Analysis on the Sierpinski Carpet

M. T. Barlow

ABSTRACT. The ‘analysis on fractals’ and ‘analysis on metric spaces’ communities have tended to work independently. Metric spaces such as the Sierpinski carpet fail to satisfy some of the properties which are generally assumed for metric spaces. This survey discusses analysis on the Sierpinski carpet, with particular emphasis on the properties of the heat kernel.

1. Background and history

Percolation was introduced by Broadbent and Hammersley [BH] in 1957. The easiest version to describe is bond percolation. Take any graph $G = (V, E)$ and a probability $p \in (0, 1)$. For each edge $e$ keep it with probability $p$ and delete it with probability $1 - p$, independently of all other edges. This is one of the core models of statistical physics, and has many applications in other areas – one is contact networks for infectious diseases. For books on percolation see [BR, Gri].

Write $G_p = (V, E_p)$ for the (random) graph obtained by the percolation process. The connected component of $x$ is called the cluster containing $x$ and is denoted $C(x)$. Percolation on $\mathbb{Z}^d$ with $d \geq 2$ has a phase transition. Set

$$\theta(p) = \mathbb{P}_p(\{|C(0)| = \infty\}), \quad p_c = p_c(d) = \inf\{p : \theta(p) > 0\}.$$ 

Theorem 1.1. [BH] For the lattice $\mathbb{Z}^d$, $p_c \in (0, 1)$.

When $p$ is small $G_p$ consists of a number of (mainly small) finite clusters. For large $p$ the cluster $C(0)$ may or may not be infinite, but a zero-one law combined with the stationary ergodic nature of the percolation process gives that infinite clusters exist with probability 1. A less elementary argument gives that if $p > p_c$ then there is exactly one infinite cluster. (As always in such contexts, this statement has to be qualified by ‘with probability 1’.)

In spite of its 50 year history, many open problems remain. The most fundamental of these is what happens at $p_c$ in low dimensions:

Open Problem 1. Is $\theta(p_c) = 0$ if $3 \leq d \leq 18$?

It is known that $\theta(p_c) = 0$ if $d = 2$ or $d \geq 19$ – see [Ke1, HS2].

Physicists are interested in ‘transport’ problems of percolation clusters, that is, in the behaviour of solutions to the wave or heat equation. A percolation cluster is
a graph, so one can define the graph (discrete) Laplacian
\[ \Delta G_p f(x) = \frac{1}{N_p(x)} \sum_{y \sim_p x} (f(y) - f(x)). \]
Here \( y \sim_p x \) means that \( \{x, y\} \) is an edge in \( G_p \), and \( N_p(x) \) is the number of neighbours of \( x \). One can then look at the heat equation on \( G_p \) (discrete space, continuous time):
\[ \frac{\partial u}{\partial t} = \Delta G_p u. \]
(The wave equation would also be of great interest, but I do not know of any work on this for percolation clusters.) For the heat equation the broad situation for the three phases of percolation is as follows:

- \( p < p_c \) (subcritical). No large scale structure.
- \( p > p_c \) (supercritical). For HE get Gaussian limits, homogenization.
- \( p = p_c \) (critical). Hard and interesting, only a few results.

The supercritical case is now well understood – see [SS, Ba2, MP, BeB]. In the critical case (or as \( p \to p_c \)) the main open problem is the existence and values of critical exponents. For example it is conjectured that there exist exponents \( \gamma, \beta \) such that
\[
E_p |C(0)| \approx (p_c - p)^{-\gamma} \quad \text{as } p \uparrow p_c, \\
\theta(p) \approx (p - p_c)^{\beta} \quad \text{as } p \downarrow p_c.
\]
The existence of these exponents has not been established, with two exceptions. The first is in high dimensions, where they take the ‘mean field’ values – see [HS1]. The other is for one particular percolation model in \( d = 2 \) (site percolation on the triangular lattice) where connections with SLE give some of these exponents – see [SW].

Physicists believe that these exponents are ‘universal’. This means that it conjectured that they should depend on the dimension \( d \), but not on the particular lattice or specific local features of the model. At first sight this is a surprise to a mathematician, since coarser features of the model, such as the value of \( p_c \), do depend on the lattice. ‘Universality’ is crucial for the physical interest of percolation and other models in statistical physics, since the main point of these models (for a physicist) is their capacity to account for the behaviour of real systems.

In 1976 P. De Gennes in a survey [Ge] suggested the study of random walks on critical percolation clusters as a tool to study their electrical resistance properties. It was believed then (and has now been proved in some cases – see [HJ]) that critical percolation clusters are in some senses ‘fractal’. To help understand how random walks would behave on such ‘fractal’ graphs, in the early 1980s mathematical physicists began to study random walks on regular exact fractals, such as the Sierpinski gasket.

While this subject was called by mathematical physicists ‘diffusion on fractals’, really they looked at random walks on various fractal graphs, mainly those with considerable regularity. In the late 1980s mathematicians began to study analysis on true fractals, starting with the easiest case, the Sierpinski gasket – see [Ku1, Go, BP, Kig1].
In the period 1990–2005 mathematical work diverged from physics. Mathematicians obtained quite detailed information about solutions of the heat equation on regular exact fractals such as the Sierpinski gasket and Sierpinski carpet. This work gave quite accurate results, but needed very strong regularity for the space. Of more interest for physics applications such as percolation would have been cruder results obtained under weaker hypotheses.

From about 2005 the mathematical theory developed tools which can handle some problems for random walks on critical percolation clusters – see for example [BJKS, KN]. (Note also the early paper [Ke2], which rather remarkably predates work on the ultimately much easier case of supercritical percolation.)

In these notes, I will discuss analysis on true (regular) fractals, which of course are also metric spaces. For other surveys or books on this topic see [Ba1, Ki2, Str].

The Sierpinski gasket (SG) shown above is rather special because of existence of ‘cut points’. This feature of the SG enables one to separate exactly the different levels of the set, and so allows various exact calculations to be performed. For example it is not hard to show that the mean time for a random walk on the Sierpinski gasket graph to cross a triangle of side $2^n$ is exactly $5^n$. The resulting space-time scaling index, denoted by physicists $d_w = d_w(SG)$ is then $\log 5/\log 2$, and this number arises in the subsequent analysis of the heat equation on the fractal SG.

While useful, the existence of these cuts points may in some sense have proved to be a snare, since it allowed questions to be solved by complicated calculations, rather than by seeking a deeper insight into the underlying issues. The Sierpinski
carpet does not have such cut points, and so is rather harder to study, but has proved its value by forcing the development of more robust tools.

2. Sierpinski carpet

The basic Sierpinski carpet (SC) is a fractal subset of \([0,1]^2\) defined in a similar way to the classical Cantor set, except that one removes the middle square out of a \(3 \times 3\) block. It can be regarded as a model space for studying diffusion in irregular media, with irregularities at many different length scales.

One can define the basic SC in \(d\) dimensions, by starting with \([0,1]^d\), dividing each cube into \(3^d\) subcubes and removing the middle cube. One can also define Generalized Sierpinski carpets in \(d \geq 2\) by removing other (symmetric) patterns of cubes. One example is the Menger Sponge – good pictures of this are easily found on the internet. For simplicity, in these notes I will mainly focus on the basic SC in \(d\) dimensions.

It is useful to have a more systematic description of these sets. A map \(\psi : \mathbb{R}^d \to \mathbb{R}^d\) is a similitude if there exists \(L > 1\) such that \(|\psi(x) - \psi(y)| = L^{-1}|x - y|\) for all \(x, y \in \mathbb{R}^d\). We call \(L\) the contraction factor of \(\psi\). Let \(M \geq 1\), and let \(\psi_1, \ldots, \psi_M\) be similitudes with contraction factors \(L_i\). For \(A \subset \mathbb{R}^d\) set

\[
\Psi(A) = \bigcup_{i=1}^M \psi_i(A),
\]

and let \(\Psi^{(n)}\) denote the \(n\)-fold composition of \(\Psi\).

Let \(\mathcal{K}\) be the set of non-empty compact subsets of \(\mathbb{R}^d\). For \(A \subset \mathbb{R}^d\) set \(A^{(\varepsilon)} = \{x : |x - a| \leq \varepsilon\ \text{for some } a \in A\}\). The Hausdorff metric \(d\) on \(\mathcal{K}\) is defined by

\[
d_{\mathcal{K}}(A, B) = \inf \{\varepsilon > 0 : A \subset B^{(\varepsilon)} \ and \ B \subset A^{(\varepsilon)}\};
\]

\(d_{\mathcal{K}}\) is a metric on \(\mathcal{K}\), and \((\mathcal{K}, d_{\mathcal{K}})\) is complete.

**Theorem 2.1.** (See [Hu, Fa]) Let \((\psi_1, \ldots, \psi_M)\) be as above, with \(L_i > 1\) for each \(1 \leq i \leq M\). Then there exists a unique \(F \in \mathcal{K}\) such that \(F = \Psi(F)\). Further, if \(G \in \mathcal{K}\) then \(\Psi^n(G) \to F\) in \(d_{\mathcal{K}}\). If \(G \in \mathcal{K}\) satisfies \(\Psi(G) \subset G\) then 

\[
F = \cap_{n=0}^{\infty} \Psi^{(n)}(G).
\]
To obtain the Hausdorff dimension $d_f(F)$ of the limiting set $F$ from the contraction factors $L_i$ a “non-overlap” condition is necessary. We say that $(\psi_1, \ldots, \psi_M)$ satisfies the open set condition if there exists an open set $U$ such that $\psi_i(U)$, $1 \leq i \leq M$, are disjoint, and $\Psi(U) \subset U$. Note that, since $\Psi(U) \subset U$, then the fixed point $F$ of $\Psi$ is given by $F = \cap \Psi^n(U)$.

