

Sobolev Inequalities in Familiar and Unfamiliar Settings

Laurent Saloff-Coste

Abstract The classical Sobolev inequalities play a key role in analysis in Euclidean spaces and in the study of solutions of partial differential equations. In fact, they are extremely flexible tools and are useful in many different settings. This paper gives a glimpse of assortments of such applications in a variety of contexts.

1 Introduction

There are few articles that have turned out to be as influential and truly important as S.L. Sobolev 1938 article [93] (the American translation appeared in 1963), where he introduces his famed inequalities. It is the idea of a functional inequality itself that Sobolev brings to life in his paper, as well as the now so familiar notion of an a priori inequality, i.e., a functional inequality established under some strong hypothesis and that might be extended later, perhaps almost automatically, to its natural domain of definition. (These ideas are also related to the theory of distributions which did not exist at the time and whose magnificent development by L. Schwartz was, in part, anticipated in the work of S.L. Sobolev.)

The most basic and important applications of Sobolev inequalities are to the study of partial differential equations. Simply put, Sobolev inequalities provide some of the very basic tools in the study of the existence, regularity, and uniqueness of the solutions of all sorts of partial differential equations, linear and nonlinear, elliptic, parabolic, and hyperbolic. I leave to others, much better qualified than me, to discuss these beautiful developments. Instead, my aim in this paper is to survey briefly an assortments of perhaps less familiar applications of Sobolev inequalities (and related inequalities) to problems and

Laurent Saloff-Coste
Cornell University, Mallot Hall, Ithaca, NY 14853, USA, e-mail: lsc@math.cornell.edu

in settings that are not always directly related to PDEs, at least not in the most classical sense. The inequalities introduced by S.L. Sobolev have turned out to be extremely useful flexible tools in surprisingly diverse settings. My hope is to be able to give to the reader a glimpse of this diversity. The reader must be warned that the collection of applications of Sobolev inequalities described below is very much influenced by my own interest, knowledge, and limitations. I have not tried at all to present a complete picture of the many different ways Sobolev inequalities have been used in the literature. That would be a very difficult task.

2 Moser's Iteration

2.1 The basic technique

This section is included mostly for those readers that are not familiar with the use of Sobolev inequalities. It illustrates some aspects of one of the basic techniques associated with their use. To the untrained eyes, the fundamental nature of Sobolev inequalities is often lost in the technicalities surrounding their use. Indeed, outside analysis, L^p spaces other than L^1 , L^2 , and L^∞ still appear quite exotic to many. As the following typical example illustrates, they play a key role in extracting the information contained in Sobolev inequalities.

Recall that Hölder's inequality states that

$$\int |fg| dx \leq \|f\|_p \|g\|_q$$

as long as $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$ (these are called conjugate exponents). A somewhat clever use of this inequality yields

$$\|f\|_r \leq \|f\|_s^\theta \|f\|_t^{1-\theta}$$

as long as $1 \leq r, s, t \leq \infty$ and $1/r = \theta/s + (1-\theta)/t$. These basic inequalities are used extensively in conjunction with Sobolev inequalities.

Let $\Delta = \sum (\partial/\partial x_i)^2$ be the Laplacian in $\mathbb{R}^n$. Consider a bounded domain $\Omega \subset \mathbb{R}^n$, $\lambda \geq 0$, and the (Dirichlet) eigenfunction/eigenvalue problem:

$$\Delta u = -\lambda u \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0. \quad (2.1)$$

Our goal is to show how the remarkable inequality

$$\sup_{\Omega} \{|u|^2\} \leq A_n \lambda^{n/2} \int_{\Omega} |u|^2 dx \quad (2.2)$$

(for solutions of (2.1)) follows from the most classical Sobolev inequality, namely, the inequality (2.5) below. For a normalized eigenfunction u with $\|u\|_2 = 1$ the inequality (2.2) bounds the size of u in terms of the associated eigenvalue. The technique illustrated below is extremely flexible and can be adapted to many situations.

In fact, we only assume that $u \in H_0^1(\Omega)$, i.e., u is the limit of smooth compactly supported functions in Ω in the norm

$$\|u\| = \left(\int_{\Omega} [|u|^2 + \sum_1^n |\partial u / \partial x_i|^2] dx \right)^{1/2}$$

and that

$$\int_{\Omega} \sum_1^n \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} dx = \lambda \int_{\Omega} u v dx \quad (2.3)$$

for any $v \in H_0^1(\Omega)$. We set $\nabla u = (\partial u / \partial x_i)_1^n$, $|\nabla u|^2 = \sum_1^n |\partial u / \partial x_i|^2$. To avoid additional technical arguments, we assume a priori that u is bounded on Ω . For $1 \leq p < \infty$ we take $v = |u|^{2p-1}(u/|u|)$ in (2.3). This yields

$$\lambda \int_{\Omega} |u|^{2p} = (2p-1) \int_{\Omega} |u|^{2p-2} |\nabla u|^2 dx = \frac{2p-1}{p^2} \int_{\Omega} |\nabla |u|^p|^2 dx. \quad (2.4)$$

As our starting point, we take the most basic Sobolev inequality

$$\forall f \in H^1(\mathbb{R}^n), \quad \left(\int |f|^{2q_n} dx \right)^{1/q_n} \leq C_n^2 \int |\nabla f|^2 dx, \quad q_n = n/(2-n). \quad (2.5)$$

If $k_n = (1 + 2/n)$, then $1/k_n = \theta_n/q_n + (1 - \theta_n)$ with $\theta_n = n/(n+2)$, and Hölder's inequality yields

$$\int |f|^{2k_n} dx \leq \left(\int |f|^{2q_n} dx \right)^{1/q_n} \left(\int |f|^2 dx \right)^{2/n}.$$

Together with the previous Sobolev inequality, we obtain

$$\int |f|^{2(1+2/n)} dx \leq C_n^2 \int |\nabla f|^2 dx \left(\int |f|^2 dx \right)^{2/n}. \quad (2.6)$$

For a discussion of this type of “multiplicative” inequality see, for example, [75, Sect. 2.3].

Now, for a solution u of (2.1) the inequalities (2.6) and (2.4) yield

$$\int |u|^{2p(1+2/n)} dx \leq C_n^2 p \lambda \left(\int |u|^{2p} dx \right)^{1+2/n}.$$

This inequality can obviously be iterated by taking $p_i = (1 + 2/n)^i$, and we get

$$\left(\int |u|^{2p_i} dx \right)^{1/p_i} \leq (1 + 2/n)^{\sum_1^i (j-1)p_j^{-1}} (3C_n^2 \lambda)^{\sum_1^i p_j^{-1}} \int |u|^2 dx.$$

Note that $\sum_1^\infty p_j^{-1} = n/2$ and $\lim_{p \rightarrow \infty} \| |u|^2 \|_p = \|u^2\|_\infty$. The desired conclusion (2.2) follows.

2.2 Harnack inequalities

The technique illustrated above is the simplest instance of what is widely known as Moser's iteration technique. In a series of papers [77]–[80], Moser developed this technique as the basis for the study of divergence form uniformly elliptic operators in $\mathbb{R}^n$, i.e., operators of the form (we use $\partial_i = \partial/\partial x_i$)

$$L_a = \sum_{i,j} \partial_i (a_{i,j}(x) \partial_j)$$

with real matrix-valued function a satisfying the ellipticity condition

$$\forall x \in \Omega, \quad \begin{cases} \sum_{i,j} a_{i,j}(x) \xi_i \xi_j \geq \varepsilon |\xi|^2, \\ \sum_{i,j} a_{i,j}(x) \xi_i \xi'_j \leq \varepsilon^{-1} |\xi| |\xi'|, \end{cases}$$

where $\varepsilon > 0$ and the coefficients $a_{i,j}$ are simply bounded real measurable functions. Because of the low regularity of the coefficients, the most basic question in this context is that of the boundedness and continuity of solutions of the equation $L_a u = 0$ in the interior of an open set Ω . This was solved earlier by De Giorgi [34] (and by Nash [81] in the parabolic case), but Moser proposed an alternative method, squarely based on the use of Sobolev inequality (2.5). To understand why one might hope this is possible, observe that the argument given in the previous section works without essential changes if, in (2.1), one replaces the Laplacian Δ by L_a .

Let u be a solution of $L_a u = 0$ in a domain Ω , in the sense that for any open relatively compact set Ω_0 in Ω , $u \in H^1(\Omega_0)$ and for all $v \in H_0^1(\Omega_0)$

$$\int_\Omega \sum_{i,j} a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx = 0.$$

In [77], Moser observed that the interior boundedness and continuity of such a solution follow from the Harnack inequality that provides a constant $C(n, \varepsilon)$

such that if u as above is nonnegative in Ω and the ball B satisfies $2B \subset \Omega$, then

$$\sup_B \{u\} \leq C(n, \varepsilon) \inf_B \{u\} \quad (2.7)$$

(a priori, the supremum and infimum should be understood here as essential supremum and essential infimum. The ball $2B$ is concentric with B with twice the radius of B). He then proceeded to prove this Harnack inequality by variations on the argument outlined in the previous section. In his later papers [78]–[80], Moser obtained a parabolic version of the above Harnack inequality. Namely, he proved that there exists a constant $C(n, \varepsilon)$ such that any nonnegative solution u of the heat equation $(\partial_t - L_a)u = 0$ in a time-space cylinder $Q = (s - 4r^2, s) \times 2B$ satisfies

$$\sup_{Q_-} \{u\} \leq C(n, \varepsilon) \inf_{Q_+} \{u\}, \quad (2.8)$$

where $Q_- = (s - 3r^2, s - 2r^2) \times B$ and $Q_+ = (s - r^2/2, s) \times B$.

Moser's iteration technique has been adapted and used in hundreds of papers studying various PDE problems. Some early examples are [2, 3, 90]. The books [42, 69, 76] contain many applications of this circle of ideas, as well as further references. The survey paper [83] deals specifically with the heat equation and is most relevant for the purpose of the present paper.

The basic question we want to explore in the next two subsections is: what exactly are the crucial ingredients of Moser's iteration? This question is motivated by our desire to use this approach in other settings such as Riemannian manifolds or more exotic spaces. Early uses of Moser's iteration technique on manifolds as in the influential papers [22, 23] were actually limited by a misunderstanding of what is really needed to run this technique successfully. Interesting early works that explored the flexibility of Moser's iteration beyond the classical setting are related to degenerated elliptic operators as in [56]–[58] (see also [39] and the references therein).

2.3 Poincaré, Sobolev, and the doubling property

Moser's technique in $\mathbb{R}^n$ uses only three crucial ingredients:

- (1) The Sobolev inequality in the form (2.6), i.e.,

$$\forall f \in C_0^\infty(\mathbb{R}^n), \quad \int |f|^{2(1+2/n)} dx \leq C_n^2 \int |\nabla f|^2 dx \left(\int |f|^2 dx \right)^{2/n}.$$

- (2) The Poincaré inequality in the unit ball B , i.e.,

$$f \in \mathcal{C}^\infty(B), \quad \int_B |f - f_B|^2 dx \leq P_n \int_B |\nabla f|^2 dx,$$

where f_B stands for the average of f over B .

(3) Translations and dilations.

Some might be surprised that the interesting John–Nirenberg inequality that appears to be a crucial tool in [77, 78] is not mentioned above. However, as Moser himself pointed out in [80], it can be avoided altogether by using a clever, but very elementary observation of Bombieri and Giusti. Somewhat unfortunately, this important simplification has been ignored by a large part of the later literature!

Obviously, in order to use the method in a larger context, one wants to replace the use of translations and dilations by hypotheses that are valid at all scales and locations. For instance, the needed Poincaré inequality takes the form

$$\forall z, \forall r > 0, f \in \mathcal{C}^\infty(B(z, r)), \quad \int_{B(z, r)} |f - f_{B(z, r)}|^2 dx \leq P_n r^2 \int_{B(z, r)} |\nabla f|^2 dx,$$

where f_B stands for the average of f over B . A correct generalization is less obvious in the case of the Sobolev inequality. As stated, the inequality (2.6) turns out to be too restrictive and not strong enough, both at the same time!

For instance, consider a complete Riemannian manifold (M, g) of dimension n . We set $|\nabla f|^2 = g(\nabla f, \nabla f)$, where the gradient ∇f is the vector field defined by $g_x(\nabla f, X) = df(X)$ for any tangent vector X at x . Let μ be the Riemannian measure, $B(x, r)$ the geodesic ball with center x and radius r , and

$$V(x, r) = \mu(B(x, r)).$$

If we assume that the inequality analogous to (2.6) holds on M , i.e.,

$$\forall f \in \mathcal{C}_0^\infty(M), \quad \int |f|^{2(1+2/n)} d\mu \leq C_M^2 \int |\nabla f|^2 d\mu \left(\int |f|^2 d\mu \right)^{2/n}, \quad (2.9)$$

then it turns out that this implies the existence of a constant $c_M > 0$ such that

$$\forall x \in M, \quad \forall r > 0, \quad \mu(B(x, r)) = V(x, r) \geq c_M r^n$$

(see [17, Proposition 2.4] and [87, Theorem 3.15]). This rules out simple manifolds such as $\mathbb{R}^{n+k}/\mathbb{Z}^k$ or $\mathbb{R}^{n-k} \times \mathbb{S}^k$ (on which, for other reasons, one knows that the above-mentioned analogs of the Harnack inequalities (2.7), (2.8) hold). Let us observe that when $n \geq 3$, (2.9) is, in fact, equivalent to the more standard Sobolev inequality

$$\forall f \in C_0^\infty(M), \quad \left(\int |f|^{2q_n} d\mu \right)^{1/q_n} \leq C_M^2 \int |\nabla f|^2 d\mu, \quad q_n = n/(n-2), \quad (2.10)$$

where the constant C_M may be different in (2.9) and in (2.10).

In the other direction, (2.10) and thus (2.9) holds in the case of hyperbolic spaces (with dimension $n > 2$ for (2.10)), but the desired Harnack inequalities fail to hold uniformly at large scale in such spaces.

Definition 2.1. We say that a complete Riemannian manifold M satisfies a *scale invariant family* of Sobolev inequalities if there is a constant C_M and a real number $q = \nu/(\nu-2) > 1$ such that for any $x \in M$, $r > 0$, and $B = B(x, r)$ we have

$$\forall f \in C_0^\infty(B), \quad \left(\int_B |f|^{2q} d\mu \right)^{1/q} \leq \frac{C_M r^2}{\mu(B)^{2/\nu}} \int_B [|\nabla f|^2 + r^{-2}|f|^2] d\mu. \quad (2.11)$$

Remark 2.1. The inequality (2.11) can be written in the form: for all $f \in C_0^\infty(B)$

$$\left(\frac{1}{\mu(B)} \int_B |f|^{2q} d\mu \right)^{1/q} \leq C_M r^2 \left(\frac{1}{\mu(B)} \int_B [|\nabla f|^2 + r^{-2}|f|^2] d\mu \right).$$

Remark 2.2. In this definition, the exact value of q is not very important and ν appears here as a technical parameter. If (2.11) holds for some $q = \nu/(\nu-2) > 1$, then the Jensen inequality shows that it also holds for all $1 < q' = \nu'/(\nu'-2) \leq q$, i.e., for all finite $\nu' \geq \nu$.

