Introduction to stochastic integration: Math 642:592, Spring 2008

I. Bounded variation functions and Lebesgue-Stieltjes integrals.

As a preliminary to the theory of stochastic integration, we recall the theory of Lebesgue-Stieltes integrals and its relation to bounded variation.

Let $G : [0, \infty) \to \mathbb{R}$ be an increasing, right-continuous function. Recall that there is a unique measure μ_G on the Borel subsets of $(0, \infty)$ such that $\mu_G((a, b]) = G(b) - G(a)$ whenever $0 < a < b$. This measure μ_G is constructed in the same way as Lebesgue measure. The requirement, $\mu_G((a, b]) = G(b) - G(a)$, defines a finitely additive measure on the algebra of subsets which are finite disjoint unions of intervals of the form $(a, b]$, $0 < a < b \leq \infty$. Then this finitely additive measure is extended to the Borel sets by Caratheodory's extension theorem. Now having the measure μ_G, we can consider the integral, $\int_{(0,\infty)} \psi(s)\, \mu_G(\mathrm{d}s)$. We prefer to write this integral with the more suggestive notation,

$$\int_{(0,\infty)} \psi(s)\, \mathrm{d}G(s).$$

If ψ is the *simple* function, $\psi(t) = \sum \psi_i \mathbf{1}_{(t_i, t_{i+1}]}(t)$,

$$\int_{(0,\infty)} \psi(s)\, \mathrm{d}G(s) = \sum_i \psi_i \left[G(t_{i+1}) - G(t_i)\right]. \tag{1}$$

Let $\Pi_n = \left\{t_i^{(n)}\right\}$ be a sequence of partitions of $[0, \infty)$ with $0 = t_0^{n)} < t_1^{(n)} < t_2^{(n)} < \cdots$ and $\lim_{i\to\infty} t_i^{(n)} = \infty$. Setting, $\|\Pi_n\| \triangleq \sup_i (t_{i+1}^{(n)} - t_i^{(n)})$, suppose that $\|\Pi_n\| \to 0$ as $n \to \infty$. When ψ is a locally bounded (bounded on each compact set) function that has left limits, then

$$\begin{aligned}
\int_{(0,t)} \psi(s-)\, \mathrm{d}G(s) &= \lim_{n\to\infty} \int_{(0,t]} \sum_i \psi(t_i^{(n)}) \mathbf{1}_{(t_i^{(n)}, t_{i+1}^{(n)}]}(s)\, \mathrm{d}G(s) \\
&= \lim_{n\to\infty} \sum_i \psi(t_i^{(n)}) \left[G(t_{i+1}^{(n)} \wedge t) - G(t_i^{(n)} \wedge t)\right].
\end{aligned} \tag{2}$$

Let $f : [0, \infty) \to I\!R$. Then the total variation of f on $[0, t]$ is

$$TV(f)(t) \triangleq \sup \left\{ \sum_i \mid f(t_{i+1}) - f(t_i) \mid ;\ 0 = t_0 < t_1 < \cdots < t_N = t \ \text{ is a partition of } [0, t]. \right\}$$

One says that f is a function of bounded variation if $TV(f)(t) < \infty$ for all t. When f is a bounded variation function, $F_1 \triangleq (TV(f) + f)/2$ and $F_2 \triangleq (TV(f) - f)/2$ define two increasing functions such that

$$f = F_1 - F_2 \ \text{ and the measures } \mu_{F_1} \text{ and } \mu_{F_2} \text{ are mutually singular.} \tag{3}$$

In fact the decomposition of (3) is unique up to an additive constant. When f is right continuous, so is $TV(f)$, hence so are F_1 and F_2, and we can define,

$$\int_{(0,\infty)} \psi(s)\, \mathrm{d}f(s) \triangleq \int_{(0,\infty)} \psi(s)\, \mathrm{d}F_1(s) - \int_{(0,\infty)} \psi(s)\, \mathrm{d}F_2(s) \tag{4}$$

The following summation formulas will be important to the discussion of stochastic integration and give us some immediate insight into issues with defining integrals against integrators f of unbounded variation. The formulas are derived by straightforward algebraic manipulation. The first is is the sum analogue of integration by parts:

$$\sum_{i=0}^{\infty} \psi(t_i^{(n)} \wedge t) \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right] = f(t)\psi(t) - f(0)\psi(0) - \sum_{i=0} f(t_{i+1}^{(n)} \wedge t) \left[\psi(t_{i+1}^{(n)} \wedge t) - \psi(t_i^{(n)} \wedge t) \right] \quad (5)$$

The second is an alternative expression for the first sum when $\psi = f$:

$$\sum_{i=0}^{\infty} f(t_i^{(n)} \wedge t) \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right] = \frac{1}{2} f^2(t) - \frac{1}{2} f^2(0) - \frac{1}{2} \sum_{i=0}^{\infty} \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right]^2. \quad (6)$$

Let $f : [0, \infty) \to \mathbb{R}$ be general, right-continuous function. The point of departure for defining integrals of the form $\int_{(0,\infty)} \psi(s) \, \mathrm{d}f(s)$ is the formula we encountered in (1) when ψ is a simple function, taken as a *definition*:

$$\int_{(0,\infty)} \sum_i \psi_i \mathbf{1}_{(t_i, t_{i+1}]}(s) \, \mathrm{d}f(s) \triangleq \sum_i \psi_i [f(t_{i+1}) - f(t_i)]. \quad (7)$$

This definition does not lead to any ambiguity because, when f is a function of bounded variation, it coincides with the Lebesgue-Stieltjes integral of ψ. However, if f is not of bounded variation, there is not in general a definition of a measure and an integral associated to f that extends to all Borel measurable sets or integrands. For such f we have to narrow the class of integrands to have a successful theory. While it may seem perverse to want to handle such ill-behaved integrators as unbounded functions, non-trivial, path-continuous martingales have paths of unbounded variation quite generally.

