4 The Itô integral (1 lecture)

Recall the $B_t = K_0 e^{rt}$ defined in (1.14). B_t satisfies the following ordinary differential equation

$$dB_t = rB_t dt, \qquad B_0 = K_0 \tag{4.1}$$

which can also be written into integral form

$$B_t = B_0 + \int_0^t rB_s ds.$$

The last term is a standard integral which is defined by

$$\int_0^t f(x)dx = \lim_{\max_i \Delta s_i \to 0} \sum_{i=1}^N f(s_i^*)\Delta s_i$$

where $0 = s_0 < s_1 < \cdots s_N = t$ form a partition of interval $[0, t]$, $\Delta s_i = s_i - s_{i-1}$, s_i^* is any point in the i-th interval $[s_{i-1}, s_i]$. The value of the integral does not depend on how we choose s_i^*.

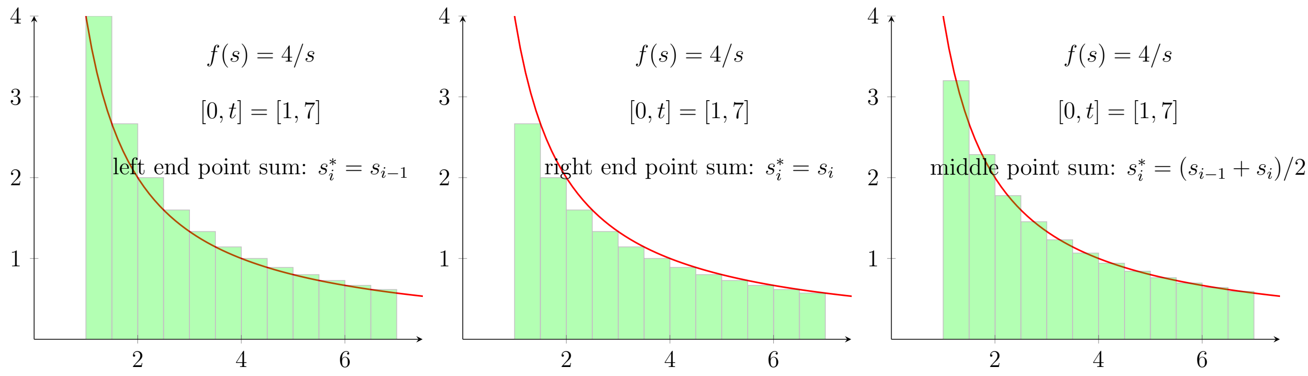

Figure 4.1: $\int_1^7 f(s)ds$ (area under the red line) and $\sum_{i=1}^{12} f(s_i^*)\Delta s_i$ (green area)

We want to introduce Itô integral so that the stochastic process $S_t \in \mathbb{R}^m$ satisfying

$$S_t(\omega) = S_0(\omega) + \int_0^t b(t, S_s)ds + \underbrace{\int_0^t \sigma(s, S_s)dW_s}_{\text{Itô integral}}$$

can be used as a model for prices of m different assets. Here $S_t, b \in \mathbb{R}^m$, $\sigma \in \mathbb{R}^{m\times n}$, and $W_s \in \mathbb{R}^n$. (You won't miss anything important if you think $m = 1 = n$.) We will know (Question 16 of Homework IV) that W_s is not differentiable. Hence we cannot write $\frac{dW_s}{ds}$ and

$$\int_0^t \sigma(s, S_s)dW_s \neq \int_0^t \sigma(s, S_s)\frac{dW_s}{ds}ds.$$

It turns out under some mild condition on f,

$$\int_0^t f(s,\omega)dW_s = \lim_{\max_i \Delta s_i \to 0} \sum_{i=1}^{N} f(s_{i-1},\omega)(W_{s_i}(\omega) - W_{s_{i-1}}(\omega)), \tag{4.2}$$

where $0 = s_0 < s_1 < \cdots < s_N = t$ form a partition of $[0,t]$. We should always choose s_i^* to be the left end point of the interval $[s_{i-1}, s_i]$. See (1.6). The issue is how you pass to the limit and how you justify the existence of such a limit.

Example 4.1 *We generate a single sample path of $\int_0^t W_s dW_s$ and compare it with the theoretical result $\frac{1}{2}W_t^2 - \frac{1}{2}t$ which will be proved later. We have to stress again that in the definition of the Itô integral, we take the left endpoint from the subinterval $[s_{i-1}, s_i]$*

```
T= 3.0; N = 300;
dt = T/N;
t = 0:dt:T;
dW = sqrt(dt)*randn(1,N);
W = zeros(1,N+1);
Integral = zeros(1,N+1);
Exact = zeros(1,N+1);

for i = 1:N
    W(i+1) = W(i) + dW(i);
    Integral(i+1) = Integral(i) + W(i)*dW(i); % Take the left endpoint.
    Exact(i+1)=W(i+1)^2/2 - i*dt/2;
end
plot(t,Integral,'r-',t,Exact,'k-.');
xlabel('t');
hlegend=legend('approx','exact');
```

Definition 4.1 *(Definition 3.1.2 of Oksendal) Let $W_t(\omega)$ be n-dimensional Brownian motion. Then we define $\mathcal{F}_t$ to be the smallest σ-algebra (Definition 2.5) containing all sets of the form*

$$\{\omega : W_{t_1}(\omega) \in \Gamma_1, \cdots, W_{t_k}(\omega) \in \Gamma_k\}, \tag{4.3}$$

where $t_j \le t$ and $\Gamma_j \subset \mathbb{R}^n$ are Borel sets (you can think of them as intervals for $n = 1$ and rectangular boxes for $n \ge 2$), k can be any positive integer. For technical reasons, we also assume that all sets of measure zero are included in $\mathcal{F}_t$.

Note that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \le t$. By Definition 2.6, $\{\mathcal{F}_t : 0 \le t \le T\}$ is a filtration. In addition to $\{\mathcal{F}_t\}$, we also have an $\mathcal{F}$ which is the σ-algebra on Ω. The existence of Ω and $\mathcal{F}$ is guaranteed by Theorem 3.1. Recall (2.15) and Example 2.4 to see what Ω looks like.

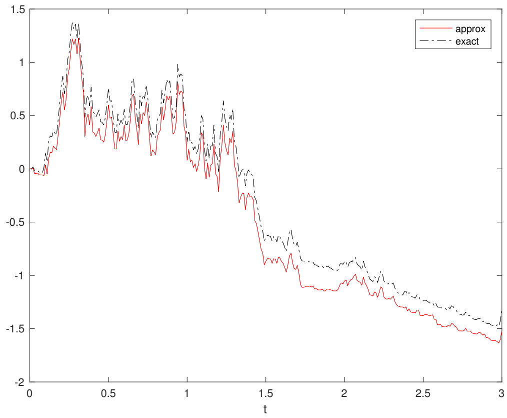

Figure 4.2: Simulation of the Itô integral $\int_0^t W_s dW_s$ and the exact answer $\frac{1}{2}W_t^2 - \frac{1}{2}t$.

Since Ω is larger than any set (representing information about ω) that generates $\mathcal{F}_t$, $\mathcal{F} \supset \mathcal{F}_t$ for all t. You should now recall the $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ that we have introduced in Section 2.9 and see that $\{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3\}$ is just a discrete in time version of $\{\mathcal{F}_t : 0 \leq t \leq T\}$.

Now we follow the presentation of Oksendal. One often thinks of $\mathcal{F}_t$ as the history of W_s up to time t. Recall Definition 2.8 on measurability. It turns out that (which follows from (4.3) in the definition of $\mathcal{F}_t$. But don't worry about its proof.) a function $X(\omega)$ is $\mathcal{F}_t$ measurable if and only if X can be written as the pointwise limit (for almost all ω) of summations of functions of the form

$$g_1(W_{t_1})g_2(W_{t_2})\cdots g_k(W_{t_k})$$

where $g_1, \cdots, g_k$ are bounded continuous functions and $t_j \leq t$ for $j \leq k$, k can be any positive integer. Intuitively, X is $\mathcal{F}_t$ measurable means that to know the value of $X(\omega)$, we do not have to know the full ω, we just need to know $W_s(\omega)$ for $s \leq t$. The value of $X(\omega)$ can be derived from the value of $W_s(\omega)$ for $s \leq t$. For example, the random variable $X(\omega) = (W_{t/2}(\omega))^3 + 7W_t(\omega)$ is $\mathcal{F}_t$-measurable, while $Y(\omega) = \sqrt{\sin(W_{2t}(\omega)) + 5}$ is not $\mathcal{F}_t$-measurable.

