# Feller Processes and Their Applications

Kihyun Han

August 23, 2021

## 1 Introduction

### 1.1 Purpose

In this report, we analyze Feller processes, which are kinds of Markov processes that have special useful properties. First, we define Feller-Dynkin semigroups and demonstrate their relation with the transition function of Feller processes. Next, utilizing the generator and the resolvent of a Feller process, we explain the inverse property and the Hille-Yosida theorem. After describing the construction of a Feller process from a Feller-Dynkin semigroup, we prove the strong Markov property of Feller processes, which results in more practical usage of stopping time, such as the Blumenthal's 0-1 law and the Dynkin formula. Also, the discussion with characteristic functions enables us to obtain the Feller-Dynkin diffusions. Lastly, we apply the overall properties and theorems to Brownian motions and Lévy processes, which are specific examples of Feller processes.

### 1.2 Methods

In this report, we summarize the sections "TRANSITION FUNCTIONS AND RESOLVENTS" and "FELLER-DYNKIN PROCESSES" (pp. 227-263) in the book "Diffusions, Markov Processes and Martingales" by L. Rogers and D. Williams. Furthermore, we rearrange the contents of the book in order to offer the better understanding to the concept and the application of Feller processes, and we provide more rigorous proofs for the theorems in the book to resolve several discrepancies that can confuse the readers.

## 2 Results

### 2.1 Feller-Dynkin semigroups and Feller processes

We consider $E$ as a locally compact Hausdorff space with countable base. This basic condition allows us to attain two basic results, the uniqueness of the limit of a sequence and separability. Also, the coffin state or the infinity state $\partial$ enables us to attain $E_\partial := E \cup \partial$, the one-point compactification of $E$.

*Notation.* Here, we use the following notation for the function spaces on $E$.
(1) $C(E)$ is the space of all (real-valued) continuous functions on $E$.
(2) $C_b(E)$ is the space of all bounded and continuous functions on $E$.
(3) $C_0(E)$ is the space of all bounded and continuous functions on $E$ which vanish at infinity.
(4) $C_K(E)$ is the space of all continuous functions on $E$ with compact support.

**Definition 2.1** (Feller-Dynkin semigroup). *A Feller-Dynkin(FD) semigroup $\{P_t : t \geq 0\}$ is defined by a semigroup of linear operators on $C_0(E)$ with following properties;*
*(i) $P_t : C_0(E) \to C_0(E)$;*
*(ii) (sub-Markov property) $\forall f \in C_0(E),\ 0 \leq f \leq 1$, then $0 \leq P_t f \leq 1$;*
*(iii) (strong continuity) $\forall f \in C_0(E),\ \|P_t f - f\| \to 0$ as $t \to 0+$;*
*(iv) (semigroup property) $\forall s, t \geq 0, P_s P_t = P_{s+t}$, where $P_0 = I$, the identity on $C_0$.*

Note that the norm $\|\cdot\|$ for functions is the supremum norm. Hence, the convergence of functions in the norm $\|\cdot\|$ is automatically uniform. Also, sub-Markov property implies the monotonicity,

$$f \leq g \ \Rightarrow\ P_t f \leq P_t g,$$

because if $f = g$, then $P_t(g - f) = P_t 0 = 0$ by linearity, and otherwise,

$$P_t g - P_t f = \|g - f\| P_t\left(\frac{g - f}{\|g - f\|}\right) \geq 0.$$

Due to sub-Markov property, the contraction property also holds:

$$\forall f \in C_0(E),\ \|P_t f\| \leq \|f\|.$$

This is because $P_t 0 = 0$, and if $\|f\| > 0$, then

$$0 \leq \frac{|f|}{\|f\|} \leq 1 \ \Rightarrow\ 0 \leq P_t\left(\frac{|f|}{\|f\|}\right) \leq 1 \ \Rightarrow\ \|P_t f\| \leq \|P_t |f|\| = \|f\| P_t\left(\frac{|f|}{\|f\|}\right) \leq \|f\|.$$

Generally, we may refer to a Feller-Dynkin semigroup $\{P_t : t \geq 0\}$ as a strongly continuous contraction semigroup (SCCSG) if the function space $C_0(E)$ in Definition 2.1 is replaced by a general Banach space. In addition, we have to realize that there is a correspondence between a FD semigroup and a transition function $P_t : E \times \mathscr{E} \to [0, 1]$ of a Feller process. For $x \in E$ and $\Gamma \in \mathscr{E}$, $P_t(x, \Gamma)$ can be interpreted as a probability of the process that starts from $x$ to lie in $\Gamma$ at time $t$, so this is an intuitive method to understand the process. The following relation can justify the discussion with FD semigroups instead of transition functions of Feller processes;

$$P_t f(x) = \int_E P_t(x, dy) f(y).$$

Now, we define the (infinitesimal) generator and the resolvent of a FD semigroup, which play an essential role in the overall discussion in this report.

**Definition 2.2** ((infinitesimal) generator of a FD semigroup). *For a FD semigroup $\{P_t : t \geq 0\}$, we define $\mathscr{D}(\mathscr{G})$ as the set of all functions $f \in C_0(E)$ for which there exists a function $g \in C_0(E)$ satisfying*

$$\lim_{t \to 0+} \|t^{-1}(P_t(f) - f) - g\| = 0.$$

*We then set $\mathscr{G} f = g$, and the (infinitesimal) generator $\mathscr{G} : \mathscr{D}(\mathscr{G}) \to C_0(E)$ is a mapping $f \mapsto \mathscr{G} f$.*

Conceptually, the generator of a FD semigroup can be identified as the derivative of $P_t$ with respect to $t$. Thus, in case $E$ is a finite state space, it is easily shown by solving a differential equation that $P_t$ simply becomes an exponential function with repect to $t$, more specifically,

$$P_t(x, y) = \exp(t\mathscr{G}(1_y)(x)).$$

**Definition 2.3** (resolvent). *We call $\{R_\lambda : \lambda > 0\}$, a family of bounded linear operators $R_\lambda : C_0(E) \to C_0(E)$, the resolvent of the FD semigroup $\{P_t : t \geq 0\}$ if*

$$R_\lambda f = \int_0^\infty e^{-\lambda t} P_t f dt.$$

**Theorem 2.1** (strongly continuous contraction resolvent). *Since a FD semigroup $\{P_t : t \geq 0\}$ is a SCCSG, the resolvent $\{R_\lambda\}$ of $\{P_t : t \geq 0\}$ is a strongly continuous contraction resolvent (SCCR) in the sense that*
*(i) (contraction) $\forall \lambda > 0$, $\|\lambda R_\lambda\| \leq 1$*
*(ii) (strong continuity) $\forall f \in C_0(E)$, $\lim_{\lambda\to\infty} \|\lambda R_\lambda f - f\| = 0$*
*(iii) (resolvent equation) $\forall \lambda, \mu > 0$, $R_\lambda - R_\mu + (\lambda - \mu) R_\lambda R_\mu = 0$*

*Proof.* The contraction property of the resolvent holds since for any $f \in C_0(E)$,

$$\|\lambda R_\lambda f\| \leq \int_0^\infty \lambda e^{-\lambda t} \|P_t f\| dt \leq \sup_{t>0} \|P_t f\| \leq \|f\|.$$

Moreover, the strong continuity comes from the fact that by Lebesgue's dominated convergence theorem,

$$\|\lambda R_\lambda f - f\| \leq \int_0^\infty \lambda e^{-\lambda t} \|P_t f - f\| dt = \int_0^\infty e^{-s} \|P_{s/\lambda} f - f\| ds \to 0 \text{ as } \lambda \to \infty.$$

Finally, the resolvent equation results from the following equations with the help of Fubini theorem.

$$\begin{aligned}
(\lambda - \mu) R_\lambda R_\mu f(x) &= (\lambda - \mu) \int_0^\infty e^{-\lambda s} \int_E P_s(x, dy) \int_0^\infty e^{-\mu t} P_t f(y) dt\ ds \\
&= (\lambda - \mu) \int_0^\infty \int_0^\infty e^{-\lambda s - \mu t} \int_E P_s(x, dy) P_t f(y) dt\ ds \\
&= (\lambda - \mu) \int_0^\infty \int_0^\infty e^{-\lambda s - \mu t} P_{s+t} f(x) dt\ ds \\
&= (\lambda - \mu) \int_0^\infty \int_0^u e^{-\lambda u} e^{(\lambda-\mu)t} P_u f(x) dt\ du \\
&= \int_0^\infty e^{-\lambda u} (e^{(\lambda-\mu)u} - 1) P_u f(x) du \\
&= R_\lambda f(x) - R_\mu f(x).
\end{aligned}$$

□

The following theorem links the generator and the resolvent of a FD semigroup.

**Theorem 2.2** (inverse property). *Suppose that $\{P_t : t \geq 0\}$ is a FD semigroup. Then, for any $\lambda > 0$, $R_\lambda : C_0(E) \to \mathscr{D}(\mathscr{G})$ and $(\lambda - \mathscr{G}) : \mathscr{D}(\mathscr{G}) \to C_0(E)$ are inverses.*

*Proof.* First of all, we prove that for any $g \in C_0(E)$, $(\lambda - \mathscr{G}) R_\lambda g = g$. Then, it directly follows that $R_\lambda g$ lies in $\mathscr{D}(\mathscr{G})$. Note that

$$\begin{aligned}
t^{-1}(R_\lambda g - e^{-\lambda t} P_t R_\lambda g)(x) &= t^{-1}[\int_0^\infty e^{-\lambda s} P_s g ds - e^{-\lambda t} P_t \int_0^\infty e^{-\lambda s} P_s g ds](x) \\
&= t^{-1}[\int_0^\infty e^{-\lambda s} P_s g(x) ds - e^{-\lambda t} \int_0^\infty \int_E P_t(x, dy) e^{-\lambda s} P_s g(y) ds]
\end{aligned}$$

$$
\begin{aligned}
&= t^{-1}[\int_0^\infty e^{-\lambda s} P_s g(x) ds - \int_0^\infty P_{s+t} g(x) e^{-\lambda(s+t)} ds] \\
&= t^{-1}[\int_0^\infty e^{-\lambda s} P_s g(x) ds - \int_t^\infty P_s g(x) e^{-\lambda s} ds] \\
&= t^{-1}[\int_0^t e^{-\lambda s} P_s g ds](x).
\end{aligned}
$$

Also, by the strong continuity of $P_t$ , we have, as $t \to 0+$,

$$
\|t^{-1} \int_0^t e^{-\lambda s} P_s g ds - g\| \le t^{-1} \int_0^t e^{-\lambda s} \|P_s g - g\| ds \ \le t^{-1} \int_0^t \|P_s g - g\| ds.
$$

Hence we obtain that

$$
\begin{aligned}
(\lambda - \mathscr{G}) R_\lambda g &= \lim_{t \to 0+} [\lambda R_\lambda g - t^{-1}(P_t R_\lambda g) e^{-\lambda t}] \\
&= \lim_{t \to 0+} [\lambda R_\lambda g + t^{-1}(R_\lambda g - e^{-\lambda t} P_t R_\lambda g) + t^{-1}(e^{-\lambda t} - 1) R_\lambda g] \\
&= g,
\end{aligned}
$$

where $\lim_{t \to 0+}$ above is the uniform limit, that is, the limit in norm $\| \cdot \|$.

In order to prove the converse, suppose that $f \in \mathscr{D}(\mathscr{G})$. Then for any $x \in E$, we have that

$$
\begin{aligned}
R_\lambda t^{-1}(P_t f - f)(x) &= [\int_0^\infty P_s[t^{-1}(P_t f - f) e^{-\lambda s}] ds](x) \\
&= [\int_0^\infty t^{-1}(P_{t+s} f - P_s f) e^{-\lambda s} ds](x) \\
&= t^{-1}[e^{\lambda t} \int_t^\infty e^{-\lambda s} P_s f ds - \int_0^\infty e^{-\lambda s} P_s f ds](x) \\
&= t^{-1}[-e^{\lambda t}\{\int_0^t e^{-\lambda s}(P_s f - f)(x) ds + \int_0^t e^{-\lambda s} f(x)\} + (e^{\lambda t} - 1) R_\lambda f(x) ds] \\
&\to 0 - f(x) + \lambda R_\lambda f(x),
\end{aligned}
$$

as $t \to 0+$ due to the strong continuity of $\{P_t : t \ge 0\}$. Since the limit of the first equation as $t \to 0+$ is $R_\lambda \mathscr{G} f$, we obtain that

$$
R_\lambda \mathscr{G} f = -f + \lambda R_\lambda f \ \Rightarrow \ R_\lambda(\lambda - \mathscr{G}) f = f.
$$

□

## 2.2 Hille-Yosida theorem

As can be seen in Theorem 2.1, a FD semigroup determines a SCCR, which exactly is its resolvent. Conversely, in the absense of a FD semigroup, the Hille-Yosida theorem is a method to obtain a FD semigroup from a resolvent. Note that the Hille-Yosida theorem holds for general Banach function space $B_0$ instead of $C_0(E)$.

