# Introduction

**Example.** Suppose that we have a gambler who repeatedly tosses a fair coin, betting £1 on getting a heads for each toss. Let

$$\xi_k = \begin{cases} 1 & \text{heads on } k\text{th toss} \\ -1 & \text{otherwise} \end{cases}$$

so $(\xi_k)_{k\geq 1}$ is an iid Bern(1/2) sequence. Let $X_n = \sum_{k=1}^n \xi_k$ be the net winnings of the gambler and $X_0 = 0$. Note $(X_n)_{n\geq 0}$ is a simple random walk on $\mathbb{Z}$, hence is a martingale (MG) with respect to $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$. Suppose that at the $m$th toss, they bet £$H_m$ on heads. Then the net winnings at time $n$ are

$$(H \cdot X)_n = \sum_{k=1}^n H_k(X_k - X_{k-1}).$$

Assume $(H_m)_{m\geq 1}$ is deterministic. We claim $H \cdot X$ is an $\mathcal{F}_n$-MG. Indeed:

(a) Integrability: obvious;

(b) Adapted: obvious;

(c) $\mathbb{E}[(H \cdot X)_{n+1} - (H \cdot X)_n|\mathcal{F}_n] = H_{n+1}\mathbb{E}[X_{n+1} - X_n|\mathcal{F}_n] = 0.$

More generally, the same is true if $H_{n+1}$ is integrable and $\mathcal{F}_n$ measurable for each $n$. This is called a *previsible process*. As before, $H \cdot X$ gives the winnings of the gambler. This is called a *martingale transform*.

The goal for the first part of the course: extend this reasoning to define

$$(H \cdot X)_t = \int_0^t H_s \mathrm{d}X_s \tag{$*$}$$

where $H$ is previsible and $X$ is a continuous martingale (e.g Brownian motion).

We cannot use the Lebesgue-Stieljes integral to define $(*)$ since this requires $X$ to have finite variation, and the only continuous martingales with finite variation are constant (see later in course). Our strategy to define the Itô integral: set

$$(H \cdot X)_t \text{``}=\text{''} \lim_{\varepsilon \to 0} \sum_{k=1}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}(X_{(k+1)\varepsilon} - X_{k\varepsilon}).$$

However we need to be careful about the type of limit since $X$ in general will be rough (not differentiable), like Brownian motion. To get convergence, we need to take advantage of cancellations. For example, if $X$ is a Brownian motion and

$H$ is a deterministic and continuous process we have

$$\begin{aligned}
&\mathbb{E}\left[\left[\sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}(X_{(k+1)\varepsilon} - X_{k\varepsilon})\right]^2\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}^2 (X_{(k+1)\varepsilon} - X_{k\varepsilon})^2 + \sum_{j \neq k} H_{k\varepsilon} H_{j\varepsilon} (X_{(k+1)\varepsilon} - X_{k\varepsilon})(X_{(j+1)\varepsilon} - X_{j\varepsilon})\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}^2 (X_{(k+1)\varepsilon} - X_{k\varepsilon})^2\right] \\
&= \sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}^2 \cdot \varepsilon \\
&\xrightarrow{\varepsilon \to 0} \int_0^t H_s^2 \mathrm{d}s.
\end{aligned}$$

The cancellations that make this work come from MG orthogonality and are what makes it possible to define the Itô integral.

After this we will learn about properties of the Itô integral:

- Stochastic analogue of the chain rule;
- Stochastic analogue of integration by parts.

The formulas will look like those in regular calculus, but with an extra term to reflect that $X$ is rough (quadratic variation). We write

$$Y_t = \int_0^t H_s \mathrm{d}X_s \iff \mathrm{d}Y_t = H_t \mathrm{d}X_t.$$

Itô's formula tells use how to write $\mathrm{d}f(Y_t)$ in terms of $\mathrm{d}Y_t$ for $f \in C^2$. This has many applications, for example

**Theorem** (Dubins-Schwarz theorem)**.** *Any continuous martingale is a time-change of a Brownian motion.*

Then we will look at Stochastic Differential Equations (SDEs), i.e

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

where $b, \sigma$ are "nice" and $B$ is a Brownian motion. For $\sigma = 0$ this is just an ODE. For $\sigma \neq 0$ this corresponds to adding noise depending on the time and state of the the system.

Last part of the course: diffusion processes and how they are related to SDEs, as well as how they can be used to solve PDEs involving 2nd order elliptic operators.

# 1 Preliminaries

Recall that $a : [0, \infty) \to \mathbb{R}$ is *càdlag* if it is right-continuous and has left limits. Let $a(x^-) = \lim_{y \to x^-} a(y)$ and $\Delta a(x) = a(x) - a(x^-)$. Suppose $a$ is non-decreasing, càdlag, $a(0) = 0$. Then there exists a unique Borel measure $\mathrm{d}a$ on $[0, \infty)$ such that $\mathrm{d}((s, t]) = a(t) - a(s)$ for all $0 \leq s < t$ (see Part II Probability & Measure).