Let $\{\mathcal{F}_t\}_{t\geq 0}$ be an increasing family of σ-algebras of subsets of Ω. Recall Definition 2.9 on adapted stochastic process, we see that a process $g(t, \omega) : [0, \infty) \times \Omega \to \mathbb{R}^n$ is called $\mathcal{F}_t$-adapted if for each $t \geq 0$ the function

$$\omega \to g(t, \omega)$$

is $\mathcal{F}_t$-measureable. Sometimes we write $g(t, \omega)$ as $g_t(\omega)$. In the lecture notes, X_t and $X(t)$ are interchangeable.

Example 4.2 *The process* $X(t, \omega) = (W_{t/2}(\omega))^3 + 7W_t(\omega)$ *is* $\mathcal{F}_t$*-adapated while* $Y(t, \omega) = \sqrt{\sin(W_{2t}(\omega)) + 5}$ *is not.*

We have mentioned near the end of Section 2.9 why we need a stochastic process to be $\mathcal{F}_t$-adapted. It is not surprising that Itô integral is based on this kind of stochastic processes (Shreve II Theorem 4.3.1, Oksendal Definition 3.1.4, and Appendix D of Duffie's "Dynamic Asset Pricing Theory"):

Definition 4.2 *Let $\mathcal{V} = \mathcal{V}(0, T)$ be the class of functions*

$$f(t, \omega) : [0, \infty) \times \Omega \to \mathbb{R}$$

(meaning a stochastic process) such that

i) $(t, \omega) \to f(t, \omega)$ *is* $\mathcal{B} \times \mathcal{F}$*-measurable, where* $\mathcal{B}$ *denotes the Borel* σ*-algebra on* $[0, \infty)$.

ii) $f(t, \omega)$ *is* $\mathcal{F}_t$*-adapted.*

iii) $\mathbb{E}[\int_0^T f(t, \omega)^2 dt] < \infty$.

For functions $f \in \mathcal{V}$, we can define its Itô integral

$$\mathcal{I}[f](\omega) = \int_0^T f(t, \omega) dW_t(\omega)$$

where W_t is 1-dimensional Brownian motion.

The idea is natural: First we define $\mathcal{I}[\phi]$ for simple class of functions ϕ (elementary functions). Then we show that each $f \in \mathcal{V}$ can be approximated by such ϕ_k's and we use this to define $\int_0^T f dW_t$ as the limit of $\int_0^T \phi_k dW_t$ for any $\phi_k \to f$.

Definition 4.3 *A function $\phi \in \mathcal{V}(0, T)$ is called elementary if it has the form*

$$\phi(t, \omega) = \sum_{j=0}^{N-1} e_j(\omega) \chi_{[t_j, t_{j+1})}(t) \tag{4.4}$$

where $0 = t_0 < t_1 < \cdots < t_N = T$ form a partition of $[0, T]$.

Remark: Note that since $\phi \in \mathcal{V}$, each function $e_j = \phi(t_j, \cdot)$ must be $\mathcal{F}_{t_j}$-measurable. This means that the value of e_j only depends on $\{W_s, s \le t_j\}$ and is hence independent of $W_{t_{j+1}} - W_{t_j}$ (recall (3.40)).

For elementary function ϕ in (4.4), define

$$\int_0^T \phi(t, \omega) dW_t(\omega) = \sum_{j=0}^{N-1} e_j(\omega)[W_{t_{j+1}} - W_{t_j}](\omega). \tag{4.5}$$

Now, we have the following important observation:

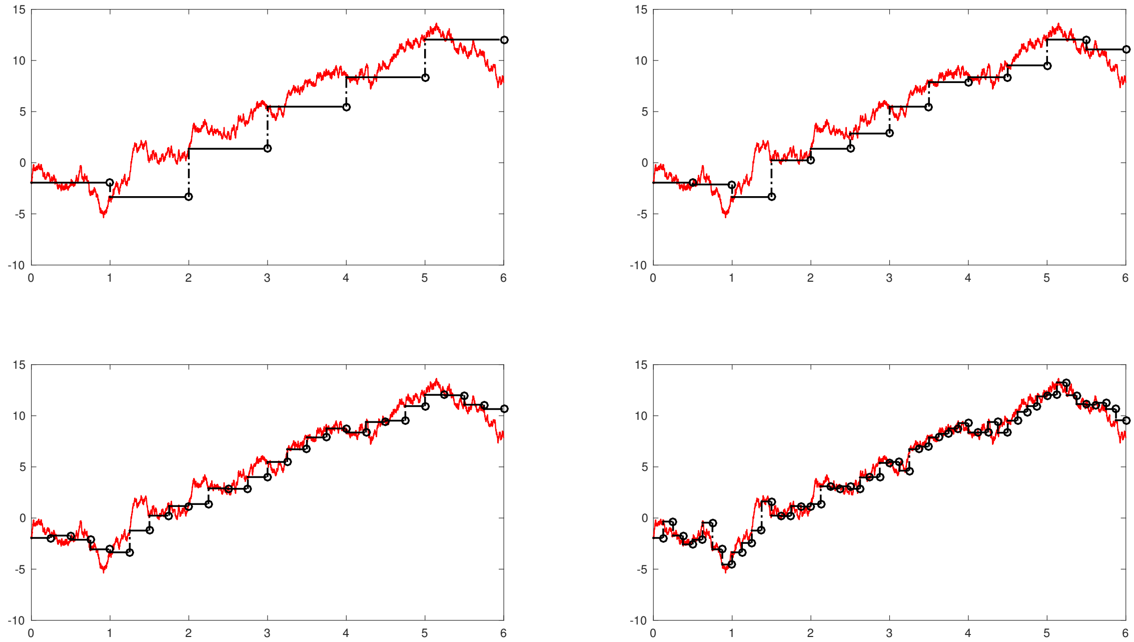

Figure 4.3: A function f (red curve) in $\mathcal{V}$ and its elementary function approximation ϕ_k (black curve) for a fixed ω. x-axis is for the t-variable.

Lemma 4.1 *(The Itô isometry. Theorem 4.2.2 of Shreve II, Lemma 3.1.5 of Oksendal) If $\phi(t,\omega) \in \mathcal{V}$ is bounded and elementary, then*

$$\mathbb{E}\left[\left(\int_0^T \phi(t,\omega)dW_t(\omega)\right)^2\right] = \mathbb{E}\left[\int_0^T \phi(t,\omega)^2 dt\right]. \tag{4.6}$$

Proof: Let $\Delta W_{t_j} = W_{t_{j+1}} - W_{t_j}$. Recall the remark after Defintion 4.3 and note two things: (i) $e_i e_j \Delta W_{t_i}$ and ΔW_{t_j} are independent if $i < j$; (ii) e_i and ΔW_{t_i} are independent. Hence

$$\mathbb{E}[e_i e_j \Delta W_{t_i} \Delta W_{t_j}] = \begin{cases} \mathbb{E}[e_i e_j \Delta W_{t_i}]\mathbb{E}[\Delta W_{t_j}] = 0 & \text{if } i \neq j, \text{ say, } i < j, \\ \mathbb{E}[e_j^2]\mathbb{E}[\Delta W_{t_j}^2] = \mathbb{E}[e_j^2](t_{j+1} - t_j) & \text{if } i = j. \end{cases}$$

