Course: Theory of Probability II
Term: Spring 2015
Instructor: Gordan Zitkovic

Lecture 19

$\mathbb{L}^2$-Stochastic integration

The stochastic integral for processes of finite variation

For $T > 0$, let $F : [0, T] \to \mathbb{R}$ be càdlàg function of finite variation, and let $|F| : [0, T] \to [0, \infty)$ be its total variation. For $0 \leq a < b \leq T$ we define

$$\mu((a, b]) = F(b) - F(a), \ \mu(\{0\}) = F(0).$$

A version of the Caratheodory extension theorem guarantees that the functional μ can be extended in a unique way to a signed (real) measure on $\mathcal{B}([0, T])$; this measure is called the **Stieltjes measure** induced by F, and it is typically denoted by dF. It is not hard to see (we leave the details to the reader) that there exist unique càdlàg functions $F_+, F_- : [0, T] \to \mathbb{R}$ of finite variation such that

1. dF_+ and dF_- are finite positive measures (do nottake negative values) with $dF = dF_+ - dF_-$, and
2. dF_+ and dF_- are absolutely continuous with respect to the measure $d|F|$ (the Stieltjes measure induced by the total variation $|F|$ of F) and $dF_+ + dF_- = d\,|F|$.

The decomposition $dF = dF_+ - dF_-$ is called the **Hahn-Jordan decomposition** of the Stieltjes measure F.

Definition 19.1 (Lebesgue-Stieltjes integral)**.** A function $h : [0, T] \to \mathbb{R}$ is said to be dF-integrable (denoted by $h \in \mathbb{L}^1(dF)$), if

$$\int_0^T |h(t)| \ d\,|F|\,(t) < \infty.$$

For $h \in \mathbb{L}^1(dF)$, its **Stieltjes integral** with respect to dF is defined by

$$\int_0^T h(u)\,dF(u) = \int_0^T h(u)\,dF_+(u) - \int_0^T h(u)\,dF_-(u).$$

When the function h is continuous, one can approximate the Lebesgue-Stieltjes integral by Riemann-like sums. More precisely, let $0 = t_0^n < t_1^n < \cdots < t_{d_n}^n < t_{d_n+1}^n = T$ be a sequence of partitions of $[0, T]$ such that

$$\lim_{n\to\infty} \sup_{0\le k\le d_n} \left|t_{k+1}^n - t_k^n\right| = 0,$$

and suppose that $h : [0, T] \to \mathbb{R}$ is a continuous function. We leave it to the reader to show that $h \in \mathbb{L}^1(dF)$ and that

$$\int_0^T h(x)\, dF(x) = \lim_{n\to\infty} \sum_{k=0}^{d_n} h(t_k^n)(F(t_{k+1}^n) - F(t_k^n)).$$

If, additionally, $F \in C^1([0, T])$, we have

$$\int_0^T h(t)\, dF(t) = \int_0^T h(t)F'(t)\, dt,$$

for all $h \in \mathbb{L}^1(dF)$.

We turn to processes, now. Let $\{X_t\}_{t\in[0,\infty)}$ be an adapted càdlàg process of finite variation, and let $\{H_t\}_{t\in[0,\infty)}$ be a progressively measurable process such that

$$\int_0^T |H_t(\omega)|\ d\,|X|_t\,(\omega) < \infty, \text{ for all } T \ge 0, \text{ for almost all } \omega,$$

where $d\,|X|_t\,(\omega)$ is the Stieltjes measure induced by the non-decreasing function $t \mapsto |X|_t\,(\omega)$ (the total-variation function of $X(\omega)$). Therefore, for almost all ω, and all $t \ge 0$, we can define the Stieltjes integral

$$Y_t(\omega) = \int_0^t H_t(\omega)\, dX_t(\omega),$$

where $dX_t(\omega)$ is the Stieltjes measure induced by the FV function $t \mapsto X_t(\omega)$. We set $Y_t(\omega) = 0$, for all $t \ge 0$ for ω in the exceptional set. It is not hard to show that the process $\{Y\}_{t\in[0,\infty)}$ is an adapted and càdlàg process of finite variation, and that it is continuous when X is continuous. The process $\{Y_t\}_{t\in[0,\infty)}$ is called the **stochastic integral** of $\{H_t\}_{t\in[0,\infty)}$ with respect to $\{X_t\}_{t\in[0,\infty)}$ and is sometimes denoted by $Y = (H \cdot X)$.

The importance of being of finite variation

Unfortunately, the Lebesgue-style integration stops with functions of finite variation. Here is why.