For $f$ measurable and integrable then the Lebesgue-Stieljes integral $f \cdot a$ is defined by

$$(f \cdot a)(t) = \int_{(0,t]} f(s)\mathrm{d}a(s) \ \forall t \geq 0.$$

Then $(f \cdot a)$ is right-continuous. Moreoever if $a$ is continuous then $(f \cdot a)$ is continuous and so we can write

$$\int_{(0,t]} f(s)\mathrm{d}a(s) = \int_0^t f(s)\mathrm{d}a(s).$$

We want to integrate against a wider class of functions. Suppose that $a^+, a^-$ are functions satisifying the same conditions as from before (i.e non-decreasing and càdlag) and set $a = a^+ - a^-$. Define

$$(f \cdot a)(t) = (f \cdot a^+)(t) - (f \cdot a^-)(t)$$

for all $f$ measurable and such that both terms on the RHS are finite. The class of functions which are a difference of càdlag non-decreasing functions coincides with the class of càdlag functions of *finite variation.*

**Definition.** Let $a : [0, \infty) \to \mathbb{R}$ be càdlag. For each $n \in \mathbb{N}$, $t \geq 0$, let

$$v^n(t) = \sum_{k=0}^{\lceil 2^n t \rceil - 1} |a((k+1)2^{-n}) - a(k2^{-n})|. \tag{$*$}$$

Then the limit $v(t) := \lim_{n \to \infty} v^n(t)$ exists and is called the *total variation* of $a$ on $(0, t]$. If $v(t) < \infty$ then we say that $a$ has *finite variation* on $(0, t]$. If $a$ has finite variation on $(0, t]$ for all $t \geq 0$, we say that $a$ is of *finite variation.*

To see that $\lim_{n \to \infty} v^n(t)$ exists, fix $t > 0$ and let $t_n^+ = 2^{-n}\lceil 2^n t \rceil$, $t_n^- = 2^n(\lceil 2^n t \rceil - 1)$ so that $t_n^+ \geq t \geq t_n^-$ for all $n$ and

$$v^n(t) = \int_{k=0}^{2^n t_n^- - 1} |a((k+1)2^{-n}) - a(k2^{-n})| + |a(t_n^+) - a(t_n^-)|.$$

The triangle inequality implies that the sum is non-decreasing in $n$, so converges. The càdlag property tells us that the second term on the RHS converges to $|\Delta a(t)|$, so $v^n(t)$ does indeed converge.

**Lemma.** *Let $a$ be a càdlag function of finite variation. Then $v$ is càdlag of finite variation with $\Delta v(t) = |\Delta a(t)|$ for all $t \geq 0$, and $v$ is non-decreasing. In particular, if $a$ is continuous then $v$ is also continuous.*

*Proof.* See Example Sheet. $\square$

**Proposition.** A càdlag function can be written as a difference of two right-continuous non-decreasing if and only if it has finite variation.

*Proof.* First assume $a = a^+ - a^-$ for $a^+, a^-$ càdlag and non-decreasing. We show $a$ has finite variation. Note

$$|a(t) - a(s)| \leq (a^+(t) - a^+(s)) + (a^-(t) - a^-(s)) \ \forall 0 \leq s < t.$$

Plugging this into $(*)$ and using the fact the sum telescopes for monotone functions to get

$$v^n(t) \leq (a^+(t_n^+) - a^+(0)) + (a^-(t_n^+) - a^-(0)).$$

Since $a^+, a^-$ are right-continuous, the RHS converges to $(a^+(t) - a^+(0)) + (a^-(t) - a^-(0))$.

Now we show the reverse direction. Assume $a$ has finite variation $v(t) < \infty$ for all $t > 0$. Set $a^+ = \frac{1}{2}(v + a)$ and $a^- = \frac{1}{2}(v - a)$. Then $a = a^+ - a^-$ and $a^+, a^-$ are càdlag since $v, a$ are càdlag (by the above lemma). We show $a^+, a^-$ are non-decreasing. For $0 \leq s < t$ define $t_n^+, t_n^-$ as before and $s_n^+, s_n^-$ analogously. Then

$$\begin{aligned}
&a^+(t) - a^+(s) \\
&= \lim_{n\to\infty} \frac{1}{2}(v^n(t) - v^n(s) + a(t) - a(s)) \\
&= \lim_{n\to\infty} \frac{1}{2} \left[ \sum_{k=2^n s_n^+}^{2^n t_n^- - 1} \big( |a((k+1)2^{-n}) - a(k2^{-n})| + a((k+1)2^{-n}) - a(k2^{-n}) \big) \right. \\
&\qquad\qquad \left. + |a(t_n^+) - a(t_n^-)| + (a(t_n^+) - a(t_n^-)) \right] \\
&\geq 0.
\end{aligned}$$

The same argument works for $a^-$. $\square$

**Random integrators**: now we discuss integration against random fuctions of finite variations. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. Recall a stochastic process $X : \Omega \times [0, \infty) \to \mathbb{R}$ is *adapated* if $X_t = X(\cdot, t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$. We also say $X$ is *càdlag* if $X(\omega, \cdot)$ is càdlag for all $\omega \in \Omega$.

**Definition.** Given a càdlag adapted process $A : \Omega \times [0, \infty) \to \mathbb{R}$, its *total variation process* $V : \Omega \times [0, \infty) \to \mathbb{R}$ is defined pathwise by setting $V(\omega, \cdot)$ to be the total variation of $A(\omega, \cdot)$.

**Lemma.** *If $A$ is càdlag, adapted and of finite variation, then $V$ is càdlag, adapted and non-decreasing.*

*Proof.* We just need to show $V$ is adapted (the rest follows by previous results). For $t \geq 0$, set as before $t_n^- = 2^{-n}(\lceil 2^n t \rceil - 1)$. Then define

$$\tilde{V}_t^n = \sum_{k=0}^{2^n t_n^- - 1} |A_{(k+1)2^{-n}} - A_{k2^{-n}}|$$

so $\tilde{V}^n$ is adapted for all $n$ as $t_n^- \leq t$. Then

$$V_t = \lim_{n \to \infty} \tilde{V}_t^n + |\Delta A(t)|$$

is $\mathcal{F}_t$-measurable as a limit/sum of $\mathcal{F}_t$-measurable functions. □