–26].

Pieces with uniform growth. Let $D = B(0,1)$ and $d > 0$. For a probability measure μ on D and an integer $N \ge 1$, put

$$K_N(\mu) = \{z \in B(0,1/2) : \ \mu(B(z,2^{-j})) \le N\, 2^{-jd} \text{ for all integers } j \ge 0\}\,,$$
$$J_d(\mu) = \inf_{N \ge 1} \frac{N}{\mu(K_N(\mu))}\,,$$

with $1/0 = \infty$. When $\mu(K_N(\mu)) > 0$, the normalized restriction ρ of μ to $K_N(\mu)$ satisfies

$$\rho(B(z,t)) \le C_d \frac{N}{\mu(K_N(\mu))}\, t^d \qquad (0 < t \le 1\,,\ z \in \mathbb{R}^2)\,. \tag{14.10}$$

Indeed, if the ball meets $K_N(\mu)$, it is contained in a ball of radius at most $4t$, with a dyadic radius, centred at a point of the intersection. Thus $J_d(\mu) < \infty$ means that a piece of μ of positive mass satisfies the growth condition (14.10), with a constant comparable to $J_d(\mu)$.

For $x \in \mathbb{R}^2$ and $0 < r \le 1$, let $\mu_{x,r}$ be the probability measure on D obtained from the restriction of M_γ to $B(x,r)$ by the change of variables $y = x + rz$ and normalization.

Lemma 14.3. *Let $0 < d < 2 - \gamma^2/2$.*

(i) *The random variables $J_d(\mu_{x,r})$, $x \in \mathbb{R}^2$, $0 < r \le 1$, are bounded in probability, uniformly in x and r.*

(ii) *Almost surely, for Lebesgue-almost every $x \in \mathbb{R}^2$,*

$$\liminf_{k \to \infty} J_d(\mu_{x,2^{-k}}) < \infty\,.$$

The exceptional event can be chosen independent of the rational numbers $d \in (0, 2-\gamma^2/2)$.

Proof. (i) *Step 1: a reference chaos.* After the change of variables $y = x + rz$, the component of the field from the scales $s < r^2$ has the covariance

$$\frac{1}{2}\int_0^1 e^{-r^2 t/2}\, e^{-|z-w|^2/(2t)}\, \frac{dt}{t}\,.$$

Add an independent Gaussian field Y_r whose covariance is obtained by replacing $e^{-r^2t/2}$ in this integral by $1-e^{-r^2t/2}$. The law of the sum does not depend on x and r. Let μ_0 be the chaos of the sum, restricted to D and normalized. The covariance of the sum truncated at the scale ε differs from $\log^+(1/(|z-w|+\varepsilon))$ by a uniformly bounded function, and its increments in space have a variance at most $C|z-w|^2/\varepsilon^2$; thus the truncations satisfy the hypotheses of [138, Section 2.1 and Theorems 4.1–4.2], and μ_0 has full support and exact dimension $2-\gamma^2/2$. Since $d<2-\gamma^2/2$, at μ_0-almost every point z the quantities $2^{jd}\mu_0(B(z,2^{-j}))$, $j\geq 0$, are bounded. Hence $\mu_0(K_N(\mu_0))>0$ for some N, and $J_d(\mu_0)<\infty$ almost surely.

Step 2: comparison. Let $C_{x,r}$ be the component of the field from the scales $s\geq r^2$, after the change of variables. For $z,w\in\overline{D}$, the two continuous fields satisfy

$$\mathbb{E}|C_{x,r}(z)-C_{x,r}(w)|^2\leq C\,|z-w|^2\,,\qquad \mathbb{E}|Y_r(z)-Y_r(w)|^2\leq C\,|z-w|\,.$$

The first bound follows by integrating $1-e^{-|z-w|^2/(2t)}\leq |z-w|^2/(2t)$ over $t\geq 1$. For the second, use $1-e^{-r^2t/2}\leq t/2$ and $1-e^{-|z-w|^2/(2t)}\leq C|z-w|/\sqrt{t}$. The bounds for the Gaussian moments and the Kolmogorov continuity theorem give bounds in probability for the oscillations of the two fields on $\overline{D}$, uniformly in x and r. Adding an independent continuous Gaussian field multiplies the chaos by the exponential of the field, corrected by its variance, as one sees on the regularizations; the variances of $C_{x,r}$ and Y_r are constant in space. Thus, in this coupling,

$$A^{-1}\mu_0\leq\mu_{x,r}\leq A\,\mu_0\,,\qquad A=\exp\bigl(\gamma\,(\mathrm{osc}_{\overline{D}}\,C_{x,r}+\mathrm{osc}_{\overline{D}}\,Y_r)\bigr)\,.$$

For each N, $K_N(\mu_0)\subset K_{\lceil AN\rceil}(\mu_{x,r})$, so that $J_d(\mu_{x,r})\leq 2A^2J_d(\mu_0)$. This proves (i).

(ii) For any fixed x and any $\lambda>0$, the probability that $\liminf_k J_d(\mu_{x,2^{-k}})$ exceeds λ is at most $\liminf_k\mathbb{P}(J_d(\mu_{x,2^{-k}})>\lambda)$, which tends to zero as $\lambda\to\infty$, by (i). Hence the lower limit is finite almost surely. The sets K_N and the quantities J_d are measurable in the pair formed by x and the realization, since the conditions defining them involve countably many radii. Fubini's theorem gives the assertion for Lebesgue-almost every x, on a single event for all rational d. □

A trace inequality.

Lemma 14.4. *Let $1\leq p<2$ and $2-p<d\leq 2$. Let ρ be a probability measure carried by D such that $\rho(B(z,t))\leq A\,t^d$ for all $z\in\mathbb{R}^2$ and $0<t\leq 1$. Then, for every Lipschitz function u on D,*

$$\int_D|u-u_{D,dx}|\,d\rho\leq C_{d,p}\,(1+A)\,\|\nabla u\|_{L^p(D,dx)}\,,\tag{14.11}$$

where $u_{D,dx}$ is the average of u on D for the Lebesgue measure.

Proof. The pointwise potential estimate on the ball bounds $|u(z)-u_{D,dx}|$ by $C\int_D|z-w|^{-1}|\nabla u(w)|\,dw$ [124]. By Fubini's theorem and Hölder's inequality, it suffices to bound the norm in $L^{p'}(D,dw)$ of $w\mapsto\int_D|z-w|^{-1}\,d\rho(z)$. For the part of the kernel with $2^{-j-1}<|z-w|\leq 2^{-j}$, $j\geq 0$, this function has a norm in $L^1(dw)$ at most $C\,2^{-j}$, and a norm in $L^\infty(dw)$ at most $C\,A\,2^{j(1-d)}$. Its norm in $L^{p'}$ is therefore at most $C(1+A)\,2^{-j(1-(2-d)/p)}$, which is summable since $p>2-d$. The part with $|z-w|>1$ is bounded. This includes $p=1$, with $p'=\infty$. □

14.3 Proof of Proposition 14.2

Condition (A1) holds for every p: since $\nu=dx$ is a Radon measure, $\mathcal{C}_p=\mathrm{Lip}_c(\mathbb{R}^2)$, which contains $C_c^\infty(\mathbb{R}^2)$. Thus p-admissibility is equivalent to the closability (A2).

Closability for $\gamma^2/2 < p < 2$. Let $1 \le p < 2$ with $p > \gamma^2/2$, and choose a rational d with $2 - p < d < 2 - \gamma^2/2$. Work on the event of Lemma 14.3(ii).

Step 1: an inequality for averages. Let x be a point at which the lower limit of Lemma 14.3(ii) is finite. Then there are a constant A_x and a sequence of radii $r \downarrow 0$ along which the measures $\mu_{x,r}$ have normalized restrictions satisfying the hypothesis of Lemma 14.4 with $A = A_x$. Let $\rho_{x,r}$ be the image of such a restriction on $B = B(x, r)$; it has a bounded density with respect to the restriction of M_γ to B, for each fixed r. By a change of variables in (14.11), along this sequence,

$$|u_{B,dx}| \le \int_B |u| \, d\rho_{x,r} + C_x \, r^{1-2/p} \, \|\nabla u\|_{L^p(B,dx)} \qquad (u \in \mathrm{Lip}_c(\mathbb{R}^2)) \,. \tag{14.12}$$

Step 2: local convergence. Suppose that $u_j \in \mathrm{Lip}_c(\mathbb{R}^2)$ tends to zero in $L^p(M_\gamma)$ and that $\nabla u_j \to g$ in $L^p(dx)$. In any ball, fix a smaller ball for which (14.12) holds. Applied to the differences $u_j - u_k$, it shows that their Lebesgue averages on the smaller ball tend to zero: the density of $\rho_{x,r}$ with respect to M_γ is bounded, and (∇u_j) is a Cauchy sequence in $L^p(dx)$. The Euclidean Poincaré inequality on the larger ball then makes (u_j) a Cauchy sequence in $W^{1,p}(dx)$ there. Let u be its local limit; then $\nabla u = g$.

Step 3: the limit vanishes. For each fixed x and r, the first term on the right-hand side of (14.12), applied to u_j, tends to zero, so that

$$|u_{B(x,r),dx}| \le C_x \, r^{1-2/p} \, \|g\|_{L^p(B(x,r),dx)} \,.$$

At a Lebesgue point of u and of $|g|^p$, the left-hand side tends to $|u(x)|$, and the right-hand side is $O(r)$, as $r \downarrow 0$ along the sequence. Hence $u = 0$ almost everywhere, and $g = \nabla u = 0$. Since the rational exponents d were handled on one event, the argument applies on that event to every p in the range considered.

The case $p \ge 2$ and the domain. The case $p = 2$ is Proposition 14.1. Let $p > 2$. Morrey's inequality [79, Theorem 5.1(3)] bounds the oscillation of a function on a ball by the norm of its gradient in L^p on a larger ball, and its average against M_γ on the ball controls the additive constant. A sequence which is Cauchy for the graph norm therefore converges locally uniformly, and the limit in $L^p(dx)$ of its gradients is the distributional gradient of the limit. If the sequence tends to zero in $L^p(M_\gamma)$, its continuous limit vanishes M_γ-almost everywhere, hence everywhere, because M_γ has full support. This proves the closability, and the inclusion of $W^{1,p}(\mathbb{R}^2, M_\gamma)$ in the right-hand side of (14.8).

For the reverse inclusion, let u belong to the right-hand side of (14.8). Truncate u at the heights $\pm L$, and multiply by a smooth cutoff equal to one on $B(0, R)$ and zero outside $B(0, 2R)$. The p-energy of the term containing the gradient of the cutoff is at most $C \, L^p \, R^{2-p}$, which tends to zero as $R \to \infty$, for fixed L, because $p > 2$. Mollification of the resulting compactly supported continuous function gives convergence in the graph norm. Let R and then L tend to infinity.

A carrier of zero Hausdorff measure. It is known that M_γ is carried by a set of Hausdorff dimension $2 - \gamma^2/2$. For the endpoint $p = \gamma^2/2$ we need the following more precise statement.

Lemma 14.5. *Let $s = 2 - \gamma^2/2$. Almost surely, there is a Borel set $S \subset \mathbb{R}^2$ with*

$$M_\gamma(\mathbb{R}^2 \setminus S) = 0 \qquad \text{and} \qquad \mathcal{H}^s(S) = 0 \,.$$

Proof. Let H_k be the field truncated to the scales $s' \ge e^{-2k}$ in (14.1). Its variance is $v_k = k + O(1)$.

Step 1: thick points for the rooted measure. Consider the measure $\mathbb{P}(d\omega)\, M_\gamma(\omega, dx)$, restricted to a bounded square. Under this measure, the conditional law of the sequence $(H_k(x) - \gamma v_k)_k$ given x is that of a Brownian motion at the times v_k. This follows from the Gaussian change of measure for the martingale of the scales [153, Theorem 17], [138, Section 4.1]. The law of the iterated logarithm for Brownian motion, together with Gaussian bounds on the bounded gaps between the times v_k, gives $H_k(x) \geq \gamma k + 2\sqrt{k}$ for infinitely many k, almost surely.

Step 2: discretization. Partition the plane into squares of side e^{-k}, and let $z_k(x)$ be the centre of the square which contains x. Under the same measure, the Gaussian variable $H_k(z_k(x)) - H_k(x)$ has a uniformly bounded mean and variance. Indeed its variance before the change of measure is

$$\int_{e^{-2k}}^{\infty} e^{-s'/2}\bigl(1 - e^{-|z_k(x)-x|^2/(2s')}\bigr)\, \frac{ds'}{s'} \leq C\,,$$

and its mean after the change of measure is $-\gamma/2$ times this variance. Gaussian tails and the Borel–Cantelli lemma give $|H_k(z_k(x)) - H_k(x)| = O(\sqrt{\log k})$. It follows that M_γ-almost every x belongs to infinitely many squares whose centres satisfy $H_k(z_k(x)) \geq \gamma k + \sqrt{k}$.

Step 3: the covering. Under $\mathbb{P}$, the probability that a given centre satisfies this inequality is at most $C_\gamma\, e^{-\gamma^2 k/2 - c_\gamma\sqrt{k}}$, for some $c_\gamma > 0$, and a bounded region meets at most $C\, e^{2k}$ squares of side e^{-k}. The expectation of the sum, over all k, of the sth powers of the diameters of the selected squares in that region is therefore at most $C \sum_k e^{-c_\gamma\sqrt{k}} < \infty$. The set of the points which belong to infinitely many selected squares is covered by the selected squares of the levels $k \geq k_0$, for every k_0; hence its measure $\mathcal{H}^s$ is zero. Exhausting the plane by bounded squares gives the set S. □

Vanishing of the relaxed energy for $p \leq \gamma^2/2$. Let $1 \leq p \leq \gamma^2/2$, and let S be the set of Lemma 14.5. Then $2 - p \geq s$ and $\mathcal{H}^{2-p}(S) = 0$.

Step 1: cutoffs of small norm. Every compact set $K \subset S$ admits functions equal to one near K, with values in $[0, 1]$, whose norm in $W^{1,p}(dx)$ is arbitrarily small. Indeed, cover K by finitely many balls of radii $r_i < 1$ with $\sum_i r_i^{2-p}$ arbitrarily small, and take the maximum of the distance cutoffs of these balls. Its p-energy is at most $C \sum_i r_i^{2-p}$, and its norm in $L^p(dx)$ is at most $(C \sum_i r_i^2)^{1/p}$; compare [124].

Step 2: the relaxed energy. Let $\varphi \in C_c^\infty(\mathbb{R}^2)$. Choose compact sets $K_j \subset S \cap \operatorname{supp}\varphi$ with $M_\gamma(\operatorname{supp}\varphi \setminus K_j) \to 0$, and functions χ_j as in Step 1 for K_j whose norms in $W^{1,p}(dx)$ tend to zero. Then

$$\varphi\chi_j \to \varphi \quad \text{in } L^p(M_\gamma)\,, \qquad \nabla(\varphi\chi_j) \to 0 \quad \text{in } L^p(dx)\,.$$

Hence $\overline{\mathcal{E}}_p(\varphi) = 0$. Since $C_c^\infty(\mathbb{R}^2)$ is dense in $L^p(M_\gamma)$ and $\overline{\mathcal{E}}_p$ is lower semicontinuous, (14.9) follows.

Step 3: the gradient is not closable. For $\varphi \neq 0$, the sequence $\varphi(1 - \chi_j) \in \mathrm{Lip}_c(\mathbb{R}^2)$ tends to zero in $L^p(M_\gamma)$, while its gradients tend to $\nabla\varphi \neq 0$ in $L^p(dx)$. Thus (A2) fails. This completes the proof of Proposition 14.2. □

14.4 Local Poincaré and cutoff inequalities

In the next two propositions the radii are sufficiently small, and the constants may depend on the realization and on a fixed compact set which contains the balls, the annuli and their enlargements. Both estimates use the bounds (14.6) for the volume: the upper bound gives the Poincaré inequality, and the lower bound gives the cutoff inequalities. Since the two bounds have different exponents, so have the resulting scales.