**Theorem 2.3** (Hille-Yosida). *Let $\{R_\lambda : \lambda > 0\}$ be a SCCR family on $B_0$. Then there exists a unique SCCSG $\{P_t : t \ge 0\}$ on $B_0$ such that*

$$
\int_{[0,\infty)} e^{-\lambda t} P_t f dt = R_\lambda f, \qquad \forall \lambda > 0, \forall f \in B_0.
$$

*Indeed, if we define*

$$G_\lambda = \lambda(\lambda R_\lambda - I),$$

$$P_{t,\lambda} = \exp(tG_\lambda) = e^{-\lambda t}\sum_{n=0}^{\infty}(\lambda t)^n(\lambda R_\lambda)^n/n!,$$

*then for each $f$ in $B_0$,*

$$P_t f = \lim_{\lambda\to\infty} P_{t,\lambda} f, \;\forall t \geq 0.$$

Before we provide the proof of Theorem 2.3, we may mention some preliminaries for the proof of the theorem. First, the range $R_\lambda(B_0)$ of $R_\lambda$ is a space $\mathscr{R}$ independent of $\lambda$. This follows from the resolvent equation since

$$\forall\lambda, \mu > 0, \; R_\mu f = R_\lambda(f + (\lambda-\mu)R_\mu f) \in R_\lambda(B_0)$$

$$\Rightarrow \; R_\mu(B_0) \subset R_\lambda(B_0).$$

In addition, $\mathscr{R}$ is dense in $B_0$ because for all $f \in B_0$, $\mu R_\mu f \to f$ as $\mu \to \infty$. Moreover, the map $R_\lambda : B_0 \to \mathscr{R}$ is a bijection due to the fact that for any $h \in B_0$ with $R_\mu h = 0$, the resolvent equation implies that

$$R_\lambda h = R_\mu h + (\lambda - \mu)R_\lambda R_\mu h = 0,$$

so $h = \lim_{\lambda\to\infty} \lambda R_\lambda h = 0$.

Finally, since there exists the inverse $R_\lambda^{-1}$ of $R_\lambda$ for any $\lambda > 0$, we can define $\mathscr{G} : \mathscr{R} \to B_0$ by $\mathscr{G} = \lambda - R_\lambda^{-1}$ for some $\lambda > 0$. By the resolvent equation, $\mathscr{G}$ is independent of $\lambda > 0$. Thus, $\mathscr{D}(\mathscr{G}) = \mathscr{R}$ and $(\lambda - \mathscr{G})^{-1} = R_\lambda$.

**Lemma 2.1.** *For any $f \in B_0$, $f \in \mathscr{D}(\mathscr{G})$ if and only if $g := \lim_{\lambda\to\infty} G_\lambda f$ exists. Also, if either of the equivalent conditions holds, then $\mathscr{G}f = g$.*

*Proof.* First, suppose that $f \in \mathscr{D}(\mathscr{G})$, then $R_\lambda(\lambda - \mathscr{G}) = I$. Thus, $\lambda R_\lambda - I = R_\lambda\mathscr{G}$. Hence, $G_\lambda f = \lambda R_\lambda \mathscr{G} f \to \mathscr{G} f$ as $\lambda \to \infty$.
For the opposite direction, $g = \lim_{\lambda\to\infty} G_\lambda f$. Then

$$R_\mu G_\lambda f = \lambda(\frac{\mu R_\mu - \lambda R_\lambda}{\lambda - \mu})f \to \mu R_\mu f - f$$

as $\lambda \to \infty$. Also, due to the contraction property of $R_\mu$, we have $R_\mu G_\lambda f \to R_\mu g$ as $\lambda \to \infty$. Hence, $\mu R_\mu f - f = R_\mu g$, so $f = R_\mu(\mu f - g) \in \mathscr{D}(\mathscr{G})$. From $f = R_\mu(\mu - \mathscr{G})f$, we attain that $\mathscr{G}f = g$. □

*Proof of Theorem 2.3.* Note that
$(i) P_{s,\lambda}P_{t,\lambda} = \exp(G_\lambda t + G_\lambda s) = P_{s+t,\lambda}$.
$(ii) \lim_{h\to 0+} h^{-1}(P_{h,\lambda} - I) = \lim_{h\to 0+} h^{-1}(\exp(hG_\lambda) - I) = G_\lambda$ ($G_\lambda$ is a bounded operator).
$(iii) P_{t,\lambda} - I = \int_0^t \exp(G_\lambda s)G_\lambda ds = \int_0^t P_{s,\lambda}G_\lambda ds.$

Since $\|\lambda R_\lambda\| \leq 1$, it is clear that

$$\|P_{t,\lambda}\| \leq e^{-\lambda t}\sum_{n=0}^{\infty}(\lambda t)^n\|\lambda R_\lambda\|^n/n! \leq 1.$$

Also, by exchanging $\lambda$ and $\mu$ in the resolvent equation, we have that

$$R_\lambda R_\mu = R_\mu R_\lambda \ \Rightarrow \ G_\lambda G_\mu = G_\mu G_\lambda \ \Rightarrow \ P_{t,\lambda} P_{t,\mu} = P_{t,\mu} P_{t,\lambda}$$

Then,

$$P_{t,\lambda} f - P_{t,\mu} f = \sum_{k=1}^{n} P_{\frac{k-1}{n}t,\lambda} P_{\frac{n-k}{n}t,\mu} [P_{\frac{t}{n},\lambda} - P_{\frac{t}{n},\mu}]$$

By taking norm in the above equation, since $\|P_{t,\lambda}\| \le 1$, we have that

$$\|P_{t,\lambda} f - P_{t,\mu} f\| = n \|P_{\frac{t}{n},\lambda} - P_{\frac{t}{n},\mu}\|.$$

Letting $n \to \infty$ and combining with (ii), we attain that

$$\|P_{t,\lambda} f - P_{t,\mu} f\| \le t \|G_\lambda f - G_\mu f\|.$$

This indicates that if $t$ lies in any compact interval, then for any $f \in \mathscr{D}(\mathscr{G})$, $P_{t,\lambda} f$ as a function of $t$ is uniformly Cauchy, thus there exists $P_t f = \lim_{\lambda \to \infty} P_{t,\lambda} f$ as the uniform limit. Since $P_{t,\lambda}$ was defined by the exponential function, which is continuous, the mapping $t \mapsto P_t f$ is also continuous. This can be extended to any $f \in B_0$ because $\mathscr{D}(\mathscr{G})$ is dense in $B_0$. Finally, by letting $\mu \to \infty$ in the following equation

$$\int_e^{-\lambda t} P_{t,\mu} f dt = (\lambda - G_\mu)^{-1} f = \frac{\mu^2}{(\lambda + \mu)^2} R_{\frac{\lambda \mu}{\lambda + \mu}} + (\lambda + \mu)^{-1} I,$$

we attain that $\int_0^\infty e^{-\lambda t} P_t f dt = R_\lambda f$. In other words, combining the result with (i), (ii), and $\|P_{t,\lambda}\| \le 1$, we have that $\{P_t : t \ge 0\}$ is a Feller semigroup with its resolvent $R_\mu$. We add a final note that by Theorem 2.2, $\{P_t\}$ is uniquely determined by $(\mathscr{G}, \mathscr{D}(\mathscr{G}))$. $\square$

The following corollary provides an easier criterion to determine whether a semigroup enjoys strong continuity.

**Corollary 2.1.** *If $\{P_t : t \ge 0\}$ on $C_0(E)$ satisfies the sub-Markov property and semigroup property, then the following apparently weaker condition, the pointwise continuity,*

$$\forall f \in C_0(E), \ \forall x \in E, \ \lim_{t \to 0+} P_t f(x) = f(x)$$

*implies strong continuity.*

*Proof.* Note that we do not know if $\{P_t\}$ is an SCCSG. For $f \in C_0(E)$, since $\|e^{-\lambda t} P_t f\| \le \|P_t f\| \le \|f\|$, we have

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x) dt \ \to \ \int_0^\infty e^{-\lambda t} P_t f(y) dt = R_\lambda f(y)$$

as $x \to y$ by the dominated convergence theorem. Thus, $R_\lambda f$ is a continuous function. Since $P_t : C_0(E) \to C_0(E)$, $P_t f \in C_0(E)$. Hence, by dominated convergence theorem, as $x \to \infty$,

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x) dt \to 0.$$

That is, $R_\lambda f \in C_0(E)$, so $R_\lambda : C_0(E) \to C_0(E)$ is a contraction resolvent.

Then by Hille-Yosida theorem, we may find a subspace $B_0$ where $\{P_t : t \ge 0\}$, $\{R_\lambda : \lambda > 0\}$ are strongly continuous, and $B_0$ is actually given by $B_0 = \overline{R_\lambda C_0(E)}$, which is equal

for any $\lambda > 0$ If $B_0 = C_0(E)$, then the proof is done. Assume that $B_0 \neq C_0(E)$. Then by the Hahn-Banach theorem, we may find a nonzero linear functional $\phi : C_0(E) \to \mathbb{R}$ such that $\phi(B_0) = \{0\}$, that is, $\phi$ annihilates $B_0$. Then for any $\lambda > 0$ and $f \in C_0(E)$, we have $\lambda R_\lambda f \in B_0$, so $\phi(\lambda R_\lambda f) = 0$
However,

$$\lambda R_\lambda f(x) = \int_0^\infty \lambda e^{-\lambda t} P_t f(x) dt = \int_0^\infty e^{-s} P_{\frac{s}{\lambda}} f(x) ds.$$

By assumption, $P_{\frac{s}{\lambda}} f(x) \to f(x)$ as $\lambda \to \infty$, and $|P_{\frac{s}{\lambda}} f(x)| \leq \|f\|$. Thus, by dominated convergence theorem, $\lambda R_\lambda f(x) \to f(x)$, and applying the dominated convergence theorem once again, we obtain a contradiction that

$$0 = \phi(\lambda R_\lambda f) = \int \lambda R_\lambda f(x) \mu(dx) \to \int f(x) \mu(dx) \phi(f)$$

where $\mu$ is obtained from the Riesz representation theorem.

□

### 2.3 Construction of Feller processes

In this section, we may turn a FD semigroup $\{P_t : t \geq 0\}$ on $C_0(E)$ into a Feller process in the sense that there exists a strong Markov, $E_\partial$-valued R-process $X$ with transition function $\{P_t\}$. The strong Markov property is dealt with in meticulous detail in the next section. There are two steps for achieving this goal.