Thus

$$\begin{aligned} \mathbb{E}\left[\left(\int_0^T \phi dW_t\right)^2\right] &\overset{(4.5)}{=} \mathbb{E}\left[\left(\sum_i e_i \Delta W_{t_i}\right)\left(\sum_i e_j \Delta W_{t_j}\right)\right] \\ &= \sum_{i,j} \mathbb{E}[e_i e_j \Delta W_{t_i} \Delta W_{t_j}] = \sum_j \mathbb{E}[e_j^2](t_{j+1} - t_j) \\ &= \mathbb{E}[\sum_j e_j^2 (t_{j+1} - t_j)] \overset{(4.4)}{=} \mathbb{E}\left[\int_0^T \phi^2 dt\right]. \qquad \square \end{aligned}$$

Following Oksendal, Page 27, or Duffie, Page 335, we have the following technical results whose proof is skipped for simplicity (won't be tested. But see Figure 4.3 for an illustration) [33]:

Lemma 4.2 *(Oksendal, Page 28 or Duffie, equation (D.1)) Let $f \in \mathcal{V}$. Then there exist elementary functions $\phi_n \in \mathcal{V}$ such that*

$$\mathbb{E}\left[\int_0^T (f-\phi_n)^2 dt\right] \to 0 \qquad \text{as } n \to \infty. \tag{4.7}$$

The following discussion up to Defintion 4.4 are rather technical. Feel free to skip all of them and they won't be tested. By inequality $\mathbb{E}[(X+Y)^2] \le 2\mathbb{E}[X^2] + 2\mathbb{E}[Y^2]$, we know that as long as $\{\phi_n\}$ satisfies $\mathbb{E}\left[\int_0^T (f-\phi_n)^2 dt\right] \to 0$,

$$\mathbb{E}\left[\int_0^T (\phi_m-\phi_n)^2 dt\right] \le 2\mathbb{E}\left[\int_0^T (\phi_m-f)^2 dt\right] + 2\mathbb{E}\left[\int_0^T (f-\phi_n)^2 dt\right] \to 0 \tag{4.8}$$

when $m, n \to \infty$.

By the isometry (4.6), we know the left hand side of (4.8) is

$$\mathbb{E}\left[\left(\int_0^T [\phi_m(t,\omega) - \phi_n(t,\omega)] dW_t(\omega)\right)^2\right], \tag{4.9}$$

which is also $\mathbb{E}[(Z_m - Z_n)^2]$ where $Z_n(\omega) \stackrel{\text{def}}{=} \int_0^T \phi_n(t,\omega) dW_t(\omega)$ is a random variable by (4.5). Hence $\lim_{m,n\to 0} \mathbb{E}[(Z_m - Z_n)^2] = 0$ and therefore $\{Z_n\}$ forms a Cauchy sequence in $L^2(\mathbb{P}) \stackrel{\text{def}}{=} \{X(\omega) : \int_\Omega X^2(\omega) d\mathbb{P}(\omega) \stackrel{(3.9)}{=} \mathbb{E}[X^2] < \infty\}$ where the distance between X and Y is defined to be $\sqrt{\mathbb{E}[(X-Y)^2]} \stackrel{\text{denoted as}}{=} \|X - Y\|$ [34].

One can learn from an advanced probability course that $L^2(\mathbb{P})$ is a complete space which means that as long as $\{Z_n : n = 1, 2, 3\cdots\}$ is a Cauchy sequence in $L^2(\mathbb{P})$ (meaning $\lim_{n,m\to 0} \|Z_n - Z_n\| = 0$), then there exists an $Z \in L(\mathbb{P})$ so that $\lim_{n\to\infty} \|Z_n - Z\| = 0$. Hence $\{Z_n\}$ has a limit in $L^2(\mathbb{P})$. This limit, which is in $L^2(\mathbb{P})$ (meaning its a random variable with finite mean and variance), is defined to be $\int_0^T f(t,\omega) dW_t(\omega)$:

---

[33] If you read Oksendal, you will see that his proof contains two preliminary steps before he can prove (4.7):

a) Let $g \in \mathcal{V}$ be bounded and $g(\cdot, \omega)$ be continuous (in t) for each ω. Then the elementary function $\phi_n = \sum_j g(t_j, \omega)\chi_{[t_j, t_{j+1})}(t)$ is in $\mathcal{V}$ and $\mathbb{E}\left[\int_0^T (g-\phi_n)^2 dt\right] \to 0$ as $n \to \infty$.

b) Let $h \in \mathcal{V}$ be bounded. Then there exist bounded functions $g_n \in \mathcal{V}$ such that $g_n(\cdot, \omega)$ is continuous (in t) for any ω and n, and $\mathbb{E}\left[\int_0^T (h-g_n)^2 dt\right] \to 0$ as $n \to \infty$.

[34] Recall that a sequence of real numbers $\{a_n\}$ is a Cauchy sequence if $\lim_{m,n\to\infty} |a_n - a_m| = 0$. It is a result of advanced calculus that a sequence of real numbers has a limit if and only of it is a Cauchy sequence. This result can be extended to sequences of random variables in $L^2(\mathbb{P})$ which, according to its definition, is the collection of all random variables having finite mean and finite variance since $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

Definition 4.4 *(The Itô integral) Let $f \in \mathcal{V}(0,T)$. Then the Itô integral of f from 0 to T is defined by*

$$\int_0^T f(t,\omega)dW_t(\omega) = \lim_{n\to\infty} \int_0^T \phi_n(t,\omega)dW_t(\omega) \qquad \text{(limit in the mean square sense[35])} \quad (4.10)$$

where $\{\phi_n\}$ is any [36] sequence of elementary functions such that

$$\mathbb{E}\left[\int_0^T (f-\phi_n)^2 dt\right] \to 0 \qquad \text{as } n\to\infty. \qquad (4.11)$$

Remark: It turns out, we can choose f from a larger class and the limit in (4.10) can be a limit in probability sense. See Section 3.3 of Oksendal or Appendix D of Duffie.

Lemma 4.1 immediately leads to

Corollary 4.1 *(The Itô isometry)*

$$\mathbb{E}\left[\left(\int_0^T f(t,\omega)dW_t(\omega)\right)^2\right] = \mathbb{E}\left[\int_0^T f(t,\omega)^2 dt\right] \qquad \text{for all } f\in\mathcal{V}(0,T). \qquad (4.12)$$

Proof: (This proof is not required for the exam. But the conclusion is.) We can simply let $n\to\infty$ in $\mathbb{E}\left[\left(\int_0^T \phi_n(t,\omega)dW_t(\omega)\right)^2\right] = \mathbb{E}\left[\int_0^T \phi_n(t,\omega)^2dt\right]$ and notice that (4.10) implies [37] $\lim_{n\to\infty}\mathbb{E}\left[\left(\int_0^T \phi_n(t,\omega)dW_t(\omega)\right)^2\right] = \mathbb{E}\left[\left(\int_0^T f(t,\omega)dW_t(\omega)\right)^2\right]$ while (4.11) implies $\lim_{n\to\infty}\mathbb{E}\left[\int_0^T \phi_n(t,\omega)^2dt\right] = \mathbb{E}\left[\int_0^T f(t,\omega)^2dt\right]$.

Corollary 4.2 *For any $f,g\in\mathcal{V}(0,T)$,*

$$\mathbb{E}\left[\left(\int_0^T f(t,\omega)dW_t(\omega)\right)\left(\int_0^T g(t,\omega)dW_t(\omega)\right)\right] = \mathbb{E}\left[\int_0^T f(t,\omega)g(t,\omega)dt\right]. \qquad (4.13)$$

Proof: (This proof is not required for the exam. But the conclusion is.) Simply consider the identity $\mathbb{E}\left[\left(\int_0^T (f(t,\omega)+g(t,\omega))dW_t(\omega)\right)^2\right] \overset{(4.12)}{=} \mathbb{E}\left[\int_0^T (f(t,\omega)+g(t,\omega))^2dt\right]$ and then expand the square terms from both sides. □

[35] $X_n\to X$ in the mean square sense if and only if $\lim_{n\to\infty}\mathbb{E}[(X_n-X)^2]=0$.