Remark 2.3. In general, (2.10) does not imply (2.11). However, (2.10) does imply (2.11) with $\nu = n$ when the manifold M has an Euclidean type volume growth, i.e., there exists $0 < v_M \leq V_M < \infty$ such that $v_M r^n \leq V(x, r) \leq V_M r^n$ for all $x \in M$ and $r > 0$. This is obviously a very restrictive and undesirable hypothesis. This is exactly the point that restricted the use of Moser's iteration technique to very local results in some early applications of the technique to analysis on Riemannian manifolds as in [22, 23].

Remark 2.4. There are many equivalent forms of (2.11). We mention three. The first one is analogous to (2.6) and reads

$$\int_B |f|^{2(1+2/\nu)} d\mu \leq \frac{C_M r^2}{\mu(B)^{2/\nu}} \int_B [|\nabla f|^2 + r^{-2}|f|^2] d\mu \left(\int_B |f|^2 d\mu \right)^{2/\nu}.$$

The second is in the form of the so-called *Nash inequality* and reads

$$\int_B |f|^{2(1+2/\nu)} d\mu \leq \frac{C_M r^2}{\mu(B)^{2/\nu}} \int_B [|\nabla f|^2 + r^{-2}|f|^2] d\mu \left(\int_B |f| d\mu \right)^{4/\nu}$$

(see [81] and [75, Sect. 2.3]). The third is often referred to as a *Faber–Krahn inequality* (see [44]) and reads

$$\lambda_D(\Omega) \geq \frac{c_M}{r^2} \left(\frac{\mu(\Omega)}{\mu(B)} \right)^{2/\nu},$$

where $\lambda_D(\Omega)$ is the lowest Dirichlet eigenvalue in Ω , an arbitrary subset of the ball B of radius r . In each case, r is the radius of B and the inequality must hold uniformly for all geodesic balls B . The exact value of the constants varies from one type of inequality to another. Many results in the spirit of these equivalences can be found in [75] in the context of Euclidean domains. A discussion in a very general setting is in [4] (see also [87, Chapt. 3]).

The following theorem describes some of the noteworthy consequences of (2.11). Let Δ_M be the Laplace operator on M , and let $h(t, x, y)$ be the (minimal) fundamental solution of the heat equation $(\partial_t - \Delta_M)u = 0$ on M , i.e., the kernel of the heat semigroup $e^{t\Delta_M}$. For complete discussions, surveys, and variants, see [43, 44, 46, 45, 48, 85, 86, 87].

Theorem 2.1. *Assume that (M, g) is a complete Riemannian manifold which satisfies the scale invariant family of Sobolev inequalities (2.11) (with some parameter $\nu > 2$). Then the following properties hold.*

- *There exists a constant V_M such that for any two concentric balls $B \subset B'$ with radii $0 < r < r' < \infty$*

$$\mu(B') \leq V_M (r'/r)^\nu \mu(B). \quad (2.12)$$

- *There exists a constant C_M such that for all $x \in M$ and $r > 0$ any positive subsolution u of the heat equation in a time-space cylinder $Q = (s - 4r^2, s) \times B(x, 2r)$ satisfies*

$$\sup_{Q'} \{u^2\} \leq C_M \frac{1}{r^2 \mu(B)} \int_Q |u|^2 d\mu ds, \quad (2.13)$$

where $Q' = (s - r^2, s) \times B(x, r)$.

- *For any integer $k \geq 0$ there is a constant $A(M, k)$ such that for all points $x, y \in M$ and $t > 0$*

$$|\partial_t^k h(t, x, y)| \leq \frac{A(M, k)}{t^k V(x, \sqrt{t})} (1 + d(x, y)^2/t)^{\nu+k} \exp\left(-\frac{d(x, y)^2}{4t}\right). \quad (2.14)$$

Remark 2.5. The inequality (2.13) can be obtained by a straightforward application of Moser's iteration technique. One of many possible applications of (2.13) is (2.14).

Remark 2.6. The volume inequality (2.12), together with the heat kernel bound

$$\forall x \in M, t > 0, \quad h(t, x, x) < \frac{A_M}{V(x, \sqrt{t})},$$

implies the Sobolev inequality (2.11).

Definition 2.2. A complete Riemannian manifold has the *doubling volume property* if there exists a constant V_M such that

$$\forall x \in M, r > 0, \quad V(x, 2r) \leq V_M V(x, r).$$

Remark 2.7. It is easy to see that the doubling property implies (2.12) with $\nu = \log_2 V_M$.

Definition 2.3. A complete Riemannian manifold admits a *scale invariant Poincaré inequality* (in L^2) if there exists a constant P_M such that

$$\forall x \in M, r > 0, \quad \int_B |f - f_B|^2 d\mu \leq P_M r^2 \int_B |\nabla f|^2 d\mu, \quad (2.15)$$

where $B = B(x, r)$ and f_B is the average of f over B .

Remark 2.8. This Poincaré inequality can be stated in terms of the spectrum of minus the Neumann Laplacian in geodesic balls. For minus the Neumann Laplacian (understood in an appropriate sense) in a ball B , the lowest eigenvalue is 0 (associated with constant functions). The L^2 Poincaré inequality above is equivalent to say that the second eigenvalue $\lambda_N(B(x, r))$ is bounded from below by cr^{-2} , where r is the radius of B .

Remark 2.9. Keeping Moser's iteration in mind, it is a very important and remarkable fact that if M satisfies both the doubling property and a scale invariant Poincaré inequality, then it satisfies (2.11) (see [85]–[87]). In this case, one can take ν to be an arbitrary number greater than 2 and such that (2.12) holds.

Definition 2.4. A complete Riemannian manifold admits a *scale invariant parabolic Harnack inequality* if there exists a constant C_M such that for any $x \in M$, $r > 0$, and $s \in \mathbb{R}$ and for any nonnegative solution u of the heat equation in the time-space cylinder $Q = (s - 4r^2, s) \times B(x, 2r)$

$$\sup_{Q_-} \{u\} \leq C_M \inf_{Q_+} \{u\}$$

with $Q_- = (s - 3r^2, s - 2r^2) \times B(x, r)$ and $Q_+ = (s - r^2, s) \times B(x, r)$.

In this setting, a version of Moser's iteration methods gives one half of the following result (see [43, 85] and a detailed discussion in [87, Sect. 5.5]).

Theorem 2.2. *Let (M, g) be a complete Riemannian manifold. The following properties are equivalent.*

- *The doubling property and a scale invariant L^2 Poincaré inequality.*
- *The scale invariant parabolic Harnack inequality.*
- *The two-sided heat kernel bound*

$$\frac{c}{V(x, \sqrt{t})} \exp\left(-A \frac{d(x, y)^2}{t}\right) \leq h(t, x, y) \leq \frac{C}{V(x, \sqrt{t})} \exp\left(-a \frac{d(x, y)^2}{t}\right)$$

for constants $0 < a, A, c, C < \infty$.

One may ask how the above properties are related to the elliptic version of Harnack inequality. This is not entirely understood, but the following result involving the Sobolev inequality (2.11) sheds some light on this question (see [61]).

Theorem 2.3. *Let M be a complete manifold satisfying the Sobolev inequality (2.11) for some $q > 1$. Then the following properties are equivalent.*

- *The scale invariant L^2 Poincaré inequality.*
- *The scale invariant elliptic Harnack inequality.*
- *The scale invariant parabolic Harnack inequality.*

We conclude with results concerning global harmonic functions.

Theorem 2.4. *Let M be a manifold satisfying the doubling volume property and a scale invariant L^2 Poincaré inequality.*

- *Any harmonic functions on M that is bounded from below must be constant.*
- *There exists $a_0 > 0$ such that for any fixed point $x \in M$ any harmonic function satisfying $\sup_y \{u(y)/(1 + d(x, y))^{a_0}\} < \infty$ must be constant.*
- *For any $a > 0$ and a fixed point $x \in M$ the space of harmonic functions on M satisfying $\sup_y \{u(y)/(1 + d(x, y))^a\} < \infty$ is finite dimensional.*

Remark 2.10. The first two statements are standard consequences of the (scale invariant) elliptic Harnack inequality which follows from the assumptions of the theorem. The last statement is a recent result due to Colding and Minicozzi [24, 25, 72, 71]. The proof of the last statement makes explicit the use of the Poincaré inequality and the doubling volume property. A number of interesting variations on this result are discussed in [24, 25, 72, 71]. A different viewpoint concerning Liouville theorems, restricted to some special circumstances, but very interesting nonetheless is developed in [68].

Example 2.1. Euclidean spaces are the model examples for manifolds that satisfy both the doubling condition and the Poincaré inequality. Larger classes of examples will be described in the next section. Interesting examples where the Poincaré inequality fails are obtained by considering manifolds M that are the connected sum of two (or more) Euclidean spaces. Here, we write $M = \mathbb{R}^n \# \mathbb{R}^n$ to mean a complete Riemannian manifold that can be decomposed in the disjoint union $E_1 \cup K \cup E_2$, where E_1, E_2 are each isometric to the outside of some compact domain with smooth boundary in $\mathbb{R}^n$ and K is a smooth compact manifold with boundary. In words, $\mathbb{R}^n \# \mathbb{R}^n$ is made of two copies of $\mathbb{R}^n$ smoothly attached together through a compact “collar.” The following facts (that are not too difficult to check) make these examples interesting.

- $M = \mathbb{R}^n \# \mathbb{R}^n$ has the doubling property. In fact, obviously, $V(x, r) \simeq r^n$.
- $M = \mathbb{R}^n \# \mathbb{R}^n$ satisfies (2.11) with ν being any positive real that is both at least n and greater than 2. In fact, (2.9) holds on $\mathbb{R}^n \# \mathbb{R}^n$ for any n , and (2.10) holds if $n > 2$. This means that Theorem 2.1 applies.
- Except for the trivial case $n = 1$, the scale invariant Poincaré inequality (2.15) does not hold on $\mathbb{R}^n \# \mathbb{R}^n$. More precisely, if o is a fixed point in the collar of $\mathbb{R}^n \# \mathbb{R}^n$ and $B_r = B(o, r)$, then for large $r \gg 1$, we have

$$\lambda_N(B_r) \simeq \begin{cases} (r^2 \log r)^{-1} & \text{if } n = 2, \\ r^{-n} & \text{if } n > 2, \end{cases}$$

where $\lambda_N(B_r)$ is the second lowest eigenvalue of the Neumann Laplacian in B_r . This means that the best Poincaré inequality in B_r has a constant that is in $r^2 \log r$ if $n = 2$ and r^n if $n \geq 3$ (instead of the desired r^2).

- For $n > 1$, $M = \mathbb{R}^n \# \mathbb{R}^n$ does not satisfy the elliptic Harnack inequality (again, it fails for nonnegative harmonic functions in the balls B_r as above, when r tends to infinity).
- For $n \leq 2$ there are no nonconstant positive harmonic functions, but for $n \geq 3$ there are nonconstant bounded harmonic function on $M = \mathbb{R}^n \# \mathbb{R}^n$.
- For $n \geq 2$ let o be a point in the collar of $M = \mathbb{R}^n \# \mathbb{R}^n$, and let x and y be, respectively, in the first and second copies of $\mathbb{R}^n$ constituting M , at distance about $r = \sqrt{t}$ from o . Then the heat kernel $h(t, x, y)$ satisfies $h(t, x, y) \simeq t^{-n+1}$. This should be compare with the Euclidean heat kernel at time t for points x, y about $\sqrt{t}$ apart which is of size about $t^{-n/2}$. For more on this we refer the reader to [49]–[51].

2.4 Examples

We briefly discuss various examples that illustrate the above-described results.

Example 2.2 (manifolds with nonnegative Ricci curvature). The Ricci curvature Ric is a symmetric $(0, 2)$ -tensor (obtained by contraction of the full curvature tensor) that contains a lot of useful information. Two well-known early examples of that are:

(1) Meyers' theorem (more on this later) stating that a complete Riemannian manifold with $\text{Ric} \geq Kg$ with $K > 0$ must be compact and

(2) Bishop's volume inequality asserting that if $\text{Ric} \geq Kg$ for some $k \in \mathbb{R}$, then the volume function on M , $V(x, r)$, is bounded from above by the volume function $V_{K/(n-1)}(r)$ of the simply connected space of the same dimension and constant sectional curvature $K/(n-1)$ (see, for example, [20, p.73], [21, Theorem 3.9] and [41, 3.85; 3.101]).

Theorem 2.5 ([14, 22, 74]). *A complete Riemannian manifold (M, g) with nonnegative Ricci curvature satisfies the equivalent properties of Theorem 2.2.*

It is interesting to note that the equivalent properties of Theorem 2.2 were proved independently for manifolds with nonnegative Ricci curvature. The doubling property follows from the more precise Bishop–Gromov volume inequality of [22]. Namely, if $\text{Ric} \geq k(n-1)g$, then

$$\forall x \in M, s > r > 0, \quad \frac{V(x, s)}{V(x, r)} \leq \frac{V_k(s)}{V_k(r)}. \quad (2.16)$$

If $k = 0$, this gives $V(x, s) \leq (s/r)^n V(x, r)$ for all $x \in M, s > r > 0$. The Poincaré inequality follows from the result in [14] (see also [21, Theorems 3.10 and 6.8] and [87, Theorem 5.6.5]). The Harnack inequality and two-sided heat kernel estimate follow from the gradient estimate of Li and Yau [74]. Of course, these results imply that the various conclusions of Theorem 2.4 hold for Riemannian manifolds with nonnegative Ricci curvature. In this setting, the last statement in Theorem 2.4 (due to Colding and Minicozzi) solves a conjecture of Yau (see [24, 25, 72, 71]).

Example 2.3. Let G be a connected real Lie group equipped with a left-invariant Riemannian metric g . Note that the Riemannian measure is also a left-invariant Haar measure. We say that G has polynomial volume growth if there exist $C, a \in (0, \infty)$ such that $V(e, r) \leq Cr^a$ for all $r \geq 1$. A group G with polynomial volume growth must be unimodular (left-invariant Haar measures are also right-invariant) and, by a theorem of Guivarc'h [55], there exists an integer N such that $c_0 \leq r^{-N} V(e, r) \leq C_0$ for all $r \geq 1$. It follows that (G, g) satisfies the volume doubling property. By a simple direct argument (see, for example, [87, Theorem 5.6.1]), the scale invariant Poincaré inequality also

holds. Hence one can apply Theorem 2.2. In fact, in this setting, one has the following result.