The first step is to construct the Daniell-Kolmogorov process Y associated with $\{P_t\}$. Consider $\omega : [0, \infty] \to E_\partial$ and $Y_t(\omega) = \omega(t)$. By the Daniell-Kolmogorov theorem, for any initial distribution measure $\mu$, there exists the unique probability measure $\mathbf{P}^\mu$ satisfying $\forall n \in \mathbb{N},\ 0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n,\ x_0, x_1, \cdots, x_n \in E_\partial$,

$$\begin{aligned} &\mathbf{P}^\mu[Y(i) \in dx_i\ \forall i = 0, 1, \cdots, n] \\ &= \mu(dx_0) P_{t_1}^{+\partial}(x_0, dx_1) P_{t_2 - t_1}^{+\partial}(x_1, dx_2) \cdots P_{t_n - t_{n-1}}^{+\partial}(x_{n-1}, dx_n) \end{aligned}$$

with the assistance of compactness of $E_\partial$, a one point compactification of $E$. Using $Y_t$, we define two provisional filtrations $\mathscr{G} := \sigma\{Y_s : s \geq 0\}$ and $\mathscr{G}_t^0 := \sigma\{Y_s : s \leq t\}$. Let $b\mathscr{G}$ be the space of all $\mathscr{G}$ measurable bounded functions. Then by applying monotone class lemma for functions, we obtain the Markov property that for any $t \geq 0$, $\eta \in b\mathscr{G}$, and $\mu \in Pr(E_\partial)$ we have, $\mathbf{P}^\mu$ almost surely,

$$E^\mu[\eta \circ \theta_t | \mathscr{G}_t^0] = E^{Y_t} \eta$$

where $\theta_t : \Omega \to \Omega$ is the shift operator, i.e. $\theta_t(\omega(s)) = \omega(t + s)$.
The second step is to regularize the path of $Y_t$. The final result $X_t$ is an R-process, that is, for all $\omega \in \Omega$, $x_t := X_t(\omega)$ is an R-function, a function of $t$ satisfying

$$x_t = \lim_{h \to t+} x_h \text{ for every } t \geq 0,$$

$$x_{t-} = \lim_{s \to t-} x_s \text{ for finite } t > 0.$$

We take $g \in C_0^+ := \{g \in C_0 : g \geq 0\}$ and $h = R_1 g$ (the resolvent with $\lambda = 1$). Due to the convergence of a non-negative supermartingale $\{e^{-t} P_t h : t \geq 0\}$, we know that for all $t$, we can obtain an R-function of $t$ by taking the limit $\lim_{\mathbb{Q} \ni q \to t+} e^{-q} h(Y_q)$ along rational numbers $q$. By separability of $E$, we have the separability of $C_0^+$.

Now, let $\{g_0, g_1, \cdots\}$ be a countable dense subset of $C_0^+$. Also, let $h_n = R_1 g_n$ and $\mathscr{H} = \{h_0, h_1, \cdots\}$. Note that for any $f \in \mathscr{D}(\mathscr{G})$, there exists $g \in C_0$ with $f = R_1 g$, and by denseness, we may find $g_k, g_l$ with $g_k - g_l \to g$. Thus, by continuity of $R_1$, we have $h_k - h_l \to f$, which implies that $\mathscr{H} - \mathscr{H}$ is dense in $\mathscr{D}(\mathscr{G})$. Combining with the fact that $\mathscr{D}(\mathscr{G})$ is dense in $C_0$ as in the Hille-Yosida theorem, we get the denseness of $\mathscr{H} - \mathscr{H}$ in $C_0$, which indicates that $h_n \geq 0$ on $E$ and that $\mathscr{H}$ separates points of $E_\partial$.

For a fixed rational number $t$, assume that $\lim_{\mathbb{Q} \ni q \to t+} Y_q$ does not exist $\mathbf{P}^\mu$-almost surely. In other words, there exists different sequences $q_n^{(1)}, q_n^{(2)}$ with rational numbers and $X_t^{(1)}, X_t^{(2)}$ such that

$$X_t^{(1)} := \lim_{\mathbb{Q} \ni q_n^{(1)} \to t+} Y_{q_n^{(1)}} \neq \lim_{\mathbb{Q} \ni q_n^{(2)} \to t+} Y_{q_n^{(2)}} =: X^{(2)}$$

$\mathbf{P}^\mu$-almost surely. Let

$$\Omega* := \{\omega : X_t^{(1)}(\omega) \neq X_t^{(2)}(\omega)\},$$

then $\mathbf{P}^\mu(\Omega*) \neq 0$. Since $\mathscr{H}$ separate points, for each $\omega \in \Omega*$, we may find $h \in \mathscr{H}$ such that $h(X_t^{(1)}(\omega)) \neq h(X_t^{(2)}(\omega))$. However, by continuity of $h_k \in \mathscr{H}$,

$$h_k(X_t^{(1)}) := \lim_{\mathbb{Q} \ni q_n^{(1)} \to t+} h_k(Y_{q_n^{(1)}}) = \lim_{\mathbb{Q} \ni q_n^{(2)} \to t+} h_k(Y_{q_n^{(2)}}) = h_k(X_t^{(2)})$$

$\mathbf{P}^\mu$-almost surely, so $\Omega_k := \{\omega : h_k(X_t^{(1)}(\omega)) \neq h_k(X_t^{(2)}(\omega))\}$ is $\mathbf{P}^\mu$ measure 0. Since $\mathscr{H}$ is a countable set, we obtain a contradiction that

$$\Omega* \subset \bigcup_{k \geq 1} \Omega_k \;\Rightarrow\; 0 < \mathbf{P}^\mu[\Omega*] \subset \mathbf{P}^\mu[\bigcup_{k \geq 1} \Omega_k] = 0.$$

Then for any $\mu$, we can conclude that $\mathbf{P}^\mu$-a.s. $X_t := \lim_{\mathbb{Q} \ni q \to t+} Y_q$ exists. Since we took $t$ as countable rational numbers, $X_t$ exists for any rational $t$, $\mathbf{P}^\mu$-almost surely. Now, we need to fill $X_t$ for irrational numbers. This is obvious by taking right limit over $X_t$ with rational $t$ for $\mathbf{P}^\mu$ almost surely. This method guarantees $\{X_t\}$ to be an R-process, which involves the right-continuity of its path. The case $\omega$ where the limit does not exist is easily resolved if we simply assign $X_t(\omega) = \partial$.

Then for all $f_1, f_2 \in C_0(E)$, as $\mathbb{Q} \ni q \to t+$ , $f_2(Y_q) \to f_2(X_t)$ pointwisely. Then by dominated convergence theorem,

$$\begin{aligned} E^\mu[f_1(Y_t) f_2(X_t)] &= \lim_{\mathbb{Q} \ni q \to t+} E^\mu[f_1(Y_t) f_2(Y_q)] \\ &= \lim_{\mathbb{Q} \ni q \to t+} E^\mu[f_1(Y_t) P_{q-t} f_2(Y_t)] \\ &= E^\mu[f_1(Y_t) f_2(Y_t)] \end{aligned}$$

By monotone-class arguments for functions, this can be extended to $f \in b(\mathscr{E}_\partial \times \mathscr{E}_\partial)$ that $E^\mu[f(Y_t, X_t)] = E^\mu[f(Y_t, Y_t)]$. Hence, $\mathbf{P}^\mu[X_t = Y_t] = 1$ by putting $f(p, q) = 1_{p=q}$, which implies the Markov property for $X_t$ as well as $Y_t$.

Note that the set $C_0(E)$ can be substituted with $C_0(E_\partial)$ by imposing 0 for any $f(\partial) = 0$. Then our new process $X$ satisfies

$$X_{t-} = \partial \text{ or } X_t = \partial \Rightarrow X_u = \partial, \; \forall u \geq t$$

in accordance with the conceptual meaning of the coffin state $\partial$. From now on, we use the following set-up.

**Definition 2.4** (canonical Feller processes)**.** *Let $\Omega$ be the space of paths of R-processes $\omega : [0, \infty) \to E_\partial$. For $\omega \in \Omega$ and $t \geq 0$, define $X_t(\omega) = \omega(t)$ and*

$$\mathscr{F}^0 := \sigma\{X_s : 0 \leq s < \infty\} = \sigma\{X_s : 0 \leq s \leq\}$$

$$\mathscr{F}_t^0 := \sigma\{X_s : s \leq t\}.$$

*Then, $X = (\Omega, \mathscr{F}, X_t : 0 \leq t \leq \infty, \{\mathbf{P}^\mu : \mu \in Pr(E_\partial)\})$ is called the canonical Feller process associated with the FD semigroup $\{P_t\}$.*

**Definition 2.5** (lifetime)**.** *The random time $\zeta := \inf\{t : X_t = \partial\}$ is called the lifetime of $X$.*

## 2.4 Markov properties

In the previous section, it is guaranteed that the Feller process made from a FD semigroup has a right continuity and the simple Markov property. We refer to the simple Markov property as weak in order to place emphasis on the difference of those two properties.

**Theorem 2.4.** *(weak Markov property) For $t \geq 0$, $\xi \in b\mathscr{F}$, and $\eta \in b\mathscr{F}$,*

$$E^\mu[\theta_t\eta|\mathscr{F}] = E^{X_t}[\eta] \qquad \mathbf{P}^\mu\text{-a.s. and}$$

$$E^\mu[\xi\theta_t\eta] = E^\mu[\xi E^{X_t}\eta]$$

*In particular, $\forall f \in C_0, s \geq 0$,*

$$E^\mu[\xi f(X_{t+s})] = E^\mu[\xi P_s f(X_t)]$$

The strong Markov property has a similar equation as the weak Markov property, but the only difference is that we put the stopping time $T$ in place of $t$. We need to define a $\{\mathscr{F}_t+^0\}$ stopping time $T$ and the filtration $\mathscr{F}^0_{T+}$.

**Definition 2.6.** *A map $T : \Omega \to [0, \infty]$ is a $\{\mathscr{F}_+\}$ stopping time if for any $0 \leq t < \infty$, $\{\omega : T(\omega) \leq t\} \in \mathscr{F}_+$*

**Theorem 2.5.** *(strong Markov property) Let $T$ be an $\{\mathscr{F}^0_{T+}\}$ stopping time. For $\xi \in b\mathscr{F}_+$, and $\eta \in b\mathscr{F}$,*

$$E^\mu[\theta_T\eta|\mathscr{F}_{\mathscr{T}+}] = E^{X_T}[\eta] \qquad \mathbf{P}^\mu\text{-a.s. and}$$

$$E^\mu[\xi\theta_T\eta] = E^\mu[\eta E^{X_T}\eta]$$

*Proof.* For a given stopping time $T$, we set $T^{(n)}$ that can right-converge to $T$ in the following manner. Without doubt, $T^{(n)} := \infty$ if $T = \infty$, and

$$T^{(n)}(\omega) = \frac{k}{2^n}, \quad \text{if } \frac{k-1}{2^n} \leq T(\omega) < \frac{k}{2^n}.(k \in \mathbb{N})$$

Suppose $\Lambda \in \mathscr{F}^0_{T+}$ and let

$$\Lambda_{n,k} = \{\omega : T^{(n)}(\omega) = \frac{k}{2^n}\} \cap \Lambda \in \mathscr{F}^0_{\frac{k}{2^n}}$$

Put $\xi_{n,k} = 1_{\Lambda_{n,k}}$. For a given initial distribution $\mu$ on $E_\partial$ and $f \in C_0$, by applying the weak Markov property, we have

$$\begin{aligned} E^\mu[f(X_{T^{(n)}+s});\Lambda] &= \sum_{k\le\infty} E^\mu[f(X_{\frac{k}{2^n}+s});\Lambda_{n,k}] \\ &= \sum_{k\le\infty} E^\mu[P_s f(X_{\frac{k}{2^n}});\Lambda_{n,k}] \\ &= E^\mu[P_s f(X_{T^{(n)}});\Lambda] \end{aligned}$$

As we let $n\to\infty$, by the right continuity of paths of $X_t$, $T^{(n)} \to T+$ implies that

$$X(T^{(n)}+s) \to X(T+s),\ X(T^{(n)}) \to X(T).$$

Note that $P_s f \in C_0$ by the definition of FD semigroup. Then as $n \to \infty$

$$f(X_{T^{(n)}+s}) \to f(X_{T+s}),\ P_s f(X_{T^{(n)}}) \to P_s f(X_T).$$

Since all elements of $C_0$ are bounded, by the Lebesuge's dominated convergence theorem, the following holds as the limit of the expectations above.

