642: 621 MATHEMATICAL FINANCE I

NOTES FOR LECTURE 8

I <u>Quadratic</u> <u>Variation</u> <u>of</u> <u>Brownian</u> <u>Motion</u>.

A) <u>Fact</u>: If $X \sim N(\mu, \sigma^2)$ then $E[(X-\mu)^4] = 3\sigma^4$.
Showing this is Exercise 3.3 in Shreve and was assigned as a homework problem.

B) <u>Another</u> <u>Fact</u>: If $X \sim N(\mu, \sigma^2)$, then

$$E\left[\left((X-\mu)^2 - \sigma^2\right)^2\right] = 2\sigma^4$$

Proof:

$$E\left[\left((X-\mu)^2 - \sigma^2\right)^2\right] = E\left[(X-\mu)^4 - 2\sigma^2(X-\mu)^2 + \sigma^4\right]$$

$$= E[(X-\mu)^4] - 2\sigma^2 E[(X-\mu)^2] + \sigma^4$$

$$= 3\sigma^4 - 2\sigma^2\sigma^2 + \sigma^4 \quad \text{by Fact A and } \sigma^2 = \mathrm{Var}(X) = E[(X-\mu)^2]$$

$$= 2\sigma^4$$

C) Let W be a Brownian motion. Let $T > 0$. For a partition Π, defined by $0 = t_0 < t_1 < \cdots < t_n = T$, let

$$\|\Pi\| := \max_{1 \le i \le n} t_i - t_{i-1}$$

and let

$$[W,W]_T^\Pi := \sum_{j=0}^{n-1} \left[W(t_{j+1}) - W(t_j)\right]^2$$

<u>Theorem</u> <u>1</u>

$$\lim_{\|\Pi\| \to 0} E\left[\left([W,W]_T^\Pi - T\right)^2\right] = 0$$

Thus the quadratic variation of W on $[0,T]$ is $[W,W](T) = T$.

This is Theorem 3.4.3 in Shreve. To prove, observe first that $T = \sum_{j=0}^{n-1} t_{j+1} - t_j$ for any partition Π: $0 = t_0 < t_1 < \dots < t_n = T$. Thus

$$[W,W]_T^\Pi - T = \sum_{j=0}^{n-1} \left[\left(W(t_{j+1}) - W(t_j)\right)^2 - (t_{j+1} - t_j)\right]$$

and so

$$\left([W,W]_T^\Pi - T\right)^2 = \sum_{j=0}^{n-1} \left[\left(W(t_{j+1}) - W(t_j)\right)^2 - (t_{j+1} - t_j)\right]^2$$

$$+ \sum_{\substack{i \ne j \\ 0 \le i,j \le n-1}} \left[\left(W(t_{j+1}) - W(t_j)\right)^2 - (t_{j+1} - t_j)\right]\left[\left(W(t_{i+1}) - W(t_i)\right)^2 - (t_{i+1} - t_i)\right] \qquad (1)$$

If $i \ne j$ $W(t_{j+1}) - W(t_j)$ and $W(t_{i+1}) - W(t_i)$ are increments of Brownian motion on disjoint intervals and are thus independent. Also, since $W(t_{j+1}) - W(t_j) \sim N(0, t_{j+1} - t_j)$, $E[(W(t_{j+1}) - W(t_j))^2] = \mathrm{Var}(W(t_{j+1}) - W(t_j)) = t_{j+1} - t_j$ and so $E[(W(t_{j+1}) - W(t_j))^2 - (t_{j+1} - t_j)] = 0$. It follows that if $i \ne j$

$$E\left[\left[\left(W(t_{j+1}) - W(t_j)\right)^2 - (t_{j+1} - t_j)\right]\left[\left(W(t_{i+1}) - W(t_i)\right)^2 - (t_{i+1} - t_i)\right]\right]$$

$$= E\left[\left(W(t_{j+1}) - W(t_j)\right)^2 - (t_{j+1} - t_j)\right] \times E\left[\left(W(t_{i+1}) - W(t_i)\right)^2 - (t_{i+1} - t_i)\right]$$

$$= 0$$

Also, since $W(t_{j+1}) - W(t_j) \sim N(0, t_{j+1} - t_j)$, fact B implies

$$E\left[\left[\left(W(t_{j+1}) - W(t_j)\right)^2 - (t_{j+1} - t_j)\right]^2\right] = 2(t_{j+1} - t_j)^2$$

Putting these results together in (1) leads to

$$E\left[\left([W,W]_T^{\Pi}-T\right)^2\right] = \sum_{j=0}^{n-1} E\left[\left[(W(t_{j+1})-W(t_j))^2-(t_{j+1}-t_j)\right]^2\right]$$

$$= \sum_{j=0}^{n-1} 2\,(t_{j+1}-t_j)^2$$

But $t_{j+1}-t_j \le \|\Pi\|$ and so $(t_{j+1}-t_j)^2 \le \|\Pi\|\,(t_{j+1}-t_j)$

Thus

$$E\left[\left([W,W]_T^{\Pi}-T\right)^2\right] \le 2\,\|\Pi\| \sum_{j=0}^{n-1}(t_{j+1}-t_j) = 2\,\|\Pi\|\,T$$

Hence

$$\lim_{\|\Pi\|\to 0} E\left[\left([W,W]_T^{\Pi}-T\right)^2\right] = 0. \qquad \blacksquare$$

II. Volatility in the Black-Scholes price model.

The Black-Scholes model for an asset price takes the form

$$S(t) = S(0)\, e^{\alpha t + \sigma W(t) - \frac{1}{2}\sigma^2 t}$$

where α and σ are parameters with $\sigma > 0$. The log return over a period $[t_j, t_{j+1}]$ is

$$\ln\left(\frac{S(t_{j+1})}{S(t_j)}\right) = \sigma\left[W(t_{j+1})-W(t_j)\right] + (\alpha-\tfrac{1}{2}\sigma^2)(t_{j+1}-t_j) \qquad (2)$$

