Stochastic Calculus

Jan Olucha Fuentes

May 18, 2025

Dear Reader,

This is a set of lecture notes typed for the course *Stochastic Calculus with applications to Finance* taught at the University of Cambridge during the academic year 2024-2025. Prerequisites for this document include the contents of a second course in Probability Theory: Conditional Expectation, Martingales, Brownian Motion, etc. in addition to a strong will to live.
Yours falsely,

JOF.

Contents

Notation and how to study these notes

Difficulty notations:
☕: denotes a proof or idea that is hard or hard to reproduce.
♪: denotes a proof or idea that requires partially unmotivated tools or machinery that is hard to pull out of thin air, but the rest of the argument is easy.
♩: denotes a proof or idea that is easy to see knowing some other results that help motivate it.
♫: denotes a proof or idea that is easy and no particularly clever ideas are needed. For revision purposes reading it a couple times will suffice.

Other notation:
$\mathrm{m}\mathscr{F}$: the set of $\mathscr{F}$-measurable functions.
$\mathrm{b}\mathscr{F}$: the set of $\mathscr{F}$-measurable and bounded functions.
$\mathbf{R}_+$: the positive real numbers and zero.
"increasing": non-decreasing, we will use "strictly increasing" when we mean that.
$\{\mathscr{F}_t\}_t$: a filtration. I am inconsistent with notation so sometimes I refer to $(\mathscr{F}_t)_t$ as a filtration, same thing for processes or any sequence really.
$\boldsymbol{B}(x,r)$: the open ball of radius r and center x.

Chapter 1

Motivation: towards the Stochastic Integral

Definition 1.0.1 (Markov Process, Transition Kernels) A Markov Process $(X_t : t \geqslant 0)$ is a Real valued $\mathscr{F}_t$ adapted stochastic process that satisfies the **Markov Property**, i.e: for any $f \in \mathrm{b}\mathscr{F}$ and any $0 \leqslant s \leqslant t$,

$$\boldsymbol{E}[f(X_t) \mid \mathscr{F}_s] = \boldsymbol{E}[f(X_t) \mid X_s]. \tag{MP}$$

For a Markov Process (X_t), we define the transition operator P_t as the map given by

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

Where $p_t(x, \cdot) = \boldsymbol{P}[X_t \in \cdot \mid X_0 = x]$ is the **transition kernel**.

Explanation as to why we want to give a meaning to

$$\int_0^t f(X_s) d W_s$$

An analogy of this integral would be to consider a "discrete stochastic integral", of the following shape, let $a \in \mathbf{R}^\infty$ be a sequence, and (ξ_i) be a sequence of independent and identically distributed Ber$(1/2)$ random variables when does the stochastic sum

$$S_n = \sum_{i=1}^{n} a_i \xi_i$$

converge? Of course a trivial answer would be if $a \in \ell^1$, then

$$\left| \sum_{i=1}^{\infty} a_i \xi_i \right| \leqslant \sum_{i=1}^{\infty} |a_i \xi_i| = \sum_{i=1}^{\infty} |a_i| < \infty.$$

However, we want some slightly more general results to motivate the construction of our stochastic integral.

Proposition 1.0.2 Let $a \in \ell^2$ be a deterministic sequence, then S_n as defined above converges in $\mathscr{L}^2$ and almost surely.

Proof. First note that (S_n) is an $\mathscr{F}_n$ Martingale, where $\{\mathscr{F}_n\}$ is the filtration generated by $\xi_1, \cdots, \xi_n$. Moreover, we claim that (S_n) is $\mathscr{L}^2$ bounded. This is not hard to see:

$$\boldsymbol{E}[|S_n|^2] = \boldsymbol{E}[S_n^2] \overset{(!)}{=} \sum_{i=1}^{n} (a_i \xi_i)^2 = \sum_{i=1}^{n} a_i^2 < \sum_{i=1}^{\infty} a_i < \infty$$

for all n, and so (S_n) is $\mathscr{L}^2$ bounded and so by Doob's $\mathscr{L}^p$ Convergence Theorem, we have that (S_n) converges in $\mathscr{L}^2$ and almost surely. (In step $(!)$ we used the fact that $\boldsymbol{E}[\xi_i \xi_j] = 0$ for $i \neq j$) ♡

Here's a Theorem that previews an argument that will be often used: **the trick of localisation**.

Theorem 1.0.3 Let (a_n) be an $\mathscr{F}_n$-previsible process with respect to the natural filtration $\mathscr{F}^{\xi}$. Moreover suppose that $\sum_n a_n^2 < \infty$ almost surely. Then S_n converges almost surely.

Proof. Consider the following random variable:

$$T_N = \inf \left\{ n \geqslant 1 : \sum_{k=1}^{n+1} a_k^2 > N \right\}$$

Since the event $\{T_N \leqslant n\}$ can be determined by looking at $a_1, \cdots, a_{n+1}$ and (a_n) is a previsible process, all these $a_k's$ are $\mathscr{F}_n$-measurable and so T_N is an $\mathscr{F}_n$ stopping time. Recall that (S_n) is an $\mathscr{F}_n$-Martingale (Indeed, from previsibility, we have that $\boldsymbol{E}[a_n \xi_n \mid \mathscr{F}_{n-1}] = a_n \boldsymbol{E}[\xi_n \mid \mathscr{F}_{n-1}]$) so by the first part of the OST, it follows that $S_n^{T_N} := S_{n \wedge T_N}$ is also a Martingale. We are going to show that $\lim_{n \to \infty} S_{n \wedge T_N} = S_{\infty \wedge T_N}$ exists. For this purpose, we simply show that $S_n^{T_N}$ is $\mathscr{L}^2$-bounded.

$$\begin{aligned}
\boldsymbol{E}\left[\left(S_n^{T_N}\right)^2\right] &= \boldsymbol{E}\left[\left(\sum_{k=1}^{n} a_k \xi_k \mathbf{1}(k \leqslant T_n)\right)^2\right] && (1.1) \\
&= \boldsymbol{E}\left[\sum_{k=1}^{n} a_k^2 \mathbf{1}(k \leqslant T_n)\right] && (1.2) \\
&\leqslant N && (1.3)
\end{aligned}$$

Where step (1) comes from definition, step (2) comes from the fact that a_k and ξ_k are independent as well as the fact that the (ξ_k) are independent. Step (3) comes from the fact that on the event $\mathbf{1}(k \leqslant T_n)$, you have that $a_1^2 + \cdots + a_k^2 \leqslant N$. (Here is why we needed to define T_N as a sum to $n+1$

and hence require previsibility). Therefore the stopped Martingale is bounded and hence we have that $S_{n \wedge T_N} \to S_{\infty \wedge T_N}$ almost surely and in $\mathscr{L}^2$. Now, if we could show that for some N, T_N was equal to infinity, then we would actually have that $S_n \to S_\infty$ almost surely and we would be done. However, this is actually possible because by assumption:

$$1 = \boldsymbol{P}\left[\sum_n a_n^2 < \infty\right]$$

But the event $\left\{\sum_n a_n^2 < \infty\right\}$ equals the event $\left\{\bigcup_N \sum_n a_n^2 \leqslant N\right\}$ which in turn equals the event $\{\bigcup_N T_N = \infty\}$. This means that with probability 1, there exists some N for which $T_N = \infty$, thus finishing our claim. ♡

Chapter 2

Finite variation and the Lebesgue-Stieltjes Integral

2.1 The Lebesgue-Stieltjes Integral

Let's start by recalling an elementary concept:

Definition 2.1.1 (Distribution Function) A function $F : \mathbf{R}_+ \to \mathbf{R}$ is said to be a distribution function if it is right continuous and non-decreasing/increasing.

Of course we have the following "obvious correspondence": if μ is a measure on $\mathbf{R}$, then we can build a distribution function as $F(x) = \mu(-\infty, x]$, the converse also holds.

Proposition 2.1.2 (Correspondence between measures and distribution functions) Let $F(x)$ be a distribution function, then there is a unique measure μ with $\mu(0, x] = F(x) - F(0)$.

Main idea: We can "reverse engineer" the measure μ. Since we have the existence of Lebesgue measure, we can try to set $\mu(0, x] = \text{Leb}(0, F(x) - F(0)]$, it may not be immediately clear this is a measure, but since the interval $(0, F(x) - F(0)]$ can be written as $G^{-1}(0, x]$ where $G(y)$ is the generalised inverse of F, and G is measurable, we are done.

Proof. The proof goes by considering the generalised inverse of F, i.e:

$$G(y) = \inf\{t : F(t) - F(0) \geqslant y\}$$

Then it is clear that $t \geqslant G(y)$ if and only if $F(t) - F(0) \geqslant y$. Then we can construct μ as follows:

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

Then

$$\mu(0,t] = \text{Leb}\{y : 0 < G(y) \leqslant t\} = \text{Leb}\{0 < y < F(t) - F(0)\} = F(t) - F(0)$$

μ is obviously a measure as it is the composition of the set inverse of a measurable function and a measure, and it it unique as it is defined a the generating π-system. ♡

Definition 2.1.3 (Local integrability) Let μ be a measure, a function $g : \mathbf{R}_+ \to \mathbf{R}$ is said to be locally μ integrable if for any $t \geqslant 0$, the function $g\,\mathbf{1}_{[0,t]} \in \mathscr{L}^1(\mu)$.

Definition 2.1.4 (Lebesgue-Stieltjes integral) Let F be a distribution function with corresponding measure μ, then a function g is said to be (locally) F-integrable if and only if it is (locally) μ-integrable, and in this case, we define (or rather use the notation)

$$\int_0^t g(s)\,\mathrm{d}F(s) = \int_0^t g(s)\mu(\mathrm{d}y)$$

Proposition 2.1.5 Let F be a distribution function and g a locally F-integrable function, then

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

is cadlag. Moreover, if F is continuous, then so is I.

Main idea: Simple application of the DCT.

Proof. Let us show right continuity of $I(t)$. The goal is to show that $\lim_{\epsilon\downarrow 0} I(t+\epsilon) = I(t)$. Since we are taking $\epsilon \downarrow 0$, without loss of generality, we can find some T large enough such that $t + \epsilon < T$. And since

$$|I(t+\epsilon)| = \left|\int_0^{t+\epsilon} g(s)\,\mathrm{d}F(s)\right| = \left|\int_{\mathbf{R}} g(s)\,\mathbf{1}([0,t+\epsilon])(s)\,\mathrm{d}F(s)\right| \leqslant \int_{\mathbf{R}} |g(s)\,\mathbf{1}([0,T])|(s)\,\mathrm{d}F(s) < \infty,$$

we have by the Dominated Convergence Theorem that

$$\lim_{\epsilon\downarrow 0} I(t+\epsilon) = \int_{\mathbf{R}} g(s)\,\mathbf{1}([0,t])(s)\,\mathrm{d}F(s) = I(t).$$

Then to show the existence of left-limits, we perform a similar argument: For this, we note that

$$I(t-\epsilon) = \int g\,\mathbf{1}_{[0,t]} - g\,\mathbf{1}_{(t-\epsilon,t]}\,\mathrm{d}F(s)$$

we can brutally bound the integrand say by $2g\mathbf{1}_{[0,t]}$, which is integrable by local integrability of g so we can pass the limit again using the DCT, and we get that

$$\lim_{\epsilon\to 0} I(t-\epsilon) = \int_0^t g(s)\mathrm{d}F(s) - g(t)\mu\{t\}$$

thus the left limit exists, and moreover, if F is continuous, then it has no atoms, and so the measure of singletons is zero, hence I is continuous. ♡

2.2 Functions of bounded Variation

Definition 2.2.1 (Function of finite variation) For a function $f:\mathbf{R}_+\to\mathbf{R}$, we define its variation $V_f(t)$ as

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

Where $x_k^N = k2^{-N}$. We say that a function f is of finite variation if $V_f(t) < \infty$ for all t, and we say that it is of bounded variation if $\sup_t V_f(t) < \infty$.

The intuition is that $V_f(t)$ measures the total oscillation of the function, by partitioning the range $[0,t]$ into a mesh of step-size 2^{-N} and then measuring the variation of the function along that step, and then adding up over all steps, and then maximising over mesh size.

Theorem 2.2.2 (Variation of càdlàg functions of finite variation is a distribution function) Let $f:\mathbf{R}_+\to\mathbf{R}$ be a cadlag function of finite variation, then $V=V_f$ is a distribution function that satisfies the bound $(\star)$:

$$V(t) - V(s) \geqslant |f(t) - f(s)|$$

Main idea: We first show that $V^N(t)$ is increasing in N, this allows us to show inequality $(\star)$ by consider how the sums along the dyadic partition telescope. Inequality $(\star)$ implies the increasing property. To show right continuity, you use inequality $(\star)$ in the telescoped sum $V^N(t) - V^N(s)$ as well as increasing property.

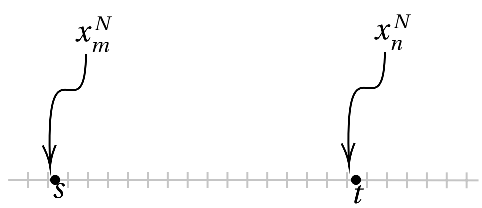
Proof. We first prove that V is a distribution function, for this, let V^N denote the variation along the mesh of size 2^{-N}, then:

$$\begin{aligned}
V^N(t) &:= \sum_{k=1}^{\infty} \left| f\left(x_k^N \wedge t\right) - f\left(x_{k-1}^N \wedge t\right) \right| && (2.1)\\
&\leqslant \sum_{k=1}^{\infty} \left| f\left(x_k^N \wedge t\right) - f\left(x_{2k-1}^{N+1} \wedge t\right) \right| + \left| f\left(x_{2k-1}^{N+1} \wedge t\right) - f\left(x_{k-1}^N \wedge t\right) \right| && (2.2)\\
&= \sum_{k=1}^{\infty} \left| f\left(x_{2k}^{N+1} \wedge t\right) - f\left(x_{2k-1}^{N+1} \wedge t\right) \right| + \left| f\left(x_{2k-1}^{N+1} \wedge t\right) - f\left(x_{2k-2}^{N+1} \wedge t\right) \right| && (2.3)\\
&= V^{N+1}(t) && (2.4)
\end{aligned}$$

Where step (2) is the triangle inequality, step (3) is due to the fact that $x_k^N := k2^{-N} = 2k2^{-N-1}$, and step (4), which requires a bit more thinking comes from the fact that the sum in (3) is accounting on the one hand, the variation over the even intervals, whereas the second term accounts for the variation on the odd intervals. We thus have shown that $V^N \leqslant V^{N+1}$ which as a consequence implies that $V = \lim_{N\to\infty} V^N$, this will help us in showing the inequality $(\star)$. The idea is that since

$$V(t) - V(s) = \lim_{N\to\infty} V^N(t) - V^N(s),$$

and these latter terms are just sums, there will be a good deal of cancelling out of terms, and the few terms that remain, we will be able to control using the fact that f is cadlag. Observe the diagram that says it all:



We consider the dyadic partition of our interval, and without loss of generality suppose that for a fixed N, we have some m and n, such that $x_m^N \leqslant s < x_{m+1}^N$ and $x_n^N \leqslant t < x_{n+1}^N$. The key is that in $V^N(t) - V^N(s)$, (almost) everything to the left of s will get cancelled out. In fact, by staring at the diagram for a bit and looking at the definition of V^N, we have that

$$V^N(t) - V^N(s) = \left| f(t) - f\left(x_n^N\right) \right| + \sum_{k=m+2}^{n} \left| f\left(x_k^N\right) - f\left(x_{k-1}^N\right) \right| + \left| f\left(x_{m+1}^N\right) - f\left(x_m^N\right) \right| - \left| f(s) - f\left(x_m^N\right) \right|$$

However, note that in the limit,

$$\lim_{N\to\infty} \left| f\left(x_{m+1}^N\right) - f\left(x_m^N\right) \right| - \left| f(s) - f\left(x_m^N\right) \right| = 0$$

because $\lim_{N\to\infty} f\left(x_{m+1}^N\right) = f(s)$ due to right continuity, and $\lim_{N\to\infty} f\left(x_m^N\right)$ is some finite number due to left limits existing. This gives that

$$\lim_{N\to\infty} V^N(t) - V^N(s) = \lim_{N\to\infty} \left|f(t) - f\left(x_n^N\right)\right| + \sum_{k=m+2}^{n} \left|f\left(x_k^N\right) - f\left(x_{k-1}^N\right)\right| \tag{2.5}$$

$$\geqslant \lim_{N\to\infty} \left|f(t) - f\left(x_{m+1}^N\right)\right| = |f(t) - f(s)| \tag{2.6}$$

Where step (6) comes from the triangle inequality, that allows us to bash all the terms into a single absolute value sign, and then telescoping occurs in there. The final equality comes again due to right continuity of f. We have established the inequality $(\star)$, and incidently, we have shown how V is increasing.

All there is left is to show that V is right continuous. To do this, we employ $(\star)$, and so by looking at equation (2.5) and plugging in $(\star)$, we have a telescoping that reads

$$V(t) - V(s) = \lim_{N\to\infty} \left|f(t) - f\left(x_n^N\right)\right| + \sum_{k=m+2}^{n} \left|f\left(x_k^N\right) - f\left(x_{k-1}^N\right)\right| \leqslant V(t) - \lim_{N\to\infty} V(x_{m+1}^N)$$

Rearranging we have that

$$\lim_{N\to\infty} V(x_{m+1}^N) \leqslant V(s)$$

but since V is increasing, we also have that $V(s) \leqslant \lim_{N\to\infty} V(x_{m+1}^N)$, because by assumption x_{m+1}^N is chosen to be strictly greater than s. Thus

$$\lim_{N\to\infty} V(x_{m+1}^N) = V(s)$$

and as such we have right continuity. (I mean, if you want full details now you should pick an ϵ and use the limit above to find a dyadic that's close enough to s so that the difference is at most ϵ and then use that distance from the dyadic to s as your δ) ♡

Remark 2.2.3 (Total Variation) Let $f : \mathbf{R}_+ \to \mathbf{R}$, we define its total variation as

$$\|f\|_{\mathrm{TV}} = \sup_{0\leqslant x_0 < x_1 < \cdots < x_k = t} \left\{ \sum_{n=1}^{k} |f(x_n) - f(x_{n-1})| \right\}$$

i.e: its highest variation over all partitions of the interval $[0, t]$. Naturally, since in $V_f(t)$ we only consider one partition, we have that $V_f(t) \leqslant \|f\|_{\mathrm{TV}}(t)$. However, from the inequality that we derived before:

$$V(t) - V(s) \geqslant |f(t) - f(s)|$$

We see by combining it into the definition of total variation, that in fact we will get a telescoping sum and get

$$\|f\|_{\mathrm{TV}}(t) \leqslant V_f(t) - V_f(0) = V_f(t)$$

Therefore we have that in fact the total variation is equal to the variation we used above, which is a powerful observation as it tells us that it is in fact enough to consider the variation along dyadic intervals. This will have major measurability benefits later on.

Recall that our goal here is to make sense of integrals with respect to some random process. For this we have first defined an integral with respect to a distribution function, but in general, the stochastic processes we want to integrate against are not distribution functions. However, as we will see now, we can make sense of some distribution function out of càdlàg functions with finite variation, and hence meaningfully define the integral against a càdlàg function

Proposition 2.2.4 (Càdlàg functions and distribution functions) If a function f is càdlàg and of finite variation then it can be written as

$$f = f^{\uparrow} - f^{\downarrow}$$

Where $f^{\uparrow}$ and $f^{\downarrow}$ are both distribution functions.

Proof. First suppose that f is càdlàg and of finite variation, then, from the inequality obtained in Theorem 2.2.2, we have that for $0 \leqslant s \leqslant t$:

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

That is to say, one the one hand:

$$V_f(t) - V_f(s) \geqslant f(t) - f(s)$$

showing that $V_f(t) - f(t)$ is increasing in t, and similarly, by choosing the other item in the $\max$, we have that $V_f(t) + f(t)$ is increasing in t. Since $V_f(t)$ is right continuous and $f(t)$ is càdlàg, it follows that both $V_f(t) + f(t)$ and $V_f(t) - f(t)$ are right continuous. We may then define

$$f^{\uparrow} = \frac{1}{2}\left(V_f(t) + f(t)\right) \qquad f^{\downarrow} = \frac{1}{2}\left(V_f(t) - f(t)\right)$$

We have just argued that $f^{\uparrow}$ and $f^{\downarrow}$ are increasing and right continuous, i.e: distribution functions, and moreover f is clearly equal to their difference. ♡

We now show a Proposition which will make our integral against a càdlàg function well-defined:

Proposition 2.2.5 (Integral with respect to $f^\uparrow - f^\downarrow$) Let f be a càdlàg function of finite variation with decomposition $f = f^\uparrow - f^\downarrow$. Then if a function $g : \mathbf{R}_+ \to \mathbf{R}$ is locally V_f integrable, then it is both locally $f^\uparrow$ and $f^\downarrow$ integrable.

Proof. This is just a one-line proof that follows after noting that $V_f = f^\uparrow + f^\downarrow$:

$$\max\left\{\int |g|\,\mathrm{d}f^\uparrow, \int |g|\,\mathrm{d}f^\downarrow\right\} \leqslant \int |g|\,\mathrm{d}f^\uparrow + \int |g|\,\mathrm{d}f^\downarrow = \int |g|\,\mathrm{d}V_f$$

♡

Now we can finally define the following:

Definition 2.2.6 (Integral with respect to càdlàg function of finite variation) Let f be càdlàg function of finite variation, then we say that a function g is locally f integrable if it is locally V_f integrable. In this case, we write

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

and also

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

We now need one last ingredient before we can finish this deterministic part of the theory and move to the random part

Theorem 2.2.7 Let f be a càdlàg function of total variation, and g a locally f integrable function, then we have that the function $I(t)$ defined as

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

is itself càdlàg of finite variation.

Proof. Since we can expand

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

and both $f^\uparrow$ and $f^\downarrow$ are distribution functions, one readily verifies from Proposition 2.1.5 that $I(t)$ is càdlàg. We now just need to show that it is of finite variation. We can now do the following cheeky trick:

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

since all g^+ and g^- are positive, it is clear that $I(t)$ is a difference of distribution functions (Indeed the things inside the parenthesis are both càdlàg and non-decreasing by positivity of g^+ and g^-). It is now easy to check that if a function $I(t)$ can be written as a difference of two distribution functions $F_1(t) - F_2(t)$, then $I(t)$ is of finite variation. This is because by the triangle inequality

$$V_I(t) \leqslant V_{F_1}(t) + V_{F_2}(t)$$

and since both F_1 and F_2 are non-decreasing, it will follow that $V_{F_1}(t) = F_1(t) - F_1(0)$ as all but two of the terms in the sum that defines variation will telescope. ♡

2.3 Finite Variation Processes and Previsible Processes

We are ready to make sense of a sort of proto-stochastic integral. After all, a stochastic process is nothing but a random function $\mathbf{R}^+ \to \mathbf{R}$, and we have seen that if we want to integrate against functions, the right class of functions to integrate against are càdlàg functions of finite variation. It is of no surprise then that the stochastic processes we will use as integrators will have sample paths that satisfy these two properties. From now on fix a background filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_t, \boldsymbol{P})$.

Definition 2.3.1 (Finite variation process) A stochastic process (Z_t) is called a finite variation process if each trajectory is càdlàg and of finite variation. That is to say, for all outcomes ω in our sample space Ω, we have that the map $t \mapsto Z_t(\omega)$ is càdlàg and of finite variation.

For reasons that will become clear in due course, the right class of integrands to consider are previsible processes.

Definition 2.3.2 (Previsible σ-algebra, previsible process) The previsible σ-algebra $\mathscr{P}$ is the σ-algebra generated by sets of the form

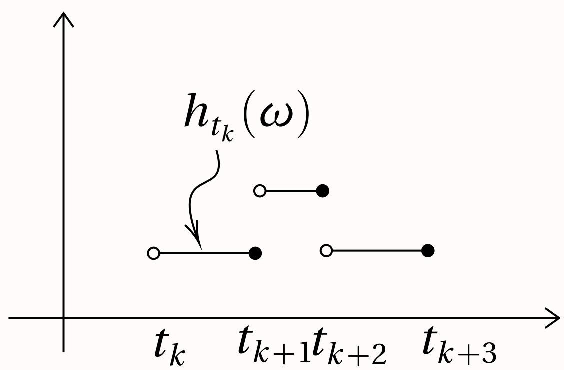
$$(s,t] \times A$$

For some value of $s < t$ and $A \in \mathscr{F}_s$. A process $(H_t)_t$ is called previsible if the map $(t, \omega) \mapsto H_t(\omega)$ is $\mathscr{P}$-measurable.

Remark 2.3.3 (Huh?) It is probably very unclear from this definition why this σ-algebra should bear the adjective "previsible", or why processes that are adapted to this σ-algebra should also be called previsible. The best way to understand this is by looking at the kind of processes that live in this sigma-algebra. Let us illustrate this via an example. Let (h_t) be an $\mathscr{F}_t$-adapted process, and let $0 < t_0 < t_1 < \cdots$ be a sequence of deterministic times. The claim is that the process

$$H_t(\omega) = \sum_{k=0}^{\infty} h_{t_k}(\omega)\,\mathbf{1}(t \in (t_k, t_{k+1}])$$

is previsible. This should not be too hard to see, for (h_t) being $\mathscr{F}_t$ adapted means that we can express h_{t_k} as a pointwise limit of indicator functions of $\mathscr{F}_{t_k}$-measurable events, and and so when multiplying out we will write $H_t(\omega)$ as a pointwise limit of a weighted sum indicator functions of the form $\mathbf{1}(A \cap (t_k, t_{k+1}]$ for $A \in \mathscr{F}_{t_k}$. Each of these sets is in $\mathscr{P}$. Now that we have established that H_t is previsible, we see that actually the behavior of H_t coincides with our intuition of what previsible should mean, because say by looking at how the graph of $H_t(\omega)$ could look like, we see that the value of H_t can be determined by looking at times just slightly to the left of t.



Moreover, it is easy to convince yourself that any left-continuous function can be realised as a pointwise limit of such things, and since measurability is preserved under pointwise limits, it follows that adapted processes with left continuous sample paths are in fact previsible, and this agrees with our intuition, because if the sample path is left continuous, then I can determine the value of the process at time t by knowing what happens slightly just before! It turns out, although we will not prove this, that $\mathscr{P}$ can be equivalently defined by the sigma algebra generated by all left-continuous adapted processes.

Theorem 2.3.4 (Integral of previsible process against process of finite variation) Let Z be a finite variation process and let H be a previsible process such that $H(\omega)$ is locally integrable against $Z_s(\omega)$ for all ω. Then the process

$$X_t(\omega) = \int_0^t H_s(\omega)\, \mathrm{d}Z_s(\omega) \tag{2.7}$$

is a process of finite variation.

