Stochastic Processes in Continuous Time

Joseph C. Watkins

December 14, 2007

Contents

1 Basic Concepts **3**
1.1 Notions of equivalence of stochastic processes . . . 3
1.2 Sample path properties . . . 4
1.3 Properties of filtrations . . . 6
1.4 Stopping times . . . 7
1.5 Examples of stopping times . . . 10

2 Lévy Processes **12**

3 Martingales **16**
3.1 Regularity of Sample Paths . . . 18
3.2 Sample Path Regularity and Lévy Processes . . . 23
3.3 Maximal Inequalities . . . 27
3.4 Localization . . . 28
3.5 Law of the Iterated Logarithm . . . 30

4 Markov Processes **33**
4.1 Definitions and Transition Functions . . . 33
4.2 Operator Semigroups . . . 35
4.2.1 The Generator . . . 35
4.2.2 The Resolvent . . . 38
4.3 The Hille-Yosida Theorem . . . 40
4.3.1 Dissipative Operators . . . 40
4.3.2 Yosida Approximation . . . 41
4.3.3 Positive Operators and Feller Semigroups . . . 45
4.3.4 The Maximum Principle . . . 46
4.4 Strong Markov Property . . . 47
4.5 Connections to Martingales . . . 49
4.6 Jump Processes . . . 53
4.6.1 The Structure Theorem for Pure Jump Markov Processes . . . 53
4.6.2 Construction in the Case of Bounded Jump Rates . . . 55
4.6.3 Birth and Death Process . . . 57
4.6.4 Examples of Interacting Particle Systems . . . 59

4.7 Sample Path Regularity . . . 61
4.7.1 Compactifying the State Space . . . 62
4.8 Transformations of Markov Processes . . . 66
4.8.1 Random Time Changes . . . 66
4.8.2 Killing . . . 67
4.8.3 Change of Measure . . . 69
4.9 Stationary Distributions . . . 71
4.10 One Dimensional Diffusions. . . . 74
4.10.1 The Scale Function . . . 76
4.10.2 The Speed Measure . . . 77
4.10.3 The Characteristic Operator . . . 80
4.10.4 Boundary Behavior . . . 82

5 Stochastic Integrals **84**
5.1 Quadratic Variation . . . 85
5.2 Definition of the Stochastic Integral . . . 90
5.3 The Itô Formula . . . 95
5.4 Stochastic Differential Equations . . . 98
5.5 Itô Diffusions . . . 101

1 Basic Concepts

We now consider stochastic processes with index set $\Lambda = [0, \infty)$. Thus, the process

$$X : [0, \infty) \times \Omega \to S$$

can be considered as a random function of time via its *sample paths* or *realizations*

$$t \to X_t(\omega), \text{ for each } \omega \in \Omega.$$

Here S is a metric space with metric d.

1.1 Notions of equivalence of stochastic processes

As before, for $m \geq 1$, $0 \leq t_1 < \cdots < t_m$, $B \in \mathcal{B}(S^m)$, *the Borel σ-algebra*, call

$$\mu_{t_1,\ldots,t_m}(B) = P\{(X_{t_1}, \ldots, X_{t_m}) \in B\}$$

the *finite dimensional distributions* of X.

We also have a variety of notions of equivalence for two stochastic processes, X and Y.

Definition 1.1. *1. Y is a* version *of X if X and Y have the same finite dimensional distributions.*

2. Y is a modification *of X if for all $t \geq 0$, (X_t, Y_t) is an $S \times S$-random variable and*

$$P\{Y_t = X_t\} = 1.$$

3. Y is indistinguishable *from X if*

$$P\{\omega; Y_t(\omega) \neq X_t(\omega) \text{ for some } t \geq 0\} = 0.$$

Recall that the Daniell-Kolmogorov extension theorem guarantees that the finite dimensional distributions uniquely determine a probability measure on $S^{[0,\infty)}$.

Exercise 1.2. *If $A \in \sigma\{X_t; t \geq 0\}$, then there exists a countable set $C \subset [0, \infty)$ so that $A \in \sigma\{X_t; t \in C\}$.*

Thus, many sets of interest, e.g., $\{x; \sup_{0 \leq s \leq T} f(x_s) > 0\}$ or $\{x : \int_0^T f(x_s)\, ds < \infty\}$ are not measurable subsets of $S^{[0,\infty)}$. Thus, we will need to find methods to find versions of a stochastic process so that these sets are measurable.

Exercise 1.3. *If X and Y are indistinguishable, then X is a modification of Y. If X is a modification of Y, then X and Y have the same finite dimensional distributions.*

1.2 Sample path properties

Definition 1.4. *We call X*

1. measurable *if X is $\mathcal{B}[0,\infty)\times\mathcal{F}$-measurable,*
2. *(almost surely)* continuous, (left continuous, right continuous) *if for (almost) all $\omega\in\Omega$ the sample path is continuous, (left continuous, right continuous).*

Focusing on the regularity of sample paths, we have

Lemma 1.5. *Let $x:[0,\infty)\to S$ and suppose*

$$x_{t+}=\lim_{s\to t+}x_s \text{ exists for all } t\geq 0,$$

and

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

Then there exists a countable set C such that for all $t\in[0,\infty)\backslash C$,

$$x_{t-}=x_t=x_{t+}.$$

Proof. Set

$$C_n=\{t\in[0,\infty);\max\{d(x_{t-},x_t),d(x_{t-},x_{t+}),d(x_t,x_{t+})\}>\frac{1}{n}\}$$

If $C_n\cap[0,m]$ is infinite, then by the Bolzano-Weierstrass theorem, it must have a limit point $t\in[0,m]$. In this case, either x_{t-} or x_{t+} would fail to exist. Now, write

$$C=\bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty}(C_n\cap[0,m]),$$

the countable union of finite sets, and, hence, countable. □

Lemma 1.6. *Let D be a dense subset of $[0,\infty)$ and let $x:D\to S$. If for each $t\geq 0$,*

$$x_t^+=\lim_{s\to t+,s\in D}x_s$$

exists, then x^+ is right continuous. If for each $t>0$,

$$x_t^-=\lim_{s\to t-,s\in D}x_s$$

exists, then x^- is left continuous.

Proof. Fix $t_0>0$. Given $\epsilon>0$, there exists $\delta>0$ so that

$$d(x_{t_0}^+,x_s)\leq\epsilon \text{ whenever } s\in D\cap(t_0,t_0+\delta).$$

Consequently, for all $s\in(t_0,t_0+\delta)$

$$d(x_{t_0}^+,x_s^+)=\lim_{u\to s+,u\in D}d(x_{t_0}^+,x_u)\leq\epsilon.$$

and x^+ is right continuous.

The left continuity is proved similarly. □

Exercise 1.7. *If both limits exist in the previous lemma for all t, then $x_t^- = x_{t-}^+$ for all $t > 0$.*

Definition 1.8. *Let $C_S[0,\infty)$ denote the space of continuous S-valued functions on $[0,\infty)$. Let $D_S[0,\infty)$ denote the space of right continuous S-valued functions having left limits on $[0,\infty)$.*

Even though we will not need to use the metric and topological issue associated with the spaces $C_S[0,\infty)$ and $D_S[0,\infty)$, we proved a brief overview of the issues.

If we endow the space $C_S[0,\infty)$ with the supremum metric $\rho_\infty(x,y) = \sup_{0\le s\le\infty} d(x_s, y_s)$, then the metric will not in general be separable. In analogy with the use of seminorms in a Frechet to define a metric, we set, for each $t > 0$,

$$\rho_t(x,y) = \sup_s d(x_{\max\{s,t\}}, y_{\max\{s,t\}}).$$

Then, ρ_T satisfies the symmetric and triangle inequality properties of a metric. However, if x and y agree on $[0,T]$, but $x_s \neq y_s$ for some $s > T$, then $x \neq y$ but $\rho_t(x,y) = 0$. However, if $\rho_t(x,y) = 0$ for all t, then $x = y$. Consider a bounded metric $\bar{d}(x_0,y_0) = \max\{d(x_0,y_0),1\}$ and set $\bar{\rho}_t(x,y) = \sup_{0\le t\le t} \bar{d}(x_s,y_s)$, then we can create a metric on $C_S[0,\infty)$ which is separable by giving increasingly small importance to large values of t. For example, we can use

$$\bar{\rho}(x,y) = \int_0^\infty e^{-t}\bar{\rho}_t(x,y)\ dt.$$

Then, $(C_S[0,\infty),\bar{\rho})$ is separable whenever (S,d) is separable and complete whenever (S,d) is complete.

For the space $D_S[0,\infty)$, then unless the jumps match up *exactly* then the distance from x to y might be large. To match up the jumps, we introduce a continuous strictly increasing function $\gamma : [0,\infty) \to [,\infty)$ that is one to one and onto and define

$$\rho_t^\gamma(x,y) = \sup_s d(x_{\max\{s,t\}}, y_{\max\{\gamma(s),t\}}).$$

and

$$\rho_(x,y) = \inf_\gamma\{\max\{\text{ess sup}_t |\log\gamma_t'|, \int_0^\infty e^{-t}\rho_t^\gamma(x,y)\ dt\}\}.$$

As before, $(D_S[0,\infty),\bar{\rho})$ is separable whenever (S,d) is separable and complete whenever (S,d) is complete.

Exercise 1.9. *$C_\mathbb{R}[0,\infty)$ with the ρ_∞ metric is not separable.* Hint: *Find an uncountable collection of real-valued continuous functions so that the distance between each of them is 1.*

With a stochastic process X with sample paths in $D_S[0,\infty)$, we have the following moment condition that guarantee that X has a $C_S[0,\infty)$ version.

Proposition 1.10. *If (S,d) be a separable metric space and set $d_1(x,y) = \min\{d(x,y),1\}$. Let X be a process with sample paths in $D_S[0,\infty)$. Suppose for each $T > 0$, there exist $\alpha > 1$, $\beta > 0$, and $C > 0$ such that*

$$E[d_1(X_t,X_s)^\beta] \le C(t-s)^\alpha \quad 0 \le s \le t \le T. \tag{1.1}$$

Then almost all sample paths of X belong to $C_S[0,\infty)$.

Proof. Let T be a positive integer.

Claim.

$$\sum_{0<t\le T} d_1(X_t, X_{t-})^\beta \le \liminf_{N\to\infty} \sum_{k=1}^{2^N T} d_1(X_{k2^{-N}}, X_{(k-1)2^{-N}})^\beta.$$

First note that by Lemma 1.5, the left side is the sum of a countable number of terms. In addition, note that for each $n \ge 1$, $\{t \in [0,T]; d_1(X_t, X_{t-}) > 1/n\}$ is a finite set. Thus, for N sufficiently large, these points are isolated by the 2^{-N} partition of $[0,T]$. Thus, in the limit, these jumps are captured. Consequently,

$$\sum_{0<t\le T} d_1(X_t, X_{t-})^\beta I_{\{d_1(X_t,X_{t-})>1/n\}} \le \liminf_{N\to\infty} \sum_{k=1}^{2^N T} d_1(X_{k2^{-N}}, X_{(k-1)2^{-N}})^\beta. \tag{1.2}$$

Note that the left side of expression (1.2) increases as n increases. Thus, let $n \to \infty$ to establish the claim.

By Fatou's lemma, and the moment inequality (1.1),

$$\begin{aligned} E[\sum_{0<t\le T} d_1(X_t, X_{t-})^\beta] &\le \liminf_{n\to\infty} \sum_{k=1}^{2^n T} E[d_1(X_{k2^{-n}}, X_{(k-1)2^{-n}})^\beta] \\ &\le \liminf_{n\to\infty} \sum_{k=1}^{2^n T} C2^{-n\alpha} \\ &= \liminf_{n\to\infty} CT2^{n(1-\alpha)} = 0. \end{aligned}$$

Thus, for each T, with probability 1, X has no discontinuities on the interval $[0,T]$ and consequently almost all sample paths of X belong to $C_S[0,\infty)$. □

1.3 Properties of filtrations

Definition 1.11. *A collection of sub-σ-algebras $\{\mathcal{F}_t; t \ge 0\}$ of the σ-algebra $\mathcal{F}$ is called a* filtration *if $s \le t$ implies that $\mathcal{F}_s \subset \mathcal{F}_t$.*

For a given stochastic process X, write $\mathcal{F}_t^X$ for the filtration $\sigma\{X_s; 0 \le s \le t\}$. Call $\{\mathcal{F}_t^X; t \ge 0\}$ is called the natural filtration *associated to the process X.*

The filtration $\{\mathcal{F}_t; t \ge 0\}$ in continuous time may have some additional structure

Definition 1.12. *1. $\{\mathcal{F}_t; t \ge 0\}$ is* right continuous *if for each $t \ge 0$,*

$$\mathcal{F}_t = \mathcal{F}_{t+} = \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon}.$$

2. $\{\mathcal{F}_t; t \ge 0\}$ is complete *if $(\Omega, \mathcal{F}, P)$ is complete and $\mathcal{F}_0$ contains all events having probability zero.*

Combining the process X and the filtration $\{\mathcal{F}_t; t \ge 0\}$, we have, as before:

Definition 1.13. *X is adapted if X_t is $\mathcal{F}_t$-measurable. In addition, a process X is $\{\mathcal{F}_t\}$-*progressive *if, for every $t \geq 0$,*

$$X|_{[0,t]\times\Omega} \text{ is } \mathcal{B}[0,t] \times \mathcal{F}_t\text{-measurable.}$$

Exercise 1.14. *1. The filtration $\{\mathcal{F}_{t+}; t \geq 0\}$ is right continuous.*

2. If X is $\{\mathcal{F}_t\}$-progressive, then X is $\{\mathcal{F}_t\}$-adapted and measurable. Hint. *Approximate X on $[0,t] \times \Omega$ by $X^n_s(\omega) = X_{\min\{t,([ns]+1)/n)\}}(\omega)$.*

3. If X is $\{\mathcal{F}_t\}$-adapted and right continuous, then is $\{\mathcal{F}_t\}$-progressive.

4. Let X be $\{\mathcal{F}_t\}$-progressive and let $f : S \to R$ be bounded and measurable. Show that $f \circ X$ is $\{\mathcal{F}_t\}$-progressive and $A_t = \int_0^t f(X_s)\, ds$ is $\{\mathcal{F}_t\}$-adapted.

1.4 Stopping times

A stopping time, τ, and the corresponding stopped σ-algebra $\mathcal{F}_\tau$ are defined for continuous time processes in a manner analogous to the discrete time case.

Definition 1.15. *A non-negative (possibly infinite) random variable is an $\{\mathcal{F}_t\}$-*stopping time *if for every $t \geq 0$,*

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

The stopped σ-algebra

$$\mathcal{F}_\tau = \{A \in \mathcal{F}; A \cap \{\tau \leq t\} \in \mathcal{F}_t, \text{ for all } t \geq 0\}.$$

Many of the same properties hold with the same proof as for discrete time processes.

Proposition 1.16. *1. The supremum of a countable number of stopping times is a stopping time.*

2. The minimum of a finite number of stopping times is a stopping time.

3. $\mathcal{F}_\tau$ is a σ-algebra

4. For a second stopping time σ, $\min\{\tau, \sigma\}$ is $\mathcal{F}_\tau$-measurable.

5. If $\sigma \leq \tau$, then $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$

The following proposition holds in discrete time. However the proof in continuous is more involved.

Proposition 1.17. *Let X be an $\mathcal{F}_t$-progressive process and τ a finite stopping time. Then X_τ is $\mathcal{F}_\tau$ measurable.*

Proof. Pick $\Gamma \in \mathcal{B}(S)$. We must show that $\{X_\tau \in B\} \in \mathcal{F}_\tau$. Consequently, for each $t \geq 0$, we must show that.

$$\{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t.$$

We shall accomplish this using a compositions of maps. Note that by the previous proposition, $\min\{\tau, t\}$ is $\mathcal{F}_t$-measurable. Thus

$$\omega \mapsto (\min\{\tau(\omega), t\}, \omega)$$

is a measurable mapping from $(\Omega, \mathcal{F}_t)$ to $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t)$.

Next, because X is an $\mathcal{F}_t$-progressive process,

$$(s,\omega) \mapsto X_s(\omega)$$

is a measurable mapping from $([0,t] \times \Omega, \mathcal{B}([0,t]) \times \mathcal{F}_t)$ to $(S, \mathcal{B}(S))$.

Thus, the composition

$$\omega \mapsto X_{\min\{\tau(\omega),t\}}(\omega)$$

is a measurable mapping from $(\Omega, \mathcal{F}_t)$ to $(S, \mathcal{B}(S))$ and $\{X_{\min\{\tau,t\}} \in \Gamma\} \in \mathcal{F}_t$.

Finally, $\{X_\tau \in B\} \cap \{\tau \leq t\} = \{X_{\min\{\tau,t\}} \in B\} \cap \{\tau \leq t\}$. Because both of these events are elements of the σ-algebra $\mathcal{F}_t$, then so is their intersection. □

Some additional properties are connected to the continuity of time.

Lemma 1.18. *A $[0,\infty]$-valued random variable τ is an $\mathcal{F}_{t+}$-stopping time if and only if $\{t < \tau\} \in \mathcal{F}_t$ for every $t \geq 0$.*

Proof. (necessity)

$$\{\tau < t\} = \bigcup_{n=1}^{\infty} \{\tau \leq t - \frac{1}{n}\} \in \mathcal{F}_t \subset \mathcal{F}_{t+}.$$

(sufficiency) For $n \geq m$,

$$\{\tau < t + \frac{1}{n}\} \in \mathcal{F}_{t+1/m}.$$

Therefore,

$$\{\tau \leq t\} = \bigcap_{n=1}^{\infty} \{\tau < t + \frac{1}{n}\} \in \bigcap_{m=1}^{\infty} \mathcal{F}_{t+1/m} \subset \mathcal{F}_{t+}.$$

□

Corollary 1.19. *For an $\mathcal{F}_{t+}$-stopping time τ,*

$$\{\tau = t\} \in \mathcal{F}_t.$$

Proof. $\{\tau = t\} = \{\tau \leq t\} \backslash \{\tau < t\} \in \mathcal{F}_t$. □

Proposition 1.20. *Let $\{\tau_n; n \geq 1\}$. If $\mathcal{F}_t$ is right continuous, then $\inf_n \tau_n, \liminf_{n\to\infty} \tau_n$, and*

$$\limsup_{n\to\infty} \tau_n$$

are $\mathcal{F}_t$-stopping times.

Proof. Use the lemma above and note that

$$\{\inf_n \tau_n < t\} = \bigcap_{n=1}^{\infty} \{\tau_n < t\}.$$

Now use

$$\liminf_{n\to\infty} \tau_n = \sup_m \inf_{n\geq m} \tau_n \text{ and } \limsup_{n\to\infty} \tau_n = \inf_m \sup_{n\geq m} \tau_n.$$

□

Proposition 1.21. *Every $\mathcal{F}_t$-stopping time is the nonincreasing limit of bounded stopping times.*

Proof. Take the stopping times $\tau_n = \min\{\tau, n\}$ for $n = 1, 2, \dots$. □

Definition 1.22. *Call a stopping time* discrete *if its range is almost surely countable.*

Proposition 1.23. *Every $\mathcal{F}_t$-stopping time is the nonincreasing limit of discrete stopping times.*

Proof. For $n = 1, 2, \dots$, choose the nested collection of sequences $0 = t_0^n < t_1^n < \cdots$ with $\lim_{k\to\infty} t_k^n = \infty$, and $\lim_{n\to\infty} \sup_k (t_{k+1}^n - t_k^n) = 0$. For an $\mathcal{F}_{t+}$-stopping time τ, define

$$\tau_n = \begin{cases} t_k^n & \text{if } t_{k-1}^n \leq \tau < t_k^n \\ \infty & \text{if } \tau = \infty. \end{cases}$$

Then clearly τ_n is discrete and $\tau = \lim_{n\to\infty} \tau_n$. To see that τ_n is a stopping time, set $\gamma_n(t) = \max\{t_k^n; t_k^n \leq t\}$. Then

$$\{\tau_n \leq t\} = \{\tau_n \leq \gamma_n(t)\} = \{\tau \leq \gamma_n(t)\} \in \mathcal{F}_{\gamma_n(t)} \subset \mathcal{F}_t.$$

□

Exercise 1.24. *Let $\{\tau_n; n \geq 1\}$ be a sequence of $\mathcal{F}_t$-stopping times and let $\tau = \lim_{n\to\infty} \tau_n$.*

1. *If $\{\tau_n; n \geq 1\}$ is nondecreasing, then τ is an $\mathcal{F}_t$-stopping time.*

2. *If $\{\tau_n; n \geq 1\}$ is nonincreasing, then τ is an $\mathcal{F}_{t+}$-stopping time.*

Definition 1.25. *For a stochastic process X and a finite stopping time τ Define the* stopped process

$$X_t^\tau = X_{\min\{\tau, t\}}$$

and the shifted process

$$\theta^\tau X_t = X_{\tau+t}.$$

Proposition 1.26. *Let X be an $\mathcal{F}_t$-progressive process and let τ be a finite stopping time. Define the collection of σ-algebras*

$$\mathcal{G}_t = \mathcal{F}_{\min\{\tau,t\}} \text{ and } \mathcal{H}_t = \mathcal{F}_{\tau+t}.$$

Then

1. *$\{\mathcal{G}_t; t \geq 0\}$ is a filtration and X^τ is both $\mathcal{G}_t$-progressive and $\mathcal{F}_t$-progressive.*

2. *$\{\mathcal{H}_t; t \geq 0\}$ is a filtration and $\theta^\tau X$ is $\mathcal{H}_t$-progressive.*

Proof. For the first statement, because $\{\min\{\tau, t\}; t \geq 0\}$ is an increasing set of stopping times, $\{\mathcal{G}_t; t \geq 0\}$ is a filtration. Because $\mathcal{G}_t \subset \mathcal{F}_t$, then if X^τ is $\mathcal{G}_t$-progressive, then it is automatically $\mathcal{F}_t$-progressive. With this in mind, choose $\Gamma \in \mathcal{B}(S)$ and $t \geq 0$. We show that

$$\{(s, \omega) \in [0, t] \times \Omega; X_{\min\{\tau(\omega), s\}}(\omega) \in \Gamma\} \in \mathcal{B}[0, t] \times \mathcal{G}_t.$$

Now, define

$$\mathcal{C}_{s,t} = \{C \cap ([0, t] \times \{\min\{\tau, t\} \geq s\}); C \in \mathcal{B}[0, t] \times \mathcal{F}_s\}.$$

Claim: $\mathcal{C}_{s,t} \subset \mathcal{B}([0,t]) \times \mathcal{G}_t$.

Clearly, $\mathcal{C}_{s,t}$ is a σ-algebra. Thus, it suffices to choose C of the form $B \times A$ where $B \in \mathcal{B}([0,t])$ and $A \in \mathcal{F}_s$ noting that the collection of such $B \times A$ generates $\mathcal{B}[0,t] \times \mathcal{F}_s$.

$A \cap \{\min\{\tau, t\} \geq s\} \in \mathcal{G}_t$. Thus, for $B \in \mathcal{B}[0,t]$,

$$(B \times A) \cap ([0,t] \times \{\min\{\tau, t\} \geq s\}) = B \times (A \cap \{\min\{\tau, t\} \geq s\}) \in \mathcal{B}[0,t] \times \mathcal{G}_t.$$

Thus, $B \times A \in \mathcal{C}_{s,t}$.

Now, write

$$\{(s,\omega) \in [0,t] \times \Omega; X_{\min\{\tau(\omega),s\}}(\omega) \in \Gamma\}$$

$$= \{(s,\omega) \in [0,t] \times \Omega; X_{\min\{\tau(\omega),s\}}(\omega) \in \Gamma, \min\{\tau(\omega), t\} \leq s \leq t\} \cup \{(s,\omega); X_s(\omega) \in \Gamma, s < \min\{\tau, t\}\}.$$

To see that the first term is in $\mathcal{B}[0,t] \times \mathcal{G}_t$, write it as

$$\{(s,\omega); \min\{\tau(\omega), t\} \leq s \leq t\} \cap ([0,t] \times \{X_{\min\{\tau,s\}}(\omega) \in \Gamma\})$$

and note that

$$\{(s,\omega); \min\{\tau(\omega), t\} \leq s \leq t\} = \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} ([\frac{k}{n}, t] \times \{\frac{k}{n} \leq \min\{\tau, t\} < \frac{k+1}{n}\}).$$

Similarly, the second term equals

$$\bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \{(s,\omega); X_s(\omega) \in \Gamma, s < \frac{k}{n}\} \cap \{(s,\omega); \frac{k}{n} \leq \min\{\tau(\omega), t\}\}$$

and use the claim.

For the second statement, first note that by the same argument as before, $\{\mathcal{H}_t; t \geq 0\}$ is a filtration. By the first part, the mapping

$$(s,\omega) \to X_{\min\{\tau(\omega)+t,s\}}$$

from $([0,\infty] \times \Omega, \mathcal{B}[0,\infty] \times \mathcal{F}_{\tau+t})$ to $(S, \mathcal{B}(S))$ is measurable. We also have the measurable mapping

$$(u,\omega) \to (\tau(\omega) + u, \omega)$$

from $([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{F}_{\tau+t})$ to $([0,\infty] \times \Omega, \mathcal{B}[0,\infty] \times \mathcal{F}_{\tau+t})$. The mapping

$$(u,\omega) \to X_{\tau+u}(\omega)$$

is the composition of these two mappings, so it is measurable. Consequently, $\theta^\tau X$ is $\mathcal{H}_t$-progressive. $\square$

1.5 Examples of stopping times

Definition 1.27. *Let $x : [0,\infty) \to S$ and let $s \geq 0$.*

1. *The* first entrance time into A *after time s is defined by*

$$\tau_e(A,s) = \inf\{t \geq s; x_t \in A\}.$$

2. *The* first contact time into A *after time* s *is defined by*

$$\tau_c(A, s) = \inf\{t \geq s; cl\{x_u; s \leq u \leq t\} \cap A \neq \emptyset\}.$$

   Here $cl(B)$ *is the closure of the set* B.

The first exit *from* B *after time* s *is the first entrance to* B^c. *If* $s = 0$, *then it is suppressed in the notation.*

Proposition 1.28. *Suppose that* X *is a right continuous* $\mathcal{F}_t$*-adapted process and that* σ *is an* $\mathcal{F}_t$*-stopping time.*

1. *If* A *is closed and* X *has left limits at each* $t > 0$ *or if* A *is compact then* $\tau_c(A, \sigma)$ *is an* $\mathcal{F}_t$*-stopping time.*

2. *If* A *is open then* $\tau_e(A, \sigma)$ *is an* $\mathcal{F}_{t+}$*-stopping time.*

Proof. If A is open, we have by the right continuuity of X,

$$\{\tau_e(A, \sigma) < t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t]} \{X_s \in A\} \cap \{\sigma < s\} \in \mathcal{F}_t$$

proving 2.

Under the conditions of part 1,

$$\tau_c(A, \sigma) = \lim_{n \to \infty} \tau_e(A^{1/n}, \sigma)$$

where $A^\epsilon = \{x; d(x, A) < \epsilon\}$. Thus

$$\{\tau_c(A, \sigma) \leq t\} = (\{\sigma \leq s\} \cap \{X_t \in A\} \cup \bigcap_{n=1}^{\infty} \{\tau_e(A^{1/n}, \sigma) < t\} \in \mathcal{F}_t$$

proving 1. $\square$

2 Lévy Processes

We begin our study of continuous time stochastic process with the continuous time analog of random walks.

Definition 2.1. *An $\mathbb{R}^d$-valued process X is call a* Lévy process *or a* process with stationary independent increments *if*

1. *(stationary increments) The distribution of $X_{t+s} - X_t$ does not depend on t.*
2. *(independent increments) Let $0 = t_0 < t_1 < \cdots < t_n$. Then the random variables*
$$\{X_{t_j} - X_{t_{j-1}}; j = 1, \ldots, n\}$$
are independent.
3. *(stochastic continuity)*
$$X_t - X_0 \to^P 0 \text{ as } t \to 0.$$

Exercise 2.2. 1. *Any linear combination of independent Lévy processes is a Lévy process.*

2. *If X has finite mean, $X_0 = 0$ and $\mu = EX_1$, then $EX_t = \mu t$.*
3. *If, in addition, X has finite variance and $\sigma^2 = Var(X_1)$, then $Var(X_t) = \sigma^2 t$.*

Definition 2.3. *A* Poisson process *N with parameter λ is a Lévy process with $N_0 = 0$ and*

$$P\{N_t = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

Exercise 2.4. *Show that the definition of a Poisson process satisfies the compatibility condition in the Daniell-Kolmogorov extension theorem.*

Definition 2.5. *Let $\{Y_k, k \geq 1\}$ be independent and identically distributed $\mathbb{R}^d$-valued random variables and let N be a Poisson process, then*

$$X_t = \sum_{k=1}^{N_t} Y_k$$

is called a compound Poisson process.

Exercise 2.6. *Show that a compound Poisson process is a Lévy process.*

Exercise 2.7. *Let Y_k take on the values ± 1 each with probability $1/2$ and let X be the associated compound Poisson process. Assume that N has parameter λ. Then*

$$P\{X_t = m\} = e^{-\lambda t} I_m(\lambda t)$$

where I_m is the modified Bessel function of order m. Hint. *The modified Bessel functions of integer order $I_m(x)$ have generating function* $\exp(x(z + z^{-1})/2)$.

If Y_k is not integer valued, then we use the characteristic function ϕ_s for X_s

$$\begin{aligned}
\phi_s(u) &= Ee^{i\langle u,X_s\rangle} = E[E[e^{i\langle u,X_s\rangle}|N_s]] \\
&= \sum_{n=0}^{\infty} E[e^{i\langle u,(Y_1+\cdots+Y_n)\rangle}]P\{N_t = n\} \\
&= \sum_{n=0}^{\infty} E[e^{i\langle u,Y_1\rangle}]^n \frac{(\lambda t)^n}{n!} e^{\lambda t} \\
&= \exp(\lambda t(\phi_Y(u) - 1)) = \exp(\lambda t \int (e^{i\langle u,y\rangle} - 1)\ \nu(dy))
\end{aligned}$$

where ν and ϕ_Y are, respectively, the distribution and the characteristic function for the Y_k.

Exercise 2.8. *For a Lévy process X let ν_s and $\phi_s(u)$ denote, respectively, the distribution and the characteristic function for the X_s. Then, $\phi_{s+t} = \phi_s\phi_t$ and consequently $\nu_{s+t} = \nu_s * \nu_t$.*

Use the dominated convergence theorem to show that the stochastic continuity of a Lévy process implies that ϕ_s is right continuous and so, by the functional relationship in the exercise above,

$$\phi_s(u) = \exp s\psi(u),$$

for some ψ.

Definition 2.9. *The function ψ is called* the characteristic exponent *for the Lévy process.*

Definition 2.10. *A one dimensional* Brownian motion *B with parameters μ and σ^2 is a Lévy process with $B_0 = 0$ and*

$$P\{B_t \in A\} = \frac{1}{\sigma\sqrt{2\pi t}} \int_A \exp\left(-\frac{(x-\mu t)^2}{2\sigma^2 t}\right)\ dx.$$

If $\mu = 0$ and $\sigma^2 = 1$, the B is called a standard *Brownian motion.*

For d-dimensional Brownian motion B, we require $\mu \in \mathbb{R}^d$ and a symmetric, positive semi-definite matrix Σ, Then, B is a Lévy process and if $B_0 = 0$, then

$$P\{B_t \in A\} = \frac{1}{\sqrt{det(\Sigma)(2\pi t)^n}} \int_A \exp\left(-\frac{(x-\mu t)^T\Sigma^{-1}(x-\mu t)}{2t}\right)\ dx.$$

Exercise 2.11. *1. Show that Brownian motion as defined above satisfies the consistency condition in the extension theorem.*

2. Find the characteristic exponent for Brownian motion.

Exercise 2.12. *Let B be a d-dimensional Brownian motion. Show that there exists an invertible matrix C , a vector μ, and a d-dimensional process $\tilde{B}_t = (\tilde{B}_t^1, \ldots, \tilde{B}_t^d)$ in which the component are independent standard Brownian motions so that $B_t = C\tilde{B}_t + \mu t$.*

As a consequence, we sometimes call B a Brownian motion if $B_t = C\tilde{B}_t$ for some $d \times d$-matrix C.

Exercise 2.13. *Let B be a standard Brownian motion. Then the following are also standard Brownian motions.*

1. $-B$,
2. $\{B_{t+t_0} - B_{t_0}; t \geq 0\}$ *for* $t_0 > 0$.
3. $\{aB_{t/a^2}; t \geq 0\}$ *for* $a > 0$.

Proposition 2.14. *Let B be a standard Brownian motion. Then,*

$$P\{\sup_t B_t = +\infty,\ \inf_t B_t = -\infty\} = 1.$$

Proof. Set $B^* = \sup\{B_t; t \in \mathbb{Q} \cap [0,\infty)\}$. Then, for any $a \in \mathbb{Q}^+$, B^* and aB^* have the same distribution. Consequently, B^* is concentrated on the set $\{0, \infty\}$. Set $p^* = P\{B^* = 0\}$. Then

$$\begin{aligned}
p^* &\leq P\{B_1 \leq 0, B_u \leq 0, u \in \mathbb{Q} \cap [1,\infty)\} \\
&\leq P\{B_1 \leq 0, \sup\{B_{1+t} - B_1; t \in \mathbb{Q}^+\} < \infty\} \\
&= P\{B_1 \leq 0\} P\{\sup\{B_{1+t} - B_1; t \in \mathbb{Q}^+\} < \infty\} \\
&= P\{B_1 \leq 0\} P\{\sup\{B_t; t \in \mathbb{Q}^+\} < \infty\} = \frac{1}{2} p^*
\end{aligned}$$

Thus, $p^* = 0$ and $P\{\sup_t B_t = +\infty\} = 1$. Because $-B$ is a standard Brownian motion, $P\{\inf_t B_t = -\infty\} = 1$. Because the intersection of a pair of probability 1 events has probability 1, the proposition follows. ☐

Take $X_0 = 0$. Then, we can write

$$X_t = (X_t - X_{(n-1)t/n}) + (X_{(n-1)t/n} - X_{(n-1)t/n}) + \cdots + (X_{t/n} - X_0).$$

Thus, for any n, X_t can be written as the sum of n independent identically distributed random variables. This motivates the following:

Definition 2.15. *A random variable Y is said to have an infinitely divisible distribution if for every n, there exists n independent identically distributed random variables, $Y_{1,n}, \ldots, Y_{n,n}$ such that*

$$Y = Y_{1,n} + \cdots + Y_{n,n}.$$

The *Lévy-Khintchine formula* gives the characteristic function for any infinitely divisible distribution on $\mathbb{R}^d$,

$$\exp \psi(u) = \exp\left(i\langle \beta, u\rangle - \frac{1}{2}\langle u, \Sigma u\rangle + \int \left(e^{i\langle u,y\rangle} - 1 - \frac{i\langle u, y\rangle}{1+|y|^2} \right)\ \nu(dy) \right).$$

For this formula, there is some flexibility in the last function $\chi(y) = y/(1+|y|^2)$. We need χ to be antisymmetric, bounded and have derivative 1 at the origin. We can choose β to be any d-dimensional vector and Σ to be any symmetric non-negative definite $d \times d$-matrix. The measure ν, called the *Lévy measure* can be any Radon measure satisfying

1. $\nu\{0\} = 0$, and
2. $\int \frac{|y|^2}{1+|y|^2}\ \nu(dy) < \infty$.