$$E^\mu[f(X_{T+s});\Lambda] = E^\mu[P_s f(X_T);\Lambda].$$

Then for all $\xi \in b\mathscr{F}^0_{T+}$, there exist simple functions $\xi_n$ with $\xi_n \to \xi$, thus by the dominated convergence,

$$E^\mu[\xi f(X_{T+s})] = E[\xi P_s f(X_T)]. \tag{1}$$

For $f, g \in C_0$ and $\xi \in b\mathscr{F}^0_{T+}$, we have $\xi f(X_{T+s}) \in b\mathscr{F}^0_{T+s}$. Since $T+s$ is $\mathscr{F}^0_{t+}$ stopping time, by putting $s \leftarrow u$ and $\xi \leftarrow \xi f(X_T + s)$ in (1), we have

$$E^\mu[\xi f(X_{T+s})g(X_{T+s+u})] = E^\mu[\xi f(X_{T+s})P_u g(X_{T+s})]$$

Also, plugging in $f(x) \leftarrow f(x)P_u g(x)$ in (1), we obtain that

$$E^\mu[\xi f(X_{T+s})P_u g(X_{T+s})] = E^\mu[\xi P_s f(X_T)P_u g(X_T))]$$

Note that

$$\begin{aligned} P_s(f(z)P_u g(z)) &= \int_E f(y)P_u g(y)P_s(z,dy) \\ &= E^z[f(X_s)P_u g(X_s)] \end{aligned}$$

Combining all the results, we may deduce that

$$E^\mu[\xi f(X_{T+s})g(X_{T+s+u})] = E^\mu[\xi E^{X_T}[f(X_s)g(X_{s+u})]],$$

that is, for $\eta = f(X_s)g(X_s+u)$, $E^\mu[\xi\theta_T\eta] = E^\mu[E^{X_T}\eta]$. This can be extended to $\eta = f_1(X_{s_1})f_2(X_{s_2})\cdots f_n(X_{s_n})$ in the similar manner. Then $\eta$ can be generalized as the indicator function of a special cylinder $1_S$, then $1_E$ for any $E \in \mathscr{F}^0$. By the simple convergence theorem, for any $\eta \in b\mathscr{F}^0$, we may find simple functions $\{\eta_n\}$ that converges to $\eta$, thus by dominated convergence theorem,

$$\xi\theta_T\eta_n \to \xi\theta_T\eta \text{ and } \xi E^{X_T}\eta_n \to \xi E^{X_T}\eta.$$

Hence, we may conclude that

$$\forall \eta \in b\mathscr{F}^0, \qquad E^\mu[\xi\theta_T\eta] = E^\mu[E^{X_T}\eta].$$

□

We introduce two theorems, Blumenthal 0-1 law and Dynkin's formula, which can be proved with the help of the strong Markov property. Here, $\{\mathscr{F}_t\}$ is the intersection of all complete filtrations with respect to any initial distribution $\mu$. This was constructed in order to regard the debut time and the hitting time as stopping times.

**Definition 2.7** (debut time and hitting time)**.** *For a Borel set $B \subset E_\partial$, let*

$$D_B = \inf\{t \geq 0 : X_t \in B\} \text{ and } \; H_B = \inf\{t > 0 : X_t \in B\}.$$

*Respectively, these two are called the debut time and the hitting time of $B$ for $X$.*

**Theorem 2.6** (Blumenthal's 0-1 law)**.** *For all $x \in E_\partial$ and $\Lambda \in \mathscr{F}_0$, $\mathbf{P}^x(\Lambda) \in \{0,1\}$.*

*Proof.* Put $T = 0, \xi = 1_\Lambda$, and $\eta = 1_\Lambda$ in the strong Markov theorem. Then

$$\begin{aligned} E^x[1_\Lambda] &= E^x[1_\Lambda \theta_0 1_\Lambda] \\ &= E^x[1_\Lambda E^{X_0}[1_\Lambda]] \\ &= E^x[1_\Lambda E^x[1_\Lambda]] \\ &= (E^x[1_\Lambda])^2. \end{aligned}$$

Hence, $\mathbf{P}^x[\Lambda] = E^x[1_\Lambda] \in \{0,1\}$. □

**Corollary 2.2.** *If for $x \in E_\partial$, $T$ is an $\{\mathscr{F}^x_t\}$ stopping time, then $\mathbf{P}^x[T = 0] \in \{0,1\}$.*

*Proof.* It is clear since $\{T = 0\} \in \mathscr{F}_0$. □

**Definition 2.8.** *Let $b\mathscr{E}^\star_0$ be the space of measurable functions $f : E_\partial \to \mathbb{R}$ with $f(\partial) = 0$. For an $\mathscr{F}_t$ stopping time $T$,*

$$P^\lambda_T f(x) = E^x[e^{-\lambda T} f(X_T)], \qquad P_T f(x) = E^x[f(X_T)].$$

*Also, $P^\lambda_t = e^{-\lambda t} P_t$.*

Since $P_t f(x) = \int_E P_t(x, dy) f(y) = E^x f(X_t)$, we have $P_t : b\mathscr{E}^\star_0 \to b\mathscr{E}^\star_0$. The following lemma helps us prove the Dynkin's formula.

**Lemma 2.2.** *If $T$ is an $\{\mathscr{F}^0_t\}]$ stopping time, then $P^\lambda_T P^\lambda_t = P^\lambda_{T+t}$*

*Proof.* For any bounded function $f \in b\mathscr{E}^\star_0$ and $x \in E_\partial$, we have

$$\begin{aligned} P^\lambda_T P^\lambda_t f(x) &= E^x[e^{-\lambda T} P^\lambda_t f(X_T); T < \infty] + E^x[e^{-\lambda T} P^\lambda_t f(X_T); T = \infty] \\ &= E^x[e^{-\lambda(T+t)} E^{X(T)}[f(X_t)]; T < \infty] \\ &= E^x[e^{-\lambda(T+t)} \theta_T(f(X_t)); T < \infty] \;\; \text{(Strong Markov Thm)} \\ &= E^x[e^{-\lambda(T+t)} f(X_{T+t}); T < \infty] \\ &= P^\lambda_{T+t} f(x) \end{aligned}$$

Thus, the result follows. □

**Remark 2.1.** *For an $\{\mathscr{F}_t\}]$ stopping time $T$, in general, $P^\lambda_t P^\lambda_T \neq P^\lambda_{T+t}$*

*Proof.* It is enough to prove that $P_T^\lambda P_t^\lambda f(x) \neq P_t^\lambda P_T^\lambda f(x)$ for some Feller process $\{X_t\}$. Consider a one-dimensional canonical Brownian motion $X_t$, and let $T$ be the hitting time $H_1$. Then $\mathbf{P}^x[T < 1] = 1, \mathbf{P}^x[X_t = 1] = 1$and

$$
\begin{aligned}
P_T^\lambda P_t^\lambda f(x) &= E^x[e^{-\lambda T} P_t^\lambda f(1); T < \infty] \\
&\to E^x[P_t f(1)] \\
&= P_t f(1)
\end{aligned}
$$

as $\lambda \to 0+$, by the dominated convergence theorem. Also, we have

$$
\begin{aligned}
P_t^\lambda P_T^\lambda f(x) &= P_t^\lambda E^x[e^{-\lambda T} f(1); T < \infty] \\
&= f(1) P_t^\lambda E^x[e^{-\lambda T}; T < \infty] \\
&\to f(1) P_t(1) \\
&= f(1)
\end{aligned}
$$

as $\lambda \to 0+$, by the dominated convergence theorem again. Since $P_t f(1) = \int P_t(1, dy) f(y) \neq f(1)$ for some function $f$, we attain the desired result. □

The Dynkin, which is given below helps us to compute the difference of the resolvents of $X_t$ and $Y_t := X_t 1_{T>t}$.

**Theorem 2.7** (Dynkin's formula)**.** *If $T$ is an $\{\mathscr{F}_t\}$ stopping time, for $g \in C_0, \lambda > 0, x \in E$,*

$$
R_\lambda g(x) = E^x \int_0^T e^{-\lambda t} g(X_t) dt + P_T^\lambda R_\lambda g(x)
$$

*Proof.* The following computation demonstrates the theorem.

$$
\begin{aligned}
P_T^\lambda R_\lambda g(x) &= E^x[e^{-\lambda T} R_\lambda g(X_T)] \\
&= E^x[e^{-\lambda T} \int_0^\infty e^{\lambda t} P_t g(X_T) dt] \\
&= E^x[e^{-\lambda(T+t)} \int_0^\infty E^{X_T}[g(X_t)] dt] \\
&= E^x[e^{-\lambda(T+t)} \int_0^\infty g(X_{T+t}) dt] \\
&= E^x[e^{-\lambda s} \int_0^\infty g(X_s) dt] \\
&= R_\lambda g(x) - E^x \int_0^T e^{-\lambda t} g(X_t) dt.
\end{aligned}
$$

□

There are a couple of alternative forms of Dynkin's formula, which will be given in the following. While the equation above deals with the resolvent, the properties of martingales are the key ingredients for the new forms.

**Theorem 2.8** (alternative forms of Dynkin's formula)**.** *The Dynkin's formula in Theorem 2.7 implies the following two equations.*

$$
E^x e^{-\lambda T} f(X_T) - f(x) = E^x \int_0^T e^{-\lambda x} (\mathscr{G} - \lambda) f(X_s) ds,
$$

*If $E^x(T) < \infty$ for some $x$, then*

$$E^x f(X_T) - f(x) = E^x \int_0^T \mathscr{G} f(X_s) ds.$$

*Proof.* For fixed $g \in C_0$, we let $\eta = \int_0^\infty e^{-\lambda s} g(X_s) ds$, which is a bounded and measurable function. Note that for any $z \in E$

$$E^z[\eta] = \int_0^\infty \int_E e^{-\lambda s} g(y) P_s(z, dy) ds = R_\lambda g(z)$$

by Fubini theorem. Then

$$\begin{aligned}
\eta &= \int_0^t e^{-\lambda s} g(X_s) ds + \int_t^\infty e^{-\lambda s} g(X_s) ds \\
&= \int_0^t e^{-\lambda s} g(X_s) ds + e^{-\lambda t} \int_0^\infty e^{-\lambda u} g(X_u) du \qquad (u = s - t) \\
&= \int_0^t e^{-\lambda s} g(X_s) ds + e^{-\lambda t} \eta \circ \theta_t
\end{aligned}$$

By applying the weak Markov property, we have that $P^\mu$ almost surely,

$$\begin{aligned}
E^x[\eta | \mathscr{F}_t] &= \int_0^t e^{-\lambda s} g(X_s) ds + e^{-\lambda t} E^x[\eta \circ \theta_t | \mathscr{F}_t] \\
&= \int_0^t e^{-\lambda s} g(X_s) ds + e^{-\lambda t} E^x[\eta \circ \theta_t | \mathscr{F}_t] \\
&= \int_0^t e^{-\lambda s} g(X_s) ds + e^{-\lambda t} E^{X_t}[\eta] \\
&= \int_0^t e^{-\lambda s} g(X_s) ds + e^{-\lambda t} R_\lambda g(X_t).
\end{aligned}$$

Note that the first formula is a uniformly integrable martingale, and that the last formula is an R-process.