Theorem 2.6. *Let G be a connected real unimodular Lie group equipped with a Riemannian metric g . The following properties are equivalent.*

- *The group G has polynomial volume growth.*
- *Any positive harmonic function on G is constant.*
- *The scale invariant elliptic Harnack inequality holds.*
- *The scale invariant Sobolev inequality (2.11) holds for some $q > 1$.*
- *The scale invariant parabolic Harnack inequality holds.*

Proof. Connected Lie groups have either strict polynomial growth $V(e, r) \simeq r^N$ for all $r \geq 1$ for some integer N or exponential volume growth (see [55]). Thus, if the volume growth is polynomial, it must be strictly polynomial and the doubling volume property follows. As already mentioned, it is also very easy to prove the scale invariant Poincaré inequality on a connected Lie group of polynomial growth (see, for example, [87, Theorem 5.6.2]). By Theorem 2.2, this shows that polynomial volume growth implies the parabolic Harnack inequality in this context. The parabolic Harnack inequality implies all the other mentioned properties (see Theorem 2.2 and the various remarks in the previous section). The Sobolev inequality (2.11) implies the doubling volume property, hence polynomial volume growth in this context. The elliptic Harnack property implies the triviality of positive harmonic functions. This, in turns, implies polynomial volume growth by [13, Theorem 1.4 or 1.6]. The stated theorem follows. $\square$

For more general results in this setting see [103]. Harmonic functions of polynomial growth on Lie groups of polynomial growth are studied in [1].

Example 2.4 (coverings of compact manifolds). Let (M, g) be a complete Riemannian manifold such that there exists a discrete group of isometries Γ acting freely and properly on (M, g) with compact quotient N . The discrete group Γ must be finitely generated. Its volume growth is defined by using the word metric and counting measure.

Theorem 2.7. *Let (M, g) be a complete Riemannian manifold such that there exists a discrete group of isometries Γ acting freely and properly on (M, g) with compact quotient M/Γ . The following properties are equivalent.*

- *The group Γ has polynomial volume growth.*
- *The scale invariant elliptic Harnack inequality holds.*
- *The scale invariant Sobolev inequality (2.11) holds for some $q > 1$.*
- *The scale invariant parabolic Harnack inequality holds.*

For a complete discussion see [88, Theorem 5.15]. Note the similarity and differences between this result and Theorem 2.6. The main difference is that, for coverings of a compact manifold, there is no known criterion based on the triviality of positive harmonic functions. This is due to the fact that the group Γ may not be linear (or close to a linear group) (see [13, 88]).

3 Analysis and Geometry on Dirichlet Spaces

3.1 *First order calculus*

One of the recent developments in the theory of Sobolev spaces concerns the definitions and properties of such spaces under minimal hypotheses. The most general setting is that of metric measure spaces. There are very good reasons to try to understand what can be done in that setting including important applications to problems coming from different areas of mathematics and even to questions concerning classical Sobolev spaces. In what follows, I only discuss a very special class of metric measure spaces, but it is useful to keep in mind the more general setting. Indeed, the theory of Sobolev spaces on metric measure spaces is also of interest because of the many similar, but different setting it unifies. We refer the reader to the entertaining books [59, 62, 89] and the review paper [63] for glimpses of the general viewpoint on “first order calculus.”

There are many interesting natural metric spaces (of finite dimension type) on which one wants to do some analysis and that are not Riemannian manifolds. Some appear as limit of Riemannian manifolds, for example, manifolds equipped with sub-Riemannian structures and more exotic objects appearing through various geometric precompactness results. Others are very familiar (polytopal complexes seem to appear in real life as often, if not more often, than true manifolds), but have not been studied in much detail as far as analysis is concerned. One natural structure that captures a good number of such examples and provides many natural analytic objects to study (beyond first order calculus) is the structure of Dirichlet spaces. The earliest detailed reference on Dirichlet spaces is [36]. We refer the reader to [40] for a detailed introduction to Dirichlet spaces.

3.2 *Dirichlet spaces*

This subsection describes a restricted class of Dirichlet spaces that provides nice metric measure spaces. There are several interesting possible variations on this theme, and we only discuss here the strongest possible version.

We start with a locally compact separable metric space M equipped with a Radon measure μ such that any open relatively compact nonempty set has positive measure. The original metric will not play any important role.

In addition, we are given a symmetric bilinear form $\mathcal{E}$ defined on a dense subset $\mathcal{D}(\mathcal{E})$ of $L^2(M, d\mu)$ such that $(u, u) \geq 0$ for any $u \in \mathcal{D}(\mathcal{E})$. We assume that $\mathcal{E}$ is closed, i.e., $\mathcal{D}(\mathcal{E})$ equipped with the norm

$$\mathcal{E}_1(u, u)^{1/2} = \sqrt{\|u\|_2^2 + \mathcal{E}(u, u)}$$

is complete (i.e., is a Hilbert space). In addition, we assume that the unit contraction

$$u \mapsto v_u = \inf\{1, \sup\{0, u\}\}$$

operates on $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ in the sense that

$$u \in \mathcal{D}(\mathcal{E}) \implies v_u \in \mathcal{D}(\mathcal{E}) \text{ and } \mathcal{E}(v_u, v_u) \leq \mathcal{E}(u, u).$$

Such a form is called a *Dirichlet form* and is associated with a self-adjoint strongly continuous semigroup of contractions H_t , $t > 0$, on $L^2(M, d\mu)$ with the additional property that $0 \leq u \leq 1$ implies $0 \leq H_t u \leq 1$. Namely, if A is the infinitesimal generator so that $H_t = e^{tA}$ (in the sense of spectral theory, say), then $\mathcal{D}(\mathcal{E}) = \text{Dom}((-A)^{1/2})$ and $\mathcal{E}(u, v) = \langle (-A)^{1/2}u, (-A)^{1/2}v \rangle$.

We assume that the form $\mathcal{E}$ is strongly local, i.e., $\mathcal{E}(u, v) = 0$ if $u, v \in \mathcal{D}(\mathcal{E})$ have compact support and v is constant on a neighborhood of the support of u . Finally, we assume that $(M, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ is regular. This means that the space $\mathcal{C}_c(M)$ of continuous compactly supported functions on M has the property that $\mathcal{D}(\mathcal{E}) \cap \mathcal{C}_c(M)$ is dense in $\mathcal{C}_c(M)$ in the sup norm $\|u\|_\infty = \sup_M \{|u|\}$ and is

dense in $\mathcal{D}(\mathcal{E})$ in the norm $\mathcal{E}_1^{1/2}$. Note that this is a hypothesis that concerns the interaction between $\mathcal{E}$ and the topology of M . We call $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ a *strictly local regular Dirichlet space*.

Under these hypotheses, there exists a bilinear form Γ defined on $\mathcal{D}(\mathcal{E}) \times \mathcal{D}(\mathcal{E})$ with the values in signed Radon measures on M such that

$$\mathcal{E}(u, v) = \int_M d\Gamma(u, v).$$

For $u \in \mathcal{D}(\mathcal{E}) \cap L^\infty(M)$, $\Gamma(u, u)$ is defined by

$$\forall \varphi \in \mathcal{D}(\mathcal{E}) \cap \mathcal{C}_c(M), \quad \int_M \varphi d\Gamma(u, u) = \mathcal{E}(u, \varphi u) - (1/2)\mathcal{E}(u^2, \varphi).$$

Although the measure $\Gamma(u, v)$ might be singular with respect to μ , it behaves much like $g(\nabla u, \nabla v)dx$ on a Riemannian manifold. For instance, versions of the chain rule and Leibnitz rule apply. In what follows, we work under

additional assumptions that imply that the set of those u in $\mathcal{D}(\mathcal{E})$ such that $d\Gamma/d\mu$ exists is rich enough (see [10, 97] for further details).

We now introduce a key ingredient to our discussion: the intrinsic distance.

Definition 3.1. Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a strictly local regular Dirichlet space. For x, y in M we set

$$\rho(x, y) = \sup\{u(x) - u(y) : u \in \mathcal{D}(\mathcal{E}) \cap \mathcal{C}_c(M), d\Gamma(u, u) \leq d\mu\}.$$

Here, the condition $d\Gamma(u, u) \leq d\mu$ means that the measure $\Gamma(u, u)$ is absolutely continuous with respect to μ with Radon–Nykodim derivative bounded by 1 almost everywhere. It is obvious that ρ is symmetric in x, y and satisfies the triangle inequality. It might well be either 0 or ∞ for some x, y . If ρ is finite and $\rho(x, y) = 0$ only if $x = y$, then ρ is a distance function.

Qualitative hypotheses.

Throughout the paper, we assume that

(A1) The function $\rho : M \times M \rightarrow [0, \infty]$ is finite, continuous, satisfies

$$\rho(x, y) = 0 \Rightarrow x = y,$$

and defines the topology of M .

(A2) The metric space (M, ρ) is a complete metric space.

With these hypotheses, one can show that the metric space (M, ρ) is a length space (i.e., $\rho(x, y)$ can be computed as the minimal length of continuous curves joining x to y , where the length of a curve is defined using ρ in a natural manner). Denote by $B(x, r)$ the open balls in (M, ρ) . Each $B(x, r)$ is precompact with compact closure given by the associated closed ball. Set $V(x, r) = \mu(B(x, r))$. For each fixed $x \in M$, $r > 0$ the function $\delta(y) = \max\{0, r - \rho(x, y)\}$ is in $\mathcal{D}(\mathcal{E}) \cap \mathcal{C}_c(M)$ and satisfies $d\Gamma(\delta, \delta) \leq d\mu$ (see [10, 11, 12, 94, 95, 96, 97] for details).

3.3 Local weak solutions of the Laplace and heat equations

Recall that A is the infinitesimal generator of the semigroup of operators associated to our Dirichlet form. Identify $L^2(M, \mu)$ with its dual using the scalar product.

Let V be a nonempty open subset of M . Consider the subspace $\mathcal{F}_c(V) \subset \mathcal{D}(\mathcal{E})$ of those functions with compact support in V . Note that $\mathcal{F}_c(V) \subset$

$\mathcal{D}(\mathcal{E}) \subset L^2(M, \mu)$ and consider their duals $L^2(X, \mu) \subset \mathcal{D}(\mathcal{E})' \subset \mathcal{F}_c(V)'$. We use the brackets $\langle \cdot, \cdot \rangle$ to denote duality pairing between these spaces. Let $\mathcal{F}_{\text{loc}}(V)$ be the space of functions $u \in L^2_{\text{loc}}(V)$ such that for any compact set $K \subset V$ there exists a function $u_K \in \mathcal{D}(\mathcal{E})$ that coincides with u almost everywhere on K .

Definition 3.2. Let V be a nonempty open subset of X . Let $f \in \mathcal{F}_c(V)'$. A function $u : V \mapsto \mathbb{R}$ is a *weak* (local) solution of $Au = f$ in V if

1. $u \in \mathcal{F}_{\text{loc}}(V)$;
2. for any function $\varphi \in \mathcal{F}_c(V)$ we have $\mathcal{E}(\varphi, u) = \langle \varphi, f \rangle$.

Remark 3.1. If f can be represented by a locally integrable function in V and u is such that there exists a function $u^* \in \text{Dom}(A)$ (the domain of the infinitesimal generator A) satisfying $u = u^*|_V$, then u is a weak local solution of $Au = f$ if and only if $Au^*|_V = f$ a.e in V .

Remark 3.2. The notion of weak local solution defined above may contain implicitly a Neumann type boundary condition if M has a natural boundary. Consider, for example, the case where M is the closed upper-half plane $P_+ = \overline{\mathbb{R}^2_+}$ equipped with its natural Dirichlet form

$$\mathcal{E}(f, f) = \iint_{\mathbb{R}^2_+} \left(\left| \frac{\partial f}{\partial x} \right|^2 + \left| \frac{\partial f}{\partial y} \right|^2 \right) dx dy, \quad f \in W^1(\mathbb{R}^2_+).$$

Let $V = \{z = (x, y) : x^2 + y^2 < 1; y \geq 0\} \subset P_+$. Note that V is open in P_+ . Let u be a local weak solution of $\Delta u = 0$ in V . Then it is easy to see that u is smooth in V and must have vanishing normal derivative along the segment $(-1, 1)$ of the real axis.

Next, we discuss local weak solutions of the heat equation $\partial_t u = Au$ in a time-space cylinder $I \times V$, where I is a time interval and V is a nonempty open subset of X . Given a Hilbert space H , let $L^2(I \rightarrow H)$ be the Hilbert space of functions $v : I \mapsto H$ such that

$$\|v\|_{L^2(I \rightarrow H)} = \left(\int_I \|v(t)\|_H^2 dt \right)^{1/2} < \infty.$$

Let $W^1(I \rightarrow H) \subset L^2(I \rightarrow H)$ be the Hilbert space of those functions $v : I \mapsto H$ in $L^2(I \rightarrow H)$ whose distributional time derivative v' can be represented by functions in $L^2(I \rightarrow H)$, equipped with the norm

$$\|v\|_{W^1(I \rightarrow H)} = \left(\int_I (\|v(t)\|_H^2 + \|v'(t)\|_H^2) dt \right)^{1/2} < \infty.$$

Given an open time interval I , we set

$$\mathcal{F}(I \times X) = L^2(I \rightarrow \mathcal{D}(\mathcal{E})) \cap W^1(I \rightarrow \mathcal{D}(\mathcal{E}')).$$

Given an open time interval I and an open set $V \subset X$ (both nonempty), let

$$\mathcal{F}_{\text{loc}}(I \times V)$$

be the set of all functions $v : I \times V \rightarrow \mathbb{R}$ such that for any open interval $I' \subset I$ relatively compact in I and open subset V' relatively compact in V there exists a function $u^\# \in \mathcal{F}(I \times X)$ satisfying $u = u^\#$ a.e. in $I' \times V'$. Finally, let

$$\mathcal{F}_c(I \times V) = \{v \in \mathcal{F}(I \times X) : v(t, \cdot) \text{ has compact support in } V \text{ for a.a. } t \in I\}.$$

Definition 3.3. Let I be an open time interval. Let V be an open subset in X , and let $Q = I \times V$. A function $u : Q \rightarrow \mathbb{R}$ is a *weak* (local) solution of the heat equation $(\partial_t - A)u = 0$ in Q if

1. $u \in \mathcal{F}_{\text{loc}}(Q)$;
2. for any open interval J relatively compact in I and $\varphi \in \mathcal{F}_c(Q)$

$$\int_J \int_V \varphi \partial_t u d\mu dt + \int_J \mathcal{E}(\varphi(t, \cdot), u(t, \cdot)) dt = 0.$$

As noted in the elliptic case, this definition may contain implicitly some Neumann type boundary condition along a natural boundary of X (see [94, 96] for a detailed discussion).

3.4 Harnack type Dirichlet spaces

The following is the main definition of this section.

Definition 3.4. We say that a regular strictly local Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(M, \mu)$ is of *Harnack type* if the distance ρ satisfies the qualitative conditions (A1), (A2), and the following scale invariant parabolic Harnack inequality holds. There exists a constant C such that for any $z \in M$, $r > 0$ and weak nonnegative solution u of the heat equation $(\partial_t - A)u = 0$ in $Q = (s - 4r^2, s) \times B(z, 2r)$ we have

$$\sup_{(t,x) \in Q_-} u(t, x) \leq C \inf_{(t,x) \in Q_+} u(t, x), \quad (3.1)$$

where $Q_- = (s - 3r^2, s - 2r^2) \times B(z, r)$, $Q_+ = (s - r^2, s) \times B(z, r)$ and both sup and inf are essential, i.e. are computed up to sets of measure zero.

