MATH 56A: STOCHASTIC PROCESSES
CHAPTER 9

9. Stochastic integration

I will continue with the intuitive description of stochastic integrals that I started last week.

9.0. the idea.

I already talked about the probabilitistic and analytic approach to Brownian motion. Stochastic integrals combine these methods. A key idea is Lévy's *quadratic variation* which is used in Kunita and Watanabe's [2] reformulation of stochastic integration.

9.0.1. *quadratic variation.* We want to define the stochastic integral

$$Z_t = \int_0^t Y_s\, dX_s$$

where X_s is Brownian motion in $\mathbb{R}^1$. However, there is a big problem because dX_s has *unbounded variation*[1]. In other words,

$$\int |dX_s| := \lim_{\delta t \to 0} \sum |X_{t_{i+1}} - X_{t_i}| = \infty.$$

Fortunately, we can still define the stochastic integral because the "quadratic variation" of X_t (denoted by $\langle X \rangle_t$) is bounded:

Theorem 9.1 (Lévy)**.**

$$\langle X \rangle_t = \int_0^t (dX_s)^2 := \lim_{\delta t \to 0} \sum (X_{t_{i+1}} - X_{t_i})^2 = t$$

with probability one.

Proof. (p. 207) It is easy to see that the quadratic variation is approximately equal to t since the summands have expected value:

$$\mathbb{E}((X_{t_{i+1}} - X_{t_i})^2) = t_{i+1} - t_i = \delta t$$

So the sum has expected value:

$$\mathbb{E}\left(\sum (X_{t_{i+1}} - X_{t_i})^2\right) = \sum \delta t = t$$

Date: December 4, 2006.

[1]The Riemann sum converges if and only the function has bounded variation.

The variance of each summand is[2]:

$$Var((X_{t_{i+1}} - X_{t_i})^2) = \mathbb{E}((X_{t_{i+1}} - X_{t_i})^4) - \mathbb{E}((X_{t_{i+1}} - X_{t_i})^2)^2 = 2\delta t^2$$

So, the sum have variance:

$$Var(\sum (X_{t_{i+1}} - X_{t_i})^2) = \sum 2\delta t^2 = 2t\delta t \to 0$$

This means that, in the limit, the sum has zero variance and is therefore not random. The value of this limit is almost sure equal to its expected value which is t. □

This theorem is usually written in the differential form

$$(dX_t)^2 = dt \tag{9.1}$$

For arbitrary increments δt of t this is

$$(\delta X_t)^2 := (X_{t+\delta t} - X_t)^2 = \delta t + o^{eff}(\delta t) \tag{9.2}$$

where I labeled the error term as an *effective little-oh*. Usual: $o(\delta t)/\delta t \to 0$ as $\delta t \to 0$. But effective little-oh means: If you take $N \approx 1/\delta t$ independent copies of $o^{eff}(\delta t)$ you get:

$$\sum_{1/\delta t\ copies} o^{eff}(\delta t) \to 0 \quad \text{as } \delta t \to 0 \tag{9.3}$$

These three equations (9.1), (9.2), (9.3) summarize the statement and proof of Lévy's theorem on quadratic variation of Brownian motion.

9.0.2. *Itō's formula.* Using quadratic variation we can "prove" Itō's formula.

Suppose that we have a particle density function $f(x)$ for $x \in \mathbb{R}$ and X_t is Brownian motion. The probabilistic argument said that we should look for the expected *present value* of f at the future position

[2]This is an easy calculation. The moment generating function for the standard normal distribution is

$$\mathbb{E}(e^{Xt}) = \int e^{xt - x^2/2} dx/\sqrt{2\pi} = e^{t^2/2} \int e^{-(x-t)^2/2} dx/\sqrt{2\pi} = e^{t^2/2}$$

The coefficient of t^{2n} in $\mathbb{E}(e^{Xt})$ is $\mathbb{E}(X^{2n})/(2n)!$ and the coefficient of t^{2n} in $e^{t^2/2}$ is $1/n!2^n$. Therefore, for $X \sim N(0,1)$,

$$\mathbb{E}(X^{2n}) = \frac{(2n)!}{n!2^n} = (2n-1)!! := 1 \cdot 3 \cdot 5 \cdot 7 \cdots (2n-1)$$

You need to multiply by σ^{2n} when $X \sim N(0, \sigma^2)$.

X_t. So, we assume that $f(x)$ is *not time dependent*. It only varies with position x. Do you remember the following formula?

$$f(X_{t+\delta t}) - f(X_t) = f'(X_t)(X_{t+\delta t} - X_t) + \frac{1}{2} f''(X_t)(X_{t+\delta t} - X_t)^2 + o(\delta t)$$

This can be abbreviated:

$$\delta f(X_t) = f'(X_t)\delta X_t + \frac{1}{2} f''(X_t)(\delta X_t)^2 + o(\delta t)$$

Use quadratic variation: $(\delta X_t)^2 = \delta t + o^{eff}(\delta t)$. Then:

$$\delta f(X_t) = f'(X_t)\delta X_t + \frac{1}{2} f''(X_t)\delta t + o^{eff}(\delta t)$$

Now take the sum from 0 to t. (We need to change t above to s so that s can be the variable going from 0 to t: $0 \leq s \leq t$.)

$$f(X_t) - f(X_0) = \sum f'(X_s)\delta X_s + \sum \frac{1}{2} f''(X_s)\delta s + \sum o^{eff}(\delta s)$$

Now, take the limit as $\delta s \to 0$. Then the last term goes to zero by (9.3) and we get *Itō's formula*:

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)\, dX_s + \int_0^t \frac{1}{2} f''(X_s)\, ds \tag{9.4}$$

Here the stochastic integral is

$$\int_0^t f'(X_s)\, dX_s := \lim_{\delta s \to 0} \sum f'(X_s)\delta X_s$$

9.0.3. *discussion.* Why is this not a proof of Itō's formula? The main thing is that we haven't defined the stochastic integral:

$$Z_t = \int_0^t Y_s\, dX_s$$

We only showed that the traditional "limit of Riemann sum" definition makes sense and gives something which satisfies Itō's formula in the special case when $Y_t = f'(X_t)$ is the derivative of a twice differentiable function of standard Brownian motion X_t. In general we need the integral defined for *predictable* stochastic processes Y_s. This means Y_s must be $\mathcal{F}_s$-measurable and *left continuous*. Some people (e.g., our book) take Y_s to be right continuous. However, following my "bible" [4], it makes more intuitive sense to have information (X_t and $\mathcal{F}_t$) be right continuous and processes Y_t based on this information should be predictable.

