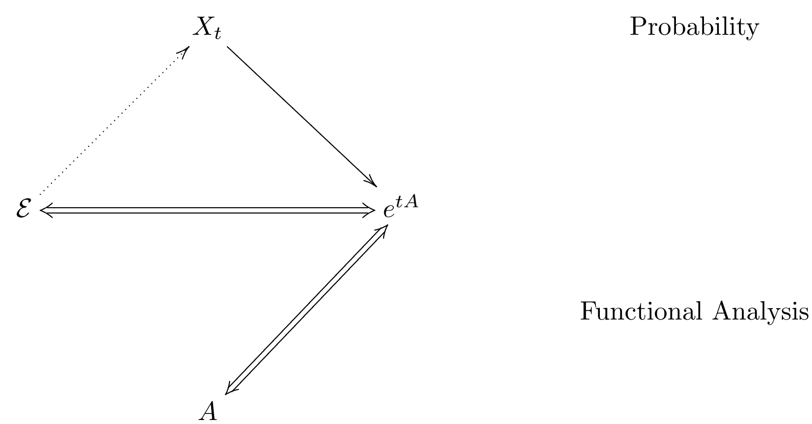
Self-adjoint operators, semigroups, Dirichlet forms and Hunt processes

Evan Randles

May 3, 2013



Contents

1 Introduction **1**

2 Functional Analysis **1**
2.1 Semigroups and their infinitesimal generators 1
2.2 The Hille-Yosida Theorem 8
2.3 The Hille-Yosida theorem for self-adjoint operators 11
2.4 Symmetric forms and self-adjoint operators 16
2.5 Dirichlet forms and Markovian semigroups 19

3 Probability **25**
3.1 Hunt Processes 25
3.2 Hunt Processes to Dirichlet forms 27
3.3 Some fun facts 29

A Appendix **30**

1 Introduction

This article is about the diagram on the title page. Our goal is to explore the connection between semigroups, self-adjoint operators, Dirichlet forms and Hunt Processes. This connection is the idea behind the probabilistic/diffusive interpretation of the heat equation. Every arrow in the diagram is shown below except for the dotted one, this is [6, Theorem 6.2.1]. A less general version can be found in [1].

2 Functional Analysis

The aim of this section is to explore the interplay between semigroups, self-adjoint operators and Dirichlet forms. Our first goal is to establish the connection between semigroups and closed, densely defined operators on Banach spaces; this is the Hille-Yosida Theorem. Our presentation of the Hille-Yosida Theorem comprises the first two subsections below and, to some extent, follows [11]. After this, we move into the setting of Hilbert spaces wherein we will prove a refinement of the Hille-Yosida theorem for self-adjoint operators and then show that there is a one to one correspondence between closed symmetric forms and self-adjoint operators. Finally, we discuss Dirichlet forms and Markovian semigroups.

2.1 Semigroups and their infinitesimal generators

In this subsection and the next, our setting is a Banach space X with norm $\|\cdot\|$. Unless otherwise stated, convergence means strong convergence. Before introducing semigroups and their generators, we recall a few standard definitions, c.f. [7, 11].

Definition 1. *Suppose that X is a Banach space and $D(A) \subseteq X$ is a linear subspace of X. By a linear operator A on X with domain $D(A)$, we mean a function $A : D(A) \to X$ that is $\mathbb{C}$-linear. We shall say that A is densely defined if $D(A)$ is a dense subset of X with respect to the norm topology.*

Definition 2 (Closed operator)**.** *A linear operator C on X with domain $D(C)$ is said to be closed if for all $\{x_n\}_n \subseteq D(C)$ such that*

$$x_n \to x \qquad \text{and} \qquad Cx_n \to y$$

as $n \to \infty$, we have

$$x \in D(C) \qquad \text{and} \qquad Cx = y.$$

We shall denote by $\mathcal{B}(X)$ the set of (everywhere defined) bounded linear operators on X and, as usual, $I \in \mathcal{B}(X)$ is the identity operator. The operator norm on $\mathcal{B}(X)$ is denoted by $\|\cdot\|_{op}$.

Definition 3 (Spectrum and resolvent set)**.** *Suppose that A is a linear operator with domain $D(A)$. The resolvent set of A is the set $\rho(A)$ consisting of all $\lambda \in \mathbb{C}$ for which $(\lambda I - A) : D(A) \to X$ is bijective and such that $(\lambda I - A)^{-1} \in \mathcal{B}(X)$. The spectrum of A denoted by $\sigma(A)$ is the complement of the resolvent set.*

Definition 4 (Operator extension)**.** *Suppose A and B are linear operators on X with respective domains $D(A)$ and $D(B)$. We say that B is an extension of A and write $A \subset B$ if $D(A) \subseteq D(B)$ and $Bx = Ax$ for all $x \in D(A)$.*

Definition 5 (Semigroup)**.** *A family $\{T_t\}_{t\geqslant 0}$ of bounded and everywhere defined operators on X is called a semigroup if*

i. $T_0 = I$,

ii. for all $t, s \geqslant 0$,

$$T_t T_s = T_s T_t = T_{t+s}$$

and

iii. for each $x \in X$,

$$\lim_{t\downarrow 0} \|T_t x - x\| = 0.$$

Property *ii* above is often referred to as the *semigroup property.* In the literature, property *iii* is not ubiquitously included in the definition of semigroup. When it is omitted, our definition coincides with that of a strongly continuous semigroup [6]. We shall not have use for such a distinction.

Definition 6 (Infinitesimal generator of a semigroup)**.** *Let $\{T_t\}_{t\geqslant 0}$ be a semigroup on X. Define*

$$A_t = \frac{1}{t}(T_t - I)$$

for $t > 0$,

$$D(A) = \left\{ x \in X : \lim_{t\downarrow 0} A_t x \text{ exists} \right\}$$

and put

$$Ax = \lim_{t\downarrow 0} A_t x \text{ whenever } x \in D(A).$$

The map $A : D(A) \to X$ is called the infinitesimal generator of the semigroup $\{T_t\}_{t\geqslant 0}$.

Remark 1. *The limit defining A is to be understood in the sense of the norm topology on X. More explicitly, for each $x \in D(A)$, there exists $Ax \in X$ for which*

$$\lim_{t\downarrow 0} \|A_t x - Ax\| = \lim_{t\downarrow 0} \left\| \frac{T_t x - x}{t} - Ax \right\| = 0.$$

Because A_t is a linear operator for each $t > 0$ it follows that A is a linear operator with domain $D(A)$.

Proposition 1 (Basic semigroup facts)**.** *Let $\{T_t\}_{t\geqslant 0}$ be a semigroup on X.*

1. *There are constants $C \geqslant 1$ and $\gamma \geqslant 0$ such that*
$$\|T_t\|_{op} \leqslant Ce^{t\gamma}$$
for all $t \geqslant 0$.

2. *For each $x \in X$, the map $t \to T_t x$ from $[0,\infty)$ into X is continuous.*

Proof. 1. First we establish the existence of C. Suppose that for some sequence of non-negative real numbers $t_n \to 0$ we have $\|T_{t_n}\|_{op} \to \infty$. Then by the uniform boundedness principle, Fact 9 of the appendix, there is $x \in X$ for which
$$\lim_{n\to\infty} \|T_{t_n} x\| = \infty.$$
This cannot be true in view of Property *iii* of Definition 5. Consequently, there must be $C \geqslant 1$ and $\delta > 0$ for which
$$\|T_t\|_{op} \leqslant C \tag{1}$$
for all $t \in [0,\delta]$. Using the semigroup property, it follows that for any $t \geqslant 0$ and natural number n,
$$T_t = T_{nt/n} = (T_{t/n})^n$$
and therefore
$$\|T_t\|_{op} \leqslant \|T_{t/n}\|_{op}^n. \tag{2}$$
So for any $t \in [0,\infty)$ choose a natural number n for which $(n-1)\delta \leqslant t < n\delta$. Combining (1) and (2) we have
$$\|T_t\|_{op} \leqslant \|T_{t/n}\|_{op}^n \leqslant C^n = CC^{n-1} \leqslant CC^{t/\delta} = Ce^{\gamma t}$$
where $\gamma = (\log(C))/\delta \geqslant 0$. This proves the first part of the proposition.

For the second, observe that for any $x \in X$, $t \in [0,\infty)$ and $h > 0$,
$$\|T_{t+h}x - T_t x\| = \|T_t(T_h x - x)\| \leqslant Ce^{\gamma t}\|T_h x - x\|$$
where we have used the semigroup property. By an appeal to Property *iii*. of Definition 5, the proof is complete. □

Definition 7. *Suppose that $\{T_t\}_{t\geqslant 0}$ is a semigroup on X. In view of Proposition 1, let $\gamma \geqslant 0$ be the minimal such constant for which conclusion* 1 *of the proposition holds. We then say that $\{T_t\}$ is γ-contractive.*

We say that $\{T_t\}$ is a contraction semigroup if it is 0*-contractive and $C = 1$. That is, $\{T_t\}$ is a contraction semigroup if*
$$\|T_t\|_{op} \leqslant 1 \;\; \textit{for all } t \geqslant 0.$$

To introduce the next example, we recall a standard definition from linear algebra: For $A \in \mathcal{B}(X)$, the *exponential of* A is the operator $e^A = \exp(A) \in \mathcal{B}(X)$ defined by

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

The defining series is the limit of its corresponding partial sums in sense of the operator norm.

Example 1. *If $A \in \mathcal{B}(X)$,*

$$T_t = e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \tag{3}$$

defines a semigroup with infinitesimal generator A where $D(A) = X$. It is easily shown that

$$\lim_{t \downarrow 0} \|T_t - I\|_{op} = 0 \tag{4}$$

from which Property iii of Definition 5 follows trivially. The constants C and γ of the above proposition are 1 and $\|A\|_{op}$ respectively.

