Widder’s representation theorem for symmetric local Dirichlet spaces

Nathaniel Eldredge and Laurent Saloff-Coste

May 10, 2014

Abstract

In classical PDE theory, Widder’s theorem gives a representation for nonnegative solutions of the heat equation on $\mathbb{R}^n$. We show that an analogous theorem holds for local weak solutions of the canonical “heat equation” on a symmetric local Dirichlet space satisfying a local parabolic Harnack inequality.

Contents

1 Introduction 2

2 Definition of local weak solutions 3

3 Assumptions 6

4 Examples 8

4.1 Intrinsic distance 8

4.2 Fractal-like spaces 11

4.3 Locally compact but infinite dimensional examples 11

5 Properties of local weak solutions 13

6 Widder’s theorem 20

7 Conditions for uniqueness of nonnegative solutions 22

8 An application to projections 23

9 Acknowledgements 26

A Energy measures 26
1 Introduction

The purpose of this article is to give an extension of Widder’s theorem [44], which gives a representation for nonnegative solutions of the heat equation, to the general setting of symmetric, strictly local, regular Dirichlet spaces.

To motivate the theorem, consider the Cauchy problem for the classical heat equation on $\mathbb{R}^n$:

$$\frac{\partial u(t, x)}{\partial t} - \frac{1}{2} \Delta u(t, x) = 0, \quad x \in \mathbb{R}^n, \quad t > 0$$

$$u(0, x) = f(x), \quad x \in \mathbb{R}^n.$$ (1.1)

When (1.1) is introduced in an elementary PDE course, an immediate question is whether solutions of (1.1) are unique. The answer, of course, is no: explicit examples of nonzero functions $u$ satisfying (1.1) with $f = 0$ are known, and can be found in [34] and references therein. However, such counterexamples generally have some sort of bad behavior. For instance, they grow rapidly as $|x| \to \infty$; if one requires certain growth conditions, such as the requirement that

$$|u(t, x)| \leq C e^{c|x|^2},$$ (1.2)

then it is well known that there is a unique solution of (1.1) which satisfies (1.2). See, e.g. [14, Theorem 2.3.6].

Another sort of bad behavior that these counterexamples exhibit is that they are unbounded below. If we think of the heat equation as a model for heat flow, then such solutions are non-physical, since temperatures cannot be less than absolute zero. Thus, it is natural to restrict our attention to nonnegative solutions of (1.1), and ask whether this is sufficient to ensure uniqueness. In 1944, D. Widder showed in [44] that the answer is affirmative. This is a satisfying result, since the hypothesis of nonnegativity seems more natural and less arbitrary than growth conditions such as (1.2).

Widder also showed that every nonnegative solution $u$ of the classical heat equation in $\mathbb{R}^n$ for times $t > 0$ is of the form

$$u(t, x) = P_t \nu(x) := \int_{\mathbb{R}^n} p(t, x, y) \nu(dy)$$ (1.3)

for some unique positive Radon measure $\nu$, where $p(t, x, y) = \frac{1}{(2\pi t)^{n/2}} e^{-|x-y|^2/2t}$ is the classical heat kernel. One can interpret this result as saying that any nonnegative solution of the heat equation for times $t > 0$ must have evolved from some initial temperature distribution $\nu$ (which may be singular). This result was later extended to nonnegative classical solutions [23] and weak solutions [2, 3] of more general second-order parabolic equations on $\mathbb{R}^n$.

To extend Widder’s result to more general spaces than $\mathbb{R}^n$, one must first notice that uniqueness of nonnegative solutions of the Cauchy problem may fail. For example, if $\Omega$ is a bounded open subset
of $\mathbb{R}^n$, then there are many nonnegative classical solutions of
\begin{equation}
\partial_t u(t, x) - \frac{1}{2} \Delta u(t, x) = 0, \quad x \in \Omega, \ t > 0
\end{equation}
\begin{equation}
\quad u(0, x) = f(x), \quad x \in \mathbb{R}^n.
\end{equation}
For instance, we might take
\begin{equation}
u(t, x) = \begin{cases}
p(t - t_0, x, x_0), & t > t_0 \geq 0 \\
0, & t \leq t_0
\end{cases}
\end{equation}
for some $x_0 \notin \Omega$. The problem is that (1.4) does nothing to exclude the possibility of heat entering through the boundary of $\Omega$. Thus we certainly cannot hope to represent a solution $u$ as $u(t, x) = P_t \nu(x)$. A more appropriate representation for $u$ might be
\begin{equation}
u(t, x) = P_t \nu(x) + h(t, x)
\end{equation}
where $h$ is another nonnegative solution of the heat equation which vanishes at time 0. This is the type of result we shall prove in the present paper. We shall return to the question of uniqueness in Section 7.

In 1992, Ancona and Taylor [1] proved a representation result of the form (1.5) in the language of abstract potential theory, where solutions of a parabolic equation on a manifold are considered as a sheaf of functions satisfying certain properties. A key ingredient in their proof is the assumption that solutions satisfy a parabolic Harnack inequality. It is worth noting that their results are able to include solutions to equations of the form $\partial_t u(t, x) - Lu(t, x) = 0$ where the second-order operator $L$ need not be elliptic but can be hypoelliptic.

In recent years, attention has focused on the notion of Dirichlet spaces (see section 2) as a setting for the study of potential theory. In this setting, one takes as the underlying space a metric measure space or similar object; in particular, no differentiable structure is assumed. However, the space carries enough structure to allow one to define a notion of a solution to a canonical “Laplace equation” or “heat equation,” and in particular to study functional inequalities for such solutions, such as Harnack and Poincaré inequalities. Since a Dirichlet space also carries a canonical stochastic process, one is also able to exploit tools from probability and probabilistic potential theory.

The main result of this paper is to prove that so-called local weak solutions of the heat equation on a Dirichlet space, under certain assumptions, are given by a representation of the form (1.5). The proof is along similar lines to that of [1], but the details are quite different. Along the way, we shall obtain several useful properties of local weak solutions.

For related results in a variety of other settings, see [26, 37, 27, 32].

## 2 Definition of local weak solutions

Let $(X, d, \mu)$ be a metric measure space: $(X, d)$ is a metric space, and $\mu$ is a positive Radon measure on $X$. We further assume that $X$ is separable and locally compact; it follows that $X$ is Polish and that every finite Borel measure is automatically Radon.
Let \((\mathcal{E}, \mathcal{D})\) be a **Dirichlet form** on \(L^2(X, \mu)\): a closed, Markovian, positive, bilinear form. (We refer the reader to [16] for the definition of a Dirichlet form and of the following properties.) The quintuple \((X, d, \mu, \mathcal{E}, \mathcal{D})\) will be called a **Dirichlet space**. We assume that \((\mathcal{E}, \mathcal{D})\) is **symmetric**, **regular**, and **strictly local**. We shall write \(\mathcal{E}(f)\) for \(\mathcal{E}(f, f)\).

To each \(f \in \mathcal{D} \cap L^\infty\) there is associated a Radon measure \(\Gamma(f)\), called the **energy measure** of \(f\), defined by

\[
\int_X \phi \, d\Gamma(f) = 2\mathcal{E}(\phi f, f) - \mathcal{E}(f^2, \phi)
\]

for \(\phi \in C_c(X) \cap \mathcal{D}\), where \(C_c(X)\) denotes the space of continuous, compactly supported functions on \(X\). (It is worth recalling that \(\mathcal{D} \cap L^\infty\) is an algebra; see [16, Theorem 1.4.2 (iii)].) One may then define the signed measure \(\Gamma(f, g) = \frac{1}{2}(\Gamma(f + g) - \Gamma(f) - \Gamma(g))\) by polarization. For the classical Dirichlet form on \(\mathbb{R}^n\), we have \(d\Gamma(f, g) = \nabla f \cdot \nabla g \, dm\). We have collected some useful properties of \(\Gamma\) in Appendix A; further information can be found in [15].

Let \(L\) denote the self-adjoint generator of \((\mathcal{E}, \mathcal{D})\), with its domain \(D(L)\). We will take the sign convention that \(L\) is a **positive** semidefinite operator, so that \(\mathcal{E}(f, g) = (f, Lg)_{L^2(X, \mu)}\) for \(f \in \mathcal{D}\), \(g \in D(L)\). Let \(P_t\) denote the associated strongly continuous, symmetric, Markovian semigroup.

Note that since \((\mathcal{E}, \mathcal{D})\) is closed, \(\mathcal{D}\) is a Hilbert space under the inner product \(\mathcal{E}_1(f, g) = \mathcal{E}(f, g) + \int_X fg \, d\mu\). (As before, \(\mathcal{E}_1(f)\) is short for \(\mathcal{E}_1(f, f)\).) The inclusion \(\mathcal{D} \hookrightarrow L^2(X, \mu)\) is obviously continuous, 1-1, and has dense image, so taking its adjoint gives another inclusion \(L^2(X, \mu) \hookrightarrow \mathcal{D}^*\) which is also continuous, 1-1, and has dense image. (So \((\mathcal{D}, L^2(X, \mu), \mathcal{D}^*)\) is a Gelfand triple.) Under these identifications, the pairing \((\ell, f)_{\mathcal{D}^*, \mathcal{D}}\) is given by \(\int_X \ell(x) f(x) \mu(dx)\) when \(\ell \in L^2(X, \mu) \subset \mathcal{D}^*\).

We will try to keep denoting this pairing by \((\ell, f)_{\mathcal{D}^*, \mathcal{D}}\), but it would be a permissible abuse of notation to just write \(\int_X \ell(x) f(x) \mu(dx)\) in all cases.

In the classical case when \(X = \mathbb{R}^n, \mu = m\) is Lebesgue measure and \(\mathcal{E}(f, g) = \int_{\mathbb{R}^n} \nabla f \cdot \nabla g \, dm\) on the domain \(\mathcal{D} = H^1(\mathbb{R}^n)\) (so \(L = -\frac{1}{2} \Delta\)) we have \(\mathcal{D}^* = H^{-1}(\mathbb{R}^n)\). So it is helpful to think of \(\mathcal{D}^*\) as some space of distributions on \(X\).

We note that the assumption that \((\mathcal{E}, \mathcal{D})\) is regular may impose some implicit “boundary conditions” on functions \(f \in \mathcal{D}\). For example, if \(X = \Omega\) is a bounded open subset of \(\mathbb{R}^n\) and \(\mathcal{E}\) is the classical Dirichlet form, then \(\mathcal{D} = H^1_0(\Omega)\), so functions in the domain \(\mathcal{D}\) must satisfy Dirichlet boundary conditions on \(\partial \Omega\).

Intuitively, we want to consider solutions \(u : (0, T) \times X \to \mathbb{R}\) of the heat equation

\[
\partial_t u + Lu = 0.
\]

However, the implicit assumption in \((2.2)\) that \(u(t, \cdot) \in D(L)\) for each \(t\) is much too strong. In particular, it is a global condition: it requires that \(u(t, \cdot)\) is in \(L^2(X, \mu)\) and satisfies certain boundary conditions. We want something more analogous to the classical heat equation

\[
\partial_t u - \frac{1}{2} \Delta u = 0
\]

on an open set \(\Omega \subset \mathbb{R}^n\), which is only a local equation. It makes sense for any \(u \in C^2\), for instance, and imposes no global conditions such as integrability or behavior at the boundary.

The resulting notion of “local weak solution” has a more complicated definition than \((2.2)\), but comes closer to the intuition of \((2.3)\) in that it avoids global conditions.
Notation 2.1. Let \( L^2([a, b]; \mathbb{D}) \) denote the Hilbert space of (strongly measurable) square-integrable vector-valued functions \( u : [a, b] \rightarrow \mathbb{D} \) under the norm
\[
\|u\|_{L^2([a, b]; \mathbb{D})}^2 := \int_a^b \mathcal{E}_1(u(t), u(t)) \, dt.
\]
\( m \) will denote Lebesgue measure on the time interval \([a, b]\).

Notation 2.2. Let \( W^{1,2}([a, b]; \mathbb{D}, \mathbb{D}^*) \) denote the vector-valued Sobolev space of functions \( u \in L^2([a, b]; \mathbb{D}) \) with one time derivative \( u' \in L^2([a, b]; \mathbb{D}^*) \). We will write \( W^{1,2} \) for short. \( W^{1,2} \) is a Hilbert space under the norm
\[
\|u\|^2_{W^{1,2}([a, b]; \mathbb{D}, \mathbb{D}^*)} := \int_a^b (\mathcal{E}_1(u(t), u(t)) + \|u'(t)\|^2_{\mathbb{D}^*}) \, dt.
\]
A standard “Sobolev embedding theorem” gives that \( W^{1,2}([a, b]; \mathbb{D}, \mathbb{D}^*) \subset C([a, b]; L^2(X, \mu)) \), so that \( u(t) \) is well-defined as an element of \( L^2(X, \mu) \) for every \( t \in [a, b] \), and \( u \) is also well-defined as an element of \( L^2([a, b] \times X) \). Also, we recall the following “product rule”: if \( u \in W^{1,2} \), then \( t \mapsto \|u(t)\|^2_{L^2(X, \mu)} \) is an absolutely continuous function on \([a, b]\), and
\[
\frac{d}{dt}\|u(t)\|^2_{L^2(X, \mu)} = (u'(t), u(t))_{\mathbb{D}^*, \mathbb{D}}. \tag{2.4}
\]
We refer the reader to [45, §25] for proofs.

One can show that any function \( u : [a, b] \rightarrow \mathbb{D} \) which is in \( W^{1,2} \) is represented by an (almost everywhere defined) function from \([a, b] \times X\) to \( \mathbb{R} \), so we may write either \( u(t) \) or \( u(t, x) \) depending on whether we prefer to think of \( u \) as a curve in a function space or a real-valued function of time and space.

Definition 2.3 (See, e.g., [41, 42]). Let \( U \subset X \) be open. A function \( u : (0, T) \times U \rightarrow \mathbb{R} \) is said to be a local weak solution (of the heat equation 2.2) if for every compact \( K \subset (0, T) \times U \), there exists a function \( u_K \in W^{1,2}([0, T]; \mathbb{D}, \mathbb{D}^*) \) such that \( u = u_K \) (\( m \times \mu \)-a.e.) on \( K \), and such that for every \( \phi \in W^{1,2}([0, T]; \mathbb{D}, \mathbb{D}^*) \) with compact support inside \( K \), we have
\[
\int_0^T (u_K'(t), \phi(t))_{\mathbb{D}^*, \mathbb{D}} \, dt + \int_0^T \mathcal{E}(u_K(t), \phi(t)) \, dt = 0. \tag{2.5}
\]
Note that, because the Dirichlet form \( (\mathcal{E}, \mathbb{D}) \) is local, the expression in (2.5) does not depend on the choice of \( u_K \). Therefore, for test functions \( \phi \) which are compactly supported inside \((0, T) \times U\), we can interpret an expression like \( \int_0^T \mathcal{E}(u(t), \phi(t)) \, dt \) as shorthand for \( \int_0^T \mathcal{E}(u_K(t), \phi(t)) \, dt \) where \( K \subset (0, T) \times U \) is any compact set containing the support of \( \phi \). We shall henceforth commit this abuse of notation.