Remark 2.3.5 Of course, by the deterministic theory we developed earlier, we already know that for a fixed ω, $X_t(\omega)$ is càdlàg and has finite variation, this was what Theorem 2.2.7 told us, the only thing that is left for us to prove is that this process is actually adapted! And this is where the previsibility of H will come into play. The proof relies on the Monotone Class Theorem, which we add in a Measure-Theoretic Appendix

Proof. We will start by showing that if H is previsible and bounded, then the integral in (3.1) is $\mathrm{m}\mathscr{F}_t$. Then we will extend to the general case. As hinted in the remark, the key is the use of the Monotone Class Theorem, for this we define the class of functions

$$\mathscr{H} = \left\{ H : \mathbf{R}_+ \times \Omega \to \mathbf{R} \text{ bounded with } \int_0^t H_s\, \mathrm{d}Z_s \in \mathrm{m}\mathscr{F}_t \right\}$$

The Monotone Class Theorem will allow us to concludet that actually $\mathrm{b}\mathscr{P} \subseteq \mathscr{H}$, thus concluding the first part of the proof. For this, we now tick the boxes of the Montone Class Theorem:

- The fact that $\mathscr{H}$ is a vector-space is obvious.
- Let $A \in \mathscr{G}$, where $\mathscr{G}$ is the generating π-system of $\mathscr{P}$. then we can write $A = (t_0, t_1] \times B$ for some $B \in \mathscr{F}_{t_0}$, then:

$$\begin{aligned}
\int_0^t \mathbf{1}_{(t_0,t_1]\times B}\, \mathrm{d}Z_s &= \mathbf{1}_B \int_{t_0\wedge t}^{t_1\wedge t} \mathrm{d}Z_s \\
&= \mathbf{1}_B \left(\int_{t_0\wedge t}^{t_1\wedge t} \mathrm{d}Z_s^{\uparrow} - \int_{t_0\wedge t}^{t_1\wedge t} \mathrm{d}Z_s^{\downarrow} \right) \\
&\frac{\mathbf{1}_B}{2}(V_f(t_1\wedge t) + Z_{t_1\wedge t} - V_f(t_0\wedge t) - Z(t_0\wedge t) \\
&- V_f(t_1\wedge t) + Z(t_1\wedge t) + V_f(t_0\wedge t) - Z(t_0\wedge t)) \\
&= \mathbf{1}_B \left(Z_{t_1\wedge t} - Z_{t_0\wedge t} \right)
\end{aligned}$$

  Now we observe that the resulting $\mathbf{1}_B\left(Z_{t_1\wedge t} - Z_{t_0\wedge t}\right)$ is $m\mathscr{F}_t$, because $B \in \mathscr{F}$, and $(Z_s)_s$ is

$(\mathscr{F}_s)_s$-adapted.

- Let $(H^n)_n$ be a sequence of bounded processes with $H^n \in \mathscr{H}$ for all n, such that they increase to a bounded process H, let us show that $H \in \mathscr{H}$. For this we very simply note that

$$\int_0^t H_s \, \mathrm{d}Z_s = \lim_{n\to\infty} \int_0^t H_s^n \, \mathrm{d}Z_s$$

Where we pulled the limit out because the sequence (H_s) is uniformly bounded by H which is itself a bounded process. Therefore we applied the DCT. Now finally, each of these integrals on the right hand side are an $\mathscr{F}_t$ measurable object by hypothesis on (H^n). Therefore we see that $H \in \mathscr{H}$.

Thus we have ticket all the boxes of the Monotone Convergence Theorem and it follows that whenever H is a bounded previsible process, we have that the integral in (3.1) is an $\mathscr{F}_t$ measurable object. Now we extend this to unbounded previsible processes H. The idea here is simply to define $H^n = (H \wedge n) \vee (-n)$, this makes H^n a bounded processes, and so for each n, we have that

$$\int_0^t H_s^n \, \mathrm{d}Z_s \in \mathrm{m}\mathscr{F}_t \, .$$

Moreover, since $|H_s^n| \leqslant |H_s|$ for all s, and $(H_s(\omega))$ is $(\boldsymbol{Z}_s(\omega))_s$ integrable by hypothesis, we can apply the Dominated Convergence Theorem, and we have that

$$\int_0^t H_s \, \mathrm{d}Z_s = \lim_{n\to\infty} \int_0^t H_s^n \, \mathrm{d}Z_s$$

and since the integral of the right hand side is $\mathrm{m}\mathscr{F}_t$, and measurability is preserved by pointwise limits, we are done. ♡

Chapter 3

The Stochastic Integral

We have so far constructed a primitive proto-stochastic integral:

$$I(t) = \int_0^t H_s \, \mathrm{d}Z_s$$

where (Z_s) is a process of finite variation and H_s is a previsible process, and we saw that $I(t)$ satisfies some desirable properties, i.e: being itself a process of finite variation. As it will turn out, most Martingales actually have infinite variation, so we will have a bit of a hard time defining the stochastic integral this way. As we are going to see in the next chapter, the right integrators will be local continuous Martingales M, and the right integrands will be previsible process K. In particular, we will see that if K and M satisfy some particular compatibility conditions, the stochastic integral

$$\int_0^t K_s \, \mathrm{d}M_s$$

can be defined as itself a continuous local Martingale. In this chapter we will also see that every continuous local Martingale M has associated to it an adapted continuous increasing process $[M]$, called the quadratic variation, which captures the "roughness accumulated through time" by the Martingale. It turns out that the compatibility condition we will seek is that

$$\int_0^t K_s^2 \mathrm{d}[M]_s < \infty$$

and this integral can be interpreted using the Lebesgue-Stieltjes Theory we developed earlier. That's the outline of the chapter, we will start by making sense of local Martingales, but we first need to make a slight measurability detour.

3.1 The Usual Conditions

Local Martingales are all about stopping times, so let us talk about them for a bit.

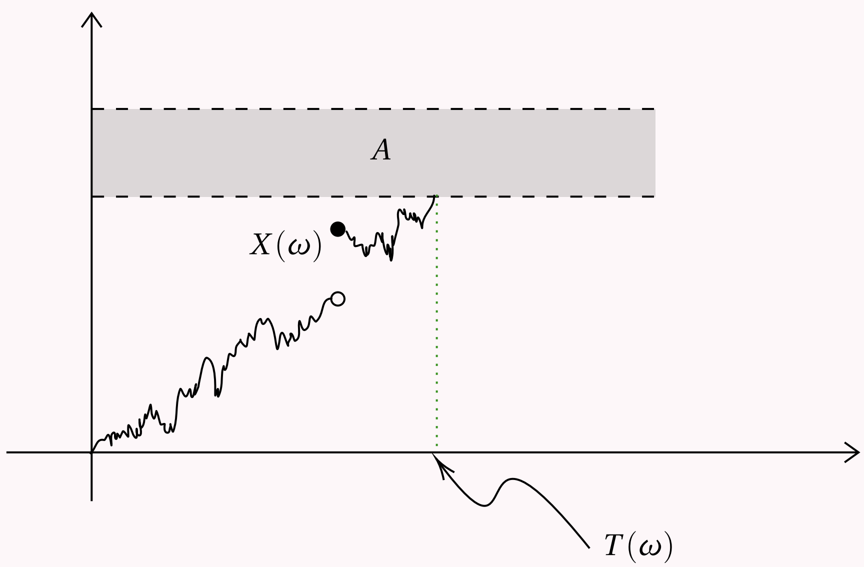
Definition 3.1.1 (Stopping times) A random variable $T : \Omega \to [0, \infty]$ is called a stopping time with respect to the filtration $\{\mathscr{F}_s\}_s$ if $\{T \leqslant t\} \in \mathscr{F}_t$.

i.e: informally, a random clock for which we can know if it has rung by time t by looking at all the information available to us at that time contained in the filtration. From a course in Probability we know how insanely useful stopping times are to gain information about all kinds of characteristics of our stopping process. The problem is however, is that when we deal with continuous time, unless we make some slight readjustments to our sigma algebra, we lose a lot of stopping times. Let us see an example which the reader may have seen already:

Example 3.1.2 (A lost stopping time) Let $(X_s)_s$ be a right continuous process and let $A \subseteq \mathbf{R}$ be an open set. Define the random variable

$$T = \inf\{t \geqslant 0 : X_t \in A\}$$

Look at the following diagram :



Then it is clear that T equals to the value pointed out in the diagram, **because it is defined as an infimum**. However, since A is open, the process (X_s) actually doesn't know it has hit A by that time! But rather, (X_s) is just standing right in front of it. Therefore T is in fact not a stopping time!

What could be an easy fix to make sure that in this example T is in fact a stopping time? Well we have two options, one - forcing A to be closed - which might seem like cheating because we are going to rid ourselves to plenty of scenarios where we could need an open set! The other solution, which will surely not leave the reader unfazed, is actually to allow the process to look just very slightly into the future! Now I bet this feels like cheating, but as it turns out, we will be able to prove that if X is a Martingale with respect to some filtration $\{\mathscr{F}\}$, then it will also still be a Martingale with respect to our "cheater's" sigma-algebra, thus in fact, we will not have broken the rules! Let us formalise this notion of the "cheater's sigma algebra".

Definition 3.1.3 (Usual conditions) A filtration $(\mathscr{F}_t)_t$ satisfies the usual conditions if:

- $\mathscr{F}_0$ contains all $\boldsymbol{P}-$null sets. That is to say:

$$\{A \in \mathscr{F} : \boldsymbol{P}[A] = 0\} \subseteq \mathscr{F}_0.$$

- It is right continuous:

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

Remark 3.1.4 Of course notice that the part that incorporates our cheater's behavior is the inclusion

$$\mathscr{F}_t \supseteq \bigcap_{\epsilon>0} \mathscr{F}_{t+\epsilon},$$

as the reverse inclusion is trivially satisfied by definition of a filtration. Once again, this means that if an event depends on what happens just immediately after time t, then this event can also be decided with the information available up to time t. For a very neat and simple example of a filtration that is not right-continuous consult **??**Example 3.17]ap-notes.

We now see the Theorem that formalises the intuition we had, namely that if we actually had allowed our process in example 3.1.2 to look slightly into the future, we could have in fact decided whether T had occurred.

Theorem 3.1.5 (Stopping times in usual conditions) Let $\{\mathscr{F}_t\}$ be a filtration the satisfies the usual conditions. Then a random time T is a stopping time if and only if $\{T < t\} \in \mathscr{F}_t$ for all $t \geqslant 0$.

Proof. The sufficient condition is trivial, just note that:

$$\{T < t\} = \bigcap_{n \geqslant 0} \left\{T \leqslant t - \frac{1}{n}\right\}.$$

In fact the necessary condition is also quite easy, just let $\epsilon > 0$, then it is clear that

$$\{T \leqslant t\} = \bigcap_{n \geqslant 1/\epsilon} \underbrace{\left\{T < t + \frac{1}{n}\right\}}_{\in \mathscr{F}_{t+\frac{1}{n}} \subseteq \mathscr{F}_{t+\epsilon}} \in \mathscr{F}_{t+\epsilon}$$

Therefore it readily follows that

$$\{T \leqslant t\} \in \bigcap_{\epsilon > 0} \mathscr{F}_{t+\epsilon} = \mathscr{F}_t$$

where this last step holds by assumption of right-continuity. ♡

Remark 3.1.6 Note that this proof only really required right-continuity.

Let us now do a thorough job and show that this modification to our sigma-algebra actually does fix the problem we had in Example 3.1.2.

Theorem 3.1.7 (Hitting times of open sets in usual conditions are stopping times) Let X be a right-continuous process taking values in $\mathbf{R}^d$ adapted to a filtration that satisfies the usual conditions. It follows that for an open set $A \subseteq \mathbf{R}^d$, we have that

$$T = \inf\{t > 0 : X_t \in A\}$$

is in fact a stopping time.

Proof. In light of Theorem 3.1.5, it suffices to show that $\{T < t\} \in \mathscr{F}_t$ for all $t \geqslant 0$. This follows after the following two steps, the last of which we will justify in a moment:

$$\{T < t\} = \{X_s \in A \text{ for some } s \in [0, t)\} \tag{3.1}$$
$$= \bigcup_{q \in [0,t) \cap \boldsymbol{Q}} \{X_q \in A\} \tag{3.2}$$

To justify (3.3) we note the following: If $X_s \in A$ for some $0 \leqslant s < t$, since A is open, there is some $\delta > 0$ small enough so that $\boldsymbol{B}(X_s, \delta) \subseteq A$. Moreover, by left continuity of our process, we can find some $0 < \epsilon_0 < t - s$ small enough such that for any $\epsilon < \epsilon_0$, one has that $\|X_{s+\epsilon} - X_s\| < \delta$. In particular, for all rationals $q \in [s, s+\epsilon_0) \cap \boldsymbol{Q}$, we have that $X_q \in A$. In conclusion, if $X_s \in A$ for some s, then we can find some rational q for which $X_q \in A$. Now the claim follows because looking at (3.3) we see that since $q < t$, then $\{X_q \in A\} \in \mathscr{F}_q \subseteq \mathscr{F}_t$. ♡

Right, recall that we are trying to move towards an integration theory in which our integrators can be Martingales. Our motivation for this detour on usual conditions stems from the fact that we want to have a good amount of stopping times, but notice how in the proof of Theorem 3.1.7 we also required X to be right-continuous. As the reader may already know from a course in probability, it turns out this is not a stringent prerequisite at all, due to the fact that if we have usual conditions, we can "regularise" our Martingale and make it right-continuous. Actually something better holds:

Theorem 3.1.8 (Martingale Regularisation Theorem) Let X be a Martingale with respect to a filtration $\{\mathscr{F}_t\}$ which satisfies the usual conditions. Then X has a cadlag modification. That is to say: there exists a cadlag process X^* with the property that for any $t \geqslant 0$,

$$\boldsymbol{P}[X_t = X_t^*] = 1.$$

Now we are in much better shape. But there may be one more thing that is bugging the reader: we still haven't explained why it is in fact justified to perform the "cheap trick" of letting the process look slightly into the future. Does this alter the properties of our process, or how it talks to our new filtration?

Theorem 3.1.9 (Martingales and usual conditions) Let X be a càdlàg Martingale with respect to a filtration $\{\mathscr{F}_t\}$, then X is also a Martingale with respect to the filtration

$$\mathscr{F}_t^* = \sigma\left(\bigcap_{\epsilon>0} \mathscr{F}_{t+\epsilon} \cup \{\boldsymbol{P} - \text{null sets}\}\right).$$

i.e: its still a Martingale if we enhance our filtration just enough to satisfy usual conditions.

Proof. We need to show that for any set $A \in \mathscr{F}_s^*$ where $0 \leqslant s \leqslant t$, we have that

$$\boldsymbol{E}\left[(X_t - X_s)\mathbf{1}_A\right] = 0.$$

It turns out that if $\mathscr{G}$ is a σ-algebra and $\mathscr{N}$ denotes the collection of $\boldsymbol{P}$-null sets, then the collection of sets of the form $A \cup M$ for $A \in \mathscr{G}$ and $M \in \mathscr{N}$ is also a σ-algebra. However, $\mathscr{F}_t^*$ is the smallest sigma-algebra that contains from $\mathscr{F}_t^+$ and $\mathscr{N}$, so in particular it is a subset of the sigma algebra of unions described above, so that $A = B \cup C$ for some set $B \in \mathscr{F}_t^+$ and $C \in \mathscr{N}$. Since $\mathbf{1}_A = \mathbf{1}_{B\cup C}$ but C is of measure zero, then $\mathbf{1}_A = \mathbf{1}_B$ almost surely. Thus we have reduced our task to the case in which $A \in \mathscr{F}_s^+$. Naturally, since $A \in \mathscr{F}_s^+$, we also have that for any $\epsilon > 0$, $A \in \mathscr{F}_{s+\epsilon}$. So by the Martingale property of X, we have that

$$\boldsymbol{E}[(X_t - X_{s+\epsilon})\mathbf{1}_A] = 0.$$

However, note that $(X_{s+\epsilon})_{\epsilon\in[0,t-s]}$ is UI, because for any such ϵ, we have that $X_{s+\epsilon} = \boldsymbol{E}[X_t \mid \mathscr{F}_{s+\epsilon}]$, and so it follows by the UI property of conditional expectation. With this out of the way, we apply Vitali's Theorem: since X is càdlàg then $X_{s+\epsilon} \to X_S$ as $\epsilon \downarrow 0$, and since almost sure convergence implies convergence in probability and the family $(X_{s+\epsilon})_{\epsilon\in[0,t-s]}$ is UI, we have that $X_{s+\epsilon} \to X_s$ in

$\mathscr{L}^1$, which means that we can pass in the limit $\epsilon \downarrow 0$ into the integral and we conclude that

$$\boldsymbol{E}[(X_t - X_s)\mathbf{1}A] = 0.$$

♡

3.2 Local Martingales

From now on we assume that all Martingales are cadlag and all filtrations satisfy the usual conditions.

Definition 3.2.1 (Local Martingale) A **local Martingale** is a cadlag adapted process X such that there exists an increasing sequence of stopping times (T_n) with $T_n \uparrow \infty$ with the property that the stopped process $X^{T_n} - X_0$ is a Martingale, which as a reminder of notation, means:

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

The sequence of stopping times (T_n) is referred to as a localising sequence, and we also say that (T_n) reduces X to a Martingale. We will always work with continuous local Martingales unless otherwise specified and we denote by $\mathscr{M}_{\text{loc}}$ the set of continuous local Martingales.

Remark 3.2.2 Note that in general the requirement of being local is not the same as that of there being a sequence of stopping times (T_n) for which X^{T_n} is itself a Martingale. For the two requirements to coincide, we would additionally need to impose that $\boldsymbol{E}[X_0 \mid \mathscr{F}_s] = X_0$, which can be achieved for example, if X_0 is a constant, or in more generality, if $\mathscr{F}_0$ is a trivial σ-algebra, which doesn't seem that far of a reach since at time 0 we don't have any information.

In light of this remark, we make the assumption from now on that all filtrations we consider have $\mathscr{F}_0$ to be trivial, thus simplifying our requirement for a local Martingale.

Remark 3.2.3 If X_0 is also measurable with respect to the "parent" sigma algebra $\mathscr{F}$, then we will also get the same result because $\boldsymbol{E}[X_0 \mid \mathscr{F}_s] = X_0$.

Proposition 3.2.4 Let X be a continuous local Martingale and set (T_n) to be the following sequence of stopping times:

$$T_n = \inf\{t > 0 : |X_t| > n\}.$$

Then (T_n) also localises X.

The moral of this proposition is that when we are dealing with local Martingales we can always take this familiar sequence of stopping times as our localising sequence.

Proof. Start by fixing a localising sequence (U_n). By virtue of X being continuous, we have that $|X_{t\wedge T_n}| \leqslant n$ for all $t \geqslant 0$, $n \in \boldsymbol{N}$, and so we can emply the DCT in the following way:

$$\begin{aligned}
\boldsymbol{E}[X_{t\wedge T_n} \mid \mathscr{F}_s] &= \boldsymbol{E}\left[\lim_{k\to\infty} X_{t\wedge T_n \wedge U_k} \middle| \mathscr{F}_s\right] && (3.3)\\
&= \lim_{k\to\infty} \boldsymbol{E}\left[X_{t\wedge T_n\wedge U_k} \middle| \mathscr{F}_s\right] && (3.4)\\
&= \lim_{k\to\infty} X_{s\wedge T_n\wedge U_k} && (3.5)\\
&= X_{s\wedge T_n} && (3.6)
\end{aligned}$$

In step (3.4) we just used the fact that (U_k) by assumption has that $U_k \uparrow \infty$. In step (3.5) we used the observation we did earlier regarding DCT. In step (3.6) we used the fact that U_k turns $M_t = (X_{t\wedge U_k})$ into a Martingale, and in line (3.5) we effectively have $M_{t\wedge T_n}$, and since a stopped Martingale is still a Martingale, then we get the desired result in (3.6). ♡

Example 3.2.5 (Geometric Brownian Motion with drift) Naturally, all Martingales local Martingales, say by choosing the sequence of stopping times $T_n = n$. However, the converse is not true, and it is not immediately obvious from the definition which counterexamples we can use. Here is a way we can approach to construct one. First consider the so-called Geometric Brownian Motion:

$$M_t = \exp\left(W_t - \frac{t}{2}\right)$$

Where W is Standard Brownian Motion. Then M is a Martingale with respect to the natural filtration generated by the Brownian Motion, indeed:

$$\begin{aligned}
\boldsymbol{E}\left[\exp\left(W_t - \frac{t}{2}\right)\middle|\mathscr{F}_s\right] &= \boldsymbol{E}\left[\exp\left(W_t - W_s + W_s - \frac{t}{2} - \frac{s}{2} + \frac{s}{2}\right)\middle|\mathscr{F}_s\right] && (3.7)\\
&= \exp\left(W_s - \frac{s}{2}\right)\boldsymbol{E}\left[W_t - W_s - \frac{t-s}{2}\right] && (3.8)\\
&= \exp\left(W_s - \frac{s}{2}\right) && (3.9)
\end{aligned}$$

Where in step (3.9) we used the fact that W_s is $\mathscr{F}_s$-measurable, as well as the Markov Property. In step (3.10), we used the form of the Moment Generating Function of the Normal Distribution. As a fun fact, note that this Martingale is not Uniformly Integrable: That is because M_t converges to 0 almost surely:

$$M_t = \exp\left(\frac{W}{t} - \frac{1}{2}\right)^t \to 0 \tag{3.10}$$

Where in the limit we used the Law of Large Numbers for Brownian Motion. However,

$$\boldsymbol{E}\left|M_t - 0\right| = \boldsymbol{E}\exp\left(W_t - \frac{t}{2}\right) = 1$$

which certainly does not go to zero, and so we have almost sure convergence yet no $\mathscr{L}^1$ convergence. Regardless, let us now construct an example of a non-Martingale that is a local Martingale.

3.3 Class D and Class DL

Recall that we said that a process X was UI if the family of random variables $\{X_t : t \in \mathbf{R}_+\}$ was UI. Even if X was UI, it does not follow that the family of random variables consisting of X sampled at random times is UI. We make this into a definition now:

Definition 3.3.1 (Doob Class) A cadlag process X is said to be in the Doob Class, or just Class D if the family of random variables $\{X_T : T \text{ is a finite stopping time}\}$ is UI.

Definition 3.3.2 (Locally in Doob Class) A cadlag process X is said to be locally in Doob CLass, or just Class DL if for all $t \geqslant 0$, the family of random variables

$$\{X_{T\wedge t} : T \text{ is a stopping time}\} = \{X_{T\wedge t} : T \text{ is a bounded stopping time}\}$$

is Uniformly Integrable.

Remark 3.3.3 (Subtlety: finiteness vs boundedness) A random variable T is finite if for all $\omega \in \Omega$, $T(\omega) < \infty$. This is a weaker condition than T being bounded, which means that there exists some $M > 0$ such that for all $\omega \in \Omega$, $T(\omega) < M$.

Now we have a Theorem that characterises when a Local Martingale is a Martingale.

Proposition 3.3.4 (Local Martingales, Martingales and Class DL) A local Martingale is a (true) Martingale if and only if it belongs to class DL.

Proof. Start by assuming that a local Martingale X is a Martingale, let us show that X is in class DL. Let T be a bounded stopping time, then by the Optional Stopping Theorem, we have that

$$\boldsymbol{E}[X_t \mid \mathscr{F}_T] = X_{t\wedge T}.$$

From which we conclude that under our assumption,

$$\{X_{T\wedge t} : T \text{ is a bounded stopping time}\} = \{\boldsymbol{E}[X_t \mid \mathscr{F}_T] : T \text{ is a bounded stopping time}\}.$$

The proof now follows by the UI property of Conditional Expectation. Conversely, suppose that X is a local Martingale that belongs to class DL, let us show that X is a Martingale. Start by choosing a sequence of increasing stopping times (T_n) for which X^{T_n} is a Martingale. We have the following observations: **(OBS 1)** By our choice of (T_n) we trivially have that for any fixed $t \geqslant 0$, $X_t = \lim_{n\to\infty} X_{t\wedge T_n}$ almost surely. In particular, since we are assuming X is in class DL, we also have that the family $\{X_{t\wedge T_n} : n \in \boldsymbol{N}\}$ is UI, and so by Vitali's Theorem our convergence gets upgraded to convergence in $\mathscr{L}^1$ **(OBS 2)**. Putting all this together:

$$\begin{aligned}
\boldsymbol{E}[X_t \mid \mathscr{F}_s] &= \boldsymbol{E}\left[\lim_{n\to\infty} X_{t\wedge T_n} \middle| \mathscr{F}_s\right] && (3.11)\\
&= \lim_{n\to\infty} \boldsymbol{E}\left[X_{t\wedge T_n} \middle| \mathscr{F}_s\right] && (3.12)\\
&= \lim_{n\to\infty} X_{s\wedge T_n} && (3.13)\\
&= X_s && (3.14)
\end{aligned}$$

Where in step (3.12) we have used **(OBS 1)**, in step (3.13) we have used **(OBS 2)** (technically, something deeper is going on here, because the usual trick of pulling limits in and out from expectations only works, well for expectations, but it turns out that this can also be upgraded to work with conditional expectations). In step (3.14) we have used the fact that X is a local Martingale and (T_n) is a localising sequence, and finally in step (3.15) we have used **(OBS 1)**. ♡

We have one more result, which follows immediately from the OST for UI cadlag Martingales.

Proposition 3.3.5 A uniformly integrable (cadlag) Martingale is in class D.

