III Stochastic Calculus

Ishan Nath, Lent 2024

Based on Lectures by Dr. Mike Tehranchi

March 19, 2025

Contents

0 Introduction

This course is interested in stochastic integration with respect to continuous martingales, which consist of a finite variation process and local martingales.

This has applications to Brownian motion. For example,

Proposition 0.1 (Lévy). *Suppose W_t and $W_t^2 - t$ are continuous local martingales. Then W is a Brownian motion.*

We also introduce stochastic differential equations, and some notions of existence and uniqueness.

The Markov processes turn out to have a relationship with partial differential equations.

We will then look at some applications to finance, for example arbitrage and utility maximisation.

0.1 Motivation

Suppose (X_t) is a real-valued Markov process, and let P_t be the operator

$$(P_t f)(x) = \mathbb{E}[f(X_t)|X_0 = x] = \int f(y)\, p_t(x, \mathrm{d}y),$$

the transition probability measure. From the Chapman-Kolmogorov equations

$$P_s \circ P_t = P_{s \circ t}.$$

We can introduce the *generator*, which for now we will write as

$$\mathcal{L} = \lim_{s \downarrow 0} \frac{P_s - I}{s}.$$

For example if $X \sim \text{Poisson}(\lambda)$, then

$$(\mathcal{L}f)(x) = \lim \mathbb{E}\left[\frac{f(X_s) - f(x)}{s} \mid X_0 = x\right] = \lambda(f(x+1) - f(x)).$$

Now suppose that X is a continuous Markov process, so

$$\begin{aligned} \mathbb{E}[X_t|X_0 = x] &= x + b(x)t + o(t), \\ \mathrm{Var}(X_t|X_0 = x) &= \sigma(x)^2 t + o(t), \end{aligned}$$

for some functions b, σ. Then we can think of Taylor expanding,

$$\begin{aligned}(\mathcal{L}f)(x) &= \lim_{t\downarrow 0} \mathbb{E}\left[\frac{f(X_t) - f(x)}{t}\right] \\ &= \lim_{t\downarrow 0} \mathbb{E}\left[f'(x)\frac{X_t - x}{t} + \frac{f''(x)}{2}\frac{(X_t - x)^2}{t} + \cdots\right] \\ &- b(x)f'(x) + \frac{1}{2}\sigma(x)^2 f''(x),\end{aligned}$$

which tells us that

$$\mathcal{L} = b\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}.$$

If we let

$$u(t, x) = \mathbb{E}\left[f(X_t)|X_0 = x\right] = (P_t f)(x),$$

then

$$\begin{aligned}\frac{\partial u}{\partial t} &= \lim_{s\downarrow 0} \frac{(P_{s+t}f)(x) - (P_t f)(x)}{s} \\ &= \lim_{s\downarrow 0}\left(\frac{P_s - I}{s}\right) \circ P_t f(x) \to \mathcal{L}u.\end{aligned}$$

This gives us a connection from continuous Markov processes to parabolic PDEs.

So to solve certain PDEs, we have the following problem. Given b, σ, construct a Markov process (X_t) with drift b and volatility σ.

If b, σ are constant, then we can simply set

$$X_t = X_0 + bt + \sigma W_t,$$

where W_t is the (standard) Brownian motion:

- $W_0 = 0$,
- $t \mapsto W_t(\omega)$ is continuous.
- $W_t - W_s \sim N(0, t - s)$, and is independent of $(W_u)_{0\le u\le s}$.

Theorem 0.1 (Wiener)**.** *Brownian motion exists.*

So an idea we could take is to approximate

$$X_{t+\Delta} \approx X_t + b(X_t)\Delta + \sigma(X_t)(W_{t+\Delta} - W_t).$$

Turning this into a differential equation,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \dot{X} = b(X) + \sigma(X)\dot{W}.$$

Unfortunately $\dot{W}$ does not have a classical meaning - Brownian motion is almost surely nowhere differentiable. Instead we can turn this into an integral:

$$X_t = X_0 + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}W_s,$$

where we still need to make sense of $\mathrm{d}W_s$.

0.2 Introducing Stochastic Integration

An analogy we will take is that of random sums.

Suppose that $(\xi_n)_{\mathbb{N}}$ are IID with $\mathbb{P}(\xi_N = \pm 1) = 1/2$. When does

$$\sum_{n=1}^{\infty} a_n \xi_n$$

make sense? If $(a_1, a_2, \ldots) = a \in \ell^1$, then this works, but does not satisfy continuity properties.

Proposition 0.2. *Suppose that $a \in \ell^2$ be deterministic. Let*

$$S_n = \sum_{k=1}^{n} a_k \xi_k.$$

Then (S_n) converges almost-surely and in ℓ^2.

Proof: The expectation is

$$\mathbb{E}[S_n^2] = \sum_{k=1}^{n} a_k^2 \leq \|a\|_{\ell^2}^2.$$

So (S_n) is ℓ^2 bounded and a martingale in the filtration generated by (ξ_n). By the martingale convergence theorem,

$$\mathbb{P}(S_n \text{ converges}) = 1,$$

for all $a \in \ell^2$.

But this is strange. Recall from linear analysis that, if

$$\left(\sum_{k=1}^{n} a_k x_k\right)$$

is bounded for all $a \in \ell^2$, then $x \in \ell^2$. This is the principle of uniform boundedness. But $(\xi_n(\omega)) \notin \ell^2$ for all ω, so there exists $a(\omega) \in \ell^2$ such that

$$\sum_{k=1}^{n} a_k(\omega)\xi_k(\omega) \text{ diverges.}$$

Indeed, we can just take $a_n(\omega) = \xi_n(\omega)/n$.

Theorem 0.2. *Suppose that (a_n) is previsible, and*

$$\sum a_n^2 < \infty \text{ almost surely.}$$

Then

$$\sum_{k=1}^{n} a_k \xi_k$$

converges almost surely.

Proof: When $\mathbb{E} \sum a_n$, then (S_n) is an ℓ^2 martingale as before.

In the general case, let

$$T_N = \inf \left\{ n \geq 1 \mid \sum_{k=1}^{n+1} a_k^2 > N \right\}.$$

This is a stopping time as (a_k) is previsible. So $S^{T_N} = (S_{N \wedge T_N})$ is a martingale, and

$$\mathbb{E}[(S_n^{T_N})^2] = \sum_{n=1}^{\infty} \mathbb{E} a_n^2 \mathbb{1}_{\{n \leq T_N\}} \leq N.$$

Hence (S^{T_N}) is an ℓ^2 martingale for all N. By martingale convergence theorem,

$$S_n^{T_N} \to S_\infty^{T_N} = S_{T_N}.$$

Let

$$A_N = \left\{ \sum_{k=1}^{\infty} a_k^2 \leq N \right\} = \{T_N = \infty\}.$$

Then $(S_n(\omega))$ converges for all $\omega \in A_N$. But,

$$\mathbb{P}\left(\bigcup A_N \right) = 1,$$

so we are done.

So let us try to write

$$X_t = X_0 + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}W_s = \int_0^t H_s \cdot \mathrm{d}Z_s,$$

where $Z_s = (t, W_t)^T$.

1 Finite Variation Lebesgue-Stieltjes Integration

Definition 1.1. $F : \mathbb{R}_+ \to \mathbb{R}$ is a distribution function if and only if F is right-continuous and increasing.

Example 1.1.

Let μ be a σ-finite measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, and $F(x) = \mu(0, x]$. Then F is a distribution function.

Proposition 1.1. *If F is a distribution function, then there exists a unique μ such that*

$$F(x) = F(0) + \mu(0, x]$$

for all $x \geq 0$.

Proof: We will assume the existence of the Lebesgue measure. Define

$$G(y) = \inf\{x \geq 0 \mid F(x) \geq F(0) + y\}.$$

Then $F(x) \geq F(0) + y \iff x \geq G(y)$. Let

$$\mu = \mathrm{Leb} \circ G^{-1}.$$

Note that

$$\begin{aligned} \mu(0, x] &= \mathrm{Leb}\{y \mid G(y) \leq x\} = \mathrm{Leb}\{y \mid y \leq F(x) - F(0)\} \\ &= F(x) - F(0). \end{aligned}$$

Uniqueness follows from Dynkin's lemma and uniqueness of extension, since $\{(0, x] \mid x \geq 0\}$ is a ring that generates the σ-algebra.

Definition 1.2. Suppose that F is a distribution function, and its corresponding measure is μ. If g is (locally) integrable, meaning

$$\int_{(0,t]} |g| \,\mathrm{d}\mu < \infty,$$

then we say that g is (locally) F-integrable, and write

$$\int_0^t g \,\mathrm{d}F = \int_{(0,t]} g \,\mathrm{d}\mu.$$

Proposition 1.2. *Given distribution F and locally F-integrable g, let*

$$I(t) = \int_0^t g \,\mathrm{d}F.$$

Then, I is right-continuous and $\lim_{s\uparrow t} I(s)$ exists for all $t \geq 0$.

Definition 1.3. $f : \mathbb{R}_+ \to \mathbb{R}$ is càdlàg if it is right-continuous, and left limits exist.

Proof: Fix $T > 0$, and $t < T$. Then

$$|g\mathbb{1}_{(0,t+\varepsilon]}| \leq |g|\mathbb{1}_{(0,T]}$$

for $\varepsilon > 0$ small, and

$$g\mathbb{1}_{(0,t+\varepsilon]} \to g\mathbb{1}_{(0,t]}$$

as $\varepsilon \downarrow 0$ pointwise. Hence by DCT,

$$\int_0^{t+\varepsilon} g \,\mathrm{d}F \to \int_0^t g \,\mathrm{d}F.$$

Moreover,

$$|g\mathbb{1}_{(0,t-\varepsilon]}| \leq |g|\mathbb{1}_{(0,T]},$$

and also

$$g\mathbb{1}_{(0,t-\varepsilon]} \to g(\mathbb{1}_{(0,t]} - \mathbb{1}_{\{t\}}),$$

hence again by DCT,

$$\int_0^{t-\varepsilon} g \,\mathrm{d}F \to \int_0^t g \,\mathrm{d}F - g(t)\mu\{t\}.$$

Remark. If F is continuous, then $\int g \,\mathrm{d}F$ is continuous.

Definition 1.4. Let f be càdlàg, and set

$$V_f(t) = \sup_N \sum_{k=1}^{\infty} \left| f(t_k^N \wedge t) - f(t_{k-1}^N \wedge t) \right|,$$

where $t_k^N = k2^{-N}$.

We say f is of *finite variation* if and only if $V_f(t) < \infty$ for all $t \geq 0$, and it is of *bounded variation* if $\sup V_f(t) < \infty$.

Theorem 1.1. *If f is finite variation, then V_f is a distribution function, and $V_f(t) - V_f(s) \geq |f(t) - f(s)|$ for all $0 \leq s \leq t$.*

Proof: We fix f, and drop it from the notation, so

$$V^N(t) = \sum \left| f(t_k^N \wedge t) - f(t_{k-1}^N \wedge t) \right|,$$

$$V^{N+1}(t) = \sum \left| f(t_k^N \wedge t) - f(t_{2k-1}^{N+1} \wedge t) \right| + \left| f(t_{2k-1}^{N+1} \wedge t) - f(t_{k-1}^N \wedge t) \right| \geq V^N(t),$$

by the triangle inequality. So we may take

$$V(t) = \lim_N V^N(t).$$

Now for $0 \leq s \leq t$,

$$V^N(t) - V^N(s) = |f(t) - f(t_n)| + \cdots + |f(t_{m+1} - f(t_m)| - |f(s) - f(t_m)|,$$

where $t_m < s \leq t_{m+1}$, $t_n < t \leq t_{n+1}$. By the triangle inequality,

$$V^N(t) - V^N(s) \geq |f(t) - f(t_{m+1}^-)| + |f(t_{m+1}) - f(t_m)| - |f(t_m) - f(s)|.$$

Taking $N \to \infty$,

$$V(t) - V(s) \geq |f(t) - f(s^+)| + |f(s^+) - f(s^-)| - |f(s^-) - f(s)|,$$

but by càdlàg, $f(s^+) = f(s)$, so these cancel and we get our bound $V(t) - V(s) \geq |f(t) - f(s)|$. Now,

$$V^N(t) - V^N(s) \leq V(t) - V(t_{m+1}) + |f(t_{m+1}) - f(t_m)| - |f(s) - f(t_m)|,$$

so

$$V(t_{m+1}) \leq V(t) - V^N(t) + V^N(s) + |f(t_{m+1}) - f(t_m)| - |f(s) - f(t_m)|.$$

Since V is increasing,

$$\limsup_{\varepsilon} V(s + \varepsilon) \leq V(s),$$

so V is right-continuous.

Remark. We can define the *total variation* as

$$\|f\|_{\text{tvar}} = \sup_{0 = t_0 < \cdots < t_n = t} \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})|.$$

If f is càdlàg and finite variation, then by $V(t) - V(s) \geq |f(t) - f(s)|$,

$$V_f(t) \leq \|f\|_{\text{tvar}} \leq V_f(t).$$

Proposition 1.3. *f is càdlàg and of finite variation if and only if $f = f^\uparrow - f^\downarrow$, where both $f^\uparrow$ and $f^\downarrow$ are distribution functions.*

If f is of finite variation, then it is possible to pick

$$f^\uparrow = \frac{V_f + f}{2}, \qquad f^\downarrow = \frac{V_f - f}{2}.$$

Proof: If f is a distribution function, then

$$V_f(t) - V_f(s) = f(t) - f(s).$$

If $f = f^\uparrow - f^\downarrow$, then by triangle inequality,

$$V_f \leq f^\uparrow + f^\downarrow.$$

If f is of finite variation, then

$$V_f(t) - V_f(s) \geq |f(t) - f(s)| = \max\{f(t) - f(s), f(s) - f(t)\},$$

so both $V_f + f$ and $V_f - f$ are increasing and càdlàg by assumption.

Proposition 1.4. *If g is locally V_f integrable, then g is both locally $f^\uparrow$ and $f^\downarrow$ integrable.*

Proof: Note

$$\int |g| \, \mathrm{d}V_f = \int |g| \, \mathrm{d}f^\uparrow + \int |g| \, \mathrm{d}f_\downarrow.$$

If the LHS is finite, then both terms on the right hand side are as well.

Definition 1.5. If g is locally V_f-integrable, we say that g is locally f-integrable, and write

$$\int g \, \mathrm{d}f = \int g \, \mathrm{d}f^\uparrow - \int g \, \mathrm{d}f^\downarrow.$$

Theorem 1.2. *Let f be càdlàg, finite variation and g locally f-integrable. Let*

$$I(t) = \int_0^t g \, \mathrm{d}f.$$

Then I is càdlàg and finite variation.

Proof: We can write

$$I(t) = \left(\int_0^t g^{\uparrow} \,\mathrm{d}f^{\uparrow} + \int_0^t g^{\downarrow} \,\mathrm{d}f^{\downarrow}\right) - \left(\int_0^t g^{-} \,\mathrm{d}f^{\uparrow} + \int_0^t g^{+} \,\mathrm{d}f^{\downarrow}\right),$$

a difference of distribution functions.

1.1 Finite Variation and Previsible Processes

Introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a filtration $(\mathcal{F}_t)_{t\geq 0}$ with

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

for $s \leq t$.

Definition 1.6. A *finite variation process* Z is such that $t \mapsto Z_t(\omega)$ is càdlàg and of finite variation for all $\omega \in \Omega$, and Z_t is $\mathcal{F}_t$-measurable.

Definition 1.7. A *previsible* or *predictable* σ-algebra on $\mathbb{R}_+ \times \Omega$ is generated by the set $(s, t] \times A$ for all $A \in \mathcal{F}_s$.

$H : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is a *previsible process* if it is measurable with respect to the previsible σ-algebra $\mathcal{P}$.

Remark. If

$$H_t(\omega) = \sum_{k=1}^{n} h_k(\omega) \mathbb{1}_{(t_{k-1}, t_k]}(t),$$

where h_k is $\mathcal{F}_{t_{k-1}}$-measurable for $0 < t_0 < \cdots < t_n$ not random, then H is previsible.

If H is left-continuous and adapted, then H is previsible.

Theorem 1.3. *Let H be previsible, Z be finite variation. Suppose that $H(\omega)$ is $Z(\omega)$-locally integrable for all ω. Let*

$$X_t = \int_0^t H_s \,\mathrm{d}Z_s.$$

Then X is a finite variation process.

Proof: From before, $t \mapsto X_t(\omega)$ is càdlàg and of finite variation for all ω. So we only need to check adaptedness.

First let Z be increasing. Let

$$\mathcal{H} = \left\{ H \mid \text{previsible, bounded, } \int H \, \mathrm{d}Z \text{ adapted} \right\}.$$

Now $h\mathbb{1}_{(t_0,t_1]} \in \mathcal{H}$, where h is $\mathcal{F}_{t_0}$ measurable, for $0 \le t_0 < t_1$ deterministic, since

$$\int_0^t h\mathbb{1}_{(t_0,t_1]} \, \mathrm{d}Z = h(Z_{t\wedge t_1} - Z_{t\wedge t_0})$$

is $\mathcal{F}_t$-measurable for all t. Also if $H^n \in \mathcal{H}$ for all n, and $H^n \to H$ is bounded, then $H \in \mathcal{H}$, since measurability is preserved by pointwise limits.