The above example comprises the class of semigroups for which (4) holds, c.f. [7, p. 66]. As we shall see, (4) fails precisely when $D(A) \neq X$. We can however say the following:

Proposition 2 (Properties of semigroups and their infinitesimal generators)**.** *Let $\{T_t\}_{t \geqslant 0}$ be a semigroup on X with infinitesimal generator A.*

1. *$D(A)$ is a dense linear subspace of X.*

2. *A is a closed operator.*

3. *For each $x \in D(A)$,*

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

Proof. It is clear the $D(A)$ is a linear subspace of X. The rest of this proof relies on one crucial idea: Averages play well with derivatives. For $x \in X$ and $t > 0$ define

$$M_t x = \frac{1}{t} \int_0^t T_s x ds$$

where the integral exists in the sense of the norm-limit of Riemann sums. Note that $M_t x \in X$. As in the standard real-valued construction of the Riemann integral, the same arguments show that $M_t \in \mathcal{B}(X)$ and $\|M_t\|_{op} \leqslant \sup_{s \leqslant t} \|T_s\|_{op}$ for all $t > 0$. Thus by Proposition 1,

$$\|M_t\|_{op} \leqslant C e^{\gamma t}$$

where $C \geqslant 1$ and $\gamma \geqslant 0$. In fact, most of the basic properties of the Riemann integral hold in this setting and are verified by precisely the same arguments;

we will use them abundantly and without explicit mention. For the precise statements the reader is encouraged to see [4, pps 3. 21-22]. With this in mind,

$$\begin{aligned} \|M_t x - x\| &= \frac{1}{t}\left\|\int_0^t T_s x ds - \int_0^t x ds\right\| \\ &= \frac{1}{t}\left\|\int_0^t (T_s x - x) ds\right\| \\ &\leqslant \frac{t}{t}\sup_{s\leqslant t}\|T_s x - x\|. \end{aligned}$$

In view of Property *iii* of Definition 5 this proves that for any $x \in X$,

$$\lim_{t\downarrow 0} M_t x = x. \tag{5}$$

We claim that for $x \in X$ and $t, \epsilon > 0$,

$$A_t M_\epsilon x = A_\epsilon M_t x. \tag{6}$$

To see this observe that

$$\int_\epsilon^{t+\epsilon} T_s x ds - \int_0^t T_s x ds = \int_t^{t+\epsilon} T_s x ds - \int_0^\epsilon T_s x ds$$

and so

$$\begin{aligned} \epsilon t A_\epsilon M_t x &= (T_\epsilon - I)\int_0^t T_s x ds = \int_0^t T_\epsilon T_s x ds - \int_0^t T_s x ds \\ &= \int_0^t T_{\epsilon+s} x ds - \int_0^t T_s x ds = \int_\epsilon^{t+\epsilon} T_s x ds - \int_0^t T_s x ds \\ &= \int_t^{t+\epsilon} T_s x ds - \int_0^\epsilon T_s x ds = \int_0^\epsilon T_t T_s x ds - \int_0^\epsilon T_s x ds \\ &= (T_t - I)\int_0^\epsilon T_s x ds = t\epsilon A_t M_\epsilon. \end{aligned}$$

This proves (6). Let us point out that in the above computation the operator T_ϵ was passed under the integral sign. This operation is justified whenever the operator is closed, see [5, Exercise 7.5.13]. Now for fixed $\epsilon > 0$, (5) and (6) show that

$$\lim_{t\downarrow 0} A_t M_\epsilon x = \lim_{t\downarrow 0} A_\epsilon M_t x = A_\epsilon x$$

and so $M_\epsilon x \in D(A)$ and

$$AM_\epsilon x = A_\epsilon x. \tag{7}$$

Thus any $x \in X$ is an accumulation point of the collection $\{M_t x\}_{t>0} \subseteq D(A)$. Hence, $D(A)$ is dense in X.

Before demonstrating that A is a closed operator we make an observation. For any $t, s > 0$, it is clear that M_t and A_s commute. We observe that for any $x \in D(A)$, $M_t x \in D(A)$ and so

$$AM_t x = \lim_{s\downarrow 0} A_s M_t x = \lim_{s\downarrow 0} M_t A_s x = M_t A x.$$

Thus for each $t > 0$, A and M_t commute.

Now we show that A is closed. Take $\{x_n\}_n \subseteq D(A)$ for which

$$x_n \to x \qquad \text{and} \qquad Ax_n \to y$$

as $n \to \infty$. Using (7) and the observation made in the previous paragraph,

$$A_s x = \lim_{n\to\infty} A_s x_n = \lim_{n\to\infty} AM_s x_n = \lim_{n\to\infty} M_s A x_n = M_s y$$

for each $s > 0$. Thus by (5), $\lim_s A_s x = y$ which proves that $x \in D(A)$ and $Ax = y$.

We now prove item 3. First observe that for $x \in D(A)$,

$$\frac{d}{dt} T_t x = \lim_{h\downarrow 0} \frac{T_{t+h}x - T_t x}{h} = T_t \lim_{h\downarrow 0} \frac{T_h x - x}{h} = T_t \lim_{h\downarrow 0} A_h x = T_t A x$$

where we have used the fact that $T_t \in \mathcal{B}(X)$. It remains to show that $AT_t x = T_t A x$ for $x \in D(A)$ and $t \geqslant 0$. This is easy to see; it follows by precisely the same argument we used to show that $M_t A x = AM_t x$ for $x \in D(A)$. The proposition is proved. $\square$

Definition 8 (Resolvent of a semigroup)**.** *Let $\{T_t\}_t$ be a γ-contractive semigroup on X. For $x \in X$ and $\lambda \in \mathbb{C}$ for which* $\operatorname{Re}\lambda > \gamma$ *put*

$$R_\lambda x = \int_0^\infty e^{-\lambda t} T_t x dt. \tag{8}$$

For each such λ, R_λ is called a resolvent (or resolvent operator) of the semigroup $\{T_t\}$.

The above integral is to be understood as an X-valued (improper) Riemann integral. Formally, it is the Laplace transform of the semigroup $\{T_t\}$. The following proposition deals, in particular, with its existence.

Proposition 3. *Suppose that $\{T_t\}_t$ is a γ-contractive semigroup. For each λ for which* $\operatorname{Re}(\lambda) > \gamma$*, $R_\lambda \in \mathcal{B}(X)$, $D(A) = R_\lambda(X)$ and $R_\lambda = (\lambda I - A)^{-1}$.*

Proof. For any $x \in X$ and λ with $\operatorname{Re}(\lambda) > \gamma$,

$$\begin{aligned}
\|R_\lambda x\| &= \left\|\int_0^\infty e^{-\lambda t} T_t x dt\right\| \leqslant \int_0^\infty \|e^{-\lambda t} T_t x\| dt \\
&\leqslant \|x\| \int_0^\infty e^{-\operatorname{Re}(\lambda) t} \|T_t\|_{op} dt \\
&\leqslant \|x\| C \int_0^\infty e^{-(\operatorname{Re}(\lambda)-\gamma)t} dt = \frac{C}{\operatorname{Re}(\lambda - \gamma)} \|x\| < \infty
\end{aligned}$$

where we have used Conclusion 1. of Proposition 1. Thus $R_\lambda \in \mathcal{B}(X)$.

Now for $x \in X$ and $\epsilon > 0$, the semigroup property guarantees that

$$\begin{aligned}
\epsilon A_\epsilon R_\lambda x &= \int_0^\infty e^{-\lambda t} T_{t+\epsilon} x dt - \int_0^\infty e^{-\lambda t} T_t x dt \\
&= \int_\epsilon^\infty e^{\lambda\epsilon} e^{-\lambda t} T_t x dt - \int_0^\infty e^{-\lambda t} T_t x dt \\
&= (e^{\lambda\epsilon} - 1) R_\lambda x - e^{\lambda\epsilon} \int_0^\epsilon e^{-\lambda t} T_t x dt \\
&= (e^{\lambda\epsilon} - 1) R_\lambda x - e^{\lambda\epsilon} \epsilon M_\epsilon x + e^{\lambda\epsilon} \int_0^\epsilon (1 - e^{\lambda t}) T_t x dt.
\end{aligned}$$

By the same argument we used to show that $M_t x \to x$ as $t \to 0$, one sees that

$$\lim_{\epsilon \downarrow 0} \frac{e^{\lambda\epsilon}}{\epsilon} \int_0^\epsilon (1 - e^{\lambda t}) T_t x dt = 0$$

and therefore

$$\lim_{\epsilon \downarrow 0} A_\epsilon R_\lambda x = \lambda R_\lambda x - x.$$

This proves that $R_\lambda x \in D(A)$ and

$$(\lambda I - A) R_\lambda x = x.$$

Now, for any $x \in D(A)$ and λ for which $\mathrm{Re}(\lambda) > \gamma$,

$$\begin{aligned}
R_\lambda(\lambda I - A) x &= \int_0^\infty e^{-\lambda t} T_t (\lambda x - A x) dt \\
&= \lambda \int_0^\infty e^{-\lambda t} T_t x dt - \int_0^\infty e^{-\lambda t} T_t A x dt \\
&= \lambda \int_0^\infty e^{-\lambda t} T_t x dt - \int_0^\infty e^{-\lambda t} \frac{d}{dt} T_t x dt \\
&= e^{-\lambda 0} T_0 x = x
\end{aligned}$$

where we used Property 3. of Proposition 2 and partial integration. Therefore

$$D(A) = R_\lambda(X) \qquad \text{and} \qquad R_\lambda = (\lambda I - A)^{-1}$$

as desired. □

An immediate corollary is the following:

Corollary 1. *Suppose that $\{T_t\}_t$ is a γ-contractive semigroup on X. If A is the infinitesimal generator of $\{T_t\}$ then*

$$(\gamma, \infty) \subseteq \rho(A).$$

Remark 2. *The notion of a resolvent class of operators, $\{R_\lambda\} \subseteq \mathcal{B}(X)$ exists abstractly in its own right, i.e. it is defined without explicit use of semigroups. With it, one can define a corresponding infinitesimal generator A. This is the tack taken in [6] in the Hilbert space setting. Under certain conditions, it is seen that the infinitesimal generator of a resolvent class $\{R_\lambda\}$ is the infinitesimal generator of a semigroup $\{T_t\}$ on X and $\{R_\lambda\}$ and $\{T_t\}$ are then related via* (8)*. This equivalence, and that given by the Hille-Yosida theorem is summarized in the following diagram:*

$$\begin{tikzcd}
 & T_t \arrow[dl, Leftrightarrow] \arrow[dr, Leftrightarrow] & \\
A \arrow[rr, Leftrightarrow] & & R_\lambda
\end{tikzcd}$$

2.2 The Hille-Yosida Theorem

In the previous subsection, we started with a semigroup and derived from it a densely defined closed linear operator which satisfied some certain nice properties. In the present subsection we determine that exact class of densely defined linear operators that are generators of semigroups. This characterization, known as the Hille-Yosida Theorem, completes the first part of our diagram:

$$\begin{tikzcd}
 & & e^{tA} \\
 & \text{Hille-Yosida} \arrow[ur, Rightarrow] \arrow[dl, Rightarrow] & \\
A & &
\end{tikzcd}$$

Theorem 1 (Hille-Yosida)**.** *A closed, densely defined linear operator A on a Banach space X is the infinitesimal generator of a semigroup $\{T_t\}$ if and only if there are constants C and γ such that for every $\lambda > \gamma$, $(\lambda I - A) : D(A) \to X$ is invertible with*

$$\|(\lambda I - A)^{-m}\|_{op} \leqslant C(\lambda - \gamma)^{-m} \tag{9}$$

for all $m \in \mathbb{N}$.

