# Mathematical Foundations for Finance

## Chapter V: Stochastic Integration

Prof. Dr. Beatrice Acciaio
beatrice.acciaio@math.ethz.ch

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

# Table of contents

- From the previous chapters, we know that the trading gains or losses from a self-financing strategy $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ are described by the *stochastic integral*

$$G(\vartheta) = \vartheta \cdot S = \int \vartheta \, dS = \sum_j \vartheta_j^{\mathrm{tr}} \Delta S_j = \sum_j \vartheta_j^{\mathrm{tr}} (S_j - S_{j-1}).$$

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

- From the previous chapters, we know that the trading gains or losses from a self-financing strategy $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ are described by the *stochastic integral*

$$G(\vartheta) = \vartheta \cdot S = \int \vartheta \, dS = \sum_j \vartheta_j^{\mathrm{tr}} \Delta S_j = \sum_j \vartheta_j^{\mathrm{tr}} (S_j - S_{j-1}).$$

- We want to develop an analogous theory in continuous time. For this we need to understand how to define and how to work with a continuous-time stochastic integral process $\int \vartheta \, dS$.

- From the previous chapters, we know that the trading gains or losses from a self-financing strategy $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ are described by the *stochastic integral*

$$G(\vartheta) = \vartheta \cdot S = \int \vartheta \, dS = \sum_j \vartheta_j^{\mathrm{tr}} \Delta S_j = \sum_j \vartheta_j^{\mathrm{tr}} (S_j - S_{j-1}).$$

- We want to develop an analogous theory in continuous time. For this we need to understand how to define and how to work with a continuous-time stochastic integral process $\int \vartheta \, dS$.

- The obvious idea is to start with Riemann sums of the form $\sum \vartheta_{t_i}^{\mathrm{tr}} (S_{t_{i+1}} - S_{t_i})$ and then pass to the limit in a suitable sense.

- The simplest idea, in this context, would be to fix $\omega$, look at the trajectories $t \mapsto S_t(\omega)$ and $t \mapsto \vartheta_t(\omega)$ and take limits of

$$\sum \vartheta_{t_i}(\omega)\big(S_{t_{i+1}}(\omega) - S_{t_i}(\omega)\big)$$

like in courses on measure and integration theory.

- The simplest idea, in this context, would be to fix $\omega$, look at the trajectories $t \mapsto S_t(\omega)$ and $t \mapsto \vartheta_t(\omega)$ and take limits of

$$\sum \vartheta_{t_i}(\omega)\big(S_{t_{i+1}}(\omega) - S_{t_i}(\omega)\big)$$

  like in courses on measure and integration theory.

**However:** This works well (i.e. for many integrands $\vartheta$) only if the function $t \mapsto S_t(\omega)$ is of *finite variation* — and this would immediately exclude a process like Brownian motion as integrator.

- ▷ The simplest idea, in this context, would be to fix $\omega$, look at the trajectories $t \mapsto S_t(\omega)$ and $t \mapsto \vartheta_t(\omega)$ and take limits of

$$\sum \vartheta_{t_i}(\omega)\big(S_{t_{i+1}}(\omega) - S_{t_i}(\omega)\big)$$

  like in courses on measure and integration theory.

**However:** This works well (i.e. for many integrands $\vartheta$) only if the function $t \mapsto S_t(\omega)$ is of *finite variation* — and this would immediately exclude a process like Brownian motion as integrator.

- ▷ Therefore, one must use a different approach, and this will be explained in this chapter.

- The simplest idea, in this context, would be to fix $\omega$, look at the trajectories $t \mapsto S_t(\omega)$ and $t \mapsto \vartheta_t(\omega)$ and take limits of

$$\sum \vartheta_{t_i}(\omega)\big(S_{t_{i+1}}(\omega) - S_{t_i}(\omega)\big)$$

  like in courses on measure and integration theory.

**However:** This works well (i.e. for many integrands $\vartheta$) only if the function $t \mapsto S_t(\omega)$ is of *finite variation* — and this would immediately exclude a process like Brownian motion as integrator.

- Therefore, one must use a different approach, and this will be explained in this chapter.

- For a proof of the fact that "naive stochastic integration is impossible", we refer to Section I.8 of Protter (2005).

**Remark.** To avoid misunderstandings later, let us clarify that defining stochastic integrals as above in a pathwise manner (i.e. $\omega$ by $\omega$) may well be possible if the integrator $S$ and the integrand $\vartheta$ match up nicely enough, even if $t \mapsto S_t(\omega)$ is not of finite variation.

**Remark.** To avoid misunderstandings later, let us clarify that defining stochastic integrals as above in a pathwise manner (i.e. $\omega$ by $\omega$) may well be possible if the integrator $S$ and the integrand $\vartheta$ match up nicely enough, even if $t \mapsto S_t(\omega)$ is not of finite variation.

But if we want to fix $S$ and allow many $\vartheta$ without imposing undue restrictions, an $\omega$-wise approach leads to problems.

**Remark.** To avoid misunderstandings later, let us clarify that defining stochastic integrals as above in a pathwise manner (i.e. $\omega$ by $\omega$) may well be possible if the integrator $S$ and the integrand $\vartheta$ match up nicely enough, even if $t \mapsto S_t(\omega)$ is not of finite variation.

But if we want to fix $S$ and allow many $\vartheta$ without imposing undue restrictions, an $\omega$-wise approach leads to problems.

- ▷ We shall see this later in the context of Itô's formula, where $\vartheta$ has the form $\vartheta_t = g(S_{t-})$ for some $C^1$-function $g$.

# Basic Construction

*Throughout this chapter*, we work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the *usual conditions of right-continuity and P-completeness*. If needed, we define $\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t$.

We also fix a (real-valued) *local martingale* $M = (M_t)_{t \geq 0}$ null at 0 and having *RCLL (right-continuous with left limits) trajectories*.

*Throughout this chapter*, we work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the *usual conditions of right-continuity and P-completeness*. If needed, we define $\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t$.

We also fix a (real-valued) *local martingale* $M = (M_t)_{t \geq 0}$ null at 0 and having *RCLL (right-continuous with left limits) trajectories*.

For any process $Y = (Y_t)_{t \geq 0}$ with RCLL trajectories, we denote by

$$\Delta Y_t := Y_t - Y_{t-} := Y_t - \lim_{\substack{s \to t, \\ s < t}} Y_s$$

the *jump of $Y$ at time $t > 0$*.

**Remark.**

1) As mentioned earlier, assuming that $M$ has RCLL trajectories is not a restriction. We can always find an RCLL version of $M$ thanks to the usual conditions on $\mathbb{F}$.

**Remark.**

1) As mentioned earlier, assuming that $M$ has RCLL trajectories is not a restriction. We can always find an RCLL version of $M$ thanks to the usual conditions on $\mathbb{F}$.

2) As we want to define stochastic integrals $\int H\,dM$, and they are always over half-open intervals of the form $(a, b]$ with $0 \leq a < b \leq \infty$, the value of $M$ at 0 is irrelevant and it is enough to look at processes defined for $t > 0$.

The *simplest example* is when $M = W$ is a Brownian motion. From the previous chapter, we know that:

The *simplest example* is when $M = W$ is a Brownian motion. From the previous chapter, we know that:

- both $W$ itself and $(W_t^2 - t)_{t \geq 0}$ are then martingales,

The *simplest example* is when $M = W$ is a Brownian motion. From the previous chapter, we know that:

- both $W$ itself and $(W_t^2 - t)_{t \geq 0}$ are then martingales,
- the quantity $t$ we subtract from $W_t^2$ is the quadratic variation of $W$, which can be obtained as a pathwise limit of sums of squared increments of $W$.

The *simplest example* is when $M = W$ is a Brownian motion. From the previous chapter, we know that:

- both $W$ itself and $(W_t^2 - t)_{t\geq 0}$ are then martingales,
- the quantity $t$ we subtract from $W_t^2$ is the quadratic variation of $W$, which can be obtained as a pathwise limit of sums of squared increments of $W$.

Similar results are true for general local martingales, and this is the key for constructing stochastic integrals.