Integration theory for stochastic integrators with unbounded variation, such as Brownian motion, will take advantage of probabilistic features, such as martingale properties. But before moving on, it will be useful to make a few remarks about defining integrals more general than (7) for deterministic or stochastic f of unbounded variation, by way of limits. For example, suppose that f is continuous and locally bounded and that ψ is right-continuous and of bounded variation. Then from (5) and the theory for bounded variation integrators,

$$\lim_{n \to \infty} \sum_{i=0}^{\infty} \psi(t_i^{(n)} \wedge t) \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right] = f(t)\psi(t) - f(0)\psi(0) - \int_{(0,t]} f(s) \, \mathrm{d}\psi(s) \quad (8)$$

We take the right-hand side of (8) as a *definition* of the integral $\int_{(0,t]} \psi(s) \, \mathrm{d}f(s)$. This definition will not apply to $\int_0^t f(s) \, \mathrm{d}f(s)$, when f is càdlàg, but of unbounded variation. However, in this case it is still possible that $V(t) = \lim_{n \to \infty} \sum_i \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right]^2$ exists. (We shall see that such a limit exists in probability when f is a martingale.) Then, refering to (6),

$$\int_0^t f(s-)\,\mathrm{d}f(s) \triangleq \lim_{n\to\infty} \sum_{i=0}^{\infty} f(t_i^{(n)} \wedge t) \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right] = \frac{1}{2}[f^2(t) - f^2(0)] - \frac{1}{2}V(t). \quad (9)$$

(We have written the integrand as $f(s-)$ because $f(s-)$ is the point-wise limit of $\sum_i f(t_i^{(n)})\mathbf{1}_{(t_i^{(n)}, t_{i+1}^{(n)}]}(s)$, as $n \to \infty$.)

Definitions (8) and (9) apply to a limited class of integrands. However the ideas contained in these definitions, especially (9), will be important for defining stochastic integrals.

II. Quadratic variation of Brownian motion and of Poisson processes.

Throughout this section $\{\Pi_n\}$ is a sequence of partitions $0 = t_0^{(n)} < t_1^{(n)} < \cdots$ of $[0, \infty)$ with $\lim_{i\to\infty} t_i^{(n)} = \infty$, for every n, and $\|\Pi_n\| \to 0$ as $n \to \infty$. The notation Π refers just to an arbitrary partition $0 = t_0 < t_1 < \cdots$ of $[0, \infty)$.

Let X be a stochastic process. For a partition Π, define

$$V(\Pi, X)(t) = \sum_{i=0}^{\infty} \left[X(t_{i+1} \wedge t) - X(t_i \wedge t) \right]^2.$$

We shall be interested in limits of $V(\Pi, X)(t)$ as $\|\Pi\| \to 0$. Such a limit, if it exists, is called the quadratic variation of X, although it will be convenient to give the precise definition of quadratic variation in a different way later on. It will turn out that martingales have quadratic variation processes and this fact and it consequences are central to stochastic integration theory, as one can begin to see by studying the special integral defined in equation (9). In this section, we will first compute quadratic variations for Brownian motions and compound Poisson processes.

Theorem 1 *Let W be a standard Brownian motion. Then for every $T > 0$,*

$$\sup_{[0,T]} \left| V(\Pi, W)(t) - t \right| \to 0 \quad \text{in probability as } \|\Pi\| \to 0 \text{ for every } T > 0. \quad (10)$$

Proof: What (10) says is that for every $c > 0$ and $\epsilon > 0$ there is a $\delta > 0$ such that if $\|\Pi\| < \delta$, then $I\!P\left(\sup_{[0,T]} \left| V(\Pi, W)(t) - t \right| > c) < \right) \epsilon$.

In this proof we will use the fact that if Z is a normal random variable with mean 0 and variance σ^2,

$$\mathbb{E}\left[(Z^2 - \sigma^2)^2 \right] = 2\sigma^4 \quad (11)$$

This is proved easily by a direct computation By Chebyshev's inequality, to show convergence in probability, it suffices to show that

$$\mathbb{E}\left[\sup_{[0,T]} \left| V(\Pi, W)(t) - t \right|^2 \right] \to 0 \quad \text{as } \|\Pi\| \to 0. \quad (12)$$

Observe that

$$V(\Pi, W)(t) - t = \sum_{i=0}^{\infty} [W(t_{i+1} \wedge T) - W(t_i \wedge T)]^2 - [t_{i+1} \wedge T - t_i \wedge T]. \tag{13}$$

Recall that $W^2(t) - t$ is a martingale. The same elementary argument used to show this, shows also that $[W(t_{i+1} \wedge t) - W(t_i \wedge t)]^2 - [t_{i+1} \wedge t - t_i \wedge t]$ is a martingale for any fixed $t_i < t_{i+1}$. It follows that $V(\Pi, W)(t) - t$, being a sum of these terms, is also a martingale, and it has continuous paths. Therefore, by Doob's inequality,

$$\mathbb{E}\left[\sup_{[0,T]} \left| V(\Pi, W)(t) - t \right|^2\right] \leq 4\mathbb{E}\left[\left| V(\Pi, W)(T) - T \right|^2\right].$$

But, since (13) is a sum of independent, zero mean terms, it follows, using (11), that

$$\begin{aligned}
\mathbb{E}\left[\left| V(\Pi, W)(T) - T \right|^2\right] &= \sum_i \mathbb{E}\left[\left([W(t_{i+1} \wedge T) - W(t_i \wedge T)]^2 - [t_{i+1} \wedge T - t_i \wedge T]\right)^2\right] \\
&= \sum_i 2[t_{i+1} \wedge T - t_i \wedge T]^2 \\
&\leq 2\|\Pi\| T
\end{aligned}$$

Thus, (12) holds and the proof is complete. ◇

Exercise 1. Use the Borel-Cantelli lemma to show that if $\{\Pi_n\}$ is the sequence of partitions defined by letting $\Pi_n = \{k2^{-n}\,;\, k \geq 0\}$, then $\sup_{[0,T]} \left| V(\Pi, W)(t) - t \right| \to 0$ almost-surely, as $n \to \infty$.

An immediate consequence of the last theorem is that Brownian paths are a.s. of unbounded variation on any interval of positive length. To state this more precisely, let $TV(W)(s,t)(\omega)$ be the variation of the path $W(\cdot)(\omega)$ over the interval $[s,t]$.

Corollary 1 *Let W be Brownian motion. Then*

$$I\!P\,(\exists\; 0 \leq s < t \text{ such that } TV(W)(s,t) < \infty) = 0.$$

Exercise 2. Prove this result.

