Martingale problems and stochastic equations

- Characterizing stochastic processes by their martingale properties
- Markov processes and their generators
- Forward equations
- Other types of martingale problems
- Change of measure
- Martingale problems for conditional distributions
- Stochastic equations for Markov processes
- Filtering
- Time change equations
- Bits and pieces
- Supplemental material
- Exercises
- References
Supplemental material

- Background: Basics of stochastic processes
- Stochastic integrals for Poisson random measures
- Technical lemmas
1. Characterizing stochastic processes by their martingale properties

- Lévy’s characterization of Brownian motion
- Definition of a counting process
- Poisson processes
- Martingale properties of the Poisson process
- Strong Markov property for the Poisson process
- Intensities
- Counting processes as time changes of Poisson processes
- Martingale characterizations of a counting process
- Multivariate counting processes

Brownian motion

A Brownian motion is a continuous process with independent, Gaussian increments. A Brownian motion $W$ is standard if the increments $W(t + r) - W(t)$, $t, r \geq 0$, have mean zero and variance $r$.

$W$ is a martingale:

$$E[W(t + r)|\mathcal{F}_t^W] = E[W(t + r) - W(t)|\mathcal{F}_t^W] + W(t) = W(t),$$

$\mathcal{F}_t^W = \sigma(W(s) : s \leq t)$.

$W$ has quadratic variation $t$, that is

$$[W]_t = \lim_{\sup|t_{i+1}-t_i|=0} \sum_i (W(t \land t_{i+1}) - W(t \land t_i))^2 = t$$

in probability, or more precisely,

$$\lim_{\sup|t_{i+1}-t_i|=0} \sup_{t \leq T} \left| \sum_i (W(t \land t_{i+1}) - W(t \land t_i))^2 - t \right| = 0$$

in probability for each $T > 0$.

$W(t)^2 - t$

is a martingale.
Lévy’s characterization of Brownian motion

**Theorem 1.1** Let $M$ be a continuous local martingale with $[M]_t = t$. Then $M$ is standard Brownian motion.

**Remark 1.2** Note that

$$E[M^{\tau_n}(t)^2] = E[[M^{\tau_n}]_t] = E[\tau_n \wedge t] \leq t$$

and $E[\sup_{s \leq t} M^{\tau_n}(s)^2] \leq 4E[M^{\tau_n}(t)^2] \leq 4t$ and it follows by the dominated convergence theorem and Fatou’s lemma that $M$ is a square integrable martingale.
Proof. Applying Itô’s formula,

\[ e^{i\theta M(t)} = 1 + \int_0^t i\theta e^{i\theta M(s)} dM(s) - \frac{1}{2} \theta^2 \int_0^t e^{i\theta M(s)} ds, \]

where the second term on the right is a martingale. Consequently,

\[ E[e^{i\theta M(t+r)} | \mathcal{F}_t] = e^{i\theta M(t)} - \frac{1}{2} \theta^2 \int_t^{t+r} E[e^{i\theta M(s)} | \mathcal{F}_t] ds \]

and

\[ \varphi_t(\theta, r) \equiv E[e^{i\theta (M(t+r) - M(t))} | \mathcal{F}_t] = 1 - \frac{1}{2} \theta^2 \int_t^{t+r} E[e^{i\theta M(s) - M(t)} | \mathcal{F}_t] ds \]

\[ = 1 - \frac{1}{2} \theta^2 \int_0^r \varphi_t(\theta, u) du \]

so \( \varphi_t(\theta, r) = e^{-\frac{1}{2} \theta^2 r} \). It follows that for \( 0 = t_0 < t_1 < \cdots < t_m \),

\[ E[\prod_{k=1}^m e^{i\theta_k (M(t_k) - M(t_{k-1}))}] = \prod e^{-\frac{1}{2} \theta_k^2 (t_k - t_{k-1})} \]

and hence \( M \) has independent Gaussian increments. \( \square \)
Definition of a counting process

Definition 1.3  \( N \) is a counting process if \( N(0) = 0 \) and \( N \) is constant except for jumps of \( +1 \).

\( D^c[0, \infty) \) will denote the space of possible counting paths that are finite for all time. \( D^c_\infty[0, \infty) \supset D^c[0, \infty) \) is the larger space allowing the paths to hit infinity in finite time.
**Poisson processes**

A Poisson process is a model for a series for random observations occurring in time. For example, the process could model the arrivals of customers in a bank, the arrivals of telephone calls at a switch, or the counts registered by radiation detection equipment.

Let $N(t)$ denote the number of observations by time $t$. In the figure above, $N(t) = 6$. Note that for $t < s$, $N(s) - N(t)$ is the number of observations in the time interval $(t, s]$. We make the following assumptions about the model.

1) Observations occur one at a time.

2) Numbers of observations in disjoint time intervals are independent random variables, i.e., if $t_0 < t_1 < \cdots < t_m$, then $N(t_k) - N(t_{k-1}), k = 1, \ldots, m$ are independent random variables.

3) The distribution of $N(t + a) - N(t)$ does not depend on $t$. 
Characterization of a Poisson process

**Theorem 1.4** Under assumptions 1), 2), and 3), there is a constant $\lambda > 0$ such that, for $t < s$, $N(s) - N(t)$ is Poisson distributed with parameter $\lambda(s - t)$, that is,

$$P\{N(s) - N(t) = k\} = \frac{(\lambda(s - t))^k}{k!}e^{-\lambda(s-t)}.$$ 

**Proof.** Let $N_n(t)$ be the number of time intervals $(\frac{k}{n}, \frac{k+1}{n}]$, $k = 0, \ldots, \lfloor nt \rfloor$ that contain at least one observation. Then $N_n(t)$ is binomially distributed with parameters $n$ and $p_n = P\{N(\frac{1}{n}) > 0\}$. Then

$$P\{N(1) = 0\} = P\{N_n(1) = 0\} = (1 - p_n)^n$$

and $np_n \to \lambda \equiv -\log P\{N(1) = 0\}$, and the rest follows by standard Poisson approximation of the binomial. $\square$
Interarrival times

Let $S_k$ be the time of the $k$th observation. Then

$$P\{S_k \leq t\} = P\{N(t) \geq k\} = 1 - \sum_{i=0}^{k-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t}, \quad t \geq 0.$$ 

Differentiating to obtain the probability density function gives

$$f_{S_k}(t) = \begin{cases} 
\frac{1}{(k-1)!} \lambda (\lambda t)^{k-1} e^{-\lambda t} & t \geq 0 \\
0 & t < 0 
\end{cases}$$

**Theorem 1.5** Let $T_1 = S_1$ and for $k > 1$, $T_k = S_k - S_{k-1}$. Then $T_1, T_2, \ldots$ are independent and exponentially distributed with parameter $\lambda$. 
Martingale properties of the Poisson process

**Theorem 1.6 (Watanabe)** If $N$ is a Poisson process with parameter $\lambda$, then $N(t) − \lambda t$ is a martingale. Conversely, if $N$ is a counting process and $N(t) − \lambda t$ is a martingale, then $N$ is a Poisson process.

**Proof.**

\begin{align*}
E[e^{i\theta(N(t+r)−N(t))} | F_t] &= 1 + \sum_{k=0}^{n-1} E[(e^{i\theta(N(s_{k+1})−N(s_k))} − 1 − (e^{i\theta} − 1)(N(s_{k+1}) − N(s_k)))e^{i\theta(N(s_k)−N(t))} | F_t] \\
&\quad + \sum_{k=0}^{n-1} \lambda(s_{k+1} − s_k)(e^{i\theta} − 1)E[e^{i\theta(N(s_k)−N(t))} | F_t]
\end{align*}

The first term converges to zero by the dominated convergence theorem, so we have

\begin{align*}
E[e^{i\theta(N(t+r)−N(t))} | F_t] &= 1 + \lambda(e^{i\theta} − 1) \int_0^r E[e^{i\theta(N(s)−N(t))} | F_t] ds \\
\text{and } E[e^{i\theta(N(t+r)−N(t))} | F_t] &= e^{\lambda(e^{i\theta}−1)t}.
\end{align*}
Strong Markov property

A Poisson process $N$ is compatible with a filtration $\{F_t\}$, if $N$ is $\{F_t\}$-adapted and $N(t + \cdot) - N(t)$ is independent of $F_t$ for every $t \geq 0$.

**Lemma 1.7** Let $N$ be a Poisson process with parameter $\lambda > 0$ that is compatible with $\{F_t\}$, and let $\tau$ be a $\{F_t\}$-stopping time such that $\tau < \infty$ a.s. Define $N_{\tau}(t) = N(\tau + t) - N(\tau)$. Then $N_{\tau}$ is a Poisson process that is independent of $F_{\tau}$ and compatible with $\{F_{\tau+t}\}$.

**Proof.** Let $M(t) = N(t) - \lambda t$. By the optional sampling theorem,

$$E[M((\tau + t + r) \wedge T)|F_{\tau+t}] = M((\tau + t) \wedge T),$$

so

$$E[N((\tau + t + r) \wedge T) - N((\tau + t) \wedge T)|F_{\tau+t}] = \lambda((\tau + t + r) \wedge T - (\tau + t) \wedge T).$$

By the monotone convergence theorem

$$E[N(\tau + t + r) - N(\tau + t)|F_{\tau+t}] = \lambda r$$

which gives the lemma. \qed
Intensity for a counting process

If $N$ is a Poisson process with parameter $\lambda$ and $N$ is compatible with $\{\mathcal{F}_t\}$, then

$$P\{N(t + \Delta t) > N(t) | \mathcal{F}_t\} = 1 - e^{-\lambda \Delta t} \approx \lambda \Delta t.$$ 

For a general counting process $N$, at least intuitively, a nonnegative, $\{\mathcal{F}_t\}$-adapted stochastic process $\lambda(\cdot)$ is an $\{\mathcal{F}_t\}$-intensity for $N$ if

$$P\{N(t + \Delta t) > N(t) | \mathcal{F}_t\} \approx E\left[ \int_t^{t+\Delta t} \lambda(s) ds | \mathcal{F}_t \right] \approx \lambda(t) \Delta t.$$ 

Let $S_n$ be the $n$th jump time of $N$.

**Definition 1.8** $\lambda$ is an $\{\mathcal{F}_t\}$-intensity for $N$ if and only if for each $n = 1, 2, \ldots$

$$N(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda(s) ds$$

is a $\{\mathcal{F}_t\}$-martingale.
Modeling with intensities

Let $Z$ be a stochastic process (cadlag, $E$-valued for simplicity) that models “external noise.” Let $D^e[0, \infty)$ denote the space of counting paths (zero at time zero and constant except for jumps of $+1$).

**Condition 1.9**

$$\lambda : [0, \infty) \times D_E[0, \infty) \times D^c[0, \infty) \to [0, \infty)$$

is measurable and satisfies $\lambda(t, z, v) = \lambda(t, z^t, v^t)$, where $z^t(s) = z(s \land t)$ ($\lambda$ is nonanticipating), and

$$\int_0^t \lambda(s, z, v) ds < \infty$$

for all $t \geq 0$, $z \in D_E[0, \infty)$ and $v \in D^c[0, \infty)$.

Let $Y$ be a unit Poisson process that is $\{G_u\}$-compatible and assume that $Z(s)$ is $G_0$-measurable for every $s \geq 0$. (In particular, $Z$ is independent of $Y$.) Consider

$$N(t) = Y(\int_0^t \lambda(s, Z, N) ds).$$

(1.1)
Solution of the stochastic equation

**Theorem 1.10** There exists a unique solution of (1.1) up to time $S_\infty \equiv \lim_{n \to \infty} S_n$,

$$\tau(t) = \int_0^t \lambda(s, Z, N) ds$$

is a $\{G_u\}$-stopping time, and for each $n = 1, 2, \ldots$,

$$N(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda(s, Z, N) ds$$

is a $\{G_{\tau(t)}\}$-martingale. (Let $F_t = G_{\tau(t)}$.)
Proof. Existence and uniqueness follows by solving from one jump to the next. Let
\( Y^r(u) = Y(r \land u) \) and let
\[
N^r(t) = Y^r(\int_0^t \lambda(s, Z, N^r) ds).
\]
Then \( N^r(t) = N(t) \), if \( \tau(t) = \int_0^t \lambda(s, Z, N) ds \leq r \). Consequently,
\[
\{ \tau(t) \leq r \} = \{ \int_0^t \lambda(s, Z, N^r) ds \leq r \} \in \mathcal{F}_r,
\]
as is \( \{ \tau(t \land S_n) \leq r \} \). Since \( M(u) = Y(u) - u \) is a martingale, by the optional sampling theorem
\[
E[M(\tau((t + v) \land S_n) \land T)|\mathcal{F}_\tau(t)] = M(\tau((t + v) \land S_n)) \land \tau(t) \land T) = M(\tau(t \land S_n) \land T).
\]
\( \Box \)
Martingale problems for counting processes

Definition 1.11 (Jacod (1974/75)) Let $Z$ be a cadlag, $E$-valued stochastic process, and let $\lambda$ satisfy Condition 1.9. A counting process $N$ is a solution of the martingale problem for $(\lambda, Z)$ if

$$N(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda(s, Z, N)ds$$

is a martingale with respect to the filtration

$$\mathcal{F}_t = \sigma(N(s), Z(r) : s \leq t, r \geq 0)$$

Theorem 1.12 If $N$ is a solution of the martingale problem for $(\lambda, Z)$, then $N$ has the same distribution as the solution of the stochastic equation (1.1).
\textbf{Proof.} Suppose $\lambda$ is an intensity for a counting process $N$ and $\int_0^\infty \lambda(s)ds = \infty$ a.s. Let $\gamma(u)$ satisfy

$$\gamma(u) = \inf\{t : \int_0^t \lambda(s)ds \geq u\}.$$  

Then, since $\gamma(u + v) \geq \gamma(u)$,

$$E[N(\gamma(u+v) \wedge S_n \wedge T) - \int_0^{\gamma(u+v) \wedge S_n \wedge T} \lambda(s)ds | \mathcal{F}_{\gamma(u)}] = N(\gamma(u) \wedge S_n \wedge T) - \int_0^{\gamma(u) \wedge S_n \wedge T} \lambda(s)ds.$$  

The monotone convergence argument lets us send $T$ and $n$ to infinity. We then have

$$E[N(\gamma(u + v)) - (u + v) | \mathcal{F}_{\gamma(u)}] = N(\gamma(u)) - u,$$

so $Y(u) = N(\gamma(u))$ is a Poisson process. But $\gamma(\tau(t)) = t$, so (1.1) is satisfied.

If $\int_0^\infty \lambda(s)ds < \infty$ with positive probability, then let $Y^*$ be a unit Poisson process that is independent of $\mathcal{F}_t$ for all $t \geq 0$ and consider $N^\epsilon(t) = N(t) + Y^*(\epsilon t)$. $N^\epsilon$ has intensity $\lambda(t) + \epsilon$, and $Y^\epsilon$, obtained as above, converges to

$$Y(u) = \begin{cases} 
N(\gamma(u)) & u < \tau(\infty) \\
N(\infty) + Y^*(u - \tau(\infty)) & u \geq \tau(\infty) 
\end{cases}$$

(except at points of discontinuity). \qed
Multivariate counting processes

$D_{d}^{c}[0, \infty)$: The collection of $d$-dimensional counting paths

**Condition 1.13** $\lambda_k : [0, \infty) \times D_{d}^{c}[0, \infty) \times D_{E}[0, \infty) \to [0, \infty)$, measurable and nonanticipating with

$$
\int_0^t \sum_k \lambda_k(s, z, v) ds < \infty, \quad v \in D_{d}^{c}[0, \infty), z \in D_{E}[0, \infty).
$$

Here, $Z$ is cadlag, $E$-valued and independent of independent Poisson processes $Y_1, \ldots, Y_d$.

$$N_k(t) = Y_k(\int_0^t \lambda_k(s, Z, N) ds), \quad (1.2)$$

where $N = (N_1, \ldots, N_d)$. Existence and uniqueness holds (including for $d = \infty$) and

$$N_k(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda_k(s, Z, N) ds$$

is a martingale for $S_n = \inf\{t: \sum_k N_k(t) \geq n\}$, but what is the correct filtration?
Multiparameter optional sampling theorem

\( \mathcal{I} \) is a directed set with partial ordering \( t \leq s \). If \( t_1, t_2 \in \mathcal{I} \), there exists \( t_3 \in \mathcal{I} \) such that \( t_1 \leq t_3 \) and \( t_2 \leq t_3 \).

\( \{ \mathcal{F}_t, t \in \mathcal{I} \}, s \leq t \) implies \( \mathcal{F}_s \subset \mathcal{F}_t \).

A stochastic process \( X(t) \) indexed by \( \mathcal{I} \) is a martingale if and only if for \( s \leq t \),

\[
E[X(t)|\mathcal{F}_s] = X(s).
\]

An \( \mathcal{I} \) valued random variable is a stopping time if and only if \( \{ \tau \leq t \} \in \mathcal{F}_t, t \in \mathcal{I} \).

\( \mathcal{F}_\tau = \{ A \in \mathcal{F} : A \cap \{ \tau \leq t \} \in \mathcal{F}_t, t \in \mathcal{I} \} \)

**Lemma 1.14** Let \( X \) be a martingale and let \( \tau_1 \) and \( \tau_2 \) be stopping times assuming countably many values and satisfying \( \tau_1 \leq \tau_2 \) a.s. If there exists a sequence \( \{ T_m \} \subset \mathcal{I} \) such that \( \lim_{m \to \infty} P\{ \tau_2 \leq T_m \} = 1 \), \( \lim_{m \to \infty} E[|X(T_m)|1_{\{\tau_2 \leq T_m\}}] = 0 \), and \( E[|X(\tau_2)|] < \infty \), then

\[
E[X(\tau_2)|\mathcal{F}_{\tau_1}] = X(\tau_1)
\]
Proof. (Kurtz (1980b)) Define

$$
\tau_i^m = \begin{cases} 
\tau_i & \text{on } \{\tau_i \leq T_m\} \\
T_m & \text{on } \{\tau_i \leq T_m\}^c 
\end{cases}
$$

Then $\tau_i^m$ is a stopping time, since

$$
\{\tau_i^m \leq t\} = (\{\tau_i^m \leq t\} \cap \{\tau_i \leq T_m\}) \cup (\{\tau_i^m \leq t\} \cap \{\tau_i \leq T_m\}^c) = (\bigcup_{s \in \Gamma, s \leq t, s \leq T_m} \{\tau_i = s\}) \cup (\{T_m \leq t\} \cap \{\tau_i \leq T_m\}^c)
$$

Let $\Gamma \subset \mathcal{I}$ be countable and satisfy $P\{\tau_i \in \Gamma\} = 1$ and $\{T_m\} \subset \Gamma$. For $A \in \mathcal{F}_{\tau_1}$,

$$
\int_{A \cap \{\tau_1^m = t\}} X(\tau_2^m) dP = \sum_{s \in \Gamma, s \leq T_m} \int_{A \cap \{\tau_1^m = t\} \cap \{\tau_2^m = s\}} X(s) dP = \sum_{s \in \Gamma, s \leq T_m} \int_{A \cap \{\tau_1^m = t\} \cap \{\tau_2^m = s\}} X(T_m) dP
$$

$$
= \int_{A \cap \{\tau_1^m = t\}} X(T_m) dP = \int_{A \cap \{\tau_1^m = t\}} X(t) dP = \int_{A \cap \{\tau_1^m = t\}} X(\tau_1^m) dP \quad \square
$$
Multiple time change

\(\mathcal{I} = [0, \infty)^d, u \in \mathcal{I}, \mathcal{F}_u = \sigma(Y_k(s_k) : s_k \leq u_k, k = 1, \ldots, d).\) Then

\[ M_k(u) \equiv Y_k(u_k) - u_k \]

is a \(\{\mathcal{F}_u\}\)-martingale. For

\[ N_k(t) = Y_k\left(\int_0^t \lambda_k(s, Z, N)ds\right), \]

define \(\tau_k(t) = \int_0^t \lambda_k(s, Z, N)ds\) and \(\tau(t) = (\tau_1(t), \ldots, \tau_d(t))\). Then \(\tau(t)\) is a \(\{\mathcal{F}_u\}\)-stopping time.

**Lemma 1.15** Let \(G_t = \mathcal{F}_{\tau(t)}\). If \(\sigma\) is a \(\{G_t\}\)-stopping time, then \(\tau(\sigma)\) is a \(\{\mathcal{F}_u\}\)-stopping time.
Approximation by discrete stopping times

Lemma 1.16 If $\tau$ is a $\{F_u\}$-stopping time, then $\tau^{(n)}$ defined by

$$\tau^{(n)}_k = \left\lceil \frac{\tau_k 2^n}{2^n} \right\rceil + 1$$

is a $\{F_u\}$-stopping time.

Proof.

$$\{\tau^{(n)} \leq u\} = \bigcap_k \{\tau^{(n)}_k \leq u_k\} = \bigcap_k \left\{\left\lceil \frac{\tau_k 2^n}{2^n} \right\rceil + 1 \leq \left\lceil \frac{u_k 2^n}{2^n} \right\rceil \right\} = \bigcap_k \{\tau_k < \frac{\left\lceil \frac{u_k 2^n}{2^n} \right\rceil}{2^n}\}$$

Note that $\tau^{(n)}_k$ decreases to $\tau_k$. 

\[ \square \]
Martingale problems for multivariate counting processes

Let \( S_n = \inf \{ t : \sum_k N_k(t) \geq n \} \).

**Theorem 1.17** Let Condition 1.13 hold. For \( n = 1, 2, \ldots \), there exists a unique solution of (1.2) up to \( S_n \), \( \tau_k(t) = \int_0^t \lambda_k(s, Z, N)ds \) defines a \( \{ F_u \} \)-stopping time, and

\[
N_k(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda_k(s, Z, N)ds
\]

is a \( \{ F_{\tau(t)} \} \)-martingale.

**Definition 1.18** Let \( Z \) be a cadlag, \( E \)-valued stochastic process, and let \( \lambda = (\lambda_1, \ldots, \lambda_d) \) satisfy Condition 1.13. A multivariate counting process \( N \) is a solution of the martingale problem for \( (\lambda, Z) \) if for each \( k \)

\[
N_k(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda_k(s, Z, N)ds
\]

is a martingale with respect to the filtration

\[
G_t = \sigma(N(s), Z(r) : s \leq t, r \geq 0)
\]
Existence and uniqueness for the martingale problem

**Theorem 1.19** Let $Z$ be a cadlag, $E$-valued stochastic process, and let $\lambda = (\lambda_1, \ldots, \lambda_d)$ satisfy Condition 1.13. Then there exists a unique solution of the martingale problem for $(\lambda, Z)$.

**Proof.** Existence follows from the time-change equation, and uniqueness follows by a result in Meyer (1971) which essentially shows that every solution of the martingale problem can be written as a solution of the time-change equation. (See also Kurtz (1980c).) Jacod (1974/75) gave the first characterization of a multivariate counting processes as a solution of a martingale problem. □
2. Markov processes and their generators

- Markov chains
- Diffusion processes
- Definition of a martingale problem
- Equivalent formulations
- Uniqueness of 1-dimensional distributions implies uniqueness of fdd
- Markov property
- Building Markov generators (sums of generators are generators—usually)
- Uniqueness under the Hille-Yosida conditions
- Quasi-left continuity

Dynkin (1965); Meyer (1967); Stroock and Varadhan (1979); Ethier and Kurtz (1986)
Markov chains

Consider \( X \in \mathbb{Z}^d \) satisfying

\[
X(t) = X(0) + \sum_{k=1}^{m} Y_k(\int_0^t \beta_k(s) ds)\zeta_k = X(0) + \sum_{k=1}^{m} R_k(t)\zeta_k
\]

where the \( Y_k \) are independent unit Poisson processes and the \( \zeta_k \) are distinct. Note that

\[
P\{X(t + \Delta t) = X(t) + \zeta_k | \mathcal{F}_t^X\} \approx P\{R_k(t + \Delta t) > R_k(t) | \mathcal{F}_t^X\} \approx \beta_k(X(t))\Delta t.
\]

Let \( \tilde{R}_k(t) = R_k(t) - \int_0^t \beta_k(X(s)) ds \) (which will typically be a martingale)

\[
f(X(t)) = f(X(0)) + \sum_{k=1}^{m} (f(X(s-) + \zeta_k) - f(X(s-))dR_k(s)
\]

\[
= f(X(0)) + \sum_{k=1}^{m} (f(X(s-) + \zeta_k) - f(X(s-))d\tilde{R}_k(s)
\]

\[
+ \int_0^t \sum_{k=1}^{m} \beta_k(X(s))(f(X(s) + \zeta_k) - f(X(s)))ds.
\]
Martingale properties of Markov chains

Define

\[ Af(x) = \sum_{k=1}^{m} \beta_k(x)(f(x + \zeta_k) - f(x)). \]

Then, at least for \( f \) with compact support,

\[ f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds \]

will be a \( \{\mathcal{F}_t^X\} \)-martingale.
Diffusion processes

Consider the Itô equation,

\[ X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds, \]

and for \( a(x) = \sigma(x)\sigma(x)^T \), define

\[ Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \partial_i \partial_j f(x) + \sum_i b_i(x) \partial_i f(x). \]

Then for \( C^2 \) functions \( f \), by Itô’s formula

\[ f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds = \int_0^t \nabla f(X(s))^T \sigma(X(s))dW(s) \]

which, at least for \( f \) with compact support, is a \( \{\mathcal{F}_t^{X,W}\} \)-martingale (and hence a \( \{\mathcal{F}_t^X\} \)-martingale).
Martingale problems: Definition

\( E \) state space (a complete, separable metric space)

\( A \) generator (usually a linear operator with domain and range in \( B(E) \))

\( \mu \in \mathcal{P}(E) \)

\( X \) is a solution of the martingale problem for \((A, \mu)\) if and only if \( \mu = PX(0)^{-1} \) and there exists a filtration \( \{\mathcal{F}_t\} \) such that

\[
M_f(t) = f(X(t)) - f(X(0)) - \int_0^t A f(X(s)) ds
\]

is an \( \{\mathcal{F}_t\} \)-martingale for each \( f \in \mathcal{D}(A) \).

In these lectures, we will assume that \( \mathcal{D}(A) \subset C_b(E) \) and that all processes are cadlag.
Examples of generators

Standard Brownian motion \((E = \mathbb{R}^d)\)

\[ Af = \frac{1}{2} \Delta f, \quad \mathcal{D}(A) = C_c^2(\mathbb{R}^d) \]

Poisson process \((E = \{0, 1, 2 \ldots\}, \mathcal{D}(A) = B(E))\)

\[ Af(k) = \lambda(f(k + 1) - f(k)) \]

Pure jump process \((E \text{ arbitrary})\)

\[ Af(x) = \lambda(x) \int_E (f(y) - f(x)) \mu(x, dy) \]

Diffusion \((E = \mathbb{R}^d, \mathcal{D}(A) = C_c^2(\mathbb{R}^d))\)

\[ Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} f(x) \quad (2.1) \]
Equivalent formulations

Suppose, without loss of generality, that $\mathcal{D}(A)$ is closed under addition of constants ($A1 = 0$). Then the following are equivalent:

a) $X$ is a solution of the martingale problems for $(A, \mu)$.

b) $PX(0)^{-1} = \mu$ and there exists a filtration $\{\mathcal{F}_t\}$ such that for each $\lambda > 0$ and each $f \in \mathcal{D}(A)$,

$$e^{-\lambda t}f(X(t)) - f(X(0)) + \int_0^t e^{-\lambda s}(\lambda f(X(s)) - Af(X(s)))ds$$

is a $\{\mathcal{F}_t\}$-martingale.

c) $PX(0)^{-1} = \mu$ and there exists a filtration $\{\mathcal{F}_t\}$ such that for each $f \in \mathcal{D}(A)$ with $\inf_{x \in \mathbb{E}} f(x) > 0$,

$$R_f(t) = \frac{f(X(t))}{f(X(0))} \exp\left\{-\int_0^t \frac{Af(X(s))}{f(X(s))}ds\right\}$$

is a $\{\mathcal{F}_t\}$-martingale.
Proof.

\[
f(X(t)) \exp\left\{ - \int_0^t \frac{Af(X(s))}{f(X(s))} ds \right\} \\
= f(X(0)) + \int_0^t \exp\left\{ - \int_0^r \frac{Af(X(s))}{f(X(s))} ds \right\} df(X(r)) \\
- \int_0^t f(X(r)) \frac{Af(X(r))}{f(X(r))} \exp\left\{ - \int_0^r \frac{Af(X(s))}{f(X(s))} ds \right\} dr \\
= f(X(0)) + \int_0^t \exp\left\{ - \int_0^r \frac{Af(X(s))}{f(X(s))} ds \right\} dM_f(r)
\]

so if \(M_f\) is a martingale, then \(R_f\) is a martingale.

Conversely, if \(R_f\) is a martingale, then

\[
f(X(0)) \int_0^t \exp\left\{ \int_0^r \frac{Af(X(s))}{f(X(s))} ds \right\} dR_f(r) = M_f(t)
\]

is a martingale.

Note that considering only \(f\) that are strictly positive is no restriction since we can always add a constant to \(f\). \(\square\)
Conditions for the martingale property

**Lemma 2.1** For \((f, g) \in A, h_1, \ldots, h_m \in \overline{C}(E),\) and \(t_1 \leq t_2 \leq \cdots \leq t_{m+1},\) let

\[
\eta(Y) \equiv \eta(Y, (f, g), \{h_i\}, \{t_i\})
\]

\[
= (f(Y(t_{m+1}) - f(Y(t_m)) - \int_{t_m}^{t_{m+1}} g(Y(s)ds) \prod_{i=1}^{m} h_i(Y(t_i)).
\]

Then \(Y\) is a solution of the martingale problem for \(A\) if and only if \(E[\eta(Y)] = 0\) for all such \(\eta.\)

The assertion that \(Y\) is a solution of the martingale problem for \(A\) is an assertion about the finite dimensional distributions of \(Y.\)
Uniqueness of 1-dimensional distributions implies uniqueness of fdd

**Theorem 2.2** If any two solutions of the martingale problem for $A$ satisfying $PX_1(0)^{-1} = PX_2(0)^{-1}$ also satisfy $PX_1(t)^{-1} = PX_2(t)^{-1}$ for all $t \geq 0$, then the f.d.d. of a solution $X$ are uniquely determined by $PX(0)^{-1}$

**Proof.** If $X$ is a solution of the MGP for $A$ and $X_a(t) = X(a + t)$, then $X_a$ is a solution of the MGP for $A$. Further more, for positive $f_i \in B(E)$ and $0 \leq t_1 < t_2 < \cdots < t_m = a$, define

$$Q(B) = \frac{E[1_B(X_a) \prod_{i=1}^m f_i(X(t_i))]}{E[\prod_{i=1}^m f_i(X(t_i))]}$$

defines a probability measure on $\mathcal{F} = \sigma(X_a(s), s \geq 0)$ and under $Q$, $X_a$ is a solution of the martingale problem for $A$ with initial distribution

$$\mu(\Gamma) = \frac{E[1_\Gamma(X(a)) \prod_{i=1}^m f_i(X(t_i))]}{E[\prod_{i=1}^m f_i(X(t_i))]}.$$
Proceeding by induction, fix $m$ and suppose $E[\prod_{i=1}^{m} f_i(X(t_i))]$ is uniquely determined for all $0 \leq t_1 < t_2 < \cdots < t_m$ and all $f_i$. Then $\mu$ is uniquely determined and the one dimensional distributions of $X_a$ under $Q$ are uniquely determined, that is

$$E[\frac{f_{m+1}(X(t_{m+1})) \prod_{i=1}^{m} f_i(X(t_i))}{E[\prod_{i=1}^{m} f_i(X(t_i))]}]$$

is uniquely determined for $t_{m+1} \geq a$. Since $a$ is arbitrary and the denominator is uniquely determined, the numerator is uniquely determined completing the induction step. $\square$
Time homogeneous Markov processes

A process $X$ is Markov with respect to a filtration $\{\mathcal{F}_t\}$ provided

$$E[f(X(t + r))|\mathcal{F}_t] = E[f(X(t + r))|X(t)]$$

for all $t, r \geq 0$ and all $f \in B(E)$.

The conditional expectation on the right can be written as $g_{f,t,r}(X(t))$ for a measurable function $g_{f,t,r}$ depending on $f$, $t$, and $r$.

If the function can be selected independently of $t$, that is

$$E[f(X(t + r))|X(t)] = g_{f,r}(X(t)),$$

then the Markov process is time homogeneous. A time inhomogeneous Markov process can be made time homogeneous by including time in the state. That is, set $Z(t) = (X(t), t)$.