Any Harnack type Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ obviously satisfies the following elliptic Harnack inequality (with the same constant C as in (3.1)). For any $z \in X$ and $r > 0$ and weak nonnegative solution u of the equation $Lu = 0$ in $B(z, 2r)$ we have

$$\sup_{B(z,r)} u \leq C \inf_{B(z,r)} u. \quad (3.2)$$

This elliptic Harnack inequality is weaker than its parabolic counterpart.

One of the simple, but important consequences of the Harnack inequality (3.1) is the following quantitative Hölder continuity estimate (see, for example, [87, Theorem 5.4.7] and [94]).

Theorem 3.1. *Assume that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a Harnack type Dirichlet form on $L^2(M, \mu)$. Then there exists $\alpha \in (0, 1)$ and $A > 0$ such that any local (weak) solution of the heat equation $(\partial_t - A)u = 0$ in $Q = (s - 4r^2, s) \times B(x, 2r)$, $x \in X$, $r > 0$ has a continuous representative and satisfies*

$$\sup_{(t,y),(t',y') \in Q'} \left\{ \frac{|u(y,t) - u(y',t')|}{[|t - t'|^{1/2} + \rho_{\mathcal{E}}(y, y')]^{\alpha}} \right\} \leq \frac{A}{r^{\alpha}} \sup_Q |u|,$$

where $Q' = (s - 3r^2, s - r^2) \times B(x, r)$.

A crucial consequence of this is that, on a Harnack type Dirichlet space, local weak solutions of the Laplace equation or the heat equation are continuous functions (in the sense that they admit a continuous representative).

Definition 3.5. Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular strictly local Dirichlet space satisfying the qualitative conditions (A1), (A2).

- We say that the *doubling volume property* holds if there is a constant D_0 such that $V(x, 2r) \leq D_0 V(x, r)$ for all $x \in M$ and $r > 0$.
- We say that the *scale invariant L^2 Poincaré inequality* holds if there is a constant P_0 such that for any ball $B = B(x, r)$ in (M, ρ)

$$\forall u \in \mathcal{F}_{\text{loc}}(B(x, r)), \quad \int_B |u - u_B|^2 d\mu \leq P_0 r^2 \int_B d\Gamma(u, u),$$

where u_B denotes the average of u over B .

- We say that these properties hold *uniformly at small scales* if they hold under the restriction that $r \in (0, 1)$.

We can now state the main result of this section which is a direct generalization of Theorem 2.2 in the setting of strictly local regular Dirichlet spaces (see [94]).

Theorem 3.2. *Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular strictly local Dirichlet space satisfying the qualitative conditions (A1), (A2). The following properties are equivalent.*

- The space $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a Harnack type Dirichlet space.
- The doubling volume property and the scale invariant Poincaré inequality are satisfied on $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$.
- The heat semigroup e^{tA} admits a transition kernel $h(t, x, y)$ satisfying the two-sided bound

$$\frac{c}{V(x, \sqrt{t})} \exp\left(-A \frac{\rho(x, y)^2}{t}\right) \leq h(t, x, y) \leq \frac{C}{V(x, \sqrt{t})} \exp\left(-a \frac{\rho(x, y)^2}{t}\right)$$

for constants $0 < a, A, c, C < \infty$.

As in the classical case, if one uses Moser's iteration techniques, one of the first steps of the proof that the doubling property and Poincaré inequality imply the parabolic Harnack inequality is that they imply the family of Sobolev inequalities

$$\forall f \in \mathcal{F}_c(B), \quad \left(\int_B |f|^{2q} d\mu \right)^{1/q} \leq \frac{C_M r^2}{\mu(B)^{2/\nu}} \left(\int_B d\Gamma(f, f) + \int_B r^{-2} |f|^2 d\mu \right) \quad (3.3)$$

for some $q > 1$ and $\nu > 2$ related to q by $q = \nu/(\nu - 2)$. This inequality implies the volume estimate

$$\forall x \in M, r > s, 0, \quad V(x, r) \leq C(r/s)^\nu V(x, s).$$

Furthermore, a precise analog of Theorem 2.1 holds in this setting, as well as the following version of Theorem 2.3 (see [61]).

Theorem 3.3. *Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular strictly local Dirichlet space satisfying the qualitative conditions (A1), (A2) and (3.3). The following properties are equivalent.*

- The scale invariant L^2 Poincaré inequality.
- The scale invariant elliptic Harnack inequality.
- The scale invariant parabolic Harnack inequality.

3.5 Imaginary powers of $-A$ and the wave equation

This section is merely a pointer to some interesting related results and literature regarding the wave equation. In the classical setting of $\mathbb{R}^n$, the wave equation is the PDE $(\partial_t^2 - \Delta)u = 0$. One of its main properties is the finite propagation speed property which asserts that if a solution u has support in the ball $B(x_0, r_0)$ at time t_0 , then, at time t , its has support in

$B(x_0, r_0 + (t - t_0))$. Although this property can be proved in a number of elegant ways in $\mathbb{R}^n$, its generalization to other settings is not quite straightforward. Basic solutions of the wave equation can be obtained as follows. Using Fourier transform, consider the operator $\cos(t\sqrt{-\Delta})$ acting on $L^2(\mathbb{R}^n)$. Then for any smooth φ with compact support

$$u(t, \cdot) = \cos(t\sqrt{-\Delta})\varphi$$

is a solution of the wave equation with $u(0, \cdot) = \varphi$. This construction generalizes using spectral theory to any (nonpositive) self-adjoint operator, in particular, to the infinitesimal generator A of a Markov semigroup associated with a strictly local regular Dirichlet space $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$. In this general setting, it is not entirely clear how to discuss the finite speed propagation property of the wave equation

$$(\partial_t^2 - A)u = 0.$$

Given a distance function d on $M \times M$ (assumed, at the very least, to be a measurable function on $M \times M$), one says that the wave equation (associated to A) has unit propagation speed with respect to d if for any functions $u_1, u_2 \in L^2(M, \mu)$ compactly supported in S_1, S_2 , respectively, with

$$d(S_1, S_2) = \min\{d(s_1, s_2) : s_1 \in S_1, s_2 \in S_2\} > t$$

we have

$$\langle \cos(t\sqrt{-A})u_1, u_2 \rangle_\mu = 0.$$

The following theorem follows from the techniques and results in [91, 92].

Theorem 3.4. *Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular strictly local Dirichlet space satisfying the qualitative conditions (A1), (A2). Then the associated wave equation has unit propagation speed with respect to the distance ρ introduced in Definition 3.1.*

This result plays an important role in the study of continuity properties on L^p spaces of various operators defined via spectral theory by the functional calculus formula

$$m(-A) = \int_0^\infty m(\lambda) dE_\lambda,$$

where E_λ stands for a spectral resolution of the self-adjoint operator $-A$. This formula defines a bounded operator on $L^2(M, \mu)$ for any bounded function m . The question then is to examine what further properties of m imply additional continuity properties of $m(-A)$. The finite speed propagation property is very helpful in the study of these questions. We refer the reader to [37, 38, 92], where earlier references and detailed discussions of the literature can be found. As an illustrative example, we state the following result. For a

function m defined on $[0, \infty)$ we set $m_t(u) = m(tu)$ and $\|m\|_{(s)} = \|(I - (d/du)^2)^{s/2}m\|_\infty$.

Theorem 3.5. *Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular strictly local Dirichlet space satisfying the qualitative conditions (A1), (A2). Assume that the Sobolev inequality (3.3) holds for some $q > 1$ and ν given by $q = \nu/(\nu - 2)$.*

Fix a function $\eta \in \mathcal{C}_c^\infty((0, \infty))$, not identically 0. If m is a bounded function such that

$$\sup_{t>0} \|\eta m_t\|_{(s)} < \infty$$

for some $s > \nu/2$, then the operator $m(-A)$ is bounded on $L^p(M, \mu)$ for each $p \in (1, \infty)$. The operators $(-A)^{i\alpha}$, $\alpha \in \mathbb{R}$, are all bounded on $L^p(M, \mu)$, $1 < p < \infty$, and there exists a constant C such that the norm of $(-A)^{i\alpha}$ on $L^p(M, \mu)$ is at most $C(1 + |\alpha|)^{\nu/2}$, for all $\alpha \in \mathbb{R}$ and $1 < p < \infty$.

3.6 Rough isometries

One of the strengths of the techniques and results discussed in this paper is their robustness. In the present context, the idea of rough isometry was introduced by Kanai [64, 66, 65] and developed further in [32]. It has also been made very popular by the work of M. Gromov. Note that rough isometries as defined below do not preserve the small scale structure of the space.

Definition 3.6. Let (M_i, ρ_i, μ_i) , $i = 1, 2$, be two measure metric spaces. We say that they are *roughly isometric* (or *quasiisometric*) as metric measure spaces if there are two maps $\varphi_k : M_i \rightarrow M_j$, $k = (i, j) \in \{(1, 2), (2, 1)\}$ and a constant A such that for $k' = (j, i)$ we have the following.

1. $\forall x \in M_i, \rho_i(x, \varphi_{k'} \circ \varphi_k(x)) \leq A$.
2. $M_j = \{y \in M_j : \rho_j(y, \varphi_k(M_i)) \leq A\}$.
3. $\forall x, x' \in M_i, A^{-1}(\rho_i(x, x') - A) \leq \rho_j(\varphi_k(x), \varphi_k(x')) \leq A(1 + \rho_i(x, x'))$.
4. $\forall x \in M_i, A^{-1}V(x, 1) \leq V(\varphi_k(x), 1) \leq AV(x, 1)$.

Condition 3 requires that each of the maps φ_k roughly preserves large enough distances (larger than $2A$, say). Condition 2 requires that each of the maps φ_k is almost surjective, in a quantitative metric sense. The first condition says that the maps φ_k and $\varphi_{k'}$ are almost inverse of each other. The last condition concerns volume transport and is obviously specific to the setting of *measure* metric spaces. This definition is nicely symmetric (as an equivalence relation should be!), but is redundant. It is enough to require the existence of one map, say from M_1 to M_2 with the last three properties. The existence of an almost inverse with the desired properties follows from the axiom of choice.

The relevance of rough isometries in the study of Harnack type Dirichlet space lies in the following stability theorem from [32, Theorem 8.3] (although [32] does not explicitly cover the setting of Dirichlet spaces, the same proof applies).

Theorem 3.6. *Let $(M_i, \mu_i, \mathcal{E}_i, \mathcal{D}(\mathcal{E}_i))$, $i = 1, 2$, be two regular strictly local Dirichlet spaces satisfying the qualitative conditions (A1), (A2). Assume further that these two spaces satisfy the volume doubling property and the L^2 Poincaré inequality, uniformly at small scales. If (M_1, ρ_1, μ_1) and (M_2, ρ_2, μ_2) are roughly isometric as metric measure spaces, then $(M_1, \mu_1, \mathcal{E}_1, \mathcal{D}(\mathcal{E}_1))$ is of Harnack type if and only if $(M_2, \mu_2, \mathcal{E}_2, \mathcal{D}(\mathcal{E}_2))$ is of Harnack type.*

Example 3.1. In a sense, the following example illustrates in the simplest non-trivial possible way the results of this section. Consider the two-dimensional cubical complex obtained as the subset M of $\mathbb{R}^3$ of those point (x, y, z) with at least one coordinate in $\mathbb{Z}$. In other words, M is the union of the planes $\{x = k\}$, $\{y = k\}$, $\{z = k\}$, $k \in \mathbb{Z}$. It is also the union $M = \bigcup_{\mathbf{k}} Q_{\mathbf{k}}$, where $Q_{\mathbf{k}}$ is the two-dimensional boundary of the unit cube with lower left back corner $\mathbf{k} \in \mathbb{Z}^3$. This space is equipped with its natural measure μ (Lebesgue measure on each of the planes above). To describe the natural Dirichlet form and its domain, we recall that if F is a face on a unit cube Q_k and if a function f in $L^2(F)$ has distributional first order partial derivatives in $L^2(F)$ (i.e., is in the Sobolev space $H^1(F)$), then the trace of f along the one dimensional edges of the face F are well defined, say, as an L^2 function on the edges. Taking into account this remark, we set (the factor of 1/2 is to account for the appearance of each face in exactly two cubes)

$$\mathcal{E}(f, g) = \frac{1}{2} \sum_{\mathbf{k}} \int_{Q_{\mathbf{k}}} \nabla f \cdot \nabla g \, d\mu$$

for all $f, g \in \mathcal{D}(\mathcal{E})$, where $\mathcal{D}(\mathcal{E})$ is the space of those functions $f \in L^2(M)$ which have distributional first order partial derivatives in $L^2(F)$ on each face F of any cube $Q_{\mathbf{k}}$, satisfy $\mathcal{E}(f, f) < \infty$, and have the property that for each pair of faces F_1, F_2 sharing an edge I , the restrictions of $f|_{F_1}$ and $f|_{F_2}$ to the edge I coincide. In the above formula, ∇f refers to the Euclidean gradient of f viewed as a function defined on each of the square faces of the cube $Q_{\mathbf{k}}$. Because of the above-mentioned trace theorem for Sobolev functions, it is easy to see that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a Dirichlet space. It is local, and one can show (although this is not entirely obvious) that it is regular (see, for example, [82]). The distance ρ associated to this Dirichlet form on M coincides with the natural shortest path distance on this cubical complex. It is not hard to check that

- The uniform small scale doubling property holds.
- The uniform small scale Poincaré inequality holds.

- The metric measure space (M, ρ, μ) is roughly isometric to $\mathbb{R}^3$.

Thus, from Theorem 3.6 it follows that this Dirichlet space is a Harnack type Dirichlet space.

4 Flat Sobolev Inequalities

In the previous sections, we discussed the role of the family of localized Sobolev inequalities (2.11) in Moser's iteration and related techniques. In some sense, the need to consider (2.11) instead of the more classical inequality (2.10) comes from looking at situations that are inhomogeneous either at the level of location or at the level of scales, or both. Because of this one sometimes refers to a global Sobolev inequality that do not require localization as a “flat” Sobolev inequality. For instance, one might ask: What complete n -dimensional Riemannian manifolds satisfy a Sobolev inequality of the form

$$\forall f \in \mathcal{C}_c(M), \quad \|f\|_{2n/(n-2)} \leq S \|\nabla f\|_2?$$

It turns out that this inequality is satisfied by a variety of manifolds not having much in common with each others, including manifolds with nonnegative Ricci curvature and maximal volume growth, as well as simply connected manifolds with nonpositive sectional curvature (see [60, Theorem 8.3] and the references therein for this result).

In this section, we discuss such inequalities: how to prove them and what are they good for?

4.1 How to prove a flat Sobolev inequality?

There are many interesting approaches to proving Sobolev inequalities and we will, essentially, discuss only one of them here. One useful aspect of this approach is its robustness. One weakness, among others, is that it never produces best constants.

Definition 4.1. Let (M, g) be a complete Riemannian manifold. We say that it satisfies an L^p pseudo-Poincaré inequality if

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \|f - f_r\|_p \leq Ar \|\nabla f\|_p$$

for all $r > 0$, where f_r is a function such that $f_r(x)$ is the average of f over the ball $B(x, r)$.

