ull4b.tex **Week 4: pm, 17.10.2018**

§3. Brownian Motion.

The Scottish botanist Robert Brown observed pollen particles in suspension under a microscope in 1828 (though others had observed the phenomenon before him),[1] and saw that they were in constant irregular motion.

In 1900 L. Bachelier considered Brownian motion a possible model for stock-market prices:

BACHELIER, L. (1900): Théorie de la spéculation. *Ann. Sci. Ecole Normale Supérieure* **17**, 21-86

– the first time Brownian motion had been used to model financial or economic phenomena, and before a mathematical theory had been developed.

In 1905 Albert Einstein considered Brownian motion as a model of particles in suspension, and used it to estimate *Avogadro's number* ($N \sim 6 \times 10^{23}$), based on the diffusion coefficient D in the *Einstein relation*

$$var X_t = Dt \qquad (t > 0).$$

In 1923 Norbert WIENER defined and constructed Brownian motion rigorously for the first time. The resulting stochastic process is often called the *Wiener process* in his honour, and its probability measure (on path-space) is called *Wiener measure*.

We define *standard Brownian motion* on $\mathbb{R}$, BM or $BM(\mathbb{R})$, to be a stochastic process $X = (X_t)_{t \geq 0}$ such that

1. $X_0 = 0$,
2. X has *independent increments*: $X_{t+u} - X_t$ is independent of $\sigma(X_s : s \leq t)$ for $u \geq 0$,
3. X has *stationary increments*: the law of $X_{t+u} - X_t$ depends only on u,
4. X has *Gaussian increments*: $X_{t+u} - X_t$ is normally distributed with mean 0 and variance u,

$$X_{t+u} - X_t \sim N(0, u),$$

5. X has *continuous paths*: X_t is a continuous function of t, i.e. $t \mapsto X_t$ is continuous in t.

For time t in a finite interval – $[0, 1]$, say – we use the following filtered space:

$\Omega = C[0, 1]$, the space of all continuous functions on $[0, 1]$.

[1]The Roman author Lucretius observed this phenomenon in gases – dust particles dancing in sunbeams – in antiquity: *De rerum natura* [The nature of things], c. 50 BC.

The points $\omega \in \Omega$ are thus random functions, and we use the coordinate mappings: X_t, or $X_t(\omega)$, $= \omega_t$.
The filtration is given by $\mathcal{F}_t := \sigma(X_s : 0 \leq s \leq t)$, $\mathcal{F} := \mathcal{F}_1$.
P is the measure on $(\Omega, \mathcal{F})$ with finite-dimensional distributions specified by the restriction that the increments $X_{t+u} - X_t$ are stationary independent Gaussian $N(0, u)$.

Theorem (WIENER, 1923). Brownian motion exists.

The best way to prove this is by construction, and one that reveals some properties. The proof that follows is originally due to Paley, Wiener and Zygmund (1933) and Lévy (1948), but is re-written in the modern language of *wavelet* expansions. We omit details; for these, see e.g. [BK] 5.3.1, or SP L20-22. The Haar system $(H_n) = (H_n(.))$ is a complete orthonormal system (cons) of functions in $L^2[0,1]$. The Schauder System (Δ_n) is obtained by integrating the Haar system. Consider the triangular ('tent') function:

$$\Delta(t) = \begin{cases} 2t & \text{on } [0, \frac{1}{2}), \\ 2(1-t) & \text{on } [\frac{1}{2}, 1], \\ 0 & \text{else.} \end{cases}$$

Write $\Delta_0(t) := t$, $\Delta_1(t) := \Delta(t)$ ('mother wavelet'), and define the nth *Schauder function* Δ_n ('daughter wavelets') by 'dilation and translation':

$$\Delta_n(t) := \Delta(2^j t - k) \qquad (n = 2^j + k \geq 1).$$

Note that Δ_n has support $[k/2^j, (k+1)/2^j]$ (so is 'localized' on this dyadic interval, which is small for n, j large). Then

$$\int_0^t H(u)du = \frac{1}{2}\Delta(t),$$

and

$$\int_0^t H_n(u)du = \lambda_n \Delta_n(t),$$

where $\lambda_0 = 1$ and for $n \geq 1$,

$$\lambda_n = \frac{1}{2} \times 2^{-j/2} \qquad (n = 2^j + k \geq 1).$$

The Schauder system (Δ_n) is again a complete orthogonal system on $L^2[0,1]$. We can now formulate the next result; for proof, see the references above.