Proof. If A is the infinitesimal generator of a semigroup $\{T_t\}$, we demonstrated in Propositions 1 and 3 that there are constants $C \geqslant 1$ and $\gamma \geqslant 0$ such that for all $\lambda > \gamma$, $(\lambda I - A)$ is invertible and (9) holds for $m = 1$. We also found that for any such λ and $x \in X$,

$$(\lambda I - A)^{-1}x = R_\lambda x = \int_0^\infty e^{-\lambda t} T_t x dt.$$

Using standard properties of the Laplace transform and the semigroup property it follows immediately that

$$R_\lambda^m x = \frac{1}{(m-1)!} \int_0^\infty t^{m-1} e^{-\lambda t} T_t x dt$$

for all $m \in \mathbb{N}$. Now by partial integration we have that

$$\|R_\lambda^m\|_{op} \leqslant \frac{C}{(m-1)!}\int_0^\infty t^{m-1}e^{(\gamma-\lambda)t}dt = C(\lambda-\gamma)^{-m}$$

from which (9) follows.

For the converse, set

$$S(\epsilon) = (I-\epsilon A)^{-1} \in \mathcal{B}(X).$$

In this notation (9) becomes

$$\|S(\epsilon)^m\|_{op} \leqslant C(1-\epsilon\gamma)^{-m} \tag{10}$$

for all $m \in \mathbb{N}$ and for all $0 < \epsilon < (1/\gamma)$. Moreover the hypothesis guarantees that

$$S(\epsilon)(I-\epsilon A)x = x \tag{11}$$

for all $x \in D(A)$ and

$$(I-\epsilon A)S(\epsilon)x = x \tag{12}$$

for all $x \in X$ whenever $0 < \epsilon < (1/\gamma)$. Rearranging things a bit we see that

$$\lim_{\epsilon\downarrow 0} S(\epsilon)x = x \tag{13}$$

for all $x \in D(A)$. Note however that estimate (10) guarantees that $\|S(\epsilon)\|_{op} \leqslant C$ for all sufficiently small ϵ and so by using the fact that $D(A)$ is dense in X we may immediately conclude that (13) holds for all $x \in X$.

Observe that by (12), for each $0 < \epsilon < (1/\gamma)$ the operator $AS(\epsilon) \in \mathcal{B}(X)$ and so we may define

$$T(t,\epsilon) = \exp(tAS(\epsilon)).$$

Using (10) and (12) we may write

$$\begin{aligned}
\|T(t,\epsilon)\|_{op} &= \|\exp(-(t/\epsilon)I)\exp((t/\epsilon)S(\epsilon))\|_{op} \\
&\leqslant e^{-t/\epsilon}\left\|\sum_{n=0}^\infty \frac{t^nS(\epsilon)^n}{\epsilon^n n!}\right\|_{op} \\
&\leqslant e^{-t/\epsilon}\sum_{n=0}^\infty \frac{t^nC(1-\epsilon\gamma)^{-n}}{\epsilon^n n!} \\
&\leqslant C\exp\left(\frac{\gamma t}{1-\epsilon\gamma}\right).
\end{aligned} \tag{14}$$

We also observe that for $\epsilon, \delta > 0$ and $x \in X$,

$$\begin{aligned}
T(t,\epsilon)x - T(t,\delta)x &= \int_0^t \frac{d}{ds}(T(s,\epsilon)T(t-s,\delta)x)ds \\
&= \int_0^t T(s,\epsilon)T(t-s,\delta)(S(\epsilon)-S(\delta))Axds.
\end{aligned}$$

Appealing to (13) with Ax in place of x, we see that by letting $\epsilon, \delta \to 0$,

$$\|T(t,\epsilon)x - T(t,\delta)x\| \to 0.$$

We may therefore define

$$T_t x = \lim_{\epsilon \downarrow 0} T(t,\epsilon)x \tag{15}$$

for all $x \in D(A)$. Using our estimate (14) and the hypothesis that $D(A)$ is dense, we deduce that this limit not only holds for all $x \in X$, but it holds uniformly for t in any compact subset of $[0,\infty)$ for all $x \in X$. Moreover we get the estimate

$$\|T_t\|_{op} \leqslant Ce^{\gamma t}. \tag{16}$$

It now follows from our definition of $T(t,\epsilon)$ and the convergence indicated above that T_t forms a semigroup on X. Let B denote the closed, densely defined infinitesimal generator of T_t. It only remains to show that $A = B$.

To this end, we first establish the equality:

$$T_t x - x = \int_0^t T_s Ax ds \tag{17}$$

for all $x \in D(A)$. Indeed, we know that for each $t, \epsilon > 0$ and $x \in D(A)$,

$$T(t,\epsilon)x - x = \int_0^t T(s,\epsilon)AS(\epsilon)xds = \int_0^t T(s,\epsilon)S(\epsilon)Axds \tag{18}$$

where we have used (11) and (12) in interchanging A and $S(\epsilon)$. Using (14), we observe that for any $s, \epsilon > 0$ and $y \in X$,

$$\begin{aligned} \|T(s,\epsilon)S(\epsilon)y - T_s y\| &\leqslant \|T(s,\epsilon)S(\epsilon)y - T(s,\epsilon)y\| + \|T(s,\epsilon)y - T_s y\| \\ &\leqslant C\exp(\gamma s/(1-\epsilon\gamma))\|S(\epsilon)y - y\| + \|T(s,\epsilon)y - T_s y\|. \end{aligned}$$

It follows from the above estimate and the fact that (15) holds uniformly on compact sets that

$$\lim_{\epsilon \downarrow 0} T(s,\epsilon)S(\epsilon)y = T_s y$$

also holds uniformly for s in any compact subset of $[0,\infty)$. So for any $x \in D(A)$,

$$\begin{aligned} T_t x - x &= \lim_{\epsilon \downarrow 0}(T(t,\epsilon)x - x) = \lim_{\epsilon \downarrow 0} \int_0^t T(s,\epsilon)S(\epsilon)Axds \\ &= \int_0^t \lim_{\epsilon \downarrow 0} T(s,\epsilon)S(\epsilon)Axds = \int_0^t T_s Axds \end{aligned}$$

as claimed. Recalling (5) and the notation from Proposition 2 it follows that for $x \in D(A)$,

$$Ax = \lim_{t \downarrow 0} M_t Ax = \lim_{t \downarrow 0} \frac{1}{t} \int_0^t T_s Axds = \lim_{t \downarrow 0} \frac{T_t - x}{t} = Bx.$$

We have proved that $A \subset B$ in the sense of operators. To complete the proof, it remains to show that $D(B) \subseteq D(A)$. To this end, select $x \in D(B)$. For sufficiently large λ we know that

$$(\lambda I - A) : D(A) \to X \text{ and } (\lambda I - B) : D(B) \to X$$

are invertible maps. For such a λ put

$$y = (\lambda I - B)x$$

and

$$\tilde{x} = (\lambda I - A)^{-1} y \in D(A).$$

Using the fact that B extends A it follows that

$$(\lambda I - B)\tilde{x} = (\lambda I - B)(\lambda I - A)^{-1} y = (\lambda I - A)(\lambda I - A)^{-1} y = y = (\lambda I - B)x.$$

Since $(\lambda I - B)$ is injective, $x = \tilde{x} \in D(A)$ whence $D(B) \subseteq D(A)$ and the theorem is proved. $\square$

Notation 1. *In light of the Hille-Yosida theorem and Example 1, for a densely defined operator A, its corresponding semigroup is henceforth denoted by* $\{e^{tA}\}_{t \geqslant 0}$.

The following theorem is the version of the Hille-Yosida theorem most often treated in textbooks [1, 5]. We do not state it as a corollary to the above theorem because it makes use of said theorem's proof.

Theorem 2 (Hille-Yosida for contraction semigroups)**.** *A closed, densely defined operator A on a Banach space X is the infinitesimal generator of a contraction semigroup $T_t (= e^{tA})$ if and only if*

$$(0, \infty) \subseteq \rho(A) \text{ and } \|(\lambda I - A)^{-1}\|_{op} \leqslant \frac{1}{\lambda} \text{ for all } \lambda > 0. \tag{19}$$

Proof. If A is the infinitesimal generator of a contraction semigroup $T_t = e^{tA}$, the result follows immediately from Proposition 2, Proposition 3 and Corollary 1.

Conversely, if (19) is satisfied for a closed, densely defined operator A, it is clear that (9) is satisfied for $\gamma = 0$ and $C = 1$. By an appeal to Theorem 1 there is a semigroup T_t of which A is its infinitesimal generator. Moreover, by (16)

$$\|T_t\|_{op} \leqslant 1e^{0t} = 1$$

and so T_t is contractive. $\square$

2.3 The Hille-Yosida theorem for self-adjoint operators

For the remainder of this article, we move into the setting of Hilbert spaces. In what follows, H will denote a separable real Hilbert space with inner product

$(\cdot,\cdot)$. For a densely defined operator $A : H \to H$ with domain $D(A)$ there is a unique closed operator $A^* : H \to H$, called the *adjoint* of A, with domain

$$D(A^*) = \{y \in H : x \to (Ax, y) \text{ is continuous from } D(A) \text{ to } \mathbb{R}\}$$

such that

$$(Ax, y) = (x, A^*y)$$

for all $x \in D(A)$ and $y \in D(A^*)$.

Definition 9 (Symmetric and self-adjoint)**.** *Let A and A^* be as above.*

- *A is called symmetric if $A \subset A^*$.*
- *A is called self-adjoint if $A = A^*$.*

We note that A is symmetric if and only if

$$(Ax, y) = (x, Ay)$$

for all $x, y \in D(A)$. Also, a symmetric operator A is self-adjoint if and only if $D(A^*) \subseteq D(A)$. Therefore everywhere defined symmetric operators are self-adjoint.