Note that $g_{f,r}$ will be linear in $f$, so we can write $g_{f,r} = T(r)f$, where $T(r)$ is a linear operator on $B(E)$ (the bounded measurable functions on $E$). The Markov property then implies $T(r + s)f = T(r)T(s)f$. 
Adding a time component

Lemma 2.3 Suppose that \( g(t, x) \) has the property that \( g(t, \cdot) \in \mathcal{D}(A) \) for each \( t \) and that \( g, \partial_t g, \) and \( Ag \) are all bounded in \( t \) and \( x \) and are continuous functions of \( t \). If \( X \) is a solution of the martingale problem for \( A \), then

\[
g(t, X(t)) - \int_0^t (\partial_s g(x, X(s)) + Ag(s, X(s))) \, ds
\]

is a martingale.
Proof.

\[ E[g(t + r, X(t + r)) - g(t, X(t)) | \mathcal{F}_t] \]
\[ = \sum_k E[g(t + s_{k+1}, X(t + s_{k+1})) - g(t + s_k, X(t + s_k)) | \mathcal{F}_t] \]
\[ = \sum_k E[g(t + s_{k+1}, X(t + s_{k+1})) - g(t + s_{k+1}, X(t + s_k)) | \mathcal{F}_t] \]
\[ + \sum_k E[g(t + s_{k+1}, X(t + s_k)) - g(t + s_k, X(t + s_k)) | \mathcal{F}_t] \]
\[ = \sum_k E[\int_{t+s_k}^{t+s_{k+1}} A g(t + s_{k+1}, X(t + r)) dr | \mathcal{F}_t] \]
\[ + \sum_k E[\int_{t+s_k}^{t+s_{k+1}} \partial_r g(t + r, X(t + s_k)) dr | \mathcal{F}_t] \]

To complete the proof, see Exercise 14. □
Markov property

**Theorem 2.4** Suppose the conclusion of Theorem 2.2 holds. If $X$ is a solution of the martingale problem for $A$ with respect to a filtration $\{\mathcal{F}_t\}$, then $X$ is Markov with respect to $\{\mathcal{F}_t\}$.

**Proof.** Assuming that $P(F) > 0$, let $F \in \mathcal{F}_r$ and for $B \in \mathcal{F}$, define

$$P_1(B) = \frac{E[1_F E[1_B | \mathcal{F}_r]]}{P(F)}, \quad P_2(B) = \frac{E[1_F E[1_B | X(r)]]}{P(F)}.$$

Define $Y(t) = X(r + t)$. Then

$$P_1\{Y(0) \in \Gamma\} = \frac{E[1_F E[1_{\{Y(0) \in \Gamma\}} | \mathcal{F}_r]]}{P(F)} = \frac{E[1_F E[1_{\{X(r) \in \Gamma\}} | \mathcal{F}_r]]}{P(F)}$$

$$= \frac{E[1_F 1_{\{X(r) \in \Gamma\}}]}{P(F)} = \frac{E[1_F E[1_{\{X(r) \in \Gamma\}} | X(r)]]}{P(F)} = P_2\{Y(0) \in \Gamma\}$$
Check that $E^{P_1}[\eta(Y)] = E^{P_2}[\eta(Y)] = 0$ for all $\eta(Y)$ as in Lemma 2.1. Therefore

$$E[1_F E[f(X(r + t)) | \mathcal{F}_r]] = P(F)E^{P_1}[f(Y(t))]$$
$$= P(F)E^{P_2}[f(Y(t))]$$
$$= E[1_F E[f(X(r + t)) | X(r)]] .$$

Since $F \in \mathcal{F}_r$ is arbitrary, $E[f(X(r + t)) | \mathcal{F}_r] = E[f(X(r + t) | X(r)]$ and the Markov property follows. □
Building Markov generators
(sums of generators are generators–usually)

Suppose $A$ and $B$ are generators. Let $X_n$ be a stochastic process such that for $k = 0, 2, 4, \ldots$ $X_n$ evolves as if it has generator $A$ on the time interval $\left[\frac{k}{n}, \frac{k+1}{n}\right)$ and as if it has generator $B$ on the time interval $\left[\frac{k+1}{n}, \frac{k+2}{n}\right)$. Then for $f \in \hat{D} = \mathcal{D}(A) \cap \mathcal{D}(B)$,

$$f(X_n(t)) - f(X_n(0)) - \int_0^t \left( \frac{1 + (-1)^{[ns]}}{n} Af(X_n(s)) + \frac{1 - (-1)^{[ns]}}{n} Bf(X_n(s)) \right) ds$$

is a martingale. Letting $n \to \infty$, at least along a subsequence $X_n$ should converge to a process such that

$$f(X(t)) - f(X(0)) - \int_0^t \frac{1}{2} (A + B)f(X(s)) ds$$

is a martingale for $f \in \hat{D}$. (True, for example, if $E$ is compact, $A, B \subset C(E) \times C(E)$, and $\hat{D}$ is dense in $C(E)$.)
Markov processes and semigroups

\{T(t) : B(E) \to B(E), t \geq 0\} is an operator semigroup if \(T(t)T(s)f = T(t+s)f\)

\(X\) is a Markov process with operator semigroup \(\{T(t)\}\) if and only if

\[E[f(X(t+s))|\mathcal{F}^X_t] = T(s)f(X(t)), \quad t, s \geq 0, f \in B(E).\]

\[
T(s + r)f(X(t)) = E[f(X(t + s + r))|\mathcal{F}^X_t] \\
= E[E[f(X(t + s + r))|\mathcal{F}^X_{t+s}]|\mathcal{F}^X_t] \\
= E[T(r)f(X(t + s))|\mathcal{F}^X_t] \\
= T(s)T(r)f(X(t))
\]
Hille-Yosida theorem

**Theorem 2.5** The closure of $A$ is the generator of a strongly continuous contraction semigroup on $L_0$ if and only if

- $\mathcal{D}(A)$ is dense in $L_0$.
- $\|\lambda f - Af\| \geq \lambda \|f\|$, $f \in \mathcal{D}(A)$, $\lambda > 0$.
- $\mathcal{R}(\lambda - A)$ is dense in $L_0$.

**Proof.** (of sufficiency) Assuming $A$ is closed (otherwise, replace $A$ by its closure), the conditions imply $\mathcal{R}(\lambda - A) = L_0$ and the semigroup is obtained by

$$T(t)f = \lim_{n \to \infty} (I - \frac{1}{n}A)^{-[nt]}f.$$

(One must show that the right side is Cauchy.)
Uniqueness under the Hille-Yosida conditions

**Theorem 2.6** If $A$ satisfies the conditions of Theorem 2.5 and $\mathcal{D}(A)$ is separating, then there is at most one solution to the martingale problem.

**Proof.** If $X$ is a solution of the martingale problem for $A$, then by Lemma 2.3, for each $t > 0$ and each $f \in \mathcal{D}(A)$, $T(t - s)f(X(s))$ is a martingale. This martingale property extends to all $f$ in the closure of $\mathcal{D}(A)$. Consequently,

$$E[f(X(t))|\mathcal{F}_s] = T(t - s)f(X(s)),$$

and $E[f(X(t))] = E[T(t)f(X(0))]$ which determines the one dimensional distributions implying uniqueness. 

□
Cadlag versions

Lemma 2.7 Suppose $E$ is compact and $A \subset \overline{C}(E) \times B(E)$. If $\mathcal{D}(A)$ is separating, then any solution of the martingale problem for $A$ has a cadlag modification.

Proof. By the upcrossing inequality, there exists a set $\Omega_f \subset \Omega$ with $P(\Omega_f) = 1$ such that for $\omega \in \Omega_f$, $\lim_{s \to t^+, s \in \mathbb{Q}} f(X(s, \omega))$ exists for each $t \geq 0$ and $\lim_{s \to t^-, s \in \mathbb{Q}} f(X(s, \omega))$ exists for each $t > 0$. \qed
Quasi-left continuity

$X$ is quasi-left continuous if and only if for each sequence of stopping times $\tau_1 \leq \tau_2 \leq \cdots$ such that $\tau \equiv \lim_{n\to\infty} \tau_n < \infty$ a.s.,

$$\lim_{n\to\infty} X(\tau_n) = X(\tau) \ a.s.$$ 

Lemma 2.8 Let $A \subset \overline{C}(E) \times B(E)$, and suppose that $D(A)$ is separating. Let $X$ be a cadlag solution of the martingale problems for $A$. Then $X$ is quasi-left continuous.

Proof. For $(f, g) \in A$,

$$\lim_{n\to\infty} f(X(\tau_n \wedge t)) = \lim_{n\to\infty} E[f(X(\tau \wedge t)) - \int_{\tau_n \wedge t}^{\tau \wedge t} g(X(s))ds | F_{\tau_n}]$$

$$= E[f(X(\tau \wedge t)) | \bigvee_n F_{\tau_n}].$$

Since $X$ is cadlag,

$$\lim_{n\to\infty} X(\tau_n \wedge t) = \begin{cases} X(\tau \wedge t) & \text{if } \tau_n \wedge t = \tau \wedge t \text{ for } n \text{ sufficiently large} \\ X(\tau \wedge t-) & \text{if } \tau_n \wedge t < \tau \wedge t \text{ for all } n \end{cases}$$

To complete the proof, see Exercise 6. \qed
Continuity of diffusion process

Lemma 2.9 Suppose $E = \mathbb{R}^d$ and

$$Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} f(x), \quad \mathcal{D}(A) = C^2_c(\mathbb{R}^d).$$

If $X$ is a solution of the martingale problem for $A$, then $X$ has a modification that is cadlag in $\mathbb{R}^d \cup \{\infty\}$. If $X$ is cadlag, then $X$ is continuous.

**Proof.** The existence of a cadlag modification follows by Lemma 2.7. To show continuity, it is enough to show that for $f \in C^\infty_c(\mathbb{R}^d)$, $f \circ X$ is continuous. To show $f \circ X$ is continuous, it is enough to show

$$\lim_{\max |t_{i+1} - t_i| \to 0} \sum (f(X(t_{i+1} \wedge t)) - f(X(t_i \wedge t)))^4 = 0.$$
From the martingale properties,

\[
E[(f(X(t + h)) - f(X(t)))^4] = \int_t^{t+h} E\left[ A f^4(X(s)) - 4 f(X(t)) A f^3(X(s)) + 6 f^2(X(t)) A f^2(X(s)) - 4 f^3(X(t)) A f(X(s)) \right] ds
\]

Check that

\[
A f^4(x) - 4 f(x) A f^3(x) + 6 f^2(x) A f^2(x) - 4 f^3(x) A f(x) = 0
\]

(2.2)
3. Forward equations

- Conditions for relative compactness for cadlag processes
- Forward equations
- Uniqueness of the forward equation under a range condition
- Construction of a solution of a martingale problem
- Stationary distributions
- Echeverria’s theorem
- Equivalence of the forward equation and the MGP
Conditions for relative compactness

Let \((E, r)\) be complete, separable metric space, and define \(q(x, y) = 1 \wedge r(x, y)\) (so \(q\) is an equivalent metric under which \(E\) is complete). Let \(\{X_n\}\) be a sequence of cadlag processes with values in \(E\), \(X_n\) adapted to \(\{\mathcal{F}_t^n\}\).

**Theorem 3.1** Assume the following:

a) For \(t \in \mathcal{T}_0\), a dense subset of \([0, \infty)\), \(\{X_n(t)\}\) is relatively compact (in the sense of convergence in distribution).

b) For \(T > 0\), there exist \(\beta > 0\) and random variables \(\gamma_n(\delta, T)\) such that for \(0 \leq t \leq T, 0 \leq u \leq \delta\),

\[
E[q^\beta(X_n(t + u), X_n(t)) | \mathcal{F}_t^n] \leq E[\gamma_n(\delta, T) | \mathcal{F}_t^n]
\]

and

\[
\lim_{\delta \to 0} \limsup_{n \to \infty} E[\gamma_n(\delta, T)] = 0.
\]

Then \(\{X_n\}\) is relatively compact in \(D_E[0, \infty)\).
Relative compactness for martingale problems

**Theorem 3.2** Let $E$ be compact, and let $\{A_n\}$ be a sequence of generators. Suppose there exists a dense subset $\mathcal{D} \subset C(E)$ such that for each $f \in \mathcal{D}$ there exist $f_n \in \mathcal{D}(A_n)$ such that $\lim_{n \to \infty} \|f_n - f\| = 0$ and $C_f = \sup_n \|A_n f_n\| < \infty$. If $\{X_n\}$ is a sequence of cadlag processes in $E$ such that for each $n$, $X_n$ is a solution of the martingale problem for $A_n$, then $\{X_n\}$ is relatively compact in $D_E[0, \infty)$.

**Remark 3.3** The assumption of compactness is not as much of a restriction as it might appear. For processes in $\mathbb{R}^d$, typically $A \subset \hat{C}(\mathbb{R}^d) \times \hat{C}(\mathbb{R}^d)$. Replace $\mathbb{R}^d$ by the one point compactification $\mathbb{R}^d \cup \{\infty\}$ and add $(1, 0)$ to $A$. 
Proof. Let \( f \in D \) and for \( \delta > 0 \), let \( h^\delta \in D \) be such that there exist 
\[
\limsup_{\delta \to 0} \sqrt{\delta} C_{h^\delta} < \infty \quad \text{and} \quad \lim_{\delta \to 0} \| h^\delta - f^2 \| = 0.
\]
Then for \( 0 \leq u \leq \delta \),
\[
E[(f(X_n(t + u)) - f(X_n(t))^2 | \mathcal{F}_t^n]
\]
\[
= E[f^2(X_n(t + u)) - f^2(X_n(t)) | \mathcal{F}_t^n]
\]
\[
- 2f(X_n(t))E[f(X_n(t + u)) - f(X_n(t)) | \mathcal{F}_t^n]
\]
\[
\leq \limsup_{n \to \infty} (2\|f^2 - h^\delta_n\| + \| \int_t^{t+u} A_n h^\delta_n(X_n(s))ds \|)
\]
\[
+ 4\|f\|\|f - f_n\| + 2\|f\|\| \int_t^{t+u} A_n f_n(X_n(s))ds \|
\]
\[
\leq 2\|f^2 - h^\delta\| + C_{h^\delta} \delta + 2\|f\|C_f \delta \equiv \gamma_f(\delta).
\]
Since \( \lim_{\delta \to 0} \gamma_f(\delta) = 0 \), relative compactness for \( \{f(X_n)\} \) follows. Since the collection of such \( f \) is dense in \( C(E) \), relative compactness of \( \{X_n\} \) follows. \( \square \)
The forward equation for a general Markov process

Let $A \subset B(E) \times B(E)$. If $X$ is a solution of the martingale problem for $A$ and $\nu_t$ is the distribution of $X(t)$, then

$$0 = E[f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds] = \nu_t f - \nu_0 f - \int_0^t \nu_s Af ds$$

so

$$\nu_t f = \nu_0 f + \int_0^t \nu_s Af ds, \quad f \in \mathcal{D}(A).$$

(3.1)

(3.1) gives the weak form of the forward equation.

**Definition 3.4** A measurable mapping $t \in [0, \infty) \rightarrow \nu_t \in \mathcal{P}(E)$ is a solution of the forward equation for $A$ if (3.1) holds for all $f \in \mathcal{D}(A)$. 
Fokker-Planck equation

Let \( Af = \frac{1}{2}a(x)f''(x) + b(x)f'(x) \), \( f \in \mathcal{D}(A) = C^2_c(\mathbb{R}) \). If \( \nu_t \) has a \( C^2 \) density, then

\[
\nu_t Af = \int_{-\infty}^{\infty} Af(x) \nu(t, x) ds
= \int_{-\infty}^{\infty} f(x) \left( \frac{1}{2} \frac{\partial^2}{\partial x^2} (a(x) \nu(t, x)) - \frac{\partial}{\partial x} (b(x) \nu(t, x)) \right) dx
\]

and the forward equation is equivalent to

\[
\frac{\partial}{\partial t} \nu(t, x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} (a(x) \nu(t, x)) - \frac{\partial}{\partial x} (b(x) \nu(t, x)),
\]

known as the Fokker-Planck equation in the physics literature.
Forward/master equation for Markov chains

By the martingale property \( E[f(X(t))] = E[f(X(0))] + \int_0^t E[Af(X(s))] ds \)
and taking \( f(y) = 1_{\{x\}}(y) \), we have

\[
P\{X(t) = x\} = P\{X(0) = x\} + \int_0^t \left( \sum_l \beta_l(x - \zeta_l) P\{X(s) = x - \zeta_l\} \right.
\]

\[
- \left. \sum_l \beta_l(x) P\{X(s) = x\} \right) ds
\]

giving the Kolmogorov forward or master equation for the distribution of \( X \). In particular, defining \( p_x(t) = P\{X(t) = x\} \) and \( \nu_x = P\{X(0) = x\} \), \( \{p_x\} \) satisfies the system of differential equations

\[
\dot{p}_x(t) = \sum_l \lambda_l(x - \zeta_l)p_{x-\zeta_l}(t) - \left( \sum_l \lambda_l(x) \right)p_x(t), \quad (3.2)
\]

with initial condition \( p_x(0) = \nu_x \).
Uniqueness for the forward equation

Lemma 3.5 If \( \{\nu_t\} \) and \( \{\mu_t\} \) are solutions of the forward equation for \( A \) with \( \nu_0 = \mu_0 \) and \( \mathcal{R}(\lambda - A) \) is separating for each \( \lambda > 0 \), then \( \int_0^\infty e^{-\lambda t} \nu_t dt = \int_0^\infty e^{-\lambda t} \mu_t dt \) and \( \mu_t = \nu_t \) for almost every \( t \). Consequently, if \( \nu \) and \( \mu \) are weakly right continuous or if \( \mathcal{D}(A) \) is separating, \( \nu_t = \mu_t \) for all \( t \geq 0 \).

Proof.

\[
\lambda \int_0^\infty e^{-\lambda t} \nu_t f dt = \nu_0 f + \lambda \int_0^\infty e^{-\lambda t} \int_0^t \nu_s A f ds dt \\
= \nu_0 f + \lambda \int_0^\infty \int_s^\infty e^{-\lambda t} \nu_s A f dt ds \\
= \nu_0 f + \int_0^\infty e^{-\lambda s} \nu_s A f ds \\
\]

and hence

\[
\int_0^\infty e^{-\lambda t} \nu_t (\lambda f - Af) dt = \nu_0 f.
\]

Since \( \mathcal{R}(\lambda - A) \) is separating, the result holds. \( \Box \)
Uniqueness for the FEQ implies uniqueness for the MGP

Lemma 3.6 If uniqueness holds for the forward equation, then uniqueness holds for the martingale problem.

Proof. Uniqueness for the forward equation gives uniqueness for the one-dimensional distributions which by Theorem 2.2 gives uniqueness for the martingale problem. □
The semigroup and the forward equation

If $A$ generates a semigroup in the Hille-Yosida sense, then

$$T(t)f = f + \int_0^t T(s)Af\,ds = f + \int_0^t AT(s)f\,ds.$$  

If $\nu_0 \in \mathcal{P}(E)$, then

$$\nu_0 T(t)f = \nu_0 f + \int_0^t \nu_0 T(s)Af\,ds$$

and if $\{T(t)\}$ is given by a transition function, $\nu_t = \int_E P(t, x, \cdot)\nu_0(dx)$ satisfies

$$\nu_t f = \nu_0 f + \int_0^t \nu_s Af\,ds, \quad f \in \mathcal{D}(A).$$

If $A$ satisfies the conditions of the Hille-Yosida theorem and $\mathcal{D}(A)$ is separating, then uniqueness follows by Lemma 3.5.
Dissipativity and the positive maximum principle

For $\lambda > 0$,

\[
\|\lambda f - t^{-1}(T(t)f - f)\| \geq (\lambda + t^{-1})\|f\| - t^{-1}\|T(t)f\| \geq \lambda\|f\|
\]

so $A$ is dissipative

\[
\|\lambda f - Af\| \geq \lambda\|f\|, \quad \lambda > 0.
\]

**Definition 3.7** A satisfies the positive maximum principle if $f(x) = \|f\|$ implies $Af(x) \leq 0$. (For our operators, we can always add a constant to $f$ making it positive. Consequently $f(x) = \sup_{y \in E} f(y)$ implies $Af(x) \leq 0$.)

**Lemma 3.8** Let $E$ be compact and $\mathcal{D}(A) \subset C(E)$. If $A$ satisfies the positive maximum principle, then $A$ is dissipative.
Digression on the proof of the Hille-Yosida theorem

The conditions of the Hille-Yosida theorem imply \((I - n^{-1}A)^{-1}\) exists and

\[
\left\| (I - n^{-1}A)^{-1} f \right\| \leq \| f \|. 
\]

In addition

\[
\left\| (I - n^{-1}A)^{-1} f - f \right\| \leq \frac{1}{n} \| Af \|. 
\]

One proof of the Hille-Yosida theorem is to show that

\[
T_n(t)f = (I - n^{-1}A)^{-[nt]}f
\]

is a Cauchy sequence and to observe that

\[
T_n(t)f = f + \frac{1}{n} \sum_{k=1}^{[nt]} (I - n^{-1}A)^{-k}Af = f + \int_0^{[nt]} T_n(s + n^{-1})Af \, ds 
\]
Probabilistic interpretation

\[(n - A)^{-1} = \int_{0}^{\infty} e^{-nt}T(t)dt \quad \text{and} \quad (I - n^{-1}A)^{-1} = n \int_{0}^{\infty} e^{-nt}T(t)dt.\] If \(\{T(t)\}\) is given by a transition function, then

\[\eta_n(x, dy) = n \int_{0}^{\infty} e^{-nt}P(t, x, dy)dt\]

is a discrete time transition function. If \(\{Y^n_k\}\) is a Markov chain with transition function \(\eta_n\), then

\[E[f(Y^n_k)] = E[(I - n^{-1}A)^{-k}f(Y_0)] = E[f(X(\frac{\Delta_1 + \cdots + \Delta_k}{n}))],\]

where the \(\Delta_i\) are independent unit exponentials, and \(X_n(t) = Y^n_{[nt]}\) can be written as

\[X_n(t) = X(\frac{1}{n} \sum_{k=1}^{[nt]} \Delta_k) \rightarrow X(t)\]
Construction of a solution of a martingale problem

**Theorem 3.9** Assume that $E$ is compact, $A \subset C(E) \times C(E)$, $(1, 0) \in A$, $A$ is linear, and $\mathcal{D}(A)$ is dense in $C(E)$. Assume that $A$ satisfies the positive maximum principle (and is consequently dissipative). Then there exists a transition function $\eta_n$ such that

$$
\int_E f(y)\eta_n(x, dy) = (I - n^{-1}A)^{-1} f(x) \quad (3.3)
$$

for all $f \in \mathcal{R}(I - n^{-1}A)$. 
Proof. Note that $\mathcal{D}((I - n^{-1}A)^{-1}) = \mathcal{R}(I - n^{-1}A)$.

For each $x \in E$, 
$$\eta_x h = (I - n^{-1}A)^{-1}h(x)$$

is a linear functional on $\mathcal{R}(I - n^{-1}A)$. If 
$$h(x) = f(x) - \frac{1}{n}Af(x)$$

for some $f \in \mathcal{D}(A)$, then $\eta_x h = f(x)$. Since $A$ satisfies the positive maximum principle, $\eta_x$ is a positive linear functional on $\mathcal{R}(I - \frac{1}{n}A)$, $|\eta_x h| = |f(x)| \leq \|h\|$ and $\eta_x 1 = 1$. The Hahn-Banach theorem implies $\eta_x$ extends to a positive linear functional on $C(E)$ (hence a probability measure).

$$\Gamma_x = \{ \eta \in \mathcal{P}(E) : \eta f = (I - n^{-1}A)^{-1}f(x), f \in \mathcal{R}(I - n^{-1}A) \}$$

is closed and $\limsup_{y \to x} \Gamma_y \subset \Gamma_x$. The measurable selection theorem implies the existence of $\eta$ satisfying (3.3). □
Approximating Markov chain

For $\eta_n$ as in (3.3), define

$$A_n f = n \left( \int_{E} f(y) \eta_n(x, dy) - f(x) \right)$$

Then $A$ is the generator of a pure-jump Markov process of the form

$$X_n(t) = Y^n_{N_n(t)}$$

where $\{Y^n_k\}$ is a Markov chain with transition function $\eta_n$ and $N_n$ is a Poisson process with parameter $n$.

Then

$$f(X_n(t)) - f(X_n(0)) - \int_0^t A_n f(X_n(s)) ds$$

is a martingale, and in particular, if $f \in \mathcal{D}(A)$ and $f_n = f - n^{-1}Af$, then $A_n f_n = Af$ and

$$M^n_f(t) = f_n(X_n(t)) - f_n(X_n(0)) - \int_0^t Af(X_n(s)) ds$$

is a martingale.
Theorem 3.2 implies \( \{X_n\} \) is relatively compact, and (see Lemma 2.1), if \( X \) is a limit point of \( \{X_n\} \), for \( 0 \leq t_1 < \cdots < t_{m+1} \),

\[
0 = E[(f_n(X_n(t_{m+1})) - f_n(X_n(t_m)) - \int_{t_m}^{t_{m+1}} A_n f(X_n(s)) ds \prod_{i=1}^{m} h_i(X_n(t_i)))]
\]

\[
\rightarrow E[(f(X(t_{m+1}) - f(X(t_m)) - \int_{t_m}^{t_{m+1}} A f(X(s)) ds \prod_{i=1}^{m} h_i(X(t_i))],
\]

at least if the \( \{t_i\} \) are selected outside the at most countable set of times at which \( X \) has a fixed point of discontinuity. Since \( X \) is right continuous, the right side is in fact zero for all choices of \( \{t_i\} \), so \( X \) is a solution of the martingale problem for \( A \).
Stationary distributions

Definition 3.10 A stochastic process $X$ is stationary if the distribution of $X_t \equiv X(t + \cdot)$ does not depend on $t$.

Definition 3.11 $\mu$ is a stationary distribution for the martingale problem for $A$ if there exists a stationary solution of the martingale problem for $A$ with marginal distribution $\mu$.

Theorem 3.12 Suppose that $D(A)$ and $R(\lambda - A)$, $\lambda > 0$, are separating and that for each $\nu \in P(E)$, there exists a solution of the martingale problem for $(A, \nu)$. If $\mu \in P(E)$ satisfies

$$\int_{E} Af \, d\mu = 0, \quad f \in D(A),$$

then $\mu$ is a stationary distribution for $A$. (See Lemma 3.5.)
Echeverria’s theorem

Theorem 3.13

Let \( E \) be compact, and let \( A \subset C(E) \times C(E) \) be linear and satisfy the positive maximum principle. Suppose that \( D(A) \) is closed under multiplication and dense in \( C(E) \). If \( \mu \in \mathcal{P}(E) \) satisfies

\[
\int_E Af \, d\mu = 0, \quad f \in D(A),
\]

then \( \mu \) is a stationary distribution of \( A \).

Example 3.14

\( E = [0, 1] \), \( Af(x) = \frac{1}{2} f''(x) \)

\( D(A) = \{ f \in C^2[0, 1] : f'(0) = f'(1) = 0, f'(\frac{1}{3}) = f'(\frac{2}{3}) \} \)

Let \( \mu(dx) = 3I_{[\frac{1}{3}, \frac{2}{3}]}(x) \, dx \).
Outline of proof

In the proof of Theorem 3.9, we constructed $\eta_n$ so that

$$\int_E f_n(y)\eta_n(x, dy) = \int_E (f(y) - \frac{1}{n}Af(y))\eta_n(x, dy) = f(x)$$

Consequently,

$$\int_E \int_E f_n(y)\eta_n(x, dy)\mu(dx) = \int_E f(x)\mu(dx)$$

$$= \int_E (f(x) - \frac{1}{n}Af(x))\mu(dx) = \int_E f_n(x)\mu(dx),$$

and $\mu$ should be a stationary distribution for $\eta_n$. 
For \( F(x, y) = \sum_{i=1}^{m} h_i(x) (f_i(y) - \frac{1}{n} Af_i(y)) + h_0(y), \quad f_i \in \mathcal{D}(A), \) define

\[
\Lambda_n F = \int \left[ \sum_{i=1}^{m} h_i(x) f_i(x) + h_0(x) \right] \mu(dx).
\]

We want to apply the Hahn Banach theorem to \( \Lambda_n \) considered as a linear functional on \( C(E \times E) \).

If \( \Lambda_n \) is given by a measure \( \nu_n \), then both marginals are \( \mu \), and letting \( \eta_n \) satisfy \( \nu_n(dx, dy) = \eta_n(x, dy)\mu(dx) \), for \( f \in \mathcal{D}(A) \),

\[
\int (f(y) - \frac{1}{n} Af(y))\eta_n(x, dy) = f(x), \quad \mu - a.s.
\]

The work is to show that \( \Lambda_n \) is a positive linear functional.
Extensions

**Theorem 3.15** Let $E$ be locally compact (e.g., $E = \mathbb{R}^d$), and let $A \subset \hat{C}(E) \times \hat{C}(E)$ satisfy the positive maximum principle. Suppose that $\mathcal{D}(A)$ is an algebra and dense in $\hat{C}(E)$. If $\mu \in \mathcal{P}(E)$ satisfies

$$\int_E Af d\mu = 0, \quad f \in \mathcal{D}(A),$$

then $\mu$ is a stationary distribution of $A$.

**Proof.** Let $\hat{E} = E \cup \{\infty\}$ and extend $A$ to include $(1, 0)$. There exists an $\hat{E}$-valued stationary solution $X$ of the martingale problem for the extended $A$, but $P\{X(t) \in E\} = \mu(E)$.

Let $E = \mathbb{R}$ and $Af(x) = a(x)f''(x)$, $f \in C_c^2$, where $a(x) > 0$ and $\int_{-\infty}^{\infty} \frac{1}{a(x)} dx = 1$. Then $\mu(dx) = \frac{1}{a(x)} dx$ is a stationary measure for $A$. 
Complete, separable $E$  

Bhatt and Karandikar (1993)

$E$ complete, separable. $A \subset C_b(E) \times C_b(E)$.

Assume that $\{g_k\}$ is closed under multiplication. Let $\mathcal{I}$ be the collection of finite subsets of positive integers, and for $I \in \mathcal{I}$, let $k(I)$ satisfy $g_{k(I)} = \prod_{i \in I} g_i$. For each $k$, there exists $a_k \geq |g_k|$. Let

$$\hat{E} = \{z \in \prod_{i=1}^{\infty} [-a_i, a_i] : z_{k(I)} = \prod_{i \in I} z_i, I \in \mathcal{I}\}.$$

Note that $\hat{E}$ is compact, and define $G : E \to \hat{E}$ by $G(x) = (g_1(x), g_2(x), \ldots)$. Then $G$ has a measurable inverse defined on the (measurable) set $G(E)$.

**Lemma 3.16** Let $\mu \in \mathcal{P}(E)$. Then there exists a unique measure $\nu \in \mathcal{P}(\hat{E})$ satisfying $\int_E g_k d\mu = \int_{\hat{E}} z_k \nu(dz)$. In particular, if $Z$ has distribution $\nu$, then $G^{-1}(Z)$ has distribution $\mu$. 

Equivalence of the forward equation and the MGP

Suppose

\[ \nu_t f = \nu_0 f + \int_0^t \nu_s A f \, ds \]

Define

\[ B_\lambda f(x, \theta) = A f(x, \theta) + \lambda \left( \int_E f(y, -\theta) \nu_0(dy) - f(x, \theta) \right) \]

and

\[ \mu_\lambda = \lambda \int_0^\infty e^{-\lambda t} \nu_t dt \times \left( \frac{1}{2} \delta_1 + \frac{1}{2} \delta_{-1} \right). \]

Then

\[ \int_E B_\lambda f \, d\mu_\lambda = 0, \quad f(x, \theta) = f_1(x) f_2(\theta), \quad f_1 \in D(A). \]

Let \( \tau_1 = \inf\{t > 0 : \Theta(t) \neq \Theta(0)\} \), \( \tau_{k+1} = \inf\{t > \tau_k : \Theta(t) \neq \Theta(\tau_k)\} \).
Theorem 3.17 Let \((Y, \Theta)\) be a stationary solution of the martingale problem for \(B_\lambda\) with marginal distribution \(\mu_\lambda\). Let \(\tau_1 = \inf\{t > 0 : \Theta(t) \neq \Theta(0)\}\), \(\tau_{k+1} = \inf\{t > \tau_k : \Theta(t) \neq \Theta(\tau_k)\}\). Define \(X(t) = Y(\tau_1 + t)\). Then conditioned on \(\tau_2 - \tau_1 > t_0\), \(X\) is a solution of the martingale problem for \(A\) and the distribution of \(X(t)\) is \(\nu_t\) for \(0 \leq t \leq t_0\).

Extension of Echeverria’s result to the forward equation is given in Theorem 4.9.19 of Ethier and Kurtz (1986) for locally compact spaces and in Theorem 3.1 of Bhatt and Karandikar (1993) for general complete separable metric spaces.
4. Other types of martingale problems

- Stopped martingale problems
- Local martingale problems
- Notions of uniqueness
- Controlled martingale problems
- Constrained martingale problems

Kurtz and Nappo (2011), Kurtz and Stockbridge (2001)
Stopped martingale problems

We weaken the assumption that $A \subset C_b(E) \times B(E)$ and consider $A \subset C_b(E) \times M(E)$, but assume the existence of a $\psi \geq 1$ such that for each $f \in \mathcal{D}(A)$, there exists $a_f > 0$ such that

$$|Af(x)| \leq a_f \psi(x).$$

**Definition 4.1** An $E$-valued process $X$ and a nonnegative random variable $\tau$ are a solution of the stopped martingale problem for $A$, if there exists a filtration $\{\mathcal{F}_t\}$ such that $X$ is $\{\mathcal{F}_t\}$-adapted, $\tau$ is a $\{\mathcal{F}_t\}$-stopping time,

$$E[\int_0^{t \wedge \tau} \psi(X(s)) ds] < \infty, \quad t \geq 0,$$

(4.1)

and for each $f \in \mathcal{D}(A)$,

$$f(X(t \wedge \tau)) - f(X(0)) - \int_0^{t \wedge \tau} Af(X(s)) ds$$

(4.2)

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

Definition 4.2 An $E$-valued process $X$ is a solution of the local-martingale problem for $A$, if there exists a filtration $\{\mathcal{F}_t\}$ such that $X$ is $\{\mathcal{F}_t\}$-adapted and a sequence $\{\tau_n\}$ of $\{\mathcal{F}_t\}$-stopping times such that $\tau_n \to \infty$ a.s. and for each $n$, $(X, \tau_n)$ is a solution of the stopped martingale problem for $A$ using the filtration $\{\mathcal{F}_t\}$.