Theorem (PWZ theorem: Paley-Wiener-Zygmund, 1933). For $(Z_n)_0^\infty$ independent $N(0,1)$ random variables, λ_n, Δ_n as above,

$$W_t := \sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(t)$$

converges uniformly on $[0,1]$, a.s. The process $W = (W_t : t \in [0,1])$ is Brownian motion.

Thus the above description does indeed define a stochastic process $X = (X_t)_{t\in[0,1]}$ on $(C[0,1], \mathcal{F}, (\mathcal{F}_t), P)$. The construction gives X on $C[0,n]$ for each $n = 1, 2, \cdots$, and combining these: X exists on $C[0,\infty)$. It is also unique (a stochastic process is uniquely determined by its finite-dimensional distributions and the restriction to path-continuity).

No construction of Brownian motion is easy: one needs both some work and some knowledge of measure theory. However, *existence* is really all we need, and this we shall take for granted. For background, see any measure-theoretic text on stochastic processes. The classic is Doob's book, quoted above (see VIII.2 there). Excellent modern texts include Karatzas & Shreve [KS] (see particularly §2.2-4 for construction and §5.8 for applications to economics), Revuz & Yor [RY], Rogers & Williams [RW1] (Ch. 1), [RW2] Itô calculus – below).

We shall henceforth denote standard Brownian motion $BM(\mathbb{R})$ – or just BM for short – by $B = (B_t)$ (B for Brown), though $W = (W_t)$ (W for Wiener) is also common. Standard Brownian motion $BM(\mathbb{R}^d)$ in d dimensions is defined by $B(t) := (B_1(t), \cdots, B_d(t))$, where $B_1, \cdots, B_d$ are *independent* standard Brownian motions in one dimension (*independent copies* of $BM(\mathbb{R})$).

Zeros.

It can be shown that Brownian motion *oscillates*:

$$\limsup_{t\to\infty} X_t = +\infty, \qquad \liminf_{t\to\infty} X_t = -\infty \qquad a.s.$$

Hence, for every n there are zeros (times t with $X_t = 0$) of X with $t \geq n$

(indeed, infinitely many such zeros). So if

$$Z := \{t \geq 0 : X_t = 0\}$$

denotes the zero-set of $BM(\mathbb{R})$:
1. Z is an *infinite* set.
Next, if t_n are zeros and $t_n \to t$, then by path-continuity $B(t_n) \to B(t)$; but $B(t_n) = 0$, so $B(t) = 0$:
2. Z is a *closed* set (Z contains its limit points).
Less obvious are the next two properties:
3. Z is a *perfect* set: every point $t \in Z$ is a limit point of points in Z. So there are *infinitely many* zeros in *every* neighbourhood of *every* zero (so the paths must oscillate amazingly fast!).
4. Z is a (Lebesgue) *null* set: Z has Lebesgue measure zero.

In particular, *any* diagram one attempts to draw of Brownian motion grossly distorts Z: *it is impossible to draw a realistic picture of a Brownian path.*

Brownian Scaling.

For each $c \in (0, \infty)$, $X(c^2 t)$ is $N(0, c^2 t)$, so $X_c(t) := c^{-1} X(c^2 t)$ is $N(0, t)$. Thus X_c has all the defining properties of a Brownian motion (Problems 9 Q2). So, X_c **IS** a Brownian motion:

Theorem. If X is BM and $c > 0$, $X_c(t) := c^{-1} X(c^2 t)$, then X_c is again a BM.

Corollary. X is *self-similar* (reproduces itself under scaling), so a Brownian path $X(.)$ is a *fractal*. So too is the zero-set Z.

Brownian motion owes part of its importance to belonging to *all* the important classes of stochastic processes: it is (strong) Markov, a (continuous) martingale, Gaussian, a diffusion, a Lévy process (process with stationary independent increments), etc.

Simulation of Brownian motion.

By the PWZ Theorem, all we need to simulate BM is a sequence of independent standard normal Z_n (how many depends on our required degree of accuracy). The most basic simulation is from the uniform distribution on

$(0,1)$, $U(0,1)$, directly from a random number generator. We can then use the Probability Integral Transformation to transform $U \sim U(0,1)$ to standard normal, $Z := \Phi(U) \sim N(0,1)$. But there is no explicit form for Φ. Because of this, it is easier to use the *Box-Muller method*: use plane polar coordinates, and generate pairs of standard normals. See e.g. my homepage, Introductory Statistics, I (under SMF: Statistical Methods for Finance).