Definition 10. *Let $B : H \to H$ be densely defined with domain $D(B)$.*

1. *If $(Bx, x) \geqslant 0$ for all $x \in D(B)$, we call B non-negative and write $B \geqslant 0$.*
2. *If $(Bx, x) \leqslant 0$ for all $x \in D(B)$, we call B non-positive and write $B \leqslant 0$.*
3. *If $(Bx, x) \geqslant a$ for all $x \in D(B)$, we write $B \geqslant a$.*

The following facts are standard but a little too far afield for us. They can both be found in [11]:

Fact 1. *Let $A : H \to H$ be self-adjoint. Then $A \leqslant 0$ if and only if $\sigma(A) \subseteq (-\infty, 0]$.*

Fact 2. *If $T : H \to H$ is self-adjoint and injective then the range of T is dense in H and T^{-1} is self-adjoint.*

We now address that so-called Hille-Yosida theorem for self-adjoint operators, it corresponds, partly, to [6, Lemmas 1.3.1 and 1.3.2]. The proof in [6] makes use of the projection-valued measure form of the spectral theorem and its corresponding functional calculus; the existence of such machinery is highly non-trivial. Our proof avoids the spectral theorem all together. The two perspectives will be connected after the proof.

Theorem 3 (Hille-Yosida for self-adjoint operators)**.** *A self-adjoint, non-positive operator A on a Hilbert space H is the infinitesimal generator of contraction semigroup $\{e^{tA}\}$ of self-adjoint operators. Conversely, the generator A a contraction semigroup $\{e^{tA}\}$ of self-adjoint operators is non-positive and self-adjoint.*

Proof. Let A be a self-adjoint, non-positive operator on H. In view of Fact 1, $(0,\infty) \subseteq \rho(A)$. Observe that for any $\lambda > 0$ and $y \in D(A)$,

$$((\lambda I - A)y, y) = \lambda(y, y) - (Ay, y) \geqslant \lambda(y, y)$$

in view of the fact that A is non-positive. So for any $\lambda > 0$ and $x \in X$, $(\lambda I - A)^{-1}x \in D(A)$ and therefore

$$\begin{aligned}
\lambda\|(\lambda I - A)^{-1}x\|^2 &= \lambda((\lambda I - A)^{-1}x, (\lambda I - A)^{-1}x) \\
&\leqslant ((\lambda I - A)(\lambda I - A)^{-1}x, (\lambda I - A)^{-1}x) \\
&= (x, (\lambda I - A)^{-1}x) \\
&\leqslant \|x\|\|(\lambda I - A)^{-1}x\|.
\end{aligned}$$

Of course, this implies that $\|(\lambda I - A)^{-1}\|_{op} \leqslant \lambda^{-1}$ for all $\lambda > 0$ whence (19) of Theorem 2 is satisfied. An appeal to the theorem gives a contraction semigroup, $\{e^{tA}\}$ whose infinitesimal generator is A. Since semigroups are everywhere defined, it remains to be shown that each e^{tA} is symmetric. To this end, we note that $\lambda I - A$ is self-adjoint for each $\lambda > 0$. An appeal to Fact 2, shows that $(\lambda I - A)^{-1}$ is also self-adjoint for each $\lambda > 0$ and so the operator $S(\epsilon)$ defined in the proof of Theorem 1 is self-adjoint for each $\epsilon > 0$. From (11) and (12), we recall that for each $x \in D(A)$ and $\epsilon > 0$, $S(\epsilon)Ax = S(\epsilon)Ax$. Therefore, for each $x \in D(A)$ and $y \in X$,

$$(AS(\epsilon)y, x) = (S(\epsilon)y, Ax) = (y, S(\epsilon)Ax) = (y, AS(\epsilon)x).$$

Using the fact that $AS(\epsilon) \in \mathcal{B}(X)$ and $D(A)$ is a dense subset of X, it now follows that

$$(AS(\epsilon)y, x) = (y, AS(\epsilon)x)$$

for all $x, y \in X$ and therefore $AS(\epsilon)$ is self-adjoint for each $\epsilon > 0$. From this and in view of Example 1, it follows that for each $t > 0$ and $\epsilon > 0$, the operator $T(t, \epsilon)$ is self-adjoint and everywhere defined. Thus, for any $x, y \in X$ and $t > 0$,

$$(e^{tA}x, y) = \lim_{\epsilon \downarrow 0}(T(t,\epsilon)x, y) = \lim_{\epsilon \downarrow 0}(x, T(t,\epsilon)y) = (x, e^{tA}y)$$

and therefore e^{tA} is indeed symmetric for each $t \geqslant 0$.

For the converse direction, let $\{e^{tA}\}$ be a contraction semigroup of self-adjoint operators. In view of Theorem 2, the infinitesimal generator A is densely defined with $(0,\infty) \subseteq \rho(A)$. Using Fact 1, we deduce that A is non-positive. It remains to show that A is self-adjoint. Again, by a limiting argument and the fact that each e^{tA} is self-adjoint it follows that R_λ, defined by (8) is also self-adjoint for all $\lambda > 0$. By Proposition 3, $(\lambda I - A)^{-1}$ is therefore self-adjoint for some $\lambda > 0$, i.e. $(\lambda I - A)^{-1}$ is self-adjoint and injective. An appeal to Fact 2 guarantees that $(\lambda I - A)$, which we know has domain $D(A)$, is self-adjoint, whence A is is self-adjoint. This proves the theorem. □

Let's discuss the framework under which the theorem is usually proven. We begin by discussing the projection-valued measure version of the spectral theorem for unbounded operators. We will take the results herein for granted, for details see [7] or [11].

Fact 3 (Spectral Theorem)**.** *Let A be a self-adjoint operator on H. Then there exists a unique projection-valued measure, $E(\cdot)$ from the Borel σ-field on $\mathbb{R}$ such that*

$$A = \int_{-\infty}^{\infty} \lambda dE(\lambda).$$

This is called the spectral resolution of A. Also, the measure E is supported on the spectrum of A in the sense that $E(\sigma(A)) = I$ and so

$$A = \int_{\sigma(A)} \lambda dE(\lambda).$$

Moreover, for any $x, y \in H$, $E_{x,y} = (E(\cdot)x, y)$ is a Borel measure on $\mathbb{R}$ for which

$$D(A) = \left\{ x \in H : \int_{-\infty}^{\infty} \lambda^2 dE_{x,x}(\lambda) < \infty \right\}$$

and

$$(Ax, y) = \int_{-\infty}^{\infty} \lambda dE_{x,y}(\lambda) \tag{20}$$

when $x \in D(A)$.

For a self-adjoint operator A on H and its projection valued measure $E(\cdot)$, one can define for any Borel-measurable function $f : \mathbb{R} \to \mathbb{C}$ the operator

$$f(A) = \int_{-\infty}^{\infty} f(\lambda) dE(\lambda) \tag{21}$$

with domain

$$D(f(A)) = \left\{ x \in H : \int_{-\infty}^{\infty} |f(\lambda)|^2 dE_{x,x}(\lambda) < \infty \right\}. \tag{22}$$

For $f(A)$ an equality analogous to (20) holds.

Fact 4. *Let $f(A) : D(f(A)) \to H$ be as defined above. $f(A)$ is self-adjoint if and only if f is real valued.*

Fact 5. *For Borel-measurable functions f and g, let $f(A) : D(f(A)) \to H$, $g(A) : D(g(A)) \to H$ and $(fg)(A) : D((fg)(A)) \to H$ be as defined in* (21) *and* (22)*. Then*

$$\begin{aligned} f(A)g(A)x &= \left(\int_{-\infty}^{\infty} f(\lambda) dE(\lambda) \right) \left(\int_{-\infty}^{\infty} g(\lambda) dE(\lambda) \right) x \\ &= \int_{-\infty}^{\infty} (fg)(\lambda) dE(\lambda) x \\ &= (fg)(A)x \end{aligned}$$

for all $x \in D((fg)(A))$.

Example 2. *Let A be a non-positive self-adjoint operator and denote by $E(\cdot)$ its projection valued measure. Note that $\sigma(A) \subseteq (-\infty, 0]$. For each $t \geqslant 0$ define the self-adjoint operator,*

$$e^{tA} = \int_{-\infty}^{\infty} e^{t\lambda} dE(\lambda) = \int_{-\infty}^{0} e^{t\lambda} dE(\lambda).$$

We note that for each $t \geqslant 0$ and $\lambda \leqslant 0$, $|e^{t\lambda}|^2 \leqslant 1$ and so for any $x \in H$,

$$\int_{-\infty}^{\infty} |e^{t\lambda}|^2 dE_{x,x}(\lambda) = \int_{-\infty}^{0} |e^{t\lambda}|^2 dE_{x,x}(\lambda) \leqslant (E(\sigma(A))x, x) = \|x\|^2 < \infty.$$

Therefore $D(e^{tA}) = H$. Moreover, using the spectral calculus, it follows that $\{e^{tA}\}$ is a contraction semigroup on H.

Remark 3. *The careful reader should notice that we have defined the symbol e^{tA} in two different ways: one via the Hille-Yosida theorem and one by means of the spectral calculus. There is justice in the world; these constructions do indeed agree. In fact, the proof of Hille-Yosida is significantly shorter in this context [6].*

Example 3. *Let A be a non-positive self-adjoint operator with projection valued measure $E(\cdot)$. Define the self-adjoint operator*

$$\sqrt{-A} = \int_{-\infty}^{0} \sqrt{-\lambda} dE(\lambda).$$

In view of the spectral calculus, for any $x \in D(A)$,

$$\begin{aligned}
\int_{-\infty}^{0} |\sqrt{-\lambda}|^2 dE_{x,x}(\lambda) &= \int_{-\infty}^{0} (-\lambda) dE_{x,x}(\lambda) \\
&= \int_{-1}^{0} dE_{x,x}(\lambda) + \int_{-\infty}^{-1} (-\lambda) dE_{x,x}(\lambda) \\
&\leqslant (E((-1,0)x, x) + \int_{-\infty}^{-1} |\lambda|^2 dE_{x,x}(\lambda) \\
&\leqslant \|x\|^2 + \int_{-\infty}^{\infty} |\lambda|^2 dE_{x,x}(\lambda) < \infty.
\end{aligned}$$

Thus $x \in D(\sqrt{-A})$. Also for any $x \in D(A)$ observe that

$$\begin{aligned}
\sqrt{-A}\sqrt{-A}x &= \left(\int_{-\infty}^{\infty} \sqrt{-\lambda} dE(\lambda)\right)\left(\int_{-\infty}^{\infty} \sqrt{-\lambda} dE(\lambda)\right) x \\
&= \int_{-\infty}^{\infty} (\sqrt{-\lambda})^2 dE(\lambda) x \\
&= -\int_{-\infty}^{\infty} \lambda dE(\lambda) x = -Ax
\end{aligned}$$

where we have made use of Fact 5. For this reason we call $\sqrt{-A}$ the square root of $-A$. It can be shown that this is the unique non-negative operator on H whose square is $-A$. We also have that

$$(-Ax, y) = (\sqrt{-A}x, \sqrt{-A}y)$$

for all $x \in D(A)$ and $y \in D(\sqrt{-A})$.