9.1. **discrete stochastic integrals.** Stochastic integrals are constructed in three steps. First you have discrete time and finite state space (a finite Markov process). Then you have continuous time and finite state space (a continuous Markov chain). Then you take a limit.

The important properties of the construction are visible at each step:

- The construction is linear.
- The result is a martingale Z_t.
- $Z_t^2 - \langle Z \rangle_t$ is also a martingale where $\langle Z \rangle_t$ is the *quadratic variation* of Z_t.

Compare this with what you know about Brownian motion:

(1) X_t is a martingale.
(2) $X_t^2 - t$ is also a martingale.
(3) $\langle X \rangle_t = t$ by Lévy's theorem which we just proved.

9.1.1. *set up.* Take simple random walk on $\mathbb{Z}$. This gives a martingale X_n with $X_0 = 0$ and increments $X_{n+1} - X_n = \pm 1$ with equal probability. Suppose that Y_n is a *predictable process*, i.e., Y_n is $\mathcal{F}_{n-1}$-measurable. The discrete integral is

$$Z_n := \sum_{i=1}^{n} Y_i(X_i - X_{i-1}) = \sum_{i=1}^{n} Y_i \delta X_i$$

(This is supposed to resemble $\int Y\,dX$.)

The idea is that, at time n, you place a bet Y_{n+1} that X_n will increase. The money that you win or lose at that step is

$$Y_{n+1}(X_{n+1} - X_n)$$

Since you cannot see the future, Y_{n+1} is only $\mathcal{F}_n$-measurable.

9.1.2. *linearity.* This construction satisfies the following linearity condition:

$$\sum (aY_i + bV_i)\delta X_i = a \sum Y_i \delta X_i + b \sum V_i \delta X_i$$

In short, Z_n is a linear function of $\{Y_i\}$.

9.1.3. *martingale.*

Theorem 9.2. *Z_n is a martingale and $Z_0 = 0$.*

Proof. This is easy to verify:

$$\mathbb{E}(Z_n \,|\, \mathcal{F}_{n-1}) = Z_{n-1} + \mathbb{E}(Y_n(X_n - X_{n-1}) \,|\, \mathcal{F}_{n-1})$$

Since Y_n is $\mathcal{F}_{n-1}$-meaaurable, the last term vanishes:

$$\mathbb{E}(Y_n(X_n - X_{n-1}) \,|\, \mathcal{F}_{n-1}) = Y_n \mathbb{E}(X_n - X_{n-1} \,|\, \mathcal{F}_{n-1}) = 0$$

So,

$$\mathbb{E}(Z_n \mid \mathcal{F}_{n-1}) = Z_{n-1}$$

□

9.1.4. *quadratic variation.* The *quadratic variation* of Z_n is just the sum of squares of differences:

$$\langle Z \rangle_n := \sum_{i=1}^{n} (Z_i - Z_{i-1})^2 = \sum Y_i^2$$

since these differences are

$$Z_i - Z_{i-1} = Y_i(X_i - X_{i-1}) = \pm Y_i$$

Theorem 9.3. *Suppose that* $\mathbb{E}(Y_i^2) < \infty$ *for each* i. *Then* $(Z_n)^2 - \langle Z \rangle_n$ *is a martingale. In particular,*

$$Var(Z_n) = \mathbb{E}(Z_n^2) = \sum_{i=1}^{n} \mathbb{E}(Y_i^2)$$

Proof. The difference between Z_n^2 and the quadratic variation of Z_n is just the sum of the cross terms:

$$Z_n^2 - \langle Z \rangle_n = 2 \sum_{1 \le i < j \le n} Y_i Y_j (X_i - X_{i-1})(X_j - X_{j-1})$$

$$= Z_{n-1}^2 - \langle Z \rangle_{n-1} + 2 \sum_{i=1}^{n-1} \underbrace{Y_i Y_n (X_i - X_{i-1})}_{\mathcal{F}_{n-1}\text{-measurable}} \underbrace{(X_n - X_{n-1})}_{\mathbb{E}=0}$$

So,

$$\mathbb{E}(Z_n^2 - \langle Z \rangle_n \mid \mathcal{F}_{n-1}) = Z_{n-1}^2 - \langle Z \rangle_{n-1}$$

□

9.2. **Integration wrt Brownian motion.** We take W_t to be standard Brownian motion. This is also called the *Wiener process*, which might explain the use of the letter "W." We want to define the integral

$$Z_t = \int_0^t Y_s \, dW_s$$

where Y_s is a predictable process (left continuous $\mathcal{F}_t$-measurable) which we need to assume is *square summable* in the sense that

$$\int_0^t \mathbb{E}(Y_s^2) \, ds < \infty \tag{9.5}$$

for all t.

9.2.1. *simple processes.* The first step is to take a *step function* Y_t. This is also called a *simple predictable process.* The book calls it a "simple strategy" to emphasize the assumption that Y_t is given by a formula. "Simple" means that Y_t takes only a finite number of values: $0, Y_0, Y_1, \cdots, Y_n$

$$Y_t = \begin{cases} 0 & \text{if } t = 0 \\ Y_0 & \text{if } 0 < t \le t_1 \\ Y_1 & \text{if } t_1 < t \le t_2 \\ & \cdots \\ Y_n & \text{if } t_n < t \end{cases}$$

The stochastic integral is the function

$$Z_t = \int_0^t Y_s \, dW_s := \sum_{i=1}^k Y_{i-1}(W_{t_i} - W_{t_{i-1}}) + Y_k(W_t - W_{t_k}) \tag{9.6}$$

if $t_k < t \le t_{k+1}$.

Remark 9.4. *You can subdivide the intervals $(t_{i-1}, t_i]$ and the integral Z_t remains the same. For example, if you insert $t_{3/2}$ between t_1 and t_2 and put $Y_{3/2} = Y_1$ then*

$$Y_1(W_2 - W_1) = Y_1(W_{3/2} - W_1) + Y_{3/2}(W_2 - W_{3/2})$$

So the sum (9.6) remains the same after subdivision.