Now, clearly $\exp(\psi(u)/n)$ is the characteristic function of a random variable. The sum of n such independent random varibles has characteristic function $\exp \psi(u)$. Thus, we have the converse that any infinitely divisible distribution can be realized by X_1 for some Lévy process X.

If $\int \frac{|y|}{1+|y|^2}\ \nu(dy) < \infty$, then we can write

$$\exp \psi(u) = \exp(i\langle\tilde{\beta}, u\rangle - \frac{1}{2}\langle u, \Sigma u\rangle + \int (e^{i\langle u,y\rangle} - 1)\ \nu(dy)).$$

where

$$\tilde{\beta} = \beta + \int \frac{y}{1+y^2}\ \nu(dy)$$

This is the characteric function of a random variable that is the independent sum of a normal random variable and a compound Poisson process evaluated at time 1.

If $\lambda = \nu(\mathbb{R}^d\backslash\{0\}) < \infty$, and if $CC^T = \Sigma$, then the process

$$X_t = \beta t + C\tilde{B}_t + \sum_{n=1}^{N_t} Y_n - \mu\lambda t$$

the X is a Lévy process with characteristic exponent ψ where $\tilde{B}_t = (\tilde{B}^1_t, \ldots, \tilde{B}^d_t)$ with components that are independent standard Brownian motions, N is a Poisson process with parameter λ and $\{Y_n; n \geq 1\}$ is an independent sequence of random variables with common distrobution ν/λ.

F

3 Martingales

The definition of martingale in continuous time is exactly the analog of the discete time definition.

Definition 3.1. *A real valued process X with $E|X_t| < \infty$ for all $t \geq 0$ and adapted to a filtration $\{\mathcal{F}_t; t \geq 0\}$ is an $\mathcal{F}_t$-*martingale *if*
$$E[X_t|\mathcal{F}_s] = X_s \quad \text{for all } t > s,$$
*an $\mathcal{F}_t$-*submartingale *if*
$$E[X_t|\mathcal{F}_s] \geq X_s \quad \text{for all } t > s,$$
and an $\mathcal{F}_t$-;gale *if the inequality above is reversed.*

Remark 3.2. *Many previous facts about discrete martingales continue to hold for continuous time martingales.*

1. *X is an $\mathcal{F}_t$ martingale if and only if for every $s < t$, $E[X_t; A] = E[X_s; A]$, for all $A \in \mathcal{F}_s$.*
2. *Suppose X is an $\mathcal{F}_t$-martingale, ϕ is convex and $E|\phi(X_t)| < \infty$ for all $t \geq 1$. Then $\phi \circ X$ is an $\mathcal{F}_t$-submartingale.*
3. *Suppose X is an $\mathcal{F}_t$-submartingale, ϕ is convex and non-decreasing, $E|\phi(X_t)| < \infty$ for all $t \geq 0$. Then $\phi \circ X$ is an $\mathcal{F}_t$-submartingale.*

Exercise 3.3. *An adapted stochastic process M is a martingale if and only if*
$$EM_\tau = EM_0$$
for every bounded stopping time τ.

Example 3.4. *Let X be a Lévy process, $X_0 = 0$.*

1. *If X has finite mean, $\mu = EX_1$. Then $X_t - \mu t$ is a martingale.*
2. *If X has mean zero and finite variance, $\sigma^2 = Var(X_1)$. Then $X_t^2 - \sigma^2 t$ is a martingale.*
3. *Let X have characteristic exponent ψ. Then $\exp(iuX_t - t\psi(u))$ is a complex-valued martingale.*

Exercise 3.5. *Compute the martingales above for the Poisson process and one-dimensional Brownian motion.*

The optional sampling theorem continues to hold. In this context, let $\{\tau_t; t \geq 0\}$ be an increasing collection of stopping times and set
$$Y_t = X_{\tau_t}, \quad \text{and} \quad \mathcal{G}_t = \mathcal{F}_{\tau_t}.$$
Assume that each τ_t satsifies the sampling integrability conditions for X. If X is an $\mathcal{F}_t$-submartingale, then Y is a $\mathcal{G}_t$-submartingale.

We have similar statements in the discrete time case with $0 = \tau_0 \leq \tau_1 \leq \cdots$ and $Y_n = X_{\tau_n}$. One important special case of this is the choice of $\tau_n = t_n$ where $\{t_n; n \geq 1\}$ is a set of non-random times.

Proposition 3.6. *Let τ and σ be two $\mathcal{F}_t$-stopping times taking values in a finite set $F = \{t_1 < \cdots < t_m\}$. If X is an $\mathcal{F}_t$-submartingale, then*

$$E[X_\tau|\mathcal{F}_\sigma] \geq X_{\min\{\tau,\sigma\}}.$$

Proof. We show that for every $A \in \mathcal{F}_\sigma$,

$$E[X_\tau; A] \geq E[X_{\min\{\tau,\sigma\}}; A].$$

Write $A = \bigcup_{i=1}^m (A \cap \{\sigma = t_i\})$. to see that it is sufficient to show that for each i,

$$E[X_\tau; A \cap \{\sigma = t_i\}] \geq E[X_{\min\{\tau,\sigma\}}; A \cap \{\sigma = t_i\}] \geq E[X_{\min\{\tau,t_i\}}; A \cap \{\sigma = t_i\}].$$

Because $A \cap \{\sigma = t_i\} \in \mathcal{F}_{t_i}$, the inequality above holds provided that

$$E[X_\tau|\mathcal{F}_{t_i}] \geq X_{\min\{\tau,t_i\}}.$$

Claim. $E[X_{\min\{\tau,t_{k+1}\}}|\mathcal{F}_{t_k}] \geq X_{\min\{\tau,t_k\}}$.

$$\begin{aligned} E[X_{\min\{\tau,t_{k+1}\}}|\mathcal{F}_{t_k}] &= E[X_{\min\{\tau,t_{k+1}\}} I_{\{\tau > t_k\}}|\mathcal{F}_{t_k}] + E[X_{\min\{\tau,t_{k+1}\}} I_{\{\tau \leq t_k\}}|\mathcal{F}_{t_k}] \\ &= E[X_{t_{k+1}}|\mathcal{F}_{t_k}] I_{\{\tau > t_k\}} + X_{\min\{\tau,t_k\}} I_{\{\tau \leq t_k\}} \\ &\geq X_{t_k} I_{\{\tau > t_k\}} + X_{\min\{\tau,t_k\}} I_{\{\tau \leq t_k\}} = X_{\min\{\tau,t_k\}} \end{aligned}$$

Thus, $E[X_\tau|\mathcal{F}_{t_i}] = E[X_{\min\{\tau,t_m\}}|\mathcal{F}_{t_i}] \geq X_{\min\{\tau,t_{m-1}\}}$. Use this for the frist step and the tower property of σ-algebras in the induction step to complete the proof. □

Example 3.7. *Let N be a Poisson process with parameter λ. Let $\tau_n = \inf\{t \geq 0; N_t = n\}$.*

1. *Fix a time t, then $\min\{\tau_n, t\}$is bounded, so by the optional sampling theorem,*

$$0 = EN_{\min\{\tau_n,t\}} - \lambda E \min\{\tau_n, t\}.$$

Now, use the monotone convergence theorem on each of these terms and the fact that $\tau_n < \infty$ almost surely to obtain

$$0 = EN_{\tau_n} - \lambda E\tau_n \text{ and, therefore, } E\tau_n = \frac{n}{\lambda}$$

2. *Consider*

$$\exp(iuN_t - \lambda t(e^{iu} - 1)),$$

the exponential martingale for the Poisson process. Then

$$1 = E[\exp(iuN_{\tau_n} - \lambda\tau_n(e^{iu} - 1)],$$

and thus

$$e^{-iun} = E[\exp(-\lambda\tau_n(e^{iu} - 1)].$$

Set

$$v = \lambda(e^{iu} - 1) \text{ or } u = -i\log(\frac{v + \lambda}{\lambda})$$

yielding

$$Ee^{-v\tau} = (\frac{\lambda}{\lambda+v})^n$$

This is the n-th power of the Laplace transform of an exponential random variable and, hence, the Laplace transform of a $\Gamma(n,\lambda)$ random variable.

Exercise 3.8. *Using N and τ_n as in the example above. Show that $\tau_{n+1}-\tau_n$ is an exponential random variable independent of $\mathcal{F}^N_{\tau_n}$ t see that $\{\tau_{n+1}-\tau_n; n \geq 1\}$ is an independent and identically distributed sequence of exponential random variables.*

Exercise 3.9. *Note that $\{N_t < n\} = \{\tau_n > t\}$. Use this to obtain the density of a $\Gamma(n,\lambda)$ random variable.*

3.1 Regularity of Sample Paths

Lemma 3.10. *Let X be a submartingale, $t > 0$ and F a finite subset of $[0,t]$. Then for each $x > 0$,*

$$P\{\max_{s\in F} X_s \geq x\} \leq \frac{1}{x} EX_t^+,$$

and

$$P\{\min_{s\in F} X_s \leq -x\} \leq \frac{1}{x}(EX_t^+ - EX_0).$$

Proof. The first statement follows immediately from the corresponding result for discrete time submartingales. For the second statement, define

$$\tau = \min\{u \in F; X_u \leq -x\}.$$

By the proposition above, $E[X_t|\mathcal{F}_\tau] \geq X_{\min\{\tau,t\}}$. In addition, if $\tau < \infty$, then $\min \tau, t\} = \tau$.

$$\begin{aligned} EX_0 &\leq EX_{\min\{\tau,t\}} = E[X_{\min\{\tau,t\}}; \{\tau < \infty\}] + E[X_{\min\{\tau,t\}}; \{\tau = \infty\}] \\ &\leq E[X_\tau; \{\tau < \infty\}] + E[X_t; \{\tau = \infty\}] \leq -xP\{\tau < \infty\} + EX_t^+. \end{aligned}$$

Now simplify. □

Corollary 3.11. *Let X be a submartingale, $t > 0$ and C a countable subset of $[0,t]$. Then for each $x > 0$,*

$$P\{\max_{s\in C} X_s \geq x\} \leq \frac{1}{x} EX_t^+,$$

and

$$P\{\min_{s\in C} X_s \leq -x\} \leq \frac{1}{x}(EX_t^+ - EX_0).$$

Proof. Choose finite subests $F_1 \subset F_2 \subset \cdots$ so that $C = \cup_{n=1}^\infty F_n$. Then, for $0 < \tilde{x} < x$,

$$P\{\max_{s\in C} X_s \geq x\} \leq \lim_{n\to\infty} P\{\max_{s\in F_n} X_s \geq \tilde{x}\} \leq \frac{1}{\tilde{x}} EX_t^+.$$

Now, let $\tilde{x} \to x$. □

Corollary 3.12. *Let D be a countable dense subset of $[0,\infty)$. Then*

$$P\{\sup_{s\in D\cap[0,t]} X_s < \infty\} = P\{\inf_{s\in D\cap[0,t]} X_s \geq -\infty\} = 1.$$

We can now make similar analogies to the discrete time case for upcrossings. For a stochastic process X, define $U(a,b,F)$ to be the number of upcrossings of the interval (a,b) by X restricted to a finite set F. For C countable, as before, write $C = \cup_{n=1}^{\infty} F_n$ and define Then $U(a,b,F_n)$ is a monotone increasing sequence. Call its limit $U(a,b,C)$.

Exercise 3.13. *Show that the definition above is independent of the choice of the finite sets F_n.*

Theorem 3.14 (Doob's upcrossing inequality)**.** *Let X be a submartingale and let $t > 0$. For a countable subset $C \subset [0,t]$,*

$$EU(a,b,C) \leq \frac{E(X_t - a)^+}{b-a}.$$

Proof. Choose F_n as above, then by the discrete time version of the uncrossing inequality

$$EU(a,b,F_n) \leq \frac{E(X_t - a)^+}{b-a}.$$

Now, use the monotone convergence theorem. □

Corollary 3.15. *Let D be a countable dense subset of $[0,\infty)$. Then*

$$P\{U(a,b,D\cap[0,t]) < \infty\} = 1.$$

Remark 3.16. *Set*

$$\Omega_0 = \bigcap_{n=1}^{\infty}\left(\{\sup_{s\in D\cap[0,n]} X_t < \infty\} \cap \{\inf_{s\in D\cap[0,n]} X_t \geq -\infty\} \cap \bigcap_{a<b,a,b\in D} \{U(a,b,D\cap[0,n]) < \infty\}\right).$$

Then, $P(\Omega_0) = 1$.

Exercise 3.17. *For $\omega \in \Omega_0$,*

$$X_t^+(\omega) = \lim_{s\to t+, s\in D} X_s(\omega)$$

exists for all $t \geq 0$.

$$X_t^-(\omega) = \lim_{s\to t-, s\in D} X_s(\omega)$$

exists for all $t > 0$. Furthermore, $X^+(\omega) \in D_S[0,\infty)$ and $X_{t-}^+(\omega) = X_t^-(\omega)$.

Set $X_t^+(\omega) = 0$ for all $\omega \notin \Omega_0$.

Proposition 3.18. *Let X be a submartingale and define X^+ as above. Then $\Gamma = \{t \geq 0; P\{X_t^+ \neq X_{t-}^+\} > 0\}$ is countable. $P\{X_t = X_t^+\} = 1$ for $t \notin \Gamma$ and*

$$\tilde{X}_t(\omega) = \begin{cases} X_t(\omega) & \text{if } t \in \Gamma \\ X_t^+(\omega) & \text{if } t \notin \Gamma \end{cases}$$

defines a modification of X almost all of whose sample paths have right and left limits at all $t \geq 0$ and are right continuous at all $t \notin \Gamma$.

Proof. Consider the space $L^\infty(\Omega)$ of real-valued random variables on Ω and define the metric

$$\rho(\xi_1, \xi_2) = E[\min\{|\xi_1 - \xi_2|, 1\}].$$

We can consider the stochastic process X as a map

$$[0, \infty) \to (L^\infty(\Omega), \rho)$$

Now X^+ has right and left hand limits in this metric and so by the lemma Γ is countable.

Choose a real number a, then $\phi_a(x) = \max\{x, a\}$ is an increasing convex function and so $\phi_a(X_t)$ is a submartingale. Consequently,

$$a \leq \phi_a(X_s) \leq E[\phi_a(X_t)|\mathcal{F}_s^X], \quad 0 \leq s \leq t.$$

Because $\{E[\phi_a(X_t)|\mathcal{F}_s^X]; 0 \leq s \leq t\}$ is uniformly integrable, and

$$|\phi_a(X_s)| \leq |a| + |E[\phi_a(X_t)|\mathcal{F}_s^X]|$$

we see that $\{\phi_a(X_s); 0 \leq s \leq t\}$ is also uniformly integrable. Therefore, almost sure convergence implies convergence in L^1 and

$$a \leq \phi_a(X_s) \leq \lim_{u \to s+, u \in D} E[\phi_a(X_u)|\mathcal{F}_s^X] = E[\phi_a(X_s^+)|\mathcal{F}_s^X].$$

For $s \notin \Gamma$, $X_s^+ = X_{s-}^+$ a.s and

$$E[E[\phi_a(X_s^+)|\mathcal{F}_s^X] - \phi_a(X_s)] \leq \lim_{u \to s-, u \in D} E[\phi_a(X_s^+) - \phi_a(X_u)] = 0.$$

Thus, the non-negative random variable $E[\phi_a(X_s^+)|\mathcal{F}_s^X] - \phi_a(X_s)$ has zero expectation and is consequently 0 almost surely.

Using again that $X_s^+ = X_{s-}^+$ a.s. and that X_{s-}^+ is $\mathcal{F}_s^X$-measurable,

$$\phi_a(X_s) = E[\phi_a(X_s^+)|\mathcal{F}_s^X] = \phi_a(X_s^+) \quad \text{almost surely.}$$

Thus, for almost all ω, if $X_s(\omega) > a$, then $X_s(\omega) = X_s^+(\omega)$. Because this holds for all $a \in \mathbb{R}$,

$$P\{X_s = X_s^+\} = 1 \quad \text{for all } s \notin \Gamma$$

i.e., X is right continuous for $s \notin \Gamma$ and $\tilde{X}$ is a modification of X

To see that $\tilde{X}$ has both left and right limits for all $t \geq 0$, replace D with $D \cup \Gamma$ in the construction of Ω_0 and call this new set $\Omega_\Gamma \subset \Omega_0$. For $\omega \in \Omega_\Gamma$, and $t \geq 0$,

$$X_t^+(\omega) = \lim_{s \to t+} X_s^+(\omega) = \lim_{s \to t+, s \in D \cup \Gamma} X_s(\omega).$$

Thus,

$$X_t^+(\omega) = \lim_{s \to t+} \tilde{X}_s(\omega)$$

and $\tilde{X}$ has right limits for every t.

Repeat the procedure with X^- to show the existence of left limits. $\square$

Remark 3.19. *In most of the situations that we will encounter, $\Gamma = \emptyset$ and if the process is defined, as in the case of Lévy processes, by its finite dimensional distributions, then we can take the version with sample paths in $D_S[0,\infty)$.*

Exercise 3.20. *The convergence under the metric ρ is equivalent to convergence in probability.*

Corollary 3.21. *Let $Z \in L^1$, then for any filtration $\{\mathcal{F}_t; t \geq 0\}$, and for all $t \geq 0$,*

$$E[Z|\mathcal{F}_s] \to^{L^1} E[Z|\mathcal{F}_{t+}] \quad \text{as } s \to t+.$$

Proof. Let $X_t = E[Z|\mathcal{F}_{t+}]$. Then X is a martingale. By the proposition above, X has right limits a.s. at each $t \geq 0$. X is uniformly integrable. Consequenty,

$$X_s \to X_{t+} \text{ a.s. and in } L^1 \text{ as } s \to t+.$$

Note that X_{t+} is $\mathcal{F}_{t+}$-measurable. Take $A \in \mathcal{F}_{t+}$, then for $s > t$, $\mathcal{F}_{t+} \subset \mathcal{F}_s$ and $E[X_s; A] = E[Z; A]$. Consequently,

$$E[X_{t+}; A] = \lim_{s \to t+} E[X_s; A] = E[Z; A].$$

Thus, $X_{t+} = E[Z|\mathcal{F}_{t_+}]$. □

Theorem 3.22 (Doob's regularity theorem)**.** *If X is a $\mathcal{F}_t$-submartingale, then the process X^+ is also a $\mathcal{F}_t$-submartingale. Moreover, X^+ is a modification of X if and only if the map*

$$t \to X_t^+$$

is right continuous from $[0,\infty)$ to $L^1(\Omega)$, that is for every $t \geq 0$,

$$\lim_{s \to t+} E[|X_s - X_t|] = 0.$$

Proof. □

Theorem 3.23 (Optional sampling formula)**.** *Let X be a right continuous $\mathcal{F}_t$-submartingale and let τ and σ be $\mathcal{F}_t$-stopping tmes. Then for each $t > 0$,*

$$E[X_{\min\{\tau,t\}}|\mathcal{F}_\sigma] \geq X_{\min\{\sigma,\tau,t\}}.$$

If, in addition, τ satisfies the sampling integrability conditions for X, then

$$E[X_\tau|\mathcal{F}_\sigma] \geq X_{\min\{\sigma,\tau\}}.$$

Proof. We have the theorem in the case that τ and σ are discrete stopping times. Now, let $\{\sigma_n, \tau_n; n \geq 1\}$ be a sequence of nonincreasing discrete stopping times that, converge, respectively to τ and σ.

As before for a, set $\phi_a(x) = \max\{x, a\}$ and note that$\phi_a(X_t)$ is a submartingale. Then,

$$E[\phi_a(X_{\min\{\tau_n,t\}})|\mathcal{F}_{\sigma_n}] \geq \phi_a(X_{\min\{\sigma_n,\tau_n,t\}}.$$

Use the fact that $\mathcal{F}_\sigma \subset \mathcal{F}_{\sigma_n}$ and the tower property to conclude that

$$E[\phi_a(X_{\min\{\tau_n,t\}})|\mathcal{F}_\sigma] \geq E[\phi_a(X_{\min\{\sigma_n,\tau_n,t\}}|\mathcal{F}_\sigma].$$

Note that

$$a \le \phi_a(X_{\min\{\tau_n,t\}}) \le E[\phi_a(X_t)|\mathcal{F}_{\tau_n}] \quad \text{and} \quad a \le \phi_a(X_{\min\{\sigma_n,\tau_n,t\}}) \le E[\phi_a(X_t)|\mathcal{F}_{\min\{\sigma_n,\tau_n\}}].$$

Consequently, as before, $\{\phi_a(X_{\min\{\tau_n,t\}}); n \ge 1\}$ and $\{\phi_a(X_{\min\{\sigma_n,\tau_n,t\}})\}$ are uniformly integrable. Now let $n \to \infty$ and use the right continuity of X to obtain

$$E[\phi_a(X_{\min\{\tau,t\}})|\mathcal{F}_\sigma] \ge E[\phi_a(X_{\min\{\sigma,\tau,t\}}|\mathcal{F}_\sigma]. = \phi_a(X_{\min\{\sigma,\tau,t\}})$$

Now let, $a \to -\infty$. The second part follows are in the discrete time case. □

We have similar theorems to the discrete time case on convergence as $t \to \infty$.

Theorem 3.24. *(Submartingale convergence theorem) Let X be a right continuous submartingale. If*

$$\sup_t EX_t^+ < \infty,$$

then

$$\lim_{t\to\infty} X_t = X_\infty$$

exists almost surely with $E|X_\infty| < \infty$.

Theorem 3.25. *Let X be a right continuous $\mathcal{F}_t$-submartingale. Then X is uniformly integrable if and only if there exists a random variable X_∞ such that*

$$X_t \to^{L^1} X_\infty.$$

Furthermore, when this holds

$$X_t \to X_\infty \quad a.s.$$

Corollary 3.26. *A nonnegative supermartingale converges almost surely.*

Proposition 3.27. *Let X be a right continuous non-negative $\mathcal{F}_t$-;gale. Let $\tau_c(0)$ be the first contact time with 0. Then, with probability 1, $X_t = 0$ for all $t \ge \tau_c(0)$.*

Proof. For $n = 1, 2, \ldots$, let $\tau_n = \tau_e([0, n^{-1}))$, the first entrance time into $[0, n^{-1})$. Then τ_n is an $\mathcal{F}_{t+}$-stopping time and $\tau_c(0) = \lim_{n\to\infty} \tau_n$. If $\tau_n < \infty$, $X_{\tau_n} \le n^{-1}$. Consequently, for every $t \ge 0$,

$$E[X_t|\mathcal{F}_{\tau_n+}] \le X_{\min\{t,\tau_n\}}$$

and hence

$$E[X_t|\mathcal{F}_{\tau_n+}]I_{\{\tau_n \le t\}}, \le \frac{1}{n},$$

and, taking $n \to \infty$, we have upon taking expected values that

$$E[X_t I_{\{\tau_c(0) \le t\}}] \le 0.$$

Now, use the non-negativity and right continuity of X. □

3.2 Sample Path Regularity and Lévy Processes

From Doob's regularilty theorem, we have that any Lévy process with finite mean has a version in $D_{\mathbb{R}^d}[0,\infty)$. For Brownian motion, we have the following.

Proposition 3.28. *Brownian motion has a version in $C_{\mathbb{R}^d}[0,\infty)$.*

Proof. $B_t = C\tilde{B}_t + \mu t$ where C is a $d \times d$ matrix, μ is a d-dimensional vector and $\tilde{B}$ is a vector consisting of d independent standard Brownian motions. Thus, if standard Brownian motion has a continuous version, so does B Thus, let $\tilde{B}$ be standard Brownian motion on $\mathbb{R}$. Then $\tilde{B}_t$ is a normal random variable, mean zero, variance t. Thus, $E\tilde{B}_t^4 = 3t^2$ or

$$E[|\tilde{B}_t - \tilde{B}_s]^4] = 3(t-s)^2.$$

Because B has a version in $D_{\mathbb{R}^d}[0,\infty)$, we can apply the moment theorem on continuous versions with $C = 3$, $\beta = 4$ and $\alpha = 2$. □

Example 3.29. *Let B be a continuous version of standard Brownian motion. Recall that for this process $P\{\sup_t B_t = \infty, \inf_t B_t = -\infty\} = 1$. In particular, because B is continuous, for each $a \in \mathbb{R}$, the stopping time $\tau_a = \inf\{t \geq 0; B_t = a\}$ is finite with probability 1.*

1. *Choose $a, b > 0$ and define $\tau = \inf\{t \geq 0; B_t \notin (-a,b)\} = \min\{\tau_{-a}, \tau_b\}$. Because each point in $\mathbb{R}$ is recurrent a.s., τ is almost surely finite. Thus, by the optional sampling theorem,*

$$0 = EB_{\min\{\tau,n\}} = bP\{B_\tau = b, \tau \leq n\} - aP\{B_\tau = a, \tau \leq n\} + E[B_n; \{\tau > n\}].$$

Now, let $n \to \infty$. For the third term, use the bounded convergence theorem to obtain

$$0 = bP\{B_\tau = b\} - aP\{B_\tau = a\}$$

or

$$E\tau = P\{B_\tau = b\} = \frac{a}{b+a}.$$

Now, use the martingale $B_t^2 - t$ to obtain

$$0 = EB^2_{\min\{\tau,n\}} - E\min\{\tau, n\}.$$

For the first term use the bounded convergence theorem and for the second use the monotone convergence theorem as $n \to \infty$ to see that

$$EB_\tau^2 = b^2\frac{a}{b+a} + a^2\frac{b}{b+a} = ab.$$

For the second term use the monotone convergence theorem as $n \to \infty$ to see that

$$E\tau = ab.$$

2. *Set $X_t = B_t + \mu t$. For $x > 0$, define the stopping times*

$$\tau_x = \inf\{t \geq 0; X_t = x\}$$

Now

$$\exp(uX_t - \alpha t) = \exp(uB_t - (\alpha - u\mu)t)$$

is a martingale provided that

$$\alpha - u\mu = \frac{1}{2}u^2,$$

that is,

$$u_\pm = -\mu \pm \sqrt{\mu^2 + 2\alpha}.$$

Note that if $\alpha > 0$, $u_- < 0 < u_+$. *Thus the martingale* $\exp(u_+X_t - \alpha t)$ *is bounded on* $[0, \tau_x]$. *The optional sampling theorem applies and*

$$1 = E[\exp(u_+X_{\tau_x} - \alpha\tau_x)] = e^{u_+x}E[e^{-\alpha\tau_x}].$$

Consequently,

$$E[e^{-\alpha\tau_x}] = \exp(-x(\sqrt{\mu^2 + 2\alpha} - \mu)).$$

Take $\alpha \to 0$, *then* $\exp(-\alpha\tau_n) \to I_{\{\tau_n < \infty\}}$. *Therefore,*

$$P\{\tau_x < \infty\} = \begin{cases} 1 & \mu \geq 0 \\ e^{2\mu x} & \mu < 0 \end{cases}$$

In addition, the Laplace transform can be inverted to see that τ_x *has density*

$$f_{\tau_x}(t) = \frac{x}{\sqrt{2\pi t^3}}\exp(-\frac{(x-\mu t)^2}{2t}).$$

Exercise 3.30. *Let* B *be standard Brownian motion,* $a > 0$, *and* $\tau = \inf\{t \geq 0; |B_t| = a\}$.

1. $E[\exp(-\alpha\tau)] = \text{sech}(a\sqrt{2\alpha})$.
2. $E\tau^2 = 5a^4/3$. Hint: *Prove that* $B_t^4 - 6tB_t^2 + 3t^2$ *is a martingale.*

Proposition 3.31. *Let* X *be a right continuous Lévy process. Then for each* $s \geq 0$, *the process* $\tilde{X}_t = X_{s+t} - X_s$ *is a Lévy process with the same finite dimensional distribution as* X *that is independent of* $\mathcal{F}^X_{s+}$.

Proof. For each $\epsilon > 0$, $\mathcal{F}^X_{s+} \subset \mathcal{F}^X_{s+\epsilon}$, $X_{s+t+\epsilon} - X_{s+\epsilon}$ is a Lévy process independent of $\mathcal{F}^X_{s+}$. Let $A \in \mathcal{F}^X_{s+}$ Choose a right continuous modification of X, times $0 = t_0 < t_1 < \cdots < t_n$, and bounded continuous functions f_j.

$$\begin{aligned}
E[\prod_{j=1}^n f_j(X_{s+t_j} - X_{s+t_{j-1}}); A] &= \lim_{\epsilon\to 0+} E[\prod_{j=1}^n f(X_{s+t_j+\epsilon} - X_{s+t_{j-1}+\epsilon}); A] \\
&= \lim_{\epsilon\to 0+} E[\prod_{j=1}^n f(X_{s+t_j+\epsilon} - X_{s+t_{j-1}+\epsilon})]P(A) \\
&= \lim_{\epsilon\to 0+} E[\prod_{j=1}^n f(X_{t_j} - X_{t_{j-1}})]P(A) \\
&= E[\prod_{j=1}^n f(X_{t_j} - X_{t_{j-1}})]P(A)
\end{aligned}$$

□

Proposition 3.32. *Let X be a right continuous Lévy process and τ an almost surely finite stopping time. Then the process $\tilde{X}_s = X_{s+\tau} - X_\tau$ is a Lévy process with respect to the filtration $\{\mathcal{F}^X_{\tau+s}; s \geq 0\}$ having the same finite dimensional distributions as $X_s - X_0$ and is independent of $\mathcal{F}^X_{\tau+}$.*

Proof. Consider times $0 = t_0 < t_1 < \cdots < t_n$, bounded continuous functions $f_1, f_2, \ldots, f_n$, and $A \in \mathcal{F}^X_{\tau+}$. For the case that τ is a discrete stopping time with range $\{s_k; k \geq 1\}$, we have

$$\begin{aligned}
E[\prod_{j=1}^n f_j(X_{\tau+t_j} - X_{\tau+t_{j-1}}); A] &= \sum_{k=1}^\infty E[\prod_{j=1}^n f_j(X_{\tau+t_j} - X_{\tau+t_{j-1}}); A \cap \{\tau = s_k\}] \\
&= \sum_{k=1}^\infty E[\prod_{j=1}^n f_j(X_{s_k+t_j} - X_{s_k+t_{j-1}}); A \cap \{\tau = s_k\}] \\
&= \sum_{k=1}^\infty E[\prod_{j=1}^n f_j(X_{t_j} - X_{t_{j-1}})] P(A \cap \{\tau = s_k\}) \\
&= E[\prod_{j=1}^n f_j(X_{t_j} - X_{t_{j-1}})] P(A)
\end{aligned}$$

Note that $A \cap \{\tau = s_k\} \in \mathcal{F}_{s_k+}$ and thus the third line follows from the second by the previous proposition.

For the general case, let $\{\tau_m; m \geq 1\}$ be a decreasing sequence of stopping times with limit τ. Pick $A \in \mathcal{F}_\tau \subset \mathcal{F}_{\tau_m}$. Now, apply the identity above to the τ_m, use right continuity and the bounded convergence theorem to obtain the result. $\square$

Corollary 3.33 (Blumenthal 0-1 law)**.** *For a Lévy process X, with $X_0 = 0$ every event in $\mathcal{F}^X_{0+}$ has probability 0 or 1.*

Proof. Fix $A \in \mathcal{F}^X_{0+}$, then, in addition, $A \in \sigma\{X_t; t \geq 0\} = \sigma\{X_t - X_0; t \geq 0\}$. Then $\mathcal{F}^X_{0+}$ and $\sigma\{X_t - X_0; t \geq 0\}$ are independent σ-algebras and thus $P(A) = P(A \cap A) = P(A)P(A)$ and so $P(A) = 0$ or 1. $\square$

In particular, if τ is an $\mathcal{F}^X_{t+}$-stopping time then

$$P\{\tau = 0\} = 0 \text{ or } 1.$$

Example 3.34.

1. *Suppose that $f(t) > 0$ for all $t > 0$. Then*

$$\limsup_{t \to 0+} \frac{B_t}{f(t)} = c \text{ a.s.}$$

for some $c \in [0, \infty]$. If $f(t) = \sqrt{t}$, the $B_t/f(t)$ is a standard normal and so the limit above cannot be finite. Consequently, $c = \infty$. If $f(t) = t^\alpha, \alpha > 1/2$, then $B_t/f(t)$ is normal, mean 0, variance $t^{1-2\alpha}$ and so the limit above cannot be positive. Consequently, $c = 0$.

2. *Let $\tau = \inf\{t > 0; B_t > 0\}$. Then $P\{\tau \leq t\} \geq P\{B_t > 0\} = 1/2$. Thus,*

$$P\{\tau = 0\} = \lim_{t \to 0+} P\{\tau \leq t\} \neq 0,$$

thus it must be 1.

Theorem 3.35 (Reflection principle)**.** *Let $a \in \mathbb{R}$ and let B be a continuous standard Brownian motion. Define the stopping time $\tau_a = \inf\{t > 0; B_t > a\}$. The process*

$$\tilde{B}_t = \begin{cases} B_t & \text{if } t < \tau_a, \\ 2a - B_t & \text{if } t \geq \tau_a. \end{cases}$$

is a continuous standard Brownian motion.

Proof. $\tau_a < \infty$ with probability 1. Therefore, $X_t = B_{\tau_a+t} - a$ and $-X$ are standard Brownian motions independent of B^{τ_a}. Consequently (B^{τ_a}, X) and $(B^{\tau_a}, -X)$ have the same distribution. Define the continuous map

$$\phi : C_{\mathbb{R}}[0,\infty) \times C_{\mathbb{R}}[0,\infty) \to C_{\mathbb{R}}[0,\infty)$$

by

$$\phi(b,x)_t = b_t I_{\{\tau_a \leq t\}} + (x_{\tau_a+t} - a) I_{\{\tau_a > t\}}.$$

Then $\phi(B^{\tau_a}, X) = B$ and $\phi(B^{\tau_a}, -X) = \tilde{B}$. □

Exercise 3.36. *State and prove a similar theorem for symmetric Lévy processes and general finite stopping times.*

Corollary 3.37. *Define*

$$B_t^* = \sup\{B_s; s \leq t\}.$$

Then, for all $a, x \geq 0$ and $t \geq 0$,

$$P\{B_t^* \geq a, B_t \leq a - x\} = P\{B_t \geq a + x\}.$$

Proof.

$$P\{B_t^* \geq a, B_t \leq a-x\} = P\{B_t^* \geq a, \tilde{B}_t \leq a-x\} = P\{B_t^* \geq a, B_t \geq a+x\} = P\{B_t \geq a+x\}.$$

□

Use the corollary in the case $a > 0$ and $x = 0$ to obtain.

$$\begin{aligned} P\{\tau_a \leq t\} &= P\{B_t^* \geq a\} = P\{B_t^* \geq a, B_t \leq a\} + P\{B_t^* \geq a, B_t \geq a\} \\ &= 2P\{B_t \geq a\} = \frac{1}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/2t}\, dx \end{aligned}$$

Exercise 3.38. 1. *Show that τ_a has density*

$$f_{\tau_n}(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-a^2/2t}.$$

2. *Show that (B_t^*, τ_a) has joint density*

$$f_{(B_t^*,\tau_a)}(x,t) = \frac{2(2a-x)}{\sqrt{2\pi t^3}} \exp(-\frac{(2a-x)^2}{2t}), \quad a \geq 0, x \leq a.$$

3. *Show that $\{\tau_a; a \geq 0\}$ is a Lévy process that does not have a finite stopping time.*

3.3 Maximal Inequalities

Proposition 3.39 (Doob's submartingale maximal inequality)**.** *Let X be a right continuous submartingale.*

1. *Then for each $x > 0$ and $t > 0$,*
$$P\{\sup_{0\le s\le t} X_s \ge x\} \le \frac{1}{x}EX_t^+,$$
and
$$P\{\inf_{0\le s\le t} X_s \le -x\} \le \frac{1}{x}(EX_t^+ - EX_0).$$

2. *If X is non-negative. Then for $\alpha > 1$ an d $t > 0$,*
$$E[\sup_{0\le s\le t} X_s^\alpha] \le \left(\frac{\alpha}{\alpha-1}\right)^\alpha EX_t^\alpha.$$

Proof. The first statement follows form the corresponding result for a countable collection of times in $[0,t]$ and right continuity.