Because of Corollary 1, we cannot define $\int_0^t X(s)\,\mathrm{d}B(s)$ path-wise using Lebesgue-Stieltjes integration theory. Nevertheless, we will be able to develop a theory in which such integrals are defined by limiting procedures for a large class of integrands X. In this theory, the quadratic variation process will play a central role.

Let $N(t)$ be a Poisson process with rate λ. Let $\{\xi_i\}$ be i.i.d. random variables with finite mean μ and variance σ^2. Consider the Lévy processes

$$Z(t) \triangleq \sum_{1}^{N(t)} \xi_i, \qquad Y(t) \triangleq \sum_{1}^{N(t)} \xi_i - \mu\lambda t.$$

A simple calculation verifies that Y has zero mean and variance $\mathbb{E}\left[Y^2(t)\right] = \lambda(\sigma^2 + \mu^2)t$.

Exercise 2. Show $Y(t)$ and $Y^2(t) - \lambda(\sigma^2 + \mu^2)t$ are martingales.

With Brownian motion, we saw that $W^2(t) - t$ is a martingale and that t is the quadratic variation of Brownian motion. With reference to Exercies 2, One might expect then that the quadratic variation process for Y is $\lambda(\sigma^2 + \mu^2)t$. However the fact that Y is a sum of jumps plus a nice, differentiable, even linear function of t implies the following.

Theorem 2

$$V(Y)(t) \triangleq \lim_{n\to\infty} V(\Pi_n, Y)(t) = \lim_{n\to\infty} [Y(t_{i+1}^{(n)} \wedge t) - Y(t_i^{(n)} \wedge t)]^2 = \sum_1^{N(t)} \xi_i^2, \quad \textit{almost surely.}$$

Notice that this is also the quadratic variation of $Z(t)$, which moves purely by jumps:

$$\lim_{n\to\infty} [Z(t_{i+1}^{(n)} \wedge t) - Z(t_i^{(n)} \wedge t)]^2 = \sum_1^{N(t)} \xi_i^2, \quad \text{almost surely.}$$

Exercise 3. Prove this last formula and use it to probe Theorem 2. (The second implies the first.)

We saw that the square of a Brownian motion minus its quadratic variation is a martingale. This is true for Y as well:

$$Y^2(t) - V(Y)(t) = Y^2(t) - \sum_1^{N(t)} \xi_i^2 \ \text{ is a martingale.}$$

To see this observe that

$$Y^2(t) - \sum_1^{N(t)} \xi_i^2 = Y^2(t) - \lambda(\sigma^2 + \mu^2)t - \left[\sum_1^{N(t)} \xi_i^2 - \lambda(\sigma^2 + \mu^2)t\right],$$

and both terms are martingales, the second because $\mathbb{E}\left[\xi_i^2\right] = \sigma^2 + \mu^2$.

III. Stochastic integration: simple integrands and predictable processes

Stochastic integral theory is about defining integrals of the form

$$\int_0^t H(s)\,\mathrm{d}X(s)$$

These integrals will be extensions of the natural definition of integral for simple integrands. A stochastic process H is called *simple* if there is a sequence of random or deterministic times $0 = T_0 < T_1 < T_2 < \cdots$ with $\lim_{i\to\infty} T_i(\omega) = \infty$ for all ω such that

$$H(s) = \sum_{i=0}^{\infty} \eta_i \mathbf{1}_{(T_i, T_{i+1}]}(s),$$

Then we define,

$$\int_0^t H(s)\,\mathrm{d}X(s) \triangleq \sum_{i=0}^{\infty} \eta_i \left[X(T_{i+1}\wedge t) - X(T_i \wedge t)\right]. \tag{14}$$

This is defined for any process X. The goal of stochastic integration theory is, for a given random integrator X, to extend the integral to more general classes of integrands Y, in a way that is mathematically natural, mathematically useful, and consistent with having limit theorems such as dominated convergence. For many a given integrator X, it is not possible to do this for even so restricted a class of integrands as bounded measurable processes.

Out first step in this introductory discussion will be to restrict our ambitions greatly, but in a way guided by potential applications. We will pose the problem of constructing stochastic integrals in a martingale framework. Recall that if $\{X_n\}$ is a discrete-time martingale with respect to some filtration $\{\mathcal{F}_n\}$ and if $\{H_n\}$ is a bounded, $\{\mathcal{F}_n\}$-predictable process, which means that H_n is $\mathcal{F}_{n-1}$ measurable for all n, then the discrete-time stochastic integral,

$$(H.X)_n = \sum_{i=1}^{n} H_i(X_i - X_{i-1})$$

is a martingale with respect to $\{\mathcal{F}_n\}$ also. This fact was fundamental to the development of discrete-time, martingale theory. Also $(H.X)_n$ has a nice intuitive interpretation; we think of X_n as the current fortune in a fair game in which a dollar is bet on every play, $X_i - X_{i-1}$ being the gain or loss on play i. Then $(H.X)_n$ is the fortune after n plays when h_i is the amount bet on play i for each i. The predictability of $\{H_n\}$ is a condition excluding clairvoyance, which explains intuitively that $H.X$ must be a martingale as well. The first stochastic integration theory will be a continuous time extension in which X is a continuous time, càdlàg martingale with respect to a filtration $\mathbb{F}$, and the integrands H are restricted so that $\int_0^t H(s)\,\mathrm{d}X(s)$ are martingales. The restriction on integrands is again called *predictability*. In the remainder of this section we develop these ideas.