I want to go over the basic properties. Maybe I won't prove them.

(1) $Z_t = \int_0^t Y_s \, dW_s$ is linear in Y_t. I.e.,

$$\int_0^t (aX_s + bY_s) \, dW_s = a \int_0^t X_s \, dW_s + b \int_0^t Y_s \, dW_s$$

(2) Z_t is a martingale which is *square summable*, i.e., $\mathbb{E}(Z_t^2) < \infty$.

(3) $Z_t^2 - \langle Z \rangle_t$ is a martingale.

(4)
$$\mathbb{E}(Z_t^2) = \int_0^t \mathbb{E}(Y_s^2)\, ds$$
(So, Z_t is square summable if and only if Y_t is square summable.)

Here the *quadratic variation* $\langle Z \rangle_t$ is given by
$$\langle Z \rangle_t = \int_0^t Y_s^2\, ds$$
So, (3) $\Rightarrow$ (4): If $Z_t^2 - \langle Z \rangle_t$ is a martingale, then
$$\mathbb{E}(Z_t^2 - \langle Z \rangle_t) = Z_0^2 - \langle Z \rangle_0 = 0$$
So,
$$\mathbb{E}(Z_t^2) = \mathbb{E}(\langle Z \rangle_t) = \mathbb{E}\left(\int_0^t Y_s^2\, ds\right) = \int_0^t \mathbb{E}(Y_s^2)\, ds$$

Now, I am going to verify properties (2) and (3) (at least on paper). The key point is that *all cross terms have expectation zero.*

9.2.2. *vanishing expectation of cross terms.*

Theorem 9.5. *Z_t is a martingale for simple processes.*

Proof. The definition of a martingale is that $\mathbb{E}(Z_t \mid \mathcal{F}_s) = Z_s$ if $s < t$. By subdividing the interval, we can always assume that s is a jump time (Remark 9.4). By induction it suffices to show this when $s = t_k$ is the last jump time before t. In other words, we have to show that the last term in Equation (9.6) has expectation 0:
$$\mathbb{E}(Y_k(W_t - W_{t_k}) \mid \mathcal{F}_{t_k}) = Y_k \mathbb{E}(W_t - W_{t_k} \mid \mathcal{F}_{t_k}) = 0$$
The point is that this is a product where the first factor Y_k is determined when the second factor is still random with zero expectation. □

And now, here is a wonderful theorem that will save us a lot of time:

Theorem 9.6 (Meyer)**.** *Suppose that Z_t is a square summable martingale wrt $\mathcal{F}_t$ with $Z_0 = 0$. Then $Z_t^2 - \langle Z \rangle_t$ is also a martingale.*

In other words, (2) $\Rightarrow$ (3)!

Proof. The idea is summarized in the following motto (from [1], 1.5.8). *"When squaring sums of martingale increments and taking the expectation, one can neglect the cross-product terms."*

This theorem is supposed to proved property (3) in all cases simultaneously. So, Z_t could be anything. However, we can always subdivide the interval $[0, t]$ into parts of length δt and get Z_t as a sum of increments:

$$Z_t = Z_t - Z_0 = \sum Z_{t_i} - Z_{t_{i-1}} = \sum \delta_i Z_t$$

The increments $\delta_i Z_t = Z_{t_i} - Z_{t_{i-1}}$ have expectation zero since Z_t is a martingale. When you square Z_t you get:

$$Z_t^2 = \sum (\delta_i Z_t)^2 + 2 \sum_{i<j} \delta_i Z_t \delta_j Z_t$$

The sum of squares converges to the quadratic variation by definition:

$$\langle Z \rangle_t := \lim_{\delta t \to 0} \sum (\delta_i Z_t)^2$$

and the cross terms $\delta_i Z_t \delta_j Z_t$ have expectation zero because the first term is determined when the second term is random with expectation zero.

$$\mathbb{E}(\delta_i Z_t \delta_j Z_t \,|\, \mathcal{F}_{t_{j-1}}) = \delta_i Z_t \mathbb{E}(\delta_j Z_t \,|\, \mathcal{F}_{t_i}) = 0$$

and by the rule of iterated expectation,

$$\mathbb{E}(\delta_i Z_t \delta_j Z_t \,|\, \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(\delta_i Z_t \delta_j Z_t \,|\, \mathcal{F}_{t_{j-1}}) \,|\, \mathcal{F}_s) = 0$$

for any $s \leq t_{j-1}$. □

9.2.3. *general stochastic integral (wrt W_t).* Now suppose that Y_t is any (square summable) predictable process. Then we convert to a simple process by letting $Y_s^{(n)}$ be the average value of Y_t over the inverval $(\frac{k-1}{n}, \frac{k}{n}]$ if s lies in the next interval $(\frac{k}{n}, \frac{k+1}{n}]$. This is to insure that, at time $t = k/n$ when we choose $Y_s^{(n)}$, we only use information from the past and not from the future, i.e., it is predictable.

Since $Y_s^{(n)}$ is a simple predictable process, we can define

$$Z_t^{(n)} := \int_0^t Y_s^{(n)} \, dW_s$$

Without saying it, the book is using the following well-known theorem in real analysis applied to the measure space $(\Omega, \mathcal{F}_t, \mathbb{P})$.

Theorem 9.7. *The space of square summable real valued functions on any measure space is complete in the L^2 metric.*

The L^2 metric is just

$$||Z_t||_2 := \mathbb{E}(Z_t^2)$$

The martingales $Z_t^{(n)}$ form a Cauchy sequence in the L^2 norm. I.e.,

$$\mathbb{E}((Z_t^{(n)} - Z_t^{(m)})^2) = \mathbb{E}(\langle Z^{(n)} - Z^{(m)} \rangle_t) = \int_0^t (Y_s^{(n)} - Y_s^{(m)})^2 \, ds \to 0$$

as $n, m \to \infty$. The book then uses the theorem about completeness of L^2 to concluded that the martingales $Z_t^{(n)}$ converge to some square summable process Z_t

$$Z_t = \int_0^t Y_s \, dW_s := \lim Z_t^{(n)}$$

Since the limit of martingales is a martingale, Z_t is a martingale. By Theorem 9.6 that is all we have to show (linearity being obvious).

9.3. **Itō's formula.** I will repeat the formula and do the examples in the book but I won't go over the proof since we already did it.