We also note that the left side of (2.5) is continuous in \( \phi \) with respect to the \( L^2([0, T]; \mathbb{D}) \) topology. Therefore, since \( W^{1,2}([0, T]; \mathbb{D}, \mathbb{D}^*) \) is dense in \( L^2([0, T]; \mathbb{D}) \), (2.5) holds for all \( \phi \in L^2([0, T]; \mathbb{D}) \) which have compact support inside \( K \).

Observe that we do not assume that \( u(t) \in \mathbb{D} \) or even \( u(t) \in L^2(X, \mu) \). This definition gives no control on the behavior of \( u \) “at infinity,” i.e. away from compact sets, and \( u \) need not satisfy any implicit Dirichlet boundary conditions. Each \( u_K \) does need to satisfy them, but \( u \) can be totally different from \( u_K \) outside the compact set \( K \). There is also no assumption as to what happens near \( t = 0 \) and \( t = T \).
3 Assumptions

In this section, we collect, and discuss, the hypotheses under which we shall prove our results.

Assumption 1. $(X, d)$ is a separable, locally compact, connected, locally connected metric space. $\mu$ is a positive Radon measure on $X$ with full support. $(\mathcal{E}, \mathcal{D})$ is a symmetric, regular, strictly local Dirichlet form on $L^2(X, \mu)$.

We shall now assume that local weak solutions satisfy a parabolic Harnack inequality. This inequality was developed by Moser, extending previous work of Hadamard and Pini; see [28, 29] and references therein for further history.

Assumption 2. Nonnegative local weak solutions satisfy the following parabolic Harnack inequality. Let $U \subset X$ be open and connected, let $K \subset U$ be compact, and fix $0 < a < b < c < d < T$. There exists a finite constant $C = C(U, K, a, b, c, d)$ such that for every nonnegative local weak solution $u$ on $(0, T) \times U$,

$$\text{esssup}_{[a, b] \times K} u \leq C \text{essinf}_{[c, d] \times K} u.$$  \hspace{1cm} (3.1)

We make no assumption as to the exact dependence of $C$ on $U, K, a, b, c, d$, so this Harnack inequality is local and in no way scale-invariant.

In the context of local Dirichlet spaces, various forms of the parabolic Harnack inequality are known to be related to other functional inequalities, such as Poincaré inequalities and heat kernel estimates. See, for example, [41, 42, 6]. We shall discuss this further in Section 4.

Assumption 3. The semigroup $P_t$ admits a continuous heat kernel. That is, there is a continuous $p : (0, \infty) \times X \times X \rightarrow \mathbb{R}$ such that, for all $f \in L^2(X, \mu)$,

$$P_t f(x) = \int_X p(t, x, y) f(y) \mu(dy) \quad \text{for } \mu\text{-a.e } x \in X.$$  \hspace{1cm} (3.2)

Since the semigroup $P_t$ is Markovian, it follows that $p \geq 0$. In fact, since $u(t, x) = p(t, x_0, x)$ is a local weak solution, the Harnack inequality (Assumption 2) implies that $p > 0$. We will make use of this later.

Notation 3.1. For any positive Radon measure $\nu$ on $X$, set

$$P_t \nu(x) = \int_X p(t, x, y) \nu(dy).$$  \hspace{1cm} (3.2)

Assumption 4. For $f \in C_c(X)$ and every $x \in X$, we have $P_t f(x) \rightarrow f(x)$ pointwise as $t \rightarrow 0$. (Note that since the heat kernel $p$ is assumed to be continuous, $P_t f$ is continuous and hence $P_t f(x)$ is well-defined for every $x \in X$.)

Assumption 4 is very mild. Indeed, we are not aware of any examples which satisfy our previous assumptions but not Assumption 4; however, we have not been able to show that Assumption 4 follows from them.
We emphasize that in Assumption 4, we require that the convergence hold for every \( x \in X \). Under our other assumptions, the convergence will certainly hold for quasi-every \( x \in X \) (i.e. except on a set which, in the terminology of [16], is zero capacity, or equivalently is exceptional). The Dirichlet form \( (\mathcal{E},\mathcal{D}) \) is associated with a continuous Hunt process \( X_t \) which is normal in the sense that \( \mathbb{P}^x(X_0 = x) = 1 \). So by dominated convergence, we have \( \mathbb{E}_x[f(X_t)] \to f(x) \) for every \( x \in X \). On the other hand, it also holds that for quasi-every \( x \), \( P_t f(x) = \mathbb{E}_x[f(X_t)] \) for every \( t > 0 \) and every \( f \in C_c(X) \).

The Hunt process \( X_t \) is not unique (one can change its behavior on an exceptional set of initial points), and in many cases, there is a canonical choice of \( X_t \) for which \( P_t f(x) = \mathbb{E}_x[f(X_t)] \) holds everywhere. One could circumvent Assumption 4 by taking as the starting point a particular continuous Hunt process \( X_t \) with a continuous transition density \( p(t,x,y) \) and taking \( (\mathcal{E},\mathcal{D}) \) to be the corresponding Dirichlet form.

It is also worth noting that Assumption 4 holds automatically in the common case that \( P_t \) is a Feller semigroup, i.e. a strongly continuous semigroup on \( C_0(X) \). In this case the convergence holds not only pointwise but uniformly. Another situation where the assumption holds is when the state space \( X \) is a homogeneous space such as a Lie group, and the Dirichlet form is translation invariant.

In Definition 2.3, the existence of the “localized” functions \( u_K \) is merely asserted, and we have little control over them. To produce them more constructively, we would like to use some “smooth” cutoff functions, as follows.

**Definition 3.2.** Let \( U \) be open and \( K \subset U \) compact. We will say a function \( \psi : X \to [0,1] \) is a nice cutoff function for \( K \) inside \( U \) if:

1. \( \psi = 1 \) on \( K \);
2. \( \psi \) is compactly supported inside \( U \);
3. For every \( f \in \mathcal{D}, \psi f \in \mathcal{D} \). (It follows from the closed graph theorem that there is a constant \( C_\psi \) such that \( \mathcal{E}_1(\psi f) \leq C_\psi \mathcal{E}_1(f) \). And in particular, \( \psi \in \mathcal{D} \).)

**Definition 3.3.** We shall say an open set \( U \subset X \) is nice if for every compact \( K \subset U \) there exists a nice cutoff function for \( K \) inside \( U \).

**Assumption 5.** There is a sequence of open, precompact, connected, nice sets \( U_n \uparrow X \).

Some sufficient conditions for this assumption to hold are as follows:

1. If \( X \) is compact, then \( X \) itself is nice, since \( \psi = 1 \) is a nice cutoff function of any closed set inside \( X \). So we can take \( U_n = X \). Actually, when \( X \) is compact, several of the arguments in this paper become trivial.
2. If the intrinsic pseudo-distance defined by (4.1) below is a genuine metric and generates the topology of \( X \), then every open set is nice: we can use the metric to construct cutoff functions that are Lipschitz. This is a common assumption in the theory of Dirichlet spaces. See Section 4.1 for more on this notion.
This reduces Assumption 5 to the statement that we can exhaust $X$ by open, precompact, connected sets (which we shall abbreviate OPC for this paragraph). But this follows from our topological assumptions on $X$. Indeed, if we write $x \sim y$ whenever there is an OPC set containing $x, y$, then by local compactness and local connectedness $\sim$ is an equivalence relation on $X$, and every equivalence class is open. But then by connectedness of $X$ there is only one equivalence class. Thus any pair of points, and moreover any finite set of points, is contained in an OPC set. The same holds for any compact set $K$: we can cover $K$ by a finite number of OPC sets $U_n$. Picking one $x_n$ from each $U_n$, we can find an additional OPC set $U_0$ which contains all the $x_n$. Now the union of $U_0$ and the $U_n$ is an OPC set which contains $K$. Since $X$ is $\sigma$-compact, this suffices.

3. If the Dirichlet space satisfies a local cutoff Sobolev inequality $CS(\beta)_{\text{loc}}$ in the sense of [6], then all open sets are nice, and so Assumption 5 holds. See Appendix B. Note that $CS(\beta)_{\text{loc}}$ is shown in [6] to follow from a uniform parabolic Harnack inequality. The authors discuss examples of spaces, such as certain fractals, that are known to satisfy this condition, although they do not have a well-behaved intrinsic distance. See also Section 4.2.

The useful consequence of Assumption 5 is that if $U$ is nice and $u$ is a local weak solution on $(0, T) \times U$, then for compact $K \subset (0, T) \times U$, we can produce the function $u_K$ explicitly as follows. We may enclose $K$ inside some set $[a, b] \times K_1$ for a compact $K_1 \subset U$. Let $\psi$ be a nice cutoff function for $K_1$ inside $U$, and let $\chi \in C_c^\infty((0, T))$ be a cutoff function which equals 1 on $[a, b]$. Set $K' = \text{supp} \chi \times \text{supp} \psi$ and let $u_K(t, x) = \chi(t)\psi(x)u_{K'}(t, x)$. This gives the same function $u_K$ no matter which of the many possible choices for $u_{K'}$ is used, and it is not hard to see that this $u_K$ is in $W^{1, 2}([0, T]; D, D^*)$.

4 Examples

This section discusses examples of spaces satisfying the assumptions listed in Section 3. As long as one is interested in local self-adjoint Markov operators defined over a locally compact metric space, it should be clear that these assumptions are quite mild and are satisfied for a great many interesting examples. We will work under the very basic Assumption 1 and focus on various contexts where the other assumptions listed in Section 3 are satisfied.

4.1 Intrinsic distance

Under Assumption 1 the intrinsic distance $d_\mathcal{E}$ is defined by

$$d_\mathcal{E}(x, y) = \inf \{ f(y) - f(x) : f \in C_c(X) \cap \mathbb{D}, d\Gamma(f) \leq d\mu \}, \quad x, y \in X. \tag{4.1}$$

Here $d\Gamma(f) \leq d\mu$ means that the Radon measure $\Gamma(f)$ is absolutely continuous with respect to $\mu$ with Radon–Nikodym derivative bounded by 1. The function $(x, y) \mapsto d_\mathcal{E}(x, y)$ may be 0 even if $x \neq y$ or may be infinity for some $x, y$ but is symmetric and satisfies the triangle inequality. Loosely speaking, the condition $d\Gamma(f) \leq d\mu$ can be understood as requiring that $f$ is “Lipschitz” with constant 1. On a complete Riemannian manifold equipped with its natural Dirichlet space structure, $d_\mathcal{E}$ equals the Riemannian distance. See [40, 53, 10].
Fractals such as the Sierpinski gasket and carpet provide examples of Dirichlet spaces with very interesting properties but where the intrinsic distance $d_E$ is identically 0 (the only functions with “bounded gradient” are the constant functions). The infinite dimensional torus $T^\infty$, the countable product of circles with its normalized Haar measure $\mu$, equipped with the Dirichlet form $E_A$ associated with an infinite symmetric positive definite matrix $A = (a_{i,j})$,

$$E_A(f) = \int_{T^\infty} \sum_{i,j} a_{i,j} \partial_i f \partial_j f d\mu,$$

provide examples where, depending on $A$, the intrinsic distance may be finite and continuous or infinite except on a dense set of measure 0. See [7].

We define the balls $B(x,r)$ and the \textbf{volume growth function} $V(x,r)$ by setting

$$B(x,r) = \{y : d_E(x,y) < r\} \text{ and } V(x,r) = \mu(B(x,r)).$$

Please note that these balls are relative to the intrinsic distance $d_E$.

Consider the following properties that may or may not be satisfied:

\textbf{(ID1)} The intrinsic distance $d_E$ is continuous and defines the topology of $X$.

\textbf{(ID2)} Property (ID1) is satisfied and for any compact set $K \subset X$ there are constants $r_K \in (0, \infty)$ and $D_K, P_K \in (0, \infty)$ such that:

- (Compact balls) For all $x \in K$ and $r < r_K$, the closed ball $\overline{B(x,2r)}$ is compact.
- (Doubling) For all any $x \in K$ and $r \in (0,r_K)$, we have $V(x,2r) \leq D_K V(x,r)$.
- (Poincaré inequality) For all any $x \in K$ and $r \in (0,r_K)$,

$$\forall f \in \mathbb{D}, \int_{B(x,r)} |f - f_{B(x,r)}|^2 d\mu \leq P_K r^2 \int_{B(x,2r)} d\Gamma(f)$$

where $f_B$ denotes the mean of $f$ over $B$.

We say that \textbf{(ID2) holds locally uniformly} if there exists $r_0, D_0, P_0 \in (0, \infty)$ such that (ID2) holds true with $r_{\{x\}} \geq r_0$, $D_{\{x\}} \leq D_0$ and $P_{\{x\}} \leq P_0$ for all $x \in X$. Note that this implies that $\overline{B(x,r_0)}$ is compact for every $x$, and in particular $d_E$ is a complete metric on $X$.

We say that \textbf{(ID2) holds uniformly at all scales and positions} if there are constants $D_0$ and $P_0$ such that (ID2) holds with $r_K = \infty$, $D_K \leq D_0$ and $P_K \leq P_0$, for any compact $K \subset X$. Spaces where this holds are sometimes said to be of \textbf{Harnack type}.

As explained in [40], property (ID1) implies the existence of an abundance of cutoff functions. In particular, under (ID1), every open set is nice, and so Assumption 5 is satisfied as argued in the previous section.

By the work of K.T. Sturm [42] extending to Dirichlet spaces earlier work by Grigoryan [17] and Saloff-Coste [35], the following theorem holds.
Theorem 4.1. Assuming (ID1), property (ID2) holds if and only if for any compact $K$ there exists a constant $H_K$ such that any nonnegative local weak solution $u$ of the heat equation \(^{(2.2)}\) in $Q = (0, r^2) \times B(x, r)$, $x \in K$, $r \in (0, r_K)$, satisfies

$$
\text{esssup}_{Q_-} \{u\} \leq H_K \text{essinf}_{Q_+} \{u\},
$$

(4.2)

where $Q_- = (r^2/4, r^2/2) \times B(x, r/2)$ and $Q_+ = (3r^2/4, r^2) \times B(x, r/2)$.

Further, if (ID2) holds locally uniformly and $r_0$ is as described above, then $r_{(x)} \geq r_0 > 0$ and there is a constant $H_0$ such that $H_{(x)} \leq H_0$ for all $x \in X$. If (ID2) holds uniformly at all scales and locations then $r_K = \infty$ and there is a constant $H_0$ such that $H_K \leq H_0$ for all compact $K \subset X$.