**Theorem 2.2.** (See [Fa], p. 119.) Let $(\psi_1, \ldots, \psi_M)$ satisfy the open set condition, and let $F$ be the fixed point of $\Psi$. Let $\alpha$ be the unique real such that

$$\sum_{i=1}^{M} L_i^{-\alpha} = 1.$$  

Then the Hausdorff dimension of $F$ is given by $d_f(F) = \alpha$. Further the $x^\alpha$-Hausdorff measure of $F$ is strictly positive and finite.

In these notes I will only be interested in the case when $L_i$ all take the same value $L$. Then it is easy to see that

$$\alpha = d_f(F) = \frac{\log M}{\log L}.$$  

We call $L$ the length scale factor, and $M$ the mass scale factor of $F$.

Let $L = 3$, $M = 3^d - 1$, and let $\psi_i$, $i = 1, \ldots, M$ be the (rotation free) similitudes which map $F_0 = [0, 1]^d$ onto each of the $M$ sub-cubes of side $L^{-1}$. We assume that $\psi_1(x) = L^{-1}x$ is the similitude which maps $[0, 1]^d$ onto $[0, L^{-1}]^d$. The open set condition clearly holds – one can take $U = (0, 1)^d$. If $F_n = \Psi^n(F_0)$ then $F_n$ is the $n$th approximation to the SC, which is given by

$$F = \cap_{n=0}^\infty F_n.$$  

Note that $F_n$ is the union of $M^n$ cubes each of side $L^{-n}$. Let $\mu_n$ be Lebesgue measure on $F_n$ renormalized to have mass 1 and $\mu$ be Hausdorff $x^\alpha$-measure on $F$, normalised so that $\mu(F) = 1$. We have that $\mu = \lim_n \mu_n$.

We will wish to consider two other sets, based on the same construction, but which give unbounded subsets of $\mathbb{R}^d$. Let

$$\tilde{F}_n = L^n F_n = \{L^n x, x \in F_n\};$$

thus $\tilde{F}_n$ is a subset of $[0, L^n]^d \subset \mathbb{R}_+^d$ and is the union of $M^n$ unit cubes. We have

$$\tilde{F}_n \cap [0, L^{n-1}]^d = \tilde{F}_{n-1}.$$  

So set

$$\tilde{F} = \cup_{n=0}^\infty \tilde{F}_n.$$  

This is called the pre-carpet – see [O1]. It is the closure of an open domain in $\mathbb{R}^d$ with piecewise smooth boundary, so is locally regular. It has fractal structure ‘at infinity’, which mimics the local structure of the compact Sierpinski carpet $F$.

Define also:

$$\hat{F} = \cap_{n=0}^\infty L^{-n} \tilde{F}.$$  

This is the unbounded fractal Sierpinski carpet; it is the union of a countable number of copies of $F$. Note that $\psi_1(\hat{F}) = \hat{F}$.

We next look at the metric space structure of $F$. Using the fact that the boundary of a cube is never removed, it is not hard to prove the following
Lemma 2.3. Let \( x, y \in F \). Then there exists a path \( \gamma \) connecting \( x, y \) consisting of countably many line segments, with length 
\[
L(\gamma) \leq C|x - y|.
\]
Thus if \( d(x, y) \) is the shortest path (geodesic) distance on \( F \) then 
\[
|x - y| \leq d(x, y) \leq C|x - y|.
\]

In these notes I will always use the geodesic or shortest path metric \( d \) when working in subsets of \( \mathbb{R}^d \), and will write \( B(x, r) \) for balls in this metric.

Lemma 2.4. \((F, d, \mu)\) is a metric measure space, and if \( \alpha = df(F) \) one has
\[
(\text{2.2}) \quad \mu(B_d(x, r)) \asymp r^\alpha, \quad x \in F, \quad 0 < r < 1.
\]

Thus the Sierpinski carpet is Ahlfors regular – see [HK].

For analysis on the Sierpinski carpet \( F \), a natural starting point is to consider upper gradients.

Definition 2.5. Let \( f \in C(F) \). A function \( g : F \to \mathbb{R} \) is an upper gradient for \( f \) if for any \( x, y \in F \) and (rectifiable) path \( (\gamma(t), t \in [0, 1]) \) connecting \( x \) and \( y \) then 
\[
|f(y) - f(x)| \leq \int_\gamma g = \int_0^1 g(\gamma(t))|\gamma'(t)|dt.
\]

We write \( |\nabla f| \) for an upper gradient of \( f \).

In many situations a (weak) \((q,p)\)-Poincaré inequality of the type
\[
(\text{2.3}) \quad \inf_a \int_{B(x, r)} |f(x) - a|^q \mu(dx) \leq C(r) \int_{B(x, \lambda r)} |\nabla f|^p \mu(dx),
\]
where \( |\nabla f| \) is any upper gradient of \( f \), gives useful information about the space – see for example [HK, He].

Theorem 2.6. Let \( H_i, i = 0, 1 \) be the left and right edges of the SC in \( d = 2 \): \( H_i = \{x = (x_1, x_2) \in F : x_1 = i\} \). Let 
\[
\mathcal{A} = \{f \in C(F) : f|_{H_i} = i, i = 0, 1\}.
\]
Let \( p > 0 \). Then there exists a sequence of functions \( f_n \in \mathcal{A} \) and upper gradients \( |\nabla f_n| \) such that 
\[
\int_F |\nabla f_n|^p \mu(dx) \to 0.
\]
Consequently no weak Poincaré inequality holds on the Sierpinski carpet.

Proof. Recall that \( F_0 = [0, 1]^2 \) and that \( F_n \) is the \( n \)th step in the approximation of the SC. Let 
\[
f_0(x_1, x_2) = x_1; \quad \text{so} \quad \int_F |\nabla f_0|^p d\mu = 1.
\]

Choose \( h_1 : [0, 1] \to [0, 1] \) to be continuous and linear on each interval \( [0, \frac{1}{3}] \), \( [\frac{1}{3}, \frac{2}{3}] \), \( [\frac{2}{3}, 1] \), with \( h(0) = 0, h(1) = 1, \) and \( h' = b \) on \( (0, \frac{1}{3}) \cup (\frac{2}{3}, 1) \), and \( h' = 3 - 2b \) on \( (\frac{1}{3}, \frac{2}{3}) \). Set 
\[
f_1(x_1, x_2) = h_1(x_1); \quad \text{note} \quad f \in \mathcal{A}.
\]
Then $|\nabla f_1|$ is equal to $b$ on 6 out of the 8 sub-squares, and $(3 - 2b)$ on 2 sub-squares. Hence
\[
\int_F |\nabla f_1|^p \mu(dx) = \frac{6}{8} b^p + \frac{2}{8} (3 - 2b)^p = \frac{1}{4} (3b^p + (3 - 2b)^p) = \varphi(b).
\]
Then $\varphi(1) = 1$ and it is easy to check that $\varphi'(1) = p/4 > 0$. We could minimise over $b$, but it is enough to note that there exists $b \in (0, 1)$ with $\varphi(b) < 1$.

With this choice of $b$ we have
\[
\int_F |\nabla f_1|^p d\mu = \varphi(b) < 1.
\]
We now iterate this construction, to obtain a sequence $f_k$ such that on each square $S$ of side $3^{-k}$ we then have
\[
\int_S |\nabla f_k|^p d\mu = \varphi(b)^k \to 0 \text{ as } k \to \infty.  \tag{1}
\]
We define $f_k(x_1, x_2) = h_k(x_1)$ where $h_k$ are defined by
\[
h_{k+1}(t) = \begin{cases} (b/3)h_k(3t), & 0 \leq t \leq \frac{1}{3} \\ (b/3) + (1 - 2b/3)h_k(3t - 1), & \frac{1}{3} \leq t \leq \frac{2}{3} \\ (1 - b/3) + (b/3)h_k(3t - 2), & \frac{2}{3} \leq t \leq 1. \end{cases}
\]
On each square $S$ of side $3^{-k}$ we then have
\[
\int_S |\nabla f_{k+1}|^p d\mu = \varphi(b) \int_S |\nabla f_k|^p d\mu.
\]
Hence
\[
\int_F |\nabla f_k|^p d\mu = \varphi(b)^k \to 0 \text{ as } k \to \infty.
\]
The functions $f_k$ constructed above are piecewise linear; if we wished we could then smooth them to obtain a sequence $\tilde{f}_k \in C^1(F) \cap A$ with the same properties.

To see that no weak Poincaré inequality can hold, we can use the functions $f_k$ to construct functions $g_k$ such that $g_k(x, y) = 1$ if $x \in [0, \frac{1}{3}], g_k(x, y) = -1$ if $x \in [\frac{2}{3}, 1], |g_k| \leq 1$ on $F$, and
\[
\int_F |\nabla g_k|^p d\mu \to 0.
\]
Since for each $k, a \in \mathbb{R}$,
\[
\int_F |g_k(x) - a|^q \mu(dx) \geq \frac{3}{8}(|1 - a|^q + |1 + a|^q) \geq \frac{3}{8},
\]
the weak Poincaré inequality fails.

We can define a quadratic form $Q$ by
\[
Q(f, f) = \int_F |\nabla f|^2 d\mu,
\]
and (if one had not seen the result above) one might hope that this would yield a suitable Dirichlet form on some subspace of $C(F)$.

**Corollary 2.7.** $Q$ is not closeable; there exists $f_n \to 0$ uniformly such that $Q(f_n - f_m, f_n - f_m) \to 0$ but $Q(f_n, f_n) \geq 1$ for all $n$. 


Choosing Proof. Let \( f \) be 0. However, \( f_n \) do not converge to 0 in \( H \).

\[ Q(f, f) + |f|^2 \]

The functions \((f_n)\) given above are Cauchy in \( H \), and since they converge to 0 in \( L^2 \), if they had a limit in \( H \) it would have to be 0. However, \( f_n \) do not converge to 0 in \( H \).

**Proof.** Let \( f \) be continuous on \( F \). Using Theorem 2.6 we can find for each \( n \) a function \( h_n \) such that

\[ ||f - h_n||_\infty \leq \frac{1}{2} ||f||_\infty, \quad Q(h_n, h_n) \leq 2^{-n}. \]

(So continuous functions can be closely approximated by functions of low energy.)