A sequence $0 = t_0 < t_1 < \cdots < t_k < t_{k+1} < \cdots < \infty$ of real numbers with the property that $\lim_{k\to\infty} t_k = +\infty$, is called a **partition of** $[0, \infty)$, and the set of all partitions of $[0, \infty)$ is denoted by $P_{[0,\infty)}$.

The elements $t_0, t_1, \ldots$ of a partition are referred to as its **nodes**. For a partition $\Delta = \{t_k\}_{k\in\mathbb{N}} \in P_{[0,\infty)}$, and $t \geq 0$, we define

$$|\Delta|_{[0,t]} = \sup_{k\in\mathbb{N}_0} |t_{k+1} \wedge t - t_k \wedge t|$$

and note that $|\Delta|_{[0,t]} = \max(|t_1 - t_0|, |t_2 - t_1|, \ldots, |t - t_{k(t)}|)$, where $k(t) = \sup\{k \in \mathbb{N} : t_k < t\}$. A sequence $\{\Delta_n\}_{n\in\mathbb{N}}$ in $P_{[0,\infty)}$ is said to **converge to identity**, denoted by $\Delta_n \to \mathrm{Id}$ if $|\Delta_n|_{[0,t]} \to 0$, for each $t \geq 0$.

Let $F : [0,\infty) \to \mathbb{R}$ be a continuous function, and let $\mathcal{D}_{simp}$ denote the set of all functions $h : [0,\infty) \to \mathbb{R}$ for which there exists a partition $\Delta = \{t_k\}_{k\in\mathbb{N}} \in P_{[0,\infty)}$, such that

$$h(t) = \alpha_0 \mathbf{1}_{\{0\}}(t) + \sum_{k=1}^{\infty} \alpha_k \mathbf{1}_{(t_k,t_{k+1}]}(t),\ n \in \mathbb{N},\ \alpha_k \in \mathbb{R},\ k \in \mathbb{N}. \tag{19.1}$$

For $h \in \mathcal{S}$ and $t > 0$ we naturally define

$$\int_0^t h(u)\, dF(u) = \alpha_0 F(0) + \sum_{k=1}^{\infty} \alpha_k (F(t \wedge t_{k+1}) - F(t \wedge t_k)). \tag{19.2}$$

Note that the sum is really finite (we could have stopped at the index of the first $t_k > t$). For a continuous function $h : [0,\infty) \to \mathbb{R}$ and a partition $\Delta \in P_{[0,\infty)}$, we define the approximation h^Δ to h in $\mathcal{D}_{simp}$ by

$$h^\Delta(t) = h(0)\mathbf{1}_{\{0\}}(t) + \sum_{k=1}^{\infty} h(t_k)\mathbf{1}_{(t_k,t_{k+1}]}(t),\ n \in \mathbb{N}, \tag{19.3}$$

and note that $h^\Delta \to h$ uniformly on compact sets. We ask the following question

Question 19.2. What properties does F need to have for the limit

$$\lim_{\Delta\to\mathrm{Id}} \int_0^t h^\Delta(u)\, dF(u),$$

to exist for each $t \geq 0$ and each continuous h?

Remark 19.3. The reader should note that the question we are asking above really deals with a weak continuity property for the eventual extension of the simple integral (19.2) to a larger class of integrands h. Indeed, if a functional $h \mapsto \int h\, dF$ is to be called an integral, it should be linear, and somewhat continuous. The requirement that the uniform convergence of integrands $h^\Delta \to h$ implies convergence of the integrals is a particularly weak form of the dominated convergence theorem. Nevertheless, as we will see, it restricts the class of functions F considerably.

The main analytic tool we are going to use is the celebrated Banach-Steinhaus (aka uniform-boundedness theorem). Remember that for a linear operator $T : X \to Y$ between two normed spaces $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$, we define its **operator norm** $||A||$ by

$$||T|| = \sup\{||Tf||_Y \,:\, f \in X,\ ||f||_X \leq 1\}.$$

Theorem 19.4 (Banach, Steinhaus)**.** *Let X be a Banach space, let Y be a normed space, and let $(B_\alpha)_{\alpha\in A}$ be a family of continuous linear operators $B_\alpha : X \to Y$. Then*

$$\sup_{\alpha\in B} ||B_\alpha|| < \infty \ \textit{ if and only if } \ \forall f \in X,\ \sup_{\alpha\in B} ||B_\alpha f||_Y < \infty.$$