Theorem 9.8 (Itō's first formula)**.** *Suppose that W_t is standard Brownian motion and $f(x)$ is C^2, i.e., twice continuously differentiable. Then*

$$f(W_t) - f(W_0) = \int_0^t f'(W_s) \, dW_s + \frac{1}{2} \int_0^t f''(W_s) \, ds$$

I just want to point out that the naïve definition of the stochastic integral that we used in the earlier proof of this formula is equivalent to the rigorous definition that I just explained because $Y_t = f'(W_t)$ is a continuous function of t. Continuity implies that the average value over an interval (used in the rigorous definition) converges to the actual value at one end (used in our naïve definition).

9.3.1. *example 1.* Let $f(t) = t^2$. Then $f'(t) = 2t$ and $f''(t) = 2$. So,

$$f(W_t) - f(W_0) = W_t^2 = \int_0^t 2W_s \, dW_s + \frac{1}{2} \int_0^t 2 \, ds$$

$$\int_0^t W_s \, dW_s = \frac{1}{2} W_t^2 - \frac{1}{2} t$$

9.3.2. *geometric Brownian motion.* Take $f(t) = e^t = f'(t) = f''(t)$. Then Itō's formula is:

$$e^{W_t} - 1 = \int_0^t e^{W_s} \, dW_s + \frac{1}{2} \int_0^t e^{W_s} \, ds$$

If we write $X_t := e^{W_t}$ this becomes:

$$X_t - 1 = \int_0^t X_s \, dW_s + \frac{1}{2} \int_0^t X_s \, ds$$

or

$$dX_t = X_t \, dW_t + \frac{1}{2} X_t \, dt$$

9.4. **Extensions of Itō's formula.** The key ideas are covariation and the product rule.

9.4.1. *covariation.* This is also called the *covariance process.*

Definition 9.9. *The* covariation *of A_t and B_t is defined to be*

$$\langle A, B\rangle_t := \lim_{\delta t\to 0}\sum \delta_i A\,\delta_i B$$

$$= \lim_{\delta t\to 0}\sum (A_{t_i} - A_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})$$

Properties

$$\langle A, A\rangle_t = \langle A\rangle_t \quad \text{(quadratic variation)}$$

$$d\langle A, B\rangle_t = dA_t dB_t \quad \text{(by definition)}$$

$$\langle A + B\rangle + t = \langle A\rangle_t + \langle B\rangle_t + 2\langle A, B\rangle_t$$

Quick proof:

$$\sum(\delta_i A + \delta_i B)^2 = \sum(\delta_i A)^2 + \sum(\delta_i B)^2 + 2\sum \delta_i A\delta_i B$$

9.4.2. *product rule.* The following formula holds without error. (See picture.)

$$\delta AB = A\delta B + B\delta A + \delta A\delta B$$

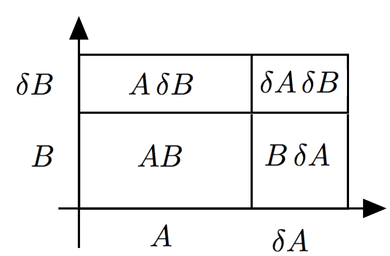

FIGURE 1. The term $\delta A\delta B$ becomes the covariation.

The infinitesmal version is

$$dAB = AdB + BdA + d\langle A, B\rangle_t$$

Example 9.10. *This is example 1 on p.209 which I also did at the end of the last section although I didn't have time to do it in class. Let* $A_t = B_t = W_t$*. Then*

$$d\langle W, W\rangle_t = d\langle W\rangle_t = dt$$

So, the product rule gives:

$$dW_t^2 = 2W_t\, dW_t + dt$$

9.4.3. *quadratic variation of* Z_t*.* is equal to the quadratic variation of the stochastic part of Z_t.

Lemma 9.11 (Lemma 1)**.** *If* f *is continuous with bounded variation (e.g. if* f *is differentiable) then*

$$\langle f\rangle_t = 0$$

(The quadratic variation of f *is zero.)*

Proof.

$$\langle f\rangle_t = \lim_{\substack{\delta t \to 0 \\ \delta f \to 0}} \sum (\delta f)^2 = \lim_{\delta f \to 0} \underbrace{|\delta f|}_{\to 0} \underbrace{\sum |\delta f|}_{\text{variation is bdd}} = 0 \cdot bdd = 0$$

□

Lemma 9.12 (Lemma 2)**.** $\langle f\rangle = 0 \Rightarrow \langle f, X\rangle = 0\ \forall X$.

Proof. If $\langle f, X\rangle_t > 0$ then

$$\langle X - af\rangle_t = \langle X\rangle_t - 2a \underbrace{\langle f, X\rangle_t}_{\text{fixed} > 0} + \underbrace{a^2 \langle f\rangle_t}_{0}$$

If we make a really big then we can make $\langle X - af\rangle_t < 0$. But this is impossible because quadratic variations are sums of squares! □

Here is the theorem we need:

Theorem 9.13. *Suppose that*

$$Z_t = \int_0^t X_s\, ds + \int_0^t Y_s\, dW_s$$

where X_s *integrable (i.e.,* $\int_0^t |X_s|\, ds < \infty$*). The second integral is the "stochastic part" of* Z_t*. Written infinitesmally:*

$$dZ_t = X_t\, dt + Y_t\, dW_t$$

The theorem is:

$$d\langle Z\rangle_t = Y_t^2 dt$$

Proof. Let $f = \int_0^t X_s\,ds$ and $g = \int_0^t Y_s\,dW_s$. Then $Z_t = f + g$. So, using the properties of covariation,

$$\langle Z\rangle_t = \langle f\rangle + \langle g\rangle + 2\langle f, g\rangle$$

But, $\langle f\rangle = 0$ by Lemma 1 (9.11) since f has bounded variation. And $\langle f, g\rangle = 0$ by Lemma 2 (9.12). So, $\langle Z\rangle_t = \langle g\rangle$ and

$$d\langle Z\rangle_t = (Y_t\,dW_t)^2 = Y_t^2\,dt$$

□

9.4.4. *Itō's second formula.* Suppose that $f(x)$ is C^2, i.e., twice continuously differentiable. Then, the Taylor series of f gives is

$$f(x+\delta x) - f(x) = f'(x)\delta x + \frac{1}{2}f''(x)(\delta x)^2 + o((\delta x)^2) \tag{9.7}$$