2.4 Symmetric forms and self-adjoint operators

In this subsection, our discussion turns to quadratic forms. Quadratic forms are of fundamental importance to both partial differential equations and Markov processes. In the case of Markov processes, the quadratic forms of interest are called Dirichlet forms; they will be studied in the next subsection. In this subsection, our goal is to show that there is a one-to-one correspondence between closed symmetric forms and non-positive self-adjoint operators. Again, our setting is a real Hilbert space H with inner product $(\cdot, \cdot)$.

Definition 11 (Symmetric form)**.** *Let $D(Q) \subseteq H$ be a dense, linear subspace of H. A symmetric form on H is a map $Q : D(Q) \times D(Q) \to \mathbb{R}$ that is*

- $\mathbb{R}$-*Bilinear,*
- *Symmetric: $Q(x, y) = Q(y, x)$ for all $x \in D(Q)$,*

and

- *Non-negative: $Q(x, x) \geqslant 0$ for all $x \in D(Q)$.*

We will call $D(Q)$ the domain of Q.

Remark 4. *For purposes of functional analysis, the requirements of the above definition are rather strong. When put in the context of complex Hilbert spaces, one can weaken the requirements and much of the theory still goes through [7,11]. We will not need this generality, but it's nice to know it's there.*

Definition 12. *A symmetric form Q with domain $D(Q)$ is said to be closed if for any sequence $\{x_n\} \subseteq D(Q)$ such that*

$$x_n \to x \quad \text{and} \quad Q(x_n - x_k, x_n - x_k) \to 0 \quad \text{as} \quad n, k \to \infty$$

we have

$$x \in D(Q) \quad \text{and} \quad Q(x_n - x, x_n - x) \to 0 \quad \text{as} \quad n \to \infty.$$

Lemma 1. *Let Q be a closed symmetric form on H such that*

$$Q(x, x) \geqslant \|x\|^2 \text{ for all } x \in D(Q). \tag{23}$$

Then $D(Q)$ is a Hilbert space in the inner product $Q(\cdot, \cdot)$.

Remark 5. *The converse of Lemma 1 is also true.*

Proof of lemma. Because Q is a form it satisfies all of the requirements of an inner product except possibly positive definiteness. Positive definiteness follows immediately from (23) (and the fact that $D(Q)$ is a linear subspace of H). Thus Q is an inner product on $D(Q)$. The substance of this proof comes in showing $D(Q)$ is complete in the metric given by Q. To this end, let $\{x_n\} \subseteq D(Q)$ be a Cauchy sequence in $D(Q)$ with respect to Q, i.e.,

$$Q(x_n - x_k, x_n - x_k) \text{ as } n, k \to \infty.$$

By (23), $\{x_n\}$ must also be a Cauchy sequence in H and since H is a Hilbert space, it has a limit in H, let's call it x. Using the fact that Q is closed, it must be true that $x \in D(Q)$ and that

$$\lim_n Q(x_n - x, x_n - x) = 0$$

whence $D(Q)$ is complete. □

Lemma 2. *Let Q be a closed symmetric form on H satisfying* (23)*. Then there exists a unique self-adjoint operator A on H such that*

1. *$A \geqslant 1$,*
2. *$D(A) \subseteq D(Q)$,*
3. *$Q(x, y) = (x, Ay)$ for all $x \in D(Q)$ and $y \in D(A)$.*

Proof. We first show the existence of A. For $y \in H$ consider the map $D(Q) : x \mapsto (x, y)$. By (23)

$$|(x, y)| \leqslant \|x\|\|y\| \leqslant \|y\|(Q(x, x))^{1/2}$$

whence $D(Q) : x \mapsto (x, y)$ is a continuous linear functional on the Hilbert space $D(Q)$ with inner product Q. In view of Lemma 1, an appeal to the so-called Riesz representation theorem for Hilbert spaces, [5, Theorem 2, Section D.3], gives a vector $Ty \in D(Q)$ such that

$$Q(x, Ty) = (x, y) \text{ for all } x \in D(Q).$$

Clearly, this produces a linear map $T : H \to D(Q)$ such that

$$Q(x, Ty) = (x, y) \text{ for all } y \in H \text{ and } x \in D(Q). \tag{24}$$

Using 23, we see immediately that $\|T\|_{op} \leqslant 1$. If $Ty = 0$ for some $y \in H$, the above equation implies that $(x, y) = 0$ for all $x \in D(Q)$. But since Q is densely defined, $(x, y) = 0$ for all $x \in H$ whence $y = 0$. Therefore T is injective. T is also symmetric. To see this observe that

$$(Tx, y) = Q(Tx, Ty) = Q(Ty, Tx) = (Ty, x) = (x, Ty) \tag{25}$$

for all $x, y \in D(Q)$ where we have used the fact that $Q(\cdot,\cdot)$ and $(\cdot,\cdot)$ are both symmetric. Now from Fact 2 it follows that the inverse of T, let's call it $A: D(A) \to H$, is self-adjoint and densely defined. Necessarily, $D(A) \subseteq D(Q)$ which establishes 2. Although both domains are dense in H, they cannot be equal unless T is surjective.

In view of (24) and our definition of A it follows that

$$Q(x, y) = Q(x, T(A)y) = (x, Ay)$$

for $x \in D(Q)$ and $y \in D(A)$. This proves 3. To see 2., we observe that for any $x \in D(A)$

$$(x, Ax) = Q(x, T(A)x) = Q(x, x) \geqslant \|x\|^2$$

and therefore, $A \geqslant 1$.

It remains to show that A is unique. Suppose that B is another self-adjoint operator satisfying conditions $1-3$. If $Bx = 0$ for some $x \in D(B)$ then by 2.,

$$Q(y, x) = (y, Bx) = 0 \text{ for all } y \in D(Q).$$

Using Lemma 1, we may conclude that $x = 0$ whence $B : D(B) \to H$ is injective. Let's denote the inverse of B by $\tilde{T}$. Using Fact 2 again we see that $D(\tilde{T})$ is dense in H. Using 1. it follows immediately that $\|\tilde{T}\|_{op} \leqslant 1$. Since a densely defined operator can be extended uniquely to all of H with the same bound, we extended it and by an abuse of notation denote the extension by $\tilde{T}$. Using the fact that B is closed (it is self-adjoint), it follows quickly that $B\tilde{T} = I$. In fact, this shows that B was surjective to begin with and extension was really unnecessary.

Consequently, for each $y \in H$, we have that

$$Q(x, Ty) = (x, y) = Q(x, \tilde{T}y)$$

for all $x \in D(Q)$. Again by Lemma 1 we conclude that $Ty = \tilde{T}y$ for all $y \in H$ which proves that $D(B) = \tilde{T}(H) = T(H) = D(A)$ and that $B = A$ as desired. □

Theorem 4. *There is a one-to-one correspondence between the closed symmetric forms on H and the non-positive self-adjoint operators on H. The correspondence is given by*

$$\begin{cases} Q(x,x) = (\sqrt{-A}x\sqrt{-A}x) \\ D(Q) = D(\sqrt{-A}). \end{cases} \tag{26}$$

Proof. Let A be a non-positive self-adjoint operator on H and define Q by (26). In view of Fact 4 and Example 3, $\sqrt{-A}$ is a non-negative self-adjoint operator on H with $D(A) \subseteq D(\sqrt{-A})$. It is clear that Q is a symmetric form on H. Let us show Q is closed. Let $\{x_n\} \subseteq D(Q)$ be a Cauchy sequence with respect to Q and such that $x_n \to x$ for some $x \in H$. Now

$$\lim_{n,k\to\infty} Q(x_n - x_k, x_n - x_k) = \lim_{n,k\to 0} \|\sqrt{-A}(x_n - x_k)\|^2 = 0$$

and because H is complete, $\sqrt{-A}x_n \to y$ as $n \to \infty$ for some $y \in H$. Since $\sqrt{-A}$ is self-adjoint, it is closed from which we deduce that

$$x \in D(\sqrt{-A}) \text{ and } \lim_{n\to\infty} \sqrt{-A}x_n = y = \sqrt{-A}x.$$

Therefore

$$x \in D(Q) \text{ and } \lim_{n\to\infty} Q(x_n - x, x_n - x) = \lim_{n\to\infty} \|\sqrt{-A}(x_n - x)\|^2 = 0.$$

It remains to show that every closed symmetric form comes about in this form. For this we will appeal to Lemma 2. Let Q be a closed symmetric form with domain $D(Q)$ and define $Q_1 : D(Q) \times D(Q) \to \mathbb{R}$ by

$$Q_1(x, y) = Q(x, y) + (x, y)$$

for $x, y \in D(Q)$. Because Q is a closed symmetric form on H is follows immediately that Q_1 is a closed symmetric form and satisfies (23). By Lemma 2, there is a unique self-adjoint operator B for which conditions $1-3$. of the lemma are satisfied. Set $A = I - B$ and $D(A) = D(B) \subseteq D(Q)$. A is clearly self-adjoint. We observe that for all $x \in D(A)$

$$(Ax, x) = (x, x) - (Bx, x) \leqslant \|x\|^2 - \|x\|^2 = 0$$

in view of condition 1. of the lemma. Therefore, A is non-positive. Also notice that for all $x \in D(A)$ and $y \in D(Q)$,

$$Q(x, y) = Q_1(x, y) - (x, y) = (Bx, y) - (x, y) = ((B - I)x, y) = (-Ax, y).$$

Using the fact that Q is closed and the operator defined in Example 3 agrees with Q in the sense defined above, it follows that Q is given by (26) as desired. □

2.5 Dirichlet forms and Markovian semigroups

In this section, we complete the second part of our diagram:

$$\mathcal{E} \Longleftrightarrow e^{tA}$$

We now focus on a particular class of Hilbert spaces and apply the theory of the last section. Let X be a locally compact separable Hausdorff space and m a positive Radon measure on X such that $\mathrm{Supp}(m) = X$. Our space of interest is the Hilbert space $H = L^2(X, m)$ with inner product

$$(f, g) = \int_X fg\,dm.$$

We will take by definition each $f \in L^2(X, m)$ to be real valued. Unless otherwise mentioned, almost everywhere means m-almost everywhere.

Definition 13 (Markovian)**.** *A bounded linear operator T on $L^2(X,m)$ is said to be Markovian if*

$$0 \leqslant Tf \leqslant 1 \ \ \textit{almost everywhere}$$

whenever $f \in L^2(X,m)$ and $0 \leqslant f \leqslant 1$ almost everywhere.