Proof. (*) One direction is trivial. For the other, suppose that

$$\sup_{\alpha\in B} ||B_\alpha f||_Y < \infty \text{ for each} f \in X.$$

Define the sequence $\{D_n\}_{n\in\mathbb{N}}$ of subsets of X by

$$D_n = \bigcap_{\alpha\in A} \{f \in X \,:\, ||B_\alpha f|| \leq n\},$$

which, as intersections of closed sets are themselves closed. By the assumption, $\cup_{n\in\mathbb{N}} D_n = X$. Since X is complete, Baire's theorem[1] implies that at least one of $\{D_n\}_{n\in\mathbb{N}}$ has nonempty interior. More explicitly, there exists $n \in \mathbb{N}$, $f_0 \in D_n$ and $\varepsilon > 0$ such that $f_0 + \varepsilon B_1(X) \subseteq D_n$, where $B_1(X) = \{f \in X \,:\, ||f||_X \leq 1\}$ denotes the unit ball of X. Consequently,

[1] Baire's (category) theorem states that the intersection of a countable collections of dense open sets in a complete metric space is dense itself.

$$||B_\alpha(\tfrac{1}{\varepsilon}f_0 + f)||_Y \leq \tfrac{1}{\varepsilon}n \text{ for all } f \in B_1(X) \text{ and all } \alpha \in A.$$

The triangle inequality now implies that

$$||B_\alpha f||_Y \leq \tfrac{1}{\varepsilon}n + ||B_\alpha \tfrac{1}{\varepsilon}f_0|| \leq \tfrac{n}{\varepsilon} + \tfrac{1}{\varepsilon}\sup_{\alpha\in A} ||B_\alpha f_0||, \text{ for all } \alpha \in A,\ f \in B_1(X),$$

and so $\sup_{\alpha\in A} ||B_\alpha|| \leq \frac{n}{\varepsilon} + \frac{1}{\varepsilon}\sup_{\alpha\in A} ||B_\alpha f_0|| < \infty$. □

Proposition 19.5. *Let F be a continuous function on $[0,1]$, such that for each continuous function $h : [0,1] \to \mathbb{R}$, the family*

$$\int_0^1 h^{\Delta_n}(u)\, dF(u), n \in \mathbb{N},$$

is bounded in $\mathbb{R}$ for each sequence $\{\Delta_n\}_{n\in\mathbb{N}}$ in $P_{[0,\infty)}$ such that $\Delta_n \to \mathrm{Id}$. Then F is a function of finite variation on $[0,1]$.

Proof. (*) Let $X = C[0,1]$ be the Banach space of all continuous functions on $[0,1]$ normed with $||h|| = \sup_{t\in[0,1]} |h(t)|$, for $h \in C[0,1]$, and

let $Y = \mathbb{R}$. We fix a sequence $\{\Delta_n\}_{n\in\mathbb{N}}$ in $P_{[0,\infty)}$ with $\Delta_n \to \mathrm{Id}$, and assume, without loss of generality that $1 \in \Delta_n$, for all $n \in \mathbb{N}$. A sequence of operators can be defined by

$$B_n : X \to Y,\ B_n h = \int_0^1 h^{\Delta_n}(u)\, dF(u),\ n \in \mathbb{N}.$$

These operators are continuous (why?), and, by the assumption, the sequence $\{B_n h\}_{n\in\mathbb{N}}$ is bounded, for each $h \in X$. Therefore, by Theorem 19.4, there exists a constant $K \geq 0$ such that

$$\left| \int_0^1 h^{\Delta_n}(u)\, dF(u) \right| \leq K||h||,\ \forall\, h \in C[0,1],\ \forall\, n \in \mathbb{N}. \tag{19.4}$$

For each $n \in \mathbb{N}$, one can construct a function $h_n \in C[0,1]$ such that $h_n(0) = \mathrm{sgn}(F(0))$ and $h_n(t_k^n) = \mathrm{sgn}(F(t_{k+1}^n) - F(t_k^n))$, for all $k \leq k^{\Delta_n}(1)$, where $\Delta_n = \{t_k^n\}_{k\in\mathbb{N}}$. Moreover, such a function can be constructed with $||h_n|| = 1$. Using (19.4), we get

$$K \geq \left| \int_0^1 h_n^{\Delta_n}(u)\, dF(u) \right| = |F(0)| + \sum_{k=0}^{k^{\Delta_n}(1)} \left| F(t_{k+1}^n) - F(t_k^n) \right|.$$

We have proved that the sequence of variations of F along any sequence of partitions converging to identity remains bounded. It follows (why?) that F must be a function of finite variation. □

We can repeat the above argument on each $[0,t]$ instead on $[0,1]$, and, since any convergent sequence is bounded, the answer to Question 19.2 must be *It is necessary that F be of finite variation*. On the other hand, when F is of finite variation, the dominated convergence theorem implies that the sequence in Question 19.2 indeed converges towards $\int_0^t h(u)\, dF(u)$.

