Semimartingales and stochastic integration
Spring 2011
Sergio Pulido*

Chris Almost†

Contents

*spulido@andrew.cmu.edu
†cdalmost@cmu.edu

Chapter 0

Motivation

Why stochastic integration with respect to semimartingales with jumps?

- To model "unpredictable" events (e.g. default times in credit risk theory) one needs to consider models with jumps.
- A lot of interesting stochastic processes jump, e.g. Poisson process, Lévy processes.

This course will closely follow the textbook, *Stochastic integration and differential equations* by Philip E. Protter, second edition. We will not cover every chapter, and some proofs given in the course will differ from those in the text. The following numbers correspond to sections in the textbook.

I Preliminaries.
  1. Filtrations, stochastic processes, stopping times, path regularity, "functional" monotone class theorem, optional σ-algebra.
  2. Martingales.
  3. Poisson processes, Brownian motion.
  4. Lévy processes.
  6. Localization procedure for stochastic processes.
  7. Stieltjes integration
  8. Impossibility of naïve stochastic integration (via the Banach-Steinhaus theorem).

II Semimartingales and stochastic integrals.
  1–3. Definition of the stochastic integral with respect to processes in $\mathbb{L}$.
  5. Properties of the stochastic integral.
  6. Quadratic variation.
  7. Itô's formula.
  8. Stochastic exponential and Lévy's characterization theorem.

III Bichteler-Dellacherie theorem.
  (NFLVR) implies S is a semimartingale
  (NFLVR) and little investment if and only if S is a semimartingale

IV Stochastic integration with respect to predictable processes and martingale representation theorems (i.e. market completeness).

For more information on the history of the development of stochastic integration, see the paper by Protter and Jarrow on that topic.

Chapter 1

Preliminaries

1.1 Review of stochastic processes

The standard setup we will use is that of a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a filtration $\mathbb{F} = (\mathscr{F}_t)_{0 \le t \le \infty}$ of sub-σ-algebras of $\mathscr{F}$. The filtration can be thought of as the flow of information. Expectation $\mathbb{E}$ will always be with respect to $\mathbb{P}$ unless stated otherwise.

Notation. We will use the convection that t, s, and u will always be real variables, not including ∞ unless it is explicitly mentioned, e.g. $\{t | t \ge 0\} = [0, \infty)$. On the other hand, n and k will always be integers, e.g. $\{n : n \ge 0\} = \{0, 1, 2, 3, \dots\} =: \mathbb{N}$.

1.1.1 Definition. $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$ satisfies the *usual conditions* if
(i) $\mathscr{F}_0$ contains all the $\mathbb{P}$-null sets.
(ii) $\mathbb{F}$ is right continuous, i.e. $\mathscr{F}_t = \bigcap_{s<t} \mathscr{F}_s$.

1.1.2 Definition. A *stopping time* is a *random time*, i.e. a measurable function $T : \Omega \to [0, \infty]$, such that $\{T \le t\} \in \mathscr{F}_t$ for all t.

1.1.3 Theorem. *The following are equivalent.*
(i) *T is a stopping time.*
(ii) *$\{T < t\} \in \mathscr{F}_t$ for all $t > 0$.*

Proof: Assume T is a stopping time. Then for any $t > 0$,

$$\{T < t\} = \bigcup_{n \ge 1} \{T \le t - \tfrac{1}{n}\} \in \bigvee_{s<t} \mathscr{F}_s \subseteq \mathscr{F}_t.$$

Conversely, since the filtration is assumed to be right continuous,

$$\{T \le t\} = \bigcap_{n \ge 1} \{T < t + \tfrac{1}{n}\} \in \bigcap_{s>t} \mathscr{F}_s = \mathscr{F}_{t+} = \mathscr{F}_t,$$

so T is a stopping time. □

1.1.4 Theorem.

(i) *If* $(T_n)_{n\geq 1}$ *is a sequence of stopping times then* $\bigwedge_n T_n$ *and* $\bigvee_n T_n$ *are stopping times.*

(ii) *If* T *and* S *are stopping times then* $T + S$ *is a stopping time.*

1.1.5 Exercises.

(i) If $T \geq S$ then is $T - S$ a stopping time?

(ii) For which constants α is αT a stopping time?

SOLUTION: Clearly αT need not be a stopping time if $\alpha < 1$. If $\alpha \geq 1$ then, for any $t \geq 0$, $t/\alpha \leq t$ so $\{\alpha T \leq t\} = \{T \leq t/\alpha\} \in \mathscr{F}_{t/\alpha} \subseteq \mathscr{F}_t$ and αT is a stopping time.

Let H be a stopping time for which $H/2$ is not a stopping time (e.g. the first hitting time of Brownian motion at the level 1). Take $T = 3H/2$ and $S = H$, both stopping times, and note that $T - S = H/2$ is not a stopping time. ✠

1.1.6 Definition. Let T be a stopping time. The σ-algebras of *events before time* T and *events strictly before time* T are

(i) $\mathscr{F}_T = \{A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_t$ for all $t\}$.

(ii) $\mathscr{F}_{T-} = \mathscr{F}_0 \vee \sigma\{A \cap \{t < T\} : t \in [0,\infty), A \in \mathscr{F}_t\}$.

1.1.7 Definition.

(i) A *stochastic process* X is a collection of $\mathbb{R}^d$-valued r.v.'s, $(X_t)_{0\leq t<\infty}$. A stochastic process may also be thought of as a function $X : \Omega \times [0,\infty) \to \mathbb{R}^d$ or as a random element of a space of paths.

(ii) X is *adapted* if $X_t \in \mathscr{F}_t$ for all t.

(iii) X is *càdlàg* $(X_t(\omega))_{t\geq 0}$ has left limits and is right continuous for almost all ω. X is *càglàd* if instead the paths have right limits and are left continuous.

(iv) X is a *modification* of Y if $\mathbb{P}[X_t \neq Y_t] = 0$ for all t. X is *indistinguishable* from Y if $\mathbb{P}[X_t \neq Y_t$ for some $t] = 0$.

(v) $X_- := (X_{t-})_{t\geq 0}$ where $X_{0-} := 0$ and $X_{t-} := \lim_{s\uparrow t} X_s$.

(vi) For a càdlàg process X, $\Delta X := (\Delta X_t)_{t\geq 0}$ is the *process of jumps* of X, where $\Delta X_t := X_t - X_{t-}$.

1.1.8 Theorem.

(i) *If* X *is a modification of* Y *and* X *and* Y *are left- (right-) continuous then they are indistinguishable.*

(ii) *If* $\Lambda \subseteq \mathbb{R}^d$ *is open and* X *is continuous from the right (càd) and adapted then* $T := \inf\{t > 0 : X_t \in \Lambda\}$ *is a stopping time.*

(iii) *If* $\Lambda \subseteq \mathbb{R}^d$ *is closed and* X *is càd and adapted then* $T := \inf\{t > 0 : X_t \in \Lambda$ *or* $X_{t-} \in \Lambda\}$ *is a stopping time.*

(iv) *If* X *is càdlàg and adapted and* $\Delta X_T \mathbf{1}_{T<\infty} = 0$ *for all stopping times* T *then* ΔX *is indistinguishable from the zero process.*

PROOF: Read the proofs of these facts as an exercise. □

1.1.9 Definition. $\mathscr{O} = \sigma(X : X$ is adapted and càdlàg) is the *optional σ-algebra*. A stochastic process X is an *optional process* if X is $\mathscr{O}$-measurable.

1.1.10 Theorem (Début theorem). *If $A \in \mathscr{O}$ then $T(\omega) := \inf\{t : (\omega, t) \in A\}$, the début time of A, is a stopping time.*

Remark. This theorem requires that the filtration is right continuous. For example, suppose that T is the hitting time of an open set by a left continuous process. Then you can prove $\{T < t\} \in \mathscr{F}_t$ for all t without using right continuity of the filtration, but you cannot necessarily prove that $\{T = t\} \in \mathscr{F}_t$ without it. You need to "look into the future" a little bit.

1.1.11 Corollary. *If X is optional and $B \subseteq \mathbb{R}^d$ is a Borel set then the hitting time $T := \inf\{t > 0 : X_t \in B\}$ is a stopping time.*

1.1.12 Theorem. *If X is an optional process then*
(i) *X is $(\mathscr{F} \otimes \mathscr{B}([0,\infty)))$-measurable and*
(ii) *$X_T \mathbf{1}_{T<\infty} \in \mathscr{F}_T$ for any stopping time T.*
In particular, $(X_{t\wedge T})_{t\geq 0}$ is also an optional process, i.e. optional processes are "stable under stopping".

1.1.13 Theorem (Monotone class theorem).
Suppose that H is collection of bounded $\mathbb{R}$-valued functions such that
(i) *H is a vector space.*
(ii) *$\mathbf{1}_\Omega \in H$, i.e. constant functions are in H.*
(iii) *If $(f_n)_{n\geq 0} \subseteq H$ is monotone increasing and $f := \lim_{n\to\infty} f_n$ (pointwise) is bounded then $f \in H$.*
(In this case H is called a monotone vector space.*) Let M be a multiplicative collection of bounded functions (i.e. if $f, g \in M$ then $fg \in M$). If $M \subseteq H$ then H contains all the bounded functions that are measurable with respect to $\sigma(M)$.*

Proof (of Theorem 1.1.12): Use the Monotone Class Theorem. Define

$$M := \{X : \Omega \times [0,\infty) \to \mathbb{R} \mid X \text{ is càdlàg, adapted, and bounded}\}$$
$$H := \{X : \Omega \times [0,\infty) \to \mathbb{R} \mid X \text{ is bounded and (i) and (ii) hold}\}$$

It can be checked that H is a monotone vector space, M is a multiplicative collection, and $\sigma(M) = \mathscr{O}$. If we prove that $M \subseteq H$ then we are done. Let $X \in M$ and define

$$X^{(n)} := \sum_{k=1}^{\infty} X_{\frac{k}{2^n}} \mathbf{1}_{[\frac{k-1}{2^n}, \frac{k}{2^n})}.$$

Since X is right continuous, $X^{(n)} \to X$ pointwise. Let B be a Borel set.

$$\{X^{(n)} \in B\} = \bigcup_{k=1}^{\infty} \left\{X_{\frac{k}{2^n}} \in B\right\} \times \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \in \mathscr{F} \otimes \mathscr{B}([0,\infty))$$

This proves that X satisfies (i). To prove (ii), let T be a stopping time and define

$$T_n := \begin{cases} \frac{k}{2^n} & \text{if } \frac{k-1}{2^n} \le T < \frac{k}{2^n} \\ \infty & \text{if } T = \infty. \end{cases}$$

Then the T_n's are stopping times and $T_n \downarrow T$.

$$\{X_{T_n} \in B\} \cap \{T_n \le t\} = \bigcup_{\substack{k=1 \\ \frac{k}{2^n} \le t}}^{\infty} \left\{X_{\frac{k}{2^n}} \in B\right\} \cap \left\{T_n = \frac{k}{2^n}\right\} \in \mathscr{F}_t$$

This shows that $X_{T_n} \in \mathscr{F}_{T_n}$. Since X is right continuous,

$$X_T \mathbf{1}_{T<\infty} = \lim_{n\to\infty} X_{T_n} \mathbf{1}_{T_n<\infty}.$$

It can be shown that, since the filtration is right continuous, $\bigcap_{n=1}^{\infty} \mathscr{F}_{T_n} = \mathscr{F}_T$, so $X_T \mathbf{1}_{T<\infty} \in \mathscr{F}_T$ and X satisfies (ii). Therefore $X \in H$. □

If X is càdlàg and adapted then an argument similar to that in the previous proof shows that $X|_{\Omega\times[0,t]} \in \mathscr{F}_t \otimes \mathscr{B}([0,t])$ for all t, i.e. X is *progressively measurable*. By a similar monotone class argument, it can be shown that every optional process is a *progressive process*.

1.1.14 Definition. $\mathscr{V} := \sigma(X : X$ is progressively measurable$)$ is the *progressive σ-algebra*.

1.1.15 Corollary. $\mathscr{O} \subseteq \mathscr{V}$.

1.2 Review of martingales

1.2.1 Theorem. *Let X be a (sub-, super-) martingale and assume that $\mathbb{F}$ satisfies the usual conditions. Then X has a right continuous modification if and only if the function $t \mapsto \mathbb{E}[X_t]$ is right continuous. Furthermore, this modification has left limits everywhere.*

Proof (sketch): The process $\widetilde{X}_t := \lim_{\substack{s\downarrow t \\ s\in\mathbb{Q}}} X_s$ is the correct modification. □

1.2.2 Corollary. *Every martingale has a càdlàg modification, unique up to indistinguishability.*

1.2.3 Theorem. *Let X be a right continuous sub-martingale with $\sup_{t\ge0} \mathbb{E}[X_t^+] < \infty$. Then $X_\infty := \lim_{t\to\infty} X_t$ exists a.s. and $X_\infty \in L^1$.*

1.2.4 Definition. A collection of random variables $(U_\alpha)_{\alpha\in A}$ is *uniformly integrable* or *u.i.* if

$$\lim_{n\to\infty} \sup_{\alpha\in A} \mathbb{E}[\mathbf{1}_{|U_\alpha|>n} U_\alpha] = 0.$$

1.2.5 Theorem. *The following are equivalent for a family $(U_\alpha)_{\alpha\in A}$.*

(i) $(U_\alpha)_{\alpha\in A}$ *is u.i.*

(ii) $\sup_{\alpha\in A}\mathbb{E}[|U_\alpha|]<\infty$ *and for all* $\varepsilon>0$ *there is* $\delta>0$ *such that if* $\mathbb{P}[\Lambda]<\delta$ *then* $\sup_{\alpha\in A}\mathbb{E}[\mathbf{1}_\Lambda U_\alpha]<\varepsilon$.

(iii) *(de la Vallée-Poussin criterion) There is a positive, increasing, convex function* G *on* $[0,\infty)$ *such that* $\lim_{x\to\infty}\frac{G(x)}{x}=\infty$ *and* $\sup_{\alpha\in A}\mathbb{E}[G(|U_\alpha|)]<\infty$.

1.2.6 Theorem. *The following are equivalent for a martingale* X.

(i) $(X_t)_{t\geq 0}$ *is u.i.*

(ii) X *is closable, i.e. there is an integrable r.v.* Z *such that* $X_t=\mathbb{E}[Z|\mathscr{F}_t]$.

(iii) X *converges in* L^1.

(iv) $X_\infty=\lim_{t\to\infty}X_t$ *a.s. and in* L^1 *and* X_∞ *closes* X.

Remark.

(i) If X is u.i. then X is bounded in L^1 (but not *vice versa*), so Theorem 1.2.3 implies that $X_t\to X_\infty$ a.s. Theorem 1.2.6 upgrades this to convergence in L^1, i.e. $\mathbb{E}[|X_t-X_\infty|]\to 0$.

(ii) If X is closed by Z then $\mathbb{E}[Z|\mathscr{F}_\infty]$ also closes X, where $\mathscr{F}_\infty:=\bigvee_{t\geq 0}\mathscr{F}_t$. Furthermore, $X_\infty=\mathbb{E}[Z|\mathscr{F}_\infty]$.

1.2.7 Example (Simple random walk). Let $(Z_n)_{n\geq 1}$ be i.i.d. with

$$\mathbb{P}[Z_n=1]=\mathbb{P}[Z_n=-1]=\frac{1}{2}.$$

Take $\mathscr{F}_t:=\sigma\{Z_k:k\leq t\}$ and $X_t:=\sum_{k=1}^{\lfloor t\rfloor}Z_k$. Then X is not a closable martingale.

1.2.8 Example. Let $M_t:=\exp(B_t-\frac{1}{2}t)$, the stochastic exponential of Brownian motion, a martingale. Then $M_t\to 0$ a.s. by Theorem 1.2.3 because it is a positive valued martingale (hence $-M$ is a sub-martingale with no positive part). But $\mathbb{E}[|M_t|]=1$ for all t so M is not u.i. by Theorem 1.2.6.

1.2.9 Theorem.

(i) *If* X *is a closable (sub-, super-) martingale and* $S\leq T$ *are stopping times then* $\mathbb{E}[X_T|\mathscr{F}_S]=X_S$ $(\geq,\leq)$.

(ii) *If* X *is a (sub-, super-) martingale and* $S\leq T$ *are bounded stopping times then* $\mathbb{E}[X_T|\mathscr{F}_S]=X_S$ $(\geq,\leq)$.

(iii) *If* X *is a right continuous sub-martingale and* $p>1$ *then*

$$\|\sup_{t\geq 0}X_t\|_{L^p}\leq\left(\frac{p}{p-1}\right)\sup_{t\geq 0}\|X_t\|_{L^p}.$$

In particular if $p=2$ *then* $\mathbb{E}[\sup_{t\geq 0}X_t^2]\leq 4\sup_{t\geq 0}\mathbb{E}[X_t^2]$.

(iv) *(Jensen's inequality) If* φ *is a convex function,* Z *is an integrable r.v., and* $\mathscr{G}$ *is a sub-*σ*-algebra of* $\mathscr{F}$ *then* $\varphi(\mathbb{E}[Z|\mathscr{G}])\leq\mathbb{E}[\varphi(Z)|\mathscr{G}]$.

1.2.10 Definition. Let X be a process and T be a random time. The *stopped process* is $X_t^T := X_t \mathbf{1}_{t\le T} + X_T \mathbf{1}_{t>T,T<\infty} = X_{T\wedge t}$.

If T is a stopping time and X is càdlàg and adapted then so is X^T.

1.2.11 Theorem.
(i) *If X is a u.i. martingale and T is a stopping time then X^T is a u.i. martingale.*
(ii) *If X is a martingale and T is a stopping time then X^T is a martingale.*
That is, martingales are "stable under stopping".

1.2.12 Definition. Let X be a martingale.
(i) If $X_t \in L^2$ for all t then X is called a *square integrable martingale*.
(ii) If $(X_t)_{t\in[0,\infty)}$ is u.i. and $X_\infty \in L^2$ then X is called an L^2-martingale.

1.2.13 Exercise. Any L^2-martingale is a square integrable martingale, but not conversely.

SOLUTION: Let X be an L^2-martingale. Then $X_\infty \in L^2$ and so, by the conditional version of Jensen's inequality,

$$\mathbb{E}[X_t^2] = \mathbb{E}[(\mathbb{E}[X_\infty|\mathscr{F}_t])^2] \le \mathbb{E}[\mathbb{E}[X_\infty^2|\mathscr{F}_t]] = \mathbb{E}[X_\infty^2] < \infty.$$

Therefore $X_t \in L^2$ for all t. We have already seen that the stochastic exponential of Brownian motion is not u.i. (and hence not an L^2-martingale) but it is square integrable because the normal distribution has a finite valued moment generating function. ✠

1.3 Poisson process and Brownian motion

1.3.1 Definition. Suppose that $(T_n)_{n\ge1}$ is a strictly increasing sequence of random times with $T_1 > 0$ a.s. The process $N_t = \sum_{n\ge1} \mathbf{1}_{T_n\le t}$ is the *counting process* associated with $(T_n)_{n\ge1}$. The random time $T := \sup_{n\ge1} T_n$ is the *explosion time*. If $T = \infty$ a.s. then N is a counting process without explosion.

1.3.2 Theorem. *A counting process is an adapted process if and only if T_n is a stopping time for all n.*

PROOF: If N is adapted then $\{t < T_n\} = \{N_t < n\} \in \mathscr{F}_t$ for all t and all n. Therefore, for all n, $\{T_n \le t\} \in \mathscr{F}_t$ for all t, so T_n is a stopping time. Conversely, if all the T_n are stopping times then, for all t, $\{N_t \le n\} = \{t \le T_n\} \in \mathscr{F}_t$ for all n. Since N takes only integer values this implies that $N_t \in \mathscr{F}_t$. □

1.3.3 Definition. An adapted process N is called a *Poisson process* if
() N is a counting process.
(i) $N_t - N_s$ is independent of $\mathscr{F}_s$ for all $0 \le s < t < \infty$ (independent increments).

(ii) $N_t - N_s \overset{(d)}{=} N_{t-s}$ for all $0 \le s < t < \infty$ (stationary increments).

Remark. Implicit in this definition is that a Poisson process does not explode. The definition can be modified slightly to allow this possibility, and then it can be proved as a theorem that a Poisson process does not explode, but the details are very technical and relatively unenlightening.

1.3.4 Theorem. *Suppose that N is a Poisson process.*
(i) *N is continuous in probability.*
(ii) *$N_t \overset{(d)}{=}$ Poisson(λt) for some $\lambda \ge 0$, called the intensity or arrival rate of N. In particular, N_t has finite moments of all orders for all t.*
(iii) *$(N_t - \lambda t)_{t\ge 0}$ and $((N_t - \lambda t)^2 - \lambda t)_{t \ge 0}$ are martingales.*
(iv) *If $\mathscr{F}_t^N := \sigma(N_s : s \le t)$ and $\mathbb{F}^N = (\mathscr{F}_t^N)_{t\ge 0}$ then $\mathbb{F}^N$ is right continuous.*

Proof: Let $\alpha(t) := \mathbb{P}[N_t = 0]$ for $t \ge 0$. For all $s < t$,

$$\begin{aligned}
\alpha(t+s) &= \mathbb{P}[N_{t+s} = 0] \\
&= \mathbb{P}[\{N_s = 0\} \cap \{N_{t+s} - N_s = 0\}] && \text{non-decreasing and non-negative} \\
&= \mathbb{P}[N_s = 0]\, \mathbb{P}[N_{t+s} - N_s = 0] && \text{independent increments} \\
&= \mathbb{P}[N_s = 0]\, \mathbb{P}[N_t = 0] && \text{stationary increments} \\
&= \alpha(t)\alpha(s)
\end{aligned}$$

If $t_n \downarrow t$ then $\{N_{t_n} = 0\} \nearrow \{N_t = 0\}$, so α is right continuous and decreasing. It follows that either $\alpha \equiv 0$ or $\alpha(t) = e^{-\lambda t}$ for some $\lambda \ge 0$. By the definition of counting process $N_0 = 0$, so α is cannot be the zero function.

(i) Observe that, given $\varepsilon > 0$ small, for all $s < t$,

$$\begin{aligned}
\mathbb{P}[|N_t - N_s| > \varepsilon] &= \mathbb{P}[|N_{t-s}| > \varepsilon] && \text{stationary increments} \\
&= \mathbb{P}[N_{t-s} > \varepsilon] && N \text{ is non-decreasing} \\
&= 1 - \mathbb{P}[N_{t-s} = 0] && N \text{ is integer valued} \\
&= 1 - e^{-\lambda(t-s)} \\
&\to 0 \text{ as } s \to t.
\end{aligned}$$

Therefore N is left continuous in probability. The proof of continuity from the right is similar.

(ii) First we need to prove that $\lim_{t\to 0} \frac{1}{t}\mathbb{P}[N_t = 1] = \lambda$. Towards this, let $\beta(t) := \mathbb{P}[N_t \ge 2]$ for $t \ge 0$. If we can show that $\lim_{t\to 0} \beta(t)/t = 0$ then we would have

$$\lim_{t\to 0} \frac{\mathbb{P}[N_t = 1]}{t} = \lim_{t \to 0} \frac{1 - \alpha(t) - \beta(t)}{t} = \lambda.$$

It is enough to prove that $\lim_{n\to\infty} n\beta(1/n) = 0$. Divide $[0,1]$ into n equal subintervals and let S_n be the number of subintervals with at least two arrivals. It can be see that $S_n \overset{(d)}{=}$ Binomial$(n, \beta(1/n))$ because of the stationary and independent increments of N. In the definition of counting process the sequence of jump times is strictly increasing, so $\lim_{n\to\infty} S_n = 0$ a.s. Clearly $S_n < N_1$, so $\lim_{n\to\infty} n\beta(1/n) = \lim_{n\to\infty} \mathbb{E}[S_n] = 0$ provided that $\mathbb{E}[N_1] < \infty$. We are going to gloss over this point.

Let $\varphi(t) := \mathbb{E}[\gamma^{N_t}]$ for $0 < \gamma < 1$.

$$\varphi(t+s) = \mathbb{E}[\gamma^{N_{t+s}}] = \mathbb{E}[\gamma^{N_s}\gamma^{N_{t+s}-N_s}] = \mathbb{E}[\gamma^{N_s}]\,\mathbb{E}[\gamma^{N_t}] = \varphi(t)\varphi(s)$$

and φ is right continuous because N has right continuous paths. Therefore $\varphi(t) = e^{t\psi(\gamma)}$ for some function ψ of γ. By definition,

$$\begin{aligned} \varphi(t) &= \sum_{n\geq 0} \gamma^n \,\mathbb{P}[N_t = n] \\ e^{t\psi(\gamma)} &= \alpha(t) + \gamma\,\mathbb{P}[N_t = 1] + \sum_{n\geq 2} \gamma^n\,\mathbb{P}[N_t = n]. \end{aligned}$$

Differentiating, $\psi(\gamma) = \lim_{t\to 0}(\varphi(t) - 1)/t = -\lambda + \gamma\lambda$ by the computations above. Comparing coefficients of γ^n in the power series shows that N_t has a Poisson distribution with rate λt, i.e. $\mathbb{P}[N_t = n] = e^{-\lambda t}(\lambda t)^n/n!$.

(iii) Exercise.

(iv) See the textbook. □

1.3.5 Definition. An adapted process B is called a *Brownian motion* if

(i) $B_t - B_s$ is independent of $\mathscr{F}_s$ for all $0 \leq s < t < \infty$ (independent increments).

(ii) $B_t - B_s \overset{(d)}{=} \text{Normal}(0, t-s)$ for all $0 \leq s < t < \infty$ (stationary, normally distributed, increments).

If $B_0 \equiv 0$ then B is a *standard Brownian motion*.

1.3.6 Theorem. *Let B be a Brownian motion.*

(i) *If $\mathbb{E}[|B_0|] < \infty$ then B is a martingale.*

(ii) *There is a modification of B with continuous paths.*

(iii) *$(B_t^2 - t)_{t\geq 0}$ is a martingale when B is standard BM.*

(iv) *Let Π_n be a refining sequence of partitions of the interval $[a, a+t]$ with $\lim_{n\to\infty} mesh(\Pi_n) = 0$. Then $\Pi_n B := \sum_{t_n\in\Pi_n}(B_{t_{i+1}} - B_{t_i})^2 \to t$ in L^2 and a.s. when B is standard BM.*

(v) *For almost all ω, the function $t \mapsto B_t(\omega)$ is of unbounded variation.*

Remark. To prove (ii) with the additional assumption that $\sum_{n\geq 0} \text{mesh}(\Pi_n) < \infty$, but dropping the assumption that the partitions are refining, you can use the Borel-Cantelli lemma. The proof of (iv) uses the backwards martingale convergence theorem.

1.4 Lévy processes

1.4.1 Definition. An adapted, real-valued process X is called a *Lévy process* if

() $X_0 = 0$

(i) $X_t - X_s$ is independent of $\mathscr{F}_s$ for all $0 \le s < t < \infty$ (independent increments).

(ii) $X_t - X_s \stackrel{(d)}{=} X_{t-s}$ for all $0 \le s < t < \infty$ (stationary increments).

(iii) $(X_t)_{t\ge 0}$ is continuous in probability.

If only (i) and (iii) hold then X is an *additive process*. If $(X_t)_{t\ge 0}$ is non-decreasing then X is called a *subordinator*.

Remark. If the filtration is not specified then we assume that $\mathbb{F} = \mathbb{F}^X$ is the natural (not necessarily complete) filtration of X. In this case X might then be called an *intrinsic Lévy process*.

1.4.2 Example. The Poisson process and Brownian motion are both Lévy processes. Let W be a standard BM and define $T_b := \inf\{t > 0 : W_t \ge b\}$. Then $(T_b)_{b\ge 0}$ is an intrinsic Lévy process and subordinator for the filtration $(\mathscr{F}_{T_b})_{b\ge 0}$. Indeed, if $0 \le a < b < \infty$ then $T_a - T_b$ is independent of $\mathscr{F}_{T_a}$ and distributed as T_{b-a} by the strong Markov property of W. We will see that T is a *stable process* with parameter $\alpha = 1/2$.