§4. Quadratic Variation (QV) of Brownian Motion; Itô's Lemma

Recall that for ξ $N(\mu, \sigma^2)$, ξ has moment-generating function (MGF)

$$M(t) := E\exp\{t\xi\} = \exp\{\mu t + \frac{1}{2}\sigma^2 t^2\}.$$

Take $\mu = 0$ below; for ξ $N(0, \sigma^2)$,

$$\begin{aligned} M(t) := E\exp\{t\xi\} &= \exp\{\frac{1}{2}\sigma^2 t^2\} \\ &= 1 + \frac{1}{2}\sigma^2 t^2 + \frac{1}{2!}\left(\frac{1}{2}\sigma^2 t^2\right)^2 + O(t^6) \\ &= 1 + \frac{1}{2!}\sigma^2 t^2 + \frac{3}{4!}\sigma^4 t^4 + O(t^6). \end{aligned}$$

So as the Taylor coefficients of the MGF are the moments (hence the name MGF!),

$$E[\xi^2] = var\xi = \sigma^2, \qquad E[\xi^4] = 3\sigma^4, \qquad \text{so} \qquad var(\xi^2) = E[\xi^4] - (E[\xi^2])^2 = 2\sigma^4.$$

For B BM, this gives in particular

$$E[B_t] = 0, \qquad var B_t = t, \qquad E[(B_t)^2] = t, \qquad var[(B_t)^2] = 2t^2.$$

In particular, for $t > 0$ *small*, this shows that the variance of B_t^2 is negligible compared with its expected value. Thus, the *randomness* in B_t^2 is negligible compared to its mean for t small.

This suggests that if we take a fine enough partition $\mathcal{P}$ of $[0,T]$ – a finite set of points

$$0 = t_0 < t_1 < \cdots < t_k = T$$

with $|\mathcal{P}| := \max |t_i - t_{i-1}|$ small enough – then writing

$$\Delta B(t_i) := B(t_i) - B(t_{i-1}), \qquad \Delta t_i := t_i - t_{i-1},$$

$\Sigma(\Delta B(t_i))^2$ will closely resemble $\Sigma E[(\Delta B(t_i)^2]$, which is $\Sigma\Delta t_i = \Sigma(t_i - t_{i-1}) = T$. This is in fact true a.s.:

$$\Sigma(\Delta B(t_i))^2 \to \Sigma\Delta t_i = T \qquad \text{as} \qquad \max|t_i - t_{i-1}| \to 0.$$

This limit is called the *quadratic variation* V_T^2 of B over $[0, T]$:

Theorem (Lévy). The quadratic variation of a Brownian path over $[0, T]$ exists and equals T, a.s.

For details of the proof, see e.g. [BK], §5.3.2, SP L22, SA L7,8.

If we increase t by a small amount to $t + dt$, the increase in the QV can be written symbolically as $(dB_t)^2$, and the increase in t is dt. So, formally we may summarise the theorem as

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

Suppose now we look at the *ordinary* variation $\Sigma|\Delta B_t|$, rather than the *quadratic* variation $\Sigma(\Delta B_t)^2$. Then instead of $\Sigma(\Delta B_t)^2 \sim \Sigma\Delta t \sim t$, we get $\Sigma|\Delta B_t| \sim \Sigma\sqrt{\Delta t}$. Now for Δt small, $\sqrt{\Delta t}$ is of a larger order of magnitude that Δt. So if $\Sigma\Delta t = t$ converges, $\Sigma\sqrt{\Delta t}$ diverges to $+\infty$. This suggests – what is in fact true – the

Corollary. The paths of Brownian motion are of infinite variation - their variation is $+\infty$ on every interval, a.s.

The QV result above leads to Lévy's 1948 result, the Martingale Characterization of BM. Recall that B_t is a continuous martingale with respect to its natural filtration $(\mathcal{F}_t)$ and with QV t. There is a remarkable converse; we give two forms.

Theorem (Lévy; Martingale Characterization of Brownian Motion). If M is any continuous local $(\mathcal{F}_t)$-martingale with $M_0 = 0$ and quadratic variation t, then M is an $(\mathcal{F}_t)$-Brownian motion.