Let \( f_0 \) satisfy \( Q(f_0, f_0) = 2, \quad ||f_0||_\infty = C < \infty \). Choose \((g_n)\) such that \( Q(g_n, g_n) \leq \delta^2 4^{-n} \) and if \( f_{n+1} = f_n - g_n \) then \( ||f_{n+1}||_\infty \leq \frac{1}{2} ||f_n||_\infty \). Then \( f_n \to 0 \) and

\[ Q(f_{n+1}, f_{n+1}) = Q(f_n, g_n) - 2Q(f_n, g_n) + Q(g_n, g_n) \]

\[ \geq Q(f_n, f_n) - 2\delta^2 Q(f_n, f_n)^{1/2}. \]

Choosing \( \delta \) small enough we have \( 1 \leq Q(f_n, f_n) \leq 3 \) for all \( n \).

This calculation suggests that the Dirichlet forms

\[ Q_n(f, f) = \int_{F_n} |\nabla f|^2 d\mu_n \]

are ‘too small’. If we are to obtain a useful limit some renormalisation is necessary; so we seek constants \( a_n \uparrow \infty \) such that if

\[ \mathcal{E}_n(f, f) = a_n Q_n(f, f), \]

then \( \mathcal{E}_n \) has (at least) subsequential limits. The obvious choice (which works) is to choose \( a_n \) so that

\[ \inf\{\mathcal{E}_n(f, f) : f \in A\} = O(1). \]

The renormalization argument involves various constants, and these turn out to have a more intuitive form if we work with the big sets \( \tilde{F}_n = L^n F_n \) rather than the small sets \( F_n \). Let \( H^n_0, H^n_1 \) be the left and right sides of \( \tilde{F}_n \):

\[ H^n_i = \{x = (x_1, \ldots, x_d) \in \tilde{F}_n : x_1 = L^n i\}, \quad i = 0, 1, \]

and

\[ \tilde{A}_n = \{f \in C(\tilde{F}_n) : f|_{H^n_i} = i, \quad i = 0, 1\} \]

be the set of functions which are zero on the LHS and 1 on the RHS. Let

\[ R_n = \inf\{\int_{\tilde{F}_n} |\nabla f|^2 dx : f \in \tilde{A}_n\}. \]

This is the minimal energy of functions in \( \tilde{A}_n \); and so \( R_n \) can be interpreted as the ‘effective resistance’ in \( F_n \) between \( H^n_0 \) and \( H^n_1 \). So it is natural to try

\[ \mathcal{E}_n(f, f) = L^{(d-2)} R_n \int_{F_n} |\nabla f|^2 dx. \]

(The term \( L^{n(d-2)} \) arises from rescaling the integral of the gradient.) We then have that \( \mathcal{E}_n(f, f) \geq 1 \) for all \( f \in C(F_n) \) with \( f = 0 \) on the left side of \( F_n \) and \( f = 1 \) on the right side.
We now wish to use the \((E_n)\) to construct a limiting Dirichlet form \(E\) on \(F\) or \(\hat{F}\). Unfortunately, the argument does not take what I would now regard as conceptually the clearest course. It would be nice to be able to proceed as follows:

1. Using some (as minimal as possible) properties of \(F\), prove that the sequence \(E_n\) has subsequential limits, in the sense of Mosco convergence.
2. Using some stronger regularity results on \(F\), obtain 'good properties' of any subsequential limit \(E\).
3. Prove the uniqueness of the limit, and that \(E_n \to E\).

It is quite reasonable to expect that, at least for (1) and (2), such a program would be possible. See Remark 2.13 below for comments on the situation as far as (3) is concerned. The first construction of limiting processes on Sierpinski carpets was done in [BB1], and used probabilistic methods (tightness and weak convergence) rather than Dirichlet form methods. An alternative approach using Dirichlet forms can be found in [KZ, HKKZ]. However, all these papers use quite strong regularity properties of the approximating forms \(E_n\) in order to show the existence of subsequential limits. Thus these arguments combine the 'existence' and 'regularity' parts of the proof in a fashion which makes it unclear exactly which properties of \(F\) are needed at each step of the argument.

Since (so far) the ideal argument suggested in (1) and (2) above does not exist, I will sketch what has been done. The proof has two main inputs:

1. Control of the constants \(R_n\) involved in the renormalization.
2. A Harnack inequality which gives enough regularity of \(E_n\) to give good properties of the limit.

**Theorem 2.8.** [BB3, McG]. There exists \(C\) such that

\[
C^{-1} R_n R_m \leq R_{n+m} \leq CR_n R_m, \quad n, m \geq 0.
\]

Hence there exists \(\rho = \lim_n R_n^{1/n}\) such that

\[
C^{-1} \rho^n \leq R_n \leq C \rho^n, \quad n \geq 0.
\]

**Proof:** \(R_n^{-1}\) is given as the energy of a function \(h_n\) in \(\tilde{A}_n\). Using copying and pasting of the functions \(h_n\) and \(h_m\) one can construct a function \(f^{n,m} \in \tilde{A}_{n+m}\) with energy less than \(cR_n^{-1} R_m^{-1}\). (See the papers cited above for the full argument, which requires a comparison between \(R_n\) and discrete network approximations.) Thus

\[
R_{n+m}^{-1} \leq E_{n+m}(f^{n,m}, f^{n,m}) \leq cR_n^{-1} R_m^{-1},
\]

which gives the first inequality. A dual characterization of resistance as the minimal energy of a unit flow (see [DS] for the classical case of electrical networks) gives a similar upper bound on \(R_{n+m}\).

A sequence \(x_n\) is subadditive if \(x_{n+m} \leq x_n + x_m\), and Fekete’s theorem states that for a subadditive sequence,

\[
\lim_{n} \frac{x_n}{n} = \inf_n \frac{x_n}{n} \in [-\infty, x_1].
\]

Since \(\log(C R_n)\) is subadditive, and \(\log(C^{-1} R_n)\) is superadditive, (2.5) follows. \(\Box\)

The subadditivity used to prove the existence of the limit \(\rho\) gives no information on its value. For the basic SC with \(d = 2\) shorting and cutting arguments give \(7/6 \leq \rho \leq 3/2\), and numerical calculations suggest that \(\rho \simeq 1.251\). For the shorting
argument, recall that ‘shorts decrease resistance’ (see [DS]), and starting with the set \( \tilde{F}_n \) impose shorts on each of the lines \( x_1 = \frac{1}{3}L^n, x_1 = \frac{2}{3}L^n \). (In terms of the definition (2.4) this corresponds to restricting the minimisation to functions which are constant on those lines.) This replaces the resistance problem for the set \( \tilde{F}_n \) with that for 8 copies of \( \tilde{F}_n-1 \), and one obtains

\[
R_n \geq \frac{1}{3}R_{n-1} + \frac{2}{3}R_{n-1} + \frac{1}{3}R_{n-1} = \frac{7}{6}R_{n-1},
\]

which implies that \( \rho \geq 7/6 \). A similar argument with cuts gives the upper bound.

In general for \( d \geq 2 \) one has

\[
\frac{2}{3^{d-1}} + \frac{1}{3^{d-1} - 1} \leq \rho \leq \frac{3}{3^{d-1} - 1},
\]

– see [BB5, Remark 5.4].

The second input is an elliptic Harnack inequality (EHI) for the pre-carpet \( \tilde{F} \). Note that as the pre-carpet is the closure of a domain in \( \mathbb{R}^d \), one can define the Laplacian and harmonic functions on \( \tilde{F} \) in the usual way. We can also define a reflecting Brownian motion \( W \) on \( \tilde{F} \). This is the continuous strong Markov process which behaves locally like a Brownian motion in \( F^n \), and has normal reflection on \( \partial \tilde{F} \) – see [BH, Fu].

Recall that we are using \( B(x, r) \) to denote balls in the pre-SC with respect to the geodesic metric – if we used the Euclidean metric we might have to be careful about connectivity.

**Theorem 2.9.** [BB4] Let \( x \in \tilde{F}, r > 0, B = B(x, r) \) and \( h \) be non-negative on \( \overline{B} \) and harmonic on \( B \). There exists a constant \( C_H \) (depending only on the Sierpinski carpet) such that if \( B' = B(x, r/2) \) then

\[
\sup_{B'} h \leq C_H \inf_{B'} h.
\]

The proof uses a probabilistic coupling argument which relies on the symmetry of \( F \) very heavily. I will need some notation for hitting and exit times for a stochastic process \( X \).

**Definition 2.10.** Let \( X = (X_t, t \in \mathbb{R}_+) \) be a stochastic process on a metric space \( \mathcal{X} \). For \( A \subset \mathcal{X} \) set

\[
T_A = T_A(X) = \inf\{t \geq 0 : X_t \in A\},
\]

\[
\tau_{x,r} = \tau_{x,r}(X) = T_{B(x, r)}(X).
\]

Given two processes \( X, Y \) define their collision time by

\[
T_C = T_C(X, Y) = \inf\{t > 0 : Y_t = X_t\},
\]

with the usual convention that \( \inf \emptyset = +\infty \).

The coupling gives that there exists a constant \( p_F > 0 \), depending on the carpet \( F \), such that the following holds. Given \( x_0 \in \tilde{F}, R > 0 \), and points \( x, y \in B(x_0, R/2dL) \), there exist reflecting Brownian motions \( W^x, W^y \) on \( \tilde{F} \) with \( W^x_{t_0} = x \), \( W^y_{t_0} = y \) such that \( W^x_t = W^y_t \) for \( t \geq T_C \) and

\[
\mathbb{P}(T_C(W^x, W^y) \leq \tau_{x_0, R}(W^x) \wedge \tau_{x_0, R}(W^y)) > p_F.
\]

The processes \( W^x \) and \( W^y \) are of course not independent – the coupling uses reflection methods as in [LR].
A simpler argument, also using the (local) reflection symmetry of $\tilde{F}$, gives a weak lower bound on the probability of hitting small balls: if $y \in B(x_0, R/2dL)$, then for the process $W = W^{x_0}$,

$$P^x(T_{B(y,\lambda R)} < \tau_{x_0, R}) > p_0 \lambda^\gamma.$$  

Here $p_0$ and $\gamma > 0$ depend only on $F$.

