Heat kernel estimates and Harnack inequalities

2025 PIMS-CRM summer school in probability

Mathav Murugan

version of June 3, 2025

This notes will be updated throughout the summer school and made available at https://personal.math.ubc.ca/~mathav/teaching/notes/ssprob25.pdf. If you find any mistakes, typos, or have any other feedback, please let me know. Solutions to exercises will be posted after the end of the school in this document.

Contents

1 Introduction

The study of heat kernels and their estimates is at the interface of analysis, geometry and probability. Let us start with some examples of heat kernel. The fundamental solution of the heat equation on $\mathbb{R}^n$

$$\partial_t u = \frac{1}{2}\Delta u, \tag{1.1}$$

is given by the classical Gauss–Weierstrass kernel

$$p_t(x,y) = \frac{1}{(2\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{2t}\right). \tag{1.2}$$

That is, for any fixed $x \in \mathbb{R}^n$, the function $(t,y) \mapsto p_t(x,y)$ solves the heat equation on $(0,\infty)\times\mathbb{R}^n$ and $\lim_{t\downarrow 0} p_t(x,\cdot) = \delta_x$, where δ_x is the Dirac mass at x in the sense of distributions. In particular, for any $f \in C_c^\infty(\mathbb{R}^n)$, the function $u(t,x) = \int_{\mathbb{R}^n} p_t(x,y)f(y)\,dy, (t,x) \in (0,\infty)\times\mathbb{R}^n$ solves the Cauchy problem for the heat equation (1.1) with initial condition $u(0,x) = f(x)$ for all $x \in \mathbb{R}^n$.

The fundamental solution of the heat equation can alternately be viewed as the transition probability density of the standard Brownian motion as $p_t(x,\cdot)$ is the density of the multivariate normal random variable with mean x and covariance matrix $tI_{n\times n}$. Thus the measure $p_t(x,y)dy$ is the law of the Brownian motion B_t starting at $B_0 = x$. By extension, we use the term *heat kernel* to refer to the transition kernel of a variety of Markov processes.

It is known that the behavior heat kernel is closely related to the geometry of the underlying space. A classical result in this direction is the Varadhan's asymptotic formula [Var] (see also [Nor, HR]). It states that the fundamental solution of the heat equation $\partial_t u = \frac{1}{2}\Delta u$ on a Riemannian manifold satisfies

$$\lim_{t\downarrow 0} 2t \log p_t(x,y) = -d(x,y)^2,$$

where Δ is the Laplace-Beltrami operator and d denotes the Riemannian distance. Despite the universal nature of Varadhan's short time asymptotic mentioned above, the long time behavior of the heat kernel of a manifold can be significantly different from the Euclidean case. For instance, the heat kernel of the 3-dimensional hyperbolic space $\mathbb{H}^3$ is given by [DM, Theorem 2.1]

$$p_t(x,y) = \frac{1}{(2\pi t)^{3/2}} \frac{d(x,y)}{\sinh(d(x,y))} \exp\left(-\frac{t}{2} - \frac{d(x,y)^2}{2t}\right).$$

In the expression (1.2) for the heat kernel, the term $|x-y|^2/t$ reflects that time scales like the square of the distance (Brownian space-time scaling). There are Markov processes with other space-time scaling behavior. We describe two such examples: jump processes on $\mathbb{R}^n$ and diffusions on fractals.

Let (Y_t) be the symmetric α-stable process (where $\alpha \in (0,2)$) on $\mathbb{R}^n$; that is, (Y_t) is a Lévy process (stationary, independent increments) such that

$$\mathbb{E}[e^{i\xi\cdot(Y_{t+s}-Y_s)}] = \exp(-t|\xi|^\alpha), \quad \text{for all } t,s \geqslant 0 \text{ and } \xi \in \mathbb{R}^n.$$

Then the law of (Y_t) admits a density (heat kernel) p_t satisfying the following estimate: there exists $C \in [1,\infty)$ such that

$$C^{-1}\min\left(t^{-n/\alpha}, \frac{t}{|x-y|^{n+\alpha}}\right) \leqslant p_t(x,y) \leqslant C\min\left(t^{-n/\alpha}, \frac{t}{|x-y|^{n+\alpha}}\right), \quad \text{for all } t>0, x,y \in \mathbb{R}^n. \tag{1.3}$$

This formula suggest the space time scaling relation distance$^\alpha$ scales like time.

A rich family of heat kernels arise from diffusions on fractals. Barlow and Perkins constructed a diffusion process on the standard Sierpiński gasket [BP]. The law of this diffusion process at any time $t > 0$ admits a density with respect to the d_f-Hausdorff measure, where $d_f = \log_2 3$ is the Hausdorff dimension of the Sierpiński gasket. Let $d_w = \log_2 5$. The transition density p_t admits the following estimates: there exists $C_1, C_2, C_3, C_4 \in (0,\infty)$ such that for any x, y in the Sierpiński gasket and any $t \in (0,1)$, we have

$$\frac{C_3}{t^{d_f/d_w}}\exp\left(-C_4\left(\frac{|x-y|^{d_w}}{t}\right)^{1/(d_w-1)}\right) \leqslant p_t(x,y) \leqslant \frac{C_1}{t^{d_f/d_w}}\exp\left(-C_2\left(\frac{|x-y|^{d_w}}{t}\right)^{1/(d_w-1)}\right) \tag{1.4}$$

Bounds of the form (1.4) are called *sub-Gaussian heat kernel estimates* is now known to hold on many fractals.

Here is the outline for this course. We will begin by covering fundamental aspects of the theory of Dirichlet forms, a powerful framework for constructing and analyzing symmetric Markov processes. We will then survey key results characterizing spaces that satisfy heat kernel estimates analogous to (1.2), (1.3), and (1.4). We will illustrate these results by applying them to a number of examples. One important class of examples is the boundary trace process. Consider the reflected Brownian motion on the $n+1$-dimensional upper half space $\mathbb{R}^n \times [0,\infty)$. The boundary trace process is obtained by removing the path of the reflection Brownian motion in the interior $\mathbb{R}^n \times (0,\infty)$ in a certain sense. A classical result of Spitzer states that the resulting process is a jump process on the boundary $\partial\mathbb{H}^{n+1} = \mathbb{R}^n$ and coincides with the symmetric 1-stable process (Cauchy process). By modifying the reflected Brownian motion to a different reflected diffusion process, Molchanov and Ostrovskii [MO] discovered that any α-stable process with $\alpha \in (0,2)$ can be obtained as trace of a diffusion process on the upper half space. This was rediscovered in an analytic setting by Caffarelli and Silvestre [CS]. We will study the behavior of boundary trace process for more general diffusions and domains.

2 Dirichlet forms and symmetric Markov processes

We refer to the standard references [FOT, CF, BH91] for a comprehensive introduction to the theory of Dirichlet forms.

2.1 Semigroup and resolvent

Definition 2.1 (semigroup of operators)**.** Let $\mathcal{H}$ be a Hilbert space over $\mathbb{R}$ equipped with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. We say that a family of linear operators $\{T_t : \mathcal{H} \to \mathcal{H} | t > 0\}$ is a *semigroup* if it satisfies

(1) Each T_t is a symmetric operator; that is, $\langle T_t(f), g\rangle = \langle f, T_t(g)\rangle$ for all $t > 0$ and $f, g \in \mathcal{H}$.

(2) $\{T_t : t > 0\}$ satisfy the semigroup property; $T_{t+s} = T_t T_s$ for all $t, s > 0$.

(3) Each T_t is a contraction operator, $\|T_t(u)\| \leqslant \|u\|$ for all $t > 0, u \in \mathcal{H}$.

(4) (strongly continuous) For all $u \in \mathcal{H}$, we have $\lim_{t \downarrow 0} \|T_t(u) - u\| = 0$.

It is more precise to call it a symmetric, strongly continuous, contraction semigroup. We will use the abbreviated term semigroup.

A Markov process $(Y_t)_{t \geqslant 0}$ on a space X defines an operator

$$P_t f(x) = \mathbb{E}_x[f(Y_t)] = \mathbb{E}[f(Y_t)|Y_0 = x],$$

for all $t > 0, x \in X$ and a suitable class of functions f. By the Markov property, for any $t, s > 0, x \in X$ and $f : X \to \mathbb{R}$ bounded, we have the semigroup property

$$P_{t+s} f(x) = \mathbb{E}_x[f(Y_{t+s})] = \mathbb{E}_x\left[\mathbb{E}_x[f(Y_{t+s})|\mathcal{F}_t]\right] = \mathbb{E}_x\left[\mathbb{E}_{Y_t}[f(Y_s)]\right] = \mathbb{E}_x[(P_s f)(Y_t)] = P_t(P_s(f))(x).$$

This estimate implies that the expected distance traveled by the diffusion process (B_t) satsifes the bound $\mathbb{E}[|B_t - B_0|] \asymp t^{1/d_w}$ for all $t \in (0,1)$. This should be compared with the expected distance traveled by the Brownian motion on $\mathbb{R}^n$ which is comparable to $\sqrt{t}$.

Example 2.2. (i) The Brownian (or heat) semigroup is given by $P_t : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is

$$P_t f(x) := \int_{\mathbb{R}^n} \frac{1}{(2\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{2t}\right) f(y)\, dy, \quad \text{for all } t > 0, f \in L^2(\mathbb{R}^n).$$

This corresponds to the standard Brownian motion on $\mathbb{R}^n$.

(ii) Let γ_n denote the standard Gaussian measure on $\mathbb{R}^n$; $\gamma_n(dx) = (2\pi)^{-n/2} e^{-|x|^2/2}\, dx$. The Ornstein-Uhlenbeck process $(X_t)_{t \geqslant 0}$ on $\mathbb{R}^n$ is a Markov process that is associated with the stochastic differential equation

$$dX_t = -X_t\, dt + \sqrt{2} dB_t,$$

where (B_t) is the standard Brownian motion on $\mathbb{R}^n$. Alternately,

$$X_t = e^{-t}X_0 + e^{-t}B_{e^{2t}-1},$$

where (B_t) is again the standard Brownian motion. Then the Ornstein-Uhlenbeck semigroup is defined by $P_t : L^2(\mathbb{R}^n, \gamma_n) \to L^2(\mathbb{R}^n, \gamma_n)$ as

$$P_t f(x) := \int_{\mathbb{R}^n} f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) \gamma_n(dy), \quad \text{for all } f \in L^2(\mathbb{R}^n, \gamma_n), t > 0.$$

(iii) Let $X = \{1, \ldots, n\}$ and $c : X \times X \to [0, \infty)$ be such that $c(x,y) = c(y,x) \geqslant 0$ for all $x, y \in X$. Let $m : X \to [0, \infty)$ be

$$m(x) := \sum_{z \in X} c(x,y).$$

Let us assume that $m(x) > 0$ for all $x \in X$. This defines a *discrete time* Markov chain $(Z_n)_{n \in \mathbb{N} \cup \{0\}}$ with transition probabilities given by

$$P(x,y) := \frac{c(x,y)}{m(x)}, \quad \text{for all } x, y \in X.$$

This defines a linear operator $Q : L^2(E,m) \to L^2(E,m)$ given by

$$Qf(x) = \mathbb{E}[f(Z_1)|Z_0 = x] = \sum_{y \in X} P(x,y)f(y).$$

By the Markov property, we have $Q^k f(x) = \mathbb{E}[f(Z_k)|Z_0 = x]$ for all $k \in \mathbb{N} \cup \{0\}$ with $Q^0 = I$, where Q^k denotes the k-fold composition of Q. A standard construction of a continuous time process (Y_t) from the discrete time process $(Z_n)_{n \in \mathbb{N} \cup \{0\}}$ is obtained by waiting at every state $x \in E$ for an exponential time Exp(1) before jumping to a state y with probability $Q(x,y)$. If $(N(t))_{t \geqslant 0}$ denotes a Poisson process with rate 1 (independent of (Z_n)) that determines the waiting times, we have that the semigroup corresponding to process $Y_t := Z_{N(t)}$ is given by

$$P_t f(x) = \mathbb{E}[f(Y_t)|Y_0 = x] = \sum_{k=0}^{\infty} \mathbb{P}(N(t) = k)\mathbb{E}[f(Z_k)|Z_0 = x] = \sum_{k=0}^{\infty} e^{-t}\frac{t^k}{k!}Q^k f(x).$$

Then $(P_t)_{t \geqslant 0}$ is a semigroup on $L^2(X, m)$.

Exercise 2.3. Verify that the examples above are a semigroup (in the sense of Definition 2.1).
Hint: In order to show strong continuity, it might help to look at Lemma 2.35.

As we will see, it is often easier to construct or analyze the Laplace transform of a semigroup.

Definition 2.4 (resolvent)**.** A resolvent on $\mathcal{H}$ is a family of linear operators $\{G_\alpha : \mathcal{H} \to \mathcal{H} | \alpha > 0\}$ on $\mathcal{H}$ such that

(1) Each G_α is a symmetric operator; ; that is, $\langle G_\alpha(f), g\rangle = \langle f, G_\alpha(g)\rangle$ for all $t > 0$ and $f, g \in \mathcal{H}$.

(2) $\{G_\alpha : \alpha > 0\}$ satisfy the resolvent equation

$$G_\alpha - G_\beta + (\alpha - \beta)G_\alpha G_\beta = 0, \quad \text{for all } \alpha, \beta > 0. \tag{2.1}$$

(3) (contraction property) For any $u \in \mathcal{H}, \alpha > 0$, we have $\|\alpha G_\alpha u\| \leqslant \|u\|$.

(4) (strongly continuous) For any $u \in \mathcal{H}$, we have $\lim_{\alpha\to\infty} \|\alpha G_\alpha(u) - u\| = 0$.

The resolvent can be viewed as the Laplace transform of the semigroup $t \mapsto T_t$. To describe this we need to define integrals of Hilbert space valued function.

Definition 2.5. Let $I \subset \mathbb{R}$ be an interval and $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. We say that a function $f : I \to \mathcal{H}$ is *weakly measurable* if for any $v \in \mathcal{H}$, the functions $t \mapsto \langle v, f(t)\rangle$ and $t \mapsto \|f(t)\|$ are measurable functions on I. If a weakly measurable function $f : I \to \mathcal{H}$ satisfies $\int_I \|f(t)\|\, dt < \infty$, we say that f is *integrable.* If $f : I \to \mathcal{H}$ is weakly measurable and integrable, then by the Riesz-Fréchet representation theorem, there exists a unique $x \in \mathcal{H}$ such that

$$\langle v, x\rangle = \int_I \langle v, f(t)\rangle\, dt, \quad \text{for all } v \in \mathcal{H}. \tag{2.2}$$

We denote x as the integral $\int_I f(t)\, dt \in \mathcal{H}$.

The following basic properties of this integral are easily verified.

Exercise 2.6. Let $I \subset \mathbb{R}$ be an interval and $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space.

1. If $f : I \to \mathcal{H}$ is continuous, then f is weakly measurable.

2. If $f : I \to \mathcal{H}$ is weakly measurable and integrable, then we have the triangle inequality

$$\left\| \int_I f(t)\, dt \right\| \leqslant \int_I \|f(t)\|\, dt.$$

3. Let $\mathcal{H}_1$ be a Hilbert space and let $T : \mathcal{H} \to \mathcal{H}_1$ be a bounded linear map. If $f : I \to \mathcal{H}$ is continuous and integrable, then $T \circ f : I \to \mathcal{H}_1$ is weakly measurable and integrable. Furthermore, we have

$$T\left(\int_I f(t)\, dt\right) = \int_I (T \circ f)(t)\, dt.$$

Every semigroup defines a resolvent as its Laplace transform.

Exercise 2.7. Consider a strongly continuous contraction semigroup $\{T_t : t > 0\}$ on $\mathcal{H}$. For any $\alpha > 0, u \in \mathcal{H}$ consider the integral

$$G_\alpha(u) = \int_0^\infty e^{-\alpha t} T_t(u)\, dt.$$

Show that $\{G_\alpha : \mathcal{H} \to \mathcal{H} | \alpha > 0\}$ defines a resolvent (in the sense of Definition 2.4).
Hint: The standard approach to showing the resolvent identity is to use

$$e^{-\alpha t} - e^{-\beta t} = (\beta - \alpha) \int_0^t e^{-\alpha(t-s)} e^{-\beta s}\, ds$$

in the integral for $G_\alpha(u) - G_\beta(u)$ and interchange the order of integration and use semigroup property.

Here is another probabilistic approach to show the resolvent identity[1]: Assume $\alpha > \beta$ and consider two independent random variables ξ_β and $\xi_{\alpha-\beta}$ with exponential distributions with rate α. Then here is an outline of the probabilistic approach.

1. $\xi_\alpha := \xi_\beta \wedge \xi_{\alpha-\beta}$ is exponentially distributed with rate α.
2. $\mathbb{E}[T_{\xi_\alpha}(u)] = \alpha G_\alpha(u), \mathbb{E}[T_{\xi_\beta}(u)] = \beta G_\beta(u)$.
3. $\mathbb{P}(\xi_\alpha = \xi_\beta) = 1 - \mathbb{P}(\xi_\alpha < \xi_\beta) = \frac{\beta}{\alpha}$.
4. Conditioned on the event $\{\xi_\alpha < \xi_\beta\}$, the distribution on $\xi_\beta - \xi_\alpha$ is exponential with parameter β (memoryless property).
5. Write

$$T_{\xi_\beta} = \mathbb{1}_{\{\xi_\alpha = \xi_\beta\}} T_{\xi_\alpha} + \mathbb{1}_{\{\xi_\alpha < \xi_\beta\}} T_{\xi_\alpha} \circ T_{\xi_\beta - \xi_\alpha} = \mathbb{1}_{\{\xi_\alpha = \xi_\beta\}} T_{\xi_\alpha} + \mathbb{1}_{\{\xi_\alpha < \xi_\beta\}} T_{\xi_\beta - \xi_\alpha} \circ T_{\xi_\alpha},$$

   and take expectations on both sides to derive the resolvent identity.