It is follows from section 3.4.3 in Shreve that, because of Theorem 1

$$E\left[\left(\sum_{j=0}^{n-1}\left(\ln\frac{S(t_{j+1})}{S(t_j)}\right)^2 - T\sigma^2\right)^2\right] \to 0 \quad \text{as}$$

or more loosely speaking

$$\sum_{j=0}^{n-1}\left(\ln\left(\frac{S(t_{j+1})}{S(t_j)}\right)\right)^2 \to \sigma^2 T \quad \text{as } \|\Pi\|\to\infty \qquad (3).$$

σ is called the volatility because it measures the strength per unit time of the fluctuation of $\ln\left(\frac{S(t_{j+1})}{S(t_j)}\right)$ about its mean. Indeed

$$\mathrm{Var}\left(\ln\left(\frac{S(t_{j+1})}{S(t_j)}\right)\right) = \mathrm{Var}\left(\sigma\left(W(t_{j+1})-W(t_j)\right)\right)$$

$$= \sigma^2(t_{j+1}-t_j).$$

(3) says that σ^2 may be estimated by $\frac{1}{T}\sum_{j=0}^{n-1}\ln\left(\frac{S(t_{j+1})}{S(t_j)}\right)^2$ where $t_0 = 0 < t_1 < \dots < t_n = T$ is a partition of $[0,T]$ into many small subintervals.

III. Generalizing the definition of integration.

In advanced mathematical analysis, the procedure of integration is defined in a very general way, and the familiar Riemann integral is just a specific example of this general approach. It is worthwhile to understand this general perspective before moving on to stochastic integration theory.

First we recall the definition of the Riemann integral.

A. Riemann integration

Let f be a real-valued function defined on an interval $[a,b]$. The Riemann integral of f on $[a,b]$, if it exists, is

$$\int_a^b f(x)\,dx = \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f(\bar{t}_j)\,[t_{j+1}-t_j] \qquad (4)$$

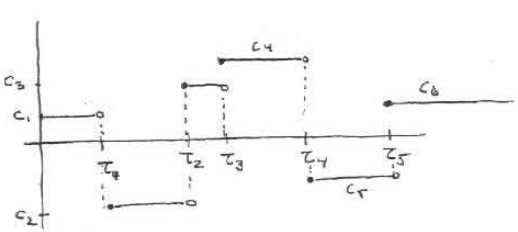
In the Riemann sum $\sum_{j=0}^{n-1} f(\bar{t}_j)(t_{j+1}-t_j)$ $\quad t_0 = a < t_1 < \cdots < t_n = b$ are the points of the partition Π and, for each j, $t_j \le \bar{t}_j \le t_{j+1}$, and the limit in (4) should be independent of the exact choice of the $\bar{t}_j$'s. When $f(\bar{t}_j) \ge 0$, $f(\bar{t}_j)(t_{j+1}-t_j)$ has the interpretation of the area of a rectangle of height $f(\bar{t}_j)$ built over the interval $[t_j, t_{j+1}]$.

The Riemann integral is especially simple to compute when the integrand f is piecewise constant; this means there are $\tau_0 = 0 < \tau_1 < \tau_2 < \cdots < \tau_M$ and constants $c_1, c_2, \ldots c_{M+1}$ so that

$$f(t) = \begin{cases} c_i, & \text{if } \tau_{i-1} \le t < \tau_i,\ 1 \le i \le M; \\ c_{M+1} & \text{if if } t \ge \tau_{M+1}. \end{cases}$$

f may also be written:

$$f(t) = \sum_{i=0}^{M-1} c_{i+1} 1_{[\tau_i, \tau_{i+1})}(t) + c_{M+1} 1_{[\tau_M, \infty)}(t).$$



Then, if $\quad \tau_{k-1} < t \le \tau_k$,

$$\int_0^t f(s)ds = \sum_{i=0}^{k-2} f(\tau_i)[\tau_{i+1}-\tau_i] + f(\tau_{k-1})[t-\tau_{k-1}]$$
$$= \sum_{i=0}^{k-2} c_{i+1}[\tau_{i+1}-\tau_i] + c_k[t-\tau_{k-1}]$$

For notational convenience, write $\quad t \wedge s = \min(t,s)$. Then this integral may be re-expressed as

$$\int_0^t f(s)ds = \sum_{i=0}^{M} f(\tau_i)[\tau_{i+1} \wedge t - \tau_i \wedge t], \qquad (4)$$

because, if $\tau_{k-1} < t \le \tau_k$, $\quad t \wedge \tau_i = \tau_i$ if $\quad i \le k-1$, and $\quad t \wedge \tau_i = t$, if $i \ge k$; if $t > \tau_M$, $t \wedge \tau = \tau_i$ for all i.

B. The Stieltjes integral

In formula (4), each term in the sum computing $\int_0^t f(s)ds$ is the value of f on the interval $[\tau_i \wedge t, \tau_{i+1} \wedge t)$ times the length of that interval. The main idea in the general notion of an integral is to replace the length of with a different 'measure' or 'weight' assignment to intervals. The Stieltjes integral is defined by replacing the length of $[\tau_i, \tau_{i+1}]$ by $G(\tau_{i+1}) - G(\tau_i)$, where G is a given function. In this discussion, to avoid technical issues, we will assume G is continuous, but, in general, it need not be. Thus if f is the piecewise-constant function defined above,

$$f(t) = \sum_{i=0}^{M-1} c_{i+1} 1_{[\tau_i, \tau_{i+1})}(t) + c_{M+1} 1_{[\tau_M, \infty)}(t)$$

the Stieltjes integral of f with respect to G over $[0,t]$ is defined by

$$\int_0^t f(s)\,dG(s) = \sum_{i=0}^{M} c_{i+1}\left[G(\tau_{i+1} \wedge t) - G(\tau_i \wedge t)\right] \qquad (5)$$
$$= \sum_{i=0}^{M} f(\tau_i)\left[G(\tau_{i+1} \wedge t) - G(\tau_i \wedge t)\right],$$

Some applications may help to grasp what is going on here. For example, suppose we replace G by the cumulative distribution function F_X of a positive random variable X, and assume

F_X is continuous. Thus $F_X(t) = \mathbb{P}(0 \le X < t)$ and $F_X(\tau_{i+1}) - F_X(\tau_i) = \mathbb{P}(\tau_i \le X < \tau_{i+1})$ (usually we define $F_X(t) = \mathbb{P}(X \le t]$ but since F_X is continuous $F_X(t) = \mathbb{P}(X<t)$ also -- $\mathbb{P}(X=t)=0$ for all t).