Square-integrable martingales as integrators

In spite of the impossibility result of the previous subsection, we will still be able to construct a satisfactory integration theory when the integrand is not of finite variation. In that case, it will be important that some randomness is present and that the class of integrands be restricted to those that do not depend on the future. This way, the terms that would otherwise accumulate and lead to the explosion in the approximating sequence will cancel each other (in the law-of-large-numbers manner). Before we move on to the general case, here is an example in a simplified (but still fully-featured) setting:

Example 19.6. Without knowing it, we have constructed a very simple version of a stochastic integral when we proved the martingale convergence theorem. Indeed, let $\{\xi_n\}_{n\in\mathbb{N}}$ be an iid sequence of coin-tosses

($\mathbb{P}[\xi_1 = 1] = \mathbb{P}[\xi_1 = -1] = \frac{1}{2}$) and let $\{M_t\}_{t\in[0,1]}$ be defined as follows:

$$M_t = \sum_{k=1}^{\infty} S_k \mathbf{1}_{[1-2^{-k+1},1-2^{-k})}(t), \text{ where } S_k = \sum_{k=1}^{n} \tfrac{1}{k}\xi_k, \text{ for } t < 1.$$

The martingale convergence theorem guarantees that the limit $M_1 = \lim_{t\to 1} M_t$ exists a.s. The reader can check that the variation of the path $M_\cdot(\omega)$ on $[0,1]$ is given by $\sum_{k=1}^{\infty} \frac{1}{k} |\xi_k(\omega)| = \sum_{k=1}^{\infty} \frac{1}{k} = +\infty$, so that no trajectory of $\{M_t\}_{t\in[0,1]}$ is of finite variation[2]. On the other hand, let $\{H_t\}_{t\in[0,1)}$ be any bounded (say by $K \geq 0$) and $\{\mathcal{F}_t^M\}_{t\in[0,1]}$-adapted left-continuous stochastic process. On each segment $[0,t]$, $t < 1$, $\{M_t\}_{t\in[0,1]}$ is a process of finite variation and so we have the following expression

$$\int_0^t H_u \, dM_u = \sum_{k=1}^{k(t)} \tfrac{h_k}{k}\xi_k,$$

where $h_k = H_{1-2^{-k+1}}$ and $k(t)$ is such that $t \in [1-2^{-k+1}, 1-2^{-k})$. Moreover, $h_k \in \sigma(\xi_1, \xi_2, \dots, \xi_{k-1})$, so that the process

$$N_n = \sum_{k=1}^{n} \tfrac{h_k}{k}\xi_k,$$

is a martingale (it is, really, a martingale transform of $\{S_n\}_{n\in\mathbb{N}}$). Moreover, it is a martingale bounded in $\mathbb{L}^2$; indeed, by the orthogonality of martingale increments, we have

$$\mathbb{E}[N_n^2] = \sum_{k=1}^{n} \mathbb{E}[\tfrac{h_k^2}{k^2}\xi_k^2] = \sum_{k=1}^{n} \tfrac{1}{k^2}K^2 \leq K^2 \sum_{k=1}^{\infty} \tfrac{1}{k^2} < \infty,$$

so that $\sup_n \mathbb{E}[N_n^2] < \infty$. Consequently, the martingale convergence theorem implies that the limit $N_\infty = \lim_{n\in\mathbb{N}} N_n$ exists a.s., and in $\mathbb{L}^2$. Equivalently, the random variable

$$\int_0^1 H_u \, dM_u = \lim_{t\nearrow 1} \int_0^t H_u \, dM_u = \lim_n N_n = N_\infty,$$

is well defined, even though no path of M is of finite variation.

It is important to note that the sequence $\{N_n\}_{n\in\mathbb{N}}$ only converges a.s. There is no way to conclude that the limit exists for each ω. Indeed, take $h_n = 1$, for all $n \in \mathbb{N}$, and note that for those ω for which $\xi_k(\omega) = 1$, for all $k \in \mathbb{N}$, the sequence

$$\sum_{k=1}^{\infty} \tfrac{1}{k}\xi_k(\omega) = \sum_{k=1}^{\infty} \tfrac{1}{k} = +\infty,$$

does not converge. Moreover, the series $\sum_{k=1}^{\infty} \frac{1}{k}\xi_k(\omega)$ will *never* converge absolutely. It is the fact that the sequence $\{\xi_k\}_{k\in\mathbb{N}}$ fluctuates so much and the fact that h_k does not "see" ξ_k that guarantee convergence $N_n \to N_\infty$, and, equivalently, allow for the integral $\int_0^1 H_u \, dM_u$ to be well-defined.