Proposition 14.6. *Let*

$$0 < a < \frac{(2-\gamma)^2}{2}\,, \qquad b > \frac{(2+\gamma)^2}{2}\,. \tag{14.13}$$

Then $\mathrm{B}(r^b)$ *and* $\mathrm{PI}_2(r^a)$ *hold locally, uniformly on each fixed compact set. More precisely, for* $B = B(x,r)$ *and for every function* u *which belongs locally to* $\mathcal{F}$*,*

$$\int_B |u - u_{B,M_\gamma}|^2 \, dM_\gamma \le C\, M_\gamma(B) \left[1 + \log \frac{Cr^a}{M_\gamma(B)}\right] \int_{2B} |\nabla u|^2 \, dx \tag{14.14}$$

$$\le C\, M_\gamma(B)\, \log \frac{e}{r} \int_{2B} |\nabla u|^2 \, dx\,. \tag{14.15}$$

The argument of the logarithm in (14.14) *is at least one, and the coefficient in* (14.14) *is at most* $C\, r^a$*.*

The inequality (14.15) shows that the Poincaré inequality holds, up to a logarithm, with the volume of the ball itself in place of a power of the radius.

Proof. Write $\mu = M_\gamma$.

The cutoff condition. The lower bound for the volume, and the bounds for the support and the Lipschitz constant of the Euclidean distance cutoff, give

$$\int_{\mathbb{R}^2} |\nabla \psi_x^r| \, dx \le C\, r \le C\, r^{1-b}\, \mu(B(x,r))\,, \tag{14.16}$$

which is (3.9) with $\Psi_{C,2}(r) = r^b$.

A bound for the Green potential. Let g_{2B} be the Green function of $-\Delta$ on $2B$ with the Dirichlet boundary condition, and put $v = \mu(B)$. The logarithmic bound for g_{2B} and the upper bound for the volume give

$$\sup_{z \in 2B} \int_B g_{2B}(z,w)\, d\mu(w) \le C \int_0^{4r} \min\{v, Cs^a\}\, \frac{ds}{s} \le C\, v \left[1 + \log \frac{Cr^a}{v}\right]. \tag{14.17}$$

The trace inequality. By Schur's test, the last coefficient bounds the norm of the Green operator on $L^2(B,\mu)$. For a bounded function h, the energy of the Green potential of $h\,\mu|_B$ is therefore at most this coefficient times $\|h\|^2_{L^2(B,\mu)}$. Pairing this potential with a function w in the Dirichlet form, and taking the supremum over $\|h\|_{L^2(B,\mu)} \le 1$, gives the trace inequality of [11, Proposition 5.2], [105, Section 9]:

$$\int_B w^2 \, d\mu \le C\, v \left[1 + \log \frac{Cr^a}{v}\right] \int_{2B} |\nabla w|^2 \, dx \qquad (w \in H_0^1(2B))\,.$$

The Poincaré inequality. Apply the trace inequality to the product of a cutoff of $2B$, equal to one on B, with $u - c$, where c is the Lebesgue average of u on $2B$. The Euclidean Poincaré inequality bounds the energy of this product by $C \int_{2B} |\nabla u|^2 \, dx$, and minimizing over the constants gives (14.14) for Lipschitz u. The lower bound for the volume gives (14.15), and $v\,[1 + \log(Cr^a/v)] \le C\, r^a$ because $v \le Cr^a$. Approximation in the local norm of the form extends the estimates to the stated domain. □

Proposition 14.7. *For every* $\beta > (2+\gamma)^2/2$*, the form* (14.5) *satisfies* $\mathrm{CSA}(r^\beta)$ *locally, uniformly on each fixed compact set: for* $A = B(x, R+r) \setminus \overline{B(x,R)}$*, there is a cutoff function* χ *for* $B(x,R) \subset B(x,R+r)$ *such that*

$$\int_A f^2 |\nabla \chi|^2 \, dx \le \frac{1}{8} \int_A \chi^2 |\nabla f|^2 \, dx + C\, r^{-\beta} \int_A f^2 \, dM_\gamma \qquad (f \in \mathcal{F})\,. \tag{14.18}$$

Proof. Choose b with $(2+\gamma)^2/2 < b < \beta$.

Step 1: a local inequality. Let $B(z,2s)$ be a small ball, and let c be the Lebesgue average of f on $B(z,2s)$. Average the inequality $c^2 \le 2(f^2 + |f-c|^2)$ over $B(z,s)$ with respect to M_γ. The trace inequality of the preceding proof, the Euclidean Poincaré inequality and the lower bound for the volume give

$$\int_{B(z,s)} f^2\,dx \le C\,s^2 \log\frac{e}{s} \int_{B(z,2s)} |\nabla f|^2\,dx + C\,s^{2-b} \int_{B(z,s)} f^2\,dM_\gamma\,. \tag{14.19}$$

Step 2: the unweighted inequality. Take a radial cutoff function θ for $B(x,R) \subset B(x,R+r)$ with $|\nabla\theta| \le C/r$, whose gradient is supported in the middle half of A. Cover this support by balls of radius $s \le r/16$ whose doubles lie in A and have bounded overlap. Summing (14.19) yields

$$\int_A f^2|\nabla\theta|^2\,dx \le C\,r^{-2}s^2 \log\frac{e}{s} \int_A |\nabla f|^2\,dx + C\,r^{-2}s^{2-b} \int_A f^2\,dM_\gamma\,.$$

Choose $k > 1$ with $2 + k(b-2) < \beta$, and take $s = c\,r^k$ with c fixed and small. The coefficient of the energy is then bounded, and the coefficient of the second term is at most $C\,r^{-\beta}$. This is the unweighted inequality (7.7) at the scale r^β, on every such annulus.

Step 3: the weighted inequality. Apply the construction of Steps 1 and 2 of the proof of Lemma 7.8, which uses only the strong locality and the scaling of the power r^β: the widths of the successive annuli are proportional to $r\,q^{j/\beta}$, and the weights of their cutoff functions are $(1-q)q^{j-1}$. On the jth annulus the sum χ is at least q^j. Taking q sufficiently close to one gives the coefficient $1/8$ in (14.18), and the coefficient of the second term remains at most $C\,r^{-\beta}$. The same estimates, with f equal to one on the annuli, give the convergence of the sum in the norm of the form. This summation does not use volume doubling [10, Lemma 5.1]. □

Thus, locally, the Liouville form satisfies a Poincaré inequality and a cutoff Sobolev inequality at two different scales, r^a with $a < (2-\gamma)^2/2 < 2$ and r^β with $\beta > (2+\gamma)^2/2 > 2$, in agreement with the order of the scales of Proposition 7.4. The gap between the two exponents is at least 4γ; it reflects the multifractal structure of the measure.

14.5 Volume, exit times and heat kernels

Failure of global doubling.

Proposition 14.8. *For each fixed $r > 0$, almost surely*

$$\sup_{x\in\mathbb{R}^2} \frac{M_\gamma(B(x,2r))}{M_\gamma(B(x,r))} = \infty\,. \tag{14.20}$$

In particular, VD *fails globally.*

Proof. The ratio at the origin is unbounded. Choose a smooth function $h \ge 0$ with compact support in $B(0,2r) \setminus \overline{B(0,r)}$, positive on a ball. The shift $H \mapsto H + th$ multiplies M_γ by $e^{\gamma th}$ [34]. It leaves the denominator of the ratio at $x = 0$ unchanged, and makes the numerator tend to infinity as $t \to \infty$. By the Cameron–Martin theorem [38, Section 2.4], the laws of H and $H + th$ are equivalent, so the ratio at the origin exceeds every integer N with positive probability.

Ergodicity. The covariance (14.1) tends to zero at infinity. The Gaussian variables obtained by testing the field against compactly supported functions therefore become asymptotically independent under translations, which gives the mixing of the field under the integer translations [123]. Its

chaos, which is covariant under translations, is therefore ergodic. By the ergodic theorem, applied to the event that the ratio at the origin exceeds N, almost surely there are centres at which the ratio exceeds N, for every N. □

For finer estimates on the masses of adjacent balls, see [33, Appendix A].

Exit times.

Proposition 14.9. *Let $\tau_{x,r}$ be the first exit time of the Liouville Brownian motion from $B(x,r)$, for a fixed realization of M_γ. Then*

$$\mathbb{E}_x \tau_{x,r} = \frac{1}{2\pi} \int_{B(x,r)} \log \frac{r}{|x-y|}\, dM_\gamma(y)\,. \tag{14.21}$$

Locally uniformly in x, for small r,

$$\frac{\log 2}{2\pi}\, M_\gamma(B(x,r/2)) \le \mathbb{E}_x \tau_{x,r} \le C\, M_\gamma(B(x,r))\, \log \frac{e}{r}\,. \tag{14.22}$$

At every point where the local dimension of M_γ exists, the exponent of the mean exit time is equal to this dimension. In particular,

$$\lim_{r\downarrow 0} \frac{\log \mathbb{E}_x \tau_{x,r}}{\log r} = 2 - \frac{\gamma^2}{2} \qquad \text{for } M_\gamma\text{-almost every } x\,. \tag{14.23}$$

Proof. The occupation formula for the time change expresses the mean exit time as the Green potential of M_γ in the ball, which is (14.21) [11, Lemma 6.3]. Restricting the integral to $B(x,r/2)$ gives the lower bound. The upper bound follows from (14.17) and the lower bound for the volume, as in [11, Lemma 6.4]. Taking logarithms gives the exponent. □

Thus the mean exit time from a ball is governed by the volume of concentric balls for M_γ, in agreement with the constant capacity (14.7), and not by a function of the radius: the time scale of the process depends on the point.

Distances and harmonic functions. Since $d\Gamma(u) = |\nabla u|^2\, dx$ and M_γ is singular, the intrinsic pseudo-distance of the Dirichlet form, defined by the constraint $d\Gamma(u) \le M_\gamma$, vanishes identically [70, Proposition 3.1]. In contrast, the constraint $|\nabla u| \le 1$ with respect to the gradient measure dx, as in Proposition 1.11, recovers the Euclidean distance. This illustrates the role of the gradient measure in Definition 1.10. By (14.5) and Weyl's lemma, the harmonic functions of the form have representatives which are harmonic in the Euclidean sense, and they satisfy the Euclidean elliptic Harnack inequality. The measure M_γ enters the parabolic equation only: the derivative in time is integrated against M_γ.

The heat kernel. The Liouville heat kernel is jointly continuous [11, Theorem 1.1]. For a compact set K and $\beta > (2+\gamma)^2/2$, Andres and Kajino prove the upper bound [11, Theorem 1.2]

$$p_t(x,y) \le \frac{C \log(1/t)}{t}\, \exp\Big[-c\Big(\frac{|x-y|^\beta \wedge 1}{t}\Big)^{1/(\beta-1)}\Big] \qquad (x \in \mathbb{R}^2\,,\ y \in K\,,\ 0 < t \le 1/4)\,. \tag{14.24}$$

The exponent β ranges over the same interval as the exponent of the cutoff Sobolev inequality of Proposition 14.7. For M_γ-almost every x and every $\eta > 18$ there are $c_x, t_x > 0$ such that [11, Theorem 1.3]

$$p_t(x,x) \ge \frac{c_x}{t\,(\log(1/t))^\eta} \qquad (0 < t \le t_x)\,. \tag{14.25}$$

Consequently the pointwise spectral dimension is two at M_γ-almost every point [11, Corollary 1.4]. On the $\sqrt{8/3}$-Liouville quantum gravity sphere, Andres, Kajino, Kavvadias and Miller obtain two-sided bounds for the heat kernel in terms of the quantum distance d_h, with a decay away from the diagonal governed by $(d_h(x,y)^4/t)^{1/3}$, up to logarithmic factors in the exponent, and bounds on the diagonal at every point [12, Theorems 1.1–1.2].

15 Riemannian gluings and the tree of cubes

When Riemannian pieces are glued along parts of their boundaries, the gradient is the Riemannian gradient on each piece and the gradient measure is the sum of the Riemannian volumes. The reference measure is a separate datum. If it is the sum of the volumes, the gluing behaves like a manifold (Section 15.1). If it is singular with respect to the volumes, the scales change. The simplest instance is a local tree, glued from intervals, with a reference measure which may be singular with respect to the length (Section 15.2). The main example of the chapter, the *tree of cubes* of Sections 15.4 to 15.8, is its analogue in rank $n \geq 2$: cubes are attached to cubes through small windows, the gradient measure is the sum of the Lebesgue measures of the cubes, and the reference measure is a measure of dimension $Q > n$ which gives no mass to any cube. For this space we determine

- the exponents of admissibility: $p \geq p_*$, for an explicit threshold $p_* \leq 2$ (Theorem 15.8);
- the cutoff scale r^{Q+p-n} and the sharp Poincaré scale, which coincide exactly for $p > n$ (Propositions 15.7 and 15.10);
- the critical Korevaar–Schoen spaces, which are equal to $W^{1,p}(X, m)$ for $p > n$ and are proper dense subspaces for admissible $1 < p \leq n$ (Theorem 15.11).

In particular the tree of cubes is an example with constant rank n in which the two scales of Part I are distinct at the exponent two, and in which the bound $n_* < p$ of Theorem 5.2 at matching scales is attained.

15.1 Finite Riemannian gluings

The following example isolates the effect of a regular interface. For background on Riemannian polyhedra and on diffusions on complexes, see [62, 44]; the piecewise Dirichlet form and the local Poincaré inequalities on Euclidean complexes are treated in [135, Sections 1.2 and 2].

Proposition 15.1. *Let $M_1, \dots, M_k$ be compact connected smooth Riemannian manifolds of dimension n with nonempty smooth boundary. Identify their entire boundaries with the same compact Riemannian manifold Σ, by isometries of the induced metrics. Equip the quotient X with its length metric and with the sum m of the Riemannian volumes. Then:*

(i) *the piecewise Riemannian gradient, with $\nu = m$, gives a uniform gradient structure of rank n, which is p-admissible for every $1 \leq p < \infty$;*

(ii) *in the product of the Sobolev spaces of the pieces, $W^{1,p}(X, m)$ consists of the families $(u_i)_{i=1}^k$ whose traces on Σ agree;*

(iii) *at small radii, $\mathrm{V}(r^n)$, $\mathrm{B}(r^2)$ and $\mathrm{PI}_p(r^p)$ hold.*

Proof. Step 1: charts. The quotient has a smooth chart at each point away from Σ. A chart at a point of the interface consists of k Euclidean half-boxes identified along their common flat face, with Riemannian metrics and densities uniformly comparable to the Euclidean ones. This description is valid for the distance of the quotient on smaller charts, by the definition of the length metric: short curves from a smaller chart stay in the larger chart, and within the larger chart the length of every piecewise smooth curve is comparable to its length in the Euclidean model. Compactness gives a finite cover and uniform constants.

Step 2: condition (G) and recovery of the metric. The piecewise gradient satisfies the Leibniz rule, the Lipschitz bound and the weak-$*$ continuity, by differentiation in the sense of distributions

in the interior of each of the finitely many pieces. Compactly supported coordinate functions in the interiors have gradients which span $\mathbb{R}^n$ almost everywhere. For (1.11), suppose that $|\nabla u| \le L$ on an open set. Smooth approximation in the charts of the interiors bounds the variation of u along every piecewise smooth curve which is contained in this set and in a single piece by L times its length. Curves along a face of the boundary are limits, in length, of curves pushed into the collar, so the same bound holds for them. A curve of length less than $2r$ starting at a point of $B(x,r)$ stays in $B(x,3r)$, and concatenations of length approaching $d(y,z)$ exist by the definition of the length metric. Adding the bounds over a concatenation and taking the infimum over the curves contained in $B(x,3r)$ proves (1.11) on $B(x,r)$, for every r.