**Theorem.** (Optional quadratic variation) For any local martingale $M = (M_t)_{t\geq 0}$ null at 0, there exists a unique adapted increasing RCLL process $[M] = ([M]_t)_{t\geq 0}$ null at 0 with $\Delta[M] = (\Delta M)^2$ and having the property that $M^2 - [M]$ is also a local martingale. $[M]$ is called the *optional quadratic variation* or *square bracket process of $M$*.

**Theorem.** (Optional quadratic variation) For any local martingale $M = (M_t)_{t \geq 0}$ null at 0, there exists a unique adapted increasing RCLL process $[M] = ([M]_t)_{t \geq 0}$ null at 0 with $\Delta[M] = (\Delta M)^2$ and having the property that $M^2 - [M]$ is also a local martingale. $[M]$ is called the *optional quadratic variation* or *square bracket process of $M$*.

This process can be obtained as the quadratic variation of $M$ in the following sense: There exists a sequence $(\Pi_n)_{n \in \mathbb{N}}$ of partitions of $[0, \infty)$ with $|\Pi_n| \to 0$ as $n \to \infty$ such that

$$P\Big[[M]_t(\omega) = \lim_{n \to \infty} \sum_{t_i \in \Pi_n} \big(M_{t_i \wedge t}(\omega) - M_{t_{i-1} \wedge t}(\omega)\big)^2 \text{ for all } t \geq 0\Big] = 1.$$

**Theorem.** (Optional quadratic variation) For any local martingale $M = (M_t)_{t\geq 0}$ null at 0, there exists a unique adapted increasing RCLL process $[M] = ([M]_t)_{t\geq 0}$ null at 0 with $\Delta[M] = (\Delta M)^2$ and having the property that $M^2 - [M]$ is also a local martingale. $[M]$ is called the *optional quadratic variation* or *square bracket process of $M$*.

This process can be obtained as the quadratic variation of $M$ in the following sense: There exists a sequence $(\Pi_n)_{n\in\mathbb{N}}$ of partitions of $[0,\infty)$ with $|\Pi_n| \to 0$ as $n \to \infty$ such that

$$P\Big[[M]_t(\omega) = \lim_{n\to\infty} \sum_{t_i\in\Pi_n} \big(M_{t_i\wedge t}(\omega) - M_{t_{i-1}\wedge t}(\omega)\big)^2 \text{ for all } t \geq 0\Big] = 1.$$

If $M$ satisfies $\sup_{0\leq s\leq T} |M_s| \in L^2$ for some $T > 0$ (and hence is in particular a true martingale on $[0,T]$), then $[M]$ is integrable on $[0,T]$ (i.e. $[M]_T \in L^1$) and $M^2 - [M]$ is also a true martingale on $[0,T]$.

**Remark.**

- Recall that Brownian motion $W$ has $[W]_t = t$.

# Optional quadratic variation - remarks



**Remark.**

- Recall that Brownian motion $W$ has $[W]_t = t$.
- Note that $[M]$ has paths of finite variation. So one can easily define integrals $\int \dots d[M]$ in a pathwise manner as usual Lebesgue–Stieltjes integrals. This does not need any new theory.

**Remark.**

- Recall that Brownian motion $W$ has $[W]_t = t$.
- Note that $[M]$ has paths of finite variation. So one can easily define integrals $\int \dots d[M]$ in a pathwise manner as usual Lebesgue–Stieltjes integrals. This does not need any new theory.

For two local martingales $M$, $N$ null at 0, we define the (optional) *covariation process* $[M, N]$ by polarisation, that is

$$[M, N] := \frac{1}{4}\big([M + N] - [M - N]\big).$$

**Remark.**

- Recall that Brownian motion $W$ has $[W]_t = t$.
- Note that $[M]$ has paths of finite variation. So one can easily define integrals $\int \dots \, d[M]$ in a pathwise manner as usual Lebesgue–Stieltjes integrals. This does not need any new theory.

For two local martingales $M$, $N$ null at 0, we define the (optional) *covariation process* $[M, N]$ by polarisation, that is

$$[M, N] := \frac{1}{4}\big([M + N] - [M - N]\big).$$

**Corollary.** From the characterisation of $[M]$ in the last theorem, it follows easily that the operation $[\,\cdot\,,\,\cdot\,]$ is bilinear, and also that $[M, N]$ is the unique adapted RCLL process $B$ null at 0, of finite variation with $\Delta B = \Delta M \Delta N$ and such that $MN - B$ is again a local martingale.

**Remark.**

1) If $M$ is (locally) square-integrable, then $[M]$ is (locally) integrable (localization will be introduced later) and therefore, by the general theory of stochastic processes, admits a so-called *predictable compensator*, i.e.

**Remark.**

1) If $M$ is (locally) square-integrable, then $[M]$ is (locally) integrable (localization will be introduced later) and therefore, by the general theory of stochastic processes, admits a so-called *predictable compensator*, i.e.

- There exists a unique increasing predictable process $\langle M \rangle = (\langle M \rangle_t)_{t \geq 0}$ null at 0 such that $[M] - \langle M \rangle$, and therefore also $M^2 - \langle M \rangle$, is a (local) martingale.

## Remark.

1) If $M$ is (locally) square-integrable, then $[M]$ is (locally) integrable (localization will be introduced later) and therefore, by the general theory of stochastic processes, admits a so-called *predictable compensator*, i.e.

- There exists a unique increasing predictable process $\langle M \rangle = (\langle M \rangle_t)_{t \geq 0}$ null at 0 such that $[M] - \langle M \rangle$, and therefore also $M^2 - \langle M \rangle$, is a (local) martingale.
- The process $\langle M \rangle$ is called the *sharp bracket* (or sometimes the *predictable variance*) process of $M$. Note that we still need to define what "predictable" means in continuous time.

2) We already point out here that any adapted process which is continuous is automatically locally bounded (see later for the definition) and therefore also locally square-integrable.

2) We already point out here that any adapted process which is continuous is automatically locally bounded (see later for the definition) and therefore also locally square-integrable.

3) If $M$ is *continuous*, then so is $[M]$, because $\Delta[M] = (\Delta M)^2 = 0$. This implies then also that $[M] = \langle M \rangle$. In particular, for a Brownian motion $W$, we have $[W]_t = \langle W \rangle_t = t$ for all $t \geq 0$.

2) We already point out here that any adapted process which is continuous is automatically locally bounded (see later for the definition) and therefore also locally square-integrable.

3) If $M$ is *continuous*, then so is $[M]$, because $\Delta[M] = (\Delta M)^2 = 0$. This implies then also that $[M] = \langle M \rangle$. In particular, for a Brownian motion $W$, we have $[W]_t = \langle W \rangle_t = t$ for all $t \geq 0$.

4) If both $M$ and $N$ are locally square-integrable (e.g. if they are continuous), we also get $\langle M, N \rangle$ via polarisation.

2) We already point out here that any adapted process which is continuous is automatically locally bounded (see later for the definition) and therefore also locally square-integrable.

3) If $M$ is *continuous*, then so is $[M]$, because $\Delta[M] = (\Delta M)^2 = 0$. This implies then also that $[M] = \langle M \rangle$. In particular, for a Brownian motion $W$, we have $[W]_t = \langle W \rangle_t = t$ for all $t \geq 0$.

4) If both $M$ and $N$ are locally square-integrable (e.g. if they are continuous), we also get $\langle M, N \rangle$ via polarisation.

5) If $M$ is $\mathbb{R}^d$-valued, then $[M]$ becomes a $d \times d$-matrix-valued process with entries $[M]^{ik} = [M^i, M^k]$. To work with that, one needs to establish more properties. The same applies to $\langle M \rangle$, if it exists.

**Definition.** We denote by $b\mathcal{E}$ the set of all *bounded elementary processes* of the form

$$H = \sum_{i=0}^{n-1} h_i I_{(t_i, t_{i+1}]},$$

with $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \cdots < t_n < \infty$ and each $h_i$ a bounded (real-valued) $\mathcal{F}_{t_i}$-measurable random variable.

**Definition.** We denote by $b\mathcal{E}$ the set of all *bounded elementary processes* of the form

$$H = \sum_{i=0}^{n-1} h_i I_{(t_i, t_{i+1}]},$$

with $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \cdots < t_n < \infty$ and each $h_i$ a bounded (real-valued) $\mathcal{F}_{t_i}$-measurable random variable.