Theorem 4.1 ([4, Theorem 9.1]). *Let (M, g) be a complete Riemannian manifold satisfying the L^p pseudo-Poincaré inequality. Assume that there exists*

$N > 0$ such that $V(x, r) \geq cr^N$ for all $x \in M$ and $r > 0$. Then the inequality

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \int_M |f|^{p(1+1/N)} d\mu \leq C(M, p) \left(\int_M |\nabla f|^p d\mu \right)^{p/N} \left(\int_M |f| d\mu \right)^{p/N}$$

holds. If $N > p$, then

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \|f\|_{pN/(N-p)} \leq S(M, p) \|\nabla f\|_p.$$

Remark 4.1. The paper [4] shows that a great number of other interesting Sobolev type inequalities follow as a corollary of the above result.

Remark 4.2. The above definition and theorem hold unchanged for $p = 2$ in the context of strictly local regular Dirichlet spaces satisfying the qualitative conditions (A1), (A2).

Remark 4.3. The volume condition $V(x, r) \geq cr^N$ is sharp in the sense that it follows from the validity of any of the two stated inequalities.

Remark 4.4. The same result holds if one replaces f_r in the pseudo-Poincaré inequality by $M_r f$ and replaces the volume hypothesis by $\|M_r f\|_\infty \leq Cr^{-N} \|f\|_1$. For instance, M_r could be averages over sets different from balls or some more sophisticated operators. As an example, let $M_r = H_{r^2}$ be the heat semigroup on (M, g) at time $t = r^2$. Then, if one knows that for all $t > 0$, $\|f - H_t f\|_p \leq C\sqrt{t} \|\nabla f\|_p$ and $\|H_t\|_{1 \rightarrow \infty} \leq Ct^{-N/2}$, then one can conclude that the inequalities stated in the above theorem hold on M . The first of these two hypotheses is always satisfied if $p = 2$.

Example 4.1. Riemannian manifolds with nonnegative Ricci curvature satisfy the pseudo-Poincaré inequality of Definition 4.1 for any $1 \leq p \leq \infty$. They satisfy the Sobolev inequality

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \|f\|_{pN/(N-p)} \leq S(M, p) \|\nabla f\|_p$$

if and only if $V(x, r) \geq cr^N$ and $N > p \geq 1$ (see [87, Sect. 3.3.5]). On these manifolds, the volume is bounded by $V(x, r) \leq C_n r^n$, where n is the topological dimension. Hence $V(x, r) \geq cr^N$ for all $r > 0$ is possible only if $N = n$ and $V(x, r) \simeq r^n$.

Example 4.2. Let (M, g) be a connected unimodular Lie group equipped with a left-invariant Riemannian metric. Then the pseudo-Poincaré inequality of Definition 4.1 holds for any $1 \leq p \leq \infty$ (see [31] or [87, 3.3.4]). The inequality

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \|f\|_{pN/(N-p)} \leq S(M, p) \|\nabla f\|_p$$

holds if and only if $V(r) \geq cr^N$ for all $r > 0$. For instance, if M is the group of upper-triangular 3 by 3 matrices with 1's on the diagonal (i.e., the Heisenberg

group), then for any left-invariant Riemannian metric, $V(x, r) \geq cr^N$ for all $r > 0$ and $N \in [3, 4]$.

4.2 Flat Sobolev inequalities and semigroups of operators

Sobolev inequalities can be generalized in useful ways in many contexts one of which involves the infinitesimal generator A of a strongly continuous semigroup of operator e^{tA} jointly defined on the spaces $L^p(M, \mu)$, $1 \leq p < \infty$. One of the most straightforward results in this context is the following theorem from [26] which extends an earlier result of Varopoulos [100] (see also [103]). For $\alpha > 0$ we set

$$(-A)^{-\alpha/2} = \Gamma(\alpha/2)^{-1} \int_0^\infty t^{-1+\alpha/2} e^{tA} dt.$$

Theorem 4.2. *Fix $p \in (1, \infty)$. Assume that e^{tA} is a bounded holomorphic semigroup of operator on $L^p(M, \mu)$ which extends as an equicontinuous semigroup on both $L^1(M, \mu)$ and $L^\infty(M, \mu)$. Then for any $N > 0$ the following two properties are equivalent.*

- *There exists C_1 such that*

$$\forall f \in L^1(M, \mu), \quad \|e^{tA} f\|_\infty \leq C_1 t^{-N/2} \|f\|_1.$$

- *There exists C_2 such that for one pair (equivalently, for all pairs) (α, q) with $0 < \alpha p < N$ and $1/q = 1/p - \alpha/N$, we have*

$$\forall f \in L^p(M, \mu), \quad \|(-A)^{-\alpha/2} f\|_q \leq C_2 \|f\|_p.$$

Remark 4.5. The first property is known as a form of ultracontractivity (boundedness of e^{tA} from L^1 to L^∞ for all $t > 0$). The second property states that a Sobolev type inequality holds, namely, $\|f\|_q \leq C_2 \|(-A)^{\alpha/2} f\|_p$, $f \in \text{Dom}((-A)^{\alpha/2})$.

Remark 4.6. A semigroup e^{tA} is bounded holomorphic on $L^p(X, \mu)$ if

$$t \|A e^{tA} f\|_p \leq C \|f\|_p$$

for all $f \in L^p(M, \mu)$ and $t > 0$. This implies that for any $\alpha \in (0, 1]$ and f in the domain of $(-A)^{\alpha/2}$

$$\forall t > 0, \quad \|f - e^{tA} f\|_p \leq C_\alpha t^{\alpha/2} \|(-A)^{\alpha/2} f\|_p.$$

This can be viewed as a form of pseudo-Poincaré inequality.

Theorems such as Theorem 4.2 apply nicely in the context of Dirichlet spaces because the associated semigroups are self-adjoint on $L^2(M, \mu)$ and contract each $L^p(M, \mu)$, $1 \leq p \leq \infty$. Semigroups of self-adjoint contractions on $L^2(M, \mu)$ are automatically bounded holomorphic on $L^2(M, \mu)$. Moreover, in the regular strictly local Dirichlet space context described earlier, the generator A is related to the form $\mathcal{E}$ and the energy form Γ by

$$\|(-A)^{1/2}f\|_2^2 = \mathcal{E}(f, f) = \int_M d\Gamma(f, f), \quad f \in \text{Dom}((-A)^{1/2}) = \mathcal{D}(\mathcal{E}).$$

For the following result see [15, 100, 103] and also [33].

Theorem 4.3. *Fix $N > 0$. Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a Dirichlet space with associated semigroup e^{tA} . The following properties are equivalent.*

- *There exists C_1 such that*

$$\forall f \in L^1(M, \mu), \quad t > 0, \quad \|e^{tA}f\|_\infty \leq C_1 t^{-N/2} \|f\|_1.$$

- *For one (equivalently, all) (α, q) with $1 < \alpha < N/2$ and $q = 2N/(N - 2\alpha)$ there exists $C(\alpha)$ such that*

$$\forall f \in \text{Dom}((-A)^{\alpha/2}), \quad \|f\|_q \leq C(\alpha) \|(-A)^{\alpha/2}f\|_2.$$

- *There exists C_2 such that*

$$\forall f \in L^1(M, \mu) \cap \mathcal{D}(\mathcal{E}), \quad \|f\|_2^{2(1+2/N)} \leq C_2 \mathcal{E}(f, f) \|f\|_1^{4/N}.$$

Remark 4.7. The first property is a particular type of ultracontractivity. The second property is a Sobolev type inequality. If $N > 2$, one can take $\alpha = 1$, $q = 2N/(N - 2)$ and the inequality takes the form $\|f\|_q \leq C_1 \mathcal{E}(f, f)^{1/2}$. The third property is a Nash inequality.

Example 4.3. Let (M, g) be an n -dimensional complete Riemannian manifold that is simply connected and has nonpositive sectional curvature. By a simple comparison argument (see, for example, [20, Theorem 6] and the references therein), the heat kernel on M is bounded from above by the Euclidean heat kernel. In particular, for all $t > 0$,

$$\sup_{x, y \in M} \{h(t, x, y)\} \leq c_n t^{-n/2}.$$

This implies that for all $t > 0$ we have $\|e^{t\Delta_M}f\|_\infty \leq c_n t^{-n/2} \|f\|_1$. Hence the above theorem gives the Sobolev inequality

$$\forall f \in C_c^\infty(M), \quad \|f\|_{2n/(n-2)} \leq S_M \|(-\Delta)^{1/2}f\|_2.$$

Of course, $\|(-\Delta)^{1/2}f\|_2 = \|\nabla f\|_2$, so that this inequality can be written as

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \|f\|_{2n/(n-2)} \leq S_M \|\nabla f\|_2.$$

There is an open conjecture that this inequality should hold with S_M being the same constant as in the Euclidean n -space.

The first property in Theorem 4.3 obviously calls for a more general formulation. The following general elegant result was obtained by Coulhon [27] (after many attempts by different authors). A smooth positive function Φ defined on $[0, \infty)$ satisfies condition (D) if there exists $\varepsilon \in (0, 1)$ such that $\varphi'(s) \geq \varepsilon \varphi'(t)$ for all $t > 0$ and $s \in [t, 2t]$, where $\varphi(s) = -\log \Phi(s)$.

Theorem 4.4 ([27]). *Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a Dirichlet space with associated semigroup e^{tA} . Let Φ be a positive smooth decreasing function on $[0, \infty)$ satisfying condition (D), and let $\Theta = -\Phi' \circ \Phi^{-1}$. The following properties are equivalent.*

- *There exists a constant $c_1 \in (0, \infty)$ such that*

$$\forall f \in L^1(M, \mu), \quad t > 0, \quad \|e^{tA}f\|_\infty \leq \Phi(c_1 t) \|f\|_1.$$

- *There exists a constant $C_1 \in (0, \infty)$ such that for all $f \in L^1(M, \mu) \cap \mathcal{D}(\mathcal{E})$ with $\|f\|_1 \leq 1$ we have*

$$\Theta(\|f\|_2^2) \leq C\mathcal{E}(f, f).$$

We refer the reader to [4, 8, 9, 27] for explicit examples and further results.

4.3 The Rozenblum–Cwikel–Lieb inequality

One of the surprising aspects of the Sobolev inequality

$$\|f\|_{2N/(N-2)}^2 \leq S^2 \mathcal{E}(f, f)$$

is how many different equivalent form it takes (hence the title “Sobolev inequalities in disguise” of [4]). Despite the equivalence of these different forms, some appear “stronger” than others. For instance, on one hand, deducing from the above inequality the Nash inequality

$$\|f\|_2^{2(1+2/N)} \leq S^2 \mathcal{E}(f, f) \|f\|_1^{4/N}$$

only involves a simple use of Hölder’s inequality (and the constant remains the same). On the other hand, recovering the Sobolev inequality from its Nash form involves some more technical arguments. The constant S changes

in the process (the two inequalities in $\mathbb{R}^N$ have different best constants) and one needs to assume that $N > 2$.

In 1972, Rozenblum proved a remarkable spectral inequality showing that, in $\mathbb{R}^N$ with $N \geq 3$, if V is a nonnegative measurable function and $\mathcal{N}_-(-\Delta - V)$ denotes the number of negative eigenvalues of $-\Delta - V$, then there exists a constant $C(N)$ such that

$$\mathcal{N}_-(-\Delta - V) \leq C(N) \int_{\mathbb{R}^N} V(x)^{N/2} dx.$$

Very different proofs were later given by Cwikel and by Lieb, and this inequality is known as the *Rozenblum–Cwikel–Lieb inequality*. We refer the reader to the review of the literature in [70, 84]. The following elegant result is taken from [70] and is based on the technique used in [73] in Euclidean space.

Theorem 4.5 ([70]). *Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a Dirichlet space with associated semigroup e^{tA} . Assume that the Sobolev inequality*

$$\forall f \in \mathcal{D}(\mathcal{E}), \quad \|f\|_{2N/(N-2)}^2 \leq S^2 \mathcal{E}(f, f)$$

holds for some $N > 2$. Then for any measurable function $V \geq 0$

$$\mathcal{N}_-(-A - V) \leq C(N) \int_M V^{N/2} d\mu.$$

In [84], this result is generalized in a number of useful ways. In particular, the following version related to Theorem 4.4 is obtained.

Theorem 4.6 ([84]). *Fix a nonnegative convex function Q on $[0, \infty)$, growing polynomially at infinity and vanishing in a neighborhood of 0. Set*

$$q(u) = \int_0^\infty v^{-1} Q(v) e^{-v/u} dv.$$

Let $(M, \mu, \mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a Dirichlet space with associated semigroup e^{tA} . Assume that

$$\forall f \in L^1(M, \mu), \quad t > 0, \quad \|e^{tA} f\|_\infty \leq \Phi(t) \|f\|_1$$

with Φ continuous, integrable at infinity $\left(\int_0^\infty \varphi(t) dt < \infty\right)$, and satisfying $\Phi(t) = O(t^{-\alpha})$ at 0 for some $\alpha > 0$. Then for any measurable function V

$$\mathcal{N}_-(-A - V) \leq \frac{1}{q(1)} \int_0^\infty \left(\int_M Q(tV(x)) d\mu(x) \right) \frac{\Phi(t)}{t} dt.$$

Remark 4.8. One can take $Q(u) = (u - 1)_+$. In this case,

$$\int_0^\infty \left(\int_M Q(tV(x)) d\mu(x) \right) \frac{\Phi(t)}{t} dt \leq \int_M \left(V(x) \int_{1/V(x)}^\infty \Phi(t) dt \right) d\mu(x)$$

so that if

$$\Psi(u) = \int_u^\infty \Phi(t) dt,$$

then

$$\mathcal{N}_-(-A - V) \leq C \int_M V(x) \Psi(1/V(x)) d\mu(x).$$

In particular, if $\Phi(t) \simeq t^{-N/2}$, $t > 0$ for some $N > 2$, then $\Psi(u) \simeq u^{-N/2+1}$, $u > 0$, and

$$\mathcal{N}_-(-A - V) \leq C \int_M V^{N/2} d\mu.$$

Example 4.4. Let (G, g) be an amenable connected Lie group of topological dimension n equipped with a left-invariant Riemannian metric with Laplace operator Δ . In this case, there are two possible behaviors for the function Φ . If G has polynomial volume growth, then

$$\Phi(t) \simeq \begin{cases} t^{-n/2} & \text{for } t \in (0, 1], \\ t^{-N/2} & \text{for } t \in (1, \infty), \end{cases}$$

where N is some integer. If that is not the case, then G has exponential volume growth and

$$\Phi(t) \leq C \times \begin{cases} t^{-n/2} & \text{for } t \in (0, 1] \\ e^{-ct^{1/3}} & \text{for } t \in (1, \infty) \end{cases}$$

for some $c, C \in (0, \infty)$ (a similar lower bound holds as well).