Step 3: admissibility and the domain. The piecewise gradient is closable, by testing separately in the interior of each piece; the interface is null for $m = \nu$, and Lipschitz functions are uniformly dense. The trace theorem on smooth half-boxes implies that the traces of the pieces of a limit in the graph norm agree on Σ. For the converse, localize with a partition of unity. In a chart at the interface, transport the function of the first half-box to each of the other half-boxes by the coordinates. This common part is approximated in $W^{1,p}$ by reflection across the face followed by convolution, with the same approximation on each sheet. Subtracting it leaves functions on the other half-boxes with zero trace on the face. After localization away from the other faces, extend them by zero, translate them inward by a distance tending to zero, and mollify. The resulting approximations vanish near the face and converge in $W^{1,p}$, also for $p = 1$. This produces Lipschitz functions on X which converge in the graph norm, and proves the assertion on the domain. For reflection and extension at a smooth boundary, see [42, Lemma 9.2 and Section 9.2] and [124]. The gluing of Sobolev functions with matching traces is treated more generally in [41, Section 5]; the gluing theorem proved there applies here for $p > 1$, while the argument above also covers $p = 1$.

Step 4: the scales. The charts give $m(B(x,r)) \asymp r^n$ uniformly in x, and $\mathrm{B}(r^2)$ follows from Lemma 3.8. For the Poincaré inequality near Σ, put $B(x,r)$ inside a union of k congruent half-boxes of diameter Cr in a chart at the interface, contained in a fixed enlargement of the ball. Let c be the mean of u on the first half-box. Pair each other half-box with the first one, by reflecting it into the opposite half-space. The restrictions of a Lipschitz function u fit together to give a Lipschitz function on the resulting full box. The Euclidean Poincaré inequality on this box, and the comparison of its mean with c, give

$$\int_{\text{pair}} |u - c|^p \, dm \le C \, r^p \int_{\text{pair}} |\nabla u|^p \, dm \,;$$

compare the argument with chains of bounded length for complexes in [135, Section 2]. Summing over the sheets and comparing c with u_B proves the required inequality. Balls in the interior use a full box in a smooth chart, and the estimate on a single half-box covers $k = 1$. □

For $k = 2$ this includes the double of any compact smooth manifold with boundary. For $k \ge 3$ the interface has at least three incident pieces, and the space is not a manifold. In all these examples the scales are Gaussian. In the tree of cubes below, m is singular with respect to the piecewise volume: the gauge is $\Theta(r) = r^n$ by Proposition 15.7(ii), while $\Phi(r) = r^Q$ with $Q > n$, and the Poincaré scale is computed separately in Proposition 15.10.

15.2 Local trees

Local trees are the case of rank one of the gluings considered in this chapter. We use the definitions and the notation of [22, Sections 2.1 and 2.2]. A *local tree* is a metric space (X,d) which is locally isometric to real trees. Its *skeleton* $\mathcal{S}$ is the union of the open geodesic segments of X, the *length*

measure ν is the one-dimensional Hausdorff measure restricted to $\mathcal{S}$, and the *weak gradient* ∂u of a Lipschitz function u is its derivative along the skeleton, which is defined ν-almost everywhere, up to sign. Metric graphs and compact real trees are local trees. The reference measure m is any Radon measure with full support satisfying a uniform control of the volume; it may be singular with respect to ν, and ν need not be locally finite.

Proposition 15.2. *Let X be a proper uniform local tree, and let m be a Radon measure with full support satisfying the uniform volume control of [22], with the volume function Φ. A measurable orientation of the arcs defines a gradient ∇ of rank one, with $|\nabla u| = |\partial u|$. The resulting gradient structure is p-admissible for every $1 \le p < \infty$. At sufficiently small radii,*

$$\mathrm{V}(\Phi)\,, \qquad \mathrm{B}(r\,\Phi(r))\,, \qquad \mathrm{PI}_p(r^{p-1}\Phi(r)) \quad (1 \le p < \infty)\,.$$

For $1 < p < \infty$ its Sobolev space is the Sobolev space of [22, Definition 2.13].

Thus $\Theta(r) = r$, and the scales match at every exponent: $\Psi_{P,p}(r) = \Psi_{C,p}(r) = r^{p-1}\Phi(r)$. In particular the time scale of the diffusion is $\Psi_{C,2}(r) = r\,\Phi(r)$; compare (9.11) and [22, Theorem 5.1].

Proof. Step 1: the derivation and recovery of the metric. Cover $\mathcal{S}$ by countably many oriented open arcs, and define ∇u on each arc, outside the earlier arcs, as the derivative of u in the direction of the orientation. Differentiation in one variable gives the Leibniz rule and the Lipschitz bound. The convergence of the derivatives in the sense of distributions on each arc, followed by the approximation of a field of $L^1(\nu)$ by its restriction to finitely many arcs, gives the weak-$*$ continuity. For two points in a small ball which is a tree, integrating along the arc which joins them gives (1.11).

Step 2: cutoffs. The distance cutoff is not suitable here: the distance to a point varies along the whole skeleton, and ν may be infinite on every ball. We use instead a cutoff which varies only along finitely many arcs. Take r so small that $B(x, 3r)$ is isometric to a ball of a real tree, and choose a maximal $(r/8)$-separated set in $B(x, 2r)$; the hypothesis on the volume bounds its cardinality. Let F be the convex hull, in this tree, of these points and of x. It is a union of a uniformly bounded number of arcs. The nearest-point projection π_F onto F is 1-Lipschitz, and $d(z, \pi_F z) \le r/8$ on $B(x, 2r)$. Define on this ball

$$\psi_x^r(z) = \bigl(3 - 2\,d(x, \pi_F z)/r\bigr)_+ \wedge 1\,,$$

and extend it by zero. It is equal to one on $B(x, r)$ and its support lies in $\overline{B}(x, 13r/8)$, so the extension is C/r-Lipschitz. The projection is locally constant outside F, so $\nabla\psi_x^r$ is carried by F. Each arc of F meets a ball $B(y, s)$ in a set of length at most $2s$, because the arc is isometric to an interval and the set has diameter at most $2s$. Consequently

$$\int_{B(y,s)\cap\mathcal{S}} |\nabla\psi_x^r|\,d\nu \le \frac{C}{r}\,\nu(F \cap B(y, s)) \le C\,\frac{s}{r} \qquad (y \in X\,,\ s > 0)\,. \tag{15.1}$$

For $s = 2r$ this is $\mathrm{B}(r\,\Phi(r))$. For related cutoffs on finite trees see [22, Lemma 3.3].

Step 3: the core spans the fibres. Use the arclength coordinate along the projection onto a compact subarc of one of the countably many oriented arcs, and multiply it by a finite maximum of the cutoffs of Step 2 which is equal to one near the subarc and supported in a tree neighbourhood of it. This gives a compactly supported Lipschitz function of finite 1-energy whose gradient has absolute value one in the interior of the subarc. Such subarcs cover the skeleton up to a set of zero length.

Step 4: the Poincaré inequality and closability. If $y, z \in B(x,r)$, the arc from y to z has length at most $2r$ and stays in the tree $B(x,3r)$. Integration along this arc, as in [22, Lemma 2.9 and Theorem 2.12], gives

$$|u(y) - u(z)|^p \le (2r)^{p-1} \int_{B(x,3r)\cap\mathcal{S}} |\nabla u|^p \, d\nu \,. \tag{15.2}$$

Averaging in z and integrating in y gives $\mathrm{PI}_p(r^{p-1}\Phi(r))$; the estimate is valid also at $p = 1$. The cutoffs give (A1). For (A2), let $u_j \to 0$ in $L^p(m)$ with ∇u_j convergent in $L^p(\nu)$. By (15.2), applied to the differences and averaged over a fixed small ball, the sequence is uniformly Cauchy on the ball, and its limit there is zero because m has full support. Testing on the arcs against compactly supported smooth functions of the arclength shows that the limit of the derivatives is zero.

Step 5: identification. For $p > 1$, the identification with the Sobolev space of [22] follows from the approximation by Lipschitz partitions of unity of finite energy and weak compactness, followed by convex combinations; see [22, Theorem 3.5 and the proof of Theorem 3.7]. □

The estimate (15.1) is local: together with $|\nabla \psi_x^r| \le C/r$ it gives

$$\int_{B(y,s)\cap\mathcal{S}} |\nabla \psi_x^r|^2 \, d\nu \le C \, \frac{s}{r^2} \asymp \Big(\frac{s}{r}\Big)^2 \frac{m(B(y,s))}{s\,\Phi(s)} \qquad (0 < s \le r) \,,$$

which is the condition $\mathrm{B}^+(\Psi_{C,2})$ of Definition 9.2.

15.3 Countable Riemannian gluings

We now define the class of gluings to which the tree of cubes belongs, and verify the properties which do not depend on the reference measure.

Definition 15.3. Let P_j, $j \ge 0$, be compact connected Riemannian manifolds of dimension n with Lipschitz boundary, possibly with corners. Identify compact pieces of hypersurfaces of these manifolds, or points when $n = 1$. Let $\mathcal{S}$ be the quotient, and assume the following.

(i) The quotient map is injective on each piece, and the intersection of two distinct pieces has zero volume in each of them.

(ii) The infimum of the lengths of the finite concatenations of curves in the pieces defines a finite metric d on $\mathcal{S}$. In each piece, curves admit approximations, in length and in position, by piecewise smooth curves with the same endpoints whose interiors lie in the interior of the piece.

(iii) The completion X of $(\mathcal{S}, d)$ is compact.

Let m be a finite Radon measure with full support on X. On $\mathcal{S}$ set

$$\nu = \sum_{j\ge 0} \mathrm{vol}_{P_j} \,, \qquad \nabla u|_{P_j} = \nabla^{P_j}(u|_{P_j}) \,, \qquad u \in \mathrm{Lip}(X) \,. \tag{15.3}$$

A measurable orthonormal frame on each piece identifies the gradient with a function with values in $\mathbb{R}^n$. We call (X, d, m), with $(\mathcal{S}, \nu, \nabla)$, a *countable Riemannian gluing*.

The inclusion of a piece is 1-Lipschitz for its intrinsic distance, so each piece is compact in X and $\mathcal{S}$ is a Borel set. The completion of the length metric is a compact length space, hence geodesic. The condition on curves in (ii) holds for Euclidean cubes. For finite gluings and the associated diffusions with the reference measure ν, see [62, 44, 135]. Quantitative gluing theorems

for Poincaré inequalities, under assumptions of thickness on the identified sets, are given in [82, Theorem 6.15] and [80, Theorem 3.3]. Those theorems use the measures of the pieces; they do not cover directly the singular reference measure considered below.

Proposition 15.4. *The operator ∇ of (15.3) satisfies (G). For $u \in \mathrm{Lip}(X)$, $r > 0$ and $y, z \in B(x, r)$,*

$$|u(y) - u(z)| \le d(y, z)\, \|\nabla u\|_{L^\infty(B(x,3r)\cap\mathcal{S},\nu)}\,. \tag{15.4}$$

With the minimal field of Definition 1.2, $(\mathcal{S}, \nu, \nabla)$ is a uniform gradient structure. It has constant rank n if the gradients of the functions of $\mathrm{Lip}(X)$ generate $\mathbb{R}^n$ almost everywhere. In particular, if a sequence $(u_k) \subset \mathcal{C}_1$ satisfies

$$\mathrm{span}\{\nabla u_k(x) : k \ge 1\} = \mathbb{R}^n \quad \text{for } \nu\text{-almost every } x\,, \tag{15.5}$$

then the structure has constant rank n and each $\mathcal{C}_p$ spans the fibres.

Thus the axioms and the recovery of the metric hold for every reference measure. What depends on m is the admissibility and the quantitative hypotheses; moreover, when ν is not locally finite, the existence of sufficiently many Lipschitz functions of finite energy, as in (15.5), is part of the problem.

Proof. Condition (G). The restriction of u to a piece has a Lipschitz constant at most $\mathrm{Lip}(u)$ for the intrinsic metric of the piece. The Leibniz rule and the bound $|\nabla u| \le \mathrm{Lip}(u)$ follow by differentiating inside each piece. A sequence which converges pointwise, with bounded Lipschitz constants, converges uniformly on each piece, and its derivatives converge in the sense of distributions; the uniform bound on the gradients then gives the weak-$*$ convergence on each piece. For a field of $L^1(\nu)$, the approximation by its restrictions to finitely many pieces gives the weak-$*$ convergence on $\mathcal{S}$.

Recovery of the metric. Put $L = \|\nabla u\|_{L^\infty(B(x,3r)\cap\mathcal{S},\nu)}$. In a coordinate chart, mollification and the comparison of the Riemannian metric with its value at the centre of the chart show that the variation of u along any piecewise smooth curve in the interior of a piece, contained in $B(x, 3r)$, is at most L times its length: the chart can be taken with distortion $1 + \varepsilon$, and a subdivision of the curve, followed by $\varepsilon \downarrow 0$, gives the stated constant. The approximation in (ii) gives the same bound for the curves in the piece, including its boundary. Let $y, z \in B(x, r) \cap \mathcal{S}$, and consider a concatenation of length ℓ joining them. At the arclength s its distance to x is at most $\min\{d(x, y) + s,\, d(x, z) + \ell - s\} \le (d(x, y) + d(x, z) + \ell)/2$, so it lies in $B(x, 3r)$ if ℓ is sufficiently close to $d(y, z)$. Add the bounds over the pieces, take the infimum over ℓ, and extend to X by the density of $\mathcal{S}$. This gives (15.4), hence (1.11) with any constant ι.

The rank. The assertion on the rank follows from the minimality of the fibres. A family satisfying (15.5) belongs to $\mathcal{C}_p$ for all $p \ge 1$, by (1.9). □

15.4 The tree of cubes: construction and volume growth

Hanson and Heinonen [80, Section 3] construct a metric space by the successive attachment of Euclidean cubes along sets which are thick in a quantitative sense. In the construction below, the relative sizes of the interfaces tend to zero, and the reference measure is singular with respect to the sum of the Euclidean volumes.

Parameters. Fix an integer $n \geq 2$, a real number $Q > n$, a parameter $0 < \sigma < 2$, and a constant $0 < c \leq (64\sqrt{n})^{-1}$. For $j \geq 0$ set

$$h_j = 2^{-j}\,, \qquad \delta_j = \begin{cases} c\, h_j \exp(-h_j^{-\sigma})\,, & n = 2\,, \\ c\, h_j^{1+\sigma/(n-2)}\,, & n > 2\,, \end{cases} \qquad \theta_j = \frac{\delta_j}{h_j}\,. \tag{15.6}$$

The number h_j is the side of the cubes of the level j, and δ_j is the side of the windows through which the new cubes of the level j are attached. The windows are small with respect to the cubes: $\theta_j \to 0$, at a rate governed by σ. Let

$$a = \lfloor 2^Q \rfloor\,, \qquad b = \frac{\log(2^Q/a)}{\log((a+1)/a)} \in [0,1)\,, \qquad N_j = a + \lfloor jb \rfloor - \lfloor (j-1)b \rfloor \in \{a, a+1\}\,.$$

The number N_j is the number of children of a cube of the level $j-1$. Since $Q > n$, $K_j = N_j - 2^n \geq 0$, and

$$\prod_{k=i+1}^{j} N_k \asymp 2^{(j-i)Q} \qquad (0 \leq i < j)\,. \tag{15.7}$$

Indeed the number of factors equal to $a+1$ differs from $(j-i)b$ by at most one. Moreover $K_j \geq 1$ for infinitely many j, with bounded gaps: either $a > 2^n$ and $K_j \geq 1$ for all j, or $a = 2^n$, in which case $b > 0$ and $K_j = 1$ whenever $N_j = a+1$.