By monotone class theorem, $\mathcal{H}$ contains all bounded previsible processes.

If H is not bounded, we can let

$$H^n = (H \wedge n) \vee (-n),$$

and then take limits using dominated convergence theorem.

For general Z, write

$$Z = \left(\frac{V_Z + Z}{2} \right) - \left(\frac{V_Z - Z}{2} \right).$$

We just need V_Z to be a finite variation process, i.e. to check it is adapted. But this is fine since

$$V_Z(t) = \lim_N \sum_{k=1}^{\infty} |Z_{t\wedge t_k^N} - Z_{t\wedge t_{k=1}^N}|,$$

where $t_k^N = k2^{-N}$. Each term is $\mathcal{F}_t$-measurable.

We can redefine Z to be *finite variation* if and only if

$$V_Z(t, \cdot) < \infty$$

almost surely, for all $t \ge 0$.

If H is previsible and

$$\int_0^t |H| \, \mathrm{d}V_Z = \int_0^t |H||\mathrm{d}Z| < \infty \qquad \text{a.s.}$$

for all $t \geq 0$, then

$$\int_0^t H \, \mathrm{d}Z$$

can be defined (pointwise on the almost-sure set, and 0 on the null set).

Proposition 1.5. *If $H^n \to H$ (t, ω) pointwise, and*

$$\int \sup |H^n| |\mathrm{d}Z| < \infty \qquad \textit{a.s. for all } t,$$

then

$$\int_0^t H^n \, \mathrm{d}Z \to \int_0^t H \, \mathrm{d}Z \qquad \textit{a.s. for all } t.$$

1.2 The Usual Conditions

Recall that in discrete time, if

$$T = \inf\{n \geq 0 \mid X_n \in A\}$$

for (X_n) adapted, then T is a stopping time, because

$$\{T \leq n\} = \bigcup_{k=0}^{n} \{X_k \in A\} \in \mathcal{F}_n.$$

In continuous time, if we define

$$T = \inf\{t \geq 0 \mid X_t \in A\},$$

then we have

$$\{T \leq t\} = \bigcap_{\varepsilon > 0} \bigcup_{s \leq t + \varepsilon} \{X_s \in A\},$$

which peeks into the future and hence is not a stopping time.

Definition 1.8. A filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfies the *usual conditions* if and only if:

- It is right continuous:

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

- It contains the null sets: if $\mathbb{P}(A) = 0$, then $A \in \mathcal{F}_t$ for all t.

Proposition 1.6. *If $(\mathcal{F}_t)$ satisfies the usual conditions, then $T : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ is a stopping time if and only if $\{T < t\} \in \mathcal{F}_t$ for all t.*

Proof: If T is a stopping time, then

$$\{T < t\} = \bigcup_n \{T \le t - 1/n\} \in \mathcal{F}_t.$$

In the other direction, consider $\{T \le t\}$. We can write

$$\{T \le t\} = \bigcap_{n : n \ge 1/\varepsilon} \{T < t + 1/n\} \in \mathcal{F}_{t+\varepsilon},$$

hence

$$\{T \le t\} \in \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t.$$

Saying we know measure 0 events is like saying we know a Brownian motion has supremum infinity: technically we don't know this at the start of the motion, but with probability 1 it holds.

Theorem 1.4. *If the filtration satisfies the usual conditions, X is adapted and right-continuous taking values in $\mathbb{R}^n$, and $A \subseteq \mathbb{R}^n$ is open, then*

$$T = \inf\{t \ge 0 \mid X_t \in A\}$$

is a stopping time.

Proof: We only need to check that $\{T < t\} \in \mathcal{F}_t$, by our lemma.

$\{T < t\}$ if and only if there exists $s < t$ such that $X_s \in A$.

Since A is open, there exists $\delta > 0$ such that the ball of radius δ around X_s is contained in A.

By right continuity, for close enough q to s, $X_q \in A$. We take q rational. Hence,

$$\{T < t\} = \bigcup_{\substack{q \text{ rational} \\ q < t}} \{X_q \in A\} \in \mathcal{F}_t,$$

since X is adapted.

Theorem 1.5 (Doob's Regularization)**.** *If the filtration satisfies the usual conditions, and if X is a martingale, then there exists another martingale X^* such that:*

- *X^* has càdlàg sample paths, and*
- *$\mathbb{P}(X^*_t = X_t) = 1$ for all t,*

i.e. X^* is a modification of X.

Example 1.2.

Consider
$$X_t = W_t + \mathbb{1}_{\{t=U\}},$$
for U a random variable. Then X_t is not càdàg as it jumps at some point, but $X_t = W_t$ which is càdlàg.

Theorem 1.6. *If X is a càdlàg martingale for a given filtration $(\mathcal{F}_t)$, then X is also a martingale for*
$$\mathcal{F}_t^* = \sigma(\mathcal{N}, \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}).$$

We introduce the notation
$$\mathcal{F}_t^+ = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}.$$

Proof: We know that
$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s$$
for all $0 \leq s \leq t$, as X is integrable, i.e.
$$\mathbb{E}[(X_t - X_s)\mathbb{1}_A] = 0$$
for all $A \in \mathcal{F}_s$, $0 \leq s < t$. We will show the same thing holds for all $Bin\mathcal{F}_s^*$, $0 \leq s < t$.

For any $B \in \mathcal{F}_s^*$, there exists $C \in \mathcal{F}_s^+$ such that $\mathbb{1}_B = \mathbb{1}_C$ almost surely. Hence we just check this condition for $C \in \mathcal{F}_s^+$.

If $C \in \mathcal{F}_s^+ \subseteq \mathcal{F}_{s+\varepsilon} \subseteq \mathcal{F}_t$ for all $0 \leq \varepsilon \leq t-s$, so
$$\mathbb{E}[(X_t - X_{s+\varepsilon})\mathbb{1}_C] = 0.$$

It remains to show that
$$\mathbb{E}[(X_{s+\varepsilon} - X_s)\mathbb{1}_C] = 0.$$

To show this we need to introduce uniform integrability.

1.3 Uniform Integrability

Theorem 1.7 (Vitali)**.** *The following are equivalent:*

- *$X_n \to X$ in probability and (X_n) is uniformly integrable.*
- *$X_n \to X$ in L^1.*

Definition 1.9. Let χ be a collection of random variables. They are *uniformly integrable* if one of the following hold:

- $\sup_{X\in\chi} \mathbb{E}[|X|\mathbb{1}_{|X|>k}] \to 0$ as $k \to \infty$.
- $\sup_{X\in\chi} \mathbb{E}[|X|] < \infty$ and
$$\sup_{X\in\chi} \sup_{A:\mathbb{P}(A)<\delta} \mathbb{E}[|X|\mathbb{1}_A] \to 0$$
as $\delta \to 0$.
- There exists $G : \mathbb{R}_+ \to \mathbb{R}_+$ with $G(x)/x \uparrow \infty$ and
$$\sup_{X\in\chi} \mathbb{E}[G(|X|)] < \infty,$$
for example $G(x) = x^p$ for $p > 1$.

Proof: We show that $X_n \to X$ in $\mathbb{P}$ and uniformly integrable means $X_n \to X$ in L^1. First note that X_n being UI implies that X is integrable (by taking a subsequence that is almost surely convergent) and $|X_n - X| \le |X| + |X_n|$ are UI. Then:

$$\begin{aligned}\mathbb{E}[|X_n - X|] &= \mathbb{E}[|X_n - X|\mathbb{1}_{|X_n-X|>2k}] + \mathbb{E}[|X_n - X|\mathbb{1}_{|X_n-X|\le 2k}] \\ &\le \varepsilon + 2k \cdot \mathbb{P}(|X_n - X| > \delta) + \delta \to \varepsilon + \delta\end{aligned}$$

as $n \to \infty$, and we can make ε and δ arbitrarily small.

Now if $X_n \to X$ in L^1, then $X_n \to X$ in $\mathbb{P}$ and

$$\begin{aligned}\mathbb{E}[|X_n|\mathbb{1}_A] &\le \mathbb{E}[(|X| + |X_n - X|)\mathbb{1}_A] \le \mathbb{E}[|X|\mathbb{1}_A] + \mathbb{E}[|X_n - X|] \\ &\le \varepsilon + \delta,\end{aligned}$$

by choosing A small enough (as X is UI) and $n \ge N$. Since ε, δ are arbitrary and $(X_1, \dots, X_N)$ are themselves UI, the result follows.

Proposition 1.7. *Let $\mathbb{G}$ be a collection of σ-algebras, and let X be integrable. Let*

$$\mathcal{Y} = \{Y = \mathbb{E}[X|\mathcal{G}], \mathcal{G} \in \mathbb{G}\}.$$

Then $\mathcal{Y}$ is uniformly integrable.

Proof:

$$\begin{aligned}\mathbb{E}[|Y|\mathbb{1}_{|Y|>k}] &= \mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|\mathbb{1}_{|Y|>k}] \\ &\leq \mathbb{E}[\mathbb{E}[|X||\mathcal{G}]\mathbb{1}_{|Y|\geq k}] = \mathbb{E}[|X|\mathbb{1}_{|Y|\geq k}] \\ &= \mathbb{E}[|X|\mathbb{1}_{|X|\geq r,|Y|\geq k}] + \mathbb{E}[|X|\mathbb{1}_{|X|\leq r,|Y|\geq k}] \\ &\leq \mathbb{E}[|X|\mathbb{1}_{|X|\geq r}] + r\mathbb{P}(|Y|\geq k),\end{aligned}$$

where the latter term is at most

$$\frac{\mathbb{E}[|Y|]}{k} \leq \frac{\mathbb{E}[|X|]}{k} \to 0.$$

Hence we find

$$\limsup_{k\to\infty} \sup_{Y\in\mathcal{Y}} \mathbb{E}[|X|\mathbb{1}_{|Y|>k}] \leq \mathbb{E}[|X|\mathbb{1}_{|X|\geq r}] \to 0,$$

by DCT.

Now we can finish the proof from last time.

Proof: (X_t) is a martingale. Fix t. Then for $s \leq t$,

$$X_s = \mathbb{E}[X_t|\mathcal{F}_s].$$

Now $(X_{s+\varepsilon} - X_s)$ is UI, and $X_{s+\varepsilon} \to X_s$ almost-surely by assumption, so the convergence is in L^1, as desired.

1.4 Local Martingales

Remark. Our standing assumptions are that:

- all martingales are càdlàg, and
- filtrations satisfy the usual conditions (unless otherwise stated).

Definition 1.10. (X_t) is a *local martingale* if and only if there exists a sequence of stopping times $T_n \uparrow \infty$ almost surely, such that

$$(X_{t\wedge T_n} - X_0)_{t\geq 0}$$

is a martingale for all n, and X is càdlàg and adapted.

Remark. If X_0 is integrable, then (X_t) is a local martingale if and only if there exists $T_n \uparrow \infty$ such that $(X_{t\wedge T_n})_t$ is a martingale.

We write X stopped at T as
$$X^T = (X_{t\wedge T}).$$

Remark. If $\mathcal{F}_0$ is trivial, then X_0 is almost-surely constant.

Unless otherwise explicitly stated, $\mathcal{F}_0$ is trivial.

The sequence (T_n) is called the *localising sequence* for X.

Proposition 1.8. *If X is continuous and*
$$T_n = \inf\{t \geq 0 \mid |X_t| > n\},$$
then (T_n) is a localising sequence.

Proof: We need to show that
$$\mathbb{E}[X_{t\wedge T_n}|\mathcal{F}_s] = X_{s\wedge T_n}$$
for all $0 \leq s \leq t$ and n.

Since X is a local martingale, there exists a localising sequence $(U_n) \uparrow \infty$ so that X^{U_n} is a martingale.
$$\begin{aligned}\mathbb{E}[X_{t\wedge T_n}|\mathcal{F}_s] &= \mathbb{E}[\lim_k X_{t\wedge T_n\wedge U_k}|\mathcal{F}_s]\\ &= \lim_k \mathbb{E}[X_{t\wedge T_n\wedge U_k}|\mathcal{F}_s]\\ &= \lim_k X_{s\wedge T_n\wedge U_k} = X_{s\wedge T_n},\end{aligned}$$
where we use the fact $|X_{t\wedge T_n}| \leq n$ and DCT, and that stopped martingales are martingales.

Example 1.3.

Let
$$M_t = e^{W_t - t/2},$$
where W is a Brownian motion and $\mathcal{F}$ is generated by the Brownian motion with the usual conditions.

M is a martingale, and also
$$M_t = (e^{W_t/t - 1/2})^t \to 0 \qquad \text{a.s.}$$

as $t \to \infty$, since by the Brownian LLN, $W_t/t \to 0$ almost-surely.

But this means that (M_t) is not uniformly integrable, since $\mathbb{E}[|M_t|] = 1$ for all t, so (M_t) does not converge in L^1. Define instead

$$X_s = \begin{cases} M_{s/(1-s)} & 0 \le s < 1, \\ 0 & s \ge 1, \end{cases} \qquad \mathcal{G}_s = \begin{cases} \mathcal{F}_{s/(1-s)} & 0 \le s < 1, \\ \mathcal{F}_\infty & s \ge 1. \end{cases}$$

We claim that X is a local martingale with respect to $(\mathcal{G}_s)$.

1.5 Class D and Class DL

The motivation for the next section is to deduce when local martingales are martingales.

Recall the following:

Theorem 1.8 (Martingale Convergence Theorem)**.** *Let X be a martingale, and*

$$\sup_{t>0} \mathbb{E}|X_t| < \infty,$$

i.e. L^1 bounded. Then there exists integrable X_∞ such that $X_t \to X_\infty$ almost-surely.

If (X_t) is uniformly integrable, then convergence is in L^1.

If (X_t) is bounded in L^p, then convergence happens in L^p.

Definition 1.11. Let T be a stopping time with respect to filtration $(\mathcal{F}_t)$. Let

$$\mathcal{F}_T = \{A \in \mathcal{F} \mid A \cap \{T \le t\} \in \mathcal{F}_t \text{ for all } t \ge 0\}.$$

This can be shown to be a σ-algebra.

Theorem 1.9 (Optional Stopping Theorem)**.** *If X is a martingale then*

$$\mathbb{E}[X_T|\mathcal{F}_S] = X_{S\wedge T}$$

for all bounded stopping times S, T.

We also have a converse:

Proposition 1.9. *If X_t is integrable and $\mathcal{F}_t$-measurable for all $t \ge 0$, and $\mathbb{E}[X_T] = X_0$ for all bounded stopping times T, then X is a martingale.*

Proof: Fix $0 \leq s \leq t$ and event $A \in \mathcal{F}_s$. Consider event

$$T = s\mathbb{1}_A + t\mathbb{1}_{A^c}.$$

Note that $\{T \leq u\} \in \mathcal{F}_u$ for all u (check when $u < s$, $s \leq u < t$ and $t \geq u$), so this means T is a stopping time. Hence

$$\mathbb{E}[X_t] = X_0 = \mathbb{E}[X_s\mathbb{1}_A + X_t\mathbb{1}_{A^c}] = \mathbb{E}[X_t + \mathbb{1}_A(X_s - X_t)].$$

Picking $A = \emptyset$, we see $\mathbb{E}[X_t] = X_0$. Hence also we get

$$\mathbb{E}[(X_t - X_s)\mathbb{1}_A] = 0 \implies \mathbb{E}[X_t|\mathcal{F}_s] = X_s$$

by the definition of conditional expectation.

Definition 1.12. An adapted càdlàg process X is in *class (D)* if and only if

$$\{X_T \mid T \text{ a finite stopping time}\}$$

is uniformly integrable.

It is in *class (DL)* if for all $t \geq 0$,

$$\{X_{T\wedge t} \mid T \text{ a finite stopping time}\}$$

is uniformly integrable.

Theorem 1.10. *A local martingale in class (DL) is a true martingale.*

Proof: By definition, there exists stopping times $T_n \uparrow \infty$ and X^{T_n} is a true martingale.

We know that $X_{t\wedge T_n} \to X_t$ almost-surely for all $t \geq 0$. By (DL) and Vitali, $X_{t\wedge T_n} \to X_t$ in L^1. Therefore,

$$\begin{aligned}\mathbb{E}[X_t|\mathcal{F}_s] &= \mathbb{E}[\lim_n X_{t\wedge T_n}|\mathcal{F}_s] \\ &= \lim_n \mathbb{E}[X_{t\wedge T_n}|\mathcal{F}_s] \\ &= \lim_n X_{s\wedge T_n} = X_s.\end{aligned}$$

Theorem 1.11. *If X is a martingale, then X is in class (DL). If in addition X is uniformly integrable, then X is in class (D).*

Proof: For any bounded stopping time T,

$$\mathbb{E}[X_t|\mathcal{F}_T] = X_{t\wedge T}$$

by OST. We know that a collection of conditional expectations indexed by σ-algebras is uniformly integrable, hence the conclusion follows.