Now, we consider $f \in \mathscr{D}(\mathscr{G})$. Then by putting $g = (\lambda - \mathscr{G}) f \in C_0$, we obtain a following uniformly integrable R-process martingale $C_t^{\lambda, f}$ given by

$$C_t^{\lambda, f} = e^{-\lambda t} f(X_t) - f(X_0) + \int_0^t e^{-\lambda s} (\lambda - \mathscr{G}) f(X_s) ds.$$

If $T$ is an $\{\mathscr{F}_t\}$ stopping time and $x \in E$, then by applying the Optional-Stopping theorem,

$$\begin{aligned}
0 &= E^x[C_0^{\lambda, f}] \\
&= E^x[C_T^{\lambda, f}] \\
&= E^x[e^{-\lambda T} f(X_T)] - E^x[f(X_0)] + E^x[\int_0^T e^{-\lambda s} (\lambda - \mathscr{G}) f(X_s) ds]
\end{aligned}$$

Since $E^x[f(X_0)] = f(x)$, we attain that

$$E^x e^{-\lambda T} f(X_T) - f(x) = E^x \int_0^T e^{-\lambda x} (\mathscr{G} - \lambda) f(X_s) ds.$$

Also, for some $x$ with $E^x(T) < \infty$, by letting $\lambda \to 0+$, we appeal to the dominated convergence to get

$$E^x f(X_T) - f(x) = E^x \int_0^T \mathscr{G} f(X_s) ds.$$

$\square$

**Remark 2.2.** *By plugging in deterministic stopping time $T = t$ in Theorem 2.8, we get*

$$P_t f - f - \int_0^t P_s \mathscr{G} f ds = 0$$

*with the help of the Fubini theorem.*

**Remark 2.3.** *For $f \in \mathscr{D}(\mathscr{G})$,*

$$C_t^f := f(X_t) - f(X_0) - \int_0^t \mathscr{G} f(X_s) ds$$

*defines a martingale relative to $(\{\mathscr{F}\}, \mathbf{P}^x)$. which is consistent with the proof of Theorem 2.8 with $\lambda = 0$. Note that this cannot be directly approached from the proof of Theorem 2.8 since the existence of $\eta$ is not ensured when $\lambda = 0$.*

*Proof.* For $x \in E$ and $t \geq 0$,

$$\begin{aligned}
E^x[C_t^f | \mathscr{F}_u] &= E^x[f(X_t)|\mathscr{F}_u] - f(X_0) - E^x[\int_0^t \mathscr{G} f(X_s) ds | \mathscr{F}_u] \\
&= f(X_u) - f(X_0) - \int_0^u \mathscr{G} f(X_u) du - E^{X_u}[\int_0^{t-u} \mathscr{G} f(X_s) ds] \\
&= f(X_u) - f(X_0) - \int_0^u \mathscr{G} f(X_u) du \\
&= C_u^f
\end{aligned}$$

since

$$E^{X_u}[\int_0^{t-u} \mathscr{G} f(X_s) ds] = E^{X_u}[f(X_{t-u})] - f(X_u) = 0$$

by Theorem 2.8 with $x \leftarrow X_u$ and $T = t - u$. $\square$

## 2.5 Extensions of the generator and Characteristic functions

We begin with two lemmata concerning extensions of the generator $\mathscr{G}$ of a given SCCSG on a Banach space $B_0$. For a Feller process, it is enough to consider $B_0 = C_0(E)$.

**Lemma 2.3** (Dynkin, Reuter)**.** *Let $\mathscr{C} : \mathscr{D}(\mathscr{C}) \to B_0$ be a linear operator extending $\mathscr{G}$, that is, $\mathscr{D}(\mathscr{G}) \subset \mathscr{D}(\mathscr{C}) \subset B_0$ and $\mathscr{C} f = \mathscr{G} f$ for all $f \in \mathscr{D}(\mathscr{G})$. Suppose that for $f \in \mathscr{D}(\mathscr{C})$, $\mathscr{C} f = f$ implies that $f = 0$. Then $\mathscr{C} = \mathscr{G}$; or equivalently, $\mathscr{D}(\mathscr{C}) = \mathscr{D}(\mathscr{G})$.*

*Proof.* For $f \in \mathscr{D}(\mathscr{C})$, let $R_1(f - \mathscr{C} f) = h$. Since the range of $R_1$ is $\mathscr{D}(\mathscr{G})$, then $h \in \mathscr{D}(\mathscr{G}) \in \mathscr{D}(\mathscr{C})$. Note that

$$h - \mathscr{C} h = h - \mathscr{G} h = (1 - \mathscr{G}) h = f - \mathscr{C} f,$$

thus $f - h = \mathscr{C} f - \mathscr{G} h = \mathscr{C}(f - h)$. Here, $f - h$ satisfies the condition in the given statement. Hence,

$$f - h = 0 \quad \Rightarrow \quad f = h \in \mathscr{D}(\mathscr{G}).$$

Thus, $\mathscr{D}(\mathscr{C}) \subset \mathscr{D}(\mathscr{G})$ gives the result. $\square$

The upcoming lemma, Dynkin's maximum principle, expands our given condition in a far more practical manner to obtain that the extension of the generator is exactly the same as itself.

**Lemma 2.4** (Dynkin's maximum principle)**.** *Suppose that $\mathscr{C} : \mathscr{D}(\mathscr{C}) \to B_0$ is a linear operator extending $\mathscr{G}$. Suppose that for $f \in \mathscr{D}(\mathscr{C})$, the fact that $f$ attains its maximum at $x$ with $f(x) \geq 0$ implies that $\mathscr{C}f(x) \leq 0$. Then $\mathscr{C} = \mathscr{G}$.*

*Proof.* By Lemma 2.3, it is enough to consider $f \in \mathscr{D}(\mathscr{C})$ with $\mathscr{C}f = f$ whether it follows that $f = 0$.

Suppose that $f$ attains its maximum at $x$. If $f(x) > 0$, then by the given condition, $f(x) = \mathscr{C}f(x) \leq 0$, which is a contradiction. Thus $f(x) \leq 0$ and since $f(x)$ was the maximum, this implies that $f \leq 0$. Since $\mathscr{C}(-f) = -f$, by applying the same argument to $-f$, we obtain that $-f \leq 0$. Combining these results, we may conclude that $f = 0$. $\square$

Eventually, we arrived at Dynkin's characteristic operator, which can be considered as an extension of the generator that fulfills the practical requirements when $E$ is equipped with topology. Also, we return to the specific function space $C_0(E)$ on which our focus was originally placed. In order to provide more rigorous definition of the characteristic operator, we first define the notion of an absorbing point.

**Definition 2.9** (absorbing)**.** *A point $x \in E$ is absorbing if either of the following equivalent conditions hold;*
*(i)* $\mathbf{P}^x[\forall t \geq 0, X_t = x] = 1$
*(ii)* $\forall t \geq 0, P_t(x, \{x\}) = 1$

**Definition 2.10** (Dynkin's characteristic operator)**.** *Let $d$ be a metric defining the topology of $E$. If $x$ is not absorbing, then for $\eta > 0$,*

$$V_{\eta,x} := \inf\{t : d(x, X_t) \geq \eta\}.$$

*Dynkin's characteristic operator $\mathscr{C}$ of a Feller process $X_t$ is given by the following. If $x \in E$ is absorbing, then for any $f \in C_0(E)$, $\mathscr{C}f(x) := 0$. Otherwise,*

$$\mathscr{C}f(x) := \lim_{\eta \to 0+} \frac{E^x[f(X_{V_{\eta,x}})] - f(x)}{E^x V_{\eta,x}}$$

*if the limit exits. As usual, the domain $\mathscr{D}(\mathscr{C})$ is defined by*

$$\mathscr{D}(\mathscr{C}) := \{f \in C_0 : \forall x \in E, \mathscr{C}f(x) \text{ exists, and } \mathscr{C}f \in C_0\}.$$

We check the well-definedness of the characteristic operator by showing that the denominator in the definition is finite and non-zero. Fortunately, the non-zeroness is obvious since $V_{\eta,x} > 0$ by the right-continuity of $X_t(\omega)$ with respect to $t$. Moreover, the following lemma resolves the matter.

**Lemma 2.5** (Dynkin)**.** *If $x$ is not absorbing, then for all sufficiently small $\eta > 0$, $E^x V_{\eta,x} < \infty$.*

*Proof.* Suppose that $x$ is not absorbing and let $B_\epsilon(x) = \{y : d(x, y) < \epsilon\}$ for $\epsilon > 0$ be the $\epsilon$-ball in $E$. If $P_t^{+\partial}(x, E_\partial \setminus \overline{B_{1/n}(x)}) = 0$ for all $n \in \mathbb{N}$, then the intersection for all $n$ implies that $P_t^{+\partial}(x, \{x\}^c) = 0$, which means that $x$ is absorbing. Then we may find some $\epsilon > 0$, $t > 0$, and $\alpha > 0$ that

$$P_t^{+\partial}(x, E_\partial \setminus \overline{B_\epsilon(x)}).$$

Let $G = E_\partial \setminus \overline{B_\epsilon(x)}$, which is an open set. We may find a sequence $\{h_n\}$ of non-negative continuous functions on $E_\partial$ that increases to $1_G$. Then as $n \to \infty$,

$$P_t^{+\partial} h_n(x) = \int_{E^\partial} P_t(x, dy) h_n(y)$$

also increases and converges to $P_t^{+\partial}(y, G) = \int_{E^\partial} P_t(x, dy) 1_G(y)$ by the dominated convergence theorem. Since $P_t^{+\partial} h_n$ is a continuous function with compact support, we attain that

$$\{y : P_t^{+\partial}(y, G) > \alpha\} = \bigcup_{n \geq 1} \{y : P_t^{+\partial} h_n(y) > \alpha\}$$

and that this set is open. Hence, we may find some positive $\eta$ that satisfy

$$B_\eta(x) \subset \{y : P_t^{+\partial}(y, E_\partial \setminus \overline{B_\epsilon(x)}) > \alpha\}.$$

In other words, for all $y \in B_\eta(x)$,

$$P_t^{+\partial}(y, B_\eta(x)) \leq 1 - \alpha < P_t^{+\partial}(y, \overline{B_\epsilon(x)}) < 1 - \alpha.$$

Then by the weak Markov property, we have that

$$\begin{aligned}
\mathbf{P}^x[X_{kt} \in B_\eta(x), \forall k \leq n] \quad &= \quad E^x[\mathbf{P}^{X_{(n-1)t}}[X_t \in B_\eta(x)] 1_{X_{kt} \in B_\eta(x), \forall k \leq n-1}] \\
&= \quad E^x[P_t^{+\partial}(X_{(n-1)t}, B_\eta(x)) 1_{X_{kt} \in B_\eta(x), \forall k \leq n-1}] \\
&\leq \quad E^x[(1-\alpha) 1_{X_{kt} \in B_\eta(x), \forall k \leq n-1}] \\
\leq \cdots &\leq \quad (1-\alpha)^n.
\end{aligned}$$

Since $\{V_{\eta,x} \geq nt\} \subset \{X_{kt} \in B_\eta(x), \forall k \leq n\}$, we finally obtain that

$$E^x V_{\eta,x} \leq \sum_{m \geq 1} \mathbf{P}^x[V_{\eta,x} \geq mt] \cdot t \leq \sum_{m \geq 1} (1-\alpha)^m \cdot t = \frac{t}{\alpha},$$

thus $E^x V_{\eta,x} < \infty$. □

The following theorem is the ultimate conclusion of this section. It shows that we may use $\mathscr{G}$ instead of $\mathscr{C}$ from now on. Despite its brevity, the theorem is of great importance since it clearly exhibits the role of the generator of FD processes in the description of the sample path within a local region.

**Theorem 2.9** (Dynkin's Characteristic-Operator Theorem for FD processes)**.** *We have* $\mathscr{G} = \mathscr{C}$.

*Proof.* The first step is to prove that $\mathscr{C}$ satisfies the condition of Lemma 2.4, Dynkin's maximum principle. Suppose that $f \in \mathscr{D}(\mathscr{C})$ and that $f$ attains its maximum at $x \in E$. If $x$ is absorbing, $0 =: \mathscr{C} f(x) \leq 0$. Otherwise, since $V_{\eta,x} < \infty$ and then $X_{V_{\eta,x}}$ cannot be $x$ almost surely for any $\eta > 0$, we have $f(X_{V_{\eta,x}}) \leq f(x)$ almost surely. Hence,

$$\mathscr{C} f(x) = \lim_{\eta \to 0+} \frac{E^x[f(X_{V_{\eta,x}})] - f(x)}{E^x V_{\eta,x}} \leq 0$$

.