Another way to rephrase the probabilistic approach is as follows. Let ξ_α and ξ_β be independent exponential random variables with rates α and β with $\alpha > \beta$ and let Z be an independent (of ξ_α, ξ_β) Bernoulli random variable with $\mathbb{P}(Z = 1) = 1 - \mathbb{P}(Z = 0) = \frac{\beta}{\alpha}$. Then verify that the random variable

$$\widetilde{\xi}_\beta := \mathbb{1}_{\{Z=1\}} \xi_\alpha + \mathbb{1}_{\{Z=0\}} (\xi_\alpha + \xi_\beta)$$

is also exponentially distributed with parameter β. Writing

$$T_{\widetilde{\xi}_\beta} = \mathbb{1}_{\{Z=1\}} T_{\xi_\alpha} + \mathbb{1}_{\{Z=0\}} T_{\xi_\alpha} \circ T_{\xi_\beta} = \mathbb{1}_{\{Z=1\}} T_{\xi_\alpha} + \mathbb{1}_{\{Z=0\}} T_{\xi_\beta} \circ T_{\xi_\alpha}$$

and taking expectations as above yields the resolvent identity.

[1]Thanks to Ryoichiro Noda for sharing this argument due to David Croydon.

2.2 Generators and background on self-adjoint operators

The generator provides an *infinitesimal* description of the semigroup. Let $A : \mathcal{D}(A) \to \mathcal{H}$ be a denseley defined operator on a Hilbert space $\mathcal{H}$; that is, $\mathcal{D}(A)$ is a dense subspace of $\mathcal{H}$.

Definition 2.8 (Generator of a semigroup)**.** Let $\{T_t : \mathcal{H} \to \mathcal{H} | t > 0\}$ be a semigroup. Then the *generator* A of the semigroup $\{T_t : t > 0\}$ is the operator $A : \mathcal{D}(A) \to \mathcal{H}$ is defined as

$$A(u) := \lim_{t \downarrow 0} \frac{T_t(u) - u}{t}, \quad \text{for all } u \in \mathcal{D}(A),$$

where the domain $\mathcal{D}(A)$ is

$$\mathcal{D}(A) = \{u \in \mathcal{H} : \lim_{t \downarrow 0} \frac{T_t(u) - u}{t} \text{ exists}\}.$$

Often the generator is easier to compute than the semigroup as we illustrate.

Example 2.9. We will slightly generalize Example 2.2-(iii). Let $X = \{1, \ldots, n\}$ and $c : X \times X \to [0, \infty)$ be such that $c(x, y) = c(y, x) \geqslant 0$ for all $x, y \in X$. Let $m : X \to [0, \infty)$ be

$$m(x) := \sum_{y \in X} c(x, y).$$

Assume that $m(x) > 0$ for all $x \in X$. This defines a *discrete time* Markov chain $(Z_n)_{n \in \mathbb{N} \cup \{0\}}$ with transition probabilities given by

$$P(x, y) := \frac{c(x, y)}{m(x)}, \quad \text{for all } x, y \in X.$$

Let $\lambda : X \to (0, \infty)$ be a function determining a continuous time process as follows: at each state $x \in X$, the process waits (independently of other waiting times and transitions) an exponential time with parameter $\lambda(x)$[2] before jumping to a new state y with probability $P(x, y)$. Call this process $(Y_t)_{t \geqslant 0}$ on X. If $\lambda \equiv 1$, the number of transitions (jumps) is a Poisson process which helped in the computation of semigroup. Since λ need note be a constant function, the computation of semigroup corresponding semigroup $P_t : \mathbb{R}^X \to \mathbb{R}^X$. Nevertheless, we can compute the generator by observing that

$$P_t f(x) = \mathbb{E}_x[f(Y_t)] = e^{-\lambda(x)t} f(x) + \sum_{y \in X} (1 - e^{-\lambda(x)t}) P(x, y) f(y) + O(t^2),$$

for all $f \in \mathbb{R}^X, t \in (0, 1), x \in X$. This implies

$$Lf(x) = \lim_{t \downarrow 0} \frac{P_t f(x) - f(x)}{t} = -\lambda(x) f(x) + \lambda(x) \sum_{y \in X} P(x, y) f(y) = \lambda(x)(I - Q) f(x).$$

[2] exponential random variable with paramet λ has mean λ^{-1}, so larger λ means smaller wait times

Since P is m-symmetric, it is easy to see that L is $\widetilde{m}$-symmetric with $\widetilde{m}(x) = \lambda(x)^{-1}m(x)$ as

$$\langle Lf, g\rangle_{L^2(\widetilde{m})} = \langle (I-Q)f, g\rangle_{L^2(m)} = \langle f, (I-Q)g\rangle_{L^2(m)} = \langle f, Lg\rangle_{L^2(\widetilde{m})}, \quad \text{for all } f, g \in L^2(X, \widetilde{m}).$$

Using the symmetry of L in $L^2(X, \widetilde{m})$, we can conclude that (P_t) is also symmetric in $L^2(X, \widetilde{m})$.

If the discrete time random walk on X is irreducible, then using limiting behavior of discrete time Markov chains, one can interpret $\widetilde{m}$ (normalized to be a probability measure) as the asymptotic occupation time of a Markov chain; that is

$$\lim_{t\uparrow\infty} \frac{1}{t}\int_0^t \mathbb{1}_{\{Y_s=y\}}\, ds = \frac{\widetilde{m}(y)}{\sum_{z\in X}\widetilde{m}(z)}, \quad \mathbb{P}_x\text{-a.s. for all } x, y \in X.$$

Using the contraction property of the semigroup, we see that the generator is a non-positive definite operator.

Lemma 2.10. *$\mathcal{D}(A)$ defined above a subspace of $\mathcal{H}$, $A : \mathcal{D}(A) \to \mathcal{H}$ is a linear map, and A is non-positive definite; that is,*

$$\langle A(u), u\rangle \leqslant 0, \quad \textit{for all } u \in \mathcal{D}(A).$$

Proof. The first two claims are easy consequence of the definition of A. If $u \in \mathcal{D}(A)$, then

$$\langle A(u), u\rangle = \lim_{t\downarrow 0}\frac{1}{t}\langle T_t(u) - u, u\rangle.$$

By Schwarz's inequality, for any $u \in \mathcal{H}, t > 0$, we have $\langle T_t(u) - u, u\rangle = \langle T_t(u), u\rangle - \|u\|^2 \leqslant \|T_t(u)\|\,\|u\| - \|u\|^2 \leqslant 0$ (since T is a contraction). □

Lemma 2.11. *Let $\{G_\alpha : \alpha > 0\}$ be a resolvent on $\mathcal{H}$. Then for each $\alpha > 0$, G_α is injective and non-negative definite.*

Proof. Let $\beta > 0$. Since G_β is linear, it suffices to show that G_β has a trivial kernel. Let $u \in \mathcal{H}$ be such that $G_\beta(u) = 0$. By the resolvent equation, for any $\beta > 0$ we have

$$G_\alpha(u) = G_\beta(u) + (\beta - \alpha)G_\alpha(G_\beta(u)) = 0.$$

Hence by strong continuity $u = \lim_{\alpha\to\infty}\alpha G_\alpha(u) = 0$. Hence G_β is injective for each $\beta > 0$.

To show that G_α is non-negative definite, we need to verify that $\langle u, G_\alpha u\rangle \geqslant 0$ for all $u \in \mathcal{H}$. To this end, fix $u \in \mathcal{H}$ and define $f(\alpha) = \langle u, G_\alpha u\rangle$. By the contraction property of the semigroup

$$\|G_\alpha(u) - G_\beta(u)\| \leqslant |\alpha - \beta|\alpha^{-1}\beta^{-1}\|u\|.$$

Therefore, $\lim_{\beta\to\alpha} G_\beta(u) = G_\alpha(u)$. For $\alpha \neq \beta$, by the resolvent equation and symmetry of the resolvent operators, we have

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} = -\langle u, G_\alpha(G_\beta(u))\rangle = -\langle G_\alpha(u), G_\beta(u)\rangle.$$

Letting $\beta \to \alpha$ and using $\lim_{\beta\to\alpha} G_\beta(u) = G_\alpha(u)$, we obtain that $\alpha \mapsto f(\alpha)$ is differentiable and

$$f'(\alpha) = -\left\|G_\alpha(u)\right\|^2 \leqslant 0.$$

By the strong continuity, we have $\lim_{\alpha\to\infty} \alpha f(\alpha) = \lim_{\alpha\to\infty}\langle u, \alpha G_\alpha(u)\rangle = \|u\|^2 < \infty$. Hence $\lim_{\alpha\to\infty} f(\alpha) = 0$. Since f is non-increasing, we conclude $f(\alpha) = \langle u, G_\alpha u\rangle \geqslant 0$ for all $\alpha \in (0, \infty)$. □

By Exercise 2.7, every semigroup defines a resolvent as a Laplace transform. So one might wonder if we can compute the generator of semigroup directly from the resolvent. This motivates the following definition.

Definition 2.12 (Generator of a resolvent)**.** Let $\{G_\alpha : \alpha > 0\}$ be a resolvent on $\mathcal{H}$. The *generator* of $\{G_\alpha : \alpha > 0\}$ is defined as the linear operator $A : \mathcal{D}(A) \to \mathcal{H}$ defined by

$$\mathcal{D}(A) = G_\alpha(\mathcal{H}), \quad A(u) = \alpha u - G_\alpha^{-1}(u), \quad \text{for any } \alpha > 0, u \in \mathcal{D}(A),$$

where $G_\alpha^{-1} : G_\alpha(\mathcal{H}) \to \mathcal{H}$ is the inverse of G_α (recall from Lemma 2.11 that G_α is injective).

The fact that the above operator is well-defined is an easy exercise in the use of resolvent equation that we state below.

Exercise 2.13. Let $\{G_\alpha : \alpha > 0\}$ be a resolvent on $\mathcal{H}$.

(i) For any $\alpha, \beta > 0$, show that $G_\alpha(\mathcal{H}) = G_\beta(\mathcal{H})$.

(ii) For any $\alpha, \beta > 0$ and any $u \in G_\alpha(\mathcal{H}) = G_\beta(\mathcal{H})$, show that $\alpha u - G_\alpha^{-1}(u) = \beta u - G_\beta^{-1}(v)$.

Hint: Use the resolvent equation (2.1).

Let $\mathcal{H}$ be a Hilbert space. Let $T : \mathcal{D}(T) \to \mathcal{H}$ be a densely defined operator ($\overline{\mathcal{D}(T)} = \mathcal{H}$). Then define

$$\mathcal{D}(T^*) = \{x \in \mathcal{H} | y \mapsto \langle T(y), x\rangle \text{ is a bounded operator on } \mathcal{D}(T)\}.$$

If $x \in \mathcal{H}$, then by Hahn-Banach theorem, there is a unique extension of the map $y \mapsto \langle x, A(y)\rangle$ to $\mathcal{H}$. Hence by Riesz-Fréchet representation theorem, there exists (a unique) $T^*(z) \in \mathcal{H}$ such that

$$\langle T(y), x\rangle = \langle y, T^*(x)\rangle, \quad \text{for all } x \in \mathcal{D}(T^*).$$

It is easy to verify that $T^* : \mathcal{D}(T^*) \to \mathcal{H}$ is a linear map. For an operator $T : \mathcal{D}(T) \to \mathcal{H}$, let $\mathcal{G}(T) = \{(u, T(u)) : u \in \mathcal{D}(T)\} \subset \mathcal{H} \times \mathcal{H}$ denote the graph of T. We view $\mathcal{H} \times \mathcal{H}$ is a Hilbert space with inner product $\langle (x_1, y_1), (x_2, y_2)\rangle_{\mathcal{H}\times\mathcal{H}} = \langle x_1, x_2\rangle + \langle y_1, y_2\rangle$ for all $(x_1, y_1), (x_2, y_2) \in \mathcal{H} \times \mathcal{H}$.

Definition 2.14. (i) We say an operator $T : \mathcal{D}(T) \to \mathcal{H}$ is *closed* if $\mathcal{G}(T)$ is a closed set (in $\mathcal{H} \times \mathcal{H}$).

(ii) We say a densely defined operator $T : \mathcal{D}(T) \to \mathcal{H}$ is *symmetric* if for all $x, y \in \mathcal{D}(T)$, we have
$$\langle T(x), y\rangle = \langle x, T(y)\rangle.$$
Equivalently, T^* is an extension of T ($\mathcal{D}(T^*) \supset \mathcal{D}(T)$ and $T^*(x) = T(x)$ for all $x \in \mathcal{D}(T)$).

(iii) We say that an operator $T : \mathcal{D}(T) \to \mathcal{H}$ is *self-adjoint* if $T = T^*$.

For a subspace M of a Hilbert space $\mathcal{H}$, by $M^\perp$ we denote the *orthogonal complement* defined by
$$M^\perp = \{u \in \mathcal{H} : \langle u, m\rangle = 0 \quad \text{for all } m \in M\}.$$
The following lemma expresses the graph of the adjoint T^* in terms of the graph of T.

Lemma 2.15. *Let $T : \mathcal{D}(T) \to \mathcal{H}$ be a densely defined operator and let $\mathcal{G}(T) = \{(u, T(u)) : u \in \mathcal{H}\} \subset \mathcal{H} \times \mathcal{H}$ denote the graph of T. Let $V : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ be defined by $V(x, y) = (y, -x)$ for all $(x, y) \in \mathcal{H} \times \mathcal{H}$. Then $\mathcal{G}(T^*) = V(\mathcal{G}(T))^\perp$.*

Proof. $(x, y) \in \mathcal{G}(T^*)$ if and only if
$$\langle T(u), x\rangle = \langle u, y\rangle, \quad \text{for all } u \in \mathcal{D}(T).$$
This can be rewritten as
$$\langle V(u, T(u)), (x, y)\rangle_{\mathcal{H}\times\mathcal{H}} = \langle (T(u), -u), (x, y)\rangle_{\mathcal{H}\times\mathcal{H}} = \langle T(u), x\rangle - \langle u, y\rangle = 0, \quad \text{for all } u \in \mathcal{D}(T).$$
□

Clearly, every self-adjoint operator is symmetric and closed (by Lemma 2.15).

Corollary 2.16. *Let $T : \mathcal{D}(T) \to \mathcal{H}$ be an injective self-adjoint operator with range $\mathcal{R}(T)$. Then $T^{-1} : \mathcal{R}(T) \to \mathcal{H}$ is a densely defined, self-adjoint operator.*

Proof. Let us show that $\mathcal{R}(T)$ is dense. This is equivalent to showing that $\mathcal{R}(T)^\perp = \{0\}$. Let $y \in \mathcal{R}(T)^\perp$ as $\overline{\mathcal{R}(T)} = (\mathcal{R}(T)^\perp)^\perp$. Then $(y, 0) \in V(\mathcal{G}(T))^\perp = \mathcal{G}(T^*) = \mathcal{G}(T)$. Since T is injective, $y = 0$.

Let $V, S : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ be defined by $V(x, y) = (y, -x)$ and $S(x, y) = (y, x)$ for all $(x, y) \in \mathcal{H} \times \mathcal{H}$. Then by Lemma 2.15 and the observation that $V \circ S = -S \circ V$, we have
$$\mathcal{G}((T^{-1})^*) = V\left(\mathcal{G}(T^{-1})\right)^\perp = V(S(\mathcal{G}(T)))^\perp = S(V(\mathcal{G}(T)))^\perp.$$
Since $S : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ is a unitary (inner-product preserving) bijection with $S^2 = \mathrm{Id}$, we have $S(V(\mathcal{G}(T)))^\perp = S(V(\mathcal{G}(T))^\perp)$ and hence by Lemma 2.15, we have
$$\mathcal{G}((T^{-1})^*) = S(V(\mathcal{G}(T))^\perp) = S(\mathcal{G}(T^*)) = S(\mathcal{G}(T)) = \mathcal{G}(T^{-1}).$$
□

Exercise 2.17. Let (Ω, μ) be a measure space and let $\mathcal{H} = L^2(\Omega, \mu)$. Let $\lambda : \Omega \to \mathbb{R}$ be a measurable function (not necessarily bounded). Show that the multiplication operator $M_\lambda : \mathcal{D}(M_\lambda) \to \mathcal{H}$ with domain

$$\mathcal{D}(M_\lambda) := \{f \in L^2(\Omega, \mu) : \lambda f \in L^2(\Omega, \mu)\}$$

defined by $M_\lambda(f) = f(\cdot)\lambda(\cdot)$ is a densely defined, self-adjoint operator.

Theorem 2.18 (Spectral theorem)**.** *Let A be a self-adjoint operator on a Hilbert space $\mathcal{H}$. Then there is a measure space (Ω, μ) and a unitary map $U : L^2(\Omega, \mu) \to \mathcal{H}$ and a measurable function $\lambda : \Omega \to \mathbb{R}$ such that*

$$U^{-1}AU = M_\lambda, \quad \mathcal{D}(A) = \{U(f) : f \in \mathcal{D}(M_\lambda)\},$$

where $M_\lambda : \mathcal{D}(M_\lambda) \to L^2(\Omega, \mu)$ is as defined in Exercise 2.17.

Furthermore, if A is non negative definite (respectively, non positive definite); that is, $\langle A(u), u\rangle \geqslant 0$ (respectively, $\langle A(u), u\rangle \leqslant 0$) for all $u \in \mathcal{D}(A)$, then $\lambda(x) \geqslant 0$ (respectively, $\lambda(x) \leqslant 0$) for μ-almost every $x \in \Omega$.

Definition 2.19 (Functional calculus)**.** Let A be a non-positive self-adjoint operator on a Hilbert space $\mathcal{H}$ and let $U : L^2(\Omega, \mu) \to \mathcal{H}$ and $\lambda : \Omega \to \mathbb{R}$ be as given in Theorem 2.18. Then for any Borel function $f : (-\infty, 0] \to \mathbb{R}$, we define $f(A) : \mathcal{D}(f(A)) \to \mathcal{H}$ denote the operator defined by

$$U^{-1}f(A)U = M_{f(\lambda(\cdot))}, \quad \mathcal{D}(f(A)) = \{U(g) : g \in \mathcal{D}(M_{f(\lambda(\cdot))})\}$$

It is easy to verify that $f(A)$ is a self-adjoint operator on $\mathcal{H}$. Note that if f is bounded, then $f(A)$ is a bounded operator with $\mathcal{D}(f(A)) = \mathcal{H}$. If f is non-negative (respectively, non-positive), then $f(A)$ is a non-negative definite (respectively, non-positive definite) operator.