1.4.3 Theorem. *Let X be a Lévy process. Then $f_t(z) := \mathbb{E}[e^{izX_t}] = e^{-t\Psi(z)}$ for some continuous function $\Psi = \Psi_X$. Furthermore, $M_t^z := e^{izX_t}/f_t(z)$ is a martingale for all $z \in \mathbb{R}$.*

Proof: Fix $z \in \mathbb{R}$. By the stationarity and independence of the increments,

$$f_t(z) = f_{t-s}(z) f_s(z) \text{ for all } 0 \le s < t < \infty. \tag{1.1}$$

We would like to show that $t \mapsto f_t(z)$ is right-continuous. By the multiplicative property (1.1), it suffices to show $t \mapsto f_t(z)$ is right continuous at zero. Suppose that $t_n \downarrow 0$; we need to show that $f_{t_n}(z) \to 1$. By definition, $|f_{t_n}(z)| \le 1$, so we are done if we can show that every convergent subsequence converges to 1. Suppose without loss of generality that $f_{t_n}(z) \to a \in \mathbb{C}$. Since X is continuous in probability, $e^{izX_{t_n}} \to 1$ in probability. Along a subsequence we have convergence almost surely, i.e. $f_{t_{t_{n_k}}}(z) \to 1$. Therefore $a = 1$ and we are done.

The multiplicative property and right continuity imply that $f_t(z) = e^{-t\Psi(z)}$ for some number $\Psi(z)$. Since $f_t(z)$ is a characteristic function, $z \mapsto f_t(z)$ is continuous (use dominated convergence). Therefore Ψ must be continuous as well. In particular, $f_t(z) \ne 0$ for all t and all z. Let $0 \le s < t < \infty$.

$$\begin{aligned}
\mathbb{E}\left[\frac{e^{izX_t}}{f_t(z)} \middle| \mathscr{F}_s \right] &= \frac{e^{izX_s}}{f_t(z)} \mathbb{E}[e^{iz(X_t - X_s)} | \mathscr{F}_s] && \text{independent increments} \\
&= \frac{e^{izX_s}}{f_t(z)} f_{t-s}(z) && \text{stationary increments}
\end{aligned}$$

$$= \frac{e^{izX_s}}{f_s(z)} \qquad f \text{ is multiplicative.}$$

Therefore M_t^z is a martingale. □

1.4.4 Theorem. *If X is an additive process then X is Markov with transition function $P_{s,t}(x,B) = \mathbb{P}[X_t - X_s \in B - x]$. In particular, X is spatially homogeneous, i.e. $P_{s,t}(x,B) = P_{s,t}(0,B-x)$.*

1.4.5 Corollary. *If X is additive, $\varepsilon > 0$, and $T < \infty$ then $\lim_{u\downarrow 0} \alpha_{\varepsilon,T}(u) = 0$, where $\alpha_{\varepsilon,T}$ is defined in the next theorem.*

PROOF: $P_{s,t}(x, B_\varepsilon(x)^C) = P(|X_t - X_s| \geq \varepsilon) \to 0$ as $s \to t$ uniformly on $[0,T]$ in probability. It is an exercise to show that the continuity in probability of X implies uniform continuity in probability on compact intervals. □

1.4.6 Theorem. *Let X be a Markov process on $\mathbb{R}^d$ with transition function $P_{s,t}$. If $\lim_{u\downarrow 0} \alpha_{\varepsilon,T}(u) = 0$, where*

$$\alpha_{\varepsilon,T}(u) = \sup\{P_{s,t}(x, B_\varepsilon(x)^C) : x \in \mathbb{R}^d, s,t \in [0,T], 0 \leq t-s \leq u\},$$

then X has a càdlàg modification. If furthermore $\lim_{u\downarrow 0} \alpha_{\varepsilon,T}(u)/u = 0$ then X has a continuous modification.

PROOF: See *Lévy processes and infinitely divisible distributions* by Kin-Iti Sato. The important steps are as follows.

Fix $\varepsilon > 0$ and $\omega \in \Omega$. Say that $X(\omega)$ has an ε-oscillation n-times in $M \subseteq [0,\infty)$ if there are $t_0 < t_1 < \cdots < t_n$ all in M such that $|X_{t_j}(\omega) - X_{t_{j-1}}(\omega)| \geq \varepsilon$. X has ε-oscillation infinitely often in M if this holds for all n. Define

$$\Omega_2' := \bigcap_{N=1}^{\infty} \bigcap_{k=1}^{\infty} \{\omega : X(\omega) \text{ does not have } \tfrac{1}{k}\text{-oscillation infinitely often in } [0,N] \cap \mathbb{Q}\}$$

It can be shown that $\Omega_2' \in \mathscr{F}$, and also that

$$\Omega_2' \subseteq \Big\{\omega : \lim_{s\downarrow t, s\in\mathbb{Q}} X_s \text{ exists for all } t \text{ and } \lim_{s\uparrow t, s\in\mathbb{Q}} X_s \text{ exists for all } t > 0\Big\}.$$

The hard part is to show that $\mathbb{P}[\Omega_2'] = 1$. □

Remark.

(i) If X is a *Feller process* then X has a càdlàg modification (see Revuz-Yor).

(ii) From now on we assume that we are working with a càdlàg version of any Lévy process that appears.

(iii) $\mathbb{P}[X_t \neq X_{t-}] = 0$ for any process X that is continuous in probability. Of course, this does not mean that X doesn't jump, it means that X has no fixed times of discontinuity.

1.4.7 Theorem. *Let X be a Lévy process and $\mathcal{N}$ be the collection of $\mathbb{P}$-null sets. If $\mathcal{G}_t := \mathcal{F}_t^X \vee \mathcal{N}$ for all t then $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ is right continuous.*

PROOF: Let $(s_1, \dots, s_n)$ and $(u_1, \dots, u_n)$ be arbitrary vectors in $(\mathbb{R}_+)^n$ and $\mathbb{R}^n$ respectively. We first want to show that

$$\mathbb{E}[e^{i(u_1 X_{s_1} + \cdots + u_n X_{s_n})} | \mathcal{G}_t] = \mathbb{E}[e^{i(u_1 X_{s_1} + \cdots + u_n X_{s_n})} | \mathcal{G}_{t+}].$$

It is clear that it suffices to show this with $s_1, \dots, s_n > t$. We take $n = 2$ for notational simplicity, and assume that $s_2 \geq s_1 > t$.

$$\begin{aligned}
\mathbb{E}[e^{i(u_1 X_{s_1} + u_2 X_{s_2})} | \mathcal{G}_{t+}] &= \lim_{w \downarrow t} \mathbb{E}[e^{i(u_1 X_{s_1} + u_2 X_{s_2})} | \mathcal{G}_w] && \text{exercise} \\
&= \lim_{w \downarrow t} \mathbb{E}[e^{iu_1 X_{s_1}} M_{s_2}^{u_2} | \mathcal{G}_w] f_{s_2}(u_2) \\
&= \lim_{w \downarrow t} \mathbb{E}[e^{iu_1 X_{s_1}} M_{s_1}^{u_2} | \mathcal{G}_w] f_{s_2}(u_2) && \text{tower property} \\
&= \lim_{w \downarrow t} \mathbb{E}[e^{i(u_1 + u_2) X_{s_1}} | \mathcal{G}_w] f_{s_2 - s_1}(u_2) && f \text{ is multiplicative} \\
&= \lim_{w \downarrow t} \frac{e^{i(u_1 + u_2) X_w}}{f_w(u_1 + u_2)} f_{s_1}(u_1 + u_2) f_{s_2 - s_1}(u_2) && M \text{ is a martingale} \\
&= e^{i(u_1 + u_2) X_t} f_{s_1 - t}(u_1 + u_2) f_{s_2 - s_1}(u_2) && X \text{ is càdlàg}
\end{aligned}$$

By the same steps, we obtain

$$\mathbb{E}[e^{i(u_1 X_{s_1} + u_2 X_{s_2})} | \mathcal{G}_t] = e^{i(u_1 + u_2) X_t} f_{s_1 - t}(u_1 + u_2) f_{s_2 - s_1}(u_2).$$

By a monotone class argument, $\mathbb{E}[Z|\mathcal{G}_t] = \mathbb{E}[Z|\mathcal{G}_{t+}]$ for any bounded $Z \in \mathcal{F}_\infty^X$. It follows that $\mathcal{G}_{t+} \setminus \mathcal{G}_t$ consists only of sets from $\mathcal{N}$. Since $\mathcal{N} \subseteq \mathcal{G}_t$, they must be equal. □

1.4.8 Corollary (Blumenthal 0-1 Law). *Let X be a Lévy process. If $A \in \bigcap_{t>0} \mathcal{F}_t^X$ then $\mathbb{P}[A] = 0$ or 1.*

1.4.9 Theorem. *If X is a Lévy process and T is a stopping time then, on $\{T < \infty\}$, $Y_t := X_{T+t} - X_T$ is a Lévy process with respect to $(\mathcal{F}_{t+T})_{t\geq 0}$ and Y has the same finite dimensional distributions as X.*

1.4.10 Corollary. *If X is a Lévy process then X is a strong Markov process.*

1.4.11 Theorem. *If X is a Lévy process with bounded jumps then $\mathbb{E}[|X_t|^n] < \infty$ for all t and all n.*

PROOF: Suppose that $\sup_t |\Delta X_t| \leq C$. Define a sequence of random times T_n recursively as follows. $T_0 \equiv 0$ and

$$T_{n+1} = \begin{cases} \inf\{t > T_n : |X_t - X_{T_n}| \geq C\} & \text{if } T_n < \infty \\ \infty & \text{if } T_n = \infty \end{cases}$$

By definition of the T_n, $\sup_s |X_s^{T_n}| \le 2nC$. The 2 comes from the possibility that X could jump by C just before hitting the stopping level. Since X is càdlàg, the T_n are all stopping times. Further, since X is a strong Markov process on $\{T_n < \infty\}$,

(i) $T_n - T_{n-1}$ is independent of $\mathscr{F}_{T_{n-1}}$ and

(ii) $T_n - T_{n-1} \overset{(d)}{=} T_1$.

Therefore

$$\mathbb{E}[e^{-T_n}] = \mathbb{E}\left[\prod_{k=0}^{n} e^{-(T_k - T_{k-1})}\right] \le (\mathbb{E}[e^{-T_1}])^n =: \alpha^n,$$

where $0 \le \alpha < 1$. The $\le$ comes from the fact that $e^{-\infty} = 0$ and some of the T_n may be ∞. (We interpret $T_k - T_{k-1}$ to be ∞ if both of them are ∞.) By Chebyshev's inequality,

$$\mathbb{P}[|X_t| > 2nC] \le \mathbb{P}[T_n < t] \le \frac{\mathbb{E}[e^{-T_n}]}{e^{-t}} \le \alpha^n e^t.$$

Finally,

$$\begin{aligned}
\mathbb{E}[e^{\beta|X_t|}] &\le 1 + \sum_{n=0}^{\infty} \mathbb{E}[e^{\beta|X_t|} \mathbf{1}_{2nC<|X_t|\le 2(n+1)C}] \\
&\le 1 + \sum_{n=0}^{\infty} e^{2\beta(n+1)C}) \mathbb{P}[|X_t| > 2nC] \\
&\le 1 + e^{2\beta C} e^t \sum_{n=0}^{\infty} (\alpha e^{2\beta C})^n
\end{aligned}$$

Choosing an appropriately small, positive β shows that $|X_t|$ has an exponential moment, so it has polynomial moments of all orders. □

1.5 Lévy measures

Suppose that $\Lambda \in \mathscr{B}(\mathbb{R})$ and $0 \notin \overline{\Lambda}$. Let X be a Lévy process (with càdlàg paths, as always) and define inductively a sequence of random times $T_\Lambda^0 \equiv 0$ and

$$T_\Lambda^{n+1} := \inf\{t > T_\Lambda^n : \Delta X_t \in \Lambda\}.$$

These times have the following properties.

(i) $\{T_\Lambda^n \ge t\} \in \mathscr{F}_{t+} = \mathscr{F}_t$ by the usual conditions and $(T_\Lambda^n)_{n\ge1}$ is an increasing sequence of (possibly ∞-valued) stopping times.

(ii) $T_\Lambda^1 > 0$ a.s. since $X_0 \equiv 0$, X has càdlàg paths, and $0 \notin \overline{\Lambda}$.

(iii) $\lim_{n\to\infty} T_\Lambda^n = \infty$ since X has càdlàg paths and $0 \notin \overline{\Lambda}$ (see Homework 1, problem 10).

Let N^Λ be the number of jumps with size in Λ before time t.

$$N_t^\Lambda := \sum_{0<s\le t} \mathbf{1}_\Lambda(\Delta X_s) = \sum_{n=1}^{\infty} \mathbf{1}_{T_\Lambda^n \ge t}$$

Since X is a strong Markov process, N^Λ has stationary and independent increments. It is also a counting process with no explosions, so it is a Poisson process. By Theorem 1.1.12(i) we can define $\nu(\Lambda) := \mathbb{E}[N_1^\Lambda]$, the intensity of N^Λ. We can extend this definition to the case when $0 \in \overline{\Lambda}$, so long as $0 \notin \Lambda$ by taking $\nu(\Lambda) = \infty$ whenever $\sum_{0<s\le t} \mathbf{1}_\Lambda(\Delta X_s)$ would fail to converge. Define $\nu(\{0\}) = 0$.

1.5.1 Theorem. *For each t and ω the map $\Lambda \mapsto N_t^\Lambda(\omega)$ is a Borel counting measure on $\mathbb{R} \setminus \{0\}$, and $\nu(\Lambda) := \mathbb{E}[N_1^\Lambda]$ is also a Borel measure on $\mathbb{R} \setminus \{0\}$.*

1.5.2 Definition. ν is the *Lévy measure* associated with X.

1.5.3 Example. If X is a Poisson process of rate λ then $\nu = \lambda\delta_1$.

1.5.4 Theorem. *If X is a Lévy process then it can be decomposed as $X = Y + Z$ where Y is a Lévy process and martingale with bounded jumps (so $Y_t \in \bigcap_{p\ge 1} L^p$ for all t) and Z is a Lévy process with paths of finite variation on compact subsets of $[0,\infty)$.*

1.5.5 Theorem. *Let X be a Lévy process with jumps bounded by C. The process $Z_t := X_t - \mathbb{E}[X_t]$ is a martingale that can be decomposed as $Z = Z^c + Z^d$, where Z^c and Z^d are independent Lévy processes. Moreover,*

$$Z_t^d := \int_{|x|\le C} x(N(t,dx) - t\nu(dx)),$$

where $N(t,dx)$ is the measure of Theorem 1.5.1, and the remainder Z^c has continuous paths.

1.5.6 Lemma. *If X is a subordinator with continuous paths then $X_t = ct$ for some constant c.*

PROOF: $X'(0) = \lim_{\varepsilon\downarrow 0} X_\varepsilon/\varepsilon$ exists a.s., and since $X'(0) \in \mathscr{F}_0$, it is a constant a.s. Let $\delta > 0$ and inductively define a sequence of stopping times by $T_0 = 0$ and

$$T_{k+1} := \inf\{t > T_k : X_t - X_{T_k} \ge (c+\delta)t\}.$$

Then, since X is continuous, $X_{T_1} = (c+\delta)T_1$ and $X_t \le (c+\delta)t$ for $t \le T_1$. Additionally, $X_t \le (c+\delta)t$ for $t \le T_k$. Because X is a Lévy process, $(T_{k+1} - T_k)_{k\ge 0}$ are strictly positive i.i.d. random variables. By the strong law of large numbers $T_k \to \infty$ as $k \to \infty$, so $X_t \le (c+\delta)t$ for all t. Therefore $X_t \le ct$ for all t, and the proof that $X_t \ge ct$ for all t is similar. □

1.5.7 Theorem (Lévy-Itô decomposition). *If X is a Lévy process then there are constants σ and α such that*

$$X_t \overset{(d)}{=} \sigma B_t + \alpha t + \int_{|x|<1} x(N(t,dx) - t\nu(dx)) + \sum_{0<s\le t} \Delta X_s \mathbf{1}_{|\Delta X_s|\ge 1}$$

where B is a standard Brownian motion and $\alpha = \mathbb{E}[X_1 - \int_{|x|\ge 1} xN(1,dx)]$.

Remark. X has finite variation on compacts if either
(i) $\sigma = 0$ and $\nu(\mathbb{R}) < \infty$ or
(ii) $\sigma = 0$ and $\nu(\mathbb{R}) = \infty$ but $\int_{|x|<1} |x|\nu(dx) < \infty$.
If $\sigma \neq 0$ or $\int_{|x|<1} |x|\nu(dx) = \infty$ then X has infinite variation on compacts. To prove that the continuous part of X is σB we will need Lévy's characterization of BM, proven later.

1.5.8 Lemma. *Let X be a Lévy process, $\Lambda \subseteq \mathbb{R} \setminus \{0\}$ be Borel with $0 \notin \overline{\Lambda}$, and let f be Borel measurable and finite valued on Λ.*
(i) *The process Y defined by*

$$Y_t := \int_\Lambda f(x) N(t, dx) = \sum_{0<s\le t} f(\Delta X_s) \mathbf{1}_{\Delta X_s \in \Lambda}$$

is a Lévy process. (Show as an exercise that Y is continuous in probability.)
(ii) *Let $M_t := X_t - J_t^\Lambda$ where J^Λ is defined by*

$$J_t^\Lambda := \int_\Lambda x N(t, dx) = \sum_{0<s\le t} f(\Delta X_s) \mathbf{1}_{\Delta X_s \in \Lambda}.$$

Then M is a Lévy process with jumps outside of Λ. In particular, if we take $\Lambda = \{x : |x| \ge 1\}$ then M is a Lévy process with bounded jumps.
(iii) *If $f\mathbf{1}_\Lambda \in L^1(\nu)$ then $\mathbb{E}[\int_\Lambda f(x)N(t,dx)] = t\int_\Lambda f(x)\nu(dx)]$. If $f\mathbf{1}_\Lambda \in L^2(\nu)$ then $\mathbb{E}[(\int_\Lambda f(x)N(t,dx) - t\int_\Lambda f(x)\nu(dx))^2] = t\int_\Lambda f(x)^2\nu(dx)$.*
(iv) *If $\Lambda_1 \cap \Lambda_2 = \varnothing$ then the process $(\sum_{0<s\le t} \Delta X_s \mathbf{1}_{\Lambda_1}(\Delta X_s))_{t\ge 0}$ is independent of the process $(\sum_{0<s\le t} \Delta X_s \mathbf{1}_{\Lambda_2}(\Delta X_s))_{t\ge 0}$.*

Proof (of Theorem 1.5.4): Let $J_t := \sum_{0<s\le t} \Delta X_s \mathbf{1}_{|\Delta X_s|\ge 1}$. J has finite variation on compacts since it is increasing and jumps finitely many times in any bounded interval. By lemma (ii), $X - J$ is a Lévy process with bounded jumps. It can be shown (as an exercise) that $\mathbb{E}[X_t - J_t] = \alpha t$, where $\alpha := \mathbb{E}[X_1 - J_1]$. Then $Y_t := X_t - J_t - \alpha t$ is a martingale and $Z_t := J_t + \alpha t$ is a Lévy process with paths of finite variation on compacts.

$$X_t = \underbrace{X_t - J_t - \alpha t}_{Y_t} + \underbrace{J_t + \alpha t}_{Z_t}.$$

□

Proof (of Theorem 1.5.5): That $Z_t := X_t - \mathbb{E}[X_t]$ is a martingale is an exercise. Suppose without loss of generality that the jumps of X are bounded by 1. Define $\Lambda_k := \{\frac{1}{k+1} < |x| \le \frac{1}{k}\}$ and $M_t^{\Lambda_k} := \int_{\Lambda_k} xN(t,dx) - t\int_{\Lambda_k} x\nu(dx)$. Then the M^{Λ_k} are martingales by lemma (ii) and are in L^2 by lemma (iii) and they are independent by lemma (iv). Let $M^n := \sum_{k=1}^n M^{\Lambda_k}$.

$$\mathrm{Var}(M^n) = \sum_{k=1}^n \mathrm{Var}(M^{\Lambda_k}) = \sum_{k=1}^n t\int_{\Lambda_k} x^2\nu(dx) = t\int_{\frac{1}{n+1}<|x|\le 1} x^2\nu(dx)$$

Prove as an exercise that $Z - M^n$ is independent of M^n (it is not an immediate consequence of the lemma). Therefore $\mathrm{Var}(M^n) = \mathrm{Var}(Z_t) - \mathrm{Var}(Z - M^n) < \infty$ since $\mathrm{Var}(Z_t) < \infty$ because Z has bounded jumps.

Since M^n is a sum of *independent* random variables we can take the limit as $n \to \infty$ in L^2 (fill in the details as an exercise). Let $Z^c := \lim_{n\to\infty}(Z - M^n)$ and $Z^d := \lim_{n\to\infty} M^n$. Then Z^c is independent of Z^d since $Z - M^n$ is independent of M^n for all n. The formula for Z^d is immediate.

Something is missing from the last part of this proof.

To show Z^c is continuous, note first that $Z - M^n$ has jumps bounded by $\frac{1}{n+1}$. By Doob's maximal inequality $\sup_{0<s\le t}(Z_s - M_s^n)$ converges in L^2. Along a subsequence $Z - M^n \to Z^c$ uniformly on compacts so Z^c is continuous. □

1.5.9 Theorem (Lévy-Khintchine formula). *Let X be a Lévy process with characteristic function $\mathbb{E}[e^{izX_t}] = e^{-t\Psi(z)}$. Then the Lévy-Khintchine exponent is*

$$\Psi(z) = \frac{\sigma^2 z^2}{2} - i\alpha z + \int_{\mathbb{R}} (1 - e^{izx} + izx\mathbf{1}_{|x|<1})\nu(dx).$$

Proof (Sketch): In the Lévy-Itô decomposition it can be shown that the parts are all independent, so the characteristic functions multiply. Show as an exercise that the characteristic function of

$$\sum_{0<s\le t} \Delta X_s \mathbf{1}_{|\Delta X_s|\ge 1} \quad \text{is} \quad e^{t\int_{|x|\ge 1}(e^{izx}-1)\nu(dx)}$$

and the characteristic function of

$$\int_{|x|<1} x(N(t,dx) - t\nu(dx)) \quad \text{is} \quad e^{t\int_{|x|<1}(1-e^{izx}+izx)\nu(dx)}.$$

□

1.5.10 Definition. A probability distribution F on $\mathbb{R}$ is an *infinitely divisible distribution* if for all n there are $X_1, \dots, X_n$ i.i.d. such that $X_1 + \cdots + X_n \stackrel{(d)}{=} F$.

1.5.11 Theorem.

(i) *If X is a Lévy process then, for all t, X_t has an infinitely divisible distribution.*

(ii) *Conversely, if F is an infinitely divisible distribution then there is a Lévy process X such that $X_1 \stackrel{(d)}{=} F$.*

(iii) *If F is an infinitely divisible distribution then its characteristic function has the form of the Lévy-Khintchine exponent with $t = 1$, where ν is a measure on $\mathbb{R}$ such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}}(|x|^2 \wedge 1)\nu(dx) < \infty$, and the representation of F in terms of (σ, α, ν) is unique.*

1.5.12 Examples.

(i) If $X \sim \mathrm{Poisson}(\lambda)$ then the characteristic function is $e^{\lambda t(e^{iz}-1)}$, so $\sigma = \alpha = 0$ and $\nu = \lambda\delta_1$.

(ii) The Γ-process is the Lévy process such that X_1 has a Γ-distribution, where $\mathbb{P}[X_1 \in dx] = \frac{b^c}{\Gamma(c)} x^{c-1} e^{bx} \mathbf{1}_{x>0} dx$ for some constants $b, c > 0$. The Lévy-Khintchine exponent is

$$c \int_{\mathbb{R}} (e^{izx} - 1) \frac{e^{-bx}}{x} \mathbf{1}_{x>0} dx.$$

In this case $\nu(dx) = e^{-bx}/x \mathbf{1}_{x>0} dx$, $\sigma = 0$, and $\alpha = \frac{c}{b}(1 - e^{-b}) > 0$. The paths of X are non-decreasing since it has positive drift, no volatility, and positive jumps. Because $\nu((0,1)) = \infty$, X has infinite activity.

(iii) Stable processes are those with Lévy-Khintchine exponent of the form

$$iz\delta + m_1 \int_0^\infty \left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1}{x^{1+\gamma}} dx + m_2 \int_{-\infty}^0 \left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1}{x^{1+\gamma}} dx$$

where $0 < \gamma < 2$. When $m_1 = m_2$ one can take $\nu(dx) = \frac{1}{|x|^{1+\gamma}} dx$. In this case there are $Y_1, Y_2, \dots$ i.i.d. and constants a_n, and b_n such that

$$\frac{1}{a_n} \sum_{i=1}^\infty Y_i - b_n \xrightarrow{\text{(w)}} X_1$$

and $(\beta^{-\frac{1}{\gamma}} X_{\beta t})_{t \geq 0} \overset{\text{(d)}}{=} (X_t)_{t \geq 0}$ for all $\beta > 0$. If $1 \leq \alpha < 2$ then $\int |x| \nu(dx) = \infty$ so the process would have infinite variation on compacts.

(iv) For the hitting time process $(T_b)_{b \geq 0}$ of standard Brownian motion B, we have

$$\mathbb{P}[T_b \in dt] = \frac{b}{\sqrt{2\pi t^3}} e^{-\frac{b^2}{2t}} \mathbf{1}_{t>0} dt$$

so the Lévy-Khintchine exponent is

$$\frac{b}{\sqrt{2\pi}} \int_0^\infty (e^{izx} - 1) x^{-\frac{3}{2}} dx.$$

It follows that T is a stable process with $\gamma = \frac{1}{2}$. Indeed,

$$\begin{aligned} \frac{1}{\beta^2} T_{\beta b} &= \frac{1}{\beta^2} \inf\{t : B_t = \beta b\} \\ &= \inf\{t/\beta^2 : B_t/\beta = b\} \\ &= \inf\{t : B_{\beta^2 t}/\beta = b\} \overset{\text{(d)}}{=} T_b \end{aligned}$$

since $(B_{\beta^2 t}/\beta)_{t \geq 0} \overset{\text{(d)}}{=} B$. Also note that T is non-decreasing.

1.6 Localization

1.6.1 Definition.

(i) If C is a collection of processes then the *localized class*, C_{loc}, is the collection of processes X such that there exists a sequence $(T_n)_{n\geq 1}$ of stopping times such that $T_n \uparrow \infty$ a.s. and $X^{T_n} \in C$ for all n. The sequence $(T_n)_{n\geq 1}$ is called a *localizing sequence* for X relative to C.

(ii) If C is the collection of adapted, càdlàg, uniformly integrable martingales then an element of the localized class is called a *local martingale*.

(iii) A stopping time T reduces X relative to C if $X^T \in C$.

(iv) C is *stable under stopping* if $X^T \in C$ for all $X \in C$ and all stopping times T.

1.6.2 Theorem. *Suppose that C is a collection of processes that is stable under stopping.*

(i) *C_{loc} is stable under stopping and $(C_{loc})_{loc} = C_{loc}$. In particular, a local local martingale is a local martingale.*

(ii) *If T reduces X relative to C and $S \leq T$ a.s. then S reduces X.*

(iii) *If M and N are local martingales then $M + N$ is a local martingale. If S and T reduce M then $S \vee T$ reduces M.*

(iv) *$(C \cap C')_{loc} = C_{loc} \cap C'_{loc}$.*

PROOF (OF (I)): Let $X \in C_{\text{loc}}$ and T be a stopping time. If $(T_n)_{n\geq 1}$ is a localizing sequence for X then $(X^T)^{T_n} = (X^{T_n})^T \in C$ since C is stable under stopping. Therefore $(T_n)_{n\geq 1}$ is a localizing sequence for X^T and it is seen that C_{loc} is stable under stopping.