Remark 4.3 If $X$ is a solution of the local martingale problem for $A$, then the localizing sequence $\{\tau_n\}$ can be taken to be predictable. In particular, we can take

$$\tau_n = \inf\{t : \int_0^t \psi(X(s)) \, ds \geq n\}.$$
Notions of uniqueness

**Definition 4.4** Uniqueness holds for the (local) martingale problem for \((A, \nu_0)\) if and only if all solutions have the same finite-dimensional distributions. Stopped uniqueness holds if for any two solutions, \((X_1, \tau_1), (X_2, \tau_2)\), of the stopped martingale problem for \((A, \nu_0)\), there exists a stochastic process \(\tilde{X}\) and nonnegative random variables \(\tilde{\tau}_1, \tilde{\tau}_2\) such that \((\tilde{X}, \tilde{\tau}_1 \lor \tilde{\tau}_2)\) is a solution of the stopped martingale problem for \((A, \nu_0)\), \((\tilde{X} (\cdot \land \tilde{\tau}_1), \tilde{\tau}_1)\) has the same distribution as \((X_1 (\cdot \land \tau_1), \tau_1)\), and \((\tilde{X} (\cdot \land \tilde{\tau}_2), \tilde{\tau}_2)\) has the same distribution as \((X_2 (\cdot \land \tau_2), \tau_2)\).

**Remark 4.5** Note that stopped uniqueness implies uniqueness. Stopped uniqueness holds if uniqueness holds and every solution of the stopped martingale problem can be extended (beyond the stopping time) to a solution of the (local) martingale problem. (See Lemma 4.5.16 of Ethier and Kurtz (1986) for conditions under which this extension can be done.)
Local forward equation

**Definition 4.6** A pair of measure-valued functions \( \{(\nu^0_t, \nu^1_t), t \geq 0\} \) is a solution of the stopped forward equation for \( A \) if for each \( t \geq 0 \), \( \nu_t \equiv \nu^0_t + \nu^1_t \in \mathcal{P}(E) \) and \( \int_0^t \nu^1_s \psi ds < \infty \), \( t \rightarrow \nu^0_t(C) \) is nondecreasing for all \( C \in \mathcal{B}(E) \), and for each \( f \in D(A) \),

\[
\nu_t f = \nu_0 f + \int_0^t \nu^1_s A f ds. \tag{4.3}
\]

A \( \mathcal{P}(E) \)-valued function \( \{\nu_t, t \geq 0\} \) is a solution of the local forward equation for \( A \) if there exists a sequence \( \{(\nu^{0,n}_t, \nu^{1,n}_t)\} \) of solutions of the stopped forward equation for \( A \) such that for each \( C \in \mathcal{B}(E) \) and \( t \geq 0 \), \( \{\nu^{1,n}_t(C)\} \) is nondecreasing and \( \lim_{n \to \infty} \nu^{1,n}_t(C) = \nu_t(C) \).

Any solution of the stopped martingale problem for \( A \) gives a solution of the stopped forward equation for \( A \), that is,

\[
\nu^0_t f = E[1_{[\tau, \infty)}(t)f(X(\tau))] \quad \text{and} \quad \nu^1_t f = E[1_{[0, \tau)}(t)f(X(t))].
\]
Technical conditions

Condition 4.7  1. $A : \mathcal{D}(A) \subset C_b(E) \times C(E)$ with $1 \in \mathcal{D}(A)$ and $A1 = 0$.

2. $\mathcal{D}(A)$ is closed under multiplication and separates points.

3. There exist $\psi \in C(E)$, $\psi \geq 1$, and constants $a_f$ such that $f \in \mathcal{D}(A)$ implies $|Af(x)| \leq a_f\psi(x)$.

4. Defining $A_0 = \{(f, \psi^{-1}Af) : f \in \mathcal{D}(A)\}$, $A_0$ is separable in the sense that there exists a countable collection $\{g_k\} \subset \mathcal{D}(A)$ such that every solution of the martingale problem for $A_0^r = \{(g_k, A_0g_k) : k = 1, 2, \ldots\}$ is a solution for $A_0$.

5. $A_0$ is a pre-generator, that is, $A_0$ is dissipative and there are sequences of functions $\mu_n : E \to \mathcal{P}(E)$ and $\lambda_n : E \to [0, \infty)$ such that for each $(f, g) \in A$ for each $x \in E$

\[
g(x) = \lim_{n \to \infty} \lambda_n(x) \int_E (f(y) - f(x))\mu_n(x, dy). \tag{4.4}
\]
Equivalence of forward equations and martingale problems

The primary consequence of Condition 4.7 is an extension of Theorem 3.13

**Theorem 4.8** If $A$ satisfies Condition 4.7 and $\{(\nu^0_t, \nu^1_t), t \geq 0\}$ is a solution of the stopped forward equation for $A$, then there exists a solution $(X, \tau)$ of the stopped martingale problem for $A$ such that $\nu^0_t f = E[1_{[\tau, \infty)}(t)f(X(\tau))]$ and $\nu^1_t f = E[1_{[0, \tau)}(t)f(X(t))]$.

If $A$ satisfies Condition 4.7 and $\{\nu_t, t \geq 0\}$ is a solution of the local forward equation for $A$, then there exists a solution $X$ of the local martingale problem for $A$ satisfying $\nu_t f = E[f(X(t))]$. 
Proof. First enlarge the state space $\tilde{E} = E \times \{0, 1\}$ and define $\tilde{\nu}_t h = \nu^0_t h(\cdot, 0) + \nu^1_t h(\cdot, 1)$. Setting $D(\tilde{A}) = \{f(x)g(y) : f \in D(A), g \in B(\{0, 1\})\}$, for $h = fg \in D(\tilde{A})$, define $\tilde{A}h(x, y) = yAh(x, y) = yg(y)Af(x)$ and $Bh(x, y) = y(h(x, 0) - h(x, y))$. Then

$$
0 = \nu^0_t h(\cdot, 1) + \nu^1_t h(\cdot, 1) - \nu^0_0 h(\cdot, 1) - \nu^1_0 h(\cdot, 1) - \int_0^t \tilde{\nu}_s \tilde{A}hds
$$

$$
= \tilde{\nu}_t h - \tilde{\nu}_0 h - \int_0^t \tilde{\nu}_s \tilde{A}hds + \nu^0_t h(\cdot, 1) - \nu^0_0 h(\cdot, 0) + \nu^0_0 h(\cdot, 1) - \nu^0_0 h(\cdot, 0)
$$

$$
= \tilde{\nu}_t h - \tilde{\nu}_0 h - \int_0^t \tilde{\nu}_s \tilde{A}hds - \int_{E \times \{0, 1\} \times [0, t]} Bh(x, y) \mu(dx \times dy \times ds),
$$

where, noting that $\nu^0_t (C)$ is an increasing function of $t$, $\mu$ is the measure determined by

$$
\mu(C \times \{1\} \times [0, t_2]) = \nu^0_{t_2} (C) - \nu^0_0 (C), \quad \mu(C \times \{0\} \times [t_1, t_2]) = 0.
$$
Controlled martingale problems

\[ X(t) = X(0) + \int_0^t \sigma(X(s), U(s))dW(s) + \int_0^t b(X(s), U(s))ds \]

\[ Z(t) = 1 - Y(\int_0^t Z(s)V(s)ds)ds \]

Define \( a(x, u) = \sigma(x, u)^2 \),

\[ Af(x, u) = \frac{1}{2} a(x, u) f''(x) + b(x, u) f'(x), \quad x \in \mathbb{R}, u \in \mathbb{U} \]

\[ vBg(z) = vz((g(z - 1) - g(z)), \quad z \in \{0, 1\}, v \in [0, \infty) \]

Then the following is a martingale:

\[ f(X(t))g(Z(t)) - f(X(0))g(Z(0)) - \int_0^t g(Z(s)) Af(X(s), U(s))ds \]

\[ -\int_0^t f(X(s))V(s)Bg(Z(s))ds \]
Optimal stopping

If $\tau = \inf\{t : Z(t) = 0\}$ and only $Z$ is controlled, that is, we have an optimal stopping problem,

$$f(X(t \wedge \tau))g(Z(t)) - f(X(0))g(Z(0)) - \int_0^t g(Z(s))Z(s)Af(X(s \wedge \tau))ds$$

$$- \int_0^t f(X(s \wedge \tau))V(s)Bg(Z(s))ds$$

is a martingale.

If $\tilde{\nu}_t h = E[h(X(t \wedge \tau), Z(t))]$ and $\mu_t h = E[\int_0^t V(s)h(X(s \wedge \tau), Z(s))ds]$, then for $h(x, z) = f(x)g(z)$,

$$\tilde{\nu}_t h = \tilde{\nu}_0 h + \int_0^t \tilde{\nu}_s \tilde{A}hds + \int_{E \times \{0,1\} \times [0,t]} Bh(x, y)\mu(dx \times dz \times ds)$$
Constrained martingale problems

Consider the Skorohod equation for reflecting Brownian motion

\[ X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds + \int_0^t \eta(X(s))d\lambda(s), \]

where \( X \in D \) and \( \lambda \) increases only when \( X \in \partial D \). Then defining

\[ Bf = \eta \cdot \nabla f, \]

\[ f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds - \int_0^t Bf(X(s))d\lambda(s) \]

\[ = \int_0^t \nabla f(X(s))^T \sigma(X(s))dW(s) \]

is a martingale.
Constrained forward equations

The forward equation becomes

\[ \nu_t f = \nu_0 f + \int_0^t \nu_s Af \, ds + \int_{\partial D \times [0,t]} Bf \mu(dx \times ds) \]

and a stationary distribution must satisfy

\[ \pi Af + \mu_{\pi} Bf = 0, \]

which in the queueing literature is known as the basic adjoint relationship (see Dai and Harrison (1991, 1992)).
General controlled martingale problem

\[ A \subset C_b(E) \times C(E \times U), \ B \subset C_b(E) \times C(E \times U) \]

Want

\[ f(X(t)) - f(X(0)) - \int_0^t \int_U A f(X(s), u) \Lambda_s(du) ds \]
\[ - \int_{E \times U \times [0,t]} B f(x, u) \Gamma(dx \times du \times ds) \]

to be martingales with respect to some \( \{ \mathcal{F}_t \} \). The corresponding “forward equation” becomes

\[ \nu_t^E f = \nu_0^E + \int_0^t \nu_s A f ds + \int_{E \times U \times [0,t]} B f(x, u) \mu(dx \times du \times ds), \]

where \( \nu_t \in \mathcal{P}(E \times U) \) and \( \nu_t^E \) is the \( E \) marginal.
5. Change of measure

- Absolute continuity and the Radon Nikodym theorem
- Bayes formula
- Local absolute continuity
- Martingales under a change of measure
- Change of measure for Brownian motion
- Change of measure for Poisson processes
- Applications of absolute continuity
Absolute continuity and the Radon-Nikodym theorem

**Definition 5.1** Let $P$ and $Q$ be probability measures on $(\Omega, \mathcal{F})$. Then $P$ is absolutely continuous with respect to $Q$ ($P \ll Q$) if and only if $Q(A) = 0$ implies $P(A) = 0$.

**Theorem 5.2** If $P \ll Q$, then there exists a random variable $L \geq 0$ such that

$$P(A) = E^Q[1_A L] = \int_A LdQ, \quad A \in \mathcal{F}.$$

Consequently, $Z$ is $P$-integrable if and only if $ZL$ is $Q$-integrable, and

$$E^P[Z] = E^Q[ZL].$$

Standard notation: $\frac{dP}{dQ} = L$. 
Bayes Formula

Recall the definition of conditional expectation: $Y = E[Z|\mathcal{D}]$ if $Y$ is $\mathcal{D}$-measurable and for each $D \in \mathcal{D}$, $\int_D YdP = \int_D ZdP$.

Lemma 5.3 (Bayes Formula) If $dP = LdQ$, then

$$E^P[Z|\mathcal{D}] = \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}]}.$$ \hspace{1cm} (5.1)

Proof. Clearly the right side of (5.1) is $\mathcal{D}$-measurable. Let $D \in \mathcal{D}$. Then

$$\int_D \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}]} dP = \int_D \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}]} LdQ$$

$$= \int_D \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}]} E^Q[L|\mathcal{D}]dQ$$

$$= \int_D E^Q[ZL|\mathcal{D}]dQ$$

$$= \int_D ZLdQ = \int_D ZdP$$

which verifies the identity. \qed
Examples

For real-valued random variables with a joint density \( X, Y \sim f_{XY}(x, y) \), conditional expectations can be computed by

\[
E[g(X)|Y = y] = \frac{\int_{-\infty}^{\infty} g(z)f_{XY}(z, y)dz}{f_Y(y)}
\]

that is, setting \( h(y) \) equal to the right side,

\[
E[g(X)|Y] = h(Y).
\]

For general random variables, suppose \( X \) and \( Y \) are independent on \((\Omega, \mathcal{F}, Q)\). Let \( L = H(X, Y) \geq 0 \), and \( E[H(X, Y)] = 1 \). Define

\[
\nu_X(\Gamma) = Q\{X \in \Gamma\}
\]

\[
dP = H(X, Y)dQ.
\]

Bayes formula becomes

\[
E^P[g(X)|Y] = \frac{E^Q[g(X)H(X, Y)|Y]}{E^Q[H(X, Y)|Y]} = \frac{\int g(x)H(x, Y)\nu_X(dx)}{\int H(x, Y)\nu_X(dx)}
\]
Local absolute continuity

**Theorem 5.4** Let \((\Omega, \mathcal{F})\) be a measurable space, and let \(P\) and \(Q\) be probability measures on \(\mathcal{F}\). Let \(\{\mathcal{F}_t\}\) be a filtration, and suppose \(P|_{\mathcal{F}_t} \ll Q|_{\mathcal{F}_t}, \ t \geq 0\). Define \(L(t) = dP/dQ|_{\mathcal{F}_t}\). Then \(\{L(t)\}\) is a nonnegative \(\mathcal{F}_t\)-martingale on \((\Omega, \mathcal{F}, Q)\) and \(L_\infty = \lim_{n \to \infty} L(t)\) satisfies \(E^Q[L_\infty] \leq 1\). If \(E^Q[L_\infty] = 1\), then \(P \ll Q\) on \(\mathcal{F}_\infty = \bigvee_t \mathcal{F}_t\).

**Proof.** If \(t < s\), and \(D \in \mathcal{F}_t \subset \mathcal{F}_s\), then \(P(D) = E^Q[L(t)1_D] = E^Q[L(s)1_D]\) which implies \(E[L(s)|\mathcal{F}_t] = L(t)\). If \(E[L_\infty] = 1\), then \(L(t) \to L_\infty\) in \(L^1\), so
\[
P(D) = E^Q[L_\infty 1_D], \quad D \in \bigcup_t \mathcal{F}_t,
\]
hence for all \(D \in \bigvee_t \mathcal{F}_t\). \(\square\)

**Proposition 5.5** \(P \ll Q\) on \(\mathcal{F}_\infty\) if and only if \(P\{\lim_{t \to \infty} L(t) < \infty\} = 1\).

**Proof.** The dominated convergence theorem implies
\[
P\{\sup_t L(t) \leq K\} = \lim_{r \to \infty} E^Q[1_{\{\sup_t L(t) \leq K\}} L(r)] = E^Q[1_{\{\sup_t L(t) \leq K\}} L_\infty].
\]
Letting \(K \to \infty\), we see that \(E^Q[L_\infty] = 1\). \(\square\)
Martingales and change of measure

(See Protter (2004), Section III.6.)

Let \( \{ \mathcal{F}_t \} \) be a filtration and assume that \( P|_{\mathcal{F}_t} \ll Q|_{\mathcal{F}_t} \) and that \( L(t) \) is the corresponding Radon-Nikodym derivative. Then as before, \( L \) is an \( \{ \mathcal{F}_t \} \)-martingale on \((\Omega, \mathcal{F}, Q)\).

**Lemma 5.6** \( Z \) is a \( P \)-local martingale if and only if \( LZ \) is a \( Q \)-local martingale.

**Proof.** For a bounded stopping time \( \tau \), \( Z(\tau) \) is \( P \)-integrable if and only if \( L(\tau)Z(\tau) \) is \( Q \)-integrable. Furthermore, if \( L(\tau \wedge t)Z(\tau \wedge t) \) is \( Q \)-integrable, then \( L(t)Z(\tau \wedge t) \) is \( Q \)-integrable and \( E^Q[L(\tau \wedge t)Z(\tau \wedge t)] = E^Q[L(t)Z(\tau \wedge t)] \).

By Bayes formula, \( E^P[Z(t+h) - Z(t)|\mathcal{F}_t] = 0 \) if and only if

\[
E^Q[L(t+h)(Z(t+h) - Z(t))|\mathcal{F}_t] = 0
\]

which is equivalent to

\[
E^Q[L(t+h)Z(t+h)|\mathcal{F}_t] = E^Q[L(t+h)Z(t)|\mathcal{F}_t] = L(t)Z(t),
\]

so \( Z \) is a martingale under \( P \) if and only if \( LZ \) is a martingale under \( Q \). \( \square \)
Semimartingale decompositions under a change of measure

**Theorem 5.7** If $M$ is a $Q$-local martingale, then

\[ Z(t) = M(t) - \int_0^t \frac{1}{L(s)} d[L, M]_s \]  

(5.2)

is a $P$-local martingale, where $[L, M]_t$ is the covariation of $L$ and $M$. (Note that the integrand is $\frac{1}{L(s)}$, not $\frac{1}{L(s-)}$.)

**Proof.** Note that $LM - [L, M]$ is a $Q$-local martingale. We need to show that $LZ$ is a $Q$-local martingale. But letting $V$ denote the second term on the right of (5.2), we have $L(t)V(t) = \int_0^t V(s-)dL(s) + \int_0^t L(s)dV(s)$ and hence

\[ L(t)Z(t) = L(t)M(t) - [L, M]_t - \int_0^t V(s-)dL(s). \]

Both terms on the right are $Q$-local martingales. \qed
Theorem 5.8 Let $N$ be a unit Poisson process on $(\Omega, \mathcal{F}, Q)$ that is compatible with $\{\mathcal{F}_t\}$. If $\lambda$ is nonnegative, cadlag, $\{\mathcal{F}_t\}$-adapted, and satisfies
\[ \int_0^t \lambda(s)ds < \infty \quad a.s., \quad t \geq 0, \]
then
\[ L(t) = \exp \left\{ \int_0^t \ln \lambda(s-)dN(s) - \int_0^t (\lambda(s) - 1)ds \right\} \]
satisfies
\[ L(t) = 1 + \int_0^t (\lambda(s-) - 1)L(s-)d(N(s) - s) \] (5.3)
and is a $Q$-local martingale. If $E[L(T)] = 1$ and we define $dP = L(T)dQ$ on $\mathcal{F}_T$, then $N(t) - \int_0^t \lambda(s)ds$ is a $P$-local martingale.
Proof. Note that at the jump times of $N$, $L(s) = \lambda(s-)L(s-)$. By Theorem 5.7,

$$Z(t) = N(t) - t - \int_0^t \frac{1}{L(s)}(\lambda(s-) - 1)L(s-)dN(s)$$

$$= \int_0^t \frac{1}{\lambda(s-)}dN(s) - t$$

is a local martingale under $P$. Consequently,

$$\int_0^t \lambda(s-)dZ(s) = N(t) - \int_0^t \lambda(s)ds$$

is a local martingale under $P$. □
Construction of counting processes by change of measure

Let $D^c[0, \infty)$ denote the collection of nonnegative integer-valued cadlag functions $x$ with $x(0) = 0$ and $x$ constant except for jumps of $+1$. Suppose that $\lambda : D^c[0, \infty) \times [0, \infty) \to [0, \infty)$,

$$\int_0^t \lambda(x, s) ds < \infty, \quad t \geq 0, x \in D^c[0, \infty)$$

and that $\lambda(x, s) = \lambda(x(\cdot \land s), s)$ (that is, $\lambda$ is nonanticipating). If we take $\lambda(t) = \lambda(N, t)$ and let $S_n = \inf \{ t : N(t) = n \}$, then defining $dP = L(S_n) dQ$ on $\mathcal{F}_{S_n}$,

$$N(t \land S_n) - \int_0^{t \land S_n} \lambda(N, s) ds$$

is a $\{ \mathcal{F}_{t \land S_n} \}$-martingale on $(\Omega, \mathcal{F}_{S_n}, P)$. In particular, $N^{S_n}(t) \equiv N(\cdot \land S_n)$ can be written as the solution of

$$N^{S_n}(t) = Y \left( \int_0^{t \land S_n} \lambda(N^{S_n}, s) ds \right)$$

for $Y$ a unit Poisson process.
Multivariate counting processes

Theorem 5.9 Let $\lambda_k(t, Z, N)$ be as in Condition 1.13, and let $N_k$, $k = 1, \ldots, m$, be independent unit Poisson processes that are independent of $Z$ on $(\Omega, \mathcal{F}, Q)$. Define

$$L(t) = \prod_{k=1}^{m} \exp \left\{ \int_0^t \log \lambda_k(s-, Z, N) dN_k(s) - \int_0^t (\lambda_k(s, Z, N) - 1) ds \right\},$$

and $dP_{|\mathcal{F}_{T \wedge S_n}} = L(T \wedge S_n) dQ_{|\mathcal{F}_{T \wedge S_n}}$. Then on $(\Omega, \mathcal{F}_{T \wedge S_n}, P)$, for $k = 1, \ldots, m$,

$$N_k(t \wedge S_n) - \int_0^{t \wedge S_n} \lambda_k(s, Z, N) ds,$

is a $\{\mathcal{F}_{t \wedge S_n}\}$-martingale.
Change of measure for Brownian motion

Let $W$ be standard Brownian motion, and let $\xi$ be an adapted process. Define

$$L(t) = \exp\left\{ \int_0^t \xi(s)dW(s) - \frac{1}{2} \int_0^t \xi^2(s)ds \right\}$$

and note that

$$L(t) = 1 + \int_0^t \xi(s)L(s)dW(s).$$

Then $L(t)$ is a local martingale.

Assume $E^Q[L(t)] = 1$ for all $t \geq 0$. Then $L$ is a martingale. Fix a time $T$, and restrict attention to the probability space $(\Omega, \mathcal{F}_T, Q)$. On $\mathcal{F}_T$, define $dP = L(T)dQ$.

For $t < T$, let $A \in \mathcal{F}_t$. Then

$$P(A) = E^Q[1_A L(T)] = E^Q[1_A E^Q[L(T)|\mathcal{F}_t]]$$

has no dependence on $T$

(crucial that $L$ is a martingale)
New Brownian motion

**Theorem 5.10** \( \widetilde{W}(t) = W(t) - \int_0^t \xi(s)ds \) is a standard Brownian motion on \( (\Omega, \mathcal{F}_T, P) \).

**Proof.** Since \( \widetilde{W} \) is continuous with quadratic variation \([\widetilde{W}]_t = t\) a.s., so by Lévy’s characterization of Brownian motion, it is enough to show that \( \widetilde{W} \) is a local martingale (and hence a martingale). But since \( W \) is a \( Q \)-martingale and \([L, W]_t = \int_0^t \xi(s)L(s)ds\), Theorem 5.7 gives the desired result. \( \square \)
Changing the drift of a diffusion

Suppose that

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s)$$

and set

$$\xi(s) = b(X(s)).$$

Note that $X$ is a diffusion with generator $\frac{1}{2}\sigma^2(x)f''(x)$. Define

$$L(t) = \exp\left\{ \int_0^t b(X(s))dW(s) - \frac{1}{2} \int_0^t b^2(X(s))ds \right\},$$

and assume that $E^Q[L(T)] = 1$ (e.g., if $b$ is bounded). Set $dP = L(T)dQ$ on $(\Omega, \mathcal{F}_T)$. 
Transformed SDE

Define \( \tilde{W}(t) = W(t) - \int_0^t b(X(s)) ds \). Then

\[
X(t) = X(0) + \int_0^t \sigma(X(s)) dW(s)
\]

\[
= X(0) + \int_0^t \sigma(X(s)) d\tilde{W}(s) + \int_0^t \sigma(X(s)) b(X(s)) ds
\]

so under \( P \), \( X \) is a diffusion with generator

\[
Af(x) = \frac{1}{2} \sigma^2(x) f''(x) + \sigma(x) b(x) f'(x).
\]
Conditions that imply local absolute continuity

Let \( Af(x) = \frac{1}{2} \sigma^2(x)f''(x) + \sigma(x)b(x)f'(x) \).

**Condition 5.11** If \( \nu \in P(\mathbb{R}) \) and \((X_n, \tau_n), n = 1, 2, \ldots \) satisfy \( X_n(0) \) has distribution \( \nu \),

\[
f(X_n(t \wedge \tau_n)) - f(X_n(0)) - \int_0^{t \wedge \tau_n} Af(X_n(s))ds,
\]

is an \( \{ \mathcal{F}_t^n \} \)-martingale for each \( f \in C^2_c(\mathbb{R}) \), and

\[
\tau_n = \inf\{t : \int_0^t b^2(X_n(s))ds \geq n\},
\]

then \( \lim_{n \to \infty} P\{\tau_n \leq T\} = 0 \) for each \( T > 0 \).

**Theorem 5.12** Suppose Condition 5.11 holds, and let \( W \) be a Brownian motion on \((\Omega, \mathcal{F}, Q)\). If \( X \) is a solution of

\[
X(t) = X(0) + \int_0^t \sigma(X(s))dW(s)
\]
on \((\Omega, \mathcal{F}, Q)\) and \( X(0) \) has distribution \( \nu \), then for each \( T \), there is a change of measure \( dP = L(T)dQ \) such that \( X \) on \((\Omega, \mathcal{F}_T, P)\) is a solution of the martingale problem for \( A \) on \([0, T]\).
Proof. Let
\[ L(t) = \exp\left\{ \int_0^t b(X(s))dW(s) - \frac{1}{2} \int_0^t b^2(X(s))ds \right\}, \]
and define \( \tau_n = \inf\{t : \int_0^t b^2(X(s))ds > n\} \). Then \( E^Q[L(T \wedge \tau_n)] = 1 \) and we can define \( dP = L(T \wedge \tau_n)dQ \) on \( \mathcal{F}_{T \wedge \tau_n} \). On \( (\Omega, \mathcal{F}_{T \wedge \tau_n}, P) \),
\[ \widetilde{W}(t \wedge \tau_n) = W(t \wedge \tau_n) - \int_0^{t \wedge \tau_n} b(X(s))ds \]
is a Brownian motion stopped at \( \tau_n \) and
\[ X(t) = X(0) + \int_0^t \sigma(X(s))d\widetilde{W}(s) + \int_0^t \sigma(X(s))b(X(s))ds, \]
for \( t \leq T \wedge \tau_n \). Then \((X, \tau_n), n = 1, 2, \ldots \) satisfies Condition 5.11, and since
\[ P\{L(T \wedge \tau_n) > K\} = P\left\{ \int_0^{T \wedge \tau_n} b(X(s))d\widetilde{W}(s) + \frac{1}{2} \int_0^{T \wedge \tau_n} b^2(X(s))ds > \log K \right\}, \]
we can apply Proposition 5.1 to conclude that \( P << Q \) on \( \mathcal{F}_T \), that is, \( E[L(T)] = 1 \).
\( \square \)
Maximum likelihood estimation

Suppose for each $\alpha \in \mathcal{A}$,

$$P_\alpha(\Gamma) = \int_{\Gamma} L_\alpha dQ,$$

and

$$L_\alpha = H(\alpha, X_1, X_2, \ldots X_n)$$

for random variables $X_1, \ldots, X_n$. The maximum likelihood estimate $\hat{\alpha}$ for the “true” parameter $\alpha_0 \in \mathcal{A}$ based on observations of the random variables $X_1, \ldots, X_n$ is the value of $\alpha$ that maximizes $H(\alpha, X_1, X_2, \ldots X_n)$.

For example, under certain conditions the distribution of

$$X_\alpha(t) = X(0) + \int_0^t \sigma(X_\alpha(s))dW(s) + \int_0^t b(X_\alpha(s), \alpha)ds,$$

will be absolutely continuous with respect to the distribution of $X$ satisfying

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s).$$

(5.6)
Sufficiency

If \( dP_\alpha = L_\alpha dQ \) where
\[
L_\alpha(X, Y) = H_\alpha(X)G(X, Y),
\]
then \( X \) is a sufficient statistic for \( \alpha \). Without loss of generality, we can assume \( E^Q[G(X, Y)] = 1 \) and hence \( d\hat{Q} = G(X, Y)dQ \) defines a probability measure.

**Example 5.13** If \((X_1, \ldots, X_n)\) are iid \( N(\mu, \sigma^2) \) under \( P(\mu, \sigma) \) and \( Q = P(0,1) \), then
\[
L(\mu, \sigma) = \frac{1}{\sigma^n} \exp \left\{ -\frac{1 - \sigma^2}{2\sigma^2} \sum_{i=1}^{n} X_i^2 + \frac{\mu}{\sigma^2} \sum_{i=1}^{n} X_i - \frac{\mu^2}{\sigma^2} \right\}
\]
so \((\sum_{i=1}^{n} X_i^2, \sum_{i=1}^{n} X_i)\) is a sufficient statistic for \((\mu, \sigma)\).
Parameter estimates and sufficiency

**Theorem 5.14** If \( \hat{\theta}(X, Y) \) is an estimator of \( \theta(\alpha) \) and \( \varphi \) is convex, then

\[
E^{P_\alpha}[\varphi(\theta(\alpha) - \hat{\theta}(X, Y))] \geq E^{P_\alpha}[\varphi(\theta(\alpha) - E^{Q}[\hat{\theta}(X, Y)|X])]
\]

**Proof.**

\[
E^{P_\alpha}[\varphi(\theta(\alpha) - \hat{\theta}(X, Y))] = E^{\hat{Q}}[\varphi(\theta(\alpha) - \hat{\theta}(X, Y))H_\alpha(X)]
\]

\[
= E^{\hat{Q}}[E^{\hat{Q}}[\varphi(\theta(\alpha) - \hat{\theta}(X, Y))|X]H_\alpha(X)]
\]

\[
\geq E^{\hat{Q}}[\varphi(\theta(\alpha) - E^{\hat{Q}}[\hat{\theta}(X, Y)|X])H_\alpha(X)]
\]
Other applications

**Finance:** Asset pricing models depend on finding a change of measure under which the price process becomes a martingale.

**Stochastic Control:** For a controlled diffusion process

\[ X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s), u(s))ds \]

where the control only enters the drift coefficient, the controlled process can be obtained from an uncontrolled process satisfying (5.6) via a change of measure.
6. Martingale problems for conditional distributions

- Conditional distributions for martingale problems
- Partially observed processes
- Filtered martingale problem
- Markov mapping theorem
- Burke’s theorem
A martingale lemma

Let \( \{ \mathcal{F}_t \} \) and \( \{ \mathcal{G}_t \} \) be filtrations with \( \mathcal{G}_t \subset \mathcal{F}_t \).

**Lemma 6.1** Suppose that

\[
M(t) = U(t) - \int_0^t V(s) ds
\]

is an \( \{ \mathcal{F}_t \} \)-martingale. Then

\[
E[U(t)|\mathcal{G}_t] - \int_0^t E[V(s)|\mathcal{G}_s] ds
\]

is a \( \{ \mathcal{G}_t \} \)-martingale.