Construction. Start with $X_0 = P_0 = [0,1]^n$ and $\mathcal{D}_0 = \{P_0\}$. Suppose that X_{j-1} and a family $\mathcal{D}_{j-1}$ of cubes of side h_{j-1}, the *mesh cubes* of the level $j-1$, have been constructed. For each mesh cube $D \in \mathcal{D}_{j-1}$:

- take its 2^n closed dyadic subcubes, the *ordinary children* of D;
- attach to D a number K_j of copies of the cube $[-h_j/2, h_j/2]^{n-1} \times [0, h_j]$, the *new children* of D, by identifying the centred $(n-1)$-dimensional cube of side δ_j in the lower face of each new cube with the centred $(n-1)$-dimensional cube of side δ_j in the horizontal hyperplane through the centre of D.

There are no other identifications. The quotient, with its length metric, is X_j, and $\mathcal{D}_j$ consists of the ordinary and the new children of the cubes of $\mathcal{D}_{j-1}$. Thus all the mesh cubes of the level j have the side h_j, and each cube of $\mathcal{D}_{j-1}$ has N_j children. A *piece* is either P_0 or one of the new cubes; a new cube of the level i is a piece of *generation* i. Let $\mathcal{P}$ be the collection of the pieces. For a new piece P, its *host* P^- is the piece which contains the mesh cube to which P is attached, and its *interface* F_P is the cube of dimension $n-1$ along which it is attached. An interface in a piece P is either the incoming interface F_P, in its lower face, or the interface of a piece whose host is P. Define

$$\mathcal{S} = \bigcup_j X_j\,, \qquad X = \overline{\mathcal{S}}\,, \qquad \nu = \sum_{P \in \mathcal{P}} \mathcal{L}^n \restriction_P\,, \qquad \nabla u|_P = \nabla_{\mathbb{R}^n}(u|_P)\,,$$

where the closure is the metric completion. For $D \in \mathcal{D}_j$, the *cell* U_D is the closure of the union of D and of its descendants. The reference measure m is the image of the product probability measure which chooses one child uniformly at each step; the image of a sequence of nested mesh cubes is the intersection of their cells. By (15.7), the mass given by this product measure to the sequences passing through a cube of the level j is comparable to h_j^Q.

The resulting space is a countable Riemannian gluing in the sense of Definition 15.3. Its combinatorial structure is that of a tree: each cube carries, besides its dyadic subcubes, new cubes attached through small windows. Its pieces have dimension n, and its reference measure distributes the mass equally among the $N_j \asymp 2^Q$ children of each cube, as on a regular tree.

Proposition 15.5. (i) *The space X is compact and geodesic, and each piece is isometrically embedded.*

(ii) *For every fixed $L < \infty$, a ball of radius $L\,h_j$ in X_j meets at most C_L mesh cubes of the level j.*

(iii) *There are 1-Lipschitz retractions $\pi_j : X \to X_j$ such that*

$$\pi_j(U_D) \subset D \quad (D \in \mathcal{D}_j)\,, \qquad \sup_{x\in X} d(x, \pi_j x) \le \sqrt{n}\,h_j\,. \tag{15.8}$$

(iv) *Distinct cells of the same level meet in a set which is null for m and for ν, and*

$$\operatorname{diam} U_D \le 3\sqrt{n}\,h_j\,, \qquad m(U_D) \asymp h_j^Q\,, \qquad m(B(x,r)) \asymp r^Q \quad (0 < r \le \operatorname{diam} X)\,. \tag{15.9}$$

(v) *The measures m and ν are mutually singular, and ν is σ-finite and infinite on every nonempty open ball.*

Proof. Step 1: retractions and compactness. The gluing theorem for CAT(0) spaces [43, Theorem II.11.1] applies at every attachment, since the interfaces are convex. Hence the earlier stages and the individual pieces are convex in the subsequent stages. The orthogonal projection of each new piece onto its incoming interface gives a 1-Lipschitz retraction $X_i \to X_{i-1}$ which moves every point by at most $\sqrt{n}\,h_i$. Composing these retractions, and using $\sum_{i>j} h_i = h_j$, proves (15.8) and the bound for the diameter in (15.9). The finite stages are compact nets of X of arbitrarily small mesh, so the completion is compact. It is geodesic, as the completion of a geodesic space in which the distance is convex along geodesics.

Step 2: the mesh bound. Distinct interfaces in a piece of X_j are h_j-separated. Indeed, in the coordinates of the piece, the interfaces of generation $i \le j$ have a normal coordinate which is an odd multiple of h_i, the incoming interface has the normal coordinate zero, and the interfaces of the same generation have centres which are $2h_i$-separated. Along a geodesic of length at most $L\,h_j$ issued from the centre of the ball, distinct interfaces are therefore crossed at most $\lfloor L \rfloor + 1$ times. Each piece meets a bounded number of interfaces within the ball, by packing in the Euclidean space, and each interface meets at most $1 + \sup_j K_j$ pieces. Hence the number of pieces met by the ball is bounded, and in each piece the packing gives the bound (ii).

Step 3: cells and the measure m. Distinct cells of the same level meet only in the intersection of their mesh cubes. Indeed the later attachments lie in a unique earlier cell, because the interfaces in a piece lie in the interiors of its mesh cubes of every strictly earlier level; the assertion then follows by projecting onto the finite stages and passing to the limit. A ball of radius $r \asymp h_j$ in X meets a bounded number of cells of the level j, by (15.8) and (ii), and the sequences passing through a cube of the level j have the mass $\asymp h_j^Q$, by (15.7). This proves the upper bound $m(B(x,r)) \le C\,r^Q$. Consequently each piece is m-null, since it is covered by $O(r^{-n})$ balls of radius r and $Q > n$. It follows that the mass of a cell is the mass of the sequences passing through its cube, that distinct cells meet in an m-null set, and that $m(U_D) \asymp h_j^Q$. For the lower bound for the volume of balls, take a cell which contains the centre, with $3\sqrt{n}\,h_j < r$ and $h_j \asymp r$.

Step 4: the measure ν. Distinct mesh cubes meet in ν-null sets, which gives the corresponding assertion for the cells. Since m gives no mass to the pieces, and ν is carried by their union, $m \perp \nu$. Each piece has finite volume, so ν is σ-finite. Finally, for a fixed $D \in \mathcal{D}_j$, the descendants of D of the level k tile $U_D \cap X_k$, and their total Euclidean volume is $\asymp h_j^Q\, h_k^{n-Q}$, which tends to infinity as $k \to \infty$; hence $\nu(U_D) = \infty$. Each ball contains a cell. □

Thus $\Phi(r) = r^Q$: the space is Ahlfors Q-regular, although its pieces have dimension $n < Q$.

15.5 The tree of cubes: gradient and cutoffs

By Proposition 15.4, the piecewise gradient satisfies (G) and the recovery of the metric. Since ν is infinite on every ball, a Lipschitz function has in general an infinite energy; for instance $u \circ \pi_j$, for the retractions of Proposition 15.5, varies in the whole of each later piece. The first task is therefore to construct Lipschitz functions of finite energy. We modify the retractions so that a function of the form $u \circ \widetilde{\pi}_j$ is constant on each later piece outside a small neighbourhood of its window.

Modified retractions. For the incoming interface F of a new piece P of generation i, let z_F be its centre and $R = \sqrt{n-1}\,\delta_i/2$ its circumradius. The map which fixes F pointwise and sends $P \setminus B(z_F, 2R)$ to the point z_F is 1-Lipschitz on its domain. Kirszbraun's theorem [15, Section 1], followed by the orthogonal projection onto F, extends it to a 1-Lipschitz retraction $P \to F$. Composing these modified retractions of the successive stages gives maps

$$\widetilde{\pi}_j : X \to X_j \qquad \text{with the properties (15.8) and} \quad \widetilde{\pi}_i \circ \widetilde{\pi}_j = \widetilde{\pi}_{\min\{i,j\}}\,.$$

Put

$$\mathcal{A} = \{u \circ \widetilde{\pi}_j :\ j \geq 0\,,\ u \in \mathrm{Lip}(X_j)\}\,, \qquad S_j = \sum_{k>j} \theta_k^n\,.$$

The numbers S_j tend to zero; they measure the total relative volume of the neighbourhoods of the windows of the generations after j.

Lemma 15.6. *Let $u \in \mathrm{Lip}(X_j)$ and $D \in \mathcal{D}_j$.*

(i) *In a piece of generation $i > j$, the function $u \circ \widetilde{\pi}_j$ is constant outside the ball of radius $C\delta_i$ about the centre of the incoming interface.*

(ii) *For $k \geq j$, at most $C\,(h_j/h_k)^n$ of the mesh cubes of the level k in U_D carry a nonconstant restriction of $u \circ \widetilde{\pi}_j$.*

(iii) *For $1 \leq p < \infty$,*

$$\int_{U_D \setminus X_j} |\nabla(u \circ \widetilde{\pi}_j)|^p \, d\nu \leq C\, h_j^n\, S_j\ \mathrm{Lip}(u|_D)^p\,. \tag{15.10}$$

Proof. (i) holds because the modified retraction of a new piece is constant outside the ball of radius $2R$ about the centre of its interface.

(iii) Call a piece *active* if the restriction of $u \circ \widetilde{\pi}_j$ to it is not constant, and give to a piece of generation k the weight δ_k^n. An active piece of generation $i > j$ has at most $C\,(\delta_i/h_k)^n$ active pieces of generation $k > i$ attached directly to it, and none when $h_k > C\delta_i$: their interfaces must meet the region where the function varies, and all of them are at a distance at least h_k from the incoming face. The total weight of these pieces is at most $C\,\delta_i^n\, S_i$. Since $S_i \to 0$, the total weight

of an active piece of generation i and of all the active pieces attached to it, directly or through a chain of attachments, is at most $C\,\delta_i^n$: for large i this follows from a geometric series of ratio $1/2$, and for the remaining finitely many generations by backward induction. The number of pieces of generation $k > j$ attached directly to a subcube of D is at most $C\,(h_j/h_k)^n$. Thus the total weight of the active pieces in $U_D \setminus X_j$ is at most $C\,h_j^n\,S_j$. By (i), the volume of the region where the function varies satisfies the same bound, and the Lipschitz constant of the function on U_D is at most $\mathrm{Lip}(u|_D)$, because $\widetilde{\pi}_j$ is 1-Lipschitz and maps U_D into D. This proves (15.10).

(ii) By the bounds just proved for direct attachments, the number of active pieces of generation $i > j$ in U_D is at most

$$C\,h_i^{-n}\Big(h_j^n + \sum_{P'} \delta_{j(P')}^n\Big) \le C\,\Big(\frac{h_j}{h_i}\Big)^n,$$

where the sum is over the active pieces P' attached earlier in the cell, and $j(P')$ is the generation of P'. In an active piece of generation $i \le k$, only $C\,[1 + (\delta_i/h_k)^n]$ mesh cubes of the level k meet the region where the function varies. Add over the active pieces, using the geometric sum of $(h_j/h_i)^n$ and the bound for the total weight, and include the $(h_j/h_k)^n$ subcubes of D itself. □

Proposition 15.7. (i) *The piecewise gradient defines a uniform gradient structure of constant rank n. The set $\mathcal{A}$ is an algebra, it is contained in $\mathcal{C}_p$ for every $1 \le p < \infty$, and it is uniformly dense in $C(X)$.*

(ii) *The structure satisfies* $\mathrm{B}(r^{Q+2-n})$ *and* $\mathrm{B}^+(r^{Q+2-n})$ *with a common choice of cutoffs $\psi = \psi_x^r \in \mathcal{A}$ satisfying*

$$\mathrm{Lip}(\psi) \le \frac{2}{r}, \qquad \int_{\mathcal{S}} |\nabla\psi|\,d\nu \le C\,r^{n-1}.$$

Thus $\Psi_{C,p}(r) = r^{Q+p-n}$ and $\Theta(r) = r^n$.

Proof. (i) By (15.10), summed over the finitely many cubes of $\mathcal{D}_j$, the functions of $\mathcal{A}$ have finite p-energy for every p, so $\mathcal{A} \subset \mathcal{C}_p$. The compatibility of the retractions shows that $\mathcal{A}$ is an algebra, and $u \circ \widetilde{\pi}_j \to u$ uniformly for $u \in \mathrm{Lip}(X)$, by (15.8). Extend the coordinate functions of each piece $P \subset X_j$, first to X_j by McShane's theorem [125], and then to X by composition with $\widetilde{\pi}_j$. The resulting functions belong to $\mathcal{A}$, and their gradients span $\mathbb{R}^n$ on P. Proposition 15.4 gives the remaining assertions.

(ii) Choose j with $h_j \asymp r$ and $\sqrt{n}\,h_j \le r/8$, and take

$$\psi(y) = \min\{1, (3 - 2\,d(\widetilde{\pi}_j y, \widetilde{\pi}_j x)/r)_+\}.$$

By (15.8) it is equal to one on $B(x,r)$, and its support lies in $B(x,2r)$. This ball meets a bounded number of cells of the level j. On each such cell, the energy on its mesh cube is at most $C\,h_j^n/r$, and the energy on the later pieces is at most $C\,h_j^n\,S_j/r$, by (15.10) at $p = 1$. Hence

$$\int_{\mathcal{S}} |\nabla\psi|\,d\nu \le C\,r^{n-1} \asymp \frac{r\,m(B(x,r))}{r^{Q+2-n}},$$

which is (3.9) with $\Psi_{C,2}(r) = r^{Q+2-n}$.

For $0 < s \le r$, take $k \ge j$ with $h_k \asymp s$. The compatibility of the retractions gives $\psi = (\psi|_{X_k}) \circ \widetilde{\pi}_k$. Every level-$k$ cell meeting $B(y,s)$ projects by π_k to a mesh cube meeting $B_{X_k}(\pi_k y, s)$, so the mesh bound of Proposition 15.5(ii) gives a bounded number of such cells. On each corresponding mesh cube the Lipschitz bound gives an energy of at most $C h_k^n/r^2$. The energy on the later pieces in

the cell is at most $Ch_k^n S_k/r^2$ by (15.10). Since (S_k) is bounded and $m(B(y,s)) \asymp s^Q$, summation gives

$$\int_{B(y,s)\cap\mathcal{S}} |\nabla\psi|^2 \, d\nu \le C\,\frac{s^n}{r^2} \le C'\left(\frac{s}{r}\right)^2 \frac{m(B(y,s))}{\Psi_{C,2}(s)}\,,$$

which is $\mathrm{B}^+(\Psi_{C,2})$. □

15.6 The tree of cubes: admissibility and Poincaré inequalities

A function of finite energy on the tree of cubes is a family of Euclidean Sobolev functions on the pieces which agree on the windows. The values on a new piece are related to the values on its host only through the window, and the strength of this relation is measured by the capacity of the window. This capacity, at the exponent p, is of the order of the inverse of the quantity κ_p defined below; it decides both the admissibility and the Poincaré scale.