Now let $X \in (C_{\text{loc}})_{\text{loc}}$ and (T_n) be a localizing sequence for X relative to C_{loc}, so that $X^{T_n} \in C_{\text{loc}}$ for all n. Then for each n there are $(T_{n,p})_{p\geq 1}$ such that $T_{n,p} \uparrow \infty$ a.s. as $p \to \infty$ and $(X^{T_n})^{T_{n,p}} \in C$ for all p. For all n there is p_n such that

$$\mathbb{P}[T_{n,p_n} < T_n \wedge n] \leq \frac{1}{2^n}$$

since a.s. convergence implies convergence in probability. Define a new sequence of stopping times $S_n := T_n \wedge \bigwedge_{m\geq n} T_{m,p_m}$. Then, for all n, $S_n \leq S_{n+1}$ and

$$\mathbb{P}[S_n < T_n \wedge n] \leq \sum_{m=n}^{\infty} \frac{1}{2^m} = \frac{1}{2^{n-1}}.$$

By the Borel-Cantelli lemma, $\mathbb{P}[S_n < T_n \wedge n \text{ i.o.}] = 0$, which implies that $S_n \uparrow \infty$ a.s. Finally, $X^{S_n} = ((X^{T_n})^{T_{n,p_n}})^{S_n} \in C$ since C is stable under stopping, so $(S_n)_{n\geq 1}$ is a localizing sequence for X relative to C. □

1.6.3 Corollary.

(i) *If X is a locally square integrable martingale then X is a local martingale (but the converse does not hold).*

(ii) *If M is a process and there are stopping times $(T_n)_{n\geq 1}$ such that $T_n \uparrow \infty$ a.s. and M^{T_n} is a martingale for all n then M is a local martingale.*

We will see that it is important to be able to decide when a local martingale a "true" martingale. To this end, let $X_t^* := \sup_{s\le t}|X_s|$ and $X^* = \sup_{s\ge 0}|X_s|$.

1.6.4 Theorem. *Let X be a local martingale.*
(i) *If $\mathbb{E}[X_t^*] < \infty$ for all t then X is a martingale.*
(ii) *If $\mathbb{E}[X^*] < \infty$ then X is a u.i. martingale.*

PROOF: Let $(T_n)_{n\ge 1}$ be a localizing sequence for X. For any $s \le t$,

$$\mathbb{E}[X_{T_n\wedge t}|\mathscr{F}_s] = X_{T_n\wedge t}$$

by the optional stopping theorem. Since $X_{T_n\wedge t}$ is dominated by the integrable random variable X_t^*, we may apply the (conditional) dominated convergence theorem to both sides. Therefore X is a martingale. If $\mathbb{E}[X^*] < \infty$ then the family $(X_t)_{t\ge 0}$ is dominated by the integrable random variable X^*, so it is u.i. □

1.6.5 Corollary.
(i) *If X is a bounded local martingale then it is a u.i. martingale.*
(ii) *If X is a local martingale and a Lévy process then X is a martingale.*
(iii) *If $(X_n)_{n\ge 1}$ is a discrete time local martingale and $\mathbb{E}|X_n| < \infty$ for all n then X is a martingale.*

1.6.6 Example. Suppose $(A_n)_{n\ge 1}$ is a measurable partition of Ω with $\mathbb{P}[A_n] = 2^{-n}$ for all n. Further suppose $(Z_n)_{n\ge 1}$ is a sequence of random variables independent of the A_n and such that $\mathbb{P}[Z_n = 2^n] = \mathbb{P}[Z_n = -2^{-n}] = 1/2$. Define

$$\mathscr{F}_t := \begin{cases} \sigma(A_n : n \ge 1) & \text{for } 0 \le t < 1 \\ \sigma(A_n, Z_n : n \ge 1) & \text{for } t \ge 1 \end{cases}$$

completed with respect to $\mathbb{P}$. Define $Y_n := \sum_{1\le k\le n} Z_k \mathbf{1}_{A_n}$ and $T_n := \infty \mathbf{1}_{\bigcup_{1\le k\le n} A_k}$. Let X_t be zero for $t < 1$ and Y_∞ for $t \ge 1$. Then $(T_n)_{n\ge 1}$ is a localizing sequence for X and $X_t^{T_n}$ is zero for $t < 1$ and Y_n for $t \ge 1$. Since Y_n is bounded for each n, X^{T_n} is a u.i. martingale, but X is not a martingale because $X_1 = Y_\infty \notin L^1$.

1.7 Integration with respect to processes of finite variation

1.7.1 Definition. We say that a process A is increasing or of finite variation if it has that property path-by-path for almost all ω.

1.7.2 Definition. Let A be of finite variation. The *total variation process* is

$$|A|_t := \sup_{n\ge 1} \sum_{k=1}^{2^n} |A_{t(k+1)2^{-n}} - A_{tk2^{-n}}|.$$

Note that $|A|_t < \infty$ a.s., $|A|$ is increasing, and if A is adapted then so is $|A|$.

1.7.3 Definition. Let A be of finite variation and $F(\omega, s)$ be $(\mathscr{F} \otimes \mathscr{B})$-measurable and bounded.

$$(F \cdot A)_t(\omega) := \int_0^t F(s, \omega) dA_s(\omega),$$

the path-by-path Lebesgue-Stieltjes integral (which is well-defined and finite for almost all ω).

The integral process is also of finite variation and if A is (right) continuous then the integral process is (right) continuous. If F is continuous path-by-path then then integral may be taken to be the Riemann-Stieltjes integral.

1.7.4 Theorem. *Let A and C be adapted, increasing processes such that $C - A$ is increasing. There exists H jointly measurable and adapted such that $A = H \cdot C$.*

PROOF (SKETCH): The correct process is

$$H(t, \omega) := \lim_{r \uparrow 1, r \in \mathbb{Q}_+} \frac{A(\omega, t) - A(\omega, rt)}{C(\omega, t) - C(\omega, rt)}.$$ □

1.7.5 Corollary. *If A is of finite variation then there is a jointly measurable process H with $-1 \le H \le 1$ such that $A = H \cdot |A|$ and $|A| = H \cdot A$.*

PROOF (SKETCH): Let $A^+ := (|A| + A)/2$ and $A^- := (|A| - A)/2$. These are both increasing processes, so by Theorem 1.7.4 there are H^+ and H^- such that $A^+ = H^+ \cdot |A|$ and $A^- = H^- \cdot |A|$. Take $H := H^+ - H^-$. □

1.7.6 Theorem.

(i) *Let A be of finite variation, H have continuous paths, and $(\Pi_n)_{n \ge 1}$ be a sequence of partitions of $[0, t]$ with mesh size converging to zero. For each $t_i \in \Pi_n$ let s_i denote any number in the interval $[t_i, t_{i+1}]$. Then*

$$\lim_{n \to \infty} \sum_{\Pi_n} H_{s_i}(A_{t_{i+1}} - A_{t_i}) = \int_0^t H_s dA_s.$$

(ii) *(Change of variable formula.) Let A be of finite variation with continuous paths and $f \in C^1(\mathbb{R})$. Then $(f(A_t))_{t \ge 0}$ is of finite variation and*

$$f(A_t) = f(A_0) + \int_0^t f'(A_s) dA_s.$$

(iii) *Let A be of finite variation with continuous paths g be a continuous function. Then then $\int_0^t g(A_s) dA_s = \int_0^{A_t} g(s) ds$.*

PROOF (OF (II)): Let $(\Pi_n)_{n\geq 1}$ be a sequence of partitions of $[0,t]$ as in (i) and consider

$$\begin{aligned} f(A_t)-f(A_0) &= \sum_{\Pi_n} f(A_{t_{i+1}})-f(A_{t_i}) \\ &= \sum_{\Pi_n} f'(A_{s_i})(A_{t_{i+1}}-A_{t_i}) && \text{for some } s_i \text{ by the MVT} \\ &\to \int_0^t f'(A_s)dA_s && \text{by part (i)} \end{aligned}$$

□

1.8 Naïve stochastic integration is impossible

1.8.1 Theorem (Banach-Steinhaus). *Let X be a Banach space and let Y be a normed linear space. Suppose $(T_\alpha)_{\alpha\in I}$ is a collection of continuous linear operators from X to Y such that $\sup_{\alpha\in I}\|T_\alpha x\|_Y<\infty$ for all $x\in X$. Then it is the case that $\sup_{\alpha\in I}\|T_\alpha\|_{\mathcal{L}(X,Y)}<\infty$.*

1.8.2 Theorem. *Let $x:[0,1]\to\mathbb{R}$ be right continuous and let $(\Pi_n)_{n\geq 1}$ be a sequence of partitions of $[0,1]$ with mesh size converging to zero. If*

$$S_n(h):=\sum_{\Pi_n} h(t_i)(x_{t_{i+1}}-x_{t_i})$$

converges for all $h\in C([0,1])$ then x is of finite variation.

PROOF: Note that S_n is a linear operator from $C([0,1])$ to $\mathbb{R}$, and

$$\|S_n\|\leq\sum_{\Pi_n}|x_{t_{i+1}}-x_{t_i}|.$$

This upper bound is achieved for each n by the piecewise linear function h_n with the property that $h(t_i)=\text{sign}(x_{t_{i+1}}-x_{t_i})$. If $S_n(h)$ exists for every h then the Banach-Steinhaus theorem implies that $\sup_{n\geq 1}\|S_n\|<\infty$, i.e. that the total variation of x over $[0,1]$ is finite. □

Remark. It should be possible to prove that if X is a right continuous process and $(\Pi_n)_{n\geq 1}$ is a sequence of partitions of $[0,t]$ mesh size converging to zero then if $\sum_{t_i,t_{i+1}\in\Pi_n} H_{t_i}(X_{t_i}-X_{t_{i-1}})$ converges in probability for all $H\in\mathscr{F}\otimes\mathscr{B}$ with continuous paths then X is of finite variation a.s.

Chapter 2

Semimartingales and stochastic integration

2.1 Introduction

2.1.1 Definition. A process H is a *simple predictable process* if it has the form

$$H_t = H_0 \mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{n-1} H_i \mathbf{1}_{(T_i, T_{i+1}]}(t)$$

where $0 = T_0 \leq T_1 \leq \cdots \leq T_n < \infty$ are stopping times and $H_i \in L^\infty(\mathscr{F}_{T_i})$ for all $i = 0, \ldots, n$. The collection of all simple predictable processes will be denoted $\mathcal{S}$. A norm on $\mathcal{S}$ is $\|H\|_u := \sup_{s \geq 0} \|H_s\|_\infty$. When we endow $\mathcal{S}$ with the topology induced by $\|\cdot\|_u$ we write $\mathcal{S}_u$.

Remark. $\mathcal{S} \subseteq \mathscr{P}$.

2.1.2 Definition. $L^0 := L^0(\Omega, \mathscr{F}, \mathbb{P}) := \{X \in \mathscr{F} : X \text{ is finite-valued a.s.}\}$ with the topology induced by convergence in probability under $\mathbb{P}$.

Remark. L^0 is not locally convex if $\mathbb{P}$ is non-atomic. The topological dual of L^0 in this case is $\{0\}$.

2.1.3 Definition. Let $H \in \mathcal{S}$ have its canonical representation. For an arbitrary stochastic process X, the *stochastic integral* of H with respect to X is

$$I_X(H) := H_0 X_0 + \sum_{i=1}^{n-1} H_i (X_{T_{i+1}} - X_{T_i}).$$

A càdlàg adapted process X is called a *total semimartingale* if the map $I_X : \mathcal{S}_u \to L^0$ is continuous. X is a *semimartingale* if X^t is a total semimartingale for all $t \geq 0$.

2.2 Stability properties of semimartingales

2.2.1 Theorem.

(i) {*semimartingales*} *and* {*total semimartingales*} *are vector spaces.*

(ii) *If* $\mathbb{Q} \ll \mathbb{P}$ *then any (total) semimartingale under* $\mathbb{P}$ *is also a (total) semimartingale under* $\mathbb{Q}$*.*

PROOF:

(i) Sums and scalar multiples of continuous operations are continuous.

(ii) Convergence in $\mathbb{P}$-probability implies convergence in $\mathbb{Q}$-probability. □

2.2.2 Theorem (Stricker). *Let* X *be a semimartingale with respect to the filtration* $\mathbb{F}$ *and* $\mathscr{G}$ *is a sub-filtration of* $\mathbb{F}$ *(i.e.* $\mathscr{G}_t \subseteq \mathscr{F}_t$ *for all* t*). If* X *is* $\mathbb{G}$ *adapted then* X *is a semimartingale with respect to* $\mathbb{G}$*.*

PROOF: $\mathbb{S}(\mathbb{G}) \subseteq \mathbb{S}(\mathbb{F})$ and the restriction of a continuous operation is continuous. □

2.2.3 Theorem. *Let* $\mathcal{A} = (A_\alpha)_{\alpha \in I}$ *be a collection of pairwise disjoint events from* $\mathscr{F}$ *and let* $\mathscr{H}_t = \sigma(\mathscr{F}_t \vee A)$*. If* X *is a semimartingale with respect to* $\mathbb{F}$ *then* X *is a semimartingale with respect to* $\mathbb{H}$*.*

Remark. This theorem is called *Jacod's countable expansion theorem* in the textbook.

PROOF: Note first that $\mathscr{H}_t = \sigma(\mathscr{F}_t \vee (\mathcal{A} \setminus \{A \in \mathcal{A} : \mathbb{P}[A] = 0\})$ since $\mathbb{F}$ is complete, and second that $\{A \in \mathcal{A} : \mathbb{P}[A] \geq 1/n\}$ is finite because the events are pairwise disjoint. It follows that we may assume without loss of generality that $\mathcal{A}$ is a countable collection $(A_n)_{n\geq 1}$. If $C := \bigcup_{n=1}^{\infty} A_n$ then we may assume without loss of generality that $C^C \in \mathcal{A}$, i.e. that $\mathcal{A}$ is a (countable) partition of Ω into events with positive probability.

For each n define $\mathbb{Q}_n := \mathbb{P}[\cdot | A_n]$, so that $\mathbb{Q}_n \ll \mathbb{P}$. By Theorem 2.2.1 X is an $(\mathbb{F}, \mathbb{Q}_n)$-semimartingale. Let $\mathbb{J}^n$ be the filtration $\mathbb{F}$ completed with respect to $\mathbb{Q}_n$.

Claim. X is a $(\mathbb{J}^n, \mathbb{Q}_n)$-semimartingale.

Indeed, if $H \in \mathbb{S}(\mathbb{J}^n, \mathbb{Q}_n)$ has its canonical representation then the T_i are $\mathbb{J}^n$-stopping times and the H_i are in $L^\infty(\Omega, \mathscr{J}^n_{T_i}, \mathbb{Q}_n)$.

Fact. Any $\mathbb{J}^n$-stopping time is $\mathbb{Q}_n$-a.s. equal to a $\mathbb{F}$-stopping time.

Denote the corresponding $\mathbb{F}$-stopping times by T_i'. Then any $H_i \in \mathscr{J}^n_{T_i}$ is $\mathbb{Q}_n$-a.s. equal to an r.v. in $\mathscr{F}_{T_i'}$ (technical exercise). This shows that H is $\mathbb{Q}_n$-a.s. equal to a process in $\mathbb{S}(\mathbb{F}, \mathbb{Q}_n)$, which proves the claim.

Note finally that $\mathbb{Q}_n[A_m] \in \{0, 1\}$ for all $A_m \in A$, so $\mathbb{F} \subseteq \mathbb{H} \subseteq \mathbb{J}^n$ for all n. By Stricker's theorem, X is an $(\mathbb{H}, \mathbb{Q}_n)$-semimartingale for all n. It can be shown that, since $d\mathbb{P} = \sum_{n=1}^{\infty} \mathbb{P}[A_n] d\mathbb{Q}_n$, X is an $(\mathbb{H}, \mathbb{P})$-semimartingale. □

2.2.4 Theorem. *Let X be a process and $(T_n)_{n\geq 1}$ be a sequence of random times such that X^{T_n} is a semimartingale for all n and $T_n \uparrow \infty$ a.s. Then X is a semimartingale.*

PROOF: Fix $t > 0$. There is n_0 such that $\mathbb{P}[T_n \leq t] < \varepsilon$ for all $n \geq n_0$. Suppose that $H^k \in \mathcal{S}$ are such that $\|H^k\|_\infty \to 0$. Let k_0 be such that $\mathbb{P}[|I_{X^{T_{n_0}}}(H^k)| > \alpha] < \varepsilon$ for all $k \geq k_0$. Then

$$\mathbb{P}[|I_{X^t}(H^k)| > \alpha] = \mathbb{P}[|I_{X^t}(H^k)| > \alpha; T_{n_0} \leq t] + \mathbb{P}[|I_{(X^{T_{n_0}})^t}(H^k)| > \alpha; T_{n_0} > t] < 2\varepsilon$$

for all $k \geq k_0$. Therefore $I_{X^t}(H^k) \to 0$ in probability, so X is a semimartingale. □

2.3 Elementary examples of semimartingales

2.3.1 Theorem. *Every adapted càdlàg process with paths of finite (total) variation is a (total) semimartingale.*

PROOF: Suppose that the total variation of X is finite. It is easy to see that, for all $H \in \mathcal{S}$, $|I_X(H)| \leq \|H\|_\infty \int_0^\infty d|X|_s$. □

2.3.2 Theorem. *Every càdlàg L^2-martingale is a total semimartingale.*

PROOF: Without loss of generality $X_0 = 0$. Let $H \in \mathcal{S}$ have the canonical representation.

$$\begin{aligned}
\mathbb{E}[(I_X(H))^2] &= \mathbb{E}\left[\left(\sum_{i=1}^{n-1} H_i(X_{T_{i+1}} - X_{T_i})\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n-1} H_i^2(X_{T_{i+1}} - X_{T_i})^2\right] && X \text{ is a martingale} \\
&\leq \|H\|_u^2\, \mathbb{E}\left[\sum_{i=1}^{n-1}(X_{T_{i+1}} - X_{T_i})^2\right] \\
&= \|H\|_u^2\, \mathbb{E}[X_{T_n}^2] && X \text{ is a martingale} \\
&\leq \|H\|_u^2\, \mathbb{E}[X_\infty^2] && \text{by Jensen's inequality}
\end{aligned}$$

Therefore if $H^k \to H$ in $\mathcal{S}_u$ then $I_X(H^k) \to I_X(H)$ in L^2, so also in probability. □

2.3.3 Corollary.
(i) *A càdlàg locally square integrable martingale is a semimartingale.*
(ii) *A càdlàg local martingale with bounded jumps is a semimartingale.*
(iii) *A local martingale with continuous paths is a semimartingale.*
(iv) *Brownian motion is a semimartingale.*
(v) *If $X = M + A$ is càdlàg, where M is a locally square integrable martingale and A is locally of finite variation, then X is a semimartingale. Such an X is called a decomposable process*
(vi) *Any Lévy process is a semimartingale by the Lévy-Itô decomposition.*

2.4 The stochastic integral as a process

2.4.1 Definition.
(i) $\mathbb{D} := \{$adapted, càdlàg processes$\}$
(ii) $\mathbb{L} := \{$adapted, càglàd processes$\}$
(iii) If $\mathcal{C}$ is a collection of processes then $b\mathcal{C}$ is the collection of bounded processes from $\mathcal{C}$.

2.4.2 Definition. For processes H^n and H we say that $H^n \to H$ *uniformly on compacts in probability* (or *u.c.p.*) if

$$\lim_{n\to\infty} \mathbb{P}\Big[\sup_{0\le s\le t} |H^n_s - H_s| \ge \varepsilon\Big] = 0$$

for all $\varepsilon > 0$ and all t. Equivalently, if $(H^n - H)^*_t \xrightarrow{(p)} 0$ for all t.

Remark. If we define a metric

$$d(X,Y) := \sum_{n=1}^{\infty} \frac{1}{2^n}\,\mathbb{E}[(X-Y)^*_n \wedge 1]$$

then u.c.p. convergence is compatible with this metric. We will denote by $\mathbb{D}_{\text{ucp}}$, $\mathbb{L}_{\text{ucp}}$, and $\mathcal{S}_{\text{ucp}}$ the topological spaces $\mathbb{D}$, $\mathbb{L}$, and $\mathcal{S}$ endowed with the u.c.p. topology. The u.c.p. topology is weaker than the uniform topology. It is important to note that $\mathbb{D}_{\text{ucp}}$ and $\mathbb{L}_{\text{ucp}}$ are complete metric spaces.

2.4.3 Theorem. *$\mathcal{S}$ is dense in $\mathbb{L}_{ucp}$.*

PROOF: Let $Y \in \mathbb{L}$ and define $R_n := \inf\{t : |Y_t| > n\}$. Then $Y^{R_n} \to Y$ u.c.p. (exercise) and Y^{R_n} is bounded by n because Y is left continuous. Thus $b\mathbb{L}$ is dense in $\mathbb{L}_{\text{ucp}}$, so it suffices to show that $\mathcal{S}$ is dense in $b\mathbb{L}_{\text{ucp}}$.

Assume that $Y \in b\mathbb{L}$ and define $Z_t := \lim_{u\downarrow t} Y_n$ for all $t \ge 0$, so that $Z \in \mathbb{D}$. For each $\varepsilon > 0$ define a sequence of stopping times $T^\varepsilon_0 := 0$ and

$$T^\varepsilon_{n+1} := \inf\{t > T^\varepsilon_n : |Z_t - Z_{T^\varepsilon_n}| > \varepsilon\}.$$

We have seen that the T^ε_n are stopping times because they are hitting times for the càdlàg process Z. Also because $Z \in \mathbb{D}$, $T^\varepsilon_n \uparrow \infty$ a.s. (this would not necessarily happen it T^ε were defined in the same way but with Y). Let

$$Z^\varepsilon := \sum_{n\ge 0} Z_{T^\varepsilon_n} \mathbf{1}_{[T^\varepsilon_n, T^\varepsilon_{n+1})}.$$

Then Z^ε is bounded and $Z^\varepsilon \to Z$ uniformly on compacts as $\varepsilon \to 0$. Let

$$U^\varepsilon := Y_0 \mathbf{1}_{\{0\}} + \sum_{n\ge 0} Z_{T^\varepsilon_n} \mathbf{1}_{(T^\varepsilon_n, T^\varepsilon_{n+1}]}.$$

Then $U^\varepsilon \to Y_0 \mathbf{1}_{\{0\}} + Z_-$ uniformly on compacts as $\varepsilon \to 0$. If we define

$$Y^{n,\varepsilon} := Y_0 \mathbf{1}_{\{0\}} + \sum_{k\geq 1}^{n} Z_{T_k^\varepsilon} \mathbf{1}_{(T_k^\varepsilon, T_{k+1}^\varepsilon]}$$

then $Y^{n,\varepsilon} \to Y$ u.c.p. as $\varepsilon \to 0$ and $n \to \infty$. □

2.4.4 Definition. Let $H \in \mathcal{S}$ and X be a càdlàg process. The *stochastic integral process* of H with respect to X is $J_X(H) := H_0 X_0 + \sum_{i=1}^n H_i(X^{T_{i+1}} - X^{T_i})$ when $H = H_0 \mathbf{1}_{\{0\}} + \sum_{i=1}^{n-1} H_i \mathbf{1}_{(T_i, T_{i+1}]}$.

Notation. We write $\int H_s dX_s := H \cdot X := J_X(H)$ and $\int_0^t H_s dX_s := I_{X^t}(H) = (J_X(H))_t$ and $\int_0^\infty H_s dX_s := I_X(H)$.

2.4.5 Theorem. *If X is a semimartingale then $J_X : \mathcal{S}_{ucp} \to \mathbb{D}_{ucp}$ is continuous.*

PROOF: We begin by proving the weaker statement, that $J_X : \mathcal{S}_u \to \mathbb{D}_{\text{ucp}}$ is continuous. Suppose that $\|H^k\|_\infty \to 0$. Let $\delta > 0$ and define a sequence of stopping times $T^k := \inf\{t : |(H^k \cdot X)_t| \geq \delta\}$. Then $H^k \mathbf{1}_{[0,T^k]} \in \mathcal{S}$ and $\|H^k \mathbf{1}_{[0,T^k]}\|_\infty \to 0$ because this already happens for $(H^k)_{k\geq 1}$. Let $t \geq 0$ be given.

$$\mathbb{P}[(H^k \cdot X)_t^* \geq \delta] \leq \mathbb{P}[(H^k \cdot X)_{t\wedge T^k}^* \geq \delta] = \mathbb{P}[I_{X^t}(H^k \mathbf{1}_{[0,T^k]}) \geq \delta] \to 0$$

since X is a semimartingale. Therefore $J_X(H^k) \to 0$ u.c.p.

Assume $H^k \to 0$ u.c.p. Let $\varepsilon > 0$, $t > 0$, and $\delta > 0$ be given. There is η such that $\|H\|_\infty < \eta$ implies $\mathbb{P}[J_X(H)_t^* \geq \delta] < \varepsilon$. Let $R^k := \inf\{s : |H_s^k| > \eta\}$. Then $R^k \to \infty$ a.s., $\tilde{H}^k := H^k \mathbf{1}_{[0,R^k]} \to 0$ u.c.p., and $\|\tilde{H}^k\| \leq \eta$.

$$\begin{aligned}\mathbb{P}[(H^k \cdot X)_t^* \geq \delta] &= \mathbb{P}[(H^k \cdot X)_t^* \geq \delta; R^k < t] + \mathbb{P}[(H^k \cdot X)_t^* \geq \delta; R^k \geq t] \\ &\leq \mathbb{P}[R^k < t] + \mathbb{P}[(\tilde{H}^k \cdot X)_t^* \geq \delta] < 2\varepsilon\end{aligned}$$

for k large enough. □

If $H \in \mathbb{L}$ and $(H^n)_{n\geq 1} \subseteq \mathcal{S}$ are such that $H^n \to H$ u.c.p. then $(H_n)_{n\geq 1}$ has the Cauchy property for the u.c.p. metric. Since J_X is continuous, $(J_X(H^n))_{n\geq 1}$ also has the Cauchy property. Since $\mathbb{D}_{\text{ucp}}$ is complete there is $J_X(H) \in \mathbb{D}$ such that $J_X(H^n) \to J_X(H)$ u.c.p. We say that $J_X(H)$ is the *stochastic integral* of H with respect to X.

2.4.6 Example. Let B be standard Brownian motion and let $(\Pi_n)_{n\geq 1}$ be a sequence of partitions of $[0,t]$ with mesh size converging to zero. Define

$$B^n := \sum_{\Pi_n} B_{t_i} \mathbf{1}_{(t_i, t_{i+1}]},$$

so that $B^n \to B$ u.c.p. (check this as an exercise).

$$\begin{aligned}(J_B(B^n))_t &= \sum_{\Pi_n} B_{t_i}(B_{t_{i+1}} - B_{t_i}) \\ &= \frac{1}{2}\sum_{\Pi_n}(B_{t_{i+1}}^2 - B_{t_i}^2) - \frac{1}{2}\sum_{\Pi_n}(B_{t_{i+1}} - B_{t_i})^2 \\ &= \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{\Pi_n}(B_{t_{i+1}} - B_{t_i})^2 \\ &\xrightarrow{(p)} \frac{1}{2}B_t^2 - \frac{1}{2}t\end{aligned}$$

Hence $\int_0^t B_s dB_s = \frac{1}{2}B_s - \frac{1}{2}t$.

2.5 Properties of the stochastic integral

2.5.1 Theorem. *Let $H \in \mathbb{L}$ and X be a semimartingale.*

(i) *If T is a stopping time then $(H \cdot X)^T = (H\mathbf{1}_{[0,T]} \cdot X) = H \cdot (X^T)$.*
(ii) *$\Delta(H \cdot X)$ is indistinguishable from $H \cdot (\Delta X)$.*
(iii) *If $\mathbb{Q} \ll \mathbb{P}$ then $H \cdot_{\mathbb{Q}} X$ is indistinguishable from $H \cdot_{\mathbb{P}} X$.*
(iv) *If $\mathbb{Q} = \sum_{k\geq 1} \lambda_k \mathbb{P}_k$ with $\lambda_k \geq 0$ and $\sum_{k\geq 1} \lambda_k = 1$ then $H \cdot_{\mathbb{Q}} X$ is indistinguishable from $H \cdot_{\mathbb{P}_k} X$ for any k for which $\lambda_k > 0$.*
(v) *If X is both a $\mathbb{P}$ and $\mathbb{Q}$ semimartingale then there is $H \cdot X$ that is a version of both $H \cdot_{\mathbb{P}} X$ and $H \cdot_{\mathbb{Q}} X$.*
(vi) *If $\mathbb{G}$ is another filtration and $H \in \mathbb{L}(\mathbb{G}) \cap \mathbb{L}(\mathbb{F})$ then $H \cdot_{\mathbb{G}} X = H \cdot_{\mathbb{F}} X$.*
(vii) *If X has paths of finite variation on compacts then $H \cdot X$ is indistinguishable from the path-by-path Lebesgue-Stieltjes integral.*
(viii) *$Y := H \cdot X$ is a semimartingale and $K \cdot Y = (KH) \cdot X$ for all $K \in \mathbb{L}$.*

Proof (of (viii)): It can be shown using limiting arguments that $K \cdot (H \cdot X) = (KH) \cdot X$ when $H, K \in \mathbb{S}$. To prove that $Y = H \cdot X$ is a semimartingale when $H \in \mathbb{L}$ there are two steps. First we show that if $K \in \mathbb{S}$ then $K \cdot Y = (KH) \cdot X$ (and this makes sense even without knowing that Y is a semimartingale). Fix $t > 0$ and choose $H^n \in \mathbb{S}$ with $H^n \to H$ u.c.p. We know $H^n \cdot X \to H \cdot X$ u.c.p. because X is a semimartingale. Therefore there is a subsequence such that $(Y^{n_k} - Y)_t^* \to 0$ a.s., where $Y^{n_k} := H^{n_k} \cdot X$. Because of the a.s. convergence, $(K \cdot Y)_t = \lim_{k\to\infty}(K \cdot Y^{n_k})_t$ a.s. Finally, $K \cdot Y^{n_k} = (KH^{n_k}) \cdot X$ since $H^n, K \in \mathbb{S}$, so

$$K \cdot Y = \text{u.c.p.-}\lim_{k\to\infty}(KH^{n_k}) \cdot X = (KH) \cdot X$$

since $KH^{n_k} \to KH$ u.c.p. and X is a semimartingale.