In what follows, X is a càdlàg martingale with respect to a filtration $\mathbb{F}$. For a deterministic partition $0 = t_0 < t_1 < t_2 < \cdots$, let H be a process of the form

$$H(t) = \sum_{i=0}^{\infty} \eta_i \mathbf{1}_{(t_i, t_{i+1}]}(t) \quad \text{where } \eta_i \text{ is } \mathcal{F}_{t_i} \text{ measurable for each } i. \tag{15}$$

Notice that such an H is a left-continuous, $\mathbb{F}$-progressively measurable process. Then, by definition (14),

$$(H.X)_t \triangleq \int_0^t H(s)\,\mathrm{d}X(s) = \sum_{i=1}^{\infty} \eta_i \left[X(t_{i+1}\wedge t) - X(t_i \wedge t)\right].$$

If each h_i is bounded, then this will be a càdlàg, $\mathbb{F}$-martingale. The argument is practically the same as for the discrete-time integral. It is only necessary to show that for each i,

$$\eta_i \left[X(t_{i+1}\wedge t) - X(t_i \wedge t)\right] = \begin{cases} 0, & \text{if } t < t_i; \\ \eta_i \left[X(t_{i+1}\wedge t) - X(t_i)\right], & \text{if } t \geq t_i, \end{cases} \tag{16}$$

is a martingale. But for any $t_i \le s < t$,

$$\mathbb{E}\left[\eta_i\left[X(t_{i+1}\wedge t) - X(t_i)\right] \,\middle|\, \mathcal{F}_s\right] = \eta_i \mathbb{E}\left[X(t_{i+1}\wedge t) \,\middle|\, \mathcal{G}_s\right] - \eta_i X(t_i) = \eta_i\left[X(t_{i+1}\wedge s) - X(t_i)\right]$$

since X is a martingale. And if $s < t_i \le t$.

$$\mathbb{E}\left[\eta_i\left[X(t_{i+1}\wedge t) - X(t_i)\right] \,\middle|\, \mathcal{F}_s\right] = \mathbb{E}\left[\mathbb{E}\left[\eta_i\left[X(t_{i+1}\wedge t) - X(t_i)\right] \,\middle|\, \mathcal{F}_{t_i}\right] \,\middle|\, \mathcal{F}_s\right] = 0.$$

This verifies (16).

We shall certainly want the integral to be defined for limits of integrands of the form (15). This leads to a definition of predictable processes in the continuous-time setting. Define the predictable σ-algebra $\mathcal{P}$ to be the smallest σ-algebra of subsets of $(0,\infty)\times\Omega$ such that the map

$$(t,\omega) \in (0,\infty)\times\Omega \to H(t,\omega)$$

is $\mathcal{P}$-measurable for every process H of the form (15).

Definition. A process $\{G(t); t > 0\}$ is called $\mathbb{F}$-predictable ($\mathbb{F}$-previsible is the terminology of RW) if the map $(t,\omega)\in(0,\infty)\times\Omega \to G(t,\omega)$ is $\mathcal{P}$-measurable. A process $\{G(t); t \ge 0\}$ is called $\mathbb{F}$-predictable $g(0)$ is $\mathcal{F}_0$-measurable and the process restricted to $t > 0$ is $\mathbb{F}$-predictable.

(In RW, predictable process are defined as processes on the time interval $(0,\infty)$. We allow the process to be defined at 0 since we often encounter this case. From the point of integration it does not matter much because if G is an integrand of a stochastic integral, the value of G at $t = 0$ does not affect the value of the integral.)

Lemma 1 *Every left-continuous, $\mathbb{F}$-adapted process is $\mathbb{F}$-predictable.*

The proof is left as an *Exercise 4*. The usual definition of $\mathcal{P}$ is that it is the smallest σ-algebra with respect to which every left-continuous, $\mathbb{F}$-adapted process is measurable.

Lemma 2 *The predictable σ-algebra $\mathcal{P}$ is generated by the set of events of the form*

$$(s,\infty)\times A \qquad s \ge 0, \quad A \in \mathcal{F}_s.$$

The proof is left as *Exercise 5.*

Lemma 3 *Suppose that $0 = T_0 < T_1 < T_2 < \cdots$ is a sequence of $\mathbb{F}$ stopping times. If for each i, η_i is $\mathcal{F}_{T_i}$ measurable, then*

$$H(t) = \sum_{i=0}^{\infty} \eta_i \mathbf{1}_{(T_i,T_{i+1}]}(t)$$

is predictable.

Proof: Each term $\eta_i \mathbf{1}_{(T_i,T_{i+1}]}(t)$ is left-continuous. If B is a Borel set which does not contain 0, then for each fixed t,

$$\{\eta_i \mathbf{1}_{(T_i,T_{i+1}]}(t) \in B\} = \{\eta_i \in B\} \cap \{T_i < t\} \cap \{T_{i+1} \geq t\}$$

Since η_i is $\mathcal{F}_{T_i}$-measurable, $\{\eta_i \in B\} \cap \{T_i < t\} \in \mathcal{F}_t$. But also $\{T_{i+1} \geq t\} \in \mathcal{F}_t$. Hence, $\{\eta_i \mathbf{1}_{(T_i,T_{i+1}]}(t) \in B\} \in \mathcal{F}_t$, also. If B contains 0,

$$\{\eta_i \mathbf{1}_{(T_i,T_{i+1}]}(t) \in B\} = \{\eta_i \mathbf{1}_{(T_i,T_{i+1}]}(t) \in B - \{0\}\} \cup \{t \leq T_i, t > T_{i+1}\},$$

and this again belongs to $\mathcal{F}_t$ because $\{t \leq T_i, t > T_{i+1}\}$ does and $B - \{0\}$ does not contain 0. $\diamond$

From the monotone class theorem, we have the following lemma; it's a minor variant of the one stated in RW, Volume II, Section IV.6, page 13.

Lemma 4 *Let $\mathcal{H}$ be a set of bounded process that is a vector space and contains constant functions. Suppose also that*

(a) If $\{h_n\}$ is a sequence of processes in $\mathcal{H}$ converging uniformly to h, then h in also in $\mathcal{H}$;

(b) If $\{h_n\}$ is a sequence of processes in $\mathcal{H}$, satisfying $0 \leq h_1 \leq h_2 \leq \cdots \leq K$, for some constant $K < \infty$, then $h(t,\omega) = \lim_{n\to\infty} h_n(t,\omega)$ defines a process in $\mathcal{H}$.

Then if $\mathcal{H}$ contains every process of the form $\eta \mathbf{1}_{(s,r]}(t)$ where $\eta \in \mathcal{F}_s$, then $\mathcal{H}$ contains all bounded $\mathbb{F}$-predictable processes.

Here is a consequence that lends some intuition into predictability.

Corollary 2 *(a) If H is a predictable process then for every $t > 0$, $H(t)(\cdot)$ is $\mathcal{F}_{t-} = \sigma\{\cup_{s<t}\mathcal{F}_s\}$-measurable.*

The proof is left as *Exercise 6.* Let $\mathcal{H}$ be the class of bounded processes satisfying the condition of the Corollary, to prove the Corollary for all bounded predictable H. A general predictable process may be realized as a limit of bounded predictable processes.