Proof. This is essentially an immediate application of the OST for UI cadlag Martingales, but I will reprove it for clarity. We know that since X is UI, then $X_n = \boldsymbol{E}[X_\infty \mid \mathscr{F}_n]$. Now we will prove the OST for discrete stopping times. If T is a discrete stopping time, then it is clear enough (say by the

DCT), that if $B \in \mathscr{F}_T$,

$$\boldsymbol{E}[X_T \mathbf{1}_B] = \sum_{n \in \boldsymbol{N} \cup \{\infty\}} \boldsymbol{E}[\mathbf{1}_B \mathbf{1}_{T=n} X_n] = \sum_{n \in \boldsymbol{N} \cup \{\infty\}} \boldsymbol{E}[\mathbf{1}_B \mathbf{1}_{T=n} X_\infty] = \boldsymbol{E}[\mathbf{1}_B X_\infty]$$

Where we used in the second equality that $\boldsymbol{E}[X_\infty \mid \mathscr{F}_n] = X_n$ and the fact that by definition of a stopped sigma algebra, $B \cap \{T = n\} \in \mathscr{F}_n$. Now that we have the OST for discrete time, namely that $\boldsymbol{E}[X_\infty \mid \mathscr{F}_T] = X_T$, we can generalise this to any (potentially continuous time) stopping time T, simply take a discretisation $T_k \downarrow T$, and observe that whenever $A \in \mathscr{F}_T$, then

$$A \cap \{T_k \leqslant n\} \subseteq A \cap \{T \leqslant n\} \in \mathscr{F}_n$$

and so A is in $\mathscr{F}_{T_k}$, therefore,

$$\boldsymbol{E}[X_\infty \mathbf{1}_A] = \boldsymbol{E}[X_{T_k} \mathbf{1}_A]$$

from the discrete OST, but also, since $X_{T_k} = \boldsymbol{E}[X_\infty \mid \mathscr{F}_{T_k}]$, by the UI property of conditional expectation, we have that (X_{T_k}) is UI, and so it converges in $\mathscr{L}^1$ and we can pass a limit inside the expecation so that

$$\boldsymbol{E}[\mathbf{1}_A X_\infty] = \boldsymbol{E}[\lim_{k \to \infty} X_{T_k} \mathbf{1}_A]$$

But now we use the fact that $T_k \downarrow T$ and X is cadlag to finish the claim. ♡

Alternative Proof which I don't like. An alternative proof is the following. Since $T \wedge n$ is a bounded stopping time, we have that using X is UI: $\boldsymbol{E}[X_\infty \mid \mathscr{F}_{T \wedge n}] = X_{T \wedge n}$. Therefore we have that

$$X_T = \lim_{n \to \infty} \boldsymbol{E}[X_\infty \mid \mathscr{F}_{T \wedge n}] =: Y_\infty.$$

If we can show that this right hand side equals $\boldsymbol{E}[X_\infty \mid \mathscr{F}_T]$ we will be done. Let $A \in \mathscr{F}_T$, then we can express $A = \lim_{n \to \infty} A \cap \{T \leqslant n\}$ pointwise, and one can see that $\mathbf{1}_{A \cap \{T \leqslant n\}} \to \mathbf{1}_A$ in $\mathscr{L}^1$, so that $Y_\infty \mathbf{1}_A = \lim_{n \to \infty} X_{T \wedge n} \mathbf{1}_{A \cap \{T \leqslant n\}}$ in $\mathscr{L}^1$. Moreover, $A \cap \{T \wedge n\}$ can be seen to be in $\mathscr{F}_{T \wedge n}$, and so we have that

$$\boldsymbol{E}[Y_\infty \mathbf{1}_A] = \lim_{n \to \infty} \boldsymbol{E}[X_{T \wedge n} \mathbf{1}_{A \cap \{T \leqslant n\}}] = \lim_{n \to \infty} \boldsymbol{E}[X_\infty \mathbf{1}_{A \cap \{T \leqslant n\}}] = \boldsymbol{E}[X_\infty \mathbf{1}_A]$$

Where the second inequality comes from the fact that $A \cap \{T \leqslant n\} \in \mathscr{F}_{T \wedge n}$ and $X_{T \wedge n} = \boldsymbol{E}[X_\infty \mid \mathscr{F}_{T \wedge n}]$.
♡

3.4 Square Integrable Martingales

Definition 3.4.1 (Square Integrable Martingales) We define $\mathscr{M}_2$ to be the space of square integrable Martingales:

$$\mathscr{M}_2 = \left\{ M \text{ a continuous Martingale with } \sup_t \boldsymbol{E} X_t^2 < \infty \right\}$$

Remark 3.4.2 (On the value of this supremum) Naturally, by the Martingale Convergence Theorem, whenever $M \in \mathscr{M}_2$, we have that $M_t \to M_\infty$ almost surely and in $\mathscr{L}^2$. Moreover, note that if M is a Martingale, then M^2 is a sub-Martingale, indeed: if $t \geqslant s$, then

$$\begin{aligned}
\boldsymbol{E}[M_t^2 \mid \mathscr{F}_s] &= \boldsymbol{E}\left[(M_t - M_s + M_s)^2 \mid \mathscr{F}_s\right] \\
&\geqslant 2\boldsymbol{E}\left[M_s(M_t - M_s) \mid \mathscr{F}_S\right] + \boldsymbol{E}[M_s^2 \mid \mathscr{F}_s] \\
&= M_s^2.
\end{aligned}$$

Therefore taking expectations of both sides, we get that $\boldsymbol{E}[M_t^2] \geqslant \boldsymbol{E}[M_s^2]$, meaning that the map $t \mapsto \boldsymbol{E} M_t^2$ is non-decreasing, and as such $\sup_t \boldsymbol{E} M_t^2 = \boldsymbol{E} M_\infty^2$.

Remark 3.4.3 (On a maximal inequality) Recall Doob's Maximal Inequality (for $p = 2$): if X is a cadlag process, then $\|\sup_t X_t\|_2 \leqslant 2\|X_t\|$. In terms of our situation, since M is assumed to be continuous, then of course its cadlag and so

$$\boldsymbol{E}\left[\sup_{t \geqslant 0} X_t^2\right] \leqslant 4\boldsymbol{E}[X_\infty^2].$$

We now have the following result:

Theorem 3.4.4 ($\mathscr{M}_2$ is a complete vector space) The vector space $\mathscr{M}_2$ is complete with respect to the norm

$$\|X\|_{\mathscr{M}_2}^2 = \boldsymbol{E} X_\infty^2$$

which we know is finite by assumption of $X \in \mathscr{M}_2$ and the remark of the value of the supremum.

Proof. We take a Cauchy sequence (X_n) in $\mathscr{M}_2$, i.e:

$$\boldsymbol{E}[(X_\infty^n - X_\infty^m)^2] \to 0$$

as $m, n \to \infty$, and show that it has a convergent subsequence. Meaning that there is some subsequence

(n_k) along which we have convergence to a continuous $\mathscr{M}_2$ Martingale. We begin by constructing the limit. By a standard trick in analysis, we can take a subsequence (n_k) for which

$$\boldsymbol{E}[(X_\infty^{n_k} - X_\infty^{n_{k+1}})^2] \leqslant 2^{-k}$$

Now notice that obviously we can write

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

So we need to show that this sum almost surely converges for all t as $k \to \infty$. For this, note that

$$\boldsymbol{E}\left[\sum_{i=1}^{\infty} \sup_{t\geqslant 0} |X_t^{n_i} - X_t^{n_{i-1}}|\right] = \sum_{i=1}^{\infty} \boldsymbol{E} \sup_{t\geqslant 0} |X_t^{n_i} - X_t^{n_{i-1}}| \tag{3.15}$$

$$\leqslant \sum_{i=1}^{\infty} \sqrt{\boldsymbol{E} \sup_{t\geqslant 0} |X_t^{n_i} - X_t^{n_{i-1}}|^2} \tag{3.16}$$

$$\leqslant 2\sum_{i=1}^{\infty} \sqrt{\boldsymbol{E}|X_\infty^{n_i} - X_\infty^{n_{i-1}}|^2} \tag{3.17}$$

Where (3.16) comes from the MCT, (3.17) comes from Jensen's Inequality, and (3.18) comes from Doob's Maximal Inequality. From this and our construction of the subsequence (n_k) it follows that the sum in (3.18) is summable, and so the original expression in (3.16) is finite. If the expectation of a random variable is finite, it must mean that it is almost surely finite, i.e:

$$\sum_{i=1}^{\infty} \sup_{t\geqslant 0} |X_t^{n_i} - X_t^{n_{i-1}}| < \infty \quad \text{a.s}$$

This implies that $X_t^\star = \lim_{k\to\infty} X_t^{n_k}$ exists. Now we prove that $X^\star$ is a continuous process. Note that

$$\sup_{t\geqslant 0} |X_t^{n_k} - X_t^\star| \leqslant \sum_{i=k+1}^{\infty} \sup |X_t^{n_i} - X_t^{n_{i-1}}| \to 0$$

as $k \to \infty$ almost surely. Therefore we have that X^{n_k} converges uniformly almost surely to $X^\star$, and since each X^{n_k} is continuous, (and the uniform limit of continuous functions is continuous) it follows that $X^\star$ is also continuous almost surely. Now it is left to show that $X^\star$ is $\mathscr{L}^2$ bounded and that it is a Martingale, both of this results are proven in a similar way. The important thing to note is that since $(X_\infty^{n_k})_k$ is an $\mathscr{L}^2$ bounded family, and so it converges in $\mathscr{L}^2$ and so (say by Jensen's Inequality) it also converges in $\mathscr{L}^1$, from this we have that $\boldsymbol{E}(X_\infty^\star)^2 = \boldsymbol{E}[\lim_{k\to\infty}(X_\infty^{n_k})^2] = \lim_{k\to\infty} \boldsymbol{E}[(X_\infty^{n_k})^2] < \infty$ where the last inequality comes from $\mathscr{L}^2$ boundedness. Therefore $X^\star$ has indeed finite norm, and finally we just need to show that it is a Martingale. Once again, this is a similar argument:

$$\begin{aligned}\boldsymbol{E}[X_\infty^\star|\mathscr{F}_t] &= \lim_{k\to\infty}\boldsymbol{E}[X_\infty^{n_k}\,|\,\mathscr{F}_t]\\&= \lim_{k\to\infty}X_t^{n_k}\\&= X_t^\star.\end{aligned}$$

Where the first inequality we used the $\mathscr{L}^1$ convergence discussed before, and in the second inequality we used the fact that X^{n_k} is a Martingale. Therefore by the Tower Law it now follows that $X^\star$ is indeed a Martingale. ♡

3.5 Quadratic Variation

In this section we are going to show that every local Martingale has associated to it a continuous adapted and increasing process, which captures the "accumulated roughness" of the process.

Proposition 3.5.1 Let M be a Martingale, and let K be a bounded $\mathscr{F}_{t_0}$ 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. We will use the converse of the Optional Stopping Theorem:

$$\begin{aligned}\boldsymbol{E}[X_T] &= \boldsymbol{E}[K(M_T - M_{T\wedge t_0})] &(3.18)\\&= \boldsymbol{E}[K\,\boldsymbol{E}[M_T - M_{T\wedge t_0}\,|\,\mathscr{F}_{t_0}]] &(3.19)\\&= 0. &(3.20)\end{aligned}$$

Thus satisfying the hypothesis of the Converse of the OST, and so X is a Martingale. Here (3.19) followed by the Tower Law, and (3.20) followed by the OST and the fact that M is a Martingale. ♡

We will want to talk about convergence of stochastic processes as a whole, and for this we need a new notion of convergence.

Definition 3.5.2 (UCP convergence) A sequence of processes Z^n converges Uniformly on Compacts in Probability (UCP) to a process Z if for all $t \geqslant 0$ and all $\epsilon > 0$:

$$\boldsymbol{P}\left[\sup_{s\in[0,t]} |Z_s^n - Z_s| > \epsilon\right] \to 0.$$

Theorem 3.5.3 (Quadratic variation exists) Let X be a continuous local Martingale. Let

$$[X]_t^{(n)} = \sum_{k\geqslant 0}\left(X_{t\wedge t_k^n} - X_{t\wedge t_{k-1}^n}\right)^2.$$

Then the process $[X]_t^{(n)}$ converges UCP to a continuous, adapted, and non-decreasing process $[X]_t$, which we call the Quadratic Variation of X.

Proof. Without loss of generality we assume that $X_0 = 0$. We distinguish two cases: when the process is uniformly bounded, meaning that there is some $C > 0$ for which $|X_t(\omega)| \leqslant C$ for all $(t,\omega) \in \mathbf{R}_+ \times \Omega$. Since a uniformly bounded local Martingale is certainly in class DL, then it is a true Martingale, and moreover it is in $\mathscr{M}_2$. (A simple DCT argument also shows that uniformly bounded local Martingales are true Martingales). Therefore the limit X_∞ exists almost surely and in $\mathscr{L}^2$ (and in $\mathscr{L}^1$). First of all notice the following observation that will make our lives easier: since $X_t \to X_\infty$ almost surely, we have that

$$[X]_t^{(n)} - [X]_{2^{-n}\lfloor 2^n t\rfloor}^{(n)} = \left(X_t - X_{2^{-n}\lfloor 2^n t\rfloor}\right)^2 \to 0$$

as $t \to \infty$. Therefore we can unambiguously write

$$\begin{aligned}
[X]_\infty^{(n)} &= \sup_k [X]_{t_k^n}^{(n)} \\
&= \sup_k \sum_{i=1}^k \left(X_{t_i^n} - X_{t_{i-1}^n}\right)^2 \\
&= \lim_{k\to\infty} \sum_{i=1}^k \left(X_{t_i^n} - X_{t_{i-1}^n}\right)^2
\end{aligned}$$

and so

$$\begin{aligned}
\boldsymbol{E}[X]_\infty^{(n)} &= \lim_{k\to\infty} \boldsymbol{E}\left[\sum_{i=1}^k \left(X_{t_i^n} - X_{t_{i-1}^n}\right)^2\right] && (3.21)\\
&= \lim_{k\to\infty} \boldsymbol{E}[X_{t_k^n}^2] && (3.22)\\
&= \boldsymbol{E}[X_\infty^2] \leqslant C^2 && (3.23)
\end{aligned}$$

Where (3.22) is due to the MCT, (3.23) is due to the Pythagorean Theorem, and (3.24) is due to X being in $\mathscr{M}_2$ in order to put the limit in. Finally the inequality comes from the assumption of uniform boundedness. Therefore $[X]_\infty^{(n)}$ is finite almost surely. Then we define

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

We can expand M_t^n into a series. Indeed, note that by telescoping we can write

$$X_t^2 = \sum_{k=1}^{\infty} X_{t\wedge t_k^n}^2 - X_{t\wedge t_{k-1}^n}^2$$

(indeed, there will be some k' for which $t_{k'}^n \leqslant t < t_{k'+1}^n$, and so the sum will only run until $k = k'+1$. The last term will be $X_t^2 - X_{t_k'^n}$ and all the terms before that will die). Therefore combining with the definition of $[X]_t^n$ gives:

$$\begin{aligned} M_t^n &= \frac{1}{2}\sum_{k=1}^{\infty} X_{t\wedge t_k^n}^2 - X_{t\wedge t_{k-1}^n}^2 - X_{t\wedge t_k^n}^2 + 2X_{t\wedge t_k^n}X_{t\wedge t_{k-1}^n} - X_{t\wedge t_{k-1}^n}^2 \\ &= \sum_{k=1}^{\infty} X_{t\wedge t_{k-1}^n}\left(X_{t\wedge t_k^n} - X_{t\wedge t_{k-1}^n}\right) \end{aligned}$$

We have established that X is a Martingale, so each of this summands is the product of a $\mathscr{F}_{t_{k-1}^n}$-measurable bounded random variable times $\left(X_{t\wedge t_k^n} - X_{t\wedge t_k^n \wedge t_{k-1}^n}\right)$. From Proposition 3.5.1 it follows that each of the summands is a Martingale. This sum contains only finitely many terms so it is clear that $(M_t^n)_t$ is a Martingale. The idea is now to show that M^n is in $\mathscr{M}_2$, and then moreover show that the sequence is Cauchy. From the limit we will be able to extract the quadratic variation $[X]$. To show that $M^n \in \mathscr{M}_2$ we compute its $\mathscr{M}_2$ norm and show it is finite.

$$\boldsymbol{E}(M_\infty^n)^2 \leqslant \boldsymbol{E}C^2\left(\sum_{k=1}^{\infty} X_{t_k^n} - X_{t_{k-1}^n}\right)^2 \tag{3.24}$$

$$= C^2\boldsymbol{E}\sum_{k=1}^{\infty}(X_{t_k^n} - X_{t_{k-1}^n})^2 \tag{3.25}$$

$$= C^2\boldsymbol{E}[X]_\infty^n \tag{3.26}$$

$$\leqslant C^4 \tag{3.27}$$

Where (3.25) follows from $|X_{t\wedge t_{k-1}^n}| \leqslant C$ and so we pull it out of the sum, step (3.26) follows from the Pythagorean Theorem, step (3.27) follows from the definition of $[X]_\infty^n$ and step (3.28) follows from the calculation we did in steps (3.22)-(3.24). We have now established that for each n, $M^n \in \mathscr{M}_2$ and now we proceed to show it is a Cauchy sequence. Since the t_k^n are dyadic numbers, one can get the following formula (don't worry too much about where it comes from), if $n \geqslant m$, then

$$M_\infty^n - M_\infty^m = \sum_{j=1}^{\infty} \left(X_{j2^{-n}} - X_{\lfloor j2^{-(n-m)} \rfloor 2^{-m}}\right)\left(X_{(j+1)2^{-n}} - X_{j2^{-n}}\right),$$

by the Pythagorean Theorem once again, we have that

$$\boldsymbol{E}|M_\infty^n - M_\infty^m|^2 = \sum_{j=1}^{\infty} \boldsymbol{E}\left(X_{j2^{-n}} - X_{\lfloor j2^{-(n-m)} \rfloor 2^{-m}}\right)^2 \left(X_{(j+1)2^{-n}} - X_{j2^{-n}}\right)^2 \tag{3.28}$$

$$\leqslant \boldsymbol{E} \sup_{|t-s| \leqslant 2^{-m}} (X_s - X_t)^2 [X]_\infty^n \tag{3.29}$$

$$\leqslant \left(\boldsymbol{E} \sup_{|t-s| \leqslant 2^{-m}} (X_s - X_t)^4\right)^{1/2} \left(\boldsymbol{E}([X]_\infty^n)^2\right)^{1/2} \tag{3.30}$$

Where step (3.29) comes from the Pythagorean Theorem, step (3.30) comes from a simple bound and the definition of $[X]_\infty^n$, and step (3.31) is just the Cauchy-Schwarz inequality. We are going to show that the first term in the final product goes to zero and that the second term is bounded. Since for all ω, $X_t(\omega) \to X_\infty(\omega)$ and $X(\omega)$ is continuous, we have that $X(\omega)$ is uniformly continuous for all ω. Therefore, as $m \to \infty$, the first term vanishes because the supremum is bounded by $16C^4$ and thus we can use the DCT. Now we show that the second term in the product is bounded. By definition of M_∞^n we can write

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

Where in the middle we used that $(a-b)^2 \leqslant 2(a^2+b^2)$ (this just follows from $(a+b)^2 \geqslant 0$). Thus we have shown that $(M^n)_n$ is Cauchy, so it has a limit $M^* \in \mathscr{M}_2$. Now we define

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

It is clear that $[X]$ is continuous and adapted, since the right hand side is. Now we need to show the two other requirements, namely that it is non-decreasing and the convergence. We will start by showing a convergence statement. By using the definition and Doob's maximal inequality:

$$\begin{aligned} \boldsymbol{E} \sup_{t \geqslant 0} \left([X]_t^n - [X]_t\right)^2 &= 4\boldsymbol{E} \sup_{t \geqslant 0} \left(M_t^n - M_t^*\right)^2 \\ &\leqslant 16\boldsymbol{E}(M_\infty^n - M_\infty^*)^2. \end{aligned}$$

This last quantity goes to zero because that's what convergence in $\mathscr{M}_2$ means. It follows that

$\sup_{t\geqslant 0}\left([X]^n_t - [X]_t\right) \to 0$ in $\mathscr{L}^2$, we can refer to this as $[X]^n \to [X]$ uniformly in $\mathscr{L}^2$. It is clear that this implies UCP because $\mathscr{L}^2$ implies convergence in probability. Now we show that $[X]$ is almost surely non-decreasing. For this, we note that convergence in $\mathscr{L}^2$ allows us to extract a subsequence (n_k) along which $[X]^n$ converges pointwise almost surely. It is also clear that the proto-variations $[X]^n$ are non-decreasing, because if $s < t$, then $[X]^n_t$ includes more non-negative terms in the sum than $[X]^n_s$ does. Therefore, we can say that almost surely,

$$[X]_s = \lim_{k\to\infty}[X]^{n_k}_s \leqslant \lim_{k\to\infty}[X]^{n_k}_t = [X]_t.$$

So we are done for the case where X is uniformly bounded. Now we need to relax our assumption to unbounded X. The price to pay will simply be that convergence will have to be relaxed to UCP. This is where we will use that X is a local Martingale, we will localise it with a stopping time that will make it bounded. For each $N \geqslant 1$, let

$$T_N = \inf\{t \geqslant 0 : |X_t| > N\}.$$

Then it is clear that by continuity of X and by the local Martingale property and the fact that this stopping times are always a localising sequence, X^{T_N} is a bounded Martingale. Hence the process $[X^{T_N}]$ is well-defined by the above construction and satisfies the desired properties. It is clear that

$$[X^{T_{N+1}}]^n_t - [X^{T_N}]^n_t \begin{cases} = 0 & t \leqslant T_N \\ \geqslant 0 & t > T_N \end{cases}$$

Therefore by taking $n \to \infty$ we have that

$$[X^{T_{N+1}}]_t - [X^{T_N}]_t \begin{cases} = 0 & t \leqslant T_N \\ \geqslant 0 & t > T_N \end{cases}$$

which means that we can define $[X]_t = \lim_N [X^{T_N}]_t$. (This limit could be ∞, but it exists due to monotonicity). This function is increasing and adapted as it is the supremum of increasing and adapted functions. Moreover, to show continuity, we note that for each fixed N, on the event that $\{t \leqslant T_N\}$, we have that

$$[X]_t = [X^{T_N}]_t$$

which means that $[X]$ is continuous on the event $\{t \leqslant T_N\}$, however, since $T_N \uparrow \infty$ as $N \to \infty$, we conclude that $[X]$ is continuous almost surely. Finally we need to show convergence in UCP. This follows from the following, let $\epsilon > 0$ and $t > 0$ and fix an $N > 0$, then

$$\begin{aligned}\boldsymbol{P}\left[\sup_{0\leqslant s\leqslant t}|[X]_s - [X]_s^n > \epsilon\right] &\leqslant \boldsymbol{P}\left[\left\{\sup_{0\leqslant s\leqslant t}|[X]_s - [X]_s^n > \epsilon\right\}\cap\{t\leqslant T_N\}\right] + \boldsymbol{P}[T_N < t]\\ &= \boldsymbol{P}\left[\left\{\sup_{0\leqslant s\leqslant t}|[X^{T_N}]_s - [X^{T_N}]_s^n > \epsilon\right\}\cap\{t\leqslant T_N\}\right] + \boldsymbol{P}[T_N < t]\end{aligned}$$

The first term goes to zero as $n\to\infty$ because X^{T_N} is a bounded Martingale so we use the previous result, and the second term can be made arbitrarily small by picking a large enough N. ♡

We have now shown the existence of this process that captures the accumulated roughness or volatility of the process. We can now attempt to give a more abstract characterisation, which perhaps is more useful in computation, and give an example of the quadratic variation of Brownian Motion

Proposition 3.5.4 (Finite Variation implies zero quadratic variation) Let $X\in\mathscr{M}_{\text{loc}}$ be a local Martingale of finite variation. Then $[X] = 0$.

Proof. For a fixed t, we naturally we have that $[X]_t^n \to [X]_t$ in probability (this is because of UCP), therefore we can find a subsequence (n_k) along which the convergence is almost sure. I.e: $[X]_t^{n_k}\to[X]_t$ a.s. However, we can also bound $[X]_t^{n_k}$ as follows:

$$[X]_t^{n_k} = \sum_{j=1}^{\infty}\left(X_{t\wedge t_j^{n_k}} - X_{t\wedge t_{j-1}^{n_k}}\right)^2 \leqslant \sup_{|r-s|\leqslant 2^{-n_k}}|X_r - X_s|\sum_{j=1}^{\infty}\left|X_{t\wedge t_j^{n_k}} - X_{t\wedge t_{j-1}^{n_k}}\right|$$

The first term of this product goes to zero due to uniform continuity of X (once again we are using the observation that if a function f is continuous on $[0,\infty)$ and it converges to a finite limit L, then it is actually uniformly continuous on $[0,\infty)$, this is because after some large enough M, the function has that $|f(x) - f(y)| < \epsilon$ for all $x, y > M$, and on the interval $[0, M]$ the function is uniformly continuous by Cantor's Theorem). The second term of this product is just the first variation which is assumed to be bounded. Therefore it follows that $[X]_t = 0$ for all t. ♡

Now we show that local continuous Martingales with zero quadratic variation are constant.

Proposition 3.5.5 (Local Martingales with zero quadratic variation are constant) Let $X\in\mathscr{M}_{\text{loc}}$ with $[X] = 0$. Then $X_t = X_0$ for all $t\geqslant 0$.

Proof. We recall from the Proof of Theorem 3.5.3 (Existence of Quadratic Variation) that if $X\in\mathscr{M}_{\text{loc}}$ is uniformly bounded, then the sequence of processes

$$M_n = \frac{1}{2}\left(X^n - [X]^n\right)$$

was shown to be in $\mathcal{M}_2$, and moreover the sequence was shown to be Cauchy, which meant by completeness of $\mathcal{M}_2$, that there is some limit $X^2 - [X] \in \mathcal{M}_2$. On the general case, we saw that we can just reduce X to a bounded Martingale with the stopping times $T_N = \inf\{t > 0 : |X_t| > n\}$, and so we have that by the previous explanation, that $(X^2 - [X])^{T_N} = (X^{T_N})^2 - [X^{T_N}] \in \mathcal{M}_2$. This means that in the general case, $X^2 - [X] \in \mathcal{M}_{\text{loc}}$. Now back to our proposition, since $X \in \mathcal{M}_{\text{loc}}$, then as we have just said, $X^2 - [X] \in \mathcal{M}_{\text{loc}}$, but since $[X] = 0$ by assumption, it follows that $X^2 \in \mathcal{M}_{\text{loc}}$. Let (T_N) (defined as usual) reduce both X and X^2 to bounded Martingales, then

$$\begin{aligned} \boldsymbol{E}\left(X_{t\wedge T_N} - X_0\right)^2 &= X_0^2 - 2X_0 \boldsymbol{E} X_{t\wedge T_N} + \boldsymbol{E} X_{t\wedge T_N}^2 \\ &= X_0^2 - 2X_0^2 + X_0^2 = 0 \end{aligned}$$

Therefore $X_{t\wedge T_N} = X_0$ almost surely, and by taking $N \to \infty$ we get the claim. ♡

With this we

Theorem 3.5.6 (Characterisation of Quadratic Variation) Let $X \in \mathcal{M}_{\text{loc}}$ and A an adapted continuous process of finite variation with $A_0 = 0$. If $X^2 - A \in \mathcal{M}_{\text{loc}}$, then $A = [X]$.