For the second step, proving that Y is a semimartingale, assume that $G_n \to G$ in $\mathbb{S}_u$. We must prove that $(G^n \cdot Y)_t \to (G \cdot Y)_t$ in probability for all t. We have $G^n H \to GH$ in $\mathbb{L}_{\text{ucp}}$, so

$$\lim_{n\to\infty} G^n \cdot Y = \lim_{n\to\infty}(G^n H) \cdot X = (GH) \cdot X = G \cdot Y.$$

Since this convergence is u.c.p. this proves the result. □

2.5.2 Theorem. *If X is a locally square integrable martingale and $H \in \mathbb{L}$ then $H \cdot X$ is a locally square integrable martingale.*

PROOF: It suffices to prove the result assuming X is a square integrable martingale, since if $(T_n)_{n\geq 1}$ is a localizing sequence for X then $(H\cdot X)^{T_n} = H\cdot X^{T_n}$, so we would have that $H \cdot X$ is a local locally square integrable martingale. Conclude with Theorem 1.6.2, since the collection of locally square integrable martingale is stable under stopping. Furthermore, we may assume that H is bounded because left continuous processes are locally bounded (a localizing sequence is $R_n := \inf\{t : |H_t| > n\}$) and that $X_0 = 0$.

Construct, as in Theorem 2.4.3, $H^n \in \mathcal{S}$ such that $H^n \to H$ u.c.p. By construction the H^n are bounded by $\|H\|_\infty$. Let $t > 0$ be given. It is easy to see that $H^n \cdot X$ is a martingale and

$$\begin{aligned}\mathbb{E}[(H^n \cdot X)^2] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} H_i^n (X_t^{T_{i+1}} - X_t^{T_i})\right)^2\right] \\ &\leq \|H\|_\infty^2 \,\mathbb{E}[X_t^2] < \infty.\end{aligned}$$

(The detail are as in the proof of Theorem 2.3.2.) It follows that $((H^n \cdot X)_t)_{n\geq 1}$ is u.i. Since we already know $(H^n \cdot X)_t \to (H \cdot X)_t$ in probability, the convergence is actually in L^1. This allows us to prove that $H \cdot X$ is a martingale, and by Fatou's lemma

$$\mathbb{E}[(H \cdot X)_t^2] \leq \liminf_n \mathbb{E}[(H^n \cdot X)_t^2] \leq \|H\|_\infty^2 \,\mathbb{E}[X_t^2] < \infty$$

so $H \cdot X$ is a square integrable martingale. □

2.5.3 Theorem. *Let X be a semimartingale. Suppose that $(\Sigma_n)_{n\geq 1}$ is a sequence of random partitions, $\Sigma_n : 0 = T_0^n \leq T_1^n \leq \cdots \leq T_{k_n}^n$ where the T_k^n are stopping times such that*

(i) $\lim_{n\to\infty} T_{k_n}^n = \infty$ *a.s. and*

(ii) $\sup_k |T_{k+1}^n - T_k^n| \to 0$ *a.s.*

Then if $Y \in \mathbb{L}$ (or $\mathbb{D}$) then $\sum_k Y_{T_k^n}(X^{T_{k+1}^n} - X^{T_k^n}) \to \int Y_- dX$.

Remark. The hard part of the proof of this theorem is that the approximations to Y do not necessarily converge u.c.p. to Y (if they did then the theorem would be a trivial consequence of the definition of semimartingale).

2.5.4 Example. Let $M_t = N_t - \lambda t$ be a compensated Poisson process (a martingale) and let $H = \mathbf{1}_{[0,T_1)}$ (a bounded process in $\mathbb{D}$), where T_1 is the first jump time of the Poisson process. Then $\int_0^t H_s dM_s = -\lambda(t \wedge T_1)$, which is not a local martingale. Examples of this kind are part of the reason that we want our integrands from $\mathbb{L}$.

2.6 The quadratic variation of a semimartingale

2.6.1 Definition. Let X and Y be semimartingales. The *quadratic variation* of X and Y is $[X,Y] = ([X,Y]_t)_{t\geq 0}$, defined by

$$[X,Y] := XY - \int X_- dY - \int Y_- dX.$$

Write $[X] := [X,X]$. The *polarization identity* is sometimes useful.

$$[X,Y] = \frac{1}{2}([X+Y] - [X] - [Y])$$

2.6.2 Theorem.
(i) $[X]_0 = X_0^2$ *and* $\Delta[X] = (\Delta X)^2$.
(ii) *If* $(\Sigma_n)_{n\geq 1}$ *are as in Theorem 2.5.3 then*

$$X_0^2 + \sum_k (X^{T_{k+1}^n} - X^{T_k^n})^2 \xrightarrow{u.c.p.} [X].$$

(iii) $[X]$ *is càdlàg, adapted, and increasing.*
(iv) $[X,Y]$ *is of finite variation,* $[X,Y]_0 = X_0 Y_0$, $\Delta[X,Y] = \Delta X \Delta Y$, *and*

$$X_0 Y_0 + \sum_k (X^{T_{k+1}^n} - X^{T_k^n})(Y^{T_{k+1}^n} - Y^{T_k^n}) \xrightarrow{u.c.p.} [X,Y].$$

(v) *If T is any stopping time then*

$$[X^T, Y] = [X, Y^T] = [X^T, Y^T] = [X,Y]^T.$$

PROOF:
(i) Recall that $X_{0-} := 0$, so $\int_0^0 X_{s-} dX_s = 0$ and $[X]_0 = (X_0)^2$. For any $t > 0$,

$$\begin{aligned}(\Delta X_t)^2 = (X_t - X_{t-})^2 &= X_t^2 + X_{t-}^2 - 2X_t X_{t-} \\ &= X_t^2 - X_{t-}^2 - 2X_{t-}\Delta X_t \\ &= (\Delta X^2)_t - 2\Delta(X_- \cdot X)_t \\ &= \Delta(X^2 - 2(X_- \cdot X))_t = \Delta[X]_t\end{aligned}$$

(ii) Fix n and let $R_n := \sup_k T_k^n$. Then $(X^2)^{R_n} \to X^2$ u.c.p., so apply Theorem 2.5.3 and the summation-by-parts trick.
(iii) Let's see why $[X]$ is increasing. Fix $s < t$ rational. If we can prove that $[X]_s \leq [X]_t$ a.s. then we are done since $[X]$ is càdlàg. Use partitions $(\Sigma_n)_{n\geq 1}$ in Theorem 2.5.3 that include s and t. Excluding terms from the sum makes it smaller since all of the summands are squares (hence non-negative), so $[X]_s \leq [X]_t$ a.s. □

2.6.3 Corollary.
(i) *If X is a continuous semimartingale of finite variation then $[X]$ is constant, i.e. $[X]_t = X_0^2$ for all t.*
(ii) *$[X,Y]$ and XY are semimartingales, so {semimartingales} form an algebra.*

2.6.4 Theorem (Kunita-Watanabe identity). *Let X and Y be semimartingales and $H,K : \Omega \times [0,\infty) \to \mathbb{R}$ be $\mathscr{F} \otimes \mathscr{B}$-measurable.*

$$\left(\int_0^\infty |H_s||K_s| d|[X,Y]|_s \right)^2 \le \int_0^\infty H_s^2 d[X]_s \int_0^\infty K_s^2 d[Y]_s.$$

PROOF: Use the following lemma.

2.6.5 Lemma. *Let $\alpha,\beta,\gamma : [0,\infty) \to \mathbb{R}$ be such that $\alpha(0) = \beta(0) = \gamma(0) = 0$, α is of finite variation, β and γ are increasing, and for all $s \le t$,*

$$(\alpha(t) - \alpha(s))^2 \le (\beta(t) - \beta(s))(\gamma(t) - \gamma(s)).$$

Then for all measurable functions f and g and all $s \le t$,

$$\left(\int_s^t |f_u||g_u| d|\alpha|_u \right)^2 \le \int_s^t f_u^2 d\beta_u \int_s^t g_u^2 d\gamma_u.$$

The lemma can be proved with the following version of the monotone class theorem.

2.6.6 Theorem (Monotone class theorem II). *Let $\mathcal{C}$ be an algebra of bounded real valued functions. Let $\mathcal{H}$ be a collection of bounded real valued functions closed under monotone and uniform convergence. If $\mathcal{C} \subseteq \mathcal{H}$ then $b\sigma(\mathcal{C}) \subseteq \mathcal{H}$.*

Both are left as exercises. □

2.6.7 Definition. Let X be a semimartingale. The *continuous part of quadratic variation*, $[X]^c$, is defined by

$$[X]_t = [X]_t^c + X_0^2 + \sum_{0<s\le t} (\Delta X_s)^2.$$

X is said to be a *quadratic pure jump semimartingale* if $[X]^c \equiv 0$.

2.6.8 Examples.
(i) If N is a Poisson process then $[N] = N$, so it is a quadratic pure jump semimartingale.
(ii) Any Lévy process with no Gaussian part is a quadratic pure jump semimartingale. Note that $[t] \equiv 0$.

2.6.9 Theorem. *If X is an adapted, càdlàg process of finite variation then X is a quadratic pure jump semimartingale.*

PROOF: The Lebesgue-Stieltjes integration-by-parts formula gives us

$$X^2 = \int X_- dX + \int X dX,$$

and by definition of $[X]$ we have

$$X^2 = 2\int X_- dX + [X].$$

Noting that

$$\int X dX = \int (X_- + \Delta X) dX = \int X_- dX + \sum (\Delta X)^2$$

we obtain $[X] = \sum(\Delta X)^2$, so X is a quadratic pure jump semimartingale. □

2.6.10 Theorem. *If X is a locally square integrable martingale that is not constant everywhere then $[X]$ is not constant everywhere. Moreover, $X^2 - [X]$ is a local martingale and if $[X] \equiv 0$ then $X \equiv 0$.*

Remark. This theorem holds when X is a local martingale, but it is harder to prove.

PROOF: We have that $X^2 - [X] = 2\int X_- dX$ is a locally square integrable martingale by Theorem 2.5.2. Assume that $[X]_t = 0$ for all $t \geq 0$. Then X^2 is a locally square integrable martingale, so suppose that $(T_n)_{n\geq 1}$ is a localizing sequence. Then $\mathbb{E}[X^2_{t\wedge T_n}] = X_0^2 = 0$ for all n and all t. Thus $X^2_{t\wedge T_n} = 0$ a.s. for all n and all t, so $X \equiv 0$. If X_0 is a constant other than zero then the proof applies to the locally square integrable martingale $X - X_0$. □

2.6.11 Corollary. *Let X be a locally square integrable martingale and $S \leq T$ be stopping times.*

(i) *If $[X]$ is constant on $[S, T] \cap [0, \infty)$ then X is constant there too.*
(ii) *If X has continuous paths of finite variation on (S, T) then X is constant on $[S, T] \cap [0, \infty)$.*

PROOF: Consider $M := X^T - X^S$ and apply Theorems 2.6.2 and 2.6.10. □

2.6.12 Corollary.

(i) *If X and Y are locally square integrable martingales then $[X, Y]$ is the unique adapted, càdlàg process of finite variation such that*
   a) $XY - [X, Y]$ is a local martingale, and
   b) $\Delta[X, Y] = \Delta X \Delta Y$ and $[X, Y]_0 = X_0 Y_0$.
(ii) *If X and Y are locally square integrable martingales with no common jumps and such that XY is a local martingale then $[X, Y] \equiv X_0 Y_0$.*

(iii) *If X is a continuous square integrable martingale and Y is a square integrable martingale of finite variation then $[X,Y] \equiv X_0 Y_0$ and XY is martingale.*

PROOF:
(i) By definition, $XY - [X,Y] = \int X_- dY + \int Y_- dX$, which is a local martingale. The second condition is Theorem 2.6.2(vi). Conversely, if A satisfied the two conditions then, by Theorem 2.6.2(vi), $A - [X,Y]$ is a continuous semimartingale of finite variation null at zero. By Corollary 2.2.3, $A = [X,Y]$. □

Fill in more of this proof.

2.6.13 Assumption. *$[M]$ always exists for a local martingale M.*

2.6.14 Corollary. *Let M be a local martingale. Then M is a square integrable martingale if and only if $\mathbb{E}[[M]_t] < \infty$ for all $t \geq 0$. We have $\mathbb{E}[M_t^2] = \mathbb{E}[[M]_t]$ when $\mathbb{E}[[M]_t] < \infty$.*

PROOF: Suppose M is a square integrable martingale. $M^2 - [M]$ is a local martingale null at zero by Corollary 2.6.12. Let $(T_n)_{n\geq 1}$ be a localizing sequence. For all $t \geq 0$,

$$\mathbb{E}[[M]_t] = \lim_{n\to\infty} \mathbb{E}[[M]_{t\wedge T_n}] = \lim_{n\to\infty} \mathbb{E}[M_{t\wedge T_n}^2] \leq \mathbb{E}[M_t^2] < \infty$$

The first equality is the monotone convergence theorem, the second is because $(M^2 - [M])^{T_n \wedge t}$ is a martingale, and the inequality is the "limsup" version of Fatou's lemma. For the converse, define stopping times $T_n := \inf\{t : |M_t| > n\} \wedge n$.

$$\sup_{0\leq s\leq t} |M_s^{T_n}| \leq n + |\Delta M_{T_n}| \leq n + \sqrt{[M]_n}$$

□

Hence $\sup_{s\leq t} |M_s^{T_n}| \in L^2 \subseteq L^1$ for all n and all t. By Doob's maximal inequality...

2.6.15 Example (Inverse Bessel process). Let B be a three dimensional Brownian motion that starts at $x \neq 0$. It can be shown that $M_t := 1/\|B_t\|$ is a local martingale. Furthermore, $\mathbb{E}[M_t^2] < \infty$ for all t and $\lim_{t\to\infty} \mathbb{E}[M_t^2] = 0$. This implies that M is not a martingale (because if it were then M^2 would be a submartingale, which contradicts that $\lim_{t\to\infty} \mathbb{E}[M_t^2] = 0$). Moreover, $\mathbb{E}[[M]_t] = \infty$ for all $t > 0$. It can be shown that M satisfies the SDE $dM_t = -M_t^2 dB_t$.

2.6.16 Corollary. *Let X be a continuous local martingale. Then X and $[X]$ have the same intervals of constancy a.s. (i.e. pathwise, cf. Corollary 2.6.11).*

2.6.17 Theorem. *Let X be a quadratic pure jump semimartingale. For any semimartingale Y,*

$$[X,Y] = X_0 Y_0 + \sum_{0<s\leq t} \Delta X_s \Delta Y_s.$$

PROOF (SKETCH): We know by the Kunita-Watanabe inequality that $d[X,Y]_s \ll d[X,X]_s$ a.s. Since $[X,X]^c \equiv 0$, this implies that $[X,Y]^c \equiv 0$, so $[X,Y]$ is equal to the sum of its jumps. □

2.6.18 Theorem. *Let $H,K \in \mathbb{L}$ and X and Y be semimartingales.*

(i) $[H\cdot X, K\cdot Y] = (HK)\cdot[X,Y]$.

(ii) *If H is càdlàg and $(\sigma_n)_{n\geq 1}$ is a sequence of random partitions tending to the identity then*

$$\sum H_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})(Y^{T_{i+1}^n} - Y^{T_i^n}) \xrightarrow{u.c.p.} \int H_{s-}d[X,Y]_s.$$

Remark. Part (i) is important for computations, but part (ii) is important theoretically. We will prove Itô's formula using part (ii).

PROOF (OF (I)): It is enough to prove that $[H\cdot X, Y] = H\cdot[X,Y]$ since we already have associativity. Suppose that $H^n \in \mathcal{S}$ are such that $H^n \to H$ u.c.p. Define $Z^n := H^n \cdot X$ and $Z := H\cdot X$, so that $Z^n \to Z$ u.c.p. Then

$$\begin{aligned}
[Z^n, Y] &= YZ^n - \int Y_- dZ^n - \int Z_-^n dY \\
&= YZ^n - \int (YH^n)_- dX - \int Z_-^n dY \\
&\to YZ - \int (YH)_- dX - \int Z_- dY \\
&= [Z,Y]
\end{aligned}$$

For all n, since H^n is simple predictable,

$$[Z^n, Y] = [H^n\cdot X, Y] = H^n[X,Y]$$

(check this). Finally, $[X,Y]$ is a semimartingale so $H^n[X,Y] \to H\cdot[X,Y]$. □

2.7 Itô's formula

We have seen that if A is continuous and of finite variation and if $f \in C^1$ then

$$f(A_t) = f(A_0) + \int_0^t f'(A_s)dA_s.$$

We also saw that if B is a standard Brownian motion then

$$B_t^2 = 2\int_0^t B_s dB_s + t.$$

Take $f(x)=x^2$ so that we can write

$$f(B_t)-f(B_0)=\int_0^t f'(B_s)dB_s+\frac{1}{2}\int_0^t f''(B_s)d[B]_s.$$

This expression does not coincide with the one for continuous processes of finite variation.

Notation. We write

$$\int_{(0,t]} H_s dX_s := \int_{0+}^t H_s dX_s := \int_0^t H_s dX_s - H_0X_0 = \int_0^t H_s \mathbf{1}_{(0,\infty)}(s)dX_s.$$

2.7.1 Theorem (Itô's formula). *If X is a semimartingale and $f \in C^2(\mathbb{R})$ then $(f(X_t))_{t\geq 0}$ is a semimartingale and*

$$f(X_t)-f(X_0)=\int_{0+}^t f'(X_{s-})dX_s+\frac{1}{2}\int_{0+}^t f''(X_{s-})d[X]^c_s + \sum_{0<s\leq t}(f(X_s)-f(X_{s-})-f'(X_{s-})\Delta X_s).$$

In particular, if X has continuous paths then

$$f(X_t)-f(X_0)=\int_{0+}^t f'(X_s)dX_s+\frac{1}{2}\int_0^t f''(X_s)d[X]_s.$$

Additionally, if X is of finite variation then

$$f(X_t)-f(X_0)=\int_{0+}^t f'(X_{s-})dX_s+\sum_{0<s\leq t}(f(X_s)-f(X_{s-})-f'(X_{s-})\Delta X_s),$$

and this last formula holds even if f is only C^1.

Proof: First assume that X is bounded by a constant K. By definition the quadratic variation of a semimartingale is finite valued, so $\sum_{0<s\leq t}(\Delta X_s)^2 \leq [X]_t < \infty$ a.s. Since f'' is continuous on the interval $[-K,K]$ it is bounded on that interval, so $\sum_{0<s\leq t} f''(X_{s-})(\Delta X_s)^2 < \infty$ a.s. and we may write Itô's formula in the following equivalent form.

$$f(X_t)-f(X_0)=\int_{0+}^t f'(X_{s-})dX_s+\frac{1}{2}\int_{0+}^t f''(X_{s-})d[X]_s + \sum_{0<s\leq t}(f(X_s)-f(X_{s-})-f'(X_{s-})\Delta X_s - f''(X_{s-})(\Delta X_s)^2)$$

Fix $t>0$. Taylor's theorem states that, for $x,y\in[-K,K]$,

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

where the remainder satisfies $|R(x,y)| \le r(|x-y|)(x-y)^2$, where r is an increasing function such that $\lim_{\varepsilon \downarrow 0} r(\varepsilon) = 0$.

Since $\sum_{0<s\le t}(\Delta X_s)^2$ is a.s. a convergent series, we can partition the jumps of X in $[0,t]$ into a finite set $\mathcal{A} := \mathcal{A}(\varepsilon, t)$ and the remainder $\mathcal{B} := \mathcal{B}(\varepsilon, t)$, with the property that $\sum_{s\in\mathcal{B}}(\Delta X_s)^2 \le \varepsilon^2$. $\mathcal{A}$ are the times of "large" jumps and $\mathcal{B}$ are the times of "small" jumps.

Suppose that $(\Sigma^n)_{n\ge 1}$ is a sequence of random dyadic partitions of $[0,t]$ with mesh size converging to zero, say $\Sigma^n : (0 = T_0^n \le T_1^n \le \cdots \le T_{k_n}^n = t)$. For each n, partition the indices $\{0, 1, \dots, k_n - 1\}$ into the follow sets

$$\begin{aligned}
\mathcal{L}^n &:= \{i : (T_i^n, T_{i+1}^n] \cap \mathcal{A} \neq \varnothing\} \\
\mathcal{S}^n &:= \{i \notin \mathcal{L}^n : (T_i^n, T_{i+1}^n] \cap \mathcal{B} \neq \varnothing\} \\
\mathcal{R}^n &:= \{0, 1, \dots, k_n - 1\} \setminus (\mathcal{L}^n \cup \mathcal{S}^n)
\end{aligned}$$

Then $\mathcal{L}^n$ is the collection of indices i for which the i^{th} interval in the partition contains a large jump of X, $\mathcal{S}^n$ is the collection of indices i for which the i^{th} interval contains a small jump but no large jump, and $\mathcal{S}^n$ is the collection of indices i for which X is continuous on the i^{th} interval. Write

$$\begin{aligned}
f(X_t) - f(X_0) &= \sum_{\Sigma^n}(f(X_{T_{i+1}^n}) - f(X_{T_i^n})) \\
&= \sum_{\mathcal{L}^n} + \sum_{\mathcal{S}^n} + \sum_{\mathcal{R}^n}(f(X_{T_{i+1}^n}) - f(X_{T_i^n})) \\
&=: S_L^n + S_S^n + S_R^n
\end{aligned}$$

By Taylor's theorem we can write

$$\begin{aligned}
S_S^n + S_R^n = \sum_{\Sigma^n} & f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n}) + \frac{1}{2}f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2 \\
&- \sum_{\mathcal{L}^n} f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n}) + \frac{1}{2}f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2 \\
&\quad + \sum_{\mathcal{S}^n \cup \mathcal{R}^n} R(X_{T_i^n}, X_{T_{i+1}^n}).
\end{aligned}$$

As $n \to \infty$, $\sum_{\mathcal{S}^n}(X_{T_{i+1}^n} - X_{T_i^n})^2 \to \sum_{s\in\mathcal{B}}(\Delta X_s)^2$. The remainder term in the above expression, call it R^n, satisfies

$$\begin{aligned}
|R^n| &= \left| \sum_{\mathcal{S}^n \cup \mathcal{R}^n} R(X_{T_i^n}, X_{T_{i+1}^n}) \right| \\
&\le \sum_{\mathcal{S}^n \cup \mathcal{R}^n} r(|X_{T_i^n} - X_{T_{i+1}^n}|)(X_{T_{i+1}^n} - X_{T_i^n})^2 \\
&= \sum_{\mathcal{S}^n} + \sum_{\mathcal{R}^n} r(|X_{T_i^n} - X_{T_{i+1}^n}|)(X_{T_{i+1}^n} - X_{T_i^n})^2 \\
&\le r(2K) \sum_{\mathcal{S}^n}(X_{T_{i+1}^n} - X_{T_i^n})^2 + \sup_{\mathcal{R}^n} r(|X_{T_{i+1}^n} - X_{T_i^n}|) \sum_{\mathcal{R}^n}(X_{T_{i+1}^n} - X_{T_i^n})^2
\end{aligned}$$

By the properties of r we can pick a subsequence $(\Sigma^{n_m})_{m\geq 1}$ such that
For each $m \geq 1$ there is n_m such that

$$\left| \sum_{\sigma^{n_m}\cap\mathcal{B}\neq 0} (X_{T^n_{i+1}} - X^n_{T^n_i})^2 - \sum_{x\in\mathcal{B}} (\Delta X_s)^2 \right| \leq \left(1 + \frac{1}{m}\right) \varepsilon^2.$$

You can find a sequence of partitions $\tilde{\sigma}^m$, constructed by refining σ^{n_m} outside the subintervals that don't contain jumps, such that

$$\sum_{\tilde{\sigma}^m\cap\mathcal{A}=\varnothing} R(X_{\tilde{T}^m_i}, X_{\tilde{T}^m_{i+1}}) \to 0$$

as $m \to \infty$ and $\varepsilon \to 0$. Work from then on with the refinements $\tilde{\sigma}^m$.

$$\limsup_{m\to\infty} |R^{n_m}| \leq (1 + r(2K))\varepsilon^2.$$

Hence

$$\begin{aligned} f(X_t) - f(X_0) &= S^{n_m}_L + S^{n_m}_S + S^{n_m}_R \\ &= S^{n_m}_L + \sum_{\Sigma^{n_m}} f'(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i}) + \frac{1}{2} f''(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i})^2 \\ &\qquad - \sum_{\mathcal{L}^n} f'(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i}) + \frac{1}{2} f''(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i})^2 + R^{n_m} \end{aligned}$$

As $m \to \infty$,

$$\sum_{\Sigma^{n_m}} f'(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i}) \to \int_{0+}^t f'(X_{s-}) dX_s \qquad \text{by Theorem 2.5.3}$$

$$\sum_{\Sigma^{n_m}} \frac{1}{2} f''(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i})^2 \to \frac{1}{2} \int_{0+}^t f''(X_{s-}) d[X]_s \qquad \text{by Theorem 2.6.18}$$

and similarly,

$$\begin{aligned} S^n_L = \sum_{\mathcal{L}^n} f(X_{T^n_{i+1}}) - f(X_{T^n_i}) &\to \sum_{s\in\mathcal{A}} f(X_s) - f(X_{s-}) \\ \sum_{\mathcal{L}^n} f'(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i}) &\to \sum_{s\in\mathcal{A}} f'(X_{s-})\Delta X_s \\ \sum_{\mathcal{L}^n} \frac{1}{2} f''(X_{T^n_i})(X_{T^n_{i+1}} - X_{T^n_i})^2 &\to \sum_{s\in\mathcal{A}} \frac{1}{2} f''(X_{s-})(\Delta X_s)^2 \end{aligned}$$

Finally, as $\varepsilon \to 0$, $\mathcal{A}(\varepsilon, t)$ exhausts the jumps of X. Since

$$f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s + \frac{1}{2} f''(X_{s-})(\Delta X_s)^2$$

is dominated by a constant times the sum of the jumps squared...