The parabolic Harnack inequality supplied by Theorem 4.1 can be shown to imply that of Assumption \(^{(2)}\) by covering the set $K$ that appears in Assumption \(^{(2)}\) with finitely many sufficiently small overlapping balls. Moreover, it is argued in \(^{(41)}\) that under (ID1)–(ID2) a measurable heat kernel exists. Another very important consequence of (ID1)–(ID2), as shown in \(^{(42, Corollary 3.3)}\), is that local weak solutions are locally Hölder continuous. In particular the heat kernel is continuous, so Assumption \(^{(3)}\) is also satisfied.

Using the continuity of local weak solutions, we can also show that Assumption \(^{(4)}\) is satisfied. Fix $f \in C_c(X)$, $x_0 \in X$ and $\epsilon > 0$. Using regularity and the Markovian properties of $(\mathcal{E}, \mathbb{D})$, we may choose $g \in C_c(X) \cap \mathbb{D}$ with $\|f - g\|_{\infty} < \epsilon$ and such that $g$ is constant on a neighborhood $U$ of $x_0$. Since $P_t f, P_t g$ are continuous on $(0, \infty) \times X$ and $P_t$ is Markovian, we also have $\|P_t f - P_t g\|_{\infty} < \epsilon$.

Now it is shown in \(^{(19, Lemma 2.28)}\) that if we set

$$
u(t, x) = \begin{cases} 
P_t g(x), & t > 0 \\
g(x_0), & t \leq 0 
\end{cases}
$$

then $u$ is a local weak solution on $(-\infty, \infty) \times U$. (This also follows from our extension principle, Lemma \(^{(5.2)}\) below.) Therefore $u$ is continuous on $(-\infty, \infty) \times U$, and in particular $P_t g \to g(x_0)$ uniformly on compact subsets of $U$. Letting $\epsilon \to 0$, it follows that $P_t f \to f$ uniformly on compact subsets of $U$, and in particular $P_t f(x_0) \to f(x_0)$.

To summarize:

Corollary 4.2. Let $(X, d)$ be a separable, locally compact, connected, locally connected metric space equipped with a positive Radon measure $\mu$ with full support and a symmetric, regular, strictly local Dirichlet form $(\mathcal{E}, \mathbb{D})$ on $L^2(X, \mu)$. If $(X, d, \mu, \mathcal{E}, \mathbb{D})$ satisfies (ID1)–(ID2) then the assumptions \(^{(2)}\), \(^{(3)}\), \(^{(4)}\), and \(^{(5)}\) of Section \(^{(3)}\) are all satisfied.

Among the great many concrete examples of Dirichlet spaces satisfying (ID1)–(ID2), let us mention inner uniform Euclidean domains equipped with the canonical Neumann-type Dirichlet form \(^{(19)}\), complete Riemannian manifolds, connected Lie groups equipped with an invariant sub-Riemannian structure given by a generating family of left-invariant vector fields \(^{(43)}\), the natural Dirichlet form on a polytopal complex (under mild structural assumptions) \(^{(13, 33)}\), and Alexandrov spaces with sectional curvature bounded below \(^{(24)}\). Another example is the harmonic Sierpiński gasket described in \(^{(21)}\), which is something of a bridge to the fractal-like spaces of the next subsection.
4.2 Fractal-like spaces

For the purposes of this paper, we call a Dirichlet space satisfying Assumption 1 fractal-like of type $\beta$ if the following local parabolic Harnack inequality relative to the metric balls $B_d(x, r)$ of $(X, d)$ is satisfied:

\[(H-\beta) \text{ For any compact } K, \text{ there exist constants } r_K, H_K \text{ such that for any } x \in K, \text{ we have } \overline{B(x, 2r_K)} \text{ is compact, and any nonnegative local weak solution } u \text{ of the heat equation (2.2) in } Q = (0, r^\beta) \times B(x, r), x \in K, r \in (0, r_K), \text{ satisfies} \]

\[
\text{esssup}_{Q-} \{u\} \leq H_K \text{essinf}_{Q_+} \{u\},
\]

where $Q_- = (r^\beta/4, r^\beta/2) \times B(x, r/2)$ and $Q_+ = (3r^\beta/4, r^\beta) \times B(x, r/2)$.

As in Section 4.1, we can also say that $(H-\beta)$ holds locally uniformly if $H(x), \ r(x)$ may be chosen independent of $x$, and uniformly at all scales and positions if moreover we can take $r(x) \equiv \infty$.

The parameter $\beta$ is known as the walk dimension and describes the space-time scaling in the Dirichlet space. Thanks to the work of Barlow, Bass, Kumagai, and their collaborators (see [6] and references therein), property $(H-\beta)$ can be characterized in a way that is similar in spirit to the statement provided by Theorem 4.1.

The parabolic Harnack inequality (4.3) is certainly stronger than that of Assumption 2, and continuity of the heat kernel (and other local weak solutions) follows as well, thus verifying Assumptions 3 and 4 as in Section 4.1. Condition (ID1) often fails to hold in fractal-like spaces, but a replacement is supplied by so-called CS($\beta$) cutoff Sobolev inequalities as described in [6]. The latter condition is implied by (4.3) [6, Theorem 2.16], and guarantees the existence of reasonable cutoff functions; in particular, as we show in Appendix B it implies Assumption 5. Thus, a fractal-like space in the sense introduced above satisfies all the assumptions introduced in Section 3.

Examples of fractal-like spaces in this sense include the Sierpiński gasket [5], generalized Sierpiński carpets [6], and Laakso spaces [38].

4.3 Locally compact but infinite dimensional examples

An interesting classes of examples comes from a symmetric Gaussian semigroup on the infinite dimensional torus $\mathbb{T}^\infty$ (the countable product of circles) and $\mathbb{R}^k \times \mathbb{T}^\infty$. Each of these spaces is equipped with its Haar measure $\mu$. Non-degenerate symmetric Gaussian convolution semigroups of measures, $(\mu^A_t)_{t \geq 0}$, are in one-to-one correspondence with symmetric positive definite matrices $A = (a_{i,j})$ in such a way that for any smooth function $\phi$ depending only on finitely many coordinates, we have

\[
\lim_{t \to 0} \frac{1}{t} (\mu^A_t(\phi) - \phi(0)) = \sum_{i,j} a_{i,j} \partial_i \partial_j \phi(0).
\]

Here the partial derivatives refer to the natural coordinate system in the corresponding product space. The sum on the right-hand side has only finitely many non-zero terms because $\phi$ depends only on finitely many coordinates. To say that $A$ is positive definite is to say that for any vector
\( \xi = (\xi_i) \) with finitely many non-zero coordinates, \( \langle A\xi, \xi \rangle = \sum a_{i,j} \xi_i \xi_j \geq 0 \) and \( \langle A\xi, \xi \rangle = 0 \) if and only if \( \xi = 0 \).

The family of measures \( (\mu_t^A)_{t \geq 0} \) defines a symmetric Markov semigroup \( f \mapsto f * \mu_t^A \) with associated Dirichlet form

\[
\mathcal{E}_A(f, g) = \int \sum_{i,j} a_{i,j} \partial_i f \partial_j g \, d\mu.
\]

The domain of this form is the closure of the smooth compactly supported functions depending only on finitely many coordinates in the norm \( \|f\|_2 + \mathcal{E}_A(f, f)^{1/2} \) and can be described more explicitly, see [7].

As examples of Dirichlet spaces on locally compact spaces, these examples are interesting because of the great variety of very different behaviors. For instance, depending on the matrix \( A \), the measures \( \mu_t^A \) may or may not have a density with respect to the Haar measure \( \mu \) and, if it exists, this density may or may not be continuous. Further, depending on \( A \), the intrinsic distance defined by (4.1) may or may not have property (ID1) whereas property (ID2) is never satisfied.

The following theorem addresses the question of whether or not the assumptions of Section 3 hold. We note that \( \mathbb{R}^k \times \mathbb{T}^\infty \) is locally compact, metrizable, path connected and locally path connected. The Dirichlet forms \( \mathcal{E}_A \) described above are regular and strictly local, so Assumption 1 is always satisfied. Assumption 4 is also always satisfied (indeed, in this case, if \( f \) is continuous and compactly supported, \( f * \mu_t^A(x) = \mathbb{E}_x[f(X_t)] \) everywhere, by invariance). Assumption 5 is also always satisfied (use smooth cutoff functions that depend only on finitely many coordinates).

Define

\[
W_A(s) = \# \{ \theta \in \mathbb{Z}^{(\infty)} : \langle A\theta, \theta \rangle \leq s \}
\]

where \( \mathbb{Z}^{(\infty)} \) is the set of integer valued sequences with finitely many non-zero entries. The function \( W_A \) may be infinite for some \( s \).

**Theorem 4.3** ([7]). **Referring to the above setting and notation, we have:**

- **Assumption 3** is satisfied if and only if

  \[
  \lim_{s \to \infty} \frac{1}{s} \log W_A(s) = 0.
  \]

- **Assumption 2** is satisfied if and only if

  \[
  \lim_{s \to \infty} \frac{1}{\sqrt{s}} \log W_A(s) = 0.
  \]

  This is also equivalent to the statement that \( \mu_t^A \) is absolutely continuous with respect to the Haar measure \( \mu \) and admits a continuous density \( x \mapsto \mu_t^A(x) \) such that

  \[
  \lim_{t \to 0} t \log \mu_t^A(0) = 0.
  \]

Computing the functions \( W_A \) is a very difficult task. However, the results become much more explicit in the case when \( A \) is a diagonal matrix with diagonal entries \( a_{i,i} = a_i > 0 \). In this case, set

\[
N_A(s) = \# \{ i : a_i \leq s \}.
\]
Then, Assumption 3 is satisfied if and only if
\[
\lim_{s \to \infty} \frac{1}{s} \log N_A(s) = 0.
\]
Assumption 2 is satisfied if and only if
\[
\lim_{s \to \infty} \frac{1}{s} N_A(s) = 0.
\]

5 Properties of local weak solutions

In this section, we collect a number of facts about local weak solutions that we will use in the proof of the main theorem.

We can integrate by parts in (2.5) and put the time derivative on the test function \( \phi \). This results in a statement that makes sense for \( u \) which are not a priori assumed to have an \( L^2 \) time derivative. By the following lemma, this new statement is equivalent to Definition 2.3.

**Lemma 5.1.** Let \( U \subset X \) be open and nice. A function \( u \in L^2([0,T]; \mathbb{D}) \) is a local weak solution on \((0,T) \times U\) if and only if it satisfies
\[
- \int_0^T (\phi'(t), u(t))_{\mathbb{D}^*, \mathbb{D}} \, dt + \int_0^T \mathcal{E}(\phi(t), u(t)) \, dt = 0 \tag{5.1}
\]
for all \( \phi \in W^{1,2}([0,T]; \mathbb{D}, \mathbb{D}^*) \) which are compactly supported inside \((0,T) \times U\).

**Proof.** The forward implication is trivial, using integration by parts. For the converse, let \( V \) be open and precompact with \( \nabla \subset U \). Let \( \psi \) be a nice cutoff function for \( \nabla \) inside \( U \). Then \( \phi \mapsto - \int_0^T \mathcal{E}(u(t), \psi \phi(t)) \, dt \) is a bounded linear functional on \( L^2([0,T]; \mathbb{D}) \). But \( L^2([0,T]; \mathbb{D}^*) = L^2([0,T]; \mathbb{D}^*)^{12} \) [Theorem II.13.5.8, Corollary 1], so there exists \( w \in L^2([0,T]; \mathbb{D}^*) \) such that for all \( \phi \in L^2([0,T]; \mathbb{D}) \),
\[
\int_0^T (w(t), \phi(t))_{\mathbb{D}^*, \mathbb{D}} \, dt = - \int_0^T \mathcal{E}(u(t), \psi \phi(t)) \, dt.
\]

For any \( \chi \in C_c(0,T) \) and any \( f \in \mathbb{D} \), we note that \( \chi(t)\psi f \in W^{1,2}([0,T]; \mathbb{D}^*) \) with compact support inside \((0,T) \times U\). Thus we have
\[
\left( \int_0^T w(t) \chi(t) \, dt, f \right)_{\mathbb{D}^*, \mathbb{D}} = \int_0^T (w(t), \chi(t) f)_{\mathbb{D}^*, \mathbb{D}} \, dt
\]
\[
= - \int_0^T \mathcal{E}(u(t), \chi(t) f) \, dt
\]
\[
= - \int_0^T (\chi'(t) \psi f, u(t))_{L^2} \, dt
\]
\[
= - \left( \int_0^T \chi'(t) \psi u(t) \, dt, f \right)_{L^2}.
\]

Thus \( w \) is the weak derivative of \( \psi u \), and so we have \( \psi u \in W^{1,2} \).
Now if $\phi \in W^{1,2}$ with compact support inside $(0, T) \times V$, we have
\[
\int_0^T (\psi u'(t), \phi(t))_{D^*, D} \, dt = -\int_0^T \mathcal{E}(u(t), \psi \phi(t)) \, dt
\]
\[
= -\int_0^T \mathcal{E}(\psi u(t), \phi(t)) \, dt
\]
since $\mathcal{E}$ is local and $\psi = 1$ on the support of $\phi$. Thus we have produced a function (namely $\psi u$) which equals $u$ a.e. on $(0, T) \times V$ and satisfies the necessary equation. Since $V$ was arbitrary, $u$ is a local weak solution on $(0, T) \times U$.

We now give an “extension principle” giving conditions for a local weak solution to have an extension backwards in time. In the setting of Euclidean space, a similar result was given in [2].

**Lemma 5.2 (Extension principle).** Let $U$ be open and nice, and let $u$ be a local weak solution on $(0, T) \times U$. Suppose that for any nice cutoff function $\psi$ compactly supported inside $U$, we have $\psi u \in L^2([0, T]; \mathbb{D})$, and $\psi u(t) \to 0$ weakly in $L^2(X, \mu)$ as $t \downarrow 0$. Extend $u$ by setting $u(t) = 0$ for $t \leq 0$. Then $u$ is a local weak solution on $(-\infty, T) \times U$.

**Proof.** Let $V$ be open and precompact with $\overline{V} \subset U$, and let $\psi$ be a nice cutoff function of $\overline{V}$ inside $U$. Let $\phi \in W^{1,2}((-\infty, T]; \mathbb{D}, \mathbb{D}^*)$ be compactly supported inside $(-\infty, T) \times V$. Fix any $\epsilon > 0$, and let $\chi_n \in C_c((\epsilon, T))$ be a bounded sequence of cutoff functions with $\chi_n \to 1_{[\epsilon, T]}$ pointwise. Integrating by parts, we have
\[
\int_\epsilon^T (\phi'(t), \psi u(t))_{D^*, D} \, dt = -\langle \phi(\epsilon), \psi u(\epsilon) \rangle_{L^2(X, \mu)} - \int_\epsilon^T (\psi u'(t), \phi(t))_{D^*, D} \, dt
\]
\[
= -\langle \phi(\epsilon), \psi u(\epsilon) \rangle_{L^2(X, \mu)} + \lim_{n \to \infty} \int_\epsilon^T (\psi u'(t), \chi_n(t) \phi(t))_{D^*, D} \, dt
\]
\[
= -\langle \phi(\epsilon), \psi u(\epsilon) \rangle_{L^2(X, \mu)} + \lim_{n \to \infty} \int_\epsilon^T \mathcal{E}(\psi u(t), \chi_n(t) \phi(t)) \, dt
\]
\[
= -\langle \phi(\epsilon), \psi u(\epsilon) \rangle_{L^2(X, \mu)} + \int_\epsilon^T \mathcal{E}(\psi u(t), \phi(t)) \, dt,
\]
since $u$, and hence $\psi u$, is a local weak solution on $(0, T) \times V$. The limits involving $\chi_n$ converge as desired by dominated convergence.