Definition 14 (Markovian Semigroup)**.** *Let $\{T_t\}$ be a semigroup on $L^2(X,dm)$. We say that $\{T_t\}$ is Markovian if T_t is Markovian for every $t \geqslant 0$.*

The *unit contraction* is the function $\Phi : \mathbb{R} \to \mathbb{R}$

$$\Phi(t) = (0 \vee t) \wedge 1 = \begin{cases} 0, & \text{if } t \leqslant 0 \\ t, & \text{if } 0 \leqslant t \leqslant 1 \\ 1, & \text{if } 1 \leqslant t \end{cases}.$$

It is easily seen that $|\Phi(t)| \leqslant |\Phi(s)|$ for all $t \leqslant s$ and

$$|\Phi(t) - \Phi(s)| \leqslant |t - s|$$

for all $t, s \in \mathbb{R}$.

Definition 15 (Dirichlet Form)**.** *A Dirichlet form on $L^2(X,m)$ is a symmetric closed form $\mathcal{E}$ with domain $D(\mathcal{E})$ such that $\Phi(f) \in D(\mathcal{E})$ whenever $f \in D(\mathcal{E})$ and*

$$\mathcal{E}\left(\Phi(g), \Phi(f)\right) \leqslant \mathcal{E}(f,f).$$

Remark 6. *The condition above is often verbalized by saying that the unit contraction operates on $\mathcal{E}$. There are a few other notions equivalent to this condition, [6].*

Definition 16 (Regular)**.** *Denote by $C_0(X)$ the set of continuous compactly supported functions on X. We say that a Dirichlet form $\mathcal{E}$ with domain $D(\mathcal{E})$ is regular, if there is a set $\mathcal{C} \subseteq D(\mathcal{E}) \cap C_0(X)$ such that the following two conditions are satisfied:*

1. *$\mathcal{C}$ is a dense in $D(\mathcal{E})$ in the sense of the norm on $D(\mathcal{E})$ defined by the inner product,*
$$\mathcal{E}_1(f,g) = \mathcal{E}(f,g) + (f,g).$$
2. *$\mathcal{C}$ is dense in $C_0(X)$ in the sense of the sup-norm topology.*

Before treating the main theorem of this section, let us give two examples. After the theorem's proof we shall give an important example of a closed symmetric form which is not a Dirichlet form.

Example 4. *In the present example, our setting is $\mathbb{R}^n$ equipped with Lebesgue measure. The symmetric form $\mathcal{E}$ defined by*

$$\mathcal{E}(f,g) = \frac{1}{2}\int_{\mathbb{R}^n} (\nabla f)(x) \cdot (\nabla g)(x) dx = \frac{1}{2}\int_{\mathbb{R}^n} \sum_{l=1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_i} dx$$

with domain

$$D(\mathcal{E}) = H^1(\mathbb{R}^n) = \left\{ f \in L^2(\mathbb{R}^n) : \int_{\mathbb{R}^n} (1 + |\xi|^2)|\hat{f}(\xi)|^2 d\xi < \infty \right\}$$

is a regular Dirichlet form, [6]. Here, $\hat{f}$ denotes the Fourier transform of $f \in L^2(\mathbb{R}^n)$ and $H^1(\mathbb{R}^n) = W^{2,1}(\mathbb{R}^n)$ is the standard Sobolev space. A nice proof of the characterization of H^1 via Fourier transform can be found in [8, Theorem 7.9].

Example 5. *Let G be a finitely generated group equipped with the discrete topology, i.e. a finitely generated discrete group. We will denote counting measure on G by $\#$. If S is a minimal set of generators, the function*

$$\mathcal{E}(f, g) = \sum_{x \in G, s \in S} (f(sx) - f(x))(g(sx) - g(x))$$

is a Dirichlet form with domain $L^2(G, \#)$, i.e., it is everywhere defined. See [15] for discussion.

Theorem 5. *There is a one to one correspondence between Dirichlet forms and Markovian symmetric contraction semigroups on $L^2(X, m)$.*

We first consider two technical results. The first we shall prove as a lemma and the second we leave as a fact. The fact's proof can be found in [6] and is an application of the measure theoretic Riesz representation theorem, c.f. [10, Theorem 2.1.4].

Lemma 3. *Let $\{T_t\}$ be a contraction semigroup on $L^2(X, dm)$ and let $\{R_\lambda\}_{\lambda>0}$ be its corresponding collection of resolvent operators defined by* (8)*. Then $\{T_t\}$ is Markovian if and only for all $\lambda > 0$, λR_λ is Markovian.*

Remark 7. *We note that our assumption that $\{T_t\}$ is a contraction semigroup guarantees that the Riemann integral in* (8) *converges for all $\lambda > 0$. Thus λR_λ is indeed a bounded operator for each $\lambda > 0$.*

Proof of lemma. We shall prove the forward direction; the converse is similar and can be found in [6]. Suppose that the semigroup $\{T_t\}_{t \geqslant 0}$ is Markovian and take $\lambda > 0$. Let $f \in L^2(X, dm)$ be such that $0 \leqslant f \leqslant 1$, almost everywhere. In the next step we have to be careful with null sets. Let $\mathcal{D}$ be the dyadic rationals on $[0, \infty)$ and set

$$X_g = \bigcap_{d \in \mathcal{D}} \{x \in X : 0 \leqslant (T_d f)(x) \leqslant 1\}.$$

Because $\mathcal{D}$ is countable and the semigroup is Markovian, $m(X \backslash X_g) = 0$. We consider the approximation by Riemann sums,

$$\lambda R_\lambda f = \lim_{n \to \infty} \frac{\lambda}{2^n} \sum_{l=0}^{2^{n+1}} e^{-\lambda l/2^n} T_{l/2^n} f$$

where the limit converges in the sense of L^2 in view of Proposition 3. It follows from basic measure theory that there is a subsequence $\{n_i\}_{i=1}^{\infty} \subseteq \mathbb{N}$ such that for almost every $x \in X$,

$$\lambda R_\lambda f(x) = \lim_{i\to\infty} \frac{\lambda}{2^{n_i}} \sum_{l=0}^{2^{n_i+1}} e^{-\lambda l/2^{n_i}} (T_{l/2^{n_i}} f)(x). \tag{27}$$

Let Y_g denote the set of full measure for which (27) holds. Since $l/2^{n_i} \in \mathcal{D}$ for all $i, l \in \mathbb{N}$, for any $x \in X_g \cap Y_g$ we have

$$\begin{aligned} 0 &\leqslant \lim_{i\to\infty} \frac{\lambda}{2^{n_i}} \sum_{l=0}^{2^{n_i+1}} e^{-\lambda l/2^{n_i}} (T_{l/2^{n_i}} f)(x) = (\lambda R_\lambda f)(x) \\ &\leqslant \lim_{i\to\infty} \frac{\lambda}{2^{n_i}} \sum_{l=0}^{2^{n_i+1}} e^{-\lambda l/2^{n_i}} \leqslant \lambda \int_0^\infty e^{-\lambda t} dt = 1. \end{aligned}$$

Because $m(X \backslash (X_g \cap Y_g)) = 0$, it follows that $0 \leqslant \lambda R_\lambda f \leqslant 1$ almost everywhere.

∎

Fact 6. *Let $S \in \mathcal{B}(L^2(X, dm))$ be self-adjoint. Then there exists a unique symmetric Radon measure σ on $X \times X$ such that*

$$(f, Sg) = \int_{X\times X} f(x)g(y)\sigma(dx, dy).$$

If S is Markovian the $\sigma(X \times E) \leqslant m(E)$ for all Borel sets $E \subseteq X$. Moreover,

$$0 \leqslant \frac{d\sigma(X \times (\cdot))}{dm} \leqslant 1 \quad \textit{almost everywhere.}$$

The function above denotes the Radon-Nikodym derivative of the measure $\sigma(X \times (\cdot))$ with respect to m.

We shall now prove Theorem 5.

Proof. By Theorems 3 and 4, for every contraction semigroup, $\{e^{tA}\}$ of self-adjoint operators on $L^2(X, m)$ there is a closed symmetric form $\mathcal{E}$ on $L^2(X, m)$ and conversely, to each closed symmetric form $\mathcal{E}$ on $L^2(X, m)$ there is a contraction semigroup $\{e^{tA}\}$ of self-adjoint operators. All that needs to be shown is that $\{e^{tA}\}$ is Markovian if and only if $\mathcal{E}$ is a Dirichlet form.

We suppose that $\mathcal{E}$ is a Dirichlet form and let $\lambda > 0$. We shall prove that λR_λ is Markovian where R_λ is the resolvent of the semigroup $\{e^{tA}\}$. To this end, let $g \in L^2(X, m)$ such that $0 \leqslant g \leqslant 1$, almost everywhere, and define

$$\psi(f) = \mathcal{E}(f, f) + \lambda(f - \frac{g}{\lambda}, f - \frac{g}{\lambda}) \text{ for } f \in D(\mathcal{E}).$$

In view of Proposition 3 and Theorem 4 we have

$$\mathcal{E}(R_\lambda g, f) + \lambda(R_\lambda g, f) = (g, f)$$

for all $g \in L^2(X,m)$ and $f \in D(\mathcal{E})$. From this it follows that

$$\psi(f) = \psi(R_\lambda g) + \mathcal{E}(R_\lambda g - f, R_\lambda g - f) + (R_\lambda g - f, R_\lambda - f)$$

and so $R_\lambda g$ is the minimizer of ψ; it is clearly the unique minimizer. Since $\mathcal{E}$ is a Dirichlet form, let Φ denote the unit contraction, define $\eta : \mathbb{R} \to \mathbb{R}$ by

$$\eta(t) = \frac{1}{\lambda}\Phi(\lambda t)$$

and put

$$w = \eta(R_\lambda g).$$

Our goal is to show that $w = R_\lambda g$ almost everywhere. By hypothesis,

$$\begin{aligned} \mathcal{E}(w,w) &= \frac{1}{\lambda}\mathcal{E}(\Phi(\lambda R_\lambda g), \Phi(\lambda R_\lambda g)) \\ &\leqslant \frac{1}{\lambda}\mathcal{E}(\lambda R_\lambda g, \lambda R_\lambda g) = \mathcal{E}(R_\lambda g, R_\lambda g). \end{aligned} \tag{28}$$

Also by our definition of η it follows that

$$|\eta(t) - s| \leqslant |t - s| \quad \text{for all } 0 \leqslant s \leqslant 1/\lambda.$$

Therefore

$$\left| w(x) - \frac{g(x)}{\lambda} \right| \leqslant \left| (R_\lambda g)(x) - \frac{g(x)}{\lambda} \right|$$

almost everywhere. Consequently

$$\begin{aligned} \left(w - \frac{g}{\lambda}, w - \frac{g}{\lambda} \right) &= \int_X \left(w - \frac{g}{\lambda} \right)^2 dm \\ &\leqslant \int_X \left(R_\lambda g - \frac{g}{\lambda} \right)^2 dm \\ &= \left(R_\lambda g - \frac{g}{\lambda}, R_\lambda g - \frac{g}{\lambda} \right). \end{aligned} \tag{29}$$

From (29) and (28) we have

$$\psi(w) \leqslant \psi(R_\lambda g)$$

and since $R_\lambda g$ minimized ψ, $w = R_\lambda g$ almost everywhere. Therefore

$$0 \leqslant \Phi(\lambda R_\lambda g) = \lambda R_\lambda g \leqslant 1$$

almost everywhere and so λR_λ is Markovian. With the help of Lemma 3, we conclude that $\{e^{tA}\}$ is Markovian.