Proposition 2.20. *Let $\{T_t : t > 0\}$ be a a strongly continuous contraction semigroup $\{T_t : t > 0\}$ on $\mathcal{H}$ and let us denote the associated resolvent by $\{G_\alpha : \alpha > 0\}$ defined as*

$$G_\alpha(u) := \int_0^\infty e^{-\alpha t} T_t(u)\, dt, \quad \text{for all } u \in \mathcal{H}, \alpha > 0. \tag{2.3}$$

Let A_s amd A_r denote the generators of the semigroup $\{T_t : t > 0\}$ and the resolvent $\{G_\alpha : \alpha > 0\}$ respectively. Then $A_s = A_r$ and is a non-positive definite self-adjoint operator.

Proof. Let $u \in \mathcal{D}(A_r)$. Then $u = G_\alpha(v)$ for some $v \in \mathcal{H}$ and

$$\begin{aligned}\frac{1}{t}\left(e^{-\alpha t}T_t(u) - u\right) &= \frac{1}{t}\left(e^{-\alpha t}T_t\left(\int_0^\infty e^{-\alpha s}T_s(u)\, ds\right) - \int_0^\infty e^{-\alpha s}T_s(u)\, ds\right) \\ &= \frac{1}{t}\left(\left(\int_t^\infty e^{-\alpha s}T_s(u)\, ds\right) - \int_0^\infty e^{-\alpha s}T_s(u)\, ds\right)\end{aligned}$$

$$= -\frac{1}{t}\left(\int_0^t e^{-\alpha s} T_s(v)\, ds\right) \xrightarrow{t\to 0} -v.$$

This implies $u \in \mathcal{D}(A_s)$ since

$$A_s(u) = \lim_{t\downarrow 0} \frac{T_t(u) - u}{t} = \lim_{t\downarrow 0} \frac{e^{-\alpha t}T_t(u) - u}{t} + \lim_{t\downarrow 0} \frac{T_t(u)(1 - e^{-\alpha t})}{t} \tag{2.4}$$

$$= -v + \alpha u = \alpha u - G_\alpha^{-1}(u) = A_r(u). \tag{2.5}$$

This show that A_r is an restriction of A_s.

It remains to show that $\mathcal{D}(A_s) \subset \mathcal{D}(A_r)$. To this end, let $u \in \mathcal{D}(A_s)$. Then by the calculation in (2.4)

$$v := -\lim_{t\downarrow 0} \frac{1}{t}\left(e^{-\alpha t}T_t(u) - u\right) = -A_s(u) + \alpha u.$$

Define $w := u - G_\alpha v$. We have

$$\lim_{t\downarrow 0} \frac{1}{t}\left(e^{-\alpha t}T_t(w) - w\right) = \lim_{t\downarrow 0} \frac{1}{t}\left(e^{-\alpha t}T_t(u) - u\right) - \lim_{t\downarrow 0} \frac{1}{t}\left(e^{-\alpha t}T_t(G_\alpha(v)) - G_\alpha(v)\right)$$

$$= -v + \lim_{t\downarrow 0} \frac{1}{t}\left(-\int_t^\infty e^{-\alpha s}T_s(v)\, ds + \int_0^\infty e^{-\alpha s}T_s(v)\, ds\right) = -v + v = 0.$$

Therefore $w \in \mathcal{D}(A_s)$ and

$$0 = \lim_{t\downarrow 0} \frac{1}{t}\left(e^{-\alpha t}T_t(w) - w\right) = A_s(w) - \alpha w.$$

Therefore

$$0 = \alpha\langle w, w\rangle - \langle A_s(w), w\rangle \geqslant \alpha\langle w, w\rangle,$$

and hence $w = 0$, or equivalently, $u = G_\alpha(v) \in \mathcal{D}(A_r)$.

By Corollary 2.16 G_α^{-1} is a self-adjoint operator and hence $A_r = \alpha - G_\alpha^{-1}$ is self-adjoint. By Lemma 2.10, $A_r = A_s$ is a non-positive definite operator. □

We now state a converse to Proposition 2.20. The resolvent and semigroup corresponding to the generator can be defined using functional calculus (Definition 2.19).

Proposition 2.21. *Let $\mathcal{H}$ be a Hilbert space and let $A : \mathcal{D}(A) \to \mathcal{H}$ be a non-positive definite, self-adjoint operator.*

(a) Then $\{T_t = \exp(tA) | t > 0\}$ and $\{G_\alpha = (\alpha - A)^{-1} | \alpha > 0\}$ are a semigroup and resolvent respectively. Furthermore, $\{G_\alpha : \alpha > 0\}$ is the resolvent corresponding to the semigroup $\{T_t = \exp(tA) | t > 0\}$.

(b) The generators of the semigroup $\{T_t > 0 | t > 0\}$ and the resolvent $\{G_\alpha | \alpha > 0\}$ in (a) coincide with A. Furthermore, there is a unique semigroup and a unique resolvent whose generators are A.

Proof. (a) Let $U : L^2(\Omega,\mu) \to \mathcal{H}$ be an unitary operator and $\lambda : \Omega \to (-\infty, 0]$ be as given in the spectral theorem (Theorem 2.18). Then we have

$$U^{-1}T_tU = M_{\exp(t\lambda(\cdot))}, \quad U^{-1}G_\alpha U = M_{(\alpha-\lambda(\cdot))^{-1}}, \quad \text{for all } t > 0, \alpha > 0.$$

The symmetry of T_t follows from symmetry of $M_{\exp(t\lambda)}$. For all $f, g \in \mathcal{H}, t > 0$ we have

$$\begin{aligned}\langle T_t(f), g\rangle &= \int_\Omega M_{\exp(t\lambda)}(U^{-1}(f))U^{-1}(g)\,d\mu = \int_\Omega \exp(t\lambda)U^{-1}(f)U^{-1}(g)\,d\mu \\ &= \int_\Omega U^{-1}(f)M_{\exp(t\lambda)}(U^{-1}(g))\,d\mu = \langle f, T_t(g)\rangle.\end{aligned}$$

The proof of the semigroup property is similar as for all $t, s > 0$

$$T_tT_s = UM_{\exp(t\lambda)}U^{-1}UM_{\exp(s\lambda)}U^{-1} = UM_{\exp(t\lambda)}M_{\exp(s\lambda)}U^{-1} = UM_{\exp((t+s)\lambda)}U^{-1} = T_{t+s}.$$

To show strong continuity, note that for any $f \in \mathcal{H}$ and using the fact that $\lambda \leqslant 0$ m-a.e., we have $|(\exp(t\lambda) - 1)U^{-1}(f)| \leqslant U^{-1}(f)$ m-a.e. for all $t > 0$. Hence by dominated convergence theorem, we have

$$\lim_{t\downarrow 0} \|T_t(f) - f\|^2 = \lim_{t\downarrow 0} \int_\Omega \left|(\exp(t\lambda) - 1)U^{-1}(f)\right|^2 d\mu = 0, \quad \text{for all } f \in \mathcal{H}.$$

The proof that $\{G_\alpha = UM_{(\alpha-\lambda(\cdot))^{-1}}U^{-1}|\alpha > 0\}$ defines a resolvent is similar and left as an exercise.

To verify that $\{G_\alpha : \alpha > 0\}$ is the resolvent corresponding to the semigroup $\{T_t = \exp(tA)|t > 0\}$, for any $f \in \mathcal{H}$, we have

$$\begin{aligned}\int_0^\infty e^{-\alpha t}T_t(f)\,dt &= \int_0^\infty U\left(e^{-\alpha t}M_{\exp(t\lambda)}(U^{-1}(f))\right)dt = U\left(\int_0^\infty e^{-\alpha t}M_{\exp(t\lambda)}(U^{-1}(f))\,dt\right) \\ &= U\left(\int_0^\infty \exp(-\alpha t + \lambda t)U^{-1}(f)\,dt\right) = U\left(M_{(\alpha-\lambda(\cdot))^{-1}}(U^{-1}(f))\right) = G_\alpha(f).\end{aligned}$$

(b) For any $f \in L^2(\Omega,\mu)$, we have the *pointwise* limit

$$\lim_{t\downarrow 0} \frac{M_{\exp(t\lambda(\cdot))}(f) - f}{t} = \lim_{t\downarrow 0} \frac{(\exp(t\lambda(\cdot) - 1)(f(\cdot)}{t} = M_\lambda(f),$$

and that since $\lambda \leqslant 0$ m-a.e., we have

$$\left|\frac{\exp(t\lambda(\cdot)) - 1}{t}\right| = \frac{1 - \exp(t\lambda(\cdot))}{t} \uparrow \lambda(\cdot), m\text{-a.e. as } t \downarrow 0.$$

So by the dominated convergence theorem,

$$\lim_{t\downarrow 0} \frac{M_{\exp(t\lambda(\cdot))}(f) - f}{t} = \lim_{t\downarrow 0} \frac{(\exp(t\lambda(\cdot) - 1)(f(\cdot)}{t} = M_\lambda(f),$$

in $L^2(\Omega, \mu)$ if and only if $f \in \mathcal{D}(M_\lambda) = \{f \in L^2(\Omega, \mu) : \lambda f \in L^2(\Omega, \mu)\}$. This implies that the generator of the semigroup $\{T_t = UM_{\exp(t\lambda(\cdot))}U^{-1} : t > 0\}$ is $A = UM_\lambda U^{-1}$. Since the generators of a semigroup and the corresponding resolvent coincide (by Proposition 2.20), we have that A is the generator of the resolvent $\{G_\alpha : \alpha > 0\}$.

By the injectivity of Laplace transform, it suffices to show that the resolvent is uniquely determined by the generator. Suppose $\{G_\alpha : \alpha > 0\}$ and $\{G'_\alpha : \alpha > 0\}$ are two resolvent with generator A, for any $f \in \mathcal{H}, \alpha > 0$, $w := G_\alpha(f) - G'_\alpha(f)$ satisfies

$$(\alpha - A)(w) = (\alpha - A)(G_\alpha(f) - G'_\alpha(f)) = f - f = 0.$$

Since $-A$ is non-negative definite, we have

$$0 = \langle (\alpha - A)(w), w \rangle \geqslant \alpha \langle w, w \rangle \geqslant 0,$$

and hence $w = 0$. Since $f \in \mathcal{H}$ is arbitrary, $G_\alpha = G'_\alpha$. □

Exercise 2.22. Show that $\{G_\alpha = UM_{(\alpha - \lambda(\cdot))^{-1}}U^{-1} | \alpha > 0\}$ defined in the proof above is a resolvent.

In the next exercise, we outline a direct proof of the uniqueness of the semigroup with a given generator (without relying on the injectivity of Laplace transform and the corresponding uniqueness result for the resolvent).

Exercise 2.23. Let $\mathcal{H}$ be a Hilbert space and let $A : \mathcal{D}(A) \to \mathcal{H}$ be a non-positive definite, self-adjoint operator. Let $\{T_t : t > 0\}$ be a semigroup with generator A.

(i) Show that for any $x \in \mathcal{D}(A), t > 0$, we have $T_t(x) \in \mathcal{D}(A)$ and

$$\frac{d}{dt}(T_t(x)) = A(T_t(x)) = T_t(A(x)).$$

(ii) Using (i), show that if $\{T_t : t > 0\}$ and $\{\widetilde{T}_t : t > 0\}$ are two semigroups with generator A, then for any $x \in \mathcal{D}(A)$ the function $E_x : (0, \infty) \to [0, \infty)$ defined by

$$E_x(t) = \left\| T_t(x) - \widetilde{T}_t(x) \right\|^2, \quad \text{for all } t > 0$$

satisfies

$$\frac{d}{dt} E_x(t) = 2\langle T_t(x) - \widetilde{T}_t(x), A(T_t(x) - \widetilde{T}_t(x)) \rangle \leqslant 0, \quad \text{for all } t > 0$$

and $\lim_{t \downarrow 0} E_x(t) = 0$.

(iii) Conclude from (ii) that $T_t(x) = \widetilde{T}_t(x)$ for all $x \in \mathcal{H}$ and $t > 0$.

The following exercise can be viewed as an inverse Laplace transform formula for the semigroup corresponding to a resolvent.

Exercise 2.24. Let $\{G_\alpha : \alpha > 0\}$ be a resolvent on $\mathcal{H}$. Show that for all $t > 0, f \in \mathcal{H}$

$$T_t(f) = \lim_{\alpha\to\infty} e^{-t\alpha} \sum_{n=0}^{\infty} \frac{(t\alpha)^n}{n!}(\alpha G_\alpha)^n(f) \tag{2.6}$$

defines a semigroup and that $\{G_\alpha : \alpha > 0\}$ is the resolvent corresponding to the semigroup $\{T_t : t > 0\}$. Hint: Use the proof of Proposition 2.21 to express $G_\alpha = UM_{(\alpha-\lambda(\cdot))^{-1}}U^{-1}$ and show that $T_t = UM_{\exp(t\lambda(\cdot))}U^{-1}$.

2.3 Closed quadratic forms

Definition 2.25. Let $\mathcal{H}$ be a Hilbert space with the inner product $\langle\cdot,\cdot\rangle$. A *quadratic form* $\mathcal{E} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ is a dense subspace of $\mathcal{F}$ (called the domain of the quadratic form) of $\mathcal{H}$ such that it satsfies the following properties:

(i) (bi-linearity) For all $a_1, a_2 \in \mathbb{R}$ and $f_1, f_2, g \in \mathcal{F}$, we have $\mathcal{E}(a_1 f_1 + a_2 f_2, g) = a_1\mathcal{E}(f_1, g) + a_2\mathcal{E}(f_2, g)$.

(ii) (symmetry) $\mathcal{E}(f, g) = \mathcal{E}(g, f)$ for all $f, g \in \mathcal{H}$.

(iii) (non-negative definite) $\mathcal{E}(f, f) \geqslant 0$ for all $f \in \mathcal{F}$.

We say that a quadratic form is said to *closed* if $\mathcal{F}$ is a Hilbert space equipped with the inner product $\mathcal{E}_1 : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$

$$\mathcal{E}_1(f, g) = \mathcal{E}(f, g) + \langle f, g\rangle, \quad \text{for all } f, g \in \mathcal{F}.$$

Familiar properties of inner product such as Schwarz inequality and triangle inequality hold for (non-negative definite) quadratic forms with identical proofs.

Exercise 2.26. Let $\mathcal{E} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ be a quadratic form on a Hilbert space $\mathcal{H}$. Show that for all $f, g \in \mathcal{F}$, we have the Schwarz inequality

$$|\mathcal{E}(u, v)| \leqslant \mathcal{E}(u, u)^{1/2}\mathcal{E}(v, v)^{1/2}$$

and the triangle inequality

$$\mathcal{E}(f + g, f + g)^{1/2} \leqslant \mathcal{E}(f, f)^{1/2} + \mathcal{E}(g, g)^{1/2}.$$

There is a one-to-one correspondence between closed quadratic forms and non-positive definite self-adjoint operators A.

Theorem 2.27. *Let A be a non-positive definite, self-adjoint operator on $\mathcal{H}$. Then $(\mathcal{E}, \mathcal{F})$ is a closed quadratic form, where*

$$\mathcal{E}(f, g) = \langle\sqrt{-A}(f), \sqrt{-A}(g)\rangle, \quad f, g \in \mathcal{F} := \mathcal{D}(\sqrt{-A}). \tag{2.7}$$

Conversely, any closed quadratic form on $\mathcal{H}$ arises from a non-positive definite, self-adjoint operator on $\mathcal{H}$ as given in (2.7).

Proof. Since $(\mathcal{F}, \mathcal{E}_1)$ is an inner product space, we need to verify completeness. Note that $\sqrt{-A}$ is a non-negative definite, self-adjoint operator. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $(\mathcal{F}, \mathcal{E}_1)$. This implies that $(f_n, \sqrt{-A}(f_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H} \times \mathcal{H}$ that belongs to the graph $\mathcal{G}(\sqrt{-A})$. Since the graph of every self-adjoint operator is closed (by Lemma 2.15), we conclude that there exists $f \in \mathcal{F}$ such that $\lim_{n\to\infty} \mathcal{E}_1(f - f_n, f - f_n) = 0$. Hence $(\mathcal{F}, \mathcal{E}_1)$ is Hilbert space.