Then

$$\int_0^t f(s)\,dF_X(t) = \sum_{i=0}^{M} c_{i+1}\,\mathbb{P}[\tau_i \wedge t \le X < \tau_{i+1} \wedge t]$$

$$= \sum_{i=0}^{M} c_{i+1}\,\mathbb{P}[X \le t,\ f(X) = c_{i+1}]$$

(because $f(x) = c_{i+1}$ when $\tau_i \le X < \tau_{i+1}$)

$$= E[f(X)\,1_{[X \le t]}]$$

For a general integrand f and a continuous G.

$$\int_0^t f(s)\,dG(s) = \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} f(\bar{t}_i)\left[G(t_{i+1}) - G(t_i)\right] \qquad (6)$$

where, in the sum, $t_0 = 0 < t_1 < \cdots < t_n = t$ are the points of the partitition Π, and $t_i \le \bar{t}_i \le t_{i+1}$ for every i. The integral is defined in this approach, only if the limit exists and is independent of how the partitions are chosen, as long as $\|\Pi\| \to 0$, and of how $\bar{t}_i$ in $[t_i, t_{i+1}]$ is chosen.

Existence of $\int_0^t f(s)\,dG(s)$ requires regularity properties on both f and G. For example, suppose we want,

for a given G, that $\int_0^t f(s)\,dG(s)$ be defined for all t and all continuous functions f. One can show (we won't) that a sufficient condition for this is that G have bounded variation. This means that for any t, there is a finite constant K_t such that

$$\sum_{i=0}^{n-1} \left| G(t_{i+1}) - G(t_i) \right| \le K_t \qquad (7)$$

for any partition $0 = t_0 < t_1 < \cdots < t_n = t$ of $[0,t]$.
(We bring this up because Brownian motion paths, which we will bring in later, do not have bounded variation.)

Example. Assume that G is continuously differentiable -- $G'(t)$ exists for all t and $G'(t)$ is continuous as a function of t. We will show that G has bounded variation. Since $G'(s)$ is continuous, the maximum of $|G'(s)|$ on an interval $[0,t]$ is finite. Call this maximum value M_t. Let $0 \le s < r \le t$. By the mean value theorem there is a ξ, where $s < \xi < r$ such that $G(r) - G(s) = G'(\xi)[r-s]$. Hence

$$|G(r) - G(s)| = |G'(\xi)|(r-s) \le M_t(r-s),$$

whenever $0 \le s < r \le t$. Thus, for any partitition $0 = t_0 < t_1 < \cdots < t_n = t$ of $[0,t]$

$$\sum_{i=0}^{n-1} |G(t_{i+1}) - G(t_i)| \le \sum_{i=0}^{n-1} M_t(t_{i+1} - t_i) = t \cdot M_t$$

This proves (7) with $K_t = t \cdot M_t$.

<u>Example</u> Assume that G is continuously differentiable. Then, for any t,

$$\int_0^t f(s)\,dG(s) = \int_0^t f(s)G'(s)\,ds$$

assuming the integrals are defined.

We do not show this rigorously, but only give the idea. Again, it is based on the mean value theorem. Given any interval $[t_i, t_{i+1}]$ there is a $t_i^* \in (t_i, t_{i+1})$ such that $G(t_{i+1}) - G(t_i) = G'(t_i^*)(t_{i+1} - t_i)$. Then

$$\int_0^t f(s)\,dG(s) = \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} f(\bar{t}_i)\left[G(t_{i+1}) - G(t_i)\right]$$

$$= \lim_{\|\Pi\| \to 0} \sum_{i=1}^{n-1} f(\bar{t}_i)\,G'(t_i^*)\left[t_{i+1} - t_i\right]$$

But this last expression is (almost) a limit of Riemann sums for the integrand $f(s)G'(s)$ and hence it equals

$$\int_0^t f(s)G'(s)\,ds$$

IV. <u>The Itô integral</u> <u>with respect to</u> <u>Brownian motion</u>.

Let $\{W(t)\}_{t \geq 0}$ be a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{\alpha(t)\}_{t \geq 0}$ be another stochastic process defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Our goal is to define an integral of the form $\int_0^t \alpha(s)\,dW(s)$.

It is tempting to just defined this as a pathwise, Stieltjes integral: for each ω let

$$\int_0^t \alpha(s)(\omega)\,dW(s)(\omega)$$

be the Stieltjes integral of $\{\alpha(s)(\omega)\}_{s \geq 0}$ with respect to $\{W(s)(\omega)\}_{s \geq 0}$. Unfortunately this will not work in general. The paths of $\{W(t)(\omega)\}_{t \geq 0}$ have unbounded variation and the Stieltjes integral will in general be defined only on a very restricted class of integrands. Instead we must proceed by an approximation approach and we must specialize the construction and restrict the class of integrands in a special way. (The fact that Brownian motion paths have unbounded variation is proved in Exercise 3.4 in Shreve.)

A. <u>The Itô integral for simple processes.</u>

A stochastic process $\{\alpha(t)\}_{t \geq 0}$ is <u>simple</u> if it can be written in the form

$$\alpha(t)(\omega) = \sum_{i=0}^{M-1} Y_i(\omega)\, 1_{[t_i, t_{i+1})}(t) + Y_M(\omega)\, 1_{[t_M, \infty)}(t)$$

for some sequence $0 = t_0 < t_1 < \cdots < t_M$ and random variables $Y_1, \ldots, Y_M$. Thus $\alpha(t)$ takes the random value Y_0 on the interval $t_0 \leq t < t_1$, the value Y_1 on $t_1 \leq t < t_2$, etc. Hence, for every ω, the path $\alpha(t)(\omega)$ is a piecewise constant function of t with change points at $t_1, t_2, \ldots, t_M$.

