Lectures on Stochastic Analysis

Thomas G. Kurtz
Departments of Mathematics and Statistics
University of Wisconsin - Madison
Madison, WI 53706-1388

Revised September 7, 2001

Contents

1 Introduction.

The first draft of these notes was prepared by the students in Math 735 at the University of Wisconsin - Madison during the fall semester of 1992. The students faithfully transcribed many of the errors made by the lecturer. While the notes have been edited and many errors removed, particularly due to a careful reading by Geoffrey Pritchard, many errors undoubtedly remain. Read with care.

These notes do not eliminate the need for a good book. The intention has been to state the theorems correctly with all hypotheses, but no attempt has been made to include detailed proofs. Parts of proofs or outlines of proofs have been included when they seemed to illuminate the material or at the whim of the lecturer.

2 Review of probability.

A *probability space* is a triple $(\Omega, \mathcal{F}, P)$ where Ω is the set of "outcomes", $\mathcal{F}$ is a σ-algebra of "events", that is, subsets of Ω, and $P : \mathcal{F} \to [0, \infty)$ is a measure that assigns "probabilities" to events. A (real-valued) *random variable* X is a real-valued function defined on Ω such that for every Borel set $B \in \mathcal{B}(\mathbb{R})$, we have $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$. (Note that the *Borel σ-algebra* $\mathcal{B}(\mathbb{R})$) is the smallest σ-algebra containing the open sets.) We will occasionally also consider S-valued random variables where S is a separable metric space (e.g., $\mathbb{R}^d$). The definition is the same with $\mathcal{B}(S)$ replacing $\mathcal{B}(\mathbb{R})$.

The probability distribution on S determined by

$$\mu_X(B) = P(X^{-1}(B)) = P\{X \in B\}$$

is called the *distrbution* of X. A random variable X has a *discrete distribution* if its range is countable, that is, there exists a sequence $\{x_i\}$ such that $\sum P\{X = x_i\} = 1$. The expectation of a random variable with a discrete distribution is given by

$$E[X] = \sum x_i P\{X = x_i\}$$

provided the sum is absolutely convergent. If X does not have a discrete distribution, then it can be approximated by random variables with discrete distributions. Define $\overline{X}_n = \frac{k+1}{n}$ and $\underline{X}_n = \frac{k}{n}$ when $\frac{k}{n} < X \leq \frac{k+1}{n}$, and note that $\underline{X}_n < X \leq \overline{X}_n$ and $|\overline{X}_n - \underline{X}_n| \leq \frac{1}{n}$. Then

$$E[X] \equiv \lim_{n\to\infty} E[\underline{X}_n] = \lim E[\overline{X}_n]$$

provided $E[\overline{X}_n]$ exists for some (and hence all) n. If $E[X]$ exists, then we say that X is *integrable*.

2.1 Properties of expectation.

a) Linearity: $E[aX + bY] = aE[X] + bE[Y]$

b) Monotonicity: if $X \geq Y$ a.s then $E[X] \geq E[Y]$

2.2 Convergence of random variables.

a) $X_n \to X$ a.s. iff $P\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1$.

b) $X_n \to X$ in probability iff $\forall \epsilon > 0$, $\quad \lim_{n\to\infty} P\{|X_n - X| > \epsilon\} = 0$.

c) X_n converges to X in distribution (denoted $X_n \Rightarrow X$) iff $\lim_{n\to\infty} P\{X_n \leq x\} = P\{X \leq x\} \equiv F_X(x)$ for all x at which F_X is continuous.

Theorem 2.1 *a) implies b) implies c).*

Proof. $(b \Rightarrow c)$ Let $\epsilon > 0$. Then

$$\begin{aligned} P\{X_n \leq x\} - P\{X \leq x + \epsilon\} &= P\{X_n \leq x, X > x + \epsilon\} - P\{X \leq x + \epsilon, X_n > x\} \\ &\leq P\{|X_n - X| > \epsilon\} \end{aligned}$$

and hence $\limsup P\{X_n \leq x\} \leq P\{X \leq x + \epsilon\}$. Similarly, $\liminf P\{X_n \leq x\} \geq P\{X \leq x - \epsilon\}$. Since ϵ is arbitrary, the implication follows. □

2.3 Convergence in probability.

a) If $X_n \to X$ in probability and $Y_n \to Y$ in probability then $aX_n + bY_n \to aX + bY$ in probability.

b) If $\mathcal{Q} : \mathbb{R} \to \mathbb{R}$ is continuous and $X_n \to X$ in probability then $\mathcal{Q}(X_n) \to \mathcal{Q}(X)$ in probability.

c) If $X_n \to X$ in probability and $X_n - Y_n \to 0$ in probability, then $Y_n \to X$ in probability.

Remark 2.2 *(b) and (c) hold with convergence in probability replaced by convergence in distribution; however (a) is not in general true for convergence in distribution.*

Theorem 2.3 *(Bounded Convergence Theorem) Suppose that $X_n \Rightarrow X$ and that there exists a constant b such that $P(|X_n| \leq b) = 1$. Then $E[X_n] \to E[X]$.*

Proof. Let $\{x_i\}$ be a partition of $\mathbb{R}$ such that F_X is continuous at each x_i. Then

$$\sum_i x_i P\{x_i < X_n \leq x_{i+1}\} \leq E[X_n] \leq \sum_i x_{i+1} P\{x_i < X_n \leq x_{i+1}\}$$

and taking limits we have

$$\begin{aligned} \sum_i x_i P\{x_i < X \leq x_{i+1}\} &\leq \underline{\lim}_{n\to\infty} E[X_n] \\ &\leq \overline{\lim}_{n\to\infty} E[X_n] \leq \sum_i x_{i+1} P\{x_i < X \leq x_{i+1}\} \end{aligned}$$

As $\max |x_{i+1} - x_i| \to 0$, the left and right sides converge to $E[X]$ giving the theorem. □

Lemma 2.4 *Let $X \geq 0$ a.s. Then $\lim_{M\to\infty} E[X \wedge M] = E[X]$.*

Proof. Check the result first for X having a discrete distribution and then extend to general X by approximation. □

Theorem 2.5 *(Monotone Convergence Theorem.) Suppose $0 \leq X_n \leq X$ and $X_n \to X$ in probability. Then $\lim_{n\to\infty} E[X_n] = E[X]$.*

Proof. For $M > 0$

$$E[X] \geq E[X_n] \geq E[X_n \wedge M] \to E[X \wedge M]$$

where the convergence on the right follows from the bounded convergence theorem. It follows that

$$E[X \wedge M] \leq \liminf_{n\to\infty} E[X_n] \leq \limsup_{n\to\infty} E[X_n] \leq E[X]$$

and the result follows by Lemma 2.4. □

Lemma 2.6 . *(Fatou's lemma.) If $X_n \geq 0$ and $X_n \Rightarrow X$, then* $\liminf E[X_n] \geq E[X]$.

Proof. Since $E[X_n] \geq E[X_n \wedge M]$ we have

$$\liminf E[X_n] \geq \liminf E[X_n \wedge M] = E[X \wedge M].$$

By the Monotone Convergence Theorem $E[X \wedge M] \to E[X]$ and the lemma folllows. $\square$

Theorem 2.7 *(Dominated Convergence Theorem) Assume $X_n \Rightarrow X$, $Y_n \Rightarrow Y$, $|X_n| \leq Y_n$, and $E[Y_n] \to E[Y] < \infty$. Then $E[X_n] \to E[X]$.*

Proof. For simplicity, assume in addition that $X_n + Y_n \Rightarrow X + Y$ and $Y_n - X_n \Rightarrow Y - X$ (otherwise consider subsequences along which $(X_n, Y_n) \Rightarrow (X, Y)$). Then by Fatou's lemma $\liminf E[X_n + Y_n] \geq E[X + Y]$ and $\liminf E[Y_n - X_n] \geq E[Y - X]$. From these observations $\liminf E[X_n] + \lim E[Y_n] \geq E[X] + E[Y]$, and hence $\liminf E[X_n] \geq E[X]$. Similarly $\liminf E[-X_n] \geq E[-X]$ and $\limsup E[X_n] \leq E[X]$ $\square$

Lemma 2.8 *(Markov's inequality)*

$$P\{|X| > a\} \leq E[|X|]/a, \quad a \geq 0.$$

Proof. Note that $|X| \geq aI_{\{|X|>a\}}$. Taking expectations proves the desired inequality. $\square$

2.4 Norms.

For $1 \leq p < \infty$, L_p is the collection of random variables X with $E[|X|^p] < \infty$ and the L_p-norm is defined by $||X||_p = E[|X|^p]^{1/p}$. L_∞ is the collection of random variables X such that $P\{|X| \leq c\} = 1$ for some $c < \infty$, and $||X||_\infty = \inf\{c : P\{|X| \leq c\} = 1$.

Properties of norms:

1) $||X - Y||_p = 0$ implies $X = Y$ a.s. .
2) $|E[XY]| \leq ||X||_p ||Y||_q \qquad \frac{1}{p} + \frac{1}{q} = 1.$
3) $||X + Y||_p \leq ||X||_p + ||Y||_p$

Schwartz inequality. $(p = q = \frac{1}{2})$.
Note that

$$0 \leq E[(aX + bY)^2] = a^2 E[X^2] + 2abE[XY] + b^2 E[Y^2].$$

Assume that $E[XY] \leq 0$ (otherwise replace X by $-X$) and take $a, b > 0$. Then

$$-E[XY] \leq \frac{a}{b} E[X^2] + \frac{b}{a} E[Y^2] .$$

Take $a = \|X\|$ and $b = \|Y\|$. $\square$

Triangle inequality. $(p = \frac{1}{2})$

We have

$$\begin{aligned}\|X+Y\|^2 &= E[(X+Y)^2] \\ &= E[X^2]+2E[XY]+E[Y^2] \\ &\leq \|X\|^2+2\|X\|\|Y\|+\|Y\|^2 \\ &= (\|X\|+\|Y\|)^2.\end{aligned}$$

$\square$

It follows that $r_p(X,Y)=||X-Y||_p$ defines a metric on L_p, the space of random variables satisfying $E[|X|^p]<\infty$. (Note that we identify two random variables that differ on a set of probability zero.) Recall that a sequence in a metric space is *Cauchy* if

$$\lim_{n,m\to\infty} r_p(X_n,X_m)=0$$

and a metric space is *complete* if every Cauchy sequence has a limit. For example, in the case $p=1$, suppose $\{X_n\}$ is Cauchy and let n_k satisfy

$$\sup_{m>n_k}\|X_m-X_{n_k}\|_1=E[|X_m-X_{n_k}|]\leq\frac{1}{4^k}.$$

Then, with probability one, the series

$$X\equiv X_{n_1}+\sum_{k=1}^{\infty}(X_{n_{k+1}}-X_{n_k})$$

is absolutely convergent, and it follows that

$$\lim_{m\to\infty}\|X_m-X\|=0.$$

2.5 Information and independence.

Information obtained by observations of the outcome of a random experiment is represented by a sub-σ-algebra $\mathcal{D}$ of the collection of events $\mathcal{F}$. If $D\in\mathcal{D}$, then the oberver "knows" whether or not the outcome is in D.

An S-valued random variable Y is *independent* of a σ-algebra $\mathcal{D}$ if

$$P(\{Y\in B\}\cap D)=P(\{Y\in B\}P(D),\forall B\in\mathcal{B}(S),D\in\mathcal{D}.$$

Two σ-algebras $\mathcal{D}_1,\mathcal{D}_2$ are independent if

$$P(D_1\cap D_2)=P(D_1)P(D_2),\quad\forall D_1\in\mathcal{D}_1,D_2\in\mathcal{D}_2.$$

Random variables X and Y are independent if $\sigma(X)$ and $\sigma(Y)$ are independent, that is, if

$$P(\{X\in B_1\}\cap\{Y\in B_2\})=P\{X\in B_1\}P\{Y\in B_2\}.$$

2.6 Conditional expectation.

Interpretation of conditional expectation in L_2.

Problem: Approximate $X \in L_2$ using information represented by $\mathcal{D}$ such that the mean square error is minimized, i.e., find the $\mathcal{D}$-measurable random variable Y that minimizes $E[(X-Y)^2]$.

Solution: Suppose Y is a minimizer. For any $\varepsilon \neq 0$ and any $\mathcal{D}$-measurable random variable $Z \in L_2$

$$E[|X-Y|^2] \leq E[|X-Y-\varepsilon Z|^2] = E[|X-Y|^2] - 2\varepsilon E[Z(X-Y)] + \varepsilon^2 E[Z^2].$$

Hence $2\varepsilon E[Z(X-Y)] \leq \varepsilon^2 E[Z^2]$. Since ε is arbitrary, $E[Z(X-Y)] = 0$ and hence

$$E[ZX] = E[ZY] \tag{2.1}$$

for every $\mathcal{D}$-measurable Z with $E[Z^2] < \infty$. □

With (2.1) in mind, for an integrable random variable X, the *conditional expectation* of X, denoted $E[X|\mathcal{D}]$, is the unique (up to changes on events of probability zero) random variable Y satisfying

A) Y is $\mathcal{D}$-measurable.

B) $\int_D X dP = \int_D Y dP$ for all $D \in \mathcal{D}$.

Note that Condition B is a special case of (2.1) with $Z = I_D$ (where I_D denotes the indicator function for the event D) and that Condition B implies that (2.1) holds for all bounded $\mathcal{D}$-measurable random variables. Existence of conditional expectations is a consequence of the Radon-Nikodym theorem.

The following lemma is useful in verifying Condition B.

Lemma 2.9 *Let $\mathcal{C} \subset \mathcal{F}$ be a collection of events such that $\Omega \in \mathcal{C}$ and $\mathcal{C}$ is closed under intersections, that is, if $D_1, D_2 \in \mathcal{C}$, then $D_1 \cap D_2 \in \mathcal{C}$. If X and Y are integrable and*

$$\int_D X dP = \int_D Y dP \tag{2.2}$$

for all $D \in \mathcal{C}$, then (2.2) holds for all $D \in \sigma(\mathcal{C})$ (the smallest σ-algebra containing $\mathcal{C}$).

Example: Assume that $\mathcal{D} = \sigma(D_1, D_2, \ldots,)$ where $\bigcup_{i=1}^{\infty} D_i = \Omega$, and $D_i \cap D_j = \emptyset$ whenever $i \neq j$. Let X be any $\mathcal{F}$-measurable random variable. Then,

$$E[X|\mathcal{D}] = \sum_{i=1}^{\infty} \frac{E[XI_{D_i}]}{P(D_i)} I_{D_i}$$

To see that the above expression is correct, first note that the right hand side is $\mathcal{D}$-measurable. Furthermore, any $D \,\epsilon\, \mathcal{D}$ can be written as $D = \bigcup_{i \epsilon A} D_i$, where $A \subset \{1, 2, 3, \ldots\}$. Therefore,

$$\begin{aligned}
\int_D \sum_{i=1}^{\infty} \frac{E[X \cdot I_{D_i}]}{P(D_i)} I_{D_i} dP &= \sum_{i=1}^{\infty} \frac{E[X \cdot I_{D_i}]}{P(D_i)} \int_{D \cap D_i} I_{D_i} dP \qquad \text{(monotone convergence thm)} \\
&= \sum_{i \epsilon A} \frac{E[X \cdot I_{D_i}]}{P(D_i)} P(D_i) \\
&= \int_D X dP
\end{aligned}$$

Properties of conditional expectation. Assume that X and Y are integrable random variables and that $\mathcal{D}$ is a sub-σ-algebra of $\mathcal{F}$.

1) $E[E[X|\mathcal{D}]] = E[X]$. Just take $D = \Omega$ in Condition B.

2) If $X \geq 0$ then $E[X|\mathcal{D}] \geq 0$. The property holds because $Y = E[X|\mathcal{D}]$ is $\mathcal{D}$-measurable and $\int_D Y dP = \int_D X dP \geq 0$ for every $D \,\epsilon\, \mathcal{D}$. Therefore, Y must be positive a.s.

3) $E[aX + bY|\mathcal{D}] = aE[X|\mathcal{D}] + bE[Y|\mathcal{D}]$. It is obvious that the RHS is $\mathcal{D}$-measurable, being the linear combination of two $\mathcal{D}$-measurable random variables. Also,

$$\begin{aligned}
\int_D (aX + bY) dP &= a \int_D X dP + b \int_D Y dP \\
&= a \int_D E[X|\mathcal{D}] dP + b \int_D E[Y|\mathcal{D}] dP \\
&= \int_D (aE[X|\mathcal{D}] + bE[Y|\mathcal{D}]) dP.
\end{aligned}$$

4) If $X \geq Y$ then $E[X|\mathcal{D}] \geq E[Y|\mathcal{D}]$. Use properties (2) and (3) for $Z = X - Y$.

5) If X is $\mathcal{D}$-measurable, then $E[X|\mathcal{D}] = X$.

6) If Y is $\mathcal{D}$-measurable and YX is integrable, then $E[YX|\mathcal{D}] = YE[X|\mathcal{D}]$. First assume that Y is a simple random variable, *i.e.*, let $\{D_i\}_{i=1}^{\infty}$ be a partition of Ω, $D_i \,\epsilon\, \mathcal{D}$, $c_i \in \mathbb{R}$, for $1 \leq i \leq \infty$, and define $Y = \sum_{i=1}^{\infty} c_i I_{D_i}$. Then,

$$\begin{aligned}
\int_D YX dP &= \int_D \left(\sum_{i=1}^{\infty} c_i I_{D_i} \right) X dP \\
&= \sum_{i=1}^{\infty} c_i \int_{D \cap D_i} X dP \\
&= \sum_{i=1}^{\infty} c_i \int_{D \cap D_i} E[X|\mathcal{D}] dP \\
&= \int_D \left(\sum_{i=1}^{\infty} c_i I_{D_i} \right) E[X|\mathcal{D}] dP
\end{aligned}$$

$$= \int_D YE[X|\mathcal{D}]P$$

For general Y, approximate by a sequence $\{Y_n\}_{n=1}^\infty$ of simple random variables, for example, defined by

$$Y_n = \frac{k}{n} \qquad \text{if} \quad \tfrac{k}{n} \le Y < \tfrac{k+1}{n},\ k \in \mathbb{Z}.$$

Then Y_n converges to Y, and the result follows by the Dominated Convergence Theorem.

7) If X is independent of $\mathcal{D}$, then $E[X|\mathcal{D}] = E[X]$. Independence implies that for $D \in \mathcal{D}$, $E[XI_D] = E[X]P(D)$,

$$\begin{aligned}\int_D XdP &= \int_\Omega XI_D dP \\ &= E[XI_D] \\ &= E[X]\int_\Omega I_D dP \\ &= \int_D E[X]dP\end{aligned}$$

Since $E[X]$ is $\mathcal{D}$-measurable, $E[X] = E[X|\mathcal{D}]$.

8) If $\mathcal{D}_1 \subset \mathcal{D}_2$ then $E[E[X|\mathcal{D}_2]|\mathcal{D}_1] = E[X|\mathcal{D}_1]$. Note that if $D \,\epsilon\, \mathcal{D}_1$ then $D \,\epsilon\, \mathcal{D}_2$. Therefore,

$$\begin{aligned}\int_D XdP &= \int_D E[X|\mathcal{D}_2]dP \\ &= \int_D E[E[X|\mathcal{D}_2]|\mathcal{D}_1]dP,\end{aligned}$$

and the result follows.

A function $\phi : \mathbb{R} \to \mathbb{R}$ is *convex* if and only if for all x and y in $\mathbb{R}$, and λ in $[0,1]$, $\phi(\lambda x + (1-\lambda)y) \le \lambda\phi(x) + (1-\lambda)\phi(y)$. We need the following fact about convex functions for the proof of the next property. Let $x_1 < x_2$ and $y \,\epsilon\, \mathbb{R}$. Then

$$\frac{\phi(x_2) - \phi(y)}{x_2 - y} \ge \frac{\phi(x_1) - \phi(y)}{x_1 - y}. \tag{2.3}$$

Now assume that $x_1 < y < x_2$ and let x_2 converge to y from above. The left side of (2.3) is bounded below, and its value decreases as x_2 decreases to y. Therefore, the right derivative ϕ^+ exists at y and

$$-\infty < \phi^+(y) = \lim_{x_2 \to y+} \frac{\phi(x_2) - \phi(y)}{x_2 - y} < +\infty.$$

Moreover,

$$\phi(x) \ge \phi(y) + \phi^+(y)(x - y), \qquad \forall x \in \mathbb{R}. \tag{2.4}$$

9) **Jensen's Inequality.** If ϕ is convex then

$$E[\phi(X)|\mathcal{D}] \geq \phi(E[X|\mathcal{D}]).$$

Define $M : \Omega \to \mathbb{R}$ as $M = \phi^+(E[X|\mathcal{D}])$. As measurability is preserved under composition, we can see that M is $\mathcal{D}$-measurable. From (2.4),

$$\phi(X) \geq \phi(E[X|\mathcal{D}]) + M(X - E[X|\mathcal{D}]),$$

and

$$\begin{aligned}
E[\phi(X)|\mathcal{D}] &\geq E[\phi(E[X|\mathcal{D}])|\mathcal{D}] + E[M(X - E[X|\mathcal{D}])|\mathcal{D}] \qquad \text{Properties 3 and 4} \\
&= \phi(E[X|\mathcal{D}]) + ME[(X - E[X|\mathcal{D}])|\mathcal{D}] \qquad \text{Property 6} \\
&= \phi(E[X|\mathcal{D}]) + M\{E[X|\mathcal{D}] - E[E[X|\mathcal{D}]|\mathcal{D}]\} \qquad \text{Property 3} \\
&= \phi(E[X|\mathcal{D}]) + M\{E[X|\mathcal{D}] - E[X|\mathcal{D}]\} \qquad \text{Property 8} \\
&= \phi(E[X|\mathcal{D}])
\end{aligned}$$

10) Let X be an S_1-valued, $\mathcal{D}$-measurable random variable and Y be an S_2-valued random variable independent of $\mathcal{D}$. Suppose that $\varphi : S_1 \times S_2 \to \mathbb{R}$ is a measurable function and that $\varphi(X, Y)$ is integrable. Define

$$\psi(x) = E[\varphi(x, Y)].$$

Then, $E[\varphi(X, Y)|\mathcal{D}] = \psi(X)$.

11) Let Y be an S_2-valued random variable (not necessarily independent of $\mathcal{D}$). Suppose that $\varphi : S_1 \times S_2 \to \mathbb{R}$ is a bounded measurable function. Then there exists a measurable $\psi : \Omega \times S_1 \to \mathbb{R}$ such that for each $x \in S_1$

$$\psi(\omega, x) = E[\varphi(x, Y)|\mathcal{D}](\omega) \qquad a.s.$$

and

$$E[\varphi(X, Y)|\mathcal{D}](\omega) = \psi(\omega, X(\omega)) \quad a.s.$$

for every $\mathcal{D}$-measurable random variable X.

12) Let $Y : \Omega \to \mathbb{N}$ be independent of the i.i.d random variables $\{X_i\}_{i=1}^{\infty}$. Then

$$E[\sum_{i=1}^{Y} X_i|\sigma(Y)] = Y \cdot E[X_1]. \tag{2.5}$$

Identity (2.5) follows from Property (10) by taking $\varphi(X, Y)(\omega) = \sum_{i=1}^{Y(\omega)} X_i(\omega)$ and noting that $\psi(y) = E[\sum_{i=1}^{y} X_i] = yE[X_1]$.

13) $E[|E[X|\mathcal{D}] - E[Y|\mathcal{D}]|^p] \leq E[|X - Y|^p]$, $p \geq 1$.

$$\begin{aligned}
E[|E[X|\mathcal{D}] - E[Y|\mathcal{D}]|^p] &= E[|E[X - Y]|\mathcal{D}]|^p] \quad \text{using linearity} \\
&\leq E[E[|X - Y|^p|\mathcal{D}]] \quad \text{using Jensen's inequality} \\
&= E[|X - Y|^p]
\end{aligned}$$

14) Let $\{X_n\}_{n=0}^{\infty}$ be a sequence of random variables and $p \geq 1$. If $\lim_{n\to\infty} E[|X - X_n|^p] = 0$, then $\lim_{n\to\infty} E[|E[X|\mathcal{D}] - E[X_n|\mathcal{D}]|^p] = 0$.

3 Continuous time stochastic processes.

A continuous time stochastic process is a random function defined on the time interval $[0, \infty)$, that is, for each $\omega \in \Omega$, $X(\cdot, \omega)$ is a real or vector-valued function (or more generally, E-valued for some complete, separable metric space E). Unless otherwise stated, we will assume that all stochastic processes are *cadlag*, that is, for each $\omega \epsilon \Omega$, $X(\cdot, \omega)$ is a right continuous function with left limits at each $t > 0$. $D_E[0, \infty)$ will denote the collection of cadlag E-valued functions on $[0, \infty)$. For each $\epsilon > 0$, a cadlag function has, at most, finitely many discontinuities of magnitude greater than ϵ in any compact time interval. (Otherwise, these discontinuities would have a right or left limit point, destroying the cadlag property). Consequently, a cadlag function can have, at most, a countable number of discontinuities.

If X is a cadlag process, then it is completely determined by the countable family of random variables, $\{X(t) : t \text{ rational}\}$.

It is possible to define a metric on $D_E[0, \infty)$ so that it becomes a complete, separable metric space. The distribution of an E-valued, cadlag process is then defined by $\mu_X(B) = P\{X(\cdot) \in B\}$ for $B \in \mathcal{B}(D_E[0, \infty))$. We will not discuss the metric at this time. For our present purposes it is enough to know the following.

Theorem 3.1 *Let X be an E-valued, cadlag process. Then μ_X on $D_E[0, \infty)$ is determined by its finite dimensional distributions $\{\mu_{t_1,t_2,\ldots,t_n} : 0 \leq t_1 \leq t_2 \leq \ldots t_n ; n \geq 0\}$ where $\mu_{t_1,t_2,\ldots,t_n}(\Gamma) = P\{(X(t_1), X(t_2), \ldots, X(t_n)) \in \Gamma\}$, $\Gamma \in \mathcal{B}(E^n)$.*

3.1 Examples.

1) *Standard Brownian Motion.* Recall that the density function of a normal random variable with expectation μ and variance σ^2 is given by

$$f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$$

For each integer $n > 0$, each selection of times $0 \leq t_1 < t_2 < \ldots < t_n$ and vector $x \in \mathbb{R}^n$, define the joint density function

$$f_{W(t_1),W(t_2),\ldots,W(t_n)}(x) = f_{0,t_1}(x_1) \cdot f_{0,t_2}(x_2 - x_1) \cdot \ldots \cdot f_{0,t_n}(x_n - x_{n-1}).$$

Note that the joint densities defined above are consistent in the sense that

$$f_{W(t'_1),\ldots,W(t'_{n-1})}(x'_1, \ldots, x'_{n-1}) = \int_{-\infty}^{\infty} f_{W(t_1),\ldots,W(t_n))}(x_1, \ldots, x_n) dx_i$$

where $(t'_1, \ldots, t'_{n-1})$ is obtained from $(t_1, \ldots, t_n)$ by deleting the ith entry. The Kolmogorov Consistency Theorem, assures the existence of a stochastic process with these finite dimensional distributions. An additional argument is needed to show that the process has cadlag (in fact continuous) sample paths.

2) *Poisson Process.* Again, we specify the Poisson process with parameter λ, by specifying its finite dimensional distributions. Let $h(\mu, k) = \exp\{-\mu\}\frac{\mu^k}{k!}$, that is, the Poisson($\mu$)

probability of k. For $t_1 < t_2 < \cdots < t_n$. Define

$$
\begin{aligned}
&P\{X(t_1) = k_1, X(t_2) = k_2, \ldots, X(t_n) = k_n\} \\
&\quad = \begin{cases} h(\lambda t_1, k_1) \cdot h(\lambda(t_2 - t_1), (k_2 - k_1)) \cdot \ldots \\ \qquad \cdot h(\lambda(t_n - t_{n-1}), (k_n - k_{n-1})) & \text{if } k_1 \leq k_2 \leq \ldots \leq k_n \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

3.2 Filtrations.

Let $\sigma(X(s) : s \leq t)$ denote the smallest σ-algebra such that $X(s)$ is $\sigma(X(s) : s \leq t)$-measurable for all $s \leq t$.

A collection of σ-algebras $\{\mathcal{F}_t\}$, satisfying

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$$

for all $s \leq t$ is called a *filtration*. $\mathcal{F}_t$ is interpreted as corresponding to the information available at time t (the amount of information increasing as time progresses). A stochastic process X is *adapted* to a filtration $\{\mathcal{F}_t\}$ if $X(t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$.

An E-valued stochastic process X adapted to $\{\mathcal{F}_t\}$ is a *Markov process* with respect to $\{\mathcal{F}_t\}$ if

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

for all $t, s \geq 0$ and $f \in B(E)$, the bounded, measurable functions on E.

A real-valued stochastic process X adapted to $\{\mathcal{F}_t\}$ is a *martingale* with respect to $\{\mathcal{F}_t\}$ if

$$E[X(t+s)|\mathcal{F}_t] = X(t) \tag{3.1}$$

for all $t, s \geq 0$.

Proposition 3.2 *Standard Brownian motion, W, is both a martingale and a Markov process.*

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

$$
\begin{aligned}
E[W(t+s)|\mathcal{F}_t] &= E[W(t+s) - W(t) + W(t)|\mathcal{F}_t] \\
&= E[W(t+s) - W(t)|\mathcal{F}_t] + E[W(t)|\mathcal{F}_t] \\
&= E[W(t+s) - W(t)] + E[W(t)|\mathcal{F}_t] \\
&= E[W(t)|\mathcal{F}_t] \\
&= W(t)
\end{aligned}
$$

so W is an $\{\mathcal{F}_t\}$-martingale. Similarly W is a Markov process. Define $T(s)f(x) = E[f(x + W(s))]$, and note that

$$
\begin{aligned}
E[f(W(t+s))|\mathcal{F}_t] &= E[f(W(t+s) - W(t) + W(t))|\mathcal{F}_t] \\
&= T(s)f(W(t)) \\
&= E[f(W(t+s))|W(t)]
\end{aligned}
$$

□

3.3 Stopping times.

A random variable τ with values in $[0,\infty]$ is an $\{\mathcal{F}_t\}$-*stopping time* if

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \forall\, t \geq 0.$$

Let X be a cadlag stochastic process that is $\{\mathcal{F}_t\}$-adapted. Then

$$\tau_\alpha = \inf\{t : X(t) \text{ or } X(t-) \geq \alpha\}$$

is a stopping time. In general, for $B \in \mathcal{B}(\mathbb{R})$, $\tau_B = \inf\{t : X(t) \in B\}$ is not a stopping time; however, if $(\Omega, \mathcal{F}, P)$ is complete and the filtration $\{\mathcal{F}_t\}$ is *complete* in the sense that $\mathcal{F}_0$ contains all events of probability zero and is *right continuous* in the sense that $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$, then for any $B \in \mathcal{B}(\mathbb{R})$, τ_B is a stopping time.

If τ, $\tau_1, \tau_2 \ldots$ are stopping times and $c \geq 0$ is a constant, then

1) $\tau_1 \vee \tau_2$ and $\tau_1 \wedge \tau_2$ are stopping times.

2) $\tau + c$, $\tau \wedge c$, and $\tau \vee c$ are stopping times.

3) $\sup_k \tau_k$ is a stopping time.

4) If $\{\mathcal{F}_t\}$ is right continuous, then $\inf_k \tau_k$, $\liminf_{k\to\infty} \tau_k$, and $\limsup_{k\to\infty} \tau_k$ are stopping times.

Lemma 3.3 *Let τ be a stopping time and for $n = 1, 2, \ldots$, define*

$$\tau_n = \frac{k+1}{2^n}, \quad \text{if } \frac{k}{2^n} \leq \tau < \frac{k+1}{2^n}, \; k = 0, 1, \ldots.$$

Then $\{\tau_n\}$ is a decreasing sequence of stopping times converging to τ.

Proof. Observe that

$$\{\tau_n \leq t\} = \{\tau_n \leq \frac{[2^n t]}{2^n}\} = \{\tau < \frac{[2^n t]}{2^n}\} \in \mathcal{F}_t.$$

□

For a stopping time τ, define

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t \geq 0\}.$$

Then $\mathcal{F}_\tau$ is a σ-algebra and is interpreted as representing the information available to an observer at the random time τ.

Lemma 3.4 *If τ_1 and τ_2 are stopping times and $\tau_1 \leq \tau_2$, then $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.*

Lemma 3.5 *If X is cadlag and $\{\mathcal{F}_t\}$-adapted and τ is a stopping time, then $X(\tau)$ is $\mathcal{F}_\tau$-measurable and $X(\tau \wedge t)$ is $\{\mathcal{F}_t\}$-adapted.*

3.4 Brownian motion

A process X has *independent increments* if for each choice of $0 \le t_1 < t_2 < \cdots < t_m$, $X(t_{k+1}) - X(t_k)$, $k = 1, \ldots, m-1$, are independent. X is a *Brownian motion* if it has continuous sample paths and independent, Gaussian-distributed increments. It follows that the distribution of a Brownian motion is completely determined by $m_X(t) = E[X(t)]$ and $a_X(t) = Var(X(t))$. X is *standard* if $m_X \equiv 0$ and $a_X(t) = t$. Note that the independence of the increments implies

$$Var(X(t+s) - X(t)) = a_X(t+s) - a_X(t),$$

so a_X must be nondecreasing. If X is standard

$$Var(X(t+s)) - X(t)) = s.$$

Consequently, standard Brownian motion has stationary increments. Ordinarily, we will denote standard Brownian motion by W.

If a_X is continuous and nondecreasing and m_X is continuous, then

$$X(t) = W(a_X(t)) + m_X(t)$$

is a Brownian motion with

$$E[X(t)] = m_X(t), \quad Var(X(t)) = a_X(t).$$

3.5 Poisson process

A Poisson process is a model for a series of 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. We assume that N is a counting process, that is, the observations come one at a time, so N is constant except for jumps of $+1$. 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.

0) The observations occur one at a time.

1) Numbers of observations in disjoint time intervals are independent random variables, that is, N has independent increments.

2) The distribution of $N(t+a) - N(t)$ does not depend on t.

Theorem 3.6 *Under assumptions 0), 1), and 2), there is a constant λ such that $N(s)-N(t)$ is Poisson distributed with parameter $\lambda(s-t)$, that is,*

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

If Theorem 3.6 holds, then we refer to N as a Poisson process with parameter λ. If $\lambda = 1$, we will call N the unit Poisson process.

More generally, if (0) and (1) hold and $\Lambda(t) = E[N(t)]$ is continuous and $\Lambda(0) = 0$, then

$$N(t) = Y(\Lambda(t)),$$

where Y is a unit Poisson process.

Let N be a Poisson process with parameter λ, and let S_k be the time of the kth 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

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

The Poisson process can also be viewed as the renewal process based on a sequence of exponentially distributed random variables.

Theorem 3.7 *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 λ.*

4 Martingales.

A stochastic process X adapted to a filtration $\{\mathcal{F}_t\}$ is a *martingale* with respect to $\{\mathcal{F}_t\}$ if

$$E[X(t+s)|\mathcal{F}_t] = X(t) \tag{4.1}$$

for all $t, s \geq 0$. It is a *submartingale* if

$$E[X(t+s)|\mathcal{F}_t] \geq X(t) \tag{4.2}$$

and a *supermartingale* if

$$E[X(t+s)|\mathcal{F}_t] \leq X(t). \tag{4.3}$$

4.1 Optional sampling theorem and Doob's inequalities.

Theorem 4.1 *(Optional sampling theorem.) Let X be a martingale and τ_1, τ_2 be stopping times. Then for every $t \geq 0$*

$$E[X(t \wedge \tau_2)|\mathcal{F}_{\tau_1}] = X(t \wedge \tau_1 \wedge \tau_2).$$

If $\tau_2 < \infty$ a.s., $E[|X(\tau_2)|] < \infty$ and $\lim_{t\to\infty} E[|X(t)|I_{\{\tau_2 > t\}}] = 0$, then

$$E[X(\tau_2)|\mathcal{F}_{\tau_1}] = X(\tau_1 \wedge \tau_2)\ .$$

The same results hold for sub and supermartingales with $=$ replaced by $\geq$ (submartingales) and $\leq$ (supermartingales).