If $h$ is harmonic on $B = B(x_0, R)$ then $h(W^{x_0})$ is a martingale. Hence if $T = \tau_B(W^{x_0}) \wedge \tau_B(W^y)$,

$$h(x) - h(y) = E(h(W^{x_0}_T) - h(W^{y}_T))$$

$$= E(h(W^{x_0}_T) - h(W^{y}_T); T < T_C)$$

$$\leq P(T < T_C) \sup_{x,y \in B} (h(x) - h(y)) \leq (1 - p_F) \sup_{x,y \in B} (h(x) - h(y)).$$

Hence, writing

$$\text{Osc } f = \sup_A f - \inf_A f,$$

and $B' = B(x_0, R/2dL)$, we have

$$\text{Osc } f \leq (1 - p_F) \text{Osc } h.$$  

This oscillation inequality is not quite enough on its own to give the Harnack inequality - see the example in [Ba3]. However, combined with (2.8) a standard argument (see for example [FS]) gives the elliptic Harnack Inequality Theorem 2.9.

We have three ‘scale factors’ for the SC:

1. $L = L_F$, the length scaling factor.
2. $M = M_F$, the volume scaling factor.
3. $\rho$, the ‘resistance scaling factor’.

As stated above, for the basic SC in $d$ dimensions, $L = 3$, and $M = 3^d - 1$. For $[0,1]^d$ (which can be regarded as a trivial SC) if $L = 3$ then $M = 3^d$ and $\rho = 3^{2-d}$.

Recall the definition of $d_f$ from (2.1), and set

$$d_w(F) = d_w = \frac{\log M \rho}{\log L}. 

$$

$d_w$ was called by physicists the walk dimension and is related to space time scaling of the heat equation. It turns out that one always has $d_w \geq 2$: for the SC in $d$ dimensions this follows from the lower bound in (2.6).

Given a regular fractal $F$, since $L$ and $M$ are given by the construction, one can calculate $d_f$ easily. The constant $\rho$ which gives $d_w$ is somehow deeper, and seems to require some analysis on the set or its approximations. Loosely one can say that $d_f$ is a ‘geometric’ constant, while $d_w$ is an ‘analytic’ constant. One may guess that in some sense $\rho$ or $\beta$ are in general inaccessible by any purely geometric argument. (An exception is for trees, where one has $d_w = 1 + d_f$.)

The two inputs (Theorem 2.8 and Theorem 2.9) lead to good control of the heat equation in $\tilde{F}$.

**Theorem 2.11.** [BB5] Let $p_t(x,y)$ be the heat kernel on the pre-SC $\tilde{F}$. Then writing $\beta = d_w$

$$p_t(x,y) \overset{(c)}{=} c \mu(B(x, t^{1/\beta}))^{-1} \exp(-c(d(x,y)^\beta / t)^{1/(\beta-1)}),$$

for $(t,x,y)$ such that $t \geq 1 \lor d(x,y)$.
Remark 2.12. 1. Here \( \cong \) means that upper bound holds with constants \( c_1, c_2 \) and lower bounds hold with constants \( c_3, c_4 \).
2. One gets the usual Gaussian type bounds if \( t \leq 1 \vee d(x, y) \).
3. Since \( \tilde{F} \) looks locally like \( \mathbb{R}^d \) at length scales less than 1, one expects different bounds when \( t, d(x, y) \) are small.
4. Integrating these bounds, if \( W_t \) is the associated diffusion process one gets
\[
\mathbb{E}^x d(x, W_t)^2 \cong t^{2/\beta}, \quad t \geq 1.
\]
When \( \beta \neq 2 \) this is called ‘anomalous diffusion’.

I will defer discussing the proof of this theorem until the next section, where it will follow from Theorem 2.8, Theorem 2.9, and a general result on heat kernel bounds in metric measure spaces, Theorem 3.14.

The theorem above shows that the space-time scaling involves the new parameter \( \beta = d_w \). There are several different ways of seeing why this anomalous scaling arises.

One is to look at the Poincaré inequality in a cube in \( \tilde{F} \). Let \( S \) be a cube of side \( R = L^n \) in \( \tilde{F} \); by symmetry we can assume that \( S = [0, L^n] \cap \tilde{F} = \tilde{F}_n \). Let \( P_n \) be the best constant in the PI
\[
\int_S (f - \mathbb{T})^2 dx \leq P_n \int_S |\nabla f|^2 dx.
\]
Let \( g_n \) be the function in \( \tilde{A}_n \) which attains the minimum for the resistance across \( S \). By symmetry the average value of \( g_n \) on \( \tilde{F}_n \) is \( 1/2 \). Then we can use \( g_n - 1 \) to build a function \( f_n \) on \( \tilde{F}_n \) which satisfies
\[
f_n((x_1, x')) = -1, \quad 0 \leq x_1 \leq \frac{1}{3} L^n,
\]
\[
\in [-1, 1], \quad \frac{4}{7} L^n \leq x_1 \leq \frac{2}{7} L^n,
\]
\[
= 1, \quad \frac{2}{7} L^n \leq x_1 \leq 1.
\]
Here I have used the notation \( x = (x_1, x') \) for points in \( \mathbb{R}^d = \mathbb{R} \times \mathbb{R}^{d-1} \). Then
\[
\int_S (f_n - \mathbb{T}_n)^2 \cong |S| = M^n, \quad \int_S |\nabla f_n|^2 \leq cR_{n-1}^{-1},
\]
and so
\[
P_n \geq cM^n R_{n-1} \geq (M \rho)^n = cR^{\log(M \rho)/\log L} = c R^\beta.
\]
The regularity given by EHI gives that in fact \( P_n \asymp (M \rho)^n \) – see [KZ]. Since \( P_n \) gives the time scale over which heat converges to equilibrium, this suggests that heat (or the diffusion process \( X \)) takes time order \( R^3 \) to move a distance \( R \), rather than the classical time \( O(R^2) \).

The heat kernel bounds on the pre-SC then enables one to prove that the rescaled Dirichlet forms \( \mathcal{E}_n \) on \( F_n \) have subsequential limits. Taking any one of these limits, one obtains a Dirichlet form \( (\mathcal{E}, \mathcal{F}) \) on \( L^2(F, \mu) \). Different subsequences might give different limits, but any limit satisfies symmetry conditions and has a heat kernel which satisfies (2.10).

Remark 2.13. A recent result [BBKT1] proves that any Dirichlet form which is symmetric with respect to suitable local symmetries of the Sierpinski carpet is unique, up to constants. This shows that if \( \mathcal{E}' \) and \( \mathcal{E}'' \) are both subsequential limits
of the \((E_n)\) then \(E'' = \lambda E'\) for some \(\lambda \in (C^{-1}, C)\), where \(C\) is the constant in
Theorem 2.8.

This does not seem enough to prove that \(\lim_n E_n\) exists. One difficulty is that we do not know if \(\lim_n \rho^{-n} R_n\) exists.

### 3. Heat kernels on metric measure spaces

Examples of these include manifolds, \(\mathbb{R}^d\), domains in \(\mathbb{R}^d\), ‘cable systems’ for graphs, and true fractals such as the Sierpinski carpet described above.

**Metric measure spaces.** Let \((X, d, \mu)\) be a metric measure space. We assume that \((X, d)\) is complete and connected, that \(d\) is a length metric, \(\mu\) is Radon, \(X\) has infinite radius, and the balls

\[ B(x, r) = \{ y : d(x, y) < r \}, \quad x \in X, \ r > 0 \]

are precompact.

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

**Definition 3.1.** \(X\) satisfies \(V_\alpha\) if

\[ c_1 r^\alpha \leq V(x, r) \leq c_2 r^\alpha, \quad x \in X, \ r > 0. \]

\(X\) satisfies volume doubling \(VD\) if

\[ V(x, 2r) \leq c_D V(x, r), \quad x \in X, \ r > 0. \]

It is easy to show that \(V_\alpha\) implies \(VD\), and that \(VD\) implies polynomial volume growth: there exists \(\alpha_1 < \infty\) such that for \(x \in X, 0 < r < R\) one has

\[ \frac{V(x, R)}{V(x, r)} \leq C(R/r)^{\alpha_1}. \]

An example of a set which satisfies \(VD\) but not \(V_\alpha\) for any \(\alpha\) is the pre-carpet \(\tilde{F}\), where one has

\[
\begin{align*}
(3.1) & \quad V(x, r) \asymp r^d, \quad r \in (0, 1), \\
(3.2) & \quad V(x, r) \asymp r^{d_{\mathcal{F}}(F)}, \quad r \in (1, \infty).
\end{align*}
\]

**Dirichlet forms.** Let \((\mathcal{E}, \mathcal{F})\) be a regular strongly local Dirichlet form on \(L^2(X, \mu)\) — see [FOT]. This means:

1. \(\mathcal{E}(f, g)\) is a symmetric bilinear form defined on a subspace \(\mathcal{F} \subset L^2(X)\).
2. \(\mathcal{E}\) is ‘Markov’: \(\mathcal{E}(f, f + 1, f, \wedge 1) \leq \mathcal{E}(f, f)\).
3. \(\mathcal{E}\) is closed: if \(||f||_H = \mathcal{E}(f, f) + ||f||_2^2\) then \(\mathcal{F}\) is a complete Hilbert space in \(||.||_H\).
4. \(\mathcal{E}\) is regular: \(\mathcal{F} \cap C_0(X)\) is dense in \(\mathcal{F}\) w.r.t. \(||.||_H\) and dense in \(C_0(X)\)
   w.r.t. \(||.||_\infty\).
5. \(\mathcal{E}\) is strongly local: if \(f_1, f_2\) have compact support and \(f_1\) is constant on an open set \(U_1\) containing \(\text{supp}(f_2)\), then \(\mathcal{E}(f_1, f_2) = 0\).