For the second statement. set $\tau = \inf\{t \ge 0; X_s > x\}$. Then τ is an $\mathcal{F}_{t+}$-stopping time. The right continuity of X implies that $X_\tau \ge x$ whenever $\tau < \infty$. Consequently,

$$\{\sup_{0\le s\le t} X_t > x\} \subset \{\tau \le t\} \subset \{\sup_{0\le s\le t} X_t \ge x\}.$$

Moreover, these events have equal probability for all but countably many $x > 0$ and hence

$$xP\{\tau \le t\} \le E[X_\tau; \{\tau \le t\}] \le E[X_t; \{\tau \le t\}].$$

By the optional sampling theorem,
$$EX_{\min\{\tau,t\}} \le EX_t.$$

Now the balance of the proof follows the same line as in the discrete time case. □

Proposition 3.40. *Let B be a continuous version of standard Brownian motion and define*

$$\tilde{B}_t = \begin{cases} tB_{1/t} & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases}$$

Then $\tilde{B}$ is also a continuous version of standard Brownian motion.

Proof. An easily calculation shows that $\tilde{B}$ has stationary independent increments, that $\tilde{B}_{t+s} - \tilde{B}_t$ is normally distributed with mean 0 and variance s and that $\tilde{B}$ is continouus on $(0,\infty)$.

To establish continuity at 0, first note that is equivalent to showing that

$$\lim_{t\to\infty} \frac{B_t}{t} = 0.$$

This subsequential limit along the integers is 0 by the strong law of large numbers. The following claim proves the theorem.

Claim.
$$\lim_{n\to\infty} \frac{\max_{n\le t\le n+1} |B_t - B_n|}{n} = 0 \text{ a.s.}$$

Let $\epsilon > 0$. Then
$$P\{\max_{n\le t\le n+1} |B_t - B_n| > n\epsilon\} = P\{\max_{0\le t\le 1} |B_t| > n\epsilon\} \le \frac{1}{(n\epsilon)^2} EB_1^2.$$

Thus,
$$\sum_{n=0}^{\infty} P\{\max_{n\le t\le n+1} |B_t - B_n| > n\epsilon\} = \sum_{n=0}^{\infty} \frac{1}{(n\epsilon)^2} = \frac{\pi^2}{6\epsilon} < \infty.$$

Thus, by the first Borel-Cantelli lemma,
$$P\{\max_{n\le t\le n+1} \frac{|B_t - B_n|}{n} > \epsilon \text{ i.o.}\} = 0$$

and the claim holds. □

We will need the following integration result for the next theorem.

Exercise 3.41. *Let Φ be the cumulative distribution function for the standard normal. Then, for $a > 0$*
$$\Phi(-a) = 1 - \Phi(a) \ge \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}a^2)a^{-1}(1 - a^{-2}).$$

Hint. *Use* $\exp(-z^2/2) > (1 - 3z^{-4})\exp(-z^2/2)$ *and find the antiderivative.*

3.4 Localization

To extend our results to a broader class of processes, we introduce the concept of localization.

Definition 3.42. *Let Y be a $\mathcal{F}_t$-adapted $D_{\mathbb{R}}[0,\infty)$-process,*

1. *A stopping time τ reduces Y if*
$$Y^\tau I_{\{\tau>0\}}$$
is uniformly integrable.

2. *Y is called a* local martingale *if there exists a increasing sequence of stopping times $\{\tau_n; n \ge 1\}$, $\lim_{n\to\infty} \tau_n = \infty$ so that τ_n reduces Y and Y^{τ_n} is a martingale. We shall call the sequence above a* reducing sequence *for Y.*

Exercise 3.43.
1. *If τ reduces Y and $\sigma \le \tau$, then σ reduces Y.*
2. *Local martingales form a vector space of processes.*
3. *Take $\tau_n = n$ to see that martingales are local martingales.*
4. *Let $\sigma_n = \inf\{t \ge 0 : |Y_t| \ge n\}$, then $\min\{\sigma_n, \tau_n\}$ is a reducing sequence for Y.*

This next proposition gives us our first glimpse into the stochastic calculus.

Proposition 3.44. *Suppose that Y is a right continuous $\mathcal{F}_t$-local martingale, and that V is a real-valued $\mathcal{F}_t$-adapted, bounded continuous process having locally bounded variation. Then*

$$M_t = V_t Y_t - \int_0^t Y_s \; dV_s$$

is an $\mathcal{F}_t$-local martingale.

Remark 3.45. *This uses the fact that we will later learn that integrals of martingales are martingales. Then, M can be found from an integration by parts formula*

$$\int_0^t V_s \; dV_s = V_t Y_t - V_0 Y_0 - \int_0^t Y_u \; dV_u.$$

Exercise 3.46 (summation by parts)**.** *Let $\{a_k; k \geq 0\}$ and $\{b_k; k \geq 0\}$ be two real valued sequences, then*

$$\sum_{k=m}^{n-1} a_k(b_{k+1} - b_k) = a_n b_n - a_m b_m - \sum_{k=m}^{n-1} b_{k+1}(a_{k+1} - a_k).$$

Proof. (of the proposition) Let $\{\tau_n; n \geq 1\}$ be a localizing sequence for Yso that τ_n is less than the contact time with the set $(-n, n)^c$.

Write $|V|_t$ for the variation of V on the interval $[0, t]$. Because V is continuous, σ_n, the first contact time of $|V|$ into $[n, \infty)$ is a stopping time. Because $|V|$ is continuous, $\sigma_n \to \infty$ as $n \to \infty$.

Write $\gamma_n = \min\{\sigma_n, \tau_n\}$, we show that M^{γ_n} is a martingale, i. e,

$$E\left[V_{t+s}^{\gamma_n} Y_{t+s}^{\gamma_n} - Y_s^{\gamma_n} V_s^{\gamma_n} - \int_{\min\{t,\gamma_n\}}^{\min\{t+s,\gamma_n\}} Y_u^{\gamma_n} \; dV_u^{\gamma_n} | \mathcal{F}_t\right] = 0. \tag{3.1}$$

Let $t = u_0 < u_1 < \cdots < u_m = t + s$ be a partition of the interval $[t, t + s]$. Then, by the tower property, and the fact that V is adapted and thatY_{γ_n} is a martingale, we see that

$$E[\sum_{k=0}^{m-1} V_{u_k}^{\gamma_n}(Y_{u_{k+1}}^{\gamma_n} - Y_{u_k}^{\gamma_n})|\mathcal{F}_t] = \sum_{k=0}^{m-1} E[V_{u_k}^{\gamma_n} E[(Y_{u_{k+1}}^{\gamma_n} - Y_{u_k}^{\gamma_n})|\mathcal{F}_{u_k}|\mathcal{F}_t] = 0.$$

Use the summation by parts formula to obtain that

$$E[V_{t+s}^{\gamma_n} Y_{t+s}^{\gamma_n} - Y_s^{\gamma_n} V_s^{\gamma_n} - \sum_{k=0}^{m-1} Y_{u_{k+1}}^{\gamma_n}(V_{u_{k+1}}^{\gamma_n} - V_{u_k}^{\gamma_n})|\mathcal{F}_t] = 0. \tag{3.2}$$

The sum in the conditional expectation converges almost surely to the Riemann-Stieltjes integral

$$\int_t^{t+s} Y_u^{\gamma_n} \; dV_u^{\gamma_n}.$$

To complete the proof, we must show that the process is uniformly integrable. However,

$$\left|\sum_{k=0}^{n-1} Y_{u_k}^{\gamma_n}(V_{u_{k+1}}^{\gamma_n} - V_{u_k}^{\gamma_n})\right| \leq \sum_{k=0}^{m-1} |Y_{u_k}^{\gamma_n}||V_{u_{k+1}}^{\gamma_n} - V_{u_k}^{\gamma_n}| \leq \max\{n, Y_{t+s}^{\gamma_n}\}(|V_{\gamma_n}|_{t+s} - |V^{\gamma_n}|_t) \leq \max\{n, Y_{t+s}^{\gamma_n}\}n$$

which is integrable.

Thus, the Riemann sums converge in L^1 and we have both that (3.1) holds and that M_{γ_n} is uniformly integrable. $\square$

3.5 Law of the Iterated Logarithm

Theorem 3.47. *(law of the iterated logarithm for Brownian motion) Let B be a standard Brownian motion. Then,*

$$P\left\{\limsup_{t\to 0+} \frac{B_t}{\sqrt{2t\log\log(1/t)}} = 1\right\} = 1.$$

Proof. (McKean) Write $h(t) = \sqrt{2t\log\log(1/t)}$.

Part 1.

$$P\left\{\limsup_{t\to 0+} \frac{B_t}{h(t)} \leq 1\right\} = 1.$$

Consider the exponential martingale $Z_t(\alpha) = \exp(\alpha B_t - \frac{1}{2}\alpha^2 t)$. Then by Doob's submartingale maximal inequality,

$$P\{\sup_{0\leq s\leq t}(B_s - \frac{1}{2}\alpha s) > \beta\} = P\{\sup_{0\leq s\leq t} Z_s(\alpha) > e^{\alpha\beta}\} \leq e^{-\alpha\beta}EZ_t(\alpha) = e^{-\alpha\beta}.$$

Fix the θ and δ in $(0,1)$ and apply the inequality above to the sequences

$$t_n = \theta^n,\ \alpha_n = \theta^{-n}(1+\delta)h(\theta^n),\ \beta_n = \frac{1}{2}h(\theta^n).$$

Then,

$$\alpha_n\beta_n = \frac{1}{2}\theta^{-n}(1+\delta)h(\theta^n)^2 = (1+\delta)\log\log\theta^{-n} = (1+\delta)(\log n + \log\log\frac{1}{\theta}).$$

Consequently, for $\gamma = (\log(1/\theta))^{(1+\delta)}$,

$$P\{\sup_{0\leq s\leq t_n}(B_s - \frac{1}{2}\alpha_n s) > \beta_n\} \leq \gamma n^{-(1+\delta)}.$$

Because $\delta > 0$, these probabilities have a finite sum over n. Thus, by the first Borel-Cantelli lemma

$$P\{\sup_{0\leq s\leq t_n}(B_s - \frac{1}{2}\alpha_n s) > \beta_n \text{ i.o.}\} = 0.$$

Thus, we have, for ω in an event having probability 1, there exists $N(\omega)$ so that

$$\sup_{0\leq s\leq t_n}(B_s - \frac{1}{2}\alpha_n s) > \beta_n$$

for all $n \geq N(\omega)$. For such n, consider t in the interval $(\theta^{n+1}, \theta^n]$. Then

$$B_t(\omega) \leq \sup_{0 \leq s \leq \theta^n} B_s(\omega) \leq \beta_n + \frac{1}{2}\alpha_n \theta^n = \frac{1}{2}(2+\delta)h(\theta^n) \leq \frac{1}{2}\theta^{-1/2}(2+\delta)h(t).$$

Consequently,

$$\limsup_{t \to 0+} \frac{B_t(\omega)}{h(t)} \leq \frac{1}{2}\theta^{-1/2}(2+\delta).$$

Now, use the fact $\inf_{\theta,\delta \in (0,1)} \frac{1}{2}\theta^{-1/2}(2+\delta) = 1$ to obtain

$$\limsup_{t \to 0+} \frac{B_t(\omega)}{h(t)} \leq 1.$$

Part 2.

$$P\left\{\limsup_{t \to 0+} \frac{B_t}{h(t)} \geq 1\right\} = 1.$$

Choose $\theta \in (0,1)$ and define the *independent* events

$$A_n = \{B_{\theta^n} - B_{\theta^{n+1}} > h(\theta^n)\sqrt{1-\theta}\}.$$

Noting that $B_{\theta^n} - B_{\theta^{n+1}}$ is a normal random variable with mean 0 and variance $\theta^n(1-\theta)$, we have

$$P(A_n) = 1 - \Phi(\theta^{-n/2}h(\theta^n)) \geq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\theta^{-n}h(\theta^n)^2\right) \theta^{n/2}h(\theta^n)^{-1}(1 - \theta^n h(\theta^n)^{-2}).$$

Because

$$\frac{1}{2}\theta^{-n}h(\theta^n)^2 = \log n + \log\log(\theta^{-1}),$$

we have that

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

by comparison to $\sum_{n=1}^{\infty}(n\sqrt{\log n})^{-1}$. Now, use the second Borel-Cantelli lemma to conclude that for infinitely many n

$$B_{\theta^n} - B_{\theta^{n+1}} > h(\theta^n)\sqrt{1-\theta}\}.$$

Use the first part of the theorem, applied to $-B$ to obtain that

$$B_{\theta^{n+1}} \geq -2h(\theta^{n+1}) > -4\sqrt{\theta}h(\theta^n)$$

for all sufficiently large $n > 1/\log\theta$,. Thus, for infinitely many n,

$$B_{\theta^n} > \left(\sqrt{1-\theta} - 4\sqrt{\theta}\right) h(\theta^n).$$

This proves that, with probability 1,

$$\limsup_{t \to 0} \frac{B_t}{h(t)} \geq \sqrt{1-\theta} - 4\sqrt{\theta}$$

for all $\theta \in (0,1)$. Now, part 2 follows by noting that $\sup_{\theta \in (0,1)} \sqrt{1-\theta} - 4\sqrt{\theta} = 1$. $\square$

Corollary 3.48. *Let B be a standard Brownian motion. Then,*

$$P\left\{\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 1\right\} = 1.$$

Proof. Use the fact that $tB_{1/t}$ is a Brownian motion. $\square$

4 Markov Processes

4.1 Definitions and Transition Functions

Definition 4.1. *A stochastic process X with values in a metric space S is called a* Markov process *provided that*

$$P\{X_{s+t} \in B | \mathcal{F}_t^X\} = P\{X_{s+t} \in B | X_t\} \tag{4.1}$$

for every $s, t \geq 0$ and $B \in \mathcal{B}(S)$.

If $\{\mathcal{F}_t; t \geq 0\}$ is a filtration and $\mathcal{F}_t^X \subset \mathcal{F}_t$, then we call X a Markov process with respect to the filtration *$\{\mathcal{F}_t; t \geq 0\}$ if the (4.1) holds with $\mathcal{F}_t^X$ replaced by $\mathcal{F}_t$.*

The probability measure α given by $\alpha(B) = P\{X_0 \in B\}$ is called the initial distribution *of X.*

Exercise 4.2.

1. *If X is Markov process with respect to the filtration $\{\mathcal{F}_t; t \geq 0\}$ then it is a Markov process.*

2. *Use the standard machine to show that the definition above is equivalent to*

$$E[f(X_{s+t}) | \mathcal{F}_t^X] = E[f(X_{s+t}) | X_t].$$

for all bounded measurable f.

Definition 4.3. *A function*

$$p : [0, \infty) \times S \times \mathcal{B}(S) \to [0, 1]$$

is a called time homogeneous transition function *if,*

1. *for every $(t, x) \in [0, \infty) \times S$, $p(t, x, \cdot)$ is a probability,*
2. *for every $x \in S$, $p(0, x, \cdot) = \delta_x$,*
3. *for every $B \in \mathcal{B}(S)$, $p(\cdot, \cdot, B)$ is measurable, and*
4. *(Chapman-Kolmogorov equation) for every $s, t \geq 0$, $x \in S$, and $B \in \mathcal{B}(S)$,*

$$p(t + s, x, B) = \int_S p(s, y, B) p(t, x, dy).$$

The transition function P is a transition function for a time-homogeneous Markov process *X, if, for every $s, t \geq 0$ and $B \in \mathcal{B}(S)$,*

$$P\{X_{t+s} \in B | \mathcal{F}_t^X] = p(s, X_t, B).$$

Exercise 4.4. *Show, using the standard machine, the the identity above is equivalent to*

$$E[f(X_{t+s} | \mathcal{F}_t^X] = \int_S f(y) p(s, X_t, dy)$$

for every $s, t \geq 0$ and bounded measurable f.

To see how the Chapman-Kolmgorov equation arises, note that for all $s, t, u \geq 0$ and $B \in \mathcal{B}(S)$,

$$\begin{aligned} p(t+s, X_u, B) = P\{X_{u+t+s} \in B|\mathcal{F}_u^X\} &= E[P\{X_{u+t+s} \in B|\mathcal{F}_{u+t}^X\}|\mathcal{F}_u^X] \\ &= E[p(s, X_{u+t}, B)|\mathcal{F}_u^X] = \int_S p(s, y, B)p(t, X_u, dy). \end{aligned}$$

Exercise 4.5. *Let X be a Lévy process with $\nu_t(B) = P\{X_t \in B\}$. Then the transition function*

$$p(t, x, B) = \nu_t(B - x).$$

In this case, the Chapman-Kolmogorov equations state that $\nu_{s+t} = \nu_s * \nu_t$.

The transition function and the initial distribution determine the finite dimensional distributions of X by

$$P\{X_0 \in B_0, X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\}$$

$$= \int_{B_0} \int_{B_1} \cdots \int_{B_{n-1}} p(t_n - t_{n-1}, x_{n-1}, B_n)p(t_{n-1} - t_{n-2}, x_{n-2}, dx_{n-1}) \cdots p(t_1, x_0, dx_1)\alpha(dx_0).$$

Exercise 4.6. *Show that the consistency condition in the Daniell-Kolmogorov are satisfied via the Chapman-Kolmogorov equation.*

This gives us a unique measure on $(S^{[0,\infty)}, \mathcal{B}(S^{[0,\infty)}))$. Denote probabilities with respect to this measure P_α and P_x for P_{δ_x}.

Rarely are we able to write this transition function explicitly. In two well known cases, Brownian motion and the Poisson process we can.

Example 4.7. *For Brownian motion,*

$$P(t, x, B) = \frac{1}{\sqrt{2\pi t}} \int_B \exp\left(-\frac{(y-x)^2}{2t}\right)\, dt, \qquad t \geq 0, x \in R, B \in (R).$$

For the Poisson process with parameter λ,

$$P(t, x, \{y\}) = \frac{(\lambda t)^{y-x}}{(y-x)!} e^{-\lambda t}, \qquad t \geq 0, y \geq x \geq 0, x, y \in N.$$

The kernels $P(s, \cdot, \cdot)$ are naturally associated with the operator on the bounded measurable functions.

Definition 4.8. *The* transition operator *for the transition function P is defined by*

$$T(s)f(x) = \int_S f(y)p(s, x, dy) = E_x[f(X_s)].$$

Consequently,

$$T(s)f(X_t) = \int_S f(y)p(s, X_t, dy) = E[f(X_{t+s})|\mathcal{F}_t^X].$$

Exercise 4.9. *The family of operators $\{T(t); t \geq 0\}$ satisfies*

1. $T(0)f = f$
2. $T(s+t)f = T(s)T(t)f$

for all bounded measurable f.

4.2 Operator Semigroups

Definition 4.10. *A one-paramter family* $\{T(t); t \geq 0\}$ *of bounded linear operators on a Banach space* $(L, ||\cdot||)$ *is call a* semigroup *if*

1. $T(0)f = f$
2. $T(s+t)f = T(s)T(t)f$

for all $f \in L$. *This semigroup is called* strongly continuous *if*

$$\lim_{t \to 0+} T(t)f = f.$$

for all $f \in L$. *This semigroup is called a* contraction *semigroup if*

$$||T(t)|| = \sup\{||T(t)f||; ||f|| = 1\} \leq 1.$$

Exercise 4.11. *Let* A *be a bounded linear operator and define, for* $t \geq 0$, *the linear operator*

$$\exp tA = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k, \ \exp 0 = I.$$

Then, $\exp tA$ *is a strongly continuous semigroup and* $||\exp tA|| \leq \exp t||A||$.

Proposition 4.12. *Let* T *be a strongly continuous contraction semigroup. Then, for each* $f \in L$, *the mapping*

$$t \to T(t)f$$

is a continuous function from $[0, \infty)$ *into* L.

Proof. For $t \geq 0$ and $h > 0$,

$$||T(t+h)f - T(t)f|| = ||T(t)(T(h)f - f)|| \leq ||T(t)||\ ||T(h)f - f|| \leq ||T(h)f - f||.$$

For $0 \leq h \leq t$,

$$||T(t-h)f - T(t)f|| = ||T(t-h)(T(h)f - f)|| \leq ||T(t-h)||\ ||T(h)f - f|| \leq ||T(h)f - f||.$$

Now, let $h \to 0$. □

4.2.1 The Generator

Definition 4.13. *Let* A *be a linear operator with domain* $\mathcal{D}(A)$, *a subspace of* L *and range* $\mathcal{R}(A)$.

1. *The* graph *of* A *is given by*

$$\mathcal{G}(A) = \{(f, Af) : f \in \mathcal{D}(A)\} \subset L \times L.$$

2. $L \times L$ *is a Banach space under the norm* $||(f,g)|| = ||f|| + ||g||$. A *is a* closed *operator if its graph is a closed subspace of* $L \times L$.

   Note that if $\mathcal{D}(A) = L$, *then* A *is closed.*

3. *Call $\tilde{A}$ an* extension *of A if $\mathcal{D}(A) \subset \mathcal{D}(\tilde{A})$ and $\tilde{A}f = Af$ for all $f \in \mathcal{D}(A)$.*

Definition 4.14. *The* (infinitesimal) generator *of a semigroup T is the linear operator G defined by*

$$Gf = \lim_{t \to 0} \frac{1}{t}(T(t)f - f).$$

The domain of G is the subspace of all $f \in L$ for which the limit above exists.

Exercise 4.15. 1. *Let L be the bounded real-valued functions on $\mathbb{N}$. For the Poisson process with parameter λ, show that the generator is*

$$Gf(x) = \lambda(f(x+1) - f(x))$$

and that $\mathcal{D}(G) = L$.

2. *Assume that A is a bounded operator. The semigroup* $\exp tA$ *has generator A. $\mathcal{D}(A) = L$.*

We will be examining continuous functions $u : [a,b] \to L$. Their Riemann integrals will be defined via Riemann sums and limits. Thus, we will have analogous definitions for Riemann integrable and for the improper Riemann integrals $\int_a^\infty u(s)\,ds$ and an analogous statement for the fundamental theorem

$$\lim_{h \to 0} \frac{1}{h} \int_a^{a+h} u(s)\,ds = u(a).$$

Exercise 4.16. 1. *If $u : [a,b] \to L$ is continuous and $||u||$ is Riemann integrable, then u is integrable and*

$$||\int_a^b u(s)\,ds|| \leq \int_a^b ||u(s)||\,ds.$$

2. *Let B be a closed linear operator on L and assume that $u : [a,b] \to L$ is continuous, that $u(s) \in \mathcal{D}(B)$ for all $s \in [a,b]$, and that both u and Bu are Riemann integrable. Then*

$$\int_a^b u(s)\,ds \in \mathcal{D}(B) \quad \text{and } B\int_a^b u(s)\,ds = \int_a^b Bu(s)\,ds.$$

3. *If u is continuously differentiable, then*

$$\int_a^b \frac{d}{dt}u(t)\,dt = u(b) - u(a).$$

The example of the semigroup $\exp tA$ motivates the following identities.

Proposition 4.17. *Let T be a strongly continuous semigroup on L with generator G.*

1. *If $f \in L$ and $t \geq 0$, then*

$$\int_0^t T(s)f\,ds \in \mathcal{D}(G) \quad \text{and} \quad T(t)f - f = G\int_0^t T(s)f\,ds. \tag{4.2}$$

2. *If $f \in \mathcal{D}(G)$ and $t \geq 0$, then $T(t)f \in \mathcal{D}(G)$ and*

$$\frac{d}{dt}T(t)f = GT(t)f = T(t)Gf. \tag{4.3}$$

3. *If $f \in \mathcal{D}(G)$ and $t \geq 0$, then*

$$T(t)f - f = \int_0^t GT(s)f\ ds = \int_0^t T(s)Gf\ ds. \tag{4.4}$$

Proof. 1. Observe that $(T(h) - I)/h$ is closed for all $h > 0$. Thus,

$$\frac{1}{h}(T(h) - I)\int_0^t T(s)f\ ds = \frac{1}{h}\int_0^t (T(s+h)f - T(s)f)\ ds = \frac{1}{h}\int_t^{t+h} T(s)f\ ds - \frac{1}{h}\int_0^h T(s)f\ ds.$$

Now, let $h \to 0$ to obtain (4.2).

2. Write $G_h = (T(h) - I)/h$. Then, for all $h > 0$

$$\frac{1}{h}(T(t+h)f - T(t)f) = G_hT(t)f = T(t)G_hf$$

Consequently, $T(t)f \in \mathcal{D}(G)$ and

$$\frac{d^+}{dt}T(t)f = GT(t)f = T(t)Gf.$$

To check the left derivative, use the identity

$$\frac{1}{-h}(T(t-h)f - T(t)f) - T(t)Gf = T(t-h)(G_h - G)f + (T(t-h) - T(t))Gf,$$

valid for $h \in (0, t]$.

3. This is a consequence of part 2 in this proposition and part 3 in the lemma above.

□

Exercise 4.18. *Interpret the theorem above for the semigroup*

$$T(s)f(x) = f(x + s)$$

for measurable functions $f : \mathbb{R} \to \mathbb{R}$.

Corollary 4.19. *If G is the generator of a a strongly continuous contraction semigroup T on L, then G is closed and $\mathcal{D}(G)$ is dense in L.*

Proof. Because, for every $f \in L$,

$$\lim_{h\to 0+}\frac{1}{h}\int_0^h T(s)f\ ds = f,$$

$\mathcal{D}(G)$ is dense in L.

To see that G is closed, choose $\{f_n; n \geq 1\}$ so that

$$f_n \to f \quad \text{and} \quad Gf_n \to g \quad \text{as } n \to \infty.$$

Then, by (4.3),

$$T(t)f_n - f_n = \int_0^t T(s)Gf_n \, ds \quad \text{for each } t > 0.$$

Thus, by letting $n \to \infty$, we obtain

$$T(t)f - f = \int_0^t T(s)g \, ds \quad \text{for each } t > 0.$$

Now, divide by t and let $t \to 0$ to conclude that

$$f \in \mathcal{D}(G) \quad \text{and} \quad Gf = g.$$

□

4.2.2 The Resolvent

We now give some necessary conditions for an operator to be the generator of a strongly continuous semigroup. The Hille-Yosida will show that this conditions are sufficient.

Definition 4.20. *For a strongly continuous contraction semigroup T, define the* resolvent,

$$R(\lambda)g = \int_0^\infty e^{-\lambda s} T(s)g \, ds.$$

Exercise 4.21. *If X is a time homogenuous Markov process with semigroup T and τ_λ is an independent exponential random variable with parameter λ, then*

$$\lambda R(\lambda)g = Eg(X_{\tau_\lambda}).$$

Proposition 4.22. *Let T be a strongly continuous contraction semigroup with generator G, then for all $\lambda > 0$, $(\lambda I - G)^{-1}$ is a bounded linear operator on L and*

$$R(\lambda)g = (\lambda I - G)^{-1} g$$

for all $g \in L$.

Proof. Fix $\lambda > 0$ and choose $g \in L$. Then

$$||R(\lambda)g|| = \int_0^\infty e^{-\lambda s} ||T(s)g|| \, ds \leq \frac{1}{\lambda} ||g||.$$

Also, for any $h > 0$,

$$\frac{1}{h}(T(h) - I)R(\lambda)g = \frac{1}{h}\int_0^\infty e^{-\lambda s}(T(s+h)g - T(s)g) \, dt = \frac{e^{\lambda h} - 1}{h} \int_0^\infty e^{-\lambda s} T(s)g \, ds - \frac{e^{\lambda h}}{h} \int_0^h e^{-\lambda s} T(s)g \, ds.$$

Let $h \to 0$ to see that $R(\lambda)g \in \mathcal{D}(G)$ and $GR(\lambda)g = \lambda R(\lambda)g - g$, or

$$(\lambda I - G)R(\lambda)g = g.$$

Thus, $R(\lambda)$ is a right inverse for $\lambda I - G$ and $\mathcal{R}(\lambda I - G) = L$.

If $g \in \mathcal{D}(G)$, then use the fact that G is closed to see that

$$R(\lambda)Gg = \int_0^\infty e^{-\lambda s}T(s)Gg\ ds = \int_0^\infty G(e^{-\lambda s}T(s))g\ ds = G\int_0^\infty e^{-\lambda s}T(s)g\ ds = GR(\lambda)g$$

or

$$R(\lambda)(\lambda I - G)g = g.$$

Thus, $(\lambda I - G)$ is one-to-one and its left inverse is $R(\lambda)$ □

Definition 4.23. $\rho(G) = \{\lambda; (\lambda I - G)$ *has a bounded inverse*$\}$ *is called the* resolvent set.

Thus, if G is the generator of a strongly continuous contraction semigroup, then its resolvent set $\rho(G)$ contains $(0, \infty)$.

Remark 4.24. *Let τ_λ be an exponential random variable with parameter λ, then as $\lambda \to \infty$, τ_λ converges in distribution to the degenerate random varibale that takes on the value 0 with probability 1. This motivates the following.*

Proposition 4.25. *Let $\{R(\lambda); \lambda \geq 0\}$ be the resolvent for a strongly continuous contraction semigroup with generator G, then*

$$\lim_{\lambda\to\infty} \lambda R(\lambda)f = f.$$

Proof. Note that for each $f \in \mathcal{D}(G)$,

$$\lambda R(\lambda)f - f = R(\lambda)Gf \text{ and that } ||G_\lambda Af|| \leq \lambda^{-1}||Gf||.$$

Thus, the formula holds for the dense set $f \in \mathcal{D}(G)$. Now, use the fact that $||\lambda R(\lambda) - I|| \leq 2$ for all $\lambda > 0$ and approximate. □

Exercise 4.26.

1. *(resolvent identity) Let $\lambda, \mu > 0$, then*

$$R(\lambda)R(\mu) = R(\mu)R(\lambda) = \frac{R(\mu) - R(\lambda)}{\lambda - \mu}.$$

2. *If $\lambda \in \rho(G)$ and $|\lambda - \mu| < ||R(\lambda)||^{-1}$, then*

$$R(\mu) = \sum_{n=0}^\infty (\lambda - \mu)^n R(\lambda)^{n+1}.$$

Consequently, $\rho(G)$ is open.

3. *If G is the generator of a semigroup then*

$$||\lambda f - Gf|| \geq \lambda||f|| \text{ for every } f \in \mathcal{D}(G), \text{ and } \lambda > 0.$$

4.3 The Hille-Yosida Theorem

The Hille-Yosida theorem gives a characterization for generators. We begin with a definition.

4.3.1 Dissipative Operators

Definition 4.27. *A linear operator A on L is* dissipative *if*

$$||\lambda f - Af|| \geq \lambda||f|| \text{ for every } f \in \mathcal{D}(A), \text{ and } \lambda > 0. \tag{4.5}$$

Remark 4.28. *If A is dissipative, then divide (4.5) by λ to obtain that $||f - Af/\lambda|| \geq ||f||$ or*

$$||f - cA \geq ||f|| \quad \text{for all} \quad c > 0.$$

Exercise 4.29. *1. If A is dissipative, then $\lambda I - A$ is one-to-one.*

2. If A is dissipative and $(\lambda I - A)^{-1}$ exists, then $||(\lambda I - A)^{-1}|| \leq \lambda^{-1}$.

Lemma 4.30. *Let A be a dissipative linear operator on L and let $\lambda > 0$. Then A is closed if and only if $\mathcal{R}(\lambda I - A)$ is closed.*

Proof. Suppose that A is closed. If $\{f_n; n \geq 0\} \subset \mathcal{D}(A)$, and $(\lambda I - A)f_n \to h$, then

$$||(\lambda I - A)(f_n - f_m)|| \geq \lambda||f_n - f_m||,$$

and $\{f_n; n \geq 0\}$ is a Cauchy sequence. Thus, there exists $f \in L$ such that $f_n \to f$ and hence

$$Af_n \to \lambda f - h.$$

Because A is closed, $f \in \mathcal{D}(A)$ and

$$h = (\lambda I - A)f.$$

This shows that $\mathcal{R}(\lambda I - A)$ is closed.

Suppose $\mathcal{R}(\lambda I - A)$ is closed. If $\{f_n; n \geq 0\} \subset \mathcal{D}(A)$, $f_n \to f$ and $Af_n \to g$. Consequently,

$$(\lambda I - A)f_n \to \lambda f - g.$$

Because $\mathcal{R}(\lambda I - A)$ is closed,

$$\lambda f - g = (\lambda I - A)f_0 \text{ for some } f_0 \in \mathcal{D}(A).$$

Because A is dissipative,

$$||(\lambda I - A)(f_n - f_0)|| \geq \lambda||f_n - f_0||,$$

and $f_n \to f_0$. Hence, $f = f_0 \in \mathcal{D}(A)$ and $Af = g$. This shows that A is closed. $\square$

Lemma 4.31. *Let A be a dissipative closed linear operator on L. Set $\rho^+(A) = \rho(A) \cap (0, \infty)$. If $\rho^+(A) \neq \emptyset$, then $\rho^+(A) = (0, \infty)$*

Proof. The condition is equivalent to the statement that $\rho^+(A)$ is both open and closed in $(0,\infty)$. We have shown that $\rho(A)$ is open.

Suppose that $\{\lambda_n; n \geq 1\} \subset \rho^+(A)$ with $\lambda_n \to \lambda > 0$. Given $g \in L$, define, for each n,

$$g_n = (\lambda I - A)(\lambda_n I - A)^{-1} g.$$

Because A is dissipative

$$||g_n - g|| = ||((\lambda I - A) - (\lambda_n I - A))(\lambda_n I - A)^{-1} g|| = ||(\lambda - \lambda_n)(\lambda_n I - A)^{-1} g|| \leq \frac{|\lambda - \lambda_n|}{\lambda_n} ||g||.$$

Thus, $g_n \to g$. Consequently, $\mathcal{R}(\lambda I - A)$ is dense in L. We have shown that this range is closed, thus $\mathcal{R}(\lambda I - A) = L$. Because $\lambda I - A$ is one-to-one and $||(\lambda I - A)^{-1}|| \leq \lambda^{-1}$, we have that $\lambda \in \rho^+(A)$. Thus, $\rho^+(A)$ is closed. □

4.3.2 Yosida Approximation

The generator of a strongly continuous contraction semigroup is sometimes an unbounded operator. The *Yosida approximation* gives us a method of approximating generators using bounded operators. Later, we will give a probabilistics interpretation to this approximation.