Then, the only remaining step is to show that $\mathscr{C}$ extends $\mathscr{G}$ from $\mathscr{D}(\mathscr{G})$ to $\mathscr{D}(\mathscr{C})$. Let $f \in \mathscr{D}(\mathscr{G})$ Note that $V_{\eta,x}$ is the hitting time of the outside of $\eta$-ball, thus a stopping time, with $E^x[V_{\eta,x}] < \infty$. Then by the alternative form of Dynkin's formula(Theorem 2.8),

$$E^x[f(X_{V_{\eta,x}})] - f(x) = E^x[\int_0^{V_{\eta,x}} \mathscr{G}f(X_s)ds]$$

helps us compute $\mathscr{C}f(x)$ easily by

$$\mathscr{C}f(x) = \lim_{\eta \to 0+} \frac{E^x[\int_0^{V_{\eta,x}} \mathscr{G}f(X_s)ds]}{E^x V_{\eta,x}}.$$

Since $\mathscr{G}f \in C_0$, the continuity of $\mathscr{G}f$ implies that for any $\epsilon > 0$, there exist $\delta > 0$ such that $\forall y \in B_\delta(x)$, $|\mathscr{G}f(y) - \mathscr{G}f(x)| < \epsilon$. Hence,

$$\begin{aligned} \eta < \delta \quad &\Rightarrow \quad \forall s < V_{\eta,x},\ d(X_s, x) \leq \eta < \delta \\ &\Rightarrow \quad \forall s < V_{\eta,x},\ |\mathscr{G}f(X_s) - \mathscr{G}f(x)| < \epsilon. \end{aligned}$$

Therefore,

$$\begin{aligned} |\mathscr{C}f(x) - \mathscr{G}f(x)| &= \lim_{\eta \to 0+} \left| \frac{E^x[\int_0^{V_{\eta,x}} (\mathscr{G}f(X_s) - \mathscr{G}f(x))ds]}{E^x V_{\eta,x}} \right| \\ &\leq \liminf_{\eta \to 0+} \frac{E^x[\int_0^{V_{\eta,x}} |\mathscr{G}f(X_s) - \mathscr{G}f(x)|ds]}{E^x V_{\eta,x}} \\ &\leq \liminf_{\eta \to 0+} \frac{E^x[\int_0^{V_{\eta,x}} \epsilon ds]}{E^x V_{\eta,x}} \\ &= \epsilon. \end{aligned}$$

Since $\epsilon$ is arbitrary, the value converges to 0. Thus, $\mathscr{C}$ is an extension of $\mathscr{G}$. $\square$

## 2.6 Feller-Dynkin diffusions

We look into Feller-Dynkin diffusions, a type of Feller processes satisfying more specific conditions including $E = \mathbb{R}^n$.

**Definition 2.11** (Feller-Dynkin diffuision)**.** *A Feller process $X$ is called a Feller-Dynkin diffusion (FD diffusion) if the following properties hold.*
*(i) the paths $t \mapsto X_t(\omega)$ are continuous on $[0, \zeta)$ where $\zeta$ is the lifetime of $X$.*
*(ii) the domain $D(G)$ contains $C_K^\infty := C_K^\infty(\mathbb{R}^n)$, the infinitely differentiable functions of compact support.*

**Theorem 2.10.** *If we define the restriction $\mathscr{L} := \mathscr{G}|_{C_K^\infty}$, then $\mathscr{L}$ satisfies the following.*
*(i) $\mathscr{L}$ is a linear map from $C_K^\infty$ to $C_0$.*
*(ii) $\mathscr{L}$ is local: if $f, g \in C_K^\infty$ and $f = g$ in some neighborhood of a point $x$, then $\mathscr{L}f(x) = \mathscr{L}g(x)$.*
*(iii) $\mathscr{L}$ satisfies the maximum principle: if $f \in C_K^\infty$ attains its maximum at $x$, then $\mathscr{L}f(x) \leq 0$.*

*Proof.* (i) is trivial since $\mathscr{L}$ is a restriction of $\mathscr{G} : \mathscr{D}(\mathscr{G}) \to C_0$.

For (ii), suppose that $f, g \in C_K^\infty$ and that $\delta > 0$ satisfies $f = g$ in $B_\delta(x)$. Then for any $0 < \eta < \delta$, $X_{V_{\eta,x}} \in B_\delta(x)$, so we have for any

$$(f - g)(X_{V_{\eta,x}}) = f(X_{V_{\eta,x}}) - g(X_{V_{\eta,x}}) = 0.$$

Then by definition of the charactersitic operator,

$$\mathscr{C}(f-g)(x) = \lim_{\eta \to 0+} \frac{E^x[(f-g)(X_{V_{\eta,x}})] - (f-g)(x)}{E^x V_{\eta,x}} = 0.$$

Finally, we appeal to Theorem 2.9 to obtain that

$$\mathscr{L}f(x) = \mathscr{C}f(x) = \mathscr{C}g(x) = \mathscr{L}g(x)$$

Lastly, for (iii), we assume that $f \in C_K^\infty$ attains its its maximum at $x$ with $f(x) \geq 0$. Then since $P_t(x, E) \leq 1$ for any $t > 0$, we have that

$$t^{-1}(P_t f(x) - f(x)) \leq \int_E t^{-1} P_t(x, dy)(f(y) - f(x)) \leq 0,$$

thus obtaining that $\mathscr{G}f(x) \leq 0$ by taking the limit as $t \to 0+$. Hence, $\mathscr{L}f(x) = \mathscr{G}f(x) \leq 0$. □

The following theorem shows how $\mathscr{L}$ behaves on a specific point.

**Theorem 2.11.** *Let $x$ be a fixed point in $E = \mathbb{R}^n$. Then the restiction $\mathscr{L}$ of $\mathscr{G}$ to $C_K^\infty$ is a second-order elliptic operator of the form*

$$\mathscr{L}f(x) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(x)\partial_i\partial_j f(x) + \sum_{i=1}^{n} b_i(x)\partial_i f(x) - c(x)f(x)$$

*for all $f \in C_K^\infty$ where $\partial_i$ denotes $\frac{\partial}{\partial x_i}$ and these properties hold.*
*(i) For any $i, j$ the functions $a_{ij}(\cdot), b_i(\cdot), c(\cdot)$ are continuous.*
*(ii) The matrix $(a_{ij}(x))_{1 \leq i,j \leq n}$ is non-negative definite symmetric.*
*(iii) $c(x) \geq 0$.*

*Proof.* We may develop the local-maximum principle and show that $\mathscr{L}$ follows the principle : "If $f \in \mathscr{D}(\mathscr{L})(= C_K^\infty)$ has a local maximum at $z$ and $f(z) \geq 0$, then $\mathscr{L}f(z) \leq 0$". This holds since if $f \in C_K^\infty$ attains local maximum at $z$ in $B_\delta(z)$, then by taking $0 < \eta < \delta$, $f(X_{V_{\eta,z}}) \leq f(z)$ implies that

$$\mathscr{L}f(z) = \mathscr{C}f(z) = \lim_{\eta \to 0+} \frac{E^x[f(X_{V_{\eta,z}})] - f(x)}{E^x V_{\eta,z}} \leq 0.$$

Now, we fix $x = (x_1, x_2, \cdots, x_n) \in E(= \mathbb{R}^n)$. We can find $\phi \in C_K^\infty$ with $\phi = 1$ in $B_\delta(x)$ for some $\delta > 0$. We let $c = -\mathscr{G}\phi$, then this function is continuous and

$$c(x) = \mathscr{C}\phi(x) := \lim_{\eta \to 0+} \frac{E^x[\phi(X_{V_{\eta,x}})] - \phi(x)}{E^x V_{\eta,x}} = 0.$$

Let $\psi_i \in C_K^\infty$ for $i = 1, 2, \cdots, , n$ with $\psi_i(y) = y_i - x_i$ in $B_\delta(x)$. Define $b_i := \mathscr{G}\psi_i$ and $a_{ij} := \mathscr{G}(\psi_i\psi_j)$. Then $b_i$ and $a_{ij}$ are continuous. For real numbers $\lambda_1, \lambda_2, \cdots, \lambda_n$, the function $h(y) := -(\sum_{i=1}^n \lambda_i\psi_i(y))^2$ has a local maximum at $x$ since $h(x) = 0$. Then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i j(x)\lambda_i\lambda_j = -\mathscr{G}h(x) = -\mathscr{L}h(x) \geq 0$$

Thus, $\{a_{ij}(x) : 1 \leq i, j \leq n\}$ is non-negative definite. Now we apply Taylor's formula ($y$ near $x$) to obtain that

$$f(y) = \alpha(y) + o(|y - x|^2),$$

where

$$\alpha(y) := f(x)\phi(y) + \sum_{i=1}^{n} \psi_i(y) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \partial_i\partial_j f(x)\psi_i(y)\psi_j(y),$$

Note that

$$\mathscr{L}\alpha(y) = -c(y)f(x) + \sum_{i=1}^{n} b_i(y)\partial_i f(x) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \partial_i\partial_j f(x) a_{ij}(y)$$

For $\epsilon > 0$, a function in $C_K^\infty$ defined near $x$ by

$$y \mapsto f(y) - \alpha(y) - \epsilon|y-x|^2$$

is non-negative and has a maximum 0 at $x$ in the local region. Thus,

$$\mathscr{L}f(x) - \mathscr{L}\alpha(x) - \epsilon\sum_{i=1}^{n} a_{ii} \leq 0$$

Since $\epsilon > 0$ was arbitrary, we obtain that $\mathscr{L}f(x) \leq \mathscr{L}\alpha(x) \leq 0$. By applying the same manner to $-f \in C_K^\infty$, we then attain that $\mathscr{L}(-f)(x) \leq \mathscr{L} - (\alpha)(x) \leq 0$, hence

$$\mathscr{L}f(x) = \mathscr{L}\alpha(x) = \frac{1}{2}\sum_i\sum_j a_{ij}(x)\partial_i\partial_j f(x) + \sum_i b_i(x)\partial_i f(x) - c(x)f(x)$$

□

Since $a_{i,j}, b_i, c$ were dependent on a fixed value $x$, Theorem 2.11 does not guarantee that the generator of any FD diffusion is exactly the form of elliptic operators. We now present relatively important consequence that the continuous Feller process is attained when the generator is given as an elliptic operator.

**Theorem 2.12.** *Let $\mathscr{L} : C_K^\infty \to C_0$ be an elliptic operator in the previous theorem. Suppose that $\{P_t : t \geq 0\}$ is an FD semigroup and $X$ is the associated Feller process with generator $\mathscr{G}$ extending $\mathscr{L}$. Then the paths of $X$ are continuous on* $[0, \zeta)$

*Proof.* Before we begin, we may only consider the case when $\{P_t\}$ is honest, that is,

$$\forall x \in E, P_t(x, E) = 1.$$

For a moment, we move our focus away from the generator, and look for a sufficient condition for the continuous path. it is enough to prove that for any compact set $K \in \mathbb{R}^n$, $\epsilon, u > 0$, and $x \in K$,

$$S := \mathbf{P}^x[\bigcup_{k=0}^{n-1}\{|X_{ku/n} - X_{(k+1)u/n}| > 3\epsilon; \forall s \leq u, X_s \in K\}] \to 0 \qquad \text{as } n \to \infty \tag{2}$$

The reason is as follows. For any given $\omega$, the sample path $\{X_t(\omega) : 0 \leq t \leq u\}$ is an R-function. Assume that $\{X_t(\omega) : 0 \leq t \leq u\}$ is not bounded in $\mathbb{R}^n$, then there exists $(t_n)_{n\geq 1} \in [0, u]$ with $|X_{t_n}| \geq n$. By sequential compactness, there exists a subsequence $t_{n_k}$ that converges to a point $t_0 \in [0, u]$. If there is a further subsequence $t_{n_{k_l}} \to t_0+$, then by right-continuity, $X_{t_0} = \lim_{l\to\infty} X_{t_{n_{k_l}}}$ cannot exist. Otherwise, there exists a further subsequence $t_{n_{k_l}} \to t_0-$, so $X_{t_0} = \lim_{l\to\infty} X_{t_{n_{k_l}}}$ does not exist finitely. These contradictions imply that $\{X_t(\omega) : 0 \leq t \leq u\}$ is bounded for all $\omega$.