Now, substitute $x = Z_t$ where Z_t is as in the theorem above.

$$f(Z_{t+\delta t}) - f(Z_t) = f'(Z_t)\,\delta Z_t + \frac{1}{2}f''(Z_t)(\delta Z_t)^2 + o$$

The infinitesmal version is:

$$df(Z_t) = f'(Z_t)\,dZ_t + \frac{1}{2}f''(Z_t)\,d\langle Z\rangle_t$$

Substitute $dZ_t = X_t\,dt + Y_t\,dW_t$ and $d\langle Z\rangle_t = Y_t^2 dt$ and we get:

Theorem 9.14 (Itō II)**.**

$$df(Z_t) = f'(Z_t)X_t\,dt + f'(Z_t)Y_t\,dW_t + \frac{1}{2}f''(Z_t)Y_t^2\,dt$$

9.4.5. *Itō's third formula.* For this we need the vector version of the Taylor series (9.7) and we need to apply it to $x = (t, Z_t)$. Then f' becomes the *gradient* $\nabla f = (f_1, f_2)$ and

$$f'(x)\,dx = (f_1, f_2)\begin{pmatrix} dt \\ dZ_t \end{pmatrix} = \dot{f}(t, Z_t)\,dt + f'(t, Z_t)\,(X_t\,dt + Y_t\,dW_t)$$

and f'' becomes the *Hessian* $D^2 f = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}$. So,

$$\frac{1}{2}f''(x)(dx)^2 = (dt, dZ_t)\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}\begin{pmatrix} dt \\ dZ_t \end{pmatrix}$$

$$= \frac{1}{2}f_{11}\underbrace{d\langle t\rangle}_{=0} + f_{12}\underbrace{d\langle t, Z_t\rangle}_{=0} + \frac{1}{2}f_{22}d\langle Z\rangle_t = \frac{1}{2}f''(t, Z_t)Y_t^2 dt$$

Putting these together we get:

Theorem 9.15 (Itō III)**.**

$$df(t, Z_t) = \dot{f}(t, Z_t)\, dt + f'(t, Z_t)\,(X_t\, dt + Y_t\, dW_t) + \frac{1}{2} f''(t, Z_t) Y_t^2 dt$$

At this point we jumped ahead to the last section 9.8 on the Black-Scholes formula.

9.5. **Continuous martingales.** We skipped this section. But this looks like a good place for me to put the proof of Lévy's theorem which implies that all continuous martingales are reparametrizations of Brownian motion.

Theorem 9.16 (Lévy)**.** *A continuous L^2 martingale M_t starting at 0 is standard Brownian motion if and only if $M_t^2 - t$ is a martingale.*

What follows is from Kunita and Watanabe [2] which is considered to be the most elegant. I also included proofs of the necessary lemmas.

9.5.1. *first step.*

Lemma 9.17. *Assuming the conditions of Theorem 9.16, $\langle M \rangle_t = t$.*

The proof of this lemma uses Meyer's theorem proved in class that

Theorem 9.18 (Meyer)**.** *If M_t is a continuous L^2 martingale then $\langle M \rangle_t$ is the unique continuous increasing process starting at 0 so that*

$$M_t^2 - \langle M \rangle_t$$

is a martingale.

Except that I didn't prove the uniqueness and I didn't define "increasing process."

9.5.2. *uniqueness of increasing process.*

Definition 9.19. *X_t is called an* increasing process *if*

$$t > s \Rightarrow X_t \geq X_s \quad a.s.$$

(1) Clearly, $X_t = t$ is a continuous increasing process.
(2) $\langle M \rangle_t$ is an increasing process starting at 0 since

$$\langle M \rangle_t - \langle M \rangle_s = \lim_{\delta t \to 0} \sum (M_{t_i} - M_{t_{i-1}})^2 \geq 0$$

And $\langle M \rangle_t$ is continuous if M_t is continuous and square summable (by definition of square summable).

Proof of uniqueness part of Meyer's Theorem. Suppose that A_t, B_t are continuous increasing processes starting at 0 and

$$M_t^2 - A_t, \quad M_t^2 - B_t$$

are martingales. Then the difference

$$A_t - B_t$$

is also a continuous martingale starting at 0 with bounded variation (for t bounded). By the following lemma this implies that $A_t = B_t$. So, the continuous increasing process that we need to subtract from M_t^2 to make it into a martingale is unique. □

Lemma 9.20. *Suppose M_t is a continuous martingale starting at 0 and M_t has bounded variation for bounded t. Then $M_t = 0$ for all t.*

Proof. By Lemma 1 (9.11), this implies that the quadratic variation of M_t is identically zero: $\langle M \rangle_t = 0$. Therefore, M_t^2 is also a martingale by the first part of Meyer's theorem that we already proved in (9.6). But $M_t^2 \geq 0$. So, $\mathbb{E}(M_t) = 0$ only if $M_t = 0$ almost surely. □

9.5.3. *Kunita-Watanabe.* One of the main results of [2] was to generalize Itō's formula to the case of L^2 martingales. Or perhaps it would be more fair to say that they formulated the theory of stochastic integrals in such a way that it easily extends to this case.

Theorem 9.21 (Kunita-Watanabe)**.** *If M_t is a continuous L^2 martingale and f is C^2 then*

$$f(M_t) - f(M_0) = \int_0^t f'(M_s)\,dM_s + \frac{1}{2}\int_0^t f''(M_s)\,d\langle M \rangle_s$$

Now we can prove Lévy's theorem. Suppose that $M_t^2 - t$ is a martingale. We know that $M_t^2 - \langle M \rangle_t$ is also a martingale. By Meyer's uniqueness theorem we can conclude that $\langle M \rangle_t = t$.