**Proof.** The lemma follows by the definition and properties of conditional expectations. \(\square\)
Martingale properties of conditional distributions

**Corollary 6.2** If $X$ is a solution of the martingale problem for $A$ with respect to the filtration $\{F_t\}$ and $\pi_t$ is the conditional distribution of $X(t)$ given $G_t \subset F_t$, then

$$\pi_t f - \pi_0 f - \int_0^t \pi_s Af \, ds \quad (6.1)$$

is a $\{G_t\}$-martingale for each $f \in D(A)$. 
Forward equation

Recall that a $\mathcal{P}(E)$-valued function $\{\nu_t, t \geq 0\}$ is a solution of the forward equation for $A$ if for each $t > 0$, $\int_0^t \nu_s \psi ds < \infty$ (see Condition 4.7) and for each $f \in \mathcal{D}(A)$,

$$\nu_t f = \nu_0 f + \int_0^t \nu_s Af ds.$$  \hspace{1cm} (6.2)

Note that if $\pi$ satisfies (6.1), then $\nu_t = E[\pi_t]$ satisfies (6.2).
Martingale characterization of conditional distributions

**Theorem 6.3** Suppose that \( \{\tilde{\pi}_t, t \geq 0\} \) is a cadlag, \( \mathcal{P}(E) \)-valued process with no fixed points of discontinuity adapted to \( \{\tilde{G}_t\} \) satisfying

\[
E\left[ \int_0^t \tilde{\pi}_s \psi ds \right] < \infty, \quad t > 0
\]

and that

\[
\tilde{\pi}_t f - \tilde{\pi}_0 f - \int_0^t \tilde{\pi}_s Af ds
\]

is a \( \{\tilde{G}_t\} \)-martingale for each \( f \in \mathcal{D}(A) \). Then there exists a solution \( X \) of the martingale problem for \( A \), a \( \mathcal{P}(E) \)-valued process \( \{\pi_t, t \geq 0\} \) with the same distribution as \( \{\tilde{\pi}_t, t \geq 0\} \), and a filtration \( \{G_t\} \) such that \( \pi_t \) is the conditional distribution of \( X(t) \) given \( G_t \).
Conditioning on a process

**Theorem 6.4** If $\{\tilde{G}_t\}$ in Theorem 6.3 is generated by a cadlag process $\tilde{Y}$ with no fixed points of discontinuity and $\tilde{\pi}(0)$, that is,

$$\tilde{G}_t = \mathcal{F}^\tilde{Y}_t \vee \sigma(\tilde{\pi}(0)),$$

then there exists a solution $X$ of the martingale problem for $A$, a $\mathcal{P}(E)$-valued process $\{\pi_t, t \geq 0\}$, and a process $Y$ such that $\{\pi_t, t \geq 0\}$ and $Y$ have the same joint distribution as $\{\tilde{\pi}_t, t \geq 0\}$ and $\tilde{Y}$ and $\pi_t$ is the conditional distribution of $X(t)$ given $\mathcal{F}^Y_t \vee \sigma(\pi(0))$. 
Idea of proof

Enlarge the state space so that the current state of the process contains all information about the past of the observation $\tilde{Y}$.

Let $\{b_k\}, \{c_k\} \subset C_b(E_0)$ satisfy $0 \leq b_k, c_k \leq 1$, and suppose that the spans of $\{b_k\}$ and $\{c_k\}$ are bounded, pointwise dense in $B(E_0)$.

Let $a_1, a_2, \ldots$ be an ordering of the rationals with $a_i \geq 1$ and

$$
\tilde{V}_{ki}(t) = c_k(\tilde{Y}(0)) - a_i \int_0^t \tilde{V}_{ki}(s) ds + \int_0^t b_k(\tilde{Y}(s)) ds
$$

(6.3)

$$
= c_k(\tilde{Y}(0)) e^{-a_it} + \int_0^t e^{-a_i(t-s)} b_k(\tilde{Y}(s)) ds.
$$
Set $\tilde{V}(t) = (\tilde{V}_{ki}(t) : k, i \geq 1) \in [0, 1]^{\infty}$,

$$\mathcal{D}(\hat{A}) = \{ f(x) \prod_{k,i=1}^{m} g_{ki}(v_{ki}) : f \in \mathcal{D}(A), g_{ki} \in C^1[0, 1], m = 1, 2, \ldots \}$$

and

$$\hat{A}(fg)(x, v, u) = g(v)Af(x) + f(x) \sum (-a_i v + b_k(u)) \partial_{ki} g(v),$$

For $fg \in \mathcal{D}(\hat{A})$,

$$\tilde{\pi}_t fg(\tilde{V}(t)) - \tilde{\pi}_0 fg(\tilde{V}(0))$$

$$- \int_0^t \left( g(\tilde{V}(s))\tilde{\pi}_s Af + \tilde{\pi}_s f \sum (-a_i \tilde{V}_{ki}(s) + b_k(\tilde{Y}(s))) \partial_{ki} g(\tilde{V}(s)) \right) ds$$

is a $\{\mathcal{F}_t^{\tilde{Y}}\}$-martingale and $\nu_t$ defined by

$$\nu_t(fgh) = E[\tilde{\pi}_t fg(\tilde{V}(t)) h(\tilde{Y}(t))]$$

is a solution of the controlled forward equation for $\hat{A}$. 
Partially observed processes

Let \( \gamma : E \rightarrow E_0 \) be Borel measurable.

**Corollary 6.5** If in Corollary 6.4, \( \tilde{Y} \) and \( \tilde{\pi} \) satisfy

\[
\int_E h \circ \gamma(x) \tilde{\pi}_t(dx) = h(\tilde{Y}(t)) \quad \text{a.s.}
\]

for all \( h \in B(E_0) \) and \( t \geq 0 \), then \( Y(t) = \gamma(X(t)) \).

The filtered martingale problem

**Definition 6.6** A $\mathcal{P}(E)$-valued process $\tilde{\pi}$ and an $E$-valued process $\tilde{Y}$ are a solution of the filtered martingale problem for $(A, \gamma)$ if

$$\tilde{\pi}_t f - \tilde{\pi}_0 f - \int_0^t \tilde{\pi}_s A f ds$$

is a $\{\mathcal{F}^{\tilde{Y}}_t \lor \sigma(\tilde{\pi}(0))\}$-martingale for each $f \in \mathcal{D}(A)$ and $\int_E h \circ \gamma(x) \tilde{\pi}_t(dx) = h(\tilde{Y}(t))$ a.s. for all $h \in B(E_0)$ and $t \geq 0$.

**Theorem 6.7** Let $\varphi_0 \in \mathcal{P}(\mathcal{P}(E))$ and define $\mu_0 = \int_{\mathcal{P}(E)} \mu \varphi_0(d\mu)$. If uniqueness holds for the martingale problem $(A, \mu_0)$, then uniqueness holds for the filtered martingale problem for $(A, \gamma, \varphi_0)$. If uniqueness holds for the filtered martingale problem for $(A, \gamma, \varphi_0)$, then $\{\pi_t, t \geq 0\}$ is a Markov process.
Markov mappings

Theorem 6.8 $\gamma : E \to E_0$, Borel measurable.

$\alpha$ a transition function from $E_0$ into $E$ satisfying

$$\alpha(y, \gamma^{-1}(y)) = 1$$

Let $\mu_0 \in \mathcal{P}(E_0)$, $\nu_0 = \int \alpha(y, \cdot) \mu_0(dy)$, and define

$$C = \{(\int_E f(z) \alpha(\cdot, dz), \int_E Af(z) \alpha(\cdot, dz)) : f \in \mathcal{D}(A)\}.$$  

If $\tilde{Y}$ is a solution of the MGP for $(C, \mu_0)$, then there exists a solution $Z$ of the MGP for $(A, \nu_0)$ such that $Y = \gamma \circ Z$ and $\tilde{Y}$ have the same distribution on $M_{E_0}[0, \infty)$.

$$E[f(Z(t))|\mathcal{F}_t^Y] = \int f(z) \alpha(Y(t), dz)$$

(at least for almost every $t$, all $t$ if $Y$ has no fixed points of discontinuity).
Uniqueness

**Corollary 6.9** If uniqueness holds for the MGP for \((A, \nu_0)\), then uniqueness holds for the \(M_{E_0}[0, \infty)\)-MGP for \((C, \mu_0)\). If \(\tilde{Y}\) has sample paths in \(D_{E_0}[0, \infty)\), then uniqueness holds for the \(D_{E_0}[0, \infty)\)-martingale problem for \((C, \mu_0)\).

Existence for \((C, \mu_0)\) and uniqueness for \((A, \nu_0)\) implies existence for \((A, \nu_0)\) and uniqueness for \((C, \mu_0)\), and hence that \(\tilde{Y}\) is Markov.
Intertwining condition

Let $\alpha(y, \Gamma)$ be a transition function from $E_0$ to $E$ satisfying

$$\alpha(y, \gamma^{-1}(y)) = 1,$$

and define $S(t) : B(E_0) \to B(E_0)$ by

$$S(t)g(y) = \alpha T(t)g \circ \gamma(y) \equiv \int_{E} T(t)g \circ \gamma(x) \alpha(y, dx).$$

**Theorem 6.10** (Rogers and Pitman (1981), cf Rosenblatt (1966)) If for each $t \geq 0$,

$$\alpha T(t)f = S(t)\alpha f, \quad f \in B(E), \quad (S(t) \text{ is a semigroup})$$

and $X$ is a Markov process with intial distribution $\alpha(y, \cdot)$ and semigroup $\{T(t)\}$, then $Y$ is a Markov process with $Y(0) = y$ and

$$P\{X(t) \in \Gamma | \mathcal{F}_t^Y\} = \alpha(Y(t), \Gamma).$$
Generator for $Y$

Note that

$$\alpha T(t)f = S(t)\alpha f, \quad f \in B(E),$$

suggests that the generator for $Y$ is given by

$$C\alpha f = \alpha Af.$$
Burke’s output theorem
Kliemann, Koch and Marchetti

\[ X = (Q, D), \text{ an } M/M/1 \text{ queue and its departure process} \]

\[ Af(k, l) = \lambda(f(k + 1, l) - f(k, l)) + \mu 1_{\{k > 0\}}(f(k - 1, l + 1) - f(k, l)) \]

\[ \gamma(k, l) = l \]

Assume \( \lambda < \mu \) and define

\[ \alpha(l, \{(k, l)\}) = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^{k-1}, \quad k = 0, 1, 2, \ldots \]

\[ \alpha(l, \{(k, m)\}) = 0, \quad m \neq l \]

Then

\[ \alpha Af(l) = \mu \sum_{k=1}^{\infty} (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^{k-1}(f(k - 1, l + 1) - f(k - 1, l)) \]

\[ = \lambda(\alpha f(l + 1) - \alpha f(l)) \]
Therefore, there exists a solution \((Q, D)\) of the martingale problem for \(A\) such that \(D\) is a Poisson process with parameter \(\lambda\) and

\[
P\{Q(t) = k | \mathcal{F}_t^D\} = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^{k-1},
\]

that is, \(Q(t)\) is independent of \(\mathcal{F}_t^D\) and is geometrically distributed.
Pitman’s theorem

$Z$ standard Brownian motion

$M(t) = \sup_{s \leq t} Z(s), \quad V(t) = M(t) - Z(t)$

$X(t) = (Z(t), M(t) - Z(t)) = (Z(t), V(t))$

$Y(t) = 2M(t) - Z(t) = \gamma(X(t) = 2V(t) + Z(t)$

$Af(z, v) = \frac{1}{2}f_{zz}(z, v) - f_{zv}(z, v) + \frac{1}{2}f_{vv}(z, v)$  \hspace{1cm} \text{b.c.} \quad f_v(z, 0) = 0$

$F(y) = \alpha f(y) = \frac{1}{y} \int_0^y f(y - 2v, v) dv$

$\alpha Af(y) = \frac{1}{2}F''(y) + \frac{1}{y}F'(y)$
7. **Stochastic equations for Markov processes in** $\mathbb{R}^d$

- Markov property
- Gaussian white noise
- Poisson random measure
- Pure jump processes
- Generators in $\mathbb{R}^d$
- Stochastic equations for Markov processes
- Martingale problems
- Change of measure for Poisson random measures
- Conditions for uniqueness
- Uniqueness and the Markov property
- Equations for spatial birth and death processes
Markov chains

Markov property: “Given the present, the future is independent of the past,” or “the future is a function of the present and inputs independent of the past.”

Discrete time: $X_0, \xi_1, \xi_2, \ldots$ independent

\[ X_{n+1} = H_{n+1}(X_n, \xi_{n+1}). \]

By iteration

\[ X_{n+1} = H_{k,n+1}(X_k, \xi_{k+1}, \ldots, \xi_{n+1}). \]
Continuous time: Processes with independent increments

Replace the sequence \( \{\xi_k\} \) by a process \( \xi \) with independent increments:

\[
X(s) = H_{t,s}(X(t), \xi(s \wedge (t + \cdot)) - \xi(t)).
\]

Essentially two choices for \( \xi \): Standard Brownian motion and a Poisson process.
Gaussian white noise integral

$\mu_0$ a $\sigma$-finite Borel measure on $(S_0, r_0)$

$A(S_0) = \{ A \in \mathcal{B}(U) : \mu_0(A) < \infty \}$

$W(A \times [0, t])$ normal with $E[W(A \times [0, t])] = 0$ and $Var(W(A \times [0, t])) = \mu_0(A)t$ and

$$E[W(A \times [0, t))W(B \times [0, s))] = \mu_0(A \cap B)t \land s$$

Then $W(\bigcup A_i \times [0, t]) = \sum W(A_i \times [0, t])$

$$Z(t) = \int_{U \times [0, t]} Y(u, s)W(du \times ds)$$

satisfies

$$E[Z(t)] = 0 \quad [Z]_t = \int_0^t \int_U Y^2(u, s)\mu_0(du)ds$$

$$E[Z^2(t)] = E[[Z]_t] = \int_0^t \int_U E[Y^2(u, s)]\mu_0(du)ds$$
Space-time Poisson random measures

\( \mu_1 \) a \( \sigma \)-finite, Borel measure on \((U, r_1)\)

\( \xi \) is a Poisson random measure on \( U \times [0, \infty) \) with mean measure \( \mu_1 \times \ell \) (where \( \ell \) denotes Lebesgue measure).

\[ \xi(A, t) \equiv \xi(A \times [0, t]) \] is a Poisson process with parameter \( \mu_1(A) \).

\[ \tilde{\xi}(A, t) \equiv \xi(A \times [0, t]) - \mu_1(A)t \] is a martingale.

**Definition 7.1** \( \xi \) is \( \{F_t\} \)-compatible, if for each \( A \in \mathcal{A}(U) \), \( \xi(A, \cdot) \) is \( \{F_t\} \) adapted and for all \( t, s \geq 0 \), \( \xi(A \times (t, t + s]) \) is independent of \( F_t \).
Stochastic integrals for Poisson random measures

For \( t_i < r_i, A_i \in \mathcal{B}(U) \), and \( \eta_i \mathcal{F}_{t_i} \)-measurable:

\[
X(u, t) = \sum_i \eta_i 1_{A_i(u)} 1_{[t_i, r_i)}(t), \quad \text{so} \quad X(u, t-) = \sum_i \eta_i 1_{A_i(u)} 1_{(t_i, r_i]}(t).
\]

(7.1)

Define

\[
I_\xi(X, t) = \int_{U \times [0, t]} X(u, s-) \xi(du \times ds) = \sum_i \eta_i \xi(A_i \times (t_i, r_i]).
\]

Then

\[
E[|I_\xi(X, t)|] \leq E\left[\int_{U \times [0, t]} |X(u, s-)\xi(du \times ds)\right] = \int_{U \times [0, t]} E[|X(u, s)|] \mu_1(du)ds
\]

and if the right side is finite, \( E[I_\xi(X, t)] = \int_{U \times [0, t]} E[X(u, s)] \mu_1(du)ds. \)
Stochastic integrals for centered Poisson random measures

Let \( \tilde{\xi}(du \times ds) = \xi(du \times ds) - \mu_1(du)ds \)

For

\[ X(u, t-) = \sum_i \eta_i 1_{A_i}(u) 1_{(t_i, r_i]}(t). \]

as in (7.1), define

\[ I_{\tilde{\xi}}(X, t) = \int_{U \times [0, t]} X(u, s-) \tilde{\xi}(du \times ds) = \int_{U \times [0, t]} X(u, s) \xi(du \times ds) - \int_0^t \int_U X(u, s) \mu_1(du)ds \]

and note that

\[ E \left[ I_{\tilde{\xi}}(X, t)^2 \right] = \int_{U \times [0, t]} E[X(u, s)^2] \mu_1(du)ds \]

if the right side is finite.

Then \( I_{\tilde{\xi}}(X, \cdot) \) is a square-integrable martingale.
An equation for a pure jump process

Consider a generator of the form

$$Af(x) = \lambda(x) \int_0^t (f(x + y) - f(x))\eta(x, dy)$$

for a process in $\mathbb{R}^d$. There exists a function $H : \mathbb{R}^d \times [0, 1]$ such for $V$ uniform $[0, 1]$, $H(x, V)$ has distribution $\eta(x, \cdot)$.

Let $\xi$ be a Poisson random measure on $[0, \infty) \times [0, 1] \times [0, \infty)$ with mean measure $\ell \times \ell \times \ell$. Then the solution of

$$X(t) = X(0) + \int_{[0, \infty) \times [0, 1] \times [0, t]} 1_{[0, \lambda(X(s-))]}(v)H(X(s-), u)\xi(dv \times du \times ds).$$
Markov processes in $\mathbb{R}^d$

Typically, a Markov process $X$ in $\mathbb{R}^d$ has a generator of the form

$$Af(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \hat{b}(x) \cdot \nabla f(x) + \int_{\mathbb{R}^d} (f(x+y) - f(x) - 1_{B_1}(y) y \cdot \nabla f(x)) \eta(x, dy)$$

where $B_1$ is the ball of radius 1 centered at the origin and $\eta$ satisfies

$$\int 1 \wedge |y|^2 |\eta(x, dy) < \infty$$

for each $x$. (See, for example, Stroock (1975), Çinlar, Jacod, Protter, and Sharpe (1980).)

$\eta(x, \Gamma)$ gives the “rate” at which jumps satisfying $X(s) - X(s-) \in \Gamma$ occur.

$B_1$ can be replaced by any set $C$ containing an open neighborhood of the origin provided that the drift term is replaced by

$$b_C(x) \cdot \nabla f(x) = \left( \hat{b}(x) + \int_{\mathbb{R}^d} y (1_C(y) - 1_{B_1}(y)) \eta(x, dy) \right) \cdot \nabla f(x).$$
A representation for $\eta$

We will assume that there exist $\lambda : \mathbb{R}^d \times S \rightarrow [0, 1]$, $\gamma : \mathbb{R}^d \times S \rightarrow \mathbb{R}^d$, and a $\sigma$-finite measure $\mu_1$ on a complete, separable metric space $S$ such that

$$\eta(x, \Gamma) = \int_S \lambda(x, u) 1_{\Gamma}(\gamma(x, u)) \mu_1(du).$$

This representation is always possible but in no way unique.
Reformulation of the generator

For simplicity, assume that there exists a fixed set $S_1 \in S$ such that for $S_2 = S - S_1$,

$$
\int_S \lambda(x, u)(\mathbf{1}_{S_1}(u)|\gamma(x, u)|^2 + \mathbf{1}_{S_2}(u)|\gamma(x, u)||\mu_1(du) < \infty
$$

and

$$
\int_S \lambda(x, u)|\gamma(x, u)||\mathbf{1}_{S_1}(u) - \mathbf{1}_{B_1}(\gamma(x, u))|\mu_1(du) < \infty.
$$

Then

$$
Af(x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + b(x) \cdot \nabla f(x)
$$

$$
+ \int_S \lambda(x, u)(f(x + \gamma(x, u)) - f(x) - \mathbf{1}_{S_1}(u)\gamma(x, u) \cdot \nabla f(x))\mu_1(du)
$$

where

$$
b(x) = \hat{b}(x) + \int_S \lambda(x, u)\gamma(x, u)(\mathbf{1}_{S_1}(u) - \mathbf{1}_{B_1}(\gamma(x, u)))\mu_1(du).
$$
Technical assumptions

For each compact $K \subset \mathbb{R}^d$

$$\sup_{x \in K} (|b(x)| + \int_{S_0} |\sigma(x, u)|^2 \mu_0(du) + \int_{S_1} \lambda(x, u)|\gamma(x, u)|^2 \mu_1(du)$$

$$+ \int_{S_2} \lambda(x, u)|\gamma(x, u)| \wedge 1 \mu_1(du)) < \infty.$$ 

$$a(x) = \int_{S_0} \sigma(x, u)\sigma(x, u)^T \mu_0(du)$$

Let $\mathcal{D}(A) = C^2_c(\mathbb{R}^d)$ and assume that for $f \in \mathcal{D}(A)$, $Af \in C_b(\mathbb{R}^d)$.

The continuity assumption can be removed and the boundedness assumption relaxed using existing technology.

For $x$ outside the support of $f$,

$$Af(x) = \int_{S} \lambda(x, u)(f(x + \gamma(x, u))) \mu_1(du).$$
Itô equations

$X$ should satisfy a stochastic differential equation of the form

$$X(t) = X(0) + \int_{S_0 \times [0, t]} \sigma(X(s), u)W(du \times ds) + \int_0^t b(X(s))ds \quad (7.2)$$

$$+ \int_{[0,1] \times S_1 \times [0, t]} 1_{[0, \lambda(X(s-), u)]}(v)\gamma(X(s-), u)\tilde{\xi}(dv \times du \times ds)$$

$$+ \int_{[0,1] \times S_2 \times [0, t]} 1_{[0, \lambda(X(s-), u)]}(v)\gamma(X(s-), u)\xi(dv \times du \times ds),$$

for $t < \tau_\infty \equiv \lim_{k \to \infty} \inf\{t : |X(t-)| \text{ or } |X(t)| \geq k\}$, where $W$ is Gaussian white noise determined by $\mu_0$ and $\xi$ is a Poisson random measure on $[0, 1] \times S \times [0, \infty)$ with mean measure $\ell \times \mu_1 \times \ell$. 
Martingale properties/problems

Assume that $\tau_\infty = \infty$. Applying Itô’s formula,

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds$$

$$= \int_{S_0 \times [0,t]} \nabla f(X(s))^T \sigma(X(s), u) W(du \times ds)$$

$$+ \int_{[0,1] \times S \times [0,t]} 1_{[0,\lambda(X(s-),u)]}(v)(f(X(s-)) + \gamma(X(s-), u)) - f(X(s-)))\tilde{\xi}(dv \times du \times ds)$$

The right side is a local martingale, so under the assumption that $Af$ is bounded, it is a martingale.

**Definition 7.2** $X$ is a solution of the martingale problem for $A$ if there exists a filtration $\{\mathcal{F}_t\}$ such that

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds \quad (7.3)$$

is a $\{\mathcal{F}_t\}$-martingale for each $f \in \mathcal{D}(A)$. 
Change of measure for Poisson random measures

Let $\lambda$ be a nonnegative, $\{\mathcal{F}_t\}$-adapted, cadlag (in $t$) process satisfying

$$
\int_{S \times [0, t]} (\lambda(u, s) - 1)^2 \wedge |\lambda(u, s) - 1| \mu_1(du)ds < \infty \quad \text{a.s., } t \geq 0.
$$

Let $M_\lambda(t) = \int_{S \times [0, t]} (\lambda(u, s) - 1) \tilde{\xi}(du \times ds)$ ($M_\lambda$ is a $\{\mathcal{F}_t\}$-martingale), and let $L$ be the solution (which will be a local martingale) of

$$
L(t) = 1 + \int_0^t L(s-)dM_\lambda(s) = 1 + \int_{S \times [0, t]} (\lambda(u, s) - 1)L(s-) \tilde{\xi}(du \times ds).
$$

(7.4)

If $\int_{S \times [0, t]} |\lambda(u, s) - 1| \mu_1(du)ds < \infty \quad \text{a.s., } t \geq 0$, then

$$
L(t) = \exp\left\{ \int_{S \times [0, t]} \log \lambda(u, s) \xi(du \times ds) - \int_{S \times [0, t]} (\lambda(u, s) - 1) \mu_1(du)ds \right\},
$$

and in general,

$$
L(t) = \exp\left\{ \int_{S \times [0, t]} \log \lambda(u, s) \tilde{\xi}(du \times ds) + \int_{S \times [0, t]} (\log \lambda(u, s) - \lambda(u, s) + 1) \mu_1(du)ds \right\}.
$$
Intensity for the transformed counting measure

If $E[L(T)] = 1$, then for $A \in \mathcal{B}(S)$ with $\mu_1(A) < \infty$,

$$M_A(t) = \int_{A \times [0,t]} \lambda(u, s) \tilde{\xi}(du \times ds)$$

is a local martingale under $Q$, and since

$$[M_A, L]_t = \int_{A \times [0,t]} \lambda(u, s)(\lambda(u, s) - 1)L(s-)\xi(du \times ds),$$

$$Z_A(t) = \int_{A \times [0,t]} \lambda(u, s) \tilde{\xi}(du \times ds)$$

$$- \int_{A \times [0,t]} \frac{1}{L(s)} \lambda(u, s)(\lambda(u, s) - 1)L(s-)\xi(du \times ds)$$

$$= \xi(A, t) - \int_{A \times [0,t]} \lambda(u, s)\mu_1(du)ds$$

is a local martingale under $dP = L(T)dQ$. 
Change for stochastic equations

Assuming $S = S_1 \cup S_2$, $S_1 \cap S_2 = \emptyset$, consider

$$Af(x) = \int_{S_1} (f(x + \alpha(x, u)) - f(x) - \alpha(x, u) \cdot \nabla f(x)) \mu_1(du)$$

$$+ \int_{S_2} (f(x + \alpha(x, u)) - f(x)) \mu_1(du)$$

Let $\mathcal{D}(A) = C^2_c(\mathbb{R}^d)$ and suppose that $Af$ is bounded for $f \in \mathcal{D}(A)$. Then $X$ satisfying

$$X(t) = X(0) + \int_{S_1 \times [0,t]} \alpha(X(s-), u) \tilde{\xi}(du \times ds) + \int_{S_2 \times [0,t]} \alpha(X(s-), u) \xi(du \times ds)$$

is a solution of the martingale problem for $A$. 
New martingale problem

Let \( \lambda(u, s) = \lambda(u, X(s^-)) \). Under \( Q \)

\[
f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds
\]

\[
= \int_{S \times [0, t]} (f(X(s^-) + \alpha(X(s^-), u)) - f(X(s^-)) \tilde{\xi}(du \times ds)
\]

is a local martingale, so under \( P \),

\[
f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds
\]

\[
- \int_0^t \int_S (f(X(s) + \alpha(X(s), u)) - f(X(s))(\lambda(u, X(s)) - 1)\mu_1(du)ds
\]

\[
= \int_{S \times [0, t]} (f(X(s^-) + \alpha(X(s^-), u)) - f(X(s^-))
\]

\[
\times (\xi(du \times ds) - \lambda(u, X(s))\mu_1(du)ds)
\]

is a local martingale.
New generator

Under $P$, $X$ is a solution of the martingale problem for

$$A_1 f(x) = \int_{S_1} (f(x + \alpha(x, u)) - f(x) - \alpha(x, u) \cdot \nabla f(x)) \lambda(u, x) \mu_1(du)$$

$$+ \int_{S_2} (f(x + \alpha(x, u)) - f(x)) \lambda(x, u) \mu_1(du)$$

$$+ \int_{S_1} \alpha(x, u)(\lambda(x, u) - 1) \mu_1(du) \cdot \nabla f(x)$$
Conditions for uniqueness for general SDE

In Itô (1951) as well as in later presentations (for example, Skorokhod (1965) and Ikeda and Watanabe (1989)), $L_2$-estimates are used to prove uniqueness for (7.2).

Graham (1992) points out the possibility and desirability of using $L_1$-estimates. (In fact, for equations controlling jump rates with factors like $1_{[0,\lambda(X(t),u)]}(v)$, $L_1$-estimates are essential.)

Kurtz and Protter (1996) develop methods that allow a mixing of $L_1$, $L_2$, and other estimates.

The uniqueness proof given here uses only $L_1$ estimates.
Theorem 7.3  Suppose there exists a constant $M$ such that

\[ |b(x)| + \int_{S_0} |\sigma(x, u)|^2 \mu(du) + \int_{S_1} |\gamma(x, u)|^2 \lambda(x, u) \mu_1(du) \]

\[ + \int_{S_2} \lambda(x, u) |\gamma(x, u)| \mu_1(du) < M, \]  

and

\[ \sqrt{\int_{S_0} |\sigma(x, u) - \sigma(y, u)|^2 \mu_0(du)} \leq M |x - y| \]  

\[ |b(x) - b(y)| \leq M |x - y| \]  

\[ \int_{S_1} (\gamma(x, u) - \gamma(y, u))^2 \lambda(x, u) \wedge \lambda(y, u) \mu_1(du) \leq M |x - y|^2 \]  

\[ \int_{S_2} \lambda(x, u) ||\gamma(x, u) - \gamma(y, u)|| \mu_1(du) \leq M |x - y| \]  

\[ \int_{S} |\lambda(x, u) - \lambda(y, u)||\gamma(y, u)|| \mu_1(du) \leq M |x - y|. \]  

Then there exists a unique solution of (7.2).
Proof. Suppose $X$ and $Y$ are solutions of (7.2). Then

\[
X(t) = X(0) + \int_{S_0 \times [0,t]} \sigma(X(s), u) W(du \times ds) + \int_0^t b(X(s)) ds \\
+ \int_{[0,\infty) \times S_1 \times [0,t]} 1_{[0,\lambda(X(s),u) \wedge \lambda(Y(s),u)]}(v) \gamma(X(s-), u) \tilde{\xi}(dv \times du \times ds) \\
+ \int_{[0,\infty) \times S_1 \times [0,t]} 1_{(\lambda(Y(s-),u) \wedge \lambda(X(s-),u), \lambda(X(s-),u)]}(v) \gamma(X(s-), u) \tilde{\xi}(dv \times du \times ds) \\
+ \int_{[0,\infty) \times S_2 \times [0,t]} 1_{[0,\lambda(X(s-),u)]}(v) \gamma(X(s-), u) \xi(dv \times du \times ds),
\]

and similarly with the roles of $X$ and $Y$ interchanged. Then (7.6) and (7.7) give the necessary Lipschitz conditions for the coefficient functions in the first two integrals on the right, (7.8) gives an $L_2$-Lipschitz condition for the third integral term, and (7.9) and (7.10) give $L_1$-Lipschitz conditions for the fourth and fifth integral terms. Theorem 7.1 of Kurtz and Protter (1996) gives uniqueness. □
Extension to unbounded coefficients

**Corollary 7.4** Suppose that there exists a function $M(r)$ defined for $r > 0$, such that (7.5) through (7.10) hold with $M$ replaced by $M(|x| \vee |y|)$. Then there exists a stopping time $\tau_\infty$ and a process $X(t)$ defined for $t \in [0, \tau_\infty)$ such that (7.2) is satisfied on $[0, \tau_\infty)$ and $\tau_\infty = \lim_{k \to \infty} \inf \{t : |X(t)| \text{ or } |X(t-)\} \geq k\}$. If $(Y, \tau)$ also has this property, then $\tau = \tau_\infty$ and $Y(t) = X(t)$, $t < \tau_\infty$.

**Proof.** The corollary follows by a standard localization argument. □
A martingale inequality

If $M$ is a local square integrable martingale, there exists a nondecreasing predictable process $\langle M \rangle$ such that $M(t)^2 - \langle M \rangle_t$ is a local martingale. In particular, $[M] - \langle M \rangle$ is a local martingale. Recall that a left-continuous, adapted process is predictable.

**Lemma 7.5** For $0 < p \leq 2$ there exists a constant $C_p$ such that for any local square integrable martingale $M$ with Meyer process $\langle M \rangle$ and any stopping time $\tau$

$$E[\sup_{s \leq \tau} |M(s)|^p] \leq C_p E[\langle M \rangle_{\tau}^{p/2}]$$

The discrete time version is due to Burkholder (1973) and the continuous time version to Lenglart, Lépingle, and Pratelli (1980). The following proof is due to Ichikawa (1986).
Proof. For $p = 2$ the result is an immediate consequence of Doob’s inequality. Let $0 < p < 2$. For $x > 0$, let $\sigma_x = \inf\{t : \langle M \rangle_t > x^2\}$. Since $\sigma_x$ is predictable, there exists a strictly increasing sequence of stopping times $\sigma^n_x \to \sigma_x$. Noting that $\langle M \rangle_{\sigma^n_x} \leq x^2$, we have

$$P\{\sup_{s \leq \tau} |M(s)| > x\} \leq P\{\sigma^n_x < \tau\} + P\{\sup_{s \leq \tau \wedge \sigma^n_x} |M(s)| > x\}$$

$$\leq P\{\sigma_x < \tau\} + \frac{E[\langle M \rangle_{\tau \wedge \sigma^n_x}]}{x^2} \leq P\{\sigma^n_x < \tau\} + \frac{E[x^2 \wedge \langle M \rangle_{\tau}]}{x^2},$$

and letting $n \to \infty$, we have

$$P\{\sup_{s \leq \tau} |M(s)| > x\} \leq P\{\langle M \rangle_{\tau} \geq x^2\} + \frac{E[x^2 \wedge \langle M \rangle_{\tau}]}{x^2}. \quad (7.12)$$

Using the identity $\int_0^\infty E[x^2 \wedge X^2]px^{p-3}dx = \frac{2}{2-p}E[|X|^p]$, the lemma follows by multiplying both sides of (7.12) by $px^{p-1}$ and integrating. \qed
Computation of Meyer process

Let $\xi$ be a Poisson random measure on $S \times [0, \infty)$ with mean measure $\mu_1 \times \ell$, and let $W$ be a Gaussian white noise on $S_0 \times [0, \infty)$ determined by $\mu_0 \times \ell$. Assume that $\xi$ and $W$ are compatible with a filtration $\{F_t\}$.