If X is not necessarily bounded then let $V_k := \inf\{t : |X_t| > k\}$ and $Y^k := X\mathbf{1}_{[0,V_k)}$. Then Y^k is a semimartingale, as it is a product of two semimartingales, bounded by k. (Note that Y^k is not the same as X^{V_k-}, as the former jumps to zero at time V_k.) By the previous part,

$$f(Y_t^k) - f(Y_0^k) = \int_{0+}^t f'(Y_{s-}^k)dY_s^k + \frac{1}{2}\int_{0+}^t f''(Y_{s-}^k)d[Y^k]_s^c + \sum_{0<s\le t}(f(Y_s^k) - f(Y_{s-}^k) - f'(Y_{s-}^k)\Delta Y_s^k).$$

Fix ω and t and k such that $V_k(\omega) > t$. One can rewrite the above in terms of Itô's formula for X. (More perspicaciously, for fixed t, for each ω there is k large enough so that $Y^k(\omega,t) = X(\omega,t)$ and the differentials are equal too.) □

2.7.2 Theorem.
(i) *If $X = (X^1,\dots,X^n)$ is an n-dimensional semimartingale and $f : \mathbb{R}^n \to \mathbb{R}$ has continuous second order partial derivatives then $f(X) = (f(X_t^1,\dots,X_t^n))_{t\ge0}$ is a semimartingale and*

$$f(X_t)-f(X_0) = \sum_{i=1}^n \int_{0+}^t \frac{\partial f}{\partial x_i}(X_{s-})dX_s^i + \frac{1}{2}\sum_{i=1}^n \int_{0+}^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-})d[X^i,X^j]_s^c + \sum_{0<s\le t}\left(f(X_s) - f(X_{s-}) - \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X_{s-})\Delta X_s^i\right)$$

(ii) *If X and Y are semimartingales, $Z := X + iY$, and f is analytic then*

$$f(Z_t) - f(Z_0) = \int_{0+}^t f'(Z_{s-})dZ_s + \frac{1}{2}\int_{0+}^t f''(Z_{s-})d[Z]_s^c + \sum_{0<s\le t}(f(Z_s) - f(Z_{s-}) - f'(Z_{s-})\Delta Z_s).$$

where "$dZ = dX + idY$" and $[Z] = [X] - [Y] + 2i[X,Y]$.

2.7.3 Definition. The *Fisk-Stratonovich integral* of is defined for semimartingales X and Y by $(Y_- \circ X)_t := \int_0^t Y_{s-} \circ dX_s := \int_0^t Y_{s-}dX_s + \frac{1}{2}[X,Y]_t^c$.

2.7.4 Theorem.
(i) *If X is a semimartingale and $f \in C^3(\mathbb{R})$*

$$f(X_t) - f(X_0) = \int_{0+}^t f'(X_{s-}) \circ dX_s + \sum_{0<s\le t}(f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s).$$

(ii) *If X and Y are semimartingales and at least one is continuous then*

$$XY - X_0Y_0 = X_- \circ Y + Y_- \circ X.$$

2.8 Applications of Itô's formula

2.8.1 Theorem (Stochastic exponential).
Let X be a semimartingale with $X_0 = 0$. Then there exists a unique semimartingale Z such that $Z_t = 1 + \int_0^t Z_{s-} dX_s$. Moreover,

$$Z_t = \mathscr{E}(X)_t := \exp(X_t - \tfrac{1}{2}[X]_t^c) \prod_{s \le t} (1 + \Delta X_s) \exp(-\Delta X_s).$$

PROOF: Let's see first that $\mathscr{E}(X)$ is a semimartingale. We know that $X - \frac{1}{2}[X]^c$ is a semimartingale, so $\exp(X - \frac{1}{2}[X]^c)$ is a semimartingale by Itô's theorem. We need to show that $\prod_{s \le t}(1 + \Delta X_s)e^{-\Delta X_s}$ is well-defined and is a semimartingale. Towards this, let $U_t = \Delta X_s \mathbf{1}_{|\Delta X_s| \le \frac{1}{2}}$. The set $\{s : |\Delta X_s| > \frac{1}{2}\}$ is finite a.s., so it is enough to prove that $\prod_{s \le t}(1 + U_s)e^{-U_s}$ is a semimartingale. For $|x| \le \frac{1}{2}$, it can be shown that $|\log(1 + x) - x| \le x^2$, whence

$$\begin{aligned}
\left| \log\left(\prod_{s \le t} (1 + U_s) e^{-U_s} \right) \right| &\le \sum_{s \le t} |\log(1 + U_s) - U_s| \\
&\le \sum_{s \le t} U_s^2 \le \sum_{s \le t} (\Delta X_s)^2 \le [X]_t < \infty \text{ a.s.}
\end{aligned}$$

Hence the infinite product converges, and moreover, $\prod_{s \le t}(1 + \Delta X_s)e^{-\Delta X_s}$ is of finite variation because the series of its log-jumps is absolutely summable. In particular it is a quadratic pure jump semimartingale.

Now we prove that Z satisfied the integral equation. Let $F(x, y) := ye^x$, $K_t := X - \frac{1}{2}[X]^c$, and $S_t := \prod_{s \le t}(1 + \Delta X_s)e^{-\Delta X_s}$. Then F is C^2 and $Z_t = F(K_t, S_t)$. By Itô's formula,

$$\begin{aligned}
Z_t - 1 &= \int_{0+}^t Z_{s-} dK_s + \int_{0+}^t e^{K_{s-}} dS_s + \frac{1}{2} \int_{0+} Z_{s-} d[K]_s^c \\
&\qquad + \sum_{0 < s \le t} (Z_s - Z_{s-} - Z_{s-}(\Delta K_s) - e^{K_{s-}} \Delta S_s) \\
&= \int_0^t Z_{s-} dK_s + \sum_{0 < s \le t} e^{K_{s-}} (\Delta S_s) + \frac{1}{2} \int_0 Z_{s-} d[X]_s^c \\
&\qquad + \sum_{0 < s \le t} (Z_s - Z_{s-} - Z_{s-}(\Delta X_s) - e^{K_{s-}} \Delta S_s) \\
&= \int_0^t Z_{s-} dK_s + \frac{1}{2} \int_0 Z_{s-} d[X]_s^c = \int_0^t Z_{s-} dX_s
\end{aligned}$$

The only tricky point is to note that $Z_s = Z_{s-}(1 + \Delta X_s)$. □

Remark.
(i) If X is of finite variation then so is $\mathscr{E}(X)$.

(ii) If X is a local martingale then so is $\mathscr{E}(X)$, but at this point we can only prove it if X is a locally square integrable martingale.

(iii) Let $T = \inf\{t : \Delta X_t = -1\}$.
  a) $\mathscr{E}(X) \neq 0$ on $[0, T)$
  b) $\mathscr{E}(X)_- \neq 0$ on $[0, T]$
  c) $\mathscr{E}(X) = 0$ on $[T, \infty)$

(iv) If X is continuous and $X_0 = 0$ then $\mathscr{E}(X) = \exp(X - \frac{1}{2}[X])$.

(v) If $\lambda \in \mathbb{R}$ and B is standard Brownian motion then $\mathscr{E}(\lambda B) = \exp(\lambda B_t - \frac{1}{2}\lambda^2 t)$ is a continuous martingale.

2.8.2 Theorem (Yor's formula). *If X and Y are semimartingales with $X_0 = Y_0 = 0$ then $\mathscr{E}(X)\mathscr{E}(Y) = \mathscr{E}(X + Y + [X, Y])$.*

PROOF: Integration-by-parts. □

Remark. If X is continuous then $(\mathscr{E}(X))^{-1} = \mathscr{E}(-X + [X])$.

2.8.3 Theorem. *Let $X = (X^1, \dots, X^n)$ be an n-dimensional local martingale with values in D, an open subset of $\mathbb{R}^n$. Assume $[X^i, X^j] = A\delta_{i,j}$, where A is some increasing process. If $u : D \to \mathbb{R}$ is harmonic (resp. sub-harmonic) then $u(X)$ is a local martingale (resp. sub-martingale).*

2.8.4 Theorem (Lévy's characterization of BM).
A stochastic process X is a Brownian motion if and only if X is a continuous local martingale with $[X]_t = t$ a.s. for all t.

PROOF: Only one direction requires proof, so assume X is a continuous local martingale and $[X]_t = t$ for all t. It suffices to show that $\mathbb{E}[e^{iu(X_t - X_s)} \mid \mathscr{F}_s] = e^{-\frac{1}{2}u^2(t-s)}$ for all $0 \le s < t < \infty$ (exercise). Fix $u \in \mathbb{R}$, let $F(x, t) := \exp(iux + \frac{1}{2}u^2 t)$, and let $Z_t := F(X_t, t)$. By Itô's formula,

$$Z_t - Z_0 = iu \int_0^t Z_s dX_s + \frac{u^2}{2} \int_0^t Z_s ds - \frac{u^2}{2} \int_0^t Z_s d[X]_s$$

$$Z_t = 1 + iu \int_0^t Z_s dX_s$$

Therefore Z is a local martingale by Theorem 2.5.2. But

$$\sup_{s \le t} |Z_t| = \sup_{s \le t} |\exp(iuX_t + \tfrac{1}{2}u^2 t)| = e^{\frac{1}{2}u^2 t} < \infty.$$

so Z is a martingale, and this implies what we wanted to prove. □

2.8.5 Theorem (Multidimensional Lévy's characterization). *If $X = (X^1, \dots, X^n)$ is a vector of n continuous local martingales and, for all i, j, $[X^i, X^j]_t = t\delta_{i,j}$ for all t then X is an n-dimensional Brownian motion.*

2.8.6 Theorem. *Let M be a continuous local martingale with $M_0 = 0$ and such that $\lim_{t\to\infty}[M]_t = \infty$. Let*

(i) $T_s = \inf\{t : [M]_t > s\}$,
(ii) $\mathscr{G}_s := \mathscr{F}_{T_s}$ *and* $\mathbb{G} = (\mathscr{G}_s)_{g\geq 0}$, *and*
(iii) $B_s := M_{T_s}$.

Then B is a $\mathbb{G}$-Brownian motion, $[M]_t$ is $\mathbb{G}$-stopping time for all t, and $M_t = B_{[M]_t}$ a.s. for all t.

Proof: Since $[M]_\infty = \infty$, $T_s < \infty$ for all s and B is well-defined. Also, since $T_s \leq T_u$ for $s \leq u$, $\mathscr{F}_{T_s} = \mathscr{G}_s \subseteq \mathscr{F}_{T_u} = \mathscr{G}_u$, so $\mathbb{G}$ really is a filtration. It is right continuous because T is right continuous. For all s and t,

$$\{[M]_t \leq s\} = \{T_s \geq t\} \in \mathscr{F}_{T_s} = \mathscr{G}_s,$$

so the $[M]_t$ are $\mathbb{G}$-stopping times. $\mathbb{E}[[M]_\infty^{T_s}] = \mathbb{E}[[M]_{T_s}] = s$ for all s since $[M]$ is a continuous process. By Corollary 1.4.8, M^{T_s} is a u.i. $\mathbb{F}$-martingale with $\mathbb{E}[M_{T_s}^2] = \mathbb{E}[[M]_{T_s}] = s$. By the optional sampling theorem

$$\mathbb{E}[B_s|\mathscr{G}_u] = \mathbb{E}[M_{T_s}|\mathscr{F}_{T_u}] = M_{T_u} = B_u$$

for all $u \leq s$, so B is a $\mathbb{G}$-martingale. One can prove that $(M^{T_s})^2 - [M]^{T_s}$ is a u.i. martingale for each s (exercise). Hence

$$\mathbb{E}[B_s^2 - B_u^2|\mathscr{G}_u] = \mathbb{E}[M_{T_s}^2 - M_{T_u}^2|\mathscr{G}_{T_u}] = \mathbb{E}[[M]_{T_s} - [M]_{T_u}|\mathscr{G}_{T_u}] = s - u.$$

This implies that $(B_u^2 - u)_{u\geq 0}$ is a $\mathbb{G}$-martingale. By Corollary 2.6.16 M and $[M]$ have the same intervals of constancy, so $B = M_T$ is a continuous process. Corollary 2.6.12 implies that $[B]_u = u$, so B is a Brownian motion by Lévy's characterization. (See the exercises around p.99 in the text regarding time changes.) Finally, note that $T_{[M]_t} \geq t$ because T is the right continuous inverse of $[M]$, and $T_{[M]_t} > t$ if and only if $[M]$ is constant on the interval $(t, T_{[M]_t})$, which happens if and only if M is constant on that interval. Hence $B_{[M]_t} = M_t$ for all t. □

2.8.7 Exercise. If B is a $\mathbb{G}$-Brownian motion then B is also a $\tilde{\mathbb{G}}$-Brownian motion, where $\tilde{\mathscr{G}}_s = \bigcap_{u>s} \mathscr{G}_u$. Note that $\mathbb{E}[Z|\bigcap_{u>s}\mathscr{G}_u] = \lim_{u\downarrow s}\mathbb{E}[Z|\mathscr{G}_u]$ (this equality is known as *Lévy's theorem*).

2.8.8 Theorem (Lévy's stochastic area formula). *Let $B = (X, Y)$ be a standard 2-dimensional Brownian motion and let $A_t := \int_0^t X_s dY_s - \int_0^t Y_s dX_s$. Then $\mathbb{E}[e^{iuA_t}] = 1/\cosh(ut)$ for all u and $0 \leq t < \infty$.*

Chapter 3

The Bichteler-Dellacherie Theorem and its connexions to arbitrage

3.1 Introduction

Classical semimartingales

3.1.1 Definition. An adapted càdlàg process S is a *classical semimartingale* if it can be decomposed as $S = S_0 + M + A$, where $M_0 = A_0 = 0$, M is a local martingale, and A is of finite variation.

Remark. Being a classical semimartingale is a local property, i.e. if S is a process for which there is a sequence of stopping times $(T_n)_{n\geq 1}$ such that $T_n \to \infty$ a.s. and S^{T_n} is a classical semimartingale for all n then S is a classical semimartingale. Say $S^{T_n} = M^n + A^n$. If one defines $M := \sum_{n\geq 1} M^n \mathbf{1}_{(T_{n-1},T_n]}$ then one can prove that M is a local martingale and $A := S - M = \sum_{n\geq 1} A^n \mathbf{1}_{(T_{n-1},T_n]}$ is of finite variation.

3.1.2 Theorem (Bichteler '79, Dellacherie '80). *For an adapted, càdlàg, real valued process S the following are equivalent.*
(i) *S is a semimartingale.*
(ii) *S is a classical semimartingale.*

The following, seemingly stronger, theorem is true. This is the theorem that is stated in the textbook and the one that we will eventually prove.

3.1.3 Theorem. *For an adapted, càdlàg, real valued process S the following are equivalent.*
(i) *S is a semimartingale.*
(ii) *S is decomposable (i.e. $S = M + A$ where M is a locally square integrable martingale and A is of finite variation.)*
(iii) *For all $\beta > 0$ there are M and A such that M is a local martingale with jumps bounded by β and A is of finite variation and $S = M + A$.*
(iv) *S is a classical semimartingale.*

So far we have seen that (ii) implies (i) (*cf.* Corollary 2.3.3) and (iii) implies (ii) (because a local martingale with bounded jumps is locally bounded, and in particular is a locally square integrable martingale). The theoretically important, and difficult, results are (iv) implies (iii) and (i) implies (iv). The former is important enough to be restated as a theorem.

3.1.4 Theorem (Fundamental theorem of local martingales).
If M is a local martingale then for all $\beta > 0$ there are N and B such that $M = N + B$, where N is a local martingale with jumps bounded by β and B is of finite variation.

Connexion to finance

Suppose that S is a *price process*, i.e. S_t is the price of an asset at time t, where we make no assumptions about S other than that it is càdlàg. Let H_t be the number of units of S that you hold at time t, so that H is a simple predictable process. The accumulated gains and losses after following the *strategy* H up to time t are $(H \cdot S)_t = \int_0^t H_s dS_s$ (the integral makes sense for any S when H is a simple predictable process). An *arbitrage* over the time horizon $[0, T]$ is a strategy H such that $H_0 = 0$, $(H \cdot S)_T \geq 0$ a.s., and $\mathbb{P}[(H \cdot S)_T > 0] > 0$. It is very important to be able to determine when there are no possible arbitrages.

3.1.5 Definition. We say that S allows a *free lunch with vanishing risk*, or *FLVR*, *for simple integrands* if there is a sequence of simple predictable processes $(H^n)_{n \geq 1}$ such that both of the following conditions hold.
(FL) $(H^n \cdot S)_T^+$ does not converge to zero in probability.
(VR) $\lim_{n \to \infty} \|(H^n \cdot S)^-\|_u = 0$.
Alternatively, S satisfies *no free lunch with vanishing risk*, or *NFLVR*, *for simple integrands* if, for all sequences of simple predictable processes $(H^n)_{n \geq 1}$, (VR) implies $(H^n \cdot S)_T^+ \to 0$ in probability.

It is clear that if S admits an arbitrage then S allows FLVR. In 1994 Delbaen and Schachermayer proved that NFLVR for simple integrands implies S is a semimartingale. Their proof relies on the Bichteler-Dellacherie theorem.

3.1.6 Definition. We say that S allows *free lunch with vanishing risk for simple integrands with little investment*, or *FLVR+LI* if there is a sequence of simple predictable processes $(H^n)_{n \geq 1}$ such that (FL) and (VR) hold and so does
(LI) $\lim_{n \to \infty} \|H^n\|_u = 0$.
Alternatively, S satisfies *no free lunch with vanishing risk and little investment*, or *NFLVR+LI*, *for simple intgrands* if, for all sequences of simple predictable processes $(H^n)_{n \geq 1}$, (VR) and (LI) together imply $(H^n \cdot S)_T^+ \to 0$ in probability.

3.1.7 Theorem. *For a locally bounded, adapted, càdlàg process S the following are equivalent.*
(i) *S satisfies NFLVR+LI.*
(ii) *S is a classical semimartingale.*

3.1.8 Theorem. *For an adapted càdlàg process S the following are equivalent.*

(i) *For all sequences $(H^n)_{n\geq 1}$ of simple predictable processes,*

a) $\lim_{n\to\infty} \|H^n\|_u = 0$*, i.e. (LI)*

b) $\lim_{n\to\infty} \sup_{0\leq t\leq T}(H^n \cdot S)_t^- = 0$ *in probability (this is slightly weaker than (VR)).*

together imply $(H^n \cdot S)_T^+ \to 0$ *in probability.*

(ii) *S is a classical semimartingale.*

Remark. The assertions (ii) implies (i) in both Theorems 3.1.7 and 3.1.8 follow from (iv) implies (i) in Theorem 3.1.3, because (i) in each theorem holds trivially for semimartingales. Also note that (i) implies (iv) in Theorem 3.1.3 is a consequence of Theorem 3.1.8.

3.2 Proofs of Theorems 3.1.7 and 3.1.8

First we want to show that if S is locally bounded, càdlàg, and adapted and satisfies NFLVR+LI then S is a classical semimartingale.

Note that the collection of processes satisfying NFLVR+LI is stable under stopping. Since the property of being a classical semimartingale is local and S is locally bounded, we may assume that $\|S\|_u \leq 1$. Without loss of generality we assume that $S_0 = 0$ and the time interval is $[0,1]$.

Notation. Let $\mathcal{D}_n$ denote the dyadic rationals in $[0,1]$ with denominator at most 2^n, i.e. $\mathcal{D}_n := \{j2^{-n} : j = 0,\dots,2^n\}$. Let $\mathcal{D} := \bigcup_{n\geq 1}\mathcal{D}_n$ denote the collection of dyadic rationals in $[0,1]$.

For all $n \geq 1$ define $S^n := M^n + A^n$, where M^n and A^n are defined on the set $\mathcal{D}_n$ inductively by $A_0^n := 0$, $M_0^n = S_0$, and for $j = 1,\dots,2^n$,

$$A^n_{j2^{-n}} - A^n_{(j-1)2^{-n}} := \mathbb{E}[S_{j2^{-n}} - S_{(j-1)2^{-n}} | \mathscr{F}_{(j-1)2^{-n}}] \tag{3.1}$$

$$M^n_{j2^{-n}} := S_{j2^{-n}} - A^n_{j2^{-n}}. \tag{3.2}$$

Then M^n is a martingale and A^n is predictable and of finite variation. The hope is that we may pass to the limit as $n \to \infty$ in some sense. We are assuming that S is locally bounded, so there are two cases.

CASE 1. If $(M^n)_{n\geq 1}$ and $(A^n)_{n\geq 1}$ remain bounded in some sense then, by using Komlós' lemma, one can pass to the limit to obtain a local martingale M and a predictable process A of finite variation such that $S = M + A$.

CASE 2. If $(M^n)_{n\geq 1}$ or $(A^n)_{n\geq 1}$ does not remain bounded then neither does the other, and one can construct a sequence of strategies witnessing that S allows FLVR+LI. The idea is that if a predictable process can "get close to" a martingale then this leads to an arbitrage.

3.2.1 Theorem. *Let $(f_n)_{n\geq 1}$ be a sequence of random variables.*

(KOMLÓS' LEMMA). *If $(f_n)_{n\geq 1}$ is bounded in L^1, i.e.* $\sup_{n\geq 1}\|f_n\|_1 < \infty$*, then there is a subsequence $(f_{n_k})_{k\geq 1}$ such that $\frac{1}{k}(f_{n_1} + \cdots + f_{n_k})$ converges a.s. as $k \to \infty$.*

(MAZUR'S LEMMA). *If $(f_n)_{n\geq 1}$ is bounded in L^2, i.e. $\sup_{n\geq 1}\|f_n\|_2 < \infty$, then there exist $g_n \in conv(f_n, f_{n+1}, \dots)$ such that $(g_n)_{n\geq 1}$ converges a.s. and in L^2.*

PROOF (OF MAZUR'S LEMMA): Let $\mathcal{H}$ be the Hilbert space generated by the f_n's. For all n define $K_n = \overline{\text{conv}}(f_n, f_{n+1}, \dots)$, where the closure is taken in the L^2-norm. In fact, K_n coincides with the weak closure of $\text{conv}(f_n, f_{n+1}, \dots)$ by a theorem from functional analysis. By the Banach-Alaoglu theorem we know that K_n is weakly compact for all n. Since $K_{n+1} \subseteq K_n$ for all n, $\bigcap_{n\geq 1} K_n \neq \emptyset$ by the finite intersection property of compact sets. Hence there is a sequence of $g_n \in \text{conv}(f_n, f_{n+1}, \dots)$ such that $(g_n)_{n\geq 1}$ converges in L^2 to an element of $\bigcap_{n\geq 1} K_n$. By passing to a subsequence we may assume that $(g_n)_{n\geq 1}$ converges a.s. □

Remark. We can apply Mazur's lemma to countably many sequences simultaneously. More precisely, suppose that $(f_n^m)_{n\geq 1}$ is bounded in L^2 for all m. Then for all n there are convex weights $(\lambda_n^n, \dots, \lambda_{N_n}^n)$, independent of m, such that the sequences $(\lambda_n^n f_n^m + \cdots + \lambda_{N_n}^n f_n^m)_{n\geq 1}$ converge a.s. and in L^2 for all m.

3.2.2 Proposition. *Let $S = (S_t)_{0\leq t\leq 1}$ be càdlàg and adapted, with $S_0 = 0$ and such that $\|S\|_u \leq 1$ and S satisfies NFLVR+LI. Let A^n and M^n be as defined in* (3.1) *and* (3.2). *For all $\varepsilon > 0$ there is $C > 0$ and a sequence of stopping times $(\rho_n)_{n\geq 1}$ such that, for all n,*

(i) *ρ_n takes values in $\mathcal{D}_n \cup \{\infty\}$.*
(ii) *$\mathbb{P}[\rho_n < \infty] < \varepsilon$.*
(iii) *The stopped processes A^{n,ρ_n} and M^{n,ρ_n} satisfy, for all n, $\|M_1^{n,\rho_n}\|_{L^2}^2 \leq C$ and*

$$\text{TV}(A^{n,\rho_n}) := \sum_{j=1}^{2^n} \left| A_{j2^{-n}}^{n,\rho_n} - A_{(j-1)2^{-n}}^{n,\rho_n} \right| \leq C.$$

3.2.3 Lemma. *Under the same assumptions as in Proposition 3.2.2, with*

$$Q^n := \sum_{j=1}^{2^n} (S_{j2^{-n}} - S_{(j-1)2^{-n}})^2,$$

the sequence $(Q^n)_{n\geq 1}$ is bounded in probability (i.e. for all $\alpha > 0$ there is M such that $\mathbb{P}[Q^n \geq M] < \alpha$ for all n).

PROOF: For all n, let $H^n := -\sum_{j=1}^{2^n} S_{(j-1)2^{-n}} \mathbf{1}_{((j-1)2^{-n}, j2^n]}$, a simple predictable process. Recall that $-a(b-a) = \frac{1}{2}(a-b)^2 + \frac{1}{2}(a^2 - b^2)$.

$$\begin{aligned}(H^n \cdot S)_t &= -\sum_{j=1}^{2^n} S_{t\wedge(j-1)2^{-n}}(S_{t\wedge j2^{-n}} - S_{t\wedge(j-1)2^{-n}}) \\ &= \frac{1}{2}\sum_{j=1}^{2^n}(S_{t\wedge j2^{-n}} - S_{t\wedge(j-1)2^{-n}})^2 + \frac{1}{2}(S_0^2 - S_t^2)\end{aligned}$$

$$= \frac{1}{2}Q^n + \frac{1}{2}(S_0^1 - S_t^2)$$

Clearly $\|H^n\|_u \le 1$, and $(H^n \cdot S)_t \ge -\frac{1}{2}$ for all t because $\|S\|_u \le 1$ and Q^n is nonnegative. Assume for contradiction that $(Q^n)_{n\ge1}$ is not bounded in L^0. Then there is $\alpha > 0$ such that for all $m > 0$ there is n_m such that $\mathbb{P}[(H^{n_m} \cdot S)_1 \ge m] \ge \alpha$. Whence $(H^{n_m}/m)_{m\ge1}$ would be a FLVR+LI. □

For $c > 0$ define a sequence of stopping times

$$\sigma_n(c) := \inf\left\{ \frac{k}{2^n} : \sum_{j=1}^{k} (S_{j2^{-n}} - S_{(j-1)2^{-n}})^2 \ge c - 4 \right\}.$$

Given $\varepsilon > 0$ there is c_1 such that $\mathbb{P}[\sigma_n(c_1) < \infty] < \varepsilon/2$ by Lemma 3.2.3.

3.2.4 Lemma. *Under the same assumptions as Proposition 3.2.2, the stopped martingales* $M^{n,\sigma_n(c_1)}$ *satisfy* $\|M_1^{n,\sigma_n(c_1)}\|_{L^2}^2 \le c_1$.

PROOF: For $n \ge 1$ and $k = 1, \dots, 2^n$, since the A^ns are predictable and the M^ns are martingales,

$$\begin{aligned}\mathbb{E}[(S_{k2^{-n}}^{\sigma_n(c_1)} - S_{(k-1)2^{-n}}^{\sigma_n(c_1)})^2] &= \mathbb{E}[(M_{k2^{-n}}^{\sigma_n(c_1)} - M_{(k-1)2^{-n}}^{\sigma_n(c_1)})^2] + \mathbb{E}[(A_{k2^{-n}}^{\sigma_n(c_1)} - A_{(k-1)2^{-n}}^{\sigma_n(c_1)})^2] \\ &\ge \mathbb{E}[(M_{k2^{-n}}^{\sigma_n(c_1)})^2 - (M_{(k-1)2^{-n}}^{\sigma_n(c_1)})^2]\end{aligned}$$

Write $\mathbb{E}[(M_1^{\sigma_n(c_1)})^2]$ as a telescoping series and simplify to get

$$\begin{aligned}\mathbb{E}[(M_1^{\sigma_n(c_1)})^2] &= \sum_{k2^{-n} \le \sigma_n(c_1)} \mathbb{E}[(S_{k2^{-n}}^{\sigma_n(c_1)} - S_{(k-1)2^{-n}}^{\sigma_n(c_1)})^2] + \mathbb{E}[(S_{\sigma_n(c_1)} - S_{\sigma_n(c_1)-2^{-n}})^2] \\ &\le (c_1 - 4) + 2^2 = c_1.\end{aligned}$$

□

3.2.5 Lemma. *Let* $V^n := \mathrm{TV}(A^{n,\sigma_n(c_1)}) = \sum_{i=1}^{2^n(\sigma_n(c_1)\wedge 1)} |A_{j2^{-n}}^n - A_{(j-1)2^{-n}}^n|$. *Under the assumptions of Proposition 3.2.2 the sequence* $(V^n)_{n\ge1}$ *is bounded in probability.*

PROOF: Assume for contradiction that $(V^n)_{n\ge1}$ is not bounded in probability. Then there is $\alpha > 0$ such that for all k there is n_k such that $\mathbb{P}[V^{n_k} \ge k] \ge \alpha$. For $n \ge 1$ define

$$b_{j-1}^n := \mathrm{sign}\left(A_{j2^{-n}}^{n,\sigma_n(c_1)} - A_{(j-1)2^{-n}}^{n,\sigma_n(c_1)} \right) \in \mathscr{F}_{(j-1)2^{-n}}$$

and $H^n(t) := \sum_{j=1}^{2^n} b_{j-1}^n \mathbf{1}_{((j-1)2^n, j2^{-n}]}(t)$. Then $\|H^n\|_u \le 1$ and

$$\begin{aligned}(H^{n,\sigma_n(c_1)} \cdot S)_t &= \sum_{j \le \lfloor t2^n \rfloor} b_{j2^{-n}}^n \left(S_{j2^{-n}}^{\sigma_n(c_1)} - S_{(j-1)2^{-n}}^{\sigma_n(c_1)} \right) + b_{\lfloor t2^n \rfloor}^n \left(S_t^{\sigma_n(c_1)} - S_{\lfloor t2^n \rfloor}^{\sigma_n(c_1)} \right) \\ &\ge (H^{n,\sigma_n(c_1)} \cdot A^n)_{\lfloor t2^n \rfloor 2^{-n}} + (H^{n,\sigma_n(c_1)} \cdot M^n)_{\lfloor t2^n \rfloor 2^{-n}} - 2\end{aligned}$$

and at time $t = 1$ we have $(H^{n,\sigma_n(c_1)} \cdot S)_1 = V^n + (H^{n,\sigma_n(c_1)} \cdot M^n)_1$. But the second summand is bounded in L^2 (it is at most c_1 by Lemma 3.2.4), so we conclude that $(H^{n,\sigma_n(c_1)} \cdot S)_1$ is not bounded in probability.