Now, let $f(x) = e^{ixz}, f'(x) = ize^{ixz}, f''(x) = -z^2 e^{ixz}$ where z is a formal variable. Then $f(M_t) = e^{iM_t z}$ and

$$\mathbb{E}(e^{iM_t z}) = 1 + i\mathbb{E}(M_t)z - \frac{1}{2}\mathbb{E}(M_t^2)z^2 - \frac{i}{3!}\mathbb{E}(M_t^3)z^3 + \frac{1}{4!}\mathbb{E}(M_t^4)z^4 + \cdots$$

gives all of the moments of M_t. So, it suffices to show that these moments are what they should be if M_t were Brownian motion. I.e., it suffices to show that

$$\mathbb{E}(e^{iM_t z}) = e^{-tz^2/2}$$

But the Kunita-Watanabe variation of Itō's formula gives:

$$\mathbb{E}(f(M_t) - f(M_0)) = \frac{1}{2}\int_0^t \mathbb{E}(f''(M_s))\,ds$$

since $d\langle M\rangle_s = ds$ and since anything predictable (like $f(M_s)$) times dM_s has expectation 0. Since $f(M_0) = e^0 = 1$ and $f''(M_s) = -z^2 e^{iM_t z}$ we have

$$\mathbb{E}(e^{iM_t z}) - 1 = \frac{1}{2}\int_0^t -z^2\mathbb{E}(e^{iM_s z})\,ds$$

Let $h(t) = \mathbb{E}(e^{iM_t z})$. Then

$$h(t) - 1 = \frac{-z^2}{2}\int_0^t h(s)\,ds$$

Differentiate both sides:

$$h'(t) = \frac{-z^2}{2}h(t)$$

This is just exponential growth: $h(t) = h(0)e^{-tz^2/2}$. But $h(0) = 1$ since $M_0 = 0$. So,

$$h(t) = \mathbb{E}(e^{iM_t z}) = e^{-tz^2/2}$$

as claimed.

9.6. **Girsanov transformation.** We skipped this section.

9.7. **Feynman-Kac.** The formula of Feynman and Kac gives another way to solve the Black-Scholes equation. First we need to understand how bonds grow if their rates are variable.

9.7.1. *variable bond rate.* Suppose the bond rate changes with time: $r(t)$.Then the value of your bonds will grow by

$$dY_t = r(t)Y_t\, dt \tag{9.8}$$

$$Y_t = Y_0\, exp\left(\int_0^t r(s)\, ds\right) \tag{9.9}$$

Why is (9.9) the solution of (9.8)?

$$Y_{t+dt} = Y_0\, exp\left(\int_0^t r(s)\, ds + \int_t^{t+dt} r(s)\, ds\right)$$

$$= Y_t\, exp(r(t)dt) = Y_t(1 + r(t)dt + \underbrace{r(t)^2 dt^2/2 + \cdots}_{=0})$$

$$dY_t = Y_t r(t)dt$$

If we solve (9.9) for Y_0 we get:

$$Y_0 = Y_t\, exp\left(\int_0^t -r(s)\, ds\right)$$

9.7.2. *the stochastic process.* Now suppose we have a stochastic process Z_t satisfying the stochastic differential equation:

$$dZ_t = a(Z_t)dt + b(Z_t)dW_t$$

Lemma 9.22. *Z_t is a martingale if and only if $a(Z_t)$ is identically zero a.s.*

Proof. Z_t is a martingale iff $\mathbb{E}(dZ_t \mid \mathcal{F}_t) = a(Z_t)dt = 0$ □

Let J_t be given by

$$J_t := exp\left(\int_0^t -r(s, Z_s)\, ds\right)$$

This is how much one dollar at time t was worth at time 0 if the bond rate depends on time and on Z_t.

9.7.3. *the payoff.* If the payoff function is $f(Z_T)$ then the value at time 0 of this payoff is

$$g(Z_T)J_T$$

How much of this value is determined by time t ($0 \le t \le T$)?

$$J_T = exp\left(\int_0^T -r(s,Z_s)\,ds\right)$$

$$= exp\left(\int_0^t -r(s,Z_s)\,ds + \int_t^T -r(s,Z_s)\,ds\right)$$

$$= J_t\, exp\left(\int_t^T -r(s,Z_s)\,ds\right)$$

9.7.4. *the martingale.*

$$M_t := \mathbb{E}(g(Z_T)J_T \,|\, \mathcal{F}_t)$$

$$= J_t \underbrace{\mathbb{E}(g(Z_T)exp\left(\int_t^T -r(s,Z_s)\,ds\right) \,|\, \mathcal{F}_t)}_{V(t,Z_t)}$$

$$= J_t V(t,Z_t)$$

with $M_T = g(Z_T)$. This is a martingale by the law of iterated expectation:

$$\mathbb{E}(M_t \,|\, \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(g(Z_T)J_T \,|\, \mathcal{F}_t) \,|\, \mathcal{F}_s) = \mathbb{E}(g(Z_T)J_T \,|\, \mathcal{F}_s) = M_s$$

Take the differential:

$$dM_t = dJ\,V + J\,dV$$

$$= J_t(-r(t,Z_t)V\,dt + J_t\left(\dot{V}\,dt + V'\,dZ_t + \frac{1}{2}V''\,d\langle M\rangle_t\right)$$

$$= J_t(-r(t,Z_t)V\,dt + J_t\left(\dot{V}\,dt + V'a\,dt + V'b\,dW_t + \frac{1}{2}V''b^2\,dt\right)$$

Since M_t is a martingale, the coefficient of dt must be zero (Lemma 9.22) Therefore, if $x = Z_t$then

$$-r(t,x)V(t,x) + \dot{V}(t,x) + a(x)V'(t,x) + \frac{b^2(x)}{2}V''(t,x) = 0 \tag{9.10}$$

$$V(T,x) = g(x) \tag{9.11}$$

Theorem 9.23. *The solution of the stochastic differential equation (9.10) with boundary condition (9.11) is given by*

$$V(t,x) = \mathbb{E}^x(g(Z_T)exp\left(\int_t^T -r(s,Z_s)\,ds\right))$$

where

$$dZ_t = a(Z_t)\,dt + b(Z_t)\,dW_t$$

9.7.5. *application to Black-Scholes.* If you apply this to the Black-Scholes equation (9.12) you get

$$r(t,x) = r$$
$$a(x) = rx$$
$$b(x) = \sigma x$$

So,

$$V(t,x) = \mathbb{E}^x((Z_T - K)_+ e^{-r(T-t)})$$
$$dZ_t = rZ_t\,dt + \sigma Z_t\,dW_t$$

These equations say that *the fair price of the option at time t is equal to the expected value of the option at time T adjusted for inflation assuming the stock has drift equal to the bond rate.*

9.8. **Black-Scholes.**

9.8.1. *the set-up.*

$$\begin{aligned} S_t &= \text{ value of one share of stock at time } t \\ dS_t &= \mu S_t dt + \sigma S_t\, dW_t \quad (= \text{return}) \\ \mu &= \text{drift (constant in this model)} \\ \sigma &= \text{volatility (also constant)} \\ W_t &= \text{standard Brownian motion} \end{aligned}$$

We are looking at a *European call option*. This is an option to buy one share of stock at price K at time T.

$$\begin{aligned} K &= \text{exercise price} = \text{strike price} \\ T &= \text{expiry date} \\ V(t,x) &:= \text{fair price of the option at time } t \text{ given that } S_t = x \end{aligned}$$

We want to calculate $V(t, S_t)$. We know how much it will be worth at expiry:

$$V(T,x) = (x-K)_+ := \max(x-K, 0)$$

But how much is the option worth today?