In the case of polynomial volume growth, application of Theorem 4.6 requires $N > 2$. Assuming that $N > 2$, the function Ψ introduced in the above remark is given by $\Psi(u) \simeq u^{-N/2+1} \mathbf{1}_{u>1} + (1 + u^{-n/2+1}) \mathbf{1}_{u \leq 1}$. Hence

$$\mathcal{N}_-(-\Delta - V) \leq C \left(\int_{\{V \geq 1\}} V(1 + V^{n/2-1}) d\mu + \int_{\{V < 1\}} V^{N/2} d\mu \right).$$

In the case of exponential volume growth, one gets

$$\mathcal{N}_-(-\Delta - V) \leq C \int_{\{V \geq 1\}} V(1 + V^{n/2-1}) d\mu + C \int_{\{V < 1\}} e^{-cV^{-1/3}} d\mu.$$

In this case, since the volume growth is exponential, we see that for a smooth positive potential with $V(x) \simeq (1 + \rho(e, x))^{-\gamma}$, $\mathcal{N}_-(-\Delta - V)$ is finite if $\gamma > 3$.

4.4 Flat Sobolev inequalities in the finite volume case

Recall that a flat Sobolev inequality of the form $\forall f \in \mathcal{C}_c^\infty(M)$, $\|f\|_{2q} \leq S_M \|\nabla f\|_2$, with $q > 1$, on a complete Riemannian manifold (M, g) , implies that the volume grows at least as r^ν with $q = \nu/(\nu - 2)$. In particular, the volume of M cannot be finite. In order to allow for some finite volume manifolds, one needs to consider inequalities of the form

$$\forall f \in \mathcal{C}_c^\infty, \quad \|f\|_{2q}^2 \leq a_M \|f\|_2^2 + C_M^2 \|\nabla f\|_2^2. \quad (4.1)$$

If we assume that the volume of M is finite, we can normalize the measure so that $\mu(M) = 1$ and then it is easy to see that the above inequality can hold only if $a_M \geq 1$. Moreover, if the global Poincaré inequality $\|f - f_M\|_2 \leq A_M \|\nabla f\|_2$ holds for all $f \in \mathcal{C}^\infty(M)$, then (4.1) implies

$$\forall f \in \mathcal{C}_c^\infty, \quad \|f\|_{2q}^2 \leq \|f\|_2^2 + S_M^2 \|\nabla f\|_2^2. \quad (4.2)$$

The aim of this section is to point out a beautiful consequence of this inequality obtained by Bakry and Ledoux [5]. We refer the reader to [5] for a complete discussion and detailed references.

Theorem 4.7 ([5, Theorem 2]). *Assume that (M, g) is a complete Riemannian manifold with finite volume. Assume that, equipped with its normalized Riemannian measure, (M, g) satisfies (4.2) for some $q > 1$ and $S_M \in (0, \infty)$. Then M is compact with*

$$\text{Diam}(M) \leq \pi \frac{\sqrt{q}}{q-1} S_M.$$

This result is a form of a well-known theorem of Meyers that asserts that an n -dimensional Riemannian manifold whose Ricci curvature is bounded from below by $\text{Ric} \geq kg$ with $k > 0$ must be compact with diameter at most $\pi \sqrt{(n-1)/k}$. Indeed, Ilias proved that, on a manifold of dimension n , the hypothesis $\text{Ric} \geq kg$ for some $k > 0$ implies the Sobolev inequality (4.2) with $q = n/(n-2)$ and $S_M^2 = 4(n-1)/n(n-2)k$. Hence Meyers' result follows from Ilias' inequality and the above theorem. The upper bound in the theorem is sharp and is attained when M is a sphere.

The above theorem of Bakry and Ledoux is, in fact, obtained in a much more general setting of strictly local Dirichlet spaces (see [5] for a precise description).

4.5 Flat Sobolev inequalities and topology at infinity

We complete this section on flat Sobolev inequalities by pointing out the relevance of the Sobolev inequality in some problems concerning topology. The following result due to Carron [18] is actually closely related to the results concerning the Rozenblum–Cwikel–Lieb inequality.

Theorem 4.8 ([18, Theorem 0.4]). *Let (M, g) be a complete Riemannian manifold (hence, connected) satisfying the Sobolev inequality*

$$\forall f \in \mathcal{C}_c^\infty(M), \quad \|f\|_{2\nu/(\nu-2)}^2 \leq S_M^2 \int |\nabla f|^2 d\mu$$

for some $\nu > 2$. Assume that the smallest negative eigenvalue ric_- of the Ricci tensor is in $L^{\nu/2}(M)$. Then M has only finitely many ends. In fact, there exists a constant $C(\nu)$ such that the number of ends is bounded by

$$1 + C(\nu) S_M^2 \int_M |\text{ric}_-|^{\nu/2} d\mu.$$

For more sophisticated results in this direction see, for example, [16, 18, 19] and the references therein.

5 Sobolev Inequalities on Graphs

All the ideas and techniques discussed in this paper can be developed and used in the discrete context of graphs, sometimes to great advantage. To a large extent, the context of graph is actually harder to work with than the context of manifolds (and strictly local Dirichlet spaces), but the new difficulties that appear are mostly of a technical nature and can often be overcome. This short section provides pointers to the literature and explains in some detail one of the first applications of Sobolev inequalities on graphs, namely, Varopoulos' solution of Kesten's conjecture regarding random walks on finitely generated groups. We refer to [47] for a short survey and to [98, 105] for a detailed treatment of some aspects.

5.1 Graphs of bounded degree

In what follows, a graph is a pair (V, E) , where E is a symmetric subset of $V \times V$ and V is finite or countable. Elements of V are vertices and elements of E are (oriented) edges. For $x, y \in V$ we write $x \sim y$ if $(x, y) \in E$ and we say that x, y are neighbors. A path in V is a sequence of vertices such that consecutive points are neighbors. The length of a path is the number of edges it crosses. The distance $\rho(x, y)$ between two points $x, y \in V$ is the minimal length of a path joining them. The degree $\mu(x)$ of $x \in V$ is the number of $y \in V$ such that $(x, y) \in E$. Throughout the paper, we assume that our graphs are connected, i.e., $\rho(x, y) < \infty$ for all $x, y \in V$ and have uniformly bounded degree, i.e., there exists $D \in [1, \infty)$ such that $\sup\{\mu(x)\} = D$.

Moreover, we equip V with the measure μ defined by $\mu(A) = D^{-1} \sum_{x \in A} \mu(x)$.

A graph is regular if $\mu(x) = D$ for all x . In this case, the measure μ is a counting measure. Let $B(x, r)$ be the (closed) ball of radius r around x , and let $V(x, r) = \mu(B(x, r))$. For a book treatment of various aspects of the study the volume growth in Cayley graphs see [35].

Given a function f on V , we set $df(x, y) = f(y) - f(x)$ and

$$|\nabla f(x)| = \left(\mu(x)^{-1} \sum_{y \sim x} |df(x, y)|^2 \right)^{1/2}.$$

Also, set $f_r(x) = V(x, r)^{-1} \sum_{B(x, r)} f(z) \mu(z)$.

We now have all the ingredients to consider whether or not the graph (V, E) satisfies the Sobolev inequality

$$\forall f \in \mathcal{C}_c(V), \quad \|f\|_{2q} \leq S \|\nabla f\|_2 \quad (5.1)$$

for some $q > 1$, and related inequalities. Here, $\mathcal{C}_c(V)$ is the space of functions with finite support. Moreover, according to our notation, we have

$$\|f\|_{2q}^{2q} = \sum_{x \in V} |f(x)|^{2q} \mu(x) \text{ and } \|\nabla f\|_2^2 = \sum_{x \in V} \sum_{y \sim x} |f(y) - f(x)|^2.$$

In what follows, we concentrate on the simple case of flat Sobolev inequalities because this case is quite interesting and important and does avoid most technical difficulties. For developments paralleling the ideas and results of Sect. 2 we refer the reader to [28, 29, 30, 52, 53, 98] and the references therein.

5.2 Sobolev inequalities and volume growth

We start with the following two theorems.

Theorem 5.1. *Fix $\nu > 0$. For a graph (V, E) as above, the following properties are equivalent.*

- $\forall f \in \mathcal{C}_c(V), \quad \|f\|_2^{(1+2/\nu)} \leq N \|\nabla f\|_2 \|f\|_1^{2/\nu}.$
- $\forall f \in \mathcal{C}_c(V)$ with support in a finite set Ω , $\|f\|_2 \leq C\mu(\Omega)^{1/\nu} \|\nabla f\|_2.$

Moreover, if $\nu > 2$, these properties are equivalent to (5.1) with $q = \nu/(\nu - 2)$. Finally, any of these inequalities implies the existence of $c > 0$ such that

$$\forall x \in V, r > 0, \quad V(x, r) \geq cr^\nu.$$

Remark 5.1. The first inequality is a Nash inequality, the second is a Faber–Krahn inequality. For a proof of this theorem see, for example, [4].

The next results gives two Nash inequalities under the volume growth hypothesis that $V(x, r) \geq cr^\nu$. The first inequality requires no additional hypotheses, whereas the second one depends on the validity of a pseudo-Poincaré inequality. Under that extra hypothesis, the Nash inequality one obtains is, in fact, equivalent to the volume lower bound. Both results are optimal (see [6]).

Theorem 5.2 ([6, 31]). *Fix $\nu > 0$ and assume that a graph (V, E) has volume growth bounded from below:*

$$\forall x \in V, r > 0, \quad V(x, r) \geq cr^\nu.$$

- *In all the cases,*

$$\forall f \in \mathcal{C}_c(V), \quad \|f\|_2^{(1+1/\gamma)} \leq N \|\nabla f\|_2 \|f\|_1^{1/\gamma}, \quad \gamma = \nu/(\nu + 1).$$

- *Assume that the pseudo-Poincaré inequality $\forall f \in \mathcal{C}_c(V), \quad \|f - f_r\|_2 \leq Cr \|\nabla f\|_2$ holds on (V, E) . Then*

$$\forall f \in \mathcal{C}_c(V), \quad \|f\|_2^{(1+2/\nu)} \leq N \|\nabla f\|_2 \|f\|_1^{2/\nu}.$$

Proof. First statement. Fix a finite set Ω . For each $x \in \Omega$ let $r(x)$ be the distance between x and $V \setminus (\Omega)$. If f has support in Ω , by a simple use of the Cauchy–Schwarz inequality, for all $x \in \Omega$, $|f(x)|^2 \leq r(x) \|\nabla f\|_2^2$. Also $\Omega \supset B(x, r(x) - 1)$ for each $x \in \Omega$. Hence, by hypothesis,

$$\mu(\Omega) \geq V(x, r(x) - 1) \geq c(r(x) - 1)^\nu \geq c'r(x)^\nu.$$

This yields $|f(x)|^2 \leq C\mu(\Omega)^{1/\nu} \|\nabla f\|_2^2$. Summing over Ω , we find

$$\|f\|_2 \leq C^{1/2} \mu(\Omega)^{(\nu+1)/2\nu} \|\nabla f\|_2.$$

The desired result follows from Theorem 5.1.

Second statement. Observe that the volume hypothesis yields

$$\|f_r\|_\infty \leq c^{-1} r^{-\nu} \|f\|_1.$$

Writing $\|f\|_2^2 = \langle f, f - f_r \rangle + \langle f, f_r \rangle$ and using the hypotheses, we obtain

$$\|f\|_2^2 \leq Cr \|f\|_2 \|\nabla f\|_2 + c^{-1} r^{-\nu} \|f\|_1^2.$$

Picking $r \simeq (\|f\|_1^2 \|f\|_2^{-1} \|\nabla f\|_2^{-1})^{1/(1+\nu)}$, we find

$$\|f\|_2^2 \leq C_1 \|f\|_2^{\nu/(1+\nu)} \|\nabla f\|_2^{\nu/(1+\nu)} \|f\|_1^{2/(1+\nu)}$$

or

$$\|f\|_2^{(2+\nu)/(1+\nu)} \leq C_1 \|\nabla f\|_2^{\nu/(1+\nu)} \|f\|_1^{2/(1+\nu)}.$$

Taking the $(1+\nu)/\nu$ th power of both sides, we arrive at the desired inequality. $\square$

5.3 Random walks

In the context of graphs, one of the possible natural definitions of the “Laplacian” (and the one we will use) is

$$\Delta_E f(x) = \mu(x)^{-1} \sum_{y \sim x} (f(y) - f(x)) = (K - I)f(x),$$

where I is the identity operator and K is the Markov kernel

$$K(x, y) = \begin{cases} \mu(x)^{-1} & \text{if } y \sim x, \\ 0 & \text{otherwise,} \end{cases}$$

and $Kf(x) = \mu(x)^{-1} \sum_{y \sim x} f(y)$. The random walk interpretation of K is as follows. Think of a particle whose current position at a (discrete) time $t \in \mathbb{N}$, is at $x \in V$ with some probability $\mathbf{p}(t)(\{x\}) = \mathbf{p}(t, x)$. At time $t + 1$, the particle picks uniformly one of the neighboring sites and moves there. Hence the probability of the particle to be at a site x at time $t + 1$ is

$$\mathbf{p}(t + 1, x) = \sum_{y \sim x} \mathbf{p}(t, y) \mu(y)^{-1} = \mathbf{p}(t) K(x),$$

where the action of K on a measure $\mathbf{p}$ is defined naturally by $\mathbf{p}K(f) = \mathbf{p}(Kf)$. It follows immediately that the operator K is a self-adjoint contraction on $L^2(V, \mu)$ and the function

$$u(t, x) = \mu(x)^{-1} \mathbf{p}(t, x)$$

is a solution of discrete time discrete space heat equation

$$u(t+1, \cdot) - u(t, \cdot) = \Delta_E u(t, \cdot).$$

In this context, the heat kernel $h(t, x, y)$ is obtained by setting

$$h(t, x, y) = u_x(t, y) = \mu(y)^{-1} \mathbf{p}_x(t, y), \quad \mathbf{p}_x(0, y) = \delta_x(y).$$

It is a symmetric function of x, y , and for any f with finite support on V

$$u(t, x) = \sum_{y \in V} h(t, x, y) f(y) \mu(y)$$

is a solution of the heat equation with the initial value f . Finally, by definition,

$$\mathbf{p}_x(t, y) = h(t, x, y) \mu(y)$$

is the probability that our particle is at y at (discrete) time t given that it started at x at time 0.

The idea of applying Sobolev type inequalities in this context was introduced by Varopoulos [101] and produced a remarkable breakthrough in the study of random walks on graphs and finitely generated groups. The book [105] gives a detailed treatment of many aspects of the resulting developments. The following theorem is the most basic result (see [15, 101, 103, 105]).

Theorem 5.3. *Fix $\nu > 0$. Let (V, E) be a connected graph with bounded degree as above. The following properties are equivalent.*

- $\forall f \in \mathcal{C}_c(V), \quad \|f\|_2^{(1+2/\nu)} \leq N \|\nabla f\|_2 \|f\|_1^{2/\nu}.$
- $\forall t \in \mathbb{N}, x, y \in V, \quad h(t, x, y) \leq C(1+t)^{-\nu/2}.$

Example 5.1. A rather interesting family of examples is as follows. Assume that the graph (V, E) has no loops (i.e., is a tree) and there exists $\nu > 0$ such that $V(x, r) \simeq r^\nu$. Such a tree must have many leaves (vertices of degree 1). For examples of such trees see [6]. Applying Theorems 5.2 and 5.3, we obtain the estimate $h(t, x, y) \leq C(1+t)^{-\nu/(1+\nu)}$. As is proved in [7], this estimate is optimal in the sense that

$$h(2t, x, x) \simeq (1+t)^{-\nu/(\nu+1)}.$$

Much more generally, the following assertion similar to Theorem 4.4 holds as well.