Theorem (Lévy). If M is any continuous $(\mathcal{F}_t)$-martingale with $M_0 = 0$ and $M_t^2 - t$ a martingale, then M is an $(\mathcal{F}_t)$-Brownian motion.

For proof, see e.g. [RW1], I.2. Observe that for $s < t$,

$$B_t^2 = [B_s + (B_t - B_s)]^2 = B_s^2 + 2B_s(B_t - B_s) + (B_t - B_s)^2,$$

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

$$E[B_t^2 - t|\mathcal{F}_s] = B_s^2 - s:$$

$B_t^2 - t$ is a martingale.

Quadratic Variation (QV).

The theory above extends to *continuous* martingales (bounded continuous martingales in general, but we work on a finite time-interval $[0, T]$, so continuity implies boundedness). We quote (for proof, see e.g. [RY], IV.1):

Theorem. A continuous martingale M 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.

Corollary. A continuous martingale M has infinite variation.

Quadratic Covariation. We write $\langle M, M \rangle$ for $\langle M \rangle$, and extend $\langle \ \rangle$ to a bilinear form $\langle .,. \rangle$ with two different arguments by the *polarization identity*:

$$\langle M, N \rangle := \frac{1}{4}(\langle M + N, M + N \rangle - \langle M - N, M - N \rangle).$$

If N is of *finite* variation, $M \pm N$ has the same QV as M, so $\langle M, N \rangle = 0$.

Itô's Lemma. We discuss Itô's Lemma in more detail in V.6 below; we pause here to give the link with quadratic variation and covariation. We quote: if $f(t, x_1, \cdots, x_d)$ is C^1 in its zeroth (time) argument t and C^2 in its remaining d space arguments x_i, and $M = (M^1, \cdots, M^d)$ is a continuous vector martingale, then (writing f_i, f_{ij} for the first partial derivatives of f with respect to its ith argument and the second partial derivatives with respect to the ith and jth arguments) $f(M_t)$ has stochastic differential

$$df(M_t) = f_0(M)dt + \Sigma_{i=1}^d f_i(M_t)dM_t^i + \frac{1}{2}\Sigma_{i,j=1}^d f_{ij}(M_t)d\langle M^i, M^j \rangle_t.$$

Integration by Parts. If $f(t, x_1, x_2) = x_1 x_2$, we obtain

$$d(MN)_t = NdM_t + MdN_t + d\langle M, N \rangle_t$$

(no half: two terms, (M, N) and (N, M). Similarly for stochastic integrals (defined below): if $Z_i := \int H_i dM_i$ $(i = 1, 2)$, $d\langle Z_1, Z_2\rangle = H_1 H_2 d\langle M_1, M_2\rangle$.
Note. The integration-by-parts formula – a special case of Itô's Lemma, as above – is in fact *equivalent* to Itô's Lemma: either can be used to derive the other. Rogers & Williams [RW1, IV.32.4] describe the integration-by-parts formula/Itô's Lemma as 'the cornerstone of stochastic calculus'.
Fractals Everywhere.

As we saw, a Brownian path is a *fractal* – a *self-similar* object. So too is its zero-set Z. Fractals were studied, named and popularised by the French mathematician Benôit B. Mandelbrot (1924-2010). See his books, and Michael F. Barnsley: *Fractals everywhere*. Academic Press, 1988.

Fractals *look the same at all scales* – diametrically opposite to the familiar functions of Calculus. In Differential Calculus, a differentiable function has a tangent; this means that locally, its graph *looks straight*; similarly in Integral Calculus. While most continuous functions we encounter are differentiable, at least piecewise (i.e., except for 'kinks'), there is a sense in which the typical, or generic, continuous function is *nowhere differentiable*. Thus Brownian paths may look pathological at first sight – but in fact they are typical!
Hedging in continuous time.

Imagine hedging an option in continuous time. In discrete time, this involves repeatedly rebalancing our portfolio between cash and stock; in continuous time, this has to be done continuously. The relevant stochastic processes (Ch. VI) are *geometric Brownian motion (GBM)*, relatives of BM, which, like BM, have *infinite variation* (finite QV). This makes the rebalancing problematic – indeed, impossible in these terms. Analogy: a cyclist has to rebalance continuously, but does so smoothly, not with infinite variation! Or, think of continuous-time control of a manned space-craft (Kalman filter). In practice, hedging has to be done discretely (as in Ch. IV). Or, we can use price processes with *jumps* (Ch. VI) – finite variation, but now the markets are incomplete.