Proof. First of all note that the process $A - [X]$ is of finite variation. This is because A has finite variation by assumption, and $[X]$ is an increasing function, and increasing functions have finite variation. Then We have that one can write this process as

$$(X^2 - [X]) - (X^2 - A)$$

The first term is a local Martingale by the arguments made at the start of the Proof of Proposition 3.5.5, and $X^2 - A$ is a local Martingale by hypothesis. Therefore their difference is also a local Martingale. We have thus shown that $A - [X]$ is a local Martingale with finite variation, hence zero quadratic variation by Proposition 3.5.4, and therefore $A - [X]$ is constant, but by assumption and definition of quadratic variation, $A_0 - [X]_0 = 0$, showing that $A = [X]$. ♡

Armed with this we have the following:

Example 3.5.7 (Quadratic Variation of Brownian Motion) Let B be Standard Brownian Motion, then by standard arguments, we know that the process $(B_t^2 - t)_{t\geqslant 0}$ is a Martingale, in particular a local Martingale, and since $(t)_{t\geqslant 0}$ is obviously a continuous adapted process of finite variation that starts at zero, we have that $[B]_t = t$.

This is quite nice. This example quantifies that Brownian Motion accumulates randomness at a unit rate.

3.6 The Stochastic Integral

We now build the real stochastic integral. Recall we work on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_t, \boldsymbol{P})$.

Definition 3.6.1 (Simple Previsible Process) A process H of the form

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

for a sequence of times $0 \leqslant t_0 < \cdots < t_n$ where $H_{t_k} \in \mathrm{b}\mathscr{F}_{t_{k-1}}$ is said to be a simple previsible process. We write $\mathfrak{S}$ for the set of all simple previsible process.

One can hopefully see the analogy between what's happening here and the construction of the Lebesgue integral, we first define the integral for an "elementary" class of processes that have a "natural" definition of the integral, and then extend by using some non-trivial techniques to some wider class of processes.

Definition 3.6.2 (Stochastic integral of simple previsible processes) Let $X \in \mathscr{M}_{\mathrm{loc}}$ and $H \in \mathfrak{S}$, then we define

$$(H \cdot X)_t :=: \int_0^t H \, \mathrm{d}X := \sum_{k=1}^{\infty} H_{t_k} \left(X_{t \wedge t_k} - X_{t \wedge t_{k-1}} \right)$$

Remark 3.6.3 (Waffle and intuition) Here's a very nice way to make sense of this. Think of a stock trader. The stock trader holds different amounts of stock over time. Here, the stock held by the trader at time t is given by the process $(H_t)_t$. The fact that H is previsible ensures that one can determine how much stock to hold by knowing only what's happened slightly before time t. The price of the asset we are trading evolves according to the local Martingale $(X_t)_t$. Then our integral $(H \cdot X)_t$ measures our P&L at time t. In the simple previsible process case it is quite easy to see where this comes from. Between times t_{k-1} and t_k we hold H_{t_k} units of stock. Then on that time interval, our net gain will be simply $H_{t_k}(X_{t_k} - X - t_{k-1})$.

Remark 3.6.4 (Stochastic integral of SPP is an $\mathscr{M}_{\mathrm{loc}}$.) Observe in the definition of the stochastic integral of a simple previsible process that we have a sum that contains only finitely-many terms, and each of this terms, of the shape

$$H_{t_k}(X_{t \wedge t_k} - X_{t \wedge t_{k-1}}) = H_{t_k}(X_{t \wedge t_k} - X_{(t \wedge t_k) \wedge t_{k-1}})$$

H is bounded by assumption, so by Proposition 3.5.1 this thing is a local Martingale.

The stochastic integral of a simple process is a local Martingale, so it has a quadratic variation, let us give an expression for it:

Proposition 3.6.5 (Quadratic Variation of integral and integral of quadratic variation) Let H and X be as above, then

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

Where the right hand integral is understood as a Lebesgue-Stieltjes integral (indeed, $[X]$ is a continuous increasing process, so it has a unique measure associated to it)

Proof. First of all we note that if

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

then since the intervals are disjoint,

$$H^2 = \sum_{k=1}^{n} H_{t_k}^2 \mathbf{1}(t_{k-1}, t_k]$$

which means that

$$\int H^2 \mathrm{d}[X] = \sum_{k=1}^{n} H_{t_k}^2 \int \mathbf{1}(t_{k-1}, t_k] \mathrm{d}[X] = \sum_{k=1}^{\infty} H_{t_k}^2 \left([X]_{t\wedge t_k} - [X]_{t\wedge t_{k-1}}\right)$$

The idea is going to be to use Theorem 3.5.6 and so we need to show that

$$\left(\sum_{k=1}^{\infty} H_{t_k}\left(X_{t\wedge t_k} - X_{t\wedge t_{k-1}}\right)\right)^2 - \sum_{k=1}^{\infty} H_{t_k}^2 \left([X]_{t\wedge t_k} - [X]_{t\wedge t_{k-1}}\right) \in \mathscr{M}_{\mathrm{loc}}$$

Since X is a local Martingale, we can reduce X to a bounded Martingale by stopping it with the "usual" stopping times T_N. For convenience, we will just write X instead of X^{T_N}. Then we have the following claim:

$$(X^{t_k} - X^{t_{k-1}})^2 - ([X]^{t_k} - [X]^{t_{k-1}})$$

is a Martingale. Let T be a bounded stopping time, we will verify that

$$\boldsymbol{E}\left[\left(X_T^{t_k} - X_T^{t_{k-1}}\right)^2 - \left([X]_T^{t_k} - [X]_T^{t_{k-1}}\right)\right] = 0 \tag{3.31}$$

But the key is that if we expand the square on the left,

$$2\boldsymbol{E}[X_{T\wedge t_k} X_{T\wedge t_{k-1}}] = 2\boldsymbol{E}[X_{T\wedge t_{k-1}} \boldsymbol{E}[X_{T\wedge t_k} \mid \mathscr{F}_{t_{k-1}}]]$$

But X is a Martingale, so this equals $2\boldsymbol{E}[X^2_{T\wedge t_k}]$. Therefore putting it all together, equation (3.31) becomes

$$\boldsymbol{E}\left[\left(X^2_{T\wedge t_k}-[X]_{T\wedge t_k}\right)-\left(X^2_{T\wedge t_{k-1}}-[X]_{T\wedge t_{k-1}}\right)\right]$$

And each of these things in brackets are a Martingale by the Characterisation of quadratic variation, so we get that the whole thing is equal to zero. Recall that our original goal is to show that

$$\left(\sum_{k=1}^{\infty}H_{t_k}\left(X_{t\wedge t_k}-X_{t\wedge t_{k-1}}\right)\right)^2-\sum_{k=1}^{\infty}H^2_{t_k}\left([X]_{t\wedge t_k}-[X]_{t\wedge t_{k-1}}\right)\in\mathscr{M}_{\text{loc}} \tag{3.32}$$

And so since we have already assumed X has been localised, we just need to pick a bounded stopping time T and show that the expected value of the expression in (3.32) stopped at T is equal to zero. Notice then that

$$\boldsymbol{E}\left(\sum_{k=1}^{\infty}H_{t_k}\left(X_{T\wedge t_k}-X_{T\wedge t_{k-1}}\right)\right)^2=\boldsymbol{E}\left(\sum_{k=1}^{\infty}H^2_{t_k}\left(X_{T\wedge t_k}-X_{T\wedge t_{k-1}}\right)^2\right) \tag{3.33}$$

$$=\boldsymbol{E}\left(\sum_{k=1}^{\infty}H^2_{t_k}\left([X]_{T\wedge t_k}-[X]_{T\wedge t_{k-1}}\right)\right) \tag{3.34}$$

and so we are done. The justification for this last step is similar to that of a proof that we have seen before, and it relies on the fact that H is previsible. Indeed: Let T be a bounded stopping time, then

$$\boldsymbol{E}\left[H^2_{t_k}\left(\left(X^{t_k}_T-X^{t_{k-1}}_T\right)^2-\left([H]^{t_k}_T-[X]^{t_{k-1}}_T\right)\right)\right]$$
$$=\boldsymbol{E}\left[H^2_{t_k}\boldsymbol{E}\left[\left(X^{t_k}_T-X^{t_{k-1}}_T\right)^2-\left([H]^{t_k}_T-[X]^{t_{k-1}}_T\right)\middle|\mathscr{F}_{t_{k-1}}\right]\right]$$

and the thing inside the conditional expectation is zero by the Optional Stopping Theorem. ♡

Corollary 3.6.6 (Itô's Isometry) If $H\in\mathfrak{S}$ and $X\in\mathscr{M}_2$, then

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

Recall that one of the main ideas we have been using so far is that if X is a bounded Martingale, then $X^2-[X]\in\mathscr{M}_2$, but in order to prove Itô's isometry we are going to need something stronger:

Proposition 3.6.7 If $X\in\mathscr{M}_2$, then $X-[X]$ is a UI Martingale. In particular,

$$\boldsymbol{E}X^2_{\infty}=X^2_0+\boldsymbol{E}[X]_{\infty}$$

Proof of (In particular). Let (T_N) be a sequence of stopping times that reduce X to a bounded Martingale. Then $X^2 - [X]$ is a Martingale, and so

$$\boldsymbol{E}[X_{T_N}^2] = X_0^2 + \boldsymbol{E}[X]_{T_N}$$

Now since $X \in \mathscr{M}_2$, we have that $X_{T_N} \to X_\infty$ almost surely and in $\mathscr{L}^2$ by the MGCT, so $\boldsymbol{E}X_{T_N}^2 \to \boldsymbol{E}X_\infty^2$. Now the term $\boldsymbol{E}[X]_{T_N} \to \boldsymbol{E}[X]_\infty$ by the MCT. ♡

Remark 3.6.8 Itô's isometry can be proven for general previsible processes using the machinery we will describe in a second:

This proves Itô's isometry because by an argument similar to that of Remark 3.6.4, we have that if $X \in \mathscr{M}_2$ and $H \in \mathfrak{S}$, then $\int H \,\mathrm{d}X \in \mathscr{M}_2$. With these technical tools we are ready to extend our class of integrands to all previsible process. For this we will use the Hilbert space Machinery we developed for $\mathscr{M}_2$, as well as the following measure-theoretic fact:

Lemma 3.6.9 (Approximation Lemma) Let μ be a finite on a measurable space $(E, \mathscr{E})$, and let $\mathscr{A}$ be a π-system that generates $\mathscr{E}$. Then the set of simple functions

$$\sum_{i=1}^n a_i \mathbf{1}(A_i)$$

for $a_i \in \mathbf{R}$ and and $A_i \in \mathscr{A}$ are dense in $\mathscr{L}^p(E, \mathscr{E}, \mu)$ for all $p \geqslant 1$.

Let us give some technicalities to set up this Lemma in our case. Clearly, we are trying to approximate previsible processes, which can simply be though of as a map $(s, \omega) \mapsto H_s(\omega)$ that is $\mathscr{P}/\mathscr{B}(\mathbf{R})$ measurable. Look at Itô's isometry. Why do you think it is called an isometry? Could it be that we could interpret

$$\boldsymbol{E}\int_0^\infty Y_s^2 \mathrm{d}[X]_s$$

as a norm? Well for a given $X \in \mathscr{M}_2$, we can construct a measure μ_X on the sigma field $\mathscr{P}$, given by $\mu_x(\mathrm{d}x, \mathrm{d}\omega) = [X](\mathrm{d}t, \omega)\boldsymbol{P}[\mathrm{d}\omega]$, or perhaps more clearly, if we define it on the generating sets as

$$\mu_X(s,t] \times A = \boldsymbol{E}[\mathbf{1}_A([X]_t - [X]_s]$$

Here of course $[X](dt, \omega)$ means the measure that we get from $[X]$ for free, since this is a continuous increasing function. Therefore, we can construct the following space $\mathscr{L}^2([0,\infty) \times \Omega, \mathscr{P}, \mu_X)$, in which a previsible process Y has norm

$$\|Y\|_{\mathscr{L}^2(X)}^2 = \boldsymbol{E}\int_0^\infty Y_s^2 \mathrm{d}[X]_s$$

Of course, by the Lemma, we have that one can approximate any $Y \in \mathscr{L}^2([0,\infty)\times\Omega,\mathscr{P},\mu_X)$ by simple previsible processes, and Itô's isometry becomes, for a simple previsible process H,

$$\|H\|_{\mathscr{L}^2(X)} = \left\| \int H \,\mathrm{d}X \right\|_{\mathscr{M}_2}$$

Finally, we are ready to construct the Itô integral for general previsible processes. We will take a slightly non-standard approach and instead of giving theorems and proofs, we explain it like if it were a recipe.

1. Let $Y \in \mathscr{L}^2(X)$ be a previsible process with inherent the integrability condition of being in $\mathscr{L}^2(X)$. By the approximation Lemma, there is a sequence $(Y_n)_n$ of simple processes in $\mathscr{L}^2(X)$ with $Y_n \to Y$ in $\mathscr{L}^2(X)$.

2. Since we have convergence, it must be that $(Y^n)_n$ is also Cauchy in $\mathscr{L}^2(X)$. This means that as $m,n\to\infty$,
$$\left\| \int Y^n \,\mathrm{d}X - \int Y^m \,\mathrm{d}X \right\|_{\mathscr{M}_2} = \|Y^n - Y^m\|_{\mathscr{L}^2(X)} \to 0$$
Where the key step, the middle equality is Itô's isometry. Showing that the sequence $\left(\int Y^n \,\mathrm{d}X\right)_n$ is Cauchy in $\mathscr{M}_2$.

3. Since $\mathscr{M}_2$ is complete, we have that there is some limiting Martingale $M \in \mathscr{M}_2$ for which
$$\int Y^n \,\mathrm{d}X \to M$$
in $\mathscr{M}_2$.

4. We define
$$\int Y \,\mathrm{d}X := M$$

This has proven existence of the Stochastic integral. Of course, there is a big gap in what we have done just now, we haven't shown that this is well defined, what if we obtained a different sequence $\tilde{Y}_n \to Y$ in $\mathscr{L}^2(X)$, could it be that we obtain a different end-result? No.

Proposition 3.6.10 The stochastic integral is well-defined.

Proof. Let (Y_n) and $(\tilde{Y}_n)$ be two sequences of simple previsible process that converge to Y. We will show that the $\mathscr{M}_2$ Martingales M and $\tilde{M}$ that correspond to (Y_n) and $(\tilde{Y}_n)$ as per the above construction are the same. This is a straightforward computation

$$\begin{aligned}\left\|M-\tilde{M}\right\|_{\mathcal{M}_2} &\leqslant \left\|M-M_n\right\|_{\mathcal{M}_2}+\left\|\tilde{M}-\tilde{M}_n\right\|_{\mathcal{M}_2}+\left\|\tilde{M}_n-M_n\right\|_{\mathcal{M}_2} &(3.35)\\ &=\left\|M-M_n\right\|_{\mathcal{M}_2}+\left\|\tilde{M}-\tilde{M}_n\right\|_{\mathcal{M}_2}+\left\|\tilde{Y}_n-Y_n\right\|_{\mathcal{L}^2(X)} &(3.36)\\ &\leqslant \left\|M-M_n\right\|_{\mathcal{M}_2}+\left\|\tilde{M}-\tilde{M}_n\right\|_{\mathcal{M}_2}+\left\|\tilde{Y}_n-Y\right\|_{\mathcal{L}^2(X)}+\left\|Y_n-Y\right\|_{\mathcal{L}^2(X)} &(3.37)\end{aligned}$$

and this whole thing goes to zero. ♡

Another fact that we can prove now that we have a stochastic integral is that Ito's isometry actually extends to all previsible processes, not just simple:

Theorem 3.6.11 (Ito's isometry revisited) Let $X \in \mathcal{M}_2$ and $H \in \mathcal{L}^2(X)$. We then have that

$$\left\|\int H\,\mathrm{d}X\right\|_{\mathcal{M}_2}=\|H\|_{\mathcal{L}^2(X)}$$

that is to say,

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

Proof. Follows immediately by an application of the triangle inequality, and approximating H with a sequence $(H_n)_n$ of simple processes, and $\int H\,\mathrm{d}X$ with a sequence of $\mathcal{M}_2$ Martingales $\int H_n\,\mathrm{d}X$. ♡

The journey was painful but we now we have a stochastic integral. Before we start doing things with it, we need to show some properties of it and later extend the integrators to local martingales.

Theorem 3.6.12 (Stochastic integral behaves well with stopping) Let $X \in \mathcal{M}_2$, $Y \in \mathcal{L}^2(X)$, and T be a stopping time. Then $X^T \in \mathcal{M}_2$, $Y\mathbf{1}(0,T] \in \mathcal{L}^2(X)$, $Y \in \mathcal{L}^2(X^T)$ and

1. $[X^T]=[X]^T$.

2. $$\int Y\mathbf{1}(0,T]\,\mathrm{d}X=\left(\int Y\,\mathrm{d}X\right)^T=\int Y\,\mathrm{d}X^T$$

Dear reader, before embarking on this proof please go take a cup of coffee and a snus.

Proof. Let us start by checking the preliminaries.

- To "show" $[X^T]=[X]^T$ we give some intuition: X^T is equal to X up to time T, after that, it becomes constant, once it becomes constant, it no longer accumulates any quadratic variation, therefore the quadratic variation of the process X^T up to time t is precisely the quadratic variation of X up to time $t \wedge T$.

- $X^T \in \mathscr{M}_2$. The fact that it is a Martingale comes from the Optional Stopping Theorem. To show that $\|X^T\|_{\mathscr{M}_2} < \infty$, we have the following nice trick:

$$\begin{aligned}
\left\|M^T\right\|_{\mathscr{M}_2}^2 &= \boldsymbol{E}\left[X_T^2\right] && (3.38)\\
&\leqslant \boldsymbol{E}\left[\sup_{t\geqslant 0} X_t^2\right] && (3.39)\\
&\leqslant 4\boldsymbol{E}[X_\infty^2] = 4\left\|X\right\|_{\mathscr{M}_2}^2 < \infty. && (3.40)
\end{aligned}$$

  Where (3.40) uses Doob's inequality.

- To show that $Y\,\mathbf{1}(0,T] \in \mathscr{L}^2(X)$ we simply note that $\mathbf{1}(0,T]$ is left-continuous hence previsible and Y is previsible by hypothesis. One should now check square integrability but that's obvious.

- $Y \in \mathscr{L}^2(X^T)$. For this, we note that

$$\boldsymbol{E}\int_0^\infty Y^2 \mathrm{d}[X^T] = \boldsymbol{E}\int_0^\infty Y^2 \mathrm{d}[X]^T = \boldsymbol{E}\int_0^T Y^2\ \mathrm{d}[X] \leqslant \boldsymbol{E}\int_0^\infty Y^2 \mathrm{d}[X] < \infty$$

  where the second equality comes because for a fixed ω, we have that since $[X]^{T(\omega)}$ is constant after time $T(\omega)$, the corresponding measure will assign zero mass to all time intervals with start point after $T(\omega)$.

We now go on to check that these integrals coincide. Let us check firstly that

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

We first check that this holds for simple processes Y, and then extend to the general case. For the case of a simple process Y, we have that

$$\begin{aligned}
\left(\int_0^t Y\,\mathrm{d}X\right)^T &:= \left(\sum_{k=1}^n H_k(X_{t\wedge t_k} - X_{t\wedge t_{k-1}})\right)^T && (3.41)\\
&= \sum_{k=1}^n H_k(X^T_{t\wedge t_k} - X^T_{t\wedge t_{k-1}}) && (3.42)\\
&=: \int_0^t Y\,\mathrm{d}X^T && (3.43)
\end{aligned}$$

We now extend to the general case. For this, let $Y \in \mathscr{L}^2(X)$ and let $(Y^n)_n$ be a sequence of previsible simple processes with $Y^n \to Y$ in $\mathscr{L}^2(X)$. We have already shown that $\int Y^n\,\mathrm{d}X^T = \left(\int Y^n\,\mathrm{d}X\right)^T$. So what we will show is that each of these terms converge to "what they should". In particular:

- $\int Y^n\,\mathrm{d}X^T \to \int Y\,\mathrm{d}X^T$ in $\mathscr{M}_2$. The idea is that $Y^n \to X$ also in $\mathscr{L}^2(X^T)$, then by the construction of the integral the claim will follow. To see the claim:

$$\|Y^n - Y\|_{\mathscr{L}^2(X^T)} = \boldsymbol{E}\int_0^\infty (Y^n - Y)^2 \mathrm{d}[X^T] \tag{3.44}$$

$$= \boldsymbol{E}\int_0^T (Y^n - Y)^2 \mathrm{d}[X] \tag{3.45}$$

$$\leqslant \boldsymbol{E}\int_0^t (Y^n - Y)^2 \mathrm{d}[X] \to 0. \tag{3.46}$$

- $\left(\int Y^n\,\mathrm{d}X\right)^T \to \left(\int Y\,\mathrm{d}X\right)^T$ in $\mathscr{M}_2$. This just follows from the simple fact that if a sequence $(M_n)_n$ in $\mathscr{M}_2$ converges to $M \in \mathscr{M}_2$, then $(M_n^T)_n$ converges in $\mathscr{M}_2$ to M^T. Indeed, by Doob's inequality

$$\|M_n^T - M^T\|_{\mathscr{M}_2} = \left\|(M^n - M)^T\right\|_{\mathscr{M}_2} \leqslant 2\|M^n - M\|_{\mathscr{M}_2} \to 0.$$

♡

So far we have constructed a stochastic integral of a previsible process H against an $\mathscr{M}_2$ Martingale. We now wish to increase our space of integrators, starting by enlarging it to $\mathscr{M}_{\text{loc}}$. To justify the validity of this we have the following proposition

Proposition 3.6.13 ($\mathscr{M}_{\text{loc}}$ integrators) Suppose X is a continuous local Martingale and H is a previsible process such that for any $t \geqslant 0$, almost surely

$$\int_0^t H_s^2 \mathrm{d}[X]_s < \infty$$

Let

$$T_n = \inf\left\{ t > 0 : |X_t| > n \text{ or } \int_0^t H_s^2 \mathrm{d}[X]_s > n \right\}$$

Then $X^{T_n} \in \mathscr{M}_2$, $H\,\mathbf{1}(0, T_n] \in \mathscr{L}^2(X^{T_n})$ and there is some $M^\star \in \mathscr{M}_{\text{loc}}$ for which

$$M^n := \int H\,\mathbf{1}(0, T_n]\,\mathrm{d}X^{T_n} \to M^*$$

UCP.

Proof. We start by noting that $X^{T_n} \in \mathscr{M}_2$. This is because we can think of T_n as the minimum of two stopping times, $T_n = U_n \wedge Q_n$ where U_n is the stopping time that captures when $|X_t|$ gets too large, and Q_n captures when the integral gets too large. Therefore $X^{T_n} = X^{U_n \wedge Q_n} = (X^{U_n})^{Q_n}$ now as we know, the thing on the inside is in $\mathscr{M}_2$, and so since stopping an $\mathscr{M}_2$ Martingale leaves it in $\mathscr{M}_2$ by Doob's inequality (Indeed if $M \in \mathscr{M}_2$, then $\|M^T\|_{\mathscr{M}_2} = \boldsymbol{E}[M_T^2] \leqslant \boldsymbol{E}[\sup_t M_t^2] \lesssim \boldsymbol{E}[M_\infty^2] < \infty$).

Moreover, since

$$\boldsymbol{E}\int_0^t H_s^2 \mathbf{1}(0,T_n]\mathrm{d}[X^{T_n}] = \boldsymbol{E}\left(\int_0^t H_s^2 \mathrm{d}[X]\right)^{U_n\wedge Q_n} < \infty$$

by definition of Q_n, we have that it indeed makes sense to talk about the processes M^n. We start by constructing M^*. Fix a value of t. Then observe that on the event $\{t < T_n\}$, we have that $M^n = M^N$ whenever $N \geqslant n$. Now note the following, since $T_n \to \infty$ almost surely, it follows that for all t, almost surely the sequence M_t^n will become constant and equal to some M_t^*. Furthermore, we note that if $\{t < T_n\}$ holds, then for all $s \in [0,t]$, we have that $|M_s^n - M_s^*| = 0$. In particular, for all $\epsilon > 0$, for all $t \geqslant 0$:

$$\boldsymbol{P}\left[\sup_{s\in[0,t]} |M_s^n - M_s^*|\right] \leqslant \boldsymbol{P}[T_n \leqslant t] \to 0$$

To check that the process M^* is a local Martingale, we note that if you stop M at T_n then its just M^n, which is an $\mathscr{M}_2$ Martingale (?). ♡

Definition 3.6.14 (Stochastic integral for local Martingales) Let H be previsible and $X \in \mathscr{M}_{\text{loc}}$, then

$$\int H \mathrm{d}X$$

is defined to be the UCP limit of $\int H\mathbf{1}(0,T_n]\mathrm{d}X^{T_n}$ for the stopping times defined as

$$T_n = \inf\left\{t > 0 : |X_t| > n \text{ or } \int_0^t H_s^2 \mathrm{d}[X]_s > n\right\}$$

Now we can give the final version of the stochastic integral, the integral with respect to semi-Martingales.

Definition 3.6.15 (Semi-Martingale) A continuous semi-Martingale X is a process of the form

$$X_t = X_0 + A_t + M_t$$

where A is a continuous adapted finite variation process, and M is a continuous local Martingale. We also take $A_0 = M_0 = 0$.

Proposition 3.6.16 (Uniqueness of semi-Martingale decomposition) Let X be a semi-Martingale. Then the processes A and M in its definition above are unique up to indistinguishability.

Proof. If we write

$$X_t = X_0 + A_t + M_t = X_0 + A'_t + M'_t$$

we can rearrange and obtain that

$$A - A' = M - M'$$

the right-hand-side is a continuous local Martingale, whereas the left-hand-side is a finite variation process. We know that continuous local Martingales of finite variation are constant and so it means almost surely for all t, $M_t - M'_t = M_0 - M'_0 = 0$. Hence $M = M'$ and $A = A'$. ♡

We can define therefore define integrals with respect to semi-Martingales in an unambiguous way:

Definition 3.6.17 (The Stochastic Integral) Let $X = X_0 + A + M$ be a semi-Martingale, and let H be a previsible process. Then

$$\int H \mathrm{d}X = \int H \mathrm{d}A + \int H \mathrm{d}X$$

where the first integral is interpreted as a Lebesgue-Stieltjes integral, and the second one as we have already discussed.

Definition 3.6.18 (Local Integrability) Let $X = X_0 + A + M$ be a semi-Martingale, and let H be a previsible process. We say H is locally X integrable if for all $t \geqslant 0$, we have that

$$\int_0^t |H_s| \mathrm{d}|A|_s + \int_0^t H_s^2 \mathrm{d}[M]_s < \infty \text{ a.s}$$

Definition 3.6.19 (Local Boundedness) A previsible process H is said to be locally bounded if there exists a sequence of stopping times (T_n) with $T_n \uparrow \infty$, and constants $C_n > 0$ for which, for any (t, ω), one has that

$$|H_t(\omega)| \mathbf{1}\{t \in (0, T_n(\omega)]\} \leqslant C_n$$

Remark 3.6.20 If H is locally bounded, then it is locally X-integrable for any continuous semi-Martingale X. Indeed: we know that with probability 1, there is some n large enough after which

$$\int_0^t H_s \mathrm{d}|A|_s = \int_0^t H_s \mathbf{1}(0, T_n] \mathrm{d}|A|_s$$

On this event, we clearly have that $\int_0^t |H_s| \mathrm{d}|A|_s \leqslant C_n V_A(t) < \infty$. A similar argument applies to the other condition.