Theorem 5.4 ([27]). *Let (V, E) be as above. Let Φ be a positive smooth decreasing function on $[0, \infty)$ satisfying condition (D), and let $\Theta = -\Phi' \circ \Phi^{-1}$. The following properties are equivalent.*

- *There exists a constant $c_1 \in (0, \infty)$ such that*

$$\forall t \in \mathbb{N}, x, y \in V, \quad h(t, x, y) \leq \Phi(c_1 t).$$

- *There exists a constant $C_1 \in (0, \infty)$ such that for all $f \in \mathcal{C}_c(V)$ with $\|f\|_1 \leq 1$*

$$\Theta(\|f\|_2^2) \leq C \|\nabla f\|_2^2.$$

Example 5.2. A case of interest is when $\Phi(t) = ce^{-t^\gamma}$ for some $\gamma \in (0, 1)$. Then $-\Phi'(t) = ct^{\gamma-1}e^{-t^\gamma}$, $\Phi^{-1}(s) = (c + \log 1/s)^{1/\gamma}$, and $\Theta(s) = s(c + \log 1/s)^{1-1/\gamma}$.

5.4 Cayley graphs

A *Cayley graph* is a graph (V, E) as above, where $V = G$ is a finitely generated group equipped with a finite generating set S and $(x, y) \in V \times V$ is in E if and only if $y = xs$ with $s \in S \cup S^{-1}$. Hence one can assume that S is symmetric, i.e., $S = S^{-1}$. These graphs are regular of degree $D = \#S$, and thus the measure μ used earlier is just a counting measure. Denote by e the identity element in G .

The random walk on a Cayley graph can be described as follows. Let $\xi_1, \xi_2, \dots$ be independent uniform picks in the finite symmetric generating set S . Then for $t \in \mathbb{N}$ and $x, y \in G$, $\mathbf{p}_x(t, y)$ is the probability that the product $X_t = x\xi_1 \cdots \xi_t$ is equal to y . It is obvious that $\mathbf{p}_x(t, y) = \mathbf{p}_e(t, x^{-1}y)$ (left-invariance). For general finitely generated groups the study of such random walks originated in H. Kesten's thesis. Later, Kesten considered the natural question of when such a random walk is recurrent. Recall that recurrence here means that, with probability 1, the walk returns infinitely often to its starting point. A walk that is not recurrent is called transient and has the property that, with positive probability, it never returns to its starting point. By a celebrated result of Polya, the random walk on the integer lattices $\mathbb{Z}^n$ is recurrent if $n = 0, 1, 2$ and is transient otherwise. One of Kesten's questions about the recurrence of random walks can be formulated as follows: What are the groups that admit recurrent random walks (with generating support). For a long time, the conjectural answer known as Kesten's conjecture was that the only groups that admit recurrent random walks are the finite extensions of $\mathbb{Z}^n$, $n = 0, 1, 2$ (i.e., those groups that contain $\{0\}$ or $\mathbb{Z}$ or $\mathbb{Z}^2$ with finite index).

A basic result around this question (see, for example, [105]) is that recurrence is equivalent to

$$\sum_{t=1}^{\infty} \mathbf{p}_e(t, e) = \infty.$$

Indeed, $\sum_{t=1}^{\infty} \mathbf{p}_x(t, y)$ can be understood as the mean number of returns to y starting from x . Thus, the question is really a question about the behavior of the associated heat kernel $h(t, x, x)$.

Theorem 5.5 ([31]). *Fix $p \in [1, \infty]$. Let (V, E) be the Cayley graph associated to a finitely generated group G equipped with a finite symmetric generating set S . Then the pseudo-Poincaré inequality*

$$\|f - f_r\|_p \leq Cr \|\nabla f\|_p \quad (5.2)$$

holds, as well as the Poncaré type inequality

$$\sum_B |f - f_B|^p \leq Cr^p \frac{V(2r)}{V(r)} \sum_{2B} |\nabla f|^p, \quad (5.3)$$

where $B = B(e, r)$, $2B = B(e, 2r)$, $V(r) = \#B(e, r)$, and f_B is the average of f over B .

Remark 5.2. The paper [31] treats mostly the case $p = 1$ (and the case $p = 2$, briefly, towards the end), partly because the other cases are obvious variations on the same argument. The inequality (5.2) with $p = 1$ is contained in [31, p. 296]. The inequality (5.3) with $p = 1$ and $p = 2$ is contained in [31, pp. 308-310] because, on a Cayley graph and under an invariant choice of paths, the constants $K(x, n)$ and $K_2(x, n)$ appearing in [31] can be of order $nV(2n)/V(n)$ and $n^2V(2n)/V(n)$ respectively. Below, we give a complete proof of the case $p = 2$, emphasizing the great similarity between these two inequalities.

Proof. We treat the case $p = 2$ (other cases are similar except for $p = \infty$ which is trivial and has little content). The crucial observation is that for any set $A \subset G$

$$\sum_{x \in A} \sum_{y \in B(e, s) \cap x^{-1}A} |f(xy) - f(x)|^2 \leq (\#S)s^2V(s) \sum_{A_{s/2}} |\nabla f|^2.$$

Here, $A_\tau = \{z \in G : \rho(z, A) \leq \tau\}$. To prove this inequality, for each y denote by γ_y a fixed path of minimal length from e to y and use the Cauchy-Schwarz inequality to get

$$|f(xy) - f(x)|^2 \leq (\#S)|y| \sum_{z \in \gamma_y} |\nabla f|(xz)^2,$$

where $|y| = \rho(e, y)$ is the graph distance between e and y (i.e., the length of y). Note that $\rho(xz, A) \leq \min\{\rho(e, z), \rho(z, y)\} \leq |y|/2 \leq s/2$. Moreover, and

this is the crucial point of the argument, y being fixed, a given vertex $\xi = xz$ can appear for at most $|y|$ different points x . Hence

$$\sum_{x \in A} \sum_{y \in B(e,s) \cap x^{-1}A} |f(xy) - f(x)|^2 \leq (\#S)s^2 V(s) \sum_{A_{s/2}} |\nabla f|^2.$$

Taking $A = G$, $r = s$ and dividing both sides by $V(r)$, we obtain the pseudo-Poincaré inequality $\|f_r - f\|_2 \leq (\#S)^{1/2} r \|\nabla f\|_2$. Taking $A = B(e, r)$, $s = 2r$ and dividing both sides by $V(r)$, we find

$$\sum_B |f - f_B|^2 \leq Cr^2 \frac{V(2r)}{V(r)} \sum_{2B} |\nabla f|^2.$$

□

Theorem 5.6. *Let (V, E) be the Cayley graph associated to a finitely generated group G equipped with a finite symmetric generating set S . Assume that $V(r) \geq cr^\nu$, $r > 0$. Then there are constants N and C such that*

$$\forall f \in \mathcal{C}_c(G), \quad \|f\|_2^{(1+2/\nu)} \leq N \|\nabla f\|_2 \|f\|_1^{2/\nu}$$

and

$$\forall t \in \mathbb{N}, x, y \in G, \quad h(t, x, y) \leq C(1+t)^{-\nu/2}.$$

In addition, if the doubling volume property $V(2r) \leq DV(r)$ holds, then the scale invariant Poincaré inequalities

$$\forall B = B(x, r), \quad \sum_B |f - f_B|^p \leq P_p r^p \sum_B |\nabla f|^p$$

are satisfied for all $p \in [1, \infty]$.

Proof. The first two properties are equivalent and follow from Theorems 5.2, 5.3, and 5.5. The last statement follows from Theorem 5.5 and a well-known, but somewhat subtle argument to get rid of the doubling of the ball over which one integrates the gradient (see, for example, [87, Sect. 5.3]). □

Remark 5.3. The statement that $V(r) \geq cr^\nu$ implies $h(t, x, x) \leq C_\varepsilon(1+t)^{-(\nu-\varepsilon)/2}$, $\varepsilon > 0$, was first proved by Varopoulos [104, 102] by different, but related methods.

Returning to Kesten's conjecture, let us observe that the above theorem implies that if a finitely generated group G satisfies $V(r) \geq cr^\nu$ with $\nu > 2$, then

$$\sum_{t=1}^{\infty} h(t, e, e) = \sum_{t=1}^{\infty} \mathbf{p}_e(t, e) < \infty, \quad (5.4)$$

i.e., the random walks on the Cayley graphs of G are transient (it is easy to see that different generating sets S always yield comparable growth functions

V). This means that a group carrying a recurrent random walk must have a volume growth function satisfying

$$\forall \varepsilon > 0, \quad \liminf_{r \rightarrow \infty} r^{-(2+\varepsilon)} V(r) < \infty.$$

By the celebrated theorem of Gromov [54] (and its extension in [99]), the condition

$$\exists A > 0, \quad \liminf_{r \rightarrow \infty} r^{-A} V(r) < \infty \quad (5.5)$$

implies that G contains a nilpotent subgroup of finite index. Since a subgroup of finite index in G has volume growth comparable to that of G and, by a theorem due to Bass, nilpotent groups have volume growth of type r^ν for some integer ν (see, for example, [35]), we see that a group carrying a recurrent walk must contain a nilpotent subgroup of finite index and volume growth of type r^0 or r^1 or r^2 . It is easy to check that this means that G is a finite extension of $\{0\}$ or $\mathbb{Z}$ or $\mathbb{Z}^2$, as desired.

Theorem 5.7 (solution of Kesten’s conjecture, [104]). *If a finitely generated group G admits a finite symmetric generating set S such that the associated random walk is recurrent, then G is a finite extension of $\{0\}$ or $\mathbb{Z}$ or $\mathbb{Z}^2$.*

Remark 5.4. In a recent preprint [67], Kleiner gave a new proof of Gromov’s theorem on groups of polynomial volume growth. His argument is quite significant since it avoids the use of the Montgomery–Zippin–Yamabe structure theory of locally compact groups (and of the solution of Hilbert fifth problem). It is also very significant from the viewpoint of the present paper and in relation to Theorem 5.7, as we will explain. The proof of Theorem 5.7 is based on two main results: the theorem of Gromov on groups of polynomial growth (albeit, only in the “small growth” case (5.4)) and Varopoulos’ result that links volume growth to the decay of the probability of return of a random walk as expressed in Theorem 5.6. Until Kleiner’s work on Gromov’s theorem, these two corner stones of the proof of Theorem 5.7 appeared to be rather unrelated. However, it is remarkable that one of the key ingredients of Kleiner’s proof is the Poincaré inequality (5.3). Recall that, in Theorem 2.4, we stated a result of Colding and Minicozzi to the effect that, on complete manifolds, the Poincaré inequality and the doubling property imply the *finite dimensionality* of the spaces of harmonic functions of polynomial growth. One of Kleiner’s main ideas in [67] is to show that, because one has (5.3), the Colding–Minicozzi finite dimensionality results for harmonic functions of polynomial growth does hold for Cayley graphs under the (weak) polynomial volume growth hypothesis (5.5). This makes Theorem 5.5 central for each of the two main ingredients of the proof of Kesten’s conjecture.

We complete with what can be seen as a generalization of Theorem 5.7 which involves Sobolev’s inequalities. Because of the relation between the Sobolev inequality and the Nash inequality and the decay of the probability of return in Theorem 5.3, it is possible to formulate Theorem 5.7 in

an equivalent way as follows: a Cayley graph always satisfies the inequality $\|f\|_2^{5/3} \leq N \|\nabla f\|_2 \|f\|_1^{2/3}$ unless the group is a finite extension of a nilpotent group of growth degree at most 2. More generally, the following assertion holds.

Theorem 5.8. *Fix a positive integer ν . On the Cayley graph of a finitely generated group G , the Nash inequality*

$$\forall f \in \mathcal{C}_c(G), \quad \|f\|_2^{1+2/\nu} \leq N \|\nabla f\|_2 \|f\|_1^{2/\nu}$$

always holds for some constant $N \in (0, \infty)$ (depending on ν , G and the Cayley graph structure) unless G is a finite extension of a nilpotent group of volume growth degree at most $\nu - 1$.

Similarly, the Sobolev inequality

$$\forall f \in \mathcal{C}_c(G), \quad \|f\|_{p\nu/(\nu-p)} \leq S \|\nabla f\|_p$$

always holds for all $\nu > p \geq 1$ and some constant $S \in (0, \infty)$ (depending on p, ν, G and the Cayley graph structure) unless G is a finite extension of a nilpotent group of volume growth degree at most $\nu - 1$.

Acknowledgement. Research partially supported by NSF (grant no. DMS 0603886).

References

1. Alexopoulos, G., Lohoué, N.: Sobolev inequalities and harmonic functions of polynomial growth. J. London Math. Soc. (2) **48**, no. 3, 452–464 (1993)
2. Aronson, D.G., Serrin, J.: Local behavior of solutions of quasilinear parabolic equations. Arch. Ration. Mech. Anal. **25**, 81–122 (1967)
3. Aronson, D.G., Serrin, J.: A maximum principle for nonlinear parabolic equations. Ann. Sc. Norm. Super. Pisa (3) **21**, 291–305 (1967)
4. Bakry, D., Coulhon, T., Ledoux, M., Saloff-Coste, L.: Sobolev inequalities in disguise. Indiana Univ. Math. J. **44**, no. 4, 1033–1074 (1995)
5. Bakry, D., Ledoux, M.: Sobolev inequalities and Myers's diameter theorem for an abstract Markov generator. Duke Math. J. **85**, no. 1, 253–270 (1996)
6. Barlow, M., Coulhon, T., Grigor'yan, A.: Manifolds and graphs with slow heat kernel decay. Invent. Math. **144**, no. 3, 609–649 (2001)
7. Barlow, M.T., Coulhon, T., Kumagai, T.: Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs. Commun. Pure Appl. Math. **58**, no. 12, 1642–1677 (2005)
8. Bendikov, A., Coulhon, T., Saloff-Coste, L.: Ultracontractivity and embedding into L^∞ . Math. Ann. **337**, no. 4, 817–853 (2007)
9. Bendikov, A., Maheux, P.: Nash type inequalities for fractional powers of nonnegative self-adjoint operators. Trans. Am. Math. Soc. **359**, no. 7, 3085–3097 (electronic) (2007)