In reality, markets have *transaction costs* (a form of *market friction* – see Ch. I). So even if we could rebalance continuously, we wouldn't: the transaction costs would in principle be infinite, and in practice make it uneconomic.

§5. Stochastic Integrals (Itô Calculus)

Stochastic integration was introduced by K. ITÔ in 1944, hence its name Itô calculus. It gives a meaning to $\int_0^t X dY = \int_0^t X_s(\omega) dY_s(\omega)$, for suitable

stochastic processes X and Y, the *integrand* and the *integrator*. We shall confine our attention here to the basic case with integrator Brownian motion: $Y = B$. Much greater generality is possible: for Y a continuous martingale, see [KS] or [RY]; for a systematic general treatment, see
MEYER, P.-A. (1976): Un cours sur les intégrales stochastiques. *Séminaire de Probabilités X: Lecture Notes on Math.* **511**, 245-400, Springer.

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 II.2. For, the Lebesgue-Stieltjes integrals described there have as integrators the difference of two monotone (increasing) functions (by Jordan's theorem), which are locally of *finite (bounded) variation, FV*. But we know from V.4 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 follow the procedure of Ch. II. We begin with the simplest possible integrands X, and extend successively much as we extended the measure-theoretic integral of Ch. II.

1. *Indicators.*
If $X_t(\omega) = I_{[a,b]}(t)$, there is exactly one plausible way to define $\int XdB$:

$$\int_0^t XdB, \quad \text{or} \quad \int_0^t X_s(\omega)dB_s(\omega), := \begin{cases} 0 & \text{if } t \le a, \\ B_t - B_a & \text{if } a \le t \le b, \\ B_b - B_a & \text{if } t \ge b. \end{cases}$$

2. *Simple functions.* Extend by linearity: if X is a linear combination of indicators, $X = \Sigma c_i I_{[a_i,b_i]}$, we should define

$$\int_0^t XdB := \Sigma c_i \int_0^t I_{[a_i,b_i]}dB.$$

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. It turns out that finite sums are not essential: one can have infinite sums, but now we take the c_i uniformly bounded.

We begin again, this time calling a *stochastic process X simple* if there is an infinite sequence

$$0 = t_0 < t_1 < \cdots < t_n < \cdots \to \infty$$

and uniformly bounded $\mathcal{F}_{t_n}$-measurable random variables ξ_n ($|\xi_n| \leq C$ for all n and ω, for some C) such that $X_t(\omega)$ can be written in the form

$$X_t(\omega) = \xi_0(\omega) I_{\{0\}}(t) + \Sigma_{i=0}^{\infty} \xi_i(\omega) I_{(t_i, t_{i+1}]}(t) \qquad (0 \leq t < \infty, \omega \in \Omega).$$

The only definition of $\int_0^t XdB$ that agrees with the above for finite sums is, if n is the unique integer with $t_n \leq t < t_{n+1}$,

$$\begin{aligned} I_t(X) := \int_0^t XdB &= \Sigma_0^{n-1} \xi_i (B(t_{i+1}) - B(t_i)) + \xi_n (B(t) - B(t_n)) \\ &= \Sigma_0^{\infty} \xi_i (B(t \wedge t_{i+1}) - B(t \wedge t_i)) \qquad (0 \leq t < \infty). \end{aligned}$$

We note here some properties of the stochastic integral defined so far:

A. $I_0(X) = 0 \qquad P - a.s.$

B. *Linearity.* $I_t(aX + bY) = aI_t(X) + bI_t(Y)$.
Proof. Linear combinations of simple functions are simple.

C. $E[I_t(X)|\mathcal{F}_s] = I_s(X) \qquad P - a.s. \qquad (0 \leq s < t < \infty):$
$I_t(X)$ is a *continuous martingale.*
Proof. There are two cases to consider.
(i) Both s and t belong to the same interval $[t_n, t_{n+1})$. Then

$$I_t(X) = I_s(X) + \xi_n (B(t) - B(s)).$$

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

$$E[I_t(X)|\mathcal{F}_s] = I_s(X) + \xi_n E[B(t) - B(s)|\mathcal{F}_s] = I_s(X).$$