Definition 4.32. *Let G be a dissipative closed linear operator on L, and suppose that $\mathcal{D}(G)$ is dense in L and that $(0,\infty) \subset \rho(G)$. Then the Yosida approximation of G is*

$$G_\lambda = \lambda G(\lambda I - G)^{-1}, \quad \lambda > 0.$$

Proposition 4.33. *The Yosida approximation has the following properties:*

1. *For each $\lambda > 0$, G_λ is a bounded linear operator on L and $\{T_\lambda(t) = \exp(tG_\lambda); t \geq 0\}$ generates a strongly continuous contraction semigroup on L.*
2. *$G_\lambda G_\mu = G_\mu G_\lambda$ for all $\lambda, \mu > 0$*
3. *For every $f \in \mathcal{D}(G)$,*
$$\lim_{\lambda\to\infty} G_\lambda f = Gf.$$

Proof. Because

$$I = (\lambda I - G)R(\lambda) \text{ on } L, \quad A_\lambda = \lambda^2 R(\lambda) - \lambda I \text{ on } L.$$

Because

$$R(\lambda)(\lambda I - A) = I \text{ on } \mathcal{D}(G), \quad A_\lambda = \lambda R(\lambda) G \text{ on } \mathcal{D}(G).$$

Note that

$$||T_\lambda(t)|| \leq e^{-t\lambda} || \exp(t\lambda^2 R(\lambda))|| \leq e^{-t\lambda} e^{t\lambda^2 ||R(\lambda)||} \leq 1.$$

proving part 1.

Part 2 follows from the first identity above and the resolvent identity.

Part 3 follows from the second identity above and the fact that for $f \in \mathcal{D}(G)$, $\lambda R(\lambda) f \to f$ as $\lambda \to \infty$ as $\lambda \to \infty$. □

Lemma 4.34. *For a pair of commuting bounded operators B and C on L assume that $||\exp(tB)|| \leq 1$ and $||\exp(tC)|| \leq 1$ for all $t \geq 0$. Then*

$$||\exp(tB)f - \exp(tC)f|| \leq t||Bf - Cf||.$$

Proof.

$$\begin{aligned}
\exp(tB)f - \exp(tC)f &= \int_0^t \frac{d}{dt}(\exp(sB)\exp((t-s)C))f\ ds \\
&= \int_0^t \exp(sB)(B-C)\exp((t-s)C)f\ ds \\
&= \int_0^t \exp(sB)\exp((t-s)C)(B-C)f\ ds.
\end{aligned}$$

The rest is easy. □

Theorem 4.35 (Hille-Yosida)**.** *A linear operator G is the generator of a strongly continuous contraction semigroup on L if and only if*

1. *$\mathcal{D}(G)$ is dense in L.*
2. *G is dissipative.*
3. *$\mathcal{R}(\lambda_0 I - G) = L$ for some $\lambda_0 > 0$*

Proof. The necessity of these conditions has been established.

We have shown that conditions 2 and 3 implies that G is closed. By condition 3, $\lambda_0 \in \rho(G)$ and, therefore, $(0, \infty) \subset \rho(G)$.

As before, for each $\lambda > 0$, define the Yosida approximation G_λ and its semigroup T_λ. Then by the lemma above, we

$$||T_\lambda(t)f - T_\mu(t)f|| \leq t||G_\lambda f - G_\mu f||.$$

for all $f \in L$, $t \geq 0$ and $\lambda, \mu > 0$.

Consequently, because the Yosida approximations converge to the generator G, we have, for all $f \in \mathcal{D}(G)$,

$$\lim_{\lambda \to \infty} T_\lambda(t)f$$

exists for all $t \geq 0$ uniformly on bounded intervals. Call this limit $T(t)f$. Because $\mathcal{D}(G)$ is dense, this limit holds for all $f \in L$.

Claim 1. T is a strongly continuous contraction semigroup.

The fact that $T(0)f = f$ and that $T(t)$ is a contraction is immediate. Use the identity

$$T(t)f - f = (T(t) - T_\lambda(t))f + (T_\lambda(t) - I).$$

to check the strong continuity at 0 and

$$T(s+t)f - T(s)T(t)f = (T(s+t) - T_\lambda(s+t))f + T_\lambda(s)(T_\lambda(t) - T(t))f + (T_\lambda(s) - T(s))T(t)f$$

to check the semigroup property.

Claim 2. G is the generator of T.

For every $f \in L, t \geq 0$, and $\lambda > 0$,

$$T_\lambda(t)f - f = \int_0^t T_\lambda(s)G_\lambda f \; ds.$$

For each $f \in \mathcal{D}(G)$ and $s \geq 0$,

$$T_\lambda(s)G_\lambda f \; - T(s)Gf = T_\lambda(s)(G_\lambda f - Gf) + (T_\lambda(s) - T(s))Gf.$$

Use the convergence of the Yosida approximations and the convergence of the semigroups to see that

$$T_\lambda(s)G_\lambda f \to T(s)Gf \quad \text{as } \lambda \to \infty$$

uniformly for $s \in [0,t]$. Thus the Riemann integrals converge and

$$T(t)f - f = \int_0^t T(s)Gf \; ds \quad f \in \mathcal{D}(G), \; t \geq 0.$$

Consequently, the generator $\tilde{G}$ of T is an extension of G.

With this in mind, choose $\tilde{f} \in \mathcal{D}(\tilde{G})$. Because $\lambda I - G$ is surjective, there exists $f \in \mathcal{D}(G)$ so that

$$(\lambda I - \tilde{G})\tilde{f} = (\lambda I - G)f.$$

Now, evaluate this expression for two distinct values of λ and subtract to obtain that $\tilde{f} = f$. □

Lemma 4.36. *Let A be a dissipative operator on L and suppose that $u : [0,\infty) \to \mathcal{D}(A)$ be continuous for all $s > 0$ and $Au : [0,\infty) \to L$ continuous. If*

$$u(t) = u(\epsilon) + \int_\epsilon^t Au(s) \; ds$$

for all $t > \epsilon > 0$, then $||u(t)|| \leq ||u(0)||$ for all $t \geq 0$.

Proof. Let $0 < \epsilon = t_0 < t_1 < \cdots < t_n = t$. Then

$$\begin{aligned}
||u(t)|| &= ||u(\epsilon)|| + \sum_{j=1}^n (||u(t_j)|| - ||u(t_{j-1})||) \\
&= ||u(\epsilon)|| + \sum_{j=1}^n (||u(t_j)|| - ||u(t_j) - (t_j - t_{j-1})Au(t_j)||) \\
&\quad + \sum_{j=1}^n (||u(t_j) - (t_j - t_{j-1})Au(t_j)|| - ||u(t_j) - (u(t_j) - u(t_{j-1}))||) \\
&\leq ||u(\epsilon)|| + \sum_{j=1}^n (||u(t_j) - \int_{t_j}^{t_{j-1}} Au(t_j) \; ds||) - ||u(t_j) - \int_{t_{j-1}}^{t_j} Au(s) \; ds||) \\
&\leq ||u(\epsilon)|| + \sum_{j=1}^n \int_{t_{j-1}}^{t_j} ||Au(t_j) - Au(s)|| \; ds.
\end{aligned}$$

The first sum is negative by the dissipativity of A. Now use the continuity of Au and u and let $\max(t_j - t_{j-1}) \to 0$. Then, let $\epsilon \to 0$. □

Proposition 4.37. *Let T and S be strongly continuous contraction semigroups on L with generators G_T and G_S. If these two generators are equal, the $T = S$.*

Proof. Use the lemma above with

$$u(t) = (T(t) - S(t))f$$

and note that $u(0) = 0$ and $Au(s) = 0$ for all s. □

Definition 4.38. *Call a linear operator A on L* closable *if it has a closed linear extension. The* closure *$\bar{A}$ of A is the minimal closed linear extension of A.*

Often the generator G is an unbounded operator. In these instances, establishing $\mathcal{D}(G)$ can be a tedious process. Thus, we look for an alternative form for the Hille-Yosida theorem.

Lemma 4.39. *Let G be a dissipative linear operator on L with $\mathcal{D}(G)$ dense in L. Then G is closable and $\overline{\mathcal{R}(\lambda I - G)} = \mathcal{R}(\lambda I - \bar{G})$ for every $\lambda > 0$.*

Proof. For the first assertion, we show that

$$\{f_n; n \geq 0\} \subset \mathcal{D}(G),\ f_n \to 0, Gf_n \to g \text{ implies } g = 0.$$

Choose $\{g_m; m \geq 0\} \subset \mathcal{D}(G)$ such that $g_m \to g$. Then by the dissipativity of G,

$$||(\lambda I - G)g_m - \lambda g|| = \lim_{n\to\infty} ||(\lambda I - G)(g_m + \lambda f_n)|| \geq \lim_{n\to\infty} \lambda||g_m + \lambda f_n|| = \lambda||g_m||$$

for every $\lambda > 0$. Now divide by λ and let $\lambda \to \infty$ to obtain, for each m, $||g_m - g|| \geq ||g_m||$. Letting $m \to \infty$ yields $g = 0$.

The inclusion $\mathcal{R}(\lambda I - \bar{G}) \subset \overline{\mathcal{R}(\lambda I - G)}$ is straightforward. The reverse inclusion follows from the fact that dissipative linear operator $\tilde{G}$ is closed if and only if $\mathcal{R}(\lambda I - \tilde{G})$ is closed. □

Theorem 4.40. *A linear operator G on L is closable and its closure $\bar{G}$ is the generator of a strongly continuous contraction semigroup on L if and only if*

1. *$\mathcal{D}(G)$ is dense in L.*
2. *G is dissipative.*
3. *$\mathcal{R}(\lambda_0 I - G)$ is dense in L for some $\lambda_0 > 0$*

Proof. By the lemma above G satisfies 1-3 if and only if G is closable and $\bar{G}$ satisfies 1-3 in the Hille-Yosida theorem. □

Definition 4.41. *Let A be a closed linear operator on L. A subspace D of $\mathcal{D}(A)$ is called a* core *for A if A is the closure of the restriction of A to D.*

Proposition 4.42. *Let G be the generator of a strongly continuous contraction semigroup on L. Then a subspace D of $\mathcal{D}(G)$ is a core for G if and only if D is dense in L and $\mathcal{R}(\lambda_0 I - G)$ is dense in L for some $\lambda_0 > 0$.*

4.3.3 Positive Operators and Feller Semigroups

We now return to the consideration of Markov processes.

Definition 4.43. *Let S be locally compact and let $C_0(S)$ be the Banach space of continuous functions that vanish at infinity with norm*

$$||f|| = \sup\{|f(x)| : x \in S\}.$$

Definition 4.44. *Call an operator on a function space* positive *if it maps nonnegative functions to nonnegative functions.*

To establish positivity, we will use the following proposition.

Proposition 4.45. *Let T be a strongly continuous contraction semigroup on L with generator G and resolvent R. For $M \subset L$, let*

$$\Lambda_M = \{\lambda > 0; \lambda R(\lambda) : M \to M\}.$$

If M is closed and convex and Λ_M is unbounded, then for each $t \geq 0$,

$$T(t) : M \to M.$$

Proof. For $\lambda, \mu \geq 0$ satisfying $|1 - \mu/\lambda| < 1$,

$$\mu R(\mu) = \sum_{n=0}^{\infty} \frac{\mu}{\lambda} \left(1 - \frac{\mu}{\lambda}\right)^n (\lambda R(\lambda))^{n+1}.$$

Because M is closed and convex, $\lambda \in \Lambda_M$ implies $(0, \lambda] \subset \Lambda_M$. Thus, $\Lambda_M = (0, \infty)$. Use the Yosida approximation

$$T_\lambda(t) = e^{-t\lambda} \exp(t\lambda(\lambda R(\lambda))) = e^{-t\lambda} \sum_{n=0}^{\infty} \frac{(t\lambda)^n}{n!} (\lambda R(\lambda))^n$$

to conclude that

$$T_\lambda(t) : M \to M.$$

Now use the fact that M is closed and that $T_\lambda(t)f$ converges to to $T(t)f$ as $\lambda \to \infty$. □

Definition 4.46. 1. *Call a semigroup* conservative *if there exist a sequence $f_n \in C_0(S)$ that is bounded in norm and converges pointwise to 1, and*

$$\lim_{n\to\infty} T(t)f_n = 1.$$

for all $t \geq 0$

2. *Call a conservative strongly continuous positive contraction semigroup T on $C_0(S)$ a* Feller semigroup.

Exercise 4.47. 1. *If T is conservative, then $A1 = 0$*

2. *If $A1 = 0$, then T is conservative.*

4.3.4 The Maximum Principle

Definition 4.48. *A linear operator A on $C_0(S)$ satisfies* the positive maximum principle *if whenever $f \in \mathcal{D}(A)$, $\tilde{x} \in S$, then*

$$f(\tilde{x}) = \sup\{f(x); x \in S\} \geq 0 \textit{ implies } Af(\tilde{x}) \leq 0.$$

Remark 4.49. *To see why this ought to be true for a Markov process, note that*

$$0 \geq E_{\tilde{x}}[f(X_s)] - f(\tilde{x}) = T(s)f(\tilde{x}) - f(\tilde{x}).$$

now divide by s and let $s \to 0$.

Proposition 4.50. *Let S be locally compact. A linear operator A on $C_0(S)$ satisfying the positive maximum principle is dissipative.*

Proof. Let $f \in \mathcal{D}(A)$ and $\lambda > 0$. There exists $\tilde{x} \in S$ so that $|f(\tilde{x})| = ||f||$. Suppose $f(\tilde{x}) \geq 0$. (Otherwise replace f with $-f$.) Then by the positive maximum principle, $Af(\tilde{x}) \leq 0$ and hence

$$||\lambda f - Af|| \geq \lambda f(\tilde{x}) - Af(\tilde{x}) \geq \lambda f(\tilde{x}) = \lambda ||f||.$$

□

In this context, we have the following variant of the Hille-Yosida theorem.

Theorem 4.51. *Let S be locally compact. A linear operator G on $C_0(S)$ is closable and its closure $\bar{G}$ is the generator of a positive strongly continuous contraction semigroup on $C_0(S)$ if and only if*

1. *$\mathcal{D}(G)$ is dense in $C_0(S)$.*
2. *G satisfies the positive maximum principle.*
3. *$\mathcal{R}(\lambda_0 I - G)$ is dense in $C_0(S)$ for some $\lambda_0 > 0$.*

Proof. The necessity of conditions 1 and 3 follows from the theorem above. To check the neccessity of 2, fix $f \in \mathcal{D}(G)$ and $\tilde{x} \in S$ so that $\sup\{f(x) : x \in S\} = f(\tilde{x}) \geq 0$. Then, for each $t \geq 0$,

$$T(t)f(\tilde{x}) \leq T(t)f^+(\tilde{x}) \leq ||f^+|| = f(\tilde{x})$$

and $Gf(\tilde{x}) \leq 0$.

Conversely, suppose G satisfies conditions 1-3. Because condition 2 implies that G is dissipative, $\bar{G}$ generates a strongly continuous contraction semigroup T. To prove that T is positive, we note, by the proposition above, it suffices to prove that $R(\lambda)$ maps nonnegative functions to nonegative functions.

Because $\mathcal{R}(\lambda_0 I - G)$ is dense for some $\lambda_0 > 0$, and G is dissipative, $\mathcal{R}(\lambda I - \bar{G}) = C_0(S)$ all $\lambda > 0$. Thus, it is equivalent to show that for $f \in \mathcal{D}(G)$, and $\lambda > 0$,

$$(\lambda I - \bar{G})f \geq 0 \text{ implies } f \geq 0.$$

To establish the contrapositive of this statement, choose $f \in \mathcal{D}(\bar{G})$ so that $\inf\{f(x) : x \in S\} < 0$. Thus, there exist $\{f_n; n \geq 0\} \in \mathcal{D}(G)$ so that

$$\lim_{n\to\infty} (\lambda I - G)f_n = (\lambda I - \bar{G})f.$$

Because $\bar{G}$ is dissipative, we have that $f_n \to f$ as $n \to \infty$. Let $\tilde{x}_n$ and $\tilde{x}$ be the points in which f_n and f, respectively take on their minimum values. Then

$$\inf\{(\lambda I - \bar{G})f; x \in S\} \leq \liminf_{n\to\infty}(\lambda I - \bar{G})f_n(\tilde{x}_n) \leq \liminf_{n\to\infty} \lambda f_n(\tilde{x}_n) = \lambda f(\tilde{x}) < 0.$$

□

Exercise 4.52. *Let X_1 and X_2 be independent Markov processes on a common state space S. Is (X_1, X_2) a Markov process? If so, determine its generator?*

4.4 Strong Markov Property

Definition 4.53. *Let X be a $\mathcal{F}_t$-progressive, $\mathcal{F}_t$ Markov process with associated transition function P. Let τ be an almost surely finite $\mathcal{F}_t$-stopping time. Then X is* strong Markov at τ *if for all $t \geq 0$ and $B \in \mathcal{B}(S)$,*

$$P\{X_{\tau+s} \in B|\mathcal{F}^X_\tau\} = p(s, X_\tau, B),$$

or equivalently, for any $t \geq 0$ and bounded measurable f,

$$E[f(X_{\tau+s})|\mathcal{F}^X_\tau] = \int_S f(y)p(s, X_\tau, dy).$$

X is called strong Markov *with respect to $\{\mathcal{F}_t; t \geq 0\}$ if X is strong Markov for all almost finite $\mathcal{F}_t$-stopping times.*

In reviewing the results on Lévy processes and its regularity properties, we obtain the following.

Proposition 4.54. *A right continuous Lévy process is strong Markov.*

The following proposition has essentially the same proof as in the discrete time case.

Proposition 4.55. *Let X be a $\mathcal{F}_t$-progressive, $\mathcal{F}_t$ Markov process with associated transition function P. The X is strong Markov at all discrete almost surely finite $\mathcal{F}_t$-stopping times.*

Now, using the essentially the same proof we had for Lévy processes, we have

Proposition 4.56. *Let X be a right continuous $\mathcal{F}_t$-progressive, $\mathcal{F}_t$ Markov process with associated transition function P. Then X is strong Markov at all almost surely finite $\mathcal{F}_{t+}$-stopping times.*

Theorem 4.57. *Let X be a Markov process with time homogenous transition function P. Let Y be a bounded measurable process and assume that and let τ be a $\mathcal{F}^X_t$-stopping time so that X is strong Markov at $\tau + t$ for all $t \geq 0$. Then*

$$E_\alpha[Y_\tau \circ \theta^\tau|\mathcal{F}_\tau] = \phi_\tau(X_\tau) \quad on\ \{\tau < \infty\}$$

where

$$\phi_s(x) = E_x Y_s.$$

Proof. By the standard machine, we need only consider functions of the form

$$Y_s = f_0(s)\prod_{k=1}^n f_k(X_{t_k}), \quad 0 < t_1 < \cdots < t_n.$$

$f_0, f_1 \ldots, f_n$ continuous and bounded. In this case

$$\begin{aligned}\phi_s(x) &= f_0(s)\int\int\cdots\int f_n(x_n)p(t_n - t_{n-1}, x_{n-1}, dx_n)f_{n-1}(x_{n-1})p(t_{n-1}-t_{n-2}, x_{n-2}, dx_{n-1}) \\ &\quad \cdots f_1(x_1)p(t_1, x, dx_1) = f_0(s)\psi(x).\end{aligned}$$

The proof will process by induction on n. The case $n = 0$ states that

$$E_\alpha[f_0(\tau)|\mathcal{F}_\tau] = f_0(\tau) \quad \text{on } \{\tau < \infty\}.$$

This follows form the fact that τ is $\mathcal{F}_\tau$-measurable.

Let

$$\tilde{\phi}(x) = E[\prod_{k=2}^n f_k(X_{t_k})].$$

Then,

$$\phi_s(x) = f_0(s)\int \phi(x_1)f_1(x_1)p(t_1, x, dx_1)$$

On $\{\tau < \infty\}$,

$$\begin{aligned}E_\alpha[Y_\tau \circ \theta^\tau|\mathcal{F}_\tau] &= E_\alpha[f_0(\tau)\prod_{k=1}^n f_k(X_{\tau+t_k})|\mathcal{F}_\tau] \\ &= f_0(\tau)E_\alpha[E_\alpha[\prod_{k=2}^n f_k(X_{\tau+t_k})|\mathcal{F}_{\tau+t_1}]f_1(X_{\tau+t_1})|\mathcal{F}_\tau] \\ &= f_0(\tau)E_\alpha[\tilde{\phi}(X_{\tau+t_1})f_1(X_{\tau+t_1})|\mathcal{F}_\tau] \\ &= f_0(\tau)E_\alpha[(\tilde{\phi}(X_{t_1})f_1(X_{t_1}))\circ\theta^\tau|\mathcal{F}_\tau] \\ &= f_0(\tau)\psi(X_\tau) = \phi_\tau(X_\tau)\end{aligned}$$

□

Using essentially the same proof as in the case of Lévy processes, we have the following.

Theorem 4.58 (Blumenthal 0-1 law)**.** *Let X be a right continuous $\mathcal{F}_t$-progressive, $\mathcal{F}_t$ Markov process. Then for each $x \in S$ every event in $\mathcal{F}^X_{0+}$ has probability 0 or 1.*

Corollary 4.59. *Let τ be an $\mathcal{F}^X_{t+}$-stopping time and fix $x \in S$. Then either*

$$P_x\{\tau = 0\} = 1 \quad \textit{or} \quad P_x\{\tau > 0\} = 1.$$

4.5 Connections to Martingales

Proposition 4.60. *Let X be an S-valued progressive Markov process and let T, the semigroup assocated to X, have generator G. Then for $f \in \mathcal{D}(G)$,*

$$M_t^f = f(X_t) - \int_0^t Gf(X_s)\ ds$$

is an $\mathcal{F}_t^X$-martingale.

Proof. For each $t, u \geq 0$

$$\begin{aligned} E[M_{t+u}^f|\mathcal{F}_t^X] &= T(u)f(X_t) - \int_0^t Gf(X_s)\ ds - \int_t^{t+u} T(s-t)Gf(X_t)\ ds \\ &= T(u)f(X_t) - \int_0^u T(s)Gf(X_t)\ ds - \int_0^t Gf(X_s)\ ds \\ &= f(X_t) - \int_0^t Gf(X_s)\ ds = M_s^f \end{aligned}$$

□

Exercise 4.61. *Let X be a right continuous Markov process with generator G and assume that*

$$f(X_t) - \int_0^t g(X_s)\ ds$$

is a martingale, then $f \in \mathcal{D}(G)$ and $Gf = g$.

The optional sampling formula gives us the following.

Theorem 4.62 (Dynkin formula)**.** *In addition to the conditions above, let τ be a stopping time with $E_\alpha \tau < \infty$, then*

$$E_\alpha f(X_\tau) = E_\alpha f(X_0) + E_\alpha[\int_0^\tau Gf(X_s)\ ds]$$

Analogous to the situation for discrete time Markov chains, we make the following definitions.

Definition 4.63. *Let G be the generator for a time homogeneous Markov process X and let $f \in \mathcal{D}(G)$. Then call*

1. f harmonic *if $Gf = 0$.*
2. f superharmonic *if $Gf \leq 0$.*
3. f subharmonic *if $Gf \geq 0$.*

Remark 4.64.

1. *If f harmonic, then $f(X_t)$ is a martingale.*
2. *If f superharmonic, then $f(X_t)$ is a supermartingale.*

3. *If f subharmonic, then $f(X_t)$ is a submartingale.*

Example 4.65. 1. *Let D_0, D_1 are two disjoint subsets of the state space S. If h is harmonic with respect to a Markov process X having generator G and*

$$h(x) = 0 \text{ for } x \in D_0, \ h(x) = 1 \text{ for } x \in D_1.$$

Assume that the first entrance times $\tau_{D_j}, j = 0, 1$ are stopping times and τ is their minimum. Then if τ has finite mean,

$$h(x) = E_x h(X_\tau) = P\{\tau_{D_0} > \tau_{D_1}\}.$$

2. *If we can solve the equation $Gf = -1$, then*

$$M_t^f = f(X_t) + t$$

is martingale. Consider a domain D so that τ_D is a stopping time. $Gf = -1$, $f = 0$ on D, and τ_D satisfies the sampling integrability conditions for M^f, then

$$f(x) = E_x \tau_D.$$

Lemma 4.66. *Assume that $g \geq 0$ and let $R(\lambda)$ be the resolvent for a positive continuous contraction semigroup T associated to a Markov process X. Then*

$$e^{-\lambda t} R(\lambda) g(X_t)$$

is a non-negative $\mathcal{F}_{t+}^X$-supermartingale.

Proof. For $t > s$, use the fact that the semigroup and the resolvent commute to obtain that

$$\begin{aligned} E[e^{-\lambda t} R(\lambda) g(X_t) | \mathcal{F}_{s+}^X] &= e^{-\lambda t} T(t-s) R(\lambda) g(X_s) \\ &= e^{-\lambda t} \int_0^\infty e^{-\lambda u} T(t-s+u) g(X_s) \, du \\ &= e^{-\lambda s} \int_{t-s}^\infty e^{-\lambda u} T(u) g(X_s) \, du \\ &\leq e^{-\lambda s} R(\lambda) g(X_s) \end{aligned}$$

□

Let's continue this line analysis. Consider the bounded random variable,

$$Y = \int_0^\infty e^{-\lambda s} g(X_s) \, ds$$

Then

$$Y = \int_0^t e^{-\lambda s} g(X_s) \, ds + e^{-\lambda t} Y \circ \theta^t$$

and

$$E_x Y = R(\lambda) g(x).$$

We have Doob's martingale

$$\begin{aligned} Z_t^{\lambda,g} = E[Y|\mathcal{F}_t^X] &= \int_0^t e^{-\lambda s} g(X_s)\ ds + e^{-\lambda t} E[Y \circ \theta^t | \mathcal{F}_t^X] \\ &= \int_0^t e^{-\lambda s} g(X_s)\ ds + e^{-\lambda t} R(\lambda) g(X_t) \end{aligned}$$

Now, because Doob's martingale is uniformly integrable, any stopping time τ satisfies the sampling integrability conditions for $Z^{\lambda,g}$ and therefore,

$$R(\lambda)g(x) = E_x[\int_0^\tau e^{-\lambda s} g(X_s)\ ds] + E_x[e^{-\lambda\tau} R(\lambda) g(X_\tau)].$$

Now, let $f = (\lambda I - G)g$, then the Doob's martingale above becomes

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

and the analog to Dynkin's formula becomes

$$f(x) = E_x[e^{-\lambda\tau} f(X_\tau)] + E_x[\int_0^\tau e^{-\lambda s} (\lambda I - G) f(X_s)\ ds].$$

Let f_λ satisfy the eigenvalue problem $Gf_\lambda = \lambda f_\lambda$ with $f_\lambda = 1$ on D, then

$$f_\lambda(x) = E_x e^{-\lambda \tau_D}.$$

Theorem 4.67. *Let X be a measurable $\mathcal{F}_t$-adapted process and let f and g be bounded, measurable functions,* $\inf_x f(x) > 0$. *Assume that*

$$Y_t = f(X_t) - \int_0^t g(X_s)\ ds$$

is an $\mathcal{F}_t$ martingale. Then

$$f(X_t) \exp\left(- \int_0^t \frac{g(X_v)}{f(X_v)}\ dv \right)$$

is a martingale.

Proof. Let

$$V_t = \exp\left(- \int_0^t \frac{g(X_v)}{f(X_v)}\ dv \right).$$

Then, we have the local martingale

$$\begin{aligned}
Y_tV_t - \int_0^t Y_s\, dV_s &= \left(f(X_t) - \int_0^t g(X_s)\, ds\right) \exp\left(-\int_0^t \frac{g(X_v)}{f(X_v)}\, dv\right) \\
&\quad + \int_0^t \left(f(X_s) - \int_0^s g(X_u)\, du\right) \frac{g(X_s)}{f(X_s)} \exp\left(-\int_0^s \frac{g(X_v)}{f(X_v)}\, dv\right) \\
&= f(X_t) \exp\left(-\int_0^t \frac{g(X_v)}{f(X_v)}\, dv\right) - \int_0^t g(X_s)\, ds \exp\left(-\int_0^t \frac{g(X_v)}{f(X_v)}\, dv\right) \\
&\quad + \int_0^t g(X_s) \exp\left(-\int_0^s \frac{g(X_v)}{f(X_v)}\, dv\right)\, ds \\
&\quad - \int_0^t \int_0^s g(X_u) \frac{g(X_s)}{f(X_s)} \exp\left(-\int_0^s \frac{g(X_v)}{f(X_v)}\, dv\right)\, du\, ds
\end{aligned}$$

For the double integral, reverse the order of integration, then we have, upon integrating the s variable that

$$\begin{aligned}
&-\int_0^t g(X_u) \exp\left(-\int_0^s \frac{g(X_v)}{f(X_v)}\, dv\right)\Big|_u^t\, du \\
= -\int_0^t g(X_u) &\exp\left(-\int_0^t \frac{g(X_v)}{f(X_v)}\, dv\right)\, du + \int_0^t g(X_u) \exp\left(-\int_0^u \frac{g(X_v)}{f(X_v)}\, dv\right)\, du
\end{aligned}$$

and the last three terms sum to zero.

Now use the fact that f and g are bounded and f is bounded away from 0 to see that the local martingale above is indeed a martingale. □

Exercise 4.68. *Let X be a measurable $\mathcal{F}_t$-adapted process and let f and g be bounded, measurable functions,*

1. *Assume that $\inf_x f(x) > 0$ and that*

$$f(X_t) \exp\left(-\int_0^t \frac{g(X_v)}{f(X_v)}\, dv\right)$$

   is an $\mathcal{F}_t$ martingale. Then

$$f(X_t) - \int_0^t g(X_s)\, ds$$

   is an $\mathcal{F}_t$ martingale.

2. *Assume that*

$$f(X_t) - \int_0^t g(X_s)\, ds$$

   is an $\mathcal{F}_t$ martingale. Then

$$e^{-\lambda t} f(X_t) + \int_0^t e^{-\lambda s}(\lambda f(X_s) - g(X_s))\, ds$$

   is an $\mathcal{F}_t$ martingale.

4.6 Jump Processes

Definition 4.69. *Call a time homogenuous Markov process X a* pure jump process*, if starting from any point $x \in S$, the process is right continuous and has all its sample paths constant except for isolated jumps.*

In the case of a countable state space, the stochastic nature of a pure jump process is captured by the *infinitesimal transitions rates*

$$P_x\{X_h = y\} = g(x,y)h + o(h).$$

The *infinitesimal generator* G is the rate of change of averages for a function $f : S \to \mathbb{R}$ of the process.

$$Gf(x) = \lim_{h\to 0} \frac{E_x f(X_h) - f(x)}{h}, \quad f \in \mathcal{D}(G).$$

To relate these two concepts, write

$$\begin{aligned} E_x f(X_h) &= \sum_{y\neq x} f(y)g(x,y)h + f(x)(1 - \sum_{y\neq x} g(x,y)h) + o(h) \\ E_x f(X_h) - f(x) &= \sum_{y\neq x} g(x,y)(f(y) - f(x))h + o(h) \\ Gf(x) &= \sum_{y\in S} g(x,y)(f(y) - f(x)) \end{aligned}$$

Thus, G can be written as an *infinitesimal transition matrix.* The xy-entry $x \neq y$ is $g(x,y)$. The diagonal entry

$$g(x,x) = -\sum_{y\neq x} g(x,y).$$

Thus, the off-diagonal entries of the matrix G are non-negative and the row sum is 1.

4.6.1 The Structure Theorem for Pure Jump Markov Processes

For processes that move from one state to another by jumping the exponential distribution plays an important role.

Proposition 4.70. *Let X be a pure jump process, then*

$$\tau_1 = \inf\{t \geq 0; X_t \neq X_0\}$$

is an $\mathcal{F}_{t+}$-stopping time.

Proof. $\tau_1 = \inf\{\sigma_n : n \geq 1\}$ where

$$\sigma_n = \inf\{k2^{-n}; X_{k/2^n} \neq X_0\}$$

Now

$$\{\sigma_n \leq t\} \in \mathcal{F}_t^X \subset \mathcal{F}_t.$$

The σ_n are $\mathcal{F}_t$-stopping times and, consequently, τ_1 is an $\mathcal{F}_{t+}$-stopping time. □

Theorem 4.71 (structure theorem for pure jump Markov processes)**.** *Under P_x, τ_1 and X_{τ_1} are independent and there is a $\mathcal{B}(S)$ measurable function λ on S such that*

$$P_x\{\tau_1 > t\} = \exp(-\lambda(x)t).$$

Proof. Set

$$e_x(t+s) = P_x\{\tau_1 > t+s\} = P_x\{\tau_1 > t+s, X_t = x, \tau_1 > t\} = P_x\{\tau_1 > t+s|X_t = x, \tau_1 > t\}P_x\{X_t = x, \tau_1 > t\}$$

Note that

$$P\{\tau_1 > t+s|X_t = x, \tau_1 > t\} = P_x\{\tau_1 > t+s|X_t = x\} = P_x\{\tau_1 > s\} = e_x(s).$$

and that $P_x\{X_t = x, \tau_1 > t\} = P_x\{\tau_1 > t\} = e_x(t)$. Thus,

$$e_x(t+s) = e_x(s)e_x(t)$$

or for some $\lambda \in [0,\infty]$, $e_x(t) = \exp(-\lambda(x)t)$. The function λ is measurable because $P_x\{\tau_1 > t\}$ is measurable.