Proof. See, for example, Ethier and Kurtz (1986), Theorem 2.2.13. □

Theorem 4.2 *(Doob's inequalities.) If X is a non-negative sub-martingale, then*

$$P\{\sup_{s\leq t} X(s) \geq x\} \leq \frac{E[X(t)]}{x}$$

and for $\alpha > 1$

$$E[\sup_{s\leq t} X(s)^\alpha] \leq (\alpha/\alpha - 1)^\alpha E[X(t)^\alpha].$$

Proof. Let $\tau_x = \inf\{t : X(t) \geq x\}$ and set $\tau_2 = t$ and $\tau_1 = \tau_x$. Then from the optional sampling theorem we have that

$$E[X(t)|\mathcal{F}_{\tau_x}] \geq X(t \wedge \tau_x) \geq I_{\{\tau_x \leq t\}} X(\tau_x) \geq x I_{\{\tau_x \leq t\}} \text{ a.s.}$$

so we have that

$$E[X(t)] \geq xP\{\tau_x \leq t\} = xP\{\sup_{s\leq t} X(s) \geq x\}$$

See Ethier and Kurtz, Proposition 2.2.16 for the second inequality. □

Lemma 4.3 *If M is a martingale and φ is convex with $E[|\varphi(M(t))|] < \infty$, then*

$$X(t) \equiv \varphi(M(t))$$

is a sub-martingale.

Proof.

$$E[\varphi(M(t+s))|\mathcal{F}_t] \geq \varphi(E[M(t+s)|\mathcal{F}_t])$$

by Jensen's inequality. □

From the above lemma, it follows that if M is a martingale then

$$P\{\sup_{s\leq t}|M(s)| \geq x\} \leq \frac{E[|M(t)|]}{x} \tag{4.4}$$

and

$$E[\sup_{s\leq t}|M(s)|^2] \leq 4E[M(t)^2]. \tag{4.5}$$

4.2 Local martingales.

The concept of a *local martingale* plays a major role in the development of stochastic integration. M is a *local martingale* if there exists a sequence of stopping times $\{\tau_n\}$ such that $\lim_{n\to}\tau_n = \infty$ a.s. and for each n, $M^{\tau_n} \equiv M(\cdot \wedge \tau_n)$ is a martingale.

The *total variation* of Y up to time t is defined as

$$T_t(Y) \equiv \sup \sum |Y(t_{i+1}) - Y(t_i)|$$

where the supremum is over all partitions of the interval $[0,t]$. Y is an *FV-process* if $T_t(Y) < \infty$ for each $t > 0$.

Theorem 4.4 *(Fundamental Theorem of Local Martingales.) Let M be a local martingale, and let $\delta > 0$. Then there exist local martingales $\tilde{M}$ and A satisfying $M = \tilde{M} + A$ such that A is FV and the discontinuities of $\tilde{M}$ are bounded by δ.*

Remark 4.5 *One consequence of this theorem is that any local martingale can be decomposed into an FV process and a local square integrable martingale. Specifically, if $\gamma_c = \inf\{t : |\tilde{M}(t)| \geq c\}$, then $\tilde{M}(\cdot\wedge\gamma_c)$ is a square integrable martingale. (Note that $|\tilde{M}(\cdot\wedge\gamma_c)| \leq c+\delta$.)*

Proof. See Protter (1990), Theorem III.13. □

4.3 Quadratic variation.

The *quadratic variation* of a process Y is defined as

$$[Y]_t = \lim_{\max|t_{i+1}-t_i|\to 0} \sum (Y(t_{i+1}) - Y(t_i))^2$$

where convergence is in probability; that is, for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every partition $\{t_i\}$ of the interval $[0,t]$ satisfying $\max|t_{i+1}-t_i| \leq \delta$ we have

$$P\{|[Y]_t - \sum (Y(t_{i+1}) - Y(t_i))^2| \geq \epsilon\} \leq \epsilon.$$

If Y is FV, then $[Y]_t = \sum_{s\leq t}(Y(s) - Y(s-))^2 = \sum_{s\leq t} \Delta Y(s)^2$ where the summation can be taken over the points of discontinuity only and $\Delta Y(s) \equiv Y(s) - Y(s-)$ is the jump in Y at time s. Note that for any partition of $[0,t]$

$$\sum (Y(t_{i+1}) - Y(t_i))^2 - \sum_{|Y(t_{i+1})-Y(t_i)|>\epsilon} (Y(t_{i+1}) - Y(t_i))^2 \leq \epsilon T_t(Y).$$

Proposition 4.6 *(i) If M is a local martingale, then $[M]_t$ exists and is right continous.*

(ii) If M is a square integrable martingale, then the limit

$$\lim_{\max|t_{i+1}-t_i|\to 0} \sum (M(t_{i+1}) - M(t_i))^2$$

exists in L_1 and $E[(M(t)^2] = E[[M]_t]$.

Proof. See, for example, Ethier and Kurtz (1986), Proposition 2.3.4. □

Let M be a square integrable martingale with M(0)=0. Write $M(t) = \sum_{i=0}^{m-1} M(t_{i+1}) - M(t_i)$ where $0 = t_0 < \ldots < t_m = t$. Then

$$\begin{aligned} E[M(t)^2] &= E[(\sum_{i=0}^{m-1} M(t_{i+1}) - M(t_i))^2] \qquad (4.6) \\ &= E[\sum_{i=0}^{m-1} (M(t_{i+1}) - M(t_i))^2 \\ &\qquad + \sum_{i\neq j} (M(t_{i+1}) - M(t_i))(M(t_{j+1}) - M(t_j))]. \end{aligned}$$

For $t_i < t_{i+1} \leq t_j < t_{j+1}$.

$$\begin{aligned} &E[(M(t_{i+1}) - M(t_i))(M(t_{j+1}) - M(t_j))] \qquad (4.7) \\ &= E[E[(M(t_{i+1}) - M(t_i))(M(t_{j+1}) - M(t_j))|\mathcal{F}_{t_j}]] \\ &= E[(M(t_{i+1}) - M(t_i))(E[M(t_{j+1})|\mathcal{F}_{t_j}] - M(t_j))] \\ &= 0, \end{aligned}$$

and thus the expectation of the second sum in (4.6) vanishes. By the L_1 convergence in Proposition 4.6

$$E[M(t)^2] = E[\sum_{i=0}^{m-1} (M(t_{i+1}) - M(t_i))^2] = E[[M]_t].$$

Example 4.7 *If $M(t) = N(t) - \lambda t$ where $N(t)$ is a Poisson process with parameter λ, then $[M]_t = N(t)$, and since $M(t)$ is square integrable, the limit exists in L_1.*

Example 4.8 *For standard Brownian motion W, $[W]_t = t$. To check this identity, apply the law of large numbers to*

$$\sum_{k=1}^{[nt]} (W(\frac{k}{n}) - W(\frac{k-1}{n}))^2.$$

Proposition 4.9 *If M is a square integrable $\{\mathcal{F}_t\}$-martingale, Then $M(t)^2 - [M]_t$ is an $\{\mathcal{F}_t\}$-martingale.*

Remark 4.10 *For example, if W is standard Brownian motion, then $W(t)^2 - t$ is a martingale.*

Proof. The conclusion follows from part (ii) of the previous proposition. For $t, s \geq 0$, let $\{u_i\}$ be a partition of $(0, t+s]$ with $0 = u_0 < u_1 < ... < u_m = t < u_{m+1} < ... < u_n = t+s$ Then

$$\begin{aligned}
&E[M(t+s)^2|\mathcal{F}_t] \\
&\quad = E[(M(t+s) - M(t))^2|\mathcal{F}_t] + M(t)^2 \\
&\quad = E[(\sum_{i=m}^{n-1} M(u_{i+1}) - M(u_i))^2|\mathcal{F}_t] + M(t)^2 \\
&\quad = E[\sum_{i=m}^{n-1} (M(u_{i+1}) - M(u_i))^2|\mathcal{F}_t] + M(t)^2 \\
&\quad = E[[M]_{t+s} - [M]_t|\mathcal{F}_t] + M(t)^2
\end{aligned}$$

where the first equality follows since, as in (4.7), the conditional expectation of the cross product term is zero and the last equality follows from the L_1 convergence in Proposition 4.6. □

4.4 Martingale convergence theorem.

Theorem 4.11 *(Martingale convergence theorem.) Let X be a submartingale satisfying $\sup_t E[|X(t)|] < \infty$. Then $\lim_{t\to\infty} X(t)$ exists a.s.*

Proof. See, for example, Durrett (1991), Theorem 4.2.10. □

5 Stochastic integrals.

Let X and Y be cadlag processes, and let $\{t_i\}$ denote a partition of the interval $[0,t]$. Then we define the stochastic integral of X with respect to Y by

$$\int_0^t X(s-)dY(s) \equiv \lim \sum X(t_i)(Y(t_{i+1}) - Y(t_i)) \tag{5.1}$$

where the limit is in probability and is taken as $\max |t_{i+1} - t_i| \to 0$. For example, let $W(t)$ be a standard Brownian motion. Then

$$\begin{aligned}
\int_0^t W(s)dW(s) &= \lim \sum W(t_i)(W(t_{i+1}) - W(t_i)) \qquad (5.2)\\
&= \lim \sum (W(t_i)W(t_{i+1}) - \frac{1}{2}W(t_{i+1})^2 - \frac{1}{2}W(t_i)^2) \\
&\qquad + \sum (\frac{1}{2}W(t_{i+1})^2 - \frac{1}{2}W(t_i)^2) \\
&= \frac{1}{2}W(t)^2 - \lim \frac{1}{2}\sum (W(t_{i+1}) - W(t_i))^2 \\
&= \frac{1}{2}W(t)^2 - \frac{1}{2}t.
\end{aligned}$$

This example illustrates the significance of the use of the left end point of $[t_i, t_{i+1}]$ in the evaluation of the integrand. If we replace t_i by t_{i+1} in (5.2), we obtain

$$\begin{aligned}
&\lim \sum W(t_{i+1})(W(t_{i+1}) - W(t_i)) \\
&\quad = \lim \sum (- W(t_i)W(t_{i+1}) + \frac{1}{2}W(t_{i+1})^2 + \frac{1}{2}W(t_i)^2) \\
&\qquad\qquad + \sum (\frac{1}{2}W(t_{i+1})^2 - \frac{1}{2}W(t_i)^2) \\
&\quad = \frac{1}{2}W(t)^2 + \lim \frac{1}{2}\sum (W(t_{i+1}) - W(t_i))^2 \\
&\quad = \frac{1}{2}W(t)^2 + \frac{1}{2}t.
\end{aligned}$$

5.1 Definition of the stochastic integral.

Throughout, we will be interested in the stochastic integral as a stochastic process. With this interest in mind, we will use a slightly different definition of the stochastic integral than that given in (5.1). For any partition $\{t_i\}$ of $[0,\infty)$, $0 = t_0 < t_1 < t_2 < \ldots$, and any cadlag x and y, define

$$S(t, \{t_i\}, x, y) = \sum x(t_i)(y(t \wedge t_{i+1}) - y(t \wedge t_i)).$$

For stochastic processes X and Y, define $Z = \int X_- dY$ if for each $T > 0$ and each $\epsilon > 0$, there exists a $\delta > 0$ such that

$$P\{\sup_{t \le T} |Z(t) - S(t, \{t_i\}, X, Y)| \ge \epsilon\} \le \epsilon$$

for all partitions $\{t_i\}$ satisfying $\max|t_{i+1} - t_i| \leq \delta$.

If X is piecewise constant, that is, for some collection of random variables $\{\xi_i\}$ and random variables $\{\tau_i\}$ satisfying $0 = \tau_0 < \tau_1 < \cdots$,

$$X = \sum \xi_i I_{[\tau_i, \tau_{i+1})} ,$$

then

$$\begin{aligned} \int_0^t X(s-)dY(s) &= \sum \xi_i (Y(t \wedge \tau_{i+1}) - Y(t \wedge \tau_i)) \\ &= \sum X(\tau_i)(Y(t \wedge \tau_{i+1}) - Y(t \wedge \tau_i)) . \end{aligned}$$

Our first problem is to identify more general conditions on X and Y under which $\int X_- dY$ will exist.

5.2 Conditions for existence.

The first set of conditions we will consider require that the integrator Y be of finite variation. The *total variation* of Y up to time t is defined as

$$T_t(Y) \equiv \sup \sum |Y(t_{i+1}) - Y(t_i)|$$

where the supremum is over all partitions of the interval $[0, t]$.

Proposition 5.1 *$T_t(f) < \infty$ for each $t > 0$ if and only if there exist monotone increasing functions f_1, f_2 such that $f = f_1 - f_2$. If $T_t(f) < \infty$, then f_1 and f_2 can be selected so that $T_t(f) = f_1 + f_2$. If f is cadlag, then $T_t(f)$ is cadlag.*

Proof. Note that

$$T_t(f) - f(t) = \sup \sum (|f(t_{i+1}) - f(t_i)| - (f(t_{i+1}) - f(t_i)))$$

is an increasing function of t, as is $T_t(f) + f(t)$. □

Theorem 5.2 *If Y is of finite variation then $\int X_- dY$ exists for all X, $\int X_- dY$ is cadlag, and if Y is continuous, $\int X_- dY$ is continuous. (Recall that we are assuming throughout that X is cadlag.)*

Proof. Let $\{t_i\}, \{s_i\}$ be partitions. Let $\{u_i\}$ be a refinement of both. Then there exist k_i, l_i, k_i', l_i' such that

$$\begin{aligned} Y(t_{i+1}) - Y(t_i) &= \sum_{j=k_i}^{l_i} Y(u_{j+1}) - Y(u_j) \\ Y(s_{i+1}) - Y(s_i) &= \sum_{j=k_i'}^{l_i'} Y(u_{j+1}) - Y(u_j). \end{aligned}$$

Define

$$t(u) = t_i, \quad t_i \le u < t_{i+1} \qquad s(u) = s_i, \quad s_i \le u < s_{i+1} \tag{5.3}$$

so that

$$\begin{aligned} &|S(t,\{t_i\},X,Y) - S(t,\{s_i\},X,Y)| \\ &\quad = |\sum X(t(u_i))(Y(u_{i+1}\wedge t) - Y(u_i\wedge t)) \\ &\qquad\qquad - \sum X(s(u_i))(Y(u_{i+1}\wedge t) - Y(u_i\wedge t))| \\ &\quad \le \sum |X(t(u_i)) - X(s(u_i))||Y(u_{i+1}\wedge t) - Y(u_i\wedge t)|. \end{aligned} \tag{5.4}$$

But there is a measure μ_Y such that $T_t(Y) = \mu_Y(0,t]$. Since $|Y(b) - Y(a)| \le \mu_Y(a,b]$, the right side of (5.4) is less than

$$\begin{aligned} \sum |X(t(u_i)) - X(s(u_i))|\mu_Y(u_i\wedge t, u_{i+1}\wedge t] &= \sum \int_{(u_i\wedge t, u_{i+1}\wedge t]} |X(t(u-)) - X(s(u-))|\mu_Y(du) \\ &= \int_{(0,t]} |X(t(u-)) - X(s(u-))|\mu_Y(du). \end{aligned}$$

But

$$\lim |X(t(u-)) - X(s(u-))| = 0,$$

so

$$\int_{(0,t]} |X(t(u-)) - X(s(u-))|\mu_Y(du) \to 0 \tag{5.5}$$

by the bounded convergence theorem. Since the integral in (5.5) is monotone in t, the convergence is uniform on bounded time intervals. □

Recall that the quadratic variation of a process is defined by

$$[Y]_t = \lim_{\max|t_{i+1}-t_i|\to 0} \sum (Y(t_{i+1}) - Y(t_i))^2,$$

where convergence is in probability. For example, if Y is a Poisson process, then $[Y]_t = Y(t)$ and for standard Brownian motion, $[W]_t = t$.

Note that

$$\sum (Y(t_{i+1}) - Y(t_i))^2 = Y(t)^2 - Y(0)^2 - 2\sum Y(t_i)(Y(t_{i+1}) - Y(t_i))$$

so that

$$[Y]_t = Y(t)^2 - Y(0)^2 - 2\int_0^t Y(s-)dY(s)$$

and $[Y]_t$ exists if and only if $\int Y_- dY$ exists. By Proposition 4.6, $[M]_t$ exists for any local martingale and by Proposition 4.9, for a square integrable martingale $M(t)^2 - [M]_t$ is a martingale.

If M is a square integrable martingale and X is bounded (by a constant) and adapted, then for any partition $\{t_i\}$,

$$Y(t) = S(t,\{t_i\},X,M) = \sum X(t_i)(M(t\wedge t_{i+1}) - M(t\wedge t_i))$$

is a square-integrable martingale. (In fact, each summand is a square-integrable martingale.) This observation is the basis for the following theorem.

Theorem 5.3 *Suppose M is a square integrable $\{\mathcal{F}_t\}$- martingale and X is cadlag and $\{\mathcal{F}_t\}$-adapted. Then $\int X_- dM$ exists.*

Proof. Claim: If we can prove

$$\int_0^t X(s-)dM(s) = \lim \sum X(t_i)(M(t_{i+1}\wedge t) - M(t_i \wedge t))$$

for every bounded cadlag X, we are done. To verify this claim, let $X_k(t) = (X(t)\wedge k)\vee(-k)$ and suppose

$$\lim \sum X_k(t_i)(M(t_{i+1}\wedge t) - M(t_i\wedge t)) = \int_0^t X_k(s-)dM(s)$$

exists. Since $\int_0^t X(s-)dM(s) = \int_0^t X_k(s-)dM(s)$ on $\{\sup_{s\le t}|X(s)| \le k\}$, the assertion is clear.

Now suppose $|X(t)| \le C$. Since M is a square integrable martingale and $|X|$ is bounded, it follows that for any partition $\{t_i\}$, $S(t, \{t_i\}, X, M)$ is a square integrable martingale. (As noted abover, for each i, $X(t_i)(M(t_{i+1}\wedge t) - M(t_i\wedge t))$ is a square-integrable martingale.) For two partitions $\{t_i\}$ and $\{s_i\}$, define $\{u_i\}$, $t(u)$, and $s(u)$ as in the proof of Theorem 5.2. Recall that $t(u_i), s(u_i) \le u_i$, so $X(t(u))$ and $X(s(u))$ are $\{\mathcal{F}_u\}$-adapted. Then by Doob's inequality and the properties of martingales,

$$\begin{aligned}
&E[\sup_{t\le T}(S(t,\{t_i\},X,M) - S(t,\{s_i\},X,M))^2] \qquad (5.6)\\
&\quad\le 4E[(S(T,\{t_i\},X,M) - S(T,\{s_i\},X,M))^2]\\
&\quad= 4E[(\sum(X(t(u_i)) - X(s(u_i))(M(u_{i+1}\wedge T) - M(u_i\wedge T)))^2]\\
&\quad= 4E[\sum(X(t(u_i)) - X(s(u_i))^2(M(u_{i+1}\wedge T) - M(u_i\wedge T))^2]\\
&\quad= 4E[\sum(X(t(u_i)) - X(s(u_i))^2([M]_{u_{i+1}\wedge T} - [M]_{u_i\wedge T})].
\end{aligned}$$

Note that $[M]$ is nondecreasing and so determines a measure by $\mu_{[M]}(0,t] = [M]_t$, and it follows that

$$\begin{aligned}
&E[\sum(X(t(u_i)) - X(s(u_i)))^2([M]_{u_{i+1}} - [M]_{u_i})] \qquad (5.7)\\
&\quad= E[\int_{(0,t]}(X(t(u)) - X(s(u)))^2\mu_{[M]}(du)],
\end{aligned}$$

since $X(t(u))$ and $X(s(u))$ are constant between u_i and u_{i+1}. Now

$$|\int_{(0,t]}(X(t(u)) - X(s(u)))^2\mu_{[M]}(du)| \le 4C^2\mu_{[M]}(0,t]\ ,$$

so by the fact that X is cadlag and the dominated convergence theorem, the right side of (5.7) goes to zero as $\max|t_{i+1} - t_i| \to 0$ and $\max|s_{i+1} - s_i| \to 0$. Consequently, $\int_0^t X(s-)dM(s)$ exists by the completeness of L^2. □

Corollary 5.4 *If M is a square integrable martingale and X is adapted, then $\int X_- dM$ is cadlag. If, in addition, M is continuous, then $\int X_- dM$ is continuous. If $|X| \leq C$ for some constant $C > 0$, then $\int X_- dM$ is a square integrable martingale.*

Proposition 5.5 *Suppose M is a square integrable martingale and*

$$E\left[\int_0^t X(s-)^2 d[M]_s\right] < \infty.$$

Then $\int X_- dM$ is a square integrable martingale with

$$E\left[(\int_0^t X(s-)dM(s))^2\right] = E\left[\int_0^t X(s-)^2 d[M]_s\right]. \tag{5.8}$$

Remark 5.6 *If W is standard Brownian motion, the identity becomes*

$$E\left[\left(\int_0^t X(s-)dW(s)\right)^2\right] = E\left[\int_0^t X^2(s)ds\right].$$

Proof. Suppose $X(t) = \sum \xi_i I_{[t_i,t_{i+1})}$ is a simple function. Then

$$\begin{aligned} E\left[\left(\int_0^t X(s-)dM(s)\right)^2\right] &= E\left[\sum X(t_i)^2(M(t_{i+1}) - M(t_i))^2\right] \\ &= E\left[\sum X(t_i)^2\left([M]_{t_{i+1}} - [M]_{t_i}\right)\right] \\ &= E\left[\int_0^t X^2(s-)d[M]_s\right]. \end{aligned}$$

Now let X be bounded, with $|X(t)| \leq C$, and for a sequence of partitions $\{t_i^n\}$ with $\lim_{n\to\infty}\sup_i |t_{i+1}^n - t_i^n| = 0$, define

$$X_n(t) = X(t_i^n), \text{ for } t_i^n \leq t < t_{i+1}^n.$$

Then by the argument in the proof of Theorem 5.3, we have

$$\begin{aligned} \int_0^t X_n(s-)dM(s) &= \sum X(t_i^n)\left(M(t \wedge t_{i+1}^n) - M(t \wedge t_i^n)\right) \\ &\to \int_0^t X(s-)dM(s), \end{aligned}$$

where the convergence is in L^2. Since $\int X_{n-} dM$ is a martingale, it follows that $\int X_- dM$ is a martingale, and

$$\begin{aligned} E\left[\left(\int_0^t X(s-)dM(s)\right)^2\right] &= \lim_{n\to\infty} E\left[\left(\int_0^t X_n(s-)dM(s)\right)^2\right] \\ &= \lim_{n\to\infty} E\left[\int_0^t X_n^2(s-)d[M]_s\right] \\ &= E\left[\int_0^t X^2(s-)d[M]_s\right], \end{aligned}$$

The last equality holds by the dominated convergence theorem. This statement establishes the theorem for bounded X.

Finally, for arbitrary cadlag and adapted X, define $X_k(t) = (k \wedge X(t)) \vee (-k)$. Then

$$\int_0^t X_k(s-)dM(s) \to \int_0^t X(s-)dM(s)$$

in probability, and by Fatou's lemma,

$$\liminf_{k\to\infty} E\left[\left(\int_0^t X_k(s-)dM(s)\right)^2\right] \geq E\left[\left(\int_0^t X(s-)dM(s)\right)^2\right].$$

But

$$\begin{aligned} \lim_{k\to\infty} E\left[\left(\int_0^t X_k(s-)dM(s)\right)^2\right] &= \lim_{k\to\infty} E\left[\int_0^t X_k^2(s-)d[M]_s\right] \qquad (5.9) \\ &= \lim_{k\to\infty} E\left[\int_0^t X^2(s-)\wedge k^2 d[M]_s\right] \\ &= E\left[\int_0^t X^2(s-)d[M]_s\right] < \infty, \end{aligned}$$

so

$$E\left[\int_0^t X^2(s-)d[M]_s\right] \geq E\left[\left(\int_0^t X(s-)dM(s)\right)^2\right].$$

Since (5.8) holds for bounded X,

$$\begin{aligned} E&\left[\left(\int_0^t X_k(s-)dM(s) - \int_0^t X_j(s-)dM(s)\right)^2\right] \qquad (5.10) \\ &= E\left[\left(\int_0^t (X_k(s-) - X_j(s-))dM(s)\right)^2\right] \\ &= E\left[\int_0^t |X_k(s-) - X_j(s-)|^2 d[M]_s\right] \end{aligned}$$

Since

$$|X_k(s) - X_j(s)|^2 \leq 4X(s)^2,$$

the dominated convergence theorem implies the right side of (5.10) converges to zero as $j, k \to \infty$. Consequently,

$$\int_0^t X_k(s-)dM(s) \to \int_0^t X(s-)dM(s)$$

in L_2, and the left side of (5.9) converges to $E[(\int_0^t X(s-)dM(s))^2]$ giving (5.8). $\square$

If $\int_0^t X(s-)dY_1(s)$ and $\int_0^t X(s-)dY_2(s)$ exist, then $\int_0^t X(s-)d(Y_1(s) + Y_2(s))$ exists and is given by the sum of the other integrals.

Corollary 5.7 *If $Y = M+V$ where M is a $\{\mathcal{F}_t\}$-local martingale and V is an $\{\mathcal{F}_t\}$-adapted finite variation process, then $\int X_- dY$ exists for all cadlag, adapted X, $\int X_- dY$ is cadlag, and if Y is continuous, $\int X_- dY$ is continuous.*

Proof. If M is a local square integrable martingale, then there exists a sequence of stopping times $\{\tau_n\}$ such that M^{τ_n} defined by $M^{\tau_n}(t) = M(t \wedge \tau_n)$ is a square-integrable martingale. But for $t < \tau_n$,

$$\int_0^t X(s-)dM(s) = \int_0^t X(s-)dM^{\tau_n}(s),$$

and hence $\int X_- dM$ exists. Linearity gives existence for any Y that is the sum of a local square integrable martingale and an adapted FV process. But Theorem 4.4 states that any local martingale is the sum of a local square integrable martingale and an adapted FV process, so the corollary follows. □

5.3 Semimartingales.

With Corollary 5.7 in mind, we define Y to be an $\{\mathcal{F}_t\}$-*semimartingale* if and only if $Y = M+V$, where M is a local martingale with respect to $\{\mathcal{F}_t\}$ and V is an $\{\mathcal{F}_t\}$-adapted finite variation process. By Theorem 4.4 we can always select M and V so that M is local square integrable. In particular, we can take M to have discontinuities uniformly bounded by a constant. If the discontinuities of Y are already uniformly bounded by a constant, it will be useful to know that the decomposition preserves this property for both M and V.

Lemma 5.8 *Let Y be a semimartingale and let $\delta \geq \sup |\Delta Y(s)|$. Then there exists a local square integrable martingale M and a finite variation process V such that $Y = M + V$,*

$$\sup |M(s) - M(s-)| \leq \delta$$

$$\sup |V(s) - V(s-)| \leq 2\delta.$$

Proof. Let $Y = \tilde{M} + \tilde{V}$ be a decompostion of Y into a local martingale and an FV process. By Theorem 4.4, there exists a local martingale M with discontinuities bounded by δ and an FV process A such that $\tilde{M} = M + A$. Defining $V = A + \tilde{V} = Y - M$, we see that the discontinuities of V are bounded by 2δ. □

The class of semimartingales is closed under a variety of operations. Clearly, it is linear. It is also easy to see that a stochastic integral with a semimartingale integrator is a semimartingale, since if we write

$$\int_0^t X(s-)dY(s) = \int_0^t X(s-)dM(s) + \int_0^t X(s-)dV(s),$$

then the first term on the right is a local square integrable martingale whenever M, is and the second term on the right is a finite variation process whenever V is.

Lemma 5.9 *If V is of finite variation, then $Z_2(t) = \int_0^t X(s-)dV(s)$ is of finite variation.*

Proof. For any partition $\{t_i\}$ of $[a,b]$,

$$\begin{aligned}|Z_2(b)-Z_2(a)| &= \lim\left|\sum X(t_i)\left(V(t_{i+1})-V(t_i)\right)\right| \\ &\leq \sup_{a\leq s<b}|X(s)|\lim\sum|V(t_{i+1})-V(t_i)| \\ &\leq \sup_{a\leq s<b}|X(s)|\left(T_b(V)-T_a(V)\right),\end{aligned}$$

and hence

$$T_t(Z)\leq \sup_{0\leq s<t}|X(s)|T_t(V). \tag{5.11}$$

□

Lemma 5.10 *Let M be a local square integrable martingale, and let X be adapted. Then $Z_1(t)=\int_0^t X(s-)dM(s)$ is a local square integrable martingale.*

Proof. There exist $\tau_1\leq\tau_2\leq\cdots$, $\tau_n\to\infty$, such that $M^{\tau_n}=M(\cdot\wedge\tau_n)$ is a square integrable martingale. Define

$$\gamma_n=\inf\left\{t:|X(t)|\text{ or }|X(t-)|\geq n\right\},$$

and note that $\lim_{n\to\infty}\gamma_n=\infty$. Then setting $X_n(t)=(X(t)\wedge n)\vee(-n)$,

$$Z_1(t\wedge\tau_n\wedge\gamma_n)=\int_0^{t\wedge\gamma_n}X_n(s-)dM^{\tau_n}(s)$$

is a square integrable martingale, and hence Z_1 is a local square integrable martingale. □

We summarize these conclusions as

Theorem 5.11 *If Y is a semimartingale with respect to a filtration $\{\mathcal{F}_t\}$ and X is cadlag and $\{\mathcal{F}_t\}$-adapted, then $\int X_-dY$ exists and is a cadlag semimartingale.*

The following lemma provides a useful estimate on $\int X_-dY$ in terms of properties of M and V.

Lemma 5.12 *Let $Y=M+V$ be a semimartingale where M is a local square-integrable martingale and V is a finite variation process. Let σ be a stopping time for which $E[[M]_{t\wedge\sigma}]=E[M(t\wedge\sigma)^2]<\infty$, and let $\tau_c=\inf\{t:|X(t)|\text{ or }|X(t-)|\geq c\}$. Then*

$$\begin{aligned}&P\{\sup_{s\leq t}|\int_0^s X(u-)dY(u)|>K\} \\ &\quad\leq P\{\sigma\leq t\}+P\{\sup_{s<t}|X(s)|\geq c\}+P\{\sup_{s\leq t\wedge\sigma\wedge\tau_c}|\int_0^s X(u-)dM(u)|>K/2\} \\ &\qquad\quad +P\{\sup_{s\leq t\wedge\tau_c}|\int_0^s X(u-)dV(u)|>K/2\} \\ &\quad\leq P\{\sigma\leq t\}+P\{\sup_{s<t}|X(s)|\geq c\}+\frac{16c^2E[[M]_{t\wedge\sigma}]}{K^2}+P\{T_t(V)\geq(2c)^{-1}K\}.\end{aligned}$$

Proof. The first inequality is immediate. The second follows by applying Doob's inequality to the square integrable martingale

$$\int_0^{s\wedge\sigma\wedge\tau_c} X(u-)dM(u)$$

and observing that

$$\sup_{u\le s}\left|\int_0^s X(u-)dV(u)\right| \le T_s(V)\sup_{u<s}|X(u)|.$$

□

5.4 Change of time variable.

We defined the stochastic integral as a limit of approximating sums

$$\int_0^t X(s-)dY(s) = \lim \sum X(t_i)(Y(t\wedge t_{i+1}) - Y(t\wedge t_i)),$$

where the t_i are a partition of $[0,\infty)$. By Theorem 5.20, the same limit holds if we replace the t_i by stopping times. The following lemma is a consequence of this observation.

Lemma 5.13 *Let Y be an $\{\mathcal{F}_t\}$-semimartingale, X be cadlag and $\{\mathcal{F}_t\}$ adapted, and γ be continuous and nondecreasing with $\gamma(0) = 0$. For each u, assume $\gamma(u)$ is an $\{\mathcal{F}_t\}$-stopping time. Then, $\mathcal{G}_t = \mathcal{F}_{\gamma(t)}$ is a filtration, $Y\circ\gamma$ is a $\{\mathcal{G}_t\}$ semimartingale, $X\circ\gamma$ is cadlag and $\{\mathcal{G}_t\}$-adapted, and*

$$\int_0^{\gamma(t)} X(s-)dY(s) = \int_0^t X\circ\gamma(s-)dY\circ\gamma(s). \tag{5.12}$$

(Recall that if X is $\{\mathcal{F}_t\}$-adapted, then $X(\tau)$ is $\mathcal{F}_\tau$ measurable).

Proof.

$$\begin{aligned}\int_0^t X\circ\gamma(s-)dY\circ\gamma(s) &= \lim\Sigma\{X\circ\gamma(t_i)(Y(\gamma(t_{i+1}\wedge t)) - Y(\gamma(t_i\wedge t)))\}\\ &= \lim\Sigma\{X\circ\gamma(t_i)(Y(\gamma(t_{i+1})\wedge\gamma(t)) - Y(\gamma(t_i)\wedge\gamma(t)))\}\\ &= \int_0^{\gamma(t)} X(s-)dY(s),\end{aligned}$$

where the last limit follows by Theorem 5.20. That $Y\circ\gamma$ is an $\{\mathcal{F}_{\gamma(t)}\}$-semimartingale follows from the optional sampling theorem. □

Lemma 5.14 *Let A be strictly increasing and adapted with $A(0) = 0$ and $\gamma(u) = \inf\{s : A(s) > u\}$. Then γ is continuous and nondecreasing, and $\gamma(u)$ is an $\{\mathcal{F}_t\}$-stopping time.*

Proof. Best done by picture. □

For A and γ as in Lemma 5.14, define $B(t) = A\circ\gamma(t)$ and note that $B(t)\ge t$.

Lemma 5.15 *Let A, γ, and B be as above, and suppose that $Z(t)$ is nondecreasing with $Z(0)=0$. Then*

$$\begin{aligned}\int_0^{\gamma(t)} Z(s-)dA(s) &= \int_0^t Z\circ\gamma(s-)dA\circ\gamma(s)\\ &= \int_0^t Z\circ\gamma(s-)d(B(s)-s)+\int_0^t Z\circ\gamma(s-)ds\\ &= Z\circ\gamma(t)(B(t)-t)-\int_0^t (B(s)-s)dZ\circ\gamma(s)\\ &\qquad -[B,Z\circ\gamma]_t+\int_0^t Z\circ\gamma(s)ds\end{aligned}$$

and hence

$$\int_0^{\gamma(t)} Z(s-)dA(s) \le Z\circ\gamma(t-)(B(t)-t)+\int_0^t Z\circ\gamma(s)ds.$$

5.5 Change of integrator.

Lemma 5.16 *Let Y be a semimartingale, and let X and U be cadlag and adapted. Suppose $Z(t)=\int_0^t X(s-)dY(s)$. Then*

$$\int_0^t U(s-)dZ(s) = \int_0^t U(s-)X(s-)dY(s).$$

Proof. Let $\{t_i\}$ be a partition of $[0,t]$, and define

$$t(s)=t_i \quad \text{as } t_i\le s<t_{i+1},$$

so that

$$\begin{aligned}\int_0^t U(s-)dZ(s) &= \lim\sum U(t_i)(Z(t\wedge t_{i+1})-Z(t\wedge t_i))\\ &= \lim\sum U(t_i\wedge t)\int_{t\wedge t_i}^{t\wedge t_{i+1}} X(s-)dY(s)\\ &= \lim\sum \int_{t\wedge t_i}^{t\wedge t_{i+1}} U(t_i\wedge t)X(s-)dY(s)\\ &= \lim\sum \int_{t\wedge t_i}^{t\wedge t_{i+1}} U(t(s-))X(s-)dY(s)\\ &= \lim\int_0^t U(t(s-))X(s-)dY(s)\\ &= \int_0^t U(s-)X(s-)dY(s)\end{aligned}$$

The last limit follows from the fact that $U(t(s-))\to U(s-)$ as $\max|t_{i+1}-t_i|\to 0$ by splitting the integral into martingale and finite variation parts and arguing as in the proofs of Theorems 5.2 and 5.3. □

Example 5.17 *Let τ be a stopping time (w.r.t. $\{\mathcal{F}_t\}$). Then $U(t) = I_{[0,\tau)}(t)$ is cadlag and adapted, since $\{U(t) = 1\} = \{\tau > t\} \in \mathcal{F}_t$. Note that*

$$Y^\tau(t) = Y(t \wedge \tau) = \int_0^t I_{[0,\tau)}(s-)dY(s)$$

and

$$\begin{aligned} \int_0^{t\wedge\tau} X(s-)dY(s) &= \int_0^t I_{[0,\tau)}(s-)X(s-)dY(s) \\ &= \int_0^t X(s-)dY^\tau(s) \end{aligned}$$

5.6 Localization

It will frequently be useful to restrict attention to random time intervals $[0,\tau]$ on which the processes of interest have desirable properties (for example, are bounded). Let τ be a stopping time, and define Y^τ by $Y^\tau(t) = Y(\tau \wedge t)$ and $X^{\tau-}$ by setting $X^{\tau-}(t) = X(t)$ for $t < \tau$ and $X(t) = X(\tau-)$ for $t \geq \tau$. Note that if Y is a local martingale, then Y^τ is a local martingale, and if X is cadlag and adapted, then $X^{\tau-}$ is cadlag and adapted. Note also that if $\tau = \inf\{t : X(t) \text{ or } X(t-) \geq c\}$, then $X^{\tau-} \leq c$.