For the second part, we can send $t \to \infty$ using the assumption of uniform integrability.

1.6 Square-integrable Martingales

Definition 1.13. The *square-integrable martingales* are the class

$$\mathcal{M}^2 = \{X = (X_t)_{t\geq 0} \text{ a continuous martingale with } \sup_{t\geq 0} \mathbb{E}[X_t^2] < \infty\}.$$

Remark. By martingale convergence theorem, $X_t \to X_\infty$ almost-surely and in L^2, and moreover by Jensen's $t \mapsto \mathbb{E}[X_t^2]$ is increasing, so

$$\sup_{t\geq 0} \mathbb{E}[X_t^2] = \mathbb{E}[X_\infty^2].$$

From advanced probability and Doob's maximal inequality,

$$\mathbb{E}[\sup_{t\geq 0} X_t^2] \leq 4\mathbb{E}[X_\infty^2].$$

On $\mathcal{M}^2$, we define a norm by

$$\|X\|_{\mathcal{M}^2} = \mathbb{E}[|X_\infty|^2]^{1/2}.$$

Theorem 1.12. *$\mathcal{M}^2$ is complete.*

Proof: Let (X^n) be a Cauchy sequence. We can find (n_k) such that

$$\mathbb{E}[|X_\infty^{n_k} - X_\infty^{n_{k-1}}|^2] \leq 2^{-k}.$$

By Doob's maximal L^2 inequality,

$$
\begin{aligned}
\mathbb{E}\sum_{k=1}^{\infty}\sup_{t\geq 0}|X_t^{n_k}-X_t^{n_{k-1}}| &= \sum_{k=1}^{\infty}\mathbb{E}\sup_{t\geq 0}|X_t^{n_k}-X_t^{n_{k-1}}|^2 \\
&\overset{\text{Jensen}}{\leq} \sum_{k=1}^{\infty}\sqrt{\mathbb{E}\sup_{t\geq 0}|X_t^{n_k}-X_t^{n_{k-1}}|} \\
&\overset{\text{Doob}}{\leq} \sum_{k=1}^{\infty}2\sqrt{\mathbb{E}|X_\infty^{n_k}-X_\infty^{n_{k-1}}|^2} \\
&\leq \sum_{k=1}^{\infty}2^{1-k/2}<\infty.
\end{aligned}
$$

Hence,

$$\sum_{k=1}^{\infty}\sup_t|X_t^{n_k}-X_t^{n_{k-1}}|<\infty \qquad \text{a.s.}$$

Therefore,

$$X^{n_k}=X_0+\sum_{i=1}^{k}(X^{n_i}-X^{n_{i-1}})$$

converges uniformly and almost-surely. Since X^{n_k} is continuous, so is the limit X.

To show that X is a martingale,

$$X_\infty^n\to X_\infty$$

in L^2, by the completeness of L^2. So,

$$\mathbb{E}[X_\infty|\mathcal{F}_t]=\lim\mathbb{E}[X_\infty^n|\mathcal{F}_t]=\lim X_t^n=X_t.$$

1.7 Quadratic Integration

If Z is a finite variation process, and

$$\int_0^t|H|\,|\mathrm{d}Z|<\infty \qquad \text{a.s. for all } t,$$

then we can define

$$\int_0^t H\,\mathrm{d}Z.$$

We will show that if Z is a continuous local marginale, and

$$\int_0^t |H|^2 \, |\mathrm{d}Z|^2 < \infty \qquad \text{a.s. for all } t,$$

then we can define

$$\int H \, \mathrm{d}Z.$$

We need to define what $|\mathrm{d}Z|^2$ is however.

Proposition 1.10. *Let M be a martingale, and K a bounded, $\mathcal{F}_{t_0}$-measurable random variable. Then*

$$X_t = K(M_t - M_{t\wedge t_0})$$

is a martingale.

Proof: Let T be a bounded stopping time. Then

$$\begin{aligned} \mathbb{E}[X_T] &= \mathbb{E}[K(M_T - M_{T\wedge t_0})] \\ &= \mathbb{E}\mathbb{E}[K(M_T - M_{T\wedge t_0})|\mathcal{F}_{t_0}] \\ &= \mathbb{E}[K\mathbb{E}[M_T - M_{T\wedge t_0}|\mathcal{F}_{t_0}]] = 0, \end{aligned}$$

by optional stopping theorem.

We could also do it by hand, computing $\mathbb{E}[X_t|\mathcal{F}_s]$.

Proposition 1.11 (Pythagorean Theorem)**.** *If $X \in \mathcal{M}^2$ and (t_n) is increasing, then*

$$\mathbb{E}[X_\infty^2] = \mathbb{E}[X_{t_0}^2] + \sum_{n=1}^{\infty} \mathbb{E}[(X_{t_n} - X_{t_{n-1}})^2].$$

The proof is for the example sheet.

Definition 1.14. A sequence on càdlàg processes Z^n converges *uniformly on compacts in probability* if and only if

$$\mathbb{P}\left(\sup_{0\le s\le t} |Z_s^n - Z_s| > \varepsilon \right) \to 0,$$

for all $t \ge 0$ and $\varepsilon \ge 0$.

This is also known as UCP convergence.

Theorem 1.13 (Existence of Quadratic Variation)**.** *Let X be a continuous, local martingale. Let*

$$[X]_t^n = \sum_{k=1}^{\infty} (X_{t \wedge t_k^n} - X_{t \wedge t_{k-1}^n})^2,$$

where $t_k^n = k2^{-n}$.

Then there exists a continuous, increasing, adapted process $[X]$ such that

$$[X]^n \to [X] \qquad \textit{ucp.}$$

Proof: First, there is no loss assuming $X_0 = 0$.

We begin by considering the case when X is uniformly bounded, so there exists $C > 0$ such that

$$|X_t(\omega)| < C$$

for all t and ω. Since $X \in \mathcal{M}^2$, the limit X_∞ exists, and for $t_k^n \leq t < t_{k+1}^n$,

$$[X]_t^n - [X]_{t_k^n}^n = (X_t - X_{t_k^n})^2 \to 0$$

as $t \to \infty$, since both converge to X_∞ which exists. Hence

$$[X]_\infty^n = \sup_k [X]_{t_k^n}^n.$$

To show this is a finite random variable, note

$$\begin{aligned}
\mathbb{E}[X]_\infty^n &= \mathbb{E}\left[\sum_{k=1}^{\infty} (X_{t_k^n} - X_{t_{k-1}^n})^2\right] \\
&= \mathbb{E}[X_\infty^2] < C^2,
\end{aligned}$$

so $[X]_\infty^n < \infty$ almost-surely.

We will define

$$M_t^n = \frac{1}{2}(X_t^2 - [X]_t^n).$$

We can analogously write

$$M_t^n = \sum_{k=1}^{n} X_{t_{k-1}^n} \left(X_{t_k^n \wedge t} - X_{t_{k-1}^n \wedge t} \right).$$

This is a continuous martingale, as it is a martingale transform, and there are only a finite number of terms. Moreover,

$$\begin{aligned}\mathbb{E}[(M_t^n)^2] &= \sum_{k=1}^{\infty} \mathbb{E}\left[(X_{t_{k-1}}^n)^2 \left(X_{t_k^n \wedge t} - X_{t_{k-1}^n \wedge t}\right)\right] \\ &\leq C^2 \mathbb{E}[[X_t]^n] \leq C^4,\end{aligned}$$

by Pythagoras and boundedness. Hence $M^n \in \mathcal{M}^2$. We want to take the limit, so we check that (M^n) is Cauchy. If $n > m$, then

$$\begin{aligned}M_\infty^n - M_\infty^m &= \sum_{j=1}^{\infty} \left(X_{j2^{-n}} - X_{\lfloor j2^{-(n-m)} \rfloor 2^{-m}}\right)\left(X_{(j+1)2^{-n}} - X_{j2^{-n}}\right) \\ \mathbb{E}[(M_\infty^n - M_\infty^m)^2] &= \sum_j \mathbb{E}\left[(X_{j2^{-n}} - X_{i2^{-m}})^2 \left(X_{t_{j-1}^n} - X_{t_j^n}\right)^2\right] \\ &\leq \mathbb{E}\left[\sup_{|t-s| \leq 2^{-m}} (X_s - X_t)^2 [X]_\infty^n\right] \\ &\leq \mathbb{E}\left[\sup_{|s-t| \leq 2^{-m}} (X_s - X_t)^4\right]^{1/4} \mathbb{E}[([X]_\infty^n)^2]^{1/2}.\end{aligned}$$

We need to show both of these terms are finite. Since X is continuous, $X_t \to X_\infty$, X is uniformly continuous, hence

$$\sup_{|s-t| \leq 2^{-m}} (X_s - X_t)^4 \to 0 \qquad \text{a.s.}$$

Moreover it is bounded by $16C^4$, so it goes to 0 by dominated convergence theorem. For the latter term,

$$\begin{aligned}([X]_\infty^n)^2 &= (X_\infty^2 - 2M_\infty^n)^2 \leq 2X_\infty^4 + 8(M_\infty^n)^2, \\ \mathbb{E}([X]_\infty^n)^2 &\leq 2C^4 + 8C^4 = 10C^4,\end{aligned}$$

for all n. Hence $M^n \to M^* \in \mathcal{M}^2$. Define

$$[X] = X^2 - 2M^*.$$

Then $[X]$ is continuous, since the right hand side is. Moreover

$$\begin{aligned}\mathbb{E} \sup_{t \geq 0} ([X]_t^n - [X]_t)^2 &= 4\mathbb{E} \sup_{t \geq 0} (M_t^n - M_t^*)^2 \\ &\leq 16\mathbb{E}(M_\infty^n - M_\infty^*)^2 \to 0,\end{aligned}$$

so $[X]^n \to [X]$ uniformly in L^2. It has uniform almost-sure convergence for some subsequence (n_k), and

$$[X]_t^n - [X]_{t_k^n}^n = (X_t - X_{t_k^n})^2 \overset{n\to\infty}{\to} 0.$$

If $s < t$, then

$$[X]_s = \lim [X]_{t_i^n}^n \leq \lim [X]_{t_k^n}^n = [X]_t,$$

where $t_i^n = \lfloor 2^n s \rfloor 2^{-n}$. This shows the properties when X is uniformly bounded.

In the general case, consider X a continuous local martingale, with $X_0 = 0$. Let

$$T_N = \inf\{t \geq 0 \mid |X_t| > N\}.$$

Then X^{T_N} is a bounded continuous martingale, and

$$[X^{T_N}]^n \to [X^{T_N}]$$

uniformly in L^2. Moreover,

$$[X^{T_{N+1}}]_t^n - [X^{T_N}]_t^n = \begin{cases} 0 & t \leq T_N, \\ \geq 0 & t > T_N. \end{cases}$$

Hence,

$$[X^{T_{N+1}}]_t \geq [X^{T_N}]_t \qquad \text{a.s.}$$

for all N and t, and so we can define

$$[X]_t = \sup_N [X^{T_N}]_t.$$

This is adapted, increasing and since

$$[X]_t = [X^{T_N}]_t \text{ on } \{t \leq T_N\},$$

$[X]$ is continuous almost-surely. Moreover $[X]^{T_N} = [X^{T_N}]$. Finally,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0 \leq s \leq t} |[X]_s^n - [X]_s| > \varepsilon\right) \\
&\qquad \leq \mathbb{P}\left(\sup_{0 \leq s \leq t} \left|[X^{T_N}]_s^n - [X^{T_N}]_s\right| > \varepsilon, t \leq T_N\right) + \mathbb{P}(t > T_N) \\
&\qquad \leq \frac{1}{\varepsilon^2} \mathbb{E} \sup_{t \geq 0} \left([X^{T_N}]_t^n - [X^{T_N}]_t\right)^2 + \mathbb{P}(t > T_N) \to 0,
\end{aligned}$$

as $N \to \infty$ and $n \to \infty$.

Proposition 1.12. *Let X be a continuous local martingale of finite variation. Then $X_t = X_0$ for all $t \geq 0$.*

Proof: Pick a subsequence so that

$$[X]_t^n \to [X]_t \qquad \text{a.s.},$$

but this is equal to

$$\sum (X_{t\wedge t_k^n} - X_{t\wedge t_{k-1}^n})^2 \leq \sup_{|r-s|\leq 2^{-n}} |X_r - X_s| \sum |X_{t\wedge t_k^n} - X_{t\wedge t_{k-1}^n}|$$

and the first term goes to 0 by uniform continuity of X, and the latter is at most $V_X(t)$.

We know that if X is a bounded continuous local martingale, then $X^2 - [X] \in \mathcal{M}^2$. If not necessarily bounded, then $X^2 - [X]$ is a local martingale, from the proof that we gave.

We also know that if X is a continuous local martingale of finite variation, then $[X] = 0$.

Proposition 1.13. *If X is a continuous local martingale with $[X] = 0$, then $X_t = X_0$ for all t.*

Proof: X and X^2 are both local martingales.

Let (T_n) reduce X (and X^2) to a bounded martingale. Then

$$\mathbb{E}[(X_{t\wedge T_n} - X_0)^2] = X_0^2 - 2X_0\mathbb{E}[X_{t\wedge T_n}] + \mathbb{E}[X_{t\wedge T_n}^2] = X_0^2 - 2X_0^2 + X_0^2 = 0.$$

So $X_{t\wedge T_n} = X_0$ almost surely for all t and n. Taking $n \to \infty$ we get the result.

1.8 Characterisation of Quadratic Variation

Proposition 1.14. *Let X be a continuous local martingale and A a finite variation adapted continuous process, with $A_0 = 0$.*

If $X^2 - A$ is a local martingale, then $A = [X]$

Proof: Note that $(X^2 - [X]) - (X^2 - A) = A - [X]$ is a local martingale of finite variation, so

$$A_t - [X]_t = 0.$$

Proposition 1.15. *Let W be a Brownian motion. Then*

$$[W]_t = t$$

for all t.

Proof: $W_t^2 - t$ is a martingale, hence the quadratic variation is t.

Remark. Fix $t \geq 0$, and let $L : C[0,t] \to \mathbb{R}$ be linear and bounded (i.e. continuous). Then

$$Lg = \int_0^t g \, \mathrm{d}f$$

for some finite variation f.

This is Riesz-Markov, and was proven in functional analysis. This is quite annoying given we want to be integrating over, for example Brownian motion which is infinite variation.

2 The Stochastic Integral

We let

$$\mathcal{M}_{\text{loc}} = \{\text{continuous local martingales}\},$$

the integrators.

Definition 2.1. The *simple previsible processes* are H of the form

$$H = \sum_{k=1}^{n} H_{t_k} \mathbb{1}_{(t_{k-1},t_k]}$$

for some $0 \le t_0 < \cdots < t_n$, and H_{t_k} is $\mathcal{F}_{t_{k-1}}$-measurable and bounded.

The set of simple previsible processes is $\mathcal{S}$.

These are the integrands.

Definition 2.2. For $X \in \mathcal{M}_{\text{loc}}$ and $H \in \mathcal{S}$,

$$\int_0^t H \,\mathrm{d}X = \sum_{k=1}^{n} H_{t_k}(X_{t_k \wedge t} - X_{t_{k-1} \wedge t}).$$

Remark. We know

$$\int H \,\mathrm{d}X \in \mathcal{M}_{\text{loc}}.$$

Proposition 2.1. *The quadratic variation satisfies*

$$\left[\int H \,\mathrm{d}X\right] = \int H^2 \,\mathrm{d}[X]$$

for $H \in \mathcal{S}$ and $X \in \mathcal{M}_{\text{loc}}$, where

$$\int_0^t H^2 \,\mathrm{d}[X] = \sum_{k=1}^{n} H_{t_k}^2([X]_{t_k \wedge t} - [X]_{t_{k-1} \wedge t}).$$

Proof: There are many methods. One is to show that

$$\left(\sum H_{t_k}(X_{t \wedge t_k} - X_{t \wedge t_{k-1}})\right)^2 - \sum H_{t_k}^2([X]_{t \wedge t_k} - [X_{t_{k-1} \wedge t}])$$

is a local martingale. By localisation, we can assume that X is bounded, so we may show that it is a true martingale.

Pick a bounded stopping time T. Then by the Pythagorean theorem,

$$\mathbb{E}\left(\sum H_{t_k}(X_{T\wedge t_k} - X_{T\wedge t_{k-1}})\right)^2 = \mathbb{E}\left(\sum H_{t_k}^2 (X_{T\wedge t_k} - X_{T\wedge t_{k-1}})^2\right).$$

We are done since $(X^{t_k} - X^{t_{k-1}})^2 - ([X]^{t_k} - [X]^{t_{k-1}})$ is a martingale, implying that this is

$$\mathbb{E}\sum H_{t_k}^2([X]_{T\wedge t_k} - [X]_{T\wedge t_{k-1}}).$$

2.1 Itô's Isometry

Proposition 2.2. *If $H \in \mathcal{S}$ and $X \in \mathcal{M}^2$, then*

$$\mathbb{E}\left(\int_0^\infty H \,\mathrm{d}X\right)^2 = \mathbb{E}\left(\int_0^\infty H^2 \,\mathrm{d}[X]\right).$$

To prove this, we need to strengthen our initial observations.