Let $\tau_1(t)$ be the first exit from state x after time t. Then, for $B \in \mathcal{B}(S), x \notin B$,

$$P_x\{X_{\tau_1} \in B, \tau_1 > t\} = P_x\{X_{\tau_1(t)} \in B, \tau_1 > t, X_t = x\} = P_x\{X_{\tau_1(t)} \in B|\tau_1 > t, X_t = x\}P_x\{\tau_1 > t, X_t = x\}.$$

For the first term

$$P_x\{X_{\tau_1(t)} \in B|\tau_1 > t, X_t = x\} = P_x\{X_{\tau_1(t)} \in B|X_t = x\} = P_x\{X_{\tau_1} \in B\}.$$

For the second term,

$$P_x\{\tau_1 > t, X_t = x\} = P_x\{\tau_1 > t\}.$$

Thus, the time of the first jump and the place of the first jump are independent. □

Let $\mu : S \times \mathcal{B}(S) \to [0,1]$ be a transition function and defined by $\mu(x,B) = P_x\{X_{\tau_1} \in B\}$, then

$$\begin{aligned} E_x f(X_h) &= E_x[f(X_h)|\tau_1 > h]P_x\{\tau_1 > h\} + E_x[f(X_h)|\tau_1 \leq h]P_x\{\tau_1 \leq h\} \\ &= f(x)e^{-\lambda(x)h} + \int f(y)\mu(x,dy)(1 - e^{-\lambda(x)h}) + o(h) \end{aligned}$$

Thus,

$$E_x f(X_h) - f(x) = (1 - e^{-\lambda(x)h})\left(\int f(y)\mu(x,dy) - f(x)\right) + o(h)$$

and

$$Gf(x) = \lambda(x)\int (f(y) - f(x))\mu(x,dy).$$

In the case of a countable set of states, set $\mu(x,\{y\}) = T(x,y)$.

$$Gf(x) = \lambda(x)\sum_{y\in S} T(x,y)(f(y) - f(x))$$

Equating the two expressions for the generator, we find that

$$g(x,x) = -\lambda(x)$$

and for $y \neq x$

$$g(x,y) = \lambda(x)T(x,y) \quad \text{or} \quad T(x,y) = \frac{g(x,y)}{\lambda(x)}$$

For example, let $S = \{1,2,3,4\}$,

$$G = \begin{pmatrix} -5 & 2 & 3 & 0 \\ 4 & -10 & 3 & 3 \\ 2 & 2 & -5 & 1 \\ 5 & 0 & 0 & -5 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & .4 & .6 & 0 \\ .4 & 0 & .3 & .3 \\ .4 & .4 & 0 & .2 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \lambda = \begin{pmatrix} 5 \\ 10 \\ 5 \\ 5 \end{pmatrix}$$

Exercise 4.72. *1. Let $T_1, \ldots, T_n$ be independent, exponential random variables, having respectively, parameters $\lambda_1, \ldots, \lambda_n$, then*

$$\min\{T_1, \ldots, T_n\}$$

is an exponential random variable with parameter $\lambda_1 + \cdots + \lambda_n$.

2. $P\{T_1 < \min\{T_2, \ldots, T_n\}\} = \lambda_1/(\lambda_1 + \cdots + \lambda_n)$.

Exercise 4.73. *Consider a stochastic process X with a finite state space S defined as follows.*

1. *For every $x, y \in S$, $x \neq y$, let $N^{(x,y)}$ denote a Poisson process with parameter $g(x,y)$. Assume that these processes are independent.*
2. *if $X_t = x$, then X jumps to y if*
$$y = \arg\min\{\tilde{y}; N_s^{(x,\tilde{y})}, s \geq t\}.$$
3. *The first jump after time t takes place at $\min_{\tilde{y}}\{N_s^{(x,\tilde{y})} : s \geq t\}$.*

Show that X is a Markov process whose generator G can be represented by a matrix with xy-entry $g(x,y)$.

4.6.2 Construction in the Case of Bounded Jump Rates

Let $\lambda : S \to [0, \lambda_{max}]$ be measurable. Define

$$Gf(x) = \lambda(x) \int_S (f(y) - f(x)) \, \mu(x, dy).$$

Exercise 4.74. *Show that G satisfies the maximum principle.*

We can construct the Markov process on S with generator G as follows:

Let $\{Y_k; k \geq 0\}$ be a Markov chain on S with initial distribution α and transition function μ. In addition, independent of the chain Y, let $\{\sigma_k; k \geq 0\}$ be independent exponentially distributed random variables with parameter 1. Set

$$\tau_0 = 0, \quad \tau_k = \sum_{j=0}^{k-1} \frac{\sigma_j}{\lambda(Y_j)}$$

and

$$X_t = Y_k, \quad \tau_k \le t < \tau_{k+1}.$$

Note that we can allow $\lambda(x) = 0$. In this case, once X arrives to x it remains there. Thus, if the Markov process arrives at the site x is remains there an exponential length of time, parameter $\lambda(x)$ and then jumps to a new site according to the transition function μ.

Let us consider the alternative representation for G. Define

$$\tilde{\mu}(x, B) = \left(1 - \frac{\lambda(x)}{\lambda_{max}}\right)\delta_x(B) + \frac{\lambda(x)}{\lambda_{max}}\mu(x, B).$$

Exercise 4.75. *If* $\tilde{G}f(x) = \lambda_{max}\int_S (f(y) - f(x))\ \tilde{\mu}(x, dy)$*, then* $G = \tilde{G}$

Write

$$Tf(x) = \int_S f(y)\ \tilde{\mu}(x, dy)$$

for the transition operator for the Markov chain $\tilde{Y}$. Then

$$G = \lambda_{max}(T - I).$$

The semigroup generated by G is

$$T(t) = \exp(tG) = e^{-\lambda_{max}t}\sum_{k=0}^{\infty}\frac{\lambda_{max}^k}{k!}T^k.$$

Let $\{\tilde{Y}_k; k \ge 0\}$ be a Markov chain on S with initial distribution α and transition function $\tilde{\mu}$ and let N be an independent Poisson process with parameter λ_{max}, and define

$$\tilde{X}_t = \tilde{Y}_{N_t}.$$

Proposition 4.76. *Let* $\mathcal{F}_t = \sigma\{(\tilde{X}_s, N_s); 0 \le s \le t\}$*, then* $\tilde{X}$ *is an* $\mathcal{F}_t$*-Markov process.*

Proof. We begin with a claim.
Claim. $E[f(\tilde{Y}_{k+N_t})|\mathcal{F}_t] = T^k f(\tilde{X}_t)$.

Fix $A \in \mathcal{F}_t^N$ and $B \in \mathcal{F}_m^{\tilde{Y}}$. Then

$$\begin{aligned} E[f(\tilde{Y}_{k+N_t}); A \cap B \cap \{N_t = m\}] &= E[f(\tilde{Y}_{k+m}); A \cap B \cap \{N_t = m\}] = P(A \cap \{N_t = m\})E[f(\tilde{Y}_{k+m}); B] \\ &= P(A \cap \{N_t = m\})E[T^k f(\tilde{Y}_m); B] = E[T^k f(\tilde{X}_t); A \cap B \cap \{N_t = m\}] \end{aligned}$$

However, sets of the form $A \cap B \cap \{N_t = m\}$ are closed under finite intersection, and the claim follows from the Sierpinski class theorem.

Now, use the independence of the increments for a Poisson process, to obtain that

$$\begin{aligned} E[f(\tilde{X}_{t+s})|\mathcal{F}_t] &= E[f(\tilde{Y}_{N_{t+s}})|\mathcal{F}_t] = E[f(\tilde{Y}_{N_{t+s}-N_t+N_t})|\mathcal{F}_t] \\ &= \sum_{k=0}^{\infty} e^{-\lambda_{max}s}\frac{(\lambda_{max}s)^k}{k!}E[f(\tilde{Y}_{k+N_t})|\mathcal{F}_t] = \sum_{k=0}^{\infty} e^{-\lambda_{max}s}\frac{(\lambda_{max}s)^k}{k!}T^k f(\tilde{X}_t) = T(s)f(\tilde{X}_t) \end{aligned}$$

Because $G = \tilde{G}$, X is a Markov process whose finite dimensional distributions agree with $\tilde{X}$.

Thus, all of the examples of generators with bounded rates are generators for a semigroup T for a Markov process on $C_0(S)$. □

Exercise 4.77. *For a two state Markov chain on $S = \{0, 1\}$ with generator*

$$A = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix},$$

show that

$$P(t, 0, \{0\}) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}.$$

4.6.3 Birth and Death Process

Example 4.78. *1. For a compound Poisson process,*

$$X_t = \sum_{n=1}^{N_t} Y_n,$$

$\lambda(x)$ is some constant λ, the rate for the Poisson process. If ν is the distribution of the Y_n, then

$$\mu(x, B) = \nu(B - x).$$

2. *A process is called a* birth and death process *if $S = N$,*

$$g(x, y) = \begin{cases} \lambda_x & y = x+1 \\ -(\lambda_x + \mu_x) & y = x \\ \mu_x & y = x-1 \end{cases}$$

3. *A* pure birth process *has $\mu_x = 0$ for all x.*

4. *A* pure death process *has $\lambda_x = 0$ for all x.*

5. *A* simple birth and death process *has $\mu_x = \mu$ and $\lambda_x = \lambda$ for all x.*

6. *A* linear birth and death process *has $\mu_x = x\mu$ and $\lambda_x = x\lambda$ for all x.*

7. *A* linear birth and death process with immigration *has $\mu_x = x\mu$ and $\lambda_x = x\lambda + a$ for all x.*

8. *$M/M/1$* queue *$S = \{0, 1, \cdots, N\}$, $\lambda_x = \lambda$, $x < N$, $\lambda_n = 0$, $\mu_x = \mu$. So customers arrive at an exponential rate, parameter λ until the queue reaches length N at which point, all arrivals are turned away. When they reach the front of the queue, they are served immediately. The service time is exponential, parameter μ.*

9. $M/M/\infty$ queue $S = \mathbb{N}$ $\lambda_x = \lambda$, $\lambda_n = 0$, $\mu_x = x\mu$. *So customers arrive at an exponential rate, parameter λ. All arrivals are served immediately. The service time is exponential, parameter μ.*

10. *A* logistic process $S = \{0, 1, \cdots, N\}$, $\lambda_x = rx(N-x)$, $\mu_x = x\mu$. *This can be used to model an epidemic. The susceptibility is of the $N-x$ non-infected individuals is proportional to the number x of infected individuals. An infected stays in this state for an exponential rate of time, parameter μ.*

11. *An* epidemic model with immunity *can be constructed by having $S = \{0, 1, \ldots, N\}^2$. The state (x, y) gives the number of infected and the number of immune. The generator*

$$Gf(x,y) = rx(N-x-y)(f(x+1,y)-f(x,y)) + \mu x(f(x-1,y+1)-f(x,y)) + \gamma y(f(x,y-1)-f(x,y)).$$

12. *For the* Moran model with mutation $S = \{0, 1, \cdots, N\}$, *consider a genetic model with two alleles, A_1 and A_2 X_t - number of alleles of type A_1 at time t. The process remains at a site for an exponential length of time, parameter λ, an individual is chosen at random to be replaced. A second individual is chosen at random to duplicate. A mutation may also occur at birth. The mutation $A_1 \to A_2$ has probably κ_1 and the mutation $A_2 \to A_1$ has probably κ_2 The generator*

$$\begin{aligned} Gf(x) &= \lambda(1-\frac{x}{N})\left(\frac{x}{N}(1-\kappa_1) + (1-\frac{x}{N})\kappa_2\right)(f(x+1)-f(x)) \\ &\quad + \lambda\frac{x}{N}\left((1-\frac{x}{N})(1-\kappa_2) + \frac{x}{N}\kappa_1\right)(f(x-1)-f(x)) \end{aligned}$$

13. *For a* random walk on the lattice $S = \mathbb{Z}^d$, *X_t - position of the particle at time t. The particle at a site x remains at x for an exponential length of time, parameter $\lambda(x)$, and then moves according to the kernel $p(x, y)$.*

    Thus,

$$\sum_y p(x,y) = 1$$

    and the generator

$$Gf(x) = \lambda(x)\sum_y p(x,y)(f(y)-f(x)).$$

Exercise 4.79. *For the compound Poisson process, if $\{Y_n; n \geq 1\}$ is a Bernoulli sequence, the X is a Poisson process with parameter λp*

To find harmonic functions $\tilde{h}$ for birth and death processes, note that

$$\lambda_x(\tilde{h}(x+1)-\tilde{h}(x)) + \mu_x(\tilde{h}(x-1)-\tilde{h}(x)) = 0.$$

$$(\tilde{h}(x+1)-\tilde{h}(x)) = \frac{\mu_x}{\lambda_x}(\tilde{h}(x)-\tilde{h}(x-1)), \ \lambda_x > 0.$$

Summing this on x, we obtain

$$(\tilde{h}(x+1)-\tilde{h}(x)) = \prod_{\tilde{x}=1}^{x}\frac{\mu_{\tilde{x}}}{\lambda_{\tilde{x}}}(\tilde{h}(1)-\tilde{h}(0)).$$

Assume that $\tilde{h}(0) = 0$ and $\tilde{h}(1) = 1$ and all of the $\lambda_x > 0$. Then any other non-constant harmonic function is a linear function of $\tilde{h}$. Then

$$\tilde{h}(y) = \sum_{x=0}^{y-1}(\tilde{h}(x+1) - \tilde{h}(x)) = \sum_{x=1}^{y-1}\prod_{\tilde{x}=1}^{x}\frac{\mu_{\tilde{x}}}{\lambda_{\tilde{x}}}.$$

Thus, for a simple birth and death process and for a linear birth and death process,

$$\tilde{h}(y) = \sum_{x=0}^{y-1}\left(\frac{\mu}{\lambda}\right)^x = \frac{1-(\mu/\lambda)^y}{1-(\mu/\lambda)}.$$

Let $D_0 = \{0\}$ and $D_1 = \{N\}$, then

$$h(x) = \frac{\tilde{h}(x)}{\tilde{h}(N)}$$

is a harmonic function satisfying

$$h(0) = 0, \quad h(N) = 1.$$

Thus, for $x \in \{0, \ldots, N\}$

$$h(x) = P_x\{\tau_0 > \tau_N\}$$

Exercise 4.80. *For a birth and death process, let $\tau = \min\{\tau_0, \tau_N\}$, for $x \in \{0, \ldots, N\}$.*

1. *Find $E_x\tau$.*
2. *Find $E_x e^{-\lambda\tau}$, $\lambda > 0$.*
3. *Consider what happens as $N \to \infty$.*

Exercise 4.81.
1. *For a simple birth and death process, X, find E_xX_t and $Var_x(X_t)$.*
2. *For a linear birth and death process, X, find E_xX_t and $Var_x(X_t)$*

4.6.4 Examples of Interacting Particle Systems

An *interacting particle aystems* is a Markov process $\{\eta_t; t \geq 0\}$ whose state space S is the configurations on a regular lattice Λ, i.e, the state is a mapping

$$\eta : \Lambda \to F$$

where F is a finite set. Examples of lattices are $\mathbb{Z}^d$, $\mathbb{Z}/M\mathbb{Z})^d$, hexagonal lattice, and regular trees. Examples of F are
{0,1} = {vacant, occupied}
{-1,1}= {spin up, spin down}
{0,1,$\cdots$,k} = {vacant, species 1, species 2, $\cdots$, species k}
{Δ, 0, $\cdots$, N} = {vacant, 0 individuals having allele A, $\cdots$, N individual having allele A}

Example 4.82 (interacting particle systems). *1.* Exclusion Process *(Spitzer, 1970). $F = \{0,1\}$, $\Lambda = \mathbb{Z}^d$.*

If initially we have one particle at x_0, i.e.,

$$\eta_0(x) = \begin{cases} 1, & \text{if} \quad x = x_0, \\ 0, & \text{if} \quad x \neq x_0. \end{cases}$$

Set $X_t = \eta_t^{-1}(\{1\})$. Then $\{X_t; t \geq 0\}$ is the random walk on the lattice.

When there are many particles, then each executes an independent random walk on the lattice excluding *transitions that take a particle to an occupied site. The generator*

$$\mathcal{G}F(\eta) = \sum_{x,y} \lambda(\eta, x)p(x,y)\eta(x)(1-\eta(y))(F(\eta^{xy}) - F(\eta))$$

where

$$\eta^{xy}(z) = \begin{cases} \eta(x), & \text{if} \quad z = y, \\ \eta(y), & \text{if} \quad z = x, \\ \eta(z), & \text{otherwise.} \end{cases}$$

Let's read the generator:

- *Particle at x_0 moves in time Δt with probability $\lambda(\eta, x)\Delta t + o(\Delta t)$*
- *Particle at x_0 move to y with probability $p(x_0, y)$.*
- *If site y is occupied, nothing happens.*

2. Coalescing Random Walk. *$F = \{0,1\}$, $\Lambda = \mathbb{Z}^d$.*

Each particle executes an independent random walk on the lattice until the walk takes a particle to an occupied site, then this particle disappears. The generator

$$\mathcal{G}F(\eta) = \sum_{x,y} \lambda(\eta, x)p(x,y)\eta(x)(1-\eta(y))(F(\eta^{xy}) - F(\eta)) + \eta(y)(F(\eta_x) - F(\eta))$$

where

$$\eta_x(z) = \begin{cases} 1 - \eta(x), & \text{if} \quad z = x, \\ \eta(z), & \text{otherwise.} \end{cases}$$

3. Voter Model *(Clifford and Sudbury, 1973, Holley and Liggett, 1975). $F = \{0,1\}$=$\{$no, yes$\}$, $\Lambda = \mathbb{Z}^d$.*

An individual changes opinion at a rate proportional to the number of neighbors that disagrees with the individual's present opinion. The generator

$$\mathcal{G}F(\eta) = \sum_{x,y} \lambda(\eta, x)(F(\eta_x) - F(\eta))$$

where

$$\lambda(\eta, x) = \frac{\lambda}{2d} \sum_{\{y; |y-x|=1\}} I_{\{\eta(y) \neq \eta(x)\}}.$$

4. Contact process *(Harris, 1974). $F = \{0,1\}$ =\{uninfected, infected\}, $\Lambda = \mathbb{Z}^d$.*

The generator

$$\mathcal{G}F(\eta) = \sum_{x,y} \lambda(\eta,x)(F(\eta_x) - F(\eta))$$

where

$$\lambda(\eta,x) = \begin{cases} \lambda \sum_{\{y;|y-x|=1\}} \eta(y), & \textit{if} \quad \eta(x) = 0, \\ 1, & \textit{if} \quad \eta(x) = 1. \end{cases}$$

Thus, λ is the relative infection rate.

5. For the stochastic Ising model*, $F = \{-1,1\}$. $\Lambda = \mathbb{Z}^d$.*

$$\mathcal{G}F(\eta) = \sum_{x,y} \exp\left(-\beta \sum_{y:|x-y|=1} \eta(x)\eta(y)\right)(F(\bar\eta_x) - F(\eta))$$

where

$$\bar\eta_x(z) = \begin{cases} -\eta(x), & \textit{if} \quad z = x, \\ \eta(z), & \textit{otherwise.} \end{cases}$$

4.7 Sample Path Regularity

The goal of this section is to show that for S separable, every Feller semigroup on $C_0(S)$ corresponds to a Markov process with sample paths in $D_S[0,\infty)$. In addition, we show that every generator A of a positive strongly continuous contraction semigroup on $C_0(S)$ corresponds to a Markov process.

We have seen that a jump Markov process with bounded jump rates $\lambda : S \to [0,\infty]$, corresponds to a Feller process X on S. Moreover, this process has a $D_S[0,\infty)$ version.

If the rates are bounded on compact sets, then because S is locally compact, we can write, for S separable,

$$S = \bigcup_{n=1}^{\infty} K_n$$

for an increasing sequence of compact sets $\{K_n; n \geq 1\}$ and define

$$\sigma_n = \inf\{t > 0; X_t \notin K_n\}.$$

Then, $\{\sigma_n; n \geq 1\}$ is an increasing sequence of stopping times. Define

$$\zeta = \lim_{n\to\infty} \sigma_n.$$

If this limit is finite, then we say that an *explosion* occurs. In these circumstances, we have that the transition function fails to satisfy $P(t,x,S) = 1$ for all $t > 0$, and the semigroup T is no longer conservative.

4.7.1 Compactifying the State Space

To include this situation in the discussion of Markov processes, the state space S has adjoined to it a special state Δ and we write $S^\Delta = S \cup \{\Delta\}$ and let

$$X_t = \Delta \quad \text{for } t \geq \zeta.$$

If S is not compact, then S^Δ is the one point compactification of S.

This suggest the following modification.

Theorem 4.83. *Let S be locally compact and separable and let T be a strongly continuous positive contraction semigroup on $C_0(S)$. For each $t \geq 0$, define the operator $T^\Delta(t)$ on $C(S^\Delta)$ by*

$$T^\Delta(t)f = f(\Delta) + T(t)(f - f(\Delta)).$$

Then T^Δ is a Feller semigroup on $C(S^\Delta)$.

Proof. The verification that T^Δ is a strongly continuous conservative semigroup on $C(S^\Delta)$ is straightforward.

Claim 1. T^Δ is positive. In other words, given $c \in \mathbb{R}^+$ and $f \in C_0(S)$, $c+f \geq 0$ implies that $c+T(t)f \geq 0$.

Because T is positive,

$$T(t)(f^+) \geq 0 \quad \text{and} \quad T(t)(f^-) \geq 0.$$

Hence,

$$-T(t)f \leq T(t)(f^-) \quad \text{and so} \quad (T(t)f)^- \leq T(t)(f^-).$$

Because $T(t)$ is a contraction, we obtain

$$||T(t)(f^-)|| \leq ||f^-|| \leq c.$$

Therefore,

$$T(t)(f^-)(x) \leq c \quad \text{and so} \quad c + T(t)f \geq 0.$$

Claim 2. T^Δ is a contraction.

Because T^Δ is positive,

$$|T^\Delta(t)f(x)| \leq T^\Delta(t)||f|| = ||f||$$

and $||T^\Delta(t)|| \leq 1$. ∎

With this modification, the condition of being a Feller semigroup states that the distribution of the Markov process at time t depends continuously on the initial state. In other words, for $f \in C(S^\Delta)$,

$$\lim_{x \to x_0} T(t)f(x) = T(t)f(x_0),$$

$$\lim_{x\to x_0} E_x f(X_t) = E_{x_0} f(X_t),$$

$$\lim_{x\to x_0} \int_{S^\Delta} f(y)P(t,x,dy) = \int_{S^\Delta} f(y)P(t,x_0,dy).$$

Thus $P(t,x,\cdot)$ converges in distribution to $P(t,x_0,\cdot)$.

To check for explosions, consider a pure birth process. For $\tau_x = \inf\{t \geq 0 : X_t = x\}$, we have, if $X_0 = 0$,

$$\tau_x = \sigma_1 + \cdots + \sigma_x$$

where $\sigma_{\tilde{x}}$ is the length of time between the arrival of X to $\tilde{x}-1$ and $\tilde{x}$.

Then, the random variables $\{\sigma_{\tilde{x}}; \tilde{x} \geq 1\}$ are independent exponential random variables with parameters $\{\lambda_{\tilde{x}-1}; \tilde{x} \geq 1\}$. Thus, the Laplace transform

$$E[e^{-u\sigma_{\tilde{x}}}] = \frac{\lambda_{\tilde{x}-1}}{u + \lambda_{\tilde{x}-1}},$$

and

$$E[e^{-u\tau_x}] = \prod_{\tilde{x}=1}^{x} \frac{\lambda_{\tilde{x}-1}}{\lambda_{\tilde{x}-1}+u} = \left(\prod_{\tilde{x}=1}^{x} (1 + \frac{u}{\lambda_{\tilde{x}-1}})\right)^{-1}.$$

For $u > 0$, the infinite product,

$$\prod_{\tilde{x}=1}^{\infty} (1 + \frac{u}{\lambda_{\tilde{x}-1}}),$$

converges if and only if

$$\sum_{\tilde{x}=1}^{\infty} \frac{1}{\lambda_{\tilde{x}-1}} < \infty.$$

Thus, if the sum is infinite, then we have no explosion.

In particular, a linear birth process has no explosion. However, if, for example,

$$\liminf_{x\to\infty} \frac{\lambda_x}{x^p} > 0$$

for some $p > 1$, then the birth process has an almost surely finite explosion time.

For a birth and death process, X, we consider the pure birth process, $\tilde{X}$ with the same birth rates as X. Couple the two processes together in the sense that the exponential time for a jump forward is the same for the first time each process reaches any site x. In this way,

$$\tilde{X}_t \leq X_t$$

for all $t \geq 0$. Thus, if X has no explosion, neither does $\tilde{X}$.

We can use the previous section to show that the Yosida approximation generates a Markov process. Write $A_\lambda = \lambda(\lambda R(\lambda) - I)$ We know that $\lambda R(\lambda)$ is a positive linear operator on $C_0(S)$. By the Riesz representation theorem, there exists a positive Borel measure

$$\lambda R(\lambda) f(x) = \int_S f(y)\ \mu_\lambda(x, dy).$$

Thus, for $f \in \mathcal{D}(G)$

$$\lambda f(x) = \int_S (\lambda I - G)f(y)\ \mu_\lambda(x, dy).$$

Because $||\lambda R(\lambda)|| \leq 1$, we have that $\mu_\lambda(x, S) \leq 1$. However, if G is conservative, we can choose $f(x) = 1/\lambda$ to obtain

$$1 = \mu_\lambda(x, S).$$

Note that if we have that G generates a Markov process, then

$$\lambda R(\lambda)f(x) = \int_S \lambda e^{-\lambda t} E_x f(X_t)\ dt = E_x f(X_{\tau_\lambda})$$

where τ_λ is an exponential random variable, parameter λ, and independent from X. Thus, $\mu_\lambda(x, \cdot)$ is the distribution of X_{τ_λ} under P_x.

Exercise 4.84. *let D be a dense subset of $C_0(S)$. Then a function $x \in D_S[0, \infty)$ if and only if $f(x) \in D_\mathbb{R}[0, \infty)$ for every $f \in D$.*

Theorem 4.85. *Let X be a $\mathcal{F}^X_{t+}$-Feller process, then X has a $D_S[0, \infty)$-valued modification.*

Proof. Let X have generator G and let $f \in \mathcal{D}(G)$, then the martingale

$$M^f_t = f(X_t) - \int_0^t Gf(X_s)\ ds$$

is right continuous on a set Ω_f of probability 1. Thus

$$f(X_t) = M^f_t - \int_0^t Gf(X_s)\ ds$$

is right continuous on Ω_f. Now take a countably dense set $D \subset \mathcal{D}(A)$ of f. Then

$$\bigcap_{f \in D} \Omega_f \subset \{X \in D_S[0, \infty)\}.$$

□

Theorem 4.86. *Let S be locally compact and separable. Let T be a positive strongly continuous contraction semigroup on $C_0(S)$ and define T^Δ as above. Let X be a Markov process corresponding to T^Δ with sample paths in $D_{S^\Delta}[0, \infty)$ and let*

$$\zeta = \inf\{t \geq 0 : X_t = \Delta \text{ or } X_{t-} = \Delta\}.$$

Then on the set $\{\zeta < \infty\}$,

$$X_{\zeta+s} = \Delta \text{ for all } s \geq 0.$$

Proof. By the Urysohn lemma, there exists $f \in C(S^\Delta)$ such that $f > 0$ on S and $f(\Delta) = 0$. Then the resolvent $R^\Delta_1 f > 0$ on S and $R^\Delta_1 f(\Delta) = 0$. Note that $e^{-t} R^\Delta_1 f(X_t)$ is a supermartingale. Because a right continuous nonnegative supermartingale is zero after its first contact with 0, the theorem follows. □

Proposition 4.87. *If, in addition to the conditions above, if the extention the semigroup T to the bounded measurable functions is conservative, then $\alpha(S) = P\{X_0 \in S\} = 1$ implies that $P_\alpha\{X \in D_S[0,\infty) = 1\}$.*

Proof. Let G^Δ be the generator of T^Δ, then $G^\Delta I_S = 0$, thus

$$E_\alpha I_S(X_t) = E_\alpha I_S(X_0) + E_\alpha[\int_0^t G^\Delta I_S(X_s)\ ds] = E_\alpha I_S(X_0) + 0.$$

or

$$P_\alpha\{\zeta > t\} = P_\alpha\{X_t \in S\} = P_\alpha\{X_0 \in S\} = 1.$$

□

Theorem 4.88. *Let S be locally compact and separable. Let T be a Feller semigroup on $C_0(S)$. Then for each probability measure α on S, there exists a Markov process X with initial distribution α and sample paths in $D_S[0,\infty)$. Moreover, X is strong Markov with respect to the filtration $\{\mathcal{F}^X_{t+}; t \geq 0\}$.*

Proof. Let G be the generator for T and let G_λ, $\lambda > 0$, be its Yosida approximation. Then T_λ is the semigroup for a jump Markov process X^λ. Moreover, for every $t \geq 0$ and $f \in C_0(S)$,

$$\lim_{\lambda\to\infty} T_\lambda(t)f = T(t)f.$$

Let T^Δ_λ and T^Δ be the corresponding semigroups on $C(S^\Delta)$. Then, for each $f \in C(S^\Delta)$,

$$\begin{aligned}\lim_{\lambda\to\infty} E[f(X^\lambda_t)] &= \lim_{\lambda\to\infty} E[T^\Delta_\lambda(t)f(X^\lambda_0)] \\ &= \lim_{\lambda\to\infty} \int T^\Delta_\lambda(t)f(x)\ \alpha(dx) = \int T^\Delta(t)f(x)\ \alpha(dx).\end{aligned}$$

Now $f \to \int T^\Delta(t)f(x)\ \alpha(dx)$ is a positive linear operator that maps the constant function 1 to the value 1, and thus, by the Riesz representation theorem is integration against a probability measure ν_t. This gives the one dimensional distributions of X. The proof that the finite dimensional distributions converge follows from a straightforward induction argument.

To check the strong Markov property, let τ be a discrete $\mathcal{F}^X_{t+}$-stopping time, then for each $\epsilon > 0$, and $A \in \mathcal{F}^X_\tau$, $A \cap \{\tau = t\} \in \mathcal{F}_{t+\epsilon}$. Thus, for $s > 0$,

$$\begin{aligned}E[f(X_{\tau+s}); A \cap \{\tau = t\}] &= E[f(X_{t+s}); A \cap \{\tau = t\}] \\ &= E[T(s-\epsilon)f(X_{t+\epsilon}); A \cap \{\tau = t\}] \\ &= E[T(s-\epsilon)f(X_{\tau+\epsilon}); A \cap \{\tau = t\}].\end{aligned}$$

Now use the right continuity of X and the strong continuity of T to conclude that

$$E[f(X_{\tau+s}); A \cap \{\tau = t\}] = E[T(s)f(X_\tau); A \cap \{\tau = t\}]$$

Summing over the range of τ we obtain

$$E[f(X_{\tau+s})|\mathcal{F}^X_{\tau+}] = T(s)f(X_\tau).$$

We can obtain the result for an arbitrary stopping time by realizing it as the decreasing sequence of discrete stopping times, use the identity above, the right continuity of X, and the strong continuity of T. □

Exercise 4.89. *Complete the proof above by supplying the induction argument to show that the finite dimensional distributions converge.*

In summary, we have

Theorem 4.90. *Let S be locally compact and separable. Let G be a linear operator on $C_0(S)$ satisfying 1-3 of the Hille-Yosida theorem and let T be a strongly continuous postive contraction semigroup generated by $\bar{G}$. Then for each $x \in S$, there exists a Markov process X corresponding to T with initial distribution δ_x and with sample paths in $D_S[0,\infty)$ if and only if G is conservative.*

4.8 Transformations of Markov Processes

Definition 4.91. *A continuous* additive functional *A of a Markov process X is adapted process satisfying*

1. *$A_0 = 0$.*
2. *A is continuous and nondecreasing.*
3. *For all $s, t \geq 0$, $A_{s+t} = A_s + A_t \circ \theta^s$.*
4. *A is constant on $[\zeta, \infty)$, $\zeta = \inf\{t > 0; X_t = \Delta\}$.*

The example that will occupy our attention in this section is $A_t = \int_0^t a(X_u)\ du$ for some nonnegative measureable function a. We shall use these additive functionals in two settings: random time changes, and killing.

Example 4.92. *If N is one of the queues given in the examples, then*

$$\int_0^t N_s\ ds$$

is the total amount of service time for the queue up to time t.

$$\int_0^t I_{\{0\}}(N_s)\ ds$$

is the amount of time that the queue is empty.

4.8.1 Random Time Changes

Let Y be a process with sample paths in $D_S[0,\infty)$ and choose c be a nonnegative continuous function on S so that $c \circ Y$ is bounded on bounded time intervals. Think of c as the rate that the clock for the process moves. Thus, for example, if $c(x) > 1$, the process moves more quickly through the state x

The goal is to describe solutions to

$$X_t = Y_{\int_0^t c(X_s)\ ds}$$

If X solves this equation, then set

$$\tau_t = \int_0^t c(X_s)\ ds = \int_0^t c(Y_{\tau_s})\ ds.$$

Set $\sigma_0 = \inf\{t > 0 : \int_0^t 1/c(X_{\tau_s})\ ds = \infty\}$, then for $t < \sigma_0$,

$$t = \int_0^t \frac{c(X_s)}{c(X_s)}\ ds = \int_0^{\tau_t} \frac{1}{c(Y_u)}\ du$$

and therefore, τ satisfies the ordinary differential equation

$$\tau_t' = c(Y_{\tau_t}), \quad \tau_0 = 0.$$

Because $c(Y)$ is bounded on bounded time intervals, the solution to the ordinary differential equation has a unique solution. Now,

$$X_t = Y_{\tau_t}.$$

Theorem 4.93. *Let Y be a $D_S[0,\infty)$-valued Markov process with generator G on $C_0(S)$. If a unique solution to*

$$X_t = Y_{\int_0^t c(X_s)\ ds}$$

exists, then X is a Markov process with generator cG.

Proof. Verifying that X is a Markov process is left as an exercise.

Note that $\{\tau_s \le t\} = \{\int_0^t 1/c(X_{\tau_u})\ du) \ge s\} \cap \{\sigma_0 \le t\} \in \mathcal{F}^Y_{t+}$. Thus, τ_s is an $\mathcal{F}_{t+}$-stopping time. For $f \in \mathcal{D}(A)$, the optional sampling theorem guarantees that

$$f(Y_{\tau_t}) - \int_0^{\tau_t} Gf(Y_s)\ ds = f(X_t) - \int_0^t c(X_u)Gf(X_u)\ du$$

is an $\mathcal{F}^Y_{\tau_t}$-martingale.

Thus, f in in the domain of the generator of the Markov process X and this generator is cA. □

Exercise 4.94. *Verify that Y in the theorem above is a Marlov process.*

4.8.2 Killing

We can choose to have a process X move to the state Δ. The rate that this happens can be spatially dependent. Thus, we define the hazard function

$$k(x) = \lim_{h\to 0} \frac{1}{h} P_x\{X_h = \Delta\}.$$

The analysis follows the previous analysis for the case k constant. Write the additive functional

$$K_t = \int_0^t k(X_s)\ ds$$

and define

$$Y = \int_0^\infty e^{-K_s} f(X_s)\ ds, \quad R_k f(x) = E_x Y.$$

Then

$$Y = \int_0^t e^{-K_s} f(X_s)\, ds + e^{-K_t} Y \circ \theta^t$$

Again, we have Doob's martingale

$$\begin{aligned} Z_t^{k,f} = E[Y|\mathcal{F}_t^X] &= \int_0^t e^{-K_s} f(X_s)\, ds + e^{-K_t} E[Y \circ \theta^t|\mathcal{F}_t^X] \\ &= \int_0^t e^{-K_s} f(X_s)\, ds + e^{-K_t} R_k f(X_t) \end{aligned}$$

Exercise 4.95. *Show that for $f \in \mathcal{D}(G)$, $R_k G f = R_k k f - f$ or*

$$R_k(k - G)f = f.$$

Now, let $f = (k - G)g$, then the Doob's martingale above becomes

$$C_t^{k,g} = e^{-K_t} g(X_t) + \int_0^t e^{-K_s}(k - G)g(X_s)\, ds$$

and the analog to Dynkin's formula becomes

$$g(x) = E_x[e^{-K_\tau} g(X_\tau)] + E_x[\int_0^\tau e^{-K_s}(k - G)g(X_s)\, ds.$$

Let g satisfy $Gg = kg$ with $g = h$ on D, then

$$g(x) = E_x[e^{-K_{\tau_D}} h(X_{\tau_D})].$$

Now define the stochastic process $\tilde{X}$ as follows:

Let ξ be an exponential random variable, parameter 1, independent of the process X. Then

$$\tilde{X}_t = \begin{cases} X_t & \text{if } K_t < \xi, \\ \Delta & \text{if } K_t \geq \xi. \end{cases}$$

Then, for $f \in C_0(S)$,

$$E_x f(\tilde{X}_t) = E_x[f(X_t) I_{\{K_t < \xi\}}] = E_x[f(X_t) E_x[I_{\{K_t < \xi\}}|\mathcal{F}_t^X]] = E_x[e^{-K_t} f(X_t)].$$

Exercise 4.96. *For $f \in C_0(S)$, define*

$$T(t)f(x) = E_x[e^{-K_t} f(X_t)].$$

Then T is a positive continuous constraction semigroup with generator $G - k$. Thus, $\tilde{X}$ is Markov process with state space S^Δ and generator $G - k$.