Define a sequence of stopping times

$$\eta_n(c) := \inf\left\{ \frac{j}{2^n} : |(H^{n,\sigma_n(c_1)} \cdot M^n)_{j2^{-n}} \geq c \right\}.$$

Because $\mathbb{E}[(\sup_{1\leq j\leq 2^n}((h^{n,\sigma_n(c_1)} \cdot M^n)_{j2^{-n}})^2] \leq 4c_1$ by Doob's sub-martingale inequality, $(H^{n,\sigma_n(c_1)} \cdot M^n)$ is bounded in probability. Therefore there is $c' > 0$ such that $\mathbb{P}[\eta_n(c') < \infty] \leq \alpha/2$. Note that $H^{n,\sigma_n(c_1)\wedge\eta_n(c')} \cdot S$ is (uniformly) bounded below by c'. We claim $(H^{n,\sigma_n(c_1)\wedge\eta_n(c')} \cdot S)_1$ is not bounded in probability. Indeed, for any n and any k,

$$\begin{aligned}\alpha &\leq \mathbb{P}[(H^{n,\sigma_n(c_1)\wedge\eta_n(c')} \cdot S)_1 \geq k] \\ &\leq \mathbb{P}[(H^{n,\sigma_n(c_1)} \cdot S)_1 \geq k, \eta_n(c') = \infty] + \mathbb{P}[\eta_n(c') < \infty].\end{aligned}$$

Since $\mathbb{P}[\eta_n(c') < \infty] \leq \alpha/2$, the probability of the other event is at least $\alpha/2$. This gives the desired contradiction because it is now easy to construct a FLVR+LI. □

PROOF (OF PROPOSITION 3.2.2): Define a sequence of stopping times

$$\tau_n(c) := \inf\left\{ \frac{k}{2^n} : \sum_{j=1}^{k} |A^n_{j2^{-n}} - A^n_{(j-1)2^{-n}}| \geq c \right\}.$$

By Lemma 3.2.5 there is c_2 such that $\mathbb{P}[\tau_n(c_2) < \infty] < \varepsilon/2$. Take $C := c_1 \vee c_2$ and $\rho_n := \sigma_n(c_1) \wedge \tau_n(c_2)$. □

3.2.6 Lemma. *Let $f, g : [0,1] \to \mathbb{R}$ be measurable functions, where f is left continuous and takes finitely many values. Say $f = \sum_{k=1}^{K} f(s_k)\mathbf{1}_{(s_{k-1},s_k]}$. Define*

$$(f \cdot g)(t) := \sum_{k=1}^{K} f(s_{k-1})(g(s_k) - g(s_{k-1})) + f(s_{k(t)})(g(t) - g(s_{k(t)}))$$

where $k(t)$ is the biggest of the k such that s_k less than or equal to t. Then for all partitions $0 \leq t_0 \leq \cdots \leq t_M \leq 1$,

$$\sum_{i=1}^{M} |(f \cdot g)(t_i) - (f \cdot g)(t_{i-1})| \leq 2\,\mathrm{TV}(f)\|g\|_\infty + \left(\sum_{i=1}^{M} |g(t_i) - g(t_{i-1})|\right)\|f\|_\infty.$$

3.2.7 Proposition. *Let $S = (S_t)_{0\leq t\leq 1}$ be càdlàg and adapted, with $S_0 = 0$ and such that $\|S\|_u \leq 1$ and S satisfies NFLVR+LI. For all $\varepsilon > 0$ there is C and a $[0,1] \cup \{\infty\}$ valued stopping time α such that $\mathbb{P}[\alpha < \infty] < \varepsilon$ and sequences $(\mathbf{M}^n)_{n\geq 1}$ and $(\mathbf{A}^n)_{n\geq 1}$ of continuous time càdlàg processes such that, for all n,*

(i) $\mathbf{A}^n_0 = \mathbf{M}^n_0 = 0$

(ii) $S^\alpha = \mathbf{A}^{n,\alpha} + \mathbf{M}^{n,\alpha}$
(iii) $\mathbf{M}^{n,\alpha}$ *is a martingale with* $\|\mathbf{M}_1^{n,\alpha}\|_{L^2}^2 \le C$
(iv) $\sum_{j=1}^{2^n} |\mathbf{A}^{n,\alpha}_{j2^{-n}} - \mathbf{A}^{n,\alpha}_{(j-1)2^{-n}}| \le C$.

Proof: Let $\varepsilon > 0$ be given. Let C, M^n, A^n, and ρ_n be as in Proposition 3.2.2. Extend M^n and A^n to all $t \in [0,1]$ by defining $M_t^n := \mathbb{E}[M_1^n|\mathscr{F}_t]$ and $A_t^n = S_t - M_t$. Note that the extended A^n is no longer necessarily predictable, and currently we only have control of the total variation of A^{n,ρ_n} over $\mathcal{D}_n$, i.e.

$$\sum_{j=1}^{2^n(\rho_n\wedge 1)} |A^n_{j2^{-n}} - A^n_{(j-1)2^{-n}}| \le C.$$

Notice that, for $t \in ((j-1)2^{-n}, j2^{-n}]$,

$$\begin{aligned} A_t^n &= S_t - M_t^n \\ &= S_t - \mathbb{E}[M^n_{j2^{-n}}|\mathscr{F}_t] \\ &= S_t - \mathbb{E}[S_{j2^{-n}} - A^n_{j2^{-n}}|\mathscr{F}_t] \\ &= A^n_{j2^{-n}} - (\mathbb{E}[S_{j2^{-n}}|\mathscr{F}_t] - S_t) \end{aligned}$$

From this and $\|S\|_u \le 1$ it follows that $\|A_t^n - A^n_{j2^{-n}}\|_\infty \le 2$, so $\|A^{n,\rho_n}\|_u \le C + 2$.

How do we find the "limit" of sequence of stopping times $(\rho_n)_{n\ge1}$? The trick is to define $R^n := \mathbf{1}_{[0,\rho_n\wedge 1]}$, a simple predictable process, and note that stopping at ρ_n is like integrating R^n, i.e. $A^{n,\rho_n} = R^n \cdot A^n$ and $M^{n,\rho_n} = R^n \cdot M^n$. We have that

$$1 \ge \mathbb{E}[R_1^n] = \mathbb{E}[\mathbf{1}_{\rho_n=\infty}] = 1 - \mathbb{P}[\rho_n < \infty] \ge 1 - \varepsilon.$$

Apply Komlós' lemma to $(R_1^n)_{n\ge1}$ to obtain convex weights $(\mu_n^n, \dots, \mu_{N_n}^n)$ such that

$$\mathbf{R}^n := \sum_{i=n}^\infty \mu_i^n R_1^i \to \mathbf{R}_1 \text{ a.s. as } n \to \infty$$

By the dominated convergence theorem, $\mathbb{E}[\mathbf{R}_1] \ge 1 - \varepsilon$. Observe that

$$\mathbf{R}^n \cdot S = \underbrace{\sum_{i=n}^\infty \mu_i^n (R^i \cdot M^i)}_{L^2 \text{ norm} \le \sqrt{C}} + \underbrace{\sum_{i=n}^\infty \mu_i^n (R^i \cdot A^i)}_{\text{TV over } \mathcal{D}_n \text{ is} \le C}$$

Define $\alpha_n := \inf\{t : \mathbf{R}_t^n \le \frac{1}{2}\}$. Each $\mathbf{R}^n$ is a left continuous, decreasing process. In particular, $\mathbf{R}^n_{\alpha_n} \ge \frac{1}{2} > 0$, so we can divide by this quantity. We claim that $\mathbb{P}[\alpha_n < \infty] < \varepsilon$. Indeed, on the event $[\alpha_n < \infty]$, $\mathbf{R}_1^n \le \mathbf{R}^n_{\alpha_n+} \le \frac{1}{2}$ so

$$\varepsilon \ge \mathbb{E}[1 - \mathbf{R}_1^n] \ge \mathbb{E}[(1 - \mathbf{R}^n_{\alpha_n+})\mathbf{1}_{\alpha_n<\infty}] \ge \frac{1}{2}\mathbb{P}[\alpha_n < \infty].$$

Define new processes $T^n_t := \mathbf{1}_{[0,\alpha_n]}(t)/\mathbf{R}^n_t$. Then $\|T^n\|_u \le 2$ and $T^n \cdot (\mathbf{R}^n \cdot S) = S^{\alpha_n}$. Thus we define $\mathbf{M}^n$ and $\mathbf{A}^n$ by

$$S^{\alpha_n} = T^n \cdot \left(\sum_{i=n}^{\infty} \mu^n_i (R^i \cdot M^i) \right) + T^n \cdot \left(\sum_{i=n}^{\infty} \mu^n_i (R^i \cdot A^i) \right) =: \mathbf{M}^n + \mathbf{A}^n.$$

The total variation of T^n over $\mathcal{D}_n$ is bounded by 3. By Lemma 3.2.6,

$$\begin{aligned} \sum_{j=1}^{2^n} |\mathbf{A}^n_{j2^{-n}} - \mathbf{A}^n_{(j-1)2^{-n}}| &\le 2\,\mathrm{TV}_n(T^n) \left\| \sum_{i=n}^{\infty} \mu^n_i (R^i \cdot A^i) \right\|_\infty \\ &\quad + \|T^n\|_\infty \,\mathrm{TV}_n \left(\sum_{i=n}^{\infty} \mu^n_i (R^i \cdot A^i) \right) \\ &\le 6(C+2) + 2C \end{aligned}$$

That $\|\mathbf{M}^n_1\|^2_{L^2} \le C$ follows from the fact that $\|M^{n,\alpha_n}_1\|^2_{L^2} \le C$. To finish the proof, we show that there is a subsequence $(\alpha_{n_k})_{k\ge1}$ such that $\alpha := \inf_k \alpha_{n_k}$ satisfies $\mathbb{P}[\alpha < \infty] \le 4\varepsilon$. We know $\mathbb{P}[\mathbf{R}_1 \le \frac{2}{3}] \le 3\varepsilon$ because $\mathbb{E}[\mathbf{R}_1] \ge 1 - \varepsilon$. Since $\mathbf{R}^n_1 \to \mathbf{R}_1$ a.s. there is a subsequence such that $\mathbb{P}[|\mathbf{R}^n_1 - \mathbf{R}_1| \ge \frac{1}{15}] \le \varepsilon 2^{-k}$. Finally,

$$\begin{aligned} \mathbb{P}[\alpha < \infty] &\le \mathbb{P}\left[\inf_k \mathbf{R}^{n_k}_1 \le \frac{3}{5} \right] \\ &\le 3\varepsilon + \mathbb{P}\left[\inf_k \mathbf{R}^{n_k}_1 \le \frac{3}{5}, \mathbf{R}_1 > \frac{2}{3} \right] \\ &\le 3\varepsilon + \sum_{k=1}^{\infty} \mathbb{P}\left[\mathbf{R}^{n_k}_1 \le \frac{3}{5}, \mathbf{R}_1 > \frac{2}{3} \right] \\ &\le 3\varepsilon + \sum_{k=1}^{\infty} \mathbb{P}\left[|\mathbf{R}^{n_k}_1 - \mathbf{R}_1| \ge \frac{1}{15} \right] \\ &\le 4\varepsilon \end{aligned}$$

Therefore $(\mathbf{M}^n)_{n\ge1}$, $(\mathbf{A}^n)_{n\ge1}$, and α have the desired properties. □

Remark. One thing to take away from this is that if you need to take a "limit of stopping times" then one way to do it is to turn the stopping times into processes and take the limits of the processes.

Proof (of (i) implies (ii) in Theorem 3.1.7):
We may assume the hypothesis of Proposition 3.2.7. Let $\varepsilon > 0$ and take C, α, $(\mathbf{M}^n)_{n\ge1}$, and $(\mathbf{A}^n)_{n\ge1}$ as in Proposition 3.2.7. Apply Komlós' lemma to find convex weights $(\lambda^n_n, \dots, \lambda^n_{N_n})$ such that

$$\begin{aligned} \lambda^n_n \mathbf{M}^{n,\alpha}_1 + \cdots + \lambda^n_{N_n} \mathbf{M}^{N_n,\alpha}_1 &\to \mathbf{M}_1 \\ \lambda^n_n \mathbf{A}^{n,\alpha}_t + \cdots + \lambda^n_{N_n} \mathbf{A}^{N_n,\alpha}_t &\to \mathbf{A}_t \end{aligned}$$

for all $t \in \mathcal{D}$, where the convergence is a.s. (and also in L^2 for $\mathbf{M}$). For all n,

$$\sum_{j=1}^{2^n} |\mathbf{A}^{n,\alpha}_{j2^{-n}} - \mathbf{A}^{n,\alpha}_{(j-1)2^{-n}}| \le C$$

so the total variation of $\mathbf{A}$ over $\mathcal{D}$ is bounded by C. Further, we have $S^\alpha = \mathbf{M}_t + \mathbf{A}_t$ (where $M_t = \mathbb{E}[M_1|\mathscr{F}_t]$). A is càdlàg on $\mathcal{D}$, so define it on all of $[0,1]$ to make it càdlàg. M is an L^2 martingale so it has a càdlàg modification. Since $\mathbb{P}[\alpha < \infty] < \varepsilon$ and $\varepsilon > 0$ was arbitrary, and the class of classical semimartingales is local, S must be a classical semimartingale. □

Remark. On the homework there is an example of a (not locally bounded) adapted càdlàg process that satisfies NFLVR and is not a classical semimartingale.

Proof (of (i) implies (ii) in Theorem 3.1.8):
We no longer assume that S is locally bounded. The trick is to leverage the result for locally bounded processes by subtracting the "big" jumps from S. Assume without loss of generality that $S_0 = 0$ and define $J_t := \sum_{s\le t} \Delta S_s \mathbf{1}_{|\Delta S_s|\ge 1}$. Then $X := S - J$ is an adapted, càdlàg, locally bounded process. We will show that Theorem 3.1.8(i) for S implies NFLVR+LI for X, so that we may apply Theorem 3.1.7 to X. Then since J is of finite variation, this will then imply S is a classical semimartingale.
Suppose $H^n \in \mathcal{S}$ are such that $\|H^n\|_u \to 0$ and $\|(H^n\cdot X)^-\|_u \to 0$. We need to prove that $(H^n \cdot X)_T \to 0$ in probability. First we will show that $\|(H^n \cdot S)^-\|_u \to 0$.

$$\begin{aligned}\sup_{0\le t\le T} (H^n\cdot S)^-_t &\le \sup_{0\le t\le T} (H^n\cdot X)^-_t + \sup_{0\le t\le T} |(H^n\cdot J)_t| \\ &\le \sup_{0\le t\le T} (H^n\cdot X)^-_t + (\|H^n\|_\infty \cdot \mathrm{TV}(J))_T \\ &\to 0 \text{ by the assumptions on } H^n\end{aligned}$$

By (i), $(H^n \cdot S)_T \to 0$ in probability. Since $(H^n \cdot J) \to 0$ in probability (as J is a semimartingale), we conclude that $(H^n \cdot X)_T = (H^n \cdot S)_T - (H^n \cdot J)_T \to 0$ in probability. Therefore X satisfies NFLVR+LI. □

Remark.

(i) S satisfies NFLVR if $\|(H^n\cdot S)^-\|_u \to 0$ implies $(H^n\cdot S)_T \to 0$ in probability. This is equivalent to the statement that $\{(H\cdot S)_T : H \in \mathcal{S}, H_0 = 0, \|(H\cdot S)^-\|_u \le 1\}$ is bounded in probability. The latter is sometimes called *no unbounded profit with bounded risk* or *NUPBR*.

(ii) It is important to note that this definition of NFLVR does not imply that there is no arbitrage. The usual definition of NFLVR in the literature is as follows. Let $\mathcal{K} := \{(H\cdot S)_T : H \in \mathcal{S}, H_0 = 0\}$ and let $\mathcal{C} := \{f \in L^\infty : f \le g$ for some $g \in \mathcal{K}\}$. It is said that there is NFLVR if and only if $\overline{\mathcal{C}} \cap L^\infty_+ = \{0\}$ (the closure is taken in the norm topology of L^∞). This is equivalent to NUPBR+NA, (where *NA* is *no arbitrage*, the condition that $\mathcal{K} \cap L^\infty_+ = \{0\}$).

3.3 A short proof of the Doob-Meyer theorem

3.3.1 Definition. An adapted process S is said to be of *class D* if the collection of random variables $\{S_\tau : \tau \text{ is a finite valued stopping time}\}$ is uniformly integrable.

càdlàg + predictable implies what? Note that *-martingale means càdlàg.

3.3.2 Theorem (Doob-Meyer for sub-martingales). *Let $S = (S_t)_{0\le t\le T}$ be a sub-martingale of class D. Then S can be written uniquely as $S = S_0 + M + A$ where M is a u.i. martingale with $M_0 = 0$ and A is a càdlàg, increasing, predictable process with $A_0 = 0$.*

We will require the following Komlós-type lemma to obtain a limit in the proof of the theorem.

3.3.3 Lemma. *If $(f_n)_{n\ge1}$ is a u.i. sequence of random variables then there are $g_n \in conv(f_n, f_{n+1}, \dots)$ such that $(g_n)_{n\ge1}$ converges in L^1.*

The details need to be filled in here.

Proof: Define $f_n^k = f_n \mathbf{1}_{|f_n|\le k}$ for all $n, k \in \mathbb{N}$. Then $f_n^k \in L^2$ for all n and k since they are bounded random variables. As noted before, there are convex weights $(\lambda_n^n, \dots, \lambda_{N_n}^n)$ and $f^k \in L^2$ such that $\lambda_n^n f_n^k + \cdots + \lambda_{N_n}^n f_{N_n}^k \to f^k \in L^2$. Since $(f_n)_{n\ge1}$ is u.i.,

$$\lim_{k\to\infty} \sup_n \|(\lambda_n^n f_n^k + \cdots + \lambda_{N_n}^n f_{N_n}^k) - (\lambda_n^n f_n + \cdots + \lambda_{N_n}^n f_{N_n})\|_1 = 0$$

which implies that $(\lambda_n^n f_n + \cdots + \lambda_{N_n}^n f_{N_n})_{n\ge0}$ is Cauchy in L^1. □

Proof (of the Doob-Meyer theorem for sub-martingales):
Assume without loss of generality that $T = 1$. Subtracting the u.i. martingale $(\mathbb{E}[S_1|\mathscr{F}_t])_{0\le t\le1}$ from S we may further assume that $S_1 = 0$ and $S_t \le 0$ for all t. Define $A_0^n = 0$ and, for $t \in \mathcal{D}_n$,

$$\begin{aligned} A_t^n - A_{t-2^{-n}}^n &:= \mathbb{E}[S_t - S_{t-2^{-n}}|\mathscr{F}_{t-2^{-n}}] \\ M_t^n &:= S_t - A_t^n. \end{aligned}$$

$A^n = (A_t^n)_{t\in\mathcal{D}_n}$ is predictable and it is increasing since S is a sub-martingale. Furthermore, $M_1^n = -A_1^n$ since $S_1 = 0$, and in fact $M_t^n = -\mathbb{E}[A_1^n|\mathscr{F}_t]$ for all $t \in \mathcal{D}_n$. Therefore for every $\mathcal{D}_n$ valued stopping time τ,

$$S_\tau = -\mathbb{E}[A_1^n|\mathscr{F}_\tau] + A_\tau^n.$$

We would like to show that $(A_1^n)_{n\ge1}$ is u.i. (and hence that $(M_1^n)_{n\ge1}$ is u.i.). For $c > 0$ define $\tau_n(c) := \inf\{(j-1)2^{-n} : A_{j2^{-n}}^n > c\} \wedge 1$, which is a stopping time since A^n is predictable. Then $A_{\tau_n(c)} \le c$ and $\{A_1^n > c\} = \{\tau_n(c) < 1\}$, so

$$\begin{aligned} \mathbb{E}[A_1^n \mathbf{1}_{A_1^n>c}] &= \mathbb{E}[A_1^n \mathbf{1}_{\tau_n(c)<1}] \\ &= \mathbb{E}[\mathbb{E}[A_1^n|\mathscr{F}_{\tau_n(c)}]\mathbf{1}_{\tau_n(c)<1}] \end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}[(-S_{\tau_n(c)} + A^n_{\tau_n(c)})\mathbf{1}_{\tau_n(c)<1}] \\
&\le -\mathbb{E}[S_{\tau_n(c)}\mathbf{1}_{\tau_n(c)<1}] + c\,\mathbb{P}[\tau_n(c) < 1]
\end{aligned}
$$

Also, $\{\tau_n(c) < 1\} \subseteq \{\tau_n(c/2) < 1\}$, so

$$
\begin{aligned}
-\mathbb{E}[S_{\tau_n(c/2)}\mathbf{1}_{\tau_n(c/2)<1}] &= \mathbb{E}[(A^n_1 - A^n_{\tau_n(c/2)})\mathbf{1}_{\tau_n(c/2)<1}] \\
&\ge \mathbb{E}[(A^n_1 - A^n_{\tau_n(c/2)})\mathbf{1}_{\tau_n(c)<1}] \\
&\ge \frac{c}{2}\,\mathbb{P}[\tau_n(c) < 1]
\end{aligned}
$$

Whence $\mathbb{E}[A^n_1 \mathbf{1}_{A^n_1 > c}] \le -2\,\mathbb{E}[S_{\tau_n(c/2)}\mathbf{1}_{\tau_n(c/2)<1}] - \mathbb{E}[S_{\tau_n(c)}\mathbf{1}_{\tau_n(c)<1}]$ We may now apply the class D assumption on S to show that $(A^n_1)_{n\ge1}$ is u.i., provided that we can show $\mathbb{P}[\tau_n(c) < 1] \to 0$ as $c \to \infty$ uniformly in n. But, for any n,

$$
\mathbb{P}[\tau_n(c) < 1] = \mathbb{P}[A^n_1 > c] \le \frac{1}{c}\,\mathbb{E}[A^n_1] = -\frac{1}{c}\,\mathbb{E}[M^n_1] = -\frac{1}{c}\,\mathbb{E}[S_0]
$$

which goes to zero as $c \to \infty$. Therefore $(A^n_1)_{n\ge1}$, and hence also $(M^n_1)_{n\ge1}$, are u.i. sequences of random variables.

Extend M^n to all $t \in [0,1]$ by defining $M^n_t := \mathbb{E}[M^n_1|\mathscr{F}_t]$. By Lemma 3.3.3 there are convex weights $\lambda^n_n, \dots, \lambda^n_{N_n}$ and $M \in L^1$ such that the sequence of martingales $\mathbf{M}^n := \lambda^n_n M^n + \cdots + \lambda^n_{N_n} M^{N_n}$ satisfies $\mathbf{M}^n_1 \to M$ in L^1. Then $\mathbf{M}^n_t \to \mathbb{E}[M|\mathscr{F}_t] =: M_t$ as well.

Fill in the details (use Jensen's inequality?).

Extend A^n to on $[0,1]$ by defining it to be $\sum_{t\in\mathcal{D}_n} A^n_t \mathbf{1}_{(t-2^{-n},t]}$. Note that A^n extended in this way is a predictable process. Define $\mathbf{A}^n := \lambda^n_n A^n + \cdots + \lambda^n_{N_n} A^{N_n}$ and $A_t := S_t - M_t$. For all $t \in \mathcal{D}$,

$$
\mathbf{A}^n_t := (S_t - \mathbf{M}^n_t) \to A_t \text{ in } L^1
$$

and, by passing to a subsequence, we may assume the convergence is a.s. This implies in particular that $(A_t)_{t\in\mathcal{D}}$ is increasing, so by right continuity, so is $(A_t)_{t\in[0,1]}$. Further, $\mathbf{A}^n$ is predictable for all n. If we can prove that $\limsup_n \mathbf{A}^n_t(\omega) = A_t(\omega)$ a.s. for all $t \in [0,1]$ then this would imply that A is a predictable process and we would be done.

By right continuity $\limsup_n \mathbf{A}^n_t \le A_t$ for all t and $\lim_{n\to\infty} \mathbf{A}^n_t = A_t$ if A is continuous at t. Since A is càdlàg there are countably many jump times. To prove what we want it suffices to prove, for all stopping times τ, $\limsup_n \mathbf{A}^n_\tau = A_\tau$ a.s. Fix a stopping time τ. Since $\mathbf{A}^n_\tau \le \mathbf{A}^n_1 \to A_1$ in L^1, we may apply Fatou's Lemma to get

$$
\liminf_n \mathbb{E}[A^n_\tau] \le \limsup_n \mathbb{E}[\mathbf{A}^n_\tau] \le \mathbb{E}[\limsup_n \mathbf{A}^n_\tau] \le \mathbb{E}[A_\tau]
$$

Therefore it is actually enough to show that $\lim_{n\to\infty} \mathbb{E}[A^n_\tau] = \mathbb{E}[A_\tau]$. Let $\sigma_n := \inf\{t \in \mathcal{D}_n : t \ge \tau\}$, a stopping time. Note that $\sigma_n \downarrow \tau$ as $n \to \infty$. By the definition of A^n, $A^n_\tau = A^n_{\sigma_n}$, so

$$
\lim_{n\to\infty} \mathbb{E}[A^n_\tau] = \lim_{n\to\infty} \mathbb{E}[A^n_{\sigma_n}]
$$

$$= \lim_{n\to\infty} \mathbb{E}[S_{\sigma_n}] - \mathbb{E}[M_{\sigma_n}] \qquad \text{discrete Doob-Meyer}$$
$$= \mathbb{E}[S_\tau] - \mathbb{E}[M_0] \qquad \text{since } S \text{ is class D}$$
$$= \mathbb{E}[A_\tau].$$
□

3.4 Fundamental theorem of local martingales

3.4.1 Definition. A stopping time T is *predictable* there there is a sequence of stopping times $(T_n)_{n\geq1}$, called an *announcing sequence*, such that $T_n < T$ on the event $\{T > 0\}$ and $T_n \uparrow T$. A stopping time T is said to be *totally inaccessible* if $\mathbb{P}[T = S < \infty] = 0$ for all predictable stopping times S.

3.4.2 Example. Suppose that X is a continuous process. $T := \inf\{t : X_t \geq K\}$ is predictable and $T_n := \inf\{t : X_t \geq K - \frac{1}{n}\}$ is an announcing sequence. The first jump time of a Poisson process is totally inaccessible.

3.4.3 Theorem. *If S is a càdlàg sub-martingale of class D and S jumps only at totally inaccessible stopping times then S can be written uniquely as $S = S_0+M+A$, where M is a u.i. martingale with $M_0 = 0$, and A is continuous, increasing, and $A_0 = 0$.*

PROOF: This is Theorem 10 in Chapter III of the textbook. The proof of uniqueness in this case is easy, since if $M + A = S = M' + A'$ are two decompositions then $A - A'$ is a continuous martingale of finite variation null at zero. Hence $A - A' \equiv 0$ by Theorem 2.6.10. See the textbook for the proof of existence. □

Remark. A stronger statement is true. If S is as in the statement of the Theorem 3.4.3 and S jumps at a totally inaccessible stopping time T then A in the decomposition is continuous at T.