[36] If you choose another sequence, say, $\tilde\phi_n$, that also satisfies (4.11), then we can form a new sequence $\{\int_0^T\phi_1dW_t, \int_0^T\tilde\phi_1dW_t, \int_0^T\phi_2dW_t, \int_0^T\tilde\phi_2dW_t,\cdots\}$ which is again a Cauchy sequence in $L^2(\mathbb{P})$ because of (4.8) and (4.9). Since subsequences of a Cauchy sequence have the same limit, $\lim_{n\to\infty}\int_0^T\phi_n(t,\omega)dW_t(\omega) = \lim_{n\to\infty}\int_0^T\tilde\phi_n(t,\omega)dW_t(\omega)$.

[37] Let X be in $L^2(\mathbb{P})$. Suppose $\lim_{n\to\infty}\mathbb{E}[(X_n-X)^2]=0$, then $\mathbb{E}[X_n^2]\le 2\mathbb{E}[(X_n-X)^2]+2\mathbb{E}[X^2] < C_1$ for some constant C_1. $\mathbb{E}[(X_n)^2]-\mathbb{E}[X^2] = \mathbb{E}[(X_n-X)(X_n+X)] \le \sqrt{\mathbb{E}[(X_n-X)^2]}\sqrt{\mathbb{E}[(X_n+X)^2]} \le C_2\sqrt{\mathbb{E}[(X_n-X)^2]}\to 0$ for some constant C_2. So, $\lim_{n\to\infty}\mathbb{E}[(X_n)^2]=\mathbb{E}[X^2]$.

Example 4.3 *Assume $W_0 = 0$. Then*

$$\int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}t. \tag{4.14}$$

Proof: *Put $\phi_n(s,\omega) = \sum_{j=0}^{N-1} W_{t_j}(\omega)\chi_{[t_j,t_{j+1})}(s)$ where $t_j = j\delta t$, $\delta t = t/N$. Then*

$$\mathbb{E}\left[\int_0^t (\phi_n - W_s)^2 ds\right] = \mathbb{E}\left[\sum_j \int_{t_j}^{t_{j+1}} (W_{t_j} - W_s)^2 ds\right]$$
$$= \sum_j \int_{t_j}^{t_{j+1}} (s - t_j) ds = \sum_j \frac{1}{2}(t_{j+1} - t_j)^2 = \frac{1}{2}T\delta t \to 0$$

as $\delta t \to 0$. So, by definition,

$$\int_0^t W_s dW_s = \lim_{\delta t \to 0} \int_0^t \phi_n dW_s = \lim_{\delta t \to 0} \sum_j W_{t_j}(W_{t_{j+1}} - W_{t_j})$$
$$= \lim_{\delta t \to 0} \sum_j \frac{1}{2}(W_{t_{j+1}}^2 - W_{t_j}^2) - \frac{1}{2}(W_{t_{j+1}} - W_{t_j})^2 = \frac{1}{2}W_t^2 - \frac{1}{2}t.$$

In the last step, we have used Question 6 of Homework III. One should compare (4.14) with Question 2.11 of Homework II.

Remark: If f and g are differentiable, we have the product rule $(fg)' = f'g + g'f$. We can then derive the integration by parts formula using the fundamental theorem of calculus:

$$f(t)g(t) - f(0)g(0) = \int_0^t (fg)' ds = \int_0^t f'g ds + \int_0^t fg' ds = \int_0^t g df + \int_0^t f dg.$$

Hence

$$\int_0^t f df = \frac{1}{2}\left(f^2(t) - f^2(0)\right).$$

However, because W_t is not differentiable on t, we cannot apply the above formula to conclude

$$\int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}W_0^2 = \frac{1}{2}W_t^2.$$

Indeed, there is an extra $-\frac{1}{2}t$ term on the right hand side.

Example 4.4 *We are ready to verify the Itô isometry $\mathbb{E}\left[\left(\int_0^T f(t,\omega)dW_t(\omega)\right)^2\right] = \mathbb{E}\left[\int_0^T f(t,\omega)^2 dt\right]$ for $f(t,\omega) = W_t(\omega)$:*

$$LHS = \mathbb{E}\left[\left(\int_0^T W_t(\omega)dW_t(\omega)\right)^2\right] \overset{(4.14)}{=} \mathbb{E}\left[\left(\frac{W_T^2}{2}-\frac{T}{2}\right)^2\right]$$
$$= \frac{1}{4}\left(\mathbb{E}[W_T^4]-2T\mathbb{E}[W_T^2]+T^2\right) \overset{Q2\ of\ HW\text{-}III}{=} \frac{3T^2-2T^2+T^2}{4} = \frac{T^2}{2}.$$
$$RHS = \mathbb{E}\left[\int_0^T (W_t(\omega))^2 dt\right] = \int_0^T \mathbb{E}\left[(W_t(\omega))^2\right]dt = \int_0^T t dt = \frac{T^2}{2}.$$

4.1 Properties of the Itô integral

The following properties can be proved easily for elementary functions. By taking limits, we see that they are true for all $f, g \in \mathcal{V}(0,T)$:

Theorem 4.1 *(Theorem 3.2.1 of Oksendal) Let $f, g \in \mathcal{V}(0,T)$ and let $0 < U < T$. Then*

i) $\int_0^T f dW_t = \int_0^U f dW_t + \int_U^T f dW_t$ *for a.a.* ω.

ii) $\int_0^T (c_1 f + c_2 g)dW_t = c_1\int_0^T f dW_t + c_2\int_0^T g dW_t$ *for a.a.* ω.

iii)
$$\mathbb{E}\left[\int_0^T f dW_t\right] = 0. \tag{4.15}$$

iv) $\int_0^T f dW_t$ *is* $\mathcal{F}_T$ *measurable.*

Here is the proof of (4.15) (the proof of (4.15) won't be tested): It is a one-line-proof: (4.15) holds for f being elementary functions. When you pass to the limit, it remains true for $f \in \mathcal{V}(0,T)$.

Here are the details: We know that for elementary function ϕ, we have $\phi(t,\omega) = \sum_{j=0}^{N-1} e_j(\omega)\chi_{[t_j,t_{j+1})}(t)$ with e_j only depends on $\{W_s, s \le t_j\}$. Hence

$$e_j \text{ and } W_{t_{j+1}} - W_{t_j} \text{ are independent.} \tag{4.16}$$

Recall that the Itô integral $\int_0^T \phi_n(t,\omega)dW_t(\omega)$ is defined to be $\sum_{j=0}^{N-1} e_j(\omega)(W_{t_{j+1}} - W_{t_j})$. So

$$\mathbb{E}\int_0^T \phi_n(t,\omega)dW_t(\omega) = \mathbb{E}\sum_{j=0}^{N-1} e_j(\omega)(W_{t_{j+1}} - W_{t_j}) = \sum_{j=0}^{N-1}\mathbb{E}\left(e_j(\omega)(W_{t_{j+1}} - W_{t_j})\right)$$
$$\overset{(4.16)}{=} \sum_{j=0}^{N-1}\mathbb{E}\left(e_j(\omega)\right)\mathbb{E}\left((W_{t_{j+1}} - W_{t_j})\right) = \sum_{j=0}^{N-1}\mathbb{E}\left(e_j(\omega)\right)\times 0 = 0. \tag{4.17}$$

Now, by definition (4.10), when $f \in \mathcal{V}(0, T)$,

$$\int_0^T f(t,\omega) dW_t(\omega) = \lim_{n\to\infty} \int_0^T \phi_n(t,\omega) dW_t(\omega), \tag{4.18}$$

for any elementary function sequence $\{\phi_n\}$ that satisfies (4.11). Taking expectation on both sides of (4.18), and switching the order of $\mathbb{E}$ and $\lim_{n\to\infty}$ [38], we get

$$\mathbb{E}\left(\int_0^T f(t,\omega) dW_t(\omega)\right) = \lim_{n\to\infty}\left(\mathbb{E}\int_0^T \phi_n(t,\omega) dW_t(\omega)\right) \overset{(4.17)}{=} 0. \tag{4.19}$$

This finishes the proof of (4.15).