In this case we define the Ito integral as we did the Stieltjes integral, but now of course everything is a random variable. Note that we can write the simple process above as

$$\alpha(t)(\omega) = \sum_{i=0}^{M-1} \alpha(t_i)(\omega)\, 1_{[t_i, t_{i+1})}(t) + \alpha(t_M)\, 1_{[t_M, \infty)}(t). \tag{8}$$

<u>Definition 1</u>. If α is a simple process, as in (8),

$$\left[\int_0^t \alpha(s)\,dW(s)\right](\omega) := \sum_{i=0}^{M} \alpha(t_i)(\omega)\left[W(t_{i+1}\wedge t)(\omega) - W(t_i\wedge t)(\omega)\right] \qquad 0 \le t < \infty$$

(in this sum $t_{M+1}=\infty$)

We have written all terms explicitly as functions of ω, to emphasize that the integral is a random variable. But in general, showing the dependence on ω is clumsy and we write, simply,

$$\int_0^t \alpha(s)\,dW(s) = \sum_{i=0}^{M} \alpha(t_i)\left[W(t_{i+1}\wedge t) - W(t_i\wedge t)\right]$$

Note also that $\int_0^t \alpha(s)\,dW(s)$ is being defined here for all $t \ge 0$, not just for a single fixed t.

B. <u>An interpretation</u>

It is very helpful to keep the following interpretation in mind. Suppose we are allowed to bet on the increments of a Brownian motion at fixed times $0 = t_0 < t_1 < \cdots < t_M$. The bet at time t_i is on the future increments $W(s) - W(t_i)$, $t_i < s \le t_{i+1}$, the amount of the bet is $\alpha(t_i)$ and the amount we win or lose by time s, $t_i \le s \le t_{i+1}$ on this bet is $\alpha(t_i)\left[W(s) - W(t_i)\right]$. Then, if $t_k < t \le t_{k+1}$

$$\int_0^t \alpha(s)\,dW(s) = \sum_{i=0}^{k-1} \alpha(t_i)\left[W(t_{i+1}) - W(t_i)\right] + \alpha(t_k)\left[W(t) - W(t_k)\right]$$

is the total of all that we have won or lost on the bets taken up to time t. In short, one may think of a stochastic integral as totaling the gains obtained by betting on Brownian motion increments.

C. <u>Adapted process</u>

Let W be a Brownian motion and let $\{\mathcal{F}(t)\}_{t\ge 0}$ be a filtration for W; recall that this means (i) $W(t)$ is $\mathcal{F}(t)$-measurable for each $t \ge 0$, and ii) $W(t) - W(s)$ is independent of $\mathcal{F}(s)$ for all $t > s$.

Think of $\mathcal{F}(t)$ as the information available at time t to a gambler who wants to bet on the increments of W. If $\{\alpha(t)\}_{t\ge 0}$ is a simple process representing his betting scheme it follows that for each t_i, $\alpha(t_i)$, the amount bet at t_i, is $\mathcal{F}(t_i)$-measurable. For <u>any</u> t.

$$\alpha(t) = \alpha(t_k)$$

where t_k is the largest of the numbers $t_0 < t_1 < \cdots < t_M$ such that $t_k \le t$. Thus $\alpha(t)$ is $\mathcal{F}(t_k)$-measurable, and since $t_k \le t$ and $\mathcal{F}(t_k) \subseteq \mathcal{F}(t)$, it is also $\mathcal{F}(t)$-measurable.

<u>Definition</u>. A stochastic process $\{\beta(t)\}_{t\ge 0}$ defined on the same probability space as W is said to be <u>adapted to</u> <u>$\{\mathcal{F}(t)\}$</u> if $\beta(t)$ is $\mathcal{F}(t)$-measurable for every $t \ge 0$.

Thus, according to the previous discussion, if we are interested in stochastic integrals as models of betting on the increments of a Brownian motion, it is natural to consider only adapted integrands $\alpha(\cdot)$.

Restricting to adapted integrands is the key idea that allows the Itô integral to be extended to more general integrands than simple processes. The irregularity of Brownian paths does not allow one to define stochastic integration by pathwise Stieltjes integration. But when adaptedness is imposed it is possible to develop a useful theory.

D. A study of $\int_0^t \alpha(s)\,dW(s)$ for adapted, simple integrands $\alpha(\cdot)$

(Note: Shreve includes an adaptedness condition as part of his definition of a simple process. We do not.)

Theorem 2. Let $\{\alpha(t)\}_{0\le t\le T}$ be a simple process which is adapted to a filtration $\{\mathcal{F}(t)\}_{t\ge 0}$ for a Brownian motion W. Assume

$$E\left[\int_0^T \alpha^2(s)\,ds\right] < \infty \qquad (9)$$

Then

(i) $\int_0^t \alpha(s)\,dW(s)$, $0\le t\le T$, is a martingale w.r.t. $\{\mathcal{F}(t)\}$
(In particular $E\left[\int_0^t \alpha(s)\,dW(s)\right]=0$ for all $t\le T$)

(ii) (Itô isometry)

$$\mathrm{Var}\left(\int_0^T \alpha(s)\,dW(s)\right) = E\left[\left(\int_0^T \alpha(s)\,dW(s)\right)^2\right] = E\left[\int_0^T \alpha^2(s)\,ds\right]$$

(iii) Let $X(t)=\int_0^t \alpha(s)\,dW(s)$. The quadratic variation process of X is $[X,X](t) = \int_0^t \alpha^2(s)\,ds$, $t\le T$

(iv) For each ω, $\left[\int_0^t \alpha(s)\,dW(s)\right](\omega)$ is a continuous function of t.

Remarks on and proof of this theorem.