Proposition 2.3. *If $X \in \mathcal{M}^2$, then $X^2 - [X]$ is a uniformly integrable martingale. In particular,*

$$\mathbb{E}[X_\infty^2] = X_0^2 + \mathbb{E}[X]_\infty.$$

This proves Itô, since if $X \in \mathcal{M}^2$ and $H \in \mathcal{S}$, then $\int H \,\mathrm{d}X \in \mathcal{M}^2$.

Proof: Let (T_n) reduce X to a bounded martingale. Then

$$\mathbb{E}[X_{T_n}^2] = X_0^2 + \mathbb{E}[X]_{T_n}.$$

Now $X_{T_n} \to X_\infty$ in L^2 by the martingale convergence theorem, which implies $\mathbb{E}X_{T_n}^2 \to \mathbb{E}X_\infty^2$. The other term converges by the monotone convergence theorem.

We now want to define $\int H \,\mathrm{d}X$ when H is previsible and $\int H^2 \,\mathrm{d}[X] < \infty$ for all t. Our first step is to consider when $X \in \mathcal{M}^2$ and

$$\mathbb{E}\left[\int_0^\infty H^2 \,\mathrm{d}[X]\right] < \infty.$$

Definition 2.3. Given $X \in \mathcal{M}^2$, let

$$L^2(X) = \left\{H \mid H \text{ previsible, } \mathbb{E}\left[\int_0^\infty H^2 \,\mathrm{d}[X]\right] < \infty\right\}.$$

Note $L^2(X) \to L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, \mu)$ where

$$\mu((s,t] \times A) = \mathbb{E}[\mathbb{1}_A([X]_t - [X]_s)].$$

This is an L^2 space, and from Itô's isometry, note that for $H \in \mathcal{S}$,

$$\|H\|_{L^2(X)} = \left\| \int H \,\mathrm{d}X \right\|_{\mathcal{M}^2}.$$

Proposition 2.4. *If $H^n \to H$ in $L^2(X)$ then there is $M \in \mathcal{M}^2$ such that*

$$\int H^n \,\mathrm{d}X \to M$$

in $\mathcal{M}^2$. Moreover, if $\tilde{H}^n \to H$ in $L^2(X)$, then

$$\int \tilde{H}^n \,\mathrm{d}X \to M$$

in $\mathcal{M}^2$ as well.

Proof: (H^n) is Cauchy in $L^2(X)$. By Itô's isometry, $(\int H^n \,\mathrm{d}X)$ is Cauchy in $\mathcal{M}^2$.

So we are done by the completeness of $\mathcal{M}^2$. For uniqueness, say that $\int \tilde{H}^n \,\mathrm{d}X \to \tilde{M}$. Then

$$\begin{aligned}\|M - \tilde{M}\|_{\mathcal{M}^2} &\le \left\| M - \int H^n \,\mathrm{d}X \right\| + \left\| \tilde{M} - \int \tilde{H}^n \,\mathrm{d}X \right\| + \left\| \int (H^n - \tilde{H}^n) \,\mathrm{d}X \right\| \\ &\le \|H^n - \tilde{H}^n\|_{L^2(X)} \to 0.\end{aligned}$$

We claim that $\mathcal{S}$ is dense in $L^2(X)$.

Definition 2.4. For $X \in \mathcal{M}^2$ and $H \in L^2(X)$, $\int H \,\mathrm{d}X$ is defined to be the $\mathcal{M}^2$ limit of $(\int H^n \,\mathrm{d}X)$, where (H^n) is any sequence in S converging to H in $L^2(X)$.

Theorem 2.1. *Let $X \in \mathcal{M}^2$, $H \in L^2(X)$ and T be a stopping time. Then $X^T \in \mathcal{M}^2$, $H \in L^2(X^T)$, $H\mathbb{1}_{(0,T]} \in L^2(X)$ and*

$$\int H \mathbb{1}_{(0,T]} \,\mathrm{d}X = \int H \,\mathrm{d}X^T = \left(\int H \,\mathrm{d}X \right)^T.$$

Proof: $X^T \in \mathcal{M}^2$ since

$$\mathbb{E}[X_T^2] \leq \mathbb{E}[\sup_t X_t^2] \leq 4\mathbb{E}[X_\infty^2] < \infty.$$

$H \in L^2(X^T)$ as

$$\begin{aligned}\mathbb{E} \int H^2 \,\mathrm{d}[X^T] = \mathbb{E} \int H^2 \,\mathrm{d}[X]^T &\overset{\text{pointwise}}{=} \mathbb{E} \int_0^T H^2 \,\mathrm{d}[X] \\ &\leq \mathbb{E} \int_0^\infty H^2 \,\mathrm{d}[X] < \infty.\end{aligned}$$

To show $H\mathbb{1}_{(0,T]} \in L^2(X)$, we first show that $\mathbb{1}_{(0,T]}$ is previsible. If T takes the values $t_1, t_2, \ldots, t_n$, then

$$\mathbb{1}_{(0,T]} = \sum_{k=1}^n \mathbb{1}_{\{T=t_k\}} \mathbb{1}_{(0,t_k]} = \sum_{k=1}^n \mathbb{1}_{\{T>t_{k-1}\}} \mathbb{1}_{(t_{k-1},t_k]},$$

where $t_0 = 0$. In general, let

$$T^n = (\lceil 2^n T \rceil 2^{-n}) \wedge n,$$

so T^n takes only a finite number of values, and is a stopping time, and

$$\mathbb{1}_{(0,T]} = \lim_n \mathbb{1}_{(0,T^n]}.$$

So $H\mathbb{1}_{(0,T]}$ is previsible, and

$$\int (H\mathbb{1}_{(0,T]})^2 \,\mathrm{d}[X] = \int_0^T H^2 \,\mathrm{d}[X] < \infty.$$

To show the equality of integrals, note that

$$\int H \,\mathrm{d}X^T = \left(\int H \,\mathrm{d}X\right)^T$$

is manifestly true if $H \in \mathcal{S}$ by looking at the formula. Let $H^n \to H$ in $L^2(X)$. Then

$$\begin{aligned}&\left\| \int H \,\mathrm{d}X^T - \left(\int H \,\mathrm{d}X\right)^T \right\|_{\mathcal{M}^2} \\ &\qquad \leq \left\| \int (H - H^n) \,\mathrm{d}X^T \right\| + \left\| \left(\int (H - H^n) \,\mathrm{d}X\right)^T \right\| \\ &\qquad = \|H - H^n\|_{L^2(X^T)} + 2\|H - H^n\|_{L^2(X)} \to 0,\end{aligned}$$

where the last inequality is by Doob and Itô. The other formulas are proven analogously.

Proposition 2.5. *If $X \in \mathcal{M}^2$, and $H \in L^2(X)$, and S, T are stopping times with $0 \le S \le T$ almost surely, then*

$$\int_0^t H \mathbb{1}_{(0,s]} \, \mathrm{d}X^S = \int_0^t H \mathbb{1}_{(0,T]} \, \mathrm{d}X^T$$

on the event $\{t \le S\}$.

Proof: The left hand side is

$$\left(\int H \, \mathrm{d}X \right)_{t \wedge S}.$$

The right hand side is

$$\left(\int H \, \mathrm{d}X \right)_{t \wedge T}.$$

They agree on $\{t \le S\}$.

Proposition 2.6. *Let $X \in \mathcal{M}_{\mathrm{loc}}$, H previsible, and*

$$\int_0^t H^2 \, \mathrm{d}[X] < \infty \qquad a.s.$$

for all t. Let

$$T_n = \inf\{t \ge 0 \mid |X_t| > n \text{ or } \int_0^t H^2 \, \mathrm{d}[X] > n\}.$$

Then $X^{T_n} \in \mathcal{M}^2$, and $H \mathbb{1}_{(0,T_n]} \in L^2(X^{T_n})$. Let

$$M^{(n)} = \int H \mathbb{1}_{(0,T_n)} \, \mathrm{d}X^{T_n}.$$

Then there is a continuous local martingale M such that $M^{(n)} \to M$ UCP.

Proof: From the previous proposition, on the event $\{t \le T_n\}$, we have $M_t^{(n)} = M_t^{(N)}$ for all $N \ge n$, because T_n is increasing and localizing.

Hence the sequence $M_t^{(n)}$ converges almost-surely to a random variable M_t^*, for every t. Moreover this convergence is uniformly almost-surely on $[0, t]$, as the sequence is eventually constant, and

$$\mathbb{P}\left(\sup_{0 \le s \le t} |M_s^n - M_s^*| > \varepsilon \right) \le \mathbb{P}(T_n > t) \to 0.$$

Definition 2.5. For X, H as before, define

$$\int H \,\mathrm{d}X = M^*.$$

Remark.

- M^* is a local martingale as $(M^*)^{T_n} = M^{(n)}$.
- The choice of sequence (T_n) is arbitrary: if $(U_n) \uparrow \infty$ are stopping times such that $X^{U_n \in \mathcal{M}^2}$ and $H\mathbb{1}_{(0,U_n]} \in L^2(X^{U_n})$, then

$$\int Y \mathbb{1}_{(0,U_n)} \,\mathrm{d}X^{U_n} \to \int Y \,\mathrm{d}X.$$

This is just because

$$\left(Y \mathbb{1}_{(0,U_n]} \,\mathrm{d}X^{U_n}\right)^{T_m} = \left(Y \mathbb{1}_{(0,T_m]} \,\mathrm{d}X^{T_m}\right)^{U_n} = \left(\int Y \,\mathrm{d}X\right)^{T_m \wedge U_n}.$$

2.2 Semimartingales

Definition 2.6. A continuous *semimartingale* is

$$Z_t = Z_0 + A_t + M_t$$

for all t, where Z_0 is constant, A is continuous of finite variation, and M is a continuous local martingale with $A_0 = M_0 = 0$.

Proposition 2.7. *The semimartingale decomposition is unique.*

Proof: Suppose that $Z_0 + A + M = Z_0 + A' + M'$.

Then $A - A' = M' - M$. The left hand side is of finite variation, and the right hand side is a continuous local martingale. Hence, This must be constant in t, so

$$M'_t - M_t = 0 = A_t - A'_t.$$

Definition 2.7. Let H be previsible, and Z a continuous semimartingale. We say that H is Z-integrable if and only if

$$\int_0^t |H||\mathrm{d}A| < \infty \qquad \text{a.s. for all } t,$$

$$\int_0^t H^2 \,\mathrm{d}[M] < \infty \qquad \text{a.s. for all } t,$$

where $Z = Z_0 + A + M$. Then define

$$\int H \, \mathrm{d}Z = \int H \, \mathrm{d}A + \int H \, \mathrm{d}M.$$

Definition 2.8. H previsible is *locally bounded* if and only if there exists stopping times $T_n \uparrow \infty$ such that

$$|H_t(\omega) \mathbb{1}_{\{t \le T_n(\omega)\}}| \le C_n$$

for all (t, ω) and all n.

Remark. If H is locally bounded then it is Z integrable for all continuous semimartingales Z.

Remark. If H is previsible and continuous, then it is locally bounded.

Remark. Bichteler-Dellacherie tells us that, if Z is a continuous adapted process and

$$\int_0^\infty H^n \, \mathrm{d}Z \to 0$$

in probability for any sequence $H^n \in \mathcal{S}$ such that

$$\left\| \sup_{t \ge 0} |H_t^n| \right\|_\infty \to 0,$$

then Z is a semimartingale.

Definition 2.9. Suppose Z is a semimartingale of the form $Z = Z_0 + A + M$. Then the quadratic variation of Z is the quadratic variation of M, i.e. $[Z] = [M]$.

We can show that if Z is a semimartingale, then

$$[Z]_t^n = \sum_k (Z_{t \wedge t_k^n} - Z_{t \wedge t_{k-1}^n})^2 \overset{\text{UCP}}{\to} [Z].$$

Moreover, if Z is a semimartingale and H is Z-integrable, then

$$\left[\int H \, \mathrm{d}Z \right] = \int H^2 \, \mathrm{d}[Z].$$

Definition 2.10. Let X and Y be continuous semimartingales. The *quadratic covariation* $[X, Y]$ is defined as

$$[X, Y] = \frac{1}{4} \left([X + Y] - [X - Y] \right).$$

Proposition 2.8.

- $[X,Y]$ *is a continuous finite variation process.*
- $[X,Y]=[Y,X]$.
- *Quadratic covariation is bilinear.*
- *If* X *is finite variation, then* $[X,Y]=0$.

2.3 Itô's Formula

Our goal is to find a formula for $f(X^1,\dots,X^n)$, where $X^1,\dots,X^n$ are continuous semimartingales and $f:\mathbb{R}^n\to\mathbb{R}$ is C^2.

In this setting, $f(X^1,\dots,X^n)$ is a semimartingale with decomposition

$$\begin{aligned} f(X_t^1,\dots,X_t^n) = f(X_0) &+ \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x^i}(X_s)\,\mathrm{d}X_s^i \\ &+ \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)\,\mathrm{d}[X^i,X^j]_s. \end{aligned}$$

It is an exercise to show that if Z is a continuous semimartingale, then

$$[Z]_t^n = \sum_{k=1}^\infty (Z_{t\wedge t_k^n} - Z_{t\wedge t_{k-1}^n})^2 \overset{\text{UCP}}{\to} [Z]_t.$$

By polarization,

$$[X,Y] = \frac{1}{4}([X+Y]-[X-Y]) = \sum_{k=1}^\infty (X_{t\wedge t_k^n} - X_{t\wedge t_{k-1}^n})(Y_{t\wedge t_k^n} - Y_{t\wedge t_{k-1}^n}).$$

From this form, $[\cdot,\cdot]$ is bilinear, and

$$[X,Y]_t \le \sqrt{[X]_t[Y]_t}.$$

Corollary 2.1. *If* X *is finite variation, then* $[X,Y]=0$.

So if there is a term in the Itô expansion with finite variation, we can remove its covariation contributions.

Proposition 2.9. $[X,Y]$ *is the unique continuous finite variation process* A *with* $A_0=0$ *such that* $XY-A$ *is a local martingale, when* X *and* Y *are both local martingales.*

Note by the above corollary, the quadratic covariation only cares about the local martingale parts of X and Y, hence why we need X and Y to be local martingales in the proposition.

Proof: Suppose $XY - A = M$ is a local martingale. Then

$$\begin{aligned} A = XY - M &= \frac{1}{4}((X+Y)^2 - (X-Y)^2) \\ &\quad - \frac{1}{4}([X+Y] - [X-Y]) + [X,Y] - M. \end{aligned}$$

So $A - [X,Y]$ is a local martingale. But the left hand side is of finite variation, hence constant.

Theorem 2.2 (Kunita-Watanabe's Identity)**.** *Let X, Y be continuous semimartingales, and let H be locally X-integrable. Then H is locally $[X,Y]$-integrable and*

$$\left[\int H \,\mathrm{d}X, Y\right] = \int H \,\mathrm{d}[X,Y].$$

Proof: Since $[\cdot,\cdot]$ is zero for finite variation processes, we can assume $X, Y \in \mathcal{M}_{\mathrm{loc}}$.

Through localization, assume $X, Y \in \mathcal{M}^2$ and $H \in L^2(X)$. We need to show that

$$\left(\int H \,\mathrm{d}X\right)(Y) - \int H \,\mathrm{d}[X,Y]$$

is a martingale, by uniqueness of finite variation processes such that $MN - A$ is a martingale. Note that $\int H \,\mathrm{d}[X,Y]$ is finite variation by polarization. We use the converse of optional stopping theorem. We need to show that for all bounded stopping times T that

$$\mathbb{E}\left[(H \,\mathrm{d}X)_T Y_T\right] = \mathbb{E}\left[\int_0^T H \,\mathrm{d}[X,Y]\right].$$

The left hand side is

$$\mathbb{E}\left[\int_0^\infty H \,\mathrm{d}X^T Y_\infty^T\right],$$

and the right hand side is

$$\mathbb{E}\left[\int_0^\infty H \,\mathrm{d}[X^T, Y^T]\right].$$

So we can drop the T and show

$$\mathbb{E}\left[\left(\int_0^\infty H \,\mathrm{d}X\right) Y\right] = \mathbb{E}\left[\int_0^\infty H \,\mathrm{d}[X,Y]\right].$$

Now this identity we can prove by hand for simple processes $H \in \mathcal{S}$. Let $H = \mathbb{1}_{(t_0,t_1]}K$ where K is $\mathcal{F}_{t_0}$-measurable.