→ For any stochastic process $X = (X_t)_{t \geq 0}$, the *stochastic integral* $\int H \, dX$ of $H \in b\mathcal{E}$ is defined by

$$\int_0^t H_s \, dX_s := (H \cdot X)_t := H \cdot X_t := \sum_{i=0}^{n-1} h_i \left( X_{t_{i+1} \wedge t} - X_{t_i \wedge t} \right) \qquad \text{for } t \geq 0.$$

**Definition.** We denote by $b\mathcal{E}$ the set of all *bounded elementary processes* of the form

$$H = \sum_{i=0}^{n-1} h_i I_{(t_i, t_{i+1}]},$$

with $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \cdots < t_n < \infty$ and each $h_i$ a bounded (real-valued) $\mathcal{F}_{t_i}$-measurable random variable.

→ For any stochastic process $X = (X_t)_{t \geq 0}$, the *stochastic integral* $\int H\,dX$ of $H \in b\mathcal{E}$ is defined by

$$\int_0^t H_s\,dX_s := (H \cdot X)_t := H \cdot X_t := \sum_{i=0}^{n-1} h_i \left( X_{t_{i+1} \wedge t} - X_{t_i \wedge t} \right) \qquad \text{for } t \geq 0.$$

If $X$ and $H$ are both $\mathbb{R}^d$-valued, the integral is still real-valued, and we simply replace products by scalar products everywhere. (But then the Lemma below looks more complicated.)

**Lemma.** Let $M$ be a square-integrable martingale (i.e. $M_t \in L^2$ for all $t \geq 0$, or equiv. $\sup_{0 \leq s \leq T} |M_s| \in L^2 \, \forall T$). Then, for every $H \in b\mathcal{E}$, the stochastic integral process $H \cdot M = \int H \, dM$ is also a square-integrable martingale, with $[H \cdot M] = \int H^2 \, d[M]$ and with the *isometry property*:

$$
\begin{aligned}
E\left[(H \cdot M_\infty)^2\right] &= E\left[\left(\int_0^\infty H_s \, dM_s\right)^2\right] \\
&= E\left[\sum_{i=0}^{n-1} h_i^2 \left([M]_{t_{i+1}} - [M]_{t_i}\right)\right] \\
&= E\left[\int_0^\infty H_s^2 \, d[M]_s\right].
\end{aligned}
$$

Note that the last $d[M]$-integral can be defined $\omega$ by $\omega$, because $t \mapsto [M]_t(\omega)$ is increasing and hence of finite variation. But of course it is here also just a finite sum, because $H$ has such a simple form.

Note that the last $d[M]$-integral can be defined $\omega$ by $\omega$, because $t \mapsto [M]_t(\omega)$ is increasing and hence of finite variation. But of course it is here also just a finite sum, because $H$ has such a simple form.

**Proof of Lemma.** Adaptedness of $H \cdot M$ is clear, and so is integrability because $H$ is bounded and each $H \cdot M_t$ is just a finite sum. Moreover, $H$ is identically 0 after $t_n$ so that both infinite integrals actually end at $t_n$.

Note that the last $d[M]$-integral can be defined $\omega$ by $\omega$, because $t \mapsto [M]_t(\omega)$ is increasing and hence of finite variation. But of course it is here also just a finite sum, because $H$ has such a simple form.

**Proof of Lemma.** Adaptedness of $H \cdot M$ is clear, and so is integrability because $H$ is bounded and each $H \cdot M_t$ is just a finite sum. Moreover, $H$ is identically 0 after $t_n$ so that both infinite integrals actually end at $t_n$.

We first argue the martingale property, for simplicity only for $s = t_i$, $t = t_{i+1}$. [$\rightarrow$ *exercise:* Prove this in detail for arbitrary $s, t$.] Indeed, by first using that $h_i$ is $\mathcal{F}_{t_i}$-measurable and bounded, and then that $M$ is a martingale, we get

$$
\begin{aligned}
E[H \cdot M_t - H \cdot M_s \mid \mathcal{F}_s] &= E\left[h_i \left(M_{t_{i+1}} - M_{t_i}\right) \,\middle|\, \mathcal{F}_{t_i}\right] \\
&= h_i E[M_{t_{i+1}} - M_{t_i} \mid \mathcal{F}_{t_i}] = 0.
\end{aligned}
$$

# Proof (continued)



Next, it is easy to check [$\rightarrow$ *exercise*] that for any square-integrable martingale $N$:

$$E\left[N_t^2 - N_s^2 \,\middle|\, \mathcal{F}_s\right] = E\left[(N_t - N_s)^2 \,\middle|\, \mathcal{F}_s\right] \qquad \text{for } s \leq t.$$

# Proof (continued)

Next, it is easy to check [→ *exercise*] that for any square-integrable martingale $N$:

$$E\left[N_t^2 - N_s^2 \,\middle|\, \mathcal{F}_s\right] = E\left[(N_t - N_s)^2 \,\middle|\, \mathcal{F}_s\right] \qquad \text{for } s \le t.$$

Applying this once to $H \cdot M$ and once to $M$ yields

$$\begin{aligned} E\left[(H \cdot M_{t_{i+1}})^2 - (H \cdot M_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i}\right] &= E\left[(H \cdot M_{t_{i+1}} - H \cdot M_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= E\left[h_i^2 (M_{t_{i+1}} - M_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= h_i^2 E\left[M_{t_{i+1}}^2 - M_{t_i}^2 \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= E\left[h_i^2 \big([M]_{t_{i+1}} - [M]_{t_i}\big) \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= E\left[H^2 \cdot [M]_{t_{i+1}} - H^2 \cdot [M]_{t_i} \,\middle|\, \mathcal{F}_{t_i}\right], \end{aligned}$$

where we have used twice that $h_i$ is $\mathcal{F}_{t_i}$-measurable and bounded, and in the 4th equality also that $M^2 - [M]$ is a martingale.

Next, it is easy to check [$\rightarrow$ *exercise*] that for any square-integrable martingale $N$:

$$E\left[N_t^2 - N_s^2 \,\middle|\, \mathcal{F}_s\right] = E\left[(N_t - N_s)^2 \,\middle|\, \mathcal{F}_s\right] \qquad \text{for } s \leq t.$$

Applying this once to $H \cdot M$ and once to $M$ yields

$$\begin{aligned} E\left[(H \cdot M_{t_{i+1}})^2 - (H \cdot M_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i}\right] &= E\left[(H \cdot M_{t_{i+1}} - H \cdot M_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= E\left[h_i^2 (M_{t_{i+1}} - M_{t_i})^2 \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= h_i^2 E\left[M_{t_{i+1}}^2 - M_{t_i}^2 \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= E\left[h_i^2 \left([M]_{t_{i+1}} - [M]_{t_i}\right) \,\middle|\, \mathcal{F}_{t_i}\right] \\ &= E\left[H^2 \cdot [M]_{t_{i+1}} - H^2 \cdot [M]_{t_i} \,\middle|\, \mathcal{F}_{t_i}\right], \end{aligned}$$

where we have used twice that $h_i$ is $\mathcal{F}_{t_i}$-measurable and bounded, and in the 4th equality also that $M^2 - [M]$ is a martingale.

This argument actually shows that the process $(H \cdot M)^2 - H^2 \cdot [M]$ is a martingale. [$\rightarrow$ *exercise:* Prove this in detail.]

# Proof (continued)

Summing up and taking expectations then gives the isometry.

Summing up and taking expectations then gives the isometry.

Moreover, it is not very difficult to argue that

$$\Delta\Big( \int H^2 \, d[M] \Big) = (\Delta(H \cdot M))^2 \qquad \text{for } H \in b\mathcal{E},$$

by exploiting that $H$ is piecewise constant and $\Delta[M] = (\Delta M)^2$.

# Proof (continued)



Summing up and taking expectations then gives the isometry.

Moreover, it is not very difficult to argue that

$$\Delta\Big( \int H^2 \, d[M]\Big) = (\Delta(H \cdot M))^2 \qquad \text{for } H \in b\mathcal{E},$$

by exploiting that $H$ is piecewise constant and $\Delta[M] = (\Delta M)^2$.

In view of the theorem and the uniqueness there, the combination of the above properties implies that

$$[H \cdot M] = \left[ \int H \, dM \right] = \int H^2 \, d[M] = H^2 \cdot [M] \qquad \text{for } H \in b\mathcal{E}.$$

This completes the proof.