Example 1. Let W be an $\mathbb{F}$-Brownian motion. Then W is $\mathbb{F}$-predictable, since its paths are continuous.

Example 2. Let N be a Poisson process and let $\mathbb{F}$ be the filtration generated by N and completed by the sets of probability 0. We shall see that N is not $\mathbb{F}$-predictable. However, for each t, $N(t) = N(t-)$ almost-surely and so $N(t)$ is $\mathcal{F}_{t-}$-measurable for each $t > 0$. Thus the converse of Corollary 1 does not hold. Note that $\{N(t-)\}$ is certainly a predictable process.

Recall that for simple integrands h, we *define* the stochastic integral $\int_0^t h(s)\, \mathrm{d}X(s)$ by (14). Here is the essential point of the discussion so far.

Theorem 3 *If X is a martingale and if H is a simple, predictable process then $\int_0^t h(s)\, \mathrm{d}X(s)$ is a martingale.*

We proved this when the times T_i in (14) are deterministic. We leave the generalization as *Exercise 7.* Hint: Recall optional stopping; this generalizes to continuous time for right-continuous martingales—see RW, II.5.

The theory in the martingale case will extend the integral to more general predictable processes while preserving the property that the integral be a martingale.

IV. Stochastic integration for martingales beyond simple integrands; examples.

Let X be a càdlàg process. Let $\{\Pi_n\}$ be a sequence of partitions with $\|\Pi_n\| \to 0$ as $n \to \infty$. Let

$$X^{(n)}(t) \triangleq \sum_{i=0}^{\infty} X(t_i^{(n)}) \mathbf{1}_{(t_i^{(n)}, t_{i+1}^{(n)}]}(t).$$

For each (t, ω), $t > 0$, $\lim_{n\to\infty} X^{(n)}(t)(\omega) = X(t-)(\omega)$. Notice that when X is $\mathbb{F}$-adapted, $\{X(t-); t > 0\}$ is $\mathbb{F}$-predictable.

In this section we will make the following ad-hoc definition.

$$\int_0^t X(s-)\, \mathrm{d}X(s) \triangleq \lim_{n\to\infty} \int_0^t X^{(n)}(s)\, \mathrm{d}X(s), \tag{17}$$

if the limit exists *in probability* and is independent of the sequence of partitions so long as $\|\Pi_n\| \to 0$ as $n \to \infty$. We will compute this limit for three case: (a) X is a deterministic, continuous, bounded variation function f; (b) X is a standard Brownian motion W; (c) X is the martingale $Y(t) = \sum_1^{N(t)} \xi_i - \mu\lambda t$ defined in section II above. Examples (b) and (c) are fundamental in the theory of stochastic integration with respect to a martingale. By comparing (b) and (c) to (a) we will uncover the essential way in which Itô calculus, the calculus of stochastic integrals with respect to martingales, differs from ordinary calculus.

Case (a): X is a deterministic, continuous, bounded variation function f.

In this case, noting that $f(s) = f(s-)$ for all s, if we define $\int_0^t f(s)\, \mathrm{d}f(s)$ by (17) it is just the Lebesgue-Stieltes as defined above in Section I of this lecture. We can compute it in closed form if we refer to the identity (6) or (9). Note that for any n

$$\begin{aligned} \sum_{i=0}^{\infty} \left[f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \right]^2 &\leq \max_j \mid f(t_{j+1}^{(n)} \wedge t) - f(t_j^{(n)} \wedge t) \mid \sum_{i=0}^{\infty} \mid f(t_{i+1}^{(n)} \wedge t) - f(t_i^{(n)} \wedge t) \mid \\ &\leq \max_j \mid f(t_{j+1}^{(n)} \wedge t) - f(t_j^{(n)} \wedge t) \mid TV(f)(t) \end{aligned}$$

By continuity of f and the assumption that $TV(f)(t) < \infty$, this tends to 0 as $n \to \infty$. Therefore, taking limits in (6).

$$\int_0^t f(s)\, \mathrm{d}f(s) = \lim_{n\to\infty} \int_0^t f^{(n)}(s) \mathrm{d}f(s) = \frac{1}{2}[f^2(t) - f^2(0)]. \tag{18}$$

In the formal language of differentials, this implies

$$\mathrm{d}f^2(t) = 2f(t)\, \mathrm{d}f(t), \tag{19}$$

which, when f is differentiable, reduces to the usual chain rule formula $\mathrm{d}f^2(t)/\mathrm{d}t = 2f(t)f'(t)$.

It is important to remark in this example that there is actually considerable flexibility in the choice of the approximating sequence. We would get the same limit if instead we replaced $f^{(n)}$ with

$$\widetilde{f}^{(n)} = \sum_{i=0}^{\infty} f(c_i^n)\mathbf{1}_{(t_i^{(n)},t_{i+1}^{(n)}]}(t),$$

where c_i^n is an arbitrary point of $[t_i^{(n)}, t_{i+1}^{(n)}]$ for every n and i. The student should prove this as *Exercise 8.*

Case (b): X is standard, Brownian motion W.

Let W be an $\mathbb{F}$-Brownian motion. In this case, $W^{(n)}(t) = \sum_{i=0}^{\infty} W(t_i^{(n)})\mathbf{1}_{(t_i^{(n)},t_{i+1}^{(n)}]}(t)$ is $\mathbb{F}$-predictable. It follows, by the same argument as at (16), which has to be generalized only slightly because W is not uniformly bounded, that $\int_0^t W^{(n)}(s)\,\mathrm{d}W(s)$ is a martingale for all n, in conformance with the idea of defining new martingales by stochastic integration. Moreover, we know from Theorem 1 that

$$\sum_{i=0}^{\infty}\left[W(t_{i+1}^{(n)}\wedge t) - W(t_i^{(n)}\wedge t)\right]^2 \quad \text{converges in probability to } t\text{, uniformly on compacts.}$$

Therefore by taking limits in (6) with f replace by W, $\int_0^t W^{(n)}(s)\,\mathrm{d}W(s)$ converges in probability to $(1/2)W^2(t)$. Hence,

$$\int_0^t W(s)\,\mathrm{d}W(s) = \frac{1}{2}[W^2(t) - t], \tag{20}$$

or, in differential notation, which we must stick to this time because the paths of W are almost surely nowhere differentiable,

$$\mathrm{d}W^2(t) = 2W(t)\,\mathrm{d}W(t) + \mathrm{d}t. \tag{21}$$

Here we see the first manifestation of Itô's rule; the fact that the quadratic variation of W is not zero leads to a different chain rule than that of ordinary calculus, as in (20), for Lebesgue-Stieltjes integrals.