Conversely, we assume that $\{e^{tA}\}$ is Markovian. For each $t > 0$ we shall denote by σ_t the symmetric Radon measure on $X \times X$ guaranteed by Fact 6 and define

$$\mathcal{E}^{(t)}(f,g) = \frac{1}{t}(f - e^{tA}f, g)$$

for $f, g \in L^2(X,m)$. In view of Theorem 4,

$$\lim_{t\downarrow 0} \mathcal{E}^{(t)}(f,f) = \mathcal{E}(f,f) \tag{30}$$

whenever $f \in D(A)$. Also by Fact 6,

$$\begin{aligned}\mathcal{E}^{(t)}(f,f) &= \frac{1}{t}\int_X f^2 dm - \frac{1}{t}\int_{X\times X} f(x)f(y)\sigma_t(dx,dy)\\ &= \frac{1}{2t}\int_{X\times X}(f(x)-f(y))^2\sigma_t(dx,dy)\\ &\qquad +\frac{1}{t}\int_X f(x)^2\left(1-\frac{d\sigma_t(X\times\cdot)}{dm}(x)\right)m(dx)\\ &=: \frac{1}{2t}\int_{X\times X}(f(x)-f(y))^2\sigma_t(dx,dy)+\frac{1}{t}\int_X f(x)^2 s_t(x)m(dx)\end{aligned}$$

for $f \in L^2(X,m)$ and where $0 \leqslant s_t \leqslant 1$ almost everywhere. Because both integrands are non-negative,

$$\begin{aligned}\mathcal{E}^{(t)}(\Phi(f),\Phi(f)) &= \frac{1}{2t}\int(\Phi(f(x))-\Phi(f(y)))^2\sigma_t(dx,dy)+\frac{1}{t}\int(\Phi(f))^2 s_t dm\\ &\leqslant \frac{1}{2t}\int(f(x)-f(y))^2\sigma_t(dx,dy)+\frac{1}{t}\int f^2 s_t dm = \mathcal{E}^{(t)}(f,f)\end{aligned} \tag{31}$$

for all $t > 0$ and $f \in L^2(X,m)$. In the above computation we used the two properties of the unit contraction discussed immediately after its introduction. It now follows from (30) and (31) that $\mathcal{E}$ is a Dirichlet form and the theorem is proved. $\square$

Example 6. *We return to Example 4. Again our setting is $\mathbb{R}^n$ with Lebesgue measure. The trichotomy of Theorems 4 and 5 takes form in the following way:*

$$// \ \mathcal{E} \Longleftrightarrow e^{tA} \Longleftrightarrow A$$

$$\mathcal{E} \Longleftrightarrow e^{t(\frac{1}{2}\Delta)} \Longleftrightarrow \tfrac{1}{2}\Delta$$

Here $\mathcal{E}$ is the Dirichlet form from Example 4 with domain $H^1(\mathbb{R}^n)$. Δ denotes the standard Laplacian operator on $\mathbb{R}^n$ and has the Sobolev space $H^2(\mathbb{R}^n)$ as its domain. By virtue of the Fourier transform, it is easy to see that for $f \in L^2(\mathbb{R}^n)$

$$(e^{t(\Delta/2)}f)(x) = \int_{\mathbb{R}^n} K_2^{(t/2)}(x-y)f(y)dy$$

where

$$K_2^t(x) = \frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|x|^2}{4t}\right) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}e^{-t|\xi|^2}d\xi$$

is the familiar heat kernel.

For illustrative purposes, we give an example of a closed symmetric form that isn't a Dirichlet form.

Example 7. *For simplicity, we work in $\mathbb{R}$. Define $K_4^{(\cdot)}:(0,\infty)\times\mathbb{R}$ by*

$$K_4^t(x)=\frac{1}{2\pi}\int_{\mathbb{R}}e^{-ix\xi}e^{-t|\xi|^4}d\xi.$$

As one can compute (numerically), K_4^1 is negative on a set of positive measure unlike its counterpart K_2^1. Consequently, the semigroup corresponding to the 4^{th}-order operator $(\Delta/2)^2$ on $H^4(\mathbb{R})$ and defined by

$$(e^{t(\Delta/2)^2}f)(x)=\int_{\mathbb{R}}K_4^{t/4}(x-y)f(y)dy$$

for $f\in L^2(\mathbb{R})$ is not Markovian. In view of Theorem 5, its corresponding closed symmetric form $\mathcal{E}$ is not a Dirichlet form. In this case $\mathcal{E}$ has domain $H^2(\mathbb{R})$ and is given by

$$\mathcal{E}(f,g)=\frac{1}{4}\int_{\mathbb{R}}(\Delta f)(\Delta g)dx=\frac{1}{4}\int_{\mathbb{R}}|\hat{f}(\xi)|^2|\hat{g}(\xi)|^2d\xi.$$

This is discussed more thoroughly in [3, Section 6.2]. A generalization of this example as it pertains to local limit theorems can be found in [9].

As it turns out, the probabilistic interpretation of Dirichlet forms and Markovian semigroups, discussed in the next section, is integral to much of analysis (and geometry). In the study of higher order partial differential operators, the analysis becomes substantially more difficult as many of the arguments that work for second order operators, with corresponding Markovian semigroups, fail in the context of higher order operators if they are true at all. The reader is encouraged to see [2] for an extensive discussion on this topic.

3 Probability

In this section, we discuss why semigroups, self-adjoint operators and Dirichlet forms are important in probability theory. In particular, we discuss their connection to Markov Processes. Our discussion follows [6] and [12].

3.1 Hunt Processes

The present section introduces the notion of a "nice" class of Markov processes called Hunt processes. Our presentation is similar to that of [1] although, we work under weaker topological assumptions on the state space.

Let S be a locally compact separable Hausdorff space. We denote the Borel σ-field on S by $\mathscr{B}(S)$. Adjoin a point Δ to S by setting $S_\Delta=S\cup\Delta$ and

$$\mathscr{B}_\Delta=\mathscr{B}(S)\cup\{B\cup\Delta:B\in\mathscr{B}(S)\}.$$

Topologically, S_Δ is the one point compactification of S.

Definition 17 (Markov Process)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability triple and take $S, \mathscr{B}(S), S_\Delta$ and $\mathscr{B}_\Delta$ as above. We call $M = (\Omega, \mathcal{F}, \{X_t\}_{t\in[0,\infty]}, \{\mathbb{P}^x\}_{x\in S_\Delta})$ a Markov process if*

1. *For each $x \in S_\Delta$, $M_x = (\Omega, \mathcal{F}, \{X_t\}_{t\in[0,\infty]}, \mathbb{P}^x)$ is a stochastic process with state space $(S, \mathscr{B}(S))$.*
2. *$\mathbb{P}^x$ is $\mathscr{B}(S)$-measurable as a function of $x \in S$ for each $t \in [0, \infty]$.*
3. *There is an admissible filtration $\{\mathcal{F}_t\}$ for which*
$$\mathbb{P}^x(X_{t+s} \in E|\mathcal{F}_t) = \mathbb{P}^{X_t}(X_s \in E),\ \mathbb{P}^x\text{- almost surely}$$
for all $x \in S$, $E \in \mathscr{B}(S)$ and $t, s \in [0, \infty]$.
4. *$\mathbb{P}^\Delta(X_t = \Delta) = 1$ for all $t \geqslant 0$.*

Definition 18 (Transition function)**.** *Let M be a Markov process. The function $p_{(\cdot)} : [0, \infty] \times S \times \mathscr{B}(S) \to [0, 1]$ defined by*
$$p_t(x, E) = \mathbb{P}^x(X_t \in E)$$
is called the transition function of the process M.

We would like to associate the starting point of a Markov process with a distribution. To this end, let M be a Markov process and for any probability measure μ on S_Δ put
$$\mathbb{P}^\mu(E) = \int_{S_\Delta} \mathbb{P}^x(E)\mu(dx)$$
whenever $E \in \mathcal{F}^0_\infty = \sigma(X_s, s < \infty)$.

Definition 19 (Strong Markov Process)**.** *A Markov process M is called a strong Markov process (with respect to an admissible Filtration $\{\mathcal{F}_t\}$) if for any stopping time τ, probability measure μ on S and $E \in \mathscr{B}_\Delta$,*
$$\mathbb{P}^\mu(X_{\tau+s} \in E|\mathcal{F}_t) = \mathbb{P}^{X_\tau}(X_s \in E) \qquad \mathbb{P}^\mu\text{-almost surely.}$$

We are now in the position to define Hunt Processes.

Definition 20 (Hunt Process)**.** *A strong Markov process with state space $(S, \mathscr{B}(S))$ is called a Hunt Process if the following conditions are satisfied:*

1. *$X_\infty(\omega) = \Delta$ for all $\omega \in \Omega$.*
2. *$X_t(\omega) = \Delta$ for all $t \geqslant \zeta(\omega)$ where*
$$\zeta(\omega) = \inf\{t \geqslant 0 : X_t(\omega) = \Delta\}.$$
3. *For each $t \in [0, \infty]$ there is a map $\theta_t : \Omega \to \Omega$ for which*
$$X_s \circ \theta_t = X_{t+s} \text{ for all } s \geqslant 0.$$

4. *M_x is cádlag for each $x \in S_\Delta$.*
5. *M is quasi-left continuous. This means that for any stopping time τ and any sequence, $\{\tau_n\}$, of stopping times increasing to τ,*

$$\mathbb{P}^x(\lim_{n\to\infty} X_{\tau_n} = X_\tau, \tau < \infty) = \mathbb{P}^x(\tau < \infty).$$

In this case we call $\zeta(\omega)$ the lifetime of M and θ_t the translation operator of M. A Hunt process is called a diffusion if, in addition, the paths are almost surely continuous.