Corollary 4.97. *(Feynman-Kac formula) Let X be a Feller process on S and generator G. Then the equation*

$$\frac{\partial u}{\partial t} = Gu - ku, \quad u(0,x) = f(x)$$

has solution

$$u(t,x) = E_x[e^{-K_t} f(X_t)].$$

4.8.3 Change of Measure

Proposition 4.98 (Bayes formula)**.** *On the probability space $(\Omega, \mathcal{F}, P)$, let L be a non-negative random variable satisfying $EL = 1$. Define the probability measure $Q(L) = E[L; A]$ and write $E_Q[Y] = E[YL]$. Then for any integrable random variable Z, and any σ-algebra $\mathcal{G} \subset \mathcal{F}$,*

$$E_Q[Z|\mathcal{G}] = \frac{E[ZL|\mathcal{G}]}{E[L|\mathcal{G}]}.$$

Proof. Clearly the right hand side is $\mathcal{G}$-measurable. Choose $A \in \mathcal{G}$, then

$$\begin{aligned} E_Q\left[\frac{E[ZL|\mathcal{G}]}{E[L|\mathcal{G}]}; A\right] &= E\left[\frac{E[ZL|\mathcal{G}]}{E[L|\mathcal{G}]} L; A\right] = E\left[E\left[\frac{E[ZL|\mathcal{G}]}{E[L|\mathcal{G}]} L|\mathcal{G}\right]; A\right] \\ &= E\left[\frac{E[ZL|\mathcal{G}]}{E[L|\mathcal{G}]} E\left[L|\mathcal{G}\right]; A\right] = E\left[E[ZL|\mathcal{G}]; A\right] \\ &= E[ZL; A] = E_Q Z. \end{aligned}$$

□

Lemma 4.99. *Assume $Q|_{\mathcal{F}_t} << P|_{\mathcal{F}_t}$, and assume that L_t is the Radon-Nikodym derivative, then Z is a Q-martingale if and only if ZL is a P-martingale.*

Proof. Because L is an adapted process, Z is adapted if and only if ZL is adapted. To show that L is a martingale, choose $A \in \mathcal{F}_t$, then $Q(A) = E[L_t; A]$. However, $A \in \mathcal{F}_{t+s}$ and therefore

$$E[L_t; A] = Q(A) = E[L_{t+s}; A].$$

By Bayes formula,

$$E_Q[Z_{t+s}|\mathcal{F}_t] = \frac{E[Z_{t+s}L_{t+s}|\mathcal{F}_t]}{E[L_{t+s}|\mathcal{F}_t} = \frac{E[Z_{t+s}L_{t+s}|\mathcal{F}_t]}{L_t}.$$

Thus,

$$E_Q[Z_{t+s}|\mathcal{F}_t] - Z_t = \frac{E[Z_{t+s}L_{t+s}|\mathcal{F}_t] - Z_t L_t}{L_t}.$$

and, consequently,

$$E_Q[Z_{t+s}|\mathcal{F}_t] = Z_t \quad \text{if and only if} \quad E[Z_{t+s}L_{t+s}|\mathcal{F}_t] - Z_t L_t.$$

□

Note that L is a positive mean one martingale. In addition, is L^{-1} is Q-integrable, the L^{-1} is a Q-martingale and $P|_{\mathcal{F}_t} << Q|_{\mathcal{F}_t}$

For X, a time homogeneous Markov process with generator G, we write $g = Gf$ for $f \in \mathcal{D}(G)$, with $\inf_x f(x) > 0$, recall that

$$\frac{f(X_t)}{f(X_0)} \exp\left(-\int \frac{g(X_v)}{f(X_v)}\ dv\right)$$

is a mean 1 local martingale.. Write $f(x) = \exp w(x)$ and $\mathcal{G}w = e^{-w}\ Ge^w$, then this local martingale becomes

$$L_t = \exp\left(w(X_t) - w(X_0) - \int_0^t \mathcal{G}w(X_v)\ dv\right).$$

For the case of standard Brownian motion B, take $w(x) = \mu x$. We shall soon learn that $G = \frac{1}{2}\Delta$, and

$$L_t = \exp(\mu B_t - \frac{1}{2}\mu^2 t)$$

is indeed a martingale.

Under the change of measure induced by L

$$\begin{aligned}
Q\{B_t \in A\} &= E[I_A(B_t)\exp(\mu B_t - \frac{1}{2}\mu^2 t)] = \frac{1}{\sqrt{2\pi t}}\int_A \exp(\mu x - \frac{1}{2}\mu^2 t)e^{-x^2/2t}\ dx \\
&= \frac{1}{\sqrt{2\pi t}}\int_A \exp -\frac{1}{2t}(x^2 - 2\mu x t - \mu^2 t^2)\ dx = \frac{1}{\sqrt{2\pi t}}\int_A \exp -\frac{1}{2t}(x - \mu t)^2\ dx \\
&= \frac{1}{\sqrt{2\pi t}}\int_{A-\mu t} e^{-x^2/2t}\ dx = P\{B_t + \mu t \in A\}
\end{aligned}$$

Exercise 4.100. *Let B, P and Q be defined as above. For times $0 \le t_1 < t_2 \cdots t_n$, and Borel sets $A_1, \ldots, A_n$,*

$$Q\{B_{t_1} \in A_1, \ldots B_{t_n} \in A_n\} = P\{B_{t_1} + \mu t_1 \in A_1, \ldots B_{t_n} + \mu t_n \in A_n\}.$$

Thus, by the Daniell-Kolmogorov extension theorem, under the measure Q, B is Brownian motion with drift μ.

Example 4.101. *Recall that for standard Brownian motion, and for $a \ge 0$, the hitting time for level a, $\tau_a = \inf\{t > 0; B_t > a\}$, has density*

$$f_{\tau_a}(s) = \frac{a}{\sqrt{2\pi t^3}}e^{-a^2/2s}.$$

If we consider the process $B_t + \mu t, \mu > 0$ under the measure Q, then this process is standard Brownian motion under the measure P. Thus,

$$\begin{aligned}
Q\{\tau_a \le t\} &= E[I_{\{\tau_a \le t\}}L_t] = E[E[I_{\{\tau_a \le t\}}L_t|\mathcal{F}_{\min\{\tau_a,t\}}]] = E[I_{\{\tau_a \le t\}}L_{\min\{\tau_a,t\}}] = E[I_{\{\tau_a \le t\}}L_\tau] \\
&= E[I_{\{\tau_a \le t\}}\exp(\mu a - \frac{1}{2}\mu^2\tau_a)] = \int_0^t \exp(\mu a - \frac{1}{2}\mu^2 s)\frac{a}{\sqrt{2\pi s^3}}e^{-a^2/2s}\ ds
\end{aligned}$$

Thus,

$$f_{\tau_a,\mu}(s) = \frac{a}{\sqrt{2\pi s^3}}e^{-(a-\mu t)^2/2s}.$$

gives the density of τ_a for Brownian motion with constant drift.

4.9 Stationary Distributions

Definition 4.102. *Call a stochastic process* stationary *if, for every $n \geq 1$, and $0 \leq s_1 < \cdots < s_n$, the finite dimensional distributions*

$$P\{X_{t+s_1} \in B_1, \ldots, X_{t+s_n} \in B_n\}, \quad B_1, \ldots, B_n \in \mathcal{B}(S).$$

are independent of $t \geq 0$.

Definition 4.103. *A probabiity measure μ for a Markov process X is called* stationary *if X is a stationary process under P_μ.*

Lemma 4.104. *μ is a stationary measure for a Markov process X if and only if*

$$P_\mu\{X_t \in B\} = \mu(B), \quad B \in \mathcal{B}(S).$$

for all $t \geq 0$.

Proof. The necessity is immediate. To prove sufficency, note that for each $t \geq 0$, $\theta^t X$ is a Markov process with initial distribution μ and the same generator as X, thus has the same finite dimensional distributions as X. □

Theorem 4.105. *Let G generate a strongly continuous contraction semigroup T on $C_0(S)$ corresponding to a Markov process X. In addition, assume that D is a core for G. Then for a probability measure μ on S the following are equivalent:*

1. *μ is a stationary distribution for X.*
2. *$\int T(t)f\ d\mu = \int f\ d\mu$ for $f \in C_0(S)$ and $t \geq 0$.*
3. *$\int Gf\ d\mu = 0$ for $f \in D$.*

Proof. $(1 \to 2)$

$$\int T(t)f\ d\mu = E_\mu[T(t)f(X_0)] = E_\mu[f(X_t)] = E_\mu[f(X_0)] = \int f\ d\mu$$

$(2 \to 1)$ By the lemma above, it suffices to show that the one dimensional distributions agree.

$$E_\mu[f(X_t)] = E_\mu[T(t)f(X_0)] = \int T(t)f\ d\mu = \int f\ d\mu = E_\mu[f(X_0)].$$

$(2 \to 3)$ follows from the definition of G.
$(3 \to 2)$ By the definition of core, 3 holds for all $f \in \mathcal{D}(G)$.

$$\int (T(t)f - f)\ d\mu = \int \int_0^t GT(s)f\ dsd\mu = \int_0^t \int GT(s)f\ d\mu ds = 0.$$

Because $\mathcal{D}(G)$ is dense in $C_0(S)$, 2 follows. □

Exercise 4.106. *Let α be a probability measure on S*

1. *If $\int f\ d\nu = \lim_{t\to\infty} \int T(t)f\ d\alpha$ exists for every $f \in C_0(S)$, then ν is a stationary measure.*

2. *If* $\int f\ d\nu = \lim_{n\to\infty} \frac{1}{t_n}\int_0^{t_n}\int T(t)f\ d\alpha$ *exists for every* $f \in C_0(S)$*, and some increasing sequence* $\{t_n; n \geq 1\}$*,* $\lim_{n\to\infty} t_n = \infty$ *then* ν *is a stationary measure.*

Definition 4.107. *A Markov process* X *is said to be* ergodic *if there exist a unique stationary measure* ν *and*

$$\lim_{t\to\infty} E_\alpha f(X_t) = \int f\ d\nu.$$

for every initial distribution α *and every* $f \in C_0(S)$*.*

Note that if X is a stationary process, then we may extend it to be a process $\{X_t; t \in \mathbb{R}\}$. As with Markov chains, we have:

Definition 4.108. *A probability measure* ν *is said to be reversible for a Markov process* X *if*

$$E_\nu[f_1(X_0)f_2(X_t)] = E_\nu[f_2(X_0)f_1(X_t)],$$

or in terms of the semigroup T*,*

$$\int f_1T(t)f_2\ d\nu = \int f_1T(t)f_2\ d\nu.$$

In other words, $T(t)$ is a self adjoint operator with respect to the measure ν.

Proposition 4.109. *If* T *is conservative, then a reversible measure is invariant.*

Proof. Take $f_2 = 1$ in the definition. □

Theorem 4.110. *Suppose that* T *is a Feller semigroup on* $C_0(S)$ *with generator* G *and let* ν *be a probability measure on* S*. Then the following statements are equivalent:*

1. ν *is reversible.*
2. $\int f_1Gf_2\ d\mu = \int f_1Gf_2\ d\mu$ *for all* $f_1, f_2 \in \mathcal{D}(G)$*.*
3. $\int f_1Gf_2\ d\mu = \int f_1Gf_2\ d\mu$ *for all* $f_1, f_2 \in D$*, a core for* G*.*
4. ν *is stationary. Let* X *be the Markov process obtained using the* ν *as the distribution of* X_0 *and extending the time set to* $\mathbb{R}$*. Then* $\{X_t; t \in \mathbb{R}\}$ *and* $\{X_{-t}; t \in \mathbb{R}\}$ *have the same finite dimensional distributions.*

Proof. $(1 \to 2)$ is straightforward and $(2 \to 3)$ is immediate. $(3 \to 1)$ is nearly identitical to the proof of a similar statement for stationary measures.

$(4 \to 1)$

$$E_\nu[f_1(X_0)f_2(X_t)] = E_\nu[f_2(X_0)f_1(X_{-t})] = E_\nu[f_2(X_0)f_1(X_t)].$$

$(1 \to 4)$ Now suppose that ν is reversible. Then given

$$f_0 = 1, f_1, \ldots, f_n, f_{n+1} = 1 \in C_b(S)$$

and $t_0 < \cdots < t_n$, define

$$g_\ell(x) = E_x[\prod_{k=1}^{\ell-1} f_k(X_{t_\ell - t_k})], \quad \text{and} \quad h_\ell(x) = E_x[\prod_{k=\ell+1}^{n+1} f_k(X_{t_k - t_\ell})].$$

Set

$$I_\ell = \int g_\ell(x) f_\ell(x) h_\ell(x)\ \nu(dx).$$

Then

$$I_0 = E_\nu[\prod_{k=1}^n f_k(X_{t_k - t_0})], \quad \text{and} \quad I_{n+1} = E_\nu[\prod_{k=1}^n f_k(X_{t_{n+1}-t_k})].$$

Claim. $I_0 = I_{n+1}$.
Let $1 \le \ell \le n$, Then

$$\begin{aligned} I_\ell = \int g_\ell f_\ell h_\ell\ d\nu &= \int (g_\ell f_\ell) T(t_{\ell+1} - t_\ell)(f_{\ell+1} h_{\ell+1})\ d\nu \\ &= \int T(t_{\ell+1} - t_\ell)(g_\ell f_\ell)(f_{\ell+1} h_{\ell+1})\ d\nu \\ &= \int g_{\ell+1} f_{\ell+1} h_{\ell+1}\ d\nu = I_{\ell+1} \end{aligned}$$

because

$$\begin{aligned} T(t_{\ell+1} - t_\ell)(f_{\ell+1}(x) h_{\ell+1}(x)) &= T(t_{\ell+1} - t_\ell) \left(f_{\ell+1}(x) E_x[\prod_{k=\ell+2}^{n+1} f_k(X_{t_k - t_{\ell+1}})] \right) \\ &= T(t_{\ell+1} - t_\ell) E_x[\prod_{k=\ell+1}^{n+1} f_k(X_{t_k - t_{\ell+1}})] = E_x[\prod_{k=\ell+1}^{n+1} f_k(X_{t_k - t_\ell})] = h_\ell(x), \end{aligned}$$

and

$$\begin{aligned} T(t_{\ell+1} - t_\ell)(g_\ell(x) f_\ell(x)) &= T(t_{\ell+1} - t_\ell) \left(E_x[\prod_{k=1}^{\ell-1} f_k(X_{t_\ell - t_k})] f_\ell(x) \right) \\ &= T(t_{\ell+1} - t_\ell) E_x[\prod_{k=1}^{\ell} f_k(X_{t_\ell - t_k})] = g_{\ell+1}(x). \end{aligned}$$

Interate this to obtain the claim.

However, by stationarity,

$$I_0 = E_\nu[\prod_{k=1}^n f_k(X_{t_k})], \quad \text{and} \quad I_{n+1} = E_\nu[\prod_{k=1}^n f_k(X_{-t_k})].$$

and 4 follows. □

For a countable state space, the identity $\int f_1 G f_2\ d\nu = \int f_2 G f_1 d\nu$ becomes

$$\sum_x \sum_y f_1(x) g_{xy}(f_2(y) - f_2(x)) \nu\{x\} = \sum_x \sum_y f_2(x) g_{xy}(f_1(y) - f_1(x)) \nu\{x\}.$$

Simplifying, this is equivalent to

$$\sum_x \sum_y f_1(x) g_{xy} f_2(y) \nu\{x\} = \sum_x \sum_y f_2(x) g_{xy} f_1(y) \nu\{x\}$$

or by changing indices

$$\sum_x \sum_y f_1(x) g_{xy} f_2(y) \nu\{x\} = \sum_x \sum_y f_2(y) g_{yx} f_1(x) \nu\{y\}.$$

By writing f_1 and f_2 as sums of indicator functions of sites in S, we have that this statement is equivalent to

$$\nu\{x\} g_{xy} = \nu\{y\} g_{yx}.$$

This is the equation of *detailed balance.*

Exercise 4.111. *Find a reversible measure for a birth and death process and give conditions so that this measure can be normalized to be a probability measure. Simplify this expression for the examples of birth and death processes in the examples above.*

4.10 One Dimensional Diffusions.

Let $-\infty \le r_- < r_+ \le +\infty$ and let $I = [r_-, r_+] \cap \mathbb{R}$, $\text{int}(I) = (r_-, r_+)$, and $\bar{I} = [r_-, r_+]$.

Definition 4.112. *A* diffusion *on I is a Feller process with paths in $C_I[0,\infty)$. A diffusion is called* regular *if, for every $x \in int(I)$ and $y \in I$,*

$$P_x\{\tau_y < \infty\} > 0,$$

where $\tau_y = \inf\{t > 0 : X_t = y\}$.

Example 4.113. *Brownian motion is a one dimensional diffusion on $\mathbb{R}$. The semigroup is*

$$T(t)f(x) = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(y) \exp(-\frac{(x-y)^2}{2t})\ dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x + y\sqrt{t}) \exp(-\frac{y^2}{2})\ dy.$$

Then, for $f \in C^2(\mathbb{R}) \cap C_0(R)$,

$$\begin{aligned}
\frac{1}{t}(T(t)f(x) - f(x)) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{t}((f(x + y\sqrt{t}) - f(x)) \exp(-\frac{y^2}{2})\ dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{t}(y\sqrt{t} f'(x) + \frac{1}{2} y^2 t f''(x + \theta y\sqrt{t}) \exp(-\frac{y^2}{2})\ dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{2} y^2 f''(x + \theta y\sqrt{t}) \exp(-\frac{y^2}{2})\ dy
\end{aligned}$$

for some $\theta \in (0,1)$. Thus, the generator

$$Gf(x) = \frac{1}{2} f''(x).$$

Exercise 4.114. *For $\lambda > 0$, solve $g'' = 2\lambda g$ and use this to find the Laplace transform for exit times.*

Exercise 4.115. *Let $\phi \in C^2(\mathbb{R})$ be increasing and let B be standard Brownian motion. Find the generator of $X = \phi \circ B$.*

The major simplifying circumstance for one-dimensional diffusions is the intemediate value theorem, i.e., if x_0 is between X_s and X_t, $s < t$, then $X_c = x_0$ for some $c \in (s,t)$. Regularity implies a much stronger condition.

Theorem 4.116. *Let J be a finite open interval, $\bar{J} \subset I$ and let τ_J be the time of the first exit from J, then if X is a regular diffusion on I*

$$\sup_{x \in J} E_x \tau_J < \infty.$$

Proof. We begin with a claim.

Claim. Assume that $p_J = \sup_{x \in J} P_x\{\tau_J > t\} < 1$, then

$$E_x \tau_J \leq \frac{t}{1 - p_J}.$$

Write $p_n = \sup_{x \in J} P_x\{\tau_J > nt\}$. Then

$$P_x\{\tau_J > (n+1)t\} \leq p_n P_x\{\tau_J > (n+1)t | \tau_J > nt\} \leq p_n p_J.$$

by the strong Markov property. Thus, $p_n \leq p_J^n$ and

$$E_x \tau_J = \int_0^\infty P_x\{\tau_J > s\}\ ds = \sum_{n=0}^\infty \int_{nt}^{(n+1)t} P_x\{\tau_J > s\}\ ds \leq t \sum_{n=0}^\infty p_J^n = \frac{t}{1 - p_J}.$$

To complete the proof, set $J = (a,b)$ and pick $y \in J$. Then, by the regularity of X, there exists a $t > 0$ and $p < 1$ so that

$$P_y\{\tau_a > t\} \leq p, \quad P_y\{\tau_b > t\} \leq p.$$

Thus, for $x \in [y, b)$,

$$P_x\{\tau_J > t\} \leq P_x\{\tau_b > t\} \leq P_y\{\tau_b > t\}.$$

For $x \in (a, y]$,

$$P_x\{\tau_J > t\} \leq P_x\{\tau_a > t\} \leq P_y\{\tau_a > t\}.$$

□

To set notation, for any open interval $J = (a,b)$ with $P_x\{\tau_J < \infty\} = 1$ for all $x \in J$, write

$$s_J(x) = P_x\{\tau_b < \tau_a\}.$$

Clearly, $s_J(a) = 1$ and $s_J(b) = 0$. If $a < x < y < b$, then by the strong Markov property,

$$s_J(x) = P_x\{\tau_y < \tau_a\} s_J(y).$$

Lemma 4.117. *s_J is continuous and strictly increasing.*

Proof. By the statements above, the theorem is equivalent to showing that for $a < x < y < b$, $P_x\{\tau_y < \tau_a\}$ is continuous, strictly increasing as a function of y and has limit 1 as $x \to y$. (This give right continuity. Left continuity follows from a symmetric argument.)

To this end,

$$P_x\{\tau_y < \tau_a\} = P_x\{\sup_{t \le \tau_a} X_t \ge y\} \to P_x\{\sup_{t \le \tau_a} X_t > x\} = 1$$

because $\tau_a > 0$ a.s. P_x

We now show that $P_x\{\tau_y < \tau_a\} = 1$ contradicts the regularity of X. Define stopping times $\xi_0 = 0$, $\sigma_{n+1} = \inf\{t > \xi_n : X_t = x\} = \sigma(\xi_n, \{x\})$, $\xi_{n+1} = \inf\{t > \sigma_{n+1}; X_t \in \{a, y\}\}$. By the strong Markov property at ξ_n,

$$P_y\{\xi_n < \infty, X_{\xi_n} = a\} = E_y[P_x\{\tau_a < \tau_y\}; \{\xi_n < \infty\}] = 0$$

and thus,

$$P_x\{\tau_a < \infty\} = 0.$$

□

4.10.1 The Scale Function

Definition 4.118. *Let X be a regular diffusion on I. A* scale function *for X is a function $s : I \to \mathbb{R}$ such that for all $a < x < b$, $a, b \in I$,*

$$P_x\{\tau_b < \tau_a\} = \frac{s(x) - s(a)}{s(b) - s(a)}.$$

If $s(x) = x$ is a scale function for X, then we say that X is in its natural scale. In particular, if $x = (a + b)/2$, then $P_x\{\tau_b < \tau_a\} = 1/2$.

Exercise 4.119.
1. *Standard Brownian motion is on natural scale.*
2. *s is uniquely determined up to an increasing affine transformation.*
3. *A process X is on natural scale if and only if $P_x\{\tau_b < \tau_a\} = 1/2$ for every $a, b \in I$ and $x = (a + b)/2$.*

Theorem 4.120. *Every regular diffusion on I has a scale function.*

Proof. If I is compact, then s_I is a scale function. To see this, take $I = [a, b]$ and $a < c < x < d < b$. Then, by the strong Markov property

$$P_x\{\tau_a < \tau_b\} = P_x\{\tau_d < \tau_c\}P_d\{\tau_a < \tau_b\} + P_x\{\tau_c < \tau_d\}P_c\{\tau_a < \tau_b\}$$

$$\frac{s_I(x) - s_i(a)}{s_I(b) - s_I(a)} = P_x\{\tau_d < \tau_c\}\frac{s_I(d) - s_i(a)}{s_I(b) - s_I(a)} + (1 - P_x\{\tau_c < \tau_d\})\frac{s_I(c) - s_i(a)}{s_I(b) - s_I(a)}.$$

Solving, we obtain,

$$P_x\{\tau_d < \tau_c\} = \frac{s_I(x) - s_I(c)}{s_I(d) - s_(c)}.$$

Now, we extend. Take $\{J_n; n \geq 1\}$ to be an increasing sequence of compact intervals with union I. Moreover, if I contains an endpoint, then construct each $J_n = [a_n, b_n]$ to include this endpoint. Define $\tilde{s}_n : J_n \to \mathbb{R}$ inductively. $\tilde{s}_1 = s_{J_1}$, and, for $n \geq 1$,

$$\tilde{s}_{n+1} = \tilde{s}_n(a_n) + (\tilde{s}_n(b_n) - \tilde{s}_n(a_n))\frac{s_{J_{n+1}}(x) - s_{J_{n+1}}(a_n)}{s_{J_{n+1}}(b_n) - s_{J_{n+1}}(a_n)}.$$

Clearly, $\tilde{s}_{n+1}(a_n) = \tilde{s}_n(a_n)$, and $\tilde{s}_{n+1}(b_n) = \tilde{s}_n(b_n)$. Repeat the argument above to show that $\tilde{s}_{n+1}$ is a scale function on J_n. Because $\tilde{s}_n$ is an affine transformation of s_{n+1} and the two functions agree at two points, they must be equal on J_n.

Consequently, we can define, for $x \in J_n$, $s = \tilde{s}_n$. □

Theorem 4.121. *If X is a regular diffusion on I, then $\tilde{X} = s(X)$ is a regular diffusion in natural scale on $s(I)$.*

Proof. Write $\tilde{a} = s(a)$, $\tilde{b} = s(b)$, and $\tilde{\tau}_a = \inf\{t > 0 : \tilde{X}_t = a\}$

$$P_{\tilde{x}}\{\tilde{\tau}_{\tilde{a}} < \tilde{\tau}_{\tilde{b}}\} = P_x\{\tau_a < \tau_b\} = \frac{s(x) - s(a)}{s(b) - s(a)} = \frac{\tilde{x} - \tilde{a}}{\tilde{b} - \tilde{a}}.$$

□

4.10.2 The Speed Measure

The structure threorem for jump Markov processes gives to each site x in a birth and death process, and function $\lambda(x)$, that determines the length of time at x and a probability $p_x = P_x\{X_{\tau_1} = x + 1\}$, the bias in direction.

For a one dimensional regular diffusion, the bias in direction is articulated through the scale function s. The length of time near x is described by $E_x\tau_J$. In this section we shall show how these mean values can be represented as integration of a kernel against a measure. To explain the kernel we have the following:

Exercise 4.122. *For a simple symmetric random walk Y on Z, and let $\tau = \min\{n; Y_n \in \{M_-, M_+\}\}$, $M_- < M_+$ and let N_y be the number of visits to y before τ. Then*

$$E_x\tau = (x - M_-)(M_+ - x),$$

and

$$E_x N_y = \begin{cases} \frac{(x-M_-)(M_+-y)}{M_+-M_-}, & x \leq y \\ \frac{(y-M_-)(M_+-x)}{M_+-M_-}, & x \geq y. \end{cases}$$

For a bounded open interval $J = (a, b)$ define, for $x, y \in J$

$$G_J(x, y) = \begin{cases} \frac{(x-a)(b-y)}{b-a} & x \leq y, \\ \frac{(y-a)(b-x)}{b-a} & x \geq y. \end{cases}$$

Theorem 4.123. *Let X be a one dimensional diffusion on I on natural scale. Then there is a unique measure m defined on $\mathcal{B}(int(I))$ such that*

1. *m is finite on bounded sets, B, $\bar{B} \subset int(I)$.*

2. *For a bounded open interval J,*

$$m(x,J) = E_x\tau_J = \int_J G_J(x,y)m(dy).$$

Proof. Let $\{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition $[a,b]$ into subintervals of length $h = (b-a)/n$. Thus, $x_k = a + hk$ Define $J_k = (x_{k-1}, x_{k+1})$.

The process restricted to the exit times of the intervals J_k is a simple symmetric random walk Y on $(b-a) + hZ$. Let σ_j denote the time between steps before τ_j and zero after. Then for a partition point x,

$$E_x\tau_J = \sum_{j=0}^{\infty} E_x\sigma_j.$$

and

$$E_x\sigma_j = E_x[E_x[\sigma_j|Y_j]] = \sum_{k=1}^{n-1} P_x\{Y_j = x_k\}m(x_k, J_k).$$

Let $n(x_{k_0}, x_k)$ denote the expected number of visits of the random walk to x_k starting at x_{k_0}. Then, by the exercise,

$$n(x_{k_0}, x_k) = \begin{cases} \frac{k_0(n-k)}{n} & x_{k_0} \leq x_k \\ \frac{k(n-k_0)}{n} & x_k \geq x_{k_0} \end{cases} = \frac{1}{h}G(x_{k_0}, x_k).$$

Thus,

$$\begin{aligned} E_x\tau_J &= \sum_{j=0}^{\infty}\sum_{k=1}^{n-1} P_x\{Y_j = x_k\}m(x_k, J_k) = \sum_{k=1}^{n-1}(\sum_{j=0}^{\infty} P_x\{Y_j = x_k\})m(x_k, J_k) \\ &= \sum_{k=1}^{n-1} n(x, x_k)m(x_k, J_k) = \sum_{k=1}^{n-1} \frac{1}{h}G_J(x, x_k)m(x, J_k) \end{aligned}$$

Define the measure m_n that gives mass $m(x_k, J_k)/h$ to the point x_k. Then the formula above becomes,

$$m(x_j, J) = \int_J G_J(x_j, y)\ m_n(dy).$$

Now consider succesive refinements, (say $n = 2^\ell$). Then the representation above holds whenever the ratio $(x-a)/(b-a)$ is a dyadic rational. This represenation above also implies that $\limsup_{n\to\infty} m_n(J) < \infty$. Now use a variant of the Prohorov theorem to show that $\{m_{2^\ell} : \ell \geq 1\}$ has compact closure in the topology determined by weak convergence. Let m be a limit point for the $\{m_{2^\ell} : \ell \geq 1\}$, then for any finite open interval J with dyadic rational endpoints, $\bar{J} \subset I$,

$$m(x,J) = \int_J G_J(x,y)\ m(dy), \quad \frac{x-a}{b-a} \text{ a dyadic rational.}$$

By the bounded convergence theorem, this can be extended to all such open intervals J.

To extend this to arbitrary y, use the strong Markov property to obtain the identity for $(\tilde{a}, \tilde{b}) = \tilde{J} \subset J$,

$$m(x, J) = m(x, \tilde{J}_n) + P_x\{\tau_{\tilde{a}} < \tau_{\tilde{b}}\}m(\tilde{b}, J) + P_x\{\tau_{\tilde{a}} > \tau_{\tilde{b}}\}m(\tilde{a}, J).$$

Now let $\tilde{a}_n$ and $\tilde{b}_n$ be dyadic rationals $\tilde{a}_n < x < \tilde{b}_n$, then

$$\lim_{n\to\infty} m(\tilde{a}, J) = \lim_{n\to\infty} m(\tilde{b}, J) = m(x, J)$$

and by the representation for dyadic rationals,

$$\lim_{n\to\infty} m(x, \tilde{J}) = 0.$$

□

Exercise 4.124. *The measure m is unique.*

Definition 4.125. *The measure m is called the* speed measure *for the process.*

Exercise 4.126.

1. *If X is not on its natural scale, then the speed measure m can be defined by*

$$m(x, J) = \int_J G_J(x, y)\ m(dy)$$

with

$$G_J(x, y) = \begin{cases} \frac{(s(x)-s(a))(s(b)-s(y))}{s(b)-s(a))} & x \le y \\ \frac{(s(y)-s(a))(s(b)-s(x))}{s(b)-s(a))} & x \le y \end{cases}$$

2. *Let X be standard Brownian motion, then for $J = (a, b)$,*

$$E_x\tau_J = (x - a)(x - b).$$

Thus, the speed measure is Lebesgue measure.

3. *Conversely, if the speed measure is Lebesgue measure, then for $J = (a, b)$,*

$$E_x\tau_J = (x - a)(x - b).$$

4. *Use the standard machine to prove that for a bounded measurable function,*

$$E_x[\int_0^{\tau_J} f(X_t)\ dt] = \int_J f(y)G_J(x, y)\ m(dy)$$

Theorem 4.127. *Let X be a regular diffusion on natural scale with speed measure m and let $\tilde{X}$ be the unique solution to*

$$\tilde{X}_t = X_{\int_0^t c(\tilde{X}_s)\ ds}.$$

Then, X is a Feller process on natural scale with speed measure

$$\tilde{m}(A) = \int_A c(x)^{-1}\ m(dx).$$

Proof. Set
$$\sigma_t = \int_0^t c(\tilde{X}_s)\ ds.$$
Then, as before
$$t = \int_0^{\sigma_t} c(\tilde{X}_s)\ ds.$$
Denote
$$\tau_a = \inf\{t > 0 : X_t = a\} \quad \tilde{\tau}_a = \inf\{t > 0 : \tilde{X}_t = a\}.$$
Then
$$P_x\{\tilde{\tau}_a < \tilde{\tau}_b\} = P_x\{\tau_a < \tau_b\}.$$
Thus, $\tilde{X}$ is on natural scale.

Let $\tilde{\tau}_J$ be the first exit time of $\tilde{X}$ from J. Then
$$\tau_J = \sigma_{\tilde{\tau}_J} \quad \text{or} \quad \tilde{\tau}_J = \int_0^{\tau_J} \frac{1}{c(X_s)}\ ds.$$
Taking expectations, we obtain
$$E_x\tilde{\tau}_J = E_x[\int_0^{\tau_J} c(X_s)^{-1}\ ds] = \int_J G_J(x,y)c(y)^{-1}\ m(dy).$$
Thus, $\tilde{X}$ has the asserted speed measure. □

4.10.3 The Characteristic Operator

Because we are working with exit times in the definition of the speed measure, it is convenient to continue this idea to give an operator similar to the generator.

Definition 4.128. *Let $x \in int(I)$, then*
$$\lim_{J\to\{x\}} \phi(J) = L$$
means that for any system of open sets $\{J_n; n \geq 1\}$ decreasing to the point x, we have
$$\lim_{n\to\infty} \phi(J_n) = L.$$

Definition 4.129. *Let X be a regular diffusion on I. The* characteristic operator
$$Cf(x) = \lim_{J\to\{x\}} \frac{E_x f(X_{\tau_J}) - f(x)}{E_x\tau_J}.$$
The domain of C is the subspace of all $f \in C_0(I)$ for which the limit above exists.