Assumption (9) is important for property (i) of the theorem and also, obviously, for property (ii). If $\alpha(t) = \sum_0^M \alpha(t_i)\,1_{[t_i,t_{i+1})}(t)$, where $0=t_0<\cdots<t_{M+1}$, then $\alpha^2(t) = \sum_0^M \alpha^2(t_i)\,1_{[t_i,t_{i+1})}(t)$, and so

$$E\left[\int_0^T \alpha^2(t)\,dt\right] = E\left[\int_0^T \sum_0^M \alpha^2(t_i)1_{[t_i,t_{i+1})}(t)\right]$$
$$= E\left[\sum_0^{M-1} \alpha^2(t_i)\left[T\wedge t_{i+1} - T\wedge t_i\right]\right]$$
$$= \sum_0^M E\left[\alpha^2(t_i)\right]\left(T\wedge t_{i+1} - T\wedge t_i\right), \qquad (10)$$

since $\int_0^T 1_{[t_i,t_{i+1})}(t)\,dt = \int_{T\wedge t_i}^{T\wedge t_{i+1}} dt = T\wedge t_{i+1} - T\wedge t_i$

Thus when condition (9) is true $E\left[\alpha^2(t_i)\right]<\infty$ for any $t_i<T$.

One other fact plays an important role in the Theorem and is a consequence of the assumption that α is adapted to $\{\mathcal{F}(t)\}_{t\ge 0}$; this fact is

for each i, $\alpha(t_i)$ and $W(t_{i+1}\wedge t) - W(t_i\wedge t)$ are independent. (11)

Indeed, when $t \le t_i$, $W(t_{i+1} \wedge t) - W(t_i \wedge t) = 0$ and so $\alpha(t_i)$ and $W(t_{i+1} \wedge t) - W(t_i \wedge t)$ are independent.
When $t > t_i$, $W(t_{i+1} \wedge t) - W(t_i \wedge t) = W(t_{i+1} \wedge t) - W(t_i)$, and this is independent of $\mathcal{F}(t_i)$ by the assumption $\{\mathcal{F}(t)\}_{t \ge 0}$ is a filtration for W. But since $\alpha(t_i)$ is $\mathcal{F}(t_i)$-measurable, by the assumption that the process $\{\alpha(t)\}_{t \ge 0}$ is adapted to $\{\mathcal{F}(t)\}_{t \ge 0}$, so $W(t_{i+1} \wedge t) - W(t_i \wedge t)$ must be independent of $\alpha(t_i)$.

An immediate consequence of (11) is that

$$E\left[\alpha(t_i)\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right]\right]$$
$$= E[\alpha(t_i)]\, E\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right] = 0,$$

because $E\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right] = 0$.

Another important consequence is that if $i \ne j$, $\alpha(t_i)\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right]$ and $\alpha(t_j)\left[W(t_{j+1} \wedge t) - W(t_j \wedge t)\right]$ are uncorrelated. Suppose, for example, that $i < j$, so that $i + 1 \le j$. Then since $\alpha(t_i)$ is $\mathcal{F}(t_i)$ and hence $\mathcal{F}(t_j)$-measurable, $W(t_{i+1} \wedge t) - W(t_i \wedge t)$ is $\mathcal{F}(t_{i+1})$-measurable and hence $\mathcal{F}(t_j)$-measurable, and since $\alpha(t_j)$ is $\mathcal{F}(t_j)$-measurable, $\alpha(t_i)\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right]\alpha(t_j)$ is $\mathcal{F}(t_j)$-measurable and hence independent of $W(t_{j+1} \wedge t) - W(t_j \wedge t)$.
Thus

$$E\left[\alpha(t_i)\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right] \cdot \alpha(t_j)\left[W(t_{j+1} \wedge t) - W(t_j \wedge t)\right]\right]$$
$$= E\left[\alpha(t_i)\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right]\alpha(t_j)\right] E\left[W(t_{j+1} \wedge t) - W(t_j \wedge t)\right]$$
$$= 0 \qquad (12)$$

Proof of Theorem 2

Let $\alpha(t) = \sum_{0}^{M} \alpha(t_i) 1_{[t_i, t_{i+1})}(t)$ define a simple process that is adapted to $\{\mathcal{F}(t)\}_{t \ge 0}$. Let $X(t) = \int_0^t \alpha(s)\,dW(s)$

Proof of (ii)

$$E\left[\left(\int_0^T \alpha(s)\,dW(s)\right)^2\right] = E\left[\left(\sum_{0}^{M} \alpha(t_i)\left[W(t_{i+1} \wedge T) - W(t_i \wedge T)\right]\right)^2\right]$$
$$= E\left[\sum_{i=0}^{M} \alpha^2(t_i)\left[W(t_{i+1} \wedge T) - W(t_i \wedge T)\right]^2\right]$$
$$+ E \sum_{\substack{i,j=0 \\ i \ne j}}^{M} \sum \alpha(t_i)\left[W(t_{i+1} \wedge T) - W(t_i \wedge T)\right]\alpha(t_j)\left[W(t_{j+1} \wedge T) - W(t_j \wedge T)\right]$$

By (12), the expectation of each term in the second sum is zero. By the independence of $\alpha(t_i)$ and $\left[W(t_{i+1} \wedge T) - W(t_i \wedge T)\right]$ for each i, and $E\left[\left(W(t_{i+1} \wedge T) - W(t_i \wedge T)\right)^2\right] = t_{i+1} \wedge T - t_i \wedge T$

$$E\left[\alpha^2(t_i)\left[W(t_{i+1} \wedge T) - W(t_i \wedge T)\right]^2\right] = E[\alpha^2(t_i)]\,(t_{i+1} \wedge T - t_i \wedge T)$$

Thus, using (10)

$$E\left[\left(\int_0^T \alpha(s)\,dW(s)\right)^2\right] = \sum_{i=0}^{M} E[\alpha^2(t_i)]\left[t_{i+1} \wedge T - t_i \wedge T\right]$$
$$= E \int_0^T \alpha^2(s)\,ds.$$

<u>Proof of (i)</u> $X(t) = \int_0^t \alpha(s)\,dW(s) = \sum_{i=0}^{M} \alpha(t_i)\left[W(t_{i+1}\wedge t) - W(t_i\wedge t)\right]$

Since a sum of martingales is a martingale, to show $\{X(t)\}_{0\le t\le T}$ is a martingale w.r.t. $\{\mathcal{F}(t)\}_{t\ge 0}$, it suffices to show that $Y_i(t) := \alpha(t_i)\left[W(t_{i+1}\wedge t) - W(t_i\wedge t)\right]$ is a martingale with respect to $\{\mathcal{F}(t)\}_{t\ge 0}$, for each i. For this we need to check that $E|Y_i(t)| < \infty$ for all t, and $E[Y(t)|\mathcal{F}(s)] = Y(s)$ when $0\le s<t$. Since $\alpha(t_i)$ are independent

$$E[Y_i^2(t)] = E\left[\alpha^2(t_i)\left[W(t_{i+1}\wedge t) - W(t_i\wedge t)\right]^2\right]$$

$$= E[\alpha^2(t_i)]\,(t_{i+1}\wedge t - t_i\wedge t)$$

and this is finite because of (10) above and the assumption $E\left[\int_0^T \alpha^2(s)ds\right] < \infty$. Since $E[Y_i^2(t)] < \infty$, $E|Y_i(t)| < \infty$ also.