3.4.4 Definition. Let A be an adapted process of finite variation with $A_0 = 0$. If A satisfies any of the following equivalent conditions then A is said to be a *natural process*.

(i) $\mathbb{E}[|A|_\infty] < \infty$ (i.e. A is of *integrable variation*) and $\mathbb{E}[[M,A]_\infty] = 0$ for all bounded martingales M.

(ii) $\mathbb{E}[\int_0^\infty M_{s-} dA_s] = \mathbb{E}[M_\infty A_\infty]$ for all all bounded martingales M. (This is an integration-by-parts formula, in some sense.)

(iii) $\mathbb{E}[\int_0^\infty M_{s-} dA_s] = \mathbb{E}[\int_0^\infty M_s dA_s]$ for all all bounded martingales M.

(iv) $\mathbb{E}[\int_0^\infty \Delta M_{s-} dA_s] = 0$ for all all bounded martingales M.

(v) $[M,A]$ is a martingale for all bounded martingales M.

Remark. A càdlàg process locally of integrable variation is of finite variation.

3.4.5 Theorem. *If A is a natural process and a martingale then $A \equiv 0$.*

Proof: Fix a finite valued stopping time T and let H be a bounded, non-negative martingale. It can be shown that, since A is a martingale,

$$\mathbb{E}\left[\int_0^T H_{s-} dA_s\right] = 0$$

using dominated convergence and the fact that it is trivial if A is a simple process (take the limit of the Riemann sums). This implies, since A is natural, that $\mathbb{E}[H_T A_T] = 0$. Now take $H_t := \mathbb{E}[\mathbf{1}_{A_T>0}|\mathscr{F}_t]$, a bounded, non-negative martingale. Then

$$\mathbb{E}[\mathbf{1}_{A_T>0}A_T] = \mathbb{E}[\mathbb{E}[\mathbf{1}_{A_T>0}|\mathscr{F}_T]A_T] = \mathbb{E}[H_T A_T] = 0.$$

Therefore $A_T \le 0$. But since $\mathbb{E}[A_T] = 0$, it must be the case that $A_T = 0$. Since this holds for any finite valued stopping time T, $A \equiv 0$. □

3.4.6 Theorem. *If A is a predictable process of integrable variation null at zero then A is a natural process.*

Proof: This is Theorem 14 in the textbook. There a different proof in Jacod and Shiryaev as well. □

3.4.7 Lemma. *If Y is a u.i. martingale of finite variation then Y is locally of integrable variation.*

Proof: Define $\tau_n := \inf\{t : |Y_t| > n\}$. For any t we have $|Y_{t-}| \le |Y_0| + |Y|_{t-}$ and hence

$$(\Delta|Y|)_t = |\Delta Y_t| \le |Y_t| + |Y_{t-}| \le |Y_t| + |Y_0| + |Y|_{t-}.$$

Thus

$$|Y|_{\tau_n} \le |Y|_{\tau_n-} + (\Delta|Y|)_{\tau_n} \le 2|Y|_{\tau_n-} + |Y_0| + |Y_{\tau_n}| \le 2n + |Y_0| + |Y_{\tau_n}| \in L^1.$$

Therefore $(\tau_n)_{n\ge1}$ is a localizing sequence for $|Y|$ relative to the class increasing processes integrable at ∞. □

3.4.8 Theorem. *If M is a predictable local martingale of finite variation then M is constant.*

Proof: Suppose that $(T_n)_{n\ge1}$ is a localizing sequence for M. If $X^n := (M - M_0)^{T_n}$ then $X^n \in \mathscr{P}$, $X_0^n = 0$, X^n is u.i., and it is of finite variation. By Lemma 3.4.7, X is locally of integrable variation. Without loss of generality assume that it is of integrable variation. By Theorem 3.4.6, X^n is natural. By Theorem 3.4.5, X^n is the zero process. □

3.4.9 Corollary. *The Doob-Meyer decomposition is unique.*

3.4.10 Theorem (Doob-Meyer without class D).
Suppose that S is a super-martingale. Then S can be decomposed uniquely as $S = S_0 + M - A$ where M is a local martingale with $M_0 = 0$ and A is càdlàg, increasing, and predictable with $A_0 = 0$.

PROOF: Define $T^m := \inf\{t : |S_t| \geq m\} \wedge m$. By the optional sampling theorem, $S_{T^m} \in L^1$ and S^{T^m} is dominated by $m \vee |S_{T^m}| \in L^1$. Hence S^{T^m} is of class D, so by the Doob-Meyer decomposition we may write $S^{T^m} = M^m - A^m$ where M^m is u.i. and A^m is increasing and predictable with $A_0^m = 0$. By the uniqueness in Doob-Meyer for class D, it must be the case that $M^m = M^{m+1}$ and $A^m = A^{m+1}$ on $[0, T^m]$. The desired processes may be defined as $M := \sum_{m=1}^{\infty} M^m \mathbf{1}_{[T_{m-1}, T_m)}$ and $A := \sum_{m=1}^{\infty} A^m \mathbf{1}_{[T_{m-1}, T_m)}$. □

3.5 Quasimartingales, compensators, and the fundamental theorem of local martingales

3.5.1 Definition. Let X be a process and Π be a partition for which $X_{t_i} \in L^1$ for all $t_i \in \Pi$. Define the *variation* of X along Π by $\operatorname{var}_\Pi(X) := \mathbb{E}[C(X, \Pi)]$, where by $C(X, \Pi) := \sum_\Pi |X_{t_i} - \mathbb{E}[X_{t_{i+1}} | \mathscr{F}_{t_i}]|$. We say that X is a *quasimartingale* if $\operatorname{var}(X) := \sup_\Pi \operatorname{var}_\Pi(X) < \infty$ (and hence $\mathbb{E}[|X_t|] < \infty$ for all $t \in [0, \infty)$).

3.5.2 Theorem (Rao). *X on $[0, \infty)$ is a quasimartingale if and only if $X = Y - Z$ where Y and Z are nonnegative super-martingales.*

PROOF: See the textbook, though the proof therein is not so clear. □

3.5.3 Theorem. *A quasimartingale X has a unique decomposition as $X = M + A$ where M is a local martingale and A is a predictable process locally of integrable variation.*

PROOF: Existence and uniqueness follow from Theorems 3.4.10 and 3.5.2. Let's see that A is locally of integrable variation. Suppose $X = Y - Z$ where $Y = M^Y - A^Y$ and $Z = M^Z - A^Z$ by the Doob-Meyer decomposition. Clearly $A = A^Y - A^Z$, so we need only show that A^Y and A^Z are locally of integrable variation. Let τ_m be a localizing sequence for M^Y.

$$\begin{aligned}
\mathbb{E}[A_\infty^{Y,\tau_m}] &= \lim_{t\to\infty} \mathbb{E}[A_t^{Y,\tau_m}] \\
&= \lim_{t\to\infty} \mathbb{E}[M_t^{Y,\tau_m}] - \mathbb{E}[Y_t^{\tau_m}] \\
&= Y_0 - \lim_{t\to\infty} \mathbb{E}[Y_t^{\tau_m}] \\
&\leq Y_0 \in L^1
\end{aligned}$$

and the same argument works for A^Z. □

3.5.4 Corollary (Existence of the compensator). *Let A be a process locally of integrable variation. There is a unique finite variation process $A^p \in \mathscr{P}$ such that $A - A^p$ is a local martingale.*

Remark. A^p is called the *compensator* of A. When $A = [X, Y]$ then $A^p =: \langle X, Y\rangle$, the *predictable variation* between X and Y.

PROOF: A is locally a quasimartingale because it is locally of finite variation. □

3.5.5 Examples.

(i) If N is a Poisson of process of rate λ then $N_t^p = \lambda t$. We know what $N - [N]$ is a local martingale, but in this case $[N] = N$. What is meaningful is that $([N]_t - \lambda t)_{t\ge 0}$ is a local martingale.

(ii) Let $A_t := \mathbf{1}_{t\ge\tau} = \mathbf{1}_{[\tau,\infty)}(t)$, where τ is a random time. Let $\mathbb{F}$ be the minimal filtration that satisfies the usual conditions and makes τ a stopping time. ($\mathbb{F}$ may be realized as follows. Let $\mathscr{F}_t^0 := \sigma(\{\tau \le s\} : s \le t) \vee \mathscr{N}$ and take $\mathscr{F}_t := \bigcap_{\varepsilon>0} \mathscr{F}_{t+\varepsilon}^0$.) The $\mathbb{F}$-compensator of A is $A_t^p = \int_0^{t\wedge\tau} \frac{dF(u)}{1-F(u-)}$, where F is the cumulative distribution function of τ on $[0,\infty]$. Note that if F is continuous then so is A^p.

3.5.6 Theorem (Existence of the predictable projection). *Let X be a bounded process that is $\mathscr{F} \otimes \mathscr{B}$ measurable. There exists a unique process ${}^pX$ such that*

(i) ${}^pX \in \mathscr{P}$

(ii) $({}^pX)_T = \mathbb{E}[X_T | \mathscr{F}_{T-}]$ *on* $\{T < \infty\}$ *for all predictable stopping times* T.

3.5.7 Theorem.

(i) *If X is a local martingale then ${}^pX = X_-$ and ${}^p(\Delta X) = 0$.*

(ii) *If A is locally of integrable variation and càdlàg then ${}^p(\Delta A) = \Delta A^p$.*

PROOF (OF (II)): By the definition of the compensator, $A - A^p$ is a local martingale. By (i), ${}^p(\Delta(A - A^p)) = 0$, so ${}^p(\Delta A) = {}^p(\Delta A^p)$. Since A^p is itself predictable, so is ΔA^p, and hence ${}^p(\Delta A^p) = \Delta A^p$. □

The following proof is based on the presentation in Jacod and Shiryaev. See the textbook for a different presentation.

PROOF (OF THE FUNDAMENTAL THEOREM OF LOCAL MARTINGALES):
Assume without loss of generality that $M_0 = 0$. By localization we may assume that M is a u.i. martingale (this requires work, but we have done it previously). Let $b = \beta/2$ and define $A_t := \sum_{s\le t} \Delta M_s \mathbf{1}_{|\Delta M_s|\ge b}$. The process A is well-defined and of finite variation since M is càdlàg. Let $T_n := \inf\{t : |A|_t > n \text{ or } |M_t| > n\}$ so that $|A|_{T_n} = |A|_{T_n-} + |\Delta A_{T_n}| \le n + |\Delta A_{T_n}|$ and $|\Delta A_{T_n}| \le |\Delta M_{T_n}| \le n + |M_{T_n}|$. Hence $|A|_{T_n} \le 2n + |M_{T_n}| \in L^1$ because M is u.i., so therefore $|A|$ is locally integrable and A^p exists. Let $B := A - A^p$ and $N := M - B$. Clearly B is of finite variation, and it is a local martingale by definition of A^p, and N is a local martingale because

it is a difference of local martingales. Define $X_t := \Delta M_t \mathbf{1}_{|\Delta M_t|<b}$, so that $\Delta A = \Delta M - X$. Then ${}^p(\Delta A) = {}^p(\Delta M) - {}^pX = -{}^pX$ by (i) of Theorem 3.5.7. Therefore $\Delta N = \Delta M - \Delta A + \Delta A^p = X - {}^pX$, so $|\Delta N| \le |X| + |{}^pX| \le 2b$, i.e. the jumps of N are bounded by $2b = \beta$. □

3.5.8 Corollary.
(i) *A classical semimartingale is a semimartingale.*
(ii) *A càdlàg sub- (super-) martingale is a semimartingale.*

PROOF: *Cf.* Theorems 3.1.4, 2.3.1, 2.3.2 and the Doob-Meyer decomposition. □

3.5.9 Theorem. *If $H \in \mathbb{L}$ and M is a local martingale then $H \cdot M$ is well-defined and is a local martingale.*

Remark. This theorem is not true if $H \notin \mathbb{L}$. (Example?) This is an important difference between the general theory of stochastic integration and the theory of stochastic integration with respect to continuous local martingales. In the latter the integral of any predictable process is a local martingale.

PROOF: Write $M = N + B$ as in the fundamental theorem of local martingales. By localization we may assume that H is bounded, M is u.i. and has bounded jumps, and that B has integrable variation (*cf.* lemma in the proof of Theorem 3.4.8). We know by Theorem 3.4.8 that $H \cdot N$ is a locally square integrable martingale. Using the dominated convergence theorem one can prove that $H \cdot B$ is a martingale (use "Riemann sums"). □

3.6 Special semimartingales and another decomposition theorem for local martingales

3.6.1 Definition. A semimartingale X is said to be a *special semimartingale* if it can be written $X = X_0 + M + A$ where $M_0 = A_0 = 0$, M is a local martingale, and A is predictable and of finite variation.

3.6.2 Examples.
(i) Quasimartingales (including sub- and super-martingales).
(ii) Local martingales
(iii) Continuous semimartingales (proved below).

3.6.3 Theorem. *If X is a special semimartingale then the decomposition in the definition is unique. This decomposition is known as the* canonical decomposition *of X.*

3.6.4 Theorem. *The following are equivalent for a semimartingale X.*
(i) *X is a special semimartingale.*

(ii) *X^*, the running maximum of X, is locally integrable.*
(iii) *$J := (\Delta X)^*$ is locally integrable.*
(iv) *Decomposing X as $X = M + A$, where M is a local martingale and A has finite variation, implies that A is locally of integrable variation.*
(v) *There exists a local martingale M and A locally of integrable variation such that $X = M + A$.*

PROOF:
(I) IMPLIES (II) Recall the following fact. If A is càdlàg, predictable, and of finite variation then A is locally of integrable variation. (This is an exercise on the homework.) Recall also that if M is a local martingale then $M_t^* = \sup_{s \le t} |M_s|$ is locally integrable. (Indeed, we may assume that M is u.i. Let $T_n := \inf\{t : M_t^* > n\}$ and note that $(M^*)^{T_n} \le n \vee |M_{T_n}|$, which is integrable.) If X is a special semimartingale then write $X = M + A$ as in the definition and note that $X^* \le M^* + A^* \le M^* + |A|$, so X_t^* is locally integrable for all t.
(II) IMPLIES (III) For all $s \le t$, $|\Delta X_s| = |X_s - X_{s-}| \le 2X_t^*$, which implies $J \le 2X^*$.
(III) IMPLIES (IV) Homework.
(IV) IMPLIES (V) Trivial.
(V) IMPLIES (I) Write $X = M + A$ where A is locally of integrable variation. Then by Corollary 3.5.4, A^p exists and $X = (M + A - A^p) + A^p$. Note that the first summand is a local martingale and the second, A^p, is predictable and of finite variation. □

3.6.5 Theorem. *If X is a continuous semimartingale then X may be written $X = M + A$, where M is a local martingale and A is continuous and of finite variation. In particular, X is a special semimartingale.*

PROOF: By Theorem 3.6.4(iii) X is special, so write $X = M + A$ in the unique way. Let's see that A is continuous. We have $A \in \mathscr{P}$, so ${}^pA = A$ and ${}^p(\Delta A) = \Delta A$. But

$$\begin{aligned}\Delta A = {}^p(\Delta A) &= {}^p(\Delta X) - {}^p(\Delta M) \\ &= {}^p(\Delta X) && {}^p(\Delta M) = 0 \text{ by Theorem 3.5.7} \\ &= 0 && X \text{ is continuous by assumption.}\end{aligned}$$
□

3.6.6 Definition. Two local martingales M and N are *orthogonal* if MN is a local martingale. In this case we write $M \perp N$. A local martingale M is *purely discontinuous* if $M_0 = 0$ and $M \perp N$ for all continuous local martingales N.

Remark. If N is a Poisson process of rate λ then $\tilde{N}_t := N_t - \lambda t$ is purely discontinuous. However, $\tilde{N}$ is not the sum of its jumps, i.e. not locally constant pathwise.

3.6.7 Theorem. *Let $\mathcal{H}^2$ be the space of L^2 martingales.*
(i) *For all $M \in \mathcal{H}^2$, $\mathbb{E}[[M]_\infty] < \infty$ and $\langle M \rangle$ exists.*

(ii) $(M,N)_{\mathcal{H}^2} := \mathbb{E}[\langle M,N\rangle_\infty] + \mathbb{E}[M_0N_0]$ *defines an inner product on* $\mathcal{H}^2$*, and* $\mathcal{H}^2$ *is a Hilbert space with this inner product.*
(iii) $M \perp N$ *if and only if* $\langle M,N\rangle \equiv 0$*, which happens if and only if* $M^T \perp N - N_0$ *for all stopping times* T*.*
(iv) *Take* $\mathcal{H}^{2,c}$ *to be the collection of martingales in* $\mathcal{H}^2$ *with continuous paths and* $\mathcal{H}^{2,d}$ *to be the collection of purely discontinuous martingales in* $\mathcal{H}^2$*. Then* $\mathcal{H}^{2,c}$ *is closed in* $\mathcal{H}^2$ *and* $(\mathcal{H}^{2,c})^\perp = \mathcal{H}^{2,d}$*. In particular,* $\mathcal{H}^2 = \mathcal{H}^{2,c} \oplus \mathcal{H}^{2,d}$*.*

3.6.8 Theorem. *Any local martingale* M *has a unique decomposition* $M = M_0 + M^c + M^d$*, where* M^c *is a continuous local martingale,* M^d *is a purely discontinuous local martingale, and* $M_0^c = M_0^d = 0$*.*

PROOF (SKETCH): Uniqueness follows from the fact that a continuous local martingale orthogonal to itself must be constant. For existence, write $M = M_0 + M' + M''$ as in the fundamental theorem of local martingales, where M' has bounded jumps and M'' is of finite variation. Then M'' is purely discontinuous and M' is locally in $\mathcal{H}^2$. But then $M' = M^c + N$ where M^c is a continuous local martingale and $N \in \mathcal{H}^{2,d}$ is purely discontinuous. Hence we may take $M^d = M'' + N$. □

Suppose that $h : \mathbb{R} \to \mathbb{R}$ and $h(x) = x$ in a neighbourhood $(-b, b)$ of zero, e.g. $h(x) = x\mathbf{1}_{|x|<1}$. Let X be a semimartingale and notice that $\Delta X_s - h(\Delta X_s) \neq 0$ only if $|\Delta X_s| \geq b$. Define

$$\check{X}(h)_t := \sum_{0<s\leq t} \Delta X_s - h(\Delta X_s).$$

Then $X(h) := X - X_0 - \check{X}(h)$ has bounded jumps, so it is a special semimartingale by Theorem 3.6.4. Write $X(h) = M(h) + B(h)$ as in that theorem, where $M(h)$ is a local martingale and $B(h)$ is predictable and of finite variation.

3.6.9 Definition. We call (B, C, ν) the *semimartingale characteristics* of X associated with h, where
(i) $B := B(h)$
(ii) $C := \langle M(h)^c\rangle = [M(h)^c]$, *cf.* Theorem 3.6.8, and note that C doesn't depend on h by the uniqueness part of that theorem.
(iii) ν is the "compensator" of the random measure μ_X associated with the jumps of X.

$$\mu_X(dt, dx) := \sum_{s\leq t} \mathbf{1}_{\Delta X_s\neq 0}\delta_{(s,\Delta X_s)}(dt, dx)$$

(Inspired by the Lévy-Khintchine formula and Lévy measures.)

Remark. $M(h)^c$ and C do not depend on the function h. We write $X^c := M(h)^c$ for the "continuous part" of the semimartingale X. (This does not mean that $X - X^c$ is locally constant.)

3.6.10 Example. If X is the Lévy process that appears in the Lévy-Khintchine formula and $h(x) = x\mathbf{1}_{|x|<1}$ then $B(h) = \alpha t$ (α depends on h), $C(h) = \sigma^2 t$ (σ^2 does not depend on h), and $\nu(dt, dx) = \nu(dx)\,dt$, where ν is the Lévy measure (which depends on h).

3.7 Girsanov's theorem

Working on the usual setup $(\Omega, \mathscr{F}, \mathbb{F} = (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, we know that if $\mathbb{Q} \ll \mathbb{P}$ and X is a $\mathbb{P}$ local martingale then X is a $\mathbb{Q}$ semimartingale (*cf.* Theorem 2.2.1). By the Bichteler-Dellacherie theorem $X = M + A = N + B$ where A and B have finite variation, M is a $\mathbb{P}$ local martingale, and N is a $\mathbb{Q}$ local martingale. The question is whether, knowing M and A, we can describe N and B.

3.7.1 Theorem. *If $\mathbb{Q}|_{\mathscr{F}_t} \ll \mathbb{P}|_{\mathscr{F}_t}$ for all t (we write $\mathbb{Q} \ll_{loc} \mathbb{P}$ in this case) and M is an adapted càdlàg process then the following hold.*

(i) *There exists a unique $Z \geq 0$ such that $Z_t = \frac{d\,\mathbb{Q}|_{\mathscr{F}_t}}{d\,\mathbb{P}|_{\mathscr{F}_t}}$ for all t.*

(ii) *MZ is a $\mathbb{P}$ martingale if and only if M is a $\mathbb{Q}$ martingale.*

(iii) *If MZ is a $\mathbb{P}$ local martingale then M is a $\mathbb{Q}$ local martingale.*

(iv) *If M is a $\mathbb{Q}$ local martingale with localizing sequence $(T_n)_{n\geq 1}$ such that $\mathbb{P}[\lim_{n\to\infty} T_n = \infty] = 1$ then MZ is a $\mathbb{P}$ local martingale.*

Remark. A property being "$\mathbb{P}$-local" subtly depends on the measure $\mathbb{P}$ via the requirement that the localizing sequence of stopping times $(T_n)_{n\geq 1}$ satisfies $T_n \uparrow \infty$ $\mathbb{P}$-a.s. If $\mathbb{Q} \sim_{\text{loc}} \mathbb{P}$ then a property holds $\mathbb{P}$-locally if and only if it holds $\mathbb{Q}$-locally.

3.7.2 Corollary. *If $\mathbb{Q} \sim_{loc} \mathbb{P}$ then MZ is a $\mathbb{P}$ local martingale if and only if M is a $\mathbb{Q}$ local martingale.*

3.7.3 Theorem (Girsanov-Meyer). *Suppose that $\mathbb{Q} \sim_{loc} \mathbb{P}$ and let Z be as in Theorem 3.7.1. If M is a $\mathbb{P}$ local martingale then $M - \int \frac{1}{Z} d[Z, M]$ is a $\mathbb{Q}$ local martingale.*

Proof: By integration-by-parts, $ZM - [Z, M] = Z_- \cdot M + M_- \cdot Z$. By Theorem 3.5.9, $ZM - [Z, M]$ is a $\mathbb{P}$ local martingale. By Corollary 3.7.2 $M - \frac{1}{Z}[Z, M]$ is a $\mathbb{Q}$ local martingale. By integration-by-parts,

$$\frac{1}{Z}[Z, M] = \frac{1}{Z_-} \cdot [Z, M] + [Z, M]_- \cdot \frac{1}{Z} + [\tfrac{1}{Z}, [Z, M]]$$

By Theorem 3.5.9 the middle term is a $\mathbb{Q}$ local martingale. We also have

$$\frac{1}{Z_-} \cdot [Z, M] + [\tfrac{1}{Z}, [Z, M]] = \frac{1}{Z_-} \cdot [Z, M] + \sum \Delta \frac{1}{Z} \Delta [Z, M] = \frac{1}{Z} \cdot [Z, M]$$

where this last integral is a Lebesgue-Stieltjes integral, since $[Z,M]$ has finite variation. Thus $\frac{1}{Z}[Z,M]$ is a $\mathbb{Q}$ local martingale plus $\frac{1}{Z}\cdot[Z,M]$. Finally,

$$M - \frac{1}{Z}\cdot[Z,M] = \underbrace{M - \frac{1}{Z}[Z,M]}_{\mathbb{Q}\text{ local mtg}} + \underbrace{\frac{1}{Z}[Z,M] - \frac{1}{Z}\cdot[Z,M]}_{\mathbb{Q}\text{ local mtg}}$$

□

Remark. $M = (M - \frac{1}{Z}\cdot[Z,M]) + \frac{1}{Z}\cdot[Z,M]$ is a decomposition showing that M is a $\mathbb{Q}$-semimartingale. If M is a special semimartingale then this might not be the canonical decomposition because it may not be the case that $\frac{1}{Z}\cdot[Z,M]$ is predictable.

3.7.4 Theorem (Girsanov's theorem, predictable version).
Suppose that $\mathbb{Q}\ll_{loc}\mathbb{P}$ and let Z be as in Theorem 3.7.1. If M is a $\mathbb{P}$ local martingale and $[M,Z]$ is $\mathbb{P}$ locally of integrable variation and $\langle M,Z\rangle$ is the $\mathbb{P}$-compensator of $[M,Z]$ then $M' = M - \frac{1}{Z}\cdot\langle Z,M\rangle$ is $\mathbb{Q}$-a.s. well-defined and is a $\mathbb{Q}$ local martingale. Moreover, $[M^c]$ is a version of $[(M')^c]$.

Proof: Let $T_n := \inf\{t : Z_t < \frac{1}{n}\}$. Then $A = \frac{1}{Z_-}\langle Z,M\rangle$ is well-defined on $[0,T_n]$. Let $T = \lim_{n\to\infty} T_n$.

$$\begin{aligned}\mathbb{Q}[T<\infty] &\le \mathbb{Q}[\bigcap_n\{T_n<\infty\}]\\ &= \lim_{n\to\infty}\mathbb{Q}[T_n<\infty]\\ &= \lim_{n\to\infty}\mathbb{E}^{\mathbb{P}}[Z_{T_n}\mathbf{1}_{T_n<\infty}]\\ &\le \lim_{n\to\infty}\frac{1}{n} = 0\end{aligned}$$

(Below all of the arguments are on $[0,T_n]$.) By integration-by-parts

$$\begin{aligned}MZ &= M_-\cdot Z + Z_-\cdot M + [Z,M]\\ &= \underbrace{M_-\cdot Z + Z_-\cdot M}_{\mathbb{P}\text{ local mtg (Thm 3.5.9)}} + \underbrace{[Z,M] - \langle Z,M\rangle}_{\mathbb{P}\text{ local mtg}} + \langle Z,M\rangle\end{aligned}$$

Hence $ZM - \langle Z,M\rangle$ is a $\mathbb{P}$-local martingale. Now, A is predictable and of finite variation.

$$AZ = A_-\cdot Z + Z_-\cdot A + [Z,A] = A\cdot Z + Z_-\cdot A$$

(Note that $A\cdot Z$ is well-defined and, by results we will see in Chapter 4, a $\mathbb{P}$ local martingale since A is locally bounded). But $Z_-\cdot A = \langle Z,M\rangle$, so $AZ - \langle Z,M\rangle$ is a $\mathbb{P}$ local martingale. Hence $ZM - ZA = ZM - \langle Z,M\rangle - (ZA - \langle Z,M\rangle)$ is a $\mathbb{P}$ local martingale. By Theorem 3.7.1 $M' = M - A$ is a $\mathbb{Q}$ local martingale.

To see that $[M^c]$ is a version of $[(M')^c]$ on uses the fact that, for a local martingale N, $[N^c] = [N]^c$, which is to be proved as an exercise. (Assume that $[N^d] = \sum(\Delta N)^2$ when $N_0 = 0$.) □

Remark. If $\mathbb{Q} \ll_{\text{loc}} \mathbb{P}$ and X is a $\mathbb{P}$-semimartingale then X is a $\mathbb{Q}$-semimartingale. Indeed, if $X = M + A$ is a decomposition under $\mathbb{P}$ then $M = M' + M''$ where M' has bounded jumps and M'' is of finite variation, so X is M' plus a finite variation process. You can prove that $[Z, M']$ is locally of integrable variation. By the previous theorem, X is $M' - \frac{1}{Z_-}\langle Z, M'\rangle$ plus a finite variation process, so is a $\mathbb{Q}$-semimartingale.