If \(D\) is a domain in \(X\) we define

\[ \mathcal{F}_D = \{ f \in \mathcal{F} : f = 0 \text{ on } D^c \}. \]

For \(f \in \mathcal{F}\) there exists a measure \(\Gamma(f, f)\) (called the ‘energy measure’) which gives \(\mathcal{E}\) by integration:

\[ \mathcal{E}(f, f) = \int_X d\Gamma(f, f). \]
For bounded $f$ the measure $\Gamma(f, f)$ is defined by requiring that
$$\int g d\Gamma(f, f) = 2\mathcal{E}(gf, f) - \mathcal{E}(g, f^2), \quad g \in C(X) \cap \mathcal{F}.$$ 

The measure $\Gamma(f, g)$ is then defined by polarisation. (See [FOT, p. 113], where the alternative notation $\mu \langle f \rangle$ is used.) For a manifold $d\Gamma(f, g) = (\nabla f, \nabla g) d\mu$, but in general $\Gamma$ need not be absolutely continuous with respect to $\mu$. $d\Gamma$ satisfies Leibnitz type rules – e.g.

$$d\Gamma(fg, h) = fd\Gamma(g, h) + gd\Gamma(f, h). \quad (3.3)$$

(Strictly speaking we may have to take a ‘quasi-continuous version’ of the functions $f, g$, but in these notes I will ignore regularity issues of this kind.)

**Definition 3.2.** We call a metric measure space with such a Dirichlet form a MMD space.

**Example 3.3.**
1. Manifolds.
2. Divergence form operators on a domain $D \subset \mathbb{R}^d$. Let $a = (a_{ij}(x))$ be bounded and measurable, and set
$$\mathcal{E}(f, f) = \int \nabla f \cdot a \nabla f. \quad (3)$$

3. The ‘cable system’ for a graph – see [Var]. If $G = (V, E)$ is a graph, let $X$ be the metric space obtained by replacing each edge $e \in E$ by a unit line segment $I_e$, joined at the vertices. Then
$$\mathcal{E}(f, f) = \sum_e \int_{I_e} f'' dx.$$ 

(These are now often, rather unfortunately, called ‘quantum graphs’).
4. Sierpinski carpets, where $\mathcal{E}$ is a subsequential limit of the approximations $\mathcal{E}_n$.

**Semigroup and heat kernel.** On a MMD space we have a Laplacian type operator $(\mathcal{L}, D(\mathcal{L}))$ which satisfies
$$\mathcal{E}(f, g) = \int (-\mathcal{L} f) g d\mu \quad \text{for } f \in D(\mathcal{L}), \quad g \in \mathcal{F}.$$ 

Associated with $\mathcal{L}$ is a semigroup, formally given by $P_t = \exp(t\mathcal{L})$. This semigroup defines a Markov process $(X_t, \mathbb{P}^x)$ on $X$, and
$$\mathbb{E}^x f(X_t) = \mathbb{E}(f(X_t) | X_0 = x) = P_t f(x), \quad x \in X, \quad t \geq 0.$$ 

If $P$ has a density $p_t(x, y)$ with respect to $\mu$ then this is the heat kernel (or transition density of the process $X$):

$$\mathbb{P}^x(X_t \in A) = P_t1_A(x) = \int_A p_t(x, y) \mu(dy). \quad (3.4)$$

The heat kernel is symmetric:

$$p_t(x, y) = p_t(y, x), \quad \mu \times \mu \text{ a.e.}, \quad (3.5)$$

satisfies the Chapman-Kolmogorov equation

$$\int p_s(x, y)p_t(y, z) \mu(dy) = p_{s+t}(x, z), \quad (3.6)$$
and the heat equation
\[ \frac{\partial}{\partial t} p_t(x, y) = \mathcal{L} p_t(x, y). \]

The definition (3.4) only defines \( p_t(x, y) \) for \( \mu \)-a.a. \( y \), but using the Chapman-Kolmogorov equation one can regularise \( p_t(x, y) \) so that (3.4)–(3.6) hold except on a set of capacity zero – see [Yan, GT3].

If \( D \subset X \) is open, we can also define the killed semigroup
\[ P_t^D f(x) = \mathbb{E}^x(f(X_t); t < \tau_D); \]
we write \( p_t^D(x, y) \) for the corresponding killed heat kernel.

**Heat kernel bounds and Hölder continuity.** This problem for divergence form operators in \( \mathbb{R}^d \) was solved by de Giorgi, Moser and Nash in the late 1950s – see in particular [Mo1, N]. Moser’s methods were extended to manifolds by Bombieri and Giusti in [BG]. These papers were followed by work of Aronson [Ar], Li and Yau [LY], Fabes and Stroock [FS], Grigoryan [Gr1], Saloff-Coste [SC1], Sturm [St3], and others. The following theorem summarises much of this work.

**Theorem 3.4.** [Gr1, SC1, St3] Let \((X, \mathcal{E})\) be a MMD space. The following are equivalent:

1. \( p_t(x, y) \) satisfies (GB) – that is Gaussian bounds,
2. A parabolic Harnack inequality (PHI) holds on \( X \),
3. \( X \) satisfies VD plus PI (Poincaré inequality).

This theorem gives the equivalence between three possible conditions relating to the heat equation on the MMD space \((X, \mathcal{E})\). Before I discuss further this result, and its usefulness, I need to define the various terms in the theorem.

**Definition 3.5.** \((p_t)\) satisfies Gaussian bounds (GB) if
\[ p_t(x, y) \overset{(c)}{=} V(x, ct^{1/2})^{-1} \exp(-cd(x, y)^2/t). \]

**Definition 3.6.** \((X, \mathcal{E})\) satisfies PI if for all balls \( B = B(x, r) \), \( f : B \to \mathbb{R} \),
\[ \min_a \int_B (f - a)^2 \, d\mu = \int_B (f - \overline{f})^2 \, d\mu \leq C \rho^2 \int_B d\Gamma(f, f). \]

For the third definition, of a Harnack inequality, we need to define harmonic functions (solutions of \( \mathcal{L} h = 0 \)) and *caloric* functions (solutions of \( \partial_t u(x, t) = \mathcal{L} u \)) in the MMD context. If \( D \subset X \) is open, we say \( h : D \to \mathbb{R} \) is harmonic in \( D \) if
\[ \mathcal{E}(f, h) = 0 \text{ for all } f \in C_0(X) \cap \mathcal{F}, \text{ supp } f \subset D. \]

There are various definitions of *caloric*, all of which essentially say that
\[ \frac{\partial}{\partial t} u(x, t) = \mathcal{L} u(x, t), (x, t) \in Q \subset X \times \mathbb{R}. \]
Lack of regularity in general means they may not be exactly equivalent, but in all cases the heat kernel \( p_t(x, y) \) and \( p_t^D(x, y) \) are caloric.
Harnack inequalities.

**Definition 3.7.** \((\mathcal{X}, \mathcal{E})\) satisfies the elliptic Harnack inequality (EHI) if whenever \(B = B(x, R)\) and \(h : \overline{B} \to \mathbb{R}_+\) is harmonic in \(B\) then if \(B' = B(x, R/2)\)
\[
\sup_{B'} h \leq C_E \inf_{B'} h.
\]
(Strictly speaking, on a MMD space we may have a small exceptional set, so should write ess sup and ess inf above.)

Easy iteration arguments show that if EHI holds then harmonic functions are Hölder continuous, with index \(\delta = \delta(C_E)\).

The parabolic Harnack inequality (PHI) is more complicated to state.

**Definition 3.8.** Let \(T = \mathbb{R}^2\), \(B = B(x, R)\), \(Q = B \times (0, T)\), and \(Q_- = B' \times [\frac{1}{4}T, \frac{1}{2}T], \quad Q_+ = B' \times [\frac{3}{4}T, T]\).

Then \((\mathcal{X}, \mathcal{E})\) satisfies PHI if whenever \(u = u(x, t)\) is non-negative and caloric in \(Q\) then
\[
\sup_{Q_-} u \leq C_H \inf_{Q_+} u.
\]

**Remark 3.9.**
1. Note that PHI implies EHI since if \(h\) is harmonic then \(u(x, t) = h(x)\) is caloric.
2. Iteration arguments show that PHI implies Hölder continuity of caloric functions (except on an exceptional set).
3. Theorem 3.4 gives necessary and sufficient conditions for PHI. It also proves that GB and PHI are stable: if \(\mathcal{E}'\) is another Dirichlet form and \(\mathcal{E}' \asymp \mathcal{E}\) then PHI holds for \((\mathcal{X}, \mathcal{E}')\) iff it holds for \((\mathcal{X}, \mathcal{E})\). The PI is clearly stable, while the stability of PHI or GB is far from evident.
4. Necessary and sufficient conditions for EHI are not known. It is also not known whether or not EHI is stable.
5. The theorem provides an effective tool for proving that GB and PHI hold for a space \((\mathcal{X}, \mathcal{E})\), since VD and PI are often quite straightforward to prove. (PI can often be obtained from an isoperimetric inequality.)

**Extensions to (fractal) MMD spaces.** We can replace the Gaussian bound on the heat kernel by more general bounds. Since I want to discuss spaces, such as the pre-Sierpinski carpet or quantum graphs, where the local and global structures are different, introduce the space-time scaling function:

\[
\Psi(r) = \begin{cases} 
R^{\beta L} & \text{if } 0 \leq r \leq 1, \\
R^\beta & \text{if } r \geq 1.
\end{cases}
\]

(If the space is locally Euclidean we will have \(\beta_L = 2\); this is the case for the pre-SC and quantum graphs).