Now let $0\le s<t$. If $t \le t_i$, $Y(t)=0$ and $Y(s)=0$ and, trivially, $E[Y(t)|\mathcal{F}(s)] = Y(s)$.

If $0\le s < t_i \le t$, use iterated conditioning, noting that $\mathcal{F}(s) \subseteq \mathcal{F}(t_i)$;

$$E[Y_i(t)|\mathcal{F}(s)] = E\left[E\left[\alpha(t_i)\left[W(t_{i+1}\wedge t) - W(t_i\wedge t)\right]\middle|\mathcal{F}(t_i)\right]\middle|\mathcal{F}(s)\right]$$

$$= E\left[\alpha(t_i)\,E\left[W(t_{i+1}\wedge t) - W(t_i\wedge t)\,|\,\mathcal{F}(t_i)\right]\middle|\mathcal{F}(s)\right]$$

(since $\alpha(t_i)$ is $\mathcal{F}(t_i)$-measurable)

$= 0$ (since by independence of $W(t_{i+1}\wedge t) - W(t_i\wedge t)$,

$= Y(s)$ $\quad E[W(t_{i+1}\wedge t) - W(t_i\wedge t)\,|\,\mathcal{F}(t_i)] = 0$

Finally, if $t_i \le s < t$, then $\alpha(t_i)$ is $\mathcal{F}(s)$-measurable and $t_i\wedge t = t_i$, so

$$E[Y_i(t)|\mathcal{F}(s)] = E\left[\alpha(t_i)\left[W(t_{i+1}\wedge t) - W(t_i)\right]\middle|\mathcal{F}(s)\right]$$

$$= \alpha(t_i)\left[E[W(t_{i+1}\wedge t)|\mathcal{F}(s)] - E[W(t_i)|\mathcal{F}(s)]\right]$$

$$= \alpha(t_i)\left[W(t_{i+1}\wedge s) - W(t_i)\right] = Y_i(s)$$

because $W(t_i)$ is $\mathcal{F}(s)$-measurable and because W is a martingale with respect to $\{\mathcal{F}(t)\}_{t\ge 0}$.

Thus we have verified $E[Y_i(t)|\mathcal{F}(s)] = Y_i(s)$ in all cases when $0\le s<t$, which shows $\{Y_i(t)\}$ is a martingale w.r.t. $\{\mathcal{F}(t)\}_{t\ge 0}$.

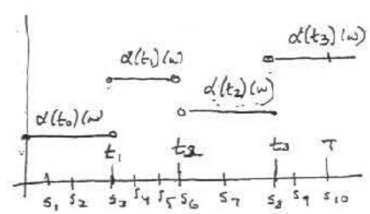
<u>Proof of (iii)</u> This 'proof' will not be fully rigorous, but it will show the reason (iii) is true. (iii) is equivalent to saying

$$[X,X]^{\Pi}(t) = \sum_{i=0}^{n-1}\left[X(s_{i+1}) - X(s_i)\right]^2 \longrightarrow \int_0^t \alpha^2(s)ds \quad \text{as } \|\Pi\|\to 0$$

(convergence in the sense of convergence in probability)

where Π denotes the partition $0 = s_0 < s_1 < \cdots < s_n = t$ of $[0,t]$. It suffices to do the case $t=T$, as the argument is the same for any t.

By adding points to Π if necessary assume it contains all the points t_j at which α changes value, as in the figure. Assume also that $t_{M+1} = T$.



Then for every i, there is some j so that $t_j \le s_i < s_{i+1} \le t_{j+1}$ and hence

$$X(s_{i+1}) - X(s_i)$$
$$= \alpha(t_j)\left[W(s_{i+1}) - W(t_j)\right] - \alpha(t_j)\left[W(s_i) - W(t_j)\right]$$
$$= \alpha(t_j)\left[W(s_{i+1}) - W(s_i)\right]$$

Therefore

$$[x,x]^{\Pi}(T) = \sum_{j=0}^{M} \sum_{\substack{i \\ t_j \le s_i < t_{j+1}}} \alpha^2(t_j)\left[W(s_{i+1})-W(s_i)\right]^2$$

$$= \sum_{j=0}^{M} \alpha^2(t_j) \sum_{\substack{i \\ t_j \le s_i < t_{j+1}}} \left[W(s_{i+1})-W(s_i)\right]^2$$

As $\|\Pi\| \to 0$,

$$\sum_{\substack{i \\ t_j \le s_i < t_{j+1}}} \left[W(s_{i+1})-W(s_i)\right]^2 \to t_{j+1} - t_j \qquad (13)$$

because the portion of Π partitioning $[t_j, t_{j+1}]$ is becoming finer and finer, and the same reasoning as in Theorem 1 of section I shows (13). Thus

$$[x,x]^{\Pi}(T) \to \sum_{j=0}^{M} \alpha^2(t_j)\left[t_{j+1}-t_j\right] = \int_0^T \alpha^2(s)\,ds$$

as $\|\Pi\| \to 0$.

