ma414l8.tex
Lecture 8. 1.3.2012

Proof.

$$\pi_n W - t = \sum_{t_i \in \pi_n} \{(W(t_{i+i}) - W(t_i))^2 - (t_{i+1} - t_i)\} = \sum_i \{(\Delta_i W)^2 - (\Delta_i t)\} = \sum_i Y_i,$$

where since $\Delta_i W \sim N(0, \Delta_i t)$, $E[(\Delta_i W)^2] = \Delta t_i$, so the Y_i have zero mean, and are independent by independent increments of W. So

$$E[(\pi_n W - t)^2] = E[(\sum_i Y_i)^2] = \sum_i E(Y_i^2),$$

since variance adds over independent summands.
Now as $\Delta_i W \sim N(0, \Delta_i t), (\Delta_i W)/\sqrt{\Delta_i t} \sim N(0,1)$, so $(\Delta_i W)^2/\Delta_i t \sim Z^2$, where $Z \sim N(0,1)$. So $Y_i = (\Delta_i W)^2 - \Delta_i t \sim (Z^2 - 1)\Delta_i t$, and

$$E[(\pi_n W - t)^2] = \sum_i E[(Z^2 - 1)^2](\Delta_i t)^2 = c \sum_i (\Delta_i t)^2,$$

writing c for $E[(Z^2 - 1)^2]$, $Z \sim N(0,1)$, a finite constant. But

$$\sum_i (\Delta_i t)^2 \leq \max_i \Delta_i t \times \sum_i \Delta_i t = |\pi_n| t,$$

giving

$$E[(\pi_n W - t)^2] \leq ct|\pi|_n \to 0 \qquad (|\pi_n| \to 0). \quad //$$

Remark. 1. From convergence in mean square, one can always extract an a.s. convergent subsequence.
2. The conclusion above extends in full generality to a.s. convergence, but an easy proof requires the reversed martingale convergence theorem, which we omit.
3. There is an easy extension to a.s. convergence under the extra restriction $\sum_n |\pi_n| < \infty$, using the Borel-Cantelli lemma and Chebychev's inequality.
4. If we consider the theorem over $[0, t + dt]$, $[0, t]$ and subtract, we can write the result formally as

$$(dW_t)^2 = dt.$$

This can be regarded either as a convenient piece of symbolism, or acronym, or as the essence of *Itô calculus* (Ch. III below).

Note.

The quadratic variation as defined above involves the limit of the quadratic variation over every sequence of partitions whose maximal subinterval length tends to zero. We stress that this is not the same as taking the supremum of the quadratic variation over *all* partitions – indeed, this would give ∞, rather than t (by the law of the iterated logarithm for Brownian motion). This second definition – strong quadratic variation – is the appropriate one in some contexts, such as Lyons' theory of rough paths, but we shall not need it, and quadratic variation will always be defined in the first sense here.

Suppose now we look at the ordinary variation $\sum |\Delta W(t)|$, rather than the quadratic variation $\sum (\Delta W(t))^2$. Then instead of $\sum (\Delta W(t))^2 \sim \sum \Delta t = t$, we get $\sum |\Delta W(t)| \sim \sum \sqrt{\Delta t}$. Now for Δt small, $\sqrt{\Delta t}$ is of a larger order of magnitude than Δt. So if $\sum \Delta t = t$ converges, $\sum \sqrt{\Delta t}$ diverges to $+\infty$. This gives:

Corollary (Lévy). The paths of Brownian motion are of unbounded variation – their variation is $+\infty$ on every interval.

Because of the above corollary, we will not be able to define integrals with respect to Brownian motion by a path-by-path procedure (for BM the relevant convergence in the above results in fact takes place with probability one). However, turning to the class of square-integrable continuous martingales $c\mathcal{M}^2$ (continuous square-integrable martingales), we find that these processes have finite quadratic variation, but all variations of higher order are zero and, except for trivial cases, all variations of lower order are infinite with positive probability. So quadratic variation is indeed the right variation to study. Returning to Brownian motion, we observe that for $s < t$,

$$E(W(t)^2|\mathcal{F}_s) = E([W(s) + (W(t) - W(s))]^2$$
$$= W(s)^2 + 2W(s)E[(W(t) - W(s))|\mathcal{F}_s] + E[(W(t) - W(s))^2|\mathcal{F}_s]$$
$$= W(s)^2 + 0 + (t - s).$$

So $W(t)^2 - t$ is a martingale. This shows that the quadratic variation is the adapted increasing process in the Doob-Meyer decomposition of W^2 (recall that W^2 is a nonnegative submartingale and thus can be written as the sum of a martingale and an adapted increasing process). This result extends to the class $c\mathcal{M}^2$ (and indeed to the broader class of local martingales – below).