Two Euclidean estimates for windows. For a Euclidean cube D, let $\langle u\rangle_D$ denote the mean of u on D for the Lebesgue measure. Let F be a flat cube of dimension $n-1$ and side δ, centred in a face or in the central hyperplane of a cube D of side comparable to h. Then

$$\left|\fint_F u\,d\mathcal{H}^{n-1} - \langle u\rangle_D\right|^p \le C\,\kappa_p(h,\delta)\int_D |\nabla u|^p\,dx\,, \qquad \kappa_p(h,\delta) = \begin{cases} \delta^{p-n}\,, & 1\le p<n\,,\\ [1+\log(h/\delta)]^{n-1}\,, & p=n\,,\\ h^{p-n}\,, & p>n\,. \end{cases} \tag{15.11}$$

Indeed, on the box of the scale δ whose cross-section is F, the fundamental theorem of calculus in the normal direction and the Euclidean Poincaré inequality bound the difference of the means on F and on the box by $C\,\delta^{1-n/p}$ times the norm of the gradient in L^p. Then compare the means on the consecutive dyadic cubical annuli from the scale δ to the scale h, by the Poincaré inequality on the union of two consecutive annuli. These annuli, and the half-annuli in the case of a face, satisfy a uniform Poincaré inequality after scaling, because they are covered by a chain of bounded length of overlapping rectangular boxes, and the resulting integrals of the energy have bounded overlap. For $p>1$, Hölder's inequality gives the coefficient

$$\Big(\sum_{\delta\lesssim 2^k\delta\lesssim h} (2^k\delta)^{(p-n)/(p-1)}\Big)^{p-1} \asymp \kappa_p(h,\delta)\,,$$

and for $p=1$ the largest coefficient is δ^{1-n}. The same orders occur in the formula for the capacity of Euclidean condensers [36, Proposition 10.8]. The argument with boxes also gives an estimate for traces on flat sections,

$$\int_F |u-\langle u\rangle_D|^p\,d\mathcal{H}^{n-1} \le C\,h^{p-1}\int_D |\nabla u|^p\,dx\,, \tag{15.12}$$

whenever F lies in a flat section of D, including a face: first take the mean in the normal coordinate, and then apply the Poincaré inequality on D.

The Poincaré inequality on cells. Put

$$\rho_{p,j} = h_j^Q\,\kappa_p(h_j,\delta_j)\,.$$

Let $u \in \mathcal{C}_p$, and let D' be a child of $D \in \mathcal{D}_{j-1}$. If D' is an ordinary child, compare its mean with that of D by the Euclidean Poincaré inequality. If D' is a new child, compare the two means with the mean on the interface, using (15.11) on each side. Since $\kappa_p(h,\delta) \geq h^{p-n}$, in both cases

$$m(U_{D'})\,|\langle u\rangle_{D'} - \langle u\rangle_D|^p \leq C\,\rho_{p,j}\left(\int_D |\nabla u|^p\,d\nu + \int_{D'} |\nabla u|^p\,d\nu\right).$$

Suppose that $\sum_{i>j} \rho_{p,i}^{1/p} \leq C\,\rho_{p,j}^{1/p}$ for all j. Then, for $D \in \mathcal{D}_j$,

$$\int_{U_D} |u - \langle u\rangle_D|^p\,dm \leq C\,\rho_{p,j} \int_{U_D} |\nabla u|^p\,d\nu\,. \tag{15.13}$$

Indeed, let $A_i u$ be the function equal to $\langle u\rangle_{D'}$ on $U_{D'}$, for each $D' \in \mathcal{D}_i$. By the preceding inequality, in which each integral of the energy appears a bounded number of times,

$$\|A_i u - A_{i-1}u\|_{L^p(U_D,m)} \leq C\,\rho_{p,i}^{1/p}\,\|\nabla u\|_{L^p(U_D,\nu)} \qquad (i > j)\,.$$

Sum over $i > j$, and use the uniform convergence of $A_i u$ to the Lipschitz function u. This telescoping argument is related to the trace estimates on weighted trees of [113, Theorem 3.2].

Admissibility.

Theorem 15.8. (i) *For $1 \leq p < \infty$, the structure is p-admissible if and only if*

$$p \geq p_* := \frac{n\sigma}{n-2+\sigma}\,. \tag{15.14}$$

(ii) *At each admissible exponent, the restriction to any piece P extends to a continuous map $W^{1,p}(X,m) \to W^{1,p}(P)$, and the Euclidean gradient of the restriction of f is $\nabla_p f|_P$.*

(iii) *For admissible $p > 1$, the algebra $\mathcal{A}$ is dense in $W^{1,p}(X,m)$ for the graph norm.*

For $n = 2$ the threshold is $p_* = 2$, and for $n > 2$ one has $p_* < 2$; in all cases $p_* < n$ or $p_* = n = 2$, and the exponent two is admissible. For $n > 2$ the threshold decreases to zero with σ: the larger the windows, the larger the range of admissibility. The mechanism is the same as in Chapter 14. If the windows have a positive capacity at the scale of the cubes, in the sense that $h_j^n\,\kappa_p(h_j,\delta_j)$ remains bounded, the values of a function on the pieces determine its values m-almost everywhere, and the gradient is closable. Otherwise a function can be modified near all the windows at a small cost in energy, so that the pieces carry a nonzero gradient while the function tends to zero in $L^p(m)$.

Proof of (i) for $p \geq p_*$, and of (ii). Since $\mathcal{A} \subset \mathcal{C}_p$ is uniformly dense in $C(X)$, (A1) holds for every p.

Step 1: recovery of the values on a piece. Let P be a piece, and let j be at least its generation. For $f \in L^p(m)$, define $R_j^P f$ on each mesh cube $D \subset P$ of the level j as the constant $\fint_{U_D} f\,dm$. For fixed j, this is a bounded map $L^p(m) \to L^p(P)$. The Euclidean Poincaré inequality and (15.13) imply, for $u \in \mathcal{C}_p$,

$$\|u|_P - R_j^P u\|_{L^p(P)}^p \leq C(h_j^p + h_j^{n-Q}\rho_{p,j})\,\|\nabla u\|_{L^p(\nu)}^p\,. \tag{15.15}$$

Step 2: the coefficient. The coefficient $h_j^{n-Q}\rho_{p,j} = h_j^n\,\kappa_p(h_j,\delta_j)$ is bounded in j if and only if $p \geq p_*$. Indeed, by (15.6):

- for $n > 2$ and $p < n$ it is comparable to $h_j^{p-\sigma(n-p)/(n-2)}$, and the exponent is nonnegative exactly when $p \geq p_*$;
- for $n = 2$ and $p < 2$ it is comparable to $h_j^p \exp((2-p)h_j^{-\sigma})$, which is unbounded;
- for $p = n$ it is comparable to $h_j^{2-\sigma}$ when $n = 2$, and to $h_j^n[\log(e/h_j)]^{n-1}$ when $n > 2$;
- for $p > n$ it is equal to h_j^p.

In all the cases where the coefficient is bounded, $\rho_{p,j} \leq C\, h_j^{Q-n}$, and $\rho_{p,j}$ has the geometric decay, up to a power of j, which was assumed for (15.13).

Step 3: closability and restriction. Fix $p \geq p_*$. The inequality (15.15) at one fixed j shows that the restriction to P is continuous from $\mathcal{C}_p$, with the graph norm, to $W^{1,p}(P)$. Let $u_k \in \mathcal{C}_p$ with $u_k \to 0$ in $L^p(m)$ and $\nabla u_k \to G$ in $L^p(\nu)$, and let v_P be the limit of $u_k|_P$ in $W^{1,p}(P)$, so that $\nabla v_P = G|_P$. Apply (15.15) to $u_k - u_\ell$ and let $k \to \infty$. This gives

$$\|v_P - u_\ell|_P + R_j^P u_\ell\|_{L^p(P)} \leq C\, \|G - \nabla u_\ell\|_{L^p(\nu)}\,,$$

with a constant independent of j. For fixed ℓ, $R_j^P u_\ell \to u_\ell|_P$ uniformly as $j \to \infty$, because the diameters of the cells tend to zero. Letting $j \to \infty$ and then $\ell \to \infty$ gives $v_P = 0$. Thus $G|_P = 0$ on every piece, which proves (A2). The continuity of the restriction on $\mathcal{C}_p$ extends to $W^{1,p}(X,m)$, which gives (ii). □

The failure of admissibility below p_* is Lemma 15.9. The density (iii) is proved in Section 15.7, after the Poincaré inequalities.

Lemma 15.9. *If $1 \leq p < p_*$, the gradient is not closable on $\mathcal{A}$, hence not on $\mathcal{C}_p$.*

Proof. Step 1: the windows are small. The formulas (15.6) give

$$\sum_i \bigl(h_i^{-n}\delta_i^{n-p}\bigr)^{1/p} < \infty\,, \qquad \delta_j^{-p}\, S_j \longrightarrow 0\,.$$

For $n > 2$ both statements follow from $(n-p)\sigma/(n-2) > p$, which is equivalent to $p < p_*$. For $n = 2$ the terms of the series decay exponentially in $h_i^{-\sigma}$, and $S_j \leq C\exp(-2h_{j+1}^{-\sigma})$ gives the second statement, since $p < 2$.

Step 2: a function on the root which vanishes on the windows. Let J be large. Choose a smooth function $v \neq 0$ supported in a corner cube of P_0 of side $h_J/4$, equal to one on the corner cube of side $h_J/8$. This corner meets no interface of generation at most J. For $i > J$, let χ_i be the sum of disjoint cutoffs equal to one on the interfaces of generation i in P_0, supported within the distance $C\delta_i$ of their centres, with gradients bounded by C/δ_i. Then

$$\|\nabla\chi_i\|_{L^p(P_0)}^p \leq C\, h_i^{-n}\delta_i^{n-p}\,,$$

and the total volume of the supports of the χ_i, $i > J$, in the corner of side $h_J/4$ is at most $C\, h_J^n\, S_J$. For large J this is smaller than half the volume of the region where $v = 1$. Hence

$$w_j = v \prod_{J<i\leq j} (1 - \chi_i)$$

converges in $W^{1,p}(P_0)$ to a function w which is not constant: for the gradients, use the product rule, dominated convergence for each term, and the summability of the norms $\|\nabla\chi_i\|_p$.

Step 3: extension. Define u_j on X_j as w_j on P_0 and zero on every other piece, and extend it to X by composition with $\widetilde{\pi}_j$. It is Lipschitz, because w_j vanishes on the interfaces of P_0 which are present in X_j, and $u_j \in \mathcal{A}$. The support of u_j lies in the cells of the h_j^{-n} mesh cubes of the root, of total m-measure at most $C\, h_j^{Q-n}$, so $u_j \to 0$ in $L^p(m)$. On the pieces other than the root, (15.10) bounds its p-energy by

$$C\, S_j\ \mathrm{Lip}(w_j)^p \le C\, S_j\, (\mathrm{Lip}(v) + \delta_j^{-1})^p \longrightarrow 0\,.$$

Thus ∇u_j converges in $L^p(\nu)$ to the field equal to ∇w on the root and to zero elsewhere, which is not zero. □

A test function at an interface. Let P be a piece of X_j, let F be an interface in P, with centre z_F, and let $0 < R \le h_j/8$. Let χ_R be a smooth radial cutoff on P, equal to one at the distances at most R from z_F and to zero at the distances at least $2R$, with $\mathrm{Lip}(\chi_R) \le C/R$. Extend it, by composition with $\widetilde{\pi}_j$, to the pieces of generation greater than j which are attached to P, directly or through a chain of attachments. Its energy on P and on these pieces satisfies

$$\int |\nabla \chi_R|^p \, d\nu \le C\, R^{n-p}\,. \tag{15.16}$$

To see this, choose $J > j$ with $2R < h_J \le 4R$. An interface in P of generation g, $j < g \le J$, is at a normal distance at least h_J from the hyperplane through z_F, by the parity of the dyadic coordinates, so χ_R vanishes on it, and the extension is zero on the pieces attached there. On P the estimate is Euclidean. For the remaining pieces, apply (15.10) to the bounded number of cubes of the level J in P which meet $B(z_F, 2R)$. For a sum $\sum_R a_R \chi_R$ over radii which are successively doubled, the gradients on the piece are supported in disjoint annuli, and the regions where the extensions vary are disjoint as well. One obtains an energy at most $C \sum_R |a_R|^p R^{n-p}$.

The Poincaré scale.

Proposition 15.10. *At an admissible exponent p, the Poincaré inequality $\mathrm{PI}_p(\Psi_{P,p})$ holds at small radii with*

$$\Psi_{P,p}(r) \asymp \begin{cases} r^{Q-\sigma}\,, & n = p = 2\,, \\ r^{Q+p-n-\sigma(n-p)/(n-2)}\,, & n > 2\,,\ p < n\,, \\ r^{Q}\,[\log(e/r)]^{n-1}\,, & n > 2\,,\ p = n\,, \\ r^{Q+p-n}\,, & p > n\,, \end{cases} \tag{15.17}$$

and this scale is sharp: if $\mathrm{PI}_p(\Psi)$ holds at small radii, then $\Psi(r) \ge c\, \Psi_{P,p}(r)$ for all sufficiently small r. Consequently the Poincaré scale and the cutoff scale $\Psi_{C,p}(r) = r^{Q+p-n}$ match if and only if $p > n$.

In all cases $\Psi_{P,p}(r) \asymp r^Q\, \kappa_p(r, \delta(r))$, where $\delta(r)$ is the side of the windows of the levels j with $h_j \asymp r$. For $p > n$ the windows play no role, because points have positive p-capacity. For admissible $p \le n$, the ratio $\Psi_{P,p}(r)/\Psi_{C,p}(r) \asymp r^{n-p}\, \kappa_p(r, \delta(r))$ tends to infinity as $r \to 0$. At the exponent two,

$$\Psi_{P,2}(r) \asymp r^{-\sigma}\, \Psi_{C,2}(r) \qquad \text{for every } n \ge 2\,.$$

Proof. Step 1: the upper bound. Choose j with $h_j \asymp r$. A bounded number of cells of the level j meet $B(x, r)$, and their mesh cubes lie in a ball of radius $C\, r$ about $\pi_j x$ in X_j. The points of the

projected ball are joined to its centre by geodesics, which yields chains of a bounded number of intersecting mesh cubes in this ball. We compare the means of u on two intersecting cubes of a chain.

- If the two cubes lie in the same piece, the Euclidean Poincaré inequality in a box which contains them bounds the pth power of the difference of their means by $C\,h_j^{p-n}$ times the energy in the box.
- If they meet across an interface of generation $i \le j$, choose a flat subcube of this interface which contains the point of intersection, of side $s = \min(\delta_i, c_n h_j/2)$ with $c_n = (64\sqrt{n})^{-1}$. On each side take a box of side comparable to h_j, centred tangentially at this subcube, with the subcube in a face or in the central section. These boxes fit in the pieces because the interface is centred. The mean on each box is compared with the mean on the nearby mesh cube by the Euclidean Poincaré inequality, and the means on the two boxes are compared by (15.11). Since $s \ge c\,\delta_j$, the resulting coefficient is at most $C\,\kappa_p(h_j, \delta_j)$.
- For two pieces which share an incoming interface, use the same comparison through their host.

All the boxes lie in a fixed dilation of $B(x, r)$. Apply (15.13) in each cell, and the comparison of the means along the chains, with $m(U_D) \asymp h_j^Q$. This gives the Poincaré inequality on the ball at the scale $\rho_{p,j}$, and (15.6) gives (15.17).