Proof of (iv). Since $W(t_i \wedge t)$ is continuous in t, the same is true for

$$\int_0^t \alpha(s)\,dW(s) = \sum_0^M \alpha(t_i)\left[W(t_{i+1} \wedge t) - W(t_i \wedge t)\right]$$

$\blacksquare$

E. A comment on property (ii) of Theorem 2, the Itô isometry property.

If X is a random variable, define

$$\|X\|_{L^2} = \sqrt{E[X^2]}$$

$\|X\|_{L^2}$ is an example of what is called a norm and may be thought of a measure of the 'size' of X. In fact, it is related mathematically to the Euclidean norm $\|x\| = \sqrt{\sum_{i=1}^{d} x_i^2}$ of a vector $(x_1, .., x_d)$.

The reason for the square root in the definition is that if α is a scalar, it is natural to require the size of αX to be $|\alpha|$ times the size of X. Indeed

$$\|\alpha X\|_{L^2} = \sqrt{E[(\alpha X)^2]} = \sqrt{\alpha^2 E[X^2]} = |\alpha|\,\|X\|_{L^2}.$$

If $\{\alpha(t)\}_{0 \le t \le T}$ is a stochastic process on $[0,T]$, define

$$\|\alpha(\cdot)\|_{\mathcal{L}^2(T)} = \sqrt{E\left[\int_0^T \alpha^2(s)\,ds\right]}$$

This is a norm on $\alpha(\cdot)$, measuring its size.

Property (ii) of Theorem 2 says

$$\left\| \int_0^T \alpha(s)\,dW(s) \right\|_{L^2} = \|\alpha(\cdot)\|_{\mathcal{L}^2(T)}. \qquad (14)$$

A map from one set to another that preserves a notion of size is called an <u>isometry</u>. In this sense, the stochastic integral, thought of as a map that takes a process $\alpha(\cdot)$ to a random variable $\int_0^t \alpha(s)\,dW(s)$, is an isometry when sizes are measured, respectively, by $\|\cdot\|_{\mathcal{L}^2(T)}$ and $\|\cdot\|_{L^2}$.

F. <u>Extending the Itô integral</u>

This talk of isometries is not just idle chatter we make up to impress ourselves. It is the key to extending the Itô integral from adapted simple integrands to more general integrands.

The set of r.v.s X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\|X\|_{L^2} < \infty$ has a property called <u>completeness</u>, which says that if $\{X_n\}$ is a sequence satisfying $\lim_{n,m\to\infty} \|X_n - X_m\| = 0$, then there exists some r.v. X, also satisfying $\|X\|_{L^2} < \infty$ such that $\lim_{n\to\infty} \|X_n - X\|_{L^2} = 0$

(In other words if $\{X_n\}$ looks like it's converging as $n \to \infty$, it does converge.) We won't go into this, but we only want to point out an important consequence.

Let $\{\alpha(t)\}_{0\le t\le T}$ be an adapted process satisfying

$$\|\alpha(\cdot)\|_{\mathcal{L}^2(T)} = \sqrt{E\left[\int_0^T \alpha^2(s)\,ds\right]} < \infty$$

It is <u>not</u> assumed that α is a simple process. However, suppose there is a sequence $\{\alpha_n(\cdot)\}$ of <u>simple</u> adapted processes such that

$$\lim_{n\to\infty} \|\alpha_n(\cdot) - \alpha(\cdot)\|_{\mathcal{L}^2(T)} \to 0$$

Then because of the Itô isometry, there is a process $X(t)$ such that

$$\left\| \int_0^t \alpha_n(s)\,dW(s) - X(t) \right\|_{L^2} \longrightarrow 0 \quad \text{for all } t \le T. \tag{15}$$

<u>Theorem and Definition 3</u>. If $\{\alpha(t)\}_{0\le t\le T}$ is adapted to $\{\mathcal{F}(t)\}$ and

$$E\left[\int_0^T \alpha^2(s)\,ds\right] < \infty \tag{16}$$

there is a sequence of simple, adapted processes $\{\alpha_n(\cdot)\}$ such that $\lim_{n\to\infty} \|\alpha_n(\cdot) - \alpha(\cdot)\|_{\mathcal{L}^2(T)} \to 0$. Let $X(t)$, $0 \le t \le T$, be defined as in (15). Then we denote $X(t)$ by

$$X(t) = \int_0^t \alpha(s)\,dW(s)$$

Moreover, $X(t)$ can be defined so that it is continuous in t and

(i) $\{X(t)\}_{0\le t\le T}$ is a martingale w.r.t. $\{\mathcal{F}(t)\}_{0\le t\le T}$

(ii) (Itô isometry)

$$E\left[\left(\int_0^t \alpha(s)\,dW(s)\right)^2\right] = E\left[\int_0^t \alpha^2(s)\,ds\right], \quad 0 \le t \le T.$$

(iii) The quadratic variation of X is

$$[X, X](t) = \int_0^t \alpha^2(s)\,ds, \quad 0 \le t \le T.$$

There are two essential points to retain from this discussion

1) $\int_0^t \alpha(s)\,dW(s)$ can be defined as limit for processes $\alpha(\cdot)$ which are not simple, as long as $\alpha(\cdot)$ is adapted to $\{\mathcal{F}(t)\}_{t\geq 0}$ and satisfies (16).

2). The integral $\int_0^t \alpha(s)\,dW(s)$ has the properties (i), (ii), (iii) which we already saw were true when α is simple and adapted.

G. <u>Examples</u> Let W be a Brownian motion and let $\{\mathcal{F}(t)\}_{t\geq 0}$ be a filtration for W.

a) W is certainly adapted and for any $T>0$

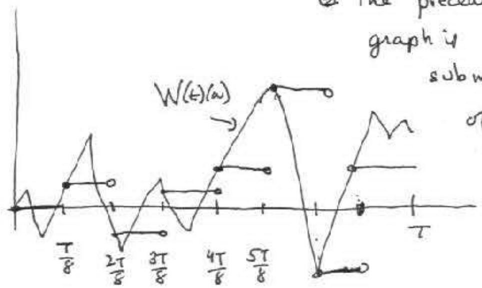
$$E\left[\int_0^T W^2(s)\,ds\right] = \int_0^T E[W^2(s)]\,ds = \int_0^T s\,ds = \tfrac{1}{2}T^2 < \infty$$

Therefore $\int_0^T W(s)\,dW(s)$ is defined.