Theorem. A martingale $M \in c\mathcal{M}^2$ is of finite quadratic variation $\langle M \rangle$, and $\langle M \rangle$ is the unique continuous increasing adapted process vanishing at zero with $M^2 - \langle M \rangle$ a martingale.

The quadratic variation result above leads to Lévy's 1948 result, the martingale characterization of Brownian motion. Recall that $W(t)$ is a continuous martingale with respect to its natural filtration $(\mathcal{F}_t)$ and with quadratic variation t. There is a remarkable converse, due to Lévy:

Theorem (Lévy's Martingale Characterization of BM). If M is any continuous, square-integrable (local) $(\mathcal{F}_t)$-martingale with $M(0) = 0$ and quadratic variation t, then M is an $(\mathcal{F}_t)$-Brownian motion.

Expressed differently this is:
If M is any continuous, square-integrable (local) $(\mathcal{F}_t)$-martingale with $M(0) = 0$ and $M(t)^2 - t$ a martingale, then M is an $(\mathcal{F}_t)$-Brownian motion.

In view of the fact that $\langle W \rangle(t) = t$, a further useful fact about Brownian motion may be guessed: If M is a continuous martingale then there exists a Brownian motion $W(t)$ such that $M(t) = W(\langle M \rangle(t))$, i.e. the martingale M can be transformed into a Brownian motion by a random time-change. These results already imply that Brownian motion is the fundamental continuous martingale.

10. Properties of Brownian Motion

Brownian Scaling.

For any $c > 0$, write

$$W_c(t) := c^{-1} W(c^2 t), \qquad t \geq 0$$

with W BM. Then W_c is Gaussian, with mean 0, variance $c^{-2} \times c^2 t = t$ and covariance

$$cov(W_c(s), W_c(t)) = c^{-2} E(W_c(s), W_c(t)) = c^{-2} \min(c^2 s, c^2 t)$$
$$= \min(s, t) = cov(W(s), W(t)).$$

Also W_c has continuous paths, as W does. So W_c has all the properties of Brownian motion. So, W_c *is* Brownian motion. It is said to be derived from W by *Brownian scaling* with *scale-factor* $c > 0$. Since

$$(W(ut) : t \geq 0) = (\sqrt{u} W(t) : t \geq 0) \qquad \text{in law, } \forall u > 0,$$

W is called *self-similar* with *index* 1/2. Brownian motion is thus a *fractal.* A piece of Brownian path, looked at under a microscope, still looks Brownian, however much we 'zoom in and magnify'. Of course, the contrast with a function f with some smoothness is stark: a differentiable function begins to look straight under repeated zooming and magnification, because it has a tangent.

Time-Inversion.

Write

$$X_t := tW(1/t).$$

Then X has mean 0 and covariance

$$cov(X_s, X_t) = st.cov(B(1/s), B(1/t)) = st. \min(1/s, 1/t) = \min(t, s) = \min(s, t).$$

Since X has continuous paths also, as above, X *is* Brownian motion. We say that X is obtained from W by *time-inversion.* This property is useful in transforming properties of BM 'in the large' ($t \to \infty$) to properties 'in the small', or local properties ($t \to 0$). For example, one can translate the law of the iterated logarithm (LIL) from global to local form.

Zero set Z of Brownian motion

This has many interesting properties. Z is:
closed (by continuity of BM);
perfect – contains all its limit point (again, by continuity of BM – these two give a sense in which Z is 'big' topologically);
uncountable (this gives a sense in which Z is big from the point of view of cardinality);
Lebesgue-null (this gives a sense in which Z is small from the point of view of Measure Theory);
a fractal – self-similar of index $\frac{1}{2}$ (from Brownian scaling). Indeed, Z has *Hausdorff dimension* $\frac{1}{2}$ (which gives a precise sense in which Z is 'half-dimensional' – and the right way in which to assess the 'size' of Z).

If BM starts at 0, with probability 1 there are *infinitely many* zeros of BM in every time-interval $(0, \epsilon)$, however small $\epsilon > 0$ is. And by the strong Markov property, if we start BM afresh at each of these, it behaves like $t = 0$ above (so there are infinitely many zeros in every interval to its right ...). One is left wondering how BM can ever escape from zero – but it does! The relevant theory here is Itô's theory of *Brownian excursions* (more generally, of excursions of a Markov process) of 1971 – but this would take us too far afield here.