3.7 Quadratic Covariation of semi-Martingales

Definition 3.7.1 Let $X = X_0 + M_t + A_t$ be a continuous semi-Martingale, then the quadratic variation of X is defined as $[X] := [M]$.

This definition is not entirely arbitrary, it is justified by the following:

Proposition 3.7.2 Let $X = X_0 + A + M$ be a continuous semi-Martingale, let $[X]^n_t$ be the proto-quadratic variation. Then $[X]^n \to [X] := [M]$ UCP.

Proof. Already have that $[M]^n \to [M]$ UCP. The idea will be to use some sort of triangle inequality argument to show that $[X^n] \to [X]$ UCP. For this we note that for any fixed $t \geqslant 0$ we have that

$$\sup_{s\in[0,t]} |[X]^n - [M]^n| = \sup_{s\in[0,t]} \left| \sum_{k\geqslant 1} \left(A_{t\wedge t_k^n} - A_{t\wedge t_{k-1}^n}\right)\left[2\left(M_{t\wedge t_k^n} - M_{t\wedge t_{k-1}^n}\right) + \left(A_{t\wedge t_k^n} - A_{t\wedge t_{k-1}^n}\right)\right]\right| \tag{3.47}$$

$$\leqslant V_A(t) \sup_{|u-v|\leqslant 2^{-n}} |2(M_u - M_v) + A_u - A_v| \tag{3.48}$$

but since the function $2M + A$ is continuous on the compact interval $[0,t]$, it's uniformly continuous, so this whole quantity goes to zero almost surely. Now putting it all together we have that

$$\sup_{s\in[0,t]} |[X]^n - [M]| \leqslant \sup_{s\in[0,t]} |[X]^n - [M]^n| + \sup_{s\in[0,t]} |[M]^n - [M]|$$

the first term goes to zero almost surely, the second one goes to zero in probability, therefore the whole thing goes to zero in probability and we have the desired UCP convergence. ♡

Theorem 3.7.3 (Quadratic variation of the stochastic integral) Let X be a semi-Martingale, and let H be X-locally integrable. Then

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

Proof. Since by definition $[X]$ ignores the finite variation part, we may as well suppose that X is a continuous local Martingale. Hence by localisation, we may assume that $X \in \mathscr{M}_2$ and $H \in \mathscr{L}^2(X)$ (This integrability condition is precisely what being locally integrable in this context means). We must show that

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

is a Martingale. This is exactly what proposition 3.6.5 says when H is a simple process. (Let us mention what this localisation trick is actually saying. To show that a process A is the quadratic variation of a local Martingale X, you need to show that $X - A$ is a local Martingale, therefore, this is the same as showing that for some sequence of stopping times (T_n), $(X-A)^{T_n} = X^{T_n} - A^{T_n}$ is a Martingale. Then one extends to the general case of previsible processes using an Example Sheet question. ♡

Definition 3.7.4 (Quadratic covariation) Let X and Y be continuous semi-Martingales. The quadratic covariation $[X,Y]$ is defined as

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

Remark 3.7.5 Note that $[X,Y]$ is the difference of two increasing (hence finite variation) process, so it is itself a finite variation process, this means that integrals of the type

$$\int H \mathrm{d}[X,Y]$$

can be made sense of as a Lebesgue-Stieltjes integral.

There is perhaps a more natural way to approach this definition:

Proposition 3.7.6 (Alternative definition of covariation) Let X and Y be continuous semi-Martingales, then

$$\sum_{k\geqslant 1}\left(X_t^{t_k^n} - X_t^{t_{k-1}^n}\right)\left(Y_t^{t_k^n} - Y_t^{t_{k-1}^n}\right) \to [X,Y]$$

UCP.

Proof. The term on the sum can easily be seen to be

$$\frac{1}{4}\left([X+Y]^n - [X-Y]^n\right)$$

both of these things converge UCP to their respective quadratic variations. Then by definition of the polarisation identity the claim follows. ♡

Now we collect some properties of this process.

Proposition 3.7.7 (Characterisation of quadratic covariation) Let X and Y be continuous local Martingales, then $[X,Y]$ is the unique finite variation process starting at zero for which $XY-[X,Y]$ is a local Martingale.

Proof. Let A be a finite variation process for which $XY-A=M$ is a local Martingale. Then we can rewrite

$$A=\frac{1}{4}\left((X+Y)^2-(X-Y)^2\right)-\frac{1}{4}([X+Y]-[X-Y])+[X,Y]-M$$

But rearranging we have that

$$A-[X,Y]=\frac{1}{4}\left(\underbrace{(X+Y)^2-[X+Y]}_{\in\mathscr{M}_{\text{loc}}}+\underbrace{(X-Y)^2-[X-Y]}_{\in\mathscr{M}_{\text{loc}}}\right)-\underbrace{M}_{\in\mathscr{M}_{\text{loc}}}$$

is a local Martingale, and since A is assumed to be of finite variation, and $[X,Y]$ being the difference of two increasing (hence finite variation) processes is itself of finite variation, it follows that $A-[X,Y]$ is a local Martingale of finite variation, hence constant and by assumption of started at zero we have that they are both equal. To show that $[X,Y]$ satisfies the claim is clear.

♡

Corollary 3.7.8 (Bilinearity of quadratic covariation) One has that

$$[X+Y,Z]=[X,Z]+[Y,Z]$$

almost surely.

Proof. A way to prove this is using the UCP limit above and then justifying why the result holding "in UCP" means that it holds almost surely. A better way to prove this is by using the characterisation of Quadratic covariation, indeed, we note that

$$(X+Y)Z-([X,Z]+[Y,Z])$$

is obviously a local Martingale, and so we are done. ♡

Corollary 3.7.9 If X is a semi-Martingale of finite variation, then $[X,Y]=0$.

Proof. We simply have that $[X+Y] = [X-Y] = [Y]$. Because the definition of the quadratic variation of a semi-Martingale only includes the local-Martingale term. ♡

Theorem 3.7.10 (Kunita-Watanabe Inequality) Let X and Y be continuous semi-Martingales and let H be locally X-integrable. Then

$$\int_0^t |H||\mathrm{d}[X,Y]| \leqslant \sqrt{[Y]_t \int_0^t H^2 \mathrm{d}[X]}$$

Theorem 3.7.11 (Kunita-Watanabe identity) Let X and Y be continuous semi-Martingales and let H be locally X-integrable. Then H is locally $[X,Y]$ integrable and

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

Proof. We can consider the case where X and Y have only their local-Martingale part. By the characterisation of quadratic covariation, we must show that

$$Z = Y \int H\,\mathrm{d}X - \int H \mathrm{d}[X,Y] \in \mathscr{M}_{\text{loc}}$$

By localisation, we can assume that $X, Y \in \mathscr{M}_2$ and so we need to show that Z is a true Martingale. We first show that this is true for simple processes. By linearity of the integral, it even suffices to show that it works for processes of the form $H = K\,\mathbf{1}(s_0, s_1]$ for $K \in \mathrm{b}\mathscr{F}_{s_0}$. If we compute the expressions above, we have that in this case

$$Z_t = K Y_t \left(X_{t\wedge s_1} - X_{t\wedge s_0}\right) - K\left([X,Y]_{t\wedge s_1} - [X,Y]_{t\wedge s_0}\right)$$

which can be conveniently rewritten as

$$Z_t = K Y_t \left(X_t^{s_1} - X_t^{s_0}\right) - K\left([X^{s_1}, Y]_t - [X^{s_0}, Y]_t\right)$$

if it weren't for the K, this would clearly be a Martingale, but as we have seen in a prior calculation, since K is $\mathscr{F}_{s_0}$ measurable, the Martingale property is still there. Indeed, let T be a bounded stopping time, then

$$\boldsymbol{E} Z_T = \boldsymbol{E}\left[K \boldsymbol{E}\left[Y_T X_T^{s_1} - [X^{s_1}, Y]_T - Y_T X_T^{s_0} - [X^{s_0}, Y]_T \middle| \mathscr{F}_{s_0}\right]\right] \tag{3.49}$$

and now one sees that this becomes zero. Therefore the converse of the OST is satisfied, and so the claim is proven for simple processes. For general processes one renames the processes H, Y, X to $H\mathbf{1}(0,T], Y^T, X^T$ for a bounded stopping time T, and so the goal is to show that

$$\boldsymbol{E}\left[Y_\infty \int_0^\infty H_s \mathrm{d}X_s - \int_0^\infty H_s \mathrm{d}[X,Y]_s\right] = 0$$

This is accomplished through a series of not very stimulating calculations. (See Lecture 12 notes) ♡

3.8 Itô's Formula

After all this dry theory, we can provide one of the fundamental results of Stochastic calculus:

Theorem 3.8.1 (Itô's formula) Let $X = (X^1, \cdots, X^n)$ be an n-dimensional continuous semi-Martingale. Let $f \in C^2(\mathbf{R}^n, \mathbf{R})$. then

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

The proof will come after a series of smaller results. The first one is an analogue to the Riemann sum:

Lemma 3.8.2 Let X be a continuous semi-Martingale, and Y a locally bounded, adapted, left continuous process, then

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

Proof. Since $X = X_0 + A + M$ where A is of finite variation, and M is a continuous local Martingale, we need to show that

$$\int_0^t Y\,\mathrm{d}A \overset{\text{UCP}}{=} \lim_{n\to\infty} \sum_{k\geqslant 1} Y_{t_{k-1}^n}\left(A_{t_k^n} - A_{t_{k-1}^n}\right)$$

and

$$\int_0^t Y\,\mathrm{d}M \overset{\text{UCP}}{=} \lim_{n\to\infty} \sum_{k\geqslant 1} Y_{t_{k-1}^n}\left(M_{t_k^n} - M_{t_{k-1}^n}\right)$$

Let us start with the dM integral. Recall that Y is locally bounded and M is a local Martingale, so we can rewrite Y for $Y\mathbf{1}(0,U_n)$ and M for M^{T_n} where (U_n) and (T_n) are a sequence of stopping times for which $Y \leqslant C_n$ and $M^{T_n} \in \mathscr{M}_2$. Now note that by left-continuity of Y, we have that

$$Y^n = \sum_{k\geqslant 1} Y_{t_k^n}\mathbf{1}(t_{k-1}^n, t_k^n]$$

converges pointwise to Y, so by the dominated convergence Theorem, (Y^n is also uniformly bounded), we have that

$$\int_{\Omega\times\mathbf{R}^+} (Y^n - Y)^2 \,\mathsf{d}[M] \otimes \mathsf{d}\boldsymbol{P} = \boldsymbol{E}\left[\int_0^\infty (Y^n - Y)^2 \,\mathsf{d}[M]\right] \to 0$$

and so $\int Y^n \,\mathsf{d}[M] \to \int Y \,\mathsf{d}[M]$ in $\mathscr{M}_2$. Now this allows us to conclude UCP convergence for these localised processes:

$$\boldsymbol{P}\left[\sup_{s\in[0,t]}\left|\int_0^s Y^n - Y \,\mathsf{d}[M]\right| > \epsilon\right] \leqslant \frac{1}{\epsilon^2}\boldsymbol{E}\left[\sup_{s\in[0,t]}\left|\int_0^s Y^n - Y \,\mathsf{d}[M]\right|^2\right] \tag{3.50}$$

$$\lesssim \boldsymbol{E}\left[\left|\int_0^t Y^n - Y \,\mathsf{d}[M]\right|^2\right] \tag{3.51}$$

$$\leqslant \boldsymbol{E}\left[\int_0^t (Y^n - Y)^2 \,\mathsf{d}[M]\right] \to 0 \tag{3.52}$$

Where (3.50) is Markov's Inequality, (3.51) is Doob's Inequality, (3.52) is Jensen's Inequality and the limit we saw just now. To recap, we have shown that for some stopping times (Q_m) with $Q_m \uparrow \infty$,

$$\boldsymbol{P}\left[\sup_{s\in[0,t]}\left|\int_0^s Y^n - Y \,\mathsf{d}[M]\right|^{Q_m} > \epsilon\right] \to 0$$

Now one needs to "remove" the Q_m. For this, let us momentarily define $Z_t^n = \sup_{s\in[0,t]}\left|\int_0^s Y^n - Y \,\mathsf{d}[M]\right|$ for convenience. Since $Q_m \uparrow \infty$ almost surely, it is true that the event $\bigcup_m \{Z_t^n = (Z_t^n)^{Q_m}\}$ has probability one. Therefore, we have that

$$\begin{aligned}
\boldsymbol{P}\left[Z_t^n > \epsilon\right] &= \boldsymbol{P}\left[\{Z_t^n > \epsilon\} \cap \bigcup_m \{Z_t^n = (Z_t^n)^{Q_m}\}\right] \\
&\leqslant \boldsymbol{P}\left[\bigcup_m \{(Z_t^n)^{Q_m} > \epsilon\}\right] \\
&= \lim_{m\to\infty} \boldsymbol{P}[(Z_t^n)^{Q_m} > \epsilon]
\end{aligned}$$

Where this last limit is due to increasing events. This means that for any arbitrarily small $\eta > 0$, there is some $M > 0$ for which

$$\boldsymbol{P}\left[\sup_{s\in[0,t]}\left|\int_0^s Y^n - Y \,\mathsf{d}[M]\right|\right] \leqslant \boldsymbol{P}\left[\sup_{s\in[0,t]}\left|\int_0^s Y^n - Y \,\mathsf{d}[M]\right|^{Q_M}\right] + \eta$$

and we have seen how the first term goes to zero. I'm not going to show the second part of the proof, i.e: the $\mathsf{d}A$ integral, but it follows by the DCT and localisation. ♡

With this out of the way, we can prove the following

Lemma 3.8.3 (Stochastic integration by parts) Let X and Y be continuous semi-Martingales, then

$$XY = \int X \, \mathrm{d}Y + \int Y \, \mathrm{d}X + [X, Y]$$

or in differential notation

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

Proof. We have the following limits in UCP sense.

$$\begin{aligned}
2\int_0^t X \, \mathrm{d}X &= 2 \lim_{n \to \infty} \sum_{k \geqslant 1} X_{t_{k-1}^n} \left(X_{t \wedge t_k^n} - X_{t \wedge t_{k-1}^n} \right) \\
&= 2 \lim_{n \to \infty} \sum_{k \geqslant 1} X^2_{t \wedge t_k^n} - X^2_{t \wedge t_{k-1}^n} - \left(X_{t \wedge t_k^n} - X_{t \wedge t_{k-1}^n} \right)^2 \\
&= X_t^2 - X_0^2 - [X]_t
\end{aligned}$$

But as we've discussed on prior occasions, since these equalities in UCP end up not depending on n, we have that $2\int_0^t X \, \mathrm{d}X = X_t^2 - X_0^2 - [X]_t$ almost surely. From this, we can apply polarisation identities to $(X + Y)$ and $(X - Y)$ using the above equation with these instead of X. ♡

With this in mind we are ready to prove Itô's formula

Proof of Itô's formula. We do for simplicity the case of dimension $n = 1$. Let us start proving the claim for polynomials. Let us prove it for monomials and then linearity does the rest. Clearly the claim holds for $f(X_t) = X_t^0$, so now suppose that for some m, one has that

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

Using the inductive hypothesis, bilinearity and Kunita-Watanabe we have that

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

where the last equality comes from the fact that since $[X]$ is of finite variation, Corollary 3.7.9 says that $[[X],X]$ vanishes. Thus in differential notation, $\mathrm{d}[X^m,X] = mX^{m-1}\mathrm{d}[X]$. From this it follows that

$$\begin{aligned}
\mathrm{d}(X^{m+1}) &= \mathrm{d}(XX^m) && (3.53)\\
&= X^m\,\mathrm{d}X + X\,\mathrm{d}(X^m) + \mathrm{d}[X^m,X] && (3.54)\\
&= X^m\,\mathrm{d}X + X\,\mathrm{d}\left(\int mX^{m-1}\,\mathrm{d}X + \int \frac{m(m-1)}{2}X^{m-2}\,\mathrm{d}[X]\right) + mX^{m-1}\,\mathrm{d}[X] && (3.55)\\
&= X^m\,\mathrm{d}X + mX^m\,\mathrm{d}X + \frac{m(m-1)}{2}X^{m-1}\,\mathrm{d}[X] + mX^{m-1}\,\mathrm{d}[X] && (3.56)\\
&= (m+1)X^m\,\mathrm{d}X + \frac{(m+1)m}{2}X^{m-1}\,\mathrm{d}[X] && (3.57)
\end{aligned}$$

Thus proving the formula for monomials. In step (3.54) we used the stochastic integration by parts, in step (3.559 we used the inductive hypothesis and the calculation done just before. In step (3.56) we used the stochastic chain rule (Example Sheet 2), that says that

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

Then step (3.57) is just putting it all together. Thus Itô's formula is proven for polynomials and so now one just needs to extend it to general C^2 functions. This is done with Weierstrass' Theorem. Finish

$\heartsuit$

Chapter 4

Applications to Brownian Motion

4.1 The Brownian Motion characterisation of $\mathscr{M}_{\text{loc}}$.

We can now use the stochsatic calculus we have developed to provide some interesting results about Brownian Motion.

Theorem 4.1.1 (Lévy's Characterisation of Brownian Motion) Let X be a continuous d-dimensional local Martingale with $X_0 = 0$ and quadratic covariation $[X^i, X^j] = t\delta_{ij}$, then X is standard Brownian Motion.

Main idea: The key is to consider the process, for $\theta \in \mathbf{R}^d$

$$M_t = \exp\left(i\theta \cdot X_t + \|\theta\|^2 t/2\right)$$

show it is a Martingale, and therefore show using the Martingale property, that the characteristic function of X is that of a Normal random variable.

Proof. Let $\theta \in \mathbf{R}^d$, and consider

$$M_t = \exp\left(i\theta \cdot X_t + \|\theta\|^2 t/2\right)$$

Then to show that M_t is a Martingale, we start by showing it is a local Martingale, for this we use Itô's formula

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

which gives that

$$\mathrm{d}M_t = iM_t\theta \cdot \mathrm{d}X_t$$

which means that M_t is a local Martingale, since it is the integral of a continuous (hence locally integrable) process against a local Martingale. Now we show that M is actually a Martingale, for this, we start by noting that for a fixed t, $\sup_{s\in[0,t]}|M_s| = \exp\left(\frac{\|\theta\|^2 t}{2}\right)$, which is integrable. We will now use this to show M is a Martingale. Let $(T_n)_n$ be a sequence of stopping times with $T_n \to \infty$ such that M^{T_n} is a Martingale. Then

$$\begin{aligned} M_s &= \lim_{n\to\infty} M_{s\wedge T_n} \\ &= \lim_{n\to\infty} \boldsymbol{E}[M_{t\wedge T_n} \mid \mathscr{F}_s] \\ &= \boldsymbol{E}[M_t \mid \mathscr{F}_s] \end{aligned}$$

Where the last equality, came from the fact that $|M_{t\wedge T_n}| \leqslant \sup_{s\in[0,t]}|M_s|$, which we showed above that is integrable, hence by the DCT we were able too introduce the limit inside the expectation. Now that we know that M is a Martingale, we have that

$$\boldsymbol{E}\left[\exp\left(i\theta\cdot(X_t - X_s)\right)|\mathscr{F}_s\right] = \exp\left(-\frac{\|\theta\|^2}{2}(t-s)\right)$$

so by taking expectation of both sides, we have that the characteristic function of the increments $X_t - X_s$ is that of a $\mathscr{N}(0, t-s)$ random variable. Moreover, since the conditional expectation given $\mathscr{F}_s$ is a non-random object, we have that $X_t - X_s$ is independent of $\mathscr{F}_s$. (for a proof of this fact consult [Teh16, Page 21]). Therefore X is Brownian Motion. ♡

The following result is a clever consequence of Lévy's Characterisation:

Theorem 4.1.2 (Dubins-Schwarz Theorem) Let X be a scalar continuous local Martingale for a filtration $(\mathscr{F}_t)_t$ so that $X_0 = 0$ and $[X]_\infty = \infty$ a.s. Define the the stopping times

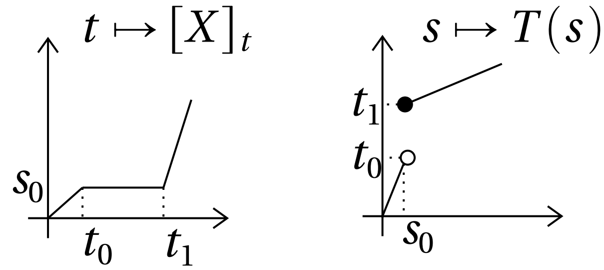
$$T(s) = \inf\{t \geqslant 0 : [X]_t = s\},$$

and the family of random variables

$$W_s = X_{T(s)}$$

with the sigma-algebras $\mathscr{G}_s = \mathscr{F}_{T(s)}$. Then $(\mathscr{G}_s)_s$ is a filtration and the process W is a Brownian Motion in $(\mathscr{G}_s)_s$. (To clarify, when we say W is a Brownian Motion in $(\mathscr{G}_s)_s$ we mean that W is adapted to this filtration and that the increments $W_t - W_s$ are independent of $\mathscr{G}_s$)

Proof. We start by noting that for a fixed $\omega \in \Omega$, the map $t \mapsto [X]_t(\omega)$ is continuous and non-decreasing. Hence the map $s \mapsto T(s,\omega)$ is increasing and right-continuous.



Let us first show that $(\mathscr{G}_s)$ is a filtration. This is immediate from the fact that T is increasing. The remaining ingredients to check are: to show that W is continuous, to check that W is a local Martingale, and to check that the quadratic variation of W is $[W]_t = t$. This last condition is immediate from the continuity of X and the definition of T. Now let us show that W is continuous. The only doubt to see that W is continuous comes from the possible jumps of T, that is to say, when the quadratic variation of X is constant. So in this regard, if we can show that whenever $[X]$ is constant then so is X we will be done. I.e. we must show that whenever $[X]_{t_1} = [X]_{t_2}$, then $X_{t_1} = X_{t_0}$. Let t_0 be given and let $U = \inf\{u > t_0 : [X]_u > [X]_{t_0}\}$. We will show that $X_U = X_{t_0}$ (note that in our diagram $U(\omega) = t_1$). For this, consider the process

$$\begin{aligned} Y_u &= X_{u \wedge U} - X_{u \wedge t_0} \\ &= \int_0^u \mathbf{1}(t_0, U](r)\, \mathrm{d}X_r \end{aligned}$$

since Y is the integral of a left-continuous process with respect to a local Martingale, we have that Y_u is itself a local Martingale. Hence it has quadratic variation

$$[Y]_\infty = \int_0^\infty \mathbf{1}(t_0, U]\, \mathrm{d}X_r = [X]_U - [X]_{t_0} = 0$$

Therefore we have that the quadratic variation of Y is always zero (because if the quadatic variation at infinity is zero, and the quadratic variation is a non-decreasing process, we must have that $[Y]_t = 0$ for all t) and since if a local Martingale has zero quadratic variation it must be almost surely constant (Indeed, if $[Y] = 0$, then by the characterisation of quadratic variation we must have that Y^2 is a Martingale, which means that $\boldsymbol{E}[Y_t^2 - Y_0^2] = 0$, and since $\boldsymbol{E}[(Y_t - Y_0)^2] = \boldsymbol{E}[Y_t^2 - Y_0^2]$, we get that $Y_t - Y_0$ almost surely). This shows that $X_U = X_{t_0}$ almost surely, and so we have almost sure continuity (This is enough for Lévy, no need for strict continuity). We now show that W is a local Martingale, then we will be done. First of all, since X is a local Martingale, let τ_N be the "canonical localising sequence", so that X^{τ_n} is a bounded Martingale (with respect to the filtration $\{\mathscr{F}_t\}$). Then define

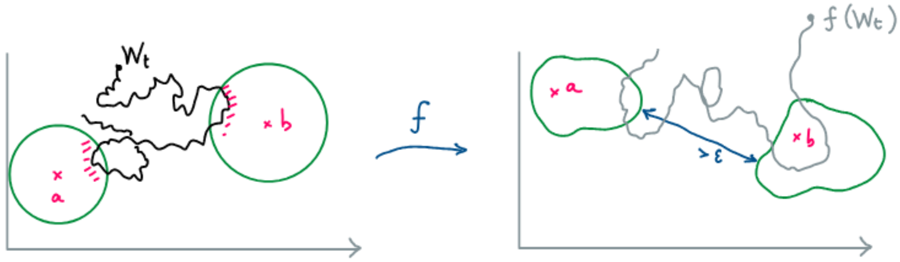


Figure 4.1: Proof by picture: conformal invariance of Brownian Motion

$\sigma_N = [X]_{\tau_N}$, so that

$$\begin{aligned}
\boldsymbol{E}[W_{\sigma_N \wedge s_1} \mid \mathscr{G}_{s_0}] &= \boldsymbol{E}[X_{T(\sigma_N \wedge s_1)} \mid \mathscr{G}_{s_0}] \\
&= \boldsymbol{E}[X_{\tau_N \wedge T(s_1)} \mid \mathscr{F}_{T(s_0)}] \\
&= X_{\tau_N \wedge T(s_0)} = W_{\sigma_N \wedge s_0}
\end{aligned}$$

Where in the first step we used the definition of W, in the second one we used the fact that $T(\sigma_N \wedge s_1) = \inf\{t : [X_t] > ([X]_{\tau_N} \wedge s_)\}$, and in the last step we used the fact that X^{τ_N} is a Martingale, along the OST. I suppose that for full clarity we can note that $[W]_s = [X]_{T(s)}$ because $W_s^2 - s = X_{T(s)}^2 - [X]_{T(s)}$ and this latter is a Martingale with respect to $(\mathscr{G}_s)_s$. ♡

Remark 4.1.3 The power of this Theorem comes from the fact that if $X \in \mathscr{M}_{\text{loc}}$ with $X_0 = 0$ and $[X]_\infty = \infty$, then X is a time-change of a Brownian Motion. Meaning, there exists a Brownian Motion for which $X_t = W_{[X]_t}$. It turns out that the assumption of infinite quadratic variation can be relaxed. This finally gives the intuition that was hinted at, that said that the quadratic variation is the clock of a Martingale, moreover, this Theorem also says that Brownian Motion is in some sense, the universal local Martingale.