9.8.2. *replicating portfolio.* The theory is that there is a portfolio O_t whose value at time T will be exactly $(S_T - K)_+$. Then the value of the option should be the present value of the portfolio. Black and Scholes assumed that there are no *arbitrage* opportunities. This implies that O_t is unique. In [1] it says that there are always sure ways to lose money. So, they don't assume that the value of O_t is unique. Instead it is proved that the fair price of the option is equal to the value of the *cheapest replicating portfolio*. Who is right is a matter of debate.

Fortunately, in this case, the replicating portfolio (also called a *hedging strategy*) is unique.

Our portfolio is just a combination of stocks and bonds but we only have one stock and one bond to choose from in this model. So,

$$O_t = X_t S_t + Y_t$$

$$\begin{aligned} X_t &= \text{number of shares of stock} \\ Y_t &= \text{money in bonds} \\ r &= \text{bond rate (assumed constant)} \end{aligned}$$

The problem is to find X_t and Y_t so that $O_t = V(t, S_t)$ for all $t \leq T$.

Note: O_t needs to be *self-financing*. This means that we have to continuously reinvest all profits. So, O_t grows in two steps:

(a) Stocks and bonds increase in value. ($\Rightarrow$ more $$)
(b) You immediately reinvest the profit. (zero net gain in this step)

Step (a):

$$dO_t = X_t dS_t + rY_t\, dt$$

Step (b): Change X_t and Y_t by dX_t, dY_t so that there is no net change over what you got in step (a). Using the product rule this means:

$$dO_t = X_t dS_t + rY_t\, dt = SdX_t + X_t dS_t + d\langle X, S\rangle_t + dY_t$$

We need to have $V(t, S_t) = O_t$. Using Itō's third formula we get:

$$dV(t, S_t) = \dot{V}(t, S_t)\, dt + V'(t, S_t)\, dS_t + \frac{1}{2}V''(t, S_t)\sigma^2 S_t^2\, dt$$

$$= \dot{V}(t, S_t)\, dt + V'(t, S_t)\mu S_t dt + \underbrace{V'(t, S_t)\sigma S_t dW_t}_{\text{stochastic part}} + \frac{1}{2}V''(t, S_t)\sigma^2 S_t^2\, dt$$

If this is equal to

$$dO_t = X_t dS_t + rY_t\, dt = X_t \mu S_t dt + \underbrace{X_t \sigma S_t dW_t}_{\text{stochastic part}} + rY_t\, dt$$

then the stochastic parts must be equal. So,

$$X_t = V'(t, S_t)$$

Some people say it this way: If you are holding the option $V(t, S_t)$ and you sell off $X_t = V'(t, S_t)$ number of shares of stock then you have hedged away all of your risk (since the stochastic parts will cancel). Therefore, the financial instrument that you are holding: $V(t, S_t) - X_t S_t$ must be increasing in value at the bond rate which is exactly the case.

Now go back to the equation $dV(t, S_t) = dO_t$ and cross out the terms which are equal ($X_t dS_t$ and $V'(t, S_t)dS_t$) and divide by dt. Then we have:

$$rY_t = \dot{V}(t, S_t) + \frac{1}{2}V''(t, S_t)\sigma^2 S_t^2$$

But we also know that

$$Y_t = O_t - X_t S_t = V(t, S_t) - V'(t, S_t)S_t$$

So,

$$\dot{V}(t, S_t) + \frac{1}{2}V''(t, S_t)\sigma^2 S_t^2 - rV(t, S_t) + rV'(t, S_t)S_t = 0$$

which we can rewrite as:

$$\dot{V}(t, x) + \frac{1}{2}\sigma^2 x^2 V''(t, x) + rxV'(t, x) - rV(t, x) = 0 \tag{9.12}$$

This is the *Black-Scholes equation.* This can be solved using Feynman-Kac. But it can also be solved directly using some tricks.

9.8.3. *simplification of Black-Scholes equation.* The first step is to notice that the drift μ does not matter. It is not part of Equation (9.12). Therefore, we may assume that $\mu = 0$ and

$$dS_t = \sigma S_t \, dW_t \tag{9.13}$$

The next step is: We may assume that $r = 0$. This is the same as measuring value in terms of value at time T adjusted for inflation. When $r = 0$ the equation becomes:

$$\dot{V}_0(t,x) + \frac{1}{2}\sigma^2 x^2 V_0''(t,x) = 0 \tag{9.14}$$

When I say "we may assume $r = 0$" I mean that, if you can solve the $r = 0$ equation then you can also solve the original equation. Suppose $V_0(t,x)$ is the solution of this equation with boundary condition $V_0(T,x) = (x-K)_+$. Then the solution of the original equation (9.12) is

$$V(t,x) = e^{-r(T-t)} V_0(t, e^{r(T-t)} x) \tag{9.15}$$

The reason is that V_0 is in terms of time T dollars and x dollars today (at time t) is worth $e^{r(T-t)}x$ dollars at time T and the output V_0 is in terms of time T dollars. So, V_0 of those dollars is worth $e^{-r(T-t)}V_0$ in today's dollars.