Now let $\epsilon \to 0$. Since $\phi \in W^{1,2}([0, T]; \mathbb{D}, \mathbb{D}^*)$ and $\psi u \in L^2([0, T]; \mathbb{D})$, we have by dominated convergence that
\[
\int_\epsilon^T (\phi'(t), \psi u(t))_{D^*, D} \, dt \to \int_0^T (\phi'(t), \psi u(t))_{D^*, D} \, dt = \int_{-\infty}^T (\phi'(t), \psi u(t))_{D^*, D} \, dt
\]
\[
\int_\epsilon^T \mathcal{E}(\psi u(t), \phi(t)) \, dt \to \int_0^T \mathcal{E}(\psi u(t), \phi(t)) \, dt = \int_{-\infty}^T \mathcal{E}(\psi u(t), \phi(t)) \, dt.
\]
Also, since $\phi \in C((-\infty, T]; L^2(X, \mu))$ we have $\phi(\epsilon) \to \phi(0)$ in $L^2(X, \mu)$, and $\psi u(\epsilon) \to 0$ weakly in $L^2(X, \mu)$, so $\langle \phi(\epsilon), \psi u(\epsilon) \rangle_{L^2(X, \mu)} \to 0$. Thus we have shown
\[
\int_{-\infty}^T (\phi'(t), \psi u(t))_{D^*, D} \, dt = \int_{-\infty}^T \mathcal{E}(\phi(t), \psi u(t)) \, dt
\]
so by Lemma 5.1 we have that \( \psi u \) (hence \( u \)) is a local weak solution on \( (-\infty, T) \times V \). Since \( V \subset U \) was arbitrary, we are done.

An important property of a local weak solution is that, locally, the size of the function controls its energy. This is the content of the following inequality. A similar inequality was used in [36, Section 5.2.2].

**Lemma 5.3** (Caccioppoli-type inequality). Let \( U \) be open, nice, and precompact, and let \( V \) be open and precompact with \( \overline{V} \subset U \). Let \( \psi \) be a nice cutoff function of \( \overline{V} \) inside \( U \). There is a constant \( C \), depending on \( U, V, \psi, T \), such that, for every non-negative local weak solution \( u \) on \( (0, T) \times U \), we have

\[
\|\psi u\|_{L^2([0,T];\mathbb{D})} \leq C \text{esssup}_{[0,T] \times U} u.
\]
(Recall that \( \|v\|_{L^2([0,T];\mathbb{D})} = \int_0^T \mathcal{E}_1(v(t), v(t)) \, dt = \int_0^T \left[ \|v(t)\|_{L^2(X, \mu)}^2 + \mathcal{E}(v(t)) \right] \, dt. \)

**Proof.** Set \( M = \text{esssup}_{[0,T] \times U} u \). Since \( \int_0^T \|\psi u(t)\|_{L^2(X, \mu)}^2 \, dt \leq TM^2 \), it will be enough to show \( \int_0^T \mathcal{E}(\psi u(t)) \, dt \leq CM^2 \).

By Proposition A.3 we have

\[
\mathcal{E}(\psi u(t), \psi u(t)) \leq 2 \int_X u(t)^2 \, d\Gamma(\psi) + 2 \int_X \psi^2 \, d\Gamma(u(t)).
\]
(5.3)

The first term is bounded by \( 2M^2 \mathcal{E}(\psi) \) so we work on the second term. By repeated application of the product rule (Proposition A.1),

\[
\int_X \psi^2 \, d\Gamma(u(t)) = \mathcal{E}(u(t), \psi^2 u(t)) - 2 \int_X u(t) \psi \, d\Gamma(u(t), \psi)
\]
\[
\leq \mathcal{E}(u(t), \psi^2 u(t)) + \left| \int_X (2u(t)) \psi \, d\Gamma(u(t), \psi) \right|
\]
\[
\leq \mathcal{E}(u(t), \psi^2 u(t)) + 2 \int_X u(t)^2 \, d\Gamma(\psi) + \frac{1}{2} \int_X \psi^2 \, d\Gamma(u(t)) \quad \text{by (A.2)}.
\]

Thus

\[
\int_X \psi^2 \, d\Gamma(u(t)) \leq 2\mathcal{E}(u(t), \psi^2 u(t)) + 4 \int_X u(t)^2 \, d\Gamma(\psi) \leq 2\mathcal{E}(u(t), \psi^2 u(t)) + 4M^2 \mathcal{E}(\psi).
\]

Now let \( \chi_n \in C^\infty_c((0, T)) \) with \( \chi_n \uparrow 1 \) pointwise and \( \int_0^T |\chi_n'(t)| \, dt \leq 5 \). By monotone convergence,

\[
\int_0^T \int_X \psi^2 \, d\Gamma(u(t)) \, dt = \lim_{n \to \infty} \int_0^T \chi_n(t) \int_X \psi^2 \, d\Gamma(u(t)) \, dt
\]
\[
\leq 4M^2 T \mathcal{E}(\psi) + \limsup_{n \to \infty} \int_0^T \mathcal{E}(u(t), \chi_n(t) \psi^2 u(t)) \, dt.
\]
However, \( \phi(t,x) = \chi_n(t)\psi(x)^2u(t,x) \) is in \( W^{1,2}([0,T];\mathbb{D},\mathbb{D}^*) \) and is compactly supported inside \((0,T) \times U\). Thus, since \( u \) is a local weak solution on \((0,T) \times U\), we have

\[
\int_0^T \mathcal{E}(u(t), \chi_n(t)\psi^2 u(t)) \, dt = -\int_0^T (u'(t), \chi_n(t)\psi^2 u(t))_{\mathbb{D}^*,\mathbb{D}} \, dt
\]

\[
= -\int_0^T \chi_n(t)((\psi u)'(t), \psi u(t))_{\mathbb{D}^*,\mathbb{D}} \, dt
\]

\[
= \int_0^T \chi_n(t)\|\psi u(t)\|_{L^2}^2 \, dt
\]

\[
\leq 5M^2\mu(U)
\]

where in the next-to-last line we used (2.1) and integration by parts. This completes the proof. □

Combining Lemmas 5.1 and 5.3 lets us show that a bounded limit of local weak solutions is another local weak solution.

**Lemma 5.4.** Let \( U \) be open and nice, and let \( u_n \) be a sequence of nonnegative local weak solutions on \((0,T) \times U\) which are uniformly bounded, i.e. \( 0 \leq u_n \leq M \) on \((0,T) \times U\). Suppose \( u_n \to u \) pointwise. Then \( u \) is a local weak solution on \((0,T) \times U\).

**Proof.** Let \( V \) be open and precompact with \( \overline{V} \subset U \), and let \( \psi \) be a nice cutoff function of \( \overline{V} \) inside \( U \). By Lemma 5.3 we have \( \|\psi u_n\|_{L^2([0,T];\mathbb{D})} \leq CM \) for all \( n \). Passing to a subsequence, we may assume \( \psi u_n \) converges weakly to \( \psi u \) in the Hilbert space \( L^2([0,T];\mathbb{D}) \). By Lemma 5.1 it immediately follows that \( \psi u \) is a local weak solution on \((0,T) \times V\) (the left side of (5.1), as a function of \( \psi u \)). Since \( V \subset U \) was arbitrary, we are done. □

We now record some properties of local weak solutions produced by the heat semigroup \( P_t \).

**Lemma 5.5.** Let \( f \in L^2(X,\mu) \), and for each \( t > 0 \) let \( u(t) = P_t f \). Then \( u(t) \in W^{1,2}([0,T];\mathbb{D},\mathbb{D}^*) \) for any \( T \); in fact, \( u(t) \in D(L) \) for every \( t > 0 \), and \( u'(t) = Lu(t) \). In particular, \( u \) is a local weak solution on \((0,\infty) \times X\).

**Proof.** This is a simple consequence of the spectral theorem. □

For \( \nu_n \) a sequence of finite positive Radon measures on \( X \), recall that \( \nu_n \to \nu \) weakly if \( \int_X f \, d\nu_n \to \int_X f \, d\nu \) for every \( f \in C_0(X) \), where \( C_0(X) \) is the space of continuous functions on \( X \) which vanish at infinity (i.e. the uniform closure of \( C_c(X) \)). (In other words, this is weak-* convergence in \( C_0(X)^* \).) Equivalently, \( \nu_n \to \nu \) weakly iff \( \{\nu_n\} \) is bounded in total variation (i.e. \( \sup_n \nu_n(X) < \infty \)) and \( \int_X f \, d\nu_n \to \int_X f \, d\nu \) for every \( f \in C_c(X) \). For a nonnegative measurable function \( f \), we identify \( f \) with the measure \( f \, d\mu \), and note that the total variation of this measure is \( \|f\|_{L^1(X,\mu)} \).

**Lemma 5.6.** If \( \nu \) is a compactly supported positive Radon measure, then \( P_t \nu \to \nu \) weakly as \( t \to 0 \). (\( P_t \nu \) is as defined in (3.3).)

**Proof.** It follows from the Markovian property of \( P_t \) that \( \|P_t \nu\|_{L^1(X,\mu)} \leq \nu(X) \), so \( \{P_t \nu\} \) is bounded in total variation.
For $f \in C_c(X)$, we have by Fubini’s theorem and the symmetry of $P_t$ that $\int_X f(x)P_t\nu(x)\mu(dx) = \int_X Pf(y)\nu(dy)$. By Assumption 4, $P_t f \to f$ pointwise, and the Markovian property gives $\|P_t f\|_\infty \leq \|f\|_\infty < \infty$. Thus by dominated convergence, $\int_X f(x)P_t\nu(x)\mu(dx) \to \int_X f(x)\nu(dx)$.

Note that Assumption 4 asserts that $P_t f(x) \to f(x)$ for every $x \in X$. “$\mu$-almost every $x$” would not be sufficient to establish Lemma 5.6, as the measure $\nu$ could charge $\mu$-null sets.

**Lemma 5.7.** If $\nu$ is a compactly supported positive Radon measure, then $u(t,x) = P_t\nu(x)$ is a local weak solution on $(0,\infty) \times X$, and moreover $P_t\nu \in \mathbb{D}$ for each $t > 0$.

**Proof.** It is sufficient to show that $P_t\nu \in L^2(X,\mu)$ for any $t > 0$, since then we can write $P_t\nu = P_{t-\epsilon}P_{\epsilon}\nu$ and apply Lemma 5.5 (replacing $t$ by $t - \epsilon$) to get that $u$ is a local weak solution on $(\epsilon,\infty) \times X$, for arbitrary $\epsilon$. If $\nu$ is supported on the compact set $K$, we have

$$\int_X |P_t\nu(x)|^2 \mu(dx) = \int_X \left| \int_K p(t,x,y)\nu(dy) \right|^2 \mu(dx)$$

$$\leq \nu(K) \int_X \left( \int_K p(t,x,y)^2 \nu(dy) \right) \mu(dx) \quad \text{(Cauchy–Schwarz)}$$

$$= \nu(K) \int_K \left( \int_X p(t,x,y)p(t,y,x)\mu(dx) \right) \nu(dy) \quad \text{(by symmetry of } p)$$

$$= \nu(K) \int_K p(2t,y,y)\nu(dy)$$

$$\leq \nu(K)^2 \sup_{y \in K} \int_X p(2t,y,y) < \infty$$

since $p$ is continuous.

**Lemma 5.8.** Let $\nu_n$ be a sequence of positive Radon measures supported in a single compact set $K \subset X$. Suppose $\nu_n \to \nu$ weakly, and $t_n \to t \in (0,\infty)$. Then $P_{t_n}\nu_n(x) \to P_t\nu(x)$ for each $x \in X$.

**Proof.** Fix $x \in X$ and $\epsilon > 0$. Write

$$P_{t_n}\nu_n(x) - P_t\nu(x) = (P_{t_n}\nu_n(x) - P_{t_n}\nu_n(x)) + (P_{t_n}\nu_n(x) - P_t\nu(x)).$$

For the first term, we have

$$|P_{t_n}\nu_n(x) - P_{t_n}\nu_n(x)| = \left| \int_K (p(t_n,x,y) - p(t,x,y))\nu_n(dy) \right| \leq \nu_n(K) \sup_{y \in K} |p(t_n,x,y) - p(t,x,y)|.$$

Since $\nu_n$ converges weakly, we have $\sup_n \nu_n(K) < \infty$. And since $p$ is continuous, we have $p(t_n,x,\cdot) \to p(t,x,\cdot)$ uniformly on $K$ as $t_n \to t$. So this term goes to zero. The second term goes to zero by definition of weak convergence, since $p(t,x,\cdot)$ is a continuous function on $K$.

An important fact about the heat semigroup is that $P_t f$ is the smallest of all nonnegative local weak solutions which equal $f$ at time $t = 0$. Intuitively, $P_t f$ is the solution which imposes Dirichlet conditions at the boundary of $X$, so that heat flows out of $X$ as much as possible, and no heat flows in.
Proposition 5.9. Let $u$ be a nonnegative local weak solution on $(-a,T) \times X$ for some $-a < 0$, and suppose $f \in L^2(X,\mu)$ satisfies $f \leq u(0)$ a.e. Then $P_t f \leq u$ a.e. on $[0,T) \times X$.

The proof requires considering what happens when we restrict our attention to some open subset $U$ and impose Dirichlet boundary conditions on $\partial U$. We make use of the following results which can be found in [18].

Definition 5.10. For $U \subset X$ open, let $\mathbb{D}(U) \subset \mathbb{D}$ denote the $\mathcal{E}_1$-closure of $\mathbb{D} \cap C_c(U)$.

Morally these are the functions from $\mathbb{D}$ satisfying Dirichlet boundary conditions on $\partial U$. There are several other possible equivalent definitions. Note in particular that if $f \in \mathbb{D}(U)$, then $f = 0$ $\mu$-a.e. on $U^c$. (This follows because $f$ is an $\mathcal{E}_1$-limit of functions $f_n \in C_c(U)$, and so a subsequence converges to $f$ almost everywhere. In fact, this can be upgraded to quasi-everywhere convergence.)

Lemma 5.11. Suppose $U \subset X$ is open, and we have $f \in \mathbb{D}$, $g \in \mathbb{D}(U)$ with $0 \leq f \leq g$ a.e. Then $f \in \mathbb{D}(U)$.