**Definition 3.10.** We say \((p_t)\) satisfies \(\text{HK}(\Psi)\) if

\[
p_t(x, y) \asymp V(x, ct^{1/\beta_L})^{-1} \exp(-c(d(x, y)^{\beta_L}/t)^{1/(\beta_L - 1)}), \quad \text{in } I_{\text{loc}}
\]

\[
p_t(x, y) \asymp V(x, ct^{1/\beta})^{-1} \exp(-c(d(x, y)^{\beta}/t)^{1/(\beta - 1)}), \quad \text{in } I_{\text{glob}}
\]
Here

\[ I_{\text{loc}} = \{ (t, x, y) : t \leq 1 \lor d(x, y) \}, \quad I_{\text{glob}} = \{ (t, x, y) : t \geq 1 \lor d(x, y) \}. \]

If \( \Psi(r) = r^\beta \) we write this condition as \( \text{HK}(\beta) \).

**Example 3.11.**

1. The pre-SC satisfies \( \text{HK}(\Psi) \) with \( \beta_L = 2, \beta = d_w \).
2. The true (infinite) SC \( \hat{F} \) satisfies \( \text{HK}(\beta) \) with \( \beta = d_w \).
3. GB are just \( \text{HK}(2) \).
4. The cable graph associated with the Sierpinski gasket (see Figure 1) satisfies \( \text{HK}(\Psi) \) with \( \beta_L = 2, \beta = \log 5 / \log 2 \).

It is natural to ask what values \( \beta \) can take.

**Theorem 3.12.** Suppose an (infinite, connected) MMD space \( (X, E) \) satisfies \( (V_\alpha) \) and \( \text{HK}(\beta) \). Then \( \alpha \geq 1, 2 \leq \beta \leq 1 + \alpha \), and all these values are possible.

**Proof.** These bounds on \( \alpha, \beta \) were given (without proof) in my Saint Flour notes [Ba1, Section 3]. See [GHL] for a proof that \( 2 \leq \beta \leq 1 + \alpha \) in the MMD context.

Hino [Hi1] proved that \( \beta \geq 2 \) by showing in general that

\[
\liminf_{t \downarrow 0} t \log p_t(x, y) \geq -C(x, y) > -\infty.
\]

If \( \text{HK}(\beta) \) holds then the LHS is \( ct \log t - c'd(x, y)^{c't^{(\beta-2)/(\beta-1)}} \), so would diverge to \(-\infty\) if \( \beta < 2 \).

An alternative argument would be to use the Davies-Gaffney bound (see [Da1]), which gives for functions \( f_i \in L^2 \) with supports \( B(x_i, r) \) that if \( R \geq 3r \) then

\[
\langle P_t f_1, f_2 \rangle \leq c_1 \| f_1 \|_2 \| f_2 \|_2 \exp(-c_2 R^2 / t).
\]

In [Ba1] I conjectured that if \( \beta = 2 \) then only \( \alpha \in \mathbb{N} \) was possible. Bourdon and Pajot [BPa], and Laakso [La], in answering a question in [HS], gave examples which showed otherwise. In the graph case I gave constructions for all \( \alpha, \beta \) satisfying the inequalities above using Laakso’s method – see [Ba4]. The same construction will work in the MMD setting.

The condition \( \beta \geq 2 \), combined with polynomial volume growth, means that heat (or the diffusion \( X_t \)) can move at most distance \( O(t^{1/2}) \) in time \( t \). So (as the Davies-Gaffney bound shows) Euclidean space gives essentially the fastest possible speed of heat diffusion. Obstacles such as the cut out cubes in the Sierpinski carpet can slow \( X \) down, but not speed it up.

**EHI and PHI.** The pre-carpet gives an example where EHI holds but PHI fails. Let \( R \gg 1, Q = B(x_0, R) \times [0, T] \) where \( T = R^2 \), and let \( d(x_1, x_0) = R/3 \). Set \( u(x, t) = p_t(x_0, x) \). Then

\[
\sup_{Q} u \asymp p_{T/4}(x_0, x) \asymp c T^{-\alpha/\beta},
\]
while
\[
\inf u \leq p_T(x_0, x_1) \leq c T^{-\alpha/\beta} \exp \left( -c(R^3/T)^{1/(\beta-1)} \right) \\
= c T^{-\alpha/\beta} \exp(-c' R^{(\beta-2)/(\beta-1)}) \ll \sup_{Q^-} u.
\]

The PHI fails in this example because heat needs time \( O(R^\beta) \) rather than \( O(R^2) \) to flow from \( x_0 \) to \( x_1 \). The fix is clear: define a modified PHI to take account of the space time scaling \( t = \Psi(r) \).

**Definition 3.13.** \((\mathcal{X}, \mathcal{E})\) satisfies PHI(\(\Psi\)) if when \( R > 0, T = \Psi(R), B = B(x,R), Q = B \times (0,T), Q^- = B' \times [\frac{1}{4}, \frac{T}{2}], Q^+ = B' \times [\frac{T}{4}, T] \), and \( u = u(x,t) \) is non-negative and caloric in \( Q \) then
\[
\sup_{Q^-} u \leq C H \inf_{Q^+} u.
\]

In the same way one defines the rescaled Poincaré inequality PI(\(\Psi\)):
\[
\min a \int_{B(x,r)} (f - a)^2 d\mu = \int_{B(x,r)} (f - \overline{f})^2 d\mu \leq C P \Psi(r) \int_{B(x,r)} d\Gamma(f,f).
\]

If \( \Psi(r) = r^\beta \) we write PHI(\(\beta\)), PI(\(\beta\)). The original PI and PHI are just PHI(2) and PI(2).

Given Theorem 3.4 a natural first guess would be that
\[
(3.9) \quad HK(\Psi) \Leftrightarrow PHI(\Psi) \Leftrightarrow VD + PI(\Psi).
\]

The first double implication in (3.9) is correct.

Introduce two additional conditions. We say T(\(\Psi\)) holds for the diffusion \( X \) if
\[
\mathbb{E}^x \tau_{x,R} \asymp \Psi(R), \quad x \in \mathcal{X}, \quad R > 0.
\]

We can define the effective resistance between two sets in the MMD context by setting
\[
R_{\text{eff}}(B_1, B_2)^{-1} = \inf \{ \mathcal{E}(f,f) : f = 1 \text{ on } B_1, f = 0 \text{ on } B_2 \}.
\]

Then we say RES(\(\Psi\)) holds for \((\mathcal{X}, \mathcal{E})\) if
\[
R_{\text{eff}}(B(x,R), B(x,2R)^c) \asymp \frac{\Psi(R)}{V(x,R)}, \quad x \in \mathcal{X}, \quad R > 0.
\]

**Theorem 3.14.** [HSC, GT3, BBKT2]. The following are equivalent:
(a) \((\mathcal{X}, \mathcal{E})\) satisfies HK(\(\Psi\)).
(b) \((\mathcal{X}, \mathcal{E})\) satisfies PHI(\(\Psi\)).
(c) \((\mathcal{X}, \mathcal{E})\) satisfies VD, EHI and T(\(\Psi\)).
(d) \((\mathcal{X}, \mathcal{E})\) satisfies VD, EHI and RES(\(\Psi\)):
These all imply PI(\(\Psi\)).

Note that this theorem does not give stability of PHI(\(\Psi\)). The conditions VD and RES(\(\Psi\)) are clearly stable, but this is not apparent for any of the other ones.

It is easy to see that one cannot have HK(\(\Psi\)) \(\Leftrightarrow\) VD + PI(\(\Psi\)) in general. If \( \Psi_1 \geq \Psi \) then PI(\(\Psi\)) implies PI(\(\Psi_1\)), while if \( \Psi_1(r)/\Psi(r) \to \infty \) then HK(\(\Psi\)) and HK(\(\Psi_1\)) cannot both hold.

The inequality PI(\(\Psi\)) tells us, roughly, that heat homogenizes in a ball size \( R \) in time at most \( \Psi(R) \), but does not preclude the possibility it might homogenize
more quickly. To ‘capture’ HK(Ψ) one needs an upper bound on the rate of heat diffusion. (This is not necessary in the classical case Ψ(r) = r^2 because this is the fastest possible.) This theorem shows that, combined with the regularity coming from EHI, control of either resistance or exit times is sufficient for this.

**Sketch proof of Theorem 3.14.** The equivalence of (a)–(d) in the graph context is given in [GT2], and except for technical issues arising from the need for regularity, the same arguments work for metric measure space. Nearly all the implications have been written out in the MMD case: the equivalence of (a) and (c) is proved in [GT3], of (a) and (b) in [HSC], and that (d) implies (c) in [BBKT2].

1. The equivalence HK(Ψ) ⇔ PHI(Ψ) is proved very much as in the classical Ψ(r) = r^2 case – see in particular [FS].
2. Given HK(Ψ) + PHI(Ψ) one gets EHI immediately, while the proof that PHI(Ψ) implies VD runs as in the classical case given in Theorem 3.4. T(Ψ) is easy since integration of the heat kernel bounds gives
   \[ \mathbb{P}^x(X_t \notin B(x, \lambda t^{1/\beta})) \leq \exp(-c\lambda^{\beta/(\beta-1)}). \]
3. A general result using potential theory gives that if \( B' = B(x, r/2), B = B(x, r) \) then there exists a probability measure \( \pi \) on \( \partial B' \) such that if \( \tau_B \) is the first exit by \( X \) from \( B \) and \( T' \) the first hit on \( B' \) then
   \[ \int_{\partial B'} \mathbb{E}^z \tau_B = R_{\text{eff}}(B', B^c) \int_{B - B'} \mu(dy) \mathbb{P}^y(T' \leq \tau_B). \]
   Using the regularity from EHI one can then show that
   \[ \mathbb{E}^z \tau_B \asymp R_{\text{eff}}(B', B^c)V(x, r), \quad z \in \partial B'. \]
   So VD + EHI + T(Ψ) ⇔ VD + EHI + RES(Ψ).
4. What remains is the hardest (and most useful) implication:
   \[ \text{VD + EHI + T(Ψ) \Rightarrow HK(Ψ).} \]
   Like most heat kernel bounds, this proceeds in several steps.