3.7.5 Theorem. *Suppose that $\mathbb{Q} \ll_{loc} \mathbb{P}$ and let Z be the density process $\frac{d\mathbb{Q}}{d\mathbb{P}}$. Let $R := \inf\{t : Z_t = 0, Z_{t-} = 0\}$, X be a $\mathbb{P}$-local martingale, and $U_t := \Delta X_R \mathbf{1}_{t \geq R}$. Then $X_t - \int_0^t \frac{1}{Z_s} d[X,Z]_s + U_t^p$ is a $\mathbb{Q}$-local martingale.*

How to use Girsanov's Theorem

Suppose that X is a $\mathbb{P}$-semimartingale and $X = X_0 + M + A$ is a decomposition with $A = J \cdot \langle M \rangle$ where $J \in \mathbb{L}$ (there may or may not be such a decomposition). Define $\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathscr{E}(-J \cdot M) - Z$, where $dZ = -JZ_- dM$. If the density process is a $\mathbb{Q}$-martingale then

$$
\begin{aligned}
&M - \int \frac{1}{Z_-} d\langle Z, M\rangle && \text{is a } \mathbb{Q}\text{-local martingale}\\
&d[Z,M] = -JZ_- d[M] &&\\
&d\langle Z,M\rangle = -JZ_- d\langle M\rangle &&\\
&M + \int J d\langle M\rangle = X - X_0 && \text{is a } \mathbb{Q}\text{-local martingale}
\end{aligned}
$$

3.7.6 Theorem. *Let M be a local martingale*

(Kazamaki's criterion). *Let $\mathbb{S}$ be the collection of bounded stopping times. If M is continuous and*

$$\sup_{T \in \mathbb{S}} \mathbb{E}[\exp(\tfrac{1}{2} M_T)] < \infty$$

then $\mathscr{E}(M)$ is a u.i. martingale.

(Novikov's criterion). *If M is continuous and $\mathbb{E}[\exp(\frac{1}{2}[M]_\infty)] < \infty$ then $\mathscr{E}(M)$ is a u.i. martingale.*

(Lépingle-Mémin). *If $\Delta M > -1$ and*

$$A_t := \tfrac{1}{2}\langle M^c\rangle + \sum_{s \leq t}((1 + \Delta M_s)\log(1 + \Delta M_s) - \Delta M_s)$$

has a compensator A^p such that $\mathscr{E}[\exp(A^p_\infty)] < \infty$ then $\mathscr{E}(M)$ is a u.i. martingale.

(Protter-Shimbo). *If M is locally square integrable such that $\Delta M > -1$ and*

$$\mathbb{E}[\exp(\tfrac{1}{2}\langle M^c\rangle_\infty + \langle M^d\rangle_\infty)] < \infty$$

then $\mathscr{E}(M)$ is a u.i. martingale.

3.7.7 Exercise. Why must the supremum in Kazamaki's criterion be taken over all bounded stopping times and not just deterministic times?

Chapter 4

General stochastic integration

We know how to integrate elements of $\mathbb{L}$ with respect to semimartingales. However, this set of integrands is not rich enough to prove "martingale representation" theorems nor many formulae regarding semimartingale local time.

4.1 Stochastic integrals with respect to predictable processes

For this section we will assume that all semimartingales that appear are null at zero.

4.1.1 Definition. Let X be a special semimartingale with canonical decomposition $X = M + A$. The $\mathcal{H}^2$-*norm* of X is $\|X\|_{\mathcal{H}^2} := \|[M]_\infty^{1/2}\|_{L^2} + \||A|_\infty\|_{L^2}$. Define $\mathcal{H}^2 := \{X : X \text{ is a special semimartingale with } \|X\|_{\mathcal{H}^2} < \infty\}$

Remark. This extends the definition given in the previous chapter. A local martingale M is in $\mathcal{H}^2$ if and only if M is an L^2-martingale.

4.1.2 Theorem. $(\mathcal{H}^2, \|\cdot\|_{\mathcal{H}^2})$ *is a Banach space.*

PROOF (SKETCH): If $X^n = M^n + A^n$ is Cauchy in $\mathcal{H}^2$ then M^n_∞ is Cauchy in L^2 and so the sequence has a limit M. As an exercise, read the rest of this proof in the textbook. □

4.1.3 Lemma. *If $H \in \mathbb{L}$ and H is bounded and $X \in \mathcal{H}^2$ then $H \cdot X \in \mathcal{H}^2$.*

PROOF: Suppose that $X = M + A$ is the canonical decomposition. By Theorem 3.5.9, $H \cdot M$ is a local martingale. To see that $H \cdot A$ is predictable, write it as a limit of Riemann sums. It is clearly of finite variation. Also notice

$$\|H \cdot X\|_{\mathcal{H}^2} \le \|H\|_u \|X\|_{\mathcal{H}^2} < \infty.$$ □

It can be seen that

$$\|H \cdot X\|_{\mathcal{H}^2} = \|(H^2 \cdot [M])_\infty^{1/2}\|_{L^2} + \||H \cdot A|_\infty\|_{L^2}.$$

The integrals on the right hand side are well-defined for any bounded process.

Notation. If $\mathscr{F}$ is a σ-algebra then $b\mathscr{F}$ denotes the collection of bounded $\mathscr{F}$-measurable functions. Let $b\mathbb{L}$ denote the collection of bounded processes in $\mathbb{L}$.

4.1.4 Definition. Let $X = M + A \in \mathcal{H}^2$ and $H, J \in b\mathscr{P}$. Define

$$d_X(H,J) := \|((H-J)^2 \cdot [M])_\infty^{1/2}\|_{L^2} + \||(H-J) \cdot A|_\infty\|_{L^2}.$$

Remark. d_X is not a metric, but it is a semi-metric.

4.1.5 Theorem. *If $X \in \mathcal{H}^2$ then*

(i) *$b\mathbb{L}$ is dense in $(b\mathscr{P}, d_X)$*

(ii) *If $(H^n)_{n\geq 1} \subseteq b\mathbb{L}$ is Cauchy in $(b\mathscr{P}, d_X)$ then $(H^n \cdot X)$ is Cauchy in $\mathcal{H}^2$.*

(iii) *If $d_X(H^n, H) \to 0$ and $d_X(J^n, H) \to 0$ with $H^n, J^n \in b\mathbb{L}$ and $H \in b\mathscr{P}$ then $\lim_{n\to\infty} H^n \cdot X = \lim_{n\to\infty} J^n \cdot X$ in $\mathcal{H}^2$.*

Remark. Theorem 4.1.5(ii) and (iii) can be more succinctly stated as follows. The mapping $(b\mathbb{L}, d_X) \to (\mathcal{H}^2, \|\cdot\|_{\mathcal{H}^2})$ that sends $H \mapsto H \cdot X$ is an isometry and can be extended to $(b\mathscr{P}, d_X)$.

Proof (of (i)): Let $\mathcal{M} := b\mathbb{L}$ and

$$\mathcal{H} := \{H \in b\mathscr{P} : \text{for all } \varepsilon > 0 \text{ there is } J \in b\mathbb{L} \text{ such that } d_X(H,J) < \varepsilon\}.$$

Then $\mathcal{M}$ is a multiplicative class and $\mathcal{H}$ is a monotone vector space. Indeed, $\mathcal{H}$ is a vector space that contains constants (exercise). Suppose that $(H^n) \subseteq \mathcal{H}$ and $H^n \uparrow H$ pointwise and H is bounded. By the dominated convergence theorem $d_X(H^n, H) \to 0$. Let $\varepsilon > 0$ and pick n_0 such that $d_X(H^{n_0}, H) < \varepsilon/2$ and pick $J \in b\mathbb{L}$ such that $d_X(H^{n_0}, J) < \varepsilon/2$. Then $d_X(H,J) < \varepsilon$, so $H \in \mathcal{H}$. Therefore $\mathcal{H}$ is a monotone vector space. By the monotone class theorem $b\mathscr{P} = b\sigma(\mathcal{M}) \subseteq \mathcal{H}$. ☐

4.1.6 Definition. Let $X \in \mathcal{H}^2$ and $H \in b\mathscr{P}$. Suppose that $(H^n)_{n\geq 1} \subseteq b\mathbb{L}$ is such that $d_X(H^n, H) \to 0$ (there is such a sequence by Theorem 4.1.5(i)). Define the *stochastic integral* of H with respect to X to be

$$H \cdot X := \mathcal{H}^2\text{-}\lim_{n\to\infty}(H^n \cdot X),$$

The limit exists by Theorem 4.1.5(ii) and $H \cdot X$ is well-defined (i.e. does not depend on the choice of sequence) by Theorem 4.1.5(iii).

4.1.7 Theorem. *If $X \in \mathcal{H}^2$ then $\mathbb{E}[(X_\infty^*)^2] = \mathbb{E}[\sup_{t\geq 0} |X_t|^2] \leq 8\|X\|_{\mathcal{H}^2}^2$.*

PROOF: Suppose that $X = M + A$. Then $X^*_\infty \le M^*_\infty + |A|_\infty$. By Doob's maximal inequality, $\mathbb{E}[(M^*_\infty)^2] \le 4\,\mathbb{E}[M^2_\infty] = 4\,\mathbb{E}[[M]_\infty]$. Whence

$$\begin{aligned}
(X^*_\infty)^2 &\le 2(M^*_\infty)^2 + 2(|A|_\infty)^2 \\
\mathbb{E}[(X^*_\infty)^2] &\le 2\,\mathbb{E}[(M^*_\infty)^2] + 2\,\mathbb{E}[(|A|_\infty)^2] \\
&\le 8\,\mathbb{E}[[M]_\infty] + 2\,\mathbb{E}[(|A|_\infty)^2] \\
&\le 8\|X\|^2_{\mathcal{H}^2}
\end{aligned}$$

□

4.1.8 Corollary. *If $X^n \to X$ in $\mathcal{H}^2$ then there is a subsequence $(X^{n_k})_{k\ge1}$ such that $(X^{n_k} - X)^*_\infty \to 0$ a.s.*

4.1.9 Theorem (Properties of the stochastic integral).
Let $X, Y \in \mathcal{H}^2$ and $H, K \in b\mathscr{P}$.

(i) $(H + K)\cdot X = H\cdot X + K\cdot X$
(ii) $H\cdot(X + Y) = H\cdot X + H\cdot Y$
(iii) *If T is a stopping time then $(H\cdot X)^T = (H\mathbf{1}_{[0,T]})\cdot X = H\cdot(X^T)$.*
(iv) $\Delta(H\cdot X) = H\Delta X$
(v) *If T is a stopping time then $H\cdot(X^{T-}) = (H\cdot X)^{T-}$.*
(vi) *If X has finite variation then $H\cdot X$ coincides with the Lebesgue-Stieltjes integral.*
(vii) $H\cdot(K\cdot X) = (HK)\cdot X$
(viii) *If X is a local martingale then $H\cdot X$ is an L^2-martingale.*
(ix) $[H\cdot X, K\cdot Y] = (HK)\cdot[X,Y]$

4.1.10 Exercise. For (ii), show that you can take the same approximating sequence for d_X and d_Y.

PROOF (SKETCH): For (iv), suppose that $(H^n)_{n\ge1} \subseteq b\mathbb{L}$ are such that $d_X(H^n, H) \to 0$ and $(H^n\cdot X - H\cdot X)^*_\infty \to 0$ a.s. Then $\lim_{n\to\infty} H^n\Delta X = \lim_{n\to\infty}\Delta(H^n\cdot X) = \Delta(H\cdot X)$. It follows that

$$\lim_{n\to\infty} H^n \mathbf{1}_{\Delta X\neq0} = \frac{\Delta(H\cdot X)}{\Delta X}\mathbf{1}_{\Delta X\neq0} \text{ a.s.}$$

Show that this implies that $\lim_{n\to\infty} H^n\Delta X = H\Delta X$.

For (ix), it is enough to show that $[H\cdot X, Y] = H\cdot[X,Y]$. Again suppose $(H^n)_{n\ge1} \subseteq b\mathbb{L}$ are such that $d_X(H^n, H) \to 0$ and $(H^n\cdot X - H\cdot X)^*_\infty \to 0$ a.s. Since Y_- is locally bounded we can assume that it is bounded.

$$\begin{aligned}
[H^n\cdot X, Y] &= H^n\cdot[X,Y] \to H\cdot[X,Y] \\
[H^n\cdot X, Y] &= (H^n\cdot X)Y - (H^n\cdot X)_-\cdot Y - Y_-\cdot(H^n\cdot X) \\
&\to (H\cdot X)Y - (H\cdot X)_-\cdot Y - Y_-\cdot(H\cdot X)
\end{aligned}$$

(This proof may require some further explanation.) □

4.1.11 Definition. A property (P) holds *prelocally* for X if there exist a sequence $(T_n)_{n\geq 1}$ of stopping times with $T_n \uparrow \infty$ such that (P) holds for $X^{T_n-} = X\mathbf{1}_{[0,T_n)} + X_{T_n-}\mathbf{1}_{[T_n,\infty)}$ (with the convention that $\infty- = \infty$).

4.1.12 Theorem. *Let X be a semimartingale with $X_0 = 0$. Then X is prelocally an element of $\mathcal{H}^2$.*

Proof: Note that X is not necessarily as special semimartingale. Write $X = M + A$ where M is a local martingale and A is of finite variation (*cf.* the Bichteler-Dellacherie Theorem). Let $T_n := \inf\{t : |M_t| > n \text{ or } |A|_t > n\}$ and define $Y = X^{T_n-}$.

$$Y_t^* = (X^{T_n-})_t^* \leq (M^{T_n-})_t^* + (A^{T_n-})_t^* \leq 2n$$

It follows that Y is a special semimartingale. We may assume, by further localization, that $[Y]$ is bounded. (Indeed, take $R_n = \inf\{t : [Y]_t > n\}$ and note that $[Y^{R_n}]^* \leq n + (\Delta Y_{R_n})^* \leq 3n$.) Suppose that $Y = N + B$ where N is a local martingale and B is predictable. We have seen that $|B|$ is locally bounded (homework), so we may assume that $B \in \mathcal{H}^2$ by localizing once more. Again from the homework, $\mathbb{E}[[Y]_\infty] = \mathbb{E}[[N]_\infty] + \mathbb{E}[[B]_\infty]$, so $\mathbb{E}[[N]_\infty] < \infty$ and hence $N \in \mathcal{H}^2$ as well. □

4.1.13 Definition. Let $X \in \mathcal{H}^2$ with canonical decomposition $X = M + A$. We say that $H \in \mathcal{P}$ is $(\mathcal{H}^2, X)$-*integrable* if

$$\mathbb{E}\left[\int_0^\infty H_s^2 d[M]_s\right] + \mathbb{E}\left[\left(\int_0^\infty |H_s| d|A|_s\right)^2\right] < \infty.$$

4.1.14 Theorem. *If $X \in \mathcal{H}^2$ and $H \in \mathcal{P}$ is $(\mathcal{H}^2, X)$-integrable then the sequence $((H\mathbf{1}_{|H|\leq n}) \cdot X)_{n\geq 1}$ is a Cauchy sequence in $\mathcal{H}^2$.*

Proof: Let $H^n := H\mathbf{1}_{|H|\leq n}$. Clearly $H^n \in b\mathcal{P}$, so $H^n \cdot X$ is well-defined.

$$\begin{aligned} \|H^n \cdot X - H^m \cdot X\|_{\mathcal{H}^2} &= d_X(H^n, H^m) \\ &= \|((H^n - H^m)^2 \cdot [M])^{1/2}\|_{L^2} + \||(H^n - H^m)^2 \cdot A|_\infty\|_{L^2} \\ &\to 0 \text{ by the dominated convergence theorem.} \end{aligned}$$

□

4.1.15 Definition. If $X \in \mathcal{H}^2$ and $H \in \mathcal{P}$ is $(\mathcal{H}^2, X)$-integrable then define $H \cdot X$ to be the limit in $\mathcal{H}^2$ of $((H\mathbf{1}_{|H|\leq n}) \cdot X)_{n\geq 1}$, which exists by Theorem 4.1.14. Yet otherwise said,

$$H \cdot X := \mathcal{H}^2\text{-}\lim_{n\to\infty}(H\mathbf{1}_{|H|\leq n}) \cdot X.$$

For any semimartingale X and $H \in \mathcal{P}$, we say that $H \cdot X$ exists if there is a sequence of stopping times $(T_n)_{n\geq 1}$ that witnesses X is prelocally in $\mathcal{H}^2$ and such that H is $(\mathcal{H}^2, X^{T_n-})$-integrable for all n. In this case we define $H \cdot X$ to be $H \cdot (X^{T_n-})$ on $[0, T_n)$ and we say that $H \in L(X)$, the set of X-*integrable processes*.

Remark.

(i) $H\cdot X$ is well-defined because $H\cdot(X^{T_m-}) = (H\cdot X^{T_n-})^{T_m-}$ for $n > m$ (*cf.* Theorem 4.1.9). By similar reasoning, $H\cdot X$ does not depend on the prelocalizing sequence chosen.

(ii) If $H \in (b\mathscr{P})_{\text{loc}}$ then $H \in L(X)$ for any semimartingale X.

(iii) Notice that if $H \in L(X)$ then $H\cdot X$ is a semimartingale. (Prove as an exercise that if $\mathcal{C}$ is the collection of semimartingales then $\mathcal{C} = \mathcal{C}_{\text{loc}} = \mathcal{C}_{\text{preloc}}$.)

(iv) *Warning:* We have not defined $H\cdot X = H\cdot M + H\cdot A$ when $X = M + A$. If it is true we will need to prove it as a theorem. Is it true?

4.1.16 Theorem. *Let X and Y be semimartingales.*

(i) *$L(X)$ is a vector space and $\Lambda : L(X) \to \{$special semimartingales$\} : H \mapsto H\cdot X$ is a linear map.*

(ii) *$\Delta(H\cdot X) = H\Delta X$ for $H \in L(X)$.*

(iii) *$(H\cdot X)^T = (H\mathbf{1}_{[0,T]})\cdot X = H\cdot(X^T)$ for $H \in L(X)$ and stopping times T.*

(iv) *If X is of finite variation and $\int_0^t |H_s|d|X|_s < \infty$ then $H\cdot X$ coincides with the Lebesgue-Stieltjes integral for all $H \in L(X)$.*

(v) *Let $H, K \in \mathscr{P}$. If $K \in L(X)$ then $H \in L(K\cdot X)$ if and only if $HK \in L(X)$, and in this case $(HK)\cdot X = H\cdot(K\cdot X)$.*

(vi) *$[H\cdot X, K\cdot Y] = (HK)\cdot[X,Y]$ for $H \in L(X)$ and $K \in L(Y)$.*

(vii) *$X \in \mathcal{H}^2$ if and only if $\sup_{\|H\|_u\le 1} \|(H\cdot X)^*_\infty\|_{L^2} < \infty$. In this case,*

$$\|X\|_{\mathcal{H}^2} \le \sup_{\|H\|_u\le 1} \|(H\cdot X)^*_\infty\|_{L^2} + 2\|[X]^{1/2}_\infty\|_{L^2} \le 5\|X\|_{\mathcal{H}^2}$$

and

$$\|X\|_{\mathcal{H}^2} \le 3 \sup_{\|H\|_u\le 1} \|(H\cdot X)^*_\infty\|_{L^2} \le 9\|X\|_{\mathcal{H}^2}.$$

PROOF (OF (VII)): To be added. □

4.1.17 Exercise. Suppose that $H \in \mathscr{P}$ and X is of finite variation and also that $\int_0^t |H_s|d|X|_s < \infty$ a.s. for all t (this is the Lebesgue-Stieltjes integral). Is it true that $H \in L(X)$? Note that this is not the same statement as Theorem 4.1.16(iv).

4.1.18 Theorem. *Suppose that X is a semimartingale and $H \in L(X)$ under $\mathbb{P}$. If $\mathbb{Q} \ll \mathbb{P}$ then $H \in L(X)$ under $\mathbb{Q}$ and $H \cdot_{\mathbb{P}} X = H \cdot_{\mathbb{Q}} X$.*

PROOF (SKETCH): Let $(T_n)_{n\ge1}$ witness that $H \in L(X)$ under $\mathbb{P}$, so that $T_n \uparrow \infty$ $\mathbb{P}$-a.s. and H is $(\mathcal{H}^2, X^{T_n-})$-integrable for all n. Let $Z_t := \mathbb{E}^{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathscr{F}_t]$, $S_n := \inf\{t : |Z_t| > n\}$ and $R_n := T_n \wedge S_n$. Since $|Z^{R_n-}| \le n$ one can show that

$$\begin{aligned}\|X\|_{\mathcal{H}^2(\mathbb{Q})} &\le \sup_{\|H\|_u\le 1} \|(H\cdot X)^*_\infty\|_{L^2(\mathbb{Q})} + 2\,\mathbb{E}^{\mathbb{Q}}[[X^{R_n-}]^{1/2}_\infty] \\ &= \sup_{\|H\|_u\le 1} \mathbb{E}_{\mathbb{P}}[Z_{R_n}((H\cdot X)^*_\infty)^2]^{1/2} + 2\,\mathbb{E}^{\mathbb{P}}[Z_{R_n}[X^{R_n-}]^{1/2}_\infty]\end{aligned}$$

$$\leq 5\sqrt{n}\|X\|_{\mathcal{H}^2(\mathbb{P})}$$

(The last inequality is an exercise. There was a hint but I missed it.) This implies that $X \in \mathcal{H}^2(\mathbb{Q})$. The proof that H is $(\mathcal{H}^2, X^{R_n-})$-integrable under $\mathbb{Q}$ is similar. □

4.1.19 Theorem. *Let M be a local martingale and $H \in \mathscr{P}$ be locally bounded. Then $H \in L(M)$ and $H \cdot M$ is a local martingale.*

PROOF: We need the following lemma.

4.1.20 Lemma. *Suppose that $X = M + A$, where $M \in \mathcal{H}^2$ and A is of integrable variation. If $H \in \mathscr{P}$ then there is a sequence $(H_n)_{n\geq 1} \in b\mathbb{L}$ such that $H^n \cdot M \to H \cdot M$ in $\mathcal{H}^2$ and $\mathbb{E}[|H^n - H| \cdot |A|_\infty] \to 0$.*

PROOF: Let $\tilde{\delta}_X(H, J) := \mathbb{E}[(H - J)^2 \cdot [M]_\infty]^{1/2} + \mathbb{E}[|H - J| \cdot |A|_\infty]$ for all $H, J \in \mathscr{P}$. Let $\mathcal{H} = \{H \in \mathscr{P}$: for all $\varepsilon > 0$ there is $J \in b\mathbb{L}$ such that $\tilde{\delta}_X(H, J) < \varepsilon\}$. One can see that $\mathcal{H}$ is a monotone vector space that contains $b\mathbb{L}$. It follows that $b\mathscr{P} \subseteq \mathcal{H}$. □

By localization we may assume that $H \in b\mathscr{P}$ and $M = M' + A$, where $M' \in \mathcal{H}^2$ and A is of integrable variation (this follows from the fundamental theorem of local martingales and a problem from the exam). By the lemma there is $(H_n)_{n\geq 1} \in b\mathbb{L}$ such that $\tilde{\delta}_X(H^n, H) \to 0$. We know that $H^n \cdot M = H^n \cdot M' + H^n \cdot A$ is a martingale for each n. Finally, $H^n \cdot M' \to H \cdot M'$ in $\mathcal{H}^2$ and $\mathbb{E}[|H^n - H| \cdot |A|_\infty] \to 0$ imply that $H \cdot M'$ and $H \cdot A$ are martingales. □

4.1.21 Theorem (Ansel-Stricker). *Let M be a local martingale and $H \in L(M)$. Then $H \cdot M$ is a local martingale if and only if there stopping times $T_n \uparrow \infty$ and $(\theta_n)_{n\geq 1}$ nonpositive integrable random variables such that $(H\Delta M)_t^{T_n} \geq \theta_n$ for all t and all n.*

Remark. If H is locally bounded and T_n is such that $M^*_{T_n}$ is integrable and H^{T_n} is bounded by n then one can take $\theta_n := -2nM^*_{T_n}$ in the theorem to show that $H \cdot M$ is a local martingale.

PROOF: If $H \cdot M$ is a local martingale then there is are $T_n \uparrow \infty$ such that $(H \cdot M)^*_{T_n}$ is integrable, and one may take $\theta_n := -2(H \cdot M)^*_{T_n}$ to see that the condition is necessary. Proof of sufficiency may be found in the textbook. □

4.1.22 Theorem. *Let $X = M + A$ be a special semimartingale and $H \in L(X)$. If $H \cdot X$ is a special semimartingale then $H \in L(M) \cap L(A)$, $H \cdot M$ is a local martingale, and $H \cdot A$ is predictable and of finite variation. In this case $H \cdot X = H \cdot M + H \cdot A$ is the canonical decomposition of $H \cdot X$.*

4.1.23 Theorem (Dominated convergence). *Let X be a semimartingale, $G \in L(X)$, $H^m \in \mathscr{P}$ such that $|H^m| \leq G$, and assume that $H^n \to H$ pointwise. Then $H^m, H \in L(X)$ for all m and $H^n \cdot X \to H \cdot X$ u.c.p.*

PROOF: Suppose that $T_n \uparrow \infty$ are such that G is $(\mathcal{H}^2, X^{T_n-})$-integrable for all n. If $X^{T_n-} = M + A$ is the canonical decomposition then

$$\mathbb{E}[(H^m)^2 \cdot [M]_\infty]^{1/2} + \mathbb{E}[|H^m| \cdot |A|_\infty^2]^{1/2} \le \mathbb{E}[G^2 \cdot [M]_\infty]^{1/2} + \mathbb{E}[|G| \cdot |A|_\infty^2]^{1/2} < \infty$$

Therefore H^m are all $(\mathcal{H}^2, X^{T_n-})$-integrable and hence so is H by the dominated convergence theorem.

For t_0 fixed

$$\| \sup_{t \le t_0} |(H^m - H) \cdot X^{T_n-}| \|_{L^2}^2 \le 8 \|(H^m - H) \cdot X^{T_n-}\|_{\mathcal{H}^2}^2 \to 0$$

by the dominated convergence theorem. Whence $H^m \cdot X^{T_n-} \to H \cdot X^{T_n-}$ uniformly on $[0, t_0]$ in probability. Given $\varepsilon > 0$ let n and m be such that $\mathbb{P}[T_n < t_0] < \varepsilon$ and $\mathbb{P}[((H^m - H) \cdot X^{T_n-})_{t_0}^* > \delta] < \varepsilon$. Whence $\mathbb{P}[((H^m - H) \cdot X)_{t_0}^* > \delta] < 2\varepsilon$. □

4.1.24 Theorem. *Let X be a semimartingale. There exists $\mathbb{Q} \sim \mathbb{P}$ such that $X^t \in \mathcal{H}^2(\mathbb{Q})$ for all t and $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is bounded.*

4.1.25 Example (Emery). Suppose that $T \sim$ Exponential(1), $\mathbb{P}[U = 1] = \mathbb{P}[U = -1] = \frac{1}{2}$, and U and T are independent. Let $X := U \mathbf{1}_{t \ge T}$ and let the filtration be the natural filtration of X, so that X is an $\mathcal{H}^2$ martingale of finite variation and $[X]_t = \mathbf{1}_{t \ge T}$. Let $H_t := \frac{1}{t} \mathbf{1}_{t > 0}$, which is predictable because it is left continuous.

But H is not $(\mathcal{H}^2, X)$-integrable.

$$\mathbb{E}[H^2 \cdot [X]_\infty] = \mathbb{E}\left[\frac{1}{T^2}\right] = \infty$$

Is $H \in L(X)$? We do know that $H \cdot X$ exists as a pathwise Lebesgue-Stieltjes integral, namely $(H \cdot X)_t = \frac{U}{T} \mathbf{1}_{t \ge T}$. We have $\mathbb{E}[|H \cdot X|_t] = \infty$ for all t (because the problem is around zero and not at ∞). It follows in particular that $H \cdot X$ is not a martingale.

Claim. If S is a finite stopping time with $\mathbb{P}[S > 0] > 0$ then $\mathbb{E}[|(H \cdot X)_S|] = \infty$.

Indeed, we have $|H \cdot X)_S| = \frac{1}{T} \mathbf{1}_{S \ge T}$. The hard part is to show that there is $\varepsilon > 0$ such that "$\{T \le \varepsilon\}$ implies $\{S \ge T\}$". If this is the case then $\mathbb{E}[|(H \cdot X)_S|] = \mathbb{E}[\frac{1}{T} \mathbf{1}_{T \le \varepsilon}] = \infty$.

(Perhaps use $\{T \le \varepsilon\} = \{T \le \varepsilon, S < T\} \cup \{T \le \varepsilon, S \ge T\}$ and $\mathbf{1}_{S<T} \in \mathscr{F}_{T-} = \sigma(T) \vee \mathscr{N}$.)

Index