Proof. See Lemma 4.4 of [18].

The following fact can easily be verified:

Proposition 5.12. The restriction $(\mathcal{E},\mathbb{D}(U))$ of $\mathcal{E}$ to $\mathbb{D}(U)$ defines a regular, strictly local, symmetric Dirichlet form on $L^2(U,\mu) \subset L^2(X,\mu)$.

Proposition 5.13. Suppose $U \subset X$ is open and nice. If $u$ is a local weak solution on $(0,T) \times U$ with respect to $(\mathcal{E},\mathbb{D})$, then it is also a local weak solution on $(0,T) \times U$ with respect to $(\mathcal{E},\mathbb{D}(U))$.

Proof. Let $K \subset (0,T) \times U$ be compact; without loss of generality we can take $K = [a,b] \times K_1$. Then there exists $u_K \in W^{1,2}([a,b];\mathbb{D},\mathbb{D}^*)$ with $u_K = u$ a.e. on $K$. Set $\tilde{u}_K = \psi u_K$, where $\psi$ is a nice cutoff function of $K_1$ inside $U$. Then it is easy to see that $\tilde{u}_K \in W^{1,2}([0,T];\mathbb{D}(U),\mathbb{D}(U)^*)$. If $\phi \in W^{1,2}([0,T];\mathbb{D}(U),\mathbb{D}(U)^*)$ with compact support inside $K$, then we also have $\phi \in W^{1,2}([0,T];\mathbb{D},\mathbb{D}^*)$, and so (2.5) holds.

$\mathbb{D}(U)$ gives us a notion of “functions vanishing on $\partial U$”, which allows one to state the following parabolic maximum principle. A similar statement is proved in [18, Proposition 4.11].

Theorem 5.14. Let $U \subset X$ be open. Suppose $u \in W^{1,2}([0,T];\mathbb{D},\mathbb{D}^*)$ satisfies:

1. $u(0) \leq 0$ a.e. (recall $u \in C([0,T];L^2(X,\mu))$ so $u(0)$ is well defined as an $L^2(X,\mu)$ function);
2. $u(t)^+ \in \mathbb{D}(U)$ for a.e. $t$;
3. For every $\phi \in L^2([0,T];\mathbb{D}(U))$ which vanishes outside some $(\delta,T-\delta)$, we have

$$
\int_0^T [(u'(t),\phi(t))_{\mathbb{D}^*,\mathbb{D}} + \mathcal{E}(u(t),\phi(t))] \, dt = 0. \tag{5.4}
$$

Then $u \leq 0$ a.e. on $[0,T] \times U$. 

18
The theorem is a fairly straightforward adaptation of the argument in \[18\] and is relegated to Appendix C.

Let \( P_t^U \) be the semigroup generated by \((\mathcal{E}, \mathbb{D}(U))\). Technically it is only a semigroup on \( L^2(U, \mu) \), but it can be extended to \( L^2(X, \mu) \) in the obvious way (by defining \( P_t^U f = P_t^U (f|_U) \)). It is strongly continuous only on \( L^2(U, \mu) \). By the spectral theorem we have \( P_t^U f \in \mathbb{D}(U) \) for all \( f \in L^2(X, \mu) \).

Regularity says that the semigroup \( P_t \) (and hence the Dirichlet form) is determined by its behavior on sufficiently large open sets, as made precise in the following lemma.

**Lemma 5.15** ([18, Lemma 4.17]). If \( \{ U_i \}_{i=1}^\infty \) is an increasing sequence of open subsets of \( X \) with \( X = \bigcup_{i=1}^\infty U_i \), then for any \( t > 0 \) and \( 0 \leq f \in L^2(X, \mu) \), we have \( P_t^{U_i} f \rightarrow P_t f \) \( \mu \)-a.e.

Combining these, we may prove Proposition 5.9.

**Proof of Proposition 5.9.** Let \( u \) be a nonnegative local weak solution on \((-a, T) \times X\), and suppose \( f \in L^2(X, \mu) \) satisfies \( f \leq u(0) \) a.e. Since \( P_t f \leq P_t f_+ \), we can assume without loss of generality that \( f \geq 0 \) a.e. (We shall only apply the theorem with nonnegative \( f \) anyway.)

Let \( U \) be an open precompact set, and let \( T' < T \) be arbitrary. Choose a nonnegative function \( u_K \in W^{1,2}([0, T'] \times \overline{U}; \mathbb{D}, \mathbb{D}^*) \) which agrees with \( u \) on some compact neighborhood of \([0, T'] \times \overline{U}\) and set \( v(t) = P_t^U f - u_K(t) \in W^{1,2}([0, T']; \mathbb{D}, \mathbb{D}^*) \) (because \( W^{1,2}([0, T']; \mathbb{D}(U), \mathbb{D}(U)^*) \subset W^{1,2}([0, T']; \mathbb{D}, \mathbb{D}^*) \)). We claim \( v \) satisfies the hypotheses of Theorem 5.14. \( v(0) \leq 0 \) is clear since \( u_K(t) \) and \( P_t^U f \) are both continuous in \( L^2 \) as \( t \downarrow 0 \), and \( f \leq u(0) \). Since \( 0 \leq v^+(t) \leq P_t^U f \) for almost every \( t \), and \( P_t^U f \in \mathbb{D}(U) \), we have \( v^+(t) \in \mathbb{D}(U) \) for almost every \( t \in [0, T'] \) by Lemma 5.11. Finally, suppose \( \phi \in L^2([0, T]; \mathbb{D}(U)) \) with \( \phi(t) = 0 \) outside some \((\delta, T'-\delta)\). Since \( \phi(t) \) is supported inside \( \overline{U} \) for almost every \( t \), \( \phi \) is compactly supported inside \((0, T') \times X\). Thus 6.1, with \( T \) replaced by \( T' \), holds for \( u_K \) (since \( u \) is a local weak solution) and also for \( P_t^U f \) (by spectral theory), so it holds for \( v \). Thus Theorem 5.14 applies and we have \( v \leq 0 \) on \([0, T'] \times X\), which is to say \( u(t) \geq P_t^U f \). Letting \( T' \rightarrow T \) we have the same on \([0, T] \times U\), and outside of \( U \) we have \( u(t) \geq 0 = P_t^U f \), so in fact \( u(t) \geq P_t^U f \) on \([0, T] \times X\).

Now take an increasing sequence of open precompact subsets \( U_i \uparrow X \) (this is possible in any separable locally compact metric space). We have \( P_t^{U_i} f \leq u \) for each \( i \), and \( P_t^{U_i} f \rightarrow P_t f \) a.e. by Theorem 5.15, so the proof is complete.

In the following lemma, we show that nonnegative local weak solutions have bounded \( L^1 \) norms on compact sets \( K \) near the initial time. This says, in some sense, that heat cannot flow out of \( K \) too rapidly. We will use this fact in conjunction with weak compactness to produce the measure \( \nu \) in the main theorem.

**Proposition 5.16.** Let \( u \) be a nonnegative local weak solution on \((0, T) \times X\), and let \( K \subset X \) be compact. Then for any \( T' < T \), we have

\[
\sup_{t \in (0, T')} \int_K u(t, y) \mu(dy) < \infty.
\]

**Proof.** Fix any \( 0 < T' < T'' < T \). For any \( t \in (0, T') \), let \( v(s, x) = u(s + t, x) \), so that \( v \) is a nonnegative local weak solution on \((-t, T-t) \times X\), and \( v(0, x) = u(t, x) \geq 1_K(x)u(t, x) \). Applying
Proposition 5.9 to \( v \) with \( f(x) = 1_K(x)u(t, x) \) gives

\[
P_s[1_Ku(t, \cdot)](x) \leq v(s, x) = u(s + t, x), \quad \mu\text{-a.e. } x
\]

for any \( 0 \leq s < T - t \). Taking \( s = T'' - t \), this reads

\[
P_{T''-t}[1_Ku(t, \cdot)](x) \leq u(T'', x), \quad \mu\text{-a.e. } x. \tag{5.5}
\]

Let \( A \subset X \) be any compact set of positive measure, and set

\[
c = \inf\{p(s, x, y) : s \in [T'' - T', T''], x \in A, y \in K\}.
\]

Since we have assumed that the heat kernel \( p \) is positive and continuous on \((0, \infty) \times X \times X\), we have \( c > 0 \). Then for \( \mu\text{-a.e. } x \in A \) we have

\[
P_{T''-t}[1_Ku(t, \cdot)](x) = \int_K u(t, y)p(T'' - t, x, y) \mu(dy)
\]

\[
\geq c \int_K u(t, y) \mu(dy)
\]

and so, combining this with \( \text{(5.5)} \)

\[
\int_K u(t, y) \mu(dy) \leq \frac{1}{c} \text{essinf}_{x \in A} u(T'', x).
\]

The right side is finite and independent of \( t \in (0, T') \), so the proof is complete. \( \square \)

6 Widder’s theorem

In this section we prove our main result.

**Theorem 6.1** (Widder’s theorem, local version). Let \( U \subset X \) be open, connected, nice and precompact, and suppose \( u \) is a nonnegative local weak solution on \((0, T) \times W\) for some open neighborhood \( W \) of \( \overline{U} \). Then, there exists a unique positive Radon measure \( \nu \) supported on \( U \), and a unique nonnegative local weak solution \( h \) on \((−∞, T) \times U\) with \( h(t, \cdot) = 0 \) for \( t \leq 0 \), such that

\[
u(t, x) = P_t\nu(x) + h(t, x), \quad t \in (0, T).
\tag{6.1}

**Proof.** For \( \epsilon > 0 \), set

\[
h_{\epsilon}(t) = \begin{cases} u(t) - P_{t-\epsilon}[1_Uu(\epsilon)], & \epsilon < t < T \\ 0, & -\infty < t \leq \epsilon. \end{cases}
\]

Since \( 1_Uu(\epsilon) \in L^2(X, \mu) \), \( h_{\epsilon} \) is a local weak solution on \((\epsilon, T) \times U\) (Lemma 5.5). By Proposition 5.9 (shifting time by \( \epsilon \)), we also have \( h_{\epsilon} \geq 0 \) on \((−\infty, T) \times U\).

For any nice cutoff function \( \psi \) supported inside \( U \), we have \( \psi u \in L^2(\epsilon, T - \epsilon; \mathbb{D}) \) by definition of local weak solution. We also have \( P_t[1_Uu(\epsilon)] \in L^2(\epsilon, T - \epsilon; \mathbb{D}) \) by Lemma 5.5, so the same holds for \( \psi P_t[1_Uu(\epsilon)] \). And by strong continuity of the heat semigroup \( P_t \), we have \( \psi h_{\epsilon}(t) \to 0 \) in \( L^2(X, \mu) \) as \( t \to \epsilon \). So by Lemma 5.2 \( h_{\epsilon} \) is a local weak solution on \((−\infty, T) \times U\).
Now by Proposition 5.16 with $K = \overline{U}$, we have that $1_U u(\epsilon)$ is bounded in $L^1$ norm as $\epsilon \downarrow 0$, or equivalently, that the Radon measures $1_U u(\epsilon) d\mu$ are bounded in total variation. Hence by compactness, there is a sequence $\epsilon_n \downarrow 0$ and a positive Radon measure $\nu$ with $1_U u(\epsilon_n) d\mu \to d\nu$ weakly. By Lemma 5.8, we have $P_{t-\epsilon_n}[1_U u(\epsilon_n)](x) \to P_t \nu(x)$ pointwise. Thus $h_{\epsilon_n}(t, x) \to h(t, x) = u(t, x) - P_t \nu(x)$ pointwise, where we take $h(t, x) = 0$ for $t \leq 0$.

We now apply the parabolic Harnack inequality (Assumption 2). Fix Theorem 6.2, which shows that $\nu$ is compact, there is a sequence $\nu_k$ of Radon measures on $U$ that converges to $\nu$ in the weak-* topology of $C_c(U)^\ast$. Consequently, there exists a unique positive Radon measure $\nu$ on $U$ such that $\int_U f d\nu = \int_U f d\mu$ for any bounded, continuous function $f$.

For uniqueness, suppose we have another decomposition $u = P_t \nu + h$. By the Harnack inequality, for each $h$, we have

$$\esssup_{[a,b] \times \overline{V}} h \leq C \essinf_{[c,d] \times \overline{V}} h \leq C \essinf_{[c,d] \times \overline{V}} u$$

since $h \leq u$. Since the bound is independent of $\epsilon$, we can apply Lemma 5.4 to find that $h = \lim_{\epsilon \to 0} h_{\epsilon_n}$ is a local weak solution on $(a, b) \times V$, and hence (since $a, b, V$ were arbitrary) on $(-\infty, T) \times U$. This completes the proof of existence.

To show uniqueness of $\nu$ and $h$, fix $f \in C_c(U)$. Lemma 5.6 says that we have $\int_X P_t \nu(x) f(x) \mu(dx) \to \int_X f d\nu$ as $t \to 0$. Since $h$ is a local weak solution on $(-\infty, T) \times U$, $h$ is continuous in $L^2_{\text{loc}}(U)$, and since $h$ vanishes for $t \leq 0$, we have $\int_X h(t, x) f(x) \mu(dx) \to 0$. Thus $\int_X u(t, x) f(x) \mu(dx) \to \int f d\nu$, which shows that $\nu$, and therefore also $h$, is uniquely determined by $u$. In fact, we have shown that $u(t) \to \nu$ weakly on $U$ (i.e. in the weak-* topology of $C_c(U)^\ast$).

\[\square\]

**Theorem 6.2 (Widder’s theorem, global version).** Let $u$ be a nonnegative local weak solution on $(0, T) \times X$. There exists a unique positive Radon measure $\nu$ (possibly infinite), and a unique nonnegative local weak solution $h$ on $(-\infty, T) \times X$ with $h(t, x) = 0$ for $t \leq 0$, such that

$$u(t, x) = P_t \nu(x) + h(t, x), \quad t \in (0, T).$$

**Proof.** Let $U_n$ be an increasing exhaustion of $X$ by open, precompact, connected, nice sets, and for each $U_n$ let $u(t) = P_t \nu + h_n(t)$ be the unique decomposition produced by Theorem 6.1. As we previously argued, $u(t) \to \nu$ weakly on $U_n$ as $t \downarrow 0$, and thus for $m > n$, we have $\nu_m = \nu_n$ on $U_n$. In particular, the measures $\nu_m$ are increasing. Their limit $\nu$ is another positive Radon measure, possibly infinite, and by monotone convergence we have $P_t \nu_m \uparrow P_t \nu$. Thus $h_n \downarrow h = u - P_t \nu$. $h$ remains a nonnegative function which vanishes for $t \leq 0$.

Moreover, if $V$ is any precompact open set, we have $\overline{V} \subset U_n$ for sufficiently large $n$. Fixing $-\infty < a < 0 < b < c < d$, we have by the Harnack inequality

$$\esssup_{[a,b] \times \overline{V}} h \leq C \essinf_{[c,d] \times \overline{V}} h \leq C \essinf_{[c,d] \times \overline{V}} u$$

Thus applying Lemma 5.4, $h$ is a local weak solution on $(a, b) \times V$, and hence on $(-\infty, T) \times X$.