**Step 1.** Obtain the upper bound
\[ \sup_x p_t^B(x, x) \leq \frac{c(\kappa)}{V(x, t^{1/\beta})}, \quad \text{(DUE)} \]
where \( B = B(x, \kappa t^{1/\beta}). \) Write \( \lambda_1(D) \) for the smallest eigenvalue in a domain \( D \subset X: \lambda_1(D) \) is given by the variational formula
\[ \lambda_1(D) = \inf \left\{ \frac{\mathcal{E}(f, f)}{||f||^2_2} : f \in \mathcal{F}_D, f \neq 0 \right\}. \]
Then the main step is to prove that there exists \( \nu > 0 \) such that if \( B = B(x_0, R) \) and \( D \subset B \) then
\[ \lambda_1(D) \geq \frac{c}{\Psi(R)} \left( \frac{\mu(B)}{\mu(D)} \right)^{\nu}. \]
This called a Faber-Krahn inequality, and this is known to imply (DUE) – see [Gr2] for the manifold case, and [GH] for the MMD case.
To prove (3.12) write \( g_B(x,y) \) for the Green function in a domain \( B \subset \mathcal{X} \). The condition \( T(\Psi) \) gives bounds on
\[
h(x) = \mathbb{P}^x \tau_{x,r} = \int_{B(x,r)} g_B(x,y) \mu(dy).
\]
The EHI enables one to pass from integrals to pointwise bounds on Green’s functions in balls, and using this one can construct a ‘candidate function’ \( f \) which when used in (3.11) gives the Faber-Krahn inequality.

**Step 2.** Establish the full upper bound in HK(\( \Psi \)). In general obtaining the full upper bound from DUE can be quite challenging, and usually requires a lengthy proof. However, in this situation the estimate \( T(\Psi) \) makes it reasonably straightforward. \( T(\Psi) \) states that the time taken to leave a ball \( B(x,r) \) is of order \( \Psi(r) \). It follows that there exists \( p_0, \delta > 0 \) such that
\[
P^x(\tau_{x,R} \leq \delta \Psi(r)) \leq \frac{1}{3}. \tag{3.13}
\]
Given a ball \( B(x,R) \) divide the ‘journey’ of \( X \) from \( x \) to \( B(x,R) \) into steps of length \( r = R/n \); clearly there have to be at least \( n \) of these. Call a part of the journey (i.e., across a ball \( B(y,r) \)) ‘quick’ if it takes time less than \( \delta \Psi(r) \), and ‘slow’ otherwise. Then the number of quick journeys is stochastically dominated by a Binomial random variable, and standard bounds give that
\[
P(\text{more than } 2n/3 \text{ of the journeys are quick}) \leq e^{-cn}. \tag{3.14}
\]
If the number of quick journeys out of the first \( n \) is less than \( 2n/3 \) then the time to exit the ball is at least
\[
t(n) = \frac{1}{3} \delta \Psi(R/n) = \frac{1}{3} \delta \Psi(R/n)n. \tag{3.15}
\]
Thus
\[
P^x(\tau_{x,R} < t(n)) \leq \exp(-cn).
\]

Given \( R \) and \( T \), we wish to choose \( n \) so that \( T = t(n) \). Using (3.13) we have \( T \asymp R^\gamma n^{1-\gamma} \), where \( \gamma = \beta \) if \( R/n \geq 1 \) and \( \gamma = \beta_L \) if \( R/n < 1 \). It follows that \( n^{\gamma-1} \asymp R^{\gamma}/T \), and so that
\[
R^{\gamma-1}/n^{\gamma-1} \asymp \frac{T}{R}. \tag{3.16}
\]

So if \( T > R \) then \( r > 1 \), and \( \gamma = \beta \), while if \( T < R \) then \( \gamma = \beta_L \). (This calculation neglects the effect of the constant \( \frac{1}{3} \delta \), but both cases give \( n = O(R) \) when \( T \asymp R \).) We thus obtain
\[
\log P^x(\tau_{x,R} < T) \leq \begin{cases} -c(R^\beta/T)^{1/(\beta-1)} & \text{if } R < T, \\ -c(R^{\beta_L}/T)^{1/(\beta_L-1)} & \text{if } R \geq T. \end{cases}
\]

One can then combine this bound with DUE to obtain the full upper bound.

**Step 3.** Obtain the lower bound. The upper bound leads easily to:
\[
p_t(x,x) \geq cV(x, t^{1/\beta})^{-1}.
\]
The key is to extend this inequality to a ball around \( x \):
\[
p_t(x,y) \geq cV(x, t^{1/\beta})^{-1}, \quad y \in B(x, \delta t^{1/\beta}). \tag{3.17}
\]
EHI gives Hölder continuity of harmonic functions. With other estimates, this extends also to give control of oscillations of \( p_t(x,y) \), and hence control of
\[
|p_t(x,x) - p_t(x,y)|.
\]
Once one has the ‘near diagonal lower bound’ (3.16) a standard chaining argument as in [FS, BB4] then gives the full lower bound.

\[\square\]

Proof of Theorem 2.11. The heat kernel bounds on the pre-SC \( \tilde{F} \) follow from the implication (d) implies (a) of Theorem 3.14. VD is immediate from (3.1) – (3.2), and EHI is proved in Theorem 2.9. Theorem 2.8 gives that if \( r = 3^n \) with \( n \geq 0 \) then the resistance across a cube of side \( r \) is \( R_n \approx \rho^n = \varphi^{d_w - df} \). Using this the condition RES(Ψ) follows, where \( \beta_L = 2 \) and \( \beta = d_w \).

\[\square\]

Stable conditions for HK(Ψ). As remarked above, the conditions in Theorem 3.14 are not stable.

**Theorem 3.15.** [BBK] The following are equivalent:
(a) \((X,\mathcal{E})\) satisfies HK(Ψ).
(b) \((X,\mathcal{E})\) satisfies VD, PI(Ψ) and CS(Ψ).

The condition CS(Ψ) states that there exist ‘low energy’ cut-off functions.

**Definition 3.16.** A function \( \varphi : X \to \mathbb{R}_+ \) is a cutoff function for \( B(x,R/2) \subset B(x,R) \) if \( \varphi(y) = 1 \) on \( B(x,r/2) \) and is zero on \( B(x,r) \).

\( (X,\mathcal{E}) \) satisfies CS(Ψ) if there exists \( \theta \in (0,1] \) such that for all \( x,R \) there exists a Hölder continuous cutoff function \( \varphi \) for \( B(x,R/2) \subset B(x,R) \) such that if \( s \in (0,R) \), \( y \in B(x,R) \), and \( f : B = B(y,s) \to \mathbb{R} \) then
\[
(3.17) \quad \int_{B(y,s/2)} f^2 d\Gamma(\varphi,\varphi) \leq c(s/R)^{2\theta} \left( \int_{B(y,s)} d\Gamma(f,f) + \Psi(s)^{-1} \int_{B(y,s)} f^2 d\mu \right).
\]

Note that if \( \Psi_1 \geq \Psi_2 \) then CS(Ψ_2) implies CS(Ψ_1): thus increasing \( \Psi \) weakens PI(·) but strengthens CS(·).

CS(2) always holds on a manifold – just take \( \varphi \) ‘linear’ between \( B(x,R/2) \) and \( B(x,R) \), so that \( ||\nabla \varphi||_\infty \leq cR^{-1} \). Then
\[
\int_{B(y,s/2)} f^2 d\Gamma(\varphi,\varphi) = \int_{B(y,s/2)} f^2 ||\nabla \varphi||^2_\infty d\mu \leq \int_{B(y,s/2)} f^2 d\mu ||\nabla \varphi||_\infty^2 \leq cR^{-2} \int_{B(y,s/2)} f^2 d\mu \leq c(s/R)^2 s^{-2} \int_{B(y,s)} f^2 d\mu.
\]

Proof of Theorem 3.15 (1) HK(Ψ) implies VD, PI(Ψ) and CS(Ψ). The first two follow by Theorem 3.14. To prove CS(Ψ) one uses properties of Green’s functions to construct a suitable \( \varphi \). Given balls \( B' = B(x_0,R/2) \subset B = B(x_0,R) \), let \( B_1 = B(x_0,2R/3) \). Define the \( \lambda \)-resolvent for the killed process by
\[
G_{\tilde{B}}^\lambda f(x) = \mathbb{E}^x \int_0^\tau f(X_s) e^{-\lambda s} ds.
\]
Set $\lambda = \Psi(R)^{-1}$ and $h(x) = G^\lambda_B 1_B(x)$. Then $h$ has support $B$, $h \leq G^\lambda 1 = \lambda^{-1}$, and the heat kernel bounds give that $h \geq e\Psi(R)$ on $B^\prime$. Thus $\varphi = 1 \wedge e\Psi(R)^{-1}h$ is a cutoff function for $B^\prime$. The parabolic Harnack inequality gives Hölder continuity of the heat kernel, and hence of $\varphi$, so it remains to verify (3.17). A starting point for this is the observation that, writing $E^\lambda(g,g) = E(g,g) + \lambda \langle g, g \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the inner product on $L^2(X, \mu)$, then if $g$ has support $B^\prime$ by [FOT, Theorem 4.4.1] $E^\lambda(g,h) = E^\lambda(g, G^\lambda_B 1_B) = \langle g, 1_B \rangle$.

We therefore have, using the inequality $2ab \leq \frac{1}{2}a^2 + 2b^2$, and the Leibnitz rule (3.3),
\[
\int f^2 d\Gamma(h,h) = \int d\Gamma(f^2 h, h) - 2 \int f h d\Gamma(f, h) \\
\leq E^\lambda(f^2 h, h) + \frac{1}{2} \int f^2 d\Gamma(h, h) + 2 \int h^2 d\Gamma(f, f) \\
= \langle fh^2, 1_B \rangle + \frac{1}{2} \int f^2 d\Gamma(h, h) + 2 \int h^2 d\Gamma(f, f).
\]
This implies (3.17) with $s = R$. Obtaining the inequality on smaller balls requires some more work: see [BBK, Proposition 3.4] for more details on this argument.