Plugging in this term, the left hand side is

$$\mathbb{E}[K(X_{t_1} - X_{t_0})Y_\infty],$$

and the right hand side is

$$\mathbb{E}[K([X,Y]_{t_1} - [X,Y]_{t_0})].$$

Expanding the LHS,

$$\begin{aligned} \text{LHS} &= \mathbb{E}[\mathbb{E}[K(X_{t_1} - X_{t_0})|\mathcal{F}_{t_1}]] \\ &= \mathbb{E}[K(X_{t_1} - X_{t_0})\mathbb{E}[Y_\infty|\mathcal{F}_{t_1}]] \\ &= \mathbb{E}[K(X_{t_1} - X_{t_0})(Y_{t_1} - Y_{t_0} + Y_{t_0})], \end{aligned}$$

so the left hand side minus the right hand side is

$$\mathbb{E}\left[K\big((X_{t_1} - X_{t_0})Y_{t_0} + (X_{t_1} - X_{t_0})(Y_{t_1} - Y_{t_0}) - ([X,Y]_{t_1} - [X,Y]_{t_0})\big)\right] = 0,$$

as these are all increments of martingales, and K, Y_{t_0}, and X_{t_0} are all $\mathcal{F}_{t_0}$-measurable.

So this is true for all $H \in \mathcal{S}$. For general $H \in L^2(X)$, take $H^n \to H$ where H^n are simple. Then the formula will still hold by dominated convergence theorem and Cauchy-Schwarz.

Suppose that X is a continuous semimartingale, A is locally X-integrable and B is locally $\int A \,\mathrm{d}X$-integrable. Then AB is locally X-integrable, and

$$\int B \left(\int A \,\mathrm{d}X\right) = \int AB \,\mathrm{d}X.$$

In our new notation, we will drop the integrals and show

$$\mathrm{d}f(X) = \sum_I \frac{\partial f}{\partial x^i}\,\mathrm{d}X^i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x^i \partial x^j}\,\mathrm{d}[X^i, X^j].$$

This just looks like a second order Taylor expansion.

Proposition 2.10. *Suppose Y is locally bounded, adapted, left-continuous (hence* *previsible) and X is a continuous semimartingale. Then*

$$\sum_{k=1}^{\infty} Y_{t_{k-1}^n}(X_{t_k^n \wedge t} - X_{t_{k-1}^n \wedge t}) \overset{\text{UCP}}{\to} \int_0^t Y \, \mathrm{d}X.$$

Remark. For Y a continuous semimartingale, if we instead chose the right endpoint,

$$\sum Y_{t_k^n}(X_{t_k^n \wedge t} - X_{t_{k-1}^n \wedge t}) \to \int_0^t Y \, \mathrm{d}X + [X, Y]_t.$$

Proof: Note that

$$Y^n = \sum_{k=1}^{4^n} Y_{t_{k-1}^n} \mathbb{1}_{(t_{k-1}^n, t_k^n]}$$

are simple, and $Y_t^n(\omega) \to Y_t(\omega)$ by left-continuity. Moreover the left hand side converges to

$$\int Y^n \, \mathrm{d}X$$

for $t < 2^n$. By localization, we can assume that Y is locally bounded, and $X = X_0 + M + A$ where M is a square-integrable martingale and A is of bounded variation.

Therefore $Y^n \to Y$ in $L^2(X)$, by dominated convergence theorem, and in $L^1(|\mathrm{d}A| \times \mathbb{P})$, also by DCT. Therefore

$$\int Y^n \, \mathrm{d}X = \int Y^n \, \mathrm{d}M + \int Y^n \, \mathrm{d}A \to \int Y \, \mathrm{d}M + \int Y \, \mathrm{d}A.$$

Now we start to prove Itô. We begin with the special case when $f(x, y) = xy$.

Proposition 2.11. *If X and Y are continuous semimartingales, then*

$$\mathrm{d}(XY) = X \, \mathrm{d}Y + Y \, \mathrm{d}X + \mathrm{d}[X, Y].$$

Proof: Let $X = Y$. We know that

$$\begin{aligned}2\int_0^t X\,\mathrm{d}X &= 2\lim\sum_k X_{t_{k-1}^n}(X_{t_k^n\wedge t} - X_{t_{k-1}^n\wedge t}) \\ &= \lim\sum_k \left(X_{t_k^n\wedge t}^2 - X_{t_{k-1}^n\wedge t}^2 - (X_{t\wedge t_k^n} - X_{t\wedge t_{k-1}^n})^2\right) \\ &= X_t^2 - X_0^2 - [X]_t.\end{aligned}$$

In the general case, we know that

$$\begin{aligned}2(X+Y)\,\mathrm{d}(X+Y) &= \mathrm{d}(X+Y)^2 - \mathrm{d}[X+Y], \\ 2(X-Y)\,\mathrm{d}(X-Y) &= \mathrm{d}(X-Y)^2 - \mathrm{d}[X-Y].\end{aligned}$$

We can the subtract and divide by 4, and use polarisation.

We will now prove the formula for monomials.

Proposition 2.12. *For integers $m \geq 1$,*

$$\mathrm{d}(X^m) = mX^{m-1}\,\mathrm{d}X + \frac{m(m-1)}{2}X^{m-2}\,\mathrm{d}[X].$$

Proof: Proof by induction, by using the product formula. This is true for $m = 2$ as we have shown. Then

$$\mathrm{d}(X^{m+1}) = \mathrm{d}(X^m X) = X\,\mathrm{d}(X^m) + X^m\,\mathrm{d}X + \mathrm{d}[X^m, X].$$

The first term is, by induction,

$$mX^m\,\mathrm{d}X + \frac{m(m-1)}{2}X^{m-1}\,\mathrm{d}[X].$$

This is by the 'chain rule' for integration. The third term is

$$\begin{aligned}[X^m, X] &= \left[X_0^m + \int mX^{m-1}\,\mathrm{d}X + \frac{1}{2}m(m-1)\int X^{m-2}\,\mathrm{d}[X], X\right] \\ &\overset{\text{Kunita-Watanabe}}{=} \int mX^{m-1}\,\mathrm{d}[X],\end{aligned}$$

since the other two terms are finite variation, so the covariation is 0. Summing everything up gives what is required:

$$\mathrm{d}(X^{m+1}) = (m+1)X^m\,\mathrm{d}X + \frac{m(m+1)}{2}X^{m-1}\,\mathrm{d}[X].$$

This shows Itô's formula is true for polynomials.

Theorem 2.3 (Itô's formula for $n = 1$)**.** *Let $f \in C^2$. Then*

$$\mathrm{d}f(X) = \frac{\partial f}{\partial x}\,\mathrm{d}X + \frac{\partial^2 f}{\partial x^2}\,\mathrm{d}[X].$$

Proof: Fix $N > 0$. By Weierstrass, we can find a polynomial p_n such that $f = p_n + h_n$, where

$$\sup_{x \in [-N,N]} (|h_n(x)| + |h_n'(x)| + |h_n''(x)|) \leq 2^{-n}.$$

We know that Itô's formula holds for p_n, so

$$\begin{aligned} f(X_t) - f(X_0) - \int_0^t f'(X)\,\mathrm{d}X &= \frac{1}{2}\int_0^t f''(X)\,\mathrm{d}[X] \\ &= h_n(X_t) - h_n(X_n) - \int_0^t h_n(X)\,\mathrm{d}X - \frac{1}{2}\int_0^t h_n''(X)\,\mathrm{d}[X]. \end{aligned}$$

By localization, we can assume that $|X_t| < N$ for all t. By familiar arguments, the right hand side tends to 0 UCP as $n \to \infty$.

3 Using the Tools

3.1 Lévy's Characterisation

Theorem 3.1 (Lévy's Charaterization of Brownian Motion). *Let X be a d-dimensional local martingale. Suppose that*

$$[X^i, X^j]_t = \begin{cases} t & i = j, \\ 0 & i \neq j, \end{cases}$$

and $X_0 = 0$. Then X is a Brownian motion in the filtration.

Proof: Fix $\theta \in \mathbb{R}^d$, and let

$$M_t = \exp\left(i\theta \cdot X_t + \frac{\|\theta\|^2}{2} t\right).$$

We would like to show that this is a local martingale. If we let

$$f(x, y) = \exp\left(i\theta \cdot x + \frac{\|\theta\|^2}{2} y\right),$$

then

$$\frac{\partial f}{\partial x^j} = i\theta^j f(x,y), \qquad \frac{\partial^2 f}{\partial x^j \partial x^k} = -\theta^j \theta^k f(x,y).$$

We find that

$$\begin{aligned} \mathrm{d}M_t = M_t &\left(i\theta \cdot \mathrm{d}X_t + \frac{1}{2}\|\theta\|^2 \, \mathrm{d}t\right) \\ &- \frac{1}{2} M_t \left(\sum_{j,k} \theta^j \theta^k \, \mathrm{d}[X^j, X^k]\right) \\ = iM_t &\theta \cdot \mathrm{d}X_t, \end{aligned}$$

using our formula for $[X^j, X^k]$. Hence M_t is a local martingale, as it an integral with respect to a local martingale. But,

$$|M_t| \le \exp\left(\frac{\|\theta\|^2}{2} t\right)$$

for all t, ω, so it is in class (DL). Therefore, it is a true martingale. So,

$$\mathbb{E}[e^{i\theta(X_t - X_s)} | \mathcal{F}_s] = e^{-\frac{\|\omega\|^2}{2}(t-s)}.$$

So, from Lévy's theorem on characterisation of characteristic functions,

$$X_t - X_s \sim N(0, (t-s)I),$$

and this is independent of $\mathcal{F}_s$. There are some measure theoretic details with respect to passing from countable θ to all that we are skipping.

Since X is continuous, this means X is a Brownian motion in the filtration.

Remark. If W^1 and W^2 are independent Brownian motions, then

$$[W^1, W^2] = 0.$$

Moreover if X is a scalar continuous local martingale, then

$$Z = \exp\left(X - \frac{1}{2}[X]\right)$$

is a local martingale. This follows from Itô:

$$\mathrm{d}Z = Z\,\mathrm{d}X + Z\left(-\frac{1}{2}\,\mathrm{d}[X]\right) + \frac{1}{2}Z\,\mathrm{d}[X] = Z\,\mathrm{d}X.$$

Theorem 3.2 (Dambis-Dubins-Schwarz Theorem)**.** *Let X be a scalar continuous local martingale in the filtration $(\mathcal{F}_t)$, and suppose that $[X]_\infty = \infty$ almost-surely. Define*

$$T(s) = \inf\{t \geq 0 \mid [X]_t > s\}.$$

These are stopping times. Define $W_s = X_{T(s)}$, and $\mathcal{G}_s = \mathcal{F}_{T(s)}$. Then W is a Brownian motion in $(\mathcal{G}_s)$.

Proof: Fix ω. Then $t \mapsto [X]_t(\omega)$ is increasing and continuous.

We want to show that if $T(s,\omega)$ jumps, then $X_{T(s,\omega)}$ does not jump. Since $T(s_0,\omega) \leq T(s_1,\omega)$ for $s_0 \leq s_1$, $\mathcal{G}_{s_0} \subseteq \mathcal{G}_{s_1}$, so we have a filtration.

If we have a jump, $[X]_{t_1} = [X]_{t_0}$. We will show that $X_{t_0} = X_{t_1}$.

Fix $t_0 \geq 0$, and set

$$T = \inf\{u \geq t_0 \mid [X]_u > [X]_{t_0}\},$$

and

$$Y_u = X_{u\wedge T} - X_{u \wedge t_0} = \int_0^u \mathbb{1}_{(t_0,T]}\,\mathrm{d}X.$$

This is a local martingale with

$$[Y]_\infty = \int_0^\infty \mathbb{1}_{(t_0,T]}\, \mathrm{d}[X] = [X]_T - [X]_{t_0} = 0.$$

Hence it must be constant. This works when t_0 is fixed. If it is not fixed, we can prove this over the rationals. Let

$$S_r = \inf \{t \geq r \mid [X]_t > [X]_r\}.$$
$$T_r = \int \{t \geq r \mid X_t \neq X_r\}.$$

We know that $T_r = S_r$ almost-surely for all r rational. But, $r \mapsto (T_r, S_r)$ is right-continuous so $S_r = T_r$ for all r. Hence W is continuous.

To show that W is a local martingale in $\mathcal{G}_s$, let τ_N be the first time that $|X_t| > N$, so X^{T_N} is a bounded martingale. Let $\sigma_N = [X]_{\tau_N}$. Then

$$\begin{aligned}\mathbb{E}[W_{\sigma_N \wedge s_1} | \mathcal{G}_{s_0}] &= \mathbb{E}[X_{\tau_N \wedge T(s_1)} | \mathcal{F}_{s_0}] \\ &\overset{\text{OST}}{=} X_{\tau_N \wedge T(s_0)} = W_{\sigma_N \wedge s_0},\end{aligned}$$

so W^{σ_N} is a martingale. Also

$$[W]_s = [X]_{T(s)} = s.$$

Then we are done by Lévy.

We can invert this to get

$$X_t = W_{[X]_t}.$$

This is up to $[X]_\infty = \infty$, but when the quadratic variation is finite this also makes sense.

3.2 Conformal Invariance of Complex Brownian Motion

Proposition 3.1. *Let X, Y be independent Brownian motions, and set*

$$W = X + iY.$$

Let f be holomorphic and non-constant on $\mathbb{C}$. Then there exists another complex Brownian motion Z and increasing process A such that

$$f(W_t) = f(0) + Z_{A_t},$$

where $A_\infty = \infty$ almost-surely.

In essence, under a holomorphic function, the paths of W and $f(W)$ are the same.

Proof: Let $U_t = u(W_t)$, $V_t = v(W_t)$ where $u + iv = f$. Then by Itô,

$$\begin{aligned}\mathrm{d}U = &\left(\frac{\partial u}{\partial x}\,\mathrm{d}X + \frac{\partial u}{\partial y}\,\mathrm{d}Y\right)\\ &+ \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\,\mathrm{d}[X] + 2\frac{\partial^2 u}{\partial x \partial y}\,\mathrm{d}[X,Y] + \frac{\partial^2 u}{\partial y^2}\,\mathrm{d}[Y]\right).\end{aligned}$$

But X, Y are independent Brownian motions so $[X] = [Y] = t$, and $[X,Y] = 0$, so the latter term becomes

$$\frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)\mathrm{d}t = 0,$$

by complex analysis (Cauchy-Riemann equations). The same holds true for V:

$$\mathrm{d}V = \frac{\partial v}{\partial x}\,\mathrm{d}X + \frac{\partial v}{\partial y}\,\mathrm{d}Y.$$

So U, V are local martingales, as they are stochastic integrals with respect to martingales. Now, from Kunita-Watanabe,

$$\begin{aligned}\mathrm{d}[U] &= \left(\frac{\partial u}{\partial x}\right)^2 \mathrm{d}[X] + 2\left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial u}{\partial y}\right)\mathrm{d}[X,Y] + \left(\frac{\partial u}{\partial y}\right)^2 \mathrm{d}[Y]\\ &= \left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right)\mathrm{d}t,\end{aligned}$$

and also

$$\mathrm{d}[V] = \left(\left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right)\mathrm{d}t.$$

But from the Cauchy-Riemann equations,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

so

$$\mathrm{d}[U] = \mathrm{d}[V] = |f'(W)|^2\,\mathrm{d}t.$$

Now let

$$A_t = \int_0^t |f'(W_s)|^2\,\mathrm{d}s.$$

Then f being non-constant means that there exists $a, b \in \mathbb{C}$ with $f(a) \neq f(b)$. Thus there are disks centred at a, b with $|f(\alpha) - f(\beta)| \geq \varepsilon$ for all α near a, β near b.

By recurrence, W visits these neighbourhoods of a and b infinitely often, almost surely. So the probability $f(W_t)$ converges is 0.

But from the example sheet, $\{A_\infty < \infty\}$ is contained in the event that f converges, so it also has probability 0. Then we are done by DDS.

3.3 Cameron-Martin-Girsanov

Definition 3.1. $\mathbb{P}$ and $\mathbb{Q}$ are *equivalent* probability measures on $(\Omega, \mathcal{F})$ if and only if

$$\mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0.$$

Theorem 3.3 (Filtered Radon-Nikodym)**.** *Let $\mathbb{P}$ and $\mathbb{Q}$ be equivalent on $(\Omega, \mathcal{F})$ with filtration $(\mathcal{F}_t)$. Then there exists $Z = (Z_t)$ a uniformly integrable martingale such that $\mathbb{P}(Z_t > 0) = 1 = \mathbb{Q}(Z_t > 0)$ and*

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[Z_t \mathbb{1}_A]$$

for any $A \in \mathcal{F}_t$.

Proof: By the unfiltered Radon-Nikodym, there exists $Z_\infty > 0$ $\mathbb{P}$-almost-surely such that

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}[Z_\infty \mathbb{1}_A] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[Z_\infty | \mathcal{F}_t] \mathbb{1}_A],$$

so we can set $Z_t = \mathbb{E}^{\mathbb{P}}[Z_\infty | \mathcal{F}_t]$.

Then $Z_0 = \mathbb{E}^{\mathbb{P}}[Z_\infty] = 1$, after assuming $\mathcal{F}_0$ is trivial.