III. Stochastic integration; Itô calculus.

1. Stochastic Integration

Stochastic integration was introduced by K. Itô in 1944, hence its name Itô calculus. It gives a meaning to

$$\int_0^t XdY = \int_0^t X(s,\omega)dY(s,\omega),$$

for suitable stochastic processes X and Y, the integrand and the integrator. We shall confine our attention here mainly to the basic case with integrator Brownian motion: $Y = W$. Much greater generality is possible; see e.g. [P] for details.

The first thing to note is that stochastic integrals with respect to Brownian motion, if they exist, must be quite different from the measure-theoretic integral of Ch. I. For, the Lebesgue-Stieltjes integrals described there have as integrators the difference of two monotone (increasing) functions, which are locally of finite variation. But we know from Ch. II that Brownian motion is of infinite (unbounded) variation on every interval. So Lebesgue-Stieltjes and Itô integrals must be fundamentally different.

In view of the above, it is quite surprising that Itô integrals can be defined at all. But if we take for granted Itô's fundamental insight that they can be, it is obvious how to begin and clear enough how to proceed. We begin with the simplest possible integrands X, and extend successively much as we extended the measure-theoretic integral of Ch. I.

Indicators.

If $X(t,\omega) = I_{[a,b]}(t)$, there is exactly one plausible way to define $\int XdW$:

$$\int_0^t X(s,\omega)dW(s,\omega) := \begin{cases} 0 & \text{if } t \le a, \\ W(t) - W(a) & \text{if } a \le t \le b, \\ W(b) - W(a) & \text{if } t \ge b. \end{cases}$$

Simple Functions.

Extend by linearity: if X is a linear combination of indicators, $X = \sum_{i=1}^n c_i I_{[a_i,b_i]}$, we should define

$$\int_0^t XdW := \sum_{i=1}^n c_i \int_0^t I_{[a_i,b_i]}dW.$$

Already one wonders how to extend this from constants c_i to suitable random variables, and one seeks to simplify the obvious but clumsy three-line expressions above.

We begin again, this time calling a stochastic process X *simple* if there is a partition $0 = t_0 < t_1 < \ldots < t_n = T < \infty$ and uniformly bounded $\mathcal{F}_{t_n}$-measurable random variables ξ_k ($|\xi_k| \leq C$ for all $k = 0, \ldots, n$ and ω, for some C) and if $X(t, \omega)$ can be written in the form

$$X(t,\omega) = \xi_0(\omega) I_{\{0\}}(t) + \sum_{i=0}^{n} \xi_i(\omega) I_{(t_i, t_{i+1}]}(t) \qquad (0 \leq t \leq T, \omega \in \Omega).$$

Then if $t_k \leq t < t_{k+1}$,

$$\begin{aligned} I_t(X) := \textstyle\int_0^t X dW &= \textstyle\sum_{i=0}^{k-1} \xi_i W(t_{i+1}) - W(t_i)) + \xi_k (W(t) - W(t_k)) \\ &= \textstyle\sum_{i=0}^{n} \xi_i (W(t \wedge t_{i+1}) - W(t \wedge t_i)). \end{aligned}$$

Note that by definition $I_0(X) = 0$ a.s. We collect some properties of the stochastic integral defined so far:

Lemma. (i) $I_t(aX + bY) = aI_t(X) + bI_t(Y)$.
(ii) $E(I_t(X)|\mathcal{F}_s) = I_s(X) \quad a.s. \quad (0 \leq s < t < \infty)$, hence $I_t(X)$ is a *continuous martingale*.

Proof. (i) follows from the fact that linear combinations of simple functions are simple.
(ii) There are two cases to consider.
(a) Both s and t belong to the same interval $[t_k, t_{k+1})$. Then

$$I_t(X) = I_s(X) + \xi_k (W(t) - W(s)).$$

But ξ_k is $\mathcal{F}_{t_k}$-measurable, so $\mathcal{F}_s$-measurable ($t_k \leq s$), so independent of $W(t) - W(s)$ (independent increments property of W). So

$$E(I_t(X)|\mathcal{F}_s) = I_s(X) + \xi_k E(W(t) - W(s)|\mathcal{F}_s) = I_s(X).$$

(b) $s < t$ belongs to a different interval from t: $s \in [t_m, t_{m+1})$ for some $m < k$. Then

$$I_t(X) = I_s(X) + \xi_m (W(t_{m+1}) - W(s)) + \sum_{i=m+1}^{k-1} \xi_i (W(t_{i+1}) - W(t_i)) + \xi_k (W(t) - W(t_k))$$

(if $k = m+1$, the sum on the right is empty, and does not appear). Take $E(.|\mathcal{F}_s)$ on the right. The first term gives $I_s(X)$. The second gives $\xi_m E[(W(t_{m+1}) - W(s))|\mathcal{F}_s] = \xi_m \cdot 0 = 0$, as ξ_m is $\mathcal{F}_s$-measurable, and similarly so do the third and fourth, completing the proof. //

Note. The stochastic integral for simple integrands is essentially a martingale transform.