**q.e.d.**

# Extension I - definitions



Our goal is now to extend the above results from $H \in b\mathcal{E}$ to a larger class of integrands.

Our goal is now to extend the above results from $H \in b\mathcal{E}$ to a larger class of integrands.

- To that end, we want to view stochastic processes as random variables on the *product space* $\overline{\Omega} := \Omega \times (0, \infty)$.

Our goal is now to extend the above results from $H \in b\mathcal{E}$ to a larger class of integrands.

- ▷ To that end, we want to view stochastic processes as random variables on the *product space* $\overline{\Omega} := \Omega \times (0, \infty)$.
- ▷ We define the *predictable $\sigma$-field* $\mathcal{P}$ on $\overline{\Omega}$ as the $\sigma$-field generated by all adapted left-continuous processes, and we call a stochastic process $H = (H_t)_{t>0}$ *predictable* if it is $\mathcal{P}$-measurable when viewed as a mapping $H : \overline{\Omega} \to \mathbb{R}$.

Our goal is now to extend the above results from $H \in b\mathcal{E}$ to a larger class of integrands.

- ▷ To that end, we want to view stochastic processes as random variables on the *product space* $\overline{\Omega} := \Omega \times (0, \infty)$.
- ▷ We define the *predictable $\sigma$-field* $\mathcal{P}$ on $\overline{\Omega}$ as the $\sigma$-field generated by all adapted left-continuous processes, and we call a stochastic process $H = (H_t)_{t>0}$ *predictable* if it is $\mathcal{P}$-measurable when viewed as a mapping $H : \overline{\Omega} \to \mathbb{R}$.
- ▷ As a consequence, every $H \in b\mathcal{E}$ is then predictable as it is adapted and left-continuous.

# Extension I - definitions

- We also define the (possibly infinite) measure $P_M := P \otimes [M]$ on $(\overline{\Omega}, \mathcal{P})$ by setting

$$E_M[Y] := E\left[\int_0^\infty Y_s(\omega)\, d[M]_s(\omega)\right] \qquad \text{for } Y \geq 0 \text{ predictable.}$$

- We also define the (possibly infinite) measure $P_M := P \otimes [M]$ on $(\overline{\Omega}, \mathcal{P})$ by setting

$$E_M[Y] := E\left[\int_0^\infty Y_s(\omega)\, d[M]_s(\omega)\right] \qquad \text{for } Y \geq 0 \text{ predictable.}$$

- The inner integral is defined $\omega$-wise as a Lebesgue–Stieltjes integral because $t \mapsto [M]_t(\omega)$ is increasing, null at 0 and RCLL. So it can be viewed as the distribution function of a (possibly infinite) $\omega$-dependent measure on $(0, \infty)$.

- Finally, we introduce the space $L^2(M) := L^2(M,P) := L^2\big(\overline{\Omega}, \mathcal{P}, P_M\big)$ as

$$L^2(M) \;:= \Big\{ \text{all (equivalence classes of) predictable } H = (H_t)_{t>0} \text{ such}$$

$$\text{that } \|H\|_{L^2(M)} := \big(E_M[H^2]\big)^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2 \, d[M]_s\right]\right)^{\frac{1}{2}} < \infty \Big\}.$$

(As usual, by taking equivalence classes, we identify $H$ and $H'$ if they agree $P_M$-a.e. on $\overline{\Omega}$.)

- Finally, we introduce the space $L^2(M) := L^2(M, P) := L^2\big(\overline{\Omega}, \mathcal{P}, P_M\big)$ as

$$L^2(M) \;:= \Big\{\text{all (equivalence classes of) predictable } H = (H_t)_{t>0} \text{ such}$$

$$\text{that } \|H\|_{L^2(M)} := \big(E_M[H^2]\big)^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2\, d[M]_s\right]\right)^{\frac{1}{2}} < \infty\Big\}.$$

(As usual, by taking equivalence classes, we identify $H$ and $H'$ if they agree $P_M$-a.e. on $\overline{\Omega}$.)

- Note that, by the previous lemma, $b\mathcal{E} \subseteq L^2(M)$ for every square-integrable martingale $M$.

Note that we can restate the first part of the Lemma as follows:

*For a fixed square-integrable martingale $M$, the mapping $H \mapsto H \cdot M$ is linear and goes from $b\mathcal{E}$ to the space $\mathcal{M}_0^2$ of all RCLL martingales $N = (N_t)_{t \geq 0}$ null at 0 which satisfy $\sup_{t \geq 0} E[N_t^2] < \infty$.*

Note that we can restate the first part of the Lemma as follows:

*For a fixed square-integrable martingale $M$, the mapping $H \mapsto H \cdot M$ is linear and goes from $b\mathcal{E}$ to the space $\mathcal{M}_0^2$ of all RCLL martingales $N = (N_t)_{t \geq 0}$ null at 0 which satisfy $\sup_{t \geq 0} E[N_t^2] < \infty$.*

The last assertion is true because each $H \cdot M$ remains constant after some $t_n$ given by $H \in b\mathcal{E}$, and because Doob's inequality gives for any martingale $N$ and any $t \geq 0$ that

$$E\left[\sup_{0 \leq s \leq t} |N_s|^2\right] \leq 4E\left[|N_t|^2\right].$$

- The martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t\to\infty} N_t$ $P$-a.s..

- The martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t\to\infty} N_t$ $P$-a.s..
- By Fatou's lemma, we have $N_\infty \in L^2$.

- The martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t\to\infty} N_t$ $P$-a.s..
- By Fatou's lemma, we have $N_\infty \in L^2$.
- The process $(N_t)_{0\le t\le\infty}$, which is defined on the closed interval $[0,\infty]$, is still a martingale.

- The martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t\to\infty} N_t$ $P$-a.s..
- By Fatou's lemma, we have $N_\infty \in L^2$.
- The process $(N_t)_{0\le t\le\infty}$, which is defined on the closed interval $[0,\infty]$, is still a martingale.
- Doob's maximal inequality implies that two martingales $N$ and $N'$ which have the same final value, i.e. $N_\infty = N'_\infty$ $P$-a.s., must coincide.

- The martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t \to \infty} N_t$ $P$-a.s..
- By Fatou's lemma, we have $N_\infty \in L^2$.
- The process $(N_t)_{0 \le t \le \infty}$, which is defined on the closed interval $[0, \infty]$, is still a martingale.
- Doob's maximal inequality implies that two martingales $N$ and $N'$ which have the same final value, i.e. $N_\infty = N'_\infty$ $P$-a.s., must coincide.

**Therefore:** We can identify $N \in \mathcal{M}_0^2$ with its limit $N_\infty \in L^2(\mathcal{F}_\infty, P)$, and so $\mathcal{M}_0^2$ becomes a *Hilbert space* with the norm

$$\|N\|_{\mathcal{M}_0^2} := \|N_\infty\|_{L^2} = \left(E[N_\infty^2]\right)^{\frac{1}{2}}$$

and the scalar product

$$(N, N')_{\mathcal{M}_0^2} := (N_\infty, N'_\infty)_{L^2} = E[N_\infty N'_\infty].$$

- So the mapping $H \mapsto H \cdot M$ from $b\mathcal{E}$ to $\mathcal{M}_0^2$ is *linear* and an *isometry* since, by the previous discussion, we have that for $H \in b\mathcal{E}$,

$$\|H \cdot M\|_{\mathcal{M}_0^2} = \left(E\left[(H \cdot M_\infty)^2\right]\right)^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2 \, d[M]_s\right]\right)^{\frac{1}{2}} = \|H\|_{L^2(M)}.$$

- So the mapping $H \mapsto H \cdot M$ from $b\mathcal{E}$ to $\mathcal{M}_0^2$ is *linear* and an *isometry* since, by the previous discussion, we have that for $H \in b\mathcal{E}$,

$$\|H \cdot M\|_{\mathcal{M}_0^2} = \left(E\left[(H \cdot M_\infty)^2\right]\right)^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2 \, d[M]_s\right]\right)^{\frac{1}{2}} = \|H\|_{L^2(M)}.$$

- This mapping can therefore be *uniquely extended to the closure of $b\mathcal{E}$ in $L^2(M)$*. In other words: we can define a stochastic integral process $H \cdot M$ for every $H$ that can be approximated, with respect to the norm $\| \cdot \|_{L^2(M)}$, by processes from $b\mathcal{E}$, and the resulting $H \cdot M$ is again a martingale in $\mathcal{M}_0^2$ and still satisfies the isometry property.