10. Biroli, M., Mosco, U.: A Saint-Venant type principle for Dirichlet forms on discontinuous media. *Ann. Mat. Pura Appl.* (4) **169**, 125-181 (1995)
11. Biroli, M., Mosco, U.: Sobolev and isoperimetric inequalities for Dirichlet forms on homogeneous spaces. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **6**, no. 1, 37-44 (1995)
12. Biroli, M., Mosco, U.: Sobolev inequalities on homogeneous spaces. *Potential Anal.* **4**, no.4, 311-324 (1995)
13. Bougerol, P., Élie, L.: Existence of positive harmonic functions on groups and on covering manifolds. *Ann. Inst. H. Poincaré Probab. Statist.* **31**, no. 1, 59-80 (1995)
14. Buser, P.: A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup.* (4) **15**, no. 2, 213-230 (1982)
15. Carlen, E.A., Kusuoka, S., Stroock, D.W.: Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.* **23**, no. 2, suppl., 245-287 (1987)
16. Carron, G.: Formes harmoniques L^2 sur les variétés non-compactes. *Rend. Mat. Appl.* (7) **21**, no. 1-4, 87-119 (2001)
17. Carron, G.: Inégalités isopérimétriques de Faber-Krahn et conséquences. In: *Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992) Sémin. Congr. Vol. 1*, pp. 205-232. Soc. Math. France, Paris (1996)
18. Carron, G.: Une suite exacte en L^2 -cohomologie. *Duke Math. J.* **95**, no. 2, 343-372 (1998)
19. Carron, G.: L^2 -cohomologie et inégalités de Sobolev. *Math. Ann.* **314**, no. 4, 613-639 (1999)
20. Chavel, I.: *Eigenvalues in Riemannian geometry* Academic Press Inc., Orlando, FL (1984)
21. Chavel, I.: *Riemannian Geometry—a Modern Introduction*. Cambridge Univ. Press, Cambridge (1993)
22. Cheeger, J., Gromov, M., Taylor, M.: Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds. *J. Differ. Geom.* **17**, no. 1, 15-53 (1982)
23. Cheng, S.Y., Li, P., Yau, Sh.T.: On the upper estimate of the heat kernel of a complete Riemannian manifold. *Am. J. Math.* **103**, no. 5, 1021-1063 (1981)
24. Colding, T.H., Minicozzi, W.P.II: Harmonic functions on manifolds. *Ann. Math.* (2) **146**, no. 3, 725-747 (1997)
25. Colding, T.H., Minicozzi, W.P.II: Weyl type bounds for harmonic functions. *Invent. Math.* **131**, no. 2, 257-298 (1998)
26. Coulhon, T.: Dimension à l'infini d'un semi-groupe analytique. *Bull. Sci. Math.* **114**, no. 4, 485-500 (1990)
27. Coulhon, T.: Ultracontractivity and Nash type inequalities. *J. Funct. Anal.* **141**, no. 2, 510-539 (1996)
28. Coulhon, T.: Analysis on infinite graphs with regular volume growth. In: *Random Walks and Discrete Potential Theory (Cortona, 1997) Sympos. Math. Vol. XXXIX*, pp. 165-187. Cambridge Univ. Press, Cambridge (1999)
29. Coulhon, T., Grigor'yan, A.: Pointwise estimates for transition probabilities of random walks on infinite graphs. In: *Fractals in Graz 2001*, pp. 119-134. Birkhäuser, Basel (2003)
30. Coulhon, T., Grigor'yan, A., Zucca, F.: The discrete integral maximum principle and its applications. *Tohoku Math. J.* (2) **57**, no. 4, 559-587 (2005)
31. Coulhon, T., Saloff-Coste, L.: Isopérimétrie pour les groupes et les variétés. *Rev. Mat. Iberoam.* **9**, no. 2, 293-314 (1993)
32. Coulhon, T., Saloff-Coste, L.: Variétés riemanniennes isométriques à l'infini. *Rev. Mat. Iberoam.* **11**, no. 3, 687-726 (1995)

33. Davies, E.B.: Heat Kernels and Spectral Theory. Cambridge Univ. Press, Cambridge (1989)
34. DeGiorgi, E.: Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3) **3**, 25-43 (1957)
35. de la Harpe, P.: Topics in Geometric Group Theory. University of Chicago Press, Chicago, IL (2000).
36. Deny, J.: Méthodes hilbertiennes en théorie du potentiel. In: Potential Theory (C.I.M.E., I Ciclo, Stresa, 1969), pp. 121-201. Edizioni Cremonese, Rome (1970)
37. Duong, T.X., Ouhabaz, E.M., Sikora, A.: Plancherel type estimates and sharp spectral multipliers. J. Funct. Anal. **196**, no. 2, 443-485 (2002)
38. Duong, T.X., Ouhabaz, E.M., Sikora, A.: Spectral multipliers for self-adjoint operators. In: Geometric Analysis and Applications (Canberra (2000)), pp. 56-66. Austral. Nat. Univ., Canberra (2001)
39. Franchi, B., Gutiérrez, C.E., Wheeden, R.L.: Weighted Sobolev-Poincaré inequalities for Grushin type operators. Commun. Partial Differ. Equ. **19**, no. 3-4, 523-604 (1994)
40. Fukushima, M., Ōshima, Yō., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes. Walter de Gruyter, Berlin (1994)
41. Gallot, S., Hulin, D., Lafontaine, J.: Riemannian Geometry. Springer-Verlag, Berlin (1990)
42. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Springer-Verlag, Berlin (2001)
43. Grigor'yan, A.: The heat equation on noncompact Riemannian manifolds (Russian). Mat. Sb. **182**, no. 1, 55-87 (1991); English transl.: Math. USSR, Sb. **72**, no.1, 47-77 (1992)
44. Grigor'yan, A.: Heat kernel upper bounds on a complete non-compact manifold. Rev. Mat. Iberoam. **10**, no. 2, 395-452 (1994)
45. Grigor'yan, A.: Analytic and geometric background of recurrence and non-explosion of the Brownian motion on Riemannian manifolds. Bull. Am. Math. Soc. (N.S.) **36**, no. 2, 135-249 (1999)
46. Grigor'yan, A.: Estimates of heat kernels on Riemannian manifolds. In: Spectral Theory and Geometry (Edinburgh (1998)), pp. 140-225. Cambridge Univ. Press, Cambridge (1999)
47. Grigor'yan, A.: Heat kernels on manifolds, graphs and fractals. In: European Congress of Mathematics, Vol. I (Barcelona (2000)), pp. 393-406. Birkhäuser, Basel (2001)
48. Grigor'yan, A.: Heat kernels on weighted manifolds and applications. In: The Ubiquitous Heat Kernel, pp. 93-191. Am. Math. Soc., Providence, RI (2006)
49. Grigor'yan, A., Saloff-Coste, L.: Heat kernel on connected sums of Riemannian manifolds. Math. Res. Lett. **6**, no. 3-4, 307-321 (1999)
50. Grigor'yan, A., Saloff-Coste, L.: Stability results for Harnack inequalities. Ann. Inst. Fourier (Grenoble) **55**, no. 3, 825-890 (2005)
51. Grigor'yan, A., Saloff-Coste, L.: Heat kernel on manifolds with ends. (2007)
52. Grigor'yan, A., Telcs, A.: Sub-Gaussian estimates of heat kernels on infinite graphs. Duke Math. J. **109**, no. 3, 451-510 (2001)
53. Grigor'yan, A., Telcs, A.: Harnack inequalities and sub-Gaussian estimates for random walks. Math. Ann. **324**, no. 3, 521-556 (2002)
54. Gromov, M.: Groups of polynomial growth and expanding maps. Inst. Hautes Études Sci. Publ. Math. **53**, 53-73 (1981)
55. Guivarc'h, Y.: Croissance polynomiale et périodes des fonctions harmoniques. Bull. Soc. Math. France **101**, 333-379 (1973)
56. Gutiérrez, C.E., Wheeden, R.: Mean value and Harnack inequalities for degenerate parabolic equations. Colloq. Math. **60/61**, no. 1, 157-194 (1990)

57. Gutiérrez, C.E., Wheeden, R.: Harnack's inequality for degenerate parabolic equations. *Commun. Partial Differ. Equ.* **16**, no. 4-5, 745-770 (1991)
58. Gutiérrez, C.E., Wheeden, R.: Bounds for the fundamental solution of degenerate parabolic equations. *Commun. Partial Differ. Equ.* **17**, no. 7-8, 1287-1307 (1992)
59. Hajlasz, P., Koskela, P.: Sobolev met Poincaré. *Mem. Am. Math. Soc.* no. 688 (2000)
60. Hebey, E.: *Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities*. Courant Lect. Notes **5**. Am. Math. Soc., Providence, RI (1999)
61. Hebisch, W., Saloff-Coste, L.: On the relation between elliptic and parabolic Harnack inequalities. *Ann. Inst. Fourier (Grenoble)* **51**, no. 5, 1437-1481 (2001)
62. Heinonen, J.: *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York (2001)
63. Heinonen, J.: Nonsmooth calculus. *Bull. Am. Math. Soc. (N.S.)* **44**, no. 2, 163-232 (electronic), (2007)
64. Kanai, M.: Rough isometries, and combinatorial approximations of geometries of noncompact Riemannian manifolds. *J. Math. Soc. Japan* **37**, no. 3, 391-413 (1985)
65. Kanai, M.: Analytic inequalities, and rough isometries between noncompact Riemannian manifolds. In: *Curvature and Topology of Riemannian Manifolds* (Katata, (1985), pp. 122-137. Springer, Berlin (1986)
66. Kanai, M.: Rough isometries and the parabolicity of Riemannian manifolds. *J. Math. Soc. Japan* **38**, no. 2, 227-238 (1986)
67. Kleiner, B.: A new proof of Gromov's theorem on groups of polynomial growth. (2007) arXiv:0710.4593v4
68. Kuchment, P., Pinchover, Y.: Liouville theorems and spectral edge behavior on abelian coverings of compact manifolds. *Trans. Am. Math. Soc.* **359**, no. 12, 5777-5815 (electronic) (2007)
69. Ladyzhenskaya, O.A., Solonnikov, V.A., Ural'tseva, N.N.: *Linear and Quasilinear Equations of Parabolic Type* Am. Math. Soc., Providence, RI (1968)
70. Levin, D., Solomyak, M.: The Rozenblum-Lieb-Cwikel inequality for Markov generators. *J. Anal. Math.* **71**, 173-193 (1997)
71. Li, P.: Harmonic functions on complete Riemannian manifolds. In: *Tsing Hua Lectures on Geometry and Analysis* (Hsinchu, (1990-1991), pp. 265-268. Int. Press, Cambridge, MA (1997)
72. Li, P.: Harmonic sections of polynomial growth. *Math. Res. Lett.* **4**, no. 1, 35-44 (1997)
73. Li, P., Yau, Sh.T.: On the Schrödinger equation and the eigenvalue problem. *Commun. Math. Phys.* **88**, no. 3, 309-318 (1983)
74. Li, P., Yau, Sh.T.: On the parabolic kernel of the Schrödinger operator. *Acta Math.* **156**, no. 3-4, 153-201 (1986)
75. Maz'ya, V.G.: *Sobolev Spaces*. Springer-Verlag, Berlin-Tokyo (1985)
76. Morrey, Ch.B., Jr.: *Multiple Integrals in the Calculus of Variations*. Springer-Verlag, New York (1966)
77. Moser, J.: On Harnack's theorem for elliptic differential equations. *Commun. Pure Appl. Math.* **14**, 577-591 (1961)
78. Moser, J.: A Harnack inequality for parabolic differential equations. *Commun. Pure Appl. Math.* **17**, 101-134 (1964)
79. Moser, J.: Correction to: "A Harnack inequality for parabolic differential equations." *Commun. Pure Appl. Math.* **20**, 231-236 (1967)
80. Moser, J.: On a pointwise estimate for parabolic differential equations. *Commun. Pure Appl. Math.* **24**, 727-740 (1971)
81. Nash, J.: Continuity of solutions of parabolic and elliptic equations. *Am. J. Math.* **80**, 931-954 (1958)

82. Pivarski, M., Saloff-Coste, L.: Small time heat kernel behavior on riemannian complexe.
83. Porper, F.O., Eidel'man, S.D.: Two-sided estimates of fundamental solutions of second-order parabolic equations, and some applications (Russian). *Uspekhi Mat. Nauk* **39**, no.3, 107-156 (1984); English transl.: *Russ. Math. Surv.* **39**, no.3, 119-178 (1984)
84. Rozenblum, G., Solomyak, M.: The Cwikel–Lieb–Rozenblum estimates for generators of positive semigroups and semigroups dominated by positive semigroups (Russian). *Algebra Anal.* **9**, no. 6, 214–236 (1997); English transl.: *St. Petersburg. Math. J.* **9**, no. 6, 1195–1211 (1998)
85. Saloff-Coste, L.: A note on Poincaré, Sobolev, and Harnack inequalities. *Int. Math. Res. Not.* **2**, 27-38 (1992)
86. Saloff-Coste, L.: Parabolic Harnack inequality for divergence form second order differential operators. *Potential Anal.* **4**, no. 4, 429-467 (1995)
87. Saloff-Coste, L.: *Aspects of Sobolev-Type Inequalities*. Cambridge Univ. Press, Cambridge (2002)
88. Saloff-Coste, L.: Analysis on Riemannian co-compact covers. In: *Surveys in Differential Geometry*. Vol. 9, pp. 351-384. Int. Press, Somerville, MA (2004)
89. Semmes, S.: *Some Novel Types of Fractal Geometry*. Clarendon Press, Oxford (2001)
90. Serrin, J.: Local behavior of solutions of quasi-linear equations. *Acta Math.* **111**, 247-302 (1964)
91. Sikora, A.: Sharp pointwise estimates on heat kernels. *Q. J. Math., Oxf. II. Ser.* **47**, no. 187, 371-382 (1996)
92. Sikora, A.: Riesz transform, Gaussian bounds and the method of wave equation. *Math. Z.* **247**, no. 3, 643-662 (2004)
93. Sobolev, S.L.: On a theorem of functional analysis (Russian). *Mat. Sb.* **46**, 471-497 (1938); English transl.: *Am. Math. Soc., Transl., II. Ser.* **34**, 39-68 (1963)
94. Sturm, K.T.: Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. *J. Math. Pures Appl. (9)* **75**, no. 3, 273-297 (1996)
95. Sturm, K.T.: Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p -Liouville properties. *J. Reine Angew. Math.* **456**, 173-196 (1994)
96. Sturm, K.T.: Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.* **32**, no. 2, 275-312 (1995)
97. Sturm, K.T.: The geometric aspect of Dirichlet forms. In: *New Directions in Dirichlet Forms*, pp. 233-277. Am. Math. Soc., Providence, RI (1998)
98. Telcs, A.: *The Art of Random Walks*. Springer-Verlag, Berlin (2006)
99. van den Dries, L., Wilkie, A.J.: Gromov's theorem on groups of polynomial growth and elementary logic. *J. Algebra* **89**, no. 2, 349-374 (1984)
100. Varopoulos, N.Th.: Hardy-Littlewood theory for semigroups. *J. Funct. Anal.* **63**, no. 2, 240-260 (1985)
101. Varopoulos, N.Th.: Isoperimetric inequalities and Markov chains. *J. Funct. Anal.* **63**, no. 2, 215-239 (1985)
102. Varopoulos, N.Th.: Convolution powers on locally compact groups. *Bull. Sci. Math. (2)* **111**, no. 4, 333-342 (1987)
103. Varopoulos, N.Th., Saloff-Coste, L., Coulhon, T.: *Analysis and Geometry on Groups*. Cambridge Univ. Press, Cambridge (1992)
104. Varopoulos, N.Th.: Théorie du potentiel sur des groupes et des variétés. *C. R. Acad. Sci., Paris Sér. I Math.* **302**, no. 6, 203-205 (1986)
105. Woess, W.: *Random Walks on Infinite Graphs and Groups*. Cambridge Univ. Press, Cambridge (2000)