Example 4.5 *Since a constant function 1 on $[0, T)$ is automatically an elementary function of the form $\phi = 1\chi_{[0,T)}(t)$ where $\chi_{[0,T)}(t) = \begin{cases} 1 & t \in [0,T) \\ 0 & \text{otherwise} \end{cases}$, by definition (4.5),*

$$\int_0^T dW_t = \int_0^T 1\chi_{[0,T)} dW_t = W_T - W_0.$$

To see that this definition makes sense, show by Itô isometry and (4.15) that $\int_0^T dW_t$ and $W_T - W_0$ have the same mean and variance.

Solution: By (4.15), $\mathbb{E}\int_0^T dW_t = 0$. By Itô isometry,

$$\mathrm{Var}\left[\int_0^T dW_t\right] = \mathbb{E}\left[\left(\int_0^T dW_t - 0\right)^2\right] = \int_0^T 1^2 dt = T.$$

At the same time, $\mathbb{E}[W_T - W_0] = 0$, $\mathrm{Var}[W_T - W_0] = T$.

Example 4.6 *Find the probability density of $Z = \exp(\int_0^T t dW_t)$.*
Solution: *Let $X = \int_0^T t dW_t = \lim_{N\to\infty} \sum_{i=0}^{N-1} t_i(W_{t_{i+1}} - W_{t_i})$. Since the sum of jointly normally distributed random variables are still normally distributed [39], and since the limit of normally distributed random variables, if exists, is still normally distributed [40], X has normal distribution. $\mathbb{E}X = 0$, $Var[X] = \mathbb{E}[X^2] = \int_0^T t^2 dt = \frac{T^3}{3}$. $X \sim N(0, \frac{T^3}{3})$. For $a \geq 0$,*

$$\mathbb{P}(Z \leq z) = \mathbb{P}(e^X \leq z) = \mathbb{P}(X \leq \log z) = \int_{-\infty}^{\log z} \frac{1}{\sqrt{2\pi}(T^3/3)^{1/2}} e^{-\frac{x^2}{2T^3/3}} dx.$$

[38] If you worry why we can switch the order, here is the explanation: In the lecture, I have explained to you that the limit in (4.18) is in the mean square sense, which means that if $Z_n(\omega) \overset{\text{def}}{=} \int_0^T \phi_n(t,\omega) dW_t(\omega)$ and $Z \overset{\text{def}}{=} \int_0^T f(t,\omega) dW_t(\omega)$, then $\mathbb{E}[|Z_n - Z|^2] \to 0$. Since $(\mathbb{E}[|X|])^2 \leq \mathbb{E}[X^2]$, we conclude $|\mathbb{E}Z_n - \mathbb{E}Z| \leq \sqrt{\mathbb{E}[|Z_n - Z|^2]} \to 0$. Hence $\lim_{n\to\infty} \mathbb{E}[Z_n] = \mathbb{E}[Z]$ which is the first equation in (4.19).

[39] See the discussion after (3.5)

[40] The proof may need to use characteristic function (a little bit like the moment generating function you have seen in Question 1 of Homework III) and is not required for this course.

We know that the pdf is the derivative of cdf, hence

$$\begin{aligned}\rho_Z(z) =& \frac{d}{dz}\mathbb{P}(Z \le z) = \frac{d}{dz}\left(\int_{-\infty}^{\log z} \frac{1}{\sqrt{2\pi}(T^3/3)^{1/2}} e^{-\frac{x^2}{2T^3/3}} dx\right) \\ =& \frac{1}{\sqrt{2\pi}(T^3/3)^{1/2}} e^{-\frac{(\log z)^2}{2T^3/3}} \frac{d}{dz}\log z = \frac{\sqrt{3}}{z\sqrt{2\pi}T^{\frac{3}{2}}} e^{-\frac{3(\log z)^2}{2T^3}}.\end{aligned}$$

Example 4.7 *Consider $Z = \int_0^T e^{-rt} dW_t$ where T and r are positive constants. Find the pdf of Z.*

Solution: By the same observation as in Example 4.6, we know Z has normal distribution. By (4.15), $\mathbb{E}Z = 0$. By Itô isometry,

$$\mathrm{Var}[Z] = \mathbb{E}[(Z-0)^2] = \mathbb{E}\left[\left(\int_0^T e^{-rt} dW_t\right)^2\right] = \mathbb{E}\int_0^T \left(e^{-rt}\right)^2 dt = \int_0^T e^{-2rt} dt = \left.\frac{e^{-2rt}}{-2r}\right|_{t=0}^{t=T} = \frac{1-e^{-2rT}}{2r}$$

Hence $Z \sim N(0, \frac{1-e^{-2rT}}{2r})$ and its pdf is

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/(2\sigma^2)}$$

with $\sigma = \sqrt{\frac{1-e^{-2rT}}{2r}}$.

Example 4.8 *Consider $Z = \int_0^T e^{\sigma W_t} dW_t$, where σ and T are positive constants and we require $W_0 = 0$ for simplicity. Find the mean and variance of Z. [Hint: In Question 1 of Homework III, you have proved/will prove that if $X \sim N(\mu, \sigma^2)$, $\mathbb{E}[e^X] = e^{\mu + \frac{1}{2}\sigma^2}$.] (If similar problem appears in the exam, similar hint will be given.)*

Solution: By (4.15), $\mathbb{E}Z = 0$. $X = 2\sigma W_t \sim N(0, 4\sigma^2 t)$

$$\begin{aligned}\mathrm{Var}[Z] =& \mathbb{E}[(Z-0)^2] = \mathbb{E}\left[\left(\int_0^T e^{\sigma W_t} dW_t\right)^2\right] = \mathbb{E}\int_0^T \left(e^{\sigma W_t}\right)^2 dt \\ =& \int_0^T \mathbb{E}\left(e^{2\sigma W_t}\right) dt = \int_0^T e^{2\sigma^2 t} dt = \frac{e^{2\sigma^2 T} - 1}{2\sigma^2}.\end{aligned}$$

Remark: Please note that unlike Example 4.5 and Example 4.7,

$$Z = \int_0^T e^{\sigma W_t} dW_t \overset{(4.2)}{=} \lim_{N\to\infty} \sum_{i=1}^N e^{\sigma W_{t_{i-1}}}(W_{t_i} - W_{t_{i-1}})$$

(with $\delta t = T/N$, $t_i = i\delta t$) is not the limit of summation of jointly normally distributed random variables (unless we change $e^{\sigma W_{t_{i-1}}}$ to some deterministic function). Hence Z in

Example 4.8 may not be normally distributed. Knowing its mean and variance may not be enough to determine its pdf.

We state without prove the following technical result. (Don't worry about those technical points in your first study. It just says that $\int_0^t f(s,\omega)dW_s$ is continuous in t, which means that if you chance t slightly, the random variable $\int_0^t f(s,\omega)dW_s$ also changes slightly.)

Theorem 4.2 *(Theorem 3.2.5 of Oksendal) Let $f \in \mathcal{V}(0,T)$. Then there exists a t-continuous version of*

$$\int_0^t f(s,\omega)dW_s(\omega), \qquad 0 \le t \le T,$$

i.e., there exists a t-continuous stochastic process J_t on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\mathbb{P}\left[J_t = \int_0^t f dB\right] = 1 \qquad \forall t \in [0,T].$$

One of the most important property of Itô integral is that it is a (cotinuous-time) Martingale.

So, recall what is a discrete-time martingale. For finite probability space $\Omega = \{$ sequence of coin tosses $\}$, a discrete-time stochastic process M_n is called a martingale if $M_n = \tilde{\mathbb{E}}_n[M_{n+1}]$ (Definition 2.4). See Defintion 2.3 for the definition of conditional expectation and Section 2.9 for $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2\, \mathcal{F}_3$. In our later discussion, we can denote $\tilde{\mathbb{E}}_n[X]$ as $\tilde{\mathbb{E}}[X|\mathcal{F}_n]$ (we have seen it in (3.56)). Please work out Question 5 of Homework IV and read its footnote.