The crucial question now is of course how we can describe the closure of $b\mathcal{E}$ and especially how big it is — the bigger the better, because we then have many integrands.

The crucial question now is of course how we can describe the closure of $b\mathcal{E}$ and especially how big it is — the bigger the better, because we then have many integrands.

**Proposition.** Suppose that $M$ is in $\mathcal{M}_0^2$. Then:

1) $b\mathcal{E}$ is dense in $L^2(M)$, i.e. the closure of $b\mathcal{E}$ in $L^2(M)$ is $L^2(M)$: for each $H \in L^2(M)$, there are $H^n \in b\mathcal{E}$ s.t. $\lim_n \|H - H^n\|_{L^2(M)} = 0$.

The crucial question now is of course how we can describe the closure of $b\mathcal{E}$ and especially how big it is — the bigger the better, because we then have many integrands.

**Proposition.** Suppose that $M$ is in $\mathcal{M}_0^2$. Then:

1) $b\mathcal{E}$ is dense in $L^2(M)$, i.e. the closure of $b\mathcal{E}$ in $L^2(M)$ is $L^2(M)$: for each $H \in L^2(M)$, there are $H^n \in b\mathcal{E}$ s.t. $\lim_n \|H - H^n\|_{L^2(M)} = 0$.
2) For every $H \in L^2(M)$, the stochastic integral process $H \cdot M = \int H\, dM$ is well defined, in $\mathcal{M}_0^2$ and satisfies the isometry property.

The crucial question now is of course how we can describe the closure of $b\mathcal{E}$ and especially how big it is — the bigger the better, because we then have many integrands.

**Proposition.** Suppose that $M$ is in $\mathcal{M}_0^2$. Then:

1) $b\mathcal{E}$ is dense in $L^2(M)$, i.e. the closure of $b\mathcal{E}$ in $L^2(M)$ is $L^2(M)$: for each $H \in L^2(M)$, there are $H^n \in b\mathcal{E}$ s.t. $\lim_n \|H - H^n\|_{L^2(M)} = 0$.
2) For every $H \in L^2(M)$, the stochastic integral process $H \cdot M = \int H\,dM$ is well defined, in $\mathcal{M}_0^2$ and satisfies the isometry property.

**Proof.** Assertion 1) uses a martingale approximation argument on $\overline{\Omega}$. Assertion 2) then follows from the discussion above. **q.e.d.**

# Remark



- By definition, $M$ in $\mathcal{M}_0^2$ means that $M$ is an RCLL martingale null at 0 with $\sup_{t\geq 0} E[M_t^2] < \infty$.

In particular, we then have $E[M_t^2] < \infty$ for every $t \geq 0$ so that every $M \in \mathcal{M}_0^2$ is also a square-integrable martingale.

# Remark



- By definition, $M$ in $\mathcal{M}_0^2$ means that $M$ is an RCLL martingale null at 0 with $\sup_{t\geq 0} E[M_t^2] < \infty$.

  In particular, we then have $E[M_t^2] < \infty$ for every $t \geq 0$ so that every $M \in \mathcal{M}_0^2$ is also a square-integrable martingale.

- However: The converse is not true! Brownian motion $W$, for example, is a martingale and has $E[W_t^2] = t$. So $\sup_{t\geq 0} E[W_t^2] = +\infty$ which means that BM is not in $\mathcal{M}_0^2$.

- By definition, $M$ in $\mathcal{M}_0^2$ means that $M$ is an RCLL martingale null at 0 with $\sup_{t\geq 0} E[M_t^2] < \infty$.

  In particular, we then have $E[M_t^2] < \infty$ for every $t \geq 0$ so that every $M \in \mathcal{M}_0^2$ is also a square-integrable martingale.

- However: The converse is not true! Brownian motion $W$, for example, is a martingale and has $E[W_t^2] = t$. So $\sup_{t\geq 0} E[W_t^2] = +\infty$ which means that BM is not in $\mathcal{M}_0^2$.

- This makes it clear that we need to extend our approach to stochastic integration further. This can be done via *localisation*.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

**Definition.**

$\mathcal{M}^2_{0,\mathrm{loc}}$: set of local martingales $M$ null at 0 for which there is a sequence of stopping times $\tau_n \nearrow \infty$ $P$-a.s. such that $M^{\tau_n} \in \mathcal{M}^2_0$ for each $n$. A local martingale $M \in \mathcal{M}^2_{0,\mathrm{loc}}$ is called *locally square-integrable*.

**Definition.**

$\mathcal{M}^2_{0,\,\mathrm{loc}}$: set of local martingales $M$ null at 0 for which there is a sequence of stopping times $\tau_n \nearrow \infty$ $P$-a.s. such that $M^{\tau_n} \in \mathcal{M}^2_0$ for each $n$. A local martingale $M \in \mathcal{M}^2_{0,\,\mathrm{loc}}$ is called *locally square-integrable*.

$L^2_{\mathrm{loc}}(M)$: set of predictable processes $H$ for which there exists a sequence of stopping times $\tau_n \nearrow \infty$ $P$-a.s. such that $HI_{]\!]0,\tau_n]\!]} \in L^2(M)$ $\forall n$, where we use the stochastic interval notation

$$]\!]0,\tau_n]\!] := \{(\omega,t) \in \overline{\Omega} \,|\, 0 < t \leq \tau_n(\omega)\}.$$

- For $M \in \mathcal{M}^2_{0,\mathrm{loc}}$ and $H \in L^2_{\mathrm{loc}}(M)$, defining the stochastic integral is straightforward; we simply set

$$H \cdot M := \left(H I_{]\!]0,\tau_n]\!]}\right) \cdot M^{\tau_n} \quad \text{on } ]\!]0,\tau_n]\!]$$

which gives a definition on all of $\overline{\Omega}$ as $\tau_n \nearrow \infty$, so that $]\!]0,\tau_n]\!]$ increases to $\overline{\Omega}$.

- For $M \in \mathcal{M}^2_{0,\,\mathrm{loc}}$ and $H \in L^2_{\mathrm{loc}}(M)$, defining the stochastic integral is straightforward; we simply set

$$H \cdot M := \left(HI_{]\!]0,\tau_n]\!]}\right) \cdot M^{\tau_n} \quad \text{on } ]\!]0,\tau_n]\!]$$

  which gives a definition on all of $\overline{\Omega}$ as $\tau_n \nearrow \infty$, so that $]\!]0,\tau_n]\!]$ increases to $\overline{\Omega}$.

- The only point we need to check is that this definition is *consistent*, i.e. that the definition on $]\!]0,\tau_{n+1}]\!] \supseteq ]\!]0,\tau_n]\!]$ does not clash with the definition on $]\!]0,\tau_n]\!]$.

- For $M \in \mathcal{M}^2_{0,\,\mathrm{loc}}$ and $H \in L^2_{\mathrm{loc}}(M)$, defining the stochastic integral is straightforward; we simply set

$$H \cdot M := \left(H I_{]\!]0,\tau_n]\!]}\right) \cdot M^{\tau_n} \quad \text{on } ]\!]0,\tau_n]\!]$$

which gives a definition on all of $\overline{\Omega}$ as $\tau_n \nearrow \infty$, so that $]\!]0,\tau_n]\!]$ increases to $\overline{\Omega}$.

- The only point we need to check is that this definition is *consistent*, i.e. that the definition on $]\!]0,\tau_{n+1}]\!] \supseteq ]\!]0,\tau_n]\!]$ does not clash with the definition on $]\!]0,\tau_n]\!]$.

- This can be done by using the (subsequently listed) properties of stochastic integrals, but we do not go into details here.

# Extension II



- For $M \in \mathcal{M}^2_{0,\,\mathrm{loc}}$ and $H \in L^2_{\mathrm{loc}}(M)$, defining the stochastic integral is straightforward; we simply set

$$H \cdot M := \left(H I_{]\!]0,\tau_n]\!]}\right) \cdot M^{\tau_n} \quad \text{on } ]\!]0,\tau_n]\!]$$

which gives a definition on all of $\overline{\Omega}$ as $\tau_n \nearrow \infty$, so that $]\!]0,\tau_n]\!]$ increases to $\overline{\Omega}$.