[2] The reader will also note that $\{M_t\}_{t\in[0,1]}$ is not continuous; a similar example with $\{M_t\}_{t\in[0,1]}$ continuous can be concocted, but it would not be as transparent as the one we give here. Also, if one prefers to work with $[0,\infty)$ instead of $[0,1]$, one can define $M_t = \sum_{k=1}^{\infty} S_k \mathbf{1}_{[k-1,k)}(t)$.

Turning to the general case, we assume that a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathbb{P})$, where $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ satisfies the usual conditions, is given and fixed throughout.

When the integrand is "simple" the definition of the stochastic integral is easy. More precisely, a stochastic process $\{H_t\}_{t\in[0,\infty)}$ is said to be **simple predictable** if there exists a partition $\Delta = \{t_n\}_{n\in\mathbb{N}} \in P_{[0,\infty)}$ and a sequence $\{K_n\}_{n\in\mathbb{N}_0}$ of random variables such that $K_n \in \mathcal{F}_{t_n}$, and

$$H_t = \sum_{n=0}^{\infty} K_n \mathbf{1}_{(t_n,t_{n+1}]}(t). \tag{19.5}$$

The set of all simple predictable processes is denoted by $\mathcal{H}_{simp}$. The subset $\mathcal{H}^{\infty}_{simp}$ of $\mathcal{H}_{simp}$ consists of all processes $\{H_t\}_{t\in[0,\infty)} \in \mathcal{H}_{simp}$ such that $||H||_{\mathcal{H}^{\infty}_{simp}} = \sup_{n\in\mathbb{N}_0} ||K_n||_{\mathbb{L}^\infty} < \infty$.

Definition 19.7 (Stochastic integration of simple processes)**.** Suppose that $\{M_t\}_{t\in[0,\infty)}$ is *an arbitrary* stochastic process, and let $\{H_t\}_{t\in[0,\infty)}$ be a simple predictable process with representation (19.5). The **(simple) stochastic integral** of H with respect to M is the stochastic process $\{(H\cdot M)_t\}_{t\in[0,\infty)}$, given by

$$(H\cdot M)_t = \sum_{n=0}^{\infty} K_n (M_{t\wedge t_{n+1}} - M_{t\wedge t_n}). \tag{19.6}$$

Remark 19.8. 1. Note that simple predictable processes are predictable (measurable in the σ-algebra generated by left-continuous adapted processes). Moreover, $\mathcal{H}_{simp}$ and $\mathcal{H}^{\infty}_{simp}$ are vector spaces and the functional $||\cdot||_{\mathcal{H}^{\infty}_{simp}}$ is a norm on $\mathcal{H}^{\infty}_{simp}$ (if indistinguishable processes are identified).

2. One of the best ways to visualize the stochastic integral is to remember the notion of the *martingale transform* from our discussion of discrete-time martingales. The gambling interpretation used there easily transfers to continuous time, at least when the integrands are simple predictable, as above. Note, though, that one has to think of "gambles" in an infinitesimal manner in continuous time.

3. The value at $t \geq 0$ of the stochastic integral is usually denoted by $(H\cdot M)_t$ or $\int_0^t H_u\, dM_u$. Sometimes the terminology is abused and the random variable $(H\cdot M)_t$ (as opposed to the whole process $\{(H\cdot M)_t\}_{t\in[0,\infty)}$ or, simply, $H\cdot M$) is called the stochastic integral.

The stochastic integral for square-integrable martingales

We turn now to a general construction. Let $\mathcal{M}^{2,c}_0$ denote the family of all continuous martingales M with $M_0 = 0$ such that $||M||_{\mathcal{M}^{2,c}_0} :=$

$\sqrt{\sup_{t\geq 0}\mathbb{E}[M_t^2]} < \infty$. Here is the fundamental property of the (simple) stochastic integral with respect to $M \in \mathcal{M}_0^{2,c}$:

Proposition 19.9 (Preservation of $\mathcal{M}_0^{2,c}$ - simple predictable integrands)**.** *For $M \in \mathcal{M}_0^{2,c}$ and $\{H_t\}_{t\in[0,\infty)} \in \mathcal{H}^\infty_{simp}$ we have $H \cdot M \in \mathcal{M}_0^{2,c}$ and*

$$\mathbb{E}[(H\cdot M)_t^2] = \sum_{n=0}^{\infty}\mathbb{E}[K_n^2(M_{t_{n+1}\wedge t} - M_{t_n\wedge t})^2]. \tag{19.7}$$