For more general Ω, we have to generalize conditional expectation (in particular, its averaging propeorty) to the following. To gain a better understanding, you should read or work out Questions 8 and 9 of Homework III.

Definition 4.5 *(Defintion 2.3.1 of Shreve II) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$* [41] *and let X be a random variable that is either nonnegative or have finite mean. The conditional expectation of X given $\mathcal{G}$, denoted $\mathbb{E}[X|\mathcal{G}]$, is any random variable that satisfies*

i) (measurability) $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable (see Definition 2.8),

ii) (partial averaging property)

$$\int_A \mathbb{E}[X|\mathcal{G}](\omega)d\mathbb{P}(\omega) = \int_A X(\omega)d\mathbb{P}(\omega) \quad \textit{for all } A \in \mathcal{G}. \tag{4.20}$$

In this more general setting, conditional expectation still have the 5 properties stated in Theorem 2.2.

[41] meaning $\mathcal{G}$ is a σ-algebra and $\mathcal{G} \subset \mathcal{F}$.

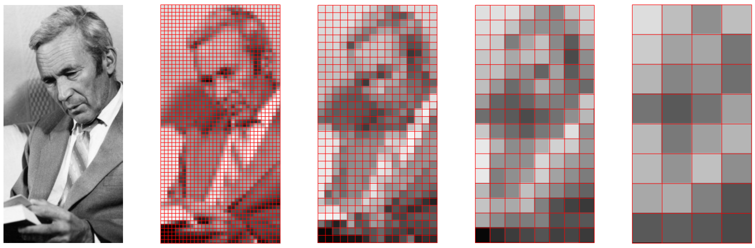

Figure 4.4: $X \leftarrow \mathbb{E}[X|\mathcal{F}_4] \leftarrow \mathbb{E}[X|\mathcal{F}_3] \leftarrow \mathbb{E}[X|\mathcal{F}_2] \leftarrow \mathbb{E}[X|\mathcal{F}_1]$.

Theorem 4.3 *(Theorem 2.3.2 of Shreve II) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Assume X, Y are random variables with finite mean.*

i) (Linearity of conditional expectations) For all constants c_1 and c_2, we have

$$\mathbb{E}[c_1X + c_2Y|\mathcal{G}] = c_1\mathbb{E}[X|\mathcal{G}] + c_2\mathbb{E}[Y|\mathcal{G}]. \tag{4.21}$$

ii) (Taking out what is known) If Z is $\mathcal{G}$-measurable, then

$$\mathbb{E}[ZX|\mathcal{G}] = Z\mathbb{E}[X|\mathcal{G}]. \tag{4.22}$$

iii) (Iterated conditioning) If $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$ (which means $\mathcal{H}$ is coarser and contains less information than $\mathcal{G}$), then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]. \tag{4.23}$$

In particular (see also (3.55)),

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]. \tag{4.24}$$

iv) (Independence) If X is independent of $\mathcal{G}$, then (see also (3.54))

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}X. \tag{4.25}$$

v) (Conditional Jensen's inequality) If $\varphi(x)$ is a convex function of the dummy variable x (e.g. $\varphi(x) = x^2$), then

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}]). \tag{4.26}$$

Definition 4.6 *(Page 312 of Oksendal, Page 74 of Shreve II, Page 332 of Duffie) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{\mathcal{F}_t : t \geq 0\}$ be an increasing family of σ-algebra (meaning $\mathcal{F}_t \subset \mathcal{F}_s$ when $t < s$). A stochastic process $\{M_t(\omega)\}$ is a martingale with respect to $\mathcal{F}_t$ if M_t is $\mathcal{F}_t$-adapted* [42] *, $\mathbb{E}[|M_t|] < \infty$ for all t, and*

$$\mathbb{E}[M_s|\mathcal{F}_t] = M_t \qquad \text{for all } s \geq t. \tag{4.27}$$

[42] Definition 2.9.

Theorem 4.4 *(Corollary 3.2.6 of Oksendal) Let $f(t,\omega) \in \mathcal{V}(0,T)$ for all T. Then*

$$M_t(\omega) = \int_0^t f(s,\omega)dW_s \tag{4.28}$$

is a martingale with respect to $\mathcal{F}_t$[43]. Furthermore,

$$\mathbb{P}\left[\sup_{0\le t\le T} |M_t| \ge \lambda\right] \le \frac{1}{\lambda^2}\mathbb{E}\left[\int_0^T f(s,\omega)^2 ds\right], \qquad \textit{for any } \lambda, T > 0. \tag{4.29}$$

Remark: In $\mathbb{E}[M_s|\mathcal{F}_t]$, we are given the value of W_τ for $\tau \le t$. $f(\tau,\omega)$ is $\mathcal{F}_\tau$ measurable means that the value of $f(\tau,\omega)$ depends on $\{W_u, u \le \tau\}$ only. So, we know the value of $\int_0^t f(\tau,\omega)dW_\tau(\omega)$. Then $\mathbb{E}[M_s|\mathcal{F}_t]$ asks, if we have already known that much of W_τ, what the average value of M_s is, after considering all possible values of the "tail" $\{W_v, v > t\}$. So, the conditional expectation $\mathbb{E}[M_s|\mathcal{F}_t]$ will average out all the "tail" part of the Brownian motion $\{W_v, v > t\}$. This is precisely (2.28).
Remark: You may compare (4.29) with Chebyshev inequality

$$\mathbb{P}\left[|X - \mathbb{E}X| \ge \lambda\right] \le \frac{1}{\lambda^2}\mathbb{E}[(X-\mathbb{E}X)^2], \qquad \text{for any } \lambda > 0 \tag{4.30}$$

which immediately implies

$$\mathbb{P}\left[|M_t| \ge \lambda\right] \le \frac{1}{\lambda^2}\mathbb{E}\left[\int_0^t f(s,\omega)^2 ds\right], \qquad \text{for any } \lambda > 0. \tag{4.31}$$

Hence (4.29) says more than what Chebyshev tells us.

Example 4.9 *(Section 3.5 of Shreve II) Let $W(t)$ be a Brownian motion and let $\mathcal{F}(t)$, $t \ge 0$, be an associated filtration (Definition 4.1). For $\alpha \in \mathbb{R}$, consider the Brownian motion with drift α:*

$$X(t) = \alpha t + W(t). \tag{4.32}$$

Show that for any function $f(y)$, and for any $0 \le s < t$, the function

$$g(x) = \frac{1}{\sqrt{2\pi(t-s)}}\int_{-\infty}^{\infty} f(y) e^{-\frac{(y-\alpha(t-s)-x)^2}{2(t-s)}} dy$$

satisfies

$$\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s)). \tag{4.33}$$

[43] This is because $M_s = M_t + \int_t^s f(\tau,\omega)dW_\tau$ and $\mathbb{E}[\int_t^s f(\tau,\omega)dW_\tau|\mathcal{F}_t]$=0.

Proof: *Let $a = \alpha t + W(s)$. a is $\mathcal{F}(s)$-measurable, meaning that a can be treated as given when evaluating $\mathbb{E}[\cdot|\mathcal{F}(s)]$. By (3.40), $W(t) - W(s) \sim N(0, t-s)$ with pdf $\frac{1}{\sqrt{2\pi(t-s)}}e^{-\frac{x^2}{2(t-s)}}$ and $W(t) - W(s)$ is independent of $\mathcal{F}(s)$.*

$$\begin{aligned}
\mathbb{E}[f(X(t))|\mathcal{F}(s)] =& \mathbb{E}[f(W(t) - W(s) + a)|\mathcal{F}(s)] \\
\overset{(3.10)}{=} & \int_{-\infty}^{\infty} f(x+a)\frac{1}{\sqrt{2\pi(t-s)}}e^{-\frac{x^2}{2(t-s)}}dx\Bigg|_{a=W(s)+\alpha t \text{ is given}} \\
\overset{y=x+a}{=} & \int_{-\infty}^{\infty} f(y)\frac{1}{\sqrt{2\pi(t-s)}}e^{-\frac{(y-a)^2}{2(t-s)}}dy \\
=& g(a - \alpha(t-s)) = g(\alpha s + W(s)).
\end{aligned} \tag{4.34}$$

Remark: (4.33) implies $X(t)$ is a Markov process according to the definition in Question 7 of Homework III, or more generally, Definition 2.3.6 Shreve II. In particular, we know Brownian motion (by setting $\alpha = 0$) is a Markov process.