- The only point we need to check is that this definition is *consistent*, i.e. that the definition on $]\!]0,\tau_{n+1}]\!] \supseteq ]\!]0,\tau_n]\!]$ does not clash with the definition on $]\!]0,\tau_n]\!]$.

- This can be done by using the (subsequently listed) properties of stochastic integrals, but we do not go into details here.

- Of course, $H \cdot M$ is then in $\mathcal{M}^2_{0,\,\mathrm{loc}}$.

- A closer look at the developments so far shows that the *definitions* (but not the preceding results and arguments) for $P_M$ and $L^2(M)$ only need $[M]$; hence one can introduce and use them for any local martingale $M$.

- A closer look at the developments so far shows that the *definitions* (but not the preceding results and arguments) for $P_M$ and $L^2(M)$ only need $[M]$; hence one can introduce and use them for any local martingale $M$.
- One can also define a stochastic integral process $H \cdot M$ for $H \in L^2_{\mathrm{loc}}(M)$ when $M$ is a general local martingale, but this requires substantially more theory.

- A closer look at the developments so far shows that the *definitions* (but not the preceding results and arguments) for $P_M$ and $L^2(M)$ only need $[M]$; hence one can introduce and use them for any local martingale $M$.
- One can also define a stochastic integral process $H \cdot M$ for $H \in L^2_{\mathrm{loc}}(M)$ when $M$ is a general local martingale, but this requires substantially more theory.
- If $M$ is $\mathbb{R}^d$-valued with components $M^i$ that are all in $\mathcal{M}^2_{0,\mathrm{loc}}$, one can also define the so-called *vector stochastic integral* $H \cdot M$ for $\mathbb{R}^d$-valued predictable processes in a suitable space $L^2_{\mathrm{loc}}(M)$; the result is then a real-valued process.

- However, one *warning* is indicated: $L^2_{\mathrm{loc}}(M)$ is not obtained by just asking that each component $H^i$ should be in $L^2_{\mathrm{loc}}(M^i)$ and then setting $H\cdot M = \sum_i H^i\cdot M^i$.
  In fact, it can happen that $H\cdot M$ is well defined whereas the individual $H^i\cdot M^i$ are not. So the intuition for the multidimensional case is that

$$\text{“}\ \int H\,dM = \int \sum_i H^i\,dM^i \neq \sum_i \int H^i\,dM^i\ \text{”}$$

- However, one *warning* is indicated: $L^2_{\mathrm{loc}}(M)$ is not obtained by just asking that each component $H^i$ should be in $L^2_{\mathrm{loc}}(M^i)$ and then setting $H \cdot M = \sum_i H^i \cdot M^i$.
  In fact, it can happen that $H \cdot M$ is well defined whereas the individual $H^i \cdot M^i$ are not. So the intuition for the multidimensional case is that

$$\text{“} \int H\,dM = \int \sum_i H^i\,dM^i \neq \sum_i \int H^i\,dM^i \text{”}$$

- One can extend the stochastic integral even further to more general integrands, but this becomes technical and also has a nontrivial *pitfall*: There are (real-valued) local martingales $M$ and predictable integrands $H$ such that the stochastic integral process $\int H\,dM$ is well defined, *but not a local martingale (!)*.

# Continuous local martingale



- To end this section on a positive note, let us consider the case where $M$ is a *continuous* local martingale null at 0, briefly written as $M \in \mathcal{M}^c_{0,\mathrm{loc}}$.
- This includes in particular the case of a Brownian motion $W$.

- To end this section on a positive note, let us consider the case where $M$ is a *continuous* local martingale null at 0, briefly written as $M \in \mathcal{M}^c_{0,\mathrm{loc}}$.
- This includes in particular the case of a Brownian motion $W$.
- Then $M$ is in $\mathcal{M}^2_{0,\,\mathrm{loc}}$ because it is even *locally bounded*: For the stopping times

$$\tau_n := \inf\left\{t \geq 0 \,\middle|\, |M_t| > n\right\} \nearrow \infty \qquad P\text{-a.s.},$$

we have by continuity that $|M^{\tau_n}| \leq n$ for each $n$, because

$$\left|M^{\tau_n}_t\right| = |M_{t\wedge\tau_n}| = \begin{cases} |M_t| \leq n & \text{if } t \leq \tau_n, \\ |M_{\tau_n}| = n & \text{if } t > \tau_n. \end{cases}$$

The set $L^2_{\mathrm{loc}}(M)$ of nice integrands for $M$ can here be explicitly described as

$$L^2_{\mathrm{loc}}(M) = \Big\{ \text{all predictable processes } H = (H_t)_{t>0} \text{ such that } \int_0^t H_s^2 \, d\langle M\rangle_s < \infty \ P\text{-a.s. for each } t \geq 0 \Big\}.$$

The set $L^2_{\mathrm{loc}}(M)$ of nice integrands for $M$ can here be explicitly described as

$$L^2_{\mathrm{loc}}(M) = \Bigg\{ \text{all predictable processes } H = (H_t)_{t>0} \text{ such that } \int_0^t H_s^2 \, d\langle M\rangle_s < \infty \ P\text{-a.s. for each } t \geq 0 \Bigg\}.$$

Finally, the resulting stochastic integral $H \cdot M = \int H \, dM$ is then (as we shall see from the properties in the next section) also a continuous local martingale, and of course null at 0.

# Properties

- If $M$ is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$, then $\int H\,dM$ is a local martingale in $\mathcal{M}^2_{0,\mathrm{loc}}$. If $H \in L^2(M)$, then $\int H\,dM$ is even a martingale in $\mathcal{M}^2_0$.

- If $M$ is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$, then $\int H\,dM$ is a local martingale in $\mathcal{M}^2_{0,\,\mathrm{loc}}$. If $H \in L^2(M)$, then $\int H\,dM$ is even a martingale in $\mathcal{M}^2_0$.

- If $M$ is a local martingale and $H$ is predictable and locally bounded (which means that there are stopping times $\tau_n \nearrow \infty$ $P$-a.s. such that $HI_{]\!]0,\tau_n]\!]}$ is bounded by a constant $c_n$, say, for each $n \in \mathbb{N}$), then $\int H\,dM$ is a local martingale.

- If $M$ is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$, then $\int H\,dM$ is a local martingale in $\mathcal{M}^2_{0,\mathrm{loc}}$. If $H \in L^2(M)$, then $\int H\,dM$ is even a martingale in $\mathcal{M}^2_0$.
- If $M$ is a local martingale and $H$ is predictable and locally bounded (which means that there are stopping times $\tau_n \nearrow \infty$ $P$-a.s. such that $HI_{]\!]0,\tau_n]\!]}$ is bounded by a constant $c_n$, say, for each $n \in \mathbb{N}$), then $\int H\,dM$ is a local martingale.
- If $M$ is a martingale in $\mathcal{M}^2_0$ and $H$ is predictable and bounded, then $\int H\,dM$ is again a martingale in $\mathcal{M}^2_0$.

- If $M$ is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$, then $\int H\,dM$ is a local martingale in $\mathcal{M}^2_{0,\,\mathrm{loc}}$. If $H \in L^2(M)$, then $\int H\,dM$ is even a martingale in $\mathcal{M}^2_0$.
- If $M$ is a local martingale and $H$ is predictable and locally bounded (which means that there are stopping times $\tau_n \nearrow \infty$ $P$-a.s. such that $HI_{]\!]0,\tau_n]\!]}$ is bounded by a constant $c_n$, say, for each $n \in \mathbb{N}$), then $\int H\,dM$ is a local martingale.
- If $M$ is a martingale in $\mathcal{M}^2_0$ and $H$ is predictable and bounded, then $\int H\,dM$ is again a martingale in $\mathcal{M}^2_0$.

**Warning**: If $M$ is a martingale and $H$ is predictable and bounded, then $\int H\,dM$ need not be a martingale; this is in striking contrast to the situation in discrete time.