Proof. Continuity, value 0 at time 0, and the martingale property of $H \cdot M$ are evident from the definition. To establish (19.7), we assume that $t = t_N$ for some N (otherwise just add it to the partition), and use the fact that the increments of square-integrable martingales are orthogonal, i.e.,

$$\mathbb{E}[K_n(M_{t_{n+1}} - M_{t_n})K_m(M_{t_{m+1}} - M_{t_m})] = 0 \text{ for } n < m.$$

Finally, $(H \cdot M) \in \mathcal{M}_0^{2,c}$ because, for $t \geq 0$, we have

$$\begin{aligned}\mathbb{E}[(H\cdot M)_t^2] &= \sum_{n\in\mathbb{N}}\mathbb{E}[K_n^2(M_{t_{n+1}} - M_{t_n})^2]\\ &\leq ||H||^2_{\mathcal{H}^\infty_{simp}}\sum_n \mathbb{E}[(M_{t_{n+1}} - M_{t_n})^2]\\ &= ||H||^2_{\mathcal{H}^\infty_{simp}}||M||^2_{\mathcal{M}_0^{2,c}}.\end{aligned}$$

□

We turn now to the notion of *quadratic variation* of a continuous martingale. The main theorem listing its properties is stated without proof, but the reader is instructed to draw the parallel with the case of the Brownian motion we covered in previous lectures. Before we start, we need to introduce some notation: let $\mathcal{A}_0^c$ denote the set of all continuous, adapted, nondecreasing processes $\{A_t\}_{t\in[0,\infty)}$ with $A_0 = 0$. We also remind the reader that, for a RCLL of RLCC process X, we define the **maximal process** X^* by $X_t^* = \sup_{s\leq t}|X_s|$. For a sequence $\{X^n\}_{n\in\mathbb{N}}$ of RCLL or RLCC processes, we say that X_n converges to X, **uniformly on compacts in probability (ucp)**, and write $X_n \overset{\text{ucp}}{\to} X$, if

$$(X^n - X)_t^* \to 0 \text{ in probability, for each } t \geq 0.$$

Theorem 19.10 (Quadratic variation of square-integrable martingales)**.** *For $M \in \mathcal{M}_0^{2,c}$, there exists a unique process $\langle M\rangle \in \mathcal{A}_0^c$ such that such that $M_t^2 - \langle M\rangle_t$ is a uniformly-integrable martingale. Furthermore, for each sequence $\{\Delta_n\}_{n\in\mathbb{N}}$ in $P_{[0,\infty)}$ with $\Delta_n \to \mathrm{Id}$, we have*

$$\langle M\rangle^{\Delta_n} \overset{\text{ucp}}{\to} \langle M\rangle,$$

where, for $\Delta = \{t_k\}_{k\in\mathbb{N}} \in P_{[0,\infty)}$, $\langle M\rangle_t^\Delta := \sum_{k=0}^\infty (M_{t_{k+1}\wedge t} - M_{t_k\wedge t})^2$.

The process $\langle M \rangle \in \mathcal{A}_0^c$ is called the **quadratic variation (process)** of M. It admits a natural interpretation as the increasing part in the (continuous-time) Doob-Meyer decomposition of M^2 and plays a central role in stochastic analysis. The reader should probably do Problem **??** at this point. Here is how quadratic variation can be used to construct the stochastic integral:

Given $M \in \mathcal{M}_0^{2,c}$, it follows from the fact that both M and $M^2 - \langle M \rangle$ are a martingales that

$$\begin{aligned}\mathbb{E}[(M_{t_{k+1}} - M_{t_k})^2 | \mathcal{F}_{t_k}] &= \mathbb{E}[M_{t_{k+1}}^2 - M_{t_k}^2 | \mathcal{F}_{t_k}] \\ &= \mathbb{E}[\langle M \rangle_{t_k+1} - \langle M \rangle_{t_k} | \mathcal{F}_{t_k}].\end{aligned}$$

Therefore, the relation (19.7) can be written as

$$\begin{aligned}\mathbb{E}[(H \cdot M)_\infty^2] &= \mathbb{E}\Big[\sum_{n=0}^{\infty} K_k^2 \, \mathbb{E}[\langle M \rangle_{t_{k+1}} - \langle M \rangle_{t_k} | \mathcal{F}_{t_k}]\Big] \\ &= \mathbb{E}\Big[\sum_{n=0}^{\infty} K_k^2 \, \big(\langle M \rangle_{t_{k+1}} - \langle M \rangle_{t_k}\big)\Big] = \mathbb{E}[\int_0^\infty H_t^2 \, d\langle M \rangle_t].\end{aligned} \tag{19.8}$$