If $X$ is predictable and $\int_{S \times [0,t]} X^2(u, s) \mu_1(du) ds < \infty$ a.s., then $M(t) = \int_{S \times [0,t]} X(u, s)\tilde{\xi}(du \times ds)$ is a local square integrable martingale with

$$\langle M \rangle_t = \int_{S \times [0,t]} X^2(u, s) \mu_1(du) ds < \infty.$$

If $X$ is adapted and $\int_0^t \int_{S_0} X(u, s)^2 \mu_0(du) ds < \infty$ a.s., then $M(t) = \int_{S_0 \times [0,t]} X(u, s)dW(du \times ds)$ is a local square integrable martingale with

$$\langle M \rangle_t = \int_0^t \int_{S_0} X(u, s)^2 \mu_0(du) ds.$$
Estimate for solutions of (7.11)

\[ E[\sup_{s \leq t} |X(s) - Y(s)|] \]
\[ \leq E[|X(0) - Y(0)|] + C_1 E\left[\left( \int_0^t \int_{S_0} |\sigma(X(s), u) - \sigma(Y(s), u)|^2 \mu_0(du)ds \right)^{\frac{1}{2}} \right] \]
\[ + C_1 E\left[\left( \int_0^t \int_{S_1} (\gamma(X(s), u) - \gamma(Y(s), u))^2 \lambda(X(s), u) \land \lambda(Y(s), u) \mu_1(du) \right)^{\frac{1}{2}} \right] \]
\[ + 2E\left[\int_0^t \int_{S_1} |\lambda(X(s), u) - \lambda(Y(s), u)|(||\gamma(X(s), u)| + |\gamma(Y(s), u)||) \mu_1(du)ds \right] \]
\[ + E\left[\int_0^t \int_{S_2} \lambda(X(s), u)|\gamma(X(s), u) - \gamma(Y(s), u)| \mu_1(du)ds \right] \]
\[ + E\left[\int_0^t \int_{S_2} |\lambda(X(s), u) - \lambda(Y(s), u)||\gamma(Y(s), u)| \mu_1(du)ds \right] \]
\[ + E\left[\int_0^t |b(X(s)) - b(Y(s))| ds \right] \]
\[ \leq E[|X(0) - Y(0)|] + D(\sqrt{t} + t) E[\sup_{s \leq t} |X(s) - Y(s)|] \]
For $t$ small enough so that $D(\sqrt{t} + t) \leq .5$, then

$$E[\sup_{s \leq t} |X(s) - Y(s)|] \leq 2E[|X(0) - Y(0)|]$$
Uniqueness and the Markov property  If $X$ is a solution of (7.2), then

$$X(t) = X(r) + \int_{S_0 \times [r,t]} \sigma(X(s), u) W(du \times ds) + \int_0^t b(X(s)) ds \quad (7.13)$$

$$+ \int_{[0,1] \times S_1 \times [r,t]} 1_{[0,\lambda(X(s-),u)]}(v) \gamma(X(s-), u) \tilde{\xi}(dv \times du \times ds)$$

$$+ \int_{[0,1] \times S_2 \times [r,t]} 1_{[0,\lambda(X(s-),u)]}(v) \gamma(X(s-), u) \xi(dv \times du \times ds),$$

Uniqueness implies $X(r)$ is independent of $W(\cdot + r) - W(r)$ and $\xi(A \times (r, \cdot))$ and that $X(t), t \geq r$ is determined by $X(r), W(\cdot + r) - W(r)$ and $\xi(A \times (r, \cdot))$, which gives the Markov property.
8. Filtering

- Bayes and Kallianpur-Striebel formula
- “Particle representations” of conditional distributions
- Continuous time filtering in Gaussian white noise
- Derivation of filtering equations
- Zakai equation
- Kushner-Stratonovich equation
- Point process observations
- Spatial observations with additive white noise
- Uniqueness
Bayes and Kallianpur Striebel formulas

Let $X$ and $Y$ be random variables defined on $(\Omega, \mathcal{F}, P)$. Suppose we want to calculate the conditional expectation

$$E^P[f(X)|Y].$$

If $P << Q$ with $dP = LdQ$, Bayes formula says

$$E^P[f(X)|Y] = \frac{E^Q[f(X)L|Y]}{E^Q[L|Y]}.$$

If $X = h(U, Y)$, $L = L(U, Y)$ and $U$ and $Y$ are independent under $Q$, then

$$E^P[f(X)|Y] = \frac{\int f(h(u, Y))L(u, Y)\mu_U(du)}{\int L(u, Y)\mu_U(du)}$$

where $\mu_U$ is the distribution of $U$.

**Method:** To compute a conditional expectation, find a reference measure under which what we don’t know is independent of what we do know.
Monte Carlo integration

Let $U_1, U_2, \ldots$ be iid with distribution $\mu_U$. Then

$$E^P[f(X)|Y] = \lim_{n \to \infty} \frac{\sum_{i=1}^{n} f(h(U_i, Y))L(U_i, Y)}{\sum_{i=1}^{n} L(U_i, Y)}$$

Note that $(U_1, Y), (U_2, Y), \ldots$ is a stationary (in fact exchangeable) sequence. Consequently,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f(h(U_i, Y))L(U_i, Y) = E^Q[f(U_1, Y)L(U_i, Y)|I]$$

$$= E^Q[f(U_1, Y)L(U_i, Y)|Y]$$

$$= \int f(h(u, Y))L(u, Y)\mu_U(dx) \quad a.s. ~ Q$$
Continuous time filtering in Gaussian white noise

To understand the intuition behind the standard “observation in Gaussian white noise” filtering model, suppose $X$ is the signal of interest and noisy observations of the form

$$O_n\left(\frac{k}{n}\right) = h(X(\frac{k}{n})) \frac{1}{n} + \frac{1}{\sqrt{n}}\zeta_k$$

are made every $n^{-1}$ time units. For large $n$, the noise swamps the signal at any one time point.

Assume that the $\{\zeta_k\}$ are iid with mean zero and variance $\sigma^2$. Then $Y_n(t) = \sum_{k=1}^{[nt]} O_n(\frac{k}{n})$ is approximately

$$Y(t) = \int_0^t h(X(s))ds + \sigma W(t). \quad (8.1)$$

Note, however, that $E[f(X(t))|\mathcal{F}_{Y_n}^t]$ does not necessarily converge to $E[f(X(t))|\mathcal{F}_t^Y]$; however, we still take (8.1) as our observation model.
Reference measure

By the Girsanov formula,

\[ E^P[g(X(t))|\mathcal{F}_t^Y] = \frac{E^Q[g(X(t))L(t)|\mathcal{F}_t^Y]}{E^Q[L(t)|\mathcal{F}_t^Y]} \]

where under \( Q \), \( X \) and \( Y \) are independent, \( Y \) is a Brownian motion with mean zero and variance \( \sigma^2 t \), and

\[ L(t) = \exp\left\{ \int_0^t \frac{h(X(s))}{\sigma^2} dY(s) - \frac{1}{2} \int_0^t \frac{h^2(X(s))}{\sigma^2} ds \right\} \]

that is,

\[ L(t) = 1 + \int_0^t \frac{h(X(s))}{\sigma^2} L(s) dY(s). \]

Under \( Q \), \( \sigma^{-1}Y \) is a standard Brownian motion, so under \( P \),

\[ W(t) = \sigma^{-1}Y(t) - \int_0^t \sigma^{-1}h(X(s)) ds \]

is a standard Brownian motion.
Monte Carlo solution

Let $X_1, X_2, \ldots$ be iid copies of $X$ that are independent of $Y$ under $Q$, and let

$$L_i(t) = 1 + \int_0^t \frac{h(X_i(s))}{\sigma^2} L_i(s) dY(s).$$

Note that

$$\phi(g, t) \equiv E^Q[g(X(s))L(s)|\mathcal{F}_s^Y] = E^Q[g(X_i(s))L_i(s)|\mathcal{F}_s^Y]$$

Claim:

$$\frac{1}{n} \sum_{i=1}^n g(X_i(s))L_i(s) \to E^Q[g(X(s))L(s)|\mathcal{F}_s^Y]$$
Zakai equation

For simplicity, assume $\sigma = 1$ and $X$ is a diffusion

$$X(t) = X(0) + \int_0^t \sigma(X(s))dB(s) + \int_0^t b(X(s))ds,$$

where under $Q$, $B$ and $Y$ are independent standard Brownian motions. Since

$$g(X(t)) = g(X(0)) + \int_0^t g'(X(s))\sigma(X(s))dB(s) + \int_0^t Ag(X(s))ds$$

$$g(X(t))L(t) = g(X(0)) + \int_0^t L(s)dg(X(s)) + \int_0^t g(X(s))dL(s)$$

$$= g(X(0)) + \int_0^t L(s)g'(X(s))\sigma(X(s))dB(s)$$

$$+ \int_0^t L(s)Ag(X(s))ds + \int_0^t g(X(s))h(X(s))L(s)dY(s)$$
Monte Carlo derivation of Zakai equation

\[ X_i(t) = X_i(0) + \int_0^t \sigma(X_i(s))dB_i(s) + \int_0^t b(X_i(s))ds, \]

where \((X_i(0), B_i)\) are iid copies of \((X(0), B)\).

\[ g(X_i(t))L_i(t) = g(X_i(0)) + \int_0^t L_i(s)g'(X_i(s))\sigma(X_i(s))dB_i(s) \]

\[ + \int_0^t L_i(s)Ag(X_i(s))ds \]

\[ + \int_0^t g(X_i(s))h(X_i(s))L_i(s)dY(s) \]

and hence

\[ \phi(g, t) = \phi(g, 0) + \int_0^t \phi(Ag, s)ds + \int_0^t \phi(gh, s)dY(s) \]
Kushner-Stratonovich equation

\[
\pi_t g = E^P[g(X(t)) | \mathcal{F}_t^Y] = \frac{\phi(g, t)}{\phi(1, t)} \\
= \frac{\phi(g, 0)}{\phi(1, 0)} + \int_0^t \frac{1}{\phi(1, s)} d\phi(g, s) - \int_0^t \frac{\phi(g, s)}{\phi(1, s)^2} d\phi(1, s) \\
+ \int_0^t \frac{\phi(g, s)}{\phi(1, s)^3} d[\phi(1, \cdot)]_s - \int_0^t \frac{1}{\phi(1, s)^2} d[\phi(g, \cdot), \phi(1, \cdot)]_s \\
= \pi_0 g + \int_0^t \pi_s Ag ds + \int_0^t (\pi_s gh - \pi_s g \pi_s h) dY(s) \\
+ \int_0^t \sigma^2 \pi_s g \pi_s h^2 ds - \int_0^t \sigma^2 \pi_s gh \pi_s h ds \\
= \pi_0 g + \int_0^t \pi_s Ag ds + \int_0^t (\pi_s gh - \pi_s g \pi_s h)(dY(s) - \pi_s h ds)
Point process observations

Model: $X$ is adapted to $\{F_t\}$, and

$$Y(t, \Gamma) - \int_0^t \int_{\Gamma} \lambda(X(s), u)\nu(du)ds$$

is an $\{F_t\}$-martingale for all $\Gamma \in \mathcal{B}(U)$ with $\nu(\Gamma) < \infty$.

$$\sup_x \int_U |\lambda(x, u) - 1| \wedge |\lambda(x, u) - 1|^2 \nu(du) < \infty.$$ 

Reference measure construction: Under $Q$, $X$ is the diffusion given by (8.2) and $Y$ is an independent, Poisson random measure with mean measure $\nu$. The change of measure is given by (7.4):

$$L(t) = 1 + \int_0^t L(s-)dM_\lambda(s) = 1 + \int_{U \times [0,t]} (\lambda(X(s), u) - 1)L(s-)\tilde{Y}(du \times ds).$$
Zakai equation

\[
g(X(t))L(t) = g(X(0)) + \int_0^t g(X(s))dL(s) + \int_0^t L(s)dg(X(s)) \\
= g(X(0)) + \int_0^t L(s)g'(X(s))\sigma(X(s))dB(s) \\
+ \int_0^t L(s)Ag(X(s))ds \\
+ \int_{U \times [0,t]} g(X(s))(\lambda(X(s),u) - 1)L(s-)\tilde{Y}(du \times ds)
\]

Assume that \(\nu(U) < \infty\). Setting \(\bar{\lambda}(x) = \int_U \lambda(x,u)\nu(du)\) and \(Cg(x) = (\bar{\lambda}(x) - 1)g(x)\), the unnormalized conditional distribution \(\phi(g,t) = E^{Q}[g(X(t))L(t)|\mathcal{F}_t^Y]\) satisfies

\[
\phi(g,t) = \phi(g,0) + \int_0^t \phi((A-C)g, s)ds + \int_0^t \phi((\lambda(\cdot,u)-1)g, s-)Y(du \times ds).
\]
Solution of the Zakai equation

Let \( \{T(t)\} \) be the semigroup given by

\[
T(t)f(x) = E[f(X(t))e^{-\int_0^t(\lambda(X(s))-1)ds} | X(0) = x]
\]

Suppose \( Y = \sum_k \delta_{(\tau_k, u_k)} \), and \( 0 < \tau_1 < \cdots < \tau_m < t < \tau_{m+1} \). Then

\[
\begin{align*}
\phi(g, t) &= \phi(T(t - \tau_m)g, \tau_m) \\
\phi(g, \tau_{k+1}-) &= \phi(T(\tau_{k+1} - \tau_k)g, \tau_k) \\
\phi(g, \tau_{k+1}) &= \phi(\lambda(\cdot, u_{k+1})g, \tau_{k+1}-) = \phi(T(\tau_{k+1} - \tau_k)\lambda(\cdot, u_{k+1})g, \tau_k).
\end{align*}
\]

If \( \phi(dx, t) = \phi(x, t)dx \), then

\[
\begin{align*}
\phi(\cdot, t) &= T^*(t - \tau_m)\phi(\cdot, \tau_m) \\
\phi(\cdot, \tau_{k+1}-) &= T^*(\tau_{k+1} - \tau_k)\phi(\cdot, \tau_k) \\
\phi(\cdot, \tau_{k+1}) &= \lambda(\cdot, u_{k+1})\phi(\cdot, \tau_{k+1}-) = \lambda(\cdot, u_{k+1})T^*(\tau_{k+1} - \tau_k)\phi(\cdot, \tau_k).
\end{align*}
\]
Counting process observations

Model: $X$ is a diffusion

$$X(t) = X(0) + \int_0^t \sigma(X(s)) dB(s) + \int_0^t b(X(s)) ds \quad (8.2)$$

and

$$Y(t) = V\left(\int_0^t \lambda(X(s)) ds\right),$$

where $V$ is unit Poisson process independent of $B$.

Reference measure construction: Under $Q$, $X$ is the diffusion given by (8.2) and $Y$ is an independent, unit Poisson process. The change of measure is given by (5.3):

$$L(t) = 1 + \int_0^t (\lambda(X(s)) - 1)L(s-)d(Y(s) - s)$$
Zakai equation

\[ g(X(t))L(t) = g(X(0)) + \int_0^t g(X(s))dL(s) + \int_0^t L(s)dg(X(s)) \]

\[ = g(X(0)) + \int_0^t L(s)g'(X(s))\sigma(X(s))dB(s) \]

\[ + \int_0^t L(s)Ag(X(s))ds \]

\[ + \int_0^t g(X(s))(\lambda(X(s)) - 1)L(s-\cdot)d(Y(s) - s) \]

The unnormalized conditional distribution \( \phi(g, t) = E_{F_t}^Q[g(X(t))L(t)|F^Y_t] \) satisfies

\[ \phi(g, t) = \phi(g, 0) + \int_0^t \phi((A - C)g, s)ds + \int_0^t \phi(Cg, s-\cdot)dY(s), \]

where \( Cg(x) = (\lambda(x) - 1)g(x). \)
Kushner-Stratonovich equation

For $\pi_t g \equiv \frac{\phi(g,t)}{\phi(1,t)}$, 

$$
\pi_t g = \pi_0 g + \int_0^t (\pi_s Ag - \pi_s \lambda g + \pi_s \lambda \pi_s g) ds
+ \int_0^t \left( \frac{\pi_s - \lambda g - \pi_s - \lambda \pi_s g}{\pi_s - \lambda} \right) dY(s)
= \pi_0 g + \int_0^t \pi_s Ag ds
+ \int_0^t \left( \frac{\pi_s - \lambda g - \pi_s - \lambda \pi_s g}{\pi_s - \lambda} \right) (dY(s) - \pi_s \lambda ds)
$$
Solution of the Zakai equation

Let \( \{T(t)\} \) be the semigroup given by

\[
T(t)f(x) = E[f(X(t))e^{-\int_0^t(\lambda(X(s))-1)ds}|X(0) = x]
\]

Suppose the jump times of \( Y \) satisfy \( 0 < \tau_1 < \cdots < \tau_m < t < \tau_{m+1} \). Then

\[
\phi(g, t) = \phi(T(t - \tau_m)g, \tau_m)
\]

\[
\phi(g, \tau_{k+1}-) = \phi(T(\tau_{k+1} - \tau_k)g, \tau_k)
\]

\[
\phi(g, \tau_{k+1}) = \phi(\lambda g, \tau_{k+1}-) = \phi(T(\tau_{k+1} - \tau_k)\lambda g, \tau_k).
\]

If \( \phi(dx, t) = \phi(x, t)dx \), then

\[
\phi(\cdot, t) = T^*(t - \tau_m)\phi(\cdot, \tau_m)
\]

\[
\phi(\cdot, \tau_{k+1}-) = T^*(\tau_{k+1} - \tau_k)\phi(\cdot, \tau_k)
\]

\[
\phi(\cdot, \tau_{k+1}) = \lambda(\cdot)\phi(\cdot, \tau_{k+1}-) = \lambda(\cdot)T^*(\tau_{k+1} - \tau_k)\phi(\cdot, \tau_k).
\]
Spatial observations with additive white noise

Signal:

\[ X(t) = X(0) + \int_0^t \sigma(X(s))dB(s) + \int_0^t b(X(s))ds + \int_{S_0 \times [0,t]} \alpha(X(s), u)W(du \times ds) \]

\[ = X(0) + \int_0^t \sigma(X(s))dB(s) + \int_{S_0 \times [0,t]} \alpha(X(s), u)Y(du \times ds) \]

\[ + \int_0^t (b(X(s)) - \int_{S_0} \alpha(X(s), u)h(X(s), u)\mu_0(du))ds \]

Observation:

\[ Y(A, t) = \int_0^t \int_A h(X(s), u)\mu_0(du)ds + W(A, t) \]

Under \( Q \), \( Y \) is Gaussian white noise on \( S_0 \times [0, \infty) \) with

\[ E[Y(A, t)Y(B, s)] = \mu_0(A \cap B)t \wedge s, \]

and \( dP|_{\mathcal{F}_t} = L(t)dQ|_{\mathcal{F}_t} \) where

\[ L(t) = 1 + \int_{S_0 \times [0,t]} L(s)h(X(s), u)Y(du \times ds) \]
Apply Itô’s formula

Under $P$, $X$ is a diffusion with generator

$$Af(x) = \frac{1}{2} \sum a_{ij}(x) \partial_i \partial_j f(x) + \sum b_i(x) \partial_i f(x)$$

where

$$a(x) = \sigma(x)\sigma(x)^T + \int_{S_0} \alpha(x,u)\alpha(x,u)^T \mu_0(du)$$

Then

$$f(X(t))L(t) = f(X(0)) + \int_0^t L(s)\nabla f(X(s))^T \sigma(X(s)) dB(s)$$

$$\quad + \int_{S_0 \times [0,t]} L(s)(\nabla f(X(s)) \cdot \alpha(X(s),u) + f(X(s))h(X(s),u))Y(du \times ds)$$

$$\quad + \int_0^t L(s)Af(X(s))ds$$
Particle representation

$B_i$ independent, standard Brownian motions, independent of $Y$ on $(\Omega, \mathcal{F}, P_0)$. Let

$$X_i(t) = X_i(0) + \int_0^t \sigma(X_i(s))dB_i(s) + \int_{S_0 \times [0,t]} \alpha(X_i(s-), u)Y(du \times ds)$$

$$+ \int_0^t \int_{S_0} (b(X_i(s)) - \alpha(X_i(s), u)h(X_i(s), u)) \mu_0(du)ds$$

$$L_i(t) = 1 + \int_{S_0 \times [0,t]} L_i(s)h(X_i(s), u)Y(du \times ds)$$

Then

$$\phi(f, t) = E^{P_0}[f(X(t))L(t)|\mathcal{F}_t^Y] = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n f(X_i(t))L_i(t)$$
Zakai equation

Since

\[ f(X_i(t))L_i(t) \]

\[ = f(X_i(0)) + \int_0^t L_i(s) \nabla f(X_i(s))^T \sigma(X_i(s)) dB_i(s) \]

\[ + \int_{S_0 \times [0,t]} L_i(s)(\nabla f(X_i(s)) \cdot \alpha(X_i(s), u) + f(X_i(s))h(X_i(s), u))Y(du \times ds) \]

\[ + \int_0^t L_i(s)Af(X_i(s)) ds, \]

\[ \phi(f, t) = \phi(f, 0) + \int_0^t \phi(Af, s) ds + \int_{S_0 \times [0,t]} \phi(\nabla f \cdot \alpha(\cdot, u) + fh(\cdot, u), s)Y(du \times ds) \]
Kushner-Stratonovich equation

It follows that

\[
\pi_t f = \frac{\phi(f, t)}{\phi(1, t)}
\]

\[
= \pi_0 f + \int_0^t \pi_s A f ds
\]

\[
+ \int_{S_0 \times [0,t]} \left( \pi_s (\nabla f \cdot \alpha(\cdot, u) + f h(\cdot, u)) \right) Y(du \times ds)
\]

\[
+ \int_0^t \int_{S_0} \left( \pi_s f \pi_s h(\cdot, u) - \pi_s (\nabla f \cdot \alpha(\cdot, u) + f h(\cdot, u)) \right) \pi_s h(\cdot, u) \mu_0(du) ds
\]

\[
= \pi_0 f + \int_0^t \pi_s A f ds
\]

\[
+ \int_{S_0 \times [0,t]} \left( \pi_s (\nabla f \cdot \alpha(\cdot, u) + f h(\cdot, u)) - \pi_s f \pi_s h(\cdot, u) \right) \tilde{Y}(du \times ds)
\]

where

\[
\tilde{Y}(A, t) = Y(A, t) - \int_0^t \int_A \pi_s h(\cdot, u) \mu_0(du) ds
\]
Uniqueness for Kushner-Stratonovich equation

Each of the Kushner-Stratonovich equations is of the form

$$\pi_t g = \pi_0 g + \int_0^t \pi_s A g ds + \int_0^t G(\pi_s_-) d(Y(s) - \pi_s h ds)$$

Each of the stochastic integrator would be a martingale if the $\pi$ was the conditional distribution. Suppose that is not true.

In each case, under some restrictions on $h$ and $\pi$, we can do a change of measure to make $Y(t) - \int_0^t \pi_s h ds$ a martingale. Under the “new” measure, $(Y, \pi)$ is a solution of the corresponding filtered martingale problem.

By uniqueness of the filtered martingale problem, $\pi_t = H(t, Y)$ for some appropriately measurable transformation. But this transformation doesn’t depend on the change of measure, so $\pi$ was already the conditional distribution process.
9. **Time change equations**

- Stochastic equations for Markov chains
- Diffusion limits??
- General multiple time-change equations
- Compatibility for multiple time-changes
- Martingale problem for the time-change equation
- Uniqueness question
Stochastic equations for Markov chains

Specify a continuous time Markov chain by specifying the intensities of its possible jumps

\[ P\{X(t + \Delta t) = X(t) + \zeta_k | \mathcal{F}_t^X\} \approx \beta_k(X(t)) \Delta t \]

Given the intensities, the Markov chain satisfies

\[ X(t) = X(0) + \sum_k Y_k(\int_0^t \beta_k(X(s)) ds) \zeta_k \]

where the \( Y_k \) are independent unit Poisson processes.

(Assume that there are only finitely many \( \zeta_k \).)
Diffusion limits??

Possible scaling limits of the form

\[
X_n(t) = X_n(0) + \frac{1}{n} \sum_k \tilde{Y}_k(n^2 \int_0^t \beta_n^k(X_n(s))ds)\zeta_k + \int_0^t F^n(X_n(s))ds
\]

where \(\tilde{Y}_k(u) = Y_k(u) - u\) and \(F^n(x) = \sum_k n \beta^n_k(x)\zeta_k\).

Note that \(\frac{1}{n} \tilde{Y}_k(n^2\cdot) \approx W_k\).

Assuming \(X_n(0) \to X(0)\), \(\beta^n_k \to \beta_k\) and \(F^n \to F\), we might expect a limit satisfying

\[
X(t) = X(0) + \sum_k W_k(\int_0^t \beta_k(X(s))ds)\zeta_k + \int_0^t F(X(s))ds
\]

\textit{Kurtz (1980a), Ethier and Kurtz (1986), Section 6.5.}
General multiple time change equations

Let $Y_1, \ldots, Y_m$ be independent, Markov processes, $Y_k$ with generator $A_k \subset C_b(E) \times C_b(E)$. Consider the system

$$X_k(t) = Y_k(\int_0^t \beta_k(X(s)) \, ds),$$

where $X = (X_1, \ldots, X_m)$ and the $\beta_k$ are continuous.
Compatibility

**Definition 9.1** A process $X$ in $D_{E_1}[0, \infty)$ is temporally compatible with $Y$ in $D_{E_2}[0, \infty)$, if for each $t \geq 0$ and $h \in B(D_{E_2}[0, \infty))$,

$$E[h(Y)|\mathcal{F}^{X,Y}_t] = E[h(Y)|\mathcal{F}^Y_t] \quad a.s. \quad (9.1)$$

(cf. (4.5) of Jacod (1980).)

If $Y$ has independent increments, then $X$ is (temporally) compatible with $Y$ if $Y(t + \cdot) - Y(t)$ is independent of $\mathcal{F}^{X,Y}_t$ for all $t \geq 0$.

If $X$ is $\{\mathcal{F}^Y_t\}$-adapted, then $X$ is compatible with $Y$. 
Compatibility for time-change equations

Let \( Y_1, \ldots, Y_m \) be independent, cadlag Markov processes, \( Y_k \) with state space \( E_k \) and generator \( A_k \). Suppose \( X = (X_1, \ldots, X_m) \) and \( E = E_1 \times \cdots \times E_m \) and \( X \) satisfies

\[
X_k(t) = Y_k\left( \int_0^t \beta_k(X(s))ds \right).
\]

Set \( \tau_k(t) = \int_0^t \beta_k(X(s))ds \), and for \( \alpha \in [0, \infty)^m \), define

\[
\mathcal{F}_\alpha^Y = \sigma(Y_k(s_k) : s_k \leq \alpha_k, k = 1, 2, \ldots) \vee \sigma(X(0))
\]

and

\[
\mathcal{F}_\alpha^\tau = \sigma(\{\tau_1(t) \leq s_1, \tau_2(t) \leq s_2, \ldots\} : s_i \leq \alpha_i, i = 1, 2, \ldots, t \geq 0).
\]

Then \( X \) is compatible with \( Y \) if for each \( h \in B(D_E[0, \infty)) \),

\[
E[h(Y)|\mathcal{F}_\alpha^\tau \vee \mathcal{F}_\alpha^Y] = E[h(Y)|\mathcal{F}_\alpha^Y]. \tag{9.2}
\]
Existence of compatible solutions

Construct a compatible approximation

\[ X^n_k(t) = Y_k \left( \int_0^t \beta_k(X\left(\frac{[ns]}{n}\right))ds \right). \]

(Essentially an Euler approximation.) Then

\[ \tau^n_k(t) = \int_0^t \beta_k(X^n\left(\frac{[ns]}{n}\right))ds \]

and

\[ \mathcal{F}_\alpha^\tau = \sigma(\{\tau^n_1(t) \leq s_1, \tau^n_2(t) \leq s_2, \ldots : s_i \leq \alpha_i, i = 1, 2, \ldots, t \geq 0\}) \subset \mathcal{F}^Y_\alpha. \]

Let \( C^1_\alpha = \{g(\cdot \land \alpha_1, \ldots, \cdot \land \alpha_m) : g \in C_\mathbb{R}([0, \infty))^m \} \) and \( C^2_\alpha = \{g(y_1(\cdot \land \alpha_1), \ldots, y_m(\cdot \land \alpha_m)) : g \in C(\prod_{k=1}^m D_{E_k}[0, \infty)) \} \).

**Lemma 9.2** Suppose \( A_k \subset C_b(E_k) \times C_b(E_k) \) and \( Y_k \) has no fixed points of discontinuity and the \( \beta_k \) are bounded and continuous. Then \( \{X^n\} \) is relatively compact. Pick a subsequence along which \( (X^n, Y) \Rightarrow (X, Y) \). Then setting \( \tau_k(t) = \int_0^t \beta_k(X(s))ds \), \( \tau \) is compatible with \( Y \).
Proof. For $f \in C_b(\prod D_{E_k}[0, \infty))$, $g_1 \in C^1_{\alpha}$ and $g_2 \in C^2_{\alpha}$

$$E[f(Y)g_1(\tau^n)g_2(Y)] = E[E[f(Y)|F^Y_{\alpha}]g_1(\tau^n)g_2(Y)]$$

For each $\alpha$ and $\epsilon > 0$ there exists $f_{\alpha,\epsilon} \in C_b(\prod D_{E_k}[0, \infty))$ such that

$$E[|E[f(Y)|F^Y_{\alpha}] - f_{\alpha,\epsilon}(Y)|] \leq \epsilon.$$

Consequently, it follows that

$$\lim_{n \to \infty} E[f(Y)g_1(\tau^n)g_2(Y)] = E[f(Y)g_1(\tau)g_2(Y)]$$

$$= \lim_{n \to \infty} E[E[f(Y)|F^Y_{\alpha}]g_1(\tau^n)g_2(Y)]$$

$$= \lim_{\epsilon \to 0} \lim_{n \to \infty} E[f_{\alpha,\epsilon}(Y)g_1(\tau^n)g_2(Y)]$$

$$= E[E[f(Y)|F^Y_{\alpha}]g_1(\tau)g_2(Y)]$$

verifying compatibility. \qed
Martingales

Assume \((1, 0) \in A_k, k = 1, \ldots, m\). If \(f_k \in \mathcal{D}(A_k)\) and \(\inf_{x \in E_k} f_k(x) > 0\), then

\[
M_f(\alpha) = \prod_{k=1}^{n} f_k(Y_k(\alpha_k)) \exp\left\{ - \int_{0}^{\alpha_k} \frac{A_k f_k(Y_k(s))}{f_k(Y_k(s))} ds \right\}
\]

is a \(\{\mathcal{F}_x^Y\}\)-martingale. If the solution \(X\) is compatible with \(Y\), then \(M_f\) is a \(\{\mathcal{G}_\alpha\} \equiv \{\mathcal{F}_x^Y \vee \mathcal{F}_x^X\}\)-martingale.

By the definition of \(\mathcal{F}_x^X\), \(\tau(t) = (\tau_1(t), \ldots, \tau_m(t))\) is a \(\{\mathcal{F}_x^Y \vee \mathcal{F}_x^X\}\)-stopping time. For simplicity, assume that the \(\beta_k\) are bounded. Then, by the optional sampling theorem

\[
M_f(\tau(t)) = \prod_{k=1}^{n} f_k(Y_k(\tau_k(t))) \exp\left\{ - \int_{0}^{\tau_k(t)} \frac{A_k f_k(Y_k(s))}{f_k(Y_k(s))} ds \right\}
\]

\[
= \prod_{k=1}^{n} f_k(X_k(t)) \exp\left\{ - \int_{0}^{t} \frac{\beta_k(X(s)) A_k f_k(X_k(s))}{f_k(X_k(s))} ds \right\}
\]

is a \(\{\mathcal{G}_{\tau(t)}\}\)-martingale.
Martingale problem for the solution

Defining $f(x) = \prod f_k(x_k)$ and

$$Bf(x) = \sum_{k=1}^{m} \beta_k(x) \prod_{l \neq k} f_l(x_l) A_k f_k(x_k),$$

we have

$$M_f(\tau(t)) = f(X(t)) \exp\left\{ - \int_0^t \frac{B f(X(s))}{f(X(s))} ds \right\},$$

and it follows that $X$ is a solution of the martingale problem for $B$. 
Uniqueness question

\[ X(t) = X(0) + \sum_{k} W_k(\int_{0}^{t} \beta_k(X(s))ds)\zeta_k + \int_{0}^{t} F(X(s))ds \]

Let \( \tau_k(t) = \int_{0}^{t} \beta_k(X(s))ds \) and \( \gamma(t) = \int_{0}^{t} F(X(s))ds \). Then

\[
\dot{\tau}_l(t) = \beta_l(X(0) + \sum_{k} W_k(\tau_k(t))\zeta_k + \gamma(t))
\]
\[
\dot{\gamma}(t) = F(X(0) + \sum_{k} W_k(\tau_k(t))\zeta_k + \gamma(t))
\]

**Problem:** Find conditions under which pathwise uniqueness holds.
Compatibility for multiple time-changes

\[ X(t) = X(0) + \sum_{k=1}^{m} W_k(\int_0^t \beta_k(X(s))ds)\zeta_k + \int_0^t F(X(s))ds \]

Set \( \tau_k(t) = \int_0^t \beta_k(X(s))ds \), and for \( \alpha \in [0, \infty)^m \), define

\[ \mathcal{F}_\alpha^W = \sigma(W_k(s_k) : s_k \leq \alpha_k, k = 1, 2, \ldots) \lor \sigma(X(0)) \]

and

\[ \mathcal{F}_\alpha^\tau = \sigma(\{\tau_1(t) \leq s_1, \tau_2(t) \leq s_2, \ldots\} : s_i \leq \alpha_i, i = 1, 2, \ldots, t \geq 0). \]

If \( \tau \) is compatible with \( W \), then \( W_k(\int_0^t \beta_k(X(s))ds), k = 1, \ldots, m \) are martingales with respect to the same filtration and hence \( X \) is a solution of the martingale problems for

\[ Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \partial_i \partial_j f(x) + F(x) \cdot \nabla f(x), \]

\[ a(x) = \sum_k \beta_k(x)\zeta_k \zeta_k^T. \]
Two-dimensional case

For $i = 1, 2$, $W_i$ standard Brownian motions.