Notice also that the integral $\int_0^t W(s)\,\mathrm{d}W(s)$ as calculated in (20) is indeed a martingale, as we want.

In this example, the predictability of the integrand is important and there is not the same flexibility as with case (a) in defining the approximating sequence. Suppose that we used insead,

$$\widetilde{W}^{(n)}(s) = \sum_{i=0}^{\infty} W(c_i^n)\mathbf{1}_{(t_i^{(n)},t_{i+1}^{(n)}]}(t)$$

where at $t_i^{(n)} < c_i^n < t_{i+1}^{(n)}$ for all n and i. Then (and the reader should verify these statements), $\widetilde{W}^{(n)}$ is no longer predictable and $\int_0^t \widetilde{W}^{(n)}(s)\,\mathrm{d}W(s)$ is not a martingale for any n. Furthermore, for a particular choice of $\{c_i^n; i \geq 0,\, n \geq 0\}$ the limit in probability of $\int_0^t \widetilde{W}^{(n)}(s)\,\mathrm{d}W(s)$ as $n \to \infty$ may exist, but this limit may not equal $(1/2)[W^2(t) - t]$, and will depend on the particular choice. In connection with this point the reader should do the following very important exercise, computing the limit when c_i^n is always chosen as the midpoint of $[t_i^{(n)}, t_{i+1}^{(n)}]$.

Exercise 9. $\lim_{n\to\infty} \sum_{i=0}^{\infty} W\left((t_i^{(n)} + t_{i+1}^{(n)})/2\right)[W(t_{i+1} \wedge t) - W(t_i \wedge t)] = (1/2)W^2(t)$, the limit being taken in the sense of probability as $n \to \infty$.

Case (c): X is a compensated compound Poisson process.

In this case, we replace X by $Y(t) = Z(t) - \mu\lambda t$, where Z is the compound process defined above in section II. That is $Z(t) = \sum_1^{N(t)} \xi_i$, where N is a Poisson process with rate λ and $\{\xi_i\}$ is a sequence of i.i.d. random variables with mean μ and variance σ^2. If we apply our computation of the quadratic variation process of Y from section II in (6), with f replaced by Y, we get

$$\int_0^t Y(s-)\,\mathrm{d}Y(s) = \frac{1}{2}\left[Y^2(t) - V(Y)(t)\right] = \frac{1}{2}\left[Y^2(t) - \sum_1^{N(t)} \xi_i^2\right] \tag{22}$$

We can express this more generically by introducing the notation $\triangle_s Y = Y(s) - Y(s-)$. Notice that $\triangle_s Y(\omega) \neq 0$ only when $s = T_i(\omega)$ for one of the arrival times T_i of N, and in this case $\triangle_s Y(\omega) = \xi_(\omega) = \triangle_s Z$. Thus, $\sum_1^{N(t)} \xi_i^2 = \sum_{s\leq t} (\triangle_s Y)^2$, and

$$\int_0^t Y(s-)\,\mathrm{d}Y(s) = \frac{1}{2}\left[Y^2(t) - \sum_{s\leq t}(\triangle_s Y)^2\right] \tag{23}$$

In the language of differentials,

$$dY^2(t) = 2Y(t-)\,\mathrm{d}Y(t) + dV(Y)(t), \quad \text{where } V(Y)(t) = \textstyle\sum_{s\leq t}(\triangle_s Y)^2.$$

Note in this case that the paths of Y are of finite variation, since Y is the sum of a compound Poisson process, which pathwise contains a finite number of jumps in any bounded interval of time, and a linear function of time. Thus,

$$\int_0^t Y(s-)\,\mathrm{d}Y(s) = \int_0^t Y(s-)\,\mathrm{d}Z(s) - \int_0^t Y(s-)\lambda\mu\,\mathrm{d}s = \sum_{s\leq t} Y(s-)\triangle_s Y(s) - \int_0^t Y(s-)\lambda\mu\,\mathrm{d}s.$$

Since the paths of Y are of finite variation, there is no problem in defining $\int_0^t H(s)\,\mathrm{d}Y(s)$ for any measurable process H, by the Lebesgue-Stieltjes integral. However we will not in general obtain a martingale as a result. For example,

$$\begin{aligned}\int_0^t Y(s)\,\mathrm{d}Y(s) &= \int_0^t Y(s-)\,\mathrm{d}Y(s) + \int_0^t \triangle_s Y\,\mathrm{d}Y(s) \\ &= \frac{1}{2}\left[Y^2(t) - \sum_{s\le t}(\triangle_s Y)^2\right] + \sum_{s\le t}(\triangle_s Y)^2 \\ &= \frac{1}{2}\left[Y^2(t) + \sum_{s\le t}(\triangle_s Y)^2\right],\end{aligned}$$

and this is not a martingale.

V. Square-integrable martingales.

Let M be a càdlàg martingale with respect to a filtration $\mathbb{F}$. (This We shall assume in this section that $\mathbb{F}$ satisfies the usual conditions. This is not needed for every step, but it's easiest to just make it a blanket hypothesis. The stochastic integral $\int_0^t H(s)\,\mathrm{d}X(s)$ for more general predictable processes will be defined as limits of stochastic integrals of simple processes which approximate H. We will first establish this theory using limits in mean square. For this reason we need to develop some facts about square-integrable martingales.

Let $\mathcal{M}_0^2$ denote the set of all càdlàg, $\mathbb{F}$-martingales M (all on a fixed probability space), such that $\sup_t \mathbb{E}\left[M_t^2\right] = \lim_{t\to\infty}\mathbb{E}\left[M_t^2\right] < \infty$. Two processes M and N are indistinguishable if

$$I\!\!P(\exists t \ge 0,\ M(t) \ne N(t)) = 0.$$

Indistinguishability is an equivalence relation, and we consider elements of $\mathcal{M}_0^2$ to the same if they are indistinguishable. That is, $\mathcal{M}_0^2$ is really the space of equivalence classes of square integrable martingales. Nevertheless, we commit the usual sin of speaking of elements of sM_0^2 as processes.