Now, we let $K_m := [-m,m]^n \in \mathbb{R}^n$ so that $K_m \subset K_{m+1} \subset\in \mathbb{R}^n$. Also, for positive integers $m, n, l$, we define

$$A_{m,n}(l) := \bigcup_{k=0}^{n-1} \{|X_{ku/n} - X_{(k+1)u/n}| > 3/l; \forall s \leq u, X_s \in K_m\}.$$

Then for any sample $\omega$, there exists $M \geq 1$ such that for all $m \geq M$, by the discussion above, $\forall s \leq u, X_s(\omega) \in K_m$. Also, if the sample path $\{X_t(\omega) : 0 \leq t \leq u\}$ is discontinuous, then the discontinuity occurs on some point $s \in [0,u]$, in other words, $X_s(\omega) \neq X_{s-}(\omega)$ due to the existence of left and right limit of R-processes. Let $|X_s(\omega) - X_{s-}(\omega)| = 2\epsilon_0$ for some $\epsilon_0 > 0$. Then for large $L \geq 1$ with $\epsilon_0 > \frac{3}{L}$, take $l \geq L$. Also, we may take large $N \geq 1$ such that for all $n \geq N$, there exists $k$ with

$$|X_{ku/n}(\omega) - X_{(k+1)u/n}(\omega)| > \epsilon_0 > 3/l.$$

To sum up, for any $\omega$ with discontinuous path, there exists $M, L, N \geq 1$ such that for all $m \geq M$, $l \geq L$, and $n \geq N$,

$$\omega \in \bigcup_{k=0}^{n-1} \{|X_{ku/n} - X_{(k+1)u/n}| > 3/l; \forall s \leq u, X_s \in K_m\}$$

This can be summarized as

$$\omega \in \bigcup_{M\geq 1} \bigcup_{L \geq 1} \bigcup_{N \geq 1} \bigcap_{m \geq M} \bigcap_{l \geq L} \bigcap_{n \geq N} A_{m,n}(l).$$

Note that $\mathbf{P}^x[A_{m,n}(l)] \to 0$ as $n \to \infty$. For an arbitrary $\delta > 0$, for fixed $m, l$, we can find $N(m,l)$ (increasing as $m$ or $l$ increases) such that

$$\mathbf{P}^x[A_{m,N(m,l)}(l)] < \frac{1}{2^{m+l}}.$$

If $N \geq N(m,l)$, then $\bigcap_{n\geq N} A_{m,n}(l) \subset A_{m,N}$, and if $N < N(m,l)$, then $\bigcap_{n \geq N} A_{m,n}(l) \subset A_{m,N(m,l)}$. In any cases,

$$\mathbf{P}^x[\bigcap_{n \geq N} A_{m,n}(l)] < \frac{1}{2^{m+l}}.$$

Thus, we attain that for any $M, L, N \geq 1$,

$$\mathbf{P}^x[\bigcap_{m\geq M} \bigcap_{l \geq L} \bigcap_{n \geq N} A_{m,n}(l)] \leq \mathbf{P}^x[\bigcap_{m\geq M} \bigcap_{l \geq L} \bigcap_{n \geq N} A_{m,n}(l)] = 0.$$

Hence, the probability of the countable union becomes 0 as well;

$$\mathbf{P}^x[\bigcup_{M\geq 1} \bigcup_{L \geq 1} \bigcup_{N \geq 1} \bigcap_{m \geq M} \bigcap_{l \geq L} \bigcap_{n \geq N} A_{m,n}(l)] = 0.$$

This implies that $\{X_t : 0 \leq t \leq u\}$ is continuous almost surely. This can be extended to any $u > 0$, which enables us to conclude that the paths of $X$ are continuous.

Now, let us prove equation (2). Note that following inequalities hold;

$$S \quad = \quad \mathbf{P}^x[\bigcup_{k=0}^{n-1} \{|X_{ku/n} - X_{(k+1)u/n}| > 3\epsilon; \forall s \leq u, X_s \in K\}]$$

$$
\begin{aligned}
&\le \sum_{k=0}^{n-1} \mathbf{P}^x[|X_{ku/n} - X_{(k+1)u/n}| > 3\epsilon; \forall s \le u, X_s \in K] \\
&\le \sum_{k=0}^{n-1} \mathbf{P}^x[|X_{ku/n} - X_{(k+1)u/n}| > 3\epsilon; X_{ku/n} \in K] \\
&= \sum_{k=0}^{n-1} \mathbf{P}^{X_{ku/n}}[|X_{u/n}| > 3\epsilon; X_0 \in K] \\
&\le n \sup_{y \in K} P_{u/n}(y, \mathbb{R}^n \ B_{3\epsilon}(y)).
\end{aligned}
$$

Thus, it suffices to show that $n \sup_{y \in K} P_{u/n}(y, \mathbb{R}^n \ B_{3\epsilon}(y)) \to 0$ as $n \to \infty$. Recall that

$$
\mathscr{L} f(x) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x) \partial_i \partial_j f(x) + \sum_{i=1}^{n} b_i(x) \partial_i f(x) - c(x) f(x)
$$

For any $x \in \mathbb{R}^n$, by putting $f \in C_K^\infty$ with $f = 1$ in some open neighborhood $x$, we have that $\partial_i \partial_j f(x) = \partial_i f(x) = 0$ and $\mathscr{L} f(x) = \mathscr{C} f(x) = 0$. Thus, $c(x) = 0$. For any given compact $K$ and open $G$ with $K \subset G$, we may find a function $f \in C_K^\infty$ with $f = 1$ on $K$, $f = 0$ on $\mathbb{R}^n \setminus G$, and $0 \le f \le 1$ with the help of Urysohn lemma. Since $c = 0$, we have $\mathscr{L} f = 0$ on $K$. Then by Remark 2.2, for any $x \in K$,

$$
P_t(x, \mathbb{R}^n \setminus G) \le 1 - P_t f(x) = f(x) - P_t f(x) = -\int_0^t P_s \mathscr{L} f(x) ds
$$

Since $\|P_s \mathscr{L} f - \mathscr{L} f\| \to 0$ as $s \to 0+$, we have that

$$
\begin{aligned}
|t^{-1} \int_0^t P_s \mathscr{L} f(x) ds| &= |t^{-1} \int_0^t P_s \mathscr{L} f(x) ds - \mathscr{L} f(x)| \\
&\le \int_0^t t^{-1} |P_s \mathscr{L} f(x) - \mathscr{L} f(x)| ds \\
&\to 0
\end{aligned}
$$

as $s \to 0+$. Hence for any compact $K$ and open set $G$ containing K,

$$
t^{-1} P_t(x, \mathbb{R}^n \setminus G) \to 0 \text{ as } t \to 0+. \tag{3}
$$

Now let $K$ be a general compact set. Then there exists $x_1, x_2, \cdots, x_r \in K$ such that $\bigcup_{j=1}^r B_\epsilon(x_j) \supset K$. Then consider (3) with $K = \overline{B_\epsilon(x_k)}$ and $G = B_{2\epsilon}(x_k))$. Then for any $\eta > 0$, there exists $\delta_k > 0$ such that $\forall t < \delta_k$,

$$
\sup_{y \in B_\epsilon(x_k)} t^{-1} P_t(y, \mathbb{R}^n \setminus B_{3\epsilon})) \le \sup_{y \in B_\epsilon(x_k)} t^{-1} P_t(y, \mathbb{R}^n \setminus \overline{B_{2\epsilon})} < \eta
$$

By taking $\delta := \min\{\delta_1, \delta_2, \cdots \delta_r\}$, we get that $\forall \eta > 0$,

$$
\sup_{y \in K} t^{-1} P_t(y, \mathbb{R}^n \setminus B_{3\epsilon})) < \eta
$$

if $n$ is large so that $\frac{t}{n} < \delta$. Thus, $X$ is continuous. □

Before we proceed to the final section, note that the Ito diffusion is defined by the solution $\{X_t\}$ of stochastic differential equations

$$
\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t
$$

where $B_t$ is the $n$-dimensional Brownian motion, and $b$ and $\sigma$ are deterministic. It is well-known that an Ito diffusion has an elliptic operator generator.

## 2.7 Applications to Brownian motions and Lévy processes

Until now, we analyzed several theoretical properties of Feller processes. However, it is as important to understand the practical features of Feller processes as setting a number of definitions and proving many theorems.

Brownian motion is one of the most typical examples of Feller processes. It is clearly observed that the probability semigroup $\{P_t\}$ of a Brownian motion is a contraction semigroup with pointwise continuity, thus strong continuous. We may consider the one-dimensional canonical Brownian motion, CBM($\mathbb{R}$), and explicitly compute its generator and the resolvent.

**Theorem 2.13** (generator of CBM($\mathbb{R}$)). *Suppose that $\{P_t\}$ is the FD semigroup of CBM($\mathbb{R}$). Then the generator of $P_t$ is exactly given by $\mathscr{G} = \frac{1}{2}\Delta$, where $\Delta$ is given by the Laplace operator. Thus, $\mathscr{D}(\mathscr{G}) = \mathscr{D}(\frac{1}{2}\Delta)$, the space of twice-differentiable functions vanishing at infinity with continuous second derivative which vanishes at infinity as well. Also, the resolvent $R_\lambda$ for $\lambda > 0$ is given by*

$$R_\lambda f(x) = \int_{\mathbb{R}} r_\lambda(x,y) f(y) dy$$

*where $r_\lambda(x,y) = \gamma^{-1} e^{\gamma|y-x|}, \gamma = \sqrt{2\lambda}$.*

*Proof.* Suppose that $f \in C_0$. By Fubini theorem,

$$\begin{aligned} R_\lambda f(x) &= \int_0^\infty e^{-\lambda t} P_t f(y) dt \\ &= \int_0 \infty \int_{\mathbb{R}} e^{-\lambda t} P_t(x,dy) f(y) dt \\ &= \int_{\mathbb{R}} \int_0^\infty e^{-\lambda t} \frac{1}{\sqrt{2\pi t}} e^{-\lambda t - \frac{1}{2t}(y-x)^2} dt f(y) dy. \end{aligned}$$

By the further computation, we may obtain that

$$\int_0^\infty e^{-\lambda t} \frac{1}{\sqrt{2\pi t}} e^{-\lambda t - \frac{1}{2t}(y-x)^2} dt = \frac{1}{\sqrt{2\lambda}} e^{-\sqrt{2\lambda}|y-x|} = r_\lambda(x,y).$$

Thus, we may gain the exact value of $R_\lambda f(x)$.

Also, for $f \in \mathscr{D}(\mathscr{G})$, we appeal to L'Hospital's rule to obtain that, as $t \to 0+$,

$$\begin{aligned} t^{-1}(P_t f(x) - f(x)) &= t^{-1} \int_{\mathbb{R}} P_t(x,dy)(f(y) - f(x)) \\ &= t^{-1} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(y-x)^2} (f(y) - f(x)) dy \\ &= t^{-1} \int_{[0,\infty)} \frac{f(x+\sqrt{t}h) - f(x)}{\sqrt{t}h} \frac{1}{\sqrt{2\pi t}} e^{-\frac{h^2}{2}} h dh \\ &+ t^{-1} \int_{[0,\infty)} \frac{f(x-\sqrt{t}h) - f(x)}{\sqrt{t}h} \frac{1}{\sqrt{2\pi t}} e^{-\frac{h^2}{2}} h dh \\ &= t^{-1} \int_{[0,\infty)} \frac{f(x+\sqrt{t}h) + f(x-\sqrt{t}h) - 2f(x)}{(\sqrt{t}h)^2} \frac{1}{\sqrt{2\pi t}} h^2 e^{-\frac{h^2}{2}} dh \\ &\to \frac{1}{2} f''(x) \end{aligned}$$

If $f \in \mathscr{D}(\mathscr{G})$, then $f \in C_0$, $\frac{1}{2}\Delta f$ exists, and $\frac{1}{2}\Delta f \in C_0$. Hence $f \in \mathscr{D}(\frac{1}{2}\Delta)$. Thus, $\frac{1}{2}\Delta$ is obviously an extension of $\mathscr{G}$. For any $f \in \mathscr{D}(\frac{1}{2}\Delta)$ where $f$ attains its maximum at $x \geq 0$, and $f(x) \geq 0$, we have $f'(x) = 0$, thus $f''(x) \leq 0$, hence $\frac{1}{2}\Delta f(x) \leq 0$. By Dynkin's maximum principle, Lemma 2.4, we obtain that

$$\mathscr{G} = \frac{1}{2}\Delta.$$

□

**Corollary 2.3.** *For any $b > 0$, the one-dimensional Brownian motion starting at 0 satisfies*

$$E^0 e^{-\lambda H_b} = e^{-b(2\lambda)^{1/2}}$$

*Proof.* Set $f(x) = e^{x(2\lambda)^{1/2}}$, then $f \in \mathscr{D}(\mathscr{G})$ and

$$(\mathscr{G} - \lambda)f = \frac{1}{2}f'' - \lambda f(\frac{1}{2}2\lambda - \lambda)f = 0.$$

Now we apply Theorem 2.8 on a stopping time $T = H_b$ and a starting point $x = 0$. Then

$$E^0[e^{-\lambda H_b} f(b)] - f(0) = 0,$$

thus $E^0[e^{-\lambda H_b}] = \frac{f(0)}{f(b)} = e^{-b(2\lambda)^{1/2}}$. □

The following example exhibits the movement of the Browian motion in $\mathbb{R}^3$ using Blumenthal's 0-1 Law (Theorem 2.6).