Define $\mathcal{E}_\alpha(f, g) := \mathcal{E}(f, g) + \alpha\langle f, g\rangle$ for all $f, g \in \mathcal{F}$. Let $\{G_\alpha = (\alpha - A)^{-1} : \alpha > 0\}$ denote the resolvent generated by A. Using the spectral theorem, it is easy to verify that

$$G_\alpha(\mathcal{H}) \subset \mathcal{F}, \quad \mathcal{E}_\alpha(G_\alpha(f), g) = \langle f, g\rangle, \quad \text{for all } f \in \mathcal{H}, g \in \mathcal{F}. \tag{2.8}$$

Conversely, let $(\mathcal{E}, \mathcal{F})$ be a closed quadratic form. Since $(\mathcal{F}, \mathcal{E}_\alpha)$ is a Hilbert space for each $\alpha > 0$ and for any $u \in \mathcal{H}$, the function $v \mapsto \langle u, v\rangle$ where $v \in \mathcal{F}$ is a bounded linear functional in the Hilbert space $(\mathcal{F}, \mathcal{E}_\alpha)$. Hence by Riesz-Fréchet representation theorem, there exists (a unique) $G_\alpha(u)$ in $\mathcal{F}$ such that

$$\mathcal{E}_\alpha(G_\alpha(u), v) = \langle u, v\rangle, \quad \text{for all } u \in \mathcal{H}, v \in \mathcal{F} \text{ and } \alpha > 0. \tag{2.9}$$

We claim that $\{G_\alpha : \alpha > 0\}$ is a resolvent. For the symmetry of G_α note that for each $u, v \in \mathcal{H}$, we have

$$\langle G_\alpha(u), v\rangle \overset{(2.9)}{=} \mathcal{E}_\alpha(G_\alpha(u), G_\alpha(v)) \overset{(2.9)}{=} \langle u, G_\alpha(v)\rangle.$$

To obtain the resolvent equation, note that for any $u \in \mathcal{H}, v \in \mathcal{F}$, we have

$$\begin{aligned}\mathcal{E}_\alpha(G_\beta(u) - (\alpha - \beta)G_\alpha G_\beta(u), v) &= \mathcal{E}_\alpha(G_\beta(u), v) - (\alpha - \beta)\mathcal{E}_\alpha(G_\alpha(G_\beta(u)), v) \\ &\overset{(2.9)}{=} \mathcal{E}_\alpha(G_\beta(u), v) - (\alpha - \beta)\langle G_\beta(u), v\rangle = \mathcal{E}_\beta(G_\beta(u), v) \\ &\overset{(2.9)}{=} \langle u, v\rangle \overset{(2.9)}{=} \mathcal{E}_\alpha(G_\alpha(u), v).\end{aligned}$$

Since $(\mathcal{F}, \mathcal{E}_\alpha)$ is a Hilbert space, we obtain the resolvent equation

$$G_\alpha(u) = G_\beta(u) - (\alpha - \beta)G_\alpha G_\beta(u), \quad \text{for all } u \in \mathcal{H}.$$

For each $\alpha > 0$, αG_α is a contraction since

$$\|(\alpha G_\alpha)(u)\| \, \|G_\alpha(u)\| = \alpha\langle G_\alpha(u), G_\alpha(u)\rangle \leqslant \mathcal{E}_\alpha(G_\alpha(u), G_\alpha(u)) = \langle u, G_\alpha(u)\rangle \leqslant \|u\| \, \|G_\alpha(u)\|,$$

for all $u \in \mathcal{H}$. For strong continuity, we use the contraction property to obtain

$$\begin{aligned}\alpha \|\alpha G_\alpha(u) - u\|^2 &\leqslant \mathcal{E}_\alpha(\alpha G_\alpha(u) - u, \alpha G_\alpha(u) - u) \overset{(2.9)}{=} \alpha^2\langle G_\alpha(u), u\rangle + \mathcal{E}(u, u) - \alpha\langle u, u\rangle \\ &\leqslant \alpha \|\alpha G_\alpha(u)\| \, \|u\| - \alpha \|u\|^2 + \mathcal{E}(u, u) \leqslant \mathcal{E}(u, u), \quad \text{for all } u \in \mathcal{F}.\end{aligned}$$

Hence

$$\lim_{\alpha\to\infty} \|\alpha G_\alpha(u) - u\| = 0, \quad \text{for all } u \in \mathcal{F}.$$

Since $\mathcal{F}$ is dense in $\mathcal{H}$, for any $u \in \mathcal{H}$, there exists a sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ such that $\lim_{n\to\infty} \|u_n - u\| = 0$. Now since αG_α is a contraction, we have

$$\|\alpha G_\alpha(u) - u\| \leqslant \|(\alpha G_\alpha)(u-u_n)\| + \|\alpha G_\alpha(u_n) - u_n\| + \|u_n - u\| \leqslant 2\|u-u_n\| + \|\alpha G_\alpha(u_n) - u_n\|.$$

Let $\alpha \to \infty$ and then $n \to \infty$ to obtain

$$\limsup_{\alpha\to\infty} \|\alpha G_\alpha(u) - u\| \leqslant 2\|u - u_n\| \xrightarrow{n\to\infty} 0.$$

This concludes the proof of strong continuity and hence $\{G_\alpha : \alpha > 0\}$ is a resolvent.

Let A be the generator of the resolvent $(\mathcal{E}', \mathcal{F}')$ closed quadratic form corresponding to the non-positive self-adjoint operator A as defined in (2.7). By (2.8), we have $G_\alpha(\mathcal{H}) \subset \mathcal{F}'$ and hence

$$\mathcal{E}'(G_\alpha(u), G_\alpha(v)) \overset{(2.8)}{=} \langle G_\alpha(u), v\rangle \overset{(2.9)}{=} \mathcal{E}_\alpha(G_\alpha(u), G_\alpha(v)), \quad \text{for all } u, v \in \mathcal{H}.$$

Hence $\mathcal{E}'$ and $\mathcal{E}$ coincide on $G_\alpha(\mathcal{H}) \times G_\alpha(\mathcal{H})$. By (2.9), $G_\alpha(\mathcal{H})$ is dense in the Hilbert space $(\mathcal{F}, \mathcal{E}_\alpha)$. Similarly, by (2.8), $G_\alpha(\mathcal{H})$ is dense in the Hilbert space $(\mathcal{F}', \mathcal{E}'_\alpha)$. Since both $(\mathcal{E}, \mathcal{F})$ and $(\mathcal{E}', \mathcal{F}')$ are closed quadratic forms that coincide on a dense set, they are equal. Hence the correspondence given in (2.7) is a bijection between non-positive definite, self-adjoint operators and closed quadratic forms. □

The quadratic form corresponding to the generator of a semigroup and resolvent can be described directly as outlined below.

Exercise 2.28. Let A be a non-positive definite, self-adjoint operator on $\mathcal{H}$ that is the generator of a semigroup $\{P_t = \exp(tA) : t > 0\}$ and resolvent $\{G_\alpha = (\alpha - A)^{-1} : \alpha > 0\}$. Let $(\mathcal{E}, \mathcal{F})$ denote the closed quadratic form corresponding to A as given by Theorem 2.27. Using the spectral theorem, show the following

(a) For any $f \in \mathcal{H}$, the function

$$t \mapsto \frac{1}{t}\langle (I - P_t)f, f\rangle$$

is non-increasing and non-negative function on $(0, \infty)$. Furthermore the quadratic form $(\mathcal{E}, \mathcal{F})$ is given by

$$\mathcal{F} = \left\{ f \in \mathcal{H} \,|\, \lim_{t\downarrow 0} \frac{1}{t}\langle (I-P_t)f, f\rangle < \infty \right\},$$

and

$$\mathcal{E}(f,f) = \lim_{t\downarrow 0} \frac{1}{t}\langle (I-P_t)f, f\rangle, \quad \text{for all } f \in \mathcal{F}.$$

(b) For any $f \in \mathcal{H}$, the function

$$\alpha \mapsto \alpha\langle (I - \alpha G_\alpha)f, f\rangle$$

is non-decreasing and non-negative function on $(0, \infty)$. Furthermore the quadratic form $(\mathcal{E}, \mathcal{F})$ is given by

$$\mathcal{F} = \left\{ f \in \mathcal{H} | \lim_{\alpha \uparrow \infty} \alpha \langle (I - \alpha G_\alpha) f, f \rangle < \infty \right\},$$

and

$$\mathcal{E}(f, f) = \lim_{\alpha \uparrow \infty} \alpha \langle (I - \alpha G_\alpha) f, f \rangle, \quad \text{for all } f \in \mathcal{F}.$$

To summarize, there is a one-to-one correspondence between

(a) Closed non-negative definite quadratic forms (Definition 2.25)

(b) Non-positive definite self-adjoint operators (Definition 2.14)

(c) Semigroups (Definition 2.1)

(d) Resolvents (Definition 2.4).

2.4 Beurling-Deny criterion and Markov operators

We introduce the definition of Dirichlet form.

Definition 2.29. Let $(X, \mathcal{M}, m)$ be a σ-finite measure space. A Dirichlet form on $L^2(X, m)$ is a *quadratic form* (in the sense of Definition 2.25; that is, bi-linear, symmetric, non-negative definite, closed, densely defined) $\mathcal{E} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ on $L^2(X, m)$ such that it satisfies the *Markov property*: for all $u \in \mathcal{F}$, we have $\tilde{u} := (0 \vee u) \wedge 1 \in \mathcal{F}$ and $\mathcal{E}(\tilde{u}, \tilde{u}) \leqslant \mathcal{E}(u, u)$.

At this point the terminology Markov property might seem strange. As we will later justify (see Theorem 2.34), the Markov property of quadratic form is equivalent to the Markovian property of the corresponding semigroup (or resolvent).

Definition 2.30. We say that a bounded linear map $T : L^2(X, m) \to L^2(X, m)$ is Markovian if for any $f \in L^2(X, m)$ such that $0 \leqslant f \leqslant 1$ m-almost everywhere, we have $0 \leqslant T(f) \leqslant 1$ m-a.e.

Exercise 2.31. Let $T : L^2(X, m) \to L^2(X, m)$ be a bounded linear operator. Then show that the following are equivalent.

(a) For any $f \in L^2(X, m)$ such that $0 \leqslant f \leqslant 1$ m-almost everywhere, we have $0 \leqslant T(f) \leqslant 1$ m-a.e.

(b) For any $f \in L^2(X, m)$ such that $f \leqslant 1$ m-almost everywhere, we have $T(f) \leqslant 1$ m-a.e.

The following exercise gives a description of all Dirichlet forms on a finite set. This special case of the *Beurling-Deny decomposition* of regular Dirichlet forms admits an elementary proof as outlined below. We use $\mathbb{1}_A$ to denote the indicator of a set A.

Exercise 2.32. Let $X = \{1, \ldots, m\}$, where $n \in \mathbb{N}$ and let m be a measure on X with $m(\{i\}) > 0$ for all $i \in X$. Let $\mathcal{E} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ be a Dirichlet form on $L^2(X, m)$.

(i) Prove that $\mathcal{F} = L^2(X, m)$.

(ii) Show that for any $i, j \in X$ with $i \neq j$, we have

$$\mathcal{E}(\mathbb{1}_{\{i\}}, \mathbb{1}_{\{j\}}) \leqslant 0.$$

Hint: Consider $f = \mathbb{1}_{\{i\}} - \epsilon \mathbb{1}_{\{j\}}$ for $\epsilon \geqslant 0$.

(iii) Show that for any $i \in X$, we have $\mathcal{E}(\mathbb{1}_{\{i\}}, \mathbb{1}_X) \geqslant 0$. **Hint**: Consider $f = \mathbb{1}_X + \epsilon \mathbb{1}_{\{i\}}$ for $\epsilon \geqslant 0$.

(iv) Using (ii) and (iii), show that there exists $c_{ij} \geqslant 0$, for all $1 \leqslant i < j \leqslant n$ and $k(i) \geqslant 0$ for all $i \in X$ such that

$$\mathcal{E}(f, f) = \sum_{1 \leqslant i < j \leqslant n} c_{ij}(f(i) - f(j))^2 + \sum_{i=1}^{n} k(i) f(i)^2, \quad \text{for all } f \in \mathcal{F} = \mathbb{R}^X.$$

Lemma 2.33. *Let $P : L^2(X, m) \to L^2(X, m)$ be a bounded, linear, m-symmetric, Markovian operator and let $F : \mathbb{R} \to \mathbb{R}$ be a 1-Lipschitz function (that is, $|F(a) - F(b)| \leqslant |a - b|$ for all $a, b \in \mathbb{R}$) such that $F(0) = 0$.*

(i) Then for any $n \in \mathbb{N}$, for all pairwise disjoint sets $A_1, \ldots, A_n$ such that $m(A_i) < \infty$ for all $i = 1, \ldots, n$, and for all $a_1, \ldots, a_n \in \mathbb{R}$, writing $f = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$, we have

$$\langle (I - P) f, f \rangle_{L^2(X,m)} = \sum_{i=1}^{n} \mu_i a_i^2 + \frac{1}{2} \sum_{1 \leqslant i < j \leqslant n} \alpha_{i,j} (a_i - a_j)^2, \tag{2.10}$$

where

$$\alpha_{i,j} := \langle \mathbb{1}_{A_i}, P \mathbb{1}_{A_j} \rangle_{L^2(X,m)} \geqslant 0, \quad \mu_i = m(A_i) - \sum_{k=1}^{n} \alpha_{i,k} \geqslant 0, \quad \text{for all } 1 \leqslant i, j \leqslant n. \tag{2.11}$$

(ii) For any $g \in L^2(X, m)$, then $\widetilde{g} := F(g)$ satisfies

$$\langle (I - P) \widetilde{g}, \widetilde{g} \rangle_{L^2(X,m)} \leqslant \langle (I - P) g, g \rangle_{L^2(X,m)}. \tag{2.12}$$

Proof. (i) Note that by the m-symmetry of P, we have $\alpha_{i,j} = \langle \mathbb{1}_{A_i}, P \mathbb{1}_{A_j} \rangle_{L^2(X,m)} = \langle P \mathbb{1}_{A_i}, \mathbb{1}_{A_j} \rangle_{L^2(X,m)} = \alpha_{j,i}$ for all $1 \leqslant i, j \leqslant n$. By linearity of P and the symmetry $\alpha_{i,j} = \alpha_{j,i}$, we have

$$\langle (I - P) f, f \rangle_{L^2(X,m)} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \langle (I - P) \mathbb{1}_{A_i}, \mathbb{1}_{A_j} \rangle_{L^2(X,m)}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} a_i^2 m(A_i) - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \alpha_{i,j} \\
&= \sum_{i=1}^{n} \mu_i a_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} \left(a_i^2 \alpha_{i,j} - a_i a_j \alpha_{i,j}\right) \\
&= \sum_{i=1}^{n} \mu_i a_i^2 + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \left(a_i^2 \alpha_{i,j} + a_j^2 \alpha_{i,j} - 2 a_i a_j \alpha_{i,j}\right) \\
&= \sum_{i=1}^{n} \mu_i a_i^2 + \frac{1}{2}\sum_{1 \leqslant i < j \leqslant n} \alpha_{i,j}(a_i - a_j)^2.
\end{aligned}
$$

Since $0 \leqslant P\mathbb{1}_{A_j} \leqslant 1$ m-a.e. by the Markovian property of P, we have $\alpha_{i,j} = \langle \mathbb{1}_{A_i}, P\mathbb{1}_{A_j}\rangle_{L^2(X,m)} \geqslant 0$. Since $A_1, \ldots, A_n$ are pairwise disjoint, by the Markovian property of P, we have

$$
\sum_{k=1}^{n} P\left(\mathbb{1}_{A_k}\right) = P\left(\sum_{k=1}^{n} \mathbb{1}_{A_k}\right) \leqslant 1, \quad m\text{-a.e.}
$$

Hence

$$
\sum_{k=1}^{n} \alpha_{i,k} = \langle \mathbb{1}_{A_i}, \sum_{k=1}^{n} P\left(\mathbb{1}_{A_k}\right)\rangle \leqslant \int_X \mathbb{1}_{A_i}\, dm = m(A_i), \quad \text{for all } 1 \leqslant i \leqslant n,
$$

or equivalently $\mu_i \geqslant 0$ for all $1 \leqslant i \leqslant n$.

(ii) If g is a simple function as given in (i), then the desired estimate (2.12) follows from (2.10) and (2.11). In general, for any $g \in L^2(X,m)$, there exists a sequence of simple function $(g_n)_{n\in\mathbb{N}}$ such that $g_n \xrightarrow{n\to\infty} g$ pointwise, $|g_n| \uparrow |g|$ pointwise, and $g_n \xrightarrow{n\to\infty} g$ in $L^2(X,m)$. Since $|F(a)| \leqslant |a|$ for all $a \in \mathbb{R}$, by the dominated convergence theorem $\widetilde{g_n} := F(g_n) \xrightarrow{n\to\infty} \widetilde{g} := F(g)$ in $L^2(X,m)$. Hence by the continuity of P, $P(\widetilde{g_n}) \xrightarrow{n\to\infty} P\widetilde{g}$ in $L^2(X,m)$ hence by using (2.12) for simple functions (since $\widetilde{g_n}$ are also simple functions) we obtain

$$
\begin{aligned}
\langle (I-P)\widetilde{g}, \widetilde{g}\rangle_{L^2(X,m)} &= \lim_{n\to\infty} \langle (I-P)\widetilde{g_n}, \widetilde{g_n}\rangle_{L^2(X,m)} \\
&\overset{(2.10),(2.11)}{\leqslant} \lim_{n\to\infty} \langle (I-P)g_n, g_n\rangle_{L^2(X,m)} = \lim_{n\to\infty} \langle (I-P)g, g\rangle_{L^2(X,m)}.
\end{aligned}
$$

□

Let $(X, \mathcal{B}, m)$ be a σ-finite measure space.

Theorem 2.34. *Let $\mathcal{E} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ be a closed quadratic form on $L^2(X,m)$. Let $\{T_t : t > 0\}$, $\{G_\alpha : \alpha > 0\}$ and $A : \mathcal{D}(A) \to L^2(X,m)$ denote the associated semigroup, resolvent and generator associated with $\mathcal{E}$ respectively. Then the following are equivalent:*

(a) T_t is a Markovian operator for each $t > 0$.

(b) αG_α is a Markovian operator for each $t > 0$.

(c) For all $f \in \mathcal{D}(A)$, we have $\langle A(f), (f-1)^+\rangle \leqslant 0$.

(d) For all $u \in \mathcal{F}$, we have $\widetilde{u} := (0 \vee u) \wedge 1 \in \mathcal{F}$ and $\mathcal{E}(\widetilde{u}, \widetilde{u}) \leqslant \mathcal{E}(u,u)$.

Proof. The implications (a) $\Longrightarrow$ (b) and (b) $\Longrightarrow$ (a) follow easily from (2.3) and (2.6).

Next, let us show that (d) $\Longrightarrow$ (b). Let $u \in \mathcal{F}$ such that $0 \leqslant u \leqslant 1$ m-a.e. and $\alpha > 0$. We consider the function $\Psi_u : \mathcal{F} \to \mathbb{R}$ defined by

$$\Psi(w) := \mathcal{E}(w,w) + \alpha\langle w - \alpha^{-1}u, w - \alpha^{-1}u\rangle.$$

For all $w \in \mathcal{F}$, we have

$$\begin{aligned}\Psi(w) - \Psi(G_\alpha(u)) &= \mathcal{E}_\alpha(w,w) - 2\langle w,u\rangle + \alpha^{-1}\langle u,u\rangle \\ &\quad - \mathcal{E}_\alpha(G_\alpha(u), G_\alpha(u)) + 2\langle G_\alpha(u), u\rangle - \alpha^{-1}\langle u,u\rangle \\ &= \mathcal{E}_\alpha(w,w) - 2\langle w,u\rangle + \mathcal{E}_\alpha(G_\alpha(u), G_\alpha(u)) \\ &= \mathcal{E}_\alpha(w,w) - 2\mathcal{E}_\alpha(G_\alpha(u), w) + \mathcal{E}_\alpha(G_\alpha(u), G_\alpha(u)) \\ &= \mathcal{E}_\alpha(G_\alpha(u) - w, G_\alpha(u) - w),\end{aligned}$$

and hence

$$\Psi(w) \leqslant \Psi(G_\alpha(u)) \quad \text{if and only if} \quad w = G_\alpha(u). \tag{2.13}$$

Now, define $v = \alpha^{-1}\left((0 \vee (\alpha G_\alpha(u))) \wedge 1\right) = (0 \vee G_\alpha(u)) \wedge \alpha^{-1}$. So by (d), we have $v \in \mathcal{F}$ and

$$\mathcal{E}(v,v) \leqslant \alpha^{-2}\mathcal{E}(\alpha G_\alpha(u), \alpha G_\alpha(u)) = \mathcal{E}(G_\alpha(u), G_\alpha(u)). \tag{2.14}$$

Since $u(x) \in [0,1]$ for m-a.e., we have for m-a.e. $x \in X$,

$$\left|v(x) - \alpha^{-1}u(x)\right| = \left|\left((0 \vee G_\alpha(u)(x)) \wedge \alpha^{-1}\right) - \alpha^{-1}u(x)\right| \leqslant \left|G_\alpha(u)(x) - \alpha^{-1}u(x)\right|. \tag{2.15}$$

By (2.14), (2.15), we have $\Psi(v) \leqslant \Psi(G_\alpha(u))$ and hence by (2.15), we conclude that $v = (0 \vee G_\alpha(u)) \wedge \alpha^{-1} = G_\alpha(u)$. Thus αG_α is a Markovian operator.