The next lemma states that it is possible to approximate an arbitrary semimartingale by semimartingales that are bounded by a constant and (consequently) have bounded discontinuities.

Lemma 5.18 *Let $Y = M + V$ be a semimartingale, and assume (without loss of generality) that $\sup_s |\Delta M(s)| \leq 1$. Let*

$$A(t) = \sup_{s\leq t}(|M(s)| + |V(s)| + [M]_s + T_s(V))$$

and

$$\sigma_c = \inf\{t : A(t) \geq c\},$$

and define $M^c \equiv M^{\sigma_c}$, $V^c \equiv V^{\sigma_c-}$, and $Y^c \equiv M^c + V^c$. Then $Y^c(t) = Y(t)$ for $t < \sigma_c$, $\lim_{c\to\infty} \sigma_c = \infty$, $|Y^c| \leq c+1$, $\sup_s |\Delta Y^c(s)| \leq 2c+1$, $[M^c] \leq c+1$, $T(V) \leq c$.

Finally, note that

$$S(t \wedge \tau, \{t_i\}, X, Y) = S(t, \{t_i\}, X^{\tau-}, Y^\tau). \tag{5.13}$$

5.7 Approximation of stochastic integrals.

Proposition 5.19 *Suppose Y is a semimartingale $X_1, X_2, X_3, \ldots$ are cadlag and adapted, and*

$$\lim_{n\to\infty} \sup_{t\leq T} |X_n(t) - X(t)| = 0$$

in probability for each $T > 0$. Then

$$\lim_{n\to\infty}\sup_{t\leq T}\left|\int_0^t X_n(s-)dY(s) - \int_0^t X(s-)dY(s)\right| = 0$$

in probability.

Proof. By linearity and localization, it is enough to consider the cases Y a square integrable martingale and Y a finite variation process, and we can assume that $|X_n| \leq C$ for some constant C. The martingale case follows easily from Doob's inequality and the dominated convergence theorem, and the FV case follows by the dominated convergence theorem. □

Theorem 5.20 *Let Y be a semimartingale and X be cadlag and adapted. For each n, let $0 = \tau_0^n \leq \tau_1^n \leq \tau_2^n \leq \cdots$ be stopping times and suppose that $\lim_{k\to\infty}\tau_k^n = \infty$ and $\lim_{n\to\infty}\sup_k|\tau_{k+1}^n - \tau_k^n| = 0$. Then for each $T > 0$*

$$\lim_{n\to\infty}\sup_{t\leq T}|S(t,\{\tau_k^n\},X,Y) - \int_0^t X(s-)dY(s)| = 0.$$

Proof. If Y is FV, then the proof is exactly the same as for Theorem 5.2 (which is an ω by ω argument). If Y is a square integrable martingale and X is bounded by a constant, then defining $\tau^n(u) = \tau_k^n$ for $\tau_k^n \leq u < \tau_{k+1}^n$,

$$\begin{aligned} E[(S(t,\{\tau_k^n\},X,Y) - \int_0^t X(s-)dY(s))^2] \\ = E[\int_0^t (X(\tau^n(u-)) - X(u-))^2 d[Y]_u] \end{aligned}$$

and the result follows by the dominated convergence theorem. The theorem follows from these two cases by linearity and localization. □

5.8 Connection to Protter's text.

The approach to stochastic integration taken here differs somewhat from that taken in Protter (1990) in that we assume that all integrands are cadlag and do not introduce the notion of *predictability*. In fact, however, predictability is simply hidden from view and is revealed in the requirement that the integrands are evaluated at the left end points in the definition of the approximating partial sums. If X is a cadlag integrand in our definition, then the left continuous process $X(\cdot-)$ is the predictable integrand in the usual theory. Consequently, our notation $\int X_- dY$ and

$$\int_0^t X(s-)dY(s)$$

emphasizes this connection.

Protter (1990) defines $H(t)$ to be simple and predictable if

$$H(t) = \sum_{i=0}^m \xi_i I_{(\tau_i,\tau_{i+1}]}(t),$$

where $\tau_0 < \tau_1 < \cdots$ are $\{\mathcal{F}_t\}$-stopping times and the ξ_i are $\mathcal{F}_{\tau_i}$ measurable. Note that H is left continuous. In Protter, $H \cdot Y$ is defined by

$$H \cdot Y(t) = \sum \xi_i \left(Y(\tau_{i+1} \wedge t) - Y(\tau_i \wedge t)\right).$$

Defining

$$X(t) = \sum \xi_i I_{[\tau_i, \tau_{i+1})}(t),$$

we see that $H(t) = X(t-)$ and note that

$$H \cdot Y(t) = \int_0^t X(s-)dY(s),$$

so the definitions of the stochastic integral are consistent for simple functions. Protter extends the definition $H \cdot Y$ by continuity, and Propositon 5.19 ensures that the definitions are consistent for all H satisfying $H(t) = X(t-)$, where X is cadlag and adapted.

6 Covariation and Itô's formula.

6.1 Quadratic covariation.

The covariation of Y_1, Y_2 is defined by

$$[Y_1, Y_2]_t \equiv \lim \sum_i (Y_1(t_{i+1}) - Y_1(t_i))\,(Y_2(t_{i+1}) - Y_2(t_i)) \tag{6.1}$$

where the $\{t_i\}$ are partitions of $[0, t]$ and the limit is in probability as $\max|t_{i+1} - t_i| \to 0$. Note that

$$[Y_1 + Y_2, Y_1 + Y_2]_t = [Y_1]_t + 2[Y_1, Y_2]_t + [Y_2]_t.$$

If Y_1, Y_2, are semimartingales, then $[Y_1, Y_2]_t$ exists. This assertion follows from the fact that

$$\begin{aligned}
[Y_1, Y_2]_t &= \lim \sum_i (Y_1(t_{i+1}) - Y_1(t_i))\,(Y_2(t_{i+1}) - Y_2(t_i)) \\
&= \lim \Big(\sum (Y_1(t_{i+1})Y_2(t_{i+1}) - Y_1(t_i)Y_2(t_i)) \\
&\qquad - \sum Y_1(t_i)(Y_2(t_{i+1}) - Y_2(t_i)) - \sum Y_2(t_i)(Y_1(t_{i+1}) - Y_1(t_i))\Big) \\
&= Y_1(t)Y_2(t) - Y_1(0)Y_2(0) - \int_0^t Y_1(s-)dY_2(s) - \int_0^t Y_2(s-)dY_1(s).
\end{aligned}$$

Recall that if Y is of finite variation, then $[Y]_t = \sum_{s \le t} (\Delta Y(s))^2$, where $\Delta Y(s) \equiv Y(s) - Y(s-)$.

Lemma 6.1 *Let Y be a finite variation process, and let X be cadlag. Then*

$$[X, Y]_t = \sum_{s \le t} \Delta X(s) \Delta Y(s).$$

Remark 6.2 *Note that this sum will be zero if X and Y have no simultaneous jumps. In particular, if either X or Y is a finite variation process and either X or Y is continuous, then $[X, Y] = 0$.*

Proof. We have that the covariation $[X, Y]_t$ is

$$\begin{aligned}
&\lim_{\max|t_{i+1}-t_i| \to 0} \sum (X(t_{i+1}) - X(t_i))(Y(t_{i+1}) - Y(t_i)) \\
&\quad = \lim_{\max|t_{i+1}-t_i| \to 0} \sum_{|X(t_{i+1})-X(t_i)| > \epsilon} (X(t_{i+1}) - X(t_i))(Y(t_{i+1}) - Y(t_i)) \\
&\qquad + \lim_{\max|t_{i+1}-t_i| \to 0} \sum_{|X(t_{i+1})-X(t_i)|] \le \epsilon} (X(t_{i+1}) - X(t_i))(Y(t_{i+1}) - Y(t_i)),
\end{aligned}$$

where the first term on the right is approximately

$$\sum_{\substack{s \le t \\ |\Delta X(s)| > \epsilon}} \Delta X(s) \Delta Y(s)$$

plus or minus a few jumps where $\Delta X(s) = \epsilon$. Since the number of jumps in X is countable, ϵ can always be chosen so that there are no such jumps. The second term on the right is bounded by

$$\epsilon \sum |Y(t_{i+1}) - Y(t_i)| \leq \epsilon T_t(Y),$$

where $T_t(Y)$ is the total variation of Y (which is bounded). □

6.2 Continuity of the quadratic variation.

Since $\sum a_i b_i \leq \sqrt{\sum a_i^2 \sum b_i^2}$ it follows that $[X,Y]_t \leq \sqrt{[X]_t[Y]_t}$. Observe that

$$\begin{aligned} [X-Y]_t &= [X]_t - 2[X,Y]_t + [Y]_t \\ [X-Y]_t + 2\left([X,Y]_t - [Y]_t\right) &= [X]_t - [Y]_t \\ [X-Y]_t + 2[X-Y,Y]_t &= [X]_t - [Y]_t. \end{aligned}$$

Therefore,

$$|[X]_t - [Y]_t| \leq [X-Y]_t + 2\sqrt{[X-Y]_t[Y]_t}. \tag{6.2}$$

Assuming that $[Y]_t < \infty$, we have that $[X-Y]_t \to 0$ implies $[X]_t \to [Y]_t$.

Lemma 6.3 *Let M_n, $n = 1,2,3,\ldots$, be square-integrable martingales with $\lim_{n\to\infty} E[(M_n(t) - M(t))^2] = 0$ for all t. Then $E\left[|[M_n]_t - [M]_t|\right] \to 0$.*

Proof. Since

$$\begin{aligned} E\left[|[M_n]_t - [M]_t|\right] &\leq E\left[[M_n - M]_t\right] + 2E\left[\sqrt{[M_n - M]_t[M]_t}\right] \\ &\leq E\left[[M_n - M]_t\right] + 2\sqrt{E\left[[M_n - M]_t\right] E\left[[M]_t\right]}, \end{aligned}$$

we have the L^1 convergence of the quadratic variation. □

Lemma 6.4 *Suppose $\sup_{s\leq t} |X_n(s) - X(s)| \to 0$ and $\sup_{s\leq t} |Y_n(s) - Y(s)| \to 0$ for each $t > 0$, and $\sup_n T_t(Y_n) < \infty$. Then*

$$\lim_n [X_n, Y_n]_t = [X,Y]_t.$$

Proof. Note that $T_t(Y) \leq \sup_n T_t(Y_n)$, and recall that

$$[X_n, Y_n]_t = \sum_{s\leq t} \Delta X_n(s) \Delta Y_n(s).$$

We break the sum into two parts,

$$\left| \sum_{\substack{s\leq t \\ |\Delta X_n(s)|\leq\epsilon}} \Delta X_n(s)\Delta Y_n(s) \right| \leq \epsilon \sum_{s\leq t} |\Delta Y_n(s)| \leq \epsilon T_t(Y_n)$$

and

$$\sum_{s\leq t, |\Delta X_n(s)|>\epsilon} \Delta X_n(s)\Delta Y_n(s).$$

Since $\Delta X_n(s) \to \Delta X(s)$ and $\Delta Y_n(s) \to \Delta Y(s)$, we have

$$\begin{aligned}\limsup |[X_n, Y_n]_t - [X, Y]_t| &= \limsup |\sum \Delta X_n(s)\Delta Y_n(s) - \sum \Delta X(s)\Delta Y(s)| \\ &\leq \epsilon(T_t(Y_n) + T_t(Y)),\end{aligned}$$

and the lemma follows. □

Lemma 6.5 *Let $Y_i = M_i + V_i$, $Y_i^n = M_i^n + V_i^n$, $i = 1, 2$, $n = 1, 2, \ldots$ be semimartingales with M_i^n a local square integrable martingale and V_i^n finite variation. Suppose that there exist stopping times γ_k such that $\gamma_k \to \infty$ as $k \to \infty$ and for each $t \geq 0$,*

$$\lim_{n\to\infty} E[(M_i^n(t \wedge \gamma_k) - M_i(t \wedge \gamma_k))^2] = 0,$$

and that for each $t \geq 0$ $\sup_{i,n} T_t(V_i^n) < \infty$ and

$$\lim_{n\to\infty} \sup_{s\leq t} |V_i^n(s) - V_i(s)| = 0.$$

Then $[Y_1^n, Y_2^n]_t \to [Y_1, Y_2]_t$.

Proof. The result follows from Lemmas 6.3 and 6.4 by writing

$$[Y_1^n, Y_2^n]_t = [M_1^n, M_2^n]_t + [M_1^n, V_2^n]_t + [V_1^n, Y_2^n]_t.$$

□

Lemma 6.5 provides the proof for the following.

Lemma 6.6 *Let Y_i be a semimartingale, X_i be cadlag and adapted, and*

$$Z_i(t) = \int_0^t X_i(s-)dY_i(s) \quad i = 1, 2$$

Then,

$$[Z_1, Z_2]_t = \int_0^t X_1(s-)X_2(s-)d[Y_1, Y_2]_s$$

Proof. First verify the identity for piecewise constant X_i. Then approximate general X_i by piecewise constant processes and use Lemma 6.5 to pass to the limit. □

Lemma 6.7 *Let X be cadlag and adapted and Y be a semimartingale. Then*

$$\lim_{\max|t_{i+1}-t_i|\to 0} \sum X(t_i)(Y(t_{i+1}\wedge t) - Y(t_i \wedge t))^2 = \int_0^t X(s-)d[Y]_s. \tag{6.3}$$

Proof. Let $Z(t) = \int_0^t 2Y(s-)dY(s)$. Observing that

$$(Y(t_{i+1}\wedge t) - Y(t_i\wedge t))^2 = Y^2(t_{i+1}\wedge t) - Y^2(t_i \wedge t) - 2Y(t_i)(Y(t_{i+1}\wedge t) - Y(t_i \wedge t))$$

and applying Lemma 5.16, we see that the left side of (6.3) equals

$$\int_0^t X(s-)dY^2(s) - \int_0^t 2X(s-)Y(s-)dY(s) = \int_0^t X(s-)d(Y^2(s) - Z(s))\ ,$$

and since $[Y]_t = Y^2(t) - Y^2(0) - \int_0^t 2Y(s-)dY(s)$, the lemma follows. □

6.3 Ito's formula.

Theorem 6.8 *Let $f \in C^2$, and let Y be a semimartingale. Then*

$$\begin{aligned} f(Y(t)) &= f(Y(0)) + \int_0^t f'(Y(s-))\,dY(s) \\ &\quad + \int_0^t \frac{1}{2} f''(Y(s-))\,d[Y]_s \\ &\quad + \sum_{s \le t} \Big(f(Y(s)) - f(Y(s-)) - f'(Y(s-))\,\Delta Y(s) \\ &\qquad - \frac{1}{2} f''(Y(s-))\,(\Delta Y(s))^2 \Big). \end{aligned} \tag{6.4}$$

Remark 6.9 *Observing that the discontinuities in $[Y]_s$ satisfy $\Delta[Y]_s = \Delta Y(s)^2$, if we define the continuous part of the quadratic variation by*

$$[Y]_t^c = [Y]_t - \sum_{s \le t} \Delta Y(s)^2,$$

then (6.4) becomes

$$\begin{aligned} f(Y(t)) &= f(Y(0)) + \int_0^t f'(Y(s-))\,dY(s) \\ &\quad + \int_0^t \frac{1}{2} f''(Y(s-))\,d[Y]_s^c \\ &\quad + \sum_{s \le t} \left(f(Y(s)) - f(Y(s-)) - f'(Y(s-))\,\Delta Y(s) \right) \end{aligned} \tag{6.5}$$

Proof. Define

$$\Gamma_f(x,y) = \frac{f(y) - f(x) - f'(x)(y-x) - \frac{1}{2} f''(x)(y-x)^2}{(y-x)^2}$$

Then

$$\begin{aligned} f(Y(t)) &= f(Y(0)) + \sum f(Y(t_{i+1})) - f(Y(t_i)) \\ &= f(Y(0)) + \sum f'(Y(t_i))\,(Y(t_{i+1}) - Y(t_i)) \\ &\quad + \frac{1}{2} \sum f''(Y(t_i))\,(Y(t_{i+1}) - Y(t_i))^2 \\ &\quad + \sum \Gamma_f(Y(t_i), Y(t_{i+1}))\,(Y(t_{i+1}) - Y(t_i))^2 . \end{aligned} \tag{6.6}$$

The first three terms on the right of (6.6) converge to the corresponding terms of (6.4) by previous lemmas. Note that the last term in (6.4) can be written

$$\sum_{s \le t} \Gamma_f(Y(s-), Y(s)) \Delta Y(s)^2. \tag{6.7}$$

To show convergence of the remaining term, we need the following lemma.

Lemma 6.10 *Let X be cadlag. For $\epsilon > 0$, let $D_\epsilon(t) = \{s \le t : |\Delta X(s)| \ge \epsilon\}$. Then*

$$\limsup_{\max|t_{i+1}-t_i|\to 0} \max_{(t_i,t_{i+1}]\cap D_\epsilon(t)=\emptyset} |X(t_{i+1}) - X(t_i)| \le \epsilon.$$

(i.e. by picking out the sub-intervals with the larger jumps, the remaining intervals have the above property.) (Recall that $\Delta X(s) = X(s) - X(s-)$)

Proof. Suppose not. Then there exist $a_n < b_n \le t$ and $s \le t$ such that $a_n \to s$, $b_n \to s$, $(a_n, b_n] \cap D_\epsilon(t) = \emptyset$ and $\limsup |X(b_n) - X(a_n)| > \epsilon$. That is, suppose

$$\limsup_{m\to\infty} \max_{(t_i^m,t_{i+1}^m]\cap D_\epsilon(t)=\emptyset} |X(t_{i+1}^m) - X(t_i)^m| > \delta > \epsilon$$

Then there exists a subsequence in $(t_i^m, t_{i+1}^m]_{i,m}$ such that $|X(t_{i+1}^m) - X(t_i^m)| > \delta$. Selecting a further subsequence if necessary, we can obtain a sequence of intervals $\{(a_n, b_n]\}$ such that $|X(a_n) - X(b_n)| > \delta$ and $a_n, b_n \to s$. Each interval satisfies $a_n < b_n < s$, $s \le a_n < b_n$, or $a_n < s \le b_n$. If an infinite subsequence satisfies the first condition, then $X(a_n) \to X(s-)$ and $X(b_n) \to X(s-)$ so that $|X(b_n) - X(a_n)| \to 0$. Similarly, a subsequence satisfying the second condition gives $|X(b_n) - X(a_n)| \to 0$ since $X(b_n) \to X(s)$ and $X(a_n) \to X(s)$. Finally, a subsequence satisfying the third condition satisfies $|X(b_n) - X(a_n)| \to |X(s) - X(s-)| = |\Delta X(s)| \ge \delta > \epsilon$, and the contradiction proves the lemma. □

Proof of Theorem 6.8 continued. Assume $f \in C^2$ and suppose f'' is uniformly continuous. Let $\gamma_f(\epsilon) \equiv \sup_{|x-y|\le\epsilon} \Gamma_f(x, y)$. Then $\gamma_f(\epsilon)$ is a continuous function of ϵ and $\lim_{\epsilon\to 0} \gamma_f(\epsilon) = 0$. Let $D_\epsilon(t) = \{s \le t : |Y(s) - Y(s^-)| \ge \epsilon\}$. Then

$$\begin{aligned} &\sum \Gamma_f(Y(t_i), Y(t_{i+1}))\,(Y(t_{i+1}) - Y(t_i))^2 \\ &\quad= \sum_{(t_i,t_{i+1}]\cap D_\epsilon(t)\neq\emptyset} \Gamma_f\,(Y(t_i), Y(t_{i+1}))\,(Y(t_{i+1}) - Y(t_i))^2 \\ &\qquad+ \sum_{(t_i,t_{i+1}]\cap D_\epsilon(t)=\emptyset} \Gamma_f\,(Y(t_i), Y(t_{i+1}))\,(Y(t_{i+1}) - Y(t_i))^2, \end{aligned}$$

where the second term on the right is bounded by

$$e(\{t_i\}, Y) \equiv \gamma_f\left(\max_{(t_i,t_{i+1}]\cap D_\epsilon(t)=\emptyset} |Y(t_{i+1}) - Y(t_i)|\right) \sum (Y(t_{i+1}) - Y(t_i))^2$$

and

$$\limsup_{\max|t_{i+1}-t_i|\to 0} e(\{t_i\}, Y) \le \gamma_f(\epsilon)[Y]_t.$$

It follows that

$$\begin{aligned} \limsup \Big|\sum \Gamma_f(Y(t_i), Y(t_{i+1}))\,(Y(t_{i+1}) - Y(t_i))^2 - \sum \Gamma_f(Y(s-), Y(s))\Delta Y(s)^2\Big| \\ \le 2\gamma_f(\epsilon)[Y]_t \end{aligned}$$

which completes the proof of the theorem. □

6.4 The product rule and integration by parts.

Let X and Y be semimartingales. Then

$$\begin{aligned}
X(t)Y(t) &= X(0)Y(0) + \sum (X(t_{i+1})Y(t_{i+1}) - X(t_i)Y(t_i)) \\
&= X(0)Y(0) + \sum X(t_i)\left(Y(t_{i+1}) - Y(t_i)\right) + \sum Y(t_i)\left(X(t_{i+1}) - X(t_i)\right) \\
&\qquad + \sum \left(Y(t_{i+1}) - Y(t_i)\right)\left(X(t_{i+1}) - X(t_i)\right) \\
&= X(0)Y(0) + \int_0^t X(s-)dY(s) + \int_0^t Y(s-)dX(s) + [X,Y]_t.
\end{aligned}$$

Note that this identity generalizes the usual product rule and provides us with a formula for integration by parts.

$$\int_0^t X(s-)dY(s) = X(t)Y(t) - X(0)Y(0) - \int_0^t Y(s-)dX(s) - [X,Y]_t. \tag{6.8}$$

Example 6.11 *(Linear SDE.) As an application of (6.8), consider the stochastic differential equation*

$$dX = -\alpha X dt + dY$$

or in integrated form,

$$X(t) = X(0) - \int_0^t \alpha X(s)ds + Y(t).$$

We use the integrating factor $e^{\alpha t}$.

$$\begin{aligned}
e^{\alpha t}X(t) &= X(0) + \int_0^t e^{\alpha t}dX(s) + \int_0^t X(s-)de^{\alpha s} \\
&= X(0) - \int_0^t \alpha X(s)e^{\alpha s}ds + \int_o^t e^{\alpha s}dY(s) + \int_0^t X(s)\alpha e^{\alpha s}ds
\end{aligned}$$

which gives

$$X(t) = X(0)e^{-\alpha t} + \int_0^t e^{-\alpha(t-s)}dY(s).$$

Example 6.12 *(Kronecker's lemma.) Let A be positive and nondecreasing, and $\lim_{t\to\infty} A(t) = \infty$. Define*

$$Z(t) = \int_0^t \frac{1}{A(s-)}dY(s).$$

If $\lim_{t\to\infty} Z(t)$ exists a.s., then $\lim_{t\to\infty} \frac{Y(t)}{A(t)} = 0$ a.s.

Proof. By (6.8)

$$A(t)Z(t) = Y(t) + \int_0^t Z(s-)dA(s) + \int_0^t \frac{1}{A(s-)}d[Y,A]_s\,. \tag{6.9}$$

Rearranging (6.9) gives

$$\frac{Y(t)}{A(t)} = Z(t) - \frac{1}{A(t)}\int_0^t Z(s-)dA(s) + \frac{1}{A(t)}\sum_{s\le t}\frac{\Delta Y(s)}{A(s-)}\Delta A(s). \tag{6.10}$$

Note that the difference between the first and second terms on the right of (6.10) converges to zero. Convergence of Z implies $\lim_{t\to\infty}\frac{\Delta Y(t)}{A(t-)} = 0$, so the third term on the right of (6.10) converges to zero giving the result. $\square$

6.5 Itô's formula for vector-valued semimartingales.

Let $Y(t) = (Y_1(t), Y_2(t), ...Y_m(t))^T$ (a column vector). The product rule given above is a special case of Itô's formula for a vector-valued semimartingale Y. Let $f \in C^2(\mathbb{R}^m)$. Then

$$\begin{aligned} f(Y(t)) &= f(Y(0)) + \sum_{k=1}^m \int_0^t \partial_k f(Y(s-))\, dY_k(s) \\ &\quad + \sum_{k,l=1}^m \frac{1}{2}\int_0^t \partial_k\partial_l f(Y(s-))\, d[Y_k, Y_l]_s \\ &\quad + \sum_{s\le t}\Big(f(Y(s)) - f(Y(s-)) - \sum_{k=1}^m \partial_k f(Y(s-))\,\Delta Y_k(s) \\ &\qquad - \sum_{k,l=1}^m \frac{1}{2}\partial_k\partial_l f(Y(s-))\,\Delta Y_k(s)\Delta Y_l(s)\Big), \end{aligned}$$

or defining

$$[Y_k, Y_l]_t^c = [Y_k, Y_l]_t - \sum_{s\le t}\Delta Y_k(s)\Delta Y_l(s), \tag{6.11}$$

we have

$$\begin{aligned} f(Y(t)) &= f(Y(0)) + \sum_{k=1}^m \int_0^t \partial_k f(Y(s-))\, dY_k(s) \\ &\quad + \sum_{k,l=1}^m \frac{1}{2}\int_0^t \partial_k\partial_l f(Y(s-))\, d[Y_k, Y_l]_s^c \\ &\quad + \sum_{s\le t}\Big(f(Y(s)) - f(Y(s-)) - \sum_{k=1}^m \partial_k f(Y(s-))\,\Delta Y_k(s)\Big). \end{aligned} \tag{6.12}$$

7 Stochastic Differential Equations

7.1 Examples.

The standard example of a stochastic differential equation is an Itô equation for a diffusion process written in differential form as

$$dX(t) = \sigma(X(t))dW(t) + b(X(t))dt$$

or in integrated form as

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds \tag{7.1}$$

If we define $Y(t) = \Big(W(t), t\Big)^T$ and $F(X) = \Big(\sigma(X), b(X)\Big)$, then (7.1) can be written as

$$X(t) = X(0) + \int_0^t F(X(s-))dY(s) . \tag{7.2}$$

Similarly, consider the stochastic difference equation

$$X_{n+1} = X_n + \sigma(X_n)\xi_{n+1} + b(X_n)h \tag{7.3}$$

where the ξ_i are iid and $h > 0$. If we define $Y_1(t) = \sum_{k=1}^{[t/h]} \xi_k$, $Y_2(t) = [t/h]h$, and $X(t) = X_{[t/h]}$, then

$$X(t) = X(0) + \int_0^t \Big(\sigma(X(s-)), b(X(s-))\Big)dY(s)$$

which is in the same form as (7.2). With these examples in mind, we will consider stochastic differential equations of the form (7.2) where Y is an $\mathbb{R}^m$-valued semimartingale and F is a $d \times m$ matrix-valued function.

7.2 Gronwall's inequality and uniqueness for ODEs.

Of course, systems of ordinary differential equations are of the form (7.2), and we begin our study by recalling the standard existence and uniqueness theorem for these equations. The following inequality will play a crucial role in our discussion.

Lemma 7.1 *(Gronwall's inequality.) Suppose that A is cadlag and non-decreasing, X is cadlag, and that*

$$0 \leq X(t) \leq \epsilon + \int_0^t X(s-)dA(s) . \tag{7.4}$$

Then

$$X(t) \leq \epsilon e^{A(t)}.$$

Proof. Iterating (7.4), we have

$$\begin{aligned} X(t) &\le \epsilon + \int_0^t X(s-)dA(s) \\ &\le \epsilon + \epsilon A(t) + \int_0^t \int_0^{s-} X(u-)dA(u)dA(s) \\ &\le \epsilon + \epsilon A(t) + \epsilon \int_0^t A(s-)dA(s) + \int_0^t \int_0^{s-} \int_0^{u-} X(\sigma-)dA(\sigma)dA(u)dA(s) \end{aligned}$$

Since A is nondecreasing, it must be of finite variation, making $[A]_t^c \equiv 0$. Ito's formula thus yields

$$\begin{aligned} e^{A(t)} &= 1 + \int_0^t e^{A(s-)}dA(s) + \Sigma_{s\le t}(e^{A(s)} - e^{A(s-)} - e^{A(s-)}\Delta A(s)) \\ &\ge 1 + \int_0^t e^{A(s-)}dA(s) \\ &\ge 1 + A(t) + \int_0^t \int_0^{s-} e^{A(u-)}dA(u)dA(s) \\ &\ge 1 + A(t) + \int_0^t A(s-)dA(s) + \int_0^t \int_0^{s-} \int_0^{u-} e^{A(v-)}dA(v)dA(u)dA(s)\ . \end{aligned}$$

Continuing the iteration, we see that $X(t) \le \epsilon e^{A(t)}$. □

Theorem 7.2 *(Existence and uniqueness for ordinary differential equations.) Consider the ordinary differential equation in* $\mathbb{R}^d$

$$\dot{X} = \frac{dX}{dt} = F(X)$$

or in integrated form,

$$X(t) = X(0) + \int_0^t F(X(s))ds. \tag{7.5}$$

Suppose F *is Lipschitz, that is,* $|F(x) - F(y)| \le L|x - y|$ *for some constant* L*. Then for each* $x_0 \in \mathbb{R}^d$*, there exists a unique solution of (7.5) with* $X(0) = x_0$*.*

Proof. (Uniqueness) Suppose $X_i(t) = X_i(0) + \int_0^t F(X_i(s))ds, \quad i = 1, 2$

$$\begin{aligned} |X_1(t) - X_2(t)| &\le |X_1(0) - X_2(0)| + \int_0^t |F(X_1(s)) - F(X_2(s))|ds \\ &\le |X_1(0) - X_2(0)| + \int_0^t L|X_1(s) - X_2(s)|ds \end{aligned}$$

By Gronwall's inequality (take $A(t) = Lt$)

$$|X_1(t) - X_2(t)| \le |X_1(0) - X_2(0)|e^{tL}.$$

Hence, if $X_1(0) = X_2(0)$, then $X_1(t) \equiv X_2(t)$. □

7.3 Uniqueness of solutions of SDEs.

We consider stochastic differential equations of the form

$$X(t) = U(t) + \int_0^t F(X(s-))dY(s). \tag{7.6}$$

where Y is an $\mathbb{R}^m$-valued semimartingale, U is a cadlag, adapted $\mathbb{R}^d$-valued process, and $F : \mathbb{R}^d \to \mathbb{M}^{d\times m}$.

We will need the following generalization of Lemma 5.12.

Lemma 7.3 *Let Y be a semimartingale, X be a cadlag, adapted process, and τ be a finite stopping time. Then for any stopping time σ for which $E[[M]_{(\tau+t)\wedge\sigma}] < \infty$,*

$$\begin{aligned} P\{\sup_{s\leq t} |\int_\tau^{\tau+s} X(u-)dY(u)| > K\} & \qquad\qquad (7.7) \\ \leq P\{\sigma \leq \tau+t\} + P\{\sup_{\tau\leq s<\tau+t} |X(s)| > c\} + \frac{16c^2 E[[M]_{(\tau+t)\wedge\sigma} - [M]_{\tau\wedge\sigma}]}{K^2} & \\ +P\{T_{\tau+t}(V) - T_\tau(V) \geq (2c)^{-1}K\}. & \end{aligned}$$

Proof. The proof is the same as for Lemma 5.12. The strict inequality in the second term on the right is obtained by approximating c from above by a decreasing sequence c_n. □

Theorem 7.4 *Suppose that there exists $L > 0$ such that*

$$|F(x) - F(y)| \leq L|x - y|.$$

Then there is at most one solution of (7.6).

Remark 7.5 *One can treat more general equations of the form*

$$X(t) = U(t) + \int_0^t F(X, s-)dY(s) \tag{7.8}$$

where $F : D_{\mathbb{R}^d}[0,\infty) \to D_{\mathbb{M}^{d\times m}}[0,\infty)$ and satisfies

$$\sup_{s\leq t} |F(x,s) - F(y,s)| \leq L \sup_{s\leq t} |x(s) - y(s)| \tag{7.9}$$

for all $x, y \in D_{\mathbb{R}^d}[0,\infty)$ and $t \geq 0$. Note that, defining x^t by $x^t(s) = x(s\wedge t)$, (7.9) implies that F is nonanticipating *in the sense that $F(x,t) = F(x^t,t)$ for all $x \in D_{\mathbb{R}^d}[0,\infty)$ and all $t \geq 0$.*

Proof. It follows from Lemma 7.3 that for each stopping time τ satisfying $\tau \leq T$ a.s. for some constant $T > 0$ and $t, \delta > 0$, there exists a constant $K_\tau(t,\delta)$ such that

$$P\{\sup_{s\leq t} |\int_\tau^{\tau+s} X(u-)dY(u)| \geq K_\tau(t,\delta)\} \leq \delta$$

for all cadlag, adapted X satisfying $|X| \leq 1$. (Take $c = 1$ in (7.7).) Furthermore, K_τ can be chosen so that for each $\delta > 0$, $\lim_{t\to 0} K_\tau(t,\delta) = 0$.

Suppose X and $\tilde{X}$ satisfy (7.6). Let $\tau_0 = \inf\{t : |X(t) - \tilde{X}(t)| > 0\}$, and suppose $P\{\tau_0 < \infty\} > 0$. Select $r, \delta, t > 0$, such that $P\{\tau_0 < r\} > \delta$ and $LK_{\tau_0\wedge r}(t,\delta) < 1$. Note that if $\tau_0 < \infty$, then

$$X(\tau_0) - \tilde{X}_0(\tau_0) = \int_0^{\tau_0} (F(X(s-)) - F(\tilde{X}(s-))dY(s) = 0. \tag{7.10}$$

Define

$$\tau_\epsilon = \inf\{s : |X(s) - \tilde{X}(s)| \geq \epsilon\}.$$

Noting that $|X(s) - \tilde{X}(s)| \leq \epsilon$ for $s < \tau_\epsilon$, we have

$$|F(X(s)) - F(\tilde{X}(s))| \leq \epsilon L,$$

for $s < \tau_\epsilon$, and

$$|\int_0^{\tau_\epsilon} (F(X(s-)) - F(\tilde{X}(s-))dY(s)| = |X(\tau_\epsilon) - \tilde{X}(\tau_\epsilon)| \geq \epsilon.$$

Consequently, for $r > 0$, letting $\tau_0^r = \tau_0 \wedge r$, we have

$$\begin{aligned} &P\{\tau_\epsilon - \tau_0^r \leq t\} \\ &\leq P\{\sup_{s\leq t\wedge(\tau_\epsilon - \tau_0^r)} |\int_0^{\tau_0^r + s} F(X(u-))dY(u) - \int_0^{\tau_0^r+s} F(\tilde{X}(u-))dY(u)| \geq \epsilon L K_{\tau_0^r}(t,\delta)\} \\ &\leq \delta. \end{aligned}$$

Since the right side does not depend on ϵ and $\lim_{\epsilon\to 0}\tau_\epsilon = \tau_0$, it follows that $P\{\tau_0 - \tau_0 \wedge r < t\} \leq \delta$ and hence that $P\{\tau_0 < r\} \leq \delta$, contradicting the assumption on δ and proving that $\tau_0 = \infty$ a.s.