$\beta_i : \mathbb{R}^2 \to (0, \infty)$, bounded

$$X_1(t) = W_1\left( \int_0^t \beta_1(X(s)) ds \right) \quad X_2(t) = W_2\left( \int_0^t \beta_2(X(s)) ds \right)$$

or equivalently

$$\dot{\tau}_i(t) = \beta_i(W_1(\tau_1(t)), W_2(\tau_2(t))), \quad i = 1, 2$$

A strong, compatible solution exists, and weak uniqueness holds by Stroock-Varadhan, so pathwise uniqueness holds.
10. **Bits and pieces**

- Equivalence of martingale problems and SDEs
  
  Kurtz (2011)

- Cluster detection

- Signal in noise with point process observations

- Filtering equations

- Computable cases

- Analysis of earthquake data

  Wu (2009)
Equivalence of SDE and MGP for diffusions

\[ X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds, \quad (10.1) \]

By Itô’s formula

\[ f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds = \int_0^t \nabla f(X(s))^T \sigma(X(s))dW(s) \]

for

\[ Af(x) = \frac{1}{2} \sum_{ij} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} f(x) \]

where \(((a_{ij})) = \sigma\sigma^T\).

If \(X\) is a solution of the MGP for \(A\) and \(\sigma\) is invertible, then

\[ W(t) = \int_0^t \sigma^{-1}(X(s))dX(s) - \int_0^t \sigma^{-1}(X(s))b(X(s))ds \]

is a standard Brownian motion and (10.1) is satisfied.
A stochastic equation and a martingale problem

Assume $W$ is one dimensional, $\lambda(x, u) \leq 1$, and $\xi$ a Poisson random measure on $[0, 1]^2 \times [0, \infty)$ with Lebesgue mean measure. Suppose $X$ satisfies

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW + \int_0^t b(X(s))ds$$

$$+ \int_1^1 \mathbf{1}_{[0, \lambda(X(s-), u)]}(v)\gamma(X(s-), u)\xi(du \times dv \times ds).$$

Then setting $a(x) = \sigma(x)^2$, $X$ will be a solution of the martingale problem for

$$Af(x) = \frac{1}{2}a(x)f''(x) + b(x)f'(x) + \int_0^1 \lambda(x, u)(f(x + \gamma(x, u)) - f(x))du.$$
Augmented process

Let $\xi$ be given by

$$\xi(C \times [0, t]) = \sum_{i=0}^{N(t)-1} 1_C(V_i), \quad C \in \mathcal{B}([0, 1]^2),$$

where $N$ is a unit Poisson process and $V_0, V_1, \ldots$ are iid and uniformly distributed over $[0, 1]^2$.

Define

$$Y(t) = Y(0) + W(t) \mod 1$$

$$Z(t) = V_{N(t)}$$
Martingale problem for \((X, Y, Z)\)

Then \((X(t), Y(t), Z(t))\) is a solution to the martingale problem for the operator \(B\) defined as follows: \(D(B)\) consists of functions of the form \(F(x, y, z) = f(x)g(y)h(z)\), \(f \in C^2_c(\mathbb{R})\), \(g \in C^2[0, 1]\) with \(g(0) = g(1), g'(0) = g'(1), g''(0) = g''(1)\), and \(h \in C([0, 1]^2)\). Then

\[
BF(x, y, z) = g(y)h(z) \left( \frac{1}{2}a(x)f''(x) + b(x)f'(x) \right) + \sigma(x)h(z)f'(x)g'(y) + \frac{1}{2}f(x)h(z)g''(y) + g(y) \left( \int_{[0,1]^2} h(u)\nu(du)f(x + 1_{[0,\lambda(x,z_1)]}(z_2)\gamma(x, z_1)) - h(z)f(x) \right)
\]

where \(\nu\) is the uniform distribution over \([0, 1]^2\).
Application of Markov mapping theorem

Let $\alpha(x, dy \times dz) = dy \nu(dz)$. Setting $\overline{g} = \int_0^1 g(y)dy$ and $\overline{h} = \int_{[0,1]^2} h(z)\nu(dz)$,

$$\alpha_{BF}(x) = \overline{g}\overline{h}Af(x).$$

By the martingale mapping theorem, every solution of the martingale problem for $A$ comes from a solution of the martingale problem for $B$ and $W$ can be recovered from $Y$ and $\xi$ can be recovered from $Z$. 
Cluster detection: Possible applications

- Internet packets that form a malicious attack on a computer system.
- Financial transactions that form a collusive trading scheme.
- Earthquakes that form a single seismic event.
The model

The observations form a marked point process $O$ with marks in $E$.

$$O(A, t) = N(A, t) + C(A, t)$$

with

$$N(A, t) = \int_{A \times [0, \infty) \times [0, t]} 1_{[0, \gamma(u)]}(v)\xi_1(du \times dv \times ds)$$

$$C(A, t) = \int_{A \times [0, \infty) \times [0, t]} 1_{[0, \lambda(u, \eta_{s-})]}(v)\xi_2(du \times dv \times ds)$$

where $\xi_1$ and $\xi_2$ are independent Poisson random measures on $E \times [0, \infty) \times [0, \infty)$ with mean measure $\nu \times \ell \times \ell$, $\ell$ denoting Lebesgue measure.

$$\eta_t(A \times [0, r]) = \int_{A \times [0, t]} 1_A(u)1_{[0, r]}(s)C(du \times ds)$$
Radon-Nikodym derivative

**Lemma 10.1** On $(\Omega, \mathcal{F}, Q)$, let $N$ and $C$ be independent Poisson random measures with mean measures $\nu_0(du \times ds) = \gamma(u)\nu(du)ds$ and $\nu_1(du \times ds) = \lambda(u)\nu(du)ds$ respectively that are compatible with $\{\mathcal{F}_t\}$. Let $L$ satisfy

$$L(t) = 1 + \int_{E \times [0,t]} \left( \frac{\lambda(u, \eta_{s-})}{\lambda(u)} - 1 \right) L(s-) (C(du \times ds) - \lambda(u)\nu(du)ds).$$

(10.2)

and assume that $L$ is a $\{\mathcal{F}_t\}$-martingale.

Define $dP|_{\mathcal{F}_t} = L(t)dQ|_{\mathcal{F}_t}$. Under $P$, for all $A$ such that $\int_0^t \int_A \lambda(u, \eta_s)\nu(du)ds < \infty$, $t > 0$,

$$C(A, t) - \int_{A \times [0,t]} \lambda(u, \eta_s)\nu(du)ds$$

is a local martingale and $N$ is independent of $C$ and is a Poisson random measure with mean measure $\nu_0$.  

The general filtering equations

**Theorem 10.2**

\[ \phi(f, t) = \phi(f, 0) - \int_{E \times [0, t]} \phi(f(\cdot))(\lambda(u, \cdot) - \lambda(u)), s)\nu(du)ds \]

\[ + \int_{E \times [0, t]} \phi(f(\cdot + \delta(u,s))\frac{\lambda(u, \cdot)}{\lambda(u)} - f(\cdot), s-)\frac{\lambda(u)}{\lambda(u) + \gamma(u)}O(du \times ds) \]

and

\[ \pi_t f = \pi_0 f \]

\[ + \int_{E \times [0, t]} \frac{\pi_{s-}(f(\cdot + \delta(u,s))\lambda(u, \cdot)) - \pi_{s-}\lambda(u, \cdot)\pi_{s-}f}{\pi_{s-}\lambda(u, \cdot) + \gamma(u)}O(du \times ds) \]

\[ - \int_{E \times [0, t]} (\pi_s(f(\cdot)\lambda(u, \cdot)) - \pi_s f \pi_s \lambda(u, \cdot))\nu(du)ds \]
Simplify

Problem: The difficulty of computing the distribution: $2^{O(E, t)}$ possible states.

Need to compromise: compute $\pi_t f = E^P[f(\eta_s)|F_s]$ for a “small” collection of $f$

Suppose one observes $u_i$ at time $\tau_i$ and $y_i = (u_i, \tau_i)$.

$\theta(y_i)(\cdot) = 1_{\{y_i \text{ is a point in the cluster}\}}$

$\theta_0(y_i)(\cdot) = 1_{\{y_i \text{ is the latest point in the cluster}\}}$

Need to be able to evaluate

$$\pi_t \lambda(u, \cdot)$$
A Markov scenario

Consider

\[
\lambda(u, \eta_t) = \sum_{i=1}^{O(E,t)} \lambda(u, y_i) \theta_0(y_i) + \epsilon(u),
\]

where \( \theta_0(y_i) = 1\{y_i \text{ is the latest point in the cluster}\} \).

Get a closed system for \( \pi_t \theta_0(y_i) \).

Let \( \theta(y)(\cdot) = 1\{y \text{ is a point in the cluster}\} \).

Get a closed system for

\[
\pi_t \theta_0(y_i), \quad \pi_t \theta(y_i), \quad \pi_t \theta(y_i) \theta_0(y_j)
\]
Earthquake declustering

- Earthquakes often occur in clusters.
- A large number of earthquakes strike without any foreshocks or aftershocks.
- Dataset: the earthquakes in the period of 1926-1995, in the rectangular area $34^\circ - 39^\circ$N and $131^\circ - 140^\circ$E, with magnitudes greater than 4.0 and depths less than 100 km.

Model for seismic events producing clusters

- $D = 1$: a cluster is active; $D = 0$: no active cluster.

- Cluster terminates at each quake in the cluster with probability $p$.

- New cluster is initiated with probability $\epsilon$. (Only one cluster is active at a time.)

- $\lambda(u, D_t, \eta_t) = 1_{\{D_t=1\}} \sum_{i=1}^{k} \lambda(u, y_i)\theta_m(y_i) + 1_{\{D_t=0\}}\epsilon$.

- $\nu$ is assumed to be the uniform measure on the rectangular region.

- $\lambda(u, y_i)$ is proportional to a 2-D normal distribution:
  \[\lambda(u, y_i) = \lambda \exp\left(-\|u - u_i\|^2 / 2d\right) / (2\pi d)\]

- $\gamma(du) = \gamma du$. 
Figure 1: time-space plot of all earthquakes (3771)
Figure 2: 1500 clustered earthquakes: (a) under ETAS model; (b) under the “mother quake” model
Basics of stochastic processes

- Filtrations
- Conditional expectations
- Stopping times
- Martingales
- Optional sampling theorem
- Doob’s inequalities
- Stochastic integrals
- Local martingales
- Semimartingales
- Computing quadratic variations
- Covariation
• Itô’s formula
Conventions and caveats

In these lectures, state spaces will always be complete, separable metric spaces (sometimes called *Polish spaces*), usually denoted \((E, r)\).

All probability spaces are complete.

All identities involving conditional expectations (or conditional probabilities) only hold almost surely (even when I don’t say so).

If the filtration \(\{\mathcal{F}_t\}\) involved is obvious, I will say adapted, rather than \(\{\mathcal{F}_t\}\)-adapted, stopping time, rather than \(\{\mathcal{F}_t\}\)-stopping time, etc.

All processes are *cadlag* (right continuous with left limits at each \(t > 0\)), unless otherwise noted.

A process is real-valued if that is the only way the formula makes sense.
Assumed background

A general knowledge of stochastic processes and martingales in continuous time and of stochastic integration will be assumed. This background is summarized in the next several slides. I will be more than happy to answer questions on any of this material. More detail can be found in the following:

The first five chapters of the notes *Lectures on Stochastic Analysis* which has the advantage of being free.

http://www.math.wisc.edu/~kurtz/m735.htm


The first two chapters of Protter, *Stochastic Integration and Differential Equations*, Second Edition
Filtrations

$(\Omega, \mathcal{F}, P)$ a probability space

Available information is modeled by a sub-$\sigma$-algebra of $\mathcal{F}$

$\mathcal{F}_t$ is the $\sigma$-algebra representing the information available at time $t$

$\{\mathcal{F}_t\}$ is a filtration, that is, $t < s$ implies $\mathcal{F}_t \subset \mathcal{F}_s$

$\{\mathcal{F}_t\}$ is complete if $\mathcal{F}_0$ contains all subsets of sets of probability zero.

A stochastic process $X$ is adapted to $\{\mathcal{F}_t\}$ if $X(t)$ is $\mathcal{F}_t$-measurable for each $t \geq 0$.

An $E$-valued stochastic process $X$ adapted to $\{\mathcal{F}_t\}$ is $\{\mathcal{F}_t\}$-Markov if

$$E[f(X(t + r))|\mathcal{F}_t] = E[f(X(t + r))|X(t)], \quad t, r \geq 0, \ f \in B(E)$$
Conditional expectations

Conditional expectation $E[X|\mathcal{D}]$ is the best estimate of a random variable $X$ using the information in $\mathcal{D}$. If the information in $\mathcal{D}$ is the information obtained by observing $Y$ and $Z$ (denoted $\mathcal{D} = \sigma(Y, Z)$), then there is a function $f_X$ such that

$$E[X|\sigma(Y, Z)] = f_X(Y, Z).$$

If $E[X^2] < \infty$, then $f_X$ is the minimizer of $E[(X - f(Y, Z))^2]$,

$$E[(X - f_X(Y, Z))^2] = \inf_{f} E[(X - f(Y, Z))^2].$$

In general, if $E|X| < \infty$, then $E[X|\mathcal{D}]$ is the $\mathcal{D}$-measurable random variable $Z$ (a random variable whose value is determined by the information in $\mathcal{D}$) satisfying

$$E[X1_D] = E[Z1_D] \quad \forall D \in \mathcal{D}.$$
Measurability for stochastic processes

A stochastic process is an indexed family of random variables, but if the index set is \([0, \infty)\), then we may want to know more about \(X(t, \omega)\) than that it is a measurable function of \(\omega\) for each \(t\). For example, for a \(\mathbb{R}\)-valued process \(X\), when are

\[
\int_a^b X(s, \omega)\,ds \quad \text{and} \quad X(\tau(\omega), \omega)
\]

random variables?

\(X\) is measurable if \((t, \omega) \in [0, \infty) \times \Omega \rightarrow X(t, \omega) \in E\) is \(\mathcal{B}([0, \infty)) \times \mathcal{F}\)-measurable.

**Lemma B.1** If \(X\) is measurable and \(\int_a^b |X(s, \omega)|\,ds < \infty\), then \(\int_a^b X(s, \omega)\,ds\) is a random variable.

If, in addition, \(\tau\) is a nonnegative random variable, then \(X(\tau(\omega), \omega)\) is a random variable.
Proof. The first part is a standard result for measurable functions on a product space. Verify the result for $X(s, \omega) = 1_A(s)1_B(\omega)$, $A \in \mathcal{B}[0, \infty)$, $B \in \mathcal{F}$ and apply the Dynkin class theorem to extend the result to $1_C$, $C \in \mathcal{B}[0, \infty) \times \mathcal{F}$.

If $\tau$ is a nonnegative random variable, then $\omega \in \Omega \rightarrow (\tau(\omega), \omega) \in [0, \infty) \times \Omega$ is measurable. Consequently, $X(\tau(\omega), \omega)$ is the composition of two measurable functions.
Measurability continued

A stochastic process $X$ is $\{\mathcal{F}_t\}$-adapted if for all $t \geq 0$, $X(t)$ is $\mathcal{F}_t$-measurable.

If $X$ is measurable and adapted, the restriction of $X$ to $[0, t] \times \Omega$ is $\mathcal{B}[0, t] \times \mathcal{F}$-measurable, but it may not be $\mathcal{B}[0, t] \times \mathcal{F}_t$-measurable.

$X$ is progressive if for each $t \geq 0$, $(s, \omega) \in [0, t] \times \Omega \rightarrow X(s, \omega) \in E$ is $\mathcal{B}[0, t] \times \mathcal{F}_t$-measurable. (Note that every cadlag and adapted process is progressive.)

Let

$$\mathcal{W} = \{A \in \mathcal{B}[0, \infty) \times \mathcal{F} : A \cap [0, t] \times \Omega \in \mathcal{B}[0, t] \times \mathcal{F}_t, t \geq 0\}.$$ 

Then $\mathcal{W}$ is a $\sigma$-algebra and $X$ is progressive if and only if $(s, \omega) \rightarrow X(s, \omega)$ is $\mathcal{W}$-measurable.

Since pointwise limits of measurable functions are measurable, point-wise limits of progressive processes are progressive.
Stopping times

Let \( \{\mathcal{F}_t\} \) be a filtration. \( \tau \) is a \( \mathcal{F}_t \)-stopping time if and only if \( \{\tau \leq t\} \in \mathcal{F}_t \) for each \( t \geq 0 \).

If \( \tau \) is a stopping time, the information available at time \( \tau \) is represented by the \( \sigma \)-algebra \( \mathcal{F}_\tau \equiv \{ A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, t \geq 0 \} \).

The definition of \( \mathcal{F}_\tau \) may look strange, but it has all the right properties. For example,

If \( \tau_1 \) and \( \tau_2 \) are stopping times with \( \tau_1 \leq \tau_2 \), then \( \mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2} \).

If \( \tau_1 \) and \( \tau_2 \) are stopping times then \( \tau_1 \) and \( \tau_1 \wedge \tau_2 \) are \( \mathcal{F}_{\tau_1} \)-measurable.
A process observed at a stopping time

If $X$ is measurable and $\tau$ is a stopping time, then $X(\tau(\omega), \omega)$ is a random variable.

**Lemma B.2** If $\tau$ is a stopping time and $X$ is progressive, then $X(\tau)$ is $\mathcal{F}_\tau$-measurable.

**Proof.** $\omega \in \Omega \to (\tau(\omega) \wedge t, \omega) \in [0, t] \times \Omega$ is measurable as a mapping from $(\Omega, \mathcal{F}_t)$ to $([0, t] \times \Omega, \mathcal{B}[0, t] \times \mathcal{F}_t)$. Consequently,

$$\omega \to X(\tau(\omega) \wedge t, \omega)$$

is $\mathcal{F}_t$-measurable, and

$$\{X(\tau) \in A\} \cap \{\tau \leq t\} = \{X(\tau \wedge t) \in A\} \cap \{\tau \leq t\} \in \mathcal{F}_t.$$

□
Right continuous processes

Most of the processes you know are either continuous (e.g., Brownian motion) or right continuous (e.g., Poisson process).

**Lemma B.3** If $X$ is right continuous and adapted, then $X$ is progressive.

**Proof.** If $X$ is adapted, then

$$
(s, \omega) \in [0, t] \times \Omega \to Y_n(s, \omega) \equiv X \left( \frac{[ns] + 1}{n} \land t, \omega \right)
$$

$$
= \sum_{k} X \left( \frac{k + 1}{n} \land t, \omega \right) 1_{[\frac{k}{n}, \frac{k+1}{n})}(s)
$$

is $\mathcal{B}[0, t] \times \mathcal{F}_t$-measurable. By the right continuity of $X$, $Y_n(s, \omega) \to X(s, \omega)$ on $\mathcal{B}[0, t] \times \mathcal{F}_t$, so $(s, \omega) \in [0, t] \times \Omega \to X(s, \omega)$ is $\mathcal{B}[0, t] \times \mathcal{F}_t$-measurable and $X$ is progressive. \qed
Examples and properties

Define $\mathcal{F}_{t^+} \equiv \bigcap_{s > t} \mathcal{F}_s$. $\{\mathcal{F}_t\}$ is right continuous if $\mathcal{F}_t = \mathcal{F}_{t^+}$ for all $t \geq 0$. If $\{\mathcal{F}_t\}$ is right continuous, then $\tau$ is a stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all $t > 0$.

Let $X$ be cadlag and adapted. If $K \subset E$ is closed, $\tau^h_K = \inf\{t : X(t) \text{ or } X(t^-) \in K\}$ is a stopping time, but $\inf\{t : X(t) \in K\}$ may not be; however, if $\{\mathcal{F}_t\}$ is right continuous and complete, then for any $B \in \mathcal{B}(E)$, $\tau_B = \inf\{t : X(t) \in B\}$ is an $\{\mathcal{F}_t\}$-stopping time. This result is a special case of the debut theorem, a very technical result from set theory. Note that

$$\{\omega : \tau_B(\omega) < t\} = \{\omega : \exists s < t \exists X(s,\omega) \in B\} = \text{proj}_\Omega\{(s,\omega) : X(s,\omega) \in B, s < t\}$$
Piecewise constant approximations

$\epsilon > 0, \tau_0^\epsilon = 0,$

$$
\tau_{i+1}^\epsilon = \inf \{ t > \tau_i^\epsilon : r(X(t), X(\tau_i^\epsilon)) \lor r(X(t-), X(\tau_i^\epsilon)) \geq \epsilon \}
$$

Define $X^\epsilon(t) = X(\tau_i^\epsilon), \tau_i^\epsilon \leq t < \tau_{i+1}^\epsilon$. Then $r(X(t), X^\epsilon(t)) \leq \epsilon$.

If $X$ is adapted to $\{\mathcal{F}_t\}$, then the $\{\tau_i^\epsilon\}$ are $\{\mathcal{F}_t\}$-stopping times and $X^\epsilon$ is $\{\mathcal{F}_t\}$-adapted. See Exercise 4.
Martingales

An $\mathbb{R}$-valued stochastic process $M$ adapted to $\{\mathcal{F}_t\}$ is an $\{\mathcal{F}_t\}$-martingale if

$$E[M(t + r)|\mathcal{F}_t] = M(t), \quad t, r \geq 0$$

Every martingale has finite quadratic variation:

$$[M]_t = \lim \sum (M(t \wedge t_{i+1}) - M(t \wedge t_i))^2$$

where $0 = t_0 < t_1 < \cdots, t_i \to \infty$, and the limit is in probability as $\max(t_{i+1} - t_i) \to 0$. More precisely, for $\epsilon > 0$ and $t_0 > 0$,

$$\lim P\{\sup_{t \leq t_0} |[M]_t - \lim \sum (M(t \wedge t_{i+1}) - M(t \wedge t_i))^2| > \epsilon\} = 0.$$

For standard Brownian motion $W$, $[W]_t = t$. 
Optional sampling theorem

A real-valued process is a submartingale if $E[|X(t)|] < \infty$, $t \geq 0$, and

$$E[X(t + s)|\mathcal{F}_t] \geq X(t), \quad t, s \geq 0.$$  

If $\tau_1$ and $\tau_2$ are stopping times, then

$$E[X(t \wedge \tau_2)|\mathcal{F}_{\tau_1}] \geq X(t \wedge \tau_1 \wedge \tau_2).$$

If $\tau_2$ is finite a.s. $E[|X(\tau_2)|] < \infty$ and $\lim_{t \to \infty} E[|X(t)|1_{\{\tau_2 > t\}}] = 0$, then

$$E[X(\tau_2)|\mathcal{F}_{\tau_1}] \geq X(\tau_1 \wedge \tau_2).$$

Of course, if $X$ is a martingale

$$E[X(t \wedge \tau_2)|\mathcal{F}_{\tau_1}] = X(t \wedge \tau_1 \wedge \tau_2).$$
Square integrable martingales

\( M \) a martingale satisfying \( E[M(t)^2] < \infty \). Then

\[
M(t)^2 - [M]_t
\]
is a martingale. In particular, for \( t > s \)

\[
E[(M(t) - M(s))^2] = E[[M]_t - [M]_s].
\]
Doob’s inequalities

Let $X$ be a submartingale. Then for $x > 0$,

\[
P\{\sup_{s \leq t} X(s) \geq x \} \leq x^{-1} E[X(t)^+] \\
P\{\inf_{s \leq t} X(s) \leq -x \} \leq x^{-1}(E[X(t)^+] - E[X(0)])
\]

If $X$ is nonnegative and $\alpha > 1$, then

\[
E[\sup_{s \leq t} X(s)^\alpha] \leq \left(\frac{\alpha}{\alpha - 1}\right)^\alpha E[X(t)^\alpha].
\]

Note that by Jensen’s inequality, if $M$ is a martingale, then $|M|$ is a submartingale. In particular, if $M$ is a square integrable martingale, then

\[
E[\sup_{s \leq t} |M(s)|^2] \leq 4E[M(t)^2].
\]
Stochastic integrals

**Definition B.4** For cadlag processes $X, Y$,

$$X_\cdot Y(t) \equiv \int_0^t X(s-)dY(s)$$

$$= \lim_{\max |t_{i+1}-t_i| \to 0} \sum X(t_i)(Y(t_{i+1} \wedge t) - Y(t_i \wedge t))$$

whenever the limit exists in probability.

**Sample paths of bounded variation:** If $Y$ is a finite variation process, the stochastic integral exists (apply dominated convergence theorem) and

$$\int_0^t X(s-)dY(s) = \int_{(0,t]} X(s-)\alpha_Y(ds)$$

$\alpha_Y$ is the signed measure with

$$\alpha_Y(0, t] = Y(t) - Y(0)$$
Existence for square integrable martingales

If $M$ is a square integrable martingale, then

$$E[(M(t + s) - M(t))^2 | \mathcal{F}_t] = E[[M]_{t+s} - [M]_t | \mathcal{F}_t]$$

For partitions $\{t_i\}$ and $\{r_i\}$

$$E \left[ \left( \sum X(t_i)(M(t_{i+1} \wedge t) - M(t_i \wedge t)) 
\quad - \sum X(r_i)(M(r_{i+1} \wedge t) - M(r_i \wedge t)) \right)^2 \right]$$

$$= E \left[ \int_0^t (X(t(s-)) - X(r(s-)))^2 d[M]_s \right]$$

$$= E \left[ \int_{(0,T]} (X(t(s-)) - X(r(s-)))^2 \alpha[M](ds) \right]$$

$t(s) = t_i \text{ for } s \in [t_i, t_{i+1})$ \quad $r(s) = r_i \text{ for } s \in [r_i, r_{i+1})$
Cauchy property

Let $X$ be bounded by a constant. As $\sup(t_{i+1} - t_i) + \sup(r_{i+1} - r_i) \to 0$, the right side converges to zero by the dominated convergence theorem.

$M \{t_i\}(t) \equiv \sum X(t_i)(M(t_{i+1} \wedge t) - M(t_i \wedge t))$ is a square integrable martingale, so

$$E \left[ \sup_{t \leq T} \left( \sum X(t_i)(M(t_{i+1} \wedge t) - M(t_i \wedge t)) ight. ight.$$

$$\left. - \sum X(r_i)(M(r_{i+1} \wedge t) - M(r_i \wedge t)) \right)^2 \right]$$

$$\leq 4E \left[ \int_{(0,t]} (X(t(s-)) - X(r(s-)))^2 \alpha[M](ds) \right]$$

A completeness argument gives existence of the stochastic integral and the uniformity implies the integral is cadlag.
Local martingales

**Definition B.5** \( M \) is a local martingale if there exist stopping times \( \{ \tau_n \} \) satisfying \( \tau_1 \leq \tau_2 \leq \cdots \) and \( \tau_n \to \infty \) a.s. such that \( M^{\tau_n} \) defined by \( M^{\tau_n}(t) = M(\tau_n \land t) \) is a martingale. \( M \) is a local square-integrable martingale if the \( \tau_n \) can be selected so that \( M^{\tau_n} \) is square integrable.

\( \{ \tau_n \} \) is called a localizing sequence for \( M \).

**Remark B.6** If \( \{ \tau_n \} \) is a localizing sequence for \( M \), and \( \{ \gamma_n \} \) is another sequence of stopping times satisfying \( \gamma_1 \leq \gamma_2 \leq \cdots \), \( \gamma_n \to \infty \) a.s. then the optional sampling theorem implies that \( \{ \tau_n \land \gamma_n \} \) is localizing.
Local martingales with bounded jumps

**Remark B.7** If $M$ is a continuous, local martingale, then

$$\tau_n = \inf\{ t : |M(t)| \geq n \}$$

will be a localizing sequence. More generally, if

$$|\Delta M(t)| \leq c$$

for some constant $c$, then

$$\tau_n = \inf\{ t : |M(t)| \lor |M(t-)| \geq n \}$$

will be a localizing sequence.

Note that $|M^{\tau_n}| \leq n + c$, so $M$ is local square integrable.
Semimartingales

**Definition B.8** $Y$ is an $\{\mathcal{F}_t\}$-semimartingale if and only if $Y = M + V$, where $M$ is a local square integrable martingale with respect to $\{\mathcal{F}_t\}$ and $V$ is an $\{\mathcal{F}_t\}$-adapted finite variation process.

In particular, if $X$ is cadlag and adapted and $Y$ is a semimartingale, then $\int X_- dY$ exists.
Computing quadratic variations

Let $\Delta Z(t) = Z(t) = Z(t-)$.  

**Lemma B.9** If $Y$ is finite variation, then  

$$[Y]_t = \sum_{s \leq t} \Delta Y(s)^2$$  

**Lemma B.10** If $Y$ is a semimartingale, $X$ is adapted, and $Z(t) = \int_0^t X(s-)dY(s)$ then  

$$[Z]_t = \int_0^t X(s-)^2d[Y]_s.$$  

**Proof.** Check first for piecewise constant $X$ and then approximate general $X$ by piecewise constant processes.  \[\square\]
Covariation

The covariation of $Y_1, Y_2$ is defined by

$$[Y_1, Y_2]_t \equiv \lim \sum_i \left( Y_1(t_{i+1} \wedge t) - Y_1(t_i \wedge t) \right) \left( Y_2(t_{i+1} \wedge t) - Y_2(t_i \wedge t) \right)$$
Itô’s formula

If \( f : \mathbb{R} \to \mathbb{R} \) is \( C^2 \) and \( Y \) is a semimartingale, then

\[
    f(Y(t)) = f(Y(0)) + \int_0^t f'(Y(s-))dY(s) + \int_0^t \frac{1}{2} f''(Y(s))d[Y]^c_s + \sum_{s \leq t} (f(Y(s)) - f(Y(s-)) - f'(Y(s-))\Delta Y(s))
\]

where \([Y]^c\) is the continuous part of the quadratic variation given by

\[
    [Y]^c_t = [Y]_t - \sum_{s \leq t} \Delta Y(s)^2.
\]
Itô’s formula for vector-valued semimartingales

If $f : \mathbb{R}^m \to \mathbb{R}$ is $C^2$, $Y_1, \ldots, Y_m$ are semimartingales, and $Y = (Y_1, \ldots, Y_m)$, then defining

$$[Y_k, Y_l]^c_t = [Y_k, Y_l]_t - \sum_{s \leq t} \Delta Y_k(s) \Delta Y_l(s),$$

$$f(Y(t)) = f(Y(0)) + \sum_{k=1}^{m} \int_{0}^{t} \partial_k f(Y(s-)) dY_k(s)$$
$$+ \sum_{k,l=1}^{m} \frac{1}{2} \int_{0}^{t} \partial_k \partial_l f(Y(s-)) d[Y_k, Y_l]^c_s$$
$$+ \sum_{s \leq t} (f(Y(s)) - f(Y(s-)) - \sum_{k=1}^{m} \partial_k f(Y(s-)) \Delta Y_k(s)).$$
Examples

$W$ standard Brownian motion

$$Z(t) = \exp\{W(t) - \frac{1}{2}t\} = \int_0^t Z(s)d(W(s) - \frac{1}{2}s) + \int_0^t \frac{1}{2}Z(s)ds$$

$$= \int_0^t Z(s)dW(s)$$
Integration with respect to Poisson random measures

- Poisson random measures
- Stochastic integrals for space-time Poisson random measures
- The predictable \( \sigma \)-algebra
- Martingale properties
- Representation of counting processes
- Stochastic integrals for centered space-time Poisson random measures
- Quadratic variation
- Lévy processes
- Gaussian white noise
Poisson distribution

**Definition P.1** A random variable $X$ has a Poisson distribution with parameter $\lambda > 0$ (write $X \sim \text{Poisson}(\lambda)$) if for each $k \in \{0, 1, 2, \ldots\}$

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

$$E[X] = \lambda \quad Var(X) = \lambda$$

and the characteristic function of $X$ is

$$E[e^{i\theta X}] = e^{\lambda(e^{i\theta} - 1)}.$$  

Since the characteristic function of a random variable characterizes its distribution, a direct computation gives
Proposition P.2 If $X_1, X_2, \ldots$ are independent random variables with $X_i \sim \text{Poisson}(\lambda_i)$ and $\sum_{i=1}^{\infty} \lambda_i < \infty$, then

$$X = \sum_{i=1}^{\infty} X_i \sim \text{Poisson} \left( \sum_{i=1}^{\infty} \lambda_i \right)$$
Poisson sums of Bernoulli random variables

**Proposition P.3** Let \( N \sim \text{Poisson}(\lambda) \), and suppose that \( Y_1, Y_2, \ldots \) are i.i.d. Bernoulli random variables with parameter \( p \in [0, 1] \). If \( N \) is independent of the \( Y_i \), then \( \sum_{i=0}^{N} Y_i \sim \text{Poisson}(\lambda p) \).