4.2 Conformal Invariance of Brownian Motion

Theorem 4.2.1 (Conformal Invariance of Brownian Motion) Let X and Y be independent standard Brownian Motions in $\mathbf{R}$ and so that $W = X + iY$ is a Complex Brownian Motion (planar Brownian Motion). Let $f : \boldsymbol{C} \to \boldsymbol{C}$ be a holomorphic function. Then

$$f(W_t) = f(0) + \widehat{W}_{A_t}$$

Where $\widehat{W}$ is a planar Brownian Motion and $(A_t)_t$ is a non-decreasing process.

Proof. Let $f(x+iy) = u(x,y) + iv(x,y)$, so that $f(W_t) = U_t + iV_t$. We can assume that U_t and V_t start at zero so that we can write $f(W_t) = f(0) + U_t + iV_t$. The goal is to show that $U_t = \widehat{X}_{A_t}$ and $V_t = \widehat{Y}_{A_t}$ for some increasing process A and a pair of independent Brownian Motions $\widehat{X}$ and $\widehat{Y}$. We first note that since $U_t = u(W_t)$ and u being the real part of a holomorphic function is C^∞, we can apply Itô's formula:

$$\mathrm{d}u(W_t) = \left(\frac{\partial u}{\partial x}\mathrm{d}X + \frac{\partial u}{\partial y}\mathrm{d}Y\right) + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\mathrm{d}[X] + 2\frac{\partial^2 u}{\partial x \partial y}\mathrm{d}[X,Y] + \frac{\partial^2 u}{\partial y^2}\mathrm{d}[Y]\right)$$

and since $[X,Y] = 0$ because they are independent, and $\Delta u = 0$ since u is harmonic, we have that

$$\mathrm{d}U_t = \frac{\partial u}{\partial x}\mathrm{d}X + \frac{\partial u}{\partial y}\mathrm{d}Y$$

In a similar fashion,

$$\mathrm{d}V_t = \frac{\partial v}{\partial x}\mathrm{d}X + \frac{\partial v}{\partial y}\mathrm{d}Y$$

From this we learn that U and V, being integrals of continuous functions against local Martingales, are themselves local Martingales, which means we are a step closer to be able to use the DDS Theorem. Next, let us compute the quadratic variations of U and V, for this, we will need the Kunita-Watanabe Theorem:

$$\begin{aligned}
[U] &= \left[\int \frac{\partial u}{\partial x}\mathrm{d}X + \int \frac{\partial u}{\partial y}\mathrm{d}Y, \int \frac{\partial u}{\partial x}\mathrm{d}X + \int \frac{\partial u}{\partial y}\mathrm{d}Y\right] \\
&= \int \left(\frac{\partial u}{\partial x}\right)^2 \mathrm{d}[X] + \int \left(\frac{\partial u}{\partial y}\right)^2 \mathrm{d}[Y] + \left[\int \frac{\partial u}{\partial x}\mathrm{d}X, \int \frac{\partial u}{\partial y}\mathrm{d}Y\right] \\
&= \int \left(\frac{\partial u}{\partial x}\right)^2 \mathrm{d}[X] + \int \left(\frac{\partial u}{\partial y}\right)^2 \mathrm{d}[Y] + \int \frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\underbrace{\mathrm{d}[X,Y]}_{=0}
\end{aligned}$$

and so

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

similarly

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

But using the fact that $\mathrm{d}[X] = \mathrm{d}[Y] = \mathrm{d}t$, and the Cauchy-Riemann equations, we have that $\mathrm{d}[U] = \mathrm{d}[V]$, and in turn, since the derivative of a complex function is the sum of the derivatives of the real and imaginary components in the real direction, we have that

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

we can call these quadratic variations say A_t, and so we have that

$$[U]_t = [V]_t = A_t = \int_0^t |f'(W_t)|^2 \, \mathrm{d}t$$

If we are able to show that $A_\infty = \infty$, then we'll have that $[U]_\infty = [V]_\infty = \infty$ and so by DDS we will have that $U_t = \widehat{X}_{A_t}$ and $V_t = \widehat{Y}_{A_t}$ for two Brownian motions $\widehat{X}$ and $\widehat{Y}$. Moreover, note that

$$\mathrm{d}[U,V] = \left(\left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial v}{\partial x}\right) + \left(\frac{\partial u}{\partial y}\right)\left(\frac{\partial v}{\partial y}\right) \right) \mathrm{d}t = 0$$

So in particular we will have that $[\widehat{X},\widehat{Y}]_{A_t} = [U,V]_t = 0$ and since A_t is increasing and continuous we have that in fact $[\widehat{X},\widehat{Y}]_t = 0$ for all t, and so the two Brownian Motions are independent. Let us then finish off the proof by showing that $A_\infty = \infty$. We will do this by using the recurrence properties of Brownian Motion. Since f is non-constant, there are points a and b such that $f(a) \neq f(b)$, and so by continuity of f there will be some $\epsilon > 0$, and disks D_1 and D_2 about a and b such that for any $\alpha \in D_1$ and all $\beta \in D_2$, $|f(\alpha) - f(\beta)| > \epsilon$. Since planar Brownian motion is neighbourhood-recurrent, it will visit D_1 and D_2 infinitely-often almost surely. Which means that $\boldsymbol{P}[f(W_t) \text{ converges}] = 0$. But since for a general local Martingale X, we have that $\{[X]_\infty < \infty\} \subseteq \{X \text{ converges}\}$, and since

$$[f(W)]_t = \int |f'(W_s)|^2 \, \mathrm{d}s,$$

(This comes from Itô's formula that says that $f(W_t) = \int_0^t f'(W_s)\,\mathrm{d}W_s + \frac{1}{2}\int_0^t f''(W_s)\,\mathrm{d}s$, and so takinig quadratic variations and using Kunita-Watanabe gives the above claim) we have that $A_\infty = \infty$. ♡

4.3 Cameron-Martin-Girsanov Theorem

In many applications of stochastic calculus, we often encounter situations where a Brownian motion acquires a drift. A key question is how this affects its distribution. The Cameron-Martin-Girsanov Theorem provides a precise answer by showing that a drifted Brownian motion can be viewed as a standard Brownian motion under an equivalent change of measure, with the Radon-Nikodym derivative given by an exponential martingale. Let us first recall some preliminary concepts:

Definition 4.3.1 (Equivalent measures) Two probability measures $\boldsymbol{P}$ and $\boldsymbol{Q}$ are said to be equivalent on $(\Omega, \mathscr{F})$ if $\boldsymbol{P}[A] = 0$ if and only if $\boldsymbol{Q}[A] = 0$. We may use the symbol $\boldsymbol{P} \sim \boldsymbol{Q}$ to denote equivalence.

We already know the Radon-Nikodym Theorem from a second course in Probability, but let us give a nicer version:

Theorem 4.3.2 (Filtered Radon-Nikodym) Let $\boldsymbol{P} \sim \boldsymbol{Q}$ be two equivalent probability measures on a filtered measurable space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_t)$. Then there exists a uniformly integrable Martingale Z such that for any t, $\boldsymbol{P}[Z_t > 0] = \boldsymbol{Q}[Z_t > 0] = 1$, and for any $A \in \mathscr{F}_t$, one has that

$$\boldsymbol{Q}[A] = \boldsymbol{E}_{\boldsymbol{P}}[Z_t \mathbf{1}_A]$$

Proof. By the Normal Radon-Nikodym, there is some integrable random variable $Z_\infty > 0$ that is positive $\boldsymbol{P}$-almost surely for which

$$\boldsymbol{Q}[A] = \boldsymbol{E}_{\boldsymbol{P}}[Z_\infty \mathbf{1}_A]$$

for any $A \in \mathscr{F}$. In particular, if we choose an $A \in \mathscr{F}_t$, then by definition of conditional expectation, we can write

$$\boldsymbol{Q}[A] = \boldsymbol{E}_{\boldsymbol{P}}[\underbrace{\boldsymbol{E}[Z_\infty \mid \mathscr{F}_t]}_{=:Z_t} \mathbf{1}_A]$$

Now since $Z = (Z_t)_t$ is $\mathscr{L}^1$ closed, it is UI, and it is also clearly a Martingale. ♡

Some immediate consequences of the Filtered Radon-Nikodym are:

Remark 4.3.3 Consequences:

1. $Z_0 = 1$: Indeed, since $Z_t = \boldsymbol{E}[Z_\infty \mid \mathscr{F}_t]$, and $\mathscr{F}_0$ is assumed to be trivial (Recall this is a running assumption throughout the entire course), we have that $Z_0 = \boldsymbol{E}[Z_\infty]$ and by choosing $A = \Omega$ we can easily see that this is equal to one.

2. If ξ is both $\boldsymbol{Q}$ and $\boldsymbol{P}$ integrable, as well as $\mathscr{F}_t$, measurable, then for all $0 \leqslant s \leqslant t$, we have that

$$\boldsymbol{E}_{\boldsymbol{Q}}[\xi \mid \mathscr{F}_s] = \frac{\boldsymbol{E}_{\boldsymbol{P}}[Z_t \xi \mid \mathscr{F}_s]}{Z_s}$$

Indeed: let $A \in \mathscr{F}_s (\subseteq \mathscr{F}_t)$, then by Radon-Nikodym, we have the following change of measure for expectations

$$\begin{aligned}
\boldsymbol{E}_{\boldsymbol{Q}}[\xi \mathbf{1}_A] &= \boldsymbol{E}_{\boldsymbol{P}}[\xi \mathbf{1}_A Z_t] && (4.1)\\
&= \boldsymbol{E}_{\boldsymbol{P}}[\mathbf{1}_A \boldsymbol{E}_{\boldsymbol{P}}[\xi Z_t \mid \mathscr{F}_s]] && (4.2)\\
&= \boldsymbol{E}_{\boldsymbol{P}}\left[\mathbf{1}_A Z_s \boldsymbol{E}_{\boldsymbol{P}}\left[\frac{\xi Z_t}{Z_s}\middle| \mathscr{F}_s\right]\right] && (4.3)\\
&= \boldsymbol{E}_{\boldsymbol{Q}}\left[\mathbf{1}_A \boldsymbol{E}_{\boldsymbol{P}}\left[\frac{\xi Z_t}{Z_s}\middle| \mathscr{F}_s\right]\right] && (4.4)
\end{aligned}$$

where (4.1) is from the fact that $A \in \mathscr{F}_t$, so we can use the Filtered Radon-Nikodym, step (4.2) is by definition of conditional expectation, step (4.3) is by putting a Z_s in the denominator inside the inner conditional expectation as it is $\mathscr{F}_s$-measurable, and step (4.4) is by the Filtered Radon-Nikodym

With this out of the way, we can now present the goal of this section: the Cameron-Martin-Girsanov Theorem.

Theorem 4.3.4 (Cameron-Martin-Girsanov) Fix a probability measure $\boldsymbol{P}$ and let W be a Brownian Motion for $\boldsymbol{P}$ in $\mathbf{R}^n$, and let $(\alpha_s)_s$ be a previsible process in $\mathbf{R}^n$ for which $\int_0^\infty \|\alpha_s\| \, \mathrm{d}s < \infty$ $\boldsymbol{P}$-almost surely. Define

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

and assume that $(Z_t)_t$ is $\boldsymbol{P}$-Uniformly Integrable Martingale. Let $\boldsymbol{Q}$ be the measure defined by

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

Then the process $\widehat{W_t} = W_t - \int_0^t \alpha_s \, \mathrm{d}s$ is a $\boldsymbol{Q}$-Brownian Motion.

The proof relies on a few Lemmas:

Definition 4.3.5 (Stochastic Exponential) Let X be a continuous-semi-Martingale, then $\mathscr{E}(X)_t = \exp\left(X_t - \frac{1}{2}[X]_t\right)$ is called the Stochastic Exponential.

Lemma 4.3.6 Let $\boldsymbol{P}$ be a probability measure with respect to which Z is a Uniformly Integrable Martingale with that is strictly positive almost surely. Let X be a continuous $\boldsymbol{P}$-local Martingale, and define a measure $\boldsymbol{Q}$ as

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

Then the process $X - [X, \log Z]$ is a $\boldsymbol{Q}$-local Martingale.

Proof. We note by Itô's formula that

$$\mathrm{d}\log Z = \frac{\mathrm{d}Z}{Z} - \frac{\mathrm{d}[Z]}{2Z^2}$$

so plugging this into quadratic covariation, and using Kunita-Watanabe, we have that

$$\begin{aligned}[X, \log Z] &= \left[X, \int \frac{\mathrm{d}Z}{Z}\right] - \left[X, \int \frac{\mathrm{d}[Z]}{2Z^2}\right] \\ &= \int \frac{\mathrm{d}[X,Z]}{Z} - \int \frac{\mathrm{d}[X,[Z]]}{2Z^2}\end{aligned}$$

and so

$$\mathrm{d}[X, \log Z] = \frac{\mathrm{d}[X,Z]}{Z}$$

Now let $\widehat{X} = X - [X, \log Z]$.By localisation suppose that $\widehat{X}$ is bounded, we will now show that it is in fact a Martingale with respect to the measure $\boldsymbol{Q}$. The key is to note that $\widehat{X}Z$ is a $\boldsymbol{P}$-local Martingale, this is because

$$\begin{aligned}\mathrm{d}(\widehat{X}Z) &= X\,\mathrm{d}Z + Z\,\mathrm{d}X + \mathrm{d}[X,Z] \\ &\quad - Z\,\mathrm{d}[X, \log Z] - [X, \log Z]\,\mathrm{d}Z \\ &= Z\,\mathrm{d}X + \widehat{X}\,\mathrm{d}Z\end{aligned}$$

and so $\widehat{X}Z$ is the sum of integrals of continuous processes against local Martingales. But since Z is UI and $\widehat{X}$ is assumed to be bounded, then $\widehat{X}Z$ is also a UI $\boldsymbol{P}$-Martingale, so that we have

$$\begin{aligned}\boldsymbol{E}_{\boldsymbol{Q}}[\widehat{X}_\infty \mid \mathscr{F}_t] &= \frac{\boldsymbol{E}_{\boldsymbol{P}}[Z_\infty \widehat{X}_\infty \mid \mathscr{F}_t]}{\boldsymbol{E}_{\boldsymbol{P}}[Z_\infty \mid \mathscr{F}_t]} \\ &= \widehat{X}_t\end{aligned}$$

Where in the first equality we used Remark 4.3.3 and in the second equality we used the fact that $Z\widehat{X}$ is a UI $\boldsymbol{P}$-Martingale, and that so is Z. ♡

In particular, if we choose Z to be a stochastic exponential $\mathscr{E}(M)$, then we have that $X - [X, M]$ is a $\boldsymbol{Q}$-local Martingale. We are ready to prove Cameron-Martin-Girsanov:

Proof of Cameron-Martin-Girsanov. Recall that we are considering

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

and we are trying to show that $\widehat{W}_t = W_t - \int_0^t \alpha_s \, \mathrm{d}s$ is Brownian Motion under $\boldsymbol{Q}$. By the previous Theorem, we see that since

$$\begin{aligned}[W, \log Z] &= \left[W, \int \alpha \cdot \mathrm{d}W\right] - \frac{1}{2}\left[W, \int \|\alpha_s\|^2 \, \mathrm{d}s\right] \\ &= \int \alpha_s \, \mathrm{d}s\end{aligned}$$

then $\widehat{W}_t$ is indeed a $\boldsymbol{Q}$-local Martingale. Continuity is obvious, so to finish off, we just need to show that the quadratic variation of $\widehat{W}$ under $\boldsymbol{Q}$ is t. For this, we note that since $\boldsymbol{P}$ and $\boldsymbol{Q}$ are equivalent measures, UCP convergence under $\boldsymbol{Q}$ is equivalent to UCP convergence under $\boldsymbol{P}$ (this is due to a homework problem that says that if $\boldsymbol{P} \sim \boldsymbol{Q}$ then convergence in $\boldsymbol{P}$-probability is the same as convergence in $\boldsymbol{Q}$-probability), and so the quadratic variation of $\widehat{W}$ under $\boldsymbol{Q}$ is the same as the quadratic variation under $\boldsymbol{P}$. Finally, we note that under $\boldsymbol{P}$,

$$\begin{aligned}[\widehat{W}^i, \widehat{W}^j] &= \left[W^i - \int \alpha_s^i \, \mathrm{d}s, W^j - \int \alpha_s^j \, \mathrm{d}s\right] \\ &= [W^i, W^j] = t\delta_{i,j}\end{aligned}$$

Where the second equality comes from the fact that those integrals are of finite variation. ♡

Remark 4.3.7 (Novikov's Condition) Naturally, the big elephant in the room is that Girsanov's Theorem requires

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

to be a Uniformly Integrable Martingale, but how can one check this without doing loads of work? It turns out that there is a condition (Novikov's condition), whose proof we omit, that guarantees it: if M is a continuous local Martingale for which

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

then

$$Z = \mathscr{E}(M)$$

is a UI Martingale. Therefore, Girsanov's Theorem is true if we merely now require that

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

There is one last important result, which we will state but not prove:

Theorem 4.3.8 (Martingale Representation Theorem) Let $(\Omega, \mathscr{F}, \boldsymbol{P})$ be a probability space in which a d-dimensional Brownian Motion W is defined, and let $(\mathscr{F}_t)_t$ be the (completed and right-continuous) filtration generated by W. Assume that $\mathscr{F} = \sigma\left(\bigcup_t \mathscr{F}_t\right)$, i.e: the information on the probability space is that generated by the Brownian Motion. Let M be a cadlag locally square integrable local Martingale. Then there exists an Leb$\otimes \boldsymbol{P}$ almost sure unique predictable d-dimensional process $\alpha = (\alpha_t)_t$ such that $\int_0^t \|\alpha_s\|^2 \, \mathrm{d}s < \infty$ and

$$M_t = M_0 + \int_0^t \alpha \cdot \mathrm{d}W$$

Chapter 5

Stochastic Differential Equations

The goal of this chapter is to make sense of objects of the form

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t \tag{5.1}$$

Where $b : \mathbf{R}^n \to \mathbf{R}$, and $\sigma : \mathbf{R}^n \to \mathbf{R}^{n\times d}$ are given and W is a d-dimensional Brownian Motion.

5.1 An application of Girsanov's Theorem

As a warmup for our study of SDE's, we use Girsanov's Theorem to show the existence of what we will define later to be a weak solution.

Proposition 5.1.1 Let $b : \mathbf{R} \to \mathbf{R}$ and $\sigma > 0$ a constant. For every constant $T > 0$ and $x \in \mathbf{R}$, there exists some probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$ on which a Brownian Motion W and a continuous semi-Martingale $(X_t)_{t\in[0,T]}$ is defined satisfying

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

and initial condition $X_0 = x$.

Proof. Let $(\Omega, \mathscr{F}, \widehat{\boldsymbol{P}})$ on which there is some Standard Brownian Motion $\widehat{W}$. Set $X = \sigma\widehat{W} + x$. Define

$$Z = \mathscr{E}\left(\frac{1}{\sigma}\int b(X)\,\mathrm{d}\widehat{W}\right)$$

Since b is bounded, then so is b^2 and so

$$\boldsymbol{E}\left[\frac{1}{2\sigma^2}\int_0^t b(X_s)^2\,\mathrm{d}s\right] < \infty$$

for all $t \geqslant 0$. And so by Novikov's condition, the process Z is in fact a UI Martingale. Therefore for a fixed $T > 0$, we can define the measure $\boldsymbol{P}$ by

$$\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}\widehat{P}} = Z_T$$

and by Girsanov's Theorem we immediately get that

$$W = \widehat{W} - \frac{1}{\sigma}\int b(X)\,\mathrm{d}t$$

is a Brownian Motion under $\boldsymbol{Q}$. So rearranging, we have that

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

♡

5.2 Definitions of solution

Let us start by introducing what a solution to an equation like 5.1 should be. It is clear that we should have the following ingredients:

1. A probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$ with a complete, right-continuous filtration $(\mathscr{F}_t)_t$,
2. a d-dimensional Brownian Motion W adapted to $(\mathscr{F}_t)_t$,
3. an adapted process X, such that $b(X)$ and $\sigma(X)$ are predictable, and

   $$\int_0^t \|b(X_s)\|\,\mathrm{d}s < \infty \quad \text{and} \quad \int_0^t \|\sigma(X_s)\|^2\,\mathrm{d}s < \infty$$

   almost surely for all t. Here $\|\sigma\|^2 = \mathrm{trace}(\sigma\sigma^T)$ is the Frobenius matrix norm. Finally, we need

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

   for all $t \geqslant 0$.

Remark 5.2.1 (Matrix integral?) In the definition above, we are talking about this thing

$$\int_0^t \sigma(X_s)\,\mathrm{d}W_s$$

which is apparently an integral of an $n \times d$ matrix against a d-dimensional Brownian Motion. The way this is defined is intuitively as a vector, so that the equation above really reads

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

Note how the item 3. really is what we mean by equation 5.1, which has no a priori meaning. The integrability conditions mentioned in item 3. ensure that the $\mathrm{d}t$ integral is well-defined as a Lebesgue-Stieltjes integral, and that the $\mathrm{d}W$ integral is well-defined according to the theory of Stochastic integration. We will now introduce two notions of solution:

Definition 5.2.2 (Strong solution) A strong solution to the SDE 5.1 takes as the data the functions b and σ **as well as** the probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$ and the Brownian Motion W (the filtration is assumed to be generated by the Brownian Motion), and the output is the process X.

Remark 5.2.3 The key to understand what this is saying is the requirement that X is adapted, and the requirement that for a strong solution, the filtration is generated by the Brownian Motion. This means that in fact X is a functional of the sample paths of the Brownian Motion. That is to say, given the dynamics b and σ, and the realisation of the noise, one can reconstruct exactly the resulting process X. From a simulation perspective, the natural thing to consider is a strong solution. Suppose we were trying to simulate the SDE 5.1 in a computer. We may discretise time in intervals $0 = t_1 < t_2 < t_3 < \cdots$ and get a family $(Z_k)_{k\geqslant 0}$ of independent $\mathscr{N}(0,1)$ random variables, then effectively, we can think of

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

Now it is clear, that the evolution of the process X is clearly driven by the values of the normal random variables $(Z_k)_{k\geqslant 0}$, so that in effect, if we were given the seed that generated the random numbers, we could always reconstruct the exact same sample path for X.

In contrast to a strong solution, we have:

Definition 5.2.4 (Weak solution) A weak solution to the SDE 5.1 takes in as the data the functions b and σ and outputs a probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$, a filtration $(\mathscr{F}_t)_t$, a Brownian Motion W and the process X.

Remark 5.2.5 (On the difference between a strong and weak solution) Recall that for the strong solution, X is adapted to the filtration generated by a given Brownian Motion, this means that we must construct an X whose randomness is exactly that of the Brownian Motion. In a weak solution, one is free to construct both X_t and W_t together in a way that the differential equation is satisfied, both X and W may have randomness of their own, instead of the strong case, where the randomness of X is constrained to the externally given W.

Let us provide an example of an SDE that admits a weak solution, that is not a strong solution.

Example 5.2.6 (Tanaka's Example) Consider the following SDE, in the case $n = d = 1$:

$$\mathrm{d}X_t = g(X_t)\,\mathrm{d}W_t, \qquad X_0 = 0.$$

where

$$g(x) = \operatorname{sgn}(x) := \begin{cases} +1 & x \geqslant 1 \\ -1 & x < 1 \end{cases}$$

This SDE has a weak solution but no strong solution

Proof. Let us show the existence of a weak solution: let $(\Omega, \mathscr{F}, \boldsymbol{P})$ be a probability space on which a Brownian Motion X is defined, let $(\mathscr{F}_t)_t$ be any filtration for which X is adapted. Define a local Martingale W by

$$W_t = \int_0^t g(X_s)\,\mathrm{d}X_s$$

note that this is indeed well defined, because even if $g(x)$ is not left-continuous, the Brownian motion is previsible and the signum function is a pointwise limit of continuous functions, which preserves measurability. Of course since X is a local Martingale, it follows that W is also a local Martingale. Now note by the chain rule that

$$\begin{aligned} \int_0^t g(X_s)\,\mathrm{d}W_s &= \int_0^t g(X_s)\,\mathrm{d}\left(\int g(X)\,\mathrm{d}X\right)_s \\ &= \int_0^t g(X_s)^2\,\mathrm{d}X_s \\ &= \int_0^t \mathrm{d}X_s = X_t \end{aligned}$$

therefore the SDE is satisfied. Now we are just left with checking that W is a Brownian Motion, but since we have seen that it is a continuous local Martingale, all left to do is simply compute its quadratic variation, but this is rather simple by using say Kunita-Watanabe:

$$[W]_t = \int_0^t \mathrm{d}[X]_s = t$$

so we indeed have a weak solution. To show that it has no strong solution, we need the following fact (Tanaka's formula) if B is a Brownian Motion, then:

$$|B_t| = \int_0^t \operatorname{sgn}(B_s)\,\mathrm{d}B_s + \lim_{\epsilon\downarrow 0}\frac{1}{2\epsilon}\mathrm{Leb}(s\in[0,t]:|B_s|\leqslant\epsilon)$$

Now suppose that X is any process that satisfies the SDE, note that X must be a Brownian Motion because it is a local Martingale whose quadratic variation is clearly seen to be equal to t. Then we see that by the chain rule, as we have done before, one can recover the Brownian Motion W in the solution from X:

$$\begin{aligned} W_t &= \int_0^t \operatorname{sgn}(X_s)\,\mathrm{d}X_s \\ &= |X_t| - \lim_{\epsilon\downarrow 0}\frac{1}{2\epsilon}\mathrm{Leb}(s\in[0,t]:|X_s|\leqslant\epsilon) \end{aligned}$$

where the last step is Tanaka's formula with the fact that we've just shown X must be Brownian Motion. In particular, W_t is a function of $|X_s|$ up to time t, which means that W_t is $\sigma(|X_s|:s\in[0,t])$ measurable. But if X were to be a strong solution, we must have that X_t is $\sigma(W_s:s\in[0,t])$ measurable, but if this were true it means that X would be a function of $|X|$, which is nonsense. ♡

5.3 Notions of uniqueness

As usual, whenever we have a solution to a differential equation, the next step is to ask whether this solution is indeed unique. But in the case of stochastic differential equations, it turns out that just as there are two different notions of solutions, there are two different notions of uniqueness:

Definition 5.3.1 (Pathwise uniqueness) The SDE 5.1 has the pathwise uniqueness property if any two solutions X and X' both defined on the same filtered space and driven by the same Brownian Motion, which moreover have the same initial conditions almost surely, then one has that

$$\boldsymbol{P}[X_t = X'_t \text{ for all } t\geqslant 0] = 1.$$

The second notion of uniqueness is the following:

Definition 5.3.2 (Uniqueness in law) The SDE 5.1 has the uniqueness in law property if any two weak solutions $(\Omega,\mathscr{F},\boldsymbol{P},X,W)$, $(\Omega',\mathscr{F}',\boldsymbol{P}',X',W')$ such that $X_0\sim X'_0$, we have that the processes

$$(X_t)_t \sim (X'_t)_t$$

Remark 5.3.3 This last statement about identical distribution is a statement about laws on the space of continuous functions, after all, recall that $(X_t)_t$ can be seen as a measurable function $\Omega \to C(\mathbf{R}; \mathbf{R}^n)$. Therefore we are asking that the laws of the two processes on $C(\mathbf{R}; \mathbf{R}^n)$ agree.