Inspired by the last expression in (19.8) above, for a progressively-measurable process H, we define

$$||H||_{\mathbb{L}^2(M)} := \sqrt{\mathbb{E}[\int_0^\infty H_t^2 \, d\langle M \rangle_t]}.$$

It is not hard to check that the family $\mathbb{L}^2(M)$ of all progressively-measurable processes H for which $||H||_{\mathbb{L}^2(M)} < \infty$ forms a vector space, and that $|| \cdot ||_{\mathbb{L}^2(M)}$ is a norm there. We also note that $\mathcal{H}_{simp}^\infty \subseteq \mathbb{L}^2(M)$, for each $M \in \mathcal{M}_0^{2,c}$ (why?). In this, new, notation, the conclusion of (19.8) looks like this:

$$||H \cdot M||_{\mathcal{M}_0^{2,c}} = ||H||_{\mathbb{L}^2(M)}, \; \forall \, H \in \mathcal{H}_{simp}^\infty, \tag{19.9}$$

and is called **Itô's isometry**. We use it to extend the domain of the stochastic integral from $\mathcal{H}_{simp}^\infty$ to a much larger set. To do that, we need a simple, but very powerful, result from real analysis:

Proposition 19.11. *Let X be a metric space and Y a* complete *metric space. Moreover, let $f : A \to Y$, with $\emptyset \neq A \subseteq X$ be a* uniformly-continuous *function. Then, there exists a unique uniformly-continuous function $\tilde{f} : \mathrm{Cl}\, A \to Y$ such that $\tilde{f}(x) = f(x)$, for all $x \in A$.*

Proof. For $x \in \mathrm{Cl}\, A$, let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in A such that $x_n \to x$, and let $y_n = f(x_n)$. Since x_n is convergent in X and f is uniformly continuous, $\{y_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in Y. Since Y is complete it admits a unique limit y, and we set

$$\tilde{f}(x) = y.$$

To show that $\tilde{f}$ is well-defined, we need to argue that y does not depend on the choice of the sequence $\{x_n\}_{n\in\mathbb{N}}$. Indeed, let $\{x'_n\}_{n\in\mathbb{N}}$ be another sequence in A with $x'_n \to x$. Then $d(x_n, x'_n) \to 0$, and, so, by uniform continuty, $d(f(x_n), f(x'_n)) = 0$, for each y. It follows that $y = \lim_n f(x_n) = \lim_n f(x'_n)$.

Next, we argue that that $\tilde{f}$ is uniformly continuous. Given $\varepsilon > 0$, let $\delta > 0$ be such that $d(x, \hat{x}) < \delta$ implies $d(f(x), f(\hat{x})) < \varepsilon$, for all $x, \hat{x} \in A$. For any $x, x' \in \operatorname{Cl} A$ with $d(x, x') < \delta$, we can choose two sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{x'_n\}_{n\in\mathbb{N}}$ in A such that $x_n \to x$ and $x'_n \to x'$ and $d(x_n, x'_n) < \delta$, for all n. It follows that $d(f(x_n), f(x'_n)) < \varepsilon$, for all n, and, so, $d(\tilde{f}(x), \tilde{f}(x')) \leq \varepsilon$, estalishing the uniform continuity of $\tilde{f}$ on $\operatorname{Cl} A$.

Finally, given $x \in A$, we may take $x_n = x$, for all n, to conclude that $\tilde{f}(x) = f(x)$. □

The completeness of the space Y is crucial for the Proposition 19.11 to work. To check that it is applicable in our case, we need the following result:

Proposition 19.12. $(\mathcal{M}_0^{2,c}, ||\cdot||_{\mathcal{M}_0^{2,c}})$ *is a Banach space.*

Proof. Simple properties of square-integrable martingales suffice to show that $\mathcal{M}_0^{2,c}$ is a linear space, and that $||\cdot||_{\mathcal{M}_0^{2,c}}$ satisfies all the axioms of a norm. The more delicate matter is completeness. We start with a Cauchy sequence $\{M^n\}_{n\in\mathbb{N}}$ and conclude immediately that their last elements $\{M^n_\infty\}_{n\in\mathbb{N}}$ form a Cauchy sequence in $\mathbb{L}^2(\mathcal{F})$. Thus, $M^n_\infty \to M_\infty$ in $\mathbb{L}^2$, for some $M_\infty \in \mathbb{L}^2(\mathcal{F})$. We define $\{M\}_{t\in[0,\infty)}$ as the RCLL version of the Lévy martingale $M_t = \mathbb{E}[M_\infty|\mathcal{F}_t]$ and use Doob's inequality to get that

$$\Big|\Big|\sup_{t\geq 0}|M^n_t - M_t|\Big|\Big|_{\mathbb{L}^2} \leq 2||M^n_\infty - M_\infty||_{\mathbb{L}^2} \to 0,$$