□

7.4 A Gronwall inequality for SDEs

Let Y be an $\mathbb{R}^m$-valued semimartingale, and let $F : \mathbb{R}^d \to \mathbb{M}^{d\times m}$ satisfy $|F(x) - F(y)| \leq L|x - y|$. For $i = 1, 2$, let U_i be cadlag and adapted and let X_i satisfy

$$X_i(t) = U_i(t) + \int_0^t F(X_i(s-))dY(s). \tag{7.11}$$

Lemma 7.6 *Let $d = m = 1$. Suppose that $Y = M + V$ where M is a square-integrable martingale and V is a finite variation process. Suppose that there exist $\delta > 0$ and $R > 0$ such that $\sup_t |\Delta M(t)| \leq \delta$, $\sup_t |\Delta V(t)| \leq 2\delta$ and $T_t(V) \leq R$, and that $c(\delta, R) \equiv (1 - 12L^2\delta^2 - 6L^2R\delta) > 0$. Let*

$$A(t) = 12L^2[M]_t + 3L^2RT_t(V) + t, \tag{7.12}$$

and define $\gamma(u)=\inf\{t:A(t)>u\}$. *(Note that the "$t$" on the right side of (7.12) only serves to ensure that A is strictly increasing.) Then*

$$E[\sup_{s\leq\gamma(u)}|X_1(s)-X_2(s)|^2]\leq\frac{3}{c(\delta,R)}e^{\frac{u}{c(\delta,R)}}E[\sup_{s\leq\gamma(u)}|U_1(s)-U_2(s)|^2]. \tag{7.13}$$

Proof. Note that

$$\begin{aligned}|X_1(t)-X_2(t)|^2 &\leq 3|U_1(t)-U_2(t)|^2 \\ &\quad +3|\int_0^t(F(X_1(s-))-F(X_2(s-)))dM(s)|^2 \\ &\quad +3|\int_0^t(F(X_1(s-))-F(X_2(s-)))dV(s)|^2.\end{aligned} \tag{7.14}$$

Doob's inequality implies

$$\begin{aligned}&E[\sup_{t\leq\gamma(u)}|\int_0^t(F(X_1(s-))-F(X_2(s-)))dM(s)|^2] \\ &\qquad\leq 4E[\int_0^{\gamma(u)}|F(X_1(s-))-F(X_2(s-))|^2d[M]],\end{aligned} \tag{7.15}$$

and Jensen's inequality implies

$$\begin{aligned}&E[\sup_{t\leq\gamma(u)}|\int_0^t(F(X_1(s-))-F(X_2(s-)))dV(s)|^2] \\ &\qquad\leq E[T_{\gamma(u)}(V)\int_0^{\gamma(u)}|F(X_1(s-))-F(X_2(s-))|^2dT_s(V)].\end{aligned} \tag{7.16}$$

Letting $Z(t)=\sup_{s\leq t}|X_1(s)-X_2(s)|^2$ and using the Lipschitz condition and the assumption that $T_t(V)\leq R$ it follows that

$$\begin{aligned}E[Z\circ\gamma(u)] &\leq 3E[\sup_{s\leq\gamma(u)}|U_1(s)-U_2(s)|^2] \\ &\quad +12L^2E[\int_0^{\gamma(u)}|X_1(s-)-X_2(s-)|^2d[M]] \\ &\quad +3L^2RE[\int_0^{\gamma(u)}|X_1(s-)-X_2(s-)|^2dT_s(V)] \\ &\leq 3E[\sup_{s\leq\gamma(u)}|U_1(s)-U_2(s)|^2] \\ &\quad +E[\int_0^{\gamma(u)}Z(s-)dA(s)] \\ &\leq 3E[\sup_{s\leq\gamma(u)}|U_1(s)-U_2(s)|^2] \\ &\quad +E[(A\circ\gamma(u)-u)Z\circ\gamma(u-)]+E[\int_0^u Z\circ\gamma(s-)ds].\end{aligned} \tag{7.17}$$

Since $0 \leq A \circ \gamma(u) - u \leq \sup_t \Delta A(t) \leq 12L^2\delta^2 + 6L^2 R\delta$, (7.12) implies

$$c(\delta, R)E[Z \circ \gamma(u)] \leq 3E[\sup_{s \leq \gamma(u)} |U_1(s) - U_2(s)|^2] + \int_0^u E[Z \circ \gamma(s-)]ds,$$

and (7.13) follows by Gronwall's inequality.

Note that the above calculation is valid only if the expectations on the right of (7.15) and (7.16) are finite. This potential problem can be eliminated by defining $\tau_K = \inf\{t : |X_1(t) - X_2(t)| \geq K\}$ and replacing $\gamma(u)$ by $\gamma(u) \wedge \tau_K$. Observing that $|X_1(s-) - X_2(s-)| \leq K$ for $s \leq \tau_K$, the estimates in (7.17) imply (7.13) with $\gamma(u)$ replaced by $\gamma(u) \wedge \tau_K$. Letting $K \to \infty$ gives (7.13) as originally stated. □

Lemma 7.6 gives an alternative approach to proving uniqueness.

Lemma 7.7 *Let $d = m = 1$, and let $U = U_1 = U_2$ in (7.11). Then there is a stopping time σ depending only on Y such that $P\{\sigma > 0\} = 1$ and $X_1(t) = X_2(t)$ for $t \in [0, \sigma]$.*

Proof Let $\tau_1 = \inf\{t > 0 : |\Delta Y(t)| > \delta\}$ and define $\hat{Y}$ by

$$\hat{Y}(t) = \begin{cases} Y(t) & t < \tau_1 \\ Y(t-) & t \geq \tau_1. \end{cases}$$

Then $\hat{Y}$ is a semimartingale satisfying $\sup_t |\Delta Y(t)| \leq \delta$ and hence by Lemma 5.8, $\hat{Y}$ can be written as $\hat{Y} = \hat{M} + \hat{V}$ where $\hat{M}$ is a local square-integrable martingale with $|\Delta \hat{M}(t)| \leq \delta$ and $\hat{V}$ is finite variation with $|\Delta \hat{V}(t)| \leq 2\delta$. Let $\tau_2 = \inf\{t : |\hat{M}(t)| \geq K\}$ and, noting that $|\hat{M}(t)| \leq K + \delta$ for $t \leq \tau_2$, we have that $\hat{M}^{\tau_2} \equiv \hat{M}(\cdot \wedge \tau_2)$ is a square-integrable martingale. Finally, let $\tau_3 = \inf\{t : T_t(\hat{V}) > R\}$ and define

$$\hat{\hat{V}}(t) = \begin{cases} \hat{V}(t) & t < \tau_3 \\ \hat{V}(t-) & t \geq \tau_3 \end{cases}$$

and $\hat{\hat{Y}} = \hat{M}^{\tau_2} + \hat{\hat{V}}$. Note that $\hat{\hat{Y}}$ satisfies the conditions of Lemma 7.6 and that $Y(t) = \hat{\hat{Y}}(t)$ for $t < \sigma \equiv \tau_1 \wedge \tau_2 \wedge \tau_3$. Setting

$$\hat{\hat{X}}_i(t) = \begin{cases} X_i(t) & t < \sigma \\ X_i(t-) & t \geq \sigma \end{cases}$$

and defining $\hat{\hat{U}}$ similarly, we see that

$$\hat{\hat{X}}_i(t) = \hat{\hat{U}}(t) + \int_0^t F(\hat{\hat{X}}_i(s-))d\hat{\hat{Y}}(s).$$

By Lemma 7.6, $X_1(t) = \hat{\hat{X}}_1(t) = \hat{\hat{X}}_2(t) = X_2(t)$ for $t < \sigma$. Since $X_i(\sigma) = X_i(\sigma-) + F(X_i(\sigma-))\Delta Y(\sigma)$, we see that $X_1(\sigma) = X_2(\sigma)$ as well. □

Proposition 7.8 . *Let $d = m = 1$, and let $U = U_1 = U_2$ in (7.11). Then $X_1 = X_2$ a.s.*

Proof. Let $\eta = \inf\{t : X_1(t) \neq X_2(t)\}$. For any $T < \infty$, $X_1(T \wedge \eta) = X_2(T \wedge \eta)$. But "starting over" at $T \wedge \eta$, Lemma 7.7 implies that there is a stopping time $\hat{\eta} > T \wedge \eta$ such that $X_1(t) = X_2(t)$ for $t \leq \hat{\eta}$, and hence $P\{\eta > T\} = 1$. But T is arbitrary, so $\eta = \infty$. □

Remark 7.9 *The proof of these results for $d, m > 1$ is essentially the same with different constants in the analogue of Lemma 7.6.*

7.5 Existence of solutions.

If X is a solution of (7.6), we will say that X is a solution of the equation (U,Y,F). Let Y^c be defined as in Lemma 5.18. If we can prove existence of a solution X^c of the equation (U,Y^c,F) (that is, of (7.6) with Y replaced by Y^c), then since $Y^c(t)=Y(t)$ for $t<\sigma_c$, we have existence of a solution of the original equation on the interval $[0,\sigma_c)$. For $c'>c$, suppose $X^{c'}$ is a solution of the equation $(U,Y^{c'},F)$. Define $\hat{X}^c(t)=X^{c'}(t)$ for $t<\sigma_c$ and $\hat{X}^c(t)=F(X^{c'}(\sigma_c-))\Delta Y^c(\sigma_c)$ for $t\geq\sigma_c$. Then $\hat{X}^c$ will be a solution of the equation (U,Y^c,F). Consequently, if for each $c>0$, existence and uniqueness holds for the equation (U,Y^c,F), then $X^c(t)=X^{c'}(t)$ for $t<\sigma_c$ and $c'>c$, and hence, $X(t)=\lim_{c\to\infty}X^c(t)$ exists and is the unique solution of the equation (U,Y,F).

With the above discussion in mind, we consider the existence problem under the hypotheses that $Y=M+V$ with $|M|+[M]+T(V)\leq R$. Consider the following approximation:

$$\begin{aligned}
X_n(0) &= X(0)\\
X_n(\frac{k+1}{n}) &= X_n(\frac{k}{n})+U(\frac{k+1}{n})-U(\frac{k}{n})+F(X(\frac{k}{n}))(Y(\frac{k+1}{n})-Y(\frac{k}{n})).
\end{aligned}$$

Let $\eta_n(t)=\frac{k}{n}$ for $\frac{k}{n}\leq t<\frac{k+1}{n}$. Extend X_n to all $t\geq 0$ by setting

$$X_n(t)=U(t)+\int_0^t F(X_n\circ\eta_n(s-))dY(s)\ .$$

Adding and subtracting the same term yields

$$\begin{aligned}
X_n(t) &= U(t)+\int_0^t(F(X_n\circ\eta_n(s-))-F(X_n(s-)))dY(s)+\int_0^t F(X_n(s-))dY(s)\\
&\equiv U(t)+D_n(t)+\int_0^t F(X_n(s-))dY(s).
\end{aligned}$$

Assume that $|F(x)-F(y)|\leq L|x-y|$, and for $b>0$, let $\gamma_n^b=\inf\{t:|F(X_n(t))|\geq b\}$. Then for $T>0$,

$$\begin{aligned}
E[\sup_{s\leq\gamma_n^b\wedge T}|D_n(s)|^2] &\leq 2E[\sup_{t\leq\gamma_n^b\wedge T}(\int_0^t(F(X_n\circ\eta_n(s-))-F(X_n(s-)))dM(s))^2]\\
&\quad +2E[\sup_{t\leq\gamma_n^b\wedge T}(\int_0^t(F(X_n\circ\eta_n(s-))-F(X_n(s-)))dV(s))^2]\\
&\leq 8L^2E[\int_0^{\gamma_n^b\wedge T}|X_n\circ\eta_n(s-)-X_n(s-)|^2d[M]_s]\\
&\quad +2RL^2E[\int_0^{\gamma_n^b\wedge T}|X_n\circ\eta_n(s-)-X_n(s-)|^2dT_s(V)]\\
&= 8L^2E[\int_0^{\gamma_n^b\wedge T}F^2(X_n\circ\eta_n(s-))(Y(s-)-Y(\eta_n(s-)))^2d[M]_s]\\
&\quad +2RL^2E[\int_0^{\gamma_n^b\wedge T}F^2(X_n\circ\eta_n(s-))(Y(s-)-Y(\eta_n(s-)))^2dT_s(V)]
\end{aligned}$$

$$
\begin{aligned}
= \quad & 8L^2b^2E[\int_0^{\gamma_n^b\wedge T}(Y(s-)-Y(\eta_n(s-)))^2d[M]_s] \\
& \qquad +2RL^2b^2E[\int_0^{\gamma_n^b\wedge T}(Y(s-)-Y(\eta_n(s-)))^2dT_s(V)] \,,
\end{aligned}
$$

so under the boundedness assumptions on Y,

$$E[\sup_{s\leq\gamma_n^b\wedge T}|D_n(s)|^2]\to 0.$$

Now assume that F is bounded, $\sup|\Delta M(s)|\leq\delta$, $\sup|\Delta V(s)|\leq 2\delta$, $T_t(V)\leq R$, and that L, δ, and R satisfy the conditions of Lemma 7.6. Since

$$X_n(t)=U(t)+D_n(t)+\int_0^t F(X_n(s-))dY(s),$$

Lemma 7.6 implies $\{X_n\}$ is a Cauchy sequence and converges uniformly in probability to a solution of

$$X(t)=U(t)+\int_0^t F(X(s-))dY(s) \,.$$

A localization argument gives the following theorem.

Theorem 7.10 *Let Y be an $\mathbb{R}^m$-valued semimartingale, and let $F:\mathbb{R}^d\to\mathbb{M}^{d\times m}$ be bounded and satisfy $|F(x)-F(y)|\leq L|x-y|$. Then for each $X(0)$ there exists a unique solution of*

$$X(t)=U(t)+\int_0^t F(X(s-))dY(s). \tag{7.18}$$

The assumption of boundedness in the above theorem can easily be weakened. For general Lipschitz F, the theorem implies existence and uniqueness up to $\tau_k=\inf\{t:|F(x(s)|\geq k\}$ (replace F by a bounded function that agrees with F on the set $\{x:|F(x)|\leq k\}$). The *global existence* question becomes whether or not $\lim_k\tau_k=\infty$? F is *locally Lispchitz* if for each $k>0$, there exists an L_k, such that

$$|F(x)-F(y)|\leq L_k|x-y| \qquad \forall|x|\leq k,|y|\leq k.$$

Note that if F is locally Lipschitz, and ρ_k is a smooth nonnegative function satisfying $\rho_k(x)=1$ when $|x|\leq k$ and $\rho_k(x)=0$ when $|x|\geq k+1$, then $F_k(x)=\rho_k(x)F(x)$ is globally Lipschitz and bounded.

Example 7.11 *Suppose*

$$X(t)=1+\int_0^t X(s)^2ds \,.$$

Then $F(x)=x^2$ is locally Lipschitz and local existence and uniqueness holds; however, global existence does not hold since X hits ∞ in finite time.

7.6 Moment estimates.

Consider the scalar Itô equation

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

Then by Itô's formula and Lemma 5.16,

$$\begin{aligned} X(t)^2 &= X(0)^2 + \int_0^t 2X(s)\sigma(X(s))dW(s) \\ &\quad + \int_0^t 2X(s)b(X(s))ds + \int_0^t \sigma^2(X(s))ds\ . \end{aligned}$$

Define $\tau_k = \inf\{t : |X(t)| \geq k\}$. Then

$$\begin{aligned} |X(t\wedge\tau_k)|^2 &= |X(0)|^2 + \int_0^{t\wedge\tau_k} 2X(s)\sigma(X(s))dW(s) \\ &\quad + \int_0^{t\wedge\tau_k} (2X(s)b(X(s)) + \sigma^2(X(s)))ds\ . \end{aligned}$$

Since

$$\int_0^{t\wedge\tau_k} 2X(s)\sigma(X(s))dW(s) = \int_0^t 1_{[0,\tau_k)} 2X(s)\sigma(X(s))dW(s)$$

has a bounded integrand, the integral is a martingale. Therefore,

$$E[|X(t\wedge\tau_k)|^2] = E[|X(0)|^2] + \int_0^t E[1_{[0,\tau_k)}(2X(s)b(X(s)) + \sigma^2(X(s)))]ds\ .$$

Assume $(2xb(x) + \sigma^2(x)) \leq K_1 + K_2|x|^2$ for some $K_i > 0$. (Note that this assumption holds if both $b(x)$ and $\sigma(x)$ are globally Lipschitz.) Then

$$\begin{aligned} m_k(t) &\equiv E[|X(t\wedge\tau_k)|^2] \\ &= E|X(0)|^2 + \int_0^t E\{1_{[0,\tau_k)}[2X(s)b(X(s)) + \sigma^2(X(s))]\}ds \\ &\leq m_0 + K_1 t + \int_0^t m_k(s)K_2 ds \end{aligned}$$

and hence

$$m_k(t) \leq (m_0 + K_1 t)e^{K_2 t}.$$

Note that

$$|X(t\wedge\tau_k)|^2 = (I_{\{\tau_k > t\}}|X(t)| + I_{\{\tau_k \leq t\}}|X(\tau_k)|)^2,$$

and we have

$$k^2 P(\tau_k \leq t) \leq E(|X(t\wedge\tau_k)|^2) \leq (m_0 + K_1 t)e^{K_2 t}.$$

Consequently, as $k \to \infty$, $P(\tau_k \leq t) \to 0$ and $X(t\wedge\tau_k) \to X(t)$. By Fatou's Lemma, $E|X(t)|^2 \leq (m_0 + K_1 t)e^{K_2 t}$.

Remark 7.12 *The argument above works well for moment estimation under other conditions also. Suppose* $2xb(x) + \sigma^2(x) \leq K_1 - \varepsilon|x|^2$*. (For example, consider the equation* $X(t) = X(0) - \int_0^t \alpha X(s)ds + W(t)$*.) Then*

$$\begin{aligned} e^{\varepsilon t}|X(t)|^2 &\leq |X(0)|^2 + \int_0^t e^{\varepsilon s} 2X(s)\sigma(X(s))dW(s) \\ &\quad + \int_0^t e^{\varepsilon s}[2X(s)b(X(s)) + \sigma^2(X(s))]ds + \int_0^t \varepsilon e^{\varepsilon s}|X(s)|^2 ds \\ &\leq |X(0)|^2 + \int_0^t e^{\varepsilon s} 2X(s)\sigma(X(s))dW(s) + \int_0^t e^{\varepsilon s} K_1 ds \\ &\leq |X(0)|^2 + \int_0^t e^{\varepsilon s} 2X(s)\sigma(X(s))dW(s) + \frac{K_1}{2}(e^{\varepsilon t} - 1), \end{aligned}$$

and hence

$$e^{\varepsilon t}E[|X(t)|^2] \leq E[|X(0)|^2] + \frac{K_1}{\varepsilon}[e^{\varepsilon t} - 1].$$

Therefore,

$$E[|X(t)|^2]] \leq e^{-\varepsilon t}E[|X(0)|^2] + \frac{K_1}{\varepsilon}(1 - e^{-\varepsilon t}),$$

which is uniformly bounded.

Consider a vector case example. Assume,

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

where σ is a $d \times m$ matrix, b is a d-dimensional vector, and W is an m-dimensional standard Brownian motion. Then

$$|X(t)|^2 = \sum_{i=1}^d |X_i(t)|^2 = |X(0)|^2 + \sum \int_0^t 2X_i(s)dX_i(s) + \sum_{i=1}^d [X_i]_t \ .$$

Define $Z_i(t) = \sum_{k=1}^m \int_0^t \sigma_{ik}(X(s))dW_k(s) \equiv \sum_{k=1}^m U_k$ where $U_k = \int_0^t \sigma_{ik}(X(s))dW_k(s)$. Then $[Z_i] = \sum_{k,l}[U_k, U_l]$, and

$$\begin{aligned} [U_k, U_l]_t &= \int_0^t \sigma_{ik}(X(s))\sigma_{il}(X(s))d[W_k, W_l]_s \\ &= \begin{cases} 0 & k \neq l \\ \int_0^t \sigma_{ik}^2(X(s))ds & k = l \end{cases} \end{aligned}$$

Consequently,

$$|X(t)|^2 = |X(0)|^2 + \int_0^t 2X(s)^T \sigma(X(s))dW(s)$$

$$
\begin{aligned}
&\quad + \int_0^t [2X(s)\cdot b(X(s)) + \sum_{i,k} \sigma_{ik}^2(X(s))]ds \\
&= |X(0)|^2 + \int_0^t 2X(s)^T \sigma(X(s))dW(s) \\
&\quad + \int_0^t (2X(s)\cdot b(X(s)) + \mathrm{trace}(\sigma(X(s))\sigma(X(s))^T))ds\ .
\end{aligned}
$$

As in the univariate case, if we assume,

$$
2x\cdot b(x) + \mathrm{trace}(\sigma(x)\sigma(x)^T) \le K_1 - \varepsilon|x|^2,
$$

then $E[|X(s)|^2]$ is uniformly bounded.

8 Stochastic differential equations for diffusion processes.

8.1 Generator for a diffusion process.

Consider

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

where X is $\mathbb{R}^d$-valued, W is an m-dimensional standard Brownian motion, σ is a $d \times m$ matrix-valued function and b is an $\mathbb{R}^d$-valued function. For a C^2 function f,

$$\begin{aligned} f(X(t)) &= f(X(0)) + \sum_{i=1}^d \int_0^t \partial_i f(X(s))dX(s) \\ &\quad + \frac{1}{2} \sum_{1 \le i,j \le d} \int_0^t \partial_i \partial_j f(X(s))d[X_i, X_j]_s. \end{aligned}$$

The covariation satisfies

$$[X_i, X_j]_t = \int_0^t \sum_k \sigma_{i,k}(X(s))\sigma_{j,k}(X(s))ds = \int_0^t a_{i,j}(X(s))ds,$$

where $a = ((a_{i,j})) = \sigma \cdot \sigma^T$, that is $a_{i,j}(x) = \sum_k \sigma_{ik}(x)\sigma_{kj}(x)$. If we denote

$$Lf(x) = \sum_{i=1}^d b_i(x)\partial_i f(x) + \frac{1}{2}\sum_{i,j} a_{i,j}(x)\partial_i\partial_j f(x),$$

then

$$\begin{aligned} f(X(t)) &= f(X(0)) + \int_0^t \nabla f^T(X(s))\sigma(X(s))dW(s) \\ &\quad + \int_0^t Lf(X(s))ds\ . \end{aligned}$$

Since

$$a = \sigma \cdot \sigma^T,$$

we have

$$\sum \xi_i \xi_j a_{i,j} = \xi^T \sigma\sigma^T \xi = |\sigma^T \xi|^2 \ge 0,$$

and hence a is nonnegative definite. Consequently, L is an elliptic differential operator. L is called the *differential generator* or simply the *generator* for the corresponding diffusion process.

Example 8.1 *If*

$$X(t) = X(0) + W(t),$$

then $((a_{i,j}(x))) = I$, *and hence* $Lf(x) = \frac{1}{2}\Delta f(x)$.

8.2 Exit distributions in one dimension.

If $d = m = 1$, then

$$Lf(x) = \frac{1}{2}a(x)f''(x) + b(x)f'(x)$$

where

$$a(x) = \sigma^2(x).$$

Find f such that $Lf(x) = 0$ (i.e., solve the linear first order differential equation for f'). Then

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))\sigma(X(s))dW(s)$$

is a local martingale. Fix $a < b$, and define $\tau = \inf\{t : X(t) \notin (a,b)\}$. If $\sup_{a<x<b} |f'(x)\sigma(x)| < \infty$, then

$$f(X(t\wedge\tau)) = f(X(0)) + \int_0^t 1_{[0,\tau)}(s)f'(X(s))\sigma(X(s))dW(s)$$

is a martingale, and

$$E[f(X(t\wedge\tau))|X(0) = x] = f(x).$$

Moreover, if we assume $\sup_{a<x<b} f(x) < \infty$ and $\tau < \infty$ a.s. then letting $t \to \infty$ we have

$$E[f(X(\tau))|X(0) = x] = f(x).$$

Hence

$$f(a)P(X(\tau) = a|X(0) = x) + f(b)P(X(\tau) = b|X(0) = x) = f(x),$$

and therefore the probability of exiting the interval at the right endpoint is given by

$$P(X(\tau) = b|X(0) = x) = \frac{f(x) - f(a)}{f(b) - f(a)}. \tag{8.1}$$

To find conditions under which $P(\tau < \infty) = 1$, or more precisely, under which $E[\tau] < \infty$, solve $Lg(x) = 1$. Then

$$g(X(t)) = g((X(0)) + \int_0^t g'(X(s))\sigma(X(s))dW(s) + t,$$

and assuming $\sup_{a<x<b} |g'(x)\sigma(x)| < \infty$, we conclude that the stochastic integral in

$$g(X(t\wedge\tau)) = g(x) + \int_0^{t\wedge\tau} g'(X(s))\sigma(X(s))dW(s) + t\wedge\tau$$

is a martingale and hence

$$E[g(X(t\wedge\tau))|X(0) = x] = g(x) + E[t\wedge\tau].$$

If

$$C = \sup_{a\le x\le b} |g(x)| < \infty,$$

then $2C \ge E[t\wedge\tau]$, so $2C \ge E[\tau]$, which implies $\tau < \infty$ a.s. By (8.1), we also have

$$\begin{aligned} E[\tau|X(0) = x] &= E[g(X(\tau))|X(0) = x] - g(x) \\ &= g(b)\frac{f(x) - f(a)}{f(b) - f(a)} + g(a)\frac{f(b) - f(x)}{f(b) - f(a)} - g(x) \end{aligned}$$

8.3 Dirichlet problems.

In the one-dimensional case, we have demonstrated how solutions of a boundary value problem for L were related to quantities of interest for the diffusion process. We now consider the more general Dirichlet problem

$$\begin{cases} Lf(x) = 0 & x \in D \\ f(x) = h(x) & x \in \partial D \end{cases} \tag{8.2}$$

for $D \subset \mathbb{R}^d$.

Definition 8.2 *A function f is* Hölder continuous *with Hölder exponent $\delta > 0$ if*

$$|f(x) - f(y)| \leq L|x - y|^{\delta}$$

for some $L > 0$.

Theorem 8.3 *Suppose D is a bounded, smooth domain,*

$$\inf_{x \in D} \sum a_{i,j}(x)\xi_i\xi_j \geq \varepsilon|\xi|^2,$$

where $\varepsilon > 0$, and $a_{i,j}$, b_i, and h are Hölder continuous. Then there exists a unique C^2 solution f of the Dirichlet problem (8.2).

To emphasize dependence on the initial value, let

$$X(t,x) = x + \int_0^t \sigma(X(s,x))dW(s) + \int_0^t b(X(s,x))ds. \tag{8.3}$$

Define $\tau = \tau(x) = \inf\{t : X(t,x) \notin D\}$. If f is C^2 and bounded and satisfies (8.2), we have

$$f(x) = E[f(X(t \wedge \tau, x))],$$

and assuming $\tau < \infty$ a.s., $f(x) = E[f(X(\tau,x))]$. By the boundary condition

$$f(x) = E[h(X(\tau,x))] \tag{8.4}$$

giving a useful representation of the solution of (8.2). Conversely, we can define f by (8.4), and f will be, at least in some weak sense, a solution of (8.2). Note that if there is a C^2, bounded solution f and $\tau < \infty$, f must be given by (8.4) proving uniqueness of C^2, bounded solutions.

8.4 Harmonic functions.

If $\Delta f = 0$ (i.e., f is harmonic) on $\mathbb{R}^d$, and W is standard Brownian motion, then $f(x+W(t))$ is a martingale (at least a local martingale).

8.5 Parabolic equations.

Suppose u is bounded and satisfies

$$\begin{cases} u_t = Lu \\ u(0,x) = f(x). \end{cases}$$

By Itô's formula, for a smooth function $v(t,x)$,

$$v(t,X(t)) = v(0,X(0)) + (local)\ martingale + \int_0^t [v_s(s,X(s)) + Lv(s,X(s))]ds.$$

For fixed $r > 0$, define

$$v(t,x) = u(r-t,x).$$

Then $\frac{\partial}{\partial t}v(t,x) = -u_1(r-t,x)$, where $u_1(t,x) = \frac{\partial}{\partial t}u(t,x)$. Since $u_1 = Lu$ and $Lv(t,x) = Lu(r-t,x)$, $v(t,X(t))$ is a martingale. Consequently,

$$E[u(r-t,X(t,x))] = u(r,x),$$

and setting $t = r$, $E[u(0,X(r,x))] = u(r,x)$, that is ,we have

$$u(r,x) = E[f(X(r,x))].$$

8.6 Properties of $X(t,x)$.

Assume now that

$$|\sigma(x) - \sigma(y)| \le K|x-y|, \quad |b(x) - b(y)| \le K|x-y|$$

for some constant K. By arguments similar to those of Section 7.6, we can obtain the estimate

$$E[|X(t,x) - X(t,y)|^n] \le C(t)|x-y|^n. \tag{8.5}$$

Consequently, we have the following

Theorem 8.4 *There is a version of $X(t,x)$ such that the mapping $(t,x) \to X(t,x)$ is contínous a.s.*

Proof. The proof is based on Kolmogorov's criterion for continuity of processes indexed by $\mathbb{R}^d$. □

8.7 Markov property.

Given a filtration $\{\mathcal{F}_t\}$, W is called an $\{\mathcal{F}_t\}$-*standard Brownian motion* if

1) W is $\{\mathcal{F}_t\}$-adapted

2) W is a standard Brownian motion

3) $W(r+\cdot) - W(r)$ is independent of $\mathcal{F}_r$.

For example, if W is an $\{\mathcal{F}_t\}$-Brownian motion, then

$$E[f(W(t+r) - W(r))|\mathcal{F}_r] = E[f(W(t))].$$

Let $W_r(t) \equiv W(r+t) - W(r)$. Note that W_r is an $\{\mathcal{F}_{r+t}\}$- Brownian motion. We have

$$\begin{aligned} X(r+t,x) &= X(r,x) + \int_r^{r+t} \sigma(X(s,x))dW(s) + \int_r^{r+t} b(X(s,x))ds \\ &= X(r,x) + \int_0^t \sigma(X(r+s,x))dW_r(s) \\ &\quad + \int_0^t b(X(r+s,x))ds. \end{aligned}$$

Define $X_r(t,x)$ such that

$$X_r(t,x) = x + \int_0^t \sigma(X_r(s,x))dW_r(s) + \int_0^t b(X_r(s,x))ds.$$

Then $X(r+t,x) = X_r(t,X(r,x))$. Intuitively, $X(r+t,x) = H_t(X(r,x),W_r)$ for some function H, and by the independence of $X(r,x)$ and W_r,

$$\begin{aligned} E[f(X(r+t,x))|\mathcal{F}_r] &= E[f(H_t(X(r,x),W_r))|\mathcal{F}_r] \\ &= u(t,X(r,x)), \end{aligned}$$

where $u(t,z) = E[H_t(z,W_r)]$. Hence

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

that is, the Markov property holds for X.

To make this calculation precise, define

$$\eta_n(t) = \frac{k}{n}, \quad \text{for } \frac{k}{n} \le t < \frac{k+1}{n},$$

and let

$$X^n(t,x) = x + \int_0^t \sigma(X^n(\eta_n(s),x))dW(s) + \int_0^t b(X^n(\eta_n(s),x))ds.$$

Suppose that $z \in C_{\mathbb{R}^m}[0,\infty)$. Then

$$H^n(t,x,z) = x + \int_0^t \sigma(H^n(\eta_n(s),x,z))dz(s) + \int_0^t b(H^n(\eta_n(s),x,z))ds$$

is well-defined. Note that $X^n(t,x) = H^n(t,x,W)$.

We also have

$$\begin{aligned} X(r+t,x) &= X_r(t,X(r,x)) \\ &= \lim_{n\to\infty} X_r^n(t,X(r,x)) \\ &= \lim_{n\to\infty} H^n(t,X(r,x),W_r). \end{aligned}$$

and it follows that

$$\begin{aligned} E[f(X(r+t,x))|\mathcal{F}_r] &= \lim_{n\to\infty} E[f(H^n(t,X(r,x),W_r)|\mathcal{F}_r] \\ &= \lim_{n\to n} E[f(H^n(t,X(r,x),W_r)|X(r,x)] \\ &= E[f(X(r+t,x))|X(r,x)]\ . \end{aligned}$$

8.8 Strong Markov property.

Theorem 8.5 *Let W be an $\{\mathcal{F}_t\}$-Brownian Motion and let τ be an $\{\mathcal{F}_t\}$ stopping time. Define $\mathcal{F}_t^\tau = \mathcal{F}_{\tau+t}$. Then $W_\tau(t) \equiv W(\tau+t) - W(\tau)$ is an $\{\mathcal{F}_t^\tau\}$ Brownian Motion.*

Proof. Let

$$\tau_n = \frac{k+1}{n}, \quad \text{when } \frac{k}{n} \le \tau < \frac{k+1}{n}.$$

Then clearly $\tau_n > \tau$. We claim that

$$E[f(W(\tau_n+t) - W(\tau_n))|\mathcal{F}_{\tau_n}] = E[f(W(t))].$$

Measurability is no problem, so we only need to check that for $A \in \mathcal{F}_{\tau_n}$

$$\int_A f(W(\tau_n+t) - W(\tau_n))dP = P(A)E[f(W(t))].$$

Observe that $A \cap \{\tau_n = k/n\} \in \mathcal{F}_{k/n}$. Thus

$$\begin{aligned} \text{LHS} &= \sum_k \int_{A\cap\{\tau_n=k/n\}} f(W(\frac{k}{n}+t) - W(\frac{k}{n}))dP \\ &= \sum_k P(A\cap\{\tau_n = k/n\})E[f(W(\frac{k}{n}+t) - W(\frac{k}{n}))] \\ &= \sum_k P(A\cap\{\tau_n = k/n\})E[f(W(t))] \\ &= E[f(W(t))]P(A). \end{aligned}$$

Note also that $\mathcal{F}_{\tau_n} \supset \mathcal{F}_\tau$. Thus

$$E[f(W(\tau_n+t) - W(\tau_n))|\mathcal{F}_\tau] = E[f(W(t))].$$

Let $n \to \infty$ to get

$$E[f(W(\tau+t) - W(\tau))|\mathcal{F}_\tau] = E[f(W(t))]. \tag{8.6}$$

Since $\tau + s$ is a stopping time, (8.6) holds with τ replaced by $\tau + s$ and it follows that W_τ has independent Gaussian increments and hence is a Brownian motion. □

Finally, consider

$$X(\tau+t,x) = X(\tau,x) + \int_0^t \sigma(X(\tau+s,x))dW_\tau(s) + \int_0^t b(X(\tau+s,x))ds.$$

By the same argument as for the Markov property, we have

$$E[f(X(\tau+t,x))|\mathcal{F}_\tau] = u(t, X(\tau,x))$$

where $u(t,x) = E[f(t,x)]$. This identity is the strong Markov property.