The material that follows is drawn from RW, volume II, II.24, pp. 42-45.

By Doob's convergence theorem $M(\infty) = \lim_{t\to\infty} M_t$ exists almost surely and in mean square and $M(t) = \mathbb{E}\left[M_\infty \mid \mathcal{F}_t\right]$. By Doob's inequality,

$$\mathbb{E}\left[\sup_{t\ge 0} M(t)^2\right] \le 4\mathbb{E}\left[M(\infty)^2\right] < \infty. \tag{24}$$

Lemma 5 *$\mathcal{M}_0^2$ is a Hilbert space when endowed with the inner product* $\langle M, N\rangle_2 \triangleq \mathbb{E}\left[M_\infty N_\infty\right]$.

Proof: Let the norm associated to the inner product be denoted $\|\cdot\|_2$. If $\|M–N\|_2 = \mathbb{E}\left[(M_\infty - N_\infty)^2\right] = 0$, then (23) implies that M and N are indistinguishable, so that $M = N$ (up to indistinguishability).

Let $M^{(n)}$ be a Cauchy sequence. There is a a random variable $M(\infty)$ such that $M^{(n)}(\infty) \to M(\infty)$ in mean square. Let $M(t) \triangleq \mathbb{E}\left[M(\infty) \mid \mathcal{F}_t\right]$, $t \ge 0$. Then $\|M^{(n)} - M\|_2 \to 0$ as $n \to \infty$ and so $\mathcal{M}_0^2$ is complete. $\diamond$

Let M be in $\mathcal{M}_0^2$. Let T be an $\mathbb{F}$-stopping time. By the optional stopping theorem $M^T(t) \triangleq M(T \wedge t)$, the process M stopped at time T, is also a martingale and $\mathbb{E}\left[(M^T)^2(\infty)\right] = \mathbb{E}\left[M^2(T)\right] \leq \mathbb{E}\left[M(\infty)^2\right] < \infty$. Therefore $\mathcal{M}_0^2$ is closed under the operation, $M \to M^T$.

Definition. A subspace $\mathcal{U}$ of $\mathcal{M}_0^2$ is called *stable* if it is closed in the Hilbert space topology and if $N^T \in \mathcal{U}$ for every $N \in \mathcal{U}$.

Let $c\,\mathcal{M}_0^2$ denote the subset of $\mathcal{M}_0^2$ consisting of all $\mathbb{F}$-martingales whose paths are almost-surely continuous. This is obviously a subspace. Let $\{M_n\}$ be a sequence of martingakes in $c\,\mathcal{M}_0^2$, and suppose $\|M_n - M\|_2 \to 0$ as $n \to \infty$. Then from (23),

$$\mathbb{E}\left[\sup_t (M_n(t) - M(t))^2\right] \to 0$$

as well. It follows that the paths of M must also be continuous almost surely. Also, if T is a stopping time M^T has continuous paths. This proves the following.

Lemma 6 *$c\,\mathcal{M}_0^2$ is a stable subspace of $\mathcal{M}_0^2$.*

Define next the subspace of $\mathcal{M}_0^2$ orthogonal to $c\,\mathcal{M}_0^2$.

$$d\,\mathcal{M}_0^2 \triangleq \left(c\,\mathcal{M}_0^2\right)^\perp \triangleq \{N \in \mathcal{M}_0^2;\ \langle M, N\rangle_2 = 0\}$$

This space is called the space of pure jump martingales.

Lemma 7 *(a) $d\,\mathcal{M}_0^2$ is stable.*

(b) $N \in d\,\mathcal{M}_0^2$ if and only if

$$\mathbb{E}\left[M(T)N(T)\right] = 0 \quad \text{for all } [0,\infty] \text{ valued, } \mathbb{F}\text{-stopping times} \tag{25}$$

for all $M \in c\,\mathcal{M}_0^2$. *(When condition (25) holds we say that M and N are* strongly orthogonal.*)*

(c) A necessary and sufficient condition that $\widetilde{M}$ and $\widetilde{N}$ in $\mathcal{M}_0^2$ are strongly orthogonal as in (25) is that $\widetilde{M}\widetilde{N}$ be a uniformly integrable martingale.

Proof: (a): Orthogonal subspace of a Hilbert space are closed and so $d\,\mathcal{M}_0^2$ is closed. If $N \mathbf{1}\,\mathcal{M}_0^2$ and if $m \in c\,\mathcal{M}_0^2$, then $\mathbb{E}\left[N^T(\infty)M(\infty)\right] = \mathbb{E}\left[\mathbb{E}\left[N(\infty) \mid \mathcal{F}_T\right] M(\infty)\right] = \mathbb{E}\left[\mathbb{E}\left[N(\infty) \mid \mathcal{F}_T\right] \mathbb{E}\left[M(\infty) \mid \mathcal{F}_T\right]\right] = \mathbb{E}\left[N(\infty)M^T(\infty)\right]$ But $M^T \in c\,\mathcal{M}_0^2$ since $c\,\mathcal{M}_0^2$ is stable and N is thus orthogonal to M^T. Hence $\mathbb{E}\left[N^T M(\infty)\right] = 0$. This is true for any $M \in c\,\mathcal{M}_0^2$, and so $N^T \in d\,\mathcal{M}_)^2$.

(b): Suppose that $N \in d\,\mathcal{M}_0^2$, and $M \in c\,\mathcal{M}_0^2$. Then from (a) and Lemma 6, $N^T \in d\,\mathcal{M}_0^2$, and $M \in c\,\mathcal{M}_0^2$, also and hence $0 = \mathbb{E}\left[M^T(\infty)N^T(\infty)\right] = \mathbb{E}\left[M(T)N(T)\right]$. Conversely if (25) holds for all $M \in c\,\mathcal{M}_0^2$, then, setting $T = \infty$ implies $\mathbb{E}\left[N(\infty)M(\infty)\right] = 0$ for all $M \in c\,\mathcal{M}_0^2$. Hence $N \in d\,\mathcal{M}_0^2$.