Step 2: test functions. Let P be a new piece of generation i, and let z_F be the centre of its incoming interface. On X_i, let u_P be the function equal to zero outside P and to $1 - \chi_R$ on P, with $\sqrt{n}\,\delta_i \le R \le h_i/8$, and extend it by composition with $\widetilde{\pi}_i$. For $p < n$ take $R = \sqrt{n}\,\delta_i$, and for $p > n$ take $R = h_i/8$. For $p = n$, take the average of these functions over the dyadic radii in this interval, and use (15.16) and the disjointness of the annuli. In the three cases

$$0 \le u_P \le 1\,, \qquad \mathcal{E}_p(u_P) \le C\,\kappa_p(h_i, \delta_i)^{-1}\,, \qquad \operatorname{supp} u_P \subset U_P\,. \tag{15.18}$$

Step 3: sharpness. The function u_P is equal to one on a cell in the half of P which is far from the interface, and to zero on the cell of an ordinary child of the mesh cube to which P is attached. Both cells have a measure $\asymp h_i^Q$ and lie in $B(z_F, C\,h_i)$. Hence the left-hand side of the Poincaré inequality for u_P on this ball is at least $c\,h_i^Q$, while the right-hand side is at most the scale times $\mathcal{E}_p(u_P)$. The same holds on every ball $B(z_F, r)$ with $r \ge C\,h_i$. By (15.18), if $\mathrm{PI}_p(\Psi)$ holds, then $\Psi(r) \ge c\,h_i^Q\,\kappa_p(h_i, \delta_i)$ for these radii. The generations at which new pieces are attached have bounded gaps, so for every sufficiently small radius r there is such an i with $C\,h_i < r \asymp h_i$. This proves the sharpness, for any fixed dilation of the balls in the Poincaré inequality. □

15.7 The tree of cubes: density of the algebra $\mathcal{A}$

We prove Theorem 15.8(iii). Let $p > 1$ be admissible, fix $u \in \mathcal{C}_p$, and write $h = h_j$. We construct $v_j \in \mathrm{Lip}(X_j)$ such that

$$\mathcal{E}_p(v_j \circ \widetilde{\pi}_j) \le C\,\mathcal{E}_p(u)\,, \qquad \sum_{P \subset X_j} \int_P |v_j - u|^p\,dx \le C\,h^p\,\mathcal{E}_p(u)\,. \tag{15.19}$$

The difficulty is that $u \circ \widetilde{\pi}_j$ itself does not approximate u in energy: on the later pieces it depends only on the values of u near the windows. The function v_j is obtained from u by a regularization at the scale h on the pieces of X_j, which makes its Lipschitz constant on each mesh cube comparable to the average of $|\nabla u|$ nearby, so that (15.10) controls the energy of the extension. All the

modifications below are made on the Euclidean pieces. Distinct interfaces in a piece of X_j are h-separated; fix $\eta > 0$, depending only on n, so small that their $2\eta h$-neighbourhoods are disjoint.

Step 1: mollification on the pieces. Reflect $u|_P$ across the faces of P and mollify at the scale $\eta h/10$, as in [42, Lemma 9.2 and Section 9.2]. For a mesh cube $D \subset P$, let $\widehat{D}$ be its h-neighbourhood in P for the maximum norm. The mollified function, still denoted by v_j, satisfies

$$\mathrm{Lip}(v_j|_D) \le C\, h^{-n/p}\, \|\nabla u\|_{L^p(\widehat{D})}\,, \qquad \|v_j - u\|_{L^p(D)} \le C\, h\, \|\nabla u\|_{L^p(\widehat{D})}\,. \tag{15.20}$$

Indeed, subtracting the mean of u on $\widehat{D}$, the oscillation of the mollified function from this mean is at most $C\, h^{1-n/p}\|\nabla u\|_{L^p(\widehat{D})}$, by the Euclidean Poincaré inequality, and the error in L^p is given by the estimate for translations in the convolution. The mollified functions of two pieces do not agree on their interface; the next two steps correct this.

Step 2: wide interfaces. On each interface of side at least $\eta h/(400\sqrt{n})$, mollify the common restriction of u tangentially at the scale $\eta h/(1600\sqrt{n})$, after reflection at the edge, and compose the resulting function with the projection onto the interface. Blend it with the mollifications of the incident pieces, using a Lipschitz cutoff equal to one within the distance ηh of the interface and to zero beyond $2\eta h$. For each mesh cube D which is affected, the estimate (15.12) on $\widehat{D}$ bounds in L^p the trace of u minus the mean of u on $\widehat{D}$. The convolution in dimension $n-1$ then bounds the oscillation from this mean by $C\, h^{1-n/p}\|\nabla u\|_{L^p(\widehat{D})}$, and the tangential gradient by $C\, h^{-n/p}\|\nabla u\|_{L^p(\widehat{D})}$. Thus (15.20) still holds, and the corrected functions agree on these interfaces.

Step 3: narrow interfaces. Let F be an interface of smaller side δ, with centre z_F. On each incident piece P, first flatten the mollification to the mean a_P of u on the box $P \cap (z_F + (-2\eta h, 2\eta h)^n)$, using a cutoff equal to one at the distances at most $\eta h/2$ from z_F and to zero at the distances at least ηh. The same estimate for convolutions on this box proves (15.20) for the flattened function. Let t be the mean of u on the interface, which is common to the incident pieces. Add to the flattened function the correction $(t - a_P)\,\chi$, where χ is the weighted average of the functions χ_R of (15.16), for the dyadic radii from $\sqrt{n}\,\delta$ to a radius comparable to $\eta h/8$, with $R \le \eta h/8$, and with weights proportional to $R^{(p-n)/(p-1)}$. The support of χ lies in the flattened region, and $\chi = 1$ on the interface. By (15.16) and the disjointness of the annuli, the energy of χ on the piece and on the later pieces attached to it is at most $C\,\kappa_p(h,\delta)^{-1}$. By (15.11) on the box of the scale ηh, $|t - a_P|^p$ is at most $C\,\kappa_p(h,\delta)$ times the energy of u on that box. Hence the energy of the correction, including its extension, is at most a constant times the energy of u on the box. Its error in $L^p(P)$ is at most $C\,h$ times the norm of its own gradient in L^p, by the Euclidean Poincaré inequality, since it vanishes on a fixed proportion of the box.

Step 4: the bounds. The modified functions agree on all the interfaces of X_j, so they define $v_j \in \mathrm{Lip}(X_j)$. The boxes $\widehat{D}$ have bounded overlap, and so have the boxes at the interfaces. By (15.10), the energy on U_D of the part which satisfies (15.20) is at most $C\,(1 + S_j)\int_{\widehat{D}} |\nabla u|^p\, dx$. The corrections at the narrow interfaces are supported in disjoint regions of a piece, and their extensions have disjoint supports on the later pieces attached to that piece. Here each part of the decomposition is composed with the retraction only on the later pieces attached to its own piece. Summing gives (15.19).

Step 5: weak limit. The inequality (15.13) on the root cell, and the bound for the mean on the root given by the second estimate of (15.19), bound $v_j \circ \widetilde{\pi}_j$ in the graph norm. Since $p > 1$, the space $W^{1,p}(X, m)$ is reflexive, and a subsequence converges weakly to some w. By the continuity of the restriction (Theorem 15.8(ii)) and the second bound of (15.19), the restriction of w to every piece P is $u|_P$. Therefore $w - u$ has zero gradient and zero mean on the root, and it vanishes, by (15.13) on the root cell, extended to $W^{1,p}(X, m)$ by convergence in the graph norm. The closure of the linear space $\mathcal{A}$ for the graph norm is weakly closed. It contains u, hence $\mathcal{C}_p$, and hence $W^{1,p}(X, m)$. □

15.8 The tree of cubes: critical Korevaar–Schoen spaces

We compute the critical exponents of Definition 4.9 for every $p \geq 1$, including the exponents at which the structure is not admissible, and we compare the critical space with the Sobolev space.

Theorem 15.11. (i) *For every $1 \leq p < \infty$,*

$$\alpha_p^{\mathrm{dens}} = \alpha_p = 1 + \frac{Q - n}{p}\,, \tag{15.21}$$

and both thresholds are attained.

(ii) *If $p > n$, the critical space $KS^{1+(Q-n)/p,p}(X)$ is equal to $W^{1,p}(X, m)$, and the weak monotonicity holds:*

$$\sup_{0<r\leq r_1} \frac{D_p(f, r)}{r^{Q+p-n}} \leq C \liminf_{r\downarrow 0} \frac{D_p(f, r)}{r^{Q+p-n}} \qquad (f \in L^p(X, m))\,. \tag{15.22}$$

(iii) *For $1 \leq p \leq n$, the estimate (15.22) fails, even for $f \in \mathcal{A}$.*

(iv) *At every admissible p,*

$$\sup_{0<r\leq r_1} \frac{D_p(f, r)}{\Psi_{P,p}(r)} \leq C \liminf_{r\downarrow 0} \frac{D_p(f, r)}{r^{Q+p-n}} \qquad (f \in L^p(X, m))\,, \tag{15.23}$$

with $\Psi_{P,p}$ given by (15.17).

(v) *For admissible $1 < p \leq n$, the critical space is a proper subspace of $W^{1,p}(X, m)$, dense for the graph norm.*

The critical exponent is the one given by the cutoff scale, $r^{p\alpha_p} = \Psi_{C,p}(r)$, for every p, as in Theorem 4.12, although the scales match only for $p > n$. The weak monotonicity (15.22) is the comparison of the supremum and of the lower limit of the Korevaar–Schoen energies at the same exponent, as in [31, Section 4.2]. In (15.23) the two sides are normalized by different scales, and their ratio measures the difference between the norm of the critical space and the graph norm of the closed gradient.

Proof. Step 1: the algebra $\mathcal{A}$ lies in the critical space. Let $u \in \mathcal{A}$ factor through X_J, and suppose that it is constant on all but M mesh cubes of the level J. By Lemma 15.6(ii), the number of mesh cubes of the level $j \geq J$ on which u is not constant is at most $C\,M\,(h_J/h_j)^n$. Let $r \asymp h_j$. If $d(x, y) < r$ and $u(x) \neq u(y)$, the geodesic between $\widetilde{\pi}_j x$ and $\widetilde{\pi}_j y$ in X_j meets a cube on which u is not constant: on a connected chain of intersecting cubes on which u is constant, u takes a single value. Thus $\widetilde{\pi}_j x$ lies within the distance r of one of these cubes, and x lies in one of a

bounded number of cells near such a cube. The set of these points x has an m-measure at most $C\,M\,(h_J/h_j)^n\,h_j^Q$. Consequently

$$D_p(u,r) \le C\,M\,h_J^n\,\mathrm{Lip}(u)^p\,r^{Q+p-n} \qquad (r \text{ sufficiently small})\,. \tag{15.24}$$

Hence $\mathcal{A} \subset KS^{1+(Q-n)/p,p}(X)$. Since $\mathcal{A}$ is uniformly dense in $C(X)$, it is dense in $L^p(X,m)$, and it contains nonconstant functions. Therefore $\alpha_p^{\mathrm{dens}} \ge 1+(Q-n)/p$, $\alpha_p \ge 1+(Q-n)/p$, and both thresholds are attained if equality holds.

Step 2: averages on the cells. Suppose that $D_p(f,r) = o(r^{Q+p-n})$; we show that f is constant. Put $a_D = \fint_{U_D} f\,dm$ for $D \in \mathcal{D}_j$, and let A_j be the function equal to a_D on U_D. The functions A_j converge to f in $L^p(m)$, because the diameters of the cells tend to zero and continuous functions are dense. On a piece P which is present at the level j, let $R_j^P f$ be equal to a_D on each mesh cube $D \subset P$ of the level j. By Jensen's inequality, for a child D' of D, and also for two intersecting mesh cubes D, D' of the same level,

$$h_j^Q\,|a_D - a_{D'}|^p \le C \int_{U_D} \fint_{B(x,C_0h_j)} |f(x)-f(y)|^p\,dm(y)\,dm(x)\,,$$

where C_0 is a sufficiently large constant; each mesh cube has a bounded number of such neighbours. Define

$$E_j = \frac{D_p(f, C_0h_j)}{h_j^{Q+p-n}}\,.$$

Summing over the ordinary subdivisions in the pieces of X_J gives

$$\Big(\sum_{P\subset X_J} \|R_{j+1}^P f - R_j^P f\|_{L^p(P)}^p\Big)^{1/p} \le C\,h_j\,E_j^{1/p} \qquad (j \ge J)\,; \tag{15.25}$$

the Lebesgue measure of a mesh cube replaces the factor h_j^Q by h_j^n. Since (E_j) is bounded, $R_j^P f$ converges in $L^p(P)$ to a function v_P, for each piece.

Step 3: the limit is constant on each piece. On each piece, interpolate the values a_D, placed at the centres of the mesh cubes, by multilinear interpolation, the values at the centres being extended as constants up to the boundary in each coordinate. Let I_j^P be the interpolating function. On each box of the interpolation, its oscillation is at most a fixed sum of differences of neighbouring values, and its gradient is at most this sum divided by h_j. With the estimate of Step 2 for intersecting mesh cubes,

$$\|I_j^P - R_j^P f\|_{L^p(P)} \le C\,h_j\,E_j^{1/p}\,, \qquad \|\nabla I_j^P\|_{L^p(P)} \le C\,E_j^{1/p}\,. \tag{15.26}$$

Since $E_j \to 0$, the functions I_j^P converge to v_P in $W^{1,p}(P)$ and $\nabla v_P = 0$, so v_P is constant.

Step 4: the constants agree. Let F be an interface between two pieces P and P^-. The values a_D on each side define a function on F, by choosing, away from the edges of the mesh, the mesh cube of that side which contains the point. The two cubes chosen at a point of F intersect, and each mesh cube meets F in a set of $(n-1)$-dimensional measure at most $C\,h_j^{n-1}$. Therefore the pth power of the norm in $L^p(F)$ of the difference of the two functions is at most $C\,h_j^{p-1}E_j$. The same bound compares each of them with the corresponding interpolating function on F. Thus the difference of the traces of I_j^P and $I_j^{P^-}$ tends to zero in $L^p(F)$, including at $p=1$. By (15.12), the traces converge to the traces of the limits, so the two constants agree. All the pieces are joined by finite chains of interfaces, which gives a single constant c.

Step 5: conclusion for the exponents. Summing (15.25) over $j \geq J$, and comparing each mass $m(U_D)$ with the volume of D at the level J, gives

$$\|A_J - c\|_{L^p(m)} \leq C\, h_J^{(Q-n)/p} \sum_{j \geq J} h_j\, E_j^{1/p} \longrightarrow 0\,.$$

Since $A_J \to f$ in $L^p(m)$, $f = c$ almost everywhere. A function of $KS^{\alpha,p}(X)$ with $\alpha > 1 + (Q-n)/p$ satisfies $D_p(f,r) = o(r^{Q+p-n})$. With Step 1, this proves (i).

Step 6: comparison at the critical exponent. Theorem 4.4 and Proposition 15.10 give (iv) at every admissible p; for $p > n$ the scales match, which gives (ii).

Step 7: failure of the weak monotonicity for $p \leq n$. Let u_P be the function of (15.18), for a new piece P of generation i. The two cells on which it takes the values zero and one give

$$\frac{D_p(u_P, C\, h_i)}{(C\, h_i)^{Q+p-n}} \geq c\, h_i^{n-p}\,. \tag{15.27}$$

For the small scales, consider first a single radial cutoff of radius R. At the level J with $2R < h_J \leq 4R$, the function $1 - \chi_R$ is constant on all the pieces of generation at most J other than P, because their interfaces miss $B(z_F, 2R)$, and on P only a bounded number of cubes of the level J meet the annulus where it varies. The function factors through $\widetilde{\pi}_J$ and is C/R-Lipschitz. By (15.24), with M bounded and $h_J \asymp R$,

$$\limsup_{r \downarrow 0} \frac{D_p(1-\chi_R, r)}{r^{Q+p-n}} \leq C\, R^{n-p}\,.$$

For $p < n$, u_P is such a function with $R = \sqrt{n}\,\delta_i$. For $p = n$, take in the definition of u_P the average over $M \asymp 1 + \log(h_i/\delta_i)$ radii which increase by the factor four, the smallest one being strictly larger than the circumradius of the interface. For a pair of points at a sufficiently small distance, at most one of the M functions has an increment: two points of the branch have radial coordinates, under $\widetilde{\pi}_i$, which differ by at most their distance, while the annuli of transition are separated; and if one of the points is outside the branch, the projection of the point inside the branch is close to the interface, where all the functions vanish. Thus $D_p(u_P, r)$, at small radii, is at most the sum of the corresponding quantities for the M cutoffs, multiplied by M^{-p}. In both cases

$$\limsup_{r \downarrow 0} \frac{D_p(u_P, r)}{r^{Q+p-n}} \leq C\, \kappa_p(h_i, \delta_i)^{-1}\,. \tag{15.28}$$

The quotient of (15.27) by this bound is at least $c\, h_i^{n-p}\, \kappa_p(h_i, \delta_i)$, which tends to infinity with i, since $\delta_i / h_i \to 0$; at $p = n$ this is due to the logarithmic factor. This proves (iii).