8.9 Equations for probability distributions.

We have now seen several formulas and assertions of the general form:

$$f\left(X(t)\right)-\int_0^t Lf\left(X(s)\right)ds \tag{8.7}$$

is a (local) martingale for all f in a specified collection of functions which we will denote $\mathcal{D}(L)$, the domain of L. For example, if

$$dX=\sigma(X)dW+b(X)dt$$

and

$$Lf(x)=\frac{1}{2}\sum_i a_{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}f(x)+\sum b_i(x)\frac{\partial}{\partial x_i}f(x) \tag{8.8}$$

with

$$((a_{ij}(x)))=\sigma(x)\sigma^T(x),$$

then (8.7) is a martingale for all $f\in C_c^2\quad(=\mathcal{D}(L))$. ($C_c^2$ denotes the C^2 functions with compact support.)

Markov chains provide another example. Suppose

$$X(t)=X(0)+\sum lY_l\left(\int_0^t\beta_l\left(X(s)\right)ds\right)$$

where $\{Y_l\}$ are independent unit Poisson processes. Define

$$Qf(x)=\beta_l(x)\sum_l\left(f(x+l)-f(x)\right).$$

Then

$$f\left(X(t)\right)-\int_0^t Qf\left(X(s)\right)ds$$

is a (local) martingale.

Since $f\left(X(t)\right)-\int_0^t Lf\left(X(s)\right)ds$ is a martingale,

$$\begin{aligned}E\left[f(X(t))\right] &= E\left[f(X(0)\right]+E\left[\int_0^t Lf(X(s))ds\right]\\ &= E\left[f(X(0))\right]+\int_0^t E\left[Lf(X(s))\right]ds.\end{aligned}$$

Let $\nu_t(\Gamma)=P\{X(t)\in\Gamma\}$. Then for all f in the domain of L, we have the identity

$$\int fd\nu_t=\int fd\nu_0+\int_0^t\int Lfd\nu_s ds, \tag{8.9}$$

which is a weak form of the equation

$$\frac{d}{dt}\nu_t=L^*\nu_t. \tag{8.10}$$

Theorem 8.6 *Let Lf be given by (8.8) with a and b continuous, and let $\{\nu_t\}$ be probability measures on $\mathbb{R}^d$ satisfying (8.9) for all $f \in C_c^2(\mathbb{R}^d)$. If $dX = \sigma(x)dW + b(x)dt$ has a unique solution for each initial condition, then $P\{X(0) \in \cdot\} = \nu_0$ implies $P\{X(t) \in \cdot\} = \nu_t$.*

In nice situations, $\nu_t(dx) = p_t(x)dx$. Then L^* should be a differential operator satisfying

$$\int_{\mathbb{R}^d} pLf\,dx = \int_{\mathbb{R}^d} fL^*p\,dx.$$

Example 8.7 *Let d=1. Integrating by parts, we have*

$$\begin{aligned}&\int_{-\infty}^{\infty} p(x)\left(\frac{1}{2}a(x)f''(x) + b(x)f'(x)\right)dx\\ &\quad= \frac{1}{2}p(x)a(x)f'(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'(x)\left(\frac{1}{2}\frac{d}{dx}(a(x)p(x)) - b(x)p(x)\right)dx.\end{aligned}$$

The first term is zero, and integrating by parts again we have

$$\int_{-\infty}^{\infty} f(x)\frac{d}{dx}\left(\frac{1}{2}\frac{d}{dx}(a(x)p(x)) - b(x)p(x)\right)dx$$

so

$$L^*p = \frac{d}{dx}\left(\frac{1}{2}\frac{d}{dx}(a(x)p(x)) - b(x)p(x)\right).$$

Example 8.8 *Let $Lf = \frac{1}{2}f''$ (Brownian motion). Then $L^*p = \frac{1}{2}p''$, that is, L is self adjoint.*

8.10 Stationary distributions.

Suppose $\int Lf\,d\pi = 0$ for all f in the domain of L. Then

$$\int f\,d\pi = \int f\,d\pi + \int_0^t \int Lf\,d\pi ds,$$

and hence $\nu_t \equiv \pi$ gives a solution of (8.9). Under appropriate conditions, in particular, those of Theorem 8.6, if $P\{X(0) \in \cdot\} = \pi$ and $f(X(t)) - \int_0^t Lf(X(s))ds$ is a martingale for all $f \in \mathcal{D}(L)$, then we have $P\{X(t)\} \in \cdot\} = \pi$, i.e. π is a stationary distribution for X.

Let $d = 1$. Assuming $\pi(dx) = \pi(x)dx$

$$\frac{d}{dx}\underbrace{\left(\frac{1}{2}\frac{d}{dx}(a(x)\pi(x)) - b(x)\pi(x)\right)}_{\substack{\text{this is a constant:}\\ \text{let the constant be 0}}} = 0,$$

so we have

$$\frac{1}{2}\frac{d}{dx}(a(x)\pi(x)) = b(x)\pi(x).$$

Applying the integrating factor $\exp(-\int_0^x 2b(z)/a(z)dz)$ to get a perfect differential, we have

$$\frac{1}{2}e^{-\int_0^x \frac{2b(z)}{a(z)}dz}\frac{d}{dx}\left(a(x)\pi(x)\right) - b(x)e^{-\int_0^x \frac{2b(z)}{a(z)}dz}\pi(x) = 0$$

$$a(x)e^{-\int_0^x \frac{2b(z)}{a(z)}dz}\pi(x) = C$$

$$\pi(x) = \frac{C}{a(x)}e^{\int_0^x \frac{2b(z)}{a(z)}dz}.$$

Assume $a(x) > 0$ for all x. The condition for the existence of a stationary distribution is

$$\int_{-\infty}^{\infty}\frac{1}{a(x)}e^{\int_0^x \frac{2b(z)}{a(z)}dz}dx < \infty.$$

8.11 Diffusion with a boundary.

(See Section 11.) Suppose

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

with $X(t) \geq 0$, and that Λ is nondecreasing and increasing only when $X(t) = 0$. Then

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

is a martingale, if $f \in C_c^2$ *and* $f'(0) = 0$.

$$\begin{aligned}\int_0^\infty p(x)Lf(x)dx &= \underbrace{\left[\frac{1}{2}p(x)a(x)f'(x)\right]_0^\infty}_{=0} - \int_0^\infty f'\left(\frac{1}{2}\frac{d}{dx}\left(a(x)p(x)\right) - b(x)p(x)\right)dx \\ &= \left[-f(x)\left(\frac{1}{2}\frac{d}{dx}\left(a(x)p(x)\right) - b(x)p(x)\right)\right]_0^\infty + \int_0^\infty f(x)L^*p(x)dx\end{aligned}$$

and hence

$$L^*p(x) = \frac{d}{dx}\left(\frac{1}{2}\frac{d}{dx}\left(a(x)p(x)\right) - b(x)p(x)\right)$$

for p satisfying

$$\left(\frac{1}{2}a'(0) - b(0)\right)p(0) + \frac{1}{2}a(0)p'(0) = 0\ . \tag{8.11}$$

The density for the distribution of the process should satisfy

$$\frac{d}{dt}p_t = L^*p_t$$

and the stationary density satisfies

$$\frac{d}{dx}\left(\frac{1}{2}\frac{d}{dx}\left(a(x)\pi(x)\right) - b(x)\pi(x)\right) = 0$$

subject to the boundary condition (8.11). The boundary condition implies

$$\frac{1}{2}\frac{d}{dx}\left(a(x)\pi(x)\right) - b(x)\pi(x) = 0$$

and hence

$$\pi(x) = \frac{c}{a(x)} e^{\int_0^x \frac{2b(z)}{a(z)} dz}, \ x \geq 0.$$

Example 8.9 *(Reflecting Brownian motion.) Let $X(t) = X(0) + \sigma W(t) - bt + \Lambda(t)$, where $a = \sigma^2$ and $b > 0$ are constant. Then*

$$\pi(x) = \frac{2b}{\sigma^2} e^{-\frac{2b}{\sigma^2} x},$$

so the stationary distribution is exponential.

9 Poisson random measures

9.1 Poisson random variables

A random variable X has a *Poisson distribution* with parameter $\lambda > \mathbf{0}$ (we write $X \sim Poisson(\lambda)$) if for each $k \in \{0, 1, 2, \ldots\}$

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

From the definition it follows that $E[X] = \lambda$, $Var(X) = \lambda$ and the characteristic function of X is given by

$$E[e^{i\theta X}] = e^{\lambda(e^{i\theta}-1)}.$$

Since the characteristic function of a random variable characterizes uniquely its distribution, a direct computation shows the following fact.

Proposition 9.1 *If $X_1, X_2, \ldots$ are independent random variables with $X_i \sim Poisson(\lambda_i)$ and $\sum_{i=1}^{\infty} \lambda_i < \infty$, then*

$$X = \sum_{i=1}^{\infty} X_i \sim Poisson\left(\sum_{i=1}^{\infty} \lambda_i\right)$$

Proof. Since for each $i \in \{1, 2, \ldots\}$, $P(X_i \geq 0) = 1$, it follows that $\sum_{i=0}^{k} X_i$ is an increasing sequence in k. Thus, $X \equiv \sum_{i=1}^{\infty} X_i$ exists. By the monotone convergence theorem

$$E[X] = \sum_{i=1}^{\infty} E[X_i] = \sum_{i=0}^{\infty} \lambda_i < \infty,$$

and X is finite almost surely. Fix $k \geq 1$. Then

$$E[e^{i\theta X}] = \lim_{k\to\infty} E\left[e^{i\theta} \sum_{i=1}^{k} X_i\right] = \lim_{k\to\infty} e^{(\sum_{i=1}^{k} \lambda_i)(e^{i\theta}-1)} = e^{(\sum_{i=1}^{\infty} \lambda_i)(e^{i\theta}-1)},$$

and hence $X \sim Poisson\,(\sum_{i=1}^{\infty} \lambda_i)$. □

Suppose in the last proposition $\sum_{i=1}^{\infty} \lambda_i = \infty$. Then

$$P\{X \leq n\} = \lim_{k\to\infty} P\left\{\sum_{i=1}^{k} X_i \leq n\right\} = \lim_{k\to\infty} \sum_{i=0}^{n} \frac{1}{i!} \left(\sum_{j=1}^{k} \lambda_j\right) \exp\left\{-\sum_{j=1}^{k} \lambda_j\right\} = 0.$$

Thus $P\{X \leq n\} = 0$ for every $n \geq 0$. In other words, $P\{X < \infty\} = 0$, and $\sum_{i=1}^{\infty} X_i \sim Poisson(\infty)$. From this we conclude the following result.

Corollary 9.2 *If $X_1, X_2, \ldots$ are independent random variables with $X_i \sim Poisson(\lambda_i)$ then*

$$\sum_{i=1}^{\infty} X_i \sim Poisson\left(\sum_{i=1}^{\infty} \lambda_i\right)$$

9.2 Poisson sums of Bernoulli random variables

A random variable Y is said to be a Bernoulli with parameter $p \in \{0,1]$ (we write $Y \sim Bernoulli(p)$) if

$$P\{Y=1\} = p\,,\ P\{Y=0\} = 1-p.$$

Proposition 9.3 *Let $N \sim Poisson(\lambda)$, and suppose that $Y_1, Y_2, \dots$ are i.i.d. Bernoulli random variables with parameter $p \in [0,1]$. If N is independent of the Y_i, then $\sum_{i=0}^{N} Y_i \sim Poisson(\lambda p)$.*

Next, we present a natural generalization of the previous fact. For $j = 1, \dots, m$, let e_j be the vector in $\mathbb{R}^m$ that has all its entries equal to zero, except for the jth which is 1.

For $\theta, y \in \mathbb{R}^m$, let

$$\langle \theta, y \rangle = \sum_{j=1}^{m} \theta_j y_j.$$

Let $Y = (Y_1, ..., Y_m)$, where the Y_j are independent and s $Y_j \sim Poisson(\lambda_j)$. Then the characteristic function of Y has the form

$$E[e^{i\langle \theta, Y\rangle}] = \exp\left\{\sum_{j=0}^{m} \lambda_j (e^{i\theta_j} - 1)\right\}$$

Noting, as before, that the characteristic function of a $\mathbb{R}^m$-valued random variable determines its distribution we have the following:

Proposition 9.4 *Let $N \sim Poisson(\lambda)$. Suppose that $Y_0, Y_1, \dots$ are independent $\mathbb{R}^m$-valued random variables such that for all $k \geq 0$ and $j \in \{1, \dots, m\}$*

$$P\{Y_k = e_j\} = p_j,$$

where $\sum_{j=1}^{m} p_j = 1$. Define $X = (X_1, ..., X_m) = \sum_{k=0}^{N} Y_k$. If N is independent of the Y_k, then $X_1, \dots, X_m$ are independent random variables and $X_j \sim Poisson(\lambda p_j)$.

Proof. Define $X = (X_1, \dots, X_m)$. Then, for arbitrary $\theta \in \mathbb{R}^m$, it follows that

$$\begin{aligned} E[e^{i\langle \theta, X\rangle}] &= \sum_{k \geq 0} E\left(\exp\left\{i \sum_{j=1}^{k} \langle Y_j, \theta \rangle\right\}\right) \cdot P\{N=k\} \\ &= e^{-\lambda} \sum_{k\geq 0} \left[E(e^{i\langle \theta, Y_1\rangle})\right]^k \cdot \frac{\lambda^k}{k!} \\ &= \exp\left\{\sum_{j=1}^{m} \lambda p_j (e^{i\theta_j} - 1)\right\} \end{aligned}$$

From the last calculation we see that the coordinates of X must be independent and $X_j \sim Poisson(\lambda p_j)$ as desired. □

9.3 Poisson random measures

Let $(E, \mathcal{E})$ be a measurable space, and let ν be a σ-finite measure on $\mathcal{E}$. Let $\mathcal{N}(E)$ be the collection of counting measures, that is, measures with nonnegative integer values, on E. ξ is an $\mathcal{N}(E)$-valued random variable on a probability space $(\Omega, \mathcal{F}, P)$ if for each $\omega \in \Omega$, $\xi(\omega, \cdot) \in \mathcal{N}(E)$ and for each $A \in \mathcal{E}$, $\xi(A)$ is a random variable with values in $\mathbb{N} \cup \{\infty\}$. (For convenience, we will write $\xi(A)$ instead of $\xi(\omega, A)$.)

An $\mathcal{N}(E)$-valued random variable is a Poisson random measure with mean measure ν if

(i) For each $A \in \mathcal{E}$, $\xi(A) \sim Poisson(\nu(A))$.

(ii) If $A_1, A_2, \ldots \in \mathcal{E}$ are disjoint then $\xi(A_1), \xi(A_2), \ldots$ are independent random variables.

Clearly, ν determines the distribution of ξ provided ξ exists. We first show existence for ν finite, and then we consider ν σ-finite.

Proposition 9.5 *Suppose that ν is a measure on $(E, \mathcal{E})$ such that $\nu(E) < \infty$. Then there exists a Poisson random measure with mean measure ν.*

Proof. The case $\nu(E) = 0$ is trivial, so assume that $\nu(E) \in (0, \infty)$. Let N be a Poisson random variable with defined on a probability space $(\Omega, \mathcal{F}, P)$ with $E[N] = \nu(E)$. Let $X_1, X_2, \ldots$ be iid E-valued random varible such that for every $A \in \mathcal{E}$

$$P\{X_j \in A\} = \frac{\nu(A)}{\nu(E)},$$

and assume that N is independent of the X_j.

Define ξ by $\xi(A) = \sum_{k=0}^{N} \mathbf{1}_{\{X_k(\omega) \in A\}}$. In other words

$$\xi = \sum_{k=0}^{N} \delta_{X_k}$$

where, for each $x \in E$, δ_x is the Dirac mass at x.

Clearly, for each ω, ξ is a counting measure on $\mathcal{E}$. To conclude that ξ is a Poisson random measure, it is enough to check that given disjoint sets $A_1, \ldots, A_m \in \mathcal{E}$ such that $\cup_{i=1}^{m} A_i = E$, $\xi(A_1), \ldots, \xi(A_m)$ are independent and $\xi(A_i) \sim Poisson(\nu(A_i))$. For this, define the $\mathbb{R}^m$-valued random vectors

$$Z_j = (\mathbf{1}_{\{X_j \in A_1\}}, \ldots, \mathbf{1}_{\{X_j \in A_m\}}).$$

Note that, for every $j \geq 0$ and $i \in \{1, \ldots, m\}$, $P\{Z_j = e_i\} = \frac{\nu(A_i)}{\nu(E)}$, since $A_1, \ldots, A_m$ partition E. Since N and the X_j are mutually independent, it follows that N and the Z_j are also. Finally, since

$$(\xi(A_1), \ldots, \xi(A_m)) = \sum_{j=1}^{N} Z_j,$$

by Proposition 9.4, we conclude that $\xi(A_1), \ldots, \xi(A_m)$ are independent random variables and $\xi(A_i) \sim Poisson(\nu(A_i))$. □

The existence of a Poisson random measure in the σ-finite case is a simple consequence of the following kind of superposition result.

Proposition 9.6 *Suppose that $\nu_1, \nu_2, \dots$ are finite measures defined on $\mathcal{E}$, and that $\nu = \sum_{i=1}^{\infty} \nu_i$ is σ-finite. For $k = 1, 2, \dots$, let ξ_k be a Poisson random measure with mean measure ν_k, and assume that $\xi_1, \xi_2, \dots$ are independent. Then $\xi = \sum_{k=1}^{n} \xi_k$ defines a Poisson random measure with mean measure ν.*

Proof. By Proposition 9.5, for each $i \geq 1$ there exists a probability space $(\Omega_i, \mathcal{F}_i, P_i)$ and a Poisson random measure ξ_i on $(\Omega_i, \mathcal{F}_i, P_i)$ with mean measure ν_i. Consider the product space $(\Omega, \mathcal{F}, P)$ where

$$\begin{aligned} \Omega &= \Omega_1 \times \Omega_2 \times \dots \\ \mathcal{F} &= \mathcal{F}_1 \times \mathcal{F}_2 \times \dots \\ P &= P_1 \times P_2 \times \dots \end{aligned}$$

Note that any random variable X_i defined on $(\Omega_i, \mathcal{F}_i, P_i)$ can be viewed as a random variable on $(\Omega, \mathcal{F}, P)$ by setting $X_i(\omega) = X_i(\omega_i)$. We claim the following:

(a) for $A \in \mathcal{E}$ and $i \geq 1$, $\xi_i(A) \sim Poisson(\nu_i(A))$.

(b) if $A_1, A_2, \dots \in \mathcal{E}$ then $\xi_1(A_1), \xi_2(A_2), \dots$ are independent random variables.

(c) $\xi(A) = \sum_{i=1}^{\infty} \xi_i(A)$ is a Poisson random measure with mean measure ν.

(a) and (b) are direct cosequences of the definitions. For (c), first note that ξ is a counting measure in $\mathcal{E}$ for each fixed ω. Moreover, from (a), (b) and Corollary 9.2, we have that $\xi(A) \sim Poisson(\nu(A))$. Now, suppose that $B_1, B_2, \dots \in \mathcal{E}$ are disjoint sets. Then, by (b) it follows that the random variables $\xi_1(B_1), \xi_2(B_1), \dots, \xi_1(B_2), \xi_2(B_2), \dots, \xi_1(B_n), \xi_2(B_n), \dots$ are independent. Consequently, $\xi(B_1), \xi(B_2), \dots$ are independet, and therefore, ξ is a Poisson random measure with mean ν. $\square$

Suppose now that ν is a σ-finite measure. By definition, there exist disjoint E_i such that $E = \cup_{i=1}^{\infty} E_i$ and $\nu(E_i) < \infty$ for all $i \geq 1$. Now, for each $i \geq 1$, consider the measure ν_i defined on $\mathcal{E}$ by the formula $\nu_i(A) = \nu(A \cap E_i)$. Clearly each measure ν_i is finite and $\nu = \sum_{i=1}^{\infty} \nu_i$. Therefore, by Proposition 9.6 we have the following

Corollary 9.7 *Suppose that ν is a σ-finite measure defined on $\mathcal{E}$. Then there exists a Poisson random measure with mean measure ν.*

9.4 Integration w.r.t. a Poisson random measure

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(E, \mathcal{E})$ be a measurable space. Let ν be a σ-finite measure on $(E, \mathcal{E})$, and let ξ be a Poisson random measure with mean measure ν. Recall that for each $\omega \in \Omega$, $\xi(\omega, \cdot)$ is a counting measure on $\mathcal{E}$. If $f : E \to \mathbb{R}$ is a measurable function with $\int |f| d\nu < \infty$, then we claim that

$$\omega \to \int_E f(x) \xi(\omega, dx)$$

is a $\mathbb{R}$-valued random variable. Consider first simple functions defined on E, that is, $f = \sum_{j=1}^{n} c_j \mathbf{1}_{A_j}$, where $n \in \mathbb{N}$, $c_1, \dots, c_n \in \mathbb{R}$, and $A_1, \dots, A_n \in \mathcal{E}$ are such that $\nu(A_j) < \infty$ for all $j \in \{1, \ldots, n\}$. Then

$$X_f(\omega) = \int_E f(x)\xi(\omega, dx) = \sum_{j=0}^{n} c_j \xi(A_j)$$

is a random variable. Note that

$$E[X_f] = \int_E f d\nu, \quad E[|X_f|] \le \int_E |f| d\nu, \tag{9.1}$$

with equality holding if $f \ge 0$. Recall that the spaces

$$L_1(\nu) = \{h : E \to \mathbb{R}\text{: } h \text{ is measurable, and } \int_E |h| d\nu < \infty\}$$

and

$$L_1(P) = \{X : \Omega \to \mathbb{R} : X \text{ is a random variable, and } E[|X|] < \infty\}$$

are Banach spaces under the norms $\|h\| = \int_E |h| d\nu$ and $\|X\| = E[|X|]$ respectively. Since the space of simple functions defined on E is dense in $L_1(\nu)$, for $f \in L_1(\nu)$, we can construct a sequence of simple funtions f_n such that $f_n \to f$ pointwise and in L_1 and $|f_n| \le |f|$. It follows that $X_f(\omega) = \int_E f(x)\xi(\omega, dx)$ is a random variable satisfying (9.1).

As convenient, we will use any of the following to denote the integral.

$$X_f = \int_E f(x)\xi(dx) = \langle f, \xi \rangle.$$

From the interpretation of X_f as an ordinary integral, we have the following.

Proposition 9.8 *Let $f, g \in L^1(\nu)$.*

(a) If $f \le g$ ν-a.s., then $X_f \le X_g$ P-a.s.

(b) If $\alpha \in \mathbb{R}$ ν-a.s., then $X_{\alpha f} = \alpha X_f$ P-a.s.

(c) $X_{f+g} = X_f + X_g$ P-a.s.

For $A \in \mathcal{E}$, let $\mathcal{F}_A = \sigma(\xi(B) : B \in \mathcal{E}, B \subset A)$. Note that if A_1 and A_2 are disjoint, then $\mathcal{F}_{A_1}$ and $\mathcal{F}_{A_2}$ are independent, that is, if $H_1 \in \mathcal{F}_{A_1}$ and $H_2 \in \mathcal{F}_{A_2}$, then $P(H_1 \cap H_2) = P(H_1)P(H_2)$. In the proof of the previous result, we have used the following result.

Proposition 9.9 *Suppose that $f, g \in L_1(\nu)$ have disjoint supports i.e. $\int_E |f| \cdot |g| d\nu = 0$. Then X_f and X_g are independent.*

Proof. Define $A := \{|f| > 0\}$ and $B := \{|f| = 0\}$. Note that

$$X_f = \int_A f(x)\xi(dx)$$

is $\mathcal{F}_A$-measurable and

$$X_g = \int g(x)\xi(dx) = \int_B g(x)\xi(dx) \quad a.s.$$

where the right side is $\mathcal{F}_B$-measurable. Since $\mathcal{F}_A$ and $\mathcal{F}_B$ are independent, it follows that X_f and X_g are independent. □

If ν has no atoms, then the support of $\xi(\omega, \cdot)$ has ν measure zero. Consequently, the following simple consequence of the above observations may at first appear surprising.

Proposition 9.10 *If $f, g \in L^1(\nu)$ then*

$$f = g \ \nu\text{-a.s. if and only if } X_f = X_g \ P\text{-a.s.}$$

Proof. That the condition is sufficient follows directly from the linearity of X_f and (9.1). For the converse, without lost of generality, we only need to prove that $X_f = 0$ P-a.s. implies that $f = 0$ ν-a.s. Since $f = f^+ - f^-$ where $f^+ = f\mathbf{1}_{\{f \geq 0\}}$ and $f^- = -f\mathbf{1}_{\{f<0\}}$, it follows that $0 = X_f = X_{f^+} - X_{f^-}$ a.s. Note that X_{f^+} and X_{f^-} are independent since the support of f^+ is disjoint from the support of f^-. Consequently, X_{f^+} and X_{f^-} must be equal a.s. to the same constant. Similar analysis demonstrates that the constant must be zero, and we have

$$\int_E |f| d\nu = \int_E f^+ d\nu + \int_E f^- d\nu = E[X_{f^+}] + E[X_{f^-}] = 0.$$

□

9.5 Extension of the integral w.r.t. a Poisson random measure

We are going to extend our definition of X_f to a larger class of functions f. As motivation, if ν is a finite measure, then, as we saw in the proof of Proposition 9.5, $\xi = \sum_{k=1}^N \delta_{X_k}$ is a Poisson random measure with mean ν, whenever $N \sim Poisson(\nu(E))$ is independent of the sequence of i.i.d. E-valued random variables $X_1, X_2, \ldots$ with $P(X_i \in A) = \frac{\nu(A)}{\nu(E)}$. Now, given any measurable function $f : E \to \mathbb{R}$, it is natural to define

$$\int_E f(x)\xi(dx) = \sum_{k=0}^N f(X_k),$$

and we want to ensure that this definition is consistent.

Proposition 9.11 *If $f : E \to \mathbb{R}$ is a simple function then for all $a, b > 0$*

$$P\{|X_f| \geq b\} \leq \frac{1}{b}\int_E |f| \wedge a \, d\nu + \left(1 - \exp\left\{-a^{-1}\int_E |f| \wedge a \, d\nu\right\}\right)$$

Proof. First, notice that

$$P\{|X_f| \geq b\} \leq P\{X_{|f|} \geq b\} \leq P\{X_{|f|\mathbf{1}_{\{|f| \leq a\}}} \geq b\} + P\{X_{|f|\mathbf{1}_{\{|f|>a\}}} > 0\}$$

But, by the Markov inequality,

$$P\{X_{|f|\mathbf{1}_{|f|\leq a}} \geq b\} \leq \frac{1}{b}\int_{\{|f|\leq a\}} |f|d\nu \leq \frac{1}{b}\int_E (|f|\wedge a)d\nu.$$

On the other hand, since $P\{X_{|f|\mathbf{1}_{|f|>a}} > 0\} = P\{\xi(|f| > a) > 0\}$ and $\xi\{|f| > a\} \sim Poisson(\nu\{|f| > a\})$, we have

$$P\{\xi(|f| > a) > 0\} = 1 - e^{-\nu\{|f|>a\}} \leq 1 - \exp\left\{-a^{-1}\int_E |f|\wedge a\, d\nu\right\},$$

giving the desired result. □

Consider the vector space

$$L_{1,0}(\nu) = \{f : E \to \mathbb{R} \text{ s.t. } |f|\wedge 1 \in L_1(\nu)\}.$$

Notice that $L_1(\nu) \subset L_{1,0}(\nu)$. In particular, $L_{1,0}(\nu)$ contains all simple functions defined on E whose support has finite ν measure. Moreover, if ν is a finite measure then this vector space is simple the space of all measurable functions. Given $f, g \in L_{1.0}(\nu)$, we define the distance between f and g as

$$d(f,g) = \int_E |f-g|\wedge 1\, d\nu.$$

The function d defines a metric on $L_{1,0}(\nu)$. Note that d does not come from any norm; however, it is easy to check that

(a) $d(f-g, p-q) = d(f-p, q-g)$

(b) $d(f-g, 0) = d(f,g)$

Before considering the next simple but useful result, note that

$$\int_E |f|\wedge 1\, d\nu = \int_{\{|f|\leq 1\}} |f|d\nu + \nu\{|f| > 1\}$$

Hence, a function f belongs to $L_{1,0}(\nu)$ if and only if both terms in the right side of the last equality are finite.

Proposition 9.12 *Under the metric d, the space of simple functions is dense in $L_{1,0}(\nu)$. Moreover, for every $f \in L_{1,0}(\nu)$, there exists a sequence of simple functions $\{f_n\}$, such that $|f_n| \leq |f|$ for every n, and $\{f_n\}$ converges pointwise and under d to f.*

Proof. Let $f \in L_{1,0}(\nu)$. First, suppose that $f \geq 0$, and define

$$f_n(x) = \sum_{k=0}^{n2^n-1} \mathbf{1}_{\{k2^{-n}\leq f<(k+1)2^{-n}\}} + \mathbf{1}_{\{n2^n\leq f\}}. \tag{9.2}$$

Then $\{f_n\}$ is a nonnegative increasing sequence that converges pointwise to f and the $range(f_n)$ is finite for all n. Consequently,

$$\int_E f_n d\nu = \int_{\{f\le 1\}} f_n d\nu + \int_{\{f>1\}} f_n \le \int_{\{f\le 1\}} f d\nu + n2^n \nu\{f>1\} < \infty,$$

and for each n,

$$0 \le (f - f_n) \wedge 1 \le (f \wedge 1) \in L_1(\nu).$$

Therefore, since $\lim_{n\to\infty}(f - f_n) \wedge 1 = 0$, by the bounded convergence theorem we can conclude that $\lim_{n\to\infty} d(f, f_n) = 0$.

For arbitrary $f \in L_{1,0}(\nu)$, write $f = f^+ - f^-$. Define f_n^+ and f_n^- as in (9.2), and set $f_n = f_n^+ - f_n^-$. Since $L_{1,0}(\nu)$ is linear and d is a metric,

$$d(f, f_n) \le d(f^+, f_n^+) + d(f^-, f_n^-),$$

and the proposition follows. □

Suppose that $f \in L_{1,0}(\nu)$. By Proposition 9.12 there exist a sequence of simple functions $\{f_n\}$ such that $\{f_n\}$ converges to f under d. But, from Proposition 9.11 with $a = 1$, we see that for all n, m and $b > 0$,

$$P\{|X_{f_n} - X_{f_m}| \ge b\} = P\{|X_{f_n - f_m}| \ge b\} \le \frac{1}{b} d(f_n, f_m) + 1 - e^{-d(f_n, f_m)},$$

and we conclude that the sequence $\{X_{f_n}\}$ is Cauchy in probability. Therefore, there exists a random variable X_f such that

$$X_f = \lim_{n\to\infty} X_{f_n},$$

where the last limit is in probability. As we showed in Section 9.4, this limit does not depend on the sequence of simple functions chosen to converge to f. Therefore, for every $f \in L_{1,0}(\nu)$, X_f is well-define and the definition is consistent with the previous definition for $f \in L_1(\nu)$.

Before continuing, let us consider the generality of our selection of the space $L_{1,0}(\nu)$. From Proposition 9.11, we could have considered the space

$$L_{1,a}(\nu) := \left\{ f : E \to \mathbb{R} : \, |f| \wedge a \in L^1(\nu) \right\}$$

for some value of a other than 1; however, $L_{1,a}(\nu) = L_{1,0}(\nu)$ and the corresponding metric

$$d_a(f, g) := \int_E (|f - g| \wedge a) d\nu$$

is equivalent to d.

Proposition 9.13 *If $f \in L_{1,0}(\nu)$, then for all $\theta \in \mathbb{R}$*

$$E[e^{i\theta X_f}] = \exp\left\{ \int_E (e^{i\theta f(x)} - 1) \, \nu(dx) \right\}$$

Proof. First consider simple f. Then, without loss of generality, we can write $f = \sum_{j=1}^n c_j \mathbf{1}_{A_j}$ with $\nu(A_j) < \infty$ for all $j \in \{1, \ldots, n\}$ and $A_1, \ldots, A_n$ disjoints. Since $\xi(A_1), \ldots, \xi(A_n)$ are independent and $\xi(A_j) \sim Poisson(\nu(A_j))$, we have that

$$\begin{aligned} E[e^{i\theta X_f}] = \prod_{j=1}^n E\left(e^{i\theta c_j \xi(A_j)}\right) &= \prod_{j=1}^n e^{\nu(A_j)(e^{i\theta c_j}-1)} \\ &= \exp\left\{\sum_{j=1}^n (e^{i\theta c_j} - 1)\nu(A_j)\right\} \\ &= \exp\left\{\int_E e^{i\theta f(x)} - 1 \, \nu(dx)\right\} \end{aligned}$$

The general case follows by approximating f by simple functions $\{f_n\}$ as in Proposition 9.12 and noting that both sides of the identity

$$E[e^{i\theta X_{f_n}}] = \exp\left\{\int_E e^{i\theta f_n(x)} - 1 \, \nu(dx)\right\}$$

converge by the bounded convergence theorem. □

9.6 Centered Poisson random measure

Let ξ be a Poisson random measure with mean ν. We define the *centered random measure* for ξ by

$$\tilde{\xi}(A) = \xi(A) - \nu(A), \quad A \in \mathcal{E}, \nu(A) < \infty.$$

Note that, for each $K \in \mathcal{E}$ with $\nu(K) < \infty$ and almost every $\omega \in \Omega$, the restriction of $\tilde{\xi}(\omega, \cdot)$ to K is a finite signed measure.

In the previous section, we defined $\int_E f(x)\xi(dx)$ for every $f \in L_{1,0}(\nu)$. Now, let $f = \sum_{j=1}^n c_j \mathbf{1}_{A_j}$ be a simple function with $\nu(A_j) < \infty$. Then, the integral of f with respect to $\tilde{\xi}$ is the random variable $\tilde{X}_f$ (sometimes we write $\int_E f(x)\tilde{\xi}(dx)$) defined as

$$\tilde{X}_f = \int_E f(x)\xi(dx) - \int_E f(x)\nu(dx) = \sum_{j=1}^n c_j(\xi(A_j) - \nu(A_j))$$

Note that $E[\tilde{X}_f] = 0$.

Clearly, from our definition, it follows that for simple functions f, g, and $\alpha, \beta \in \mathbb{R}$ that

$$\tilde{X}_{\alpha f + \beta g} = \alpha \tilde{X}_f + \beta \tilde{X}_g.$$

Therefore, the integral with respect to a centered Poisson random measure is a linear function on the space of simple functions. The next result is the key to extending our definition to the space $L^2(\nu) = \{h : E \to \mathbb{R} : h \text{ is measurable, and } \int_E h^2 d\nu < \infty\}$, which is a Banach space under the norm $\|h\|_2 = \left(\int_E h^2 d\nu\right)^{1/2}$. Similarly, $L^2(P) = \{X : \Omega \to \mathbb{R} : X \text{ is a random variable, and } \int_E X^2 dP < \infty\}$ is a Banach space under the norm $\|X\|_2 = \{E[X^2]\}^{1/2}$.