The implications (a) $\Longrightarrow$ (d) and (b) $\Longrightarrow$ (d) follows from Lemma 2.33 (by setting $F(t) = (0 \vee t) \wedge 1$ in Lemma 2.33-(ii)) and Exercise 2.28.

Next, let us show (b) $\Longrightarrow$ (c). By using Lemma 2.33 (by setting $F(t) = t \wedge 1$) and Exercise 2.28(b), we obtain that for all $f \in \mathcal{F}$, $f \wedge 1 \in \mathcal{F}$ and $\mathcal{E}(f \wedge 1, f \wedge 1) \leqslant \mathcal{E}(f,f)$. In particular for $f \in \mathcal{D}(A)$, we have $f \wedge 1 = f - (f-1)_+$ and hence $\mathcal{E}(f - (f-1)_+, f - (f-1)_+) \leqslant \mathcal{E}(f,f)$ or equivalently,

$$2\langle A(f), (f-1)_+\rangle = -2\mathcal{E}(f, (f-1)_+) \leqslant -\mathcal{E}((f-1)_+, (f-1)_+) \leqslant 0.$$

It remains to show (c) $\Longrightarrow$ (b). By Exercise 2.31, it suffices to show that if $f \in L^2(X,m)$ satisfies $f \leqslant 1$ m-a.e., then $\alpha G_\alpha(f) \leqslant 1$ for all $\alpha > 0$. Fix $\alpha > 0$ and f as above and set $g := \alpha G_\alpha(f)$ so that $g \in \mathcal{D}(A)$ and

$$\alpha g - A(g) = \alpha f. \tag{2.16}$$

By (2.16) and (c) $\langle A(g), (g-1)^+\rangle \leqslant 0$, we obtain $\langle g, (g-1)_+\rangle \leqslant \langle f, (g-1)_+\rangle$, or equivalently,

$$\int_{\{g\geqslant 1\}} (g-f)(g-1)_+\rangle \leqslant 0. = \int_X (g-f)(g-1)_+\rangle \geqslant 0.$$

Since $f \leqslant 1$ m-a.e., the above inequality implies that $g = \alpha G_\alpha(f) \leqslant 1$ m-a.e. □

In order to show the existence of heat kernel, it is useful to consider the Markovian semigroup as an operator on $L^p(X,m)$ for all $1 \leqslant p \leqslant \infty$. By the Markovian property, for any $f \in L^2(X,m) \cap L^\infty(X,m)$, we have $P_t(f) \in L^2(X,m) \cap L^\infty(X,m)$ with $\|P_t(f)\|_\infty \leqslant \|f\|_\infty$. Since $(X, \mathcal{M}, m)$ is a σ-finite space, there exists an increasing sequence of measurable sets $(A_n)_{n\in\mathbb{N}}$ with $m(A_n) < \infty$ such that $X = \bigcup_{n\in\mathbb{N}} A_n$. So for any $f \in L^\infty(X,m)$ with $f \geqslant 0$ m-a.e., we have $f\mathbb{1}_{A_n} \in L^2(X,m) \cap L^\infty(X,m)$ for all $n \in \mathbb{N}$. Hence by the Markovian property of P_t, we have

$$\|P_t(f\mathbb{1}_{A_n})\|_\infty \leqslant \|f\|_\infty, \quad P_t(f\mathbb{1}_{A_n}) \leqslant P_t(f\mathbb{1}_{A_{n+1}}) \quad m\text{-a.e., for all } n \in \mathbb{N}.$$

Hence $\lim_{n\to\infty} P_t(f\mathbb{1}_{A_n})(x)$ exists for m-a.e. $x \in X$. We set the pointwise limit as

$$P_t f := \lim_{n\to\infty} P_t(f\mathbb{1}_{A_n}), \quad \text{for all } f \in L^\infty(X,m) \text{ with } f \geqslant 0 \ m\text{-a.e.},$$

and $P_t(f) = P_t(f_+) - P_t(f_-)$ for all $f \in L^\infty(X,m)$. This defines an operator $P_t : L^\infty(X,m) \to L^\infty(X,m)$ as a contraction operator; that is $\|P_t(f)\|_\infty \leqslant \|f\|_\infty$ for all $f \in L^\infty(X,m)$. By the dominated convergence theorem and the symmetry of P_t in $L^2(X,m)$, for any $t > 0$, $g \in L^1(X,m) \cap L^\infty(X,m)$ and for any $f \in L^\infty(X,m)$, we have

$$\int_X gP_t(f)\,dm = \int_X fP_t(g)\,dm.$$

Hence $\|P_t(g)\|_1 = \sup_{f\in L^\infty(X,m), \|f\|_\infty=1} \int_X fP_t(g)\,dm \leqslant \|g\|_1$ for all $g \in L^1(X,m) \cap L^\infty(X,m)$. Since $L^1(X,m) \cap L^\infty(X,m)$ is dense in $L^1(X,m)$, by continuous extension we obtain a linear contraction $P_t : L^1(X,m) \to L^1(X,m)$. For any $p \in [1,\infty)$, by Hölder inequality we have

$$\|P_t(f)\|_p \leqslant \|f\|_p, \quad \text{for all } f \in L^1(X,m) \cap L^\infty(X,m). \tag{2.17}$$

Since $L^1(X,m) \cap L^\infty(X,m)$ is dense in $L^p(X,m)$, we obtain a linear contraction $P_t : L^p(X,m) \to L^p(X,m)$ for all $1 \leqslant p < \infty$.

The following lemma provides a convenient sufficient condition for the strong continuity property of a semigroup.

Lemma 2.35. *Let $\mathcal{H} = L^2(X,m)$ be a Hilbert space over $\mathbb{R}$, where m is a measure on X. Consider a family of linear operators $\{T_t : L^2(X,m) \to L^2(X,m) | t > 0\}$ such that each T_t is a Markovian, contraction operator. Furthermore if $\mathcal{L} \subset L^1(X,m) \cap L^\infty(X,m)$ is a dense subspace of $L^2(X,m)$ such that for any $f \in \mathcal{L}$, we have*

$$\lim_{t\downarrow 0} T_t(f)(x) = f(x) \quad \text{for } m\text{-a.e. } x \in X. \tag{2.18}$$

Then for any $u \in L^2(X,m)$, we have $\lim_{t\downarrow 0} \|T_t(u) - u\|_2 = 0$.

Proof. Since T_t is a contraction, we have

$$\|T_t(f) - f\|_2^2 = \|T_t(f)\|_2^2 + \|f\|_2^2 - 2\langle f, T_t(f)\rangle \leqslant 2\,\|f\|_2^2 - 2\langle f, T_t(f)\rangle, \quad \text{for all } f \in L^2(X,m). \tag{2.19}$$

If $f \in \mathcal{L}$, by (2.18) and the dominated convergence theorem (dominating function is $\|f\|_\infty f \in L^1(X,m)$ by the Markovian property of T_t; by Exercise 2.31 we have $\|T_t(f)\|_\infty \leqslant \|f\|_\infty$) we have

$$\lim_{t\downarrow 0}\langle f, T_t(f)\rangle = \|f\|_2^2. \tag{2.20}$$

Hence by (2.19) and (2.20) we have $\lim_{t\downarrow 0} \|T_t(f) - f\|_2 = 0$ for all $f \in \mathcal{L}$. By the density of $\mathcal{L}$ in $L^2(X,m)$ and the contraction property of T_t, we have $\lim_{t\downarrow 0} \|T_t(f) - f\|_2 = 0$ for all $f \in L^2(X,m)$ □

2.5 Regular Dirichlet forms and Fukushima's theorem

Definition 2.36. Let $(X, \mathcal{M}, m)$ be a σ-finite metric measure space, and let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(X,m)$. We say the Dirichlet form $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(X,m)$ is *regular*, if it satisfies the following properties:

(a) X is a locally compact separable metrizable topological space X, with $\mathcal{M}$ the associated Borel σ-field, a Radon measure m on X with full support (a Borel measure m on X which is finite on any compact subset of X and strictly positive on any non-empty open subset of X).

(b) The vector space $\mathcal{F} \cap C_c(X)$ is dense both in $(\mathcal{F}, \mathcal{E}_1)$ and in $(C_c(X), \|\cdot\|_\infty)$.

A fundamental theorem of Fukushima the assumption of the regularity of the Dirichlet form allows us to construct a m-symmetric Markov process on X whose semigroup coincides with the semigroup of the Dirichlet form. Assume for the moment that the semigroup corresponding to a Dirichlet form is defined over a space of pointwise well-defined functions. Then the Markov property along with Riesz–Markov–Kakutani representation theorem would then imply the existence of a sub-probability measure on X which can be made into a transition probability on $X \cup \Delta$, where Δ is an absorbing cemetery state. Fukushima observed that the assumption of regularity ensures that every function in the domain of the form can be modified to continuous outside a small set (quasi-continuous). This allows us to overcome the difficulty of functions in L^2 being not pointwise well-defined.

Theorem 2.37 (Fukushima's theorem)**.** *[FOT, Theorem 7.2.1] Let $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form on $L^2(X,m)$. Let $\mathcal{B}$ denote the Borel σ-field on X. Let $X \cup \{\Delta\}$ denote the one-point compactification of X and let $\mathcal{B}_\Delta = \mathcal{B} \cup \{B \cup \Delta | B \in \mathcal{B}\}$. For each $x \in X \cup \{\Delta\}$, there is a $X_\Delta := X \cup \{\Delta\}$-valued stochastic process $(\Omega, (\mathcal{F}_t)_{t\geqslant 0}, \{Y_t : t \in [0,\infty]\}, \{\mathbb{P}_x\}_{x \in X\cup\{\Delta\}})$ that satisfies the following properties:*

1. *$(\mathcal{F}_t)_{t\geqslant 0}$ is a right continuous filtration on Ω and (Y_t) is $(\mathcal{F}_t)_{t\geqslant 0}$-adapted; that is $Y_t : (\Omega, \mathcal{F}_t) \to (X \cup \Delta, \mathcal{B}_\Delta)$ is measurable for each $t \geqslant 0$.*

2. *For each $E \in \mathcal{B}_\Delta$ and $t > 0$, the function $x \mapsto \mathbb{P}_x(Y_t \in E)$ is measurable on $(X \cup \Delta, \mathcal{B}_\Delta)$.*

3. *(Markov property) For each $x \in X, t, s \geqslant 0$ and $E \in \mathcal{B}$ we have*

$$\mathbb{P}_x(Y_{t+s} \in E | \mathcal{F}_t) = \mathbb{P}_{Y_t}(Y_s \in E).$$

4. *(cemetery is absorbing) $Y_\infty(\omega) = \Delta$ for all $\omega \in \Omega$, $\mathbb{P}_\Delta(Y_t = \Delta) = 1$ for all $t \geqslant 0$. More generally, $\mathbb{P}_x(Y_t(\omega) = \Delta) = 1$ for all $t \geqslant \zeta(\omega)$, where $\zeta(\omega)$ denotes the* lifetime of the process

$$\zeta(\omega) := \inf\{t \geqslant 0 : Y_t(\omega) = \Delta\}.$$

5. *(normal process) For each $x \in X$, we have $\mathbb{P}_x(Y_0 = x) = 1$.*

6. *(càdlàg paths) $t \mapsto Y_t(\omega)$ is right continuous on $[0, \infty)$ and has left limits on $(0, \infty)$.*

7. *(time shift operators) For each $t \geqslant 0$, there exists a* time shift operator $\theta_t : \Omega \to \Omega$ *such that $Y_s \circ \theta_t = Y_{t+s}$.*

8. *(strong Markov property) For any probability measure μ on $(X \cup \{\Delta\}, \mathcal{B}_\Delta)$ and for any $(\mathcal{F}_t)$-stopping time T, and for all $s \geqslant 0$, we have*

$$\mathbb{P}_\mu\left(Y_{T+s} \in E | \mathcal{F}_T\right) = \mathbb{P}_{Y_T}(Y_s \in E), \quad \textit{for all } E \in \mathcal{B}_\Delta.$$

9. *The process is* quasi-left continuous *on $(0, \infty)$: for any sequence of $(\mathcal{F}_t)_{t \geqslant 0}$ stopping times $T_n \uparrow T$ and any probability measure μ on $(X \cup \{\Delta\}, \mathcal{B}_\Delta)$, we have*

$$\mathbb{P}_\mu\left(\lim_{n\to\infty} Y_{T_n} = Y_T, T < \infty\right) = \mathbb{P}_\mu(T < \infty).$$

The semigroup (P_t^Y) corresponding to the process (Y_t) defined by $P_t^Y f(x) = \mathbb{E}_x[f(Y_t)\mathbb{1}_{\{t<\zeta\}}]$ coincides with the semigroup corresponding to the Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(X, m)$ (the term $\mathbb{1}_{\{t<\zeta\}}$ is usually dropped with the convention that every function f on X is extended to $X \cup \{\Delta\}$ by setting $f(\Delta) = 0$).

The Markov process in Fukushima's theorem is not quite unique because the semigroup associated to a Dirichlet form is not well-defined pointwise. However, it is essentially unique outside a very small set as shown in [FOT, Theorem 4.2.8]. We describe the uniqueness below.

Let $X \cup \{\Delta\}$-valued stochastic process $(\Omega, (\mathcal{F}_t)_{t\geqslant 0}, \{Y_t : t \in [0, \infty]\}, \{\mathbb{P}_x\}_{x \in X \cup \{\Delta\}})$ be a stochastic process as above. We say that subset $\mathcal{N}$ of X is *properly exceptional* for the Markov process $\{Y_t : t \in [0, \infty]\}$ if $\mathcal{N}$ is a Borel set with $m(\mathcal{N}) = 0$ and

$$\mathbb{P}_x\left(\{\omega \in \Omega | Y_t(\omega) \in X_\Delta \backslash \mathcal{N}, Y_{t-}(\omega) \in X_\Delta \backslash \mathcal{N}\}, \text{for all } t \geqslant 0\right) = 1, \quad \text{for all } x \in X \backslash \mathcal{N}.$$

The process constructed in Fukushima's theorem is unique in the following sense. Suppose $\{Y_t : t \geqslant 0\}$ and $\{\widetilde{Y}_t : t \geqslant 0\}$ be two Markov processes that satisfy the conclusion of

Theorem 2.37, then there is a common properly exceptional set $\mathcal{N}$ for both $\{Y_t : t \geqslant 0\}$ and $\{\widetilde{Y}_t : t \geqslant 0\}$ such that for all $x \in X \backslash \mathcal{N}$, for all $t \geqslant 0$ and $E \in \mathcal{B}$, we have

$$\mathbb{P}_x(Y_t \in E) = \widetilde{\mathbb{P}}_x(\widetilde{Y}_t \in E).$$

Let us recall some basic facts about regular Dirichlet forms. First, we define the 1-*capacity* $\mathrm{Cap}_1(A)$ of $A \subset X$ with respect to $(X, m, \mathcal{E}, \mathcal{F})$ by

$$\mathrm{Cap}_1(A) := \inf\left\{\mathcal{E}_1(f,f) \,\middle|\, f \in \mathcal{F},\ f \geqslant 1\ m\text{-a.e. on a neighborhood of } A\right\}, \tag{2.21}$$

where $\mathcal{E}_1 := \mathcal{E} + \langle \cdot, \cdot \rangle_{L^2(X,m)}$ as defined before. A subset $\mathcal{N}$ of X is said to be $\mathcal{E}$-*polar* if $\mathrm{Cap}_1(\mathcal{N}) = 0$. For $A \subset X$ and a statement $\mathcal{S}(x)$ on $x \in A$, we say that $\mathcal{S}$ holds $\mathcal{E}$-*quasi-everywhere on* A ($\mathcal{E}$-*q.e. on* A for short), or $\mathcal{S}(x)$ holds for $\mathcal{E}$-*quasi-every* $x \in A$ ($\mathcal{E}$-*q.e.* $x \in A$ for short), if $\mathcal{S}(x)$ holds for any $x \in A \backslash \mathcal{N}$ for some $\mathcal{E}$-polar $\mathcal{N} \subset X$.

Polar sets and properly exceptional are defined using the Dirichlet form $(\mathcal{E}, \mathcal{F})$ and the corresponding Markov process $(Y_t)_{t \geqslant 0}$ respectively. These two notions are closely related as follows: any properly exceptional set $\mathcal{N} \subset X$ for $(Y_t)_{t \geqslant 0}$ is $\mathcal{E}$-polar, c any $\mathcal{E}$-polar subset of X is included in some properly exceptional set $\mathcal{N}$ for $(Y_t)_{t \geqslant 0}$.

Definition 2.38 (Quasi-continuous function)**.** Let $(\mathcal{E}, \mathcal{F})$ is a regular Dirichlet form on $L^2(X, m)$. Let $f \in L^2(X, m)$. A *quasi-continuous version* of f is a (pointwise defined) function $\widetilde{f} : X \to \mathbb{R}$, $\widetilde{f} = f$ m-a.e. and for $\epsilon > 0$ there exists an open set G of X such that $\mathrm{Cap}_1(G) < \epsilon$ and $\widetilde{f}\big|_{X \backslash G}$.

A crucial ingredient in the proof of Theorem 2.37 is the existence of quasicontinuous modification of functions in $\mathcal{F}$ [FOT, Theorem 2.1.3].