Step 8: the critical space is a proper subspace. Let $1 < p \leq n$ be admissible. The critical space embeds continuously in $W^{1,p}(X, m)$, by Theorem 4.4, and it contains $\mathcal{A}$, which is dense for the graph norm by Theorem 15.8(iii). Both spaces are Banach spaces for their norms; for the critical space this follows from Fatou's lemma. If the two spaces were equal, the bounded inverse theorem would bound the norm of the critical space by the graph norm. The functions u_P contradict this, by (15.27), (15.18) and $\|u_P\|_{L^p(m)}^p \leq C\, h_i^Q$. This proves (v). □

References

[1] P. Alonso Ruiz, F. Baudoin, L. Chen, L. Rogers, N. Shanmugalingam and A. Teplyaev, Besov class via heat semigroup on Dirichlet spaces I: Sobolev type inequalities, *J. Funct. Anal.* **278** (2020), 108459.

[2] P. Alonso Ruiz, F. Baudoin, L. Chen, L. Rogers, N. Shanmugalingam and A. Teplyaev, Besov class via heat semigroup on Dirichlet spaces II: BV functions and Gaussian heat kernel estimates, *Calc. Var. Partial Differential Equations* **59** (2020), Paper No. 103.

[3] P. Alonso Ruiz, F. Baudoin, L. Chen, L. Rogers, N. Shanmugalingam and A. Teplyaev, Besov class via heat semigroup on Dirichlet spaces III: BV functions and sub-Gaussian heat kernel estimates, *Calc. Var. Partial Differential Equations* **60** (2021), Paper No. 170.

[4] P. Alonso Ruiz and F. Baudoin, Gagliardo–Nirenberg, Trudinger–Moser and Morrey inequalities on Dirichlet spaces, *J. Math. Anal. Appl.* **497** (2021), no. 2, Paper No. 124899. doi:10.1016/j.jmaa.2020.124899.

[5] P. Alonso Ruiz and F. Baudoin, Korevaar–Schoen p-energies and their Γ-limits on Cheeger spaces, *Nonlinear Anal.* **256** (2025), 113779. doi:10.1016/j.na.2025.113779.

[6] L. Ambrosio and S. Di Marino, Equivalent definitions of BV space and of total variation on metric measure spaces, *J. Funct. Anal.* **266** (2014), 4150–4188.

[7] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, Clarendon Press, Oxford, 2000.

[8] L. Ambrosio, N. Gigli and G. Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces, *Rev. Mat. Iberoam.* **29** (2013), no. 3, 969–996. doi:10.4171/RMI/746.

[9] L. Ambrosio, N. Gigli and G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, *Invent. Math.* **195** (2014), 289–391.

[10] S. Andres and M. T. Barlow, Energy inequalities for cutoff functions and some applications, *J. Reine Angew. Math.* **699** (2015), 183–215.

[11] S. Andres and N. Kajino, Continuity and estimates of the Liouville heat kernel with applications to spectral dimensions, *Probab. Theory Related Fields* **166** (2016), 713–752, doi:10.1007/s00440-015-0670-4.

[12] S. Andres, N. Kajino, K. Kavvadias and J. Miller, Two-sided heat kernel bounds for $\sqrt{8/3}$-Liouville Brownian motion, preprint, arXiv:2507.13269v2, 15 April 2026.

[13] R. Anttila, S. Eriksson-Bique and R. Shimizu, Construction of self-similar energy forms and singularity of Sobolev spaces on Laakso-type fractal spaces, preprint, arXiv:2503.13258v2, 8 September 2026.

[14] R. Anttila, An approach to sub-Gaussian heat kernel estimates via analysis on metric spaces, preprint (2025), arXiv:2509.04155v2.

[15] D. Azagra, E. Le Gruyer and C. Mudarra, Kirszbraun's theorem via an explicit formula, *Canad. Math. Bull.* **64** (2021), 142–153.

[16] D. Bakry, T. Coulhon, M. Ledoux and L. Saloff-Coste, Sobolev inequalities in disguise, *Indiana Univ. Math. J.* **44** (1995), no. 4, 1033–1074. Author's text.

[17] M. T. Barlow, R. F. Bass and T. Kumagai, Stability of parabolic Harnack inequalities on metric measure spaces, *J. Math. Soc. Japan* **58** (2006), 485–519; see also the authors' 2020 correction to Section 3.

[18] M. T. Barlow, A. Grigor'yan and T. Kumagai, On the equivalence of parabolic Harnack inequalities and heat kernel estimates, *J. Math. Soc. Japan* **64** (2012), no. 4, 1091–1146. Author's text.

[19] D. Bate, Structure of measures in Lipschitz differentiability spaces, *J. Amer. Math. Soc.* **28** (2015), 421–482.

[20] D. Bate, Purely unrectifiable metric spaces and perturbations of Lipschitz functions, *Acta Math.* **224** (2020), 1–65.

[21] F. Baudoin and L. Chen, Sobolev spaces and Poincaré inequalities on the Vicsek fractal, *Ann. Fenn. Math.* **48** (2023), no. 1, 3–26. doi:10.54330/AFM.122168.

[22] F. Baudoin, L. Chen and M. Yang, Korevaar–Schoen and heat kernel characterizations of Sobolev and BV spaces on local trees, preprint, arXiv:2505.10177.

[23] F. Baudoin and N. Garofalo, Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries, *J. Eur. Math. Soc.* **19** (2017), no. 1, 151–219. doi:10.4171/JEMS/663. Locators refer to arXiv:1101.3590v5.

[24] F. Baudoin, M. Bonnefont and N. Garofalo, A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality, *Math. Ann.* **358** (2014), no. 3–4, 833–860. doi:10.1007/s00208-013-0961-y. Locators refer to arXiv:1007.1600v3.

[25] F. Baudoin, M. Bonnefont, N. Garofalo and I. H. Munive, Volume and distance comparison theorems for sub-Riemannian manifolds, *J. Funct. Anal.* **267** (2014), no. 7, 2005–2027. doi:10.1016/j.jfa.2014.07.030. Locators refer to arXiv:1211.0221v3.

[26] F. Baudoin and B. Kim, Sobolev, Poincaré, and isoperimetric inequalities for subelliptic diffusion operators satisfying a generalized curvature dimension inequality, *Rev. Mat. Iberoam.* **30** (2014), no. 1, 109–131. doi:10.4171/RMI/771.

[27] F. Baudoin, B. Kim and J. Wang, Transverse Weitzenböck formulas and curvature dimension inequalities on Riemannian foliations with totally geodesic leaves, *Commun. Anal. Geom.* **24** (2016), no. 5, 913–937. doi:10.4310/CAG.2016.v24.n5.a1. Locators refer to arXiv:1408.0548v3.

[28] F. Baudoin, E. Grong, K. Kuwada and A. Thalmaier, Sub-Laplacian comparison theorems on totally geodesic Riemannian foliations, *Calc. Var. Partial Differential Equations* **58** (2019), no. 4, Article 130. doi:10.1007/s00526-019-1570-8. Locators refer to the author version of 6 June 2019.

[29] F. Baudoin, E. Grong, L. Rizzi and S. Vega-Molino, Comparison theorems on H-type sub-Riemannian manifolds, *Calc. Var. Partial Differential Equations* **64** (2025), no. 5, Article 143. doi:10.1007/s00526-025-02992-w.

[30] M.-K. von Renesse, On local Poincaré via transportation, *Math. Z.* **259** (2008), no. 1, 21–31. doi:10.1007/s00209-007-0206-4. Section locators refer to the author version of 3 May 2007.

[31] F. Baudoin, Korevaar–Schoen–Sobolev spaces and critical exponents in metric measure spaces, *Ann. Fenn. Math.* **49** (2024), 487–527; arXiv:2207.12191v6. Precise theorem and lemma locators refer to this author version.

[32] O. Ben-Bassat, R. S. Strichartz and A. Teplyaev, What is not in the domain of the Laplacian on Sierpinski gasket type fractals, *J. Funct. Anal.* **166** (1999), 197–217.

[33] N. Berestycki, C. Garban, R. Rhodes and V. Vargas, KPZ formula derived from Liouville heat kernel, *J. Lond. Math. Soc. (2)* **94** (2016), 186–208, doi:10.1112/jlms/jdw031.

[34] N. Berestycki, An elementary approach to Gaussian multiplicative chaos, *Electron. Commun. Probab.* **22** (2017), paper no. 27, 1–12, doi:10.1214/17-ECP58.

[35] A. Björn and J. Björn, *Nonlinear Potential Theory on Metric Spaces*, EMS Tracts in Mathematics 17, European Mathematical Society, Zürich, 2011.

[36] A. Björn, J. Björn and J. Lehrbäck, Sharp capacity estimates for annuli in weighted $\mathbb{R}^n$ and in metric spaces, *Math. Z.* **286** (2017), 1173–1215.

[37] A. Björn, J. Björn and P. Lahti, Removable sets for Newtonian Sobolev spaces and a characterization of p-path almost open sets, *Rev. Mat. Iberoam.* **39** (2023), 1143–1180. doi:10.4171/RMI/1419.

[38] V. I. Bogachev, *Gaussian Measures*, Mathematical Surveys and Monographs 62, American Mathematical Society, Providence, RI, 1998. doi:10.1090/surv/062.

[39] E. Bombieri and E. Giusti, Harnack's inequality for elliptic differential equations on minimal surfaces, *Invent. Math.* **15** (1972), 24–46. doi:10.1007/BF01418640.

[40] J. Bourgain, H. Brezis and P. Mironescu, Another look at Sobolev spaces, in: *Optimal Control and Partial Differential Equations*, IOS Press, Amsterdam, 2001, 439–455.

[41] K. Brewster, D. Mitrea, I. Mitrea and M. Mitrea, Extending Sobolev functions with partially vanishing traces from locally (ε, δ)-domains and applications to mixed boundary problems, *J. Funct. Anal.* **266** (2014), no. 7, 4314–4421. doi:10.1016/j.jfa.2014.02.001.

[42] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer, New York, 2011. doi:10.1007/978-0-387-70914-7.

[43] M. R. Bridson and A. Haefliger, *Metric Spaces of Non-Positive Curvature*, Grundlehren der mathematischen Wissenschaften 319, Springer, Berlin, 1999.

[44] M. Brin and Y. Kifer, Brownian motion, harmonic functions and hyperbolicity for Euclidean complexes, *Math. Z.* **237** (2001), 421–468.

[45] E. Bruè and D. Semola, Constancy of the dimension for RCD(K, N) spaces via regularity of Lagrangian flows, *Comm. Pure Appl. Math.* **73** (2020), 1141–1204, doi:10.1002/cpa.21849.

[46] P. Buser, A note on the isoperimetric constant, *Ann. Sci. École Norm. Sup. (4)* **15** (1982), 213–230.

[47] M. Capolli, A. Pinamonti and G. Speight, Maximal directional derivatives in Laakso space, *Commun. Contemp. Math.* **27** (2025), no. 4, 2450017. doi:10.1142/S0219199724500172.

[48] E. A. Carlen, S. Kusuoka and D. W. Stroock, Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré Probab. Statist.* **23** (1987), suppl. no. 2, 245–287. Full text.

[49] J. Cheeger, Differentiability of Lipschitz functions on metric measure spaces, *Geom. Funct. Anal.* **9** (1999), 428–517.

[50] J. Cheeger and B. Kleiner, Inverse limit spaces satisfying a Poincaré inequality, *Anal. Geom. Metr. Spaces* **3** (2015), 15–39. doi:10.1515/agms-2015-0002.

[51] Z.-Q. Chen, P. Kim, T. Kumagai and J. Wang, Heat kernel upper bounds for symmetric Markov semigroups, *J. Funct. Anal.* **281** (2021), no. 4, 109074, doi:10.1016/j.jfa.2021.109074.

[52] A. Chen and M. Murugan, Bounding martingale dimension by spectral dimension, preprint (2026), arXiv:2610.39312v1.

[53] Z.-Q. Chen and M. Fukushima, *Symmetric Markov Processes, Time Change, and Boundary Theory*, London Mathematical Society Monographs Series 35, Princeton University Press, Princeton, 2012.

[54] F. Cipriani and J.-L. Sauvageot, Derivations as square roots of Dirichlet forms, *J. Funct. Anal.* **201** (2003), 78–120.

[55] T. Coulhon and A. Grigor'yan, On-diagonal lower bounds for heat kernels and Markov chains, *Duke Math. J.* **89** (1997), no. 1, 133–199. doi:10.1215/S0012-7094-97-08908-0.

[56] P. Creutz and E. Soultanis, Maximal metric surfaces and the Sobolev-to-Lipschitz property, *Calc. Var. Partial Differential Equations* **59** (2020), Paper No. 177. doi:10.1007/s00526-020-01843-0.

[57] E. B. Davies, Explicit constants for Gaussian upper bounds on heat kernels, *Amer. J. Math.* **109** (1987), no. 2, 319–333. doi:10.2307/2374577.

[58] M. M. Day, Some more uniformly convex spaces, *Bull. Amer. Math. Soc.* **47** (1941), 504–507.

[59] G. De Philippis and N. Gigli, Non-collapsed spaces with Ricci curvature bounded from below, *J. Éc. polytech. Math.* **5** (2018), 613–650, doi:10.5802/jep.80.

[60] G. De Philippis, A. Marchese and F. Rindler, On a conjecture of Cheeger, in *Measure Theory in Non-Smooth Spaces*, De Gruyter Open, Warsaw, 2017, 145–155.

[61] S. Di Marino, Sobolev and BV spaces on metric measure spaces via derivations and integration by parts, preprint (2014), arXiv:1409.5620.

[62] J. Eells and B. Fuglede, *Harmonic Maps between Riemannian Polyhedra*, Cambridge Tracts in Mathematics 142, Cambridge University Press, Cambridge, 2001.

[63] S. Eriksson-Bique, Alternative proof of Keith–Zhong self-improvement and connectivity, *Ann. Acad. Sci. Fenn. Math.* **44** (2019), 407–425. Published article.

[64] S. Eriksson-Bique, On the resistance conjecture, preprint (2026), arXiv:2602.05477v2.

[65] S. Eriksson-Bique and M. Murugan, On the energy image density conjecture of Bouleau and Hirsch, preprint (2025), arXiv:2510.13659; author's text dated 4 November 2025.

[66] S. Eriksson-Bique and M. Murugan, Martingale and analytic dimensions, preprint (2026), arXiv:2609.38583.

[67] G. B. Folland, Subelliptic estimates and function spaces on nilpotent Lie groups, *Ark. Mat.* **13** (1975), 161–207. doi:10.1007/BF02386204.

[68] B. Franchi, P. Hajłasz and P. Koskela, Definitions of Sobolev classes on metric spaces, *Ann. Inst. Fourier (Grenoble)* **49** (1999), no. 6, 1903–1924. doi:10.5802/aif.1742.

[69] M. Fukushima, Y. Oshima and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes*, second edition, de Gruyter Studies in Mathematics 19, de Gruyter, Berlin, 2011.

[70] C. Garban, R. Rhodes and V. Vargas, On the heat kernel and the Dirichlet form of Liouville Brownian motion, *Electron. J. Probab.* **19** (2014), no. 96, 1–25, doi:10.1214/EJP.v19-2950.