For uniqueness, suppose we have another decomposition $u(t) = P_t \tilde{\nu} + \tilde{h}(t)$. Let $f \in C_c(X)$ be nonnegative; $f$ is supported in one of the $U_n$, so as $t \to 0$ we have

$$\int_X u(t, x) f(x) \mu(dx) \to \int_{U_n} f(x) d\nu_n = \int_X f d\nu.$$ 

Thus since $\tilde{h}(t) \to 0$ in $L^2_{\text{loc}}(X)$, we also have

$$\int_X P_t f d\tilde{\nu} = \int_X P_t \tilde{\nu}(x) f(x) \mu(dx) \to \int_X f d\nu.$$
Since $P_t f \to f$ pointwise, Fatou’s lemma gives $\int_X f \, d\nu \leq \int_X f \, dv$. Thus we have $\tilde{\nu} \leq \nu$, so $\eta = \nu - \tilde{\nu}$ is a positive Radon measure. Since $P_t \eta = \tilde{h}(t) - \tilde{h}(t)$, we have that $P_t \eta \to 0$ in $L^2_{loc}(X, \mu)$ as $t \to 0$.

If $K$ is any compact set and $\eta|_K$ is the restriction of $\eta$ to $K$, then $0 \leq P_t \eta|_K \leq P_t \eta$, so we also have $P_t \eta|_K \to 0$ in $L^2_{loc}$. However, since $\eta|_K$ is a compactly supported Radon measure, Lemma 5.6 gives $P_t \eta|_K \to \eta|_K$ weakly, so $\eta|_K = 0$. Letting $K \uparrow X$, we have $\eta = 0$, and thus $\nu = \tilde{\nu}$ so the decomposition is unique. □

7 Conditions for uniqueness of nonnegative solutions

In the classical version of Widder’s theorem (for the classical heat equation on $\mathbb{R}^d$), the function $h$ appearing in (6.2) is actually zero, and the theorem just states that $u(t, x) = P_t \nu(x)$. Thus in the classical case, a nonnegative solution of the heat equation is uniquely determined by its initial values. However, in our general setting, $h$ can certainly be nonzero. For example, let $X = (0, \infty)$ be the open half-line, with the classical Dirichlet form $\mathcal{E}(f, g) = \frac{1}{2} \int_0^\infty f'g' \, dm$ with its domain $\mathcal{D} = H^1_0((0, \infty))$. Then $u(t, x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$ is a local weak solution, but it is easy to see that the decomposition in (6.2) must have $\nu = 0$ and $h = u$.

We record here some conditions that are, or are not, necessary or sufficient to guarantee that $h = 0$.

1. The Dirichlet space $(X, \mu, \mathcal{E}, \mathcal{D})$ is said to be conservative (or stochastically complete) if $P_1 = 1$, or equivalently if the corresponding Hunt process $X_t$ has an infinite lifetime, almost surely. This condition is necessary, but not sufficient, to ensure $h = 0$.

To see it is necessary, observe that $v(t, x) = 1 - (P_1)(x)$ satisfies the hypotheses of Lemma 5.2 and hence can be regarded as a nonnegative local weak solution on $(-\infty, \infty) \times X$. Applying Theorem 6.2 to $u(t, x) = v(t - 1, x)$, we have $v(t - 1, x) = P_t \nu(x) + h(t, x)$. If $h = 0$ then $v(t - 1, x) = P_t \nu(x)$, but since this vanishes for all $0 \leq t \leq 1$, we must have $\nu = 0$ and hence $v = 0$ identically.

To see it is not sufficient, see the next example.

2. A stronger condition is that the Dirichlet space, or equivalently its corresponding Hunt process, be recurrent. This is also not sufficient to ensure $h = 0$. Consider $X = \mathbb{R}^2 \setminus \{0\}$ with the classical Dirichlet form $\mathcal{E}(f, g) = \frac{1}{2} \int \nabla f \cdot \nabla g \, dm$ and its domain $\mathcal{D} = H^1_0(X)$. Since points are polar for Brownian motion in $\mathbb{R}^2$, this is a recurrent Dirichlet space. However, it is not hard to see that

$$u(t, x) = \begin{cases} \frac{1}{2\pi(t-1)} e^{-|x|^2/(2(t-1))}, & t > 1 \\ 0, & t \leq 1 \end{cases}$$

is a nonnegative local weak solution on $(0, \infty) \times X$. Since it vanishes for $0 \leq t \leq 1$, its decomposition according to Theorem 6.2 must have $\nu = 0$, so $h$ cannot be 0.

Recurrence is not necessary, as can be seen by considering the classical Dirichlet form on $\mathbb{R}^d$, $d \geq 3$. The fact that $h = 0$ in this case is Widder’s original theorem; it is also included in the Harnack-type case below.

3. Completeness under an intrinsic metric is not sufficient to ensure $h = 0$. See [4, Section 7.7] for an example of a complete two-dimensional Riemannian manifold with unbounded negative
curvature, such that the Brownian motion explodes in finite time with positive probability. This corresponds to a strictly local Dirichlet space which is not conservative.

4. If \( X \) is compact, then any local weak solution \( u \) on \( I \times X \) is actually a global weak solution, because we can take \( K = [a, b] \times X \) in Definition 2.3. In particular, we have \( u(t) \in \mathbb{D} \) for all \( t \). So we can apply the maximum principle (Theorem 5.14) to \( h \) and immediately conclude that \( h = 0 \). In fact, when \( X \) is compact, many of the arguments in this paper become much simpler.

5. Under the basic assumptions made in Section 3, if we assume further that either (ID2) holds locally uniformly or that (H-\( \beta \)) holds locally uniformly then we can follow the elegant argument of [22]. A nonnegative local weak solution \( u \) is said to be \textit{minimal} if the only local weak solutions \( v \) satisfying \( 0 \leq v \leq u \) are of the form \( v = \lambda u \); the Choquet representation theorem says that any solution has an integral representation in terms of minimal solutions. Suppose, then, that \( h \) is a nonnegative local weak solution which vanishes for \( t \leq 0 \); it has a representation in terms of minimal solutions \( \tilde{h} \) that also vanish for \( t \leq 0 \). However, it follows from the locally uniform parabolic Harnack inequality that for sufficiently small \( \epsilon \), we have \( \tilde{h}(t - \epsilon, x) \leq H_0 \tilde{h}(t, x) \); thus by minimality \( \tilde{h}(t - \epsilon, x) = \lambda \tilde{h}(t, x) \) and we conclude that \( \tilde{h} \) vanishes for \( t \leq \epsilon \). By iteration, \( \tilde{h} \) vanishes everywhere, and so the same must be true of \( h \).

In fact, [22] proves the much stronger statement that nonnegative minimal weak solutions \( u \) of (2.2) on \( (-\infty, T) \times X \) are in fact of the form
\[
  u(t, x) = e^{\alpha t} v(x)
\]
where \( v \) is a nonnegative minimal weak solution of \( L v = \alpha v \) on \( X \).

The question of whether a nonnegative solution of the heat equation is uniquely determined by its initial values has been studied by many authors in various settings. In addition to [44] and [2, 3], we mention [20, 30, 31, 22].

8 An application to projections

In this short section we outline what we think is a compelling application of our main result to the study of the projection of one Dirichlet space onto another.

Let \( (X_i, d_i, \mu_i, E_i, \mathbb{D}_i), i = 1, 2 \) be two Dirichlet spaces satisfying the assumptions of Section 3. Assume further that \( (X_i, d_i), i = 1, 2 \) are complete. We are interested in considering the situation when there exists a continuous projection map \( \pi : X_1 \to X_2 \) with the following properties:

(P1) For any \( x, y, \tilde{x} \) with \( x, y \in X_2, \tilde{x} \in X_1 \) with \( \pi(\tilde{x}) = x \),
\[
d_2(x, y) = \min\{d_1(\tilde{x}, \tilde{y}) : \tilde{y} \in \pi^{-1}\{y\}\}.
\]

(P2) If \( u \) is a local weak solution of (2.2) in \( (0, T) \times U \) on \( X_2 \) then \( v(t, x) = u(t, \pi(x)) \) is a local weak solution of (2.2) in \( (0, T) \times \tilde{U} \) on \( X_1 \) where \( \tilde{U} = \pi^{-1}(U) \).

(P3) A Borel set \( A \subset X_2 \) is \( \mu_2 \)-negligible if and only if \( \pi^{-1}(A) \) is \( \mu_1 \)-negligible.
If $B_i(x, r)$ denotes the ball of radius $r$ around $x$ in $X_i$, $i = 1, 2$, then (P1) implies that
\[ \pi^{-1}(B_2(x, r)) \supset B_1(\tilde{x}, r) \]
and that
\[ \pi(B_1(\tilde{x}, r)) \subset B_2(x, r) \]
for any $\tilde{x}$ such that $\pi(\tilde{x}) = x$.
Condition (P3) implies that we can disintegrate $\mu_1$ with respect to $\mu_2$; that is, there exists a family of measures $\nu_\pi^z$ on $X_1$, $z \in X_2$, with $\nu_\pi^z$ supported in $\pi^{-1}(z)$ and such that
\[ \int_{X_1} f(x) \, d\mu_1(x) = \int_{X_2} \int_{X_1} f(x) \, d\nu_\pi^z(x) \, d\mu_2(z) \quad (8.1) \]
for any nonnegative measurable $f$ on $X_1$. We will assume that this disintegration formula has the following continuity property.

(P4) For $f \in C_c(X_1)$, the measurable compactly supported function
\[ z \mapsto \int f \, d\nu_\pi^z \]
admits a continuous version.

**Theorem 8.1.** Referring to the setup introduced above, assume that (P1), (P2), (P3) and (P4) hold true and that $(X_1, d_1, \mu_1, \mathcal{E}_1, \mathbb{D}_1)$ satisfies the parabolic Harnack inequality (H-\(\beta\)) of Section 4 locally uniformly. Then the same is true for $(X_2, d_2, \mu_2, \mathcal{E}_2, \mathbb{D}_2)$. Furthermore, the two heat kernels are related by
\[ p_2(t, x, y) = \int_{X_1} p_1(t, \tilde{x}, z) \, d\nu_\pi^y(z) \]
where $\tilde{x}$ is such that $\pi(\tilde{x}) = x$.

**Proof.** The first assertion follows immediately by inspection, using (P1) and (P2) to lift a local solution on $X_2$ to a local solution on $X_1$.
Consider
\[ (t, \tilde{z}) \mapsto u_y(t, \tilde{z}) = p_2(t, \pi(\tilde{z}), y). \]
This is a nonnegative local weak solution of (2.2) on $(0, \infty) \times X_1$. By Theorem 6.2 (with $h \equiv 0$ because of the validity of the local uniform parabolic Harnack inequality) there exists a nonnegative Radon measure $\omega_y$ on $X_1$ such that, for all $\tilde{z} \in X_1$ and $t > 0$,
\[ u_y(t, \tilde{z}) = \int_{X_1} p_1(t, \tilde{z}, \zeta) \, d\omega_y(\zeta) \]
and, for any $f \in C_c(X_1)$, as $t$ tends to 0,
\[ \int_{X_1} u_y(t, \zeta) f(\zeta) \, d\mu_1(\zeta) = \int_{X_1} p_2(t, \pi(\zeta), y) f(\zeta) \, d\mu_1(\zeta) \to \int_{X_1} f(\zeta) \, d\omega_y(\zeta). \]

24
Now,
\[ \int_{X_1} p_2(t, \pi(\zeta), y) f(\zeta) \, d\mu_1(\zeta) = \int_{X_2} \left( \int_{X_1} f(\nu_\pi_z) \right) p_2(t, z, y) \, d\mu_2(z). \]

Since \( z \mapsto \int_{X_1} f(\nu_\pi_z) \) admits a continuous version, we see that
\[ \int_{X_1} p_2(t, \pi(\zeta), y) f(\zeta) \, d\mu_1(\zeta) \to \int_{X_1} f(\nu_\pi_y). \]

In other words, the Radon measure \( \omega_y \) is, in fact, equal to \( \nu_\pi_y \).

Theorem 8.1 is surprising and interesting even in the simplest cases. Consider for instance the case when \( X_1 \) and \( X_2 \) are complete Riemannian manifolds, each equipped with its natural Dirichlet space structure, and \( \pi \) is the projection associated with a countable group \( G \) of isometries acting properly and freely on \( X_1 \). The hypotheses (P1)–(P4) are clearly satisfied (the measure \( \nu_\pi_z \) is the counting measure on the countable set \( \pi^{-1}(z) \).) The theorem says that the heat kernels on \( X_1 \) and \( X_2 \) are related by
\[ p_2(t, \pi(x), \pi(y)) = \sum_{g \in G} p_1(t, x, gy). \]

This statement includes the non-trivial fact that the sum on the right hand side is finite.

Another illustrative application of Theorem 8.1 is to relate the Gaussian heat kernel of Brownian motion on \( \mathbb{R}^n \) to its radial part, the Bessel process, which is associated with an explicit Dirichlet space on the semi-axis (this requires a proper treatment of the point 0, depending on dimension). See [11, page 126]. In this case, the group action is the action of the rotation group.

The setting of Dirichlet spaces allow us to treat in exactly the same way the very natural case that arises when \( X_1 \) and \( X_2 \) are polytopal complexes (satisfying mild assumptions, see [33, 13]) instead of Riemannian manifolds. See also [7, 8] for examples involving the Dirichlet forms of Section 4.3 on \( \mathbb{R}^k \times \mathbb{T}_\infty \).

Theorem 8.1 can be applied in a wide variety of contexts where the projection \( \pi \) is associated to the proper continuous action of a locally compact group \( G \) on \( X_1 \) that preserves the distance \( d_1 \). See for instance [9], especially Corollary 4.6 and Section 6, for descriptions of concrete examples.

We end this section by specializing Theorem 8.1 in the context of sub-Riemannian diffusions on unimodular groups.

**Theorem 8.2.** Let \( G \) be a unimodular Lie group equipped with its Haar measure and a family \( \{X_1, \ldots, X_k\} \) of left invariant vector fields generating the Lie algebra of \( G \). Let \( H \) be a closed subgroup of \( G \) equipped with its Haar measure. Let \( \pi: G \to M = H \backslash G \) be the projection on the quotient space \( M \) of right-cosets. Equip \( M \) with its natural \( G \)-invariant measure. Let \( L_G = \sum_{i=1}^k X_i^2 \) and \( L_M = \sum_{i=1}^k [d\pi(X_i)]^2 \) be the associated hypoelliptic sub-Laplacians on \( G \) and \( M \), respectively. Then the heat kernels on \( M \) and \( G \) are related by
\[ p_M(t, x, y) = \int_H p_G(g_x, hg_y) d_H h \]
where \( \pi(g_x) = x, \pi(g_y) = y \) and \( d_H h \) is the Haar measure on \( H \). In particular, the right-hand side is finite.