Hence, possibly through a subsequence, we have

$$\sup_{t\geq 0}|M^n_t - M_t| \to 0, \text{ a.s.}$$

It follows that almost all trajectories of M are uniform limits of the corresponding trajectories of M^n. By assumption, all M^n are continuous martingales, and uniform limits of continuous functions are continuous, so M is a continuous martingale and $M = \lim_n M^n$ in $\mathcal{M}_0^{2,c}$. □

If we combine the Itô's isometry (19.9) (which yields the uniform continuity of the map $H \mapsto H \cdot M$ on $\mathcal{H}^\infty_{simp}$) with Proposition 19.12 (which asserts that the target space is complete) and use Proposition 19.11 we immediately conclude that the map

$$H \mapsto H \cdot M \in \mathcal{M}_0^{2,c}$$

admits a unique linear and continuous extension of the simple predictable integral from $\mathcal{H}^\infty_{simp}$ to its closure in $\mathbb{L}^2(M)$, with values in $\mathcal{M}^{2,c}_0$. The only question still remaining is to give a nice description of the closure of $\mathcal{H}^\infty_{simp}$ in $\mathbb{L}^2(M)$:

Proposition 19.13. *For any $M \in \mathcal{M}^{2,c}_0$, $\mathcal{H}^\infty_{simp}$ is dense in $\mathbb{L}^2(M)$.*

Proof. Without loss of generality, we assume that $\mathbb{E}[\langle M\rangle] = 1$. Let $\{\Delta_n\}_{n\in\mathbb{N}}$ be a sequence of nested partitions in $P_{[0,\infty)}$ such that $\Delta_n \to \mathrm{Id}$; for concreteness, we take $\Delta_n = \{k2^{-n} : 0 \le k \le n2^n\}$). For each n, let $\mathcal{P}_n$ be the σ-algebra on the product $[0,\infty)\times\Omega$ generated by all simple predictable processes H of the form

$$H_t = \sum_{k=0}^{n2^n} K_k \mathbf{1}_{(k2^{-n},(k+1)2^{-n}]}(t),\ K_k \in \mathcal{F}_{k2^{-n}},\ \text{for all } k. \tag{19.10}$$

In fact, it is not hard to see that a process H is measurable with respect to $\mathcal{P}_n$, precisely when it is of the form (19.10). Also, since all left-continuous processes can be obtained as pointwise limits of the processes of the form (19.10), we have

$$\mathcal{P} = \sigma(\cup_n \mathcal{P}_n),$$

where $\mathcal{P}$ denotes the predictable σ-algebra. The product space $\Omega^* = [0,\infty)\times\Omega$, together with the σ-algebra $\mathcal{P}$, the filtration $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$, and the probability measure $\mathbb{P}^*$, given by

$$\mathbb{P}^*[A] = \mathbb{E}[\int_0^\infty \mathbf{1}_A(t,\omega)\, d\langle M\rangle_t(\omega)],\ \text{for } A \in \mathcal{P},$$

forms a filtered probability space. Moreover, random variables on this space correspond precisely to the predictable processes on the original space, and their expectations to the expected $d\langle M\rangle$-integrals, i.e., if $\mathbb{E}^*$ denotes the expectation under $\mathbb{P}^*$, we have

$$\mathbb{E}^*[H] = \mathbb{E}[\int_0^T H_t\, d\langle M\rangle_t] \text{ where .}$$

In particular, we have

$$H_n \to H \text{ in } \mathbb{L}^2(\mathbb{P}^*) \text{ iff } H_n \to H \text{ in } \mathbb{L}^2(M).$$

Given $\tilde{H} \in \mathbb{L}^2(\mathbb{P}^*)$, define

$$\tilde{H}_n = \mathbb{E}^*[\tilde{H}|\mathcal{P}_n].$$

This is a square-integrable martingale and $\tilde{H} \in \mathcal{P}$, so $\tilde{H}^n \to \tilde{H}$ in $\mathbb{L}^2(\mathbb{P}^*)$. Therefore, by the above characterization of convergence in $\mathbb{L}^2(M)$, we have $\tilde{H}^n \to \tilde{H}$ in $\mathbb{L}^2(M)$, as required. □

We can now summarize all our findings in a compact statement:

Theorem 19.14. *Given $M \in \mathcal{M}_0^{2,c}$, there exist a unique linear isometry*

$$\mathbb{L}^2(M) \ni H \mapsto H \cdot M \in \mathcal{M}_0^{2,c}$$

which extends the simple predictable integral (19.6).