(ii) $s < t$ and t belong to different intervals: $s \in [t_m, t_{m+1})$ for $m < n$. Then

$$\begin{aligned} E[I_t(X)|\mathcal{F}_s] &= E[E[I_t(X)|\mathcal{F}_{t_n}]|\mathcal{F}_s] \qquad \text{(iterated conditional expectations: } s < t_n) \\ &= E[I_{t_n}(X)|\mathcal{F}_s], \end{aligned}$$

by (i). Now $I_{t_n}(X)$ is $I_{t_{m+1}}(X)$ plus a sum of products of ξs and the Brownian increment over the next interval. Taking $E[.|\mathcal{F}_s]$, the terms in the sum contribute 0, again as in (i) above. This reduces to t_{m+1}:

$$E[I_t(X)|\mathcal{F}_s] = E[I_{t_{m+1}}(X)|\mathcal{F}_s].$$

This reduces to case (i). //

Note. The stochastic integral for simple integrands is essentially a martingale transform, and the above is essentially the proof of Ch. III that martingale transforms are martingales.

We pause to note a property of martingales which we shall need below. Call $X_t - X_s$ the *increment* of X over $(s,t]$. Then for a *martingale* X (square-integrable, i.e. in L_2, to make the expectations below well-defined – see below), *the product of the increments over disjoint intervals has zero mean.* For, if $s < t \le u < v$,

$$\begin{aligned} E[(X_v - X_u)(X_t - X_s)] &= E[E[(X_v - X_u)(X_t - X_s)|\mathcal{F}_u]] \\ &= E[(X_t - X_s)E[(X_v - X_u)|\mathcal{F}_u]], \end{aligned}$$

taking out what is known (as $s, t \le u$). The inner expectation is zero by the martingale property, so the LHS is zero, as required.

D (*Itô isometry*). $E[(I_t(X))^2]$, or $E[(\int_0^t X_s dB_s)^2]$, $= E\int_0^t X_s^2 ds$.
Proof. The LHS above is $E[I_t(X).I_t(X)]$, i.e.

$$E[(\Sigma_{i=0}^{n-1}\xi_i(B(t_{i+1}) - B(t_i)) + \xi_n(B(t) - B(t_n)))^2].$$

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

$$E[\Sigma_{i=0}^{n-1}\xi_i^2(B(t_{i+i} - B(t_i))^2 + \xi_n^2(B(t) - B(t_n))^2].$$

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 $var(B(t_{i+1}) - B(t_i)) = t_{i+1} - t_i$. So we obtain

$$\Sigma_{i=0}^{n-1}E[\xi_i^2](t_{i+1} - t_i) + E[\xi_n^2](t - t_n).$$

This is $\int_0^t E[X_u^2]du = E\int_0^t X_u^2 du$, as required.

E. *Itô isometry (continued).* $I_t(X) - I_s(X) = \int_s^t X_u dB_u$ satisfies

$$E[(\int_s^t X_u dB_u)^2] = E[\int_s^t X_u^2 du] \qquad P - a.s.$$

Proof: as above.

F. *Quadratic variation.* The QV of $I_t(X) = \int_0^t X_u dB_u$ is $\int_0^t X_u^2 du$.

This is proved in the same way as the case $X \equiv 1$, that B has quadratic variation process t.

Integrands.

The properties above suggest that $\int_0^t X dB$ should be defined only for processes with

$$\int_0^t EX_u^2 du < \infty \qquad \text{for all} \quad t.$$

We shall restrict attention to such X in what follows. This gives us an L_2-theory of stochastic integration (compare the L_2-spaces introduced in Ch. II), for which Hilbert-space methods are available.

3. *Approximation.*

Recall steps 1 (indicators) and 2 (simple integrands). By analogy with the integral of Ch. II, we seek a suitable class of integrands suitably approximable by simple integrands. It turns out that:

(i) The suitable class of integrands is the class of left-continuous adapted processes X with $\int_0^t EX_u^2 du < \infty$ for all $t > 0$ (or all $t \in [0, T]$ with finite time-horizon T, as here),

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

(iii) The stochastic integral $\int_0^t X dB$ so defined still has properties A-F above.

It is not possible to include detailed proofs of these assertions in a course of this type [recall that we did not construct the measure-theoretic integral of Ch. II in detail either – and this is harder!]. The key technical ingredient needed is the *Kunita-Watanabe inequalities*. See e.g. [KS], §§3.1-2.