For background information regarding the setting of this theorem, see [43, 25].
9 Acknowledgements

The authors wish to thank Benjamin Steinhurst for many helpful discussions. We would also like to thank the anonymous referee for a very careful reading of the paper which led to numerous corrections and improvements.

N. Eldredge was partially supported by National Science Foundation grant DMS-0739164. L. Saloff-Coste was partially supported by National Science Foundation grant DMS-1004771.

A Energy measures

Let \((E, D)\) be a regular, strictly local Dirichlet form on \(L^2(X, \mu)\).

Recall that \(D \cap L^{\infty}(X, \mu)\) is an algebra, and for \(f, g \in D \cap L^{\infty}\) we have

\[
\sqrt{E(fg)} \leq \|f\|_{\infty} \sqrt{E(g)} + \|g\|_{\infty} \sqrt{E(f)}.
\]

(See [16, Theorem 1.4.2 (ii)].)

For \(f \in D \cap L^{\infty}\), we can define a Radon measure \(\Gamma(f)\) on \(X\) by taking

\[
\int_X \phi d\Gamma(f) := 2E(\phi f, f) - E(f^2, \phi)
\]

for \(\phi \in C_c(X) \cap D\). (In the classical case, \(d\Gamma(f) = |\nabla f|^2 dm\), where \(m\) is Lebesgue measure.)

To see that \((A.1)\) in fact defines a Radon measure, i.e. a continuous linear functional on \(C_c(X)\), set

\[
\alpha_t(\phi) := \frac{1}{t} [2(f\phi, P_t f - f) - (P_t(f^2) - f^2, \phi)]
\]

and note that

\[
|\alpha_t(\phi)| \leq \frac{1}{t} [2|\phi| \|P_t(f - \phi)| + \|P_t(f^2) - f^2||_{L^1}||\phi||_{\infty}
\]

where \(\|P_t(f^2) - f^2||_{L^1} < \infty\) because \(f \in L^2\), so \(f^2 \in L^1\), and \(P_t\) is a contraction on \(L^1\) (which follows from the Markovian property). So \(\alpha_t\) is a bounded linear functional on \(C_c(X)\), and as \(t \to 0\), \(\alpha_t(\phi) \to 2E(\phi f, f) - E(f^2, \phi)\). By the uniform boundedness principle, a pointwise limit of bounded linear functionals is another bounded linear functional.

One may then define the signed measure \(\Gamma(f, g) = \frac{1}{2}(\Gamma(f + g) - \Gamma(f) - \Gamma(g))\) by polarization, where \(\Gamma(f, f) = \Gamma(f)\).

Note that for \(f, g \in D \cap L^{\infty}\), the integral \(\int_X f d\Gamma(g)\) needs some care to be well-defined, since \(f\) is technically only defined up to \(\mu\)-null sets, which \(\Gamma(g)\) may charge. However, a quasi-continuous \(\mu\)-version \(\tilde{f}\) of \(f\) is uniquely defined up to polar sets, which \(\Gamma(g)\) does not charge. So \(\int_X f d\Gamma(g)\) should be interpreted as \(\int_X \tilde{f} d\Gamma(g)\).

Note that [15] discusses energy measures for additive functionals, but this is a generalization of the energy measure of a function. Note also in the strictly local case we have, in the notation of [15], \(\mathcal{E}^{\text{res}} = \mathcal{E}\).

Some properties of \(\Gamma\) which we shall use:

26
Proposition A.1 ([15, Lemma 5.4.2]). For \( f, g, h \in D \cap L^\infty \), \( d\Gamma(fg, h) = fd\Gamma(g, h) + gd\Gamma(f, h) \).

Proposition A.2 ([15, Lemma 5.4.3]). For \( f, g, h, k \in D \cap L^\infty \), we have the Cauchy–Schwarz inequality
\[
\left( \int_X |fg| \, d|\Gamma(h, k)| \right)^2 \leq \int_X f^2 \, d\Gamma(h) \int_X g^2 \, d\Gamma(k).
\]

We remark that using the AM-GM inequality, we have the useful form
\[
\left| \int_X fg \, d\Gamma(h, k) \right| \leq \frac{1}{2} \left( \int_X f^2 \, d\Gamma(h) + \int_X g^2 \, d\Gamma(k) \right).
\] (A.2)

The latter form will be more useful to us.

Corollary A.3. For \( f, g \in D \cap L^\infty \),
\[
d\Gamma(fg) \leq 2(f^2d\Gamma(g) + g^2d\Gamma(f)).
\]

Proof. Fix \( h \in D \cap C_c(X) \) with \( h \geq 0 \). By repeated application of the Leibniz rule (Proposition A.1),
\[
\int_X h \, d\Gamma(fg) = \int_X hf^2 \, d\Gamma(g) + \int_X hg^2 \, d\Gamma(f) + 2 \int_X hfg \, d\Gamma(f, g).
\]
Writing \( hfg \) as \((f\sqrt{h})(g\sqrt{h})\) and applying (A.2), we have what we want. \( \square \)

B Cutoff Sobolev inequalities

In [6], the notion of a (local) cutoff Sobolev inequality is defined. Here \( X \) is assumed to be a strictly local Dirichlet space. \( X \) was also assumed to be a metric space; we write \( d \) for the metric and \( B(x, r) \) for the open metric balls. Note that \( d \) is not assumed to be an intrinsic distance in the sense of the previous section.

Definition B.1 ([6]). \( X \) is said to satisfy a local cutoff Sobolev inequality \( \text{CS}(\beta)_{\text{loc}} \) if there exists \( \theta \in (0, 1] \) and constants \( c_1, c_2 \) such that for every \( x_0 \in X \), \( 0 < R \leq 1 \), there exists a cutoff function \( \psi \) with the properties:

1. \( \psi \geq 1 \) on \( B(x_0, R/2) \);
2. \( \psi = 0 \) on \( B(x_0, R) \); c;
3. \( |\psi(x) - \psi(y)| \leq c_1 (d(x, y)/R)^\theta \) for all \( x, y \in X \);
4. For any ball \( B(x, s) \) with \( 0 \leq s \leq R \) and \( f \in D \),
\[
\int_{B(x, s)} f^2 \, d\Gamma(\psi) \leq c_2 (s/R)^{2\theta} \left( \int_{B(x, 2s)} d\Gamma(f) + s^\beta \int_{B(x, 2s)} f^2 \, d\mu \right). \tag{B.1}
\]
Note that by replacing \( \psi \) with \( \tilde{\psi} = \psi \wedge 1 \vee 0 \), we can assume \( 0 \leq \tilde{\psi} \leq 1 \). (If \( \psi \) satisfies the above condition, so does \( \tilde{\psi} \): we have \( |\tilde{\psi}(x) - \tilde{\psi}(y)| \leq |\psi(x) - \psi(y)| \), and the Markovian properties of \( (\mathcal{E}, \mathbb{D}) \) implies \( d\Gamma(\tilde{\psi}) \leq d\Gamma(\psi) \).)

We remark that if \( \psi \) satisfies (B.1), then for any \( f \in \mathbb{D} \), we have \( f\psi \in \mathbb{D} \). We clearly have \( f\psi \in L^2(X, \mu) \) since \( \psi \) is bounded. If we take \( s = R \) in (B.1), then the left side becomes \( \int_X f^2 d\Gamma(\psi) \) and the right side is controlled by \( E_1(f) \). If \( f \in \mathbb{D} \cap L^\infty \) we can apply Corollary A.3 and see that \( E_1(f\psi) \) is controlled by \( E_1(f) \). Since by the Markovian property, \( \mathbb{D} \cap L^\infty \) is \( E_1 \)-dense in \( \mathbb{D} \), an approximation argument shows \( f\psi \in \mathbb{D} \) for any \( f \in \mathbb{D} \).

**Lemma B.2.** If \( X \) satisfies \( \text{CS}(\beta)_{\text{loc}} \), then all open sets in \( X \) are nice, and \( X \) satisfies Assumption \( \mathbb{D} \)

*Proof.* Let \( U \subset X \) be open, and \( K \subset U \) compact. By compactness of \( K \) and local compactness of \( X \), we can cover \( K \) by a finite number of balls \( B(x_i, R_i/2) \) such that \( R_i \leq 1 \) and \( B(x_i, R_i) \subset U \). Let \( \psi_i : X \to [0, 1] \) be the corresponding cutoff functions as provided by Definition B.1. If \( f \in \mathbb{D} \), then we have argued that \( f\psi_i \in \mathbb{D} \) for each \( i \). Now taking \( \psi = \max \psi_i \) gives a nice cutoff function for \( K \) inside \( U \), since \( f\psi = \max \{ f^+\psi_i \} - \max \{ f^-\psi_i \} \in \mathbb{D} \) as well.

## C Maximum principle

In this appendix we give a proof of Theorem 5.14 adapted from [18], for local weak solutions.

**Lemma C.1.** Let \( u \in C^1([0, T]; L^2(X, \mu)) \), and suppose \( \Phi \in C^1(\mathbb{R}) \) satisfies \( \Phi(0) = 0 \) and \( |\Phi'| \leq 1 \). Then \( \Phi \circ u \in C([0, T]; L^2(X, \mu)) \), and \( \Phi \circ u \) is differentiable with \( (\Phi \circ u)'(t, x) = \Phi'(u(t, x))u'(t, x) \).

*Proof.* Note that \( \Phi \) is Lipschitz, so if \( f \in L^2 \), then \( |\Phi \circ f| \leq |f| \) and so \( \Phi \circ f \in L^2 \). Also, if \( f, g \in L^2 \), then \( |\Phi \circ f - \Phi \circ g| \leq |f - g| \) and so \( f \mapsto \Phi \circ f \) is a continuous function on \( L^2 \) (indeed, Lipschitz). Thus \( \Phi \circ u \in C([0, T]; L^2) \).

For the derivative, we must show that for each \( t \in (0, T) \) and every real sequence \( \epsilon_n \to 0 \),

\[
\lim_{\epsilon_n \to 0} \frac{\Phi(u(t + \epsilon_n, \cdot)) - \Phi(u(t, \cdot))}{\epsilon_n} = \Phi'(u(t, \cdot))u'(t, \cdot)
\]

(C.1)

with the convergence in \( L^2 \). Let us write

\[
\frac{\Phi(u(t + \epsilon_n, x)) - \Phi(u(t, x))}{\epsilon_n} = \frac{\Phi(u(t, x) + \epsilon_n u'(t, x)) - \Phi(u(t, x))}{\epsilon_n} + \frac{\Phi(u(t + \epsilon_n, x)) - \Phi(u(t, x) + \epsilon_n u'(t, x))}{\epsilon_n}.
\]

The first term converges to \( \Phi'(u(t, x))u'(t, x) \) pointwise, and since \( \Phi \) is Lipschitz we also have

\[
\frac{\Phi(u(t, x) + \epsilon_n u'(t, x)) - \Phi(u(t, x))}{\epsilon_n} \leq |u'(t, x)| \in L^2
\]
We remark in particular that for \( f \) and \( v \)

Thus the lemma holds for such \( L \)

Also, it is shown in [18, Lemma 4.3] that

L

\[ \|v(t + \epsilon_n, x) - \Phi(u(t, x) + \epsilon_n u'(t, x))\|_{L^2} \leq \|u(t + \epsilon_n, x) - u(t, x) - \epsilon_n u'(t, x)\|_{L^2} \]

which goes to 0 in \( L^2 \) by differentiability of \( u \).

\[ \left( \int_{-\infty}^{s} \varphi(\xi) d\xi \right) \]

and \( \Phi(s) = \left( \int_{-\infty}^{s} \varphi(\xi) d\xi \right)^{1/2} \), so that \( \varphi = 2\Phi\Phi' \). Then one can verify that:

\[ \begin{align*}
&\bullet \varphi, \Phi \in C^1(\mathbb{R}); \\
&\bullet \varphi > 0 \text{ and } \Phi > 0 \text{ on } (0, \infty); \\
&\bullet 0 \leq \varphi' \leq 1 \text{ and } 0 \leq \Phi' \leq 1 \text{ on } \mathbb{R}.
\end{align*} \]

We remark in particular that for \( f \in \mathbb{D} \), \( \varphi \circ f \) is a normal contraction of \( f \), and thus \( \mathcal{E}_1(\varphi \circ f) \leq \mathcal{E}_1(f) \). Also, it is shown in [18, Lemma 4.3] that

\[ \mathcal{E}(f, \varphi \circ f) \geq \mathcal{E}(\varphi \circ f) \geq 0. \]  

(C.2)

**Lemma C.3.** For any \( v \in W^{1,2}([0, T]; \mathbb{D}, \mathbb{D}^*) \), the function \( t \mapsto \|\Phi(v(t, \cdot))\|^2_{L^2} \) is absolutely continuous, and

\[ \frac{d}{dt} \|\Phi(v(t, \cdot))\|^2_{L^2} = \langle v'(t), \varphi(v(t, \cdot)) \rangle_{\mathbb{D}^*, \mathbb{D}}. \]

**Proof.** Suppose first that \( v \in C^1([0, T]; \mathbb{D}) \). We have by the product rule and Lemma [C.1] that

\[ \frac{d}{dt} \|\Phi(v(t, \cdot))\|^2_{L^2} = 2 \int_X \Phi'(v(t, x)) v'(t, x) \Phi(v(t, x)) \mu(dx) \\
= \int_X v'(t, x) \varphi(v(t, x)) \mu(dx) = \langle v'(t), \varphi(v(t, \cdot)) \rangle_{\mathbb{D}^*, \mathbb{D}}. \]

Thus the lemma holds for such \( v \). Integrating by parts gives, for any \( \chi \in C^\infty_c((0, T)), \)

\[ \int_0^T \|\Phi(v(t, \cdot))\|^2_{L^2} \chi'(t) dt = - \int_0^T \chi(t) \langle v'(t), \varphi(v(t, \cdot)) \rangle_{\mathbb{D}^*, \mathbb{D}} dt. \]  

(C.3)

Now suppose \( v \in W^{1,2}([0, T]; \mathbb{D}, \mathbb{D}^*). \) Since \( C^1([0, T]; \mathbb{D}) \) is dense in \( W^{1,2} \), we can find a sequence \( v_n \in C^1([0, T]; \mathbb{D}) \) with \( v_n \to v \) in \( W^{1,2} \), and hence also in \( C([0, T]; L^2(X, \mu)), \) which is to say \( \|v(t) - v_n(t)\|_{L^2} \to 0 \) uniformly in \( t \). Since \( \Phi \) is Lipschitz, we also have \( \|\Phi(v(t, \cdot)) - \Phi(v_n(t, \cdot))\|_{L^2} \to 0 \) uniformly in \( t \), and it follows that

\[ \int_0^T \|\Phi(v_n(t, \cdot))\|^2_{L^2} \chi'(t) dt \to \int_0^T \|\Phi(v(t, \cdot))\|^2_{L^2} \chi'(t) dt. \]
We also have \( v'_{n} \to v' \) in \( L^{2}([0,T];\mathbb{D}^{*}) \) and \( v_{n} \to v \) in \( L^{2}([0,T];\mathbb{D}) \). In particular, passing to a subsequence, we have \( v_{n}(t) \to v(t) \) in \( \mathcal{E}_{1}^{1/2} \)-norm and \( v'_{n}(t) \to v'(t) \) in \( \mathbb{D}^{*} \)-norm for almost every \( t \).