For \( j = 1, \ldots, m \), let \( e_j \) be the vector in \( \mathbb{R}^m \) that has all its entries equal to zero, except for the \( j \)th which is 1.

For \( \theta, y \in \mathbb{R}^m \), let \( \langle \theta, y \rangle = \sum_{j=1}^{m} \theta_j y_j \).

**Proposition P.4** Let \( N \sim \text{Poisson}(\lambda) \). Suppose that \( Y_1, Y_2, \ldots \) are independent \( \mathbb{R}^m \)-valued random variables such that for all \( k \geq 0 \) and \( j \in \{1, \ldots, m\} \)

\[
P\{Y_k = e_j\} = p_j,
\]

where \( \sum_{j=1}^{m} p_j = 1 \). Define \( X = (X_1, \ldots, X_m) = \sum_{k=0}^{N} Y_k \). If \( N \) is independent of the \( Y_k \), then \( X_1, \ldots, X_m \) are independent random variables and \( X_j \sim \text{Poisson}(\lambda p_j) \).
Poisson random measures

Let \((U, d_U)\) be a complete, separable metric space, and let \(\nu\) be a \(\sigma\)-finite measure on \(U\). Let \(\mathcal{N}(U)\) denote the collection of counting measures on \(U\).

**Definition P.5** A Poisson random measure on \(U\) with mean measure \(\nu\) is a random counting measure \(\xi\) (that is, a \(\mathcal{N}(U)\)-valued random variable) such that

\begin{enumerate}
  \item For \(A \in \mathcal{B}(U)\), \(\xi(A)\) has a Poisson distribution with expectation \(\nu(A)\)
  \item \(\xi(A)\) and \(\xi(B)\) are independent if \(A \cap B = \emptyset\).
\end{enumerate}

For \(f \in M(U), f \geq 0\), define

\[\psi_{\xi}(f) = E[\exp\{-\int_U f(u)\xi(du)\}] = \exp\{-\int (1 - e^{-f})d\nu\}\]

(Verify the second equality by approximating \(f\) by simple functions.)
Existence

**Proposition P.6** Suppose that $\nu$ is a measure on $U$ such that $\nu(U) < \infty$. Then there exists a Poisson random measure with mean measure $\nu$.

**Proof.** The case $\nu(U) = 0$ is trivial, so assume that $\nu(U) \in (0, \infty)$. Let $N$ be a Poisson random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ with $E[N] = \nu(U)$. Let $X_1, X_2, \ldots$ be iid $U$-valued random variables such that for every $A \in \mathcal{B}(U)$,

$$P\{X_j \in A\} = \frac{\nu(A)}{\nu(U)},$$

and assume that $N$ is independent of the $X_j$.

Define $\xi$ by $\xi(A) = \sum_{k=0}^{N} 1\{X_k \in A\}$. In other words $\xi = \sum_{k=0}^{N} \delta_{X_k}$ where, for each $x \in U$, $\delta_x$ is the Dirac mass at $x$.

Extend the existence result to $\sigma$-finite measures by partitioning $U = \bigcup_i U_i$, where $\nu(U_i) < \infty$. 

□
Identities

Let $\xi$ be a Poisson random measure with mean measure $\nu$.

**Lemma P.7** Suppose $f \in M(U)$, $f \geq 0$. Then

$$E\left[ \int f(y)\xi(dy) \right] = \int f(y)\nu(dy)$$

**Lemma P.8** Suppose $\nu$ is nonatomic and let $f \in M(\mathcal{N}(U) \times U)$, $f \geq 0$. Then

$$E\left[ \int_U f(\xi, y)\xi(dy) \right] = E\left[ \int_U f(\xi + \delta_y, y)\nu(dy) \right]$$
Proof. Suppose $0 \leq f \leq 1_{U_0}$, where $\nu(U_0) < \infty$. Let $U_0 = \bigcup_k U_k^n$, where the $U_k^n$ are disjoint and $\text{diam}(U_k^n) \leq n^{-1}$. If $\xi(U_k^n)$ is 0 or 1, then

$$
\int_{U_k^n} f(\xi, y) \xi(dy) = \int_{U_k^n} f(\xi(\cdot \cap U_k^{n,c}) + \delta_y, y) \xi(dy)
$$

Consequently, if $\max_k \xi(U_k^n) \leq 1$,

$$
\int_{U_0} f(\xi, y) \xi(dy) = \sum_k \int_{U_k^n} f(\xi(\cdot \cap U_k^{n,c}) + \delta_y, y) \xi(dy)
$$

Since $\xi(U_0) < \infty$, for $n$ sufficiently large, $\max_k \xi(U_k^n) \leq 1$,

$$
E[\int_U f(\xi, y) \xi(dy)] = E[\int_{U_0} f(\xi, y) \xi(dy)] = \lim_{n \to \infty} \sum_k E[\int_{U_k^n} f(\xi(\cdot \cap U_k^{n,c}) + \delta_y, y) \xi(dy)]
$$

$$
= \lim_{n \to \infty} \sum_k E[\int_{U_k^n} f(\xi(\cdot \cap U_k^{n,c}) + \delta_y, y) \nu(dy)]
$$
\[ = E \left[ \int_U f(\xi + \delta_y, y) \nu(dy) \right]. \]

Note that the last equality follows from the fact that

\[ f(\xi(\cdot \cap U_{k,c}^n) + \delta_y, y) \neq f(\xi + \delta_y, y) \]

only if \( \xi(U_k^n) > 0 \), and hence, assuming \( 0 \leq f \leq 1_{U_0} \),

\[ | \sum_k \int_{U_k^n} f(\xi(\cdot \cap U_{k,c}^n) + \delta_y, y) \nu(dy) - \int_{U_0} f(\xi + \delta_y, y) \nu(dy) | \leq \sum_k \xi(U_k^n) \nu(U_k^n), \]

where the expectation of the right side is \( \sum_k \nu(U_k^n)^2 = \int_{U_0} \nu(U^n(y)) \nu(dy) \leq \int_{U_0} \nu(U_0 \cap B_{1/n}(y)) \nu(dy) \), where \( U^n(y) = U_k^n \) if \( y \in U_k^n \). \( \lim_{n \to \infty} \nu(U_0 \cap B_{1/n}(y)) = 0 \), since \( \nu \) is nonatomic. \( \square \)
Space-time Poisson random measures

Let $\xi$ be a Poisson random measure on $U \times [0, \infty)$ with mean measure $\nu \times \ell$ (where $\ell$ denotes Lebesgue measure).

$\xi(A, t) \equiv \xi(A \times [0, t])$ is a Poisson process with parameter $\nu(A)$.

If $\nu(A) < \infty$, $\tilde{\xi}(A, t) \equiv \xi(A \times [0, t]) - \nu(A)t$ is a martingale.

**Definition P.9** $\xi$ is $\{\mathcal{F}_t\}$ compatible, if for each $A \in \mathcal{B}(U)$, $\xi(A, \cdot)$ is $\{\mathcal{F}_t\}$ adapted and for all $t, s \geq 0$, $\xi(A \times (t, t + s])$ is independent of $\mathcal{F}_t$. 
Stochastic integrals for Poisson random measures

For $i = 1, \ldots, m$, let $t_i < r_i$ and $A_i \in \mathcal{B}(U)$, and let $\eta_i$ be $\mathcal{F}_{t_i}$-measurable. Let $X(u, t) = \sum_i \eta_i 1_{A_i}(u) 1_{(t_i, r_i)}(t)$, and note that

$$X(u, t- ) = \sum_i \eta_i 1_{A_i}(u) 1_{(t_i, r_i)}(t). \quad (P.1)$$

Define

$$I_\xi(X, t) = \int_U \int [0, t] X(u, s-) \xi(du \times ds) = \sum_i \eta_i \xi(A_i \times (t_i, r_i)).$$

Then

$$E[|I_\xi(X, t)|] \leq E \left[ \int_U \int [0, t] |X(u, s-)| \xi(du \times ds) \right]$$

$$= \int_U \int [0, t] E[|X(u, s)|] \nu(du) ds$$

and if the right side is finite, $E[I_\xi(X, t)] = \int_U \int [0, t] E[X(u, s)] \nu(du) ds$. 
Estimates in $L_{1,0}$

If

$$\int_{U \times [0,t]} |X(u, s-)| \wedge 1 \xi(du \times ds) < \infty$$

then $\xi\{(u, s) : |X(u, s-)| > 1\} < \infty$ and

$$E \left[ \sup_{t \leq T} |I_\xi(X \wedge 1, t)| \right] \leq \int_{U \times [0,T]} E[|X(u, s)| \wedge 1] \nu(du) ds$$

Definition P.10 Let $L_{1,0}(U, \nu)$ denote the space of $\mathcal{B}(U) \times \mathcal{B}[0, \infty) \times \mathcal{F}$-measurable mappings $(u, s, \omega) \rightarrow X(u, s, \omega)$ such that

$$\int_0^\infty e^{-s} \int_U E[|X(u, s)| \wedge 1] \nu(du) ds < \infty.$$

Let $S_-$ denote the collection of $\mathcal{B}(U) \times \mathcal{B}[0, \infty) \times \mathcal{F}$ measurable mappings $(u, s, t) \rightarrow \sum_{i=1}^m \eta_i(\omega) 1_{A_i}(u) 1_{(t_i, r_i]}(t)$ defined as in (P.1).
Lemma P.11

\[ d_{1,0}(X, Y) = \int_0^\infty e^{-s} \int_U E[|X(u, s) - Y(u, s)| \wedge 1] \nu(du) ds \]

defines a metric on \( \mathcal{L}_{1,0}(U, \nu) \), and the definition of \( I_\xi \) extends to the closure of \( S_- \) in \( \mathcal{L}_{1,0}(U, \nu) \).
The predictable $\sigma$-algebra

**Warning:** Let $N$ be a unit Poisson process. Then $\int_0^\infty e^{-s} E[|N(s) - N(s^-)| \wedge 1] ds = 0$, but $P\{\int_0^t N(s)dN(s) \neq \int_0^t N(s^-)dN(s)\} = 1 - e^{-t}$.

**Definition P.12** Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{\mathcal{F}_t\}$ be a filtration in $\mathcal{F}$. The $\sigma$-algebra $\mathcal{P}$ of predictable sets is the smallest $\sigma$-algebra in $\mathcal{B}(U) \times \mathcal{B}[0, \infty) \times \mathcal{F}$ containing sets of the form $A \times (t_0, t_0 + r_0] \times B$ for $A \in \mathcal{B}(U)$, $t_0, r_0 \geq 0$, and $B \in \mathcal{F}_{t_0}$.

**Remark P.13** Note that for $B \in \mathcal{F}_{t_0}$, $1_{A \times (t_0, t_0 + r_0] \times B}(u, t, \omega)$ is left continuous in $t$ and adapted and that the mapping $(u, t, \omega) \to X(u, t-, \omega)$, where $X(u, t-, \omega)$ is defined in (P.1), is $\mathcal{P}$-measurable.

**Definition P.14** A stochastic process $X$ on $U \times [0, \infty)$ is predictable if the mapping $(u, t, \omega) \to X(u, t, \omega)$ is $\mathcal{P}$-measurable.
Lemma P.15 If the mapping \((u, t, \omega) \mapsto X(u, t, \omega)\) is \(\mathcal{B}(U) \times \mathcal{B}[0, \infty) \times \mathcal{F}\)-measurable and adapted and is left continuous in \(t\), then \(X\) is predictable.

Proof. Let \(0 = t_0^n < t_1^n < \cdots\) and \(t_{i+1}^n - t_i^n \leq n^{-1}\). Define \(X_n(u, t, \omega) = X(u, t_i^n, \omega)\) for \(t_i^n < t \leq t_{i+1}^n\). Then \(X_n\) is predictable and

\[
\lim_{n \to \infty} X_n(u, t, \omega) = X(u, t, \omega)
\]

for all \((u, t, \omega)\). \(\square\)
Stochastic integrals for predictable processes

**Lemma P.16** Let $G \in \mathcal{P}$, $B \in \mathcal{B}(U)$ with $\nu(B) < \infty$ and $b > 0$. Then $1_{B \times [0,b]}(u, t)1_G(u, t, \omega)$ is a predictable process and

$$I_\xi(1_{B \times [0,b]}1_G, t)(\omega) = \int_{U \times [0,t]} 1_{B \times [0,b]}(u, s)1_G(u, s, \omega)\xi(du \times ds, \omega) \quad a.s. \quad (P.2)$$

and

$$E\left[\int_{U \times [0,t]} 1_{B \times [0,b]}(u, s)1_G(u, s, \cdot)\xi(du \times ds)\right] \quad (P.3)$$

$$= E\left[\int_{U \times [0,t]} 1_{B \times [0,b]}1_G(u, s, \cdot)\nu(du)ds\right]$$
Proof. Let 

\[ \mathcal{A} = \left\{ \bigcup_{i=1}^{m} A_i \times (t_i, t_i + r_i] \times G_i : t_i, r_i \geq 0, A_i \in \mathcal{B}(U), G_i \in \mathcal{F}_{t_i} \right\}. \]

Then \( \mathcal{A} \) is an algebra. For \( G \in \mathcal{A} \), (P.2) holds by definition, and (P.3) holds by direct calculation. The collection of \( G \) that satisfy (P.2) and (P.3) is closed under increasing unions and decreasing intersections, and the monotone class theorem (see Theorem 4.1 of the Appendix of Ethier and Kurtz (1986)) gives the lemma. □
Lemma P.17 Let $X$ be a predictable process satisfying

$$
\int_0^\infty e^{-s} \int_U E[|X(u, s)| \wedge 1] \nu(du) ds < \infty.
$$

Then $\int_{U \times [0,t]} |X(u, t)| \xi(du \times ds) < \infty \ a.s.$ and

$$
I_\xi(X, t)(\omega) = \int_{U \times [0,t]} X(u, t, \omega) \xi(du \times ds, \omega) \ a.s.
$$

Proof. Approximate by simple functions. \hfill \square
Consequences of predictability

Lemma P.18 If $X$ is predictable and $\int_{U \times [0,t]} |X(u, s)| \wedge 1 \nu(du)ds < \infty$ a.s. for all $t$, then

$$\int_{U \times [0,t]} |X(u, s)| \xi(du \times ds) < \infty \quad a.s. \quad (P.4)$$

and

$$\int_{U \times [0,t]} X(u, s) \xi(du \times ds)$$

exists a.s.

Proof. Let $\tau_c = \inf \{t : \int_{U \times [0,t]} |X(u, s)| \wedge 1 \nu(du)ds \geq c \}$, and consider $X_c(s, u) = 1_{[0,\tau_c]}(s) X(u, s)$. Then $X_c$ satisfies the conditions of Lemma P.17, so

$$\int_{U \times [0,t]} |X(u, s)| \wedge 1 \xi(du \times ds) < \infty \quad a.s.$$

But this implies $\xi\{(u, s) : s \leq t, |X(u, s)| > 1\} < \infty$, so (P.4) holds. □
Martingale properties

**Theorem P.19** Suppose $X$ is predictable and $\int_{U \times [0,t]} E[|X(u, s)|] \nu(du) ds < \infty$ for each $t > 0$. Then

$$\int_{U \times [0,t]} X(u, s) \xi(du \times ds) - \int_0^t \int_U X(u, s) \nu(du) ds$$

is a $\{\mathcal{F}_t\}$-martingale.
Proof. Let $A \in \mathcal{F}_t$ and define $X_A(u, s) = 1_A X(u, s) 1_{(t,t+r]}(s)$. Then $X_A$ is predictable and

$$E[1_A \int_{U \times (t,t+r]} X(u, s) \xi(du \times ds)] = E[\int_{U \times [0,t+r]} X_A(u, s) \xi(du \times ds)]$$

$$= E[\int_{U \times [0,t+r]} X_A(u, s) \nu(du) ds]$$

$$= E[1_A \int_{U \times (t,t+r]} X(u, s) \nu(du) ds]$$

and hence

$$E[\int_{U \times (t,t+r]} X(u, s) \xi(du \times ds) | \mathcal{F}_t] = E[\int_{U \times (t,t+r]} X(u, s) \nu(du) ds | \mathcal{F}_t].$$

□
Local martingales

Lemma P.20 If

$$\int_{U \times [0,t]} |X(u, s)| \nu(du) ds < \infty \quad a.s. \quad t \geq 0,$$

then

$$\int_{U \times [0,t]} X(u, s) \xi(du \times ds) - \int_{U \times [0,t]} X(u, s) \nu(du) ds$$

is a local martingale.
Proof. If $\tau$ is a stopping time and $X$ is predictable, then $1_{[0,\tau]}(t)X(u,t)$ is predictable. Let

$$\tau_c = \{ t > 0 : \int_{U \times [0,t]} |X(u,s)|\nu(du)ds \geq c \}.$$ 

Then

$$\int_{U \times [0,t]} X(u,s)\xi(du \times ds) - \int_{U \times [0,t\wedge \tau_c]} X(u,s)\nu(du)ds$$

$$= \int_{U \times [0,\tau_c]} 1_{[0,\tau_c]}(s)X(u,s)\xi(du \times ds) - \int_{U \times [0,\tau_c]} 1_{[0,\tau_c]}(s)X(u,s)\nu(du)ds.$$ 

is a martingale. □
Representation of counting processes

Let \( U = [0, \infty) \) and \( \nu = \ell \). Let \( \lambda \) be a nonnegative, predictable process, and define \( G = \{(u, t) : u \leq \lambda(t)\} \). Then

\[
N(t) = \int_{[0, \infty) \times [0, t]} 1_G(u, s) \xi(du \times ds) = \int_{[0, \infty) \times [0, t]} 1_{[0, \lambda(s)]}(u) \xi(du \times ds)
\]

is a counting process with intensity \( \lambda \).

Stochastic equation for a counting process

\[
\lambda : [0, \infty) \times D_E[0, \infty) \times D_c[0, \infty) \to [0, \infty), \quad \lambda(t, z, v) = \lambda(t, z^{t^-}, v^{t^-}),
\]

\( t \geq 0, Z \) independent of \( \xi \)

\[
N(t) = \int_{[0, \infty) \times [0, t]} 1_{[0, \lambda(s, Z, N)]}(u) \xi(du \times ds) \quad \text{(P.5)}
\]
Semimartingale property

Corollary P.21 If $X$ is predictable and $\int_{U \times [0,t]} |X(u,s)| \wedge 1 \nu(du)ds < \infty$ a.s. for all $t$, then $\int_{U \times [0,t]} |X(u,s)| \xi(du \times ds) < \infty$ a.s.

$$\int_{U \times [0,t]} X(u,s) \xi(du \times ds)$$

$$= \int_{U \times [0,t]} 1_{\{|X(u,s)\leq 1\}} X(u,s) \xi(du \times ds) - \int_0^t \int_U 1_{\{|X(u,s)\leq 1\}} X(u,s) \nu(du)ds$$

\underbrace{\text{local martingale}}

$$+ \int_0^t \int_U 1_{\{|X(u,s)\leq 1\}} X(u,s) \nu(du)ds + \int_{U \times [0,t]} 1_{\{|X(u,s)|>1\}} X(u,s) \xi(du \times ds)$$

\underbrace{\text{finite variation}}

is a semimartingale.
Stochastic integrals for centered Poisson random measures

Let \( \tilde{\xi}(du \times ds) = \xi(du \times ds) - \nu(du)ds \)

For

\[
X(u, t-) = \sum_i \eta_i 1_{A_i}(u) 1_{(t_i, r_i]}(t).
\]

as in (P.1), define

\[
I_{\tilde{\xi}}(X, t) = \int_{U \times [0,t]} X(u, s-) \tilde{\xi}(du \times ds) = \int_{U \times [0,t]} X(u, s-) \xi(du \times ds) - \int_0^t \int_{U} X(u, s) \nu(du)ds
\]

and note that

\[
E \left[ I_{\tilde{\xi}}(X, t)^2 \right] = \int_{U \times [0,t]} E[X(u, s)^2] \nu(du)ds
\]

if the right side is finite.

Then \( I_{\tilde{\xi}}(X, \cdot) \) is a square-integrable martingale.
Extension of integral

The integral extends to *predictable* integrands satisfying

$$\int_{U \times [0,t]} |X(u, s)|^2 \wedge |X(u, s)| \nu(du) ds < \infty \quad \text{a.s.} \quad \text{(P.6)}$$

so that

$$\int_{U \times [0,t \wedge \tau]} X(u, s) \tilde{\xi}(du \times ds) = \int_{U \times [0,t]} 1_{[0,\tau]}(s) X(u, s) \tilde{\xi}(du \times ds) \quad \text{(P.7)}$$

is a martingale for any stopping time satisfying

$$E \left[ \int_{U \times [0,t \wedge \tau]} |X(u, s)|^2 \wedge |X(u, s)| \nu(du) ds \right] < \infty,$$

and (P.7) is a local square integrable martingale if

$$\int_{U \times [0,t]} |X(u, s)|^2 \nu(du) ds < \infty \quad \text{a.s.}$$
Quadratic variation

Note that if $X$ is predictable and
\[
\int_{U \times [0, t]} |X(u, s)| \wedge 1 \nu(du)ds < \infty \quad a.s. \quad t \geq 0,
\]
then
\[
\int_{U \times [0, t]} |X(u, s)|^2 \wedge 1 \nu(du)ds < \infty \quad a.s. \quad t \geq 0,
\]
and
\[
[I_\xi(X, \cdot)]_t = \int_{U \times [0, t]} X^2(u, s) \xi(du \times ds).
\]

Similarly, if
\[
\int_{U \times [0, t]} |X(u, s)|^2 \wedge |X(u, s)| \nu(du)ds < \infty \quad a.s.,
\]
then
\[
[I_{\tilde{\xi}}(X, \cdot)]_t = \int_{U \times [0, t]} X^2(u, s) \xi(du \times ds).
\]
Semimartingale properties

**Theorem P.22** Let $Y$ be a cadlag, adapted process. If $X$ is predictable and satisfies (P.4), $I_\xi(X, \cdot)$ is a semimartingale and

$$\int_0^t Y(s-)dI_\xi(X, s) = \int_{U \times [0,t]} Y(s-)X(u, s)\xi(du \times ds),$$

and if $X$ satisfies (P.6), $I_{\tilde{\xi}}(X, \cdot)$ is a semimartingale and

$$\int_0^t Y(s-)dI_{\tilde{\xi}}(X, s) = \int_{U \times [0,t]} Y(s-)X(u, s)\tilde{\xi}(du \times ds).$$
Lévy processes

**Theorem P.23** Let $U = \mathbb{R}$ and $\int_{\mathbb{R}} |u|^2 \wedge 1 \nu(du) < \infty$. Then

$$Z(t) = \int_{[-1,1] \times [0,t]} u \tilde{\xi}(du \times ds) + \int_{[-1,1]^c \times [0,t]} u \xi(du \times ds)$$

is a process with stationary, independent increments with

$$E[e^{i\theta Z(t)}] = \exp\left\{ t \int_{\mathbb{R}} (e^{i\theta u} - 1 - i\theta u 1_{[-1,1]}(u))\nu(du) \right\}$$
Proof.

\[ e^{i\theta Z(t)} = 1 + \int_0^t i\theta e^{i\theta Z(s-)} dZ(s) + \sum_{s \leq t}(e^{i\theta Z(s)} - e^{i\theta Z(s-)} - i\theta e^{i\theta Z(s-)} \Delta Z(s)) \]

\[ = 1 + \int_{[-1,1] \times [0,t]} i\theta e^{i\theta Z(s-)} u\tilde{\xi}(du \times ds) + \int_{[-1,1]^c \times [0,t]} i\theta e^{i\theta Z(s-)} u\xi(du \times ds) \]

\[ + \int_{\mathbb{R} \times [0,t]} (e^{i\theta(Z(s-)+u)} - e^{i\theta Z(s-)} - i\theta e^{i\theta Z(s-)} u)\xi(du \times ds) \]

\[ = 1 + \int_{[-1,1] \times [0,t]} i\theta e^{i\theta Z(s-)} u\tilde{\xi}(du \times ds) \]

\[ + \int_{\mathbb{R} \times [0,t]} e^{i\theta Z(s-)}(e^{i\theta u} - 1 - i\theta u 1_{[-1,1]}(u))\xi(du \times ds) \]

Taking expectations

\[ \varphi(\theta, t) = 1 + \int_{\mathbb{R} \times [0,t]} \varphi(\theta, s)(e^{i\theta u} - 1 - i\theta u 1_{[-1,1]}(u))\nu(du)ds \]

so \( \varphi(\theta, t) = \exp\{t \int_{\mathbb{R}} (e^{i\theta u} - 1 - i\theta u 1_{[-1,1]}(u))\nu(du)\} \)

\( \square \)
Approximation of Lévy processes

For $0 < \epsilon < 1$, let

$$
Z^\epsilon(t) = \int_{[-1,-\epsilon) \cup (\epsilon,1] \times [0,t]} u\tilde{\xi}(du \times ds) + \int_{[-1,1] \times [0,t]} u\xi(du \times ds)
$$

$$
= \int_{(-\infty,-\epsilon) \cup (\epsilon,\infty) \times [0,t]} u\xi(du \times ds) - t \int_{[-1-\epsilon) \cup (\epsilon,1]} u\nu(du)
$$

that is, throw out all jumps of size less than or equal to $\epsilon$ and the corresponding centering. Then

$$
E[|Z^\epsilon(t) - Z(t)|^2] = t \int_{[-\epsilon,\epsilon]} u^2 \nu(du).
$$

Consequently, since $Z^\epsilon - Z$ is a square integrable martingale, Doob’s inequality gives

$$
\lim_{\epsilon \to 0} E[\sup_{s \leq t} |Z^\epsilon(s) - Z(s)|^2] = 0.
$$
Summary on stochastic integrals

If $X$ is predictable and $\int_{U \times [0, t]} |X(u, s)| \wedge 1 \nu(du) ds < \infty$ a.s. for all $t$, then

$$\int_{U \times [0, t]} |X(u, s)| \xi(du \times ds) < \infty \text{ a.s.}$$

$$\int_{U \times [0, t]} X(u, s) \xi(du \times ds)$$

$$= \int_{U \times [0, t]} 1_{\{ |X(u, s)| \leq 1 \}} X(u, s) \xi(du \times ds) - \int_0^t \int_{U} 1_{\{ |X(u, s)| \leq 1 \}} X(u, s) \nu(du) ds$$

local martingale

$$+ \int_0^t \int_{U} 1_{\{ |X(u, s)| \leq 1 \}} X(u, s) \nu(du) ds + \int_{U \times [0, t]} 1_{\{ |X(u, s)| > 1 \}} X(u, s) \xi(du \times ds)$$

finite variation

is a semimartingale.
If $X$ is predictable and

$$\int_{U \times [0,t]} |X(u, s)|^2 \wedge |X(u, s)| \nu(du) ds < \infty \quad \text{a.s.,}$$

then

$$\int_{U \times [0,t]} X(u, s) \tilde{\xi}(du \times ds)$$

$$= \lim_{\epsilon \to 0^+} \int_{U \times [0,t]} \mathbf{1}_{\{|X(u,s)| \geq \epsilon\}}(s) X(u, s) \tilde{\xi}(du \times ds)$$

$$= \lim_{\epsilon \to 0^+} \left( \int_{U \times [0,t]} \mathbf{1}_{\{|X(u,s)| \geq \epsilon\}} X(u, s) \xi (du \times ds) \right)$$

$$- \int_0^t \int_U \mathbf{1}_{\{|X(u,s)| \geq \epsilon\}} X(u, s) \nu(du) ds$$

exists and is a local martingale.
Gaussian white noise

\((U, d_U)\) a complete, separable metric space; \(\mathcal{B}(U)\), the Borel sets
\(\mu\) a (Borel) measure on \(U\)
\[\mathcal{A}(U) = \{A \in \mathcal{B}(U) : \mu(A) < \infty\}\]
\(W(A, t) \equiv W(A \times [0, t])\) Mean zero, Gaussian process indexed by \(\mathcal{A}(U) \times [0, \infty)\)

\[E[W(A, t)W(B, s)] = t \wedge s\mu(A \cap B),\]
\[W(\varphi, t) = \int \varphi(u)W(du, t)\]
\[\varphi(u) = \sum a_i I_{A_i}(u)\]
\[W(\varphi, t) = \sum_i a_i W(A_i, t)\]

\[E[W(\varphi_1, t)W(\varphi_2, s)] = t \wedge s \int_U \varphi_1(u)\varphi_2(u)\mu(du)\]

Define \(W(\varphi, t)\) for all \(\varphi \in L_2(\mu)\).
Definition of integral

\[ X(t) = \sum_i \xi_i(t) \varphi_i \quad \text{adapted process in } L_2(\mu) \]

\[ I_W(X, t) = \int_{U \times [0,t]} X(s, u) W(du \times ds) = \sum_i \int_0^t \xi_i(s) dW(\varphi_i, s) \]

\[ E[I_W(X, t)^2] = E\left[ \sum_{i,j} \int_0^t \xi_i(s) \xi_j(s) ds \int_U \varphi_i \varphi_j d\mu \right] \]

\[ = E\left[ \int_0^t \int_U X(s, u)^2 \mu(du) ds \right] \]

The integral extends to measurable and adapted processes satisfying

\[ \int_0^t \int_U X(s, u)^2 \mu(du) ds < \infty \quad a.s. \]

so that

\[ (I_W(X, t))^2 - \int_0^t \int_U X(s, u)^2 \mu(du) ds \]

is a local martingale.
Technical lemmas

- Some definitions
- Caratheodary extension theorem
- Dynkin class theorem
- Essential supremum
- Martingale convergence theorem
- Kronecker’s lemma
- Law of large numbers for martingales
- Geometric rates
- Uniform integrability
- Measurable functions
- Measurable selection
• Dominated convergence theorem
• Metric spaces
• Sequential compactness
• Completeness
Some definitions

Definition T.1

$D_E[0, \infty)$ is the space of cadlag functions with the Skorohod ($J_1$) topology.

A collection of functions $D \subset B(E)$ is separating if $\int f \, d\mu_1 = \int f \, d\mu_2$ for all $f \in D$ implies $\mu_1 = \mu_2$.

$t$ is a fixed point of discontinuity for a cadlag process $X$ if

$$P\{X(t) \neq X(t-)\} > 0$$
Caratheodary extension theorem

**Theorem T.2** Let $M$ be a set, and let $A$ be an algebra of subsets of $M$. If $\mu$ is a $\sigma$-finite measure on $A$, then there exists a unique extension of $\mu$ to a measure on $\sigma(A)$. 
Dynkin class theorem

A collection $\mathcal{D}$ of subsets of $\Omega$ is a Dynkin class if $\Omega \in \mathcal{D}$, $A, B \in \mathcal{D}$ and $A \subset B$ imply $B - A \in \mathcal{D}$, and $\{A_n\} \subset \mathcal{D}$ with $A_1 \subset A_2 \subset \cdots$ imply $\bigcup_{n} A_n \in \mathcal{D}$.

**Theorem T.3** Let $S$ be a collection of subsets of $\Omega$ such that $A, B \in S$ implies $A \cap B \in S$. If $\mathcal{D}$ is a Dynkin class with $S \subset \mathcal{D}$, then $\sigma(S) \subset \mathcal{D}$.

**Corollary T.4** Suppose two finite measures, $\mu$, $\nu$, have the same total mass and agree on a collection of subsets $S$ that is closed under intersection. Then they agree on $\sigma(S)$. 
Essential Supremum

Let \( \{Z_{\alpha}, \alpha \in \mathcal{I}\} \) be a collection of random variables. Note that if \( \mathcal{I} \) is uncountable, \( \sup_{\alpha \in \mathcal{I}} Z_{\alpha} \) may not be a random variable; however, we have the following:

**Lemma T.5** There exists a random variable \( \overline{Z} \) such that \( P\{Z_{\alpha} \leq \overline{Z}\} = 1 \) for each \( \alpha \in \mathcal{I} \) and there exist \( \alpha_k, k = 1, 2, \ldots \) such that \( \overline{Z} = \sup_k Z_{\alpha_k} \).