Theorem 4.130. *Let X be a regular diffusion on I with generator A. Let $f \in \mathcal{D}(A)\cap\mathcal{D}(C)$, then $Af(x) = Cf(x)$ for all $x \in int(I)$.*

Proof. Choose $f \in \mathcal{D}(A)$, and $x \in \text{int}(I)$. Set $g = Af$. Then Dynkin's formula states that

$$E_x[f(X_{\tau_J})] = f(x) + E[\int_0^{\tau_J} g(X_s)\ ds].$$

for an interval J, $\bar{J} \subset \text{int}(I)$. Choose $\epsilon > 0$ so that $g(J) \subset (g(x)-\epsilon, g(x)+\epsilon)$. Then,

$$g(x) - \epsilon < \frac{E_x f(X_{\tau_J}) - f(x)}{E_x \tau_J} < g(x) + \epsilon$$

and $Cf(x) = Af(x)$. □

Theorem 4.131. *If X is a regular diffusion on the line on natural scale and with speed measure equal to Lebesgue measure, the X is standard Brownian motion.*

Proof. Let A be the generator for X and let C be its characteristic operator. Then, because R is open, $\mathcal{D}(A) \subset \mathcal{D}(C)$. Write an interval containing x as $J(\lambda, \epsilon) = (x + \lambda\epsilon, x - (1-\lambda)\epsilon)$, $\lambda \in (0,1)$. Then,

$$E_x \tau_{J(\lambda,\epsilon)} = \lambda(1-\lambda)\epsilon^2.$$

For $f \in C^2(R)$,

$$Ef(X_{\tau_{J(\lambda,\epsilon)}}) = (1-\lambda)f(x+\lambda\epsilon) + \lambda f(x-(1-\lambda)\epsilon).$$

and

$$Ef(X_{\tau_{J(\lambda,\epsilon)}}) - f(x) = (1-\lambda)(f(x+\lambda\epsilon) - f(x)) + \lambda(f(x-(1-\lambda)\epsilon) - f(x)).$$

Now, choose sequences $\epsilon_n \to 0$ and $\lambda_n \in (0,1)$. Then, by L'Hôpital's rule

$$\begin{aligned}\lim_{n\to\infty} \frac{E_x f(X_{\tau_{J(\epsilon_n,\lambda_n)}}) - f(x)}{E_x \tau_J(\epsilon_n, \lambda_n)} &= \lim_{n\to\infty} \frac{(1-\lambda_n)(f(x+\lambda_n\epsilon_n) - f(x)) + \lambda_n(f(x-(1-\lambda_n)\epsilon_n) - f(x))}{\lambda_n(1-\lambda_n)\epsilon_n^2} \\ &= \lim_{n\to\infty} \frac{f'(x+\lambda_n\epsilon_n) + f(x-(1-\lambda_n)\epsilon_n)}{2\epsilon_n} = \frac{1}{2}f''(x).\end{aligned}$$

□

Set aside for the moment behavior in the boundary of I. Beginning with any scale s and any speed measure having a density c, we can transform it to Brownian by a change of scale and a time change. Thus, by inverting these transformations, we can transform a Brownian motion into a regular diffusion having scale and speed measure $\int_A c(x)\ dx$. This shows, among this class of scales and speed measure, that the behavior of a regular diffusion is uniquely determined in the interior of I by these two properties of the diffusion.

For diffusions X in this class, the generator takes the form

$$A = \frac{1}{2}a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}, \quad a(x) > 0.$$

Let $x_0 \in I$ and define

$$s(x) = \int_{x_0}^x \exp\left(-2\int_a^y \frac{b(z)}{a(z)^2}\ dz\right)\ dy$$

Then s is a harmonic function

For $J = (a,b) \subset \text{int}(I)$, we have by Dynkin's formula,

$$s(x) = s(a)P\{X_{\tau_J} = a\} + s(b)P\{X_{\tau_J} = b\}.$$

Therefore, s is a scale function.

Now, use the chain rule to see that $s(X)$ has generator

$$\frac{1}{2}a(x)s'(x)^2\frac{d^2}{dx^2}.$$

Thus, the speed measure has density $1/(a(x)s'(x)^2)$.

Example 4.132. 1. *The radial part of an n-dimensional Brownian motion R is called a* Bessel process. *In this case, $I = [0,\infty)$ and*

$$A = \frac{1}{2}\frac{d^2}{dr^2} + \frac{n-1}{2r}\frac{d}{dr}.$$

2. *The* Ornstein-Uhlenbeck process *is obtained by subjecting particles undergoing physical Brownian motion to an elastic force. The generator is*

$$A = \frac{\sigma^2}{2}\frac{d^2}{dx^2} - \rho x\frac{d}{dx}, \quad I = R.$$

3. *Financial models of stock prices use* geometric Brownian motion. *Here, $I = (0,\infty)$ and the generator is*

$$A = \frac{\sigma^2}{2}x\frac{d^2}{dx^2} + \rho x\frac{d}{dx}.$$

4. *A diffusion approximation of a neutral 2 allele model with scaled mutation rates μ_0 and μ_1 is*

$$A = \frac{1}{2}x(1-x)\frac{d^2}{dx^2} + (\mu_1 x - \mu_0(1-x))\frac{d}{dx}, \quad I = [0,1].$$

Exercise 4.133. 1. *Find a scale function and the speed measure for the four examples above.*

2. *Show that a normal distribution is stationary for the Ornstein-Uhlenbeck process.*

3. *Show that a beta distribution is stationary for the diffusion in example 4.*

4.10.4 Boundary Behavior

We now have a good understanding of the behavior of a regular diffusion on $\text{int}(I)$. We conclude by examining its behavior at the boundary.

Definition 4.134. *Let $I = [r_-, r_+] \cap R$. The r_+ is called an* accessible *is there exists $t > 0$ and $x \in int(I)$ such that*

$$\inf_{\{y \in (x,r_+)\}} P_x\{\tau_y \le t\} > 0.$$

If r_+ is inaccessible, call it natural *if there exists $t > 0$ and $y \in int(I)$ such that*

$$\inf_{\{y \in (x, r_+)\}} P_x\{\tau_y \le t\} > 0.$$

Call it entrance *if for all $t > 0$ and $y \in int(I)$ such that*

$$\lim_{x \to r_+ -} P_x\{\tau_y \le t\} = 0.$$

An accessible boundary may or may not be regular. If it is not, then call the boundary point an exit*. Corresponding statements hold for r_-*

In schematic form, we have

	$r_+ \to \mathrm{int}(I)$	$r_+ \not\to \mathrm{int}(I)$
$\mathrm{int}(I) \to r_+$	regular	exit
$\mathrm{int}(I) \not\to r_+$	entrance	natural

Theorem 4.135. *Fix $x_0 \in I$ and define*

$$u(x) = \int_{x_0}^x s(y)\ m(dy).$$

Then r_+ is accessible if and only if $u(r_+) < \infty$

Proof. Let $J = (x_0, r_+)$. By the lemma, for $y \in J$

$$\begin{aligned}
\infty > E_y \tau_J &= \int_{x_0}^{r_+} G_J(y,z)\ m(dz) \\
&\ge \frac{s(r_+) - s(y)}{s(r_+) - s(x_0)} \int_y^{r_+} (s(z) - s(x_0))\ m(dz) \\
&\ge \frac{s(r_+) - s(y)}{s(r_+) - s(x_0)} \left(u(r_+) - u(y) - s(x_0) m(y, r_+)\right).
\end{aligned}$$

Because $m(x,y) < \infty$ and $u(y) < \infty$, we have that $u(r_+) < \infty$.

Conversely, assume that $u(r_+) < \infty$ and choose $x \in (x_), r_+)$. Then,

$$E_x \tau_{x_0} = \lim_{y \to r_+ -} m(x, (x_0, y)) = \int_{x_0}^y G_{(x_0,y)}(x, s)\ m(dz).$$

Note that,

$$E_x \tau_{x_0} \text{ is nondecreasing as } x \text{ incresing.}$$

and that

$$\int_{x_0}^y G_{(x_0,y)}(x,s)\ m(dz) = \frac{2(s(r_+) - s(x)}{2(s(r_+) - s(x_0)} \int_{x_0}^x (s(z) - s(x_0))\ m(dz) + \frac{2(s(c) - s(x_0)}{2(s(r_+) - s(x_0)} \int_{x_0}^x (s(r_0) - s(z))\ m(dz).$$

This tends to zero as $x \to r_+$. Thus,

$$\lim_{x \to r_+} E_x \tau_{x_0} = 0$$

and r_+ is accessible. $\square$

5 Stochastic Integrals

In this section we develop the theory of stochastic integrals with the primary intent to apply them to the study of a class of Markov processes having continuous sample paths.

The theory of Riemann-Steitjes integral applies readily to continuous functions having finite variation on compact intervals. Thus, we can readily obtain a stochastic integral for

$$(X \cdot V)_t(\omega) = \int_0^t X_s(\omega)\ dV_s(\omega).$$

for

1. progressively measurable continuous V having finite variation on compact intervals, and
2. progressively measurable continuous X bounded on compact intervals.

To extend this to continuous martingales, we must take into account the following.

Exercise 5.1. *Let M be a martingale with $EM_t^2 < \infty$ for all $t > 0$. Then, for $s, t > 0$,*

$$Var(M_{s+t}|\mathcal{F}_t) = E[M_{s+t}^2|\mathcal{F}_t] - M_t^2 = E[(M_{s+t} - M_t)^2|\mathcal{F}_t].$$

Thus

$$E[(M_{s+t} - M_t)^2] = E[M_{s+t}^2 - M_t^2].$$

Definition 5.2. *A partition Π of $[0, T]$ is a strictly increasing sequence of real numbers $0 = t_0 < t_1 < \cdots < t_n = T$. A partition Π of $[0, \infty)$ is a strictly increasing unbounded sequence of real numbers $0 = t_0 < t_1 < \cdots < t_k \cdots$, $\lim_{k\to\infty} t_k = \infty$. The mesh$(\Pi) = \sup\{t_{k+1} - t_k; k \geq 0\}$*

For any process X and partition $\Pi = \{t_0 < t_1 < \cdots < t_k \cdots\}$ of $[0, \infty)$, define the simple process

$$X_t^\Pi = \sum_{k=0}^\infty X_{t_k} I_{[t_k, t_{k+1})}(t).$$

Theorem 5.3. *A continuous martingale having finite variation on compact intervals is constant.*

Proof. By considering the martingale $\tilde{M}_t = M_t - M_0$, we can assume that $M_0 = 0$. In addition, let σ_c be the time that the variation of M reaches C. Then by considering the martingale M^{σ_c}, we can assume that M has variation bounded above by c.

Now, let $\Pi = \{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t\}$, be a partitions of $[0, t]$, then

$$EM_t^2 = \sum_{j=1}^n E[M_{t_{j+1}}^2 - M_j^2] = \sum_{j=1}^n E[(M_{t_{j+1}} - M_{t_j})^2] \leq nE[\max_{1\leq j\leq k} |M_{t_{j+1}} - M_{t_j}|].$$

The random variable $max_{1\leq j\leq k}|M_{t_{j+1}} - M_{t_j}|$ is bounded above by c and converges to 0 almost surely as the mesh of the partition tends to 0. Thus,

$$EM_t^2 = 0 \text{ and } M_t = 0 \text{ a.s.}$$

□

5.1 Quadratic Variation

Definition 5.4. *For the partition* Π*,* $0 = t_0 < t_1 < \cdots < t_k \cdots,$ *of* $[0, \infty)$*, the* quadratic variation process *of a stochastic process* X *along* Π *is*

$$Q_t^\Pi(X) = \sum_{j=1}^{\infty} (M_{\min\{t,t_{j+1}\}} - M_{\min\{t,t_j\}})^2.$$

We say that X *has* finite quadratic variation *if there exists a process* $\langle X, X \rangle$ *such that*

$$Q_t^\Pi(X) \to^P \langle X, X \rangle_t \text{ as } mesh(\Pi) \to 0.$$

Theorem 5.5. *A continuous bounded martingale* M *has finite quadratic variation* $\langle M, M \rangle_t$ *that is the unique continuous increasing adapted process vanishing at zero such that*

$$M_t^2 - \langle M, M \rangle_t$$

is a martingale.

Proof. The uniquess of $\langle M, M \rangle$ is easy, let A and B be two such processes. Then,

$$A_t - B_t = (M_t^2 - B_t) - (M_t^2 - A_t)$$

as the difference of two martingales is itself a martingale. Because it has finite variation on compact intervals, it is constant. □

For existence, we will prove:

Theorem 5.6. *If* $\{\Pi_n; n \geq 1\}$ *is a sequence of partitions of* $[0, \infty)$ *with* $mesh(\Pi_n) \to 0$ *then*

$$\lim_{n\to\infty} E[\sup_{0\leq t\leq T} (Q_t^{\Pi_n}(M) - \langle M, M \rangle_t)] = 0.$$

Exercise 5.7. *Let* B *denote standard one-dimensional Brownian motion. With the partitions* Π_n *described as above, show that* $Q_t^{\Pi_n}(B) \to^{L^2} t$ *as* $n \to \infty$*.*

For this, we will introduce a second sequence I^n of process, which we shall later recognize as the approximation to the stochastic integral $\int_0^t M_s \, dM_s$. Write $\Pi_n = \{0 = t_0^n < t_1^n < \cdots < t_k^n \cdots\}$, then

$$I_t^{\Pi_n} = \sum_{k=0}^{\infty} M_{\min\{t,t_k^n\}} (M_{\min\{t,t_{k+1}^n\}} - M_{\min\{t,t_k^n\}}).$$

Exercise 5.8. *1.*

$$M_t^2 - M_0^2 = Q_t^{\Pi_n}(M) + 2I_t^{\Pi_n}. \tag{5.1}$$

2. I^{Π_n} *is a mean zero martingale. Consequently,* $M_t^2 - Q_t^{\Pi_n}(M)$ *is a mean-zero martingale.*

Lemma 5.9. *Choose $c \geq \sup_{t\geq 0}|M_t|$, then*

$$E[(I_t^{\Pi_n})^2] \leq c^4.$$

Proof. The quadratic variation of I^{Π_n} associated with Π^n

$$Q_t^{\Pi_n}(I^{\Pi_n}) = \sum_{k=0}^{\infty} M^2_{\min\{t,t_k^n\}}(M_{\min\{t,t_{k+1}^n\}} - M_{\min\{t,t_k^n\}})^2 \leq c^2 Q_t^{\Pi_n}(M).$$

Thus,

$$E[Q_t^{\Pi_n}(I^n)] \leq c^2 E[Q_t^{\Pi_n}(M)] = c^2 E[M_t^2 - M_0^2] \leq c^4.$$

□

Lemma 5.10. *With $c \geq \sup_{t\geq 0}|M_t|$,*

$$\lim_{m,n\to\infty} E[\sup_{0\leq t\leq T}(I_t^{\Pi_n} - I_t^{\Pi_m})^2] = 0. \tag{5.2}$$

Proof. Write

$$\Pi_n \cup \Pi_m = \{u_0 < u_1 < \cdots < u_k \cdots\}.$$

Then

$$I_t^{\Pi_n} - I_t^{\Pi_m} = \sum_{k=1}^{\infty}(M^{\Pi_n}_{\min\{t,u_k\}} - M^{\Pi_m}_{\min\{t,u_k\}})(M_{\min\{t,u_{k+1}\}} - M_{\min\{t,u_k\}})$$

is a mean zero martingale, and therefore, $(I_t^{\Pi_n} - I_t^{\Pi_m})^2 - Q_t^{\Pi_m\cup\Pi_m}(I^{\Pi_n} - I^{\Pi_m})$ is, by the exercise, a mean zero martingale.. Thus,

$$\begin{aligned} Q_t^{\Pi_m\cup\Pi_m}(I^{\Pi_n} - I^{\Pi_m}) &= \sum_{k=1}^{\infty}(M^{\Pi_n}_{\min\{t,u_k\}} - M^{\Pi_m}_{\min\{t,u_k\}})^2(M_{\min\{t,u_{k+1}\}} - M_{\min\{t,u_k\}})^2 \\ &\leq \sup_{0\leq s\leq t}(M_s^{\Pi_n} - M_s^{\Pi_m})^2 Q_t^{\Pi_m\cup\Pi_m}(M). \end{aligned}$$

By Doob's inequality, and then by Schwartz's inequality, we have,

$$\begin{aligned} E[\sup_{0\leq t\leq T}(I_t^{\Pi_n} - I_t^{\Pi_m})^2] &\leq 4E[(I_T^{\Pi_n} - I_T^{\Pi_m})^2] = 4E[Q_T^{\Pi_n\cup\Pi_m}(I^{\Pi_n} - I^{\Pi_m})] \\ &\leq 4E[\sup_{0\leq t\leq T}(M_t^{\Pi_n} - M_t^{\Pi_m})^4]^{1/2} E[Q_T^{\Pi_m\cup\Pi_m}(M)^2]^{1/2}. \end{aligned}$$

In addition,

$$E[Q_T^{\Pi_m\cup\Pi_m}(M)^2] \leq 2(E[(M_T^2 - M_0^2)^2] + 4E[(I_T^{\Pi_m\cup\Pi_n})^2]) \leq 12c^4.$$

Finally, $\sup_{0\leq t\leq T}(M_t^{\Pi_n} - M_t^{\Pi_m})^4$ is bounded and converges to zero almost surely as $m,n\to\infty$. Thus, by the bounded converges theorem, its expectation converges to zero and (5.2) holds. □

Returning to the proof of the theorem.

Proof. The continuous processes I^{Π_n} are Cauchy in the norm $||X-Y||_{2,T} = E[\sup_{0\le t\le T}(X_t-Y_t)^2]^{1/2}$, thus, by (5.1). so is $Q^{\Pi_n}(M)$. Because the space is complete and the convergences is uniform almost surely, the limit $Q(M)$ is continuous. Choose, if necessary, a subsequence that converges almost surely.

Write $\tilde{\Pi}_n = \cup_{k=1}^n \Pi_n$. Then for any $s,t \in \tilde{\Pi}_n$, $s < t$

$$Q_t^{\tilde{\Pi}_n}(M) \ge Q_s^{\tilde{\Pi}_n}(M).$$

Consequently, $Q(M)$ is nondecreasing on the dense set $\cup_{k=1}^\infty \Pi_n$. Because $Q(M)$ is continuous, it is nondecreasing. Finally, that fact that $M_t^2 - Q_t(M)$ is a martingale can be seen by taking limits on the expectations

$$E[M_t^2 - Q_t^{\tilde{\Pi}_n}(M); A] = E[M_s^2 - Q_s^{\tilde{\Pi}_n}(M); A], \quad A \in \mathcal{F}_s,\ s < t.$$

□

To extend this theorem to the broader class of local martingales.

Proposition 5.11. *Let M be a bounded continuous martingale and let τ be stopping time. Then*

$$\langle M^\tau, M^\tau\rangle_t = \langle M, M\rangle_t^\tau.$$

Proof. By the optional sampling theorem,

$$(M^\tau)^2 - \langle M, M\rangle^\tau.$$

is a martingale. Now use the uniqueness of the quadratic variations process. □

Theorem 5.12. *A continuous local martingale M has associated to it $\langle M, M\rangle_t$ the unique continuous increasing adapted process vanishing at zero so that*

$$M_t^2 - \langle M, M\rangle_t$$

is a local martingale.

In addition, for any $T > 0$ and any sequence of partitions $\{\Pi_n; n \ge 1\}$ of $[0,\infty)$ with mesh$(\Pi_n) \to 0$ then

$$\sup_{0\le t\le T} |Q_t^{\Pi_n}(M) - \langle M, M\rangle_t| \to^P 0$$

as $n \to \infty$.

Proof. Choose $\{\tau_m; m \ge 1\}$ be a reducing sequence for M so that $M^{\tau_m} I_{\{\tau_m > 0\}}$ is bounded. Let $Q^m(M)$ denote the unique continuously increasing process, $Q_0^m(M) = 0$ so that

$$(M^{\tau_m})^2 I_{\{\tau_m > 0\}} - Q^m(M)$$

is a bounded martingale. By the uniqueness property and the proposition above,

$$(Q^{m+1}(M))^{\tau_m} = Q^m(M)$$

on $\{\tau_m > 0\}$. This uniquely defines $\langle M, M\rangle$ on $[0, \tau_m]$.

For the second statement, let $\epsilon, \delta > 0$. Choose m_0 so that

$$P\{\tau_{m_0} > T\} < \delta.$$

Then, on the random interval $[0, \tau_{m_0}]$,

$$Q_t^{\Pi_n}(M^{\tau_{m_0}}) = Q_t^{\Pi_n}(M) \quad \text{and} \quad \langle M^{\tau_{m_0}}, M^{\tau_{m_0}}\rangle_t = \langle M, M\rangle_t.$$

Consequently,

$$P\{\sup_{0\le t\le T} |Q_t^{\Pi_n}(M) - \langle M, M\rangle_t| > \epsilon\} \le \delta + P\{\sup_{0\le t\le T} |Q_t^{\Pi_n}(M^{\tau_{m_0}}) - \langle M^{\tau_{m_0}}, M^{\tau_{m_0}}\rangle_t| > \epsilon\}.$$

The last term tends to 0 as $\text{mesh}(\Pi_n) \to 0$. □

For any quadratic form, we have a polarization identity to create a bilinear form. That is the motivation behind the next theorem.

Theorem 5.13. *Let M and N be two continuous local martingales, then there is a unique continuous adapted process $\langle M, N\rangle$ of bounded variation vanishing at zero so that*

$$M_t N_t - \langle M, N\rangle_t$$

is a local martingale.

In addition, set, for the partition Π for $0 = t_0 < t_1 < \cdots < t_k \cdots$, $\lim_{k\to\infty} t_k = \infty$.

$$C_t^{\Pi}(M, N) = \sum_{j=1}^{\infty} (M_{\min\{t,t_{j+1}\}} - M_{\min\{t,t_j\}})(N_{\min\{t,t_{j+1}\}} - N_{\min\{t,t_j\}}).$$

Then, for any $t > 0$ and any sequence of partitions $\{\Pi_n; n \ge 1\}$ of $[0, \infty)$ with mesh$(\Pi_n) \to 0$ then

$$\sup_{0\le t\le T} (C_t^{\Pi_n}(M, N) - \langle M, N\rangle_t) \to^P 0$$

as $n \to \infty$.

Proof. Set

$$\langle M, N\rangle_t = \frac{1}{4}(\langle M + N, M + N\rangle_t) - \langle M - N, M - N\rangle_t)$$

and use an appropriate reducing sequence. Uniqueness follows are before. □

Definition 5.14. *We will call $\langle M, N\rangle$ the* bracket *of M and N and $\langle M, M\rangle$ the increasing process associated with M*

Exercise 5.15.

1. *For a stopping time τ, $\langle M^\tau, N^\tau\rangle = \langle M, N^\tau\rangle = \langle M, N\rangle^\tau$.*
2. *$\langle M, M\rangle = 0$ if and only if $M_t = M_0$ a.s. for every t.*
3. *The intervals of constancy of M and $\langle M, M\rangle$ are the some. (*Hint: *For constancy on $[a, b]$, consider the martingale $\theta^a M_{\min\{t,b-a\}}$.)*
4. *The brackets process is positive definite, symmetric and bilinear.*
5. *If M and N are independent, then $\langle M, N\rangle$ is zero.*

For any positive definite, symmetric and bilinear form, we have

Proposition 5.16 (Cauchy-Schwartz). *Let M and N be continuous local martingales. Then*

$$|\langle M,N\rangle_t - \langle M,N\rangle_s| \leq (\langle M,M\rangle_t - \langle M,M\rangle_s)^{1/2}\,(\langle N,N\rangle_t - \langle N,N\rangle_s)^{1/2}.$$

Soon, we shall develop the stochastic integral. We will begin with a class of simple functions and extend using a Hilbert space isometry. The Kunita-Watanabe inequalities are central to this extension.

Proposition 5.17. *Let M and N be local martingales and let H and K be measurable processes, then*

$$|\int_0^t H_s K_s d\langle M,N\rangle_s| \leq \left(\int_0^t H_s^2 d\langle M,M\rangle_s\right)^{1/2}\left(\int_0^t K_s^2 d\langle N,N\rangle_s\right)^{1/2}.$$

Proof. By the standard machine, it suffices to consider H and K of the form

$$K = K_0 I_{\{0\}} + \sum_{j=1}^n K_j I_{(t_{j-1},t_j]},\; H = H_0 I_{\{0\}} + \sum_{j=1}^n H_j I_{(t_{j-1},t_j]}$$

with $0 = t_0 < \cdots < t_{n-1} < t_n = t$..

Use the Cauchy-Schwartz inequality for sums.

$$\begin{aligned}
|\int_0^t H_s K_s d\langle M,N\rangle_s| &\leq \sum_{j=1}^n |H_j K_j||\langle M,N\rangle_{t_{j+1}} - \langle M,N\rangle_{t_]}| \\
&\leq \sum_{j=1}^n |H_j K_j|\left(\langle M,M\rangle_{t_{j+1}} - \langle M,M\rangle_{t_]}\right)^{1/2}\left(\langle N,N\rangle_{t_{j+1}} - \langle N,N\rangle_{t_]}\right)^{1/2} \\
&\leq \left(\sum_{j=1}^n H_j^2(\langle M,M\rangle_{t_{j+1}} - \langle M,M\rangle_s)\right)^{1/2}\left(\sum_{j=1}^n K_j^2(\langle N,N\rangle_{t_{j+1}} - \langle N,N\rangle_{t_]})\right)^{1/2} \\
&= \left(\int_0^t H_s^2 d\langle M,M\rangle_s\right)^{1/2}\left(\int_0^t K_s^2 d\langle N,N\rangle_s\right)^{1/2}.
\end{aligned}$$

□

Applying the Hölder inequality yields the following.

Corollary 5.18 (Kunita-Watanabe inequality). *Let p and q be conjugate exponents. Then*

$$E\left[\left|\int_0^t H_s K_s d\langle M,N\rangle\right|\right] \leq \left\|\left(\int_0^t H_s^2 d\langle M,M\rangle\right)^{1/2}\right\|_{L^p}\left\|\left(\int_0^t K_s^2 d\langle N,N\rangle\right)^{1/2}\right\|_{L^q}.$$

We now introduce the class of processes that will become the stochastic integraters.

Definition 5.19. *An* continuous $\mathcal{F}_t$-semimartingale *is a continuous process which can be written*

$$X = M + V \tag{5.3}$$

where M is a continuous local martingale and V is a continuous adapted process having finite variation on a sequence of stopping time $\{\tau_n; n \geq 1\}, \lim_{n\to\infty} \tau_n = \infty$ bounded time intervals.

Exercise 5.20. *1. Show that V has zero quadratic variation.*

2. Using the notation above show that $\langle M, V\rangle = \lim_{mesh(\Pi_n)\to 0} C^{\Pi_n}(M, V) = 0$.

3. If $V_0 = 0$, show that the representation in (5.3) is unique.

Thus, if $X + V$ and $Y + W$ are semimartingales using the representation above with M and N local martingales and V and W processes having bounded variation on bounded intervals, then

$$\langle X, Y\rangle = \langle M, N\rangle.$$

5.2 Definition of the Stochastic Integral

Stochastic integration will begin proceed using the standard machine. The extension of the stochastic integral in the limiting process will be based on a Hilbert space isometry. Here are the spaces of interest.

Definition 5.21. *1. Define $\mathcal{H}^2$ to be the space of L^2-bounded martingales. These are martingales M, such that*

$$\sup_{t\geq 0} EM_t^2 < \infty.$$

Use $\mathcal{H}_0^2$ for those elements $M \in \mathcal{H}^2$ with $M_0 = 0$.

2. For $M \in \mathcal{H}^2$ define $\mathcal{L}^2(M)$ to be the space of progressively measurable processes K such that

$$||K||_M^2 = E[K^2 \cdot \langle M, M\rangle] = E[\int_0^\infty K_s^2 \, d\langle M, M\rangle_s] < \infty.$$

If $M \in \mathcal{H}^2$, then M is a uniformly integrable martingale, the its limit exists and is in $L^2(P)$. Thus we can place the following norm

$$||M||_{\mathcal{H}^2}^2 = EM_\infty^2.$$

Proposition 5.22. *A continuous local martingale M is in $\mathcal{H}_0^2$ if and only if $E\langle M, M\rangle_\infty < \infty$. In this case, $M^2 - \langle M, M\rangle$ is a uniformly integrable martingale.*

Proof. Let $\{\tau_n; n \geq 1\}$ be a sequence that reduces M. Then,

$$E[M^2_{\min\{t,\tau_n\}} I_{\{\tau_n>0\}}] - E[\langle M, M\rangle_{\min\{t,\tau_n\}} I_{\{\tau_n>0\}}] = E[M_0^2 I_{\{\tau_n>0\}}].$$

Because $M_\infty^* \in L^2(P)$, we can use the dominated and monotone convergence theorem to conclude that

$$E[M_\infty^2] - E[\langle M, M\rangle_\infty] = E[M_0^2]$$

and $E\langle M,M\rangle_\infty < \infty$.

Conversely, by the same equality,

$$E[M^2_{\min\{t,\tau_n\}} I_{\{\tau_n>0\}}] \le E[\langle M,M\rangle_\infty] + E[M_0^2] \le C.$$

By Fatou's lemma,

$$EM_t^2 \le \liminf_{n\to\infty} E[M^2_{\min\{t,\tau_n\}} I_{\{\tau_n>0\}}] \le C$$

Thus, $\{M_{\min\{t,\tau_n\}} I_{\{\tau_n>0\}}; n \ge 1, t \ge 0\}$ is uniformly integrable. Take $s<t$ and $A \in \mathcal{F}_s$, then

$$E[M_t; A] = \lim_{n\to\infty} E[M_{\min\{t,\tau_n\}} I_{\{\tau_n>0\}}; A] = \lim_{n\to\infty} E[M_{\min\{s,\tau_n\}} I_{\{\tau_n>0\}}; A] = E[M_s; A]$$

and M is a martingale.

Finally, note that

$$\sup_{t\ge 0} |M_t^2 - \langle M,M\rangle_t| \le (M^2)^*_\infty + \langle M,M\rangle_\infty,$$

an integrable random variable and thus $\{M_t^2 - \langle M,M\rangle_t; t \ge 0\}$ is uniformly integrable. □

Note that $||M||_{\mathcal{H}^2} = E\langle M,M\rangle_\infty$

A simple process has the form,

$$K = K_0 I_{\{0\}} + \sum_{j=1}^n K_j I_{(t_{j-1},t_j]}.$$

To be progressive, we must have that K_j is $\mathcal{F}_{t_{j-1}}$-measurable. The stochastic integral

$$(K \cdot M) = \int K_s \ dM_s = \sum_{j=1}^n K_j (M_{t_{j-1}} - M_{t_j}).$$

Note that

$$E(K \cdot M)^2 = E[\sum_{j=1}^n \sum_{k=1}^n K_j(M_{t_j} - M_{t_{j-1}}) K_k (M_{t_k} - M_{t_{k-1}})].$$

If $j<k$,

$$E[K_j(M_{t_j} - M_{t_{j-1}}) K_k (M_{t_k} - M_{t_{k-1}})] = E[K_j(M_{t_j} - M_{t_{j-1}}) K_k E[M_{t_k} - M_{t_{k-1}} | \mathcal{F}_{t_{k-1}}]] = 0.$$

Thus,

$$\begin{aligned}
E(K \cdot M)^2 &= E[\sum_{j=1}^n K_j^2 (M_{t_j} - M_{t_{j-1}})^2] = E[\sum_{j=1}^n K_j^2 E[(M_{t_j} - M_{t_{j-1}})^2 | \mathcal{F}_{t_{j-1}}]] \\
&= E[\sum_{j=1}^n K_j^2 E[(M_{t_j}^2 - M_{t_{j-1}}^2 | \mathcal{F}_{t_{j-1}}]] = E[\sum_{j=1}^n K_j^2 E[\langle M,M\rangle_{t_j} - \langle M,M\rangle_{t_{j-1}} | \mathcal{F}_{t_{j-1}}]] \\
&= E[\sum_{j=1}^n K_j^2 (\langle M,M\rangle_{t_j} - \langle M,M\rangle_{t_{j-1}})] = E[\int_0^\infty K_s^2 \ d\langle M,M\rangle_s] = ||K||_M^2
\end{aligned}$$

Thus, the mapping

$$K \to K \cdot M$$

is a Hilbert space isometry from L^2 to $\mathcal{L}^2(M)$.

Exercise 5.23. *Let N and M be square integrable martingales and let H and K be simple adapted functions, then*

$$E[(K \cdot M)(H \cdot N)] = E[(KH \cdot \langle M, N\rangle)].$$

Theorem 5.24. *Let $M \in \mathcal{H}_0^2$, then for each $K \in L^2(M)$, there is a unique element $K \cdot M \in \mathcal{H}_0^2$, such that*

$$\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle.$$

Proof. To verify uniqueness, note that if M^1 and M^2 are two martingales in $\mathcal{H}_0^2$, such that

$$\langle M^1, N\rangle = \langle M^2, N\rangle \quad \text{for every } N \in \mathcal{H}_0^2,$$

Then,

$$\langle M^1 - M^2, M^1 - M^2\rangle = 0.$$

and $M^1 - M^2$ is constant, hence 0.

By the Kunita-Watanabe inequality,

$$|E[\int_0^\infty K_s \; d\langle M, N\rangle_s]| \leq ||N||_{\mathcal{H}^2}||K||_{\mathcal{L}^2(M)}.$$

Thus the map

$$N \to E[K \cdot \langle M, N\rangle]$$

is a continuous linear functional on $\mathcal{H}^2$. Consequently, there exists an element $K \cdot M \in \mathcal{H}^2$ so that

$$E[(K \cdot M)_\infty N_\infty] = E[(K \cdot \langle M, N\rangle)_\infty].$$

Claim. $(K \cdot M)N - K \cdot \langle M, N\rangle$ is a martingale.

Recalling that element in $\mathcal{H}_0^2$ are uniformly integrable. Thus, for any stopping time τ, we have

$$\begin{aligned} E[(K \cdot M)_\tau N_\tau] &= E[E[(K \cdot M)_\infty | \mathcal{F}_\tau] N_\tau] = E[(K \cdot M)_\infty N_\tau] \\ &= E[(K \cdot M)_\infty N_\infty^\tau] = E[(K \cdot \langle M, N^\tau\rangle)_\infty] \\ &= E[(K \cdot \langle M, N\rangle^\tau)_\infty] = E[(K \cdot \langle M, N\rangle)_\tau] \end{aligned}$$

By the uniqueness of the covariation process, we have that $\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle$

□

This also shows that

$$||K \cdot M||_{\mathcal{H}^2}^2 = E[(K \cdot M)_\infty^2] = E[(K^2 \cdot \langle M, N\rangle)_\infty] = ||K||_{\mathcal{L}^2(M)}$$

and thus

$$K \to K \cdot M$$

is an isometry.

Definition 5.25. *The martingale $K \cdot M$ is called the* Itô stochastic integral *of K with respect to M. It is often also denoted by*

$$(K \cdot M)_t = \int_0^t K_s \; dM_s.$$

From out knowledge of Riemann-Stieltjes integrals, we have that

$$J_t = \int_0^t K_s \; d\langle M, N\rangle_s \quad \text{implies} \quad \int_0^t H_s \; dJ_s = \int_0^t H_s K_s \; d\langle M, N\rangle_s.$$

We would like a similar identity for stochastic integrals, namely

$$Y_t = \int_0^t K_s \; dM_s \quad \text{implies} \quad \int_0^t H_s \; dY_s = \int_0^t H_s K_s \; dM_s.$$

This is the content of the next theorem.