- *Linearity*: If $M$ is a local martingale and $H, H'$ are in $L^2_{\mathrm{loc}}(M)$ and $a, b \in \mathbb{R}$, then also $aH + bH'$ is in $L^2_{\mathrm{loc}}(M)$ and

$$(aH + bH')\cdot M = (aH)\cdot M + (bH')\cdot M = a(H\cdot M) + b(H'\cdot M).$$

- *Linearity*: If $M$ is a local martingale and $H, H'$ are in $L^2_{\text{loc}}(M)$ and $a, b \in \mathbb{R}$, then also $aH + bH'$ is in $L^2_{\text{loc}}(M)$ and

$$(aH + bH') \cdot M = (aH) \cdot M + (bH') \cdot M = a(H \cdot M) + b(H' \cdot M).$$

- *Associativity*: If $M$ is a local martingale and $H \in L^2_{\text{loc}}(M)$, then we already know that $H \cdot M$ is again a local martingale. Then a predictable process $K$ is in $L^2_{\text{loc}}(H \cdot M)$ if and only if $KH$ is in $L^2_{\text{loc}}(M)$, and then

$$K \cdot (H \cdot M) = (KH) \cdot M,$$

that means

$$\int K \, d\left( \int H \, dM \right) = \int KH \, dM.$$

- Suppose that $M$ is a local martingale, $H \in L^2_{\mathrm{loc}}(M)$ and $\tau$ is a stopping time. Then $M^\tau$ is a local martingale by the stopping theorem, $H$ is in $L^2_{\mathrm{loc}}(M^\tau)$, $HI_{]\!]0,\tau]\!]}$ is in $L^2_{\mathrm{loc}}(M)$, and we have

$$(H \cdot M)^\tau = H \cdot (M^\tau) = \left(HI_{]\!]0,\tau]\!]}\right) \cdot M = \left(HI_{]\!]0,\tau]\!]}\right) \cdot (M^\tau).$$

**In words:** A stopped stochastic integral is computed by either first stopping the integrator and then integrating, or setting the integrand equal to 0 after the stopping time and then integrating, or combining the two.

- Suppose that $M, N$ are local martingales, $H \in L^2_{\mathrm{loc}}(M)$ and $K \in L^2_{\mathrm{loc}}(N)$. Then

$$\left[\int H\, dM, N\right] = \int H\, d\,[M, N]$$

and

$$\left[\int H\, dM, \int K\, dN\right] = \int HK\, d[M, N].$$

**In words:** The covariation process of two stochastic integrals is obtained by integrating the product of the integrands with respect to the covariation process of the integrators.

- In particular, $[\int H\, dM] = \int H^2\, d[M]$. (We have seen this already for $H \in b\mathcal{E}$ in the previous section)

- Suppose $M$ is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$. Then we already know that $H \cdot M$ is in $\mathcal{M}^2_{0,\,\mathrm{loc}}$ and therefore RCLL. Its jumps are given by
$$\Delta\left(\int H\,dM\right)_t = H_t\,\Delta M_t \qquad \text{for } t > 0,$$
where $\Delta Y_t := Y_t - Y_{t-}$ again denotes the jump at time $t$ of a process $Y$ with trajectories which are RCLL (right-continuous and having left limits).

**Example.** To illustrate why the direct use of the definitions is complicated, let us compute the stochastic integral $\int W\,dW$ for a Brownian motion $W$.

This is well defined because $M := W$ is in $\mathcal{M}^2_{0,\,\mathrm{loc}}$ (it is even continuous) and $H := W$ is predictable and locally bounded, because it is adapted and continuous.

**Example.** To illustrate why the direct use of the definitions is complicated, let us compute the stochastic integral $\int W\, dW$ for a Brownian motion $W$.

This is well defined because $M := W$ is in $\mathcal{M}^2_{0,\mathrm{loc}}$ (it is even continuous) and $H := W$ is predictable and locally bounded, because it is adapted and continuous.

Because

$$2W_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right) = W^2_{t_{i+1}} - W^2_{t_i} - \left(W_{t_{i+1}} - W_{t_i}\right)^2 ,$$

by elementary algebra, we obtain by summing up that

$$\sum_{t_i \in \Pi_n} W_{t_i \wedge t}\left(W_{t_{i+1}\wedge t} - W_{t_i \wedge t}\right) = \frac{1}{2}\left(W^2_t - W^2_0\right) - \frac{1}{2}\sum_{t_i \in \Pi_n}\left(W_{t_{i+1}\wedge t} - W_{t_i\wedge t}\right)^2 .$$

If the mesh size $|\Pi_n|$ of the partition sequence $(\Pi_n)$ goes to 0, then the sum on the right-hand side converges $P$-a.s. to $t$, if the partitions are also refining.

# Example (continued)

If the mesh size $|\Pi_n|$ of the partition sequence $(\Pi_n)$ goes to 0, then the sum on the right-hand side converges $P$-a.s. to $t$, if the partitions are also refining.

We therefore expect to obtain

$$\int_0^t W_s \, dW_s = \frac{1}{2} W_t^2 - \frac{1}{2} t,$$

and we shall see later from Itô's formula that this is indeed correct.

If the mesh size $|\Pi_n|$ of the partition sequence $(\Pi_n)$ goes to 0, then the sum on the right-hand side converges $P$-a.s. to $t$, if the partitions are also refining.

We therefore expect to obtain

$$\int_0^t W_s \, dW_s = \frac{1}{2} W_t^2 - \frac{1}{2} t,$$

and we shall see later from Itô's formula that this is indeed correct.

Note that we should expect the first term $\frac{1}{2} W_t^2$ from classical calculus (where we have $\int_0^x y \, dy = \frac{1}{2} x^2$); the second-order correction term $\frac{1}{2} t$ appears due to the quadratic variation of Brownian trajectories.

# Extension to Semimartingales

**Definition.** A *semimartingale* is a stochastic process $X = (X_t)_{t \geq 0}$ that can be decomposed as

$$X = X_0 + M + A,$$

where $M$ is a local martingale null at 0 and $A$ is an adapted RCLL process null at 0 and having trajectories of finite variation.

**Definition.** A *semimartingale* is a stochastic process $X = (X_t)_{t \geq 0}$ that can be decomposed as

$$X = X_0 + M + A,$$

where $M$ is a local martingale null at 0 and $A$ is an adapted RCLL process null at 0 and having trajectories of finite variation.

A semimartingale $X$ is called *special* if there is such a decomposition where $A$ is in addition predictable.

# Remarks



1) If $X$ is a special semimartingale, the decomposition with $A$ predictable is *unique* and called the *canonical decomposition*. The uniqueness result uses that any local martingale which is predictable and of finite variation must be constant.

1) If $X$ is a special semimartingale, the decomposition with $A$ predictable is *unique* and called the *canonical decomposition*. The uniqueness result uses that any local martingale which is predictable and of finite variation must be constant.

2) If $X$ is a *continuous* semimartingale, both $M$ and $A$ can be chosen continuous as well. Therefore $X$ is special because $A$ is then predictable, because $A$ is adapted and continuous.

1) If $X$ is a special semimartingale, the decomposition with $A$ predictable is *unique* and called the *canonical decomposition*. The uniqueness result uses that any local martingale which is predictable and of finite variation must be constant.

2) If $X$ is a *continuous* semimartingale, both $M$ and $A$ can be chosen continuous as well. Therefore $X$ is special because $A$ is then predictable, because $A$ is adapted and continuous.

3) If $X$ is a semimartingale, then we define its *optional quadratic variation* or *square bracket* process $[X] = ([X]_t)_{t\geq 0}$ via

$$[X] := [M] + 2[M, A] + [A] := [M] + 2\sum \Delta M \Delta A + \sum (\Delta A)^2.$$

1) If $X$ is a special semimartingale, the decomposition with $A$ predictable is *unique* and called the *canonical decomposition*. The uniqueness result uses that any local martingale which is predictable and of finite variation must be constant.

2) If $X$ is a *continuous* semimartingale, both $M$ and $A$ can be chosen continuous as well. Therefore $X$ is special because $A$ is then predictable, because $A$ is adapted and continuous.

3) If $X$ is a semimartingale, then we define its *optional quadratic variation* or *square bracket* process $[X] = ([X]_t)_{t\geq 0}$ via

$$[X] := [M] + 2[M, A] + [A] := [M] + 2\sum \Delta M \Delta A + \sum (\Delta A)^2.$$

One can show that this is well defined and does not depend on the chosen decomposition of $X$. However, $X^2 - [X]$ is no longer a local martingale, but *only a semimartingale* in general.