[71] C. Garban, R. Rhodes and V. Vargas, Liouville Brownian motion, *Ann. Probab.* **44** (2016), no. 4, 3076–3110, doi:10.1214/15-AOP1042.

[72] N. Gigli and B.-X. Han, Independence on p of weak upper gradients on RCD spaces, *J. Funct. Anal.* **271** (2016), no. 1, 1–11. doi:10.1016/j.jfa.2016.04.014.

[73] N. Gigli and F. Nobili, A first-order condition for the independence on p of weak gradients, *J. Funct. Anal.* **283** (2022), no. 11, 109686. doi:10.1016/j.jfa.2022.109686.

[74] N. Gigli, Nonsmooth differential geometry—an approach tailored for spaces with Ricci curvature bounded from below, *Mem. Amer. Math. Soc.* **251** (2018), no. 1196.

[75] A. Grigor'yan, J. Hu and K.-S. Lau, Generalized capacity, Harnack inequality and heat kernels of Dirichlet forms on metric measure spaces, *J. Math. Soc. Japan* **67** (2015), 1485–1549.

[76] A. Grigor'yan and L. Liu, Heat kernel and Lipschitz–Besov spaces, *Forum Math.* **27** (2015), 3567–3613. Author's text.

[77] A. A. Grigor'yan, The heat equation on noncompact Riemannian manifolds, *Math. USSR-Sb.* **72** (1992), 47–77. doi:10.1070/SM1992v072n01ABEH001410.

[78] A. Grigor'yan and A. Telcs, Two-sided estimates of heat kernels on metric measure spaces, *Ann. Probab.* **40** (2012), 1212–1284.

[79] P. Hajłasz and P. Koskela, Sobolev met Poincaré, *Mem. Amer. Math. Soc.* **145** (2000), no. 688.

[80] B. Hanson and J. Heinonen, An n-dimensional space that admits a Poincaré inequality but has no manifold points, *Proc. Amer. Math. Soc.* **128** (2000), no. 11, 3379–3390. doi:10.1090/S0002-9939-00-05453-8.

[81] E. Hebey, *Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities*, Courant Lecture Notes in Mathematics 5, American Mathematical Society, Providence, 1999.

[82] J. Heinonen and P. Koskela, Quasiconformal maps in metric spaces with controlled geometry, *Acta Math.* **181** (1998), 1–61. doi:10.1007/BF02392747.

[83] J. Heinonen, *Lectures on Analysis on Metric Spaces*, Universitext, Springer, New York, 2001.

[84] M. Hino, On singularity of energy measures on self-similar sets, *Probab. Theory Related Fields* **132** (2005), 265–290.

[85] M. Hino, Energy measures and indices of Dirichlet forms, with applications to derivatives on some fractals, *Proc. Lond. Math. Soc. (3)* **100** (2010), 269–302.

[86] M. Hino, Measurable Riemannian structures associated with strong local Dirichlet forms, *Math. Nachr.* **286** (2013), 1466–1478.

[87] M. Hino, Upper estimate of martingale dimension for self-similar fractals, *Probab. Theory Related Fields* **156** (2013), 739–793.

[88] M. Hino, Martingale dimensions for a class of metric measure spaces, preprint (2025), revised 26 October 2026, arXiv:2511.09930v2.

[89] M. Hinz, D. Koch and M. Meinert, Sobolev spaces and calculus of variations on fractals, in: *Analysis, Probability and Mathematical Physics on Fractals*, World Scientific, 2020, 419–450; arXiv:1805.04456.

[90] M. Hinz, M. Röckner and A. Teplyaev, Vector analysis for Dirichlet forms and quasilinear PDE and SPDE on metric measure spaces, *Stochastic Process. Appl.* **123** (2013), 4373–4406.

[91] T. Hytönen, A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa, *Publ. Mat.* **54** (2010), 485–504.

[92] M. Ionescu, L. G. Rogers and A. Teplyaev, Derivations, Dirichlet forms and spectral analysis, *J. Funct. Anal.* **263** (2012), 2141–2169.

[93] D. Jerison, The Poincaré inequality for vector fields satisfying Hörmander's condition, *Duke Math. J.* **53** (1986), 503–523. doi:10.1215/S0012-7094-86-05329-9.

[94] R. Jiang, H. Li and H. Zhang, Heat kernel bounds on metric measure spaces and some applications, *Potential Anal.* **44** (2016), 601–627. doi:10.1007/s11118-015-9521-2.

[95] J.-P. Kahane, Sur le chaos multiplicatif, *Ann. Sci. Math. Québec* **9** (1985), no. 2, 105–150. Journal record.

[96] N. Kajino and M. Murugan, On singularity of energy measures for symmetric diffusions with full off-diagonal heat kernel estimates, *Ann. Probab.* **48** (2020), 2920–2951.

[97] N. Kajino, Heat kernel asymptotics for the measurable Riemannian structure on the Sierpinski gasket, *Potential Anal.* **36** (2012), 67–115. doi:10.1007/s11118-011-9221-5.

[98] N. Kajino and R. Shimizu, Contraction properties and differentiability of p-energy forms with applications to nonlinear potential theory on self-similar sets, arXiv:2404.13668, author revision of 28 September 2024. Author's text.

[99] N. Kajino and R. Shimizu, Korevaar–Schoen p-energy forms and associated p-energy measures on fractals, in: *Facets of Contemporary Analysis, Geometry and Non-Euclidean Statistics*, Tohoku Series in Mathematical Sciences 1, Springer, 2026, 113–182; arXiv:2404.13435v2.

[100] T. Kato, *Perturbation Theory for Linear Operators*, second edition, Grundlehren der mathematischen Wissenschaften 132, Springer, Berlin, 1976.

[101] S. Keith and X. Zhong, The Poincaré inequality is an open ended condition, *Ann. of Math. (2)* **167** (2008), 575–599. doi:10.4007/annals.2008.167.575.

[102] S. Keith, Modulus and the Poincaré inequality on metric measure spaces, *Math. Z.* **245** (2003), 255–292. doi:10.1007/s00209-003-0542-y.

[103] S. Keith, A differentiable structure for metric measure spaces, *Adv. Math.* **183** (2004), 271–315.

[104] J. Kigami, Measurable Riemannian geometry on the Sierpinski gasket: the Kusuoka measure and the Gaussian heat kernel estimate, *Math. Ann.* **340** (2008), 781–804. doi:10.1007/s00208-007-0169-0.

[105] J. Kigami, Time changes of the Brownian motion: Poincaré inequality, heat kernel estimate and protodistance, *Mem. Amer. Math. Soc.* **259** (2019), no. 1250, doi:10.1090/memo/1250.

[106] J. Kinnunen, R. Korte, J. Lehrbäck and A. V. Vähäkangas, A maximal function approach to two-measure Poincaré inequalities, *J. Geom. Anal.* **29** (2019), 1763–1810. doi:10.1007/s12220-018-0061-z.

[107] J. Kinnunen and V. Latvala, Lebesgue points for Sobolev functions on metric spaces, *Rev. Mat. Iberoam.* **18** (2002), 685–700. doi:10.4171/RMI/332.

[108] B. Kim, Poincaré inequality and the uniqueness of solutions for the heat equation associated with subelliptic diffusion operators, preprint (2013), arXiv:1305.0508v1.

[109] J. Kinnunen, J. Lehrbäck, A. V. Vähäkangas and X. Zhong, Maximal function estimates and self-improvement results for Poincaré inequalities, *Manuscripta Math.* **158** (2019), 119–147. doi:10.1007/s00229-018-1016-1.

[110] B. Kleiner and J. M. Mackay, Differentiable structures on metric measure spaces: a primer, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **16** (2016), 41–64.

[111] B. Kleiner and A. Schioppa, PI spaces with analytic dimension 1 and arbitrary topological dimension, *Indiana Univ. Math. J.* **66** (2017), 495–546. arXiv:1504.06646.

[112] N. J. Korevaar and R. M. Schoen, Sobolev spaces and harmonic maps for metric space targets, *Comm. Anal. Geom.* **1** (1993), 561–659.

[113] P. Koskela, K. N. Nguyen and Z. Wang, Trace and density results on regular trees, *Potential Anal.* **57** (2022), 101–128. doi:10.1007/s11118-021-09907-2.

[114] P. Koskela, N. Shanmugalingam and J. T. Tyson, Dirichlet forms, Poincaré inequalities, and the Sobolev spaces of Korevaar and Schoen, *Potential Anal.* **21** (2004), 241–262. doi:10.1023/B:POTA.0000033331.88514.6e.

[115] P. Koskela, N. Shanmugalingam and Y. Zhou, Intrinsic geometry and analysis of diffusion processes and L^∞-variational problems, *Arch. Ration. Mech. Anal.* **214** (2014), 99–142.

[116] P. Koskela and P. MacManus, Quasiconformal mappings and Sobolev spaces, *Studia Math.* **131** (1998), 1–17.

[117] P. Koskela and Y. Zhou, Geometry and analysis of Dirichlet forms, *Adv. Math.* **231** (2012), 2755–2801.

[118] S. Kusuoka, Dirichlet forms on fractals and products of random matrices, *Publ. Res. Inst. Math. Sci.* **25** (1989), 659–680.

[119] K. Kuwae, $(1,p)$-Sobolev spaces based on strongly local Dirichlet forms, *Math. Nachr.* **297** (2024), 3723–3740. doi:10.1002/mana.202400025; arXiv:2310.11652v2.

[120] T. J. Laakso, Ahlfors Q-regular spaces with arbitrary $Q > 1$ admitting weak Poincaré inequality, *Geom. Funct. Anal.* **10** (2000), 111–123. doi:10.1007/s000390050003. Erratum, *Geom. Funct. Anal.* **12** (2002), 650, doi:10.1007/s00039-002-8261-9.

[121] J. Lierl, Parabolic Harnack inequality on fractal-type metric measure Dirichlet spaces, *Rev. Mat. Iberoam.* **34** (2018), 687–738, doi:10.4171/RMI/1001.

[122] N. Marola, M. Miranda Jr. and N. Shanmugalingam, Characterizations of sets of finite perimeter using heat kernels in metric spaces, *Potential Anal.* **45** (2016), 609–633.

[123] G. Maruyama, The harmonic analysis of stationary stochastic processes, *Mem. Fac. Sci. Kyushu Univ. Ser. A* **4** (1949), no. 1, 45–106. doi:10.2206/kyushumfs.4.45.

[124] V. Maz'ya, *Sobolev Spaces with Applications to Elliptic Partial Differential Equations*, second edition, Grundlehren der mathematischen Wissenschaften 342, Springer, Heidelberg, 2011.

[125] E. J. McShane, Extension of range of functions, *Bull. Amer. Math. Soc.* **40** (1934), 837–842. doi:10.1090/S0002-9904-1934-05978-0.

[126] M. Miranda Jr., Functions of bounded variation on "good" metric spaces, *J. Math. Pures Appl. (9)* **82** (2003), 975–1004.

[127] J. Moser, A Harnack inequality for parabolic differential equations, *Comm. Pure Appl. Math.* **17** (1964), 101–134. doi:10.1002/cpa.3160170106. Correction, *Comm. Pure Appl. Math.* **20** (1967), 231–236.

[128] J. Moser, On a pointwise estimate for parabolic differential equations, *Comm. Pure Appl. Math.* **24** (1971), 727–740. doi:10.1002/cpa.3160240507.

[129] M. Murugan and R. Shimizu, First-order Sobolev spaces, self-similar energies and energy measures on the Sierpiński carpet, *Comm. Pure Appl. Math.* **78** (2025), 1523–1608. doi:10.1002/cpa.22247.

[130] M. Murugan and L. Saloff-Coste, Davies' method for anomalous diffusions, *Proc. Amer. Math. Soc.* **145** (2017), no. 4, 1793–1804, doi:10.1090/proc/13324.

[131] M. Murugan, Martingale and analytic dimensions coincide under Gaussian heat kernel bounds, *Bull. Sci. Math.* **204** (2025), 103710. doi:10.1016/j.bulsci.2025.103710.

[132] A. Nagel, E. M. Stein and S. Wainger, Balls and metrics defined by vector fields I. Basic properties, *Acta Math.* **155** (1985), 103–147, doi:10.1007/BF02392539.

[133] J. Nash, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.* **80** (1958), 931–954. doi:10.2307/2372841.

[134] K. Pietruska-Pałuba, Heat kernel characterisation of Besov–Lipschitz spaces on metric measure spaces, *Manuscripta Math.* **131** (2010), 199–214. doi:10.1007/s00229-009-0310-3.

[135] M. Pivarski and L. Saloff-Coste, Small time heat kernel behavior on Riemannian complexes, *New York J. Math.* **14** (2008), 459–494.

[136] T. Rajala, Approximation by uniform domains in doubling quasiconvex metric spaces, *Complex Anal. Synerg.* **7** (2021), Art. 4. doi:10.1007/s40627-021-00062-3.

[137] T. Rajala, Interpolated measures with bounded density in metric spaces satisfying the curvature-dimension conditions of Sturm, *J. Funct. Anal.* **263** (2012), no. 4, 896–924, doi:10.1016/j.jfa.2012.05.006.

[138] R. Rhodes and V. Vargas, Gaussian multiplicative chaos and applications: a review, *Probab. Surv.* **11** (2014), 315–392, doi:10.1214/13-PS218.

[139] L. Saloff-Coste, *Aspects of Sobolev-Type Inequalities*, London Mathematical Society Lecture Note Series 289, Cambridge University Press, Cambridge, 2002.

[140] L. Saloff-Coste, A note on Poincaré, Sobolev, and Harnack inequalities, *Internat. Math. Res. Notices* **1992** (1992), no. 2, 27–38. doi:10.1155/S1073792892000047.

[141] K. Sasaya, Construction of p-energy measures associated with strongly local p-energy forms, *J. Funct. Anal.* **291** (2026), 111489. doi:10.1016/j.jfa.2026.111489.

[142] K. Sasaya, The strongly local part of nonlinear Dirichlet structures and its energy measures, preprint (2026), arXiv:2611.14400v1.

[143] A. Schioppa, Examples of 2-unrectifiable normal currents, arXiv:1508.00756v2 (2016).

[144] A. Schioppa, Derivations and Alberti representations, *Adv. Math.* **293** (2016), 436–528.

[145] N. Shanmugalingam, Newtonian spaces: An extension of Sobolev spaces to metric measure spaces, *Rev. Mat. Iberoam.* **16** (2000), 243–279. doi:10.4171/RMI/275.

[146] R. Shimizu, Characterizations of Sobolev functions via Besov-type energy functionals in fractals, *Potential Anal.* **63** (2025), 2121–2156; arXiv:2410.15317v2.

[147] P. Stollmann, A dual characterization of length spaces with application to Dirichlet metric spaces, *Studia Math.* **198** (2010), 221–233.

[148] K.-T. Sturm, On the geometry of metric measure spaces. II, *Acta Math.* **196** (2006), 133–177, doi:10.1007/s11511-006-0003-7.

[149] B. Steinhurst and A. Teplyaev, Spectral analysis on Barlow and Evans' projective limit fractals, *J. Spectr. Theory* **11** (2021), 91–123. doi:10.4171/JST/337.

[150] K.-T. Sturm, Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties, *J. Reine Angew. Math.* **456** (1994), 173–196.

[151] K.-T. Sturm, Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality, *J. Math. Pures Appl. (9)* **75** (1996), 273–297.

[152] N. Weaver, Lipschitz algebras and derivations. II. Exterior differentiation, *J. Funct. Anal.* **178** (2000), 64–112.

[153] A. Shamov, On Gaussian multiplicative chaos, *J. Funct. Anal.* **270** (2016), 3224–3261, doi:10.1016/j.jfa.2016.03.001.