If you don't believe it, you can just differentiate the expression in (9.15):

$$\begin{aligned}
\dot{V} &= rV + e^{-r(T-t)}\dot{V}_0 - rxV_0' \\
V &= e^{-r(T-t)} V_0 \\
V' &= V_0' \\
V'' &= e^{r(T-t)} V_0''
\end{aligned}$$

Then

$$\begin{aligned}
&\dot{V} + \frac{1}{2}\sigma^2 x^2 V'' + rxV' - rV \\
&= (rV + e^{-r(T-t)}\dot{V}_0 - rxV_0') + \frac{1}{2}\sigma^2 x^2 e^{r(T-t)} V_0'' + rxV' - rV \\
&= e^{-r(T-t)}\dot{V}_0 + \frac{1}{2}\sigma^2 x^2 e^{r(T-t)} V_0'' = 0
\end{aligned}$$

9.8.4. *solution of Black-Scholes.* Remember that $\mu = 0 = r$. Thus

$$dS_t = \sigma S_t \, dW_t \tag{9.16}$$

$$dV_0(t, S_t) = dO_t = X_t dS_t + rY_t = X_t \sigma S_t \, dW_t$$

So, $S_t, O_t, V_0(t, S_t)$ are all martingales. Therefore,

$$V_0(t, S_t) = \mathbb{E}(V_0(T, S_T) \,|\, \mathcal{F}_t)$$

But T is the payoff date. So, we get

$$V_0(T, x) = g(x) = (x - K)_+ = \max(x - K, 0)$$

where $g(x)$ represents the payoff function. So, the value of the option is just the expected payoff which will depend on the value of the stock at time T. So, we need to calculate S_T.

Equation (9.16) can be solved for S_t as follows.

$$\begin{aligned} d \ln S_t &= \frac{dS_t}{S_t} + \frac{1}{2}\left(\frac{-1}{S_t^2}\right)\sigma^2 S_t^2 dt \\ &= \sigma \, dW_t - \frac{\sigma^2}{2} \, dt \end{aligned}$$

So,

$$\ln S_t = \sigma W_t - \frac{\sigma^2 t}{2} + C$$

Plugging in $t = 0$ gives $C = \ln S_0$. So,

$$S_t = S_0 exp\left(\sigma W_t - \frac{\sigma^2 t}{2}\right)$$

Now attach a little timer on the stock certificate and set that timer to zero at time $T - t$. When the timer reaches t the expiry date has arrived:

$$\begin{aligned} V_0(T - t, x) &= \mathbb{E}(g(S_t) \,|\, S_0 = x) \\ &= \mathbb{E}\left(g(x \, e^{\sigma W_t - \sigma^2 t/2})\right) \end{aligned}$$

Letting $y = x \, e^{-\sigma^2 t/2}$ we get:

$$V_0(T - t, y \, e^{\sigma^2 t/2}) = \mathbb{E}\left((y \, e^{\sigma W_t} - K)_+\right)$$

But $\sigma W_t \sim N(0, \sigma^2 t)$. So we can use the following lemma with $b = \sigma\sqrt{t}, a = y = x \, e^{-\sigma^2 t/2} = xe^{-b^2/2}$.

Lemma 9.24. *Suppose that $X = ae^{bZ}$ where $Z \sim N(0, 1)$. Then the expected value of $g(X) = (X - K)_+$ is*

$$\mathbb{E}((X - K)_+) = ae^{b^2/2}\Phi\left(\frac{\ln(a/K) + b^2}{b}\right) - K\Phi\left(\frac{\ln(a/K)}{b}\right)$$

where Φ is the distribution function for $N(0, 1)$.

Theorem 9.25. *The value of the European call option at time $T-t$ is*

$$V_0(T-t,x) = x\Phi\left(\frac{\ln(x/K)+\sigma^2 t/2}{\sigma\sqrt{t}}\right) - K\Phi\left(\frac{\ln(x/K)-\sigma^2 t/2}{\sigma\sqrt{t}}\right)$$

if $S_t = x$ assuming the bond rate is zero and

$$V(T-t,x) = e^{-rt}V_0(t, e^{rt}x)$$

$$= x\Phi\left(\frac{\ln(x/K)+rt+\sigma^2 t/2}{\sigma\sqrt{t}}\right) - e^{-rt}K\Phi\left(\frac{\ln(x/K)+rt-\sigma^2 t/2}{\sigma\sqrt{t}}\right)$$

if the bond rate is a constant r.

Proof of Lemma 9.24. The expected value is given by an integral which is easy to compute:

$$\mathbb{E}((ae^{bZ}-K)_+) = \int_{-\infty}^{\infty} (ae^{bz}-K)_+ \phi(z)\,dz$$

Since $ae^{bz} = K$ when $z = \frac{1}{b}\ln(K/a)$ we can write this as:

$$\int_{\frac{1}{b}\ln(K/a)}^{\infty} ae^{bz-z^2/2} - Ke^{-z^2/2}\,\frac{dz}{\sqrt{2\pi}}$$

For the first part you can change variables by $z = y+b, dz = dy$ to get

$$\int_{\frac{1}{b}\ln(K/a)}^{\infty} ae^{bz-z^2/2}\,\frac{dz}{\sqrt{2\pi}} = \int_{\frac{1}{b}\ln(K/a)-b}^{\infty} ae^{b^2/2-y^2/2}\frac{dy}{\sqrt{2\pi}}$$

Changing y to $-y$ this gives

$$ae^{b^2/2}\int_{-\infty}^{\frac{1}{b}\ln(a/K)+b} e^{-y^2/2}\frac{dy}{\sqrt{2\pi}} = ae^{b^2/2}\Phi\left(\frac{1}{b}\ln(a/K)+b\right)$$

The second term is easy:

$$\int_{\frac{1}{b}\ln(K/a)}^{\infty} -Ke^{-z^2/2}\,\frac{dz}{\sqrt{2\pi}} = -K\left(1-\Phi\left(\frac{1}{b}\ln(K/a)\right)\right)$$

$$= -K\Phi\left(\frac{1}{b}\ln(a/K)\right)$$

□

I'll stop here. (But in the lecture I have to go back to section 9.7 on the Feynman-Kac formula.)

"The one thing probabilists can do which analysts can't is *stop*..."
-Sid Port, analyst [3]

References

1. Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.
2. H. Kunita and S. Watanabe, *On square integrable martingales*, Nagoya Math. J. **30** (1967), 209–245.
3. L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 1*, second ed., Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Ltd., Chichester, 1994, Foundations.
4. L. C. G. Rogers and David Williams, *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Itô calculus, Reprint of the second (1994) edition.