**Example 2.1.** *Let $X$ be the canonical Brownian motion in $\mathbb{R}^3$. Let $W$ be the boundary of a cone $V$ in $\mathbb{R}^3$ and $b$ be the tip of $W$. Prove that $b$ is regular for $W$, that is, $\mathbf{P}^b[H_W = 0] = 1$ where $H_W := \inf\{t > 0 : X_t \in W\}$.*

*Proof.* Since $V$ is a cone and $X_t$ is the canonical Brownian motion, the direction from the starting point of a path at time $t$ is independent of $t$. Thus, we obtain that there exist a real number $\alpha$ with $0 < \alpha < 1$ such that for any $t > 0$,

$$\mathbf{P}^b[X_t \in V] = \mathbf{P}^b[X_1 \in V] =: \alpha$$

Let $n$ be any positive integer.

$$\mathbf{P}^b[H_V > \frac{1}{n}] = \mathbf{P}^b[X_s \notin V; s \leq \frac{1}{n}] \leq \mathbf{P}^b[X_{\frac{1}{n}} \notin V] = 1 - \alpha$$

Hence,

$$\mathbf{P}^b[H_V > 0] = \lim_{n\to\infty} \mathbf{P}^b[H_V > \frac{1}{n}] \leq 1 - \alpha < 1,$$

thus, $\mathbf{P}^b[H_V = 0] = 1 - \mathbf{P}^b[H_V > 0] > 0$. Finally, combining with Blumenthal's 0-1 Law, we obtain that $\mathbf{P}^b[H_V = 0] = 1$.

Now, we may consider $V^c$ as another cone, thus $\mathbf{P}^b[H_{V^c} = 0] = 1$. Since $X$ has a continuous path, if $H_G > 0$, then the path of $X$ would entirely lie within $V$ or within $V^c$ before $H_G$. Each of the cases implies that $H_{V^c} \geq H_G > 0$ and $H_V \geq H_G > 0$, respectively. Thus, we may deduce that

$$\mathbf{P}^b[H_G > 0] \leq \mathbf{P}^b[\{H_V > 0\} \cup \{H_{V^c} > 0\}] = 0,$$

thus, $\mathbf{P}^b[H_G = 0] = 1$. □

We now take a look at continuous one-dimensional Lévy processes, which are R-processes with stationary independent increments; the law of $X_{t+h} - X_t$ does not depend on t. Then the following Lévy's theorem clearly demonstrates how the processes look like. By this theorem, we may conclude that the Lévy processes is, in fact, a type of FD diffusions.

**Theorem 2.14** (Lévy). *For a continuous one-dimensional Lévy process $X_t$, there exists constants $\sigma$ and $\mu$ such that*

$$X_t = \sigma B_t + \mu t$$

*where $B$ is a Brownian motion.*

*Proof.* It is clear that the semigroup $\{P_t\}$ associated with $X$ has the FD property. Let $\mathscr{D} = \{f \in C_0 : f'' \in C_0\}$. Then $C_K^\infty \subset \mathscr{D}$. Also, for any $f \in \mathscr{D}$, there exists $\theta_x \in (0,1)$ such that

$$f'(x) = f(x+1) - f(x) - \frac{1}{2}f''(x + \theta_x),$$

hence $f' \in C_0$ For the characteristic operator $\mathscr{C}$, if we prove that $\mathscr{D} \subset \mathscr{D}(\mathscr{C})$ and that there exists some real numbers $\sigma > 0$ and $\mu$ such that $\forall f \in \mathscr{D}$,

$$\mathscr{C}f = \frac{1}{2}\sigma^2 f'' + \mu f',$$

then $\mathscr{C}$ extends the generator $\mathscr{G}_1$ of $\sigma B_t + \mu t$ from $\mathscr{D}$ to $\mathscr{D}(\mathscr{C})$. This is because $\mathscr{G}_1$ on $\mathscr{D}$ is given by for any $f \in \mathscr{D}$ and $x \in \mathbb{R}$,

$$\begin{aligned}
\mathscr{G}_1 f(x) &= \lim_{t \to 0+} t^{-1}(P_t f(x) - f(x)) \\
&= \lim_{t \to 0+} t^{-1} \int_{\mathbb{R}} (P_t(x, dy))(f(y) - f(x)) \\
&= \lim_{t \to 0+} t^{-1} \int_{\mathbb{R}} (Q_t(x, dz))(f(z + t\mu) - f(z) + f(z) - f(x)) \quad (y = z + t\mu) \\
&= \frac{1}{2}\sigma^2 f'' + \lim_{t \to 0+} t^{-1} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(z-x)^2}(f(z + t\mu) - f(z))dz \quad (z = x + \sqrt{t}w) \\
&= \frac{1}{2}\sigma^2 f'' + \lim_{t \to 0+} \int_{\mathbb{R}} Q_1(0, dw)(f(x + \sqrt{t}w + t\mu) - f(x + \sqrt{t}w))t^{-1} \\
&= \frac{1}{2}\sigma^2 + \mu f'(x)
\end{aligned}$$

by the change of variables where $Q_t(x, \Gamma)$ is a transition function of $\sigma B_t$. The last limit comes from the mean value theorem and the dominated convergence theorem. From the above equation, it is obvious that $\mathscr{D}$ is exactly the domain of

$$\mathscr{G}_1 = \frac{1}{2}\sigma^2 \frac{d^2}{dx^2} + \mu \frac{d}{dx}.$$

Since $X$ is continuous, the maximum principle for $\mathscr{C}$ follows. Then by Theorem 2.4, two generators $\mathscr{G}_1$ and $\mathscr{C}$ are identical. Thus, we may find the corresponding resolvents, and then by uniqueness in the Hille-Yosida theorem, we may conclude that

$$X_t = \sigma B_t + \mu t$$

.

It only remains to show the existence of such $\sigma$ and $\mu$. Let $H_y$ be the hitting time of $y$ and For $a < x < b$, suppose that

$$0 < \mathbf{P}^x[H_a < H_b] = 1 - \mathbf{P}^x[H_a > H_b] < 1.$$

Let $\tau_0^h = 0$ and $\tau_{n+1} = \inf\{t > \tau_n^h : |X_t - X_{\tau_n^h}| = h\}$. Then $\{X_{\tau_n^h}\}$ is merely a random walk. Therefore, when $a = x - kh$ and $b = x + mh$, an event $\{H_a < H_b\}$ becomes a simple gambler's ruin, then we may obtain by easy computation solving a system of equations that there exists real numbers $\beta > 0$ and $\gamma$ with

$$\mathbf{P}^x[H_a < H_b] = \frac{e^{\gamma b} - e^{\gamma x}}{e^{\gamma b} - e^{\gamma a}};$$

$$E^x[H_a \wedge H_b] = \frac{\beta}{\gamma} \frac{(b-x)(e^{\gamma x} - e^{\gamma a})^2 + (x-a)(e^{\gamma b} - e^{\gamma x})^2}{e^{\gamma b} - e^{\gamma a}}.$$

If $\gamma = 0$, we naturally assign the limits of the values as $\gamma \to 0$. If $a = x - kh$ and $b = x + mh$ does not hold, by monotonicity of $\mathbf{P}^x[H_a < H_.]$, we take $h \to 0+$ to obtain the exact same result as above. By the continuity of the path, $\tau_1^h = V_{1,x}$. Then for any $f \in \mathscr{D}$, we compute the value applying L'Hospital's rule;

$$\begin{aligned}
\mathscr{C}f(x) &= \lim_{\eta\to 0+} \frac{E^x[f(X_{\tau_1^\eta})] - f(x)}{E^x[\tau_1^\eta]} \\
&= \lim_{\eta\to 0+} \frac{f(x-\eta)\frac{e^{\gamma(x+\eta)} - e^{\gamma x}}{e^{\gamma(x+\eta)} - e^{\gamma(x-\eta)}} + f(x+\eta)\frac{e^{\gamma x} - e^{\gamma(x-\eta)}}{e^{\gamma(x+\eta)} - e^{\gamma(x-\eta)}} - f(x)}{\frac{\beta}{\gamma}\frac{\eta(e^{\gamma x} - e^{\gamma(x-\eta)})^2 + \eta(e^{\gamma(x+\eta)} - e^{\gamma x})^2}{e^{\gamma(x+\eta)} - e^{\gamma(x-\eta)}}} \\
&= \lim_{\eta\to 0+} \frac{\gamma}{\beta} \frac{f(x-\eta)(e^{\gamma\eta} - 1) + f(x+\eta)(1 - e^{-\gamma\eta}) - f(x)(e^{\gamma\eta} - e^{-\gamma\eta})}{\eta(e^{\gamma\eta} - 2 + e^{-\gamma\eta})} \\
&= \lim_{\eta\to 0+} \frac{\gamma}{\beta} \frac{f(x-\eta)e^{\gamma\eta} + f(x+\eta) - f(x)(e^{\gamma\eta} + 1)}{\eta(e^{\gamma\eta} - 1)} \\
&= \lim_{\eta\to 0+} \frac{-f'(x-\eta)e^{\gamma\eta} + \gamma(f(x-\eta) - f(x))e^{\gamma\eta} + f'(x+\eta)}{2\beta\eta} \\
&= \lim_{\eta\to 0+} \frac{f''(x-\eta) - \gamma f'(x-\eta)e^{\gamma\eta} - \gamma f'(x-\eta)e^{\gamma\eta} + \gamma^2(f(x-\eta) - f(x))e^{\gamma\eta} + f''(x+\eta)}{2\beta} \\
&= \frac{1}{\beta}f'' - \frac{\gamma}{\beta}f'.
\end{aligned}$$

Thus, by the existence of the limit, we have $\mathscr{D} \subset \mathscr{D}(\mathscr{C})$. In addition, by setting $\sigma^2 := \frac{2}{\beta}$ and $\mu = -\frac{\gamma}{\beta}$, we finally obtain that

$$\mathscr{C}f = \frac{1}{2}\sigma^2 f'' + \mu f'.$$

□

# References


[1] R. Dong. *Feller Processes and Semigroups*. 2003. URL: `https://www.stat.berkeley.edu/~pitman/s205s03/lecture27.pdf`.

[2] P. DuChateau. *The Hille-Yosida Theorem*. URL: `https://www.math.colostate.edu/~pauld/M645/HilleYosida.pdf`.

[3] R. Durrett. *Probability: Theory and Examples.* The Wadsworth & Brooks/Cole Statistics/Probability Series. Duxbury Press, 1996. ISBN: 9780534243180. URL: https://books.google.co.kr/books?id=kkc%5C_AQAAIAAJ.

[4] E.B. Dynkin **andothers**. *Markov Processes: Volume 1.* Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012. ISBN: 9783662000335. URL: https://books.google.co.kr/books?id=C6EynwEACAAJ.

[5] O. Kallenberg. *Foundations of Modern Probability.* Probability and Its Applications. Springer New York, 2002. ISBN: 9780387953137. URL: https://books.google.co.kr/books?id=TBgFslMy8V4C.

[6] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* **volume** 82. **january** 2000. ISBN: 978-3-540-60243-9. DOI: 10.1007/978-3-662-03185-8.

[7] L. C. G. Rogers **and** David Williams. *Diffusions, Markov Processes and Martingales.* 2 **edition**. **volume** 2. Cambridge Mathematical Library. Cambridge University Press, 2000. DOI: 10.1017/CBO9780511805141.

[8] L. C. G. Rogers **and** David Williams. *Diffusions, Markov Processes, and Martingales.* 2 **edition**. **volume** 1. Cambridge Mathematical Library. Cambridge University Press, 2000. DOI: 10.1017/CBO9781107590120.

[9] C. Shalizi. *Feller Processes.* URL: https://www.stat.cmu.edu/~cshalizi/754/2006/notes/lecture-14.pdf.