For any n, let

$$\alpha^n(t) = \sum_{i=0}^{n-1} W\left(\tfrac{iT}{n}\right) 1_{\left[\frac{iT}{n}, \frac{(i+1)T}{n}\right)}(t)$$

The figure shows how $\alpha^n(\cdot)$ is related to W on $[0,T]$



The piecewise constant graph is α^n. Over each subinterval, the value of α is the value of W at the left endpoint of the subinterval

It is very important that $\alpha^n(t) = W(\frac{iT}{n})$, the left endpoint of the interval $[\frac{iT}{n}, \frac{(i+1)T}{n}]$, when $t \in [\frac{iT}{n}, \frac{(i+1)T}{n}]$, because then $\alpha^n(\cdot)$ is adapted to $\{\mathcal{F}(t)\}_{t\geq 0}$. Indeed, if $t \in [\frac{iT}{n}, \frac{(i+1)T}{n})$ then $\alpha^n(t) = W(\frac{iT}{n})$ is $\mathcal{F}(\frac{iT}{n})$-measurable, and hence, since $\mathcal{F}(\frac{iT}{n}) \subseteq \mathcal{F}(t)$, it is $\mathcal{F}(t)$-measurable.

If we choose to let $\alpha^n(t)$ be the value at another point in $[\frac{iT}{n}, \frac{i+1}{n}T]$ when $t \in [\frac{iT}{n}, \frac{i+1}{n}T)$, $\alpha^n(\cdot)$ would no longer be adapted to $\{\mathcal{F}(t)\}_{t\geq 0}$. For example suppose we set $\tilde{\alpha}^n(t) = W(\frac{iT}{n} + \frac{T}{2n})$, the value of W at the midpoint of $[\frac{iT}{n}, \frac{i+1}{n}T]$, when $\frac{iT}{n} \leq t < \frac{(i+1)T}{n}$. Then $\tilde{\alpha}^n(\frac{iT}{n}) = W(\frac{iT}{n} + \frac{T}{2n})$ is $\mathcal{F}(\frac{iT}{n} + \frac{T}{2n})$-measurable, but <u>not</u> $\mathcal{F}(\frac{iT}{n})$-measurable. Hence it will not be adapted to $\{\mathcal{F}(t)\}_{t\geq 0}$.

Because, the paths of W are continuous,

$$\lim_{n\to\infty} \alpha^n(t)(\omega) = W(t)(\omega) \quad \text{for every } \omega \text{ and every } t<T$$

(Imagine the partition in the figure above getting finer and finer.)
It is not too hard to show in fact that

$$\lim_{n\to\infty} E\left[\int_0^T [\alpha^n(s) - W(s)]^2\,ds\right] = 0 \quad \text{for any } T>0.$$

We won't do this here. But it says that $\int_0^T W(s)\,dW(s)$ is the limit of $\int_0^T \alpha^n(s)\,dW(s) = \sum_{i=0}^{n-1} W(\frac{iT}{n})\left[W(\frac{(i+1)T}{n}) - W(\frac{iT}{n})\right]$ as $n\to\infty$.

The limit of $\sum_{i=0}^{n-1} W(\frac{iT}{n})\left[W(\frac{(i+1)T}{n}) - W(\frac{iT}{n})\right]$ is computed very carefully in Example 4.3.2 pp 134-137. You must read and understand this theorem thoroughly.

The result is

$$\int_0^T W(s)\,dW(s) = \frac{1}{2}W^2(T) - \frac{1}{2}T \qquad (17)$$

Remarks

1) Theorem 3 says that $\int_0^T W(s)\,dW(s)$ should be a martingale in the parameter T. You should check that $\frac{1}{2}W^2(T) - \frac{1}{2}T$ is indeed a martingale.

2) Let G be a continuously differentiable function and recall the definition of Stieltjes integral and the discussion in the example on page 9 of these notes. Then Since

$$\frac{d}{ds}G^2(s) = 2G(s)G'(s)$$

$$\int_0 G(s)\,dG(s) = \int_0 G(s)G'(s)\,ds = \int_0 \frac{1}{2}\frac{d}{ds}G^2(s)\,ds$$

$$= \frac{1}{2}\left[G^2(\ ) - G^2(0)\right]$$

When $G(0)=0$, we get

$$\int_0^T G(s)\,dG(s) = \frac{1}{2}G^2(T). \qquad (18)$$

Notice how this differs from (17). Itô integrals do not behave like Stieltjes integrals. The term $-\frac{1}{2}T$ appears in (17) because as you will see by examining Example 4.3.2 in Shreve, $[W,W](t)=t$. $-\frac{1}{2}T$ must appear in (17) to make $\frac{1}{2}W^2(T) - \frac{1}{2}T$ into a martingale.

b) In a homework exercise, you will use a similar procedure to show

$$\int_0^t W^2(s)\,dW(s) = \frac{1}{3}W^3(t) - \int_0^t W(s)\,ds$$

c). Let

$$\tilde{\alpha}^n(t) = \sum_0^n W\left(\frac{iT}{n} + \frac{T}{2n}\right) 1_{\left[\frac{iT}{n}, \frac{(i+1)T}{n}\right)}(t)$$

We showed above that this is not adapted.

However

$$\int_0^T \tilde{\alpha}^n(s)\,dW(s) = \sum_0^n W\left(\frac{iT}{n} + \frac{T}{2n}\right)\left[W\left(\frac{(i+1)T}{n}\right) - W\left(\frac{iT}{n}\right)\right]$$

(although this is no longer a martingale)

and

$$E\left[\int_0^T [\tilde{\alpha}^n(s) - W(s)]^2\,ds\right] \to 0 \quad \text{as } n \to \infty$$

However, you will show in Exercise 4.4 of Shreve that

$$\lim_{n\to\infty} \int_0^T \tilde{\alpha}^n(s)\,dW(s) = \frac{1}{2}W^2(T).$$

Thus, it is very important that the integral $\int_0^t W(s)\,dW(s)$ be defined by approximating with adapted integrands. Using non-adapted integrands can lead to a different answer.