Now if ξ is $\mathcal{F}_t$-measurable and $\mathbb{Q}$-integrable, to compute $\mathbb{E}^{\mathbb{Q}}[\xi | \mathcal{F}_s]$ in terms of Z and $\mathbb{E}^{\mathbb{P}}$,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[\xi \mathbb{1}_A] &= \mathbb{E}^{\mathbb{P}}[\xi Z_t \mathbb{1}_A] \\ &= \mathbb{E}^{\mathbb{P}}\left[Z_s \frac{\mathbb{E}^{\mathbb{P}}[\xi Z_t | \mathcal{F}_s]}{Z_s} \mathbb{1}_A\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\frac{\mathbb{E}^{\mathbb{P}}[\xi Z_t | \mathcal{F}_s]}{Z_s} \mathbb{1}_A\right], \end{aligned}$$

so we find that

$$\mathbb{E}^{\mathbb{Q}}[\xi|\mathcal{F}_s] = \frac{\mathbb{E}^{\mathbb{P}}[\xi Z_t|\mathcal{F}_s]}{Z_s}.$$

Theorem 3.4 (Cameron-Martin-Girsanov)**.** *Let W be a d-dimensional Brownian motion, and α be d-dimensional and previsible, with*

$$\int_0^\infty \|\alpha_s\|^2 \,\mathrm{d}s < \infty \qquad a.s.$$

Let

$$Z_t = \exp\left(\int_0^t \alpha_s \cdot \mathrm{d}W_s - \frac{1}{2}\int_0^t \|\alpha_s\|^2 \,\mathrm{d}s\right),$$

a local martingale. Suppose that Z is a uniformly integrable martingale. Let $\mathbb{Q}$ have density

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = Z_\infty.$$

Let

$$\hat{W}_t = W_t - \int_0^t \alpha_s \,\mathrm{d}s.$$

Then $\hat{W}$ is a $\mathbb{Q}$-Brownian motion.

If X is a continuous semimartingale, we let

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}[X]_t\right).$$

This is the *Dooleans-Dade stochastic exponential*. We know that if $X \in \mathcal{M}_{\mathrm{loc}}$, then $\mathcal{E}(X) \in \mathcal{M}_{\mathrm{loc}}$ by Itô since

$$\mathrm{d}\mathcal{E}(X) = \mathcal{E}(X)\,\mathrm{d}X.$$

Proposition 3.2. *Let $M \in \mathcal{M}_{\mathrm{loc}}$, $M_0 = 0$. Suppose that $\mathcal{E}(M)$ is a UI martingale, such that $\mathcal{E}(M)_\infty > 0$ almost-surely. Let $\mathbb{Q}$ be defined by*

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \mathcal{E}(M).$$

Let $X \in \mathcal{M}_{\mathrm{loc}}(\mathbb{P})$, and

$$\hat{X} = X - [X, M].$$

Then $\hat{X} \in \mathcal{M}_{\mathrm{loc}}(\mathbb{Q})$.

Proof: We claim that

$$\hat{X}\mathcal{E}(M) \in \mathcal{M}_{\text{loc}}(\mathbb{P}).$$

Indeed,

$$\begin{aligned}
\mathrm{d}(\hat{X}\mathcal{E}(M)) &= \hat{X}\,\mathrm{d}(\mathcal{E}(M)) + \mathcal{E}(M)\,\mathrm{d}\hat{X} + \mathrm{d}[\hat{X}, \mathcal{E}(M)] \\
&= \hat{X}\mathcal{E}(M)\,\mathrm{d}M + \mathcal{E}(M)(\mathrm{d}X - \mathrm{d}[X, M]) + \mathcal{E}(M)\,\mathrm{d}[X, M] \text{ (KW)},
\end{aligned}$$

so the finite variation term cancels, and this is a stochastic integral with respect to local continuous martingales.

By localisation, we can assume that $\hat{X}$ is bounded. Since $\mathcal{E}(M)$ is UI, it is of class (D), so

$$\{\mathcal{E}(M)_T \mid T \text{ a finite stopping time}\}$$

is UI. But $\hat{X}_T$ is bounded by a constant, so $\{\hat{X}_T\mathcal{E}(M)_T\}$ is UI, hence $\hat{X}\mathcal{E}(M)$ is a UI martingale, and

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[\hat{X}_\infty|\mathcal{F}_t] &= \frac{\mathbb{E}^{\mathbb{P}}[\hat{X}_\infty\mathcal{E}(M)_\infty|\mathcal{F}_t]}{\mathbb{E}^{\mathbb{P}}[\mathcal{E}(M)_\infty|\mathcal{F}_t]} \\
&= \frac{\hat{X}_t\mathcal{E}(M)_t}{\mathcal{E}(M)_t} = \hat{X}_t.
\end{aligned}$$

We now prove Cameron-Martin-Girsanov.

Proof: Recall that W is a $\mathbb{P}$-Brownian motion, and if α is previsible with

$$\int_0^\infty \|\alpha_s\|^2\,\mathrm{d}s < \infty$$

$\mathbb{P}$-almost-surely, then we can let

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \mathcal{E}\left(\int \alpha \cdot \mathrm{d}W\right)_\infty.$$

If we assume this is UI, then,

$$\hat{W}_t = W_t - \int_0^t \alpha_s\,\mathrm{d}s$$

is a $\mathbb{Q}$-Brownian motion.

By the previous proposition,

$$\begin{aligned}\hat{W}^i &= W^i - \left[W_i, \int \alpha \cdot \mathrm{d}W\right] \\ &\overset{\text{KW}}{=} W^i - \int \alpha_s \,\mathrm{d}s\end{aligned}$$

is a local martingale. Note that $\mathbb{Q} \sim \mathbb{P}$ since $\int \alpha \,\mathrm{d}W$ converges, since

$$\left[\int \alpha \,\mathrm{d}w\right]_\infty = \int_0^\infty \|\alpha\|^2 \,\mathrm{d}s < \infty$$

by assumption. So $\mathcal{E}\left(\int \alpha \,\mathrm{d}W\right) > 0$ almost-surely. Calculating quadratic variation is the same under equivalent measures (as seen in the third example sheet). Hence

$$[\hat{W}^i, \hat{W}^j]_t = [W^i, W^j]_t = t\delta^{ij},$$

hence $\hat{W}$ is a Brownian motion by Lévy.

How do we decide if $\mathcal{E}(M)$ is a UI martingale? There are several ways.

Proposition 3.3 (Novikov)**.** *Let $M \in \mathcal{M}_{\text{loc}}$ be such that*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}[M]_\infty\right)\right] < \infty.$$

Then $\mathcal{E}(M)$ is a UI martingale.

Applying this to CMG, if

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^\infty \|\alpha\|^2 \,\mathrm{d}s\right)\right] < \infty,$$

then $\mathcal{E}\left(\int \alpha \,\mathrm{d}W\right)$ is a UI martingale.

3.4 Applications of Cameron-Martin-Girsanov

Suppose we want to solve

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

where $X_0 = x$, $\sigma > 0$ is constant and $b : \mathbb{R} \to \mathbb{R}$ is bounded and measurable. Let $\tilde{W}$ be a Brownian motion on $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ and let

$$X_t = x + \sigma \tilde{W}_t.$$

Fix a constant time $T > 0$, and let

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\tilde{\mathbb{P}}} = \mathcal{E}\left(\int \frac{b(X)}{\sigma}\,\mathrm{d}\tilde{W}\right)_T$$

By CMG,

$$W_t = \tilde{W}_t - \int_0^t \frac{b(X_s)}{\sigma}\,\mathrm{d}s$$

is a $\mathbb{P}$-Brownian motion for $t \in [0, T]$. So

$$W_t = \frac{X_t - x}{\sigma} - \int_0^t \frac{b(X_s)}{\sigma}\,\mathrm{d}s,$$

and hence

$$X_t = x + \int_0^t b(X_s)\,\mathrm{d}s + \sigma W_t$$

on $t \in [0, T]$.

3.5 Itô's Martingale Representation Theorem

Suppose that the filtration is generated by a d-dimensional Brownian motion. If X is a locally square integrable local martingale, then there exists a previsible α, with

$$\int_0^t \|\alpha\|^2\,\mathrm{d}s < \infty \qquad \text{a.s.}$$

such that

$$X_t = X_0 + \int_0^t \alpha_s \cdot \mathrm{d}W_s.$$

Proof: By localisation, it is enough to show this for an $X \in L^2(\sigma(W))$. There exists a previsible α such that $\alpha \in L^2(W)$,

$$\mathbb{E}\int_0^\infty \|\alpha_s\|^2\,\mathrm{d}s < \infty$$

and

$$X = \mathbb{E}[X] + \int_0^\infty \alpha\,\mathrm{d}W.$$

By Itô's isometry, it is enough to check this for X is a dense subset of L^2. It is enough to check when

$$X = \exp\left(i\sum_{k=1}^{n}\theta_k \cdot (W_{t_k} - W_{t_{k-1}})\right),$$

where θ_k is not random. Let

$$\beta = i\sum_{k=1}^{n}\theta_k \mathbb{1}_{(t_{k-1},t_k]}.$$

Then we can let

$$\alpha = \mathcal{E}\left(\int \beta \, \mathrm{d}W\right)\beta.$$

Example 3.1. (Counterexample for Example Sheets)

Some people said that,

$$\begin{aligned} \mu(s,t] = 0 &\implies \nu(s,t] = 0, \\ \implies \mu(A) = 0 &\implies \nu(A) = 0, \end{aligned}$$

for A Borel. This is not true. Let μ be Lebesgue, and

$$\nu = \sum_{q_n \in \mathbb{Q}} 2^{-n}\delta_{q_n}.$$

Then none of these have 0 $\mu(s,t]$, but $\nu(\mathbb{R}\setminus\mathbb{Q}) = 0$. However this is true if we have

$$\mu(s,t] = \nu(s,t] + \pi(s,t],$$

for some other measure π. Then we can show

$$\mu(A) = \nu(A) + \pi(A)$$

for all A Borel, by Dynkin's. Then we get the required result, that $\mu(A) = 0 \implies \nu(A) = 0$.

4 Stochastic Differential Equations

We will consider equations of the form

$$\mathrm{d}X = b(X)\,\mathrm{d}t + \sigma(X)\,\mathrm{d}W,$$

where $X_0 = x$, $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$, where W is a d-dimensional Brownian motion.

A solution consists of:

(1) A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $(\mathcal{F}_t)$ satisfying the usual conditions.

(2) W is a Brownian motion defined on the above filtration.

(3) An adapted X such that

$$\int_0^t \|b(X_s)\|\,\mathrm{d}s < \infty \qquad \text{a.s.}$$

for all t, and

$$\int_0^t \|\sigma(X_s)\|^2\,\mathrm{d}s < \infty \qquad \text{a.s.}$$

for all t, where $\|\sigma\|^2 = \mathrm{tr}(\sigma\sigma^T)$.

(4) X is a solution to the equation

$$X_t = X_0 + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}W_s$$

for all t.

Definition 4.1. A *strong solution* is where (1) and (2) are given, and (3) is the output, where the filtration is generated by W.

This embodies the notion of causality.

Remark. Consider a simulation scheme, where we discretize time and replace $W_{t_k} - W_{t_{k-1}}$ with $\sqrt{t_k - t_{k-1}}Z_k$, where $Z \sim N(0, I)$. Then we can let

$$X_{t_k} = X_{t_{k-1}} + b(X_{t_{k-1}})(t_k - t_{k-1}) + \sigma(X_{t_{k-1}})\sqrt{t_k - t_{k-1}}Z_k,$$

so we see that

$$X_{t_k} = f(X_0, Z_1, \dots, Z_k).$$

Definition 4.2. A *weak solution* has output which is (1), (2) and (3). In particular, X may not be adapted to the filtration generated by W, $(\mathcal{F}_t^W)$.

Remark. This is also a natural interpretation. In a modelling sense, we only care about b and σ; W is auxiliary.

Example 4.1. (Tanaka's Example)

Let $n = 1$, $b = 0$ and
$$\sigma(x) = \operatorname{sgn}(x),$$
where $\sigma(0) = 0$. We claim there exists a weak solution. Let $(\Omega, \mathcal{F}, \mathbb{P})$, and $(\mathcal{F}_t)$ be supporting a Brownian motion X. Let

$$W_t = \int_0^t \operatorname{sgn}(X_s)\,\mathrm{d}X_s.$$

Then,

$$[W] = \int_0^t \mathbb{1}(X_s \neq 0)\,\mathrm{d}s = t,$$

since $\mathbb{1}(X_s = 0) = 0$ almost-surely. Hence W is a Brownian motion by Lévy. Then, by Fubini's for integration,

$$\int_0^t \operatorname{sgn}(X_s)\,\mathrm{d}W_s = \int_0^t (\operatorname{sgn} X_s)^2\,\mathrm{d}X_s = \int_0^t \mathbb{1}(X_s \neq 0)\,\mathrm{d}X_s = X_t.$$

However, there is no strong solution. For a sketch, if we applied fake Itô, then

$$|X_t| = \int_0^t \operatorname{sgn}(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t \delta_0(X_s)\,\mathrm{d}s.$$

Let W_t be defined as before. Then we can approximate

$$\operatorname{sgn}(x) \sim \tanh\left(\frac{x}{\varepsilon}\right),$$

for ε. Hence by Itô,

$$\varepsilon \log\cosh\left(\frac{X_t}{\varepsilon}\right) = \int_0^t \tanh\left(\frac{X_s}{\varepsilon}\right)\mathrm{d}X_s + \frac{1}{2\varepsilon}\int_0^t \frac{\mathrm{d}s}{\cosh(X_s/\varepsilon)^2}.$$

Taking the limit as $\varepsilon \to 0$,

$$\begin{aligned}
\varepsilon \log \cosh \left(\frac{X}{\varepsilon}\right) &\to |X|, \\
\tanh \left(\frac{X}{\varepsilon}\right) &\to \operatorname{sgn}(x), \\
\mathbb{E} \int_0^t \left(\tanh \left(\frac{X_s}{\varepsilon}\right) - \operatorname{sgn}(X_s)\right)^2 \mathrm{d}s &\to 0,
\end{aligned}$$

by DCT. If X is a solution of the equation, then

$$W_t = |X_t| - \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \frac{\mathrm{d}s}{\cosh\left(X_s/\varepsilon\right)^2},$$

so W is $\sigma(|X_s|, 0 \le s \le t)$ measurable. Hence X cannot be $\sigma(W_s, 0 \le s \le t)$.

The difference is only really prevalent for the martingale representation theorem.

The quantity at the end of the proof is

$$\begin{aligned}
\lim_{\varepsilon to 0} \frac{1}{2\varepsilon} \int_0^t \frac{\mathrm{d}s}{\cosh\left(X_s/\varepsilon\right)^2} &= \lim_{\varepsilon \to 9} \frac{1}{2\varepsilon} \mathrm{Leb}\{s \in [0, t] \mid |X_s < \varepsilon\} \\
&= |X_t| - \int_0^t \operatorname{sgn}(X_s), \mathrm{d}X_s,
\end{aligned}$$

which is the *local time* of X at $X = 0$. The above generalization of Itô's to non-C^2 functions is known as *Tanaka's formula*.

4.1 Uniqueness

Definition 4.3 (Pathwise Uniqueness)**.** A SDE is *pathwise unique* if, given the stochastic differential equation and (1) and (2), the probability space and Brownian motion, if X and X' are two solutions with $X_0' = X_0$, then

$$\mathbb{P}(X_t = X_t' \text{ for all t}) = 1.$$

Definition 4.4 (Uniqueness in Law)**.** A SDE is *unique in law* if, given two weak solutions and conditions (1), (2) and (3), and suppose that $X_0 \sim X_0'$, then

$$(X_t)_{t \ge 0} \sim (X_t')_{t \ge 0}.$$

In other words,

$$\mathbb{P} \circ X^{-1} = \mathbb{P}' \circ (X')^{-1}.$$

Example 4.2. (Tanaka Two)

Let $g(x) = \operatorname{sgn}(x) + \mathbb{1}(x = 0)$, i.e.

$$g(x) = \begin{cases} 1 & x \geq 0, \\ -1 & x < 0. \end{cases}$$

Then consider

$$\mathrm{d}X = g(X)\,\mathrm{d}W, \qquad X_0 = 0.$$

This has uniqueness in law, since any weak solution has

$$[X]_t = \int_0^t g(X_s)^2\,\mathrm{d}s = t,$$

so since X is a local martingale by Lévy it must be a Brownian motion.

But there is no pathwise uniqueness. Let X be a weak solution, and let $\hat{X} = -X$. Then,

$$\mathrm{d}\hat{X} = -g(-\hat{X})\,\mathrm{d}W = (g(\hat{X}) - 2\mathbb{1}(X_t) = 0)\,\mathrm{d}W.$$

But note that

$$\int_0^t \mathbb{1}(X_t = 0)\,\mathrm{d}W_s = 0,$$

because it is a local martingale, and it has quadratic variation

$$\left[\int \mathbb{1}(X_s = 0)\,\mathrm{d}W\right]_t = \int_0^t \mathbb{1}(X_s = 0)\,\mathrm{d}s = 0,$$

since $X_s \sim N(0, s)$. So $-X$ is also a solution.