Theorem 2.39. *Let* $(\mathcal{E}, \mathcal{F})$ *be a regular Dirichlet form on* $L^2(X, m)$*. Then every function* $f \in \mathcal{F}$ *admits a quasi-continuous modification.*

Remark 2.40. The assumption of local compactness in the definition of regular Dirichlet forms does not allow for some infinite dimensional examples. There is a fruitful generalization of regular Dirichlet form that allows for such examples called *quasi-regular Dirichlet forms*. We refer to [AM, CF] for more on this theory.

2.6 Irreducibility, recurrence and transience

Familiar probabilistic notions such as irreducibility, recurrence and transience can be defined at the level of Dirichlet forms.

Let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(X, m)$ and let $(P_t)_{t>0}$ denote the corresponding semigroup.

Definition 2.41. We say that a measurable set $A \subset X$ is $\mathcal{E}$-*invariant*, if it satisfies

$$P_t(\mathbb{1}_A f) = \mathbb{1}_A P_t(f), \quad m\text{-a.e. on } X \text{ for any } f \in L^2(X, m) \text{ and any } t \in (0, \infty).$$

Equivalently, $\mathbb{1}_A u \in \mathcal{F}$ for any $u \in \mathcal{F}$ and

$$\mathcal{E}(u,u) = \mathcal{E}(\mathbb{1}_A u, \mathbb{1}_A u) + \mathcal{E}(\mathbb{1}_{A^c} u, \mathbb{1}_{A^c} u).$$

We say that a Dirichlet form $(\mathcal{E}, \mathcal{F})$ is *irreducible* if $m(A)m(X\backslash A) = 0$ for any $\mathcal{E}$-invariant set A.

In order to define recurrence and transience, we recall the definition of extended Dirichlet space.

Definition 2.42 (Extended Dirichlet space)**.** We define the *extended Dirichlet space* $\mathcal{F}_e$ of a Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(X,m)$ as the space of m-equivalence classes of functions $f: X \to \mathbb{R}$ such that $\lim_{n\to\infty} f_n = f$ m-a.e. on X for some $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{F}$ with $\lim_{k\wedge l\to\infty} \mathcal{E}(f_k - f_l, f_k - f_l) = 0$. Then the limit $\mathcal{E}(f,f) := \lim_{n\to\infty} \mathcal{E}(f_n, f_n) \in \mathbb{R}$ exists and is independent of a choice (why?) of such $\{f_n\}_{n\in\mathbb{N}}$ for each $f \in \mathcal{F}_e$, so that $\mathcal{E}$ is canonically extended to $\mathcal{F}_e \times \mathcal{F}_e$.

We say that a Markovian semigroup $(P_t)_{t\geqslant 0}$ is *transient* if for any $f \in L^1(X,m) \cap L^\infty(X,m)$ with $f \geqslant 0$-m-a.e., we have

$$Gf := \lim_{T\to\infty} \int_0^T P_t(f)\, dm < \infty, \quad m\text{-a.e.}$$

We say that a Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(X,m)$ is *transient* if there exists a $f \in L^1(X,m) \cap L^\infty(X,m)$ such that f is strictly positive m-a.e., and satisfying

$$\int_X |u| f\, dm \leqslant \sqrt{\mathcal{E}(u,u)}, \quad \text{for all } u \in \mathcal{F}.$$

The following theorem gives equivalent definitions of transience.

Theorem 2.43. *Let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(X,m)$ and let $(P_t)_{t>0}$ denote the corresponding semigroup. The following are equivalent:*

1. *The semigroup $(P_t)_{t>0}$ is transient.*
2. *The Dirichlet form $(\mathcal{E}, \mathcal{F})$ is transient.*
3. *$(\mathcal{F}_e, \mathcal{E})$ is a Hilbert space.*

We say that a m-symmetric Markov semigroup $(P_t)_{t\geqslant 0}$ is *recurrent*, if for all non-negative $f \in L^1(X,m)$, we have

$$Gf := \lim_{T\to\infty} \int_0^T P_t f\, dt \in \{0, \infty\} \quad m\text{-a.e. on } X.$$

We say that a Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(X,m)$ is *recurrent* if $\mathbb{1}_X \in \mathcal{F}_e$ and $\mathcal{E}(\mathbb{1}_X, \mathbb{1}_X) = 0$. Analogous to Theorem 2.43, we have the following result.

Theorem 2.44. *Let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(X, m)$ and let $(P_t)_{t>0}$ denote the corresponding semigroup. The following are equivalent:*

1. *The semigroup $(P_t)_{t>0}$ is recurrent.*
2. *The Dirichlet form $(\mathcal{E}, \mathcal{F})$ is recurrent.*

The following exercise concerns the Dirichlet form for n-dimensional Bessel process.

Exercise 2.45. Let $n \in (0, \infty)$. Consider the measure $m(dx) = x^{n-1}\,dx$ on $X = (0, \infty)$. Let $\mathcal{F}$ denote the set of all functions $f : X \to \mathbb{R}$ such that f is absolutely continuous on X and satisfies

$$\int_{(0,\infty)} \left(|f'(x)|^2 + |f(x)|^2\right) m(dx) = \int_{(0,\infty)} \left(|f'(x)|^2 + |f(x)|^2\right) x^{n-1}\,dx < \infty.$$

Then $(\mathcal{E}, \mathcal{F})$ is a Dirichlet form on $L^2(X, m)$, where

$$\mathcal{E}(f, f) = \int_X f'(x) g'(x)\, m(dx), \quad \text{for all } f, g \in \mathcal{F}.$$

1. Show that $(\mathcal{E}, \mathcal{F})$ is irreducible.
2. Show that $(\mathcal{E}, \mathcal{F})$ is recurrent if $n \in (0, 2]$ and is transient if $n \in (2, \infty)$.
   Hint: For the case $n \in (2, \infty)$, the proving the following inequality is useful

   $$|f(x) - f(y)|^2 \leqslant (n-2)^{-1}(x^{2-n} - y^{2-n})\mathcal{E}(f, f), \quad \text{for all } f \in \mathcal{F}, x, y \in I \text{ with } x < y.$$

3 Heat kernel

Definition 3.1. Let $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form on $L^2(X, m)$. A family $\{p_t\}_{t>0}$ of $[0, \infty]$-valued Borel measurable functions on $X \times X$ is called the *heat kernel*, if p_t is an integral kernel of the operator P_t for any $t \in (0, \infty)$, that is, for any $t \in (0, \infty)$ and any $f \in L^2(X, m)$,

$$P_t f(x) = \int_X p_t(x, y) f(y)\, dm(y) \qquad \text{for } m\text{-a.e. } x \in X.$$

3.1 Existence of heat kernel via ultracontractivity

In general, a heat kernel need not exist (consider the semigroup consisting of identity maps). From a probabilistic perspective, if the heat kernel exists, $p_t(x, y)\, m(dy)$ can be viewed as the law of the corresponding Markov process started at x at time t. So the existence of heat kernel can be viewed as the absolute continuity (with respect to the reference measure m) of the law of a Markov process.

The following very general result is useful to show the existence of a heat kernel (see [DS, Theorem 6, VI.8.6]).

Proposition 3.2. *Suppose $(X, \mathcal{M}, m)$ be a σ-finite, separable measure space and let $T : L^1(X, m) \to L^\infty(X, m)$ be a bounded linear operator. Then there exists a jointly measurable function $K : X \times X \to \mathbb{R}$ such that $K \in L^\infty(X \times X, m \times m)$ with $\|K\|_\infty = \|T\|_{1\to\infty}$ and*

$$Tf(x) = \int_X K(x, y)f(y)\, m(dy), \quad m\text{-a.e. for each } f \in L^1(X, m).$$

A linear operator that is bounded from L^1 to L^∞ as above is said to be *ultracontractive*. Ultracontractivity of semigroup offers a way to show the existence of heat kernel.

Proposition 3.3. *Let $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form on $L^2(X, m)$ and let $\{P_t : t > 0\}$ denote the associated semigroup. Assume that there exists $C_1 \geqslant 0, C_2 > 0$ and $n \geqslant 1$ such that we have the* Nash inequality

$$\|f\|_2^{1+n/2} \leqslant \|f\|_1 \left(C_1 \|f\|_2^2 + C_2 \|\mathcal{E}(f, f)\|\right)^{n/4}, \quad \text{for all } f \in L^1(X, m) \cap \mathcal{F}. \tag{3.1}$$

Then the semigroup $\{P_t : t > 0\}$ admits a heat kernel $\{p_t(\cdot, \cdot)\}$ such that

$$\operatorname*{ess\,sup}_{x,y\in X} p_t(x, y) \leqslant \max\left(2C_1, \frac{nC_2}{t}\right)^{n/2}.$$

Proof. First we show the estimate

$$\|P_t(f)\|_2 \leqslant \max\left(2C_1, \frac{nC_2}{2t}\right)^{n/4} \|f\|_1, \quad \text{for all } t > 0 \text{ and } f \in L^1(X, m). \tag{3.2}$$

By the density of $L^1 \cap L^2$ in L^1 and linearity of P_t, it suffices to consider $f \in L^1(X, m) \cap L^2(X, m)$ with $\|f\|_1 = 1$. By the spectral theorem, we note that $P_t(f) \in \mathcal{D}(A)$ for all $t > 0$, where A is the generator and

$$\frac{d}{dt}\|P_t(f)\|^2 = 2\langle \frac{d}{dt}P_t(f), P_t(f)\rangle = 2\langle AP_t(f), P_t(f)\rangle = -2\mathcal{E}(P_t(f), P_t(f)).$$

Setting $\Psi(t) := \|P_t(f)\|_2^2$ by the above equality and (3.1), we obtain

$$\begin{aligned}\Psi(t)^{1+2/n} = \|P_t(f)\|_2^{2+4/n} &\overset{(3.1)}{\leqslant} \|P_t(f)\|_1^{4/n}\left(C_1 \|P_t(f)\|_2^2 + C_2 \|\mathcal{E}(P_t(f), P_t(f))\|\right)\\ &\overset{(2.17)}{\leqslant} C_1\Psi(t) - 2^{-1}C_2\Psi'(t).\end{aligned}$$

If $\Psi(t) \geqslant (2C_1)^{n/2}$, then $\Psi(t)^{1+2/n} - C_1\Psi(t) \geqslant \Psi(t)(\Psi(t)^{2/n} - C_1) \geqslant 2^{-1}\Psi(t)^{2/n}$. This leads to the differential inequality,

$$\Psi(t)^{1+2/n} \leqslant -C_2\Psi'(t)$$

for all $t > 0$ such that $\Psi(t) \geqslant (2C_1)^{n/2}$ (recall that Ψ is non-increasing). This can be written as

$$\frac{d}{dt}\left(\Psi(t)^{-2/n}\right) = -\frac{2}{n}\Psi(t)^{-(1+2/n)}\Psi'(t) \geqslant \frac{2}{nC_2}$$

which implies (3.2).

Now by duality of L_p spaces and symmetry of P_t, we have $\|P_t\|_{1\to 2} = \|P_t\|_{2\to\infty}$. To see this, note that

$$\begin{aligned}\|P_t\|_{2\to\infty} &= \sup_{f\in L^2, \|f\|_2=1} \|P_t(f)\|_\infty = \sup_{f\in L^1\cap L^\infty, \|f\|_2=1} \|P_t(f)\|_\infty \\ &= \sup_{\substack{f\in L^1\cap L^\infty, \|f\|_2=1,\\ g\in L^1, \|g\|_1=1}} \int_X g P_f(f)\, dm = \sup_{\substack{f\in L^1\cap L^\infty, \|f\|_2=1,\\ g\in L^1\cap L^\infty, \|g\|_1=1}} \langle P_t(f), g\rangle \\ &= \sup_{\substack{f\in L^1\cap L^\infty, \|f\|_2=1,\\ g\in L^1\cap L^\infty, \|g\|_1=1}} \langle f, P_t(g)\rangle = \sup_{g\in L^1\cap L^\infty, \|g\|_1=1} \|P_t(g)\|_2 = \|P_t\|_{1\to 2}.\end{aligned}$$

Now by the semigroup property and submultiplicativity of operator norms, we have

$$\|P_t\|_{1\to\infty} \leqslant \left\|P_{t/2}\right\|_{1\to 2} \left\|P_{t/2}\right\|_{2\to\infty} = \left\|P_{t/2}\right\|_{1\to 2}^2 \overset{(3.2)}{\leqslant} \max\left(2C_1, \frac{nC_2}{t}\right)^{n/2}.$$

The existence of heat kernel and the upper bound follows from Proposition 3.2 □

Example 3.4. J. Nash obtained the following inequality on $\mathbb{R}^n$: there exists $C > 0$ such that

$$\|f\|_2^{1+n/2} \leqslant C\, \|f\|_1 \left(\int_{\mathbb{R}^n} |\nabla f(x)|^2\, dx\right)^{n/4}, \quad \text{for all } f \in L^1(\mathbb{R}^n) \cap W^{1,2}(\mathbb{R}^n).$$

Therefore Proposition 3.3, gives the upper bound $p_t(x, y) \lesssim t^{-n/2}$ for all $x, y \in \mathbb{R}^n$ for the heat kernel for Brownian motion on $\mathbb{R}^n$. The advantage of this method is that it is robust to perturbations. Consider a measurable symmetric positive definite matrix valued function $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ such that there exists $\Lambda \in [1, \infty)$

$$\Lambda^{-1} \|\xi\|^2 \leqslant \xi^t \mathcal{A}(x)\xi \leqslant \Lambda \|\xi\|^2, \quad \text{for all } x, \xi \in \mathbb{R}^n.$$

Then the Dirichlet form $\mathcal{E}(f, f) = \int_{\mathbb{R}^n} (\nabla f(x))^t \mathcal{A}(x)\nabla(f)\, dx$ on $W^{1,2}(\mathbb{R}^n)$ corresponding to uniformly elliptic operator $\mathrm{div}(\mathcal{A}(x)\nabla(\cdot))$ also satisfies the above Nash inequality and hence the same upper bound.

Let us quickly sketch the proof of Nash for the case when $f \in C^\infty(\mathbb{R}^n)$ is smooth and well-behaved at infinity [?]. The Fourier transform $\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i x\cdot\xi}\, dx$ satisfies the Parseval's identity $\left\|\widehat{f}\right\|_2 = \|f\|_2$ and the elementary bound $\left\|\widehat{f}\right\|_\infty \leqslant \|f\|_1$. Again by Parseval's identity and the Fourier transform of derivative formula, we have

$$\int_{\mathbb{R}^n} |\nabla f(x)|^2\, dx = \int_{\mathbb{R}^n} (2\pi|\xi|)^2 \left|\widehat{f}(\xi)\right|^2 d\xi.$$

Denoting ω_n as the volume of unit ball in $\mathbb{R}^n$, we have

$$\|f\|_2^2 = \int_{\mathbb{R}^n} \left|\widehat{f}(\xi)\right|^2 d\xi$$

$$\leqslant \int_{\{|\xi|<R\}} \left\|\widehat{f}\right\|_\infty d\xi + R^{-2}\int_{\{|\xi|\geqslant R\}} |\xi|^2 \left|\widehat{f}(\xi)\right|^2 d\xi$$
$$\leqslant \|f\|_1^2 \omega_n R^n + (2\pi R)^{-2}\int_{\mathbb{R}^n} |\nabla f(x)|^2 \, dx.$$

Minimizing the above expression as a function of R, we obtain the Nash inequality when f is smooth and well-behaved at infinity. The general case follows from the same argument by viewing f as a tempered distribution.

3.2 Brownian motion on the Sierpiński gasket

Let $V_0 = \{q_1, q_2, q_3\}$, where

$$q_1 := \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad q_2 := (0,0), \quad q_3 := (1,0).$$

Let $S := \{1,2,3\}$ and set $f_j : \mathbb{R}^2 \to \mathbb{R}^2$ as $f_j(x) := \frac{1}{2}(x + q_j)$ for all $j \in S$. Define inductively for $m \in \mathbb{N}$

$$V_m = \bigcup_{j\in S} f_j(V_{m-1}).$$

By an induction argument, $V_m \subset V_{m+1}$ for all $m \in \mathbb{N} \cup \{0\}$. Set $V_* = \bigcup_{m=0}^{\infty} V_m$ and let $K = \overline{V_*}$ denote the closure of V_* in $\mathbb{R}^2$. Since $V_* = \bigcup_{j\in S} f_j(V^*)$, we have

$$K = \bigcup_{j\in S} f_j(K).$$

We set $F_j = f_j\big|_K$ for all $j \in S$. We define words of lengths with alphabet S as $W_m = S^m = w_1 \dots w_m : w_1, \dots, w_m \in S$ for all $m \in \mathbb{N} \cup \{0\}$ (with $W_0 = \{\varnothing\}$ consisting of the empty word). For $m \in \mathbb{N} \cup \{0\}$, $w = w_1 \dots w_m \in W_m$ and $F_w := F_{w_1} \circ \cdots \circ F_{w_m}$ (with $F_\varnothing = I$) and $x_0 := q_1$.

The approach behind constructing a Dirichlet form on K is to construct a limit of a sequence of Dirichlet forms on V_m as $m \to \infty$. Equivalently, from a probabilistic perspective, the diffusion on Sierpiński gasket is constructed as a limit of random walks on a sequence of graph approximations. For $x, y \in V_m$, we say that $x \overset{m}{\sim} y$ if and only if $x \neq y$ and $x, y \in F_w(V_0)$ for some $w \in W_m$.