Take $\alpha = 0$ for simplicity and introduce transition density

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}}e^{-\frac{(y-x)^2}{2\tau}}. \tag{4.35}$$

(We have seen it before in (3.28). So, the integrand in (3.29) is the production of transition densities.) Then

$$g(x) = \int_{-\infty}^{\infty} f(y)p(t-s, x, y)dy \tag{4.36}$$

and (4.33) can be rewritten as (with $t > s$)

$$\mathbb{E}[f(W(t))|\mathcal{F}(s)] = \int_{-\infty}^{\infty} f(y)p(t-s, W(s), y)dy. \tag{4.37}$$

This equation has the following interpretation. Conditioned on the information in $\mathcal{F}(s)$ (which contains all the information obtained by observing the Brownian motion up to and including time s), the conditional density of $W(t)$ is $p(t-s, W(s), y)$. This is a density in the variable y. This density is normal with mean $W(s)$ and variance $t-s$. In particular, the only information from $\mathcal{F}(s)$ that is relevant is the value of $W(s)$. The fact that only $W(s)$ is relevant is the essence of the Markov property. See Question 7 of Homework III for the finite probability space case.

4.2 Homework IV

(Only submit solutions to Questions 5,6,9,10,17.)

1. (The weak law of large number) Suppose $\{Y_i\}$ are i.i.d. (independent identically distributed) random variables with $\mathbb{E}Y_i = a$ and $\mathrm{Var}Y_i = b < \infty$. Show that $\frac{1}{N}\sum_{i=1}^{N} Y_i \to a$ (which is equivalent to $\frac{1}{N}\sum_{i=1}^{N}(Y_i - a) \to 0$) in the mean square sense as $N \to \infty$.

   Proof: $\mathbb{E}[(\frac{1}{N}\sum_{i=1}^{N} Y_i - a)^2] = \frac{1}{N^2}\mathbb{E}[(\sum_{i=1}^{N} Y_i - Na)^2] = \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}[(Y_i - a)^2] = \frac{b}{N} \to 0$ as $N \to \infty$.

2. (Another proof of Question 6 of Homework III) Let $\{W_s : 0 \le s \le T\}$ be a 1-dimensional Brownian motion with $W_{t_0} = W_0 = 0$. Given $t > 0$ and $N \in \mathbb{Z}_+$, let $\delta t = t/N$ and $t_j = j\delta t$. Use the weak law of large number to prove that

$$\lim_{\delta t \to 0} \sum_{j=0}^{N-1} (W_{t_{j+1}} - W_{t_j})^2 = t \tag{4.38}$$

   in the mean square sense. [44]

   Proof: $\mathbb{E}\left(\frac{(W_{t_{j+1}} - W_{t_j})^2}{\delta t}\right) = 1.$

$$\lim_{\delta t \to 0} \sum_{j} (W_{t_{j+1}} - W_{t_j})^2 = t \lim_{N \to \infty} \frac{\sum_{j=0}^{N-1} \frac{(W_{t_{j+1}} - W_{t_j})^2}{\delta t}}{N} = t \times 1 = 1. \tag{4.39}$$

3. Use the same notation as in Question 2. Show that W_t has unbounded first variation, which means

$$\lim_{N \to \infty} \sum_{j=0}^{N-1} |W_{t_{j+1}} - W_{t_j}| = \infty. \tag{4.40}$$

   Proof: Since

$$\sum_{j=1}^{N-1} (W_{t_{j+1}} - W_{t_j})^2 \le \left(\max_{0 \le k \le N-1} |W_{t_{k+1}} - W_{t_k}|\right)\left(\sum_{j=1}^{N-1} |W_{t_{j+1}} - W_{t_j}|\right),$$

   we know

$$\infty > \sum_{j=1}^{N-1} |W_{t_{j+1}} - W_{t_j}| \ge \frac{\sum_{j=1}^{N-1} (W_{t_{j+1}} - W_{t_j})^2}{\max_{0 \le k \le N-1} |W_{t_{k+1}} - W_{t_k}|}. \tag{4.41}$$

[44] FYI: Let Y_i, $i = 1, 2, \cdots$, be a sequence of independent and identically distributed random variables with mean $E[Y_i] = \mu$. The strong law of large number says that $\frac{\sum_{i=1}^{N}(Y_i - \mu)}{N} \to 0$ with probability 1. It means that for almost all ω, $\frac{\sum_{i=1}^{N}(Y_i(\omega) - \mu)}{N} \to 0$. So, if we use the strong law of large number, we can conclude that (4.38) happens with probability 1.

As W_t is continuous, $\lim_{N\to\infty} \max_{0\le k\le N-1} |W_{t_{k+1}} - W_{t_k}| = 0$. Using (4.38) and letting $N \to \infty$ in (4.41), we obtain

$$\lim_{N\to\infty} \sum_{j=0}^{N-1} |W_{t_{j+1}} - W_{t_j}| = \infty.$$

4. Let Y_i, $i = 1, 2, \cdots$, be a sequence of independent and identically distributed random variables with mean $E[Y_i] = \mu$ and $\text{Var}[Y_i] = \sigma^2$. The central limit theorem says that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n (Y_i - \mu)}{\sqrt{n}} \to Z \sim N(0, \sigma^2) \tag{4.42}$$

in the sense of distribution. It means that as n increases, the distribution of the random variable $\frac{\sum_{i=1}^n (Y_i-\mu)}{\sqrt{n}}$ becomes closer and closer to that of $N(0, \sigma^2)$ random variable Z. Now, consider the symmetric random walk (defined in Example 2.11)

$$M_k = \sum_{i=1}^k Z_i$$

where $M_0 = 0$, $\mathbb{P}(Z_i = 1) = \mathbb{P}(Z_i = -1) = \frac{1}{2}$. Fix t, define $W_t^{(n)} = \sqrt{\frac{t}{n}} M_n$. Show that

$$\lim_{n\to\infty} W_t^{(n)} \to Z \sim N(0, t) \tag{4.43}$$

in the sense of distribution.

Solution: $\mathbb{E}[Z_i] = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$. $\text{Var}[Z_i] = \mathbb{E}[Z_0^2] - (\mathbb{E}[Z_i])^2 = 1$. By the central limit theorem

$$\frac{\sum_{i=1}^k (Z_i - \mathbb{E}[Z_i])}{\sqrt{n}} \to Z \sim N(0, 1)$$

in the sense of distribution. Hence

$$W_t^{(n)} = \sqrt{t}\, \frac{\sum_{i=1}^k (Z_i - \mathbb{E}[Z_i])}{\sqrt{n}} \to \sqrt{t} Z \sim N(0, t)$$

in the sense of distribution.