When we do vectorized SDEs, we really mean

$$X_t^i = X_0^i + \int_0^t b^i(X_s)\,\mathrm{d}s + \sum_{k=1}^d \int_0^t \sigma^{ik}(X_s)\,\mathrm{d}W_s^k.$$

Theorem 4.1. *The stochastic differential equation*

$$\mathrm{d}X = b(X)\,\mathrm{d}t + \sigma(X)\,\mathrm{d}W \tag{$\star$}$$

has pathwise uniqueness if b and σ are locally Lipschitz, i.e. for all N, there exists $K_N > 0$ such that

$$\|b(x) - b(y)\| \leq K_N\|x - y\|, \qquad \|\sigma(x) - \sigma(y)\| \leq K_N\|x - y\|,$$

for all x, y with $\|x\|, \|y\| \leq N$.

The key lemma is the follows.

Lemma 4.1 (Grönwall's Lemma)**.** *Suppose f is locally integrable and*

$$f(t) \leq a + b \int_0^t f(s)\,\mathrm{d}s,$$

for all $t \geq 0$, for a, b constants. Then

$$f(t) \leq ae^{bt}.$$

Using this we show pathwise uniqueness.

Proof: Let X and X' be two weak solutions of $(\star)$, defined on the same set-up with $X_0 = X_0'$ almost-surely. Fix $N > 0$ and let

$$T_N = \inf\{t \geq 0 \mid \|X_t\| > N \text{ or } \|X_t'\| > N\}$$

Let

$$f(t) = \mathbb{E}\left[\|X_{t\wedge T_N} - X'_{t\wedge T_N}\|^2\right].$$

Then by Itô,

$$\begin{aligned}\mathrm{d}\|X_{t\wedge T_N} - X'_{t\wedge T_N}\|^2 = {} & 2(X_{t\wedge T_N} - X'_{t\wedge T_N}) \cdot (b(X_{t\wedge T_N}) - b(X'_{t\wedge T_N}))\,\mathrm{d}t \\ & + 2(X_{t\wedge T_N} - X'_{t\wedge T_N}) \cdot (\sigma(X_{t\wedge T_N}) - \sigma(X'_{t\wedge T_N}))\,\mathrm{d}W \\ & + \|\sigma(X_{t\wedge T_N}) - \sigma(X'_{t\wedge T_N})\|^2\,\mathrm{d}t.\end{aligned}$$

Because of our stopping time,

$$\|\sigma(x)\| \leq \|\sigma(0)\| + K_N \cdot N$$

for $\|x\| \leq N$. So the length of the vector in the $\mathrm{d}W$ term is at most

$$8N(\|\sigma(0)\| + K_N N),$$

which is uniform in (t, ω). Hence the stochastic integral is a true martingale, since the quadratic variation of the term is $[M]_t \leq Ct$. Therefore it is square integrable. So,

$$\begin{aligned}f(t) &\leq \mathbb{E}\int_0^t 2\|X_{s\wedge T_N} - X'_{s\wedge T_N}\|^2 K_N\,\mathrm{d}s \\ &\quad + \mathbb{E}\int_0^t K_N^2\|X_{s\wedge T_n} - X'_{s\wedge T_N}\|^2\,\mathrm{d}s \\ &= \int_0^t bf(s)\,\mathrm{d}s,\end{aligned}$$

where $b = 2K_N + K_N^2 < \infty$. Hence by Grönwall, $f(t) = 0$ for all t. Thus

$$\mathbb{P}\left(\sup_{0 \le s \le T_N} \|X_s - X_s'\| = 0\right) = 1.$$

Now we can send $N \to \infty$.

Theorem 4.2 (Yamada-Watanabe Theorem)**.** *If* $(\star)$ *has pathwise uniqueness, then it has uniqueness in law.*

Proof: We will sketch the proof. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t), W, X)$ be the first setup, and $(\Omega', \mathcal{F}', \dots, X')$ be the second, with $X_0 \sim X_0' \sim \lambda$, some measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Let C^j be the set of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^j$.

Let μ be the measure on $\mathbb{R}^n \times C^d \times C^n$ be defined by

$$\mu(A \times B \times C) = \mathbb{P}(X_0 \in A, W \in B, X \in C),$$

where B, C are Borel subsets of C^d, C^n respectively. Since X_0 and W are independent under $\mathbb{P}$, we can factorise

$$\mu(\mathrm{d}x, \mathrm{d}w, \mathrm{d}y) = \lambda(\mathrm{d}x)\mathbb{W}(\mathrm{d}w)\nu(x, w; \mathrm{d}y),$$

where ν is a conditional law of X given X_0 and W. Formally,

$$\nu(X_0, W; C) = \mathbb{P}(X \in C \mid X_0, W).$$

Define μ' similarly. Then

$$\mu'(\mathrm{d}x, \mathrm{d}w, \mathrm{d}y) = \lambda(\mathrm{d}x)\mathbb{W}(\mathrm{d}w)\nu'(x, w' \, \mathrm{d}y).$$

Let $\hat{\Omega} = \mathbb{R}^n \times C^d \times C^n \times C^n$, and

$$\hat{\mathbb{P}}(\mathrm{d}x, \mathrm{d}w, \mathrm{d}y, \mathrm{d}y') = \lambda(\mathrm{d}x)\mathbb{W}(\mathrm{d}w)\nu(x, w; \mathrm{d}y)\nu'(x, w; \mathrm{d}y').$$

Define

$$\hat{X}_0(x, w, y, y') = x = \hat{X}_0'(x, w, y, y'),$$

and also

$$\hat{W}_t(x, w, y, y') = w(t),$$

and

$$\hat{X}_t(x, w, y, y') = y(t), \qquad \hat{X}_t'(x, w, y, y') = y'(t).$$

Then note that $\hat{X}$ and $\hat{X}'$ are two solutions on the same set-up with $\hat{X}_0 = \hat{X}'_0$ almost-surely, so by pathwise uniqueness,

$$\hat{\mathbb{P}}(\hat{X}_t = \hat{X}'_t \text{ for all } t) = 1,$$

hence

$$\mathbb{P}(X \in C) = \hat{\mathbb{P}}(\hat{X} \in C) = \hat{\mathbb{P}}(\hat{X}' \in C) = \mathbb{P}'(X' \in C).$$

This gives uniqueness in law.

4.2 Strong Existence

Theorem 4.3 (Itô)**.** *Consider*

$$\mathrm{d}X = b(X)\,\mathrm{d}t + \sigma(X)\,\mathrm{d}W. \tag{$\star$}$$

Suppose that b, σ are globally Lipschitz. Then $(\star)$ has a unique strong solution.

Proof: We will show that there exists a solution on $[0, T]$, where T depends on the Lipschitz constant K, but not on X_0.

Build (X_t^1) with inputs X_0 and (W_t). Then we build (X_t^2) with inputs X'_T and $(W_{t+T} - W_T)$. Let

$$\tilde{X}_t = \begin{cases} X_t^1 & 0 \le t \le T, \\ X_{t-T}^2 & T \le t \le 2T. \end{cases}$$

We can check that $\tilde{X}$ solves the SDE on $[0, 2T]$. This is clear for $[0, T]$. For $T \le t \le 2T$,

$$\begin{aligned} \tilde{X}_t = X_{t-T}^2 = X_T^1 + \int_0^{t-T} b(X_s^2)\,\mathrm{d}s + \int_0^{t-T} \sigma(X_s)\,\mathrm{d}(W_{s+T} - W_T) \\ = X_0 + \int_0^T b(\tilde{X}_s)\,\mathrm{d}s + \int_0^T \sigma(\tilde{X}_s)\,\mathrm{d}W_s + \int_T^t b(\tilde{X}_s)\,\mathrm{d}s + \int_T^t \sigma(\tilde{X}_s)\,\mathrm{d}W_s. \end{aligned}$$

Recall the *Banach fixed point theorem*: Let $F : B \to B$, where B is a Banach space with norm $|||\cdot|||$, and suppose

$$|||F(x) - F(y)||| \le c|||x - y|||,$$

for all $x, y \in B$ where $0 < c < 1$. Then there exists a unique fixed point $x^* \in B$, i.e. $F(x^*) = x^*$.

We want to find Φ that takes X_0 and $(W_t)_{0\le t\le T}$ and returns $(X_t)_{0\le t\le T}$. We will find a contraction F that will build Φ. F may depend on X_0, so we can use different maps on each interval.

In our case, we will let

$$F(Y)_t = X_0 + \int_0^t b(Y_s)\,\mathrm{d}s + \int_0^t \sigma(Y_s)\,\mathrm{d}W_s.$$

We will build a norm on this space of processes. Let

$$|||Y||| = \sqrt{\mathbb{E}\sup_{t\in[0,T]}\|Y_t\|^2},$$

and our space be $B = \{Y \mid \text{continuous, adapted, } |||Y||| < \infty\}$. We assert that B is complete with respect to this norm.

We need to solve the case when X_0 is random as well, so we can glue together solutions. Assume that $\mathbb{E}\|X_0\|^2 < \infty$, else we can replace $|||\cdot|||$ with

$$\sqrt{\mathbb{E}e^{-\|X_0\|}\sup_{0\le t\le T}\|Y_t\|^2}.$$

We only need this so that $F(0)\in B$. Now we will show that F is a contraction:

$$\begin{aligned}
&|||F(X)-F(Y)|||^2\\
&= \mathbb{E}\left(\sup_{0\le t\le T}\left\|\int_0^t (b(X_s)-b(Y_s))\,\mathrm{d}s + \int_0^t (\sigma(X_s)-\sigma(Y_s))\,\mathrm{d}W_s\right\|^2\right)\\
&\le 2\mathbb{E}\sup_{0\le t\le T}\left(\int_0^t \|b(X_s)-b(Y_s)\|\,\mathrm{d}s\right)^2 + 2\mathbb{E}\left(\sup_t\left\|\int_0^t (\sigma(X_s)-\sigma(Y_s))\,\mathrm{d}W_s\right\|^2\right)
\end{aligned}$$

For the first term, we can bound it by

$$\begin{aligned}
2\mathbb{E}\left(\int_0^T \|b(X)-b(Y)\|\,\mathrm{d}s\right)^2 &\le 2K^2\mathbb{E}\left(\int_0^T \|X_s-Y_s\|\,\mathrm{d}s\right)^2\\
&\le 2K^2T^2\mathbb{E}\sup_{0\le s\le T}\|X_s-Y_s\|^2 = 2K^2T^2|||X-Y|||^2.
\end{aligned}$$

For the second term, we apply the Burkholder inequality, as found in example sheets:

$$\begin{aligned}
&2\mathbb{E}\left(\sup_{0\le t\le T}\left\|\int_0^t (\sigma(X_s)-\sigma(Y_s))\,\mathrm{d}W_s\right\|^2\right)\\
&\le 8\mathbb{E}\int_0^T \|\sigma(X_s)-\sigma(Y_s)\|^2\,\mathrm{d}s\\
&\le 8K^2T|||X-Y|||^2,
\end{aligned}$$

as before.

Finally we check that $F : B \to B$. It suffices to check that $F(0) \in B$, since

$$|||F(X)||| \le |||F(X)-F(0)||| + |||F(0)||| \le c|||X||| + |||F(0)|||.$$

Indeed,

$$F(0) = \mathbb{E}\left(\sup_{0\le t\le T}\left\|X_0 + \int_0^t b(0)\,\mathrm{d}s + \int_0^t \sigma(0)\,\mathrm{d}W_s\right\|^2\right) < \infty,$$

by the assumption on X_0 and properties of W. We can pick $c = (2T + 8)K^2T < 1$, so by Banach fixed point theorem, there is a fixed point.

We have uniqueness from the previous theorem.

We can show that if σ, b are globally Lipschitz, and $X_0 \in L^p$ for $p \ge 2$, then

$$\mathbb{E}\left(\sup_{0\le t\le T}\|X_t\|^p\right) < \infty$$

for all T. This is using Grönwall, Burkholder and

$$\|b(X)\| \le \|b(0)\| + K\|X\|, \qquad \|\sigma(X)\| \le \|\sigma(0)\| + K\|X\|.$$

4.3 Feynman-Kac Formula

We are given:

- $b : \mathbb{R}^n \to \mathbb{R}^n$,
- $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$,
- yet more functions $c : \mathbb{R}^n \to \mathbb{R}$,
- $v : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$,

- $\phi : \mathbb{R}^n \to \mathbb{R}$.

Assume that $v(0,x) = \phi(x)$ for all x, $v \in C^2$ and

$$\frac{\partial v}{\partial \tau} + c(x)v(\tau x) = \sum_i b^i \frac{\partial v}{\partial x^i} + \frac{1}{2}\sum_{i,j} a^{ij} \frac{\partial^2 v}{\partial x^i \partial x^j},$$

where $a = \sigma\sigma^T$, i.e.

$$a^{ij} = \sum_{k=1}^{d} \sigma^{ik}\sigma^{jk}.$$

Finally suppose that

$$\mathrm{d}X = b(X)\,\mathrm{d}t + \sigma(X)\,\mathrm{d}W$$

has a weak solution, where X_0 may be random. Fix $\tau > 0$ and let

$$M_t = \exp\left(-\int_0^t c(X_s)\,\mathrm{d}s\right) v(\tau - t, X_t).$$

Theorem 4.4 (Feynman-Kac Formula)**.** *$(M_t)_{0\le t\le\tau}$ is a local martingale. If M is a true martingale (e.g. if v and c are bounded), then*

$$v(\tau, X_0) = \mathbb{E}\left[\exp\left(-\int_0^\tau c(X_s)\,\mathrm{d}s\right)\phi(X_\tau) \;\middle|\; X_0\right].$$

Proof: We apply Itô: note we get the stochastic product law by applying

Itô on $f(x,y) = xy$. From this,

$$\begin{aligned}
\mathrm{d}M &= -c(X_t)\exp\left(-\int_0^t c(X_s)\,\mathrm{d}s\right) v(\tau - t, X_t)\,\mathrm{d}t \\
&\qquad + \exp\left(-\int_0^t c(X_s)\,\mathrm{d}s\right) \mathrm{d}v(\tau - t, X_t) + \underbrace{\cancel{[\exp(\cdots), v(\cdots)]}}_{\text{as first term is FV}} \\
&= -c\exp(\cdots)v\,\mathrm{d}t \\
&\qquad + \exp(\cdots)\left(-\frac{\partial v}{\partial \tau}\,\mathrm{d}t + \sum_i \frac{\partial v}{\partial x^i}\,\mathrm{d}X^i + \sum_{i,j}\frac{\partial^2 v}{\partial x^i \partial x^j}\,\mathrm{d}[X^i, X^j]\right) \\
&\overset{(1)}{=} -c\exp(\cdots)v\,\mathrm{d}t + \exp(\cdots)\Bigg(-\frac{\partial v}{\partial \tau}\,\mathrm{d}t + \sum_i \frac{\partial v}{\partial x^i} b^i\,\mathrm{d}t \\
&\qquad\qquad + \sum_{i,k}\frac{\partial v}{\partial x^i}\sigma^{ik}\,\mathrm{d}W^k + \frac{1}{2}\sum_{i,j}\frac{\partial^2 v}{\partial x^i \partial x^j}\sum_k \sigma^{ik}\sigma^{jk}\,\mathrm{d}t\Bigg) \\
&\overset{(2)}{=} \exp\left(-\int_0^t c(X_s)\,\mathrm{d}s\right)\sum_{i,k}\frac{\partial v}{\partial x^i}\sigma^{ik}\,\mathrm{d}W^k,
\end{aligned}$$

where in (1) we used the fact X solves te SDE, and in (2) we cancel everything using what we know. So this is a stochastic integral with respect to W, hence is a local martingale.

If it is a true martingale, then

$$v(\tau, X_0) = M_0 = \mathbb{E}[M_\tau | \mathcal{F}_0] = \mathbb{E}\left[\exp\left(-\int_0^\tau c(X_s)\,\mathrm{d}s\right)\phi(X_\tau) \;\middle|\; \mathcal{F}_0\right].$$

Then we can condition with respect to X_0, by using the tower property to get the required formula.

Remark. If we can solve the SDE for any $X_0 = x \in \mathbb{R}^n$, and can find a bounded solution to the PDE, then the PDE solution is unique.

Similarly, suppose we can solve the PDE with a bounded solution for any initial condition $\phi = v(0, \cdot)$. Then the law of X_τ is unique.