We define quadratic forms $E^{(m)}, \mathcal{E}^{(m)} : V_m \times V_m \to \mathbb{R}$ for all $u, v \in \mathbb{R}^{V_m}$

$$E^{(m)}(u,v) = \frac{1}{2} \sum_{\substack{x,y\in V_m,\\ x\overset{m}{\sim}y}} (u(x) - u(y)(v(x) - v(y))), \quad \mathcal{E}^{(m)}(u,v) = \left(\frac{5}{3}\right)^m E^{(m)}(u,v). \tag{3.3}$$

The reason for the re-scaling factor $\left(\frac{5}{3}\right)^m$ is due to the following fact. For any $m \in \mathbb{N} \cup \{0\}, u \in \mathbb{R}^{V_m}$ there exists a unique extension $H_{m.m+1}(u) \in \mathbb{R}^{V_{m+1}}$ of u such that

$$E^{(m+1)}(H_{m.m+1}(u), H_{m.m+1}(u)) = \min_{\substack{v\in\mathbb{R}^{V_{m+1}},\\ v|_{V_m}=u}} E^{(m+1)}(v,v) = \frac{3}{5} E^{(m)}(u,u). \tag{3.4}$$

The analysis can be reduced to the case $m = 0$. Suppose $(u(q_1), u(q_2), u(q_3)) = (a, b, c) \in \mathbb{R}^3$. For $j, k \in S$ with $j \neq k$, we set $q_{kj} = F_k(q_j)$, so that $F_k(K) \cap F_j(K) = \{q_{jk}\}$. Let $v \in \mathbb{R}^{V_1}$ be an extension of u with $(x, y, z) := (v(q_{23}), v(q_{31}), v(q_{12}))$, then

$$E^{(1)}(v,v) = (y-z)^2+(a-y)^2+(a-z)^2+(b-x)^2+(x-z)^2+(z-b)^2+(x-c)^2+(c-y)^2+(y-x)^2.$$

We need to minimize this expression as a function of x, y, z. This leads to

$$4x = b + z + y + c, \quad 4y = c + x + z + a, \quad 4z = a + y + x + b,$$

or equivalently,

$$x = \frac{a + 2b + 2c}{5}, \quad y = \frac{2a + b + 2c}{5}, \quad z = \frac{2a + 2b + c}{5}.$$

For v chosen as above, it is easy to compute that $E^{(1)}(v, v) = \frac{3}{5}E^{(0)}(u, u)$. This completes the proof of (3.4).

By (3.4), for any $u \in \mathbb{R}^{V_*}$, the sequence $\left(\mathcal{E}^{(m)}(u\big|_{V_m}, u\big|_{V_m})\right)_{m\in\mathbb{N}\cup\{0\}}$ is non-decreasing and hence $\lim_{m\to\infty} \mathcal{E}^{(m)}(u\big|_{V_m}, u\big|_{V_m}) \in [0, \infty]$ exists. Define

$$\mathcal{F}_* := \{u \in \mathbb{R}^{V_*} | \lim_{m\to\infty} \mathcal{E}^{(m)}(u\big|_{V_m}, u\big|_{V_m}) < \infty\}, \quad \mathcal{E}^{(*)}(u, u) = \lim_{m\to\infty} \mathcal{E}^{(m)}(u\big|_{V_m}, u\big|_{V_m}),$$

for all $u \in \mathcal{F}_*$. It is easy to see that $\mathcal{F}_*$ is a subspace of $\mathbb{R}^{V_*}$ since $\mathcal{E}^{(*)}(au, au) = a^2\mathcal{E}^{(*)}(u, u)$ for all $a \in \mathbb{R}, u \in \mathbb{R}^{V_*}$, and $\mathcal{E}^{(*)}(u + v)^{1/2} \leqslant \mathcal{E}^{(*)}(u, u)^{1/2} + \mathcal{E}^{(*)}(v, v)^{1/2}$ for all $u, v \in \mathcal{F}_*$ (see Exercise 2.26). By bi-linearity we have that $\lim_{m\to\infty} \mathcal{E}^{(m)}(u\big|_{V_m}, u\big|_{V_m}) \in \mathbb{R}$ exists for all $u, v \in \mathcal{F}_*$.

The following self-similarity follows easily from the definition of the energies.

Exercise 3.5. For any $u \in \mathbb{R}^{V_*}$, we have

$$\mathcal{E}^{(m+1)}(u\big|_{V_{m+1}}, u\big|_{V_{m+1}}) = \frac{5}{3}\sum_{j\in S} \mathcal{E}^{(m)}(u \circ F_j\big|_{V_m}, u \circ F_j\big|_{V_m}), \tag{3.5}$$

and hence

$$\mathcal{F}_* = \{u \in \mathbb{R}^{V_*} : u \circ F_j \in \mathcal{F}_* \text{ for all } j \in S\}, \quad \mathcal{E}^{(*)}(u, u) = \frac{5}{3}\sum_{j\in S} \mathcal{E}^{(*)}(u \circ F_j, u \circ F_j). \tag{3.6}$$

More generally, for any $m \in \mathbb{N}$, we have

$$\mathcal{F}_* = \{u \in \mathbb{R}^{V_*} : u \circ F_\tau \in \mathcal{F}_* \text{ for all } \tau \in W_m\}, \quad \mathcal{E}^{(*)}(u, u) = \left(\frac{5}{3}\right)^m \sum_{\tau\in W_m} \mathcal{E}^{(*)}(u \circ F_\tau, u \circ F_\tau). \tag{3.7}$$

Proposition 3.6. *For all $x, y \in V_*$ and $u \in \mathcal{F}_*$, we have*

$$|u(x) - u(y)|^2 \leqslant 400|x - y|^\alpha \mathcal{E}^{(*)}(u, u) \tag{3.8}$$

where $|\cdot|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $\alpha = \log_2(5/3)$. In particular, every $u \in \mathcal{F}_*$ is uniformly continuous in V_* and hence admits a unique continuous extension to its closure $K = \overline{V_*}$.

Proof. Note that the graph corresponding to V_1 with edges given by $\overset{1}{\sim}$ has diameter 2. Hence for any $y, z \in V_1$ there exists $q \in V_1$ such that one of hte following hold: $y = z = q$ or $y \overset{1}{\sim} q = z$ or $y \overset{1}{\sim} q \overset{1}{\sim} z$. In all these cases, for any $v \in \mathbb{R}^{V_1}$ we have

$$\begin{aligned}|v(y) - v(z)| &\leqslant |v(y) - v(q)| + |v(q) - v(z)| \leqslant \sqrt{2}\left(|v(y) - v(q)|^2 + |v(q) - v(z)|^2\right)^{1/2} \\ &\leqslant \sqrt{2}E^{(1)}(v, v) \overset{(3.3)}{=} \sqrt{6/5}\mathcal{E}^{(1)}(v, v). \end{aligned} \tag{3.9}$$

Now let $u \in \mathcal{F}_*, x \in V_*$ be arbitrary. Then there exists $m \in \mathbb{N}, w = w_1 \dots w_m \in W_m, j \in S$ such that $F_w(q_j)$. Set for $1 \leqslant k \leqslant m$, $x_k = F_{w_1 \dots w_{k-1}}(q_j)$ (with $F_\varnothing = I$ for $j = 0$). Then for all $k = 1, \dots, m$ we have

$$\begin{aligned}|u(x_{k-1}) - u(x_k))| &= \left|u \circ F_{w_1 \dots w_{k-1}}(F^{-1}_{w_1 \dots w_{k-1}}(x_{k-1})) - u \circ F_{w_1 \dots w_{k-1}}(F^{-1}_{w_1 \dots w_{k-1}}(x_k))\right| \\ &\overset{(3.9)}{\leqslant} \sqrt{6/5}\mathcal{E}^{(1)}(u \circ F_{w_1 \dots w_{k-1}}|_{V_1}, u \circ F_{w_1 \dots w_{k-1}}|_{V_1})^{1/2} \leqslant \sqrt{6/5}\mathcal{E}^{(*)}(u \circ F_w, u \circ F_w) \\ &\leqslant \sqrt{\frac{6}{5}}\left(\frac{3}{5}\right)^{(k-1)/2}\left(\left(\frac{5}{3}\right)^{(k-1)}\sum_{\tau \in W_{k-1}} \mathcal{E}^{(*)}(u \circ F_\tau, u \circ F_\tau)\right)^{1/2} \\ &\overset{(3.7)}{=} \sqrt{\frac{6}{5}}\left(\frac{3}{5}\right)^{(k-1)/2} \mathcal{E}^{(*)}(u, u)^{1/2}. \end{aligned} \tag{3.10}$$

Hence we obtain

$$\begin{aligned}|u(q_1) - u(x)| &\leqslant \sum_{k=1}^{m} |u(x_{k-1}) - u(x_k))| \overset{(3.10)}{\leqslant} \sum_{k=1}^{m} \sqrt{\frac{6}{5}}\left(\frac{3}{5}\right)^{(k-1)/2} \mathcal{E}^{(*)}(u, u)^{1/2} \\ &\leqslant \sum_{k=1}^{\infty} \sqrt{\frac{6}{5}}\left(\frac{3}{5}\right)^{(k-1)/2} \mathcal{E}^{(*)}(u, u)^{1/2} \leqslant 5\mathcal{E}^{(*)}(u, u)^{1/2}. \end{aligned} \tag{3.11}$$

Thus for any $x, y \in V_*$ and any $u \in \mathcal{F}_*$, we have

$$|u(x) - u(y)| \leqslant |u(x) - u(q_1)| + |u(y) - u(q_1)| \overset{(3.11)}{\leqslant} 10\mathcal{E}^{(*)}(u, u)^{1/2}. \tag{3.12}$$

Since for any $w \in W_m, m \in \mathbb{N} \cup \{0\}$, $F_w(K)$ is a subset of an equilateral triangle with side length 2^{-m}, for any $w, v \in W_m$ with $F_w(K) \cap F_v(K) \neq \varnothing$, we have

$$|x - y| \leqslant 2^{-m} + 2^{-m} = 2^{-m+1}, \quad \text{for all } x \in F_v(K), y \in F_w(K). \tag{3.13}$$

Let $x, y \in K$ be arbitrary with $x \neq y$. Let $n := n_{x,y} = \max\{m \in \mathbb{N} \cup \{0\} :$ there exist $v, w \in W_m$ with $x \in F_v(V_*), y \in F_w(V_*), F_v(K) \cap F_w(K) \neq \varnothing\}$. Let $q \in F_v(K) \cap F_w(K) \subset V_*$, where $v, w \in W_n$ with $x \in F_v(V_*), y \in F_w(V_*)$ as above. Then

$$\begin{aligned}|u(x) - u(y)| &\leqslant |u(x) - u(q)| + |u(y) - u(q)| = \left|u \circ F_v(F_v^{-1}(x)) - u \circ F_w(F_w^{-1}(y))\right| \\ &\overset{(3.12)}{\leqslant} 10\left(\mathcal{E}^{(*)}(u \circ F_v, u \circ F_v)^{1/2} + \mathcal{E}^{(*)}(u \circ F_w, u \circ F_w)^{1/2}\right) \\ &\leqslant 10\sqrt{2}\left(\mathcal{E}^{(*)}(u \circ F_v, u \circ F_v) + \mathcal{E}^{(*)}(u \circ F_w, u \circ F_w)\right)^{1/2} \end{aligned}$$

$$\overset{v\neq w}{\leqslant} 10\sqrt{2}\left(\frac{3}{5}\right)^{n/2}\left(\left(\frac{5}{3}\right)^{n}\sum_{\tau\in W_n}\mathcal{E}^{(*)}(u\circ F_\tau\big|_{V_*}, u\circ F_\tau\big|_{V_*})\right)^{1/2}$$

$$\overset{(3.7)}{=} 10\sqrt{2}\left(\frac{3}{5}\right)^{n/2}\mathcal{E}^{(*)}(u,u)^{1/2}. \tag{3.14}$$

Let $i,j \in S$ be such that $x \in F_{vi}(K), y \in F_{wj}(K)$, where $vi, wj \in W_{n+1}$ denotes concatenation of words. By the maximality on n, $F_{vi}(K) \cap F_{wj}(K) = \varnothing$ and hence $|x-y| \geqslant \frac{\sqrt{3}}{2}2^{-n-1}$ (draw a picture to see the possibilities). Hence from (3.14), we estimate

$$|u(x)-u(y)|^2 \leqslant (200)\cdot 2^{-\alpha n}\mathcal{E}^{(*)}(u,u) \leqslant 200\left(\frac{4}{\sqrt{3}}\right)^{\alpha}|x-y|^{\alpha}\mathcal{E}^{(*)}(u,u) \leqslant 400|x-y|^{\alpha}\mathcal{E}^{(*)}(u,u).$$

□

Let $\mathcal{F} \subset C(K)$ be defined as

$$\mathcal{F} := \{u \in C(K) : \lim_{m\to\infty}\mathcal{E}^{(m)}(u\big|_{V_m}, u\big|_{V_m}) < \infty\} = \{u \in C(K) : u\big|_{V_*} \in \mathcal{F}_*\}.$$

The standard Dirichlet form on Sierpińksi gasket $\mathcal{E} : \mathcal{F}\times\mathcal{F} \to \mathbb{R}$ is defined as

$$\mathcal{E}(u,v) := \lim_{m\to\infty}\mathcal{E}^{(m)}(u\big|_{V_m}, v\big|_{V_m}), \quad \text{for all } u,v \in \mathcal{F}.$$

The following properties of $(\mathcal{E},\mathcal{F})$ follow easily from the construction. From (3.6), we have the self-similarity property

$$\mathcal{F} = \{u \in C(K) : u\circ F_j \in \mathcal{F} \text{ for all } j \in S\}, \quad \mathcal{E}(u,u) = \frac{5}{3}\sum_{j\in S}\mathcal{E}(u\circ F_j, u\circ F_j). \tag{3.15}$$

By Proposition 3.6 and the density of V_* in K, we have

$$|u(x)-u(y)|^2 \leqslant 400|x-y|^{\alpha}\mathcal{E}(u,u), \quad \text{for all } u \in \mathcal{F}, \tag{3.16}$$

where $\alpha = \log_2(5/3)$. In particular, by (3.16), we have

$$\mathbb{R}\mathbb{1}_K = \{a\mathbb{1}_K | a \in \mathbb{R}\} = \{u \in \mathcal{F} | \mathcal{E}(u,u) = 0\}.$$

For any 1-Lipschitz function $F : \mathbb{R} \to \mathbb{R}$ and for any $u \in \mathcal{F}$, we have $F\circ u \in \mathcal{F}$ and $\mathcal{E}(F\circ u, F\circ u) \leqslant \mathcal{E}(u,u)$. In particular, for any $u \in \mathcal{F}$, we have $\tilde{u} := u_+ \wedge 1 \in \mathcal{F}$ and $\mathcal{E}(\tilde{u},\tilde{u}) \leqslant \mathcal{E}(u,u)$. The subspace $\mathcal{F}$ is closed under pointwise multiplication (that is, an algebra). More precisely, we have for all $u,v \in \mathcal{F}$, we have $uv \in \mathcal{F}$ and (denoting by $\|\cdot\|_\infty$, the sup norm)

$$\mathcal{E}(uv,uv) \leqslant 2\|u\|_\infty^2\mathcal{E}(v,v) + 2\|v\|_\infty^2\mathcal{E}(u,u). \tag{3.17}$$

The estimate (3.17) follows from showing an analogous estimate for $\mathcal{E}^{(m)}$ for all $m \in \mathbb{N}$ using the elementary inequality

$$((uv)(x)-(uv)(y))^2 \leqslant 2u(x)^2(v(x)-v(y))^2 + 2v(y)^2(u(x)-u(y))^2, \quad \text{for all } x,y \in K.$$

Next, we show that $\mathcal{F}$ is dense subspace of the Banach space $(C(K), \|\cdot\|_\infty)$. Since $\mathcal{F}$ is an algebra, by the Stone-Weierstrass theorem (cf. [Fol, Corollary 4.50]), it suffices to show that $\mathcal{F}$ separates points and that for any $x \in K$, there exists $v \in \mathcal{F}$ such that $v(x) \neq 0$. There it suffices to show the following: for any non-empty finite subset $V \subset K$, we have

$$\{u\big|_V : u \in \mathcal{F}\} = \mathbb{R}^V.$$

Since the inclusion $\{u\big|_V : u \in \mathcal{F}\} \subset \mathbb{R}^V$ is trivial, it suffices to show that $\{u\big|_V : u \in \mathcal{F}\} \supset \mathbb{R}^V$. To this end, it suffices to consider V with $\#V \geqslant 2$ as $\mathbb{R}\mathbb{1}_K \subset \mathcal{F}$. Let $g \in \mathbb{R}^V$ be arbitrary. Choose $m \in \mathbb{N}$ such that $2^{1-m} < \min_{\substack{x,y \in V, \\ x \neq y}} |x-y|$. Then for any $x, y \in V$ with $x \neq y$ and for any $v, w \in W_m$ with $x \in F_v(K), y \in F_w(K)$, we have

$$F_v(K) \cap F_w(K) = \varnothing. \tag{3.18}$$

For each $x \in V$, pick $w_x \in W_m$ such that $x \in F_{w_x}(K)$. Define $h \in \mathbb{R}^{V_m}$ such that for all $x \in V$ and for all $z \in V_m \cap F_{w_x}(K)$, we have $h(z) = g(x)$, and h is arbitrarily defined at other vertices (such a function h exists due to (3.18)). There exists $u \in \mathcal{F}$ such that $u\big|_{V_n} = H_{n-1,n} \circ \cdot \circ H_{m.m+1} h$. It is easy to see that $u\big|_{F_{w_x}(K)} \equiv g(x)$ for all $x \in V$ and hence $u\big|_V = g$. This concludes the proof that in the Banach space $(C(K), \|\cdot\|_\infty)$, we have

$$\overline{\mathcal{F}}^{\|\cdot\|_\infty} = C(K). \tag{3.19}$$

We summarize the construction of Dirichlet form for Brownian motion on the Sierpiński gasket below.