**Proof.** Without loss of generality, we can assume \( 0 < Z_{\alpha} < 1 \). (Otherwise, replace \( Z_{\alpha} \) by \( \frac{1}{1+e^{-Z_{\alpha}}} \).) Let \( C = \sup \{E[Z_{\alpha_1} \vee \cdots \vee Z_{\alpha_m}], \alpha_1, \ldots, \alpha_m \in \mathcal{I}, m = 1, 2, \ldots \} \). Then there exist \( (\alpha_1^n, \ldots, \alpha_m^n) \) such that

\[
C = \lim_{n \to \infty} E[Z_{\alpha_1^n} \vee \cdots \vee Z_{\alpha_m^n}].
\]

Define \( \overline{Z} = \sup \{Z_{\alpha_i^n}, 1 \leq i \leq m_n, n = 1, 2, \ldots \} \), and note that \( C = E[\overline{Z}] \) and \( C = E[\overline{Z} \vee Z_{\alpha}] \) for each \( \alpha \in \mathcal{I} \). Consequently, \( P\{Z_{\alpha} \leq \overline{Z}\} = 1 \). \[ \square \]
Martingale convergence theorem

**Theorem T.6** Suppose \{X_n\} is a submartingale and \(\sup_n E[|X_n|] < \infty\). Then \(\lim_{n \to \infty} X_n\) exists a.s.
Kronecker’s lemma

Lemma T.7 Let \( \{ A_n \} \) and \( \{ Y_n \} \) be sequences of random variables where \( A_0 > 0 \) and \( A_{n+1} \geq A_n, \ n = 0, 1, 2, \ldots \). Define \( R_n = \sum_{k=1}^{n} \frac{1}{A_{k-1}}(Y_k - Y_{k-1}) \). and suppose that \( \lim_{n \to \infty} A_n = \infty \) and that \( \lim_{n \to \infty} R_n \) exists a.s. Then, \( \lim_{n \to \infty} \frac{Y_n}{A_n} = 0 \) a.s.

Proof.

\[
A_n R_n = \sum_{k=1}^{n} (A_k R_k - A_{k-1} R_{k-1}) = \sum_{k=1}^{n} R_{k-1} (A_k - A_{k-1}) + \sum_{k=1}^{n} A_k (R_k - R_{k-1})
\]

\[
= Y_n - Y_0 + \sum_{k=1}^{n} R_{k-1} (A_k - A_{k-1}) + \sum_{k=1}^{n} \frac{1}{A_{k-1}} (Y_k - Y_{k-1})(A_k - A_{k-1})
\]

and

\[
\frac{Y_n}{A_n} = \frac{Y_0}{A_n} + R_n - \frac{1}{A_n} \sum_{k=1}^{n} R_{k-1} (A_k - A_{k-1}) - \frac{1}{A_n} \sum_{k=1}^{n} \frac{1}{A_{k-1}} (Y_k - Y_{k-1})(A_k - A_{k-1})
\]
Law of large numbers for martingales

**Lemma T.8** Suppose \( \{A_n\} \) is as in Lemma T.7 and is adapted to \( \{F_n\} \), and suppose \( \{M_n\} \) is a \( \{F_n\} \)-martingale such that for each \( \{F_n\} \)-stopping time \( \tau \), \( E[(M_\tau - M_{\tau - 1})^2 1_{\{\tau < \infty\}}] < \infty \). If

\[
\sum_{k=1}^{\infty} \frac{1}{A_{k-1}^2} (M_k - M_{k-1})^2 < \infty \quad \text{a.s.,}
\]

then \( \lim_{n \to \infty} \frac{M_n}{A_n} = 0 \) a.s.

**Proof.** Without loss of generality, we can assume that \( A_n \geq 1 \). Let

\[
\tau_c = \min\{n : \sum_{k=1}^{n} \frac{1}{A_{k-1}^2} (M_k - M_{k-1})^2 \geq c\}.
\]

Then

\[
\sum_{k=1}^{\infty} \frac{1}{A_{k-1}^2} (M_{k \wedge \tau_c} - M_{(k-1) \wedge \tau_c})^2 \leq c + (M_{\tau_c} - M_{\tau_c-1})^2 1_{\{\tau_c < \infty\}}.
\]
It follows that $R_n^c = \sum_{k=1}^{n} \frac{1}{A_{k-1}} (M_{k \wedge \tau_c} - M_{(k-1) \wedge \tau_c})$ converges a.s. and hence, by Lemma T.7, that $\lim_{n \to \infty} \frac{M_{n \wedge \tau_c}}{A_n} = 0$. \qed
Geometric convergence

Lemma T.9 Let \( \{M_n\} \) be a martingale with \(|M_{n+1} - M_n| \leq c\) a.s. for each \( n \) and \( M_0 = 0 \). Then for each \( \epsilon > 0 \), there exist \( C \) and \( \eta \) such that

\[
P\{\frac{1}{n}|M_n| \geq \epsilon\} \leq Ce^{-n\eta}.
\]

Proof. Let \( \hat{\varphi}(x) = e^{-x} + e^x \) and \( \varphi(x) = e^x - 1 - x \). Then, setting \( X_k = M_k - M_{k-1} \)

\[
E[\hat{\varphi}(aM_n)] = 2 + \sum_{k=1}^{n} E[\hat{\varphi}(aM_k) - \hat{\varphi}(aM_{k-1})]
\]

\[
= 2 + \sum_{k=1}^{n} E[\exp\{aM_{k-1}\} \varphi(aX_k) + \exp\{-aM_{k-1}\} \varphi(-aX_k)]
\]

\[
\leq 2 + \sum_{k=1}^{n} \varphi(ac) E[\hat{\varphi}(aM_{k-1})],
\]
and hence
\[ E[\hat{\varphi}(aM_n)] \leq 2e^{n\varphi(ac)}. \]

Consequently,
\[
P\{\sup_{k\leq n} \frac{1}{n}|M_k| \geq \epsilon\} \leq \frac{E[\hat{\varphi}(aM_n)]}{\hat{\varphi}(an\epsilon)} \leq 2e^{n(\varphi(ac)-a\epsilon)}.\]

Then \( \eta = \sup_a (a\epsilon - \varphi(ac)) > 0 \), and the lemma follows. \( \square \)
Uniform integrability

Lemma T.10 If $X$ is integrable, then for $\epsilon > 0$ there exists a $K > 0$ such that

$$\int_{\{|X| > K\}} |X|dP < \epsilon.$$  

Proof. $\lim_{K \to \infty} |X|1_{\{|X| > K\}} = 0$ a.s. \qed

Lemma T.11 If $X$ is integrable, then for $\epsilon > 0$ there exists a $\delta > 0$ such that $P(F') < \delta$ implies $\int_{F} |X|dP < \epsilon.$

Proof. Let $F_n = \{|X| \geq n\}$. Then $nP(F_n) \leq E[|X|1_{F_n}] \to 0$. Select $n$ so that $E[|X|1_{F_n}] \leq \epsilon/2$, and let $\delta = \frac{\epsilon}{2n}$. Then $P(F') < \delta$ implies

$$\int_{F} |X|dP \leq \int_{F_n} |X|dP + \int_{F_n^c \cap F} |X|dP < \frac{\epsilon}{2} + n\delta = \epsilon$$

\qed
Theorem T.12 Let \( \{X_\alpha\} \) be a collection of integrable random variables. The following are equivalent:

a) \( \sup E[|X_\alpha|] < \infty \) and for \( \epsilon > 0 \) there exists \( \delta > 0 \) such that \( P(F) < \delta \) implies \( \sup_\alpha \int_F |X_\alpha| dP < \epsilon \).

b) \( \lim_{K \to \infty} \sup_\alpha E[|X_\alpha| \mathbf{1}_{\{|X_\alpha| > K\}}] = 0 \).

c) \( \lim_{K \to \infty} \sup_\alpha E[|X_\alpha| - |X_\alpha| \wedge K] = 0 \)

d) There exists a convex function \( \varphi \) with \( \lim_{|x| \to \infty} \varphi(x) = \infty \) such that \( \sup_\alpha E[\varphi(|X_\alpha|)] < \infty \).
Proof. a) implies b) follows by
\[ P\{\|X_\alpha\| > K\} \leq \frac{E[\|X_\alpha\|]}{K} \]

b) implies d): Select $N_k$ such that
\[ \sum_{k=1}^{\infty} k \sup_{\alpha} E[\mathbf{1}_{\{\|X_\alpha\| > N_k\}} \|X_\alpha\|] < \infty \]

Define $\varphi(0) = 0$ and
\[ \varphi'(x) = k, \quad N_k \leq x < N_{k+1}. \]

Recall that $E[\varphi(\|X\|)] = \int_0^\infty \varphi'(x) P\{|X| > x\} dx$, so
\[ E[\varphi(\|X_\alpha\|)] = \sum_{k=1}^{\infty} k \int_{N_k}^{N_{k+1}} P\{|X_\alpha| > x\} dx \leq \sum_{k=1}^{\infty} k \sup_{\alpha} E[\mathbf{1}_{\{\|X_\alpha\| > N_k\}} \|X_\alpha\|]. \]

d) implies b): $E[\mathbf{1}_{\{\|X_\alpha\| > K\}} \|X_\alpha\|] < \frac{E[\varphi(\|X_\alpha\|)]}{\varphi(K)/K}$
b) implies a): $\int_\mathcal{F} |X_\alpha|dP \leq P(F)K + E[1_{\{|X_\alpha|>K\}}|X_\alpha|].$

To see that (b) is equivalent to (c), observe that

$$E[|X_\alpha| - |X_\alpha| \wedge K] \leq E[|X_\alpha|1_{\{|X_\alpha|>K\}}] \leq 2E[|X_\alpha| - |X_\alpha| \wedge \frac{K}{2}]$$

$\square$
Uniformly integrable families

- For $X$ integrable, $\Gamma = \{ E[X|\mathcal{D}] : \mathcal{D} \subset \mathcal{F} \}$
- For $X_1, X_2, \ldots$ integrable and identically distributed
  \[ \Gamma = \left\{ \frac{X_1 + \cdots + X_n}{n} : n = 1, 2, \ldots \right\} \]
- For $Y \geq 0$ integrable, $\Gamma = \{ X : |X| \leq Y \}$. 
Uniform integrability and $L^1$ convergence

**Theorem T.13**  \( X_n \to X \) in $L^1$ iff \( X_n \to X \) in probability and \( \{X_n\} \) is uniformly integrable.

**Proof.** If \( X_n \to X \) in $L^1$, then

\[
\lim_{n \to \infty} E[|X_n| - |X_n| \wedge K] = E[|X| - |X| \wedge K]
\]

and Part (c) of Theorem T.12 follows from the fact that

\[
\lim_{K \to \infty} E[|X| - |X| \wedge K] = \lim_{K \to \infty} E[|X_n| - |X_n| \wedge K] = 0.
\]

\( \square \)
Measurable functions

Let \((M_i, \mathcal{M}_i)\) be measurable spaces.

\(f : M_1 \rightarrow M_2\) is measurable if \(f^{-1}(A) = \{x \in M_1 : f(x) \in A\} \in \mathcal{M}_1\) for each \(A \in \mathcal{M}_2\).

**Lemma T.14** If \(f : M_1 \rightarrow M_2\) and \(g : M_2 \rightarrow M_3\) are measurable, then \(g \circ f : M_1 \rightarrow M_3\) is measurable.
Measurable selection theorem

**Theorem T.15** Let \((M, \mathcal{M})\) be a measurable space and \((S, \rho)\) a complete, separable metric space. Suppose for each \(x \in M\), \(\Gamma_x\) is a closed subset of \(S\). and that for every open set \(U \subset S\), \(\{x \in M : \Gamma_x \cap U \neq \emptyset\}\) \(\in \mathcal{M}\). Then there exist \(f_n : M \to S\), \(n = 1, 2, \ldots\), such that \(f_n\) is \(\mathcal{M}\)-measurable, \(f_n(x) \in \Gamma_x\) for every \(x \in M\), and \(\Gamma_x\) is the closure of \(\{f_1(x), f_2(x), \ldots\}\).
Domained convergence theorem

**Theorem T.16** Let $X_n \to X$ and $Y_n \to Y$ in probability. Suppose that $|X_n| \leq Y_n$ a.s. and $E[Y_n|D] \to E[Y|D]$ in probability. Then

$$E[X_n|D] \to E[X|D] \quad \text{in probability}$$

**Proof.** A sequence converges in probability iff every subsequence has a further subsequence that converges a.s., so we may as well assume almost sure convergence. Let $D_{m,c} = \{\sup_{n \geq m} E[Y_n|D] \leq c\}$. Then

$$E[Y_n 1_{D_{m,c}}|D] = E[Y_n|D] 1_{D_{m,c}} \xrightarrow{L_1} E[Y|D] 1_{D_{m,c}} = E[Y 1_{D_{m,c}}|D].$$

Consequently, $E[Y_n 1_{D_{m,c}}] \to E[Y 1_{D_{m,c}}]$, so $Y_n 1_{D_{m,c}} \to Y 1_{D_{m,c}}$ in $L_1$ by the ordinary dominated convergence theorem. It follows that $X_n 1_{D_{m,c}} \to X 1_{D_{m,c}}$ in $L_1$ and hence

$$E[X_n|D] 1_{D_{m,c}} = E[X_n 1_{D_{m,c}}|D] \xrightarrow{L_1} E[X 1_{D_{m,c}}|D] = E[X|D] 1_{D_{m,c}}.$$
Since $m$ and $c$ are arbitrary, the lemma follows.
Metric spaces

d : S × S → [0, ∞) is a metric on S if and only if d(x, y) = d(y, x),
d(x, y) = 0 if and only if x = y, and d(x, y) ≤ d(x, z) + d(z, y).

If d is a metric then d ∧ 1 is a metric.

Examples

- \( \mathbb{R}^m \) \( d(x, y) = |x - y| \)
- \( C[0, 1] \) \( d(x, y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)| \)
- \( C[0, \infty) \) \( d(x, y) = \int_0^\infty e^{-t} \sup_{s \leq t} 1 \wedge |x(s) - y(s)| \; dt \)
Sequential compactness

$K \subset S$ is sequentially compact if every sequence $\{x_n\} \subset K$ has a convergent subsequence with limit in $K$.

**Lemma T.17** If $(S, d)$ is a metric space, then $K \subset S$ is compact if and only if $K$ is sequentially compact.

**Proof.** Suppose $K$ is compact. Let $\{x_n\} \subset K$. If $x$ is not a limit point of $\{x_n\}$, then there exists $\epsilon_x > 0$ such $\max\{n : x_n \in B_{\epsilon_x}(x)\} < \infty$. If $\{x_n\}$ has no limit points, then $\{B_{\epsilon_x}(x), x \in K\}$ is an open cover of $K$. The existence of a finite subcover contradicts the definition of $\epsilon_x$.

If $K$ is sequentially compact, and $\{U_\alpha\}$ is an open cover of $K$. Let $x_1 \in K$ and $\epsilon_1 > \frac{1}{2} \sup_\alpha \sup\{r : B_r(x_1) \subset U_\alpha\}$ and define recursively, $x_{k+1} \in K \cap (\bigcup_{l=1}^k B_{\epsilon_l}(x_l))$ and $\epsilon_{k+1} > \frac{1}{2} \sup_\alpha \sup\{r : B_r(x_{k+1}) \subset U_\alpha\}$. (If $x_{k+1}$ does not exist, then there is a finite subcover in $\{U_\alpha\}$.) By sequential compactness, $\{x_k\}$ has a limit point $x$ and $x \notin B_{\epsilon_k}(x_k)$ for
any $k$. But setting $\epsilon = \frac{1}{2} \sup_{\alpha} \sup \{ r : B_{r}(x) \subset U_{\alpha} \}$, $\epsilon_k > \epsilon - d(x, x_k)$, so if $d(x, x_k) < \epsilon/2$, $x \in B_{\epsilon_k}(x_k)$. \hfill \square
Completeness

A metric space \((S, d)\) is complete if and only if every Cauchy sequence has a limit.

Completeness depends on the metric, not the topology: For example

\[
r(x, y) = \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right|
\]

is a metric giving the usual topology on the real line, but \(\mathbb{R}\) is not complete under this metric.
Exercises

1. $\tau$ is an $\{F_t\}$-stopping time if for each $t \geq 0$, $\{\tau \leq t\} \in F_t$.

For a stopping time $\tau$,

$$F_\tau = \{ A \in F : \{\tau \leq t\} \cap A \in F_t, t \geq 0 \}$$

(a) Show that $F_\tau$ is a $\sigma$-algebra.

(b) Show that for $\{F_t\}$-stopping times $\sigma, \tau$, $\sigma \leq \tau$ implies that $F_\sigma \subset F_\tau$. In particular, $F_{\tau \wedge t} \subset F_t$.

(c) Let $\tau$ be a discrete $\{F_t\}$-stopping time satisfying $\{\tau < \infty\} = \cup_{k=1}^\infty \{\tau = t_k\} = \Omega$. Show that $F_\tau = \sigma\{A \cap \{\tau = t_k\} : A \in F_{t_k}, k = 1, 2, \ldots\}$.

(d) Show that the minimum of two stopping times is a stopping time and that the maximum of two stopping times is a stopping time.

2. Prove the $W$ as defined in the background material is a $\sigma$-algebra.

3. Let $N$ be a Poisson process with parameter $\lambda$. Then $M(t) = N(t) - \lambda t$ is a martingale. Compute $[M]_t$.

4. Let $0 = \tau_0 < \tau_1 < \cdots$ be stopping times satisfying $\lim_{k \to \infty} \tau_k = \infty$, and for $k = 0, 1, 2, \ldots$, let $\xi_k \in F_{\tau_k}$.

Define

$$X(t) = \sum_{k=0}^\infty \xi_k 1_{[\tau_k, \tau_{k+1})}(t).$$

Show that $X$ is adapted.

**Example:** Let $X$ be a cadlag adapted process and let $\epsilon > 0$. Define $\tau_0^\epsilon = 0$ and for $k = 0, 1, 2, \ldots$, $\tau_{k+1}^\epsilon = \inf\{t > \tau_k^\epsilon : |X(t) - X(\tau_k^\epsilon)| \lor |X(t-) - X(\tau_k^\epsilon)| \geq \epsilon\}$.

Note that the $\tau_k^\epsilon$ are stopping times, by Problem 1. Define

$$X^\epsilon(t) = \sum_{k=0}^\infty X(\tau_k^\epsilon) 1_{[\tau_k^\epsilon, \tau_{k+1}^\epsilon)}(t).$$
Then $X^\epsilon$ is a piecewise constant, adapted process satisfying
\[
\sup_t |X(t) - X^\epsilon(t)| \leq \epsilon.
\]

5. Show that $E[f(X)|\mathcal{D}] = E[f(X)]$ for all bounded continuous functions (all bounded measurable functions) if and only if $X$ is independent of $\mathcal{D}$.

6. Let $X$ and $Y$ $S$-valued random variables defined on $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G} \subset \mathcal{F}$ be a sub-$\sigma$-algebra. Suppose that $M \subset \overline{C}(S)$ is separating and
\[
E[f(X)|\mathcal{G}] = f(Y) \quad a.s.
\]
for every $f \in M$. Show that $X = Y$ a.s.

7. Let $\tau$ be a discrete stopping time with values $\{t_i\}$. Show that
\[
E[Z|\mathcal{F}_{\tau}] = \sum_i E[Z|\mathcal{F}_{t_i}]1_{\{\tau=t_i\}}.
\]

8. Let $N$ be a stochastically continuous counting processes with independent increments. Show that the increments of $N$ are Poisson distributed.

9. Let $N_n$ and $N$ be counting processes. Suppose that $(N_n(t_1), \ldots, N_n(t_m)) \Rightarrow (N(t_1), \ldots, N(t_m))$ for all choices of $t_1, \ldots, t_m$ in $\mathcal{T}_0$, where $\mathcal{T}_0$ is dense in $[0, \infty)$. Show that $N_n \Rightarrow N$ under the Skorohod topology.

10. Suppose the $Y_1, \ldots, Y_m$ are independent Poisson random variables with $E[Y_i] = \lambda_i$. Let $Y = \sum_{i=1}^m Y_i$. Compute $P\{Y_i = 1|Y = 1\}$. More generally, compute $P\{Y_i = k|Y = m\}$.

11. Let $\xi$ be a Poisson random measure on $(U, d_U)$ with mean measure $\nu$. For $f \in M(U)$, $f \geq 0$, show that
\[
E[\exp\{-\int_U f(u)\xi(du)\}] = \exp\{-\int \{1 - e^{-f}\}d\nu\}.
\]
12. Give an example of a right continuous process $X$ such that $\int_{U \times [0,t]} |X(u, s)| \wedge 1\nu(du)ds < \infty$ a.s. but $\int_{U \times [0,t]} |X(u, s)| \xi(du \times ds) = \infty$ a.s.

13. Let $Y$, $\xi_1$, and $\xi_2$ be independent random variables with values in complete, separable metric spaces, $E$, $S_1$, and $S_2$. Suppose that $G : E \times S_1 \to E_0$ and $H : E \times S_2 \to E_0$ are Borel measurable functions and that $G(Y, \xi_1) = H(Y, \xi_2)$ a.s. Show that there exists a Borel measurable function $F : E \to E_0$ such that $F(Y) = G(Y, \xi_1) = H(Y, \xi_2)$ a.s.
14. Let $X$ be a measurable stochastic process that is bounded by a constant $c$, and let $N_n$ be a Poisson process with parameter $n$ that is independent of $X$. Let $S^n_k$ denote the jump times of $N_n$. Show that

$$\sum_k X(S^n_k) (S^n_{k+1} \land t - S^n_k \land t)$$

converges in $L_1$ to $\int_0^t X(s) \, ds$.

15. Suppose that $X$ is a Markov process with generator $A$ and semigroup $\{T(t)\}$. Let $V$ be a unit Poisson process independent of $X$, and define

$$X_n(t) = X(\frac{1}{n} V(nt)).$$

Show that $X_n$ is a Markov process. What is its generator.

16. Let $\{Y_k\}$ be a discrete time Markov chain with state space $E$ and transition operator $\mu(x, \Gamma) = P\{Y_{k+1} \in \Gamma \mid Y_k = x\}$. Let $\lambda : E \to [0, \infty)$, and let $\Delta_i, i = 0, 1, \ldots$ be independent, unit exponential random variables. Define

$$X(t) = Y_k \quad \text{for} \quad \sum_{i=0}^{k-1} \frac{\Delta_i}{\lambda(Y_i)} \leq t < \sum_{i=1}^{k} \frac{\Delta_i}{\lambda(Y_i)}.$$

Show that $X$ is Markov and compute its generator. (Assume $\lambda$ is bounded to avoid technicalities.)
17. For a one-dimensional diffusion with generator

\[ Af(x) = \frac{1}{2} a(x) f''(x) + b(x) f'(x), \]

check the claim in (2.2).

18. Suppose that \( A \subset B(E) \times B(E) \) and that a solution of the martingale problem for \((A, \delta_x)\) exists for each \( x \in E \). Show that \( A \) is dissipative. (Recall the equivalent forms of the martingale problem.)
19. Let \( E = (0, 1] \) and \( L = \{ f \in C((0, 1]) : \lim_{x \to 0} f(x) \text{ exists} \} \). Consider the operator

\[ Af(x) = -f'(x) \]

with domain

\[ \mathcal{D}(A) = \{ f \in L : f' \in L, \lim_{x \to 0} f(x) = \int_0^1 f(y)dy \} . \]

a) Show that \( \mathcal{R}(\lambda - A) = L \) (implying uniqueness for the corresponding martingale problem), but \( \mathcal{D}(A) \) is not dense in \( L \).

b) Describe the behavior of the process.

c) Is this process quasi-left continuous? Explain.

20. Give an example of a right continuous, adapted process \( X \) (not predictable!) such that

\[ \int_{U \times [0,t]} |X(u, s)| \wedge 1 \nu(du)ds < \infty \quad \text{a.s.} \]

but

\[ \int_{U \times [0,t]} |X(u, s)| \xi(du \times ds) = \infty \quad \text{a.s.} \]

21. Develop a stochastic differential equation model for the movement of a toad in a hale storm. Assume that when a hale stone lands, the toad jumps directly away from where the hale stone landed, and assume that the distance the toad jumps depends on how close the hale stone lands to the toad and how large the hale stone is.

For simplicity, assume that the toad is on an infinite, flat surface and that the statistics of the hale storm are translationally and rotationally invariant.

22. Derive the generator for the model developed in Problem 21.

23. For the model developed in Problem 21, give conditions under which the movement of the toad can be approximated by a Brownian motion.
24. Use the martingale central limit theorem to prove a central limit theorem for a continuous time finite Markov chain \( \{\xi(t), t \geq 0\} \). Specifically, assume that the chain is irreducible so that there is a unique stationary distribution \( \{\pi_k\} \) and prove convergence for the vector-valued process \( U_n = (U_n^1, \ldots, U_n^d) \) defined by

\[
U_n^k(t) = \frac{1}{\sqrt{n}} \int_0^{nt} (1_{\{\xi(s) = k\}} - \pi_k) ds.
\]

25. Show that \( U_n \) defined in Problem 24 is not a “good” sequence of semimartingales. (The easiest approach is probably to show that the conclusion of the stochastic integral convergence theorem is not valid.)

26. Show that \( U_n \) can be written as \( M_n + Z_n \) where \( \{M_n\} \) is a “good” sequence and \( Z_n \Rightarrow 0 \).

27. (Random evolutions) Let \( \xi \) be as in Problem 24, and let \( X_n \) satisfy

\[
\dot{X}_n(t) = \sqrt{n}F(X_n(s), \xi(ns)).
\]

Suppose \( \sum_i F(x, i)\pi_i = 0 \). Write \( X_n \) as a stochastic differential equation driven by \( U_n \), give conditions under which \( X_n \) converges in distribution to a limit \( X \), and identify the limit.
28. Prove the following:

**Lemma T.18** Suppose \( \{X_n\} \) is relatively compact in \( D_S[0, \infty) \) and that for each \( n \), \( \{\mathcal{F}_t^n\} \) is a filtration and \( \pi_t^{(n)} \) is the conditional distribution of \( X_n(t) \) given \( \mathcal{F}_t^n \). Then for each \( \epsilon > 0 \) and \( T > 0 \), there exists a compact \( K_{\epsilon,T} \subset S \) such that

\[
\sup_n P\{\sup_{t \leq T} \pi_t^{(n)}(K_{\epsilon,T}^c) \geq \epsilon\} \leq \epsilon
\]
29. Let \( \xi_1, \xi_2, \ldots \) be iid uniform \([0, r]\) and independent of \( N(0) \). Let \( U_i(0) = \xi_i, \ i = 1, \ldots, N(0) \), and for \( b > 0 \), let

\[
\dot{U}_i(t) = bU_i(t)
\]

so

\[
U_i(t) = U_i(0)e^{bt}.
\]

Define

\[
N(t) = \#\{i : U_i(t) < r\}.
\]

Let \( f(u, n) = \prod_{i=1}^{n} g(u_i) \), where \( 0 \leq g \leq 1 \), \( g \) is continuously differentiable, and \( g(u_i) = 1 \) for \( u_i > r \).

The generator for \( U(t) = (U_1(t), \ldots, U_{N(t)}) \) is

\[
Af(u, n) = f(u, n) \sum_{i=1}^{n} bu_i g'(u_i)/g(u_i)
\]

and

\[
f(U(t), N(t)) - f(U(0), N(0)) - \int_{0}^{t} Af(U(s), N(s))ds = 0
\]

is a martingale.

(a) Apply the Markov mapping theorem with \( \gamma(u_1, \ldots, u_n) = n \) and

\[
\alpha(n, du) = \frac{1}{r^n} du_1 \ldots du_n
\]

to show that \( N \) is a Markov process.

(b) Give a direct argument that shows why \( N \) is a Markov process.
30. Verify the following moment identities for a Poisson random measure $\xi$ with mean measure $\nu$

$$E[e^{\int f(z)\xi(dz)}] = e^{\int (e^f - 1)d\nu},$$

or letting $\xi = \sum_i \delta_{Z_i}$,

$$E[\prod_i g(Z_i)] = e^{\int (g - 1)d\nu}.$$

Similarly,

$$E[\sum_j h(Z_j) \prod_i g(Z_i)] = \int hgd\nu e^{\int (g - 1)d\nu},$$

and

$$E[\sum_{i \neq j} h(Z_i)h(Z_j) \prod_k g(Z_k)] = (\int hgd\nu)^2 e^{\int (g - 1)d\nu},$$

31. Consider a process with state space $((-1, 1) \times E \times [0, r])^n$ and generator given by $f(z) = \prod_i^n g(x_i, y_i, u_i)$, $z_i = (x_i, y_i, u_i)$, and

$$Af(z) = \sum_{i=1}^n \lambda \left( \frac{g(-x_i, y_i, u_i)}{g(x_i, y_i, u_i)} - 1 \right) + \sum_{i \neq j} \mu \mathbf{1}_{\{u_i < u_j, x_i = x_j\}} f(z) \left( \frac{g(x_j, y_i, u_j)}{g(x_j, y_j, u_j)} - 1 \right).$$

Apply the Markov mapping theorem to characterize the process $\eta^n(t) = \sum_{i=1}^n \delta(x_i, y_i)$ and, under appropriate scaling, use the lookdown model to derive a limiting model for $\eta^n$ as $r \to \infty$. 
Stochastic analysis exercises

1. Show that if $Y_1$ is cadlag and $Y_2$ is finite variation, then

$$[Y_1, Y_2]_t = \sum_{s \leq t} \Delta Y_1(s) \Delta Y_2(s).$$

2. Using the fact that martingales have finite quadratic variation, show that semimartingales have finite quadratic variation.

3. Using the above results, show that the covariation of two semimartingales exist.

4. Consider the stochastic differential equation

$$X(t) = X(0) + \int_0^t aX(s)dW(s) + \int_0^t bX(s)ds.$$  

Find $\alpha$ and $\beta$ so that

$$X(t) = X(0) \exp\{\alpha W(t) + \beta t\}$$

is a solution.

5. Let $W$ be standard Brownian motion and suppose $(X, Y)$ satisfies

$$X(t) = x + \int_0^t Y(s)ds$$

$$Y(t) = y - \int_0^t X(s)ds + \int_0^t cX(s-)dW(s)$$

where $c \neq 0$ and $x^2 + y^2 > 0$. Assuming all moments are finite, define $m_1(t) = E[X(t)^2]$, $m_2(t) = E[X(t)Y(t)]$, and $m_3(t) = E[Y(t)^2]$. Find a system of linear differential equations satisfied by $(m_1, m_2, m_3)$, and show that the expected “total energy” $(E[X(t)^2 + Y(t)^2])$ is asymptotic to $ke^{\lambda t}$ for some $k > 0$ and $\lambda > 0$. 
6. Let $X$ and $Y$ be independent Poisson processes. Show that with probability one, $X$ and $Y$ do not have simultaneous discontinuities and that $[X, Y]_t = 0$, for all $t \geq 0$.

7. Two local martingales, $M$ and $N$, are called orthogonal if $[M, N]_t = 0$ for all $t \geq 0$.

(a) Show that if $M$ and $N$ are orthogonal, then $[M + N]_t = [M]_t + [N]_t$.
(b) Show that if $M$ and $N$ are orthogonal, then $M$ and $N$ do not have simultaneous discontinuities.
(c) Show that if $M$ is finite variation, then $M$ and $N$ are orthogonal if and only if they have no simultaneous discontinuities.

8. Let $W$ be standard Brownian motion. Use Ito’s Formula to show that
$$M(t) = e^{\alpha W(t) - \frac{1}{2} \alpha^2 t}$$
is a martingale. (Note that the martingale property can be checked easily by direct calculation; however, the problem asks you to use Ito’s formula to check the martingale property.)

9. Let $N$ be a Poisson process with parameter $\lambda$. Use Ito’s formula to show that
$$M(t) = e^{\alpha N(t) - \lambda(e^\alpha - 1)t}$$
is a martingale.

10. Let $X$ satisfy
$$X(t) = x + \int_0^t \sigma X(s) dW(s) + \int_0^t bX(s) ds$$
Let $Y = X^2$.

(a) Derive a stochastic differential equation satisfied by $Y$.
(b) Find $E[X(t)^2]$ as a function of $t$. 
11. Suppose that the solution of $dX = b(X)dt + \sigma(X)dW$, $X(0) = x$ is unique for each $x$. Let $\tau = \inf\{t > 0 : X(t) \notin (\alpha, \beta)\}$ and suppose that for some $\alpha < x < \beta$, $P\{\tau < \infty, X(\tau) = \alpha|X(0)\} > 0$ and $P\{\tau < \infty, X(\tau) = \beta|X(0) = x\} > 0$.

(a) Show that $P\{\tau < T, X(\tau) = \alpha|X(0) = x\}$ is a nonincreasing function of $x$, $\alpha < x < \beta$.

(b) Show that there exists a $T > 0$ such that

$$ \inf_x \max \{P\{\tau < T, X(\tau) = \alpha|X(0) = x\}, P\{\tau < T, X(\tau) = \beta|X(0) = x\}\} > 0 $$

(c) Let $\gamma$ be a nonnegative random variable. Suppose that there exists a $T > 0$ and a $\rho < 1$ such that for each $n$, $P\{\gamma > (n + 1)T|\gamma > nT\} < \rho$. Show that $E[\gamma] < \infty$.

(d) Show that $E[\tau] < \infty$. 
References


Abstract

Martingale problems and stochastic equations

Beginning with Lvy's characterization of Brownian motion, the martingale properties of stochastic processes have proven to be fundamental in the formulation and analysis of stochastic models, in particular, in the study of Markov processes. Stochastic equations of various types provide a second fundamental approach to defining stochastic models. These lectures will explore aspects of both of these approaches and their relationship. The Martingale problem of Stroock and Varadhan will be formulated and some of its properties described. The notion of a weak solution of a stochastic differential equation will be introduced and the fact that solutions of appropriate martingale problems give weak solutions of stochastic differential equations will be discussed. Time change equations for Markov chains and diffusions will be given and the equivalence of these equations to corresponding martingale problems shown.