Proposition 9.14 *If f is a simple function, then $E[\tilde{X}_f^2] = \int_E f^2 d\nu$.*

Proof. Write $f = \sum_{j=1}^n c_j \mathbf{1}_{A_j}$, where $A_1, \ldots, A_n$ are disjoint sets with $\nu(A_j) < \infty$. Since $\xi(A_1), \ldots, \xi(A_n)$ are independent and $\xi(A_j) \sim Poisson(\nu(A_j))$, we have that

$$\begin{aligned}
E[\tilde{X}_f^2] &= E\left(\sum_{j=1}^n \sum_{i=1}^n c_j c_i (\xi(A_j) - \nu(A_j))(\xi(A_i) - \nu(A_i))\right) \\
&= \sum_{j=1}^n c_j^2 E[\xi(A_j) - \nu(A_j)]^2 \\
&= \sum_{j=1}^n c_j^2 \nu(A_j) \\
&= \int_E f^2 d\nu
\end{aligned}$$

□

The last proposition, shows that $\tilde{X}_f$ determines a linear isometry from $L^2(\nu)$ into $L^2(P)$. Therefore, since the space of simple functions is dense in $L^2(\nu)$, we can extend the definition of $\tilde{X}_f$ to all $f \in L^2(\nu)$. As in Section 9.4, if (f_n) is a sequence of simple functions that converges to f in $L^2(\nu)$, then we define

$$\tilde{X}_f = \lim_{n\to\infty} \tilde{X}_{f_n}.$$

where the limit is $L^2(P)$. Clearly, the linearity of $\tilde{X}$ over the space of simple functions is inherited by the limit. Also, for every $f \in L^2(\nu)$, we have that

$$E[\tilde{X}_f^2] = \int_E f^2 d\nu$$

and

$$E[\tilde{X}_f] = 0.$$

Before continuing, note that if f is a simple function, then $E|\tilde{X}_f| \leq 2\int_E |f| d\nu$. This inequality is enough to extend the definition of $\tilde{X}$ to the space $L^1(\nu)$, where the simple functions are also dense. Since ν is not necessarily finite, the spaces $L^1(\nu)$ and $L^2(\nu)$ are not necessarily comparable. This slitghtly different approach will in the end be irrelevant, because the space in which we are going to define $\tilde{X}$ contains both $L_2(\nu)$ and $L_1(\nu)$.

Now, we extend the definition of $\tilde{X}_f$ to a larger class of functions f. For this purpose, consider the vector space

$$L_{2,1}(\nu) = \{f : E \to \mathbb{R} : |f|^2 \wedge |f| \in L^1(\nu)\},$$

or equivalently, let $\Phi(z) = z^2 \wedge (2z - 1)$. Then

$$L_{2,1}(\nu) = L_\Phi(\nu) = \{f : E \to \mathbb{R} : \Phi(|f|) \in L^1(\nu)\}.$$

Note that $L_1(\nu) \subset L_\Phi(\nu)$ and $L_2(\nu) \subset L_\Phi(\nu)$. In particular, $L_\Phi(\nu)$ contains all the simple functions defined on E whose support has finite measure. Since Φ is convex and nondecreasing, $\Phi(0) = 0$, and Φ satisfies $\lim_{z\to\infty} z^{-1}\Phi(z) = \infty$, $L_\Phi(\nu)$ is an Orlicz space with norm

$$\|f\|_\Phi = \inf\{c : \int_E \Phi(\frac{f}{c})d\nu < 1\}.$$

As in Proposition 9.12, that the space of simple functions with support having finite measure is dense in $L_\Phi(\nu)$. The proof of this assertion follows by the same argument as in Proposition 9.12.

Proposition 9.15 *The space of simple functions is dense in $L_\Phi(\nu)$. Moreover, for every $f \in L_\Phi(\nu)$ there exists a sequence of simple functions $\{f_n\}$, such that $|f_n| \leq |f|$ for every n, and $\{f_n\}$ converges pointwise and $\|\cdot\|_\Phi$ to f.*

Proof. Take $f_n = f_n^+ - f_n^-$ as constructed in the proof of Proposition 9.12. □

Again the key to our extension is an inequality that allows us to define $\tilde{X}$ as a limit in probability instead of one in $L_2(P)$.

Proposition 9.16 *If $f : E \to \mathbb{R}$ is a simple function with support having finite measure, then*

$$E[|\tilde{X}_f|] \leq 3\|f\|_\Phi$$

Proof. Fix $a > 0$. Then, for $c > 0$ and $0 < \alpha < 1$,

$$\begin{aligned}
E[|\tilde{X}_f|] &\leq cE[|\tilde{X}_{c^{-1}f\mathbf{1}_{\{c^{-1}|f|\leq 1\}}}|] + cE[|\tilde{X}_{c^{-1}f\mathbf{1}_{\{c^{-1}|f|>1\}}}|] \\
&\leq c\sqrt{E[|\tilde{X}_{c^{-1}f\mathbf{1}_{\{c^{-1}|f|\leq 1\}}}|^2} + 2c\int_U |c^{-1}f|\mathbf{1}_{\{c^{-1}|f|>1\}}d\nu \\
&\leq c\sqrt{\int_U c^{-2}f^2\mathbf{1}_{\{c^{-1}|f|\leq 1\}}\mu(du)} + 2c\int_U |c^{-1}f|\mathbf{1}_{\{c^{-1}|f|>1\}}d\nu \\
&\leq c\sqrt{\int_U \Phi(c^{-1}|f|)d\nu} + 2c\int_U \Phi(c^{-1}|f|)d\nu
\end{aligned}$$

Taking $c = \|f\|_\Phi$, the right side is bounded by $3\|f\|_\Phi$. □

Now, suppose that $f \in L_\Phi(\nu)$. Then by Proposition 9.15 there exists a sequence $\{f_n\}$ of simple functions that converges to f in $\|\cdot\|_\Phi$. But, from Proposition 9.16, it follows that for all $n, m \geq 1$

$$P\{|\tilde{X}_{f_n} - \tilde{X}_{f_m}| \geq a\} \leq \frac{3\|f_n - f_m\|_\Phi}{a}.$$

Therefore, the sequence $\{\tilde{X}_{f_n}\}$ is Cauchy in probability, and hence, there exists a random variable $\tilde{X}_f$ such that

$$\tilde{X}_f = \lim_{n\to\infty} \tilde{X}_{f_n} \quad \text{in probability.}$$

As usual, the definition does not depend on the choice of $\{f_n\}$. Also, note that the definition of $\tilde{X}_f$ for functions $f \in L^2(\nu)$ is consistent with the definition for $f \in L_\Phi(\nu)$.

We close the present section with the calculation of the characteristic function for the random variable $\tilde{X}_f$.

Proposition 9.17 *If $f \in L_\Phi(\nu)$, then for all $\theta \in \mathbb{R}$*

$$E[e^{i\theta \tilde{X}_f}] = \exp\left\{ \int_E (e^{i\theta f(x)} - 1 - if(x))\, \nu(dx) \right\} \tag{9.3}$$

Proof. First, from Proposition 9.13, (9.3) holds for simple functions. Now, let $f \in L_\Phi(\nu)$, and let $\{f_n\}$ be as in Proposition 9.15. Since $\tilde{X}_f = \lim_{n\to\infty} \tilde{X}_{f_n}$ in probability, without lossing generality we can assume that $\{\tilde{X}_{f_n})$ converges almost surely to $\tilde{X}_f$ on Ω. Hence, for every $\theta \in \mathbb{R}$, $\lim_{n\to\infty} E[e^{i\theta \tilde{X}_{f_n}}] = E[e^{i\theta \tilde{X}_f}]$. On the other hand, since $\{f_n\}$ converges pointwise to f, it follows that for all $\theta \in \mathbb{R}$,

$$\lim_{n\to\infty} (e^{i\theta f_n} - 1 - if_n) = e^{i\theta f} - 1 - if.$$

But, there exists a constant $k > 0$ such that

$$|e^{i\theta f_n} - 1 - if_n| \le k \cdot (|f_n|^2 \wedge |f_n|) \le k \cdot (|f|^2 \wedge |f|) \in L^1(d\nu),$$

and the result follows by the dominated convergence theorem. □

9.7 Time dependent Poisson random measures

Let $(U, \mathcal{U}, \mu)$ be a measurable space where μ is a σ-finite measure, and define

$$\mathcal{A} = \{A \in \mathcal{U} : \mu(A) < \infty\}.$$

Let $\mathcal{B}[0,\infty)$ be the Borel σ-algebra of $[0,\infty)$, and denote Lebesgue measure on $\mathcal{B}[0,\infty)$ by m. Then, the product measure $\mu \times m$ is σ-finite on $\mathcal{U} \times \mathcal{B}[0,\infty)$ and therefore, by Corollary 9.7, there exists a Poisson random meaure Y, with mean measure $\mu \times m$. Denote the corresponding centered Poisson random measure by $\tilde{Y}$.

For a $A \in \mathcal{U}$ and $t \ge 0$, we write $Y(A,t)$ instead of $Y(A \times [0,t])$. Similarly, we write $\tilde{Y}(A,t)$ instead of $\tilde{Y}(A \times [0,t])$.

Proposition 9.18 *For each $A \in \mathcal{A}$, $Y(A,\cdot)$ is a Poisson process with intensity $\mu(A)$. In particular, $\tilde{Y}(A,\cdot)$ is a martingale.*

Proof. Fix $A \in \mathcal{A}$. Clearly, the process $Y(A,\cdot)$ satisfies the following properties almost surely: (i) $Y(A,0) = 0$ (ii) $Y(A,t) \sim Poisson(\mu(A)t)$ (iii) $Y(A,\cdot)$ has cadlag nondecreasing sample paths with jumps of size one. Hence, to conclude that $Y(A,\cdot)$ is a Poisson process, it is enough to check that $Y(A,t_1) - Y(A,t_0), \ldots, Y(A,t_n) - Y(A,t_{n-1})$ are independent random variables, whenever $0 = t_0 < \ldots < t_n$. But

$$Y(A,t_i) - Y(A,t_{i-1}) = Y(A \times (t_{i-1}, t_i])$$

for every $i \in \{1,\ldots,n\}$, and the sets $A \times (t_0,t_1], \ldots, A \times (t_{n-1},t_n]$ are disjoint in $U \times [0,\infty)$. Consequently, the random variables are independent, and hence $Y(A,\cdot)$ is a Poisson random process with intensity $\mu(A)$. □

Proposition 9.19 *If $A_1, A_2, \ldots \in \mathcal{A}$ are disjoint sets, then the processes $Y(A_1, \cdot), Y(A_2, \cdot), \ldots$ are independent.*

Proof. Fix $n \geq 1$ and let $0 = t_0 < \ldots < t_m$. Note that, for $i = 1, \ldots, n$ and $j = 1, \ldots m$, the random variables $Y(A_i, t_j) - Y(A_i, t_{j-1})$ are independent because the sets $A_i \times (t_{j-1}, t_j]$ are disjoint, and the independence of $Y(A_1, \cdot), Y(A_2, \cdot), \ldots$ follows. □

For each $t \geq 0$, define the σ-algebra

$$\mathcal{F}_t^Y = \sigma\left(Y(A, s) \text{ s.t. } A \in \mathcal{A} \text{ and } s \in [0, t]\right) \subset \mathcal{F}$$

By definition, $Y(A, \cdot)$ is $\{\mathcal{F}_t^Y\}$-adapted process for all $A \in \mathcal{A}$. In addition, by Proposition 9.19, for all $A \in \mathcal{A}$ and $s, t \geq 0$, $Y(A, t+s) - Y(A, t)$ is independen of $\mathcal{F}_t^Y$. This independence will play a central role in the definition of the stochastic integral with respect to Y. More generally, we will say that Y is *compatible* with the filtration $\{\mathcal{F}_t^Y\}$ if Y is adapted to $\{\mathcal{F}_t^Y\}$ and $Y(A, t+s) - Y(A, t)$ is independent of $\mathcal{F}_t^Y$ for all $A \in \mathcal{U}$, and $s.t \geq 0$.

9.8 Stochastic integrals for time-dependent Poisson random measures

Let Y be as in Section 9.7, and assume that Y is compatible with the filtration $\{\mathcal{F}_t\}$. For $\varphi \in L_{1,0}(\mu)$. define

$$Y(\varphi, t) \equiv \int_{U \times [0,t]} \varphi(u) Y(du \times ds)$$

Then $Y(\varphi, \cdot)$ is a process with independent increments and, in particular, is a $\{\mathcal{F}_t\}$-semimartingale. Suppose $\xi_1, \ldots, \xi_m$ are cadlag, $\{\mathcal{F}_t\}$-adapted processes and that $\varphi_1, \ldots, \varphi_m \in L_{1,0}(\mu)$. Then

$$Z(u, t) = \sum_{k=1}^{m} \xi_k(t) \varphi_k(u) \tag{9.4}$$

is a cadlag, $L_{1,0}(\mu)$-valued process, and we define

$$I_Z(t) = \int_{U \times [0,t]} Z(u, s-) Y(du \times ds) \equiv \sum_{k=1}^{m} \int_0^t \xi_k(s-) dY(\varphi, s). \tag{9.5}$$

Lemma 9.20 *Let $Y = \sum_i \delta_{(U_i, S_i)}$. Then for Z given by (9.4) and I_Z by (9.5), with probability one,*

$$I_Z(t) = \sum_{k=1}^{m} \sum_i \mathbf{1}_{[0,t]}(S_i) \xi_k(S_i-) \varphi_k(U_i) = \sum_i \mathbf{1}_{[0,t]}(S_i) Z(U_i, S_i-),$$

and hence,

$$I_Z(t) \leq I_{|Z|}(t). \tag{9.6}$$

Proof. Approximate φ_k by $\varphi_k^\epsilon = \varphi_k \mathbf{1}_{\{|\varphi_k| \geq \epsilon\}}$, $\epsilon > 0$. Then $Y(\{u : |\varphi_k(u)| \geq \epsilon\} \times [0, t]) < \infty$ a.s. and with φ_k replaced by φ_k^ϵ, the lemma follows easily. Letting $\epsilon \to 0$ gives the desired result. □

Lemma 9.21 *Let Z be given by (9.4) and I_Z by (9.5). If $E[\int_{U\times[0,t]}|Z(u,s)|\mu(du)ds]<\infty$, then*

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

and

$$E[|I_Z(t)|]\leq\int_{U\times[0,t]}E[|Z(u,s)|]\mu(du)ds$$

Proof. The identity follows from the martingale properties of the $Y(\varphi_i,\cdot)$, and the inequality then follows by (9.6). □

With reference to Proposition 9.11, we have the following lemma.

Lemma 9.22 *Let Z be given by (9.4) and I_Z by (9.5). Then for each stopping time τ,*

$$\begin{aligned}P\{\sup_{s\leq t}|I_Z(s)|\geq b\}\leq P\{\tau\leq t\}+\frac{1}{b}E[\int_{U\times[0,t\wedge\tau]}|Z(u,s)|\wedge a\mu(du)ds]\\ +1-E[\exp\{-a^{-1}\int_{U\times[0,t\wedge\tau]}|Z(u,s)|\wedge a\mu(du)ds\}]\end{aligned}$$

Proof. First, note that

$$P\{\sup_{s\leq t}|I_Z(s)|\geq b\}\leq P\{\tau\leq t\}+P\{\sup_{s\leq t\wedge\tau}|I_Z(s)|\geq b\}.$$

By (9.6),

$$P\{\sup_{s\leq t\wedge\tau}|I_Z(s)|\geq b\}\leq P\{I_{|Z|}(t\wedge\tau)\geq b\}\leq P\{I_{|Z|\mathbf{1}_{\{|Z|\leq a\}}}(t\wedge\tau)\geq b\}+P\{I_{|Z|\mathbf{1}_{\{|Z|>a\}}}(t\wedge\tau)>0\}.$$

But, by the Markov inequality,

$$\begin{aligned}P\{I_{|Z|\mathbf{1}_{\{|Z|\leq a\}}}(t\wedge\tau)\geq b\} &\leq \frac{1}{b}\int_{U\times[0,t]}E[|Z(u,s)|\mathbf{1}_{\{|Z(u,s)|\leq a,s\leq\tau\}}]\mu(du)ds\\ &\leq \frac{1}{b}E[\int_{U\times[0,t\wedge\tau]}|Z(u,s)|\wedge a\mu(du)ds].\end{aligned}$$

On the other hand,

$$\begin{aligned}P\{I_{|Z|\mathbf{1}_{\{|Z|>a\}}}(t\wedge\tau)>0\} &= P\{Y(\{(u,s):|Z(u,s-)|>a,s\leq\tau\})>0\}\\ &= E\left[1-\exp\left\{-\int_0^t\mu\{u:|Z(u,s)|>a\}ds\right\}\right]\end{aligned}$$

giving the desired result.

□

Lemma 9.22 gives the estimates necessary to extend the integral to cadlag and adapted, $L_{1,0}(\mu)$-valued processes.

Theorem 9.23 *If Z is a cadlag, adapted $L_{1,0}(\mu)$-valued processes, then there exist Z_n of the form (9.4) such that $\sup_{t\leq T}\int_U |Z(u,t)-Z_n(u,t)|\wedge 1\mu(du)\to 0$ in probability for each $T>0$, and there exists an adapted, cadlag process I_Z such that $\sup_{t\leq T}|I_Z(t)-I_{Z_n}(t)|\to 0$ in probability.*

Remark 9.24 *We define*

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

The estimate in Lemma 9.22 ensures that the integral is well defined.

Now consider $\tilde{Y}$. For $\varphi\in L_\Phi$, $\tilde{Y}(\varphi,t)=\int_{U\times[0,t]}\varphi(u)\tilde{Y}(du\times ds)$ is a martingale. For Z given by (9.4), but with $\varphi_k\in L_\Phi$, define

$$\tilde{I}_Z(t)=\int_{U\times[0,t]} Z(u,s-)\tilde{Y}(du\times ds)\equiv\sum_{k=1}^m\int_0^t \xi_k(s-)d\tilde{Y}(\varphi,s). \tag{9.7}$$

Then $\tilde{I}_Z$ is a local martingale with

$$[\tilde{I}_Z]_t=\int_{U\times[0,t]} Z^2(u,s-)Y(du\times ds).$$

Note that if Z has values in L_Φ, then Z^2 has values in $L_{1,0}$.

Lemma 9.25 *If $E[\int_0^t\int_U Z^2(u,s)\mu(du)ds]<\infty$, then $\tilde{I}_Z$ is a square integrable martingale with*

$$E[\tilde{I}_Z^2(t)]=E\left[\int_{U\times[0,t]} Z^2(u,s-)Y(du\times ds)\right]=E\left[\int_0^t\int_U Z^2(u,s)\mu(du)ds\right]$$

Lemma 9.26 *Let Z be given by (9.4) and I_Z by (9.5). Then for each stopping time τ,*

$$P\{\sup_{s\leq t}|\tilde{I}_Z(s)|\geq a\}\leq P\{\tau\leq t\}+\quad \frac{16}{a^2}\vee\frac{4}{a}E\left[\int_0^{t\wedge\tau}\int_U \Phi(|Z(u,s)|)\mu(du)ds\right]$$

Proof. As before,

$$P\{\sup_{s\leq t}|\tilde{I}_Z(s)|\geq a\}\leq P\{\tau\leq t\}+P\{\sup_{s\leq t\wedge\tau}|\tilde{I}_Z(s)|\geq a\}$$

Fix $a>0$. Then

$$\begin{aligned}
P\{\sup_{s\leq t\wedge\tau}|\tilde{I}_Z(s)|\geq a\} &\leq P\{\sup_{s\leq t\wedge\tau}|\tilde{I}_{Z\mathbf{1}_{\{|Z|\leq 1\}}}(s)|\geq 2^{-1}a\}+P\{\sup_{s\leq t\wedge\tau}|\tilde{I}_{Z\mathbf{1}_{\{|Z|>1\}}}(s)|\geq 2^{-1}a\}\\
&\leq \frac{16}{a^2}E\left[\int_0^{t\wedge\tau}\int_U |Z(u,s)|^2\mathbf{1}_{\{|Z(u,s)|\leq 1\}}\mu(du)\right]\\
&\qquad+\frac{4}{a}E[\int_0^{t\wedge\tau}\int_U |Z(u,s)|\mathbf{1}_{\{|Z(u,s)|>1\}}\mu(du)ds]\\
&\leq \frac{16}{a^2}\vee\frac{4}{a}E\left[\int_0^{t\wedge\tau}\int_U \Phi(|Z(u,s)|)\mu(du)ds\right].
\end{aligned}$$

□

Remark 9.27 *Lemma 9.26 gives the estimates needed to extend the definition of*

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

to all cadlag and adapted, $L_\Phi(\mu)$-valued processes Z.

Lemma 9.28 *If Z is cadlag and adapted with values in $L_1(\mu)$, then*

$$\int_{U\times[0,t]} Z(u,s-)\tilde{Y}(du\times ds) = \int_{U\times[0,t]} Z(u,s-)Y(du\times ds) - \int_0^t\int_U Z(u,s)\mu(du)ds$$

Lemma 9.29 *If $E[\int_0^t\int_U Z^2(u,s)\mu(du)ds] < \infty$, then $\tilde{I}_Z$ is a square integrable martingale with*

$$E[\tilde{I}_Z^2(t)] = E\left[\int_{U\times[0,t]} Z^2(u,s-)Y(du\times ds)\right] = E\left[\int_0^t\int_U Z^2(u,s)\mu(du)ds\right]$$

10 Limit theorems.

10.1 Martingale CLT.

Definition 10.1 *$f : D_{\mathbb{R}}[0,\infty) \to \mathbb{R}$ is continuous in the* compact uniform topology *if*

$$\sup_{t \le T} |x_n(t) - x(t)| \to 0,$$

for every $T > 0$, implies $f(x_n) \to f(x)$.

Definition 10.2 *A sequence of cadlag stochastic processes $\{Z_n\}$ converges in distribution to a continuous stochastic process Z (denoted $Z_n \Rightarrow Z$), if*

$$E[f(Z_n)] \to E[f(Z)]$$

for every bounded f that is continuous in the compact uniform topology.

Example 10.3 *Consider $g \in C_b(\mathbb{R})$, $h \in C(\mathbb{R}^d)$ and $x : [0,\infty) \to \mathbb{R}^d$, where $C_b(\mathbb{R})$ is the space of all bounded continuous function on $\mathbb{R}$. Define*

$$F(x) = g(\sup_{s \le 27} h(x(s))).$$

Then F is continous in the compact uniform topology, Note that if $x_n \to x$ in the compact uniform topology, then $h \circ x_n \to h \circ x$ in the compact uniform topology.

Example 10.4 *In the notation of the last example,*

$$G(x) = g(\int_0^{27} h(x(s))ds)$$

is also continuous in the compact uniform topology.

Theorem 10.5 *Let $\{M_n\}$ be a sequence of martingales. Suppose that*

$$\lim_{n\to\infty} E[\sup_{s \le t} |M_n(s) - M_n(s-)|] = 0 \tag{10.1}$$

and

$$[M_n]_t \to c(t) \tag{10.2}$$

for each $t > 0$, where $c(t)$ is continuous and deterministic. Then $M_n \Rightarrow M = W \circ c$.

Remark 10.6 *If*

$$\lim_{n\to\infty} E[|[M_n]_t - c(t)|] = 0, \qquad \forall t \ge 0, \tag{10.3}$$

then by the continuity of c, both (10.1) and (10.2) hold. If (10.2) holds and $\lim_{n\to\infty} E[[M_n]_t] = c(t)$ for each $t \ge 0$, then (10.3) holds by the dominated convergence theorem.

Proof. See Ethier and Kurtz (1986), Theorem 7.1.4. □

Example 10.7 *If $M_n \Rightarrow W \circ c$, then*

$$P\{\sup_{s\leq t} M_n(s) \leq x\} \to P\{\sup_{s\leq t} W(c(s)) \leq x\} = P\{\sup_{u\leq c(t)} W(u) \leq x\}.$$

Corollary 10.8 *(Donsker's invariance principle.) Let ξ_k be iid with mean zero and variance σ^2. Let*

$$M_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k.$$

Then M_n is a martingale for every n, and $M_n \Rightarrow \sigma W$.

Proof. Since M_n is a finite variation process, we have

$$\begin{aligned} [M_n]_t &= \sum_{s\leq t} (\Delta M_n(s))^2 \\ &= \frac{1}{n} \sum_{k=1}^{[nt]} \xi_k^2 \\ &= \frac{[nt]}{n[nt]} \sum_{k=1}^{[nt]} \xi_k^2 \to t\sigma^2. \end{aligned}$$

where the limit holds by the law of large numbers. Consequently, (10.2) is satisfied. Note that the convergence is in L_1, so by Remark 10.6, (10.1) holds as well. Theorem 10.5 gives $M_n \Rightarrow W(\sigma^2 \cdot)$. □

Corollary 10.9 *(CLT for renewal processes.) Let ξ_k be iid, positive and have mean m and variance σ^2. Let*

$$N(t) = \max\{k : \sum_{i=1}^{k} \xi_i \leq t\}.$$

Then

$$Z_n(t) \equiv \frac{N(nt) - nt/m}{\sqrt{n}} \Rightarrow W(\frac{t\sigma^2}{m^3}).$$

Proof. The renewal theorem states that

$$E[|\frac{N(t)}{t} - \frac{1}{m}|] \to 0$$

and

$$\frac{N(t)}{t} \to \frac{1}{m}, \qquad a.s.$$

Let $S_k = \sum_1^k \xi_i$, $M(k) = S_k - mk$ and $\mathcal{F}_k = \sigma\{\xi_1, \ldots, \xi_k\}$. Then M is an $\{\mathcal{F}_k\}$-martingale and $N(t)+1$ is an $\{\mathcal{F}_k\}$ stopping time. By the optional sampling theorem $M(N(t)+1)$ is a martingale with respect to the filtration $\{F_{N(t)+1}\}$.

Note that

$$\begin{aligned} M_n(t) &= -M(N(nt)+1)/(m\sqrt{n}) \\ &= \frac{N(nt)+1}{\sqrt{n}} - \frac{S_{N(nt)+1} - nt}{m\sqrt{n}} - \frac{nt}{m\sqrt{n}} \\ &= \frac{N(nt) - nt/m}{\sqrt{n}} + \frac{1}{\sqrt{n}} - \frac{1}{m\sqrt{n}}(S_{N(nt)+1} - nt) \ . \end{aligned}$$

So asymptotically Z_n behaves like M_n, which is a martingale for each n. Now, for M_n, we have

$$\sup_{s\leq t} |M_n(s) - M_n(s-)| = \max_{k\leq N(nt)+1} |\xi_k - m|/m\sqrt{n}$$

and

$$[M_n]_t = \frac{1}{m^2 n} \sum_1^{N(nt)+1} |\xi_k - m|^2 \to \frac{t\sigma^2}{m^3}.$$

Since

$$E[[M_n]_t] = \frac{1}{m^2 n} E[N(nt)+1] \to \frac{t\sigma^2}{m^3}$$

Remark 10.6 applies and $Z_n \Rightarrow W(\frac{t\sigma^2}{m^3})$. □

Corollary 10.10 *Let $N(t)$ be a Poisson process with parameter λ and*

$$X(t) = \int_0^t (-1)^{N(s)} ds.$$

Define

$$X_n(t) = \frac{X(nt)}{\sqrt{n}}.$$

Then $X_n \Rightarrow \lambda^{-1} W$.

Proof. Note that

$$\begin{aligned} (-1)^{N(t)} &= 1 - 2\int_0^t (-1)^{N(s-)} dN(s) \\ &= 1 - 2M(t) - 2\lambda \int_0^t (-1)^{N(s)} ds, \end{aligned}$$

where

$$M(t) = \int_0^t (-1)^{N(s-)} d(N(s) - s)$$

is a martingale. Thus

$$X_n(t) = \frac{X(nt)}{\sqrt{n}} = \frac{1 - (-1)^{N(nt)}}{2\lambda\sqrt{n}} - \frac{M(nt)}{\lambda\sqrt{n}}.$$

One may apply the martingale CLT by observing that $[M_n]_t = N(nt)/(n\lambda^2)$ and that the jumps of M_n are of magnitude $1/(\lambda\sqrt{n})$. □

The martingale central limit theorem has a vector analogue.

Theorem 10.11 *(Multidimensional Martingale CLT). Let $\{M_n\}$ be a sequence of $\mathbb{R}^d$-valued martingales. Suppose*

$$\lim_{n\to\infty} E[\sup_{s\le t} |M_n(s) - M_n(s-)|] = 0 \tag{10.4}$$

and

$$[M_n^i, M_n^j]_t \to c_{i,j}(t) \tag{10.5}$$

for all $t \ge 0$ where, $C = ((c_{i,j}))$ is deterministic and continuous. Then $M_n \Rightarrow M$, where M is Gaussian with independent increments and $E[M(t)M(t)^T] = C(t)$.

Remark 10.12 *Note that $C(t) - C(s)$ is nonnegative definite for $t \ge s \ge 0$. If C is differentiable, then the derivative will also be nonnegative definite and will, hence have a nonnegative definite square root. Suppose $C(t) = \sigma(t)^2$ where σ is symmetric. Then M can be written as*

$$M(t) = \int_0^t \sigma(s)dW(s)$$

where W is d-dimensional standard Brownian motion.

10.2 Sequences of stochastic differential equations.

Let $\{\xi_k\}$ be iid with mean zero and variance σ^2. Suppose X_0 is independent of $\{\xi_k\}$ and

$$X_{k+1} = X_k + \sigma(X_k)\frac{\xi_{k+1}}{\sqrt{n}} + \frac{b(X_k)}{n}.$$

Define $X_n(t) = X_{[nt]}$, $W_n(t) = 1/\sqrt{n}\sum_{k=1}^{[nt]} \xi_k$, and $V_n(t) = [nt]/n$. Then

$$X_n(t) = X_n(0) + \int_0^t \sigma(X_n(s-))dW_n(s) + \int_0^t b(X_n(s-))dV_n(s)$$

By Donsker's theorem, $(W_n, V_n) \Rightarrow (\sigma W, V)$, with $V(t) = t$.

More generally we have the following equation:

$$X_n(t) = X(0) + \varepsilon_n(t) + \int_0^t \sigma(X_n(s-))dW_n(s) + \int_n^t b(X_n(s-))dV_n(s)\ . \tag{10.6}$$

Theorem 10.13 *Suppose in 10.6 W_n is a martingale, and V_n is a finite variation process. Assume that for each $t \ge 0$, $\sup_n E[[W_n]_t] < \infty$ and $\sup_n E[T_t(V_n)] < \infty$ and that $(W_n, V_n, \varepsilon_n) \Rightarrow (W, V, 0)$, where W is standard Brownian motion and $V(t) = t$. Suppose that X satisfies*

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds \tag{10.7}$$

and that the solution of (10.7) is unique. Then $X_n \Rightarrow X$.

Proof. See Kurtz and Protter (1991). □

10.3 Approximation of empirical CDF.

Let ξ_i be i.i.d and uniform on $[0,1]$, let $N_n(t) = \sum_{k=1}^n I_{[\xi_k,1]}(t)$, where $0 \le t \le 1$. Define $\mathcal{F}_t^n = \sigma(N_n(u);\ u \le t)$. For $t \le s \le 1$, we have

$$\begin{aligned} E[N_n(s)|\mathcal{F}_t^n] &= E[N_n(t) + N_n(s) - N_n(t)|\mathcal{F}_t^n] \\ &= N_n(t) + E[N_n(s) - N_n(t)|\mathcal{F}_t^n] \\ &= N_n(t) + (n - N_n(t))(s-t)/(1-t)\ . \end{aligned}$$

It follows that

$$\tilde{M}_n(t) = N_n(t) - \int_0^t \frac{n - N_n(s)}{1-s} ds$$

is a martingale.

Define $F_n(t) \equiv \frac{N_n(t)}{n}$ and $B_n(t) = \sqrt{n}(F_n(t) - 1) = \frac{N_n(t) - nt}{\sqrt{n}}$. Then

$$\begin{aligned} B_n(t) &= \frac{1}{\sqrt{n}}(N_n(t) - nt) \\ &= \frac{1}{\sqrt{n}}(\tilde{M}_n(t) + nt - \sqrt{n}\int_0^t \frac{B_n(s)ds}{1-s} - nt) \\ &= M_n(t) - \int_0^t \frac{B_n(t)}{1-s} ds. \end{aligned}$$

where $M_n(t) = \frac{\tilde{M}_n(t)}{\sqrt{n}}$. Note that $[M_n]_t = F_n(t)$ and by the law of large numbers, $[M_n]_t \to t$. Since $F_n(t) \le 1$, the convergence is in L_1 and Theorem 10.5 implies $M_n \Rightarrow W$. Therefore, $B_n \Rightarrow B$ where

$$B(t) = W(t) - \int_0^t \frac{B(s)}{1-s} ds$$

at least if we restrict our attention to $[0, 1-\varepsilon]$ for some $\varepsilon > 0$. To see that convergence is on the full interval $[0,1]$, observe that

$$E\int_{1-\varepsilon}^1 \frac{|B_n(s)|}{1-s} ds = \int_{1-\varepsilon}^1 \frac{E[|B_n(s)|]}{1-s} ds \le \int_{1-\varepsilon}^1 \frac{\sqrt{E}[B_n^2(s)]}{1-s} ds \le \int_{1-\varepsilon}^1 \frac{\sqrt{s} - s^2}{1-s} ds$$

which is integrable. It follows that for any $\delta > 0$, $\sup_n P\{\sup_{1-\varepsilon \le s \le 1} |B_n(1) - B_n(s)| \ge \delta\} \to 0$. This uniform estimate ensures that $B_n \Rightarrow B$ on the full interval $[0,1]$. The process B is known as Brownian Bridge.

10.4 Diffusion approximations for Markov chains.

Let X be an integer-valued process and write

$$X(t) = X(0) + \sum_{l \in \mathbb{Z}} l N_l(t)$$

where the N_l are counting processes, that is, N_l counts the number of jumps of X of size l at or before time t. Assume that X is Markov with respect to $\{\mathcal{F}_t\}$, and suppose that $P(X(t+h)=j|X(t)=i)=q_{ij}h+o(h)$ for $i\neq j$. If we define $\beta_l(x)=q_{x,x+l}$, then

$$E[N_l(t+h)-N_l(t)|\mathcal{F}_t]=q_{X(t),X(t)+l}h+o(h)=\beta_l(X(t))+o(h)\ .$$

Our first claim is that $M_l(t)\equiv N_l(t)-\int_0^t\beta_l(X(s))ds$ is a martingale (or at least a local martingale). If we define $\tau_l(n)=\inf\{t:N_l(t)=n\}$, then for each n, $M_l(\cdot\wedge\tau_l(n))$ is a martingale.

Assume everything is nice, in particular, for each l, assume that $M_l(t)$ is an $\{\mathcal{F}_t\}$-martingale. Then

$$X(t)=X(0)+\sum_l lN_l(t)=X(0)+\sum_l lM_l(t)+\sum_l l\int_0^t\beta_l(X(s))ds\ .$$

If $\sum_l|l|\beta_l(x)<\infty$, then we can interchange the sum and integral. Let $b(x)\equiv\sum_l l\beta_l(x)$, so we have

$$X(t)=X(0)+\sum_l lM_l(t)+\int_0^t b(X(s))ds\ .$$

Note that $[M_l]_t=N_l(t)$ and $[M_l,M_k]_t=[N_l,N_k]_t=0$. Therefore $E[N_l(t)]=E[\int_0^t\beta_l(X(s))ds]$ and $E[[M_l]_t]=E[N_l(t)]=E[(M_l(t))^2]$ holds. Consequently,

$$E[(\sum_k^m lM_l(t))^2]=\sum_k^m l^2E[M_l(t)^2]=\sum[\int_0^t\sum_k^m l^2\beta_l(X(s))ds],$$

so if

$$E[\int_0^t\sum_l l^2\beta_l(X(s))ds]<\infty,$$

then $\sum lM_l(t)$ is a square integrable martingale. If we only have $\sum l^2\beta_l(x)\leq\infty$ for each x and $\sum_l N_l(t)<\infty$ a.s., then let $\tau_c=\inf\{t,\sum l^2\beta_l(t)\geq c\}$ and assume that $\tau_c\to\infty$ as $c\to\infty$. Then $\sum_l lM_l(t)$ is a local martingale.

Now consider the following sequence: $X_n(t)=X_n(0)+\frac{\sum lN_l^n(t)}{n}$ where, for example, we can take $X_n(t)=\frac{X(n^2t)}{n}$ and $N_l^n(t)=N_l(n^2t)$ with X and N defined as before. Assume

$$M_l^n(t)\equiv\frac{1}{n}(N_l^n(t)-\int_0^t n^2\beta_l^n(X_n(s))ds)$$

is a martingale, so we have $[M_l^n]_t=\frac{N_l^n(t)}{n^2}$ and $E[[M_l^n]_t]=E[\int_0^t\beta_l^n(X_n(s))ds]$.