3.2 Hunt Processes to Dirichlet forms

In this section, we consider a Hunt processes M with state space $(S, \mathscr{B}(S))$. We shall also assume that the space is equipped with a "reference" measure, m, which we will take to be a positive Radon measure with $\mathrm{Supp}(m) = S$. When S is a locally compact group m is often taken to be Haar measure, see [15].

Let $p_{(\cdot)}$ be the transition function associated to M and for each $t > 0$ define

$$(T_t f)(x) = \int_S f(y) p_t(x, dy) = \mathbb{E}^x f(X_t) \qquad (\leqslant \infty) \tag{32}$$

for any f for which the integral exists in the extended sense. In view of the simple approximation lemma and the Monotone convergence theorem, this includes all non-negative Borel-measurable functions.

Definition 21. *We say that M is an m-symmetric Hunt process if for each $t > 0$,*

$$\int_S f(x)(T_t g)(x) dm(x) = \int_S (T_t f)(x) g(x) dm(x). \tag{33}$$

for all non-negative Borel-measurable functions f and g.

Lemma 4. *Let M be an m-symmetric Hunt process with state space $(S, \mathscr{B}(S), m)$. Then for each $t > 0$, T_t, defined by (32), extends to a bounded self-adjoint operator on $L^2(S, m)$ with $\|T_t\|_{op} \leqslant 1$.*

Proof. Fix $t > 0$ and observe that for any non-negative Borel-measurable function f,

$$((T_t f)(x))^2 \leqslant (T_t \mathbb{1}_S)(T_t f^2)(x) \leqslant (T_t f^2)(x) \;\; m\text{-almost everywhere}$$

where we have used Schwarz's inequality. Now because T_t is m-symmetric,

$$\int_S |T_t f|^2 dm \leqslant \int_S (T_t f^2)(x) dm(x) \leqslant \int_S (T_t \mathbb{1}_S)(x)(f(x))^2 dm(x) \leqslant \int_s |f|^2 dm.$$

It follows immediately that we can extend T_t (linearly) to a bounded operator on $L^2(S, m)$ with $\|T_t\|_{op} \leqslant 1$. In view of (33), T_t must be self-adjoint. $\square$

The next theorem provides the final piece to our diagram:

$$\begin{tikzcd} X_t \arrow[dr] & \\ & e^{tA} \end{tikzcd}$$

Theorem 6 (Hunt processes to Dirichlet forms)**.** *Let S be a separable, locally compact, Hausdorff space and m, a positive Radon measure on X with $Supp(m) = S$. Then for any m-symmetric Hunt process M with state space $(S, \mathscr{B}(S), m)$, $\{T_t\}$, defined by (32), is a self-adjoint Markovian contraction semigroup on $L^2(S, m)$.*

The following corollary is immediate from Proposition 2:

Corollary 2. *Let M and T_t be as above and let A denote the non-positive definite self-adjoint infinitesimal generator of T_t. Then for any $f \in D(A) \subseteq L^2(S, m)$, $u(x,t) = \mathbb{E}^x(f(X_t))$ solves the heat equation:*

$$\begin{cases} (\partial_t - A)u(x,t) = 0 \\ u(x,0) = f(x). \end{cases}$$

Proof of theorem. We first define T_0 to be the identity operator on $L^2(S, m)$; the careful reader should be relieved. In view of the previous lemma, $\{T_t\}_{t \geqslant 0}$ is a collection of self-adjoint operators on $L^2(S, m)$ with $\|T_t\|_{op} \leqslant 1$ for all $t \geqslant 0$. To see that the semigroup property is satisfied, observe that for any $E \in \mathscr{B}(S)$ and any $s, t \geqslant 0$,

$$\begin{aligned} T_t T_s \mathbb{1}_E(x) &= \int_S \int_E p_s(y, dz) p_t(x, dy) \\ &= \int_S \mathbb{P}^y(X_s \in E) p_t(x, dy) \\ &= \mathbb{P}^x(X_{s+t} \in E) = T_{s+t}\mathbb{1}_E \end{aligned}$$

where in the last step we used the Markov property. Since T_t is linear, the above equation forces $T_t T_s$ and T_{t+s} to agree on a dense subset of $L^2(S, m)$, namely the simple functions. Since the operators are bounded, it follows immediately that $\{T_t\}$ satisfies the semigroup property.

To see that T_t is Markovian observe that for any $0 \leqslant f \leqslant 1$

$$0 \leqslant T_t f(x) = \int_S f(y) p_t(x, dy) \leqslant \int_S p_t(x, dy) = \mathbb{P}^x(X_t \in S) \leqslant 1.$$

It only remains to show that

$$\lim_{t \to 0} \|T_t f - f\| = 0$$

for each $f \in L^2(S,m)$. To this end, we observe that for any $E \in \mathscr{B}(S)$ and $x \in S$,

$$\begin{aligned}
\lim_{t\downarrow 0} T_t \mathbb{1}_E(x) &= \lim_{t\downarrow 0} \int_S \mathbb{1}_E(y) p_t(x,dy) \\
&= \lim_{t\downarrow 0} \int_E p_t(x,dy) \\
&= \lim_{t\downarrow 0} \mathbb{P}^x(X_t \in E) \\
&= \mathbb{P}^x(\lim_{t\downarrow 0} X_t \in E) \\
&= \mathbb{P}^x(X_0 \in E) \\
&= \mathbb{1}_E(x).
\end{aligned}$$

Above we have used the right continuity of sample paths, the strong Markov property and the bounded convergence theorem. It follows trivially that the above formula agrees for any simple function in $L^2(S,m)$. From here, a standard density argument shows that the convergence holds in the desired sense for all $f \in L^2(S,m)$. □

By virtue of the Theorem 5 and the theorem above, we have the following Corollary:

Corollary 3. *Let M be an m-symmetric Hunt process with state space $(S, \mathscr{B}(S))$. Then m has an associated Dirichlet form $\mathcal{E}$ with domain $D(\mathcal{E}) \subseteq L^2(S,m)$.*

As it takes the better part of 200 pages to prove and for cultural reasons, we state an important partial converse to the above corollary as a fact, see [6, Theorem 6.2.1]. It corresponds to the dotted arrow on the cover page.

Fact 7. *Given a regular Dirichlet form $\mathcal{E}$ on $L^2(S,m)$, there exists an m-symmetric Hunt process M with state space $(S, \mathscr{B}(S))$ whose Dirichlet form is the given one.*

3.3 Some fun facts

This short section is included to highlight a couple of ways in which Dirichlet forms can be used say something probabilistic. We take everything herein for granted. For details, see [13], [14] and references therein.

Let S be a separable locally compact Hausdorff space equipped with a positive Radon measure m with $\mathrm{Supp}(m) = S$. For an m-symmetric Hunt process M with state space $(S, \mathscr{B}(S))$, its associate Dirichlet form $\mathcal{E}$ defines and "intrinsic" metric on S by

$$\rho(x,y) = \sup\{u(x) - u(y) : u \in D(\mathcal{E}) \cap C_0(S), \mu_{\langle u \rangle} \leqslant m\}$$

where $\mu_{\langle u \rangle}$ is an "energy" measure defined in [13]. For some fixed $x_0 \in S$, let

$$v(r) = m(\{x \in S : \rho(x, x_0) \leqslant r\}).$$

We have the following:

Fact 8. *Take M and v as above. If*

$$\int_1^\infty \frac{r}{v(r)} dr = \infty$$

then each M_x is recurrent. Moreover, for any $E, F \in \mathscr{B}(S)$ and $t > 0$,

$$\int_E \mathbb{P}^x(X_t \in F) m(dx) \leqslant \sqrt{m(E)m(F)} \exp\left(-\frac{\rho^2(E,F)}{2t}\right).$$

A Appendix

Fact 9 (Uniform Boundedness Principle)**.** *Let X be a Banach space with norm $\|\cdot\|$ and $\{\Lambda_\alpha\}_{\alpha\in A} \subseteq \mathcal{B}(X)$. Then either there exists an $M < \infty$ such that*

$$\|\Lambda_\alpha\|_{op} \leqslant M$$

for all $\alpha \in A$, or

$$\sup_{\alpha\in A} \|\Lambda_\alpha x\| = \infty$$

for all x in some dense G_δ set in X.

The above fact is also called the Banach-Steinhaus theorem. See [10, Theorem 5.8] for a proof.

References

[1] Bass, R. F. *Stochastic Processes.* Cambridge University Press, Cambridge, UK, 2011.

[2] Davies, E. B. Lp spectral theory of higher-order elliptic differential operators. *Bull. London Math. Soc. 29* (1997), 513–546.

[3] Davies, E. B. *Linear Operators and their Spectra.* Cambridge University Press, Cambridge, 2007.

[4] Dynkin, E. B. *Markov Processes*, vol. I. Springer-Verlag, 1965.

[5] Evans, L. C. *Partial Differential Equations.* Amer. Math. Soc., Providence, Rhode Island, 2010.

[6] Fukushima, M. *Dirichlet Forms and Markov Processes.* Kodansha Scientific Books, Tokyo, 1980.

[7] Gross, L. Lecture notes on functional analysis. Math 713, Spring 2012, 2012.

[8] Lieb, E. H., and Loss, M. *Analysis*, second ed. Amer. Math. Soc., Providence, R.I., 2001.

[9] Randles, E., and Saloff-Coste, L. On the convolution powers of complex functions on z. *arXiv* (2012). Submitted.

[10] Rudin, W. *Real and Complex Analysis*, 3rd ed. WCB McGraw Hill, Boston, Massachusetts, 1987.

[11] Rudin, W. *Functional Analysis*, 2nd ed. McGraw-Hill, Inc., 1991.

[12] Steinhurst, B. A. Math7770f10 lecture notes.

[13] Sturm, K. Analysis on local dirichlet spaces. *I.J. Reine Angew. Math. 456* (1994), 173–196.

[14] Sturm, K. Monotone approximation of energy functionals for mappings into metric spaces. *I.J. Reine Angew. Math. 486* (1997), 129–151.

[15] Varopoulos, N., Saloff-Coste, L., and Coulhon, T. *Analysis and Geometry on Groups*, vol. 100 of *Cambridge Tracts in Math.* Cambridge University Press, Cambridge, 1992.