(c): Suppose $\mathbb{E}\left[\widetilde{M}(T)\widetilde{N}(T)\right] = 0$ for all $[0,\infty]$ valued, $\mathbb{F}$-stopping times. Let $t > 0$ and let $A \in \mathcal{F}_t$. Define the stopping time,

$$T(\omega) \triangleq \begin{cases} t, & \text{if } \omega \in A; \\ \infty, & \text{if not.} \end{cases}$$

Then,

$$0 = \mathbb{E}\left[\widetilde{M}(T)\widetilde{N}(T)\right] = \mathbb{E}\left[\mathbf{1}_A\widetilde{M}(t)\widetilde{N}(t)\right] + \mathbb{E}\left[\mathbf{1}_{A^c}\widetilde{M}(\infty)\widetilde{N}(\infty)\right].$$

But also,

$$0 = \mathbb{E}\left[\widetilde{M}(\infty)\widetilde{N}(\infty)\right] = \mathbb{E}\left[\mathbf{1}_A\widetilde{M}(\infty)\widetilde{N}(\infty)\right] + \mathbb{E}\left[\mathbf{1}_{A^c}\widetilde{M}(\infty)\widetilde{N}(\infty)\right].$$

Taing the difference, it follows that

$$\mathbb{E}\left[\mathbf{1}_A\widetilde{M}(t)\widetilde{N}(t)\right] = \mathbb{E}\left[\mathbf{1}_A\widetilde{M}(\infty)\widetilde{N}(\infty)\right] \quad \text{for any } A \in \mathcal{F}_t.$$

But this implies that $\widetilde{M}(t)\widetilde{N}(t) = \mathbb{E}\left[\widetilde{M}(\infty)\widetilde{N}(\infty) \mid \mathcal{F}_t\right]$ for all $t \geq 0$, which means that $\widetilde{M}\widetilde{N}$ is a uniformly integrable martingale.

Conversely, if $\widetilde{M}\widetilde{N}$ is a uniformly integrable martingale, then, by optional stopping $\mathbb{E}\left[M(T)N(T)\right] = \mathbb{E}\left[M(0)N(0)\right] = 0$ for any stopping time T. ◇

Some of the techniques and statements in the discussion above generalize. See RW, IV.24, and II.77.6.

Example. Let Y be a Lévy process with respect to $\mathbb{F}$ of the form, $Y(t) = \sum_1^{N(t)} \xi_i - \lambda\mu t$, as defined in section II, where N is a Poisson process with rate λ, and $\xi_1, \xi_2, \ldots$ are i.i.d. random variables in dependent of N with mean μ and variance σ^2. For any finite r, let $Y^r(t) = Y(t \wedge r)$. Then $Y^r \in d\mathcal{M}_0^2$.

To see this, first we compute the total variation of Y (*exercise 11*): $TV(Y)(t) = \sum_1^{N(t)} |\xi_i| + \lambda\mu t$. Let $\rho = E|\xi_i|$. For every t, $\mathbb{E}\left[TV(Y)^2(t)\right] = (\lambda\mu t)^2 + \lambda^2 t^2 \mu(\rho + \rho^2) + \lambda t(\sigma^2 + \mu^2) < \infty$. Let T be any $\mathbb{F}$-stopping time and let $M \in c\mathcal{M}_0^2$. Then

$$\mathbb{E}\left[Y^r(\infty)M(\infty)\right] = \mathbb{E}\left[Y(r)M(\infty)\right] = \mathbb{E}\left[Y(r)\mathbb{E}\left[M(\infty) \mid \mathcal{F}_r\right]\right] = \mathbb{E}\left[Y(r)M(r)\right].$$

Let $\{\Pi_n\}$ be a sequence $0 = t_0^{(n)} < t_1^{(n)} < \cdots t_n^{(n)} = r$ of partitions of $[0,r]$ with $\|\Pi_n\| \to 0$ as $n \to \infty$. Write

$$\begin{aligned} Y(r)M(r) &= \sum_1^n \left[Y(t_{i+1}^{(n)})M(t_{i+1}^{(n)}) - Y(t_i^{(n)})M(t_i^{(n)})\right] \\ &= \sum_1^n Y(t_i^{(n)})\left[M(t_{i+1}^{(n)}) - M(t_i^{(n)})\right] + \sum_1^n M(t_i^{(n)})\left[Y(t_{i+1}^{(n)}) - Y(t_i^{(n)})\right] \\ &\quad \sum_1^n \left[M(t_{i+1}^{(n)}) - M(t_i^{(n)})\right]\left[Y(t_{i+1}^{(n)}) - Y(t_i^{(n)})\right] \end{aligned} \tag{26}$$

By the martingale property, the expectation of the first two terms is zero. The second term is bounded by

$$\sup_i \left| M(t_{i+1}^{(n)}) - M(t_i^{(n)}) \right| TV(Y)(r) \leq 2 \sup_{[0,r]} |M_t| TV(Y)(r)$$

The first term in this sum converges to 0 almost-surely as $n \to \infty$, since the paths of M are continuous. The second term is integrable, because, by Hölder's inequality,

$$\mathbb{E}\left[\sup_{[0,r]} |M_t| TV(Y)(r)\right] \leq \left[\mathbb{E}\left[\sup_{[0,r]} |M_t|^2\right] \mathbb{E}\left[TV(Y)^2(r)\right]\right]^{1/2} < \infty.$$

Thus, using the dominated convergence theorem, we see that by taking taking expectations of both sides of (26) and letting $n \to \infty$, $\mathbb{E}\left[Y^r(\infty)M(\infty)\right] = 0$. This shows that Y is orthogonal to $c\,\mathcal{M}_0^2$.$\diamond$

If U is a closed subspace of a Hilbert space, then that Hilbert space is a direct sum of U and $U^\perp$. As a consequence, there is a unique decomposition of any martingale in $\mathcal{M}_0^2$ into a sum of a continuous and a pure jump martingale.

Theorem 4 *Any martingale $M \in \mathcal{M}_0^2$ has a unique decomposition $X = M_c + M_d$, where $M_c \in c\,\mathcal{M}_0^2$ and $M_d \in d\,\mathcal{M}_0^2$.*