**Definition** If $X$ is a semimartingale, we can define a stochastic integral $H \cdot X = \int H \, dX$ at least for any process $H$ which is predictable and locally bounded. We simply set

$$H \cdot X := H \cdot M + H \cdot A,$$

where $H \cdot M$ is defined as in the previous section and $H \cdot A$ is defined $\omega$-wise as a Lebesgue–Stieltjes integral.

**Definition** If $X$ is a semimartingale, we can define a stochastic integral $H \cdot X = \int H\, dX$ at least for any process $H$ which is predictable and locally bounded. We simply set

$$H \cdot X := H \cdot M + H \cdot A,$$

where $H \cdot M$ is defined as in the previous section and $H \cdot A$ is defined $\omega$-wise as a Lebesgue–Stieltjes integral.

- ▷ Of course, one still needs to check that this is well defined (e.g. without ambiguity if $X$ has several decompositions), but this can be done.

**Definition** If $X$ is a semimartingale, we can define a stochastic integral $H \cdot X = \int H \, dX$ at least for any process $H$ which is predictable and locally bounded. We simply set

$$H \cdot X := H \cdot M + H \cdot A,$$

where $H \cdot M$ is defined as in the previous section and $H \cdot A$ is defined $\omega$-wise as a Lebesgue–Stieltjes integral.

- Of course, one still needs to check that this is well defined (e.g. without ambiguity if $X$ has several decompositions), but this can be done.
- The resulting stochastic integral then has all the *properties* from the last section except those that rest in an essential way on the (local) martingale property.

**Definition** If $X$ is a semimartingale, we can define a stochastic integral $H \cdot X = \int H \, dX$ at least for any process $H$ which is predictable and locally bounded. We simply set

$$H \cdot X := H \cdot M + H \cdot A,$$

where $H \cdot M$ is defined as in the previous section and $H \cdot A$ is defined $\omega$-wise as a Lebesgue–Stieltjes integral.

- ▷ Of course, one still needs to check that this is well defined (e.g. without ambiguity if $X$ has several decompositions), but this can be done.
- ▷ The resulting stochastic integral then has all the *properties* from the last section except those that rest in an essential way on the (local) martingale property.
- ▷ So the isometry property for example is lost.

# Properties



For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.

For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.
- If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
  [This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.]

For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.
- If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
  [This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.]
- *linearity:* same formula as before.

# Properties



For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.
- If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
  [This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.]
- *linearity:* same formula as before.
- *associativity:* same formula as before.

# Properties



For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.
- If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
  [This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.]
- *linearity:* same formula as before.
- *associativity:* same formula as before.
- *behaviour under stopping:* same formula as before.

# Properties



For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.
- If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
  [This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.]
- *linearity:* same formula as before.
- *associativity:* same formula as before.
- *behaviour under stopping:* same formula as before.
- *quadratic variation and covariation:* same formula as before.

# Properties

For $X$ semimartingale and $H$ predictable and locally bounded, we have:

- $H \cdot X$ is a *semimartingale*.
- If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
  [This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.]
- *linearity:* same formula as before.
- *associativity:* same formula as before.
- *behaviour under stopping:* same formula as before.
- *quadratic variation and covariation:* same formula as before.
- *jumps:* same formula as before.

# Properties (continued)

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

- If $X$ is *continuous*, then so is $H \cdot X$; this is clear from $\Delta(H \cdot X) = H \Delta X = 0$.

- If $X$ is *continuous*, then so is $H \cdot X$; this is clear from $\Delta(H \cdot X) = H \Delta X = 0$.
- There is also a sort of *dominated convergence theorem*: If $H^n$, $n \in \mathbb{N}$, are predictable processes with $H^n \to 0$ pointwise on $\overline{\Omega}$ and $|H^n| \leq |H|$ for some locally bounded $H$, then $H^n \cdot X \to 0$ uniformly on compacts in probability, which means that

$$\sup_{0 \leq s \leq t} |H^n \cdot X_s| \longrightarrow 0 \qquad \text{in probability as } n \to \infty, \text{ for every } t \geq 0.$$

- If $X$ is *continuous*, then so is $H \cdot X$; this is clear from $\Delta(H \cdot X) = H\Delta X = 0$.
- There is also a sort of *dominated convergence theorem*: If $H^n$, $n \in \mathbb{N}$, are predictable processes with $H^n \to 0$ pointwise on $\overline{\Omega}$ and $|H^n| \leq |H|$ for some locally bounded $H$, then $H^n \cdot X \to 0$ uniformly on compacts in probability, which means that

$$\sup_{0 \leq s \leq t} |H^n \cdot X_s| \longrightarrow 0 \qquad \text{in probability as } n \to \infty, \text{ for every } t \geq 0.$$

- This can also be viewed as a *continuity property* of the stochastic integral operator $H \mapsto H \cdot X$, as (pointwise and locally bounded) convergence of $(H^n)$ implies convergence of $(H^n \cdot X)$, in the sense of the last equation.

The above concept is in fact very natural and has a number of very good properties:

- If $X$ is a semimartingale and $f$ is a $C^2$-function, then $f(X)$ is again a semimartingale. This will follow from *Itô's formula*, which even gives an explicit expression for $f(X)$.

The above concept is in fact very natural and has a number of very good properties:

- If $X$ is a semimartingale and $f$ is a $C^2$-function, then $f(X)$ is again a semimartingale. This will follow from *Itô's formula*, which even gives an explicit expression for $f(X)$.
- If $X$ is a semimartingale with respect to $P$, and $R$ is a probability measure equivalent to $P$, then $X$ is still a semimartingale with respect to $R$. This will follow from *Girsanov's theorem*, which even gives a decomposition of $X$ under $R$.

# Further remarks (continued)

- If $X$ is any adapted process with RC trajectories, we can always define the (elementary) stochastic integral $H \cdot X$ for processes $H$ in $b\mathcal{E}$.

- If $X$ is any adapted process with RC trajectories, we can always define the (elementary) stochastic integral $H \cdot X$ for processes $H$ in $b\mathcal{E}$.
- If $X$ is such that this mapping on $b\mathcal{E}$ also has the continuity property for any sequence $(H^n)_{n\in\mathbb{N}}$ in $b\mathcal{E}$ converging pointwise to 0 and with $|H^n| \leq 1$ for all $n$, then $X$ must in fact be a semimartingale.

## Further remarks (continued)



- If $X$ is any adapted process with RC trajectories, we can always define the (elementary) stochastic integral $H \cdot X$ for processes $H$ in $b\mathcal{E}$.
- If $X$ is such that this mapping on $b\mathcal{E}$ also has the continuity property for any sequence $(H^n)_{n\in\mathbb{N}}$ in $b\mathcal{E}$ converging pointwise to 0 and with $|H^n| \leq 1$ for all $n$, then $X$ must in fact be a semimartingale.
- This deep result is due to *Bichteler and Dellacherie* and shows that semimartingales are a *natural class of integrators*.

- One direct consequence of this result for finance is that semimartingales are the natural processes to model discounted asset prices in financial markets.

- One direct consequence of this result for finance is that semimartingales are the natural processes to model discounted asset prices in financial markets.

- In fact, the fundamental theorem of asset pricing (in a suitably general version for continuous-time models) essentially says that a suitably arbitrage-free model should be such that $S$ is a local martingale (or more generally a $\sigma$-martingale) under some $Q \approx P$. But then $S$ is a $Q$-semimartingale and thus also a $P$-semimartingale.

# Further remarks (continued)



- One direct consequence of this result for finance is that semimartingales are the natural processes to model discounted asset prices in financial markets.
- In fact, the fundamental theorem of asset pricing (in a suitably general version for continuous-time models) essentially says that a suitably arbitrage-free model should be such that $S$ is a local martingale (or more generally a $\sigma$-martingale) under some $Q \approx P$. But then $S$ is a $Q$-semimartingale and thus also a $P$-semimartingale.
- Put differently, the above result implies that if we start with any model where $S$ is *not a semimartingale*, there will be *arbitrage* of some kind. Things become different if one includes transaction costs; but in frictionless markets, one must be careful about this issue.