We pause to note a property of square-integrable martingales which we shall need below. Call $M(t) - M(s)$ the increment of M over $(s,t]$. Then for a martingale M, the product of the increments over disjoint intervals has zero mean. For, if $s < t \leq u < v$,

$$E[(M(v)-M(u))(M(t)-M(s))] = E[E((M(v)-M(u))(M(t)-M(s))|\mathcal{F}_u)]$$

$$= E[(M(t) - M(s))E((M(v) - M(u))|\mathcal{F}_u)],$$

taking out what is known (as $s, t \leq u$). The inner expectation is zero by the martingale property, so the left-hand side is zero, as required.

We now can add further properties of the stochastic integral for simple functions X.

Lemma. (i) We have the Itô isometry

$$E[(I_t(X))^2], \text{ or } E[(\int_0^t XdW)^2], = E(\int_0^t X(s)^2 ds).$$

(ii) $E((I_t(X) - I_s(X))^2|\mathcal{F}_s) = E(\int_s^t X(u)^2 du)$ a.s.

Proof. We only show (i); the proof of (ii) is similar. The left-hand side in (i) above is $E(I_t(X) \cdot I_t(X))$, i.e.

$$E([\sum_{i=0}^{k-1} \xi_i(W(t_{i+1}) - W(t_i)) + \xi_k(W(t) - W(t_k))]^2).$$

Expanding out the square, the cross-terms have expectation zero by above, leaving

$$E(\sum_{i=0}^{k-1} \xi_i^2(W(t_{i+1}) - W(t_i))^2 + \xi_k^2(W(t) - W(t_k))^2).$$

Since ξ_i is $\mathcal{F}_{t_i}$-measurable, each ξ_i^2-term is independent of the squared Brownian increment term following it, which has expectation $var(W(t_{i+1})-W(t_i)) = t_{i+1} - t_i$. So we obtain

$$\sum_{i=0}^{k-1} E(\xi_i^2)(t_{i+1} - t_i) + E(\xi_k^2)(t - t_k).$$

This is (using Fubini's theorem) $\int_0^t E(X(u)^2)du = E(\int_0^t X(u)^2 du)$, as required. //

The Itô isometry above suggests that $\int_0^t X dW$ should be defined only for processes with

$$\int_0^t E(X(u)^2)du < \infty \qquad \text{for all } t. \tag{$*$}$$

We then can transfer convergence on a suitable L^2-space of stochastic processes to a suitable L^2-space of martingales. This gives us an L^2-theory of stochastic integration, for which Hilbert-space methods are available.

Approximation.

By analogy with the integral of Ch. I, we seek a class of integrands suitably approximable by simple integrands. It turns out that:

(i) The suitable class of integrands is the class of $(\mathcal{B}([0,\infty)) \times \mathcal{F})$-measurable, $(\mathcal{F}_t)$- adapted processes X with $\int_0^t E(X(u)^2)du < \infty$ for all $t > 0$.

(ii) Each such X may be approximated by a sequence of simple integrands X_n so that the stochastic integral $I_t(X) = \int_0^t X dW$ may be defined as the limit of $I_t(X_n) = \int_0^t X_n dW$.

(iii) The properties from both lemmas above remain true for the stochastic integral $\int_0^t X dW$ defined by (i) and (ii).

We must omit detailed proofs of these assertions here. The key technical ingredients needed are Hilbert-space methods in spaces defined by integrals related to the quadratic variation of the integrator (which is just t in our Brownian motion setting here) and the Kunita-Watanabe inequalities ([P], 61).

Without $(*)$, the stochastic integral need not yield a mg, but only a *local martingale*. This is a process M such that there exists a sequence of stopping times $T_n \uparrow +\infty$ such that each of the stopped and shifted processes $M^{T_n} - M_0$ is a (true) martingale. Local mgs are much more general than (true) mgs. They are used to define *semi-martingales* – sums of a local mg and a FV process; these are the most general *stochastic integrators*.