Now by \cite{16}, Theorem 1.4.2 (v), we have \( \varphi(v_{n}(t,\cdot)) \to \varphi(v(t,\cdot)) \) \( \mathcal{E}_{1} \)-weakly for almost every \( t \). Thus \( (v'_{n}(t),\varphi(v_{n}(t,\cdot)))_{\mathbb{D}^{*},\mathbb{D}} \to (v'(t),\varphi(v(t,\cdot)))_{\mathbb{D}^{*},\mathbb{D}} \) for almost every \( t \). Now since \( \varphi(v_{n}(t,\cdot)) \) is a normal contraction of \( v_{n}(t,\cdot) \), we have

\[
|v'(t),\varphi(v_{n}(t,\cdot))|_{\mathbb{D}^{*},\mathbb{D}} \leq |v'(t)|_{\mathbb{D}^{*}} \sqrt{\mathcal{E}_{1}(\varphi(v_{n}(t,\cdot)))} \leq |v'(t)|_{\mathbb{D}^{*}} \sqrt{\mathcal{E}_{1}(v_{n}(t))}.
\]

Since \( v_{n} \to v \) in \( W^{1,2} \), it follows that the expression on the right side converges in \( L^{1}([0,T]) \). Thus by a variant of the dominated convergence theorem, we have

\[
\int_{0}^{T} \chi(t)(v'_{n}(t),\varphi(v_{n}(t,\cdot)))_{\mathbb{D}^{*},\mathbb{D}} dt \to \int_{0}^{T} \chi(t)(v'(t),\varphi(v(t,\cdot)))_{\mathbb{D}^{*},\mathbb{D}} dt.
\]

By passing to the limit, we have shown that (C.3) holds for all \( v \in W^{1,2}([0,T];\mathbb{D},\mathbb{D}^{*}) \), which implies the desired result.

\[ \square \]

Proof of Theorem 5.14. Let \( \chi \in C_{c}^{\infty}((0,T)) \), and set \( \phi(t,x) = \chi(t)\varphi(u(t,x)) \). Since \( \varphi \) is a normal contraction and \( \varphi(u) = \varphi(u^{+}) \), we have \( \phi(t) \in \mathbb{D}(U) \) for a.e. \( t \). Also, since \( \mathcal{E}_{1}(\varphi(u(t))) \leq \mathcal{E}_{1}(u(t)) \), we have \( \phi \in L^{2}([0,T];\mathbb{D}(U)) \), and \( \phi(t) \) vanishes outside the support of \( \chi \). Thus (5.4) holds for \( \phi \); that is,

\[
\int_{0}^{T} \chi(t)(u'(t),\varphi(u(t,\cdot)))_{\mathbb{D}^{*},\mathbb{D}} dt = -\int_{0}^{T} \chi(t)\mathcal{E}(u(t),\varphi(u(t,\cdot))) dt.
\]

Now \( \chi \) was arbitrary, so we must have \( (u'(t),\varphi(u(t,\cdot)))_{\mathbb{D}^{*},\mathbb{D}} = -\mathcal{E}(u(t),\varphi(u(t))) \) for a.e. \( t \). By Lemma C.3 and (C.2), if we write \( a(t) = \|\Phi(u(t,\cdot))\|_{L^{2}}^{2} \), this says \( a'(t) \leq 0 \). But \( a(0) = 0 \) since \( u(0) \leq 0 \), and \( a \geq 0 \) by definition, so we must have \( a = 0 \) identically. So we have \( \Phi(u(t,\cdot)) = 0 \) a.e. Since \( \Phi(s) > 0 \) for all \( s > 0 \), it must be that \( u \leq 0 \) a.e.

\[ \square \]

References

[1] A. Ancona and J. C. Taylor. Some remarks on Widder’s theorem and uniqueness of isolated singularities for parabolic equations. In Partial differential equations with minimal smoothness and applications (Chicago, IL, 1990), volume 42 of IMA Vol. Math. Appl., pages 15–23. Springer, New York, 1992.

[2] D. G. Aronson. Non-negative solutions of linear parabolic equations. Ann. Scuola Norm. Sup. Pisa (3), 22:607–694, 1968.

[3] D. G. Aronson. Addendum: “Non-negative solutions of linear parabolic equations” (Ann. Scuola Norm. Sup. Pisa (3) 22 (1968), 607–694). Ann. Scuola Norm. Sup. Pisa (3), 25:221–228, 1971.

[4] Robert Azencott. Behavior of diffusion semi-groups at infinity. Bull. Soc. Math. France, 102: 193–240, 1974. ISSN 0037-9484.

[5] Martin T. Barlow and Edwin A. Perkins. Brownian motion on the Sierpiński gasket. Probab. Theory Related Fields, 79(4):543–623, 1988. ISSN 0178-8051. doi: 10.1007/BF00318785. URL http://dx.doi.org/10.1007/BF00318785
[6] Martin T. Barlow, Richard F. Bass, and Takashi Kumagai. Stability of parabolic Harnack inequalities on metric measure spaces. *J. Math. Soc. Japan*, 58(2):485–519, 2006. ISSN 0025-5645. URL http://projecteuclid.org/getRecord?id=euclid.jmsj/1149166785.

[7] A. Bendikov and L. Saloff-Coste. On- and off-diagonal heat kernel behaviors on certain infinite dimensional local Dirichlet spaces. *Amer. J. Math.*, 122(6):1205–1263, 2000. ISSN 0002-9327. URL http://muse.jhu.edu/journals/american_journal_of_mathematics/v122/122.6bendikov.pdf.

[8] Alexander Bendikov and Laurent Saloff-Coste. Invariant local Dirichlet forms on locally compact groups. *Ann. Fac. Sci. Toulouse Math. (6)*, 11(3):303–349, 2002. ISSN 0240-2963. URL http://www.numdam.org/item?id=AFST_2002_6_11_3_303_0.

[9] Alexander Bendikov, Laurent Saloff-Coste, Maura Salvatori, and Wolfgang Woess. The heat semigroup and Brownian motion on strip complexes. *Adv. Math.*, 226(1):992–1055, 2011. ISSN 0001-8708. doi: 10.1016/j.aim.2010.07.014. URL http://dx.doi.org/10.1016/j.aim.2010.07.014.

[10] M. Biroli and U. Mosco. A Saint-Venant type principle for Dirichlet forms on discontinuous media. *Ann. Mat. Pura Appl. (4)*, 169:125–181, 1995. ISSN 0003-4622. doi: 10.1007/BF01759352. URL http://dx.doi.org/10.1007/BF01759352.

[11] Zhen-Qing Chen and Masatoshi Fukushima. *Symmetric Markov processes, time change, and boundary theory*, volume 35 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2012. ISBN 978-0-691-13605-9.

[12] N. Dinculeanu. *Vector measures*. International Series of Monographs in Pure and Applied Mathematics, Vol. 95. Pergamon Press, Oxford, 1967.

[13] J. Eells and B. Fuglede. *Harmonic maps between Riemannian polyhedra*, volume 142 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2001. ISBN 0-521-77311-3. With a preface by M. Gromov.

[14] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998. ISBN 0-8218-0772-2.

[15] Masatoshi Fukushima. *Dirichlet forms and Markov processes*, volume 23 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1980. ISBN 0-444-85421-5.

[16] Masatoshi Fukushima, Yōichi Ōshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1994. ISBN 3-11-011626-X.

[17] A. A. Grigor’yan. The heat equation on noncompact Riemannian manifolds. *Mat. Sb.*, 182(1):55–87, 1991. ISSN 0368-8666.

[18] Alexander Grigor’yan and Jiaxin Hu. Off-diagonal upper estimates for the heat kernel of the Dirichlet forms on metric spaces. *Invent. Math.*, 174(1):81–126, 2008. ISSN 0020-9910. doi: 10.1007/s00222-008-0135-9. URL http://dx.doi.org/10.1007/s00222-008-0135-9.
[19] Pavel Gyrya and Laurent Saloff-Coste. Neumann and Dirichlet heat kernels in inner uniform domains. *Astérisque*, 336:vii+144, 2011. ISSN 0303-1179.

[20] Kazuhiro Ishige and Minoru Murata. Uniqueness of nonnegative solutions of the Cauchy problem for parabolic equations on manifolds or domains. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 30(1):171–223, 2001. ISSN 0391-173X. URL http://www.numdam.org/item?id=ASNSP_2001_4_30_1_171_0

[21] Jun Kigami. Measurable Riemannian geometry on the Sierpinski gasket: the Kusuoka measure and the Gaussian heat kernel estimate. *Math. Ann.*, 340(4):781–804, 2008. ISSN 0025-5831. doi: 10.1007/s00208-007-0169-0. URL http://dx.doi.org/10.1007/s00208-007-0169-0

[22] A. Korányi and J. C. Taylor. Minimal solutions of the heat equation and uniqueness of the positive Cauchy problem on homogeneous spaces. *Proc. Amer. Math. Soc.*, 94(2):273–278, 1985. ISSN 0002-9939. doi: 10.2307/2045390. URL http://dx.doi.org/10.2307/2045390

[23] M. Krzyżański. Sur les solutions non négatives de l’équation linéaire normale parabolique. *Rev. Roumaine Math. Pures Appl.*, 9:393–408, 1964. ISSN 0035-3965.

[24] Kazuhiro Kuwae, Yoshihiro Machigashira, and Takashi Shioya. Sobolev spaces, Laplacian, and heat kernel on Alexandrov spaces. *Math. Z.*, 238(2):269–316, 2001. ISSN 0025-5874. doi: 10.1007/s002090100252. URL http://dx.doi.org/10.1007/s002090100252

[25] Patrick Maheux. Estimations du noyau de la chaleur sur les espaces homogènes. *J. Geom. Anal.*, 8(1):65–96, 1998. ISSN 1050-6926. doi: 10.1007/BF02922109. URL http://dx.doi.org/10.1007/BF02922109

[26] Bernard Mair and J. C. Taylor. Integral representation of positive solutions of the heat equation. In *Théorie du potentiel (Orsay, 1983)*, volume 1096 of *Lecture Notes in Math.*, pages 419–433. Springer, Berlin, 1984. doi: 10.1007/BFb0100123. URL http://dx.doi.org/10.1007/BFb0100123

[27] Pedro J. Mendez-Hernandez and Minoru Murata. Semismall perturbations, semi-intrinsic ultracontractivity, and integral representations of nonnegative solutions for parabolic equations. *J. Funct. Anal.*, 257(6):1799–1827, 2009. ISSN 0022-1236. doi: 10.1016/j.jfa.2009.05.028. URL http://dx.doi.org/10.1016/j.jfa.2009.05.028

[28] Jürgen Moser. A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.*, 17:101–134, 1964. ISSN 0010-3640.

[29] Jürgen Moser. Correction to: “A Harnack inequality for parabolic differential equations”. *Comm. Pure Appl. Math.*, 20:231–236, 1967. ISSN 0010-3640.

[30] Minoru Murata. Heat escape. *Math. Ann.*, 327(2):203–226, 2003. ISSN 0025-5831. doi: 10.1007/s00208-002-0381-x. URL http://dx.doi.org/10.1007/s00208-002-0381-x

[31] Minoru Murata. Uniqueness theorems for parabolic equations and Martin boundaries for elliptic equations in skew product form. *J. Math. Soc. Japan*, 57(2):387–413, 2005. ISSN 0025-5645. URL http://projecteuclid.org/getRecord?id=euclid.jmsj/1158242064

32
[32] Minoru Murata. Integral representations of nonnegative solutions for parabolic equations and elliptic Martin boundaries. *J. Funct. Anal.*, 245(1):177–212, 2007. ISSN 0022-1236. doi: 10.1016/j.jfa.2006.12.012. URL http://dx.doi.org/10.1016/j.jfa.2006.12.012

[33] Melanie Pivarski and Laurent Saloff-Coste. Small time heat kernel behavior on Riemannian complexes. *New York J. Math.*, 14:459–494, 2008. ISSN 1076-9803. URL http://nyjm.albany.edu:8000/j/2008/14_459.html

[34] P. C. Rosenbloom and D. V. Widder. A temperature function which vanishes initially. *Amer. Math. Monthly*, 65:607–609, 1958. ISSN 0002-9890.

[35] L. Saloff-Coste. A note on Poincaré, Sobolev, and Harnack inequalities. *Internat. Math. Res. Notices*, (2):27–38, 1992. ISSN 1073-7928. doi: 10.1155/S1073792892000047. URL http://dx.doi.org/10.1155/S1073792892000047

[36] Laurent Saloff-Coste. *Aspects of Sobolev-type inequalities*, volume 289 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2002. ISBN 0-521-00607-4.

[37] Laurent Saloff-Coste. Uniformly elliptic operators on Riemannian manifolds. *J. Differential Geom.*, 36(2):417–450, 1992. ISSN 0022-040X. URL http://projecteuclid.org/getRecord?id=euclid.jdg/1214448748

[38] Benjamin Steinhurst. Uniqueness of locally symmetric Brownian motion on Laakso spaces. *Potential Analysis*, in press. ISSN 0926-2601. URL http://dx.doi.org/10.1007/s11118-012-9273-1 10.1007/s11118-012-9273-1.

[39] Karl-Theodor Sturm. The geometric aspect of Dirichlet forms. In *New directions in Dirichlet forms*, volume 8 of *AMS/IP Stud. Adv. Math.*, pages 233–277. Amer. Math. Soc., Providence, RI, 1998.

[40] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and $L^p$-Liouville properties. *J. Reine Angew. Math.*, 456:173–196, 1994. ISSN 0075-4102. doi: 10.1515/crll.1994.456.173. URL http://dx.doi.org/10.1515/crll.1994.456.173

[41] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.*, 32(2):275–312, 1995. ISSN 0030-6126. URL http://projecteuclid.org/getRecord?id=euclid.ojm/1200786053

[42] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. *J. Math. Pures Appl. (9)*, 75(3):273–297, 1996. ISSN 0021-7824.

[43] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and geometry on groups*, volume 100 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1992. ISBN 0-521-35382-3.

[44] D. V. Widder. Positive temperatures on an infinite rod. *Trans. Amer. Math. Soc.*, 55:85–95, 1944. ISSN 0002-9947.

[45] J. Wloka. *Partial differential equations*. Cambridge University Press, Cambridge, 1987. ISBN 0-521-25914-2; 0-521-27759-0. Translated from the German by C. B. Thomas and M. J. Thomas.