We would now like to see an example of an SDE that has uniqueness in law but not pathwise uniqueness.

Example 5.3.4 (Tanaka's example continued) Tanaka's SDE:

$$\mathrm{d}X_t = \mathrm{sgn}(X_t)\,\mathrm{d}W_t$$

with $X_0 = 0$ has uniqueness in law but no pathwise uniqueness. (Recall that here sgn means the slightly unusual version of the signum function that we described in Example 5.2.6)

Proof. To check if an SDE has uniqueness in law, we have to check that for any two weak solutions, their laws agree (Here we make the distinction of weak solution because we only care about the law of the solution, not the pathwise properties which are determined by the probability measure on the space). But if X is any weak solution, as we saw before, by computing the quadratic variation and seeing $[X]_t = t$, we deduce that X is in fact a Brownian Motion, which means that all weak solutions have the Wiener measure as their law and so we have uniqueness in law. However, we are now going to show that one does not have pathwise uniqueness. Suppose that $(\Omega, \mathscr{F}, \boldsymbol{P}, X, W)$ is a weak solution. Then note that

$$\begin{aligned} \mathrm{d}(-X_t) &= -g(X_t)\,\mathrm{d}W_t \\ &= g(-X_t)\,\mathrm{d}W_t \end{aligned}$$

Where the first equality follows by Ito's Lemma. The second equality follows by inspection. It follows that the process $-X$ is also a weak solution, but since they are on the same probability space, we deduce that this SDE does not have pathwise uniqueness. ♡

It might seem natural to believe that the reason why this SDE fails to have pathwise uniqueness is the lack of smoothness of the function σ (Recall the general form of the SDE $\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$). Indeed, we now see that by fixing some smoothness condition on the function b and σ, we indeed are guaranteed to have pathwise uniqueness:

Theorem 5.3.5 (Pathwise uniqueness) The SDE 5.1 has pathwise uniqueness if the functions b and σ are locally Lipschitz, i.e: for all $N > 0$, there is some $K_N > 0$ such that whenever $\|x\|$ and $\|y\|$ are both no greater than N, we have that

$$\|b(x)-b(y)\| \leqslant K_N \|x-y\| \qquad \text{and} \qquad \|\sigma(x)-\sigma(y)\| \leqslant K_N \|x-y\|$$

(Note that the norm with the σ's is the matrix norm)

Remark 5.3.6 Note that this Lipschitz condition means that in particular the integrands are locally bounded so the Stochastic Integral Makes sense.

to prove this Theorem we will need the following lemma

Lemma 5.3.7 (Gronwall's Lemma) Suppose there are constants $a \in \mathbf{R}$ and $b > 0$ for which a locally integrable function f satisfies for all $t \geqslant 0$:

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

Then $f(t) \leqslant a\exp(bt)$ for all $t \geqslant 0$.

Proof of Theorem 5.3.5. We prove this for the case of dimension $n = d = 1$. Let X and X' be two solutions to the SDE 5.1 defined on the same probability space with $X_0 = X'_0$ almost surely. For a fixed $N > 0$, define $T_N = \inf\{t \geqslant 0 : |X_t| > N \text{ or } |X'_t| > N\}$. The goal is to show that for the function $f(t) = \boldsymbol{E}\left[\left\|X_{t\wedge T_N} - X'_{t\wedge T_N}\right\|^2\right]$, one has that $f(t) = 0$ for all $t \geqslant 0$. This will finish off the proof, because

$$\begin{aligned}
\boldsymbol{P}[X_t = X'_t \text{ for all } t] &= \boldsymbol{P}\left[\bigcap_{N=1}^{\infty}\left\{\sup_{t\leqslant T_N}\|X'_t - X_t\| = 0\right\}\right] \\
&= \lim_{N\to\infty}\boldsymbol{P}\left[\sup_{t\leqslant T_N}\|X'_t - X_t\| = 0\right]
\end{aligned}$$

and in turn, if we manage to show that $f(t) = 0$, we'll have that for each N, $\boldsymbol{E}\left[\left\|X_{t\wedge T_N} - X'_{t\wedge T_N}\right\|^2\right] = 0$ and so $\boldsymbol{P}\left[X_{t\wedge T_N} = X'_{t\wedge T_N}\right] = 1$ for all t. Now the key observation is that since X and X' are continuous (almost surely) we have that

$$\begin{aligned}
\boldsymbol{P}\left[\sup_{t\leqslant T_N}\|X_t - X'_t\| = 0\right] &= \boldsymbol{P}\left[\bigcap_{q\in\boldsymbol{Q}}\left\{\left\|X_{q\wedge T_N} - X'_{q\wedge T_N}\right\| = 0\right\}\right] \\
&\geqslant 1 - \sum_{q\in\boldsymbol{Q}}\underbrace{\boldsymbol{P}\left[X_{q\wedge T_N} \neq X'_{q\wedge T_N}\right]}_{=0}
\end{aligned}$$

So all left to do is to prove that $f(t)$ is in fact zero for all values of $t \geqslant 0$. For this we simply note that

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

and

$$X'_t = X'_0 + \int_0^t b(X'_s)\,\mathrm{d}s + \int_0^t \sigma(X'_s)\,\mathrm{d}W_s$$

then if we assume that $X_0 = X'_0$ and take the difference, we can apply the inequality $(a+b)^2 \leqslant 2a^2 + 2b^2$ and see that (Here I'm a bit careless writing $\|X - Y\|$ as $(X-Y)^2$ and viceversa, but its all good because we have assumed that we are working in dimensions one)

$$\begin{aligned}
\boldsymbol{E}\left[\left(X_{t\wedge T_N} - X'_{t\wedge T_N}\right)^2\right] &\leqslant 2\boldsymbol{E}\left[\left(\int_0^{t\wedge T_N} b(X_s) - b(X'_s)\,\mathrm{d}s\right)^2\right] + 2\boldsymbol{E}\left[\left(\int_0^{t\wedge T_N} \sigma(X_s) - \sigma(X'_s)\,\mathrm{d}W_s\right)^2\right] \\
&\leqslant 2\boldsymbol{E}\left[\int_0^{t\wedge T_N} (b(X_s) - b(X'_s)\,\mathrm{d}s)^2\right] + 2\boldsymbol{E}\left[\int_0^{t\wedge T_N} (\sigma(X_s) - \sigma(X - s'))^2\,\mathrm{d}s\right] \\
&\leqslant 4K_N^2 \int_0^t \boldsymbol{E}\left[\left\|X_{s\wedge T_N} - X'_{s\wedge T_N}\right\|^2\right]\mathrm{d}s
\end{aligned}$$

Where in the first step we used $(a+b)^2 \leqslant 2(a^2+b^2)$, in step two, we used Jensen's inequality for the first integral and Ito's isometry for the second integral, and in the third step, we used the Lipschitz condition as well as Fubini's Theorem. Now from Gronwall's Lemma, it follows that in fact $f(t) = 0$. ♡

To me it seems natural that pathwise uniqueness is a stronger condition than uniqueness in law, but proving it is not a trivial result, as we have to deal with solutions being defined in different probability spaces. This result is indeed correct, but we ommit its proof:

Theorem 5.3.8 (Yamada-Watanabe) If an SDE has the pathwise uniqueness property, then it has uniqueness in law.

5.4 Existence of strong solutions

In this section we provide a result that states that under certain smoothness conditions, one is guaranteed to have a solution to the SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t \tag{5.2}$$

Theorem 5.4.1 (Existence of strong solutions) Suppose the functions b and σ are globally Lipschitz, i.e, there is some $K > 0$ such that for all $x, y \in \mathbf{R}^n$:

$$\|b(x) - b(y)\| \leqslant K\|x - y\| \qquad \|\sigma(x) - \sigma(y)\| \leqslant K\|x - y\|$$

(Mind the blatant abuse of norm notation) Then there is a unique strong solution to the SDE 5.2.

Main idea: The main idea is to show that for some deterministic value of time T, one can show the existence of a strong solution on $[0, T]$ for any starting value X_0, and then extend by "gluing solutions" together.

Proof. Suppose we can prove the existence of a strong solution $X^{(1)}$ to the SDE 5.2 for any initial condition X_0 on an interval $[0, T]$ for some $T > 0$. Recall that the key difference between weak and strong existence is that for $X^{(1)}$ to be a strong solution, it must be indeed a function of the Brownian motion W defined on the underlying space, as well as of the initial condition X_0. This means there is a function φ such that $X_t^{(1)} := \varphi(W_t, X_0)$ satisfies

$$X_t^{(1)} = X_0 + \int_0^t b\left(X_s^{(1)}\right) \mathrm{d}s + \int_0^t \sigma\left(X_s^{(1)}\right) \mathrm{d}W_s$$

for any $t \in [0, T]$. We could then use this same function φ with initial condition $X_T^{(1)}$ and the Brownian Motion $(W_{u+T} - W_T)_{u \geqslant 0}$ to construct a process $X^{(2)}$ such that for $t \in [0, T]$

$$X_t^{(2)} = X_T^{(1)} + \int_0^t b\left(X_s^{(2)}\right) \mathrm{d}s + \int_0^t \sigma\left(X_s^{(2)}\right) \mathrm{d}(W_{s+T} - W_T)$$

And we can glue both processes together to form the following

$$X_t = X_t^{(1)} \mathbf{1}\{0 \leqslant t \leqslant T\} + X_{t-T}^{(2)} \mathbf{1}\{T < t \leqslant 2T\}$$

By construction, this process solves the SDE 5.2 in $t \in [0, T]$, let us check that it also solves it in $t \in [T, 2T]$:

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

Thus showing that the SDE is also satisfied (why is this change of variables justified?). By iterating this procedure, we could construct a solution on $[0, \infty)$. The uniqueness of the solution would follow as global Lipschitz property implies the local Lipschitz property needed for the Uniqueness Theorem to hold. We now prove that there exists a solution on some interval $[0,T]$. The key is to consider the following map F: for an adapted continuous process Y and an initial condition X_0, define

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

The key of the proof is to find a Banach space $\mathscr{B}$ of adapted processes where this map F is actually a contraction with respect to its norm. First we define the norm

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

where T is to be determined later. Consider $\mathscr{B}$ to be the vector space of continuous adapted processes Y with $|||Y||| < \infty$. It can be checked that this space is complete in a similar manner that it was shown earlier in these notes that $\mathscr{M}_2$ is complete. Assume further (it can be shown that this assumption is without loss by modifying the norm), that X_0 is square-integrable. Then we can now show that F is a contraction. Using the general fact that $(a+b)^2 \leqslant 2(a^2+b^2)$, one has that

$$|||F(X)-F(Y)||| \leqslant 2\boldsymbol{E}\left[\sup_{t\in[0,T]}\left\|\int_0^t b(X_s)-b(Y_s)\,\mathrm{d}s\right\|^2\right] + 2\boldsymbol{E}\left[\sup_{t\in[0,T]}\left\|\int_0^t \sigma(X_s)-\sigma(Y_s)\,\mathrm{d}W_s\right\|^2\right] \quad (5.3)$$

We now use Jensen's inequality combined with our Lipschitz assumption to see that

$$\begin{aligned}\boldsymbol{E}\left[\sup_{t\in[0,T]}\left\|\int_0^t b(X_s)-b(Y_s)\,\mathrm{d}s\right\|^2\right] &\leqslant \boldsymbol{E}\left[\left(\sup_{t\in[0,T]}\int_0^t \|b(X_s)-b(Y_s)\|\,\mathrm{d}s\right)^2\right] \\ &\leqslant K^2T^2\boldsymbol{E}\left[\sup_{t\in[0,t]}\|X_t-Y_s\|^2\right]\end{aligned}$$

♡

We can obtain a similar bound for the second term in 5.3 by using Burkholder's inequality

$$\boldsymbol{E}\left[\sup_{t\in[0,T]}|M_t|^2\right] \leqslant C\,\boldsymbol{E}\,[[M]_T]$$

as well as the Lipschitz assumption to say that

$$\boldsymbol{E}\left[\sup_{t\in[0,T]}\left\|\int_0^t \sigma(X_s)-\sigma(Y_s)\,\mathrm{d}W_s\right\|^2\right] \leqslant CK^2T\boldsymbol{E}\left[\sup_{t\in[0,T]}\|X_t-Y_t\|^2\right]$$

and so combining all this, we have that

$$|||F(X)-F(Y)||| \leqslant (2T+2C)K^2T|||X-Y|||^2$$

hence we can choose T small enough so that this coefficient, call it c, is less than one and F is indeed a contraction. By the Fixed Point Theorem, we have a unique solution to SDE 5.2 in the time interval $[0,T]$, provided that we can check that F actually maps $\mathscr{B}$ to $\mathscr{B}$. This is straightforward, because by a simple application of the triangle inequality, we have that

$$|||F(X)||| \leqslant |||F(X)-F(0)||| + |||F(0)|||$$

the first term is at most $c|||X|||$ by the contraction property, and since

$$F(0)_t = X_0 + t\,b(0) + \sigma(0)W_t$$

and the norm $|||\cdot|||$ is an integral over a finite time horizon, it is easy to see that $F(0)$ has finite norm, and so $F(X)$ has also finite norm and is a bunch of integrals so it is indeed an adapted continuous process that has finite norm so it belongs to $\mathscr{B}$.

5.5 Relation to PDE's and the Feynman-Kac formula

In this section we see an interesting application of SDE Theory that connects partial differential equations to stochastic differential equations. We already know from Advanced Probability that the heat equation $\Delta u = 0$ in a "good enough" domain with $\mathscr{C}^2$ boundary conditions is intimately related to the average paths of Brownian motion on that domain. This formula can be seen as a more general version of this statement. The connection is given by the following Theorem:

Theorem 5.5.1 (Feynman-Kac formula) Let $v : \mathbf{R}_+ \times \mathbf{R}^n \to \mathbf{R}$ be $\mathscr{C}^{1,2}(\mathbf{R}_+; \mathbf{R}^n)$ and satisfy the PDE

$$\frac{\partial v}{\partial t} = \sum_{i=1}^{n} b^i \frac{\partial v}{\partial x^i} + \frac{1}{2}\sum_{i,j} a^{ij} \frac{\partial^2 v}{\partial x^i \partial x^j} - c\, v \tag{5.4}$$

where $a^{ij} = \sum_{k=1}^{d} \sigma^{ik}\sigma^{jk}$, i.e: $a^{ij} = (\sigma\sigma^T)_{i,j}$ and initial conditions $v(0,x) = f(x)$, for some given functions $c : \mathbf{R}^n \to \mathbf{R}$ and $f : \mathbf{R}^n \to \mathbf{R}$. Then for every finite time horizon T, one has that the process $(M_t)_{t\in[0,T]}$ defined by

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

is a local Martingale, where X_t is the process defined by the SDE $\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$. Furthermore, if v is bounded and c is bounded from below, then

$$v(T, X_0) = \boldsymbol{E}\left[\exp\left\{-\int_0^T c(X_s)\,\mathrm{d}s\right\} f(X_T)\middle| X_0\right].$$

Proof. The argument boils down to using Ito's formula to show that M_t is an integral of a continuous process against a local Martingale. For this we perform first integration by parts:

$$\begin{aligned}\mathrm{d}M_t = {} & \mathrm{d}\left(\exp\left\{-\int_0^t c(X_s)\,\mathrm{d}s\right\}\right) v(T-t,X_t) + \exp\left\{-\int_0^t c(X_s)\,\mathrm{d}s\right\} \mathrm{d}(v(T-t,X_t)) \\ & + \mathrm{d}\left[\exp\left\{-\int_0^t c(X_s)\,\mathrm{d}s\right\}, v(T-t,X_t)\right]\end{aligned}$$

Next, we have that

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

(Notice that we leave out the term with quadratic variation, because the quadratic variation of the integral inside the exponential is of finite variation). On the other hand:

$$\mathrm{d}(v(T-t,X_t)) = \sum_{i=1}^{n} \frac{\partial v}{\partial x^i}\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial x^i \partial x^j}\,\mathrm{d}[X^i, X^j]_t - \frac{\partial v}{\partial t}\,\mathrm{d}t$$

and one now sees, by using the fact that $\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$ that

$$
\begin{aligned}
\left[X^i, X^j\right]_t &= \left[\sum_{k=1}^{d}\int \sigma^{ik}(X_s)\,\mathrm{d}W_s^k, \sum_{l=1}^{d}\int \sigma^{jl}(X_s)\,\mathrm{d}W_s^l\right]_t \\
&= \sum_{k,l}\left[\int \sigma^{ik}(X_s)\,\mathrm{d}W_s^k, \int \sigma^{jl}(X_s)\,\mathrm{d}W_s^l\right]_t \\
&= \sum_{k=1}^{d}(\sigma^{ik}\sigma^{jk})(X_s)t \\
&= a_{ij}t
\end{aligned}
$$

where in the third equality we used Kunita-Watanabe, as well as the fact that W is a Brownian Motion (Levy). Putting all of this nonsense together:

$$
\begin{aligned}
\mathrm{d}M_t = \exp&\left\{-\int_0^t c(X_s)\,\mathrm{d}s\right\}\left(\sum_{i=1}^{n}\frac{\partial v}{\partial x_i}b^i + \frac{1}{2}\sum_{i,j}a_{ij}\frac{\partial^2}{\partial x^i \partial x^j} - \frac{\partial v}{\partial t} - c\,v\right) \\
&+ \exp\left\{-\int_0^t c(X_s)\,\mathrm{d}s\right\}\sum_{ik}\sigma^{ik}\frac{\partial v}{\partial x^i}\,\mathrm{d}W^k
\end{aligned}
$$

We observe that the first term of this sum incorporates precisely the PDE 5.4, and so we are left just with the second term. In other words, M_t really is a local Martingale. Now we use the hypothesis that c is bounded from below, meaning that the exponential is bounded, and then using that v is bounded, we also get that the derivative term is bounded, and so we get that M is bounded. Indeed remember that M is a Martingale over a finite time horizon, and

$$
M_t = \exp\left\{\int_0^t c(X_s)\,\mathrm{d}s\right\}\int_0^t \sum_{i,k}\sigma^{ik}(X_s)\frac{\partial v}{\partial x_i}(T-t, X_s)\,\mathrm{d}W_s^k.
$$

moreover we see that if $f \leqslant c$ is a bounded previsible function and say W is a Brownian Motion, then the martingale Y defined by $Y_t = \int_0^t f(X_s)\,\mathrm{d}W_s$ is indeed bounded on a fixed time Horizon, for

$$
[Y]_t = \int_0^t f^2(X_s)\,\mathrm{d}s \leqslant c^2 t
$$

and so if the quadratic variation is at most $c^2 t$, then $|Y_t|$ cannot be greater than $c^2 t$, and so for a finite time horizon we are bounded . Now since $(M_t)_{t\in[0,T]}$ is a bounded local Martingale, it is indeed a true Martingale (this was an easy application of DCT after choosing a localising sequence). Now by looking again at the definition of M_t, and plugging in $t = 0$, we see that

$$
\begin{aligned}
v(T, X_0) &= M_0 \\
&= \boldsymbol{E}[M_T \mid \mathscr{F}_0] \\
&= \boldsymbol{E}\left[\exp\left\{-\int_0^T c(X_s)\,\mathrm{d}s\right\} f(X_T) \middle| \mathscr{F}_0\right]
\end{aligned}
$$

now we can use Tower law on both sides with respect to the sigma algebra $\sigma(X_0)$ and we are done. ♡

Remark 5.5.2 (Waffle) Let us take a minute to appreciate this result by putting into words what the formula is saying (in one dimension for simplicity). Suppose that you have some physical system with initial conditions $v(0, x) = f(x)$, and it evolves according to the PDE

$$
\frac{\partial v}{\partial t}(x, t) = b(x, t)\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2(x, t)\frac{\partial^2}{\partial x^2} - V(x)v(x, t)
$$

if we imagine $v(x, t)$ to indicate, say the amount of "heat" at position x and time t, this is saying that the "heat bump" (this is described by the initial condition) moves with a drift given by $b(x, t)$, and diffuses with a diffusivity given by σ. Moreover, there is a potential V that directly adds or removes heat at some points. The Feynman-Kac equation tells us that one can understand the motion of the heat blob as follows: suppose for simplicity that there is no potential, then Feynman-Kac says that to understand the heat $u(x, T)$ at time T and position x, one can "launch" many particles that randomly move according to the appropriate SDE $\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_s)\,\mathrm{d}W_t$, and see what at what temperature of the "initial heat blob" they are at time T, then one takes this average. If we included the potentials, this says that now we must moreover discount the effect that the potential has had on the path of the particle.

Chapter 6

Applications to Finance

6.1 Introduction

In this last chapter, we consider how the theory of stochastic integration and differential equations can be used in application to financial Mathematics. In this section we will consider a Market with $1+d$ assets, where the first asset is a money market account, i.e: an account that accrues interest, and d "risky assets".

Definition 6.1.1 (Market) A market with $1+d$ assets is a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}\}_t, \boldsymbol{P})$ with:

1. $\{\mathscr{F}_t\}_{t\geqslant 0}$ satisfying the usual conditions.
2. A *risk-free* asset, whose price is given by a semi-Martingale $(B_t)_{t\geqslant 0}$ with dynamics

$$\mathrm{d}B_t = B_t r_t \,\mathrm{d}t$$

   where $(r_t)_{t\geqslant 0}$ is the *interest-rate*, a previsible, locally $\mathrm{d}t$-integrable process.
3. d risky assets, with prices given by the d-dimensional continuous semi-Martingale $(S_t^1, \cdots, S_t^d)_{t\geqslant 0}$.

Remark 6.1.2 (Assumptions) We will make some simplifying assumptions about our Market to the theory:

1. There are no transaction costs.
2. There is no bid-ask spread.
3. There is no price impact.

Remark 6.1.3 (Risk-free asset) The risk free asset should be interpreted as money that one deposits in an interest-bearing bank account. The dynamics are easy to solve for, simply:

$$B_t = B_0 \exp\left\{\int_0^t r_s \,\mathrm{d}s\right\}, B_0 > 0.$$

We now introduce a trader that holds:

- $\varphi = (\varphi_t)_{t\geqslant 0}$ shares of the money market asset, and $\theta = \left(\theta_t^i\right)_{t\geqslant 0, 1\leqslant i\leqslant d}$ shares of asset i at time t. These are assumed to be previsible processes, such that ϕ and θ are B-integrable and S-integrable respectively.

We can think of these processes as being under the control of the trader, (hence the previsibility). One of the key requirements of the theory is that we do not want to allow for cases where the trader consumes from their wealth on the market, nor deposits more wealth. This is the condition of self-financing.

Definition 6.1.4 (Self-financing) A $1+d$ dimensional process (φ, θ) is called a self-financing trading strategy if it is precisible, (B, S)-integrable, and

$$\mathrm{d}(\varphi_t B_t + \theta_t \cdot S_t) = \varphi_t \,\mathrm{d}B_t + \theta_t \cdot \mathrm{d}S_t$$

Remark 6.1.5 (On the definition of self-financing portfolio) Naturally, the value V_t of our portfolio at time t is given by $V_t = \varphi_t B_t + \theta_t \cdot S_t$, no much controversy in that. If we wanted to see the infinitesimal change in value of our portfolio, we could use Ito's Lemma and obtain some extra terms in the differential of V_t, namely $B_t \,\mathrm{d}(\varphi_t) + S_t \cdot \mathrm{d}(\theta_t)$ plus quadratic covariation terms, which usually in the usual models we will consider. Therefore, the extra terms we get, correspond to the infinitesimal change of stock we hold, this would correspond to selling stock to consume from our wealth or injecting wealth by means of adding more stock, the self-financing condition simply forbids that from happening. Informally, the change of wealth comes solely from the change of price of the money market and the change of price of the risky assets.

6.2 Construction of a self-financing strategy

As it turns out, if we are given an initial wealth x and a trading strategy θ, we can construct the process φ, i.e: we can find how much money we need to put into the money-market for the strategy to be self-financing. Indeed, we start noting that by definition of the value of the portfolio, $B_t \varphi_t = V_t - \theta_t \cdot S_t$. Therefore the self-financing equation is equivalent to:

$$\mathrm{d}V_t = r_t (V_t - \theta_t \cdot S_t)\,\mathrm{d}t + \theta_t \cdot \mathrm{d}S_t.$$

The goal will be to use this equation to constrain a value of V_t that makes the portfolio self-financing, and then simply use the expression of φ in terms of B, V, θ. Once again, we can rewrite the self-financing equation equivalently as

$$\begin{aligned}
\mathrm{d}\left(\exp\left\{-\int_0^t r_s\,\mathrm{d}s\right\} V_t\right) &= -r\exp\left\{-\int_0^t r_s\,\mathrm{d}s\right\} V_t\,\mathrm{d}t + \exp\left\{-\int_0^t r_s\,\mathrm{d}s\right\}\left(r_t\left(X_t - \theta_t\cdot S_t\right)\mathrm{d}t + \theta_t\cdot\mathrm{d}S_t\right)\\
&= \theta_t\cdot\left(\exp\left\{-\int_0^t r_u\,\mathrm{d}u\right\}\mathrm{d}S_t - r_t\exp\left\{\int_0^t r_u\,\mathrm{d}u\right\} S\,\mathrm{d}t\right)\\
&= \theta_t\cdot\mathrm{d}\left(\exp\left\{-\int_0^t r_u\,\mathrm{d}u\right\} S_t\right)
\end{aligned}$$

where the first step is just an application of Ito's formula and the fact that $\exp\left\{-\int_0^t r_u\,\mathrm{d}u\right\}$ is of finite variation, (so that the quadratic covariation term dies), and then some minor manipulation on the second step, and finally, integration by parts formula with again the finite variation observation.

Remark 6.2.1 In conclusion, this computation tells us that the *discounted* wealth is equal to a stochastic integral of θ_t against the discounted asset prices.

Remark 6.2.2 (Financial jargon: discounting) If V_t is the value of our portfolio at time t, sometimes we wish to talk about the discounted value, $\exp\left\{-\int_0^t r_u\,\mathrm{d}u\right\} V_t$. Discounting is a way of measuring "how good the returns of the portfolio actually are", by adjusting for the fact that putting money in the risk-free asset would have generated some growth either way. It we look at the raw value V_t, it might seem that the portfolio is making money, but it could be partially due to overall growth in the economy via the interest rate, thus discounting eliminates the effect of the risk-free growth.