For simplicity, we assume that $\sup_n\sup_x\beta_l^n(x)<\infty$ and that only finitely many of the β_l^n are nonzero. Define $b_n(x)\equiv n\sum_l l\beta_l^n(x)$. Then we have

$$X_n(t)=X_n(0)+\sum_l lM_l^n(t)+\int_0^t b_n(X_n(s))ds$$

Assume:

1) $X_n(0) \to X(0)$,

2) $\beta_l^n(x) \to \beta_l(x)$,

3) $b_n(x) \to b(x)$,

4) $n^2 \inf_x \beta_l^n(x) \to \infty$

where the convergence in 1-3 is uniform on bounded intervals. By our assumptions, $\frac{E[(M_l^n(t))^2]}{n^2} \to 0$, so by Doob's inequality. It follows that $\sup_{s\le t} |\frac{M_l^n(t)}{n}| \to 0$. Consequently, $[M_l^n]_t \sim \int_0^t \beta_l^n(X_n(s))ds$.

Define $W_l^n(t) = \int_0^t \frac{1}{\sqrt{\beta_l^n(X_n(s-)}} dM_l^n(s)$. Then

$$[W_l^n]_t = \int_0^t \frac{d[M_l^n]_s}{\beta_l^n(X_n(s-)} = \int_0^t \frac{1}{n\beta_l^n(X_n(s-))} dM_l^n(s) + t \equiv u_l^n(t) + t.$$

Note that $u_l^n(t)$ is a martingale, and

$$\begin{aligned}
E[u_l^n(t)^2] &= E[\int_0^t \frac{d[M_n^l]_s}{n^2\beta_l^n(X_n(s-)^2}] \\
&= E[\int_0^t \frac{\beta_l^n(X_n(s))ds}{n^2\beta_l^n(X_n(s)^2}] \\
&= E[\int_0^t \frac{ds}{n^2\beta_l^n(X_n(s^-))}] \to 0.
\end{aligned}$$

Consequently, under the above assumptions, $[W_l^n]_t \to t$ and hence $W_l^n \Rightarrow W_l$, where W_l is a standard Brownian motion.

By definition, $M_l^n(t) = \int_0^t \sqrt{\beta_l^n(X_n(s-))} dW_l^n(s)$, so

$$X_n(t) = X_n(0) + \sum_l l \int_0^t \sqrt{\beta_l^n(X_n(s-))} dW_l^n(s) + \int_0^t b_n(X_n(s))ds \ .$$

Let

$$\begin{aligned}
\varepsilon_n(t) &= X_n(0) - X(0) + \sum_l \int_0^t l(\sqrt{\beta_l^n(X_n(s-))} - \sqrt{\beta_l(X(s-))}) dW_l^n(s) \\
&+ \int_0^t (b_n(X_n(s)) - b(X(s))ds
\end{aligned}$$

which converges to zero at least until $X_n(t)$ exits a fixed bounded interval. Theorem 10.13 gives the following.

Theorem 10.14 *Assume 1-4 above. Suppose the solution of*

$$X(t) = X(0) + \sum_l l \int_0^t \sqrt{\beta_l(X(s))} dW_l(s) + \int_0^t b(X(s))ds$$

exists and is unique. Then $X_n \Rightarrow X$.

10.5 Convergence of stochastic integrals.

Theorem 10.15 *Let Y_n be a semimartingale with decomposition $Y_n = M_n + V_n$. Suppose for each $t \geq 0$ that $\sup_{s \leq t} |X_n(s) - X(s)| \to 0$ and $\sup_{s \leq t} |Y_n(s) - Y(s)| \to 0$ in probability as $n \to \infty$, and that $\sup_n E[M_n(t)^2] = \sup_n E[[M_n]_t] < \infty$ and $\sup_n E[T_t(V_n)] < \infty$. Then for each $T > 0$*

$$\sup_{t \leq T} \left| \int_0^t X_n(s) dY_n(s) - \int_0^t X(s) dY(s) \right| \to 0$$

in probability.

Proof. See Kurtz and Protter (1991). □

11 Reflecting diffusion processes.

11.1 The M/M/1 Queueing Model.

Arrivals form a Poisson process with parameter λ, and the service distribution is exponential with parameter μ. Consequently, the length of the queue at time t satisfies

$$Q(t) = Q(0) + Y_a(\lambda t) - Y_d(\mu \int_0^t I_{\{Q(s)>0\}} ds) \ ,$$

where Y_a and Y_d are independent unit Poisson processes. Define the busy period $B(t)$ to be

$$B(t) \equiv \int_0^t I_{\{Q(s)>0\}} ds$$

Rescale to get

$$X_n(t) \equiv \frac{Q(nt)}{\sqrt{n}}.$$

Then $X_n(t)$ satisfies

$$X_n(t) = X_n(0) + \frac{Y_a(\lambda_n nt)}{\sqrt{n}} - \frac{1}{\sqrt{n}} Y_d(n\mu_n \int_0^t I_{\{X_n(s)>0\}} ds).$$

For a unit Poisson process Y, define $\tilde{Y}(u) \equiv Y(u) - u$ and observe that

$$\begin{aligned} X_n(t) &= X_n(0) + \frac{1}{\sqrt{n}} \tilde{Y}_a(n\lambda_n t) - \frac{1}{\sqrt{n}} \tilde{Y}_d(n\mu_n \int_0^t I_{\{X_n(s)>0\}} ds) \\ &\quad + \sqrt{n}(\lambda_n - \mu_n)t + \sqrt{n}\mu_n \int_0^t I_{\{X_n(s)=0\}} ds \end{aligned}$$

We already know that if $\lambda_n \to \lambda$ and $\mu_n \to \mu$, then

$$\begin{aligned} W_a^n(t) &\equiv \frac{1}{\sqrt{n}} \tilde{Y}_a(n\lambda_n t) \Rightarrow \sqrt{\lambda} W_1(t) \\ W_d^n(t) &\equiv \frac{1}{\sqrt{n}} \tilde{Y}_d(n\mu_n t) \Rightarrow \sqrt{\mu} W_2(t)), \end{aligned}$$

where W_1 and W_2 are standard Brownian motions. Defining

$$\begin{aligned} c_n &\equiv \sqrt{n}(\lambda_n - \mu_n) \\ \Lambda_n(t) &\equiv \sqrt{n}\mu_n(t - B_n(t)), \end{aligned}$$

we can rewrite $X_n(t)$ as

$$X_n(t) = X_n(0) + W_a^n(t) - W_d^n(B_n(t)) + c_n t + \Lambda_n(t).$$

Noting that Λ_n is nondecreasing and increases only when X_n is zero, we see that (X_n, Λ_n) is the solution of the Skorohod problem corresponding to $X_n(0) + W_a^n(t) - W_d^n(B_n(t)) + c_n t$, that is, the following:

Lemma 11.1 *For $w \in D_{\mathbb{R}}[0,\infty)$ with $w(0) \geq 0$, there exists a unique pair (x,λ) satisfying*

$$x(t) = w(t) + \lambda(t) \tag{11.1}$$

such that $\lambda(0) = 0$, $x(t) \geq 0 \forall t$, and λ is nondecreasing and increases only when $x = 0$. The solution is given by setting $\lambda(t) = 0 \vee \sup_{s \leq t}(-w(s))$ and defining x by (11.1).

Proof. We leave it to the reader to check that $\lambda(t) = 0 \vee \sup_{s\leq t}(-w(s))$ gives a solution. To see that it gives the only solution, note that for $t < \tau_0 = \inf\{s : w(s) \leq 0\}$ the requirements on λ imply that $\lambda(t) = 0$ and hence $x(t) = w(t)$. For $t \geq \tau_0$, the nonnegativity of x implies

$$\lambda(t) \geq -w(t),$$

and $\lambda(t)$ nondecreasing implies

$$\lambda(t) \geq \sup_{s\leq t}(-w(s)).$$

If t is a point of increase of λ, then $x(t) = 0$, so we must have

$$\lambda(t) = -w(t) \leq \sup_{s\leq t}(-w(s)). \tag{11.2}$$

Since the right side of (11.2) is nondecreasing, we must have $\lambda(t) \leq \sup_{s\leq t}(-w(s))$ for all $t > \tau_0$, and the result follows. $\square$

Thus in the problem at hand, we see that Λ_n is determined by

$$\Lambda_n(t) = 0 \vee (-\inf_{s\leq t}(X_n(0) + W_a^n(s) - W_d^n(s) - W_d^n(B_n(s)) + c_n s))$$

Consequently, if

$$X_n(0) + W_a^n(t) - W_d^n(B_n(t)) + c_n t$$

converges, so does Λ_n and X_n along with it. Assuming that $c_n \to c$, the limit will satisfy

$$\begin{aligned} X(t) &= X(0) + \sqrt{\lambda}W_1(t) - \sqrt{\lambda}W_2(t) + ct + \Lambda(t) \\ \Lambda(t) &= 0 \vee \sup_{s\leq t}(-(X(0) + \sqrt{\lambda}(W_1(s) - W_2(s)) + ct)). \end{aligned}$$

Recalling that $\sqrt{\lambda}(W_1 - W_2)$ has the same distribution as $\sqrt{2\lambda}W$, where W is standard Brownian motion, the limiting equation can be simplified to

$$\begin{aligned} X(t) &= X(0) + \sqrt{2\lambda}W(t) + ct + \Lambda(t) \\ X(t) &\geq 0 \quad \forall t, \end{aligned}$$

where Λ is nondecreasing and Λ increases only when $X(t) = 0$.

11.2 The G/G/1 queueing model.

Let $\eta_1, \eta_2, ...$ be i.i.d. with $\eta_i > 0$ and

$$\lambda = \frac{1}{E[\eta_i]}.$$

The η_i represent interarrival times. The service times are denoted ξ_i, which are also i.i.d. and positive with

$$\mu = \frac{1}{E[\xi_i]}.$$

The arrival process is

$$A(t) = \max\{k : \Sigma_{i=1}^{k} \eta_i \leq t\}$$

and the service process is

$$S(t) = \max\{k : \Sigma_{i=1}^{k} \xi_i \leq t\}.$$

The queue-length then satisfies

$$Q(t) = Q(0) + A(t) - S(\int_0^t I_{\{Q(s)>0\}} ds).$$

Following the approach taken with the M/M/1 queue, we can express the G/G/1 queue as the solution of a Skorohod problem and use the functional central limit theorem for the renewal processes to obtain the diffusion limit.

11.3 Multidimensional Skorohod problem.

We now consider a multidimensional analogue of the problem presented in Lemma 11.1. Let D be convex and let $\eta(x)$ denote the unit inner normal at $x \in \partial D$. Suppose w satisfies $w(0) \in D$. Consider the equation for (x, λ)

$$\begin{aligned} x(t) &= w(t) + \int_0^t \eta(x(s)) d\lambda(s) \\ x(t) &\epsilon \overline{D} \qquad \forall t \geq 0, \end{aligned}$$

where λ is nondecreasing and increases only when $x(t) \epsilon \partial D$.

Proof of Uniqueness. Let

$$x_i(t) = w(t) + \int_0^t \eta(x_i(s)) d\lambda_i(s)$$

Assume continuity for now. Since λ_i is nondecreasing, it is of finite variation. Itô's formula yields

$$\begin{aligned} (x_1(t) - x_2(t))^2 &= \int_0^t 2(x_1(s) - x_2(s)) d(x_1(s) - x_2(s)) \\ &= \int_0^t 2(x_1(s) - x_2(s)) \eta(x_1(s)) d\lambda_1(s) \\ &\qquad - \int_0^t 2(x_1(s) - x_2(s)) \eta(x_2(s)) d\lambda_2(s) \\ &\leq 0, \end{aligned}$$

where the inequality follows from the fact that λ_i increases only when x_i is on the boundary and convexity implies that for any $x \in \partial D$ and $y \in D$, $\eta(x) \cdot (y - x) \geq 0$. Consequently, uniqueness follows.

If there are discontinuities, then

$$\begin{aligned}
|x_1(t) - x_2(t)|^2 &= \int_0^t 2(x_1(s-) - x_2(s-) \cdot \eta(x_1(s))d\lambda_1(s) \\
&\quad - \int_0^t 2(x_1(s-) - x_2(s-)) \cdot \eta(x_2(s))d\lambda_2(s) + [x_1 - x_2]_t \\
&= \int_0^t 2(x_1(s-) - x_2(s-)\eta(x_1(s))d\lambda_1(s) \\
&\quad - \int_0^t 2(x_1(s-) - x_2(s-))\eta(x_2(s))d\lambda_2(s) \\
&\quad +2\sum_{s \leq t}(\Delta x_1(s) - \Delta x_2(s))(\eta(x_1(s))\Delta\lambda_1(s) - \eta(x_2(s))\Delta\lambda_2(s)) \\
&\quad -[x_1 - x_2]_t \\
&= \int_0^t 2(x_1(s) - x_2(s))\eta(x_1(s))d\lambda_1(s) \\
&\quad - \int_0^t 2(x_1(s) - x_2(s))\eta(x_2(s))d\lambda_2(s) - [x_1 - x_2]_t \\
&\leq 0,
\end{aligned}$$

so the solution is unique.

Let W be standard Brownian motion, and let (X, λ) satisfy

$$X(t) = W(t) + \int_0^t \eta(X(s))d\lambda(s)\ , \qquad X(t)\epsilon\overline{D}, \qquad \forall t \geq 0$$

λ nondecreasing, and λ increasing only when $X\epsilon\partial D$. Itô's formula yields

$$\begin{aligned}
f(X(t)) &= f(X(0) + \int_0^t \nabla f(X(s))dW(s) + \int_0^t \frac{1}{2}\Delta f(X(s))ds \\
&\quad + \int_0^t \eta(X(s))\nabla f(X(s))d\lambda(s).
\end{aligned}$$

Assume $\eta(x) \cdot \nabla f = 0$ for $x\epsilon\partial D$. If we solve

$$u_t = \frac{1}{2}\Delta u$$

subject to the Neumann boundary conditions

$$\begin{aligned}
\eta(x) \cdot \nabla u(x,t) &= 0 \qquad \forall x\epsilon\partial D, \\
u(x,0) &= f(x),
\end{aligned}$$

we see that $u(r-t, X(t))$ is a martingale and hence

$$E[f(X(t,x))] = u(t,x).$$

Similarly, we can consider more general diffusions with normal reflection corresponding to the equation

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

subject to the conditions that $X(t)\epsilon\overline{D}\ \forall t \geq 0$, λ is nondecreasing and increases only when $X(t)\epsilon\partial D$. To examine uniqueness, apply Itô's formula to get

$$\begin{aligned} |X_1(s) - X_2(s)|^2 \quad = \quad & \int_0^t 2(X_1(s) - X_2(s))^T(\sigma(X_1(s)) - \sigma(X_2(s)))dW(s) \\ & + \int_0^t (X_1(s) - X_2(s)) \cdot (b(X_1(s)) - b(X_2(s)))ds \\ & + \int_0^t trace(\sigma(X_1(s)) - \sigma(X_2(s)))(\sigma(X_1(s)) - \sigma(X_2(s)))^T ds \\ & + \int_0^t 2(X_1(s) - X_2(s)) \cdot \eta(X_1(s))d\lambda_1(s) \\ & - \int_0^t 2(X_1(s) - X_2(s)) \cdot \eta(X_2(s))d\lambda_2(s). \end{aligned} \tag{11.3}$$

The last two terms are negative as before, and assuming that σ and b are Lipschitz, we can use Gronwall's inequality to obtain uniqueness just as in the case of unbounded domains.

Existence can be proved by an iteration beginning with

$$X_0(t) \equiv X(0),$$

and then letting

$$X_{n+1}(t) = X(0) + \int_0^t \sigma(X_n(s))dW(s) + \int_0^t b(X_n(s))ds + \int_0^t \eta(X_n(s))d\lambda_{n+1}(s).$$

An analysis similar to (11.3) enables one to show that the sequence $\{X_n\}$ is Cauchy.

11.4 The Tandem Queue.

Returning to queueing models, consider a simple example of a queueing network, the tandem queue:

$$\begin{aligned} Q_1(t) &= Q_1(0) + Y_a(\lambda t) - Y_{d1}(\mu_1 \int_0^t I_{\{Q_1(s)>0\}}ds) \\ Q_2(t) &= Q_2(0) + Y_{d1}(\mu_1 \int_0^t I_{\{Q_1(s)>0\}}ds) - Y_{d2}(\mu_2 \int_0^t I_{\{Q_2(s)>0\}}ds) \ . \end{aligned}$$

If we assume that

$$\begin{aligned}
\lambda^n, \mu_1^n, \mu_2^n &\quad \to \lambda \\
c_1^n \equiv \sqrt{n}(\lambda^n - \mu_1^n) &\quad \to c_1 \\
c_2^n \equiv \sqrt{n}(\mu_1^n - \mu_2^n) &\quad \to c_2
\end{aligned}$$

and renormalize the queue lengths to define

$$\begin{aligned}
X_1^n(t) = \frac{Q_1(nt)}{\sqrt{n}} &= X_1^n(0) + W_{a1}^n(t) + c_1^n t - W_{d1}^n\left(\int_0^t I_{\{X_1^n(s)>0\}} ds\right) \\
&+ \sqrt{n}\mu_1^n \int_0^t I_{\{X_1^n(s)=0\}} ds \\
X_2^n(t) &= X_n^n(0) + W_{d1}^n\left(\int_0^t I_{\{X_1^n(s)>0\}} ds\right) - W_{d2}^n\left(\int_0^t I_{\{X_2^n(s)>0\}} ds\right) \\
&+ c_2^n t - \sqrt{n}\mu_1^n \int_0^t I_{\{X_1^n(s)=0\}} ds + \sqrt{n}\mu_2^n \int_0^t I_{\{X_2^n(s)=0\}} ds,
\end{aligned}$$

we can obtain a diffusion limit for this model. We know already that X_1^n converges in distribution to the X_1 that is a solution of

$$X_1(t) = X_1(0) + \lambda W_1(t) - \lambda W_2(t) + c_1 t + \Lambda_1(t) \ .$$

For X_2^n, we use similar techniques to show X_2^n converges in distribution to X_2 satisfying

$$X_2(t) = X_2(0) + \lambda W_2(t) - \lambda W_3(t) + c_2 t - \Lambda_1(t) + \Lambda_2(t),$$

or in vector form

$$X(t) = X(0) + \begin{pmatrix} \lambda & -\lambda & 0 \\ 0 & \lambda & -\lambda \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} t + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Lambda_1(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Lambda_2(t)$$

where Λ_1 increases only when $X_1 = 0$ and Λ_2 increases only when $X_2 = 0$.

12 Change of Measure

Let $(\Omega, \mathcal{F}, Q)$ be a probability space and let L be a non-negative random variable, such that

$$E^Q[L] = \int L dQ = 1.$$

Define $P(\Gamma) \equiv \int_\Gamma L dQ$ where $\Gamma \in \mathcal{F}$. P is a probability measure on $\mathcal{F}$. This makes P absolutely continuous with respect to Q ($P << Q$) and L is denoted by

$$L = \frac{dP}{dQ}.$$

12.1 Applications of change-of-measure.

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 α that maximizes $H(\alpha, X_1, X_2, \ldots X_n)$.

For example, let

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

We will give conditions under which the distribution of X_α is absolutely continuous with respect to the distribution of X satisfying

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s)\ . \tag{12.1}$$

Sufficiency: If $dP_\alpha = L_\alpha dQ$ where

$$L_\alpha(X, Y) = H_\alpha(X)G(Y),$$

then X is a *sufficient statistic* for α.

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 (12.1) via a change of measure.

12.2 Bayes Formula.

Assume $dP = LdQ$ on $(\Omega, \mathcal{F})$. Note that $E^P[X] = E^Q[XL]$. We want to derive the corresponding formula for conditional expectations.

Recall that $Y = E[Z|\mathcal{D}]$ if

1) Y is $\mathcal{D}$-measurable.

2) For each $D \in \mathcal{D}$, $\int_D Y dP = \int_D Z dP$.

Lemma 12.1 *(Bayes Formula)*

$$E^P[Z|\mathcal{D}] = \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}].} \tag{12.2}$$

Proof. Clearly the right side of (12.2) is $\mathcal{D}$-measurable. Let $D \in \mathcal{D}$. Then

$$\begin{aligned}
\int_D \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}].} dP &= \int_D \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}]} LdQ \\
&= \int_D \frac{E^Q[ZL|\mathcal{D}]}{E^Q[L|\mathcal{D}]} E^Q[L|\mathcal{D}]dQ \\
&= \int_D E^Q[ZL|\mathcal{D}]dQ \\
&= \int_D ZLdQ = \int_D ZdP
\end{aligned}$$

which verifies the identity. □

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

$$E[g(Y)|X = x] = \frac{\int_{-\infty}^{\infty} y f_{XY}(x,y)dy}{f_X(x)}$$

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

$$\begin{aligned}
\nu_Y(\Gamma) &= Q\{Y \in \Gamma\} \\
dP &= H(X,Y)dQ.
\end{aligned}$$

Bayes formula becomes

$$E^P[g(Y)|X] = \frac{\int g(y)H(X,y)\nu_Y(dy)}{\int H(X,y)\nu_Y(dy)}$$

The left side is equal to

$$\frac{E^Q[g(Y)H(X,Y)|X]}{E^Q[H(X,Y)|X]},$$

and the independence of X and Y gives the identity by Property 10 of the Section 2.6.

12.3 Local absolute continuity.

Let $(\Omega, \mathcal{F})$ be a measurable space, and let P and Q be probability measures on $\mathcal{F}$. Suppose $\mathcal{D}_n \subset \mathcal{D}_{n+1}$ and that for each n, $P|_{\mathcal{D}_n} << Q|_{\mathcal{D}_n}$. Define $L_n = \left.\frac{dP}{dQ}\right|_{\mathcal{D}_n}$. Then $\{L_n\}$ is a nonnegative $\{\mathcal{D}_n\}$-martingale on $(\Omega, \mathcal{F}, Q)$ and $L = \lim_{n\to\infty} L_n$ satisfies $E^Q[L] \leq 1$. If $E^Q[L] = 1$, then $P << Q$ on $\mathcal{D} = \bigvee_n \mathcal{D}_n$. The next proposition gives conditions for this absolute continuity in terms of P.

Proposition 12.2 *$P << Q$ on $\mathcal{D}$ if and only if $P\{\overline{\lim}_{n\to\infty} L_n < \infty\} = 1$.*

Proof. We have

$$P\{\sup_{n\leq N} L_n \leq K\} = \int I_{\{\sup_{n\leq N} L_n \leq K\}} L_N dQ.$$

The dominated convergence theorem implies

$$P\{\sup_n L_n \leq K\} = \int_{\{\sup_n L_n \leq K\}} L dQ.$$

Letting $K \to \infty$ we see that $E^Q[L] = 1$. □

12.4 Martingales and change of measure.

(See Protter (1991), Section III.6.) Let $\{\mathcal{F}_t\}$ be a filtration and assume that $P|_{\mathcal{F}_t} << Q|_{\mathcal{F}_t}$ 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 12.3 *Z is a P-local martingale if and only if LZ is a Q-local martingale.*

Proof. Note that for a bounded stopping time τ, $Z(\tau)$ is P-integrable if and only if $L(\tau)Z(\tau)$ is Q-integrable. 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

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

□

Theorem 12.4 *If M is a Q-local martingale, then*

$$Z(t) = M(t) - \int_0^t \frac{1}{L(s)} d[L, M]_s \tag{12.3}$$

is a P-local martingale. (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 (12.3), we have

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

and both terms on the right are Q-local martingales. □

12.5 Change of measure for Brownian motion.

Let W be standard Brownian motion, and let ξ be an adapted process. Define

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

and note that

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

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

For independent standard Brownian motions $W_1, \ldots, W_m$, and adapted processes $\{\xi_i\}$,

$$L(t) = \exp\{\sum_i \int_0^t \xi_i(s)dW(s) - \frac{1}{2}\sum_i \int_0^t \xi_i^2(s)ds\}$$

is the solution of

$$L(t) = 1 + \sum_i \int_0^t \xi_i(s)L(s)dW_i(s).$$

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

$$\begin{aligned} P(A) = E^Q[I_A L(T)] &= E^Q[I_A E^Q[L(T)|\mathcal{F}_t]] \\ &= \underbrace{E^Q[I_A L(t)]}_{\substack{\text{has no dependence on T} \\ \text{(crucial that L is a martingale)}}} \end{aligned}$$

Claim: $\tilde{W}_i(t) = W_i(t) - \int_0^t \xi_i(s)ds$ is a standard Brownian motion on $(\Omega, \mathcal{F}_T, P)$. Since $\tilde{W}_i$ is continous and $[\tilde{W}_i]_t = t$ a.s., it is enough to show that $\tilde{W}_i$ is a local martingale (and hence a martingale). But since W_i is a Q-martingale and $[L, W_i]_t = \int_0^t \xi_i(s)L(s)ds$, Theorem 12.4 gives the desired result. Since $[\tilde{W}_i, \tilde{W}_j]_t = [W_i, W_j]_t = 0$ for $i \neq j$, the $\tilde{W}_i$ are independe.

Now 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

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

and assume that $E^Q[L(T)] = 1$ (e.g., if b is bounded). Set $dP = L(T)dQ$ on $(\Omega, \mathcal{F}_T)$. Define $\tilde{W}(t) = W(t) - \int_0^t b(X(s))ds$. Then

$$\begin{aligned} X(t) &= X(0) + \int_0^t \sigma(X(s))dW(s) \qquad (12.4) \\ &= X(0) + \int_0^t \sigma(X(s))d\tilde{W}(s) + \int_0^t \sigma(X(s))b(X(s))ds \end{aligned}$$

so under P, X is a diffusion with generator

$$\frac{1}{2}\sigma^2(x)f''(x) + \sigma(x)b(x)f'(x). \qquad (12.5)$$

We can eliminate the appriori assumption that $E^Q[L(T)] = 1$ by defining $\tau_n = \inf\{t : \int_0^t b^2(X(s))ds > n\}$ and defining $dP = L(T \wedge \tau_n)dQ$ on $\mathcal{F}_{T\wedge\tau_n}$. Then on $(\Omega, \mathcal{F}_{T\wedge\tau_n}, P)$, X is a diffusion with generator (12.5) stopped at time $T \wedge \tau_n$. But if there is a unique (in distribution) such diffusion and $\int_0^t b^2(X(s))ds$ is almost surely finite for this diffusion, then we can apply Proposition 12.2 to conclude that $P << Q$ on $\mathcal{F}_T$, that is, $E[L(T)] = 1$.

12.6 Change of measure for Poisson processes.

Let N be an $\{\mathcal{F}_t\}$-unit Poisson process on $(\Omega, \mathcal{F}, Q)$, that is, N is a unit Poisson process adapted to $\{\mathcal{F}_t\}$, and for each t, $N(t + \cdot) - N(t)$ is independent of $\mathcal{F}_t$. If Z is nonnegative and $\{\mathcal{F}_t\}$-adapted, then

$$L(t) = \exp\left\{\int_0^t \ln Z(s-)dN(s) - \int_0^t (Z(s) - 1)ds\right\}$$

satisfies

$$L(t) = 1 + \int_0^t (Z(s-) - 1)L(s-)d(N(s) - s)$$

and is a Q-local martingale. If $E[L(T)] = 1$ and we define $dP = L(T)dQ$ on $\mathcal{F}_T$, the $N(t) - \int_0^t Z(s)ds$ is a P-local martingale.

If $N_1, \ldots, N_m$ are independent unit Poisson processes and the Z_i are nonneagative and $\{\mathcal{F}_t\}$-adapted

$$L(t) = \prod_{i=1}^m \exp\left\{\int_0^t \ln Z_i(s-)dN_i(s) - \int_0^t (Z_i(s) - 1)ds\right\}$$

satisfies

$$L(t) = 1 + \int_0^t (Z_i(s-) - 1)L(s-)d(N_i(s) - s).$$

Let $J[0, \infty)$ denote the collection of nonnegative integer-valued cadlag functions that are constant except for jumps of $+1$. Suppose that $\lambda_i : J[0, \infty)^m \times [0, \infty) \to [0, \infty)$, $i = 1, \ldots, m$ and that $\lambda_i(x, s) = \lambda_i(x(\cdot \wedge s), s)$ (that is, λ_i is nonanticipating). For $N = (N_1, \ldots, N_m)$, if we take $Z_i(t) = \lambda_i(N,t)$ and let $\tau_n = \inf\{t : \sum_i N_i(t) = n\}$, then defining $dP = L(\tau_n)dQ$ on $\mathcal{F}_{\tau_n}$, N on $(\Omega, \mathcal{F}_{\tau_n}, P)$ has the same distribution as the solution of

$$\tilde{N}_i(t) = Y_i(\int_0^{t\wedge\tilde{\tau}_n} \lambda_i(\tilde{N}, s)ds)$$

where the Y_i are independent unit Poisson process and $\tilde{\tau}_n = \inf\{t : \sum_i \tilde{N}_i(t) = n\}$.

13 Finance.

Consider financial activity over a time interval $[0, T]$ modeled by a probability space $(\Omega, \mathcal{F}, P)$. Assume that there is a "fair casino" or market such that at time 0, for each event $A \in \mathcal{F}$, a price $Q(A) \geq 0$ is fixed for a bet or a contract that pays one dollar at time T if and only if A occurs. Assume that the market is such that an investor can either buy or sell the policy and that $Q(\Omega) < \infty$. An investor can construct a *portfolio* by buying or selling a variety (possibly countably many) of contracts in arbitrary multiples. If a_i is the "quantity" of a contract for A_i ($a_i < 0$ corresponds to selling the contract), then the payoff at time T is

$$\sum_i a_i I_{A_i}.$$

We will require that $\sum_i |a_i| Q(A_i) < \infty$ so that the initial cost of the portfolio is (unambiguously)

$$\sum_i a_i Q(A_i).$$

The market has *no arbitrage* if no combination (buying and selling) of countably policies with a net cost of zero results in a positive profit at no risk. That is, if $\sum |a_i| Q(A_i) < \infty$,

$$\sum_i a_i Q(A_i) = 0,$$

and

$$\sum_i a_i I_{A_i} \geq 0 \quad a.s.,$$

then

$$\sum_i a_i I_{A_i} = 0 \quad a.s.$$

The no arbitrage requirement has the following implications.

Lemma 13.1 *Assume that there is no arbitrage. If $P(A) = 0$, then $Q(A) = 0$. If $Q(A) = 0$, then $P(A) = 0$.*

Proof. Suppose $P(A) = 0$ and $Q(A) > 0$. Then construct a portfolio by buying one unit of Ω and selling $Q(\Omega)/Q(A)$ units of A. Then the net cost is

$$Q(\Omega) - \frac{Q(\Omega)}{Q(A)} Q(A) = 0$$

and the payoff is

$$1 - \frac{Q(\Omega)}{Q(A)} I_A = 1 \quad a.s.$$

which contradicts the no arbitrage assumption.

Now suppose $Q(A) = 0$. Construct a portfolio by buying one unit of A. The cost of the portfolio is $Q(A) = 0$ and the payoff is $I_A \geq 0$. So by the no arbitrage assumption, $I_A = 0$ a.s., that is, $P(A) = 0$. $\square$

Lemma 13.2 *If there is no arbitrage and $A \subset B$, then $Q(A) \leq Q(B)$.*

Proof. Suppose $Q(B) < Q(A)$. Construct a portfolio by buying one unit of B and selling $Q(B)/Q(A)$ units of A. Then the net cost of the portfolio is

$$Q(B) - \frac{Q(B)}{Q(A)} Q(A) = 0$$

and the payoff is

$$I_B - \frac{Q(B)}{Q(A)} I_A = I_{B-A} + (1 - \frac{Q(B)}{Q(A)}) I_A \geq 0,$$

which is strictly positive on B. But $Q(A) > 0$ implies $P(A) > 0$, so there is a postive payoff with positive probability contradicting the no arbitrage assumption. □

Theorem 13.3 *If there is no arbitrage, Q must be a measure on $\mathcal{F}$ that is equivalent to P.*

Proof. Let $A_1, A_2, \ldots$ be disjoint and for $A = \cup_{i=1}^{\infty} A_i$, suppose that $Q(A) < \rho \equiv \sum_i Q(A_i)$. Then buy one unit of A and sell $Q(A)/\rho$ units of A_i for each i. The net cost is zero and the net payoff is

$$I_A - \frac{Q(A)}{\rho} \sum_i I_{A_i} = (1 - \frac{Q(A)}{\rho}) I_A.$$

Note that $Q(A_i) > 0$ implies $P(A_i) > 0$ and hence $P(A) > 0$, so the right side is ≥ 0 a.s. and is > 0 with positive probability, contradicting the no arbtrage assumption. It follows that $Q(A) \geq \rho$.

If $Q(A) > \rho$, then sell one unit of A and buy $Q(A)/\rho$ units of A_i for each i. □

Theorem 13.4 *If there is no arbitrage, $Q << P$ and $P << Q$. (P and Q are equivalent.)*

Proof. The result follows form Lemma 13.1. □

If X and Y are random variables satisfying $X \leq Y$ a.s., then no arbitrage should mean

$$Q(X) \leq Q(Y).$$

It follows that for any Q-integrable X, $Q(X) = \int X dQ$.

13.1 Assets that can be traded at intermediate times.

Let $\{\mathcal{F}_t\}$ represent the information available at time t. Let $B(t)$ be the price of a bond at time t that is worth \$1 at time T (e.g. $B(t) = e^{-r(T-t)}$), that is, at any time $0 \leq t \leq T$, $B(t)$ is the price of a contract that pays exactly \$1 at time T. Note that $B(0) = Q(\Omega)$, and define $\hat{Q}(A) = Q(A)/B(0)$.

Let $X(t)$ be the price at time t of another tradeable asset. For any stopping time $\tau \leq T$, we can buy one unit of the asset at time 0, sell the asset at time τ and use the money received ($X(\tau)$) to buy $X(\tau)/B(\tau)$ units of the bond. Since the payoff for this strategy is $X(\tau)/B(\tau)$, we must have

$$X(0) = \int \frac{X(\tau)}{B(\tau)} dQ = \int \frac{B(0)X(\tau)}{B(\tau)} d\hat{Q}.$$

Lemma 13.5 *If $E[Z(\tau)] = E[Z(0)]$ for all bounded stopping times τ, then Z is a martingale.*

Corollary 13.6 *If X is the price of a tradeable asset, then X/B is a martingale on $(\Omega, \mathcal{F}, \hat{Q})$.*

Consider $B(t) \equiv 1$. Let W be a standard Brownian motion on $(\Omega, \mathcal{F}, P)$ and $\mathcal{F}_t = \mathcal{F}_t^W$, $0 \le t \le T$. Suppose X is the price of a tradeable asset given as the unique solution of

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

For simplicity, assume $\sigma(X(t)) > 0$ for all $t \ge 0$. Then

$$\mathcal{F}_t^X = \mathcal{F}_t^W,$$

since, setting

$$M(t) = X(t) - X(0) - \int_0^t b(X(s))ds,$$

we have

$$W(t) = \int_0^t \frac{1}{\sigma(X(s))} dM(s).$$

Suppose $\hat{Q}$ ($= Q$ since $B(0) = 1$) is a pricing measure and

$$L = L(T) = \frac{d\hat{Q} \mid_{\mathcal{F}_T}}{dP \mid_{\mathcal{F}_T}}.$$

Then $L(t) = E[L(T)|\mathcal{F}_t]$, $0 \le t \le T$ is a martingale and

$$\tilde{W}(t) = W(t) - \int_{\mathbf{0}}^t \frac{1}{L(s)} d[L, W]_s$$

is a Brownian motion on $(\Omega, \mathcal{F}, \hat{Q})$.