Remark. Suppose that $c(x) \geq 0$ for all x. Let $Z \sim \exp(1)$, independent of (X_t). Let

$$T = \inf\left\{t \geq 0 \;\middle|\; \int_0^t c(X_s)\,\mathrm{d}s > Z\right\}.$$

We let T be the lifetime of $\tilde{X}$, where

$$\tilde{X}_t = \begin{cases} X_t & t < T, \\ \Delta & t \geq T, \end{cases}$$

where $\Delta \notin \mathbb{R}^n$. By convention $\phi(\Delta) = 0$ for any function. Then

$$\begin{aligned} \mathbb{E}[\phi(\tilde{X}_\tau)] &= \mathbb{E}\left[\phi(X_\tau)\mathbb{1}_{\{\tau<T\}}\right] \\ &= \mathbb{E}[\mathbb{E}\left[\left[\phi(X_\tau)\mathbb{1}\left(\int_0^\tau c(X_s)\,\mathrm{d}s\right) \mid X\right]\right] \\ &= \mathbb{E}\left[\phi(X_t)\mathbb{P}\left(Z > \int_0^\tau c(X_s)\,\mathrm{d}s \mid X\right)\right] \\ &= \mathbb{E}\left[\phi(X_t)\exp\left(-\int_0^\tau c(X_s)\,\mathrm{d}s\right)\right], \end{aligned}$$

so we interpret c as the 'rate of killing'. In finance, c has interpretation as an interest rate.

Remark. Suppose $f : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ (or f a distribution) satisfying

$$\frac{\partial f}{\partial \tau} = \sum \frac{\partial f}{\partial x^i} b^i + \frac{1}{2}\sum \frac{\partial^2 f}{\partial x^i \partial x^j} a^{ij},$$

with

$$f(0, x, y) = \delta_y(x).$$

Here f and the PDE is interpreted in the weak sense. From Feynman-Kac, $f(\tau, x, y)$ is the density of X_τ at y given $X_0 = x$. We can also write

$$\frac{\partial f}{\partial \tau} = -\sum_i \frac{\partial}{\partial y^i}(b^i f) + \frac{1}{2}\sum \frac{\partial^2}{\partial y^i \partial y^j}(a^{ij} f).$$

Then $f(t, x, y)$ is the density of X_t at y given $X_0 = x$. This needs some more conditions. Note

$$v(t, x) = \mathbb{E}[\phi(X_t) | X_0 = x] = \int \phi(y) f(t, x, y)\,\mathrm{d}y$$

$$\begin{aligned} v(t+\varepsilon, x) - v(t, x) &= \mathbb{E}\left[\int_t^{t+\varepsilon} \sum_i b^i \frac{\partial \phi}{\partial x^i} + \frac{1}{2}\int_t^{t+\varepsilon} \sum_{i,j} a^{ij} \frac{\partial^2 \phi}{\partial x^i \partial x^j} + \cancel{\text{martingale}}\right] \\ &= \int_t^{t+\varepsilon} \sum_i b^i(y) \frac{\partial \phi(y)}{\mathrm{d}X^i} f(t, x, y)\,\mathrm{d}y + \cdots \\ &= \sum \phi(y)\left(-\sum \frac{\partial}{\partial y^i}(b^i f)\right) + \cdots \end{aligned}$$

I have no idea what this means.

5 Continuous Time Finance

Consider a market with $1+d$ assets. The first is a bank account, which is risk-free, and there are d risky assets.

We will assume there is no transaction cost, no bid-ask spread, no price impact.

We will let B_t be the price of the bank (or money market) account at time t, which we will assume is a semimartingale of the form

$$\mathrm{d}B_t = B_t r_t \, \mathrm{d}t,$$

where (r_t) is previsible and locally $\mathrm{d}t$-integrable and $B_0 > 0$, hence

$$B_t = B_0 \exp\left(\int_0^t r_s \, \mathrm{d}s\right),$$

and we will let S_t^i be the price of the asset i at time t, which again is a continuous semimartingale.

Some questions we may have are:

- what is the optimal investment?
- how do we price contingent claims?

Introduce a trader. Suppose they hold ϕ_t shares of the bank account, and θ_t^i shares of asset i between time $t-\Delta t$ and t. This means that we assume (ϕ_t) and (θ_t^i) are previsible, and B-integrable, S^i-integrable respectively. Out wealth at time $t-\Delta t$ will be

$$X_{t-\Delta t} = \phi_t B_{t-\Delta t} + \sum \theta_t^i S_{t-\Delta t}^i.$$

We will assume self-financing, so

$$X_t = \phi_t B_t + \sum_i \theta_t^i S_t^i.$$

Definition 5.1. A $(1+d)$-dimensional process (ϕ, θ) is a self-financing trading strategy if it is previsible, (B, S)-integrable and

$$\mathrm{d}(\phi_t B_t + \theta_t \cdot S_t) = \phi_t \, \mathrm{d}B_t + \theta_t \, \mathrm{d}S_t.$$

If we let $X_t = \phi_t B_t + \theta_t \cdot S_t$, then

$$\mathrm{d}X_t = \phi_t B_t r_t \, \mathrm{d}t + \Theta_t \cdot \mathrm{d}S_t = r_t(X_t - \theta_t \cdot S_t) \, \mathrm{d}t + \theta_t \cdot \mathrm{d}S_t.$$

This can be solved:

$$
\begin{aligned}
\mathrm{d}\left(\exp\left(-\int_0^t r_s\,\mathrm{d}s\right)X_t\right) &= -re^{-\int r}X\,\mathrm{d}t + e^{-\int r}(rX - r\theta\cdot S)\,\mathrm{d}t + e^{-\int r}\theta\cdot\mathrm{d}S \\
&= \theta\cdot\left(e^{-\int r}\,\mathrm{d}S - re^{-\int r}S\,\mathrm{d}t\right) \\
&= \theta\cdot\mathrm{d}\left(\exp\left(-\int_0^t r_s\,\mathrm{d}s\right)S_t\right).
\end{aligned}
$$

As a bit of notation, write

$$
\begin{aligned}
X_t^{x,\theta} &= \exp\left(\int_0^t r_s\,\mathrm{d}s\right)\left(x + \int_0^t \theta_s\cdot\mathrm{d}\left(\exp\left(-\int_0^t r_s\,\mathrm{d}s\right)S_s\right)\right) \\
&= B_t\left(\frac{x}{B_0} + \int_0^t \theta\cdot\mathrm{d}\left(\frac{S}{B}\right)\right).
\end{aligned}
$$

Our controls are the initial wealth, and the d-dimensional process θ.

Remark. Given x, θ, then setting

$$\phi_t = \frac{X_t^{x,\theta} - \theta_t\cdot S_t}{B_t},$$

then (ϕ, θ) is self-financing.

5.1 Optimal Investment

A typical question will be: given a utility function $u : \mathbb{R}\to\mathbb{R}$ and $x\in\mathbb{R}$, try to maximize

$$\mathbb{E}[(X_T^{x,\theta})].$$

Here $T > 0$ is a given investment horizon. U, the utility function, will be assumed to be increasing and concave (i.e. risk-averse).

Consider a previsible process π such that

$$X_T^{0,\pi} \geq 0 \qquad \text{a.s.} \qquad \mathbb{P}(X_T^{0,\pi} > 0) > 0.$$

Notice that $(x,\theta)\mapsto X^{x,\theta}$ is linear, so

$$X_T^{x,\theta+\pi} = X_T^{x,\theta} + X_T^{\theta,\pi} \geq X_T^{x,\theta},$$

and strictly greater with positive probability. Hence

$$\mathbb{E}[U(X_T^{x,\theta+\pi})] \geq \mathbb{E}[U(X_T^{x,\theta})].$$

In this case there is no optimal solution.

As a warning, this is not the definition of arbitrage. Let $r = 0$, $d = 1$ and $S_t = W_t$, a Brownian motion. Then there exists $(\pi_t)_{0 \le t \le T}$ such that

$$\int_0^T \pi_s^2 \,\mathrm{d}s < \infty$$

almost-surely, and

$$\int_0^T \pi_s \,\mathrm{d}W_s = K > 0.$$

Let $f : [0, T] \to [0, \infty]$ be strictly increasing and continuous, with $f(0) = 0$ and $f(T) = \infty$, for example

$$f(t) = \frac{t}{T - t},$$

and let

$$Z_u = \int_0^{f^{-1}(u)} \sqrt{f'(s)} \,\mathrm{d}W_s.$$

Then note that

$$[Z]_u = \int_0^{f^{-1}(u)} f'(s) \,\mathrm{d}s = u,$$

hence Z is a Brownian motion by Lévy. Let

$$\tau = \inf\{u \ge 0 \mid Z_u = K\},$$

then $\tau < \infty$ by properties of Brownian motion, and let $\sigma = f^{-1}(\tau)$. Finally let

$$\pi_t = \mathbb{1}_{\{t \le \sigma\}} \sqrt{f'(t)}.$$

Then we find that

$$\int_0^T \pi_s \,\mathrm{d}W_s = \int_0^\sigma \sqrt{f'(s)} \,\mathrm{d}W_s = Z_\tau = K,$$

and indeed

$$\int_0^T \pi_s^2 \,\mathrm{d}s = \tau < \infty.$$

(CHECK THIS)

5.2 Admissibility and Arbitrage

Definition 5.2. A trading strategy θ (a d-dimensional previsible process with the appropriate integrability conditions) is *x-admissible* if

$$X_t^{x,\theta} \geq 0 \qquad \text{a.s.}$$

for all $t \geq 0$.

This rules out the doubling strategy that we saw earlier, but it does not rule out suicide strategies.

Definition 5.3. An *arbitrage* is a 0-admissible trading strategy such that there exists $T > 0$ non-random such that

$$\mathbb{P}(X_T^{0,\theta} \geq 0) = 1, \qquad \mathbb{P}(X_T^{0,\theta} > 0) > 0.$$

Definition 5.4. An *equivalent local martingale measure* $\mathbb{Q}$ is equivalent to $\mathbb{P}$, such that S^i/B are $\mathbb{Q}$-local martingales.

Theorem 5.1 (Version of Fundamental Theorem of Asset Pricing)**.** *Suppose there exists an equivalent local martingale measure. Then there is no arbitrage.*

Proof: Let θ be 0-admissible, and set

$$X_t = \int_{)}^{t} \theta_s \cdot \mathrm{d}\left(\frac{S_s}{B_s}\right).$$

Let $\mathbb{Q}$ be an equivalent local martingale measure. Then X/B is a $\mathbb{Q}$-local martingale, since it is a stochastic integral with respect to a local martingale.

X is non-negative $\mathbb{P}$-almost surely, hence $\mathbb{Q}$-almost surely, so X is a supermartingale by Fatou. Hence

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{X_T}{B_T}\right] \leq \frac{X_0}{B_0} = 0,$$

so $X_T = 0$ $\mathbb{Q}$-almost surely, hence $\mathbb{P}$-almost surely.

Here admissibility, which ruled out doubling processes, allowed us to show that X was a supermartingale.

Our main example is

$$\mathrm{d}S_t^i = S_t^i \left(u_t^i \,\mathrm{d}t + \sum_{k=1}^{n} \sigma_t^{ik} \,\mathrm{d}W_t^k\right).$$

Some motivation from linear algebra: exactly one of the following is true, for a fixed b. Either there exists x with $Ax = b$, or there exists y with $A^T y = 0$ and $y^T b = 0$.

Then note

$$\begin{aligned} \mathrm{d}\left(\frac{S^i}{B}\right) &= \frac{S^i}{B}\left((u^i - r)\,\mathrm{d}t + \sum \sigma^{ik}\,\mathrm{d}W^k\right), \\ \mathrm{d}\left(\frac{S}{B}\right) &= \mathrm{diag}\left(\frac{S}{B}\right)\left((\mu - r\mathbf{1})\,\mathrm{d}t + \sigma\,\mathrm{d}W\right). \end{aligned}$$

If we want to make this a local martingale, we will use Girsanov. In the first case, suppose there is λ such that

$$\sigma\lambda = \mu - r\mathbf{1}.$$

Then

$$\mathrm{d}\left(\frac{S}{B}\right) = \mathrm{diag}\left(\frac{S}{B}\right)\sigma(\mathrm{d}W + \lambda\,\mathrm{d}t).$$

Theorem 5.2. *Suppose that*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \|\lambda\|^2\,\mathrm{d}s\right)\right] < \infty$$

for some $T > 0$ not random. Then set

$$Z_t = \mathcal{E}\left(-\int_0^t \lambda\,\mathrm{d}W\right).$$

$(Z_t)_{0\le t\le T}$ is a true martingale, so we can set

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = Z_T.$$

Then under $\mathbb{Q}$, the $(S/B)_{0\le t\le T}$ are local martingales.

Proof: Let $\hat{W} = W + \int \lambda\,\mathrm{d}t$. Then by CMG, $\hat{W}$ is a Brownian motion on $[0, T]$ under $\mathbb{Q}$.

So S/B are stochastic integrals with respect to $\mathbb{Q}$-local martingales, hence are $\mathbb{Q}$-local martingales themselves.

Now suppose we are in the second case, so

$$\sigma^T\pi = 0, \qquad \pi^T(\mu - r\mathbf{1}) > 0.$$

Then let

$$\theta = \operatorname{diag}\left(\frac{S}{B}\right)^{-1} \pi.$$

This gives

$$\mathrm{d}\left(\frac{X}{B}\right) = \pi^T(\mu - r\mathbf{1})\,\mathrm{d}t,$$

hence

$$\frac{X^{0,\theta}}{B} = \int_0^T \pi^T(\mu - r\mathbf{1})\,\mathrm{d}t > 0.$$

5.3 Contingent Claim Pricing

Suppose we are given a market with $1 + d$ assets, and we know the law of (B, S^i).

Now introduce a new asset, a *contingent claim*.

Example 5.1.

Consider $d = 1$. A (European) *call option* is the right but not the obligation to buy the asset at a fixed time T, at a fixed price K.

Consider the payout of the call: if $S_T > K$, then it is optimal to exercise the call, in which case we buy the asset for K, then immediately sell the asset to the market. Otherwise, if $S_T \leq K$ we do not exercise the call. Then the payout is

$$\text{payout} = \begin{cases} S_T - K & S_T > K, \\ 0 & S_T \leq K. \end{cases}$$

This is generally written as $(S_T - K)^+$.

Definition 5.5. A *European claim* is specified by an expiration date (or maturity), and a $\mathcal{F}_T$-measurable random variable ξ modelling the payout at maturity.

A European claim is *vanilla* if

$$\xi = g(S_T^1, \ldots, S_T^d),$$

for a non-random payout function g, that only depends on the end prices of the assets.

Calls are vanilla since $\xi = g(S_T)$, where

$$g(x) = (x - K)^+.$$

Theorem 5.3. *Consider a market with prices (B, S^i) and let $\mathbb{Q}$ be a local martingale measure. Let (ξ_t) be the price of a contingent claim.*

The augmented market with prices (B, S, ξ) has no arbitrage if ξ/B is a $\mathbb{Q}$-local martingale.

Proof: This is just the definition of $\mathbb{Q}$ being a local martingale measure for the augmented market.

From the fundamental theorem of asset pricing, we get no arbitrage.

Example 5.2.

Consider a European claim with time T payout ξ_T. If we set

$$\xi_T = \mathbb{E}^{\mathbb{Q}}\left[\frac{\xi_T}{B_T}|\mathcal{F}_T\right],$$

then we have no arbitrage.

For terminology, we let (X/B) be X discounted by B.

Example 5.3. (Black-Scholes)

Consider $d = 1$, and

$$\begin{aligned} \mathrm{d}B &= Br\,\mathrm{d}t, \\ \mathrm{d}S &= S(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W), \end{aligned}$$

where μ, σ are constants, and W is a Brownian motion. Our goal is to price a European call with payout

$$\xi_T = (S_T - K)^+.$$

Note that

$$\begin{aligned} \mathrm{d}\left(\frac{S}{B}\right) &= \frac{S}{B}\left((\mu - r)\,\mathrm{d}t + \sigma\,\mathrm{d}W\right) \\ &= \frac{S}{B}\sigma\left(\mathrm{d}W + \left(\frac{\mu - r}{\sigma}\right)\mathrm{d}t\right), \end{aligned}$$

hence

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \exp\left(-\left(\frac{\mu - r}{\sigma}\right)W_T - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 T\right)$$

is the equivalent local martingale measure for the market (B_t, S_t), since

$$\hat{W}_t = W_t + \left(\frac{\mu - r}{\sigma}\right) t$$

is a $\mathbb{Q}$-Brownian motion by CMG.

Let g be of polynomial growth, and suppose that

$$\xi_T = g(S_T).$$

Using Itô's formula, we find

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) + \sigma W_t\right) = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right) t + \sigma \hat{W}_t\right).$$

Now we calculate

$$\begin{aligned}\xi_t &= B_t \mathbb{E}^{\mathbb{Q}}\left[\frac{g(S_T)}{B_T} \mid \mathcal{F}_t\right] \\ &= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}\left[g\left(S_t \exp\left(\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma(\hat{W}_T - \hat{W}_t)\right)\right) \mid \mathcal{F}_t\right].\end{aligned}$$

So expanding,

$$\xi_t = e^{-r(T-t)} \int g(S_t e^{(r - \frac{\sigma^2}{2})(T-t) + \sigma\sqrt{T-t} z}) \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, \mathrm{d}z = V(t, S_t).$$

Recall that, if a function V satisfies

$$\frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} = rV$$

with $V(T, s) = g(s)$ for all s, then

$$e^{-rt} V(t, S_t)$$

is a $\mathbb{Q}$-local martingale by Itô's formula. If it is a true martingale, then

$$V(t, s) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[g(S_T) | S_T = s].$$

Index