With the above computation in mind, we can redefine the value of our portfolio V_t, by removing from our controls the amount of money we have in the money market, and having our initial wealth x and our trading strategy θ as the only controls to our wealth. In summary,

Proposition 6.2.3 A portfolio with initial wealth x and trading strategy θ is self-financing if and only if its value V_t, can be expressed as $V_t = V_t^{x,\theta}$, where

$$\begin{aligned}
V_t^{x,\theta} &= \exp\left\{\int_0^t r_u\,\mathrm{d}u\right\}\left(x + \int_0^t \theta_s\cdot\mathrm{d}\left(\exp\left\{-\int_0^t r_u\,\mathrm{d}u\right\} S_s\right)\right)\\
&= B_t\left(\frac{x}{B_0} + \int_0^t \theta\cdot\mathrm{d}\left(\frac{S}{B}\right)\right)
\end{aligned}$$

Remark 6.2.4 As an immediate consequence, if we choose $V_t^{x,\theta}$ as above, setting

$$\varphi_t = \frac{X_t^{x,\theta} - \theta_t \cdot S_t}{B_t}$$

gives that (φ, θ) is a self-financing trading strategy.

6.3 Arbitrage

Now we approach the following question: given a utility function $u : \mathbf{R} \to \mathbf{R}$, and $x \in \mathbf{R}$, can we find

$$\theta^* = \text{argmax}_\theta \boldsymbol{E}[u(V_T^{x,\theta})] \tag{6.1}$$

where T is some fixed time horizon? Some usual assumptions that we make on our utility function is that it is increasing (you prefer more to less), and concave (you are risk averse). We will not actually try to solve this question, but rather use it as motivation to talk about the importance of one of the main assumptions of our theory, namely that there is indeed no free lunch. An example of such "free-lunch" strategies would be a d-dimensional previsible process π such that for a fixed time horizon T, $\boldsymbol{P}[V_T^{0,\pi} \geqslant 0] = 1$ and $\boldsymbol{P}[V_T^{0,\pi} > 0] > 0$. The key objection to allowing these kinds of strategies in our market is that θ^* in 6.1 does not exist. The key fact is that $V_t^{x,\theta}$ is sort of linear in θ. In particular,

$$V_t^{x,\theta_1+\theta_2} = V_t^{x,\theta_1} + V_t^{0,\theta_2}$$

this comes from the shape we obtained for $V_t^{x,\theta}$ in Proposition 6.2.3, so that if θ^* existed and π is an arbitrage, then

$$\boldsymbol{E}\left[u\left(V_T^{x,\theta^*+\pi}\right)\right] > \boldsymbol{E}\left[u\left(V_T^{x,\theta^*}\right)\right]$$

since $V_t^{0,\pi} \geqslant 0$ with probability one and strictly greater than zero on some set of non-zero probability. The conditions on π we gave above are related to the notion of arbitrage, and we want to rule these kind of strategies out. However, let us illustrate an example of why we actually need some more conditions on what we refer to as an arbitrage.

Example 6.3.1 (A "legit" example of free-lunch) Consider a market where the interest rate r_s is identically zero and there is one risky asset, namely a Brownian motion. Then there exists a strategy π, with in fact $\boldsymbol{P}[V_T^{0,\pi} > 0] = 1$.

Proof. This example will be so pathological that we can in fact make our gains arbitrarily big. Let $K \geqslant 0$ (this will be our gains by time T) and let $f : [0, T] \to [0, \infty]$ be a differentiable, strictly increasing function such that $f(0) = 0$ and $f(T) = +\infty$. Then define

$$Z_t = \int_0^{f^{-1}(t)} \sqrt{f'(s)}\,\mathrm{d}W_s$$

observe that

$$[Z]_t = \int_0^{f^{-1}(t)} f'(s)\,\mathrm{d}s = t$$

and so in fact Z is a Brownian motion. From this we can define

$$\tau = \inf\{t > 0 : Z_t = K\} < \infty \text{ a.s}$$

and now the idea is to let $\sigma = f^{-1}(\tau)$ and define $\pi_t = \mathbf{1}\{t \leqslant \sigma\}\sqrt{f^{-1}(t)}$. Then we observe that indeed this strategy is W-integrable, for

$$\int_0^t \pi_s^2\,\mathrm{d}s < \infty$$

and moreover,

$$V_T = \int_0^T \pi_s\,\mathrm{d}W_s = \int_0^\sigma \sqrt{f^{-1}(s)}\,\mathrm{d}W_s = Z_\tau = K \text{ a.s}$$

♡

Remark 6.3.2 (What happened here?) Morally speaking, what happened here is that we chose our strategy π in a way that our portfolio value was a Brownian motion sped up to hit K by time T. The problem in doing so, is that we are in fact allowing for times $t \leqslant T$ where our portfolio has arbitrarily large losses, but since we know that Brownian motion will eventually hit K with probability one, we just wait the losses out. This will be the final detail that we want to rule out. In real life, this kind of strategies would not be admissible, as we would be arbitrarily in the red for some time before reaching our target payout, and unless you have an infinite bankroll, you probably won't be able to execute this.

In light of this, we introduce the following definition:

Definition 6.3.3 (Admissibility) A trading strategy θ is a-admissible if

$$\boldsymbol{P}[V_t^{a,\theta} \geqslant 0] = 1.$$

and now we are able to introduce the definition of arbitrage:

Definition 6.3.4 (Arbitrage) An arbitrage is a 0-admissible trading strategy π, such that there exists some (non-random) time horizon T for which $\boldsymbol{P}[V_T^{0,\pi} \geqslant 0] = 1$ and $\boldsymbol{P}[V_T^{0,\pi} > 0] > 0$.

6.3.1 Constructing an arbitrage

Suppose that we have a $d+1$ dimensional Market with prices:

$$\begin{cases} \mathrm{d}B_t = r B_t \,\mathrm{d}t \\ \mathrm{d}S_t^{(i)} = S_t^{(i)} \left(\mu^{(i)} \,\mathrm{d}t + \sum_{k=1}^n \sigma^{ik} \,\mathrm{d}W_t^{(k)} \right) \end{cases}$$

Where $r, \boldsymbol{\mu}, \boldsymbol{\sigma}$ are sufficiently integrable previsible processes and the W^k's are Brownian Motions. Under what conditions can we hope to find an arbitrage in the Market? Note that by Ito's Formula:

$$\mathrm{d}(S^{(i)}/B) = (S^{(i)}/B) \left((\mu^{(i)} - r) \,\mathrm{d}t + \sum_{k=1}^n \sigma^{ik} \,\mathrm{d}W_t^{(k)} \right),$$

and recall that if $\phi = \left(\phi^{(1)}, \cdots, \phi^{(d)}\right)$ is our trading strategy, then the value of our portfolio, $X_t^{0,\phi}$, supposing for simplicity that $B_0 = 1$, is given by

$$\begin{aligned} \frac{X_t^{0,\phi}}{B_t} &= \int_0^t \phi \cdot \mathrm{d}\left(\frac{S}{B}\right) & (6.2) \\ &= \sum_{i=1}^d \int_0^t \phi_s^{(i)} \frac{S_s^{(i)}}{B_s} (\mu^{(i)} - r) \,\mathrm{d}t + \sum_{k=1}^n \int_0^t \sum_{i=1}^d \phi_s^{(i)} \frac{S_s^{(i)}}{B_s} \sigma_s^{ik} \,\mathrm{d}W_s^{(k)} & (6.3) \\ &= \int_0^t \boldsymbol{\psi} \cdot (\boldsymbol{\mu} - r\mathbf{1}) \,\mathrm{d}t + \int_0^t \boldsymbol{\sigma}^T \boldsymbol{\psi} \cdot \mathrm{d}W_s & (6.4) \end{aligned}$$

the last equation is just putting it into matrix form, and we have set $\phi_s^{(i)} = \frac{B_s}{S_s^{(i)}} \psi_s^{(i)}$ for process $\boldsymbol{\psi}$. This means that if we can manage to find a process $\boldsymbol{\psi}$ such that

$$\boldsymbol{\sigma}^T \boldsymbol{\psi} = 0 \qquad \text{and} \qquad \boldsymbol{\psi} \cdot (\boldsymbol{\mu} - r\mathbf{1}) > 0, \tag{6.5}$$

then we have found arbitrage.

Example 6.3.5 Suppose a three asset market has the dynamics

$$\begin{cases} \mathrm{d}B_t = 2B_t \,\mathrm{d}t, \\ \mathrm{d}S_t^{(1)} = 3 \,\mathrm{d}t + \mathrm{d}W_t^{(1)} - 2 \,\mathrm{d}W_t^{(2)} \\ \mathrm{d}S_t^{(2)} = 5 \,\mathrm{d}t - 2 \,\mathrm{d}W_t^{(1)} + 4 \,\mathrm{d}W_t^{(2)} \end{cases}$$

then we have that

$$\sigma = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \qquad \mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

thus in light of the previous discussion, we see that $\psi = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ satisfies 6.5, and so by trading the strategy

$$\phi_s = \begin{pmatrix} \frac{2B_s}{S_s^{(1)}} & \frac{B_s}{S_s^{(2)}} \end{pmatrix}$$

we have indeed a free lunch.

6.4 The Fundamental Theorem of Asset Pricing

Our next step is to study some mathematical condition that rules out the existence of arbitrage. For this we introduce the following definition

Definition 6.4.1 (Equivalent Local Martingale Measure) Let $\boldsymbol{P}$ be the "real world" probability measure, i.e: the probability measure of our market. We say that $\boldsymbol{Q}$ is an equivalent local Martingale measure (ELMM), if:

1. $\boldsymbol{Q}$ is equivalent to $\boldsymbol{P}$,
2. the discounted asset prices S^i/B form a $\boldsymbol{Q}$-local Martingale.

We will now see that the existence of an ELMM rules out any arbitrage opportunities. An informal interpretation of this is that if under $\boldsymbol{Q}$ the discounted asset prices form a Martingale, then they are a "fair game", and so it is impossible to make money out of them under the probability measure $\boldsymbol{Q}$. But since equivalence of measures preserves the "impossibility of statements", it will turn out that arbitrages will also not be possible under the real world probability measure. Let us now introduce (a version of)

Theorem 6.4.2 (The Fundamental Theorem of Asset Pricing) Suppose that in our market there exists an ELMM. Then the market is arbitrage-free.

Proof. Let θ be a 0-admissible trading strategy. Then under this strategy, we have that our wealth equation reads

$$V_t^{0,\theta} = B_t \int_0^t \theta_u \cdot \mathrm{d}\left(\frac{S_u}{B_u}\right).$$

Then if $\boldsymbol{Q}$ is an ELMM, we have that $\frac{V_t^{0,\theta}}{B_t}$ is a $\boldsymbol{Q}$-local Martingale. However, since θ is admissible, we have that $V_t^{0,\theta} \geqslant 0$ $\boldsymbol{P}$, and hence $\boldsymbol{Q}$-almost surely, and since non-negative local Martingales are super Martingales (Homework sheet 1), we have that $V_t^{0,\theta}$ is a $\boldsymbol{Q}$ super Martingale. Therefore,

$$\boldsymbol{E}_{\boldsymbol{Q}}[V_t^{0,\theta}/B_t] \leqslant V_0^{0,\theta}/B_0 = 0.$$

Therefore we have that $V_t^{0,\theta} = 0$ $\boldsymbol{Q}$-almost surely, and so also $\boldsymbol{P}$-almost surely. Hence no arbitrage is possible. ♡

6.5 Contingent Claim Pricing

In this last section we will discuss the topic of contingent claims and their pricing. The setup is as follows, we have a market with $d+1$ assets, (B,S), and we want to market a new asset whose payout ξ at some maturity, depends (is contingent) on the price of the underlying assets.

Example 6.5.1 (European call option) The canonical example of a contingent claim is the European call option. Suppose we are in a market with only one asset, namely the risky asset S. A European call option is the right but not the obligation to buy a stock of the asset for a predetermined price K at a future date T.

The big question is: how much should I charge someone to sell them this option? In order to tackle this question, we first need to observe that the payout of our European option is precisely $(K - S_T)^+$. Let us briefly give a definition to clarify what we mean by European:

Definition 6.5.2 A European claim is specified by an expiration date $T > 0$ and an $\mathscr{F}_T$ random variable ξ, that specifies the payout of the claim at expiration. A European call is called vanilla if it is a function of the underlying assets, i.e, if

$$\xi = f(S_T^1, \cdots, S_T^d)$$

We would now like to augment our market to include a new asset, whose price $(\xi_t)_{t\geqslant 0}$ represents the price of our European claim at time t.

Remark 6.5.3 If our original market admits an ELMM $\boldsymbol{Q}$, there is a natural way to price this asset, we simply want ξ/B_t to be a $\boldsymbol{Q}$-local Martingale, so we can set

$$\xi_t = B_t \boldsymbol{E}_{\boldsymbol{Q}}\left[\frac{\xi_T}{B_T}\middle|\mathscr{F}_t\right],$$

and now it follows that this augmented market has no arbitrage either. This is actually the “correct" way to price the option: even though we will not provide a formal proof, intuitively, this price will be the one required replicate the payout of the option using a portfolio of the stock and the money market. Indeed, suppose that the cost of the replicating portfolio was any lower, then I could buy the portfolio, sell the option, and replicate it and get a free lunch, which is forbidden by the fact that we have constructed this augmented market to be arbitrage-free. However it might not seem immediately clear how to actually calculate this quantity. This is what we do in the next section

6.5.1 The Black-Scholes formula

With the introduction of contingent claims out of the way, we can now talk about the famous Black-Scholes formula, which gives an explicit way of pricing European call options.

Setup of the model

Definition 6.5.4 (Black-Scholes model) Our setup is now in a $d = 1$ dimensional market, where our risk-free asset has the dynamics

$$\mathrm{d}B_t = r\,B_t\,\mathrm{d}t$$

and our risky asset S has dynamics

$$\mathrm{d}S = S\,(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t)$$

where $\sigma > 0$ and μ are constants and W is Brownian motion. All of the processes above are defined only in the finite time horizon $[0, T]$ for some $T > 0$.

We will now compute explicitly the price at time t, ξ_t according to the no-arbitrage principle of a contingent claim that pays $g(S_T)$ at time t. Of course the first step is to show that this market is arbitrage free. As it turns out, we can actually compute an equivalent local Martingale measure by using Novikov's criterion and Girsanov's Theorem. Let us present how one finds this measure, perhaps in a slightly backwards but more illustrative way:

Step 1: identify the drift

We want to find a measure $\boldsymbol{Q}$, under which S/B is indeed a $\boldsymbol{Q}$-local Martingale. Note that by Ito's formula:

$$\begin{aligned}
\mathrm{d}\left(\frac{S}{B}\right) &= \mathrm{d}\left(\frac{1}{B_0}\exp\left\{\left(\mu - \frac{\sigma^2}{2} - r\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t\right\}\right) \\
&= \frac{S}{B}\left(\left(\mu - \frac{\sigma^2}{2} - r\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t + \frac{\sigma^2}{2}\,\mathrm{d}t\right) \\
&= \frac{S}{B}\left((\mu - r)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\right). \qquad (6.6)
\end{aligned}$$

A way of verifying that this is indeed a local Martingale under $\boldsymbol{Q}$, would be to set for example

$$\widehat{W}_t = \frac{\mu - r}{\sigma}t + W_t, \qquad (6.7)$$

i.e: a Brownian motion with drift. If we can find a measure $\boldsymbol{Q}$ for which $\widehat{W}_t$ is indeed a Brownian motion, we would rewrite 6.6 as $\frac{S}{B}\sigma\,\mathrm{d}\widehat{W}_t$, which would turn S/B into a $\boldsymbol{Q}$-local Martingale.

Step 2: Using Girsanov's Theorem

Girsanov's Theorem states that under certain conditions, if W is a $\boldsymbol{P}$-Brownian motion, and $(\alpha_s)_s$ is such that $\mathscr{E}\left(\int \alpha_s \, \mathrm{d}W_s\right)$ is a UI Martingale, then $\widehat{W_t} = W_t - \int_0^t \alpha_s \, \mathrm{d}s$ is a $\boldsymbol{Q}$-Brownian Motion under a suitable measure $\boldsymbol{Q}$. Looking at 6.7 we notice that the appropriate choice of α_s is to set

$$\alpha_s \equiv -\frac{\mu - r}{\sigma},$$

and since $\boldsymbol{E} \exp\left(\int_0^T \alpha_s^2 \, \mathrm{d}s\right) < \infty$, Novikov's condition tells that in fact $\mathscr{E}(\int \alpha_s \, \mathrm{d}W_s)$ is a UI Martingale, and so we have our desired measure $\boldsymbol{Q}$ for which $\widehat{W_t} = \frac{\mu - r}{\sigma} t + W_t$ is a $\boldsymbol{Q}$-Brownian motion. Therefore we have an equivalent local Martingale measure and so the model defined in 6.5.4 is arbitrage free. Naturally, one may now rewrite the market dynamics as

$$\mathrm{d}S_t = S_t \left(r \, \mathrm{d}t + \sigma \, \mathrm{d}\widehat{W_t}\right),$$

or alternatively solve for S_t:

$$S_t = S_0 \exp\left\{\left(r - \frac{\sigma^2}{2}\right) t + \sigma \widehat{W_t}\right\} \tag{6.8}$$

Step 3: Price the contingent claim

As mentioned in Remark 6.5.3, a way to price new assets at time $t \leqslant T$ is given by

$$\xi_t / B_t = \boldsymbol{E}_{\boldsymbol{Q}}\left[\frac{\xi_T}{B_T} \middle| \mathscr{F}_t\right],$$

this way we immediately have that ξ/B is a $\boldsymbol{Q}$-local Martingale, and since $\boldsymbol{Q}$ was so that the risky asset divided by the bank account was also a $\boldsymbol{Q}$-local Martingale, we would then have that the entire market (B, S, ξ) is arbitrage-free. Naturally, if what we are trying to price is a contingent claim with payout $g(S_T)$ at time T, ξ_T should be $g(S_T)$ and so from this we get that

$$\begin{aligned}
\xi_t &= e^{-r(T-t)} \boldsymbol{E}_{\boldsymbol{Q}}\left[g(S_T) | \mathscr{F}_t\right] && (6.9)\\
&= e^{-r(T-t)} \boldsymbol{E}_{\boldsymbol{Q}}\left[g\left(S_t \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\left(\widehat{W_T} - \widehat{W_t}\right)\right\}\right) \middle| \mathscr{F}_t\right] && (6.10)\\
&= e^{-r(T-t)} \int_{\mathbf{R}} g\left(S_t \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\left(\sqrt{T-t}\, z\right)\right\}\right) \phi(z) \, \mathrm{d}z && (6.11)
\end{aligned}$$

where in this last step we have used the *independence Lemma*, and $\phi(z)$ is the density of a standard Gaussian. Thus if s is the price of the stock at time t, the price of the option will be the function

$$v(s,t) = e^{-r(T-t)} \int_{\mathbf{R}} g\left(s \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\left(\sqrt{T-t}\, z\right)\right\}\right) \phi(z) \, \mathrm{d}z \tag{6.12}$$

Replicating the payoff

We end these notes with the following result, which tells us how one may produce a trading strategy whose payout at time T is the exact same as that of the contingent claim, i.e: $g(S_T)$.

Appendix A

Measure Theory

Here is some Measure Theory that is used throughout this document. All of it is seen already in previous courses.

A.1 Big Theorems

Theorem A.1.1 (Monotone Class Theorem for functions) Let $\mathscr{A}$ be a π-system that contains Ω and let $\mathscr{H}$ be a collection of functions $\Omega \to \mathbf{R}$ with the following properties:

- For all $A \in \mathscr{A}$, we have that $\mathbf{1}(A) \in \mathscr{H}$.
- If f and g are both functions in $\mathscr{H}$, then $f + g \in \mathscr{H}$.
- If $\{f_n\}$ is a sequence of measurable functions in $\mathscr{H}$ whose monotone increasing pointwise limit is some bounded function f, then $f \in \mathscr{H}$

Then $\mathscr{H}$ contains all b$\sigma(\mathscr{A})$ functions.

A.2 Uniform Integrability

Definition A.2.1 (Uniform Integrability) A family χ of Random Variables is said to be Uniformly Integrable if any of the two equivalent characterisations holds:

- $\sup_{X\in\chi} \boldsymbol{E}\left[|X|\mathbf{1}(|X| \geqslant k)\right] \to 0$ as $k \to \infty$.
- χ is $\mathscr{L}^1$ bounded and $\sup_{X\in\chi} \sup_{A:\boldsymbol{P}[A]<\delta} \boldsymbol{E}\left[|X\mathbf{1}_A|\right] \to 0$ as $\delta \to 0$.

Theorem A.2.2 (Bounded Convergence Theorem) If (X_n) is a uniformly bounded family of random variables, i.e: $|X_n| < C$ for all n, then if $X_n \to X$ in probability, we also have that $X_n \to X$ in $\mathscr{L}^1$.

Theorem A.2.3 (Vitali's Theorem) Let (X_n) be a family of Random Variables. The following are equivalent:

- $X_n \to X$ in Probability and (X_n) is Uniformly Integrable.
- $X_n \to X$ in $\mathscr{L}^1$.

Proof of the direct part: We want to show that $\boldsymbol{E}[|X_n - X|] \to 0$ as $n \to \infty$. And we have the following:

- Since UI families are in particular $\mathscr{L}^1$ bounded, its easy to check that $\boldsymbol{E}[|X|] < \infty$ by using Fatou's Lemma and the fact that if we have $X_n \to X$ in probability there is a subsequence along which we have convergence almost surely. From this, it follows that there exists some K_1 large enough so that $\boldsymbol{E}[|X|\mathbf{1}_{|X|>K_1}] < \epsilon/3$, since a single integrable random variable is also uniformly integrable.
- By definition of (X_n) being UI, there exists some K_2 large enough so that $\sup_n \boldsymbol{E}[|X_n|\mathbf{1}(|X_n| > K_2)] < \epsilon/3$.

Finally, we define X_n^K and X^K as $(X_n \vee -K) \wedge K$ and $(X \vee -K) \wedge K$ respectively. Then we have that

$$\boldsymbol{E}|X_n - X| \leqslant \boldsymbol{E}|X_n^K - X^K| + \boldsymbol{E}|X - X^K| + \boldsymbol{E}|X_n - X_n^K|$$

Note that $|X - X^K| \leqslant |X|\mathbf{1}(|X| \geqslant K)$ (you may wish to draw a diagram) and by the bounded convergence theorem for any K, there is some n large enough, so that we have that $\boldsymbol{E}|X_n^K - X^K| < \epsilon/3$. Now just choose $K := K_1 \vee K_2$ and choose the corresponding n and so $\boldsymbol{E}|X_n - X| < \epsilon/3 + \epsilon/3 + \epsilon/3$. ♡

Theorem A.2.4 (UI Property of Conditional Expectation) Let X be an integrable random variable on $(\Omega, \mathscr{F}, \boldsymbol{P})$ and let $\mathscr{G}$ be a collection of sub-sigma-algebras of $\mathscr{F}$. Then the family of random variables

$$\{\boldsymbol{E}[X \mid \mathscr{H}] : \mathscr{H} \in \mathscr{G}\}$$

is UI

Proof. The key idea is that for any $Y = \boldsymbol{E}[X \mid \mathscr{H}]$ in the described family, we can use the Tower Property and see that

$$\boldsymbol{P}[|Y| > k] = \frac{\boldsymbol{E}|Y|}{k} = \frac{\boldsymbol{E}[|\boldsymbol{E}[X \mid \mathscr{H}]|]}{k} \leqslant \frac{\boldsymbol{E}[\boldsymbol{E}[|X| \mid \mathscr{H}]]}{k} = \frac{\boldsymbol{E}[|X|]}{k}$$

Now, we simply have that

$$\begin{aligned} \boldsymbol{E}|Y|\mathbf{1}(|Y| > K) &\leqslant \boldsymbol{E}|X|\mathbf{1}(Y > K) && \text{(A.1)} \\ &\leqslant \boldsymbol{E}|X|\mathbf{1}(Y > K, X \leqslant r) + \boldsymbol{E}|X|\mathbf{1}(Y > K, X > r) && \text{(A.2)} \\ &\leqslant r\,\boldsymbol{P}[Y > K] + \boldsymbol{E}|X|\mathbf{1}(X > r) && \text{(A.3)} \end{aligned}$$

Where in step (A.1) we have used Jensen's Inequality and the Tower Property, and in step (A.3) we have used the fact that $\mathbf{1}(A \cap B) \leqslant \mathbf{1}(A)$. Now one takes $K \to \infty$, and then $r \to \infty$ and use the fact that a single integrable random variable is itself a UI family. Note of course that (A.3) does not depend on Y, from here we determine that

$$\sup_{\mathscr{H}} \boldsymbol{E}|Y|\mathbf{1}(|Y| > K) \to 0$$

as $K \to \infty$. ♡

Appendix B

Probability

B.1 Stopping

Definition B.1.1 (Stopped σ-algebra) Let T be a stopping time for a filtration $(\mathscr{F}_t)$, the stopped sigma algebra: $\mathscr{F}_T$ is defined as the set of events that can be seen before T rings:

$$\mathscr{F}_T = \{A \in \mathscr{F} : A \cap \{T \leqslant t\} \in \mathscr{F}_t\}$$

Theorem B.1.2 (Optional Stopping Theorem for Continuous Martingales) Let X be a continuous time Martingale, S and T be bounded stopping times, then

$$\boldsymbol{E}[X_T \mid \mathscr{F}_S] = X_{T \wedge S}$$

In fact Martingales, **are** the class of adapted processes that satisfy the Optional Stopping Theorem:

Theorem B.1.3 (Converse to the Optional Stopping Theorem) Let X be a cadlag adapted process, if for any bounded stopping time T we have that $\boldsymbol{E}[X_T] = \boldsymbol{E}[X_0]$, then X is a Martingale.

Proof. Let $0 \leqslant s \leqslant t$, and $A \in \mathscr{F}_s$, our goal is to show that $\boldsymbol{E}[\mathbf{1}_A(X_t - X_s)] = 0$. For this, construct the stopping time $T = s\,\mathbf{1}_A + t\,\mathbf{1}_{A^c}$. By hypothesis, we have that

$$\boldsymbol{E}[X_0] = \boldsymbol{E}[X_T] = \boldsymbol{E}[X_s\,\mathbf{1}_A + X_t\,\mathbf{1}_{A^c}] = \boldsymbol{E}[X_t + (X_s - X_t)\,\mathbf{1}_A]$$

This holds for any $A \in \mathscr{F}_s$, so firstly, in particular choose $A = \varnothing$, this shows that $\boldsymbol{E}[X_0] = \boldsymbol{E}[X_t]$ so we may subtract that from both sides of the equation, and then we are left with our goal. ♡