Proposition 3.7. *Let m be a Radon probability measure on the Sierpiński gasket K with full support. Then $(\mathcal{E}, \mathcal{F})$ is a strongly local, regular, Dirichlet form on $L^2(K, m)$.*

Proof. Since $C(K)$ is dense in $L^2(K, m)$, by (3.19), we have that $\mathcal{F}$ is a dense subspace of $L^2(K, m)$. The bi-linearity, symmetry, non-negative definiteness and Markov property of $(\mathcal{E}, \mathcal{F})$ follow from the corresponding properties for $\mathcal{E}^{(m)}$. Next, let us verify that $(\mathcal{E}, \mathcal{F})$ is closed. To this end, let $(f_n)_{n \in \mathbb{N}}$ be an $\mathcal{E}_1$-Cauchy sequence. By (3.8), for all $x, y \in K$ and for all $g \in \mathcal{F}$, we have (denoting every $g \in \mathcal{F}$ by its continuous version)

$$|g(x)|^2 \leqslant 2(|g(y)|^2 + 400\mathcal{E}(g,g))$$

By averaging over y with respect to m, we have

$$|g(x)|^2 \leqslant 2(\int_K |g(y)|^2 \, dm + 400\mathcal{E}(g,g)) \leqslant 800\mathcal{E}_1(g,g), \quad \text{for all } x \in K, g \in \mathcal{F}. \tag{3.20}$$

By (3.20) for each $x \in K$, $(f_n(x))_{n \in \mathbb{N}}$ is a Cauchy sequence and hence $\lim_{n \to \infty} f_n(x) = f(x)$ exists. Let $\epsilon > 0$ and $n \in \mathbb{N}$ be arbitrary. Then exists $N \in \mathbb{N}$ such that for all $k, l \in \mathbb{N}$ with $k \wedge l \geqslant N$, we have

$$\mathcal{E}^{(n)}\left((f_k - f_l)\big|_{V_n}, (f_k - f_l)\big|_{V_n}\right) \leqslant \mathcal{E}(f_k - f_l, f_k - f_l) \leqslant \mathcal{E}_1(f_k - f_l, f_k - f_l) \leqslant \epsilon, \quad \text{for all } n \in \mathbb{N}.$$

By letting $l \to \infty$ and using the pointwise convergence of f_l to f, we obtain

$$\sup_{n\in\mathbb{N}} \mathcal{E}^{(n)}\left((f_k - f)\big|_{V_n}, (f_k - f)\big|_{V_n} \right) \leqslant \epsilon.$$

Hence $f \in \mathcal{F}$ and $\lim_{k\to\infty} \mathcal{E}(f_k - f, f_k - f) = 0$.

It remains to show that f_k converges to f in $L^2(K, m)$. By (3.20) and $\sup_{n\in\mathbb{N}} \mathcal{E}_1(f_n, f_n) < \infty$, we have

$$\sup_{n\in\mathbb{N}} \sup_{x\in K} |f_n(x)| < \infty.$$

Combining this with (3.8) and $\sup_{n\in\mathbb{N}} \mathcal{E}_1(f_n, f_n) < \infty$, we have that the sequence $(f_n)_{n\in\mathbb{N}}$ is uniformly boundeded and equicontinuous. Hence by the Arzela-Ascoli theorem, $(f_n)_{n\to\infty}$ converges to f in the sup norm and hence $f_n \xRightarrow[L^2(K,m)]{k\to\infty} f$. Therefore $(\mathcal{E}_1, \mathcal{F})$ is a Hilbert space. The regularity of $(\mathcal{E}, \mathcal{F})$ follows from (3.19) and $\mathcal{F} \subset C(K)$.

Let us verify the strong locality property. Let $f, g \in \mathcal{F}$ and $a \in \mathbb{R}$ be such that $\operatorname{supp}_m(f - a\mathbb{1}_K) \cap \operatorname{supp}_m(g) = \varnothing$. Then there exists $n \in \mathbb{N}$ such that

$$\operatorname{dist}(\operatorname{supp}_m(f - a\mathbb{1}_K), \operatorname{supp}_m(g)) = \inf_{\substack{x\in\operatorname{supp}_m(f-a\mathbb{1}_K),\\ y\in\operatorname{supp}_m(g)}} |x - y| > 2^{-n}.$$

Therefore for all $w \in W_n$, we have either $F_w(K) \cap \operatorname{supp}_m(f - a\mathbb{1}_K) = \varnothing$ or $F_w(K) \cap \operatorname{supp}_m(g) \neq \varnothing$ (or possibly both). Thus by the self-similarity property (3.15) we have

$$\mathcal{E}(f, g) = \sum_{w\in W_n} \left(\frac{5}{3}\right)^n \mathcal{E}(f \circ F_w, g \circ F_w) = 0.$$

□

Our next goal is to obtain the existence of heat kenrel for Brownian motion on the Sierpinski gasket. To this end, we need the following general notion of energy measures. The energy measure of a function f can be viewed as the generalization of the measure $A \mapsto \int_A |\nabla f(x)|^2\, dx$.

Definition 3.8 (Energy measure; [FOT, (3.2.13), (3.2.14) and (3.2.15)])**.** Let $(\mathcal{E}, \mathcal{F})$ be a strongly local, regular, Dirichlet form on $L^2(X, m)$. The $\mathcal{E}$*-energy measure* $\Gamma(f, f)$ of $f \in \mathcal{F}$ is defined, first for $f \in \mathcal{F} \cap L^\infty(X, m)$ as the unique ($[0, \infty]$-valued) Radon measure on X such that

$$\int_X g\, d\Gamma(f, f) = \mathcal{E}(f, fg) - \frac{1}{2}\mathcal{E}(f^2, g) \qquad \text{for all } g \in \mathcal{F} \cap C_c(X), \tag{3.21}$$

next by $\Gamma(f, f)(A) := \lim_{n\to\infty} \Gamma\big((-n) \vee (f \wedge n), (-n) \vee (f \wedge n)\big)(A)$ for each $A \in \mathcal{B}(X)$ for $f \in \mathcal{F}$.

By [FOT, (3.2.13) and (3.2.14)], we have the triangle inequality for energy measure

$$\left|\sqrt{\Gamma(f, f)(B)} - \sqrt{\Gamma(g, g)(B)}\right|^2 \leqslant \Gamma(f - g, f - g)(B) \leqslant \mathcal{E}(f - g, f - g), \tag{3.22}$$

for all Borel sets B and for all $f, g \in \mathcal{F}$.

The self-similarity property of energy (3.15) along with the definition of energy measure leads to the self-similarity of the corresponding energy measures.

Exercise 3.9. Show that for the Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(K, m)$ in Proposition 3.7, the corresponding energy measure satisfies

$$\Gamma(f, f) = \left(\frac{5}{3}\right)^m \sum_{w \in W_m} (F_w)_*(\Gamma(f \circ F_w, f \circ F_w)), \quad \text{for all } m \in \mathbb{N}, \text{ and } f \in \mathcal{F}. \tag{3.23}$$

On the Sierpiński gasket, let $B(x, r)$ denote the open ball centered at $x \in K$ with radius $r > 0$ with respect to the Euclidean metric. We show the following *Poincaré inequality* on the Sierpiński gasket.

Proposition 3.10. *Let $(\mathcal{E}, \mathcal{F})$ be the Dirichlet form on $L^2(K, m)$ in Proposition 3.7. Set $d_w = \log 5/\log 2$. There exists $C > 0, A \geqslant 1$ such that for all $x \in K, 0 < r \leqslant 1$, $f \in \mathcal{F}$, we have*

$$\int_{B(x,r)} \left|f(y) - f_{B(x,r)}\right|^2 m(dy) \leqslant C r^{d_w} \int_{B(x,Ar)} d\Gamma(f, f), \tag{3.24}$$

where $f_{B(x,r)} = \frac{1}{m(B(x,r))} \int_{B(x,r)} f(z)\, m(dz)$.

Proof. Since $m(F_w(K)) = 3^{-m}$ and $\operatorname{diam}(F_w(K)) = 2^{-m}$ for all $w \in W_m, m \in \mathbb{N}$, there exists $C_1 > 0$ such that

$$C_1^{-1} r^{d_f} \leqslant m(B(x, r)) \leqslant C_1 r^{d_f}, \quad \text{for all } x \in K, 0 < r \leqslant 1. \tag{3.25}$$

Therefore for all $x \in K, 0 < r \leqslant 1$, $f \in \mathcal{F}$,

$$\int_{B(x,r)} \left|f(y) - f_{B(x,r)}\right|^2 m(dy) \leqslant C_1 r^{d_f} \sup_{y,z \in B(x,r)} |f(y) - f(z)|^2. \tag{3.26}$$

Fix any $y, z \in B(x, r)$ and $f \in \mathcal{F}$. Let $n \in \mathbb{N} \cup \{0\}$ be largest integer such that there exist $v, w \in W_n$ such that $y \in F_v(K), z \in F_w(K)$ and $F_v(K) \cap F_w(K) \neq \varnothing$. Let $q \in F_v(K) \cap F_w(K)$, $v \neq w$. Similar to the proof of Proposition 3.6 (see (3.10)), we estimate

$$\begin{aligned} |f(y) - f(z)|^2 &\leqslant 2\left(|f(y) - f(q)|^2 + |f(z) - f(q)|^2\right) \\ &\overset{(3.8)}{\leqslant} 800\left(\mathcal{E}(f \circ F_v, f \circ F_v) + \mathcal{E}(f \circ F_w, f \circ F_w)\right) \\ &\overset{(3.23)}{\leqslant} 800\left(\frac{3}{5}\right)^n \Gamma(f, f)\left(F_v(K) \cup F_w(K)\right) \end{aligned} \tag{3.27}$$

As explained in the proof of Proposition 3.6, $\frac{\sqrt{3}}{2} 2^{-n-1} \leqslant |y - z| < 2r$ and hence by (3.26) and (3.27), we obtain the desired Poincaré inequality (3.24) with $A = 1 + \frac{8}{\sqrt{3}}$ as

$$F_v(K) \cup F_w(K) \subset B(x, r + 2^{-n}) \subset B(x, (1 + 8/\sqrt{3})r).$$

□

For a function $f \in L^1(K,m)$ and $r > 0$, by $f_r : K \to \mathbb{R}$ we denote the function

$$f_r(x) := \fint_{B(x,r)} f\,dm = \frac{1}{m(B(x,r))}\int_{B(x,r)} f\,dm, \quad \text{for all } x \in K.$$

The following estimate is called the pseudo-Poincaré inequality. The difference from Poincaré inequality is that the integrals involved are global.

Lemma 3.11. *There exists $C > 0$ such that for all $f \in \mathcal{F}, r > 0$, we have*

$$\int_K |f(x) - f_r(x)|^2\, m(dx) \leqslant C r^{d_w} \mathcal{E}(f,f),$$

where $d_w := \log_2 5$.

Proof. It suffices to assume $0 < r < 1$ as the case $r \geqslant 1$ follows from Proposition 3.10. By Jensen's inequality and the volume estimate (3.25), we have

$$\begin{aligned}\int_K |f(x) - f_r(x)|^2\, m(dx) &\lesssim \int_K \fint_{B(x,r)} |f(x) - f(y)|^2\, m(dy)\, m(dx) \\ &\lesssim r^{-d_f}\int_K\int_K |f(x) - f(y)|^2 \mathbb{1}_{\{d(x,y)\leqslant r\}}\, m(dy)\, m(dx) \qquad (3.28)\end{aligned}$$

Let N denote a r-net (a maximal r-separated subset; any two distinct points in N are at least distance r apart and any set that strictly contains N is not r-separated). The maximality of N implies that

$$\sum_{n\in N} \mathbb{1}_{B(n,r)} \geqslant \mathbb{1}_K, \quad \sum_{n\in N} \mathbb{1}_{B(n,2r)}(x)\mathbb{1}_{B(n,2r)}(y) \leqslant \mathbb{1}_{\{d(x,y)\leqslant r\}}, \quad \text{for all } x, y \in K.$$

The balls $B(n, 2Ar)$, where A is the constant in (3.24) do not overlap too much in the sense that $\sum_{n\in N} \mathbb{1}_{B(n,2Ar)} \lesssim \mathbb{1}_K$ (due to the volume estimate (3.25)). Hence by (3.28), for all $f \in \mathcal{F}$, we have

$$\begin{aligned}\int_K |f(x) - f_r(x)|^2\, m(dx) &\lesssim \int_K \fint_{B(x,r)} |f(x) - f(y)|^2\, m(dy)\, m(dx) \\ &\lesssim r^{-d_f}\int_K\int_K |f(x) - f(y)|^2 \mathbb{1}_{\{d(x,y)\leqslant r\}}\, m(dy)\, m(dx) \\ &\lesssim r^{-d_f}\sum_{n\in N}\int_{B(n,2r)}\int_{B(n,2r)} |f(x) - f(y)|^2\, m(dy)\, m(dx) \\ &\lesssim \sum_{n\in N}\int_{B(n,2r)} \left|f(x) - f_{B(n,2r)}\right|^2 m(dx) \\ &\overset{(3.24)}{\lesssim} r^{d_w}\sum_{n\in N}\int_K \mathbb{1}_{B(n,2Ar)}\, d\Gamma(f,f) \lesssim r^{d_w}\mathcal{E}(f,f). \qquad (3.29)\end{aligned}$$

□

We obtain a Nash inequality for the Brownian motion on the Sierpiński carpet.

Proposition 3.12. *There exists* $C_1, C_2 \in (0,\infty)$ *such that*

$$\|f\|_2^{1+n/2} \leqslant \|f\|_1 \left(C_1 \|f\|_2^2 + C_2 \|\mathcal{E}(f,f)\|\right)^{n/4}, \quad \textit{for all } f \in L^1(K,m) \cap \mathcal{F},$$

where $n = 2d_f/d_w = 2\log 5/\log 3$.

Proof. Let $0 \leqslant r \leqslant 1$ and $f \in L^1(K,m) \cap \mathcal{F}$. Then by Lemma 3.11,

$$\|f\|_2 \leqslant \|f - f_r\|_2 + \|f_r\|_2 \lesssim r^{d_w/2}\mathcal{E}(f,f)^{1/2} + \|f_r\|_2 . \tag{3.30}$$

By Cauchy-Schwarz inequality, for all $f \in L^1(X,m), r \in (0,1]$, we have

$$\|f_r\|_2^2 \leqslant \|f_r\|_\infty \|f_r\|_1 \lesssim r^{-d_f} \|f\|_1 \|f_r\|_1 .$$

By Jensen's inequality, we have

$$\|f_r\|_1 \leqslant \int f_r(x)\, m(dx) \lesssim r^{-d_f} \int_K \int_K f(y) \mathbb{1}_{\{d(x,y)<r\}}\, m(dy)\, m(dx) \asymp \int_K f(y)\, m(dy) \leqslant \|f\|_1 .$$

Combining the above tow estimates, we obtain

$$\|f_r\|_2 \lesssim r^{-d_f/2} \|f\|_1 , \quad \text{for all } f \in L^1(K,m) \text{ and } r \in (0,1].$$

Thus by (3.30), we obtain

$$\|f\|_2 \leqslant \|f - f_r\|_2 + \|f_r\|_2 \lesssim r^{d_w/2}\mathcal{E}(f,f)^{1/2} + r^{-d_f/2} \|f\|_1 , \quad \text{for all } f \in \mathcal{F} \cap L^1(K,m) \text{ and } r \in (0,1].$$

Optimizing over $r \in (0,1]$ yields the desired estimate. □

Hence the using the Nash inequality above, we obtain the existence of heat kernel and an upper bound using Proposition 3.3.

Proposition 3.13. *Let* m *denote the self-similar measure on the Sierpiński gasket* K *such that* $m(F_w(K)) = 3^{-k}$ *for all* $k \in \mathbb{N}$ *and* $w \in W_k$. *Then the Dirichlet form* $(\mathcal{E},\mathcal{F})$ *on* $L^2(K,m)$ *admits a heat kernel. There exists* $C > 0$ *such that*

$$p_t(x,y) \leqslant \frac{C}{t^{d_f/d_w}}, \quad \textit{for all } t \in (0,1], x,y \in K, \textit{ where } d_f = \log_2(3), d_w = \log_2(5).$$

As mentioned earlier in (1.4) sharp two-sided bounds on the heat kernel was obtained by Barlow and Perkins and is referred to as sub-Gaussian heat kernel bounds [BP].

References

[AM] S. Albeverio and Z.-M. Ma, A general correspondence between Dirichlet forms and right processes, *Bull. Amer. Math. Soc.* **26** (1992), 245–252

[BP] M. T. Barlow and E. A. Perkins, Brownian motion on the Sierpinski gasket, *Probab. Theory Related Fields* **79** (1988), no. 4, 543–623.

[BH91] N. Bouleau, F. Hirsch, Dirichlet forms and analysis on Wiener space. *De Gruyter Stud. Math.*, **14** *Walter de Gruyter & Co., Berlin*, 1991. x+325 pp.

[CS] L. Caffarelli, L. Silvestre. An extension problem related to the fractional Laplacian. *Comm. Partial Differential Equations* **32** (7) (2007) 1245–1260.

[CF] Z.-Q. Chen and M. Fukushima, *Symmetric Markov Processes, Time Change, and Boundary Theory*, London Math. Soc. Monogr. Ser., vol. 35, Princeton University Press, Princeton, NJ, 2012.

[Dav] E. B. Davies, *Linear Operators and their Spectra* Cambridge studies in advanced mathematics 106, Cambridge University Press, Cambridge, 2007.

[DM] E. B. Davies, N. Mandouvalos, Heat kernel bounds on hyperbolic space and Kleinian groups,*Proc. London Math. Soc.*(3), **52** (1988) no.1, 182– 208.

[Doo] J. L. Doob, Boundary properties for functions with finite Dirichlet integrals. *Ann. Inst. Fourier* (Grenoble) **12** (1962), 573–621.

[DS] N. Dunford and J. T. Schwartz, *Linear Operators.* Part 1. Interscience Publishers, 1958.

[FOT] Fukushima, Oshima, Takaeda.

[Fol] G. B. Folland, *Real analysis* Pure Appl. Math. (N. Y.) Wiley-Intersci. Publ. John Wiley & Sons, Inc., New York, 1999, xvi+386 pp.

[Lax] P. D. Lax, *Functional analysis.* Pure Appl. Math. (N. Y.) Wiley-Interscience [John Wiley & Sons], New York, 2002. xx+580 pp.

[HR] M. Hino, J. A. Ramírez. Small-time Gaussian behavior of symmetric diffusion semi-groups. *Ann. Probab.* **31** (2003) 1254–1295.

[MO] S. A. Molchanov, E. Ostrowski. Symmetric stable processes as traces of degenerate diffusion processes. *Theor. Prob. Appl.* **14** 128–131, 1969.

[Nor] J. R. Norris. Heat kernel asymptotics and the distance function in Lipschitz Riemannian manifolds. *Acta Math.* **179** (1997) 79–103.

[RS] M. Reed, B. Simon, *Methods of modern mathematical physics. I. Functional analysis*, Academic Press, New York-London, 1972, xvii+325 pp.

[Spi] F. Spitzer. Some theorems concerning 2-dimensional Brownian motion. *Trans. Amer. Math. Soc.* **87** (1958) 187–197.

[Var] S. R. S. Varadhan, On the behaviour of the fundamental solution of the heat equation with variable coefficients, *Comm. Pure Appl. Math.*, **20** (1967), 431–455.