(2) In many arguments one needs to control expressions of the form
\[
\int_A f^2 |\nabla \varphi|^2 d\mu,
\]
where $A = B(x, r + h) - B(x, r)$ and $\varphi$ is a cutoff function between $B(x, r)$ and $B(x, r + h)$. The CS inequalities enable one to do this. The ‘extra’ term
\[
\int_A d\Gamma(f, f)
\]
often turns out to be harmless or controllable.

In particular, CS($\Psi$) provides a family of cutoff functions which enable one to run the first (‘easy’) part of Moser’s argument. Thus one can use the methods of [Mo1, Mo2, Mo3] to show that VD, PI($\Psi$) and CS($\Psi$) imply EHI. Once one has EHI one can use Theorem 3.14, since PI($\Psi$) and CS($\Psi$) also imply RES($\Psi$). □

**Remark 3.17.** (1) The condition CS($\beta$) is unfortunately rather complicated. However Theorem 3.15 does not preclude the possibility that there might exist a simpler condition $?(\Psi)$ such that $?(\Psi) + PI(\Psi) + VD \Leftrightarrow HK(\Psi)$. One candidate for this is RES($\Psi$).

(2) If $(E_i, \mathcal{F}_i)$, $i = 1, 2$ are strongly local Dirichlet forms with
\[
C^{-1} E_1(f, f) \leq E_2(f, f) \leq CE_1(f, f), \quad f \in \mathcal{F},
\]
then by [LJ, Proposition 1.5.5(b)] their energy measures satisfy
\[
C^{-1} \Gamma_1(f, f) \leq \Gamma_2(f, f) \leq C\Gamma_1(f, f), \quad f \in \mathcal{F}.
\]
(See also [Mos], p. 389.) Thus the condition CS($\Psi$) is stable, and therefore this theorem does give the stability of HK($\Psi$) and PHI($\Psi$).

(3) A recent paper [AB] provides some simplification of CS($\Psi$).
Suppose that $V_\alpha$ holds and $\Psi(r) = r^\beta$, so that $\text{HK}(\beta)$ takes the form

$$p_t(x, y) \sim c t^{-\alpha/\beta} \exp(-c(d(x, y)^\beta/t)^{1/(\beta-1)}).$$

In particular $p_t(x, x) \sim t^{-\alpha/\beta}$. The behaviour of $X$ can be divided into two main cases.

(1) $\alpha < \beta$. In this case the process $X_t$ is recurrent, and in fact ‘hits points’, so that the range

$$R_t = \{X_s, 0 \leq s \leq t\}$$

has positive $\mu$ measure. Ultimately the process $X$ will hit every point in the space $\mathcal{X}$ infinitely often.

(2) $\alpha > \beta$. In this case the process $X_t$ is transient, and $R_t$ has zero $\mu$ measure. The probability that $X$ hits any specific point after time 0 is zero. The Green function on the whole space $\mathcal{X}$ exists and satisfies

$$g(x, y) = \int_0^\infty p_t(x, y) dt \sim \frac{1}{d(x, y)^{\alpha-\beta}}.$$  

In the classical case of $\mathbb{R}^d$ the first case (1) $\alpha < \beta$ only arises when $d = 1$. In the MMD setting we have a richer family of ‘low dimensional’ spaces; in fact nearly all the exact fractals studied in the physics literature turn out to have $\alpha < \beta$.

The resistance estimates on $R_n$ are good enough to show that the basic SC is recurrent if $d = 2$, and transient for $d \geq 3$. One can find generalised Sierpinski carpets in $d \geq 3$ for which $d_f > 2$ but $d_w < d_f$, so that $X$ is still recurrent – see [BB5].

In low dimensions one has a much nicer version of the stability result. Introduce the condition $\text{PR}(\beta)$ (for ‘pointwise resistance’):

$$R_{\text{eff}}(x, y) = R_{\text{eff}}(\{x\}, \{y\}) \sim \frac{d(x, y)^\beta}{V(x, d(x, y))}.$$  

In the recurrent case $R_{\text{eff}}(x, y)$ is a metric –see [Kig3].

**Theorem 3.18.** [Kum]. Let $\mathcal{X}$ satisfy $V_\alpha$, and let $\beta > \alpha$. Then the following are equivalent:

1. $(\mathcal{X}, \mathcal{E})$ satisfies $\text{HK}(\beta)$.
2. $(\mathcal{X}, \mathcal{E})$ satisfies $\text{PR}(\beta)$.

So in this case Harnack inequalities are no longer difficult, and one just has to calculate resistance between points.

This result was originally proved in the graph context in [BCK], and has applications to random walks on critical percolation clusters – see [BJKS, KN]. It is interesting to note that while de Gennes initially proposed the use of random walks to study resistance of percolation clusters, the mathematics has been in the opposite direction: using resistances to study random walks.

In the transient case one has $R_{\text{eff}}(x, y) = \infty$, and in spaces such as $\mathbb{R}^d$

$$\limsup_{d(x, y) \to \infty} R_{\text{eff}}(B(x, \varepsilon), B(y, \varepsilon)) < \infty.$$  

Thus resistance between points or small sets gives no useful information.
Rough isometries. Let \((X_i, d_i, \mu_i)\) be metric measure spaces. A map \(\varphi : X_1 \to X_2\) is a rough isometry if there exist constants \(C_1, C_2, C_3\) such that

\[
C_1^{-1}(d_1(x, y) - C_2) \leq d_2(\varphi(x), \varphi(y)) \leq C_1(d_1(x, y) + C_2),
\]

\[
X_2 = \bigcup_{x \in X_1} B_{d_2}(\varphi(x), C_2),
\]

\[
C_3^{-1}\mu_1(B_{d_1}(x, C_2)) \leq \mu_2(B_{d_2}(\varphi(x), C_2)) \leq C_3\mu_1(B_{d_1}(x, C_2)).
\]

If there exists a rough isometry between two spaces they are said to be roughly isometric. (One can check that this is an equivalence relation.) If two spaces are roughly isometric then they have the same large scale structure, although their local structures may be quite different. For example, \(\mathbb{Z}^d\) and \(\mathbb{R}^d\) are roughly isometric, as are the pre-carpet \(\tilde{F}\) and the infinite SC \(\hat{F}\).

Theorems 3.15 and 3.18 imply that the global part of \(\operatorname{HK}(\Psi)\), i.e. that relating to the index \(\beta\), is stable under rough isometries. (One also needs some local regularity of both spaces – see [BBK, Theorem 2.21] for more details.)

In particular if \(M\) is a manifold (with some local regularity) which is roughly isometric to the pre-SC \(\tilde{F}\) then the heat kernel on \(M\) also satisfies \(\operatorname{HK}(\Psi')\) – where \(\Psi'\) is given by (3.7) with indices \(\beta\) and 2.

**Local structure.** In the MMD setting we have a Dirichlet form \((\mathcal{E}, \mathcal{F})\) on \(L^2(X, \mu)\). Suppose that \(\operatorname{HK}(\beta)\) holds: we can then ask about the regularity properties of the functions in \(\mathcal{F}\).

**Theorem 3.19.** Let \(X\) be the Sierpinski carpet. Suppose that \(f \in C^1(\mathbb{R}^2)\) and \(g = f|_F \in \mathcal{F}\). Then \(g\) is constant.

However, \(\mathcal{F}\) can be described explicitly as a Besov space. This was done for the Sierpinski gasket in [Jo], but holds much more generally. Set

\[
N_\theta(f) = \sup_{0 < r \leq 1} r^{-\alpha} \int_F V(x, r)^{-1} \int_{B(x, r)} |f(x) - f(y)|^2 \mu(dx)\mu(dy),
\]

\[
W(\theta) = \{f \in L^2(F, \mu) : N_\theta(f) < \infty\}.
\]

**Theorem 3.20.** [GHL, KS] Suppose \(\operatorname{HK}(\beta)\) holds. Then \(\mathcal{F} = W(\beta)\), and

\[
\mathcal{E}(f, f) \asymp N_\beta(f).
\]

Further,

\[
\beta = \sup\{\theta : \dim W(\theta) = \infty\}.
\]

**Corollary 3.21.** Let \(\mathcal{E}_1, \mathcal{E}_2\) be Dirichlet forms on a metric measure space \(X\). If \((X, \mathcal{E}_i)\) satisfy \(\operatorname{HK}(\beta_i)\), then \(\beta_1 = \beta_2\).

**Open Problem 2.** This result gives a characterization of \(\beta = d_w\) which appears to be independent of the Dirichlet form \(\mathcal{E}\). Can one use it to obtain bounds on the ‘walk dimension’ \(d_w\) of the SC?
Energy measures. These are the measures $d\Gamma(f, f)$, $f \in \mathcal{F}$. Suppose that $(\mathcal{X}, \mathcal{E})$ satisfies HK($\beta$). If one had $d\Gamma(f, f) \leq C(f) d\mu$ then we could define a metric associated with $\mathcal{E}$ by setting

$$L = \{ f : d\Gamma(f, f) \leq d\mu \},$$

and

$$d\mathcal{E}(x, y) = \sup \{ f(y) - f(x) : f \in L \}.$$  

One would then expect to have Gaussian heat kernel bounds with respect to $d\mathcal{E}$, and on-diagonal heat kernel decay of the form $t^{-d_f(F)/2}$ – see [St2].

**Theorem 3.22.** [Hi2] Let $\mathcal{X}$ be a Sierpinski carpet. If $f \in \mathcal{F}$ and $d\Gamma(f, f) \ll d\mu$, then $f$ is constant.

Kusuoka [Kn2] proved a similar result for the Sierpinski gas ket (SG). Thus the energy measures and Hausdorff measure are mutually singular.

References


ANALYSIS ON THE SIERPINSKI CARPET