Theorem 13.7 *(Martingale Representation Theorem.) Suppose M is a martingale on $(\Omega, \mathcal{F}, P)$ with respect to the filtration generated by a standard Brownian motion W. Then there exists an adapted, measurable process U such that $\int_0^t U^2(s)ds < \infty$ a.s. for each $t > 0$ and*

$$M(t) = M(0) + \int_0^t U(s)dW(s).$$

Note that the definition of the stochastic integral must be extended for the above theorem to be valid. Suppose U is progressive and statisfies

$$\int_0^t |U(s)|^2 ds < \infty \quad a.s.$$

for every $t > 0$. Defining $U(s) = U(0)$, for $s < 0$, set

$$U_n(t) = n \int_{t-\frac{1}{n}}^t U(s)ds.$$

Note that U_n is continuous and adapted and that

$$\int_0^t |U(s)-U_n(s)|^2 ds \to 0.$$

It follows that the sequence

$$\int_0^t U_n(s)dW(s)$$

is Cauchy in probability,

$$P\{\sup_{s\leq t}|\int_0^t U_n(s)dW(s)-\int_0^t U_m(s)dW(s)|>\epsilon\}\leq P\{\sigma\leq t\}+\frac{4E[\int_0^{t\wedge\sigma}|U_n(s)-U_m(s)|^2ds]}{\epsilon^2},$$

and

$$\int_0^t U(s)dW(s)\equiv\lim_{n\to\infty}\int_0^t U_n(s)dW(s).$$

Let

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

Then $[L,W]_t=\int_0^t U(s)ds$ and

$$X(t)=X(0)+\int_0^t\sigma(X(s))d\tilde W(s)+\int_0^t\left(\frac{\sigma(X(s))}{L(s)}U(s)+b(X(s))\right)ds.$$

Lemma 13.8 *If M is a continuous local martingale of finite variation, then M is constant in time.*

Proof. We have

$$(M(t)-M(0))^2=\int_0^t 2(M(s)-M(0))^2dM(s)+[M]_t$$

Since $[M]_t=0$, $(M(t)-M(0))^2$ must be a local martingale and hence must be identically zero. □

Since X must be a martingale on $(\Omega,\mathcal{F},\hat Q)$, the lemma implies

$$\frac{\sigma(X(s))}{L(s)}U(s)+b(X(s))=0.$$

It follows that

$$L(t)=1-\int_0^t\frac{b(X(s))}{\sigma(X(s))}L(s)dW(s),$$

so

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

$$\tilde W(t)=W(t)+\int_0^t\frac{b(X(s))}{\sigma(X(s))}ds,$$

and

$$X(t) = X(0) + \int_0^t \sigma(X(s))d\tilde{W}(s).$$

Note that $E[L(t)] = 1$ if

$$\hat{Q}\{\left|\int_0^t \frac{b(X(s))}{\sigma(X(s))}\right| < \infty\} = 1.$$

For example, if

$$X(t) = x_0 + \int_0^t \sigma X(s)dW(s) + \int_0^t bX(s)ds,$$

that is,

$$X(t) = x_0 \exp\{\sigma W(t) - \frac{1}{2}\sigma^2 t + bt\},$$

then

$$L(t) = \exp\{\frac{b}{\sigma}W(t) - \frac{1}{2}\frac{b^2}{\sigma^2}t\}.$$

Under $d\hat{Q} = L(T)dP$, $\tilde{W}(t) = W(t) + \frac{b}{\sigma}t$ is a a standard Brownian motion and

$$E^{\hat{Q}}[f(X(T))] = \int_{-\infty}^{\infty} f(x_0 \exp\{\sigma y - \frac{1}{2}\sigma^2 T) \frac{1}{\sqrt{2\pi T}} e^{-\frac{y^2}{2T}} dy.$$

How reasonable is the assumption that there exists a pricing measure Q? Start with a model for a collection of tradeable assets. For example, let

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

or more generally just assume that X is a vector semimartingale. Allow certain trading strategies producing a payoff at time T:

$$Y(T) = Y(0) + \sum_i \int_0^t H_i(s-)dX_i(s)$$

Arbitrage exists if there is a trading strategy satisfying

$$Y(T) = \sum_i \int_0^t H_i(s-)dX_i(s) \geq 0 \quad a.s.$$

with $P\{Y(T) > 0\} > 0$.

13.2 First fundamental "theorem".

Theorem 13.9 *(Meta theorem) There is no arbitrage if and only if there exists a probability measure Q equivalent to P under which the X_i are martingales.*

Problems:

- What trading strategies are allowable?
- The definition of *no arbitrage* above is, in general, too weak to give theorem.

For example, assume that $B(t) \equiv 1$ and that there is a single asset satisfying

$$X(t) = X(0) + \int_0^t \sigma X(s)dW(s) + \int_0^t bX(s)ds = X(0) + \int_0^t \sigma X(s)d\tilde{W}(s).$$

Let $T = 1$ and for some stopping time $\tau < T$, let

$$H(t) = \frac{1}{\sigma X(t)(1-t)}, \quad 0 \le t < \tau,$$

and $H(t) = 0$ for $t \ge \tau$. Then for $t < \tau$,

$$\int_0^t H(s)dX(s) = \int_0^t \frac{1}{1-s}d\tilde{W}(s) = \hat{W}(\int_0^t \frac{1}{(1-s)^2}ds),$$

where $\hat{W}$ is a standard Brownian motion under $\hat{Q}$. Let

$$\hat{\tau} = \inf\{u : \hat{W}(u) = 1\}, \qquad \int_0^\tau \frac{1}{(1-s)^2}ds = \hat{\tau}.$$

Then with probability 1,

$$\int_0^1 H(s)dX(s) = 1.$$

Admissible trading strategies: The trading strategy denoted $\{x, H_1, \ldots, H_d\}$ is *admissible* if for

$$V(t) = x + \sum_i \int_0^t H_i(s-)dX_i(s)$$

there exists a constant a such that

$$\inf_{0\le t\le T} V(t) \ge -a, \quad a.s.$$

No arbitrage: If $\{0, H_1, \ldots, H_d\}$ is an admissible trading strategy and $\sum_i \int_0^T H_i(s-)dX_i(s) \ge 0$ a.s., then $\sum_i \int_0^T H_i(s-)dX_i(s) = 0$ a.s.

No free lunch with vanishing risk: If $\{0, H_1^n, \ldots, H_d^n\}$ are admissible trading strategies and

$$\lim_{n\to\infty} \|0 \wedge \sum_i \int_0^T H_i^n(s-)dX_i(s)\|_\infty = 0,$$

then

$$|\sum_i \int_0^T H_i^n(s-)dX_i(s)| \to 0$$

in probability.

Theorem 13.10 *(Delbaen and Schachermayer). Let $X = (X_1, \ldots, X_d)$ be a bounded semimartingale defined on $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_t = \sigma(X(s), s \le t)$. Then there exists an equivalent martingale measure defined on $\mathcal{F}_T$ if and only if there is no free lunch with vanishing risk.*

13.3 Second fundamental "theorem".

Theorem 13.11 *(Meta theorem) If there is no arbitrage, then the market is complete if and only if the equivalent martingale measure is unique.*

Problems:

- What prices are "determined" by the allowable trading strategies?
- Specifically, how can one "close up" the collection of attainable payoffs?

Theorem 13.12 *If there exists an equivalent martingale random measure, then it is unique if and only if the set of replicable, bounded payoffs is "complete" in the sense that*

$$\{x + \sum_i \int_0^T H_i(s-)dX_i(s) : H_i \text{ simple}\} \cap L_\infty(P)$$

is weak dense in* $L_\infty(P, \mathcal{F}_T)$*,*

For general B, if we assume that after time 0 all wealth V must either be invested in the assets $\{X_i\}$ or the bond B, then the number of units of the bond held is

$$\frac{V(t) - \sum_i H_i(t)X_i(t)}{B(t)},$$

and

$$V(t) = V(0) + \sum_i \int_0^t H_i(s-)dX_i(s) + \int_0^t \frac{V(s-) - \sum_i H_i(s-)X_i(s-)}{B(s-)} dB(s).$$

Applying Itô's formula, we have

$$\frac{V(t)}{B(t)} = \frac{V(0)}{B(0)} + \sum_i \int_0^t \frac{H_i(s-)}{B(s-)} dX_i(s),$$

which should be a martingale under $\hat{Q}$.

14 Filtering.

Signal:

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

Observation:

$$Y(t) = \int_0^t h(X(s))ds + \alpha V(t)$$

Change of measure

$$\begin{aligned}
\frac{dQ|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}} = L_0(t) &= 1 - \int_0^t \alpha^{-1} h(X(s)) L_0(t) dV(s) \\
&= 1 - \int_0^t \alpha^{-2} h(X(s)) L_0(s) dY(s) + \int_0^t \alpha^{-1} h^2(X(s)) L_0(s) ds \\
&= \exp\{-\int_0^t \alpha^{-1} h(X(s)) dV(s) - \frac{1}{2}\int_0^t \alpha^{-2} h^2(X(s)) ds\} \\
&= \exp\{-\int_0^t \alpha^{-2} h(X(s)) dY(s) + \frac{1}{2}\int_0^t \alpha^{-2} h^2(X(s)) ds\}.
\end{aligned}$$

Define

$$\tilde{V}(t) = V(t) + \int_0^t \frac{h(X(s))}{\alpha} ds.$$

W and $\tilde{V}$ are independent Brownian motions under Q. Therefore X and $Y = \alpha\tilde{V}$ are independent under Q.

Therefore

$$\frac{dP|_{\mathcal{F}_t}}{dQ|_{\mathcal{F}_t}} = L(t) = L_0(t)^{-1} = \exp\{\int_0^t \alpha^{-2} h(X(s)) dY(s) - \frac{1}{2}\int_0^t \alpha^{-2} h^2(X(s)) ds\}$$

and

$$L(t) = 1 + \int_0^t \alpha^{-2} h(X(s)) L(s) dY(s).$$

Set $L(t, X, Y) = L(t)$.

$$E^P[f(X(t))|\mathcal{F}_t^Y] = \frac{E^Q[f(X(t))L(t,X,Y)|\mathcal{F}_t^Y]}{E^Q[L(t,X,Y)|\mathcal{F}_t^Y]} = \frac{\int f(x(t))L(t,x,Y)\mu_X(dx)}{\int L(t,x,Y)\mu_X(dx)}$$

Let ϕ be the measure-valued process determined by

$$\langle \phi(t), f\rangle = E^Q[f(X(t))L(t)|\mathcal{F}_t^Y].$$

We want to derive a differential equation for ϕ.

$$
\begin{aligned}
f(X(t))L(t) &= f(X(0)) + \int_0^t f(X(s))dL(s) \\
&\quad + \int_0^t L(s)\sigma(X(s))f'(X(s))dW(s) + \int_0^t L(s)Lf(X(s))ds \\
&= f(X(0)) + \int_0^t f(X(s))L(s)\alpha^{-2}h(X(s))dY(s) \\
&\quad + \int_0^t L(s)\sigma(X(s))f'(X(s))dW(s) + \int_0^t L(s)Lf(X(s))ds.
\end{aligned}
$$

Lemma 14.1 *Suppose X has finite expectation and is $\mathcal{H}$-measurable and that $\mathcal{D}$ is independent of $\mathcal{G} \vee \mathcal{H}$. Then*

$$E[X|\mathcal{G} \vee \mathcal{D}] = E[X|\mathcal{G}].$$

Proof. It is sufficient to show that for $G \in \mathcal{G}$ and $D \in \mathcal{D}$,

$$\int_{D \cap G} E[X|\mathcal{G}]dP = \int_{D \cap G} XdP.$$

But the independence assumption implies

$$
\begin{aligned}
\int_{D \cap G} E[X|\mathcal{G}]dP &= E[I_D I_G E[X|\mathcal{G}]] \\
&= E[I_D]E[I_G E[X|\mathcal{G}]] \\
&= E[I_D]E[I_G X] \\
&= E[I_D I_G X] \\
&= \int_{D \cap G} XdP.
\end{aligned}
$$

$\square$

Lemma 14.2 *Suppose that Y has independent increments and is compatible with $\{\mathcal{F}_t\}$. Then for $\{\mathcal{F}_t\}$-progressive U satisfying $\int_0^t E[|U(s)|]ds < \infty$*

$$E^Q[\int_0^t U(s)ds|\mathcal{F}_t^Y] = \int_0^t E^Q[U(s)|\mathcal{F}_s^Y]ds.$$

Proof. By the Fubini theorem for conditional expectations

$$E^Q[\int_0^t U(s)ds|\mathcal{F}_t^Y] = \int_0^t E^Q[U(s)|\mathcal{F}_t^Y]ds.$$

The identity then follows by Lemma 14.1 and the fact that $Y(r) - Y(s)$ is independent of $U(s)$ for $r > s$. $\square$

Lemma 14.3 *Suppose that Y is an $\mathbb{R}^d$-valued process with independent increments that is compatible with $\{\mathcal{F}_t\}$ and that there exists $p \geq 1$ and c such that*

$$E[|\int_0^t U(s)dY(s)|] < cE[\int_0^t |U(s)|^p ds]^{p^{-1}}, \tag{14.1}$$

for each $\mathbb{M}^{m\times d}$-valued, $\{\mathcal{F}_t\}$-predictable process U such that the right side of (14.1) is finite. Then for each such U,

$$E[\int_0^t U(s)dY(s)|\mathcal{F}_t^Y] = \int_0^t E[U(s)|\mathcal{F}_s^Y]dY(s). \tag{14.2}$$

Proof. If $U = \sum_{i=1}^m \xi_i I_{(t_{i-1},t_i]}$ with $0 = t_0 < \cdots < t_m$ and ξ_i, $\mathcal{F}_{t_i}$-measurable, (14.2) is immediate. The lemma then follows by approximation. □

Lemma 14.4 *Let Y be as above, and let $\{\mathcal{F}_t\}$ be a filtration such that Y is compatible with $\{\mathcal{F}_t\}$. Suppose that M is an $\{\mathcal{F}_t\}$-martingale that is independent of Y. If U is $\{\mathcal{F}_t\}$-predictable and*

$$E[|\int_0^t U(s)dM(s)|] < \infty,$$

then

$$E[\int_0^t U(s)dM(s)|\mathcal{F}_t^Y] = 0.$$

Applying the lemmas, we have the Zakai equation:

$$\begin{aligned}\langle\phi(t), f\rangle = \langle\phi(0), f\rangle &+ \int_0^t \langle\phi(s), Lf\rangle ds \\ &+ \int_0^t \langle\phi(s), \alpha^{-2} fh\rangle dY(s)\end{aligned}$$

15 Problems.

1. Let N be a nonnegative integer-valued random variable with $E[N] < \infty$. Let $\xi_1, \xi_2, \ldots$ be iid with mean m and independent of N. Show, using the definition of conditional expectation, that
$$E[\sum_{k=1}^{N} \xi_k | N] = mN.$$

2. Let $\{\mathcal{F}_k\}$ be a discrete-time filtration, let $\{\xi_k\}$ be iid such that ξ_k is $\mathcal{F}_k$-measurable and $(\xi_{k+1}, \xi_{k+2}, \ldots)$ is independent of $\mathcal{F}_k$, and let τ be an $\{\mathcal{F}_k\}$-stopping time.
   (a) Show that $(\xi_{\tau+1}, \xi_{\tau+2}, \ldots)$ is independent of $\mathcal{F}_\tau$.
   (b) Show that if the assumption that the $\{\xi_k\}$ are identically distributed is dropped, then the assertion in part a is no longer valid.

3. Let $\xi_1, \xi_2, \ldots$ be positive, iid random variables with $E[\xi_i] = 1$, and let $\mathcal{F}_k = \sigma(\xi_i, i \leq k)$. Define
$$M_k = \prod_{i=1}^{k} \xi_i.$$
Show that $\{M_k\}$ is a $\{\mathcal{F}_k\}$-martingale.

4. Let $N(t)$ be a counting process with $E[N(t)] < \infty$, and let $\xi_1, \xi_2, \ldots$ be positive, iid random variables with $E[\xi_i] = 1$ that are independent of N. Define $M(t) = \prod_{i=1}^{N(t)} \xi_i$. Show that M is a martingale. (Justify your answer using the properties of conditional expectations.)

5. Let Y be a Poisson process with intensity λ, and let $\xi_1, \xi_2, \ldots$ be iid with mean m and variance σ^2 that are independent of Y. Define
$$X(t) = \sum_{k=1}^{Y(t)} \xi_k$$
Show that X has stationary, independent increments, and calculate $E[X(t) - X(s)]$ and $Var(X(t) - X(s))$.

6. Let τ be a discrete stopping time with range $\{t_1, t_2, \ldots\}$. Show that
$$E[Z|\mathcal{F}_\tau] = \sum_{k=1}^{\infty} E[Z|\mathcal{F}_{t_k}] I_{\{\tau = t_k\}}\ .$$

7. (a) Let W denote standard Brownian motion, and let $\{t_i\}$ be a partition of the interval $[0, t]$. What is $\lim \sum |W(t_{i+1}) - W(t_i)|^3$ as $\max |t_{i+1} - t_i| \to 0$?
   (b) What is $\lim \sum W(t_i) \left((W(t_{i+1}) - W(t_i))^2 - (t_{i+1} - t_i)\right)$ as $\max |t_{i+1} - t_i| \to 0$?

(c) Use the limits in part a) and b) to directly calculate $\int_0^t W^2(s-)dW(s)$ from the definition of the stochastic integral.

8. Let $0 \le \tau_1 \le \tau_2 \le \cdots$ be $\{\mathcal{F}_t\}$-stopping times, and for $k = 1, 2, \ldots$, let ξ_k be $\mathcal{F}_{\tau_k}$-measurable. Define

$$X(t) = \sum_{k=1}^{\infty} \xi_k I_{[\tau_k, \tau_{k+1})}(t),$$

and show that X is $\{\mathcal{F}_t\}$-adapted

9. (a) For each $n > 0$, show that $M_n(t) = \int_0^t W(s-)^n dW(s)$ is square integrable and that M_n is a martingale. (You do not need to explicitly compute M_n.)

   (b) Show that $Z(t) = \int_0^t e^{W(s-)^4} dW(s)$ is a local martingale. (It is not a martingale, since it is not integrable.) In particular, find a sequence of stopping times τ_n such that $Z(\cdot \wedge \tau_n)$ is a martingale.

10. Let Y be a Poisson process with intensity λ.

    (a) Find a cadlag process U such that

$$e^{-\alpha Y(t)} = 1 + \int_0^t U(s-)dY(s) \tag{15.1}$$

    (b) Use (15.1) and the fact that $Y(t) - \lambda t$ is a martingale to compute $E[e^{-\alpha Y(t)}]$.

    (c) Define

$$Z(t) = \int_0^t e^{-\alpha Y(s-)} dY(s).$$

    Again use the fact that $Y(t) - \lambda t$ is a martingale to calculate $E[Z(t)]$ and $Var(Z(t))$.

11. Let N be a Poisson process with parameter λ, and let $X_1, X_2, \ldots$ be a sequence of Bernoulli trials with parameter p. Assume that the X_k are independent of N. Let

$$M(t) = \sum_{k=1}^{N(t)} X_k.$$

What is the distribution of $M(t)$?

12. Let N be a Poisson process with parameter λ. For $t < s$:

    (a) What is the covariance of $N(t)$ and $N(s)$?

    (b) Calculate the probability that $P\{N(t) = 1, N(s) = 1\}$.

    (c) Give an event in terms of S_1 and S_2 that is equivalent to the event $\{N(t) = 1, N(s) = 1\}$, and use the calculation in part 12b to calculate the joint density function for S_1 and S_2.

13. Let Y be a continuous semimartingale. Solve the stochastic differential equation

$$dX = aX\,dt + bX\,dY, \qquad X(0) = x_0$$

Hint: Look for a solution of the form $X(t) = A\exp\{Bt + CY(t) + D[Y]_t\}$ for some set of constants, A, B, C, D.

14. 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 $\lambda > 0$.

15. 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 $t \geq 0$.

16. 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) Suppose that M_n are pairwise orthogonal, square integrable martingales (that is, $[M_n, M_m]_t = 0$ for $n \neq m$). Suppose that $\sum_{k=1}^{\infty} E[[M_n]_t] < \infty$ for each t. Show that

    $$M \equiv \sum_{k=1}^{\infty} M_n$$

    converges in L^2 and that M is a square integrable martingale with $[M] = \sum_{k=1}^{\infty}[M_k]$.

17. Let $X_1, X_2 \ldots$ and $Y_1, Y_2, \ldots$ be independent, unit Poisson processes. For $\lambda_k > 0$ and $c_k \in \mathbb{R}$, define

$$M_n(t) = \sum_{k=1}^{n} c_k(Y_k(\lambda_k t) - X_k(\lambda_k t))$$

    (a) Suppose $\sum c_k^2 \lambda_k < \infty$. Show that for each $T > 0$,

    $$\lim_{n,m\to\infty} E[\sup_{t\leq T}(M_n(t) - M_m(t))^2] = 0$$

    and hence we can define

    $$M(t) = \sum_{k=1}^{\infty} c_k(Y_k(\lambda_k t) - X_k(\lambda_k t))$$

(b) Under the assumptions of part (a), show that M is a square integrable martingale, and calculate $[M]$.

18. Suppose in Problem 17 that $\sum c_k^2 \lambda_k < \infty$ but $\sum |c_k \lambda_k| = \infty$. Show that for $t > 0, T_t(M) = \infty$ a.s. (Be careful with this. In general the total variation of a sum is not the sum of the total variations.)

19. 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.)

20. Let N be a Poisson process with parameter λ. Use Ito's formula to show that

$$M(t) = e^{\alpha N(t) - \lambda(e^{\alpha} - 1)t}$$

is a martingale.

21. 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.

22. 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 $\alpha < x < \beta$, $P\{\tau < \infty, X(\tau) = \alpha | X(0)\} = x\} > 0$ 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 γ 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 $E[\tau] < \infty$.

23. Let $dX = -bX^2 dt + cX dW$, $X(0) > 0$.

(a) Show that $X(t) > 0$ for all t a.s.

(b) For what values of b and c does $\lim_{t\to\infty} X(t) = 0$ a.s.?

24. Let $dX = (a - bX)dt + \sqrt{X}dW$, $X(0) > 0$ where a and b are positive constants. Let $\tau = \inf\{t > 0 : X(t) = 0\}$.

    (a) For what values of a and b is $P\{\tau < \infty\} = 1$?

    (b) For what values of a and b is $P\{\tau = \infty\} = 1$?

25. Let M be a k-dimensional, continuous Gaussian process with stationary, mean zero, independent increments and $M(0) = 0$. Let B be a $k\times k$-matrix all of whose eigenvalues have negative real parts, and let X satisfy

$$X(t) = x + \int_0^t BX(s)ds + M(t)$$

Show that $\sup_t E[|X(t)|^n] < \infty$ for all n. (Hint: Let $Z(t) = CX(t)$ for a judiciously selected, nonsingular C, show that Z satisfies an equation of the same form as X, show that $\sup_t E[|Z(t)|^n] < \infty$, and conclude that the same must hold for X.)

26. Let $X(t, x)$ be as in (8.3) with σ and b continuous. Let $D \subset \mathbb{R}^d$ be open, and let $\tau(x)$ be the exit time from D starting from x, that is,

$$\tau(x) = \inf\{t : X(t, x) \notin D\}.$$

Assume that h is bounded and continuous. Suppose $x_0 \in \partial D$ and

$$\lim_{x\to x_0} P[\tau(x) < \varepsilon] = 1, \forall \varepsilon > 0.$$

Show that

$$\lim_{x\to x_0} P[|X(\tau(x), x) - x_0| > \varepsilon] = 0 \forall \varepsilon > 0,$$

and hence, if f is defined by (8.4), $f(x) \to h(x_0)$, as $x \to x_0$.

27. (Central Limit Theorem for Random Matrices) Let $A_1, A_2, \ldots$ be independent, identically distributed, matrix-valued random variables with expectation zero and finite variance. Define

$$\begin{aligned} Y_n(t) &= \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} A_k \\ X_n(t) &= (I + \frac{1}{\sqrt{n}}A_1)(I + \frac{1}{\sqrt{n}}A_2)\cdots(I + \frac{1}{\sqrt{n}}A_{[nt]}) \end{aligned}$$

Show that X_n satisfies a stochastic differential equation driven by Y_n, conclude that the sequence X_n converges in distribution, and characterize the limit.

In Problems 28 and 29, assume that $d = 1$, that σ and β are continuous, and that

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

Recall that $Lf(x) = \frac{1}{2}\sigma(x)^2 f''(x) + b(x)f'(x)$.

28. Let $\alpha < \beta$, and suppose that there is a function f that is C^2 and satisfies $Lf(x) \geq \delta > 0$ on $[\alpha, \beta]$. Define $\tau = \inf\{t : X(t) \notin (\alpha, \beta)\}$. Show that $E[\tau] < \infty$.

29. Let $\alpha < \beta$, and suppose that $\inf_{x \in (\alpha,\beta)} \sigma(x)^2 > 0$. Define $\tau = \inf\{t : X(t) \notin (\alpha, \beta)\}$. Show that $E[\tau] < \infty$.

In Problems 30 through 33, let X be real-valued and satisfy $dX = \sigma(X)dW + b(X)dt$ where σ is bounded and strictly positive, and suppose that $v(x) = x^2$ satisfies $Lv \leq K - \epsilon v$ for some $\epsilon > 0$. Let $\tau_0 = \inf\{t : X(t) \leq 0\}$.

30. Show that if $E[X(0)^2] < \infty$, then $E[\tau_0] < \infty$.

31. Assume that $E[X(0)^2] < \infty$. Show that if f is twice continuously differentiable with bounded first derivative, then

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t Lf(X(s))ds = 0$$

where convergence is in L^2. (Convergence is also almost sure. You can have 2 points extra credit if you show that.)

32. Show that for every bounded continuous g, there exists a constant c_g such that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t g(X(s))ds = c_g$$

Hint: Show that there exists c_g and f such that $Lf = g - c_g$. Remark. In fact, under these assumptions the diffusion process has a stationary distribution π and $c_g = \int g d\pi$.

33. Show that if f is twice continuously differentiable with bounded first derivative, then

$$W_n(t) \equiv \frac{1}{\sqrt{n}} \int_0^{nt} Lf(X(s)ds$$

converges in distribution to αW for some constant α.

Ornstein Unlenbeck Process. (Problems 34-40.) The Ornstein-Uhlenbeck process was originally introduced as a model for the velocity of "physical" Brownian motion,

$$dV = -\lambda V dt + \sigma dW ,$$

where $\sigma, \lambda > 0$. The location of a particle with this velocity is then given by

$$X(t) = X(0) + \int_0^t V(s)ds.$$

Explore the relationship between this model for physical Brownian motion and the usual mathematical model. In particular, if space and time are rescaled to give

$$X_n(t) = \frac{1}{\sqrt{n}} X(nt),$$

what happens to X_n as $n \to \infty$?

34. Derive the limit of X_n directly from the stochastic differential equation. (Note: No fancy limit theorem is needed.) What type of convergence do you obtain?

35. Calculate $E[V(t)^4]$. Show (without necessarily calculating explicitly), that if $E[|V(0)|^k] < \infty$, then $\sup_{0 \le t < \infty} E[|V(t)|^k] < \infty$.

36. Compute the stationary distribution for V. (One approach is given Section 9. Can you find another.)

37. Let g be continous with compact support, and let $c_g = \int g d\pi$, where π is the stationary distribution for V. Define

$$Z_n(t) = \frac{1}{\sqrt{n}} \left(\int_0^{nt} g(V(s))ds - nC_g t \right).$$

Show that Z_n converges in distribution and identify the limit.

38. Note that Problem 34 is a result of the same form as Problem 37 with $g(v) = v$, so the condition that g be continuous with compact support in Problem 37 is not necessary. Find the most general class of g's you can, for which a limit theorem holds.

39. Consider $X(t) = X(0) + \int_0^t V(s)ds$ with the following modification. Assume that $X(0) > 0$ and keep $X(t) \ge 0$ by switching the sign of the velocity each time X hits zero. Derive the stochastic differential equation satisfied by (X, V) and prove the analogue of Problem 34. (See Lemma 11.1.)

40. Consider the Ornstein-Unlenbeck process in $\mathbb{R}^d$

$$dV = -\lambda V dt + \sigma dW$$

where W is now d-dimensional Brownian motion. Redo as much of the above as you can. In particular, extend the model in Problem 39 to convex sets in $\mathbb{R}^d$.

41. Let X be a diffusion process with generator L. Suppose that h is bounded and C^2 with $h \ge \epsilon > 0$ and that Lh is bounded. Show that

$$L(t) = \frac{h(X(t))}{h(X(0))} \exp\left\{ -\int_0^t \frac{Lh(X(s))}{h(X(s))} ds \right\}$$

is a martingale with $E[L(t)] = 1$.

42. For $x>0$, let

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

and $\tau=\inf\{t:X(t)=0\}$. Give conditions on σ, as general as you can make them, that imply $E[\tau]<\infty$.

43. Let $X(t)=X(0)+W(t)$ where $W=(W_1,W_2)$ is two-dimensional standard Brownian motion. Let $Z=(R,\Theta)$ be the polar coordinates for the point (X_1,X_2). Derive a stochastic differential equation satisfied by Z. Your answer should be of the form

$$Z(t)=Z(0)+\int_0^t\sigma(Z(s))dW(s)+\int_0^t b(Z(s))ds.$$

44. Assume that $d=1$. Let

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

where σ and b have bounded continuous derivatives. Derive the stochastic differential equation that should be satisfied by

$$Y(t,x)=\frac{\partial}{\partial x}X(t,x)$$

if the derivative exists and show that

$$Y(t,x)\equiv\lim_{h\to 0}\frac{1}{h}(X(t,x+h)-X(t,x))$$

exists in L^2 where Y is the solution of the derived equation.

45. Let N be a unit Poisson process and let

$$W_n(t)=\frac{1}{\sqrt{n}}\int_0^{nt}(-1)^{N(s)}ds$$

(Recall that $W_n\Rightarrow W$ where W is standard Brownian motion.) Show that there exist martingales M_n such that $W_n=M_n+V_n$ and $V_n\to 0$, but $T_t(V_n)\to\infty$.

46. Let W_n be as in Problem 45. Let σ have a bounded, continuous derivative, and let

$$X_n(t)=\int_0^t\sigma(X_n(s))dW_n(s).$$

Show that $X_n\Rightarrow X$ for some X and identify the stochastic differential equation satisfied by X. Note that by Problem 45, the conditions of Theorem 10.13 are not satisfied for

$$X_n(t)=\int_0^t\sigma(X_n(s-))dM_n(s)+\int_0^t\sigma(X_n(s-))dV_n(s).\tag{15.2}$$

Integrate the second term on the right of (15.2) by parts, and show that the sequence of equations that results, does satisfy the conditions of Theorem 10.13.

Central limit theorem for Markov chains. (Problems 47-54.) Let $\xi_0, \xi_1, \ldots$ be an irreducible Markov chain on a finite state space $\{1, ..., d\}$, let $P = ((p_{ij}))$ denote its transition matrix, and let π be its stationary distribution. For any function h on the state space, let πh denote $\sum_i \pi_i h(i)$.

47. Show that
$$f(\xi_n) - \sum_{k=0}^{n-1}(Pf(\xi_k) - f(\xi_k))$$
is a martingale.

48. Show that for any function h, there exists a solution to the equation $Pg = h - \pi h$, that is, to the system
$$\sum_j p_{ij} g(j) - g(i) = h(i) - \pi h.$$

49. The ergodic theorem for Markov chains states that
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} h(\xi_k) = \pi h.$$
Use the martingale central limit theorem to prove convergence in distribution for
$$W_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}(h(\xi_k) - \pi h)\ .$$

50. Use the martingale central limit theorem to prove the analogue of Problem 49 for a continuous time finite Markov chain $\{\xi(t), t \geq 0\}$. In particular, use the multidimensional theorem to prove convergence for the vector-valued process $U_n = (U_n^1, \ldots, U_n^d)$ defined by
$$U_n^k(t) = \frac{1}{\sqrt{n}}\int_0^{nt}(I_{\{\xi(s)=k\}} - \pi_k)ds$$

51. Explore extensions of Problems 49 and 50 to infinite state spaces.

Limit theorems for stochastic differential equations driven by Markov chains

52. Show that W_n defined in Problem 49 and U_n defined in Problem 50 are not "good" sequences of semimartingales, in the sense that they fail to satisfy the hypotheses of Theorem 10.13. (The easiest approach is probably to show that the conclusion is not valid.)

53. Show that W_n and U_n can be written as $M_n + Z_n$ where $\{M_n\}$ is a "good" sequence and $Z_n \Rightarrow 0$.

54. (Random evolutions) Let ξ be as in Problem 50, 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 equations driven by U_n, give conditions under which X_n converges in distribution to a limit X, and identify the limit.

55. (a) Let W be a standard Brownian motion, let σ_i, $i = 1, 2$, be bounded, continuous functions, and suppose that

$$X_i(t) = X_i(0) + \int_0^t \sigma_i(s) X_i(s) dW(s), \quad i = 1, 2.$$

Apply Itô's formula to find an SDE satisfied by $Z = X_1 X_2$.

(b) Let W_1 and W_2 be independent standard Brownian motions. Let

$$Y_i(t) = Y_i(0) + \int_0^t \sigma_i(s) Y_i(s) dW_i(s), \quad i = 1, 2.$$

Find an SDE satisfied by $U = Y_1 Y_2$, and show that U is a martingale.

56. Suppose the price X of a tradeable asset is the unique solution of

$$X(t) = X(0) + Y(\int_0^t \lambda(X(s))ds) - \int_0^t \mu(X(s))ds, \tag{15.3}$$

where Y is a unit Poisson process, $X(0)$ is independent of Y, Y and $X(0)$ are defined on $(\Omega, \mathcal{F}, P)$, and λ and μ are bounded and strictly positive. Let $\mathcal{F}_t = \sigma(X(s) : s \leq t)$. Find a measure Q equivalent to P such that $\{X(t), 0 \leq t \leq T\}$ is a martingale under Q and $X(0)$ has the same distribution under Q as under P.

57. Suppose that $\mu : \mathbb{R} \to \mathbb{R}$ is Lipshitz.

(a) Show that the solution of (15.3) is unique.

(b) Let u be cadlag. Show that

$$x(t) = u(t) - \int_0^t \mu(x(s))ds \tag{15.4}$$

has a unique solution and that x is cadlag.

(c) Let $\Gamma(t, u) = x(t)$, where x is the unique solution of (15.4). Show that Γ is nonanticipating in the sense that $\Gamma(t, u) = \Gamma(t, u^t)$, $t \geq 0$, where $u^t(s) = u(s \wedge t)$.

58. (Extra credit) Show that Q in Problem 56 is unique. (You might begin by showing that the distribution of X is the same under any Q satisfying the conditions of the problem.)

59. Let $\alpha, \beta \in (0, \infty)^2$, and let $X = (X_1, X_2)$ satisfy

$$X(t) = X(0) + \alpha Y_1(\int_0^t \lambda_1(X(s))ds) - \beta Y_2(\int_0^t \lambda_2(X(s))ds),$$

where Y_1 and Y_2 are independent, unit Poisson processes independent of $X(0)$; $Y_1, Y_2, X(0)$ are defined on $(\Omega, \mathcal{F}, P)$; α and β are linearly independent; and $0 < \epsilon \leq \lambda_1, \lambda_2 \leq \epsilon^{-1}$, for some $\epsilon > 0$. Show that there exists a probability measure Q equivalent to P under which X is a martingale and $X(0)$ has the same distribution under Q as under P.

References.

Durrett, Richard (1991). *Probability: Theory and Examples.* Wadsworth and Brooks/Cole, Belmont.

Ethier, Stewart N., and Kurtz, Thomas G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

Protter, Philip (1990). *Stochastic Integration and Differential Equations.* Springer-Verlag, Berlin.

Kurtz, Thomas G. and Protter, Philip (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* 19, 1035-1070.