Proposition 5.26. *Let $K \in \mathcal{L}^2(M)$ and $H \in \mathcal{L}^2(K \cdot M)$, then*

$$(HK) \cdot M = H \cdot (K \cdot M).$$

Proof. Apply the theorem above twice to see that $\langle K \cdot M, K \cdot M\rangle = K^2\langle M, M\rangle$. Thus, we have $H \cdot K \in \mathcal{L}^2(K \cdot M)$. For $N \in \mathcal{H}_0^2$, we also have

$$\langle (HK) \cdot M, N\rangle = (HK) \cdot \langle M, N\rangle = H \cdot (K \cdot \langle M, N\rangle) = H \cdot \langle K \cdot M, N\rangle = \langle H \cdot (K \cdot M), N\rangle.$$

Now use the uniqueness of in the definition of the stochastic integral. □

Proposition 5.27. *Let τ be a stoppping time, then*

$$K \cdot M^\tau = KI_{[0,\tau]} \cdot M = (K \cdot M)^\tau.$$

Proof. Claim. $M^\tau = I_{[0,\tau]} \cdot M$.

Choose $N \in \mathcal{H}_0^2$,

$$\langle M^\tau, N\rangle = \langle M, N\rangle^\tau = I_{[0,\tau]} \cdot \langle M, N\rangle = \langle I_{[0,\tau]} \cdot M, N\rangle.$$

Now use the preceding proposition,

$$K \cdot M^\tau = K \cdot (I_{[0,\tau]} \cdot M) = KI_{[0,\tau]} \cdot M = I_{[0,\tau]} K \cdot M = I_{[0,\tau]} \cdot (K \cdot M) = (K \cdot M)^\tau,$$

which completes the proof. □

We now we shall use localization to extend the definition of the stochastic integral.

Definition 5.28. *Let M be a continuous local martingale, then $\mathcal{L}^2_{loc}(M)$ is the space of progressively measurable functions K for which there exists an increasing sequence of stopping times $\{\tau_n; n \geq 1\}$ increasing to infinity so that*

$$E[\int_0^{\tau_n} K_s^2 \, d\langle M, M\rangle_s] < \infty.$$

Exercise 5.29. *For any $K \in \mathcal{L}^2_{loc}(M)$, there exists a continuous local martingale $K \cdot M$ such that for any continuous local martingale N,*

$$\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle.$$

Definition 5.30.

1. *Call a progressively measurable process K* locally bounded *if there exists a sequence of stopping times $\{\tau_n; n \geq 1\}$ increasing to infinity and constants C_n such that $K^{\tau_n} \leq C_n$.*
2. *Let K be locally bounded and let $X = M + V$ be a continuous semimartingale, $X_0 = 0$, the* Itô stochastic integral *of K with respect to X is the continuous semimartingle*

$$K \cdot X = K \cdot M + K \cdot V$$

where $K \cdot M$ is the stochastic integral defined above and $K \cdot V$ is the usual Stieltjes integral.

Exercise 5.31. *The map $K \to K \cdot X$ has the following properties:*

1. *$H \cdot (M \cdot X) = (HM) \cdot X$.*
2. *$(K \cdot M)^\tau = (KI_{[0,\tau]}) \cdot M) = K \cdot M^\tau$.*
3. *If X is a local martingale or a process of finite variation, so is $K \cdot X$.*
4. *If K is an progressively measurable simple function,*

$$K = K_0 I_{\{0\}} + \sum_{j=1}^\infty K_j I_{(t_{j-1}, t_j]}, \ 0 = t_0 < t_1 < \cdots, \lim_{j\to\infty} t_j = \infty,$$

then

$$(K \cdot X)_t = \sum_{j=1}^\infty K_j (X_{\min\{t,t_j\}} - X_{\min\{t,t_{j-1}\}}.$$

Theorem 5.32 (bounded convergence theorem)**.** *Let X be a continuous semimartingale and let $\{K^n; n \geq 1\}$ be locally bounded processes converging to zero pointwise with $|K^n| \leq K$ for some locally bounded process K, then for any $T > 0$ and $\epsilon > 0$,*

$$\lim_{n\to\infty} P\{\sup_{0\leq t\leq T} |(K^n \cdot X)_t| > \epsilon\} = 0.$$

Proof. If X is a process of finite variation, then we can apply the bounded convergence theorem for each $\omega \in \Omega$.

If X is a local martingale and it τ reduces X, then by the bounded convergence theorem, $(K^n)^\tau$ converges to zero in $\mathcal{L}^2(X^\tau)$ and thus $(K^n \cdot X)^\tau$ converges to zero in $\mathcal{H}^2$. Now repeat the argument on the convergence of the quadratic variation over a partition for a local martingale. ∎

Exercise 5.33. *Let K be left-continuous and let $\{\Pi^n; n \geq 1\}$ be a sequence of partitions of $[0,\infty)$, with $\Pi^n = \{0 = t_0^n < t_0^n < \cdots\}$, $mesh(\Pi^n) \to 0$, then*

$$\lim_{n\to\infty} P\{\sup_{0\leq t\leq T} |(K \cdot X)_t - \sum_{j=1}^{\infty} K_{t_{j-1}^n}(X_{\min\{t,t_j^n\}} - X_{\min\{t,t_{j-1}^n\}}| > \epsilon\} = 0.$$

Exercise 5.34. *1. Let X be a continuous semimartingale and let the stochastic integrals $H \cdot X$ and $H \cdot X$ exists, then for $a, b \in R$*

$$\int_0^t (aH_s + bK_s)\ dX_s = a\int_0^t H_s\ dX_s + b\int_0^t K_s\ dX_s.$$

2. Let X and $\tilde{X}$ be two continuous semimartingales and let K and $\tilde{K}$ be two predictable locally bounded processes. Prove that if the finite dimensional distributions of (K, X) and $(\tilde{K}, \tilde{X})$ agree, then so do the finite dimensional distributions of $(K, X, K \cdot X)$ and $(\tilde{K}, \tilde{X}, \tilde{K} \cdot \tilde{X})$.

5.3 The Itô Formula

The Itô formula is the change of variable formula for stochastic integrals. The formula gives explicitily the decomposition of functions of a semimartingale as a local martingale and a process of bounded variation.

Proposition 5.35. *Let X be a continuous semimartingale, then*

$$X_t^2 = X_0^2 + 2\int_0^t X_s\ dX_s + \langle X, X\rangle_t.$$

Proof. Let $\{\Pi^n; n \geq 1\}$ be a sequence of partitions of $[0,\infty)$, $\Pi^n = \{0 = t_0^n < t_0^n < \cdots\}$, mesh$(\Pi^n) \to 0$. Expand to obtain,

$$\sum_{j=1}^{\infty}(X_{\min\{t,t_j\}} - X_{\min\{t,t_{j-1}\}})^2 = X_t^2 - X_0^2 - 2\sum_{j=1}^{\infty} X_{\min\{t,t_{j-1}\}}(X_{\min\{t,t_j\}} - X_{\min\{t,t_{j-1}\}}).$$

Then,

$$\sum_{j=1}^{\infty}(X_{\min\{t,t_j\}} - X_{\min\{t,t_{j-1}\}})^2 \to \langle X, X\rangle_t.$$

and

$$\sum_{j=1}^{\infty} X_{\min\{t,t_{j-1}\}}(X_{\min\{t,t_j\}} - X_{\min\{t,t_{j-1}\}}) \to \int_0^t X_s\ dX_s$$

uniformly in probability on compact intervals. □

Corollary 5.36. *(Integration by parts formula) Let X and Y be continuous semimartingales, then*

$$X_tY_t = X_0Y_0 + \int_0^t X_s\ dY_s + \int_0^t Y_s\ dX_s + \langle X, Y\rangle_t.$$

Proof. The corollary follows from polarization of the identity in the previous proposition. □

Example 5.37. *Let B be standard Brownian motion, then*

$$B_t^2 - t = \int_0^t B_s \; dB_s.$$

Definition 5.38. *A d-dimensional* vector (continuous) local martingale *is an $\mathbb{R}^d$-valued stochastic process $X = (X^1, \dots, X^d)$ such that each component is a (continuous) local martingale. A* complex (continuous) local martingale *is one in which both the real and imaginary parts are (continuous) local martingales.*

Theorem 5.39 (Itô formula)**.** *Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$ and $X = (X^1, \dots, X^d)$ be a vector continuous semimartingale, then $f \circ X$ is a continuous semimartingale and*

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) \; dX_s^i + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \; d\langle X^i, X^j \rangle_s.$$

Thus, the class of semimartingales is closed under composition with C^2-functions.

Proof. If the formula holds for f, then, by the integration by parts formula, the formula holds for $g(x) = x^k f(x)$. Thus, the Itô formula holds for polynomials. Now, for any compact set K, stop the process at τ_K. By the Stone-Weierstrass theorem, any $f \in C^2(K, \mathbb{R})$ can be uniformly approximated by polynomials. Thus the formula holds up to time τ_K by the stochastic dominated convergence theorem. Now, extend the result to all of $\mathbb{R}^d$. □

From a vector semimartingale X we often write the stochastic integral

$$Y_t = Y_0 + \sum_{i=1}^d \int_0^t H_s^i \; dX_s^i$$

in *differential form*

$$dY_t = \sum_{i=1}^d H_s^i \; dX_s^i.$$

Using this notation, the Itô formula becomes

$$df(X_t) + \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_s) \; dX_s^i + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \; d\langle X^i, X^j \rangle_s.$$

Example 5.40. 1. *let $f : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{C}$ satisfy*

$$\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0$$

Then, if M is a local martingale, then so is the process $\{f(M_t, \langle M, M \rangle)_t; t \geq 0\}$. In particular,

$$\mathcal{E}_t(\lambda M) = \exp(\lambda M_t - \frac{\lambda^2}{2} \langle M, M \rangle_t), \quad \lambda \in \mathbb{C}$$

is a local martingale.

This local martingale satisfies the stochastic differential equation

$$dY_t = \lambda Y_t \ dM_t.$$

2. *Note that for standard Brownian motion,*

$$\exp(\lambda B_t - \frac{\lambda^2}{2}t)$$

is a martingale. Use the standard machine to show that

$$\mathcal{E}_t^f(B) = \exp\left(\int_0^t f(t) \ dB_s - \frac{1}{2}\int_0^t f(t)^2 \ ds\right)$$

is a martingale.

3. *If B is a d-dimensional Brownian motion and if $f \in C^2(R^+ \times R^d)$, then*

$$f(t, B_t) - \int_0^t (\frac{1}{2}\Delta f + \frac{\partial f}{\partial t})(s, B_s) \ ds$$

is a martingale. Thus, if f is harmonic in $\mathbb{R}^d$, then $f \circ B$ is a local martingale. In addition, $\frac{1}{2}\Delta$ is the restriction of the generator of B to $C^2(\mathbb{R}^d)$.

Exercise 5.41. *If $B = (B^1, \ldots, B^d)$ is d-dimensional Brownian motion then $\langle B^i, B^j \rangle = \delta_{ij} t$.*

Example 5.42. *If $f : \mathbb{R}^d \to \mathbb{R}$ is a function of the radius, then the Laplacian can be replaced by its radial component. Writing $R_t = |B_t|$, we have*

$$f(t, R_t) - \int_0^t (Gf + \frac{\partial f}{\partial t})(s, R_s) \ ds$$

is a martingale if G is the generator for the Bessel process in d dimensions.

Check that $s(r) = \ln r$ is a scale function for the Bessel process in 2 dimensions and that $s(r) = -r^{2-d}$ is a scale function for dimension $d \geq 3$. So if $r_i < r < r_o$, then

$$P_r\{\tau_{r_o} > \tau_{r_i}\} = \frac{s(r_o) - s(r)}{s(r_o) - s(r_i)}.$$

Now let $r_o \to \infty$, then for $d = 2$,

$$P_r\{\tau_{r_i} < \infty\} = 1.$$

Thus, in 2 dimensions, Brownian motion must return infinitely often to any open set containing the origin. For $d \geq 3$

$$P_r\{\tau_{r_i} < \infty\} = \left(\frac{r_i}{r}\right)^{d-2}$$

and Brownian has a positive probability of never returning to a given neighborhood of the origin. In short, Brownian motion is recurrent in dimension 2 and transient in dimensions higher than 2. This conforms with similar results for random walks.

Theorem 5.43 (Lévy's characterization theorem)**.** *Assume that X is a $\mathbb{R}^d$ valued stochastic process satisfying $X_0 = 0$. Then the following are equivalent.*

1. *X is d-dimensional standard Brownian motion.*
2. *X is a continuous local martingale and $\langle X^i, X^j \rangle = \delta_{ij} t$.*

Proof. $1 \to 2$ is the exercise above.
$(2 \to 1)$ Take $\lambda = i$ and $M = \xi \cdot X$ in the exponential martingale, then

$$\langle M, M \rangle_t = |\xi|^2 t$$

and we have the local martingale

$$\exp(i\xi \cdot X_t + \frac{1}{2}|\xi|^2 t).$$

Because it is bounded, it is a martingale. Thus, for $s < t$

$$E[\exp(i\xi \cdot X_t + \frac{1}{2}|\xi|^2 t)|\mathcal{F}_s] = \exp(i\xi \cdot X_s + \frac{1}{2}|\xi|^2 s).$$

or

$$E[\exp(i\xi \cdot (X_t - X_s)|\mathcal{F}_s] = \exp(-\frac{1}{2}|\xi|^2 (t-s)).$$

Thus, X has stationary and independent Gaussian increments and X_1 has a normal distribution with the identity matrix as its covariance matrix. Thus, it is standard Brownian motion. □

5.4 Stochastic Differential Equations

Definition 5.44. *Let f and g be adapted functions taking values in $\mathbb{R}^{d\times r}$ and $\mathbb{R}^d$ respectively and let B be a standard r-dimensional Brownian motion. Consider the* stochastic differential equation

$$X_t^i = X_0^i + \sum_{j=1}^r (\int_0^t f_{ij}(s, X)\ dB_s^j + \int_0^t g_i(s, X)\ ds)$$

or, in vector differential form,

$$dX_t = f(s, X)\ dB_s + g_i(s, X)\ ds.$$

A solution *is a pair (X, B) of $\mathcal{F}_t$-adapted processes that satisfy this equation.*

The variety of notions of equivalence for stochastic processes leads to differing definitions of uniqueness of uniqueness for solutions to stochastic differential equations.

Definition 5.45.

1. *Let (X, B) and $(\tilde{X}, B)$ be two solutions on the same probability space with the same filtration and the same initial conditions. Then the stochastic differential equation above is said to satisfy* pathwise uniqueness *if X and $\tilde{X}$ are indistinguishable.*
2. *Let (X, B) and $(\tilde{X}, \tilde{B})$ be two solutions so that the distributions of X_0 and $\tilde{X}_0$ are equal. Then the stochastic differential equation above is said to satisfy* uniqueness in law *if X and $\tilde{X}$ are two versions of the same process.*

We will be focused on pathwise uniquesness. We begin with the basic estimate used to study the dependence of a solution on initial condiations.

Theorem 5.46. *(Gronwall's inequality) Let $f, g : [0,T] \to R$ be continuous and nonnegative. Suppose*

$$f(t) \leq A + \int_0^t f(s)g(s)\ ds, \quad A \geq 0.$$

Then, for $t \in [0,T]$,

$$f(t) \leq A \exp(\int_0^t g(s)\ ds).$$

Proof. First suppose $A > 0$ and set

$$h(t) = A + \int_0^t f(s)g(s)\ ds.$$

Thus $h(t) > 0$. Then,

$$h'(t) = f(s)g(s) \leq h(t)g(t) \quad \text{or} \quad \frac{h'(t)}{h(t)} \leq g(t).$$

Integration gives

$$f(t) \leq h(t) \leq A \exp(\int_0^t g(s)\ ds).$$

For $A = 0$, take limits to see that f is the zero function. □

The basic hypothesis for existence and uniqueness that we will use is an assumption of the following Lipschitz condition.

Let $x, \tilde{x} \in C_{\mathbb{R}^r}[0,\infty)$. There exists a constant K such that for every $t > 0$,

$$|f(t,x) - f(t,\tilde{x})| + |g(t,x) - g(t,\tilde{x})| \leq K \sup_{s \leq t} |x_s - \tilde{x}_s|.$$

Theorem 5.47. *Let $\mathcal{F}_t$ be a right-continuous complete filtration on $(\Omega, \mathcal{F}P)$ and let Z be a continuous r-dimensional semimartingale. If the progressive process $f(t,X)$ satisfies the the Lipschitz condition above and if for every $t > 0$,*

$$E[\sup_{0 \leq s \leq t} \left(\int_0^s f(r,\bar{x})\ dr\right)^2] < \infty,$$

where $\bar{x}$ is a constant function for some $x_0 \in R^r$, then there is a pathwise unique process X such that

$$X_t = x + \int_0^t f(s,X)\ dZ_s.$$

Proof. Note that the Lipschitz condition guarantees that if the condition above holds for one choice x_0 then it holds for all $x \in R^r$.

Write the semimartingale $Z = M + A$ in its canonical decomposition. Let $|A|_t$ denote the variation of A up to time t

Case 1. The measures associated to $\langle M, M\rangle$ and $|A|$ are absolutely continuous with respect to Lebesgue measure and for all i and j.

$$\frac{d\langle M^i, M^j\rangle_t}{dt} \leq 1, \quad \frac{d|A|_t}{dt} \leq 1.$$

Fix $x \in R^r$ and define the mapping

$$(SU)_t = x + \int_0^t f(s, U)\ dZ_s.$$

and the norm

$$||U - V||_{2,t}^2 = E[\sup_{0\leq s\leq t} |U_s - V_s|^2].$$

Then, by the Doob and Cauchy-Schwartz inequality.

$$\begin{aligned}
||SU - SV||_{2,t}^2 &\leq 2E[\sup_{0\leq s\leq t} |\int_0^s (f(r,U) - f(r,V))\ dM_s|^2] + 2E[\sup_{0\leq s\leq t} |\int_0^s (f(r,U) - f(r,V))\ d|A|_s|^2] \\
&\leq 8E[\int_0^t (f(r,U) - f(r,V))\ dM_s|^2 + 2E[\sum_{j=1}^r |A^i|_t \int_0^t (f_j(r,U) - f_j(r,V))^2\ d|A^j|_s|] \\
&\leq 8E[\sum_{i=1}^r \sum_{j=1}^r \int_0^t (f_i(r,U) - f_i(r,V))(f_j(r,U) - f_j(r,V))\ d\langle M^i, M^j\rangle_s] \\
&\quad +2tE[\sum_{j=1}^r \int_0^t (f_j(r,U) - f_j(r,V))^2\ d|A^j|_s|] \\
&\leq 2K^2(4r^2 + rt) \int_0^t ||U - V||_{2,r}^2\ dr
\end{aligned}$$

This sets us up for the Picard-Lindelöf iteration scheme. Set

$$X^0 = x, \quad \text{and} \quad X^n = SX^{n-1}$$

and write

$$C = 2K^2(4r^2 + rT).$$

Now, check, by induction, that

$$||X^{n+1} - X^n||_{2,t}^2 \leq \frac{(Ct)^n}{n!} ||X^1 - X^0||_{2,T}^2.$$

Thus, $\{X^n; n \geq 1\}$ is a Cauchy sequence in $||\cdot||_{2,T}$. Call the limit X. Then X is a continuous process and satisfies $X = XS$. In other words, X is a solution.

To show that X is unique, let $\tilde{X}$ be a second solution. Define $\tau_k = \inf\{t > 0 : |X_t| \geq k \text{ or } |\tilde{X}_l| \geq k\}$. Use Gronwall's inequality, with $f(t) = ||X^{\tau_k} - \tilde{X}^{\tau_k}||_{2,t}$ and $g(t) = C$ to see that $\tilde{X}_0 = x$, implies $\tilde{X}^{\tau_k} = X^{\tau_k}$. Now, let T and k go to infinity.

Case 2. The general case.

Define the strictly increasing process

$$\bar{A}_t = t + \sum_{i=1}^{r}\sum_{j=1}^{r}\langle M^i, M^j\rangle_t + \sum_{i=1}^{r}|A^i|_t$$

Consider the time change $C_t = \tilde{A}_t^{-1}$ and write

$$\tilde{M}_t = M_{C_t}, \quad \tilde{A}_t = A_{C_t}, \quad \text{and} \quad \tilde{Z}_t = \tilde{M}_t + \tilde{A}_t.$$

Then $\tilde{Z}$ satisfies the circumstances in case 1 with the filtration $\tilde{\mathcal{F}}_t = \mathcal{F}_{C_t}$. Thus,

$$\tilde{X}_t = x + \int_0^t f(s, \tilde{X})\, d\tilde{Z}_s$$

has a unique solution. Now, $X_t = \tilde{X}_{\tilde{A}_t}$ is the unique solution to the equation in the statement of the theorem. □

5.5 Itô Diffusions

Definition 5.48. *A continuous stochastic process X is called a time homogeneous* Itô diffusion *is there exists measurable mappings*

1. $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times r}$, *(*the diffusion matrix*), and*
2. $b : \mathbb{R}^d \to \mathbb{R}^d$, *(*the drift*)*

and an r-dimensional Brownian motion B so that

$$X_t^i = X_0^i + \sum_{j=1}^{r} \sigma_{ij}(X_s)\, dB_s^j + \int_0^t b_i(X_s)\, ds$$

has a solution that is unique in law.

If σ and b satisfy an appropriate Lipschitz hypothesis, then we have pathwise uniqueness. To show that this suffices, we have the following.

Theorem 5.49. *Let σ and b be locally bounded and Borel measurable, and let α be a probability measure on $\mathbb{R}^d$. Then pathwise uniqueness to the Itô stochastic differential equation above implies distribution uniqueness.*

We begin with a lemma.

Lemma 5.50. *The following condition is sufficient for uniqueness in law:*

Let $x \in \mathbb{R}^d$. Whenever (X, B) and $(\tilde{X}, \tilde{B})$ are two solutions satisfying

$$X_0 = x, \quad \tilde{X}_0 = x, \quad a.s.$$

Then the distribution of X and $\tilde{X}$ are equal.

Proof. Let P be the distribution of (X, B) on the space $C_{\mathbb{R}^r}[0,\infty)$ and write the pair (ξ, β) for the coordinate functions for the first d and last r coordinates respectively. Because $C_{\mathbb{R}^{d+r}}[0,\infty)$ is a Polish space, there is a regular conditional probability

$$P(B|\mathcal{B}_0)(\xi,\beta) = P(B,(\xi,\beta)).$$

In particular, β has the distribution of r-dimensional Brownian motion. The triple $(X, B, \sigma(X)\cdot B)$ and $(\xi, \beta, \sigma(\xi)\cdot\beta)$ have the same distribution and thus (ξ, β) solves the stochastic differential equation.

Repeat this letting $\tilde{P}$ be the distribution of $(\tilde{X}, \tilde{B})$. Then, the regular conditional probabilities agree

$$\tilde{P}(B,(\xi,\beta)) = P(B,(\xi,\beta))$$

and so if X_0 and $\tilde{X}_0$ have the same distribution, then $P = \tilde{P}$. □

Theorem 5.51. *If pathwise uniqueness holds for an Itô type stochastic differential equation, then uniqueness in law holds. Moreover, there exists a measureable map*

$$F : C_{R^d}[0,\infty) \to C_{R^r}[0,\infty)$$

such that $F(\beta)$ is adapted to the natural filtration $\mathcal{B}^r_t = \sigma\{\xi_s; s \leq t\}$ and for every solution (X, B), we have that

$$X_t(\omega) = F(B(\omega))_t \quad a.s.$$

Proof. Let (X, B) be a solution on (Ω, P) and let $(\tilde{X}, \tilde{B})$ be a solution on $(\tilde{\Omega}, \tilde{P})$. With, for some $x \in R^d$,

$$P\{X_0 = x\} = \tilde{P}\{\tilde{X}_0 = x\} = 1.$$

For the solution (X, B) let Q be the image of P under the map

$$\omega \to (X(\omega), B(\omega)).$$

For the solution $(\tilde{X}, \tilde{B})$ let $\tilde{Q}$ be the image of $\tilde{P}$ under the map

$$\tilde{\omega} \to (\tilde{X}(\tilde{\omega}), \tilde{B}(\tilde{\omega})).$$

In addition, let W be Wiener measure, that is the law of r-dimensional Brownian motion, Define a probability measure Π on $C_{R^d}[0,\infty) \times C_{R^d}[0,\infty) \times C_{R^d}[0,\infty)$ by

$$\Pi(d\xi, d\tilde{\xi}, d\beta) = Q(\beta, d\xi)\tilde{Q}(\beta, d\tilde{\xi})W(d\beta)$$

and consider the filtration

$$\mathcal{F}_t = \sigma\{(\xi_s, \tilde{\xi}_s, \beta_s); 0 \leq s \leq t\}.$$

Claim 1. β is an r-dimensional $\mathcal{F}_t$ Brownian motion.

Let $A \in \sigma\{\xi_s; 0 \leq s \leq t\}$, $\tilde{A} \in \sigma\{\tilde{\xi}_s; 0 \leq s \leq t\}$, and $B \in \sigma\{\beta_s; 0 \leq s \leq t\}$ and let $s, t > 0$, then

$$\begin{aligned} E_x[\exp\langle i\theta, \beta_{t+s} - \beta_t\rangle; A \cap \tilde{A} \cap B] &= \int_B \exp\langle i\theta, \beta_{t+s} - \beta_t\rangle Q(\beta, A)\tilde{Q}(\beta, \tilde{A})W(d\beta) \\ &= \exp(-\frac{1}{2}|\theta|^2 t)\int_B Q(\beta, A)\tilde{Q}(\beta, \tilde{A})W(d\beta) \\ &= \exp(-\frac{1}{2}|\theta|^2 t)\Pi(A \times \tilde{A} \times B) \end{aligned}$$

Thus β has Gaussian independent increments with identity covariance matrix.

Claim 2. (ξ, β) and $(\tilde{\xi}, \beta)$ are two solutions to the stochastic differential equation on the space

$$(C_{R^d}[0,\infty) \times C_{R^d}[0,\infty) \times C_{R^r}[0,\infty), \Pi)$$

with filtration $\{\mathcal{F}_t; t \geq 0\}$.

The joint law of $(X, B, \sigma(X) \cdot B)$ under P is the same as the joint law of $(\xi, \beta, \sigma(\xi) \cdot \beta)$ under Π. Similarly, the joint law of $(\tilde{X}, \tilde{B}, \sigma(\tilde{X}) \cdot \tilde{B})$ under $\tilde{P}$ is the same as the joint law of $(\tilde{\xi}, \beta, \sigma(\tilde{\xi}) \cdot \beta)$ under Π. Because $\xi_0 = \tilde{\xi}_0 = x$ a.s. Π, the property of pathwise uniqueness implies that ξ and $\tilde{\xi}$ are indistinguishable. Thus, uniqueness in law holds.

Solution is Markov. □

Exercise 5.52. *If X is an Itô diffusion and if $f \in C^2(R^d, R)$, then*

$$\begin{aligned}
f(X_t) &= f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s)\, dX_s^i + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)\, d\langle X^i, X^j \rangle_s \\
&= f(X_0) + \sum_{i=1}^d \int_0^t (b_i(X_s)\frac{\partial f}{\partial x^i}(X_s) + \frac{1}{2}(\sigma\sigma^T)_{ij}(X_s))\, ds + \sum_{i=1}^d \sum_{k=1}^r \int_0^t \sigma_{ik}(X_s)\frac{\partial f}{\partial x^i}(X_s)\, dB_s^k.
\end{aligned}$$

In particular, we have the martingales

$$M_t^f = f(X_0) + \int_0^t Gf(X_s)\, ds.$$

with

$$Gf(x) = \sum_{i=1}^d b_i(x)\frac{\partial f}{\partial x^i}(x) + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d a_{ij}(x)\frac{\partial^2 f}{\partial x^i \partial x^j}(x)$$

and $a = \sigma\sigma^T$.
square roots

Definition 5.53. *A stochastic process X is called a* solution to the martingale problem *for the collection $\mathcal{A}$ if*

$$f(X_t) - \int_0^t g(X_s)\, ds$$

is a martingale for every $(f,g) \in \mathcal{A}$.

For a probability distribution α, we say that X is a solution to the martingale problem for $(\mathcal{A}, \alpha)$ if X is a solution to the martingale problem and $\alpha = PX_0^{-1}$.

Uniqueness holds for the martingale problem $(\mathcal{A}, \alpha)$ if any two solutions have the same finite dimensional distributions. In this case, we say that the martingale problem is well posed. *If uniqueness holds for any initial distribution α, then we say that the martingale problem $\mathcal{A}$ is well posed.*

The relationship to Itô diffusions is as follows.

Theorem 5.54. *Let*

$$\sigma : R^d \to R^{d\times d}, \quad \textit{and} \quad b : R^d \to R^d$$

and let α be a probability measure on R^d. Define A as above and $\mathcal{A} = \{(f, Af) : f \in C_c^\infty(R^d)\}$. Then there exists a solution to the Itô stochastic differential equation corresponding to σ and b with initial distribution α if and only if there exists a solution to the martingale problem $(\mathcal{A}, \alpha)$.

Moreover, uniqueness in law to the stochastic differential equation corresponding to σ and b with initial distribution α if and only if the martingale problem $(\mathcal{A}, \alpha)$ is well posed.

This will follow from a sequence of lemmas:

Lemma 5.55. *Assume that σ and b are locally bounded and measurable. In addition, σ^{-1} exists and is locally bounded. If X is a solution to the $C_{R^d}[0,\infty)$-martingale problem for $(\mathcal{A}, \alpha)$ with respect to a complete filtration $\{\mathcal{F}_t; t \geq 0\}$, then there exists a d-dimensional Brownian motion B such that (X, B) is a solution to the stochastic differential equation.*

Proof. For every $f \in C_c^\infty(R^d)$

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Af(X_s)\, ds$$

is a continuous martingale, we have that this expression gives a continuous local martingale for every $f \in C^\infty(R^d)$. Choose $f(x) = x^i$ to obtain the martingale

$$M_t^i = X_t^i - X_0 - \int_0^t b_i(X_s)\, ds$$

is a local martingale. Choose $f(x) = x^i x^j$,

$$\langle M^i, M^j \rangle_t = \langle X^i, X^j \rangle_t = \int_0^t a_{ij}(X_s)\, ds.$$

Claim.

$$B_t = \int_0^t \sigma^{-1}(X_s)\, dM_s$$

is a d-dimensional $\mathcal{F}_t$-adapted Brownian motion.

B is a vector valued continuous local martingale and $B_0 = 0$. Thus, by Lévy's characterization, it suffices to show that $\langle B^i, B^j \rangle_t = \delta_{ij} t$.

$$\begin{aligned}
\langle B^i, B^j \rangle_t &= \sum_{k=1}^d \sum_{\ell=1}^d \langle \sigma^{-1}(X)_{ik} \cdot M^k, \sigma^{-1}(X)_{j\ell} \cdot M^\ell \rangle_t \\
&= \sum_{k=1}^d \sum_{\ell=1}^d \int_0^t (\sigma_{ik}^{-1} a_{k\ell} (\sigma^T)_{\ell j}^{-1})(X_s)\, ds = \delta_{ij} t.
\end{aligned}$$

Thus,

$$\int_0^t \sigma(X_s)\, dB_s = \int_0^t dM_s = X_t - X_0 \int_0^t b(X_s)\, ds$$

and the stochastic differential equation has a solution. □

To consider the case of singular σ we have the following,

Lemma 5.56. *There exists Borel measurable functions*

$$\rho, \eta : R^d \to R^{d\times d}$$

such that

$$\rho a \rho^T + \eta\eta^T = I_d, \quad \sigma\eta = 0, \quad (I_d - \sigma\rho)a(I_d - \sigma\rho)^T = 0.$$

Lemma 5.57. *Let B' be a d-dimensional $\mathcal{F}'_t$-Brownian motion on (Ω', P'). Define*

$$\tilde{\Omega} = \Omega \times \Omega', \quad \tilde{P} = P \times P', \quad \text{and} \quad \tilde{X}_t(\omega, \omega') = X_t(\omega).$$

Then there exists a d-dimensional $\mathcal{F}_t \times \mathcal{F}'_t$-Brownian motion $\tilde{B}$ such that $(\tilde{X}, \tilde{B})$ is a solution to the stochastic differential equation.

Proof. With M as above, define

$$\tilde{M}_t(\omega, \omega') = M_t(\omega) \quad \text{and} \quad \tilde{B}'_t(\omega, \omega') = B'_t(\omega').$$

Because $\tilde{M}$ and $\tilde{B}$ are independent, $\langle M^k, \tilde{B}^\ell\rangle_t = 0$ for $k, \ell = 1, \ldots d$.

Claim.

$$\tilde{B}_t = \int_0^t \rho(X_s)\ d\tilde{M}_s + \int_0^t \eta(X_s)\ d\tilde{B}'_s$$

defines a d-dimensional Brownian motion.

As before, we have that $\tilde{B}$ is a vector valued continuous local martingale.

$$\begin{aligned}
\langle \tilde{B}^i, \tilde{B}^j\rangle_t &= \sum_{k=1}^d \sum_{\ell=1}^d \langle \rho(X)_{ik}\cdot \tilde{M}^k, \rho(X)_{j\ell}\cdot \tilde{M}^\ell\rangle_t + \sum_{k=1}^d \sum_{\ell=1}^d \langle \eta(X)_{ik}\cdot \tilde{B}'^k, \eta(X)_{j\ell}\cdot \tilde{B}'^\ell\rangle_t \\
&= \sum_{k=1}^d \sum_{\ell=1}^d \int_0^t (\rho_{ik} a_{k\ell}\rho_{j\ell})(X_s)\ ds + \sum_{k=1}^d \sum_{\ell=1}^d \int_0^t (\eta_{ik}\delta_{k\ell}\eta_{jl})(X_s)\ ds \\
&= \int_0^t (\rho a \rho^T + \eta\eta^T)_{ij}(X_s)\ ds = \delta_{ij} t
\end{aligned}$$

To show that this Brownian can be used to solve the stochastic differential equation, note that,

$$\begin{aligned}
\int_0^t \sigma(\tilde{X}_s)\ d\tilde{B}_s &= \int_0^t (\sigma\rho)(\tilde{X}_s)\ d\tilde{M}_s + \int_0^t (\sigma\eta)(\tilde{X}_s)\ d\tilde{B}_s \\
&= \int_0^t (\sigma\rho)(\tilde{X}_s)\ d\tilde{M}_s = \tilde{M}_t - \int_0^t (I_d - \sigma\rho)(\tilde{X}_s)\ d\tilde{M}_s \\
&= \tilde{M}_t = \tilde{X}_t \tilde{X}_0 - \int_0^t b(X_s)\ ds
\end{aligned}$$

because

$$\langle \sum_{k=1}^d (I_d - \sigma\rho)(\tilde{X}_s)_{ik}\cdot \tilde{M}^k, \sum_{\ell=1}^d (I_d - \sigma\rho)_{ik}(\tilde{X}_s)\cdot \tilde{M}^k\rangle_t = \sum_{k=1}^d \sum_{\ell=1}^d \int_0^t (I_d - \sigma\rho)_{ik} a_{k\ell}(I_d - \sigma\rho)_{i\ell}(\tilde{X}_s)\ ds = 0$$

and the lemma follows. □