5. Recall (2.28)

$$\tilde{\mathbb{E}}_n[X](\omega_1 \cdots \omega_n) = \sum_{\omega_{n+1}\cdots\omega_N} q_u^{\#H(\omega_{n+1}\cdots\omega_N)} q_d^{\#T(\omega_{n+1}\cdots\omega_N)} X(\omega_1 \cdots \omega_n \omega_{n+1} \cdots \omega_N)$$

and the S_3 defined in Example 2.4. **Show** that

$$\tilde{\mathbb{E}}_2[S_3](HH)\mathbb{P}(A_{HH}) = \sum_{\omega\in A_{HH}} S_3(\omega)\mathbb{P}(\omega), \tag{4.44}$$

where $A_{HH} = \{HHH, HHT\}$ is defined in (2.40). Since $\tilde{\mathbb{E}}_2[S_3](\omega)$ does not change value on A_{HH}, (4.44) can be rewritten as

$$\int_{A_{HH}} \tilde{\mathbb{E}}_2[S_3](\omega) d\mathbb{P}(\omega) = \int_{A_{HH}} S_3(\omega) d\mathbb{P}(\omega). \tag{4.45}$$

Similarly, **prove**

$$\int_{A_{HT}} \tilde{\mathbb{E}}_2[S_3](\omega) d\mathbb{P}(\omega) = \int_{A_{HT}} S_3(\omega) d\mathbb{P}(\omega).\text{[45]} \tag{4.47}$$

6. Let W_t be 1-dimensional Brownian motion with $W_0 = 0$ and let $X_t = e^{W_t^2}$. Show that

$$\mathbb{E}[X_t^2] = \frac{1}{\sqrt{1-4t}}, \qquad t \in [0, 1/4).$$

7. Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$, be a filtration for this Brownian motion. Show that $W^2(t) - t$ is a martingale, i.e., $\mathbb{E}[W^2(t) - t|\mathcal{F}(s)] = W^2(s) - s$ or equivalently $\mathbb{E}[W^2(t) - W^2(s)|\mathcal{F}(s)] = t - s$ for $0 \leq s \leq t$.

   Proof:

$$\begin{aligned}
&\mathbb{E}[W^2(t) - W^2(s)|\mathcal{F}(s)] \\
=&\mathbb{E}[(W(t) - W(s))^2 + 2W(t)W(s) - 2W^2(s)|\mathcal{F}(s)] \\
\overset{(4.22)}{=}&\mathbb{E}[(W(t) - W(s))^2|\mathcal{F}(s)] + 2W(s)\mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] = t - s.
\end{aligned}$$

8. Use the Itô isometry which is Corollary 4.1 and property (iii) of Theorem 4.1 to answer the following questions. Let W_t be 1-dimensional Brownian motion with $W_0 = 0$. Define

$$X = \int_0^T t dW_t, \qquad \text{and} \qquad Y = \int_0^T (T-t) dW_t.$$

   Determine $\mathbb{E}[X]$, $\mathrm{Var}[X]$, $\mathbb{E}[Y]$, $\mathrm{Var}[Y]$. Note that

$$X + Y = \int_0^T T dW_t = T W_T.$$

   Determine $\mathrm{Var}[X+Y]$ and then determine $\mathrm{Cov}(X,Y) \overset{\text{def}}{=} \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$.

   Solution: $\mathbb{E}X = \mathbb{E}Y = 0$. $\mathrm{Var}X = \mathbb{E}[X^2] = \mathbb{E}[\left(\int_0^T t dW_t\right)^2] = \mathbb{E}\int_0^T t^2 dt = \frac{T^3}{3}$.

[45] By the same method, one can prove that

$$\int_A \tilde{\mathbb{E}}_2[S_3](\omega) d\mathbb{P}(\omega) = \int_A S_3(\omega) d\mathbb{P}(\omega) \tag{4.47}$$

for $A \in \{A_{HH}, A_{HT}, A_{TH}, A_{TT}\}$ or more generally for any $A \in \mathcal{F}_2$ defined by (2.41). Hence if one recalls the standard definition of conditional expectation of $\tilde{\mathbb{E}}[S_3|\mathcal{F}_2]$ in (4.20), we have $\tilde{\mathbb{E}}_2[S_3] = \tilde{\mathbb{E}}[S_3|\mathcal{F}_2]$.

$\text{Var}Y = \mathbb{E}[Y^2] = \mathbb{E}[\left(\int_0^T (T-t)dW_t\right)^2] = \mathbb{E}\int_0^T (T-t)^2 dt = \frac{T^3}{3}$.
$\text{Var}[X+Y] = \mathbb{E}[(X+Y)^2] = T^2 T = T^3$.
Since $\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2+Y^2+2XY]$,

$$\mathbb{E}[XY] = \frac{1}{2}\left(\mathbb{E}[(X+Y)^2] - \mathbb{E}[X^2] - \mathbb{E}[Y^2]\right) = \frac{T^3}{6}.$$

$$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{T^3}{6}.$$

9. For the M_t defined in (4.28), prove that

$$\mathbb{E}[M_t] = 0.$$

(Hint: Use (4.24) with $\mathcal{G} = \mathcal{F}_0$.) Then find the variance of

$$M_t = \int_0^t e^{-\alpha s} dW_s$$

for $\alpha > 0$.

10. Let

$$Y_t = \int_0^t \sqrt{|W_s|} dW_s,$$

where $|W_s|$ denotes the absolute value of W_s. Determine Var$[Y]$.

11. Let W_t be one-dimensional Brownian motion, $\sigma \in \mathbb{R}$ be constant and $s \geq t \geq 0$. (1) Use Question 1 of Homework III to prove that

$$\mathbb{E}[e^{\sigma(W_s - W_t)}] = e^{\frac{1}{2}\sigma^2(s-t)}. \tag{4.48}$$

(2) Prove directly from the definition that

$$M_t = e^{\sigma W_t - \frac{1}{2}\sigma^2 t}, \qquad t \geq 0 \tag{4.49}$$

is a martingale with respect to $\mathcal{F}_t$ ((4.27)). Then use this result to prove that $\mathbb{E}[M_t] = 1$ for all $t \geq 0$ if $W_0 = 0$. (Hint: If $s \geq t$, then $\mathbb{E}[M_s|\mathcal{F}_t] =$
$\mathbb{E}[M_t e^{\sigma(W_s-W_t)-\frac{1}{2}\sigma^2(s-t)}|\mathcal{F}_t] \overset{(4.22)}{=} M_t e^{-\frac{1}{2}\sigma^2(s-t)}\mathbb{E}[e^{\sigma(W_s-W_t)}|\mathcal{F}_t]$.)

Proof: (1) $W_s - W_t \sim N(0, s-t)$. $\mathbb{E}[e^{\sigma(W_s-W_t)}] = e^{\frac{1}{2}\sigma^2(s-t)}$.

(2)

$$\begin{aligned}
\mathbb{E}[M_s|\mathcal{F}_t] =& M_t e^{-\frac{1}{2}\sigma^2(s-t)}\mathbb{E}[e^{\sigma(W_s-W_t)}|\mathcal{F}_t] \\
\overset{(4.25),(3.12)}{=}& M_t e^{-\frac{1}{2}\sigma^2(s-t)}\mathbb{E}[e^{\sigma(W_s-W_t)}] \\
=& M_t e^{-\frac{1}{2}\sigma^2(s-t)} e^{\frac{1}{2}\sigma^2(s-t)} = M_t.
\end{aligned}$$

$\mathbb{E}[M_t] \overset{(4.24)}{=} \mathbb{E}[\mathbb{E}[M_t|\mathcal{F}_0]] = \mathbb{E}[M_0] = \mathbb{E}[1] = 1.$

12. (Page 324 of "Dynamic Asset Pricing Theory", 3rd edition, by Duffie. An equivalent definition of conditional expectation. This is for your information only, in case you will read Duffie or other books later in your career. It won't be tested.) For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with Ω finite, if $\mathcal{G}$ is a sub-σ-algebra[46], then $\mathcal{G}$ represents in some sense "less information". The conditional expectation of X given a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$ is defined as any $\mathcal{G}$-measurable random variable denoted by $\mathbb{E}[X|\mathcal{G}]$, satisfying the property that

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]Z] = \int_\Omega \mathbb{E}[X|\mathcal{G}](\omega)Z(\omega)d\mathbb{P}(\omega) = \int_\Omega X(\omega)Z(\omega)d\mathbb{P}(\omega) = \mathbb{E}[XZ] \tag{4.50}$$

for any $\mathcal{G}$-measurable random variable Z. Please compare it with (4.20) and show that (4.50) implies (4.20). Then use (4.22) to show that (4.20) implies (4.50).

Proof: "(4.50)⇒(4.20)": For any set $A \in \mathcal{G}$, define $I_A = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise.} \end{cases}$ Then 1_A is $\mathcal{G}$-measurable since $A \in \mathcal{G}$. So, we can let $Z = 1_A$ in (4.50). This leads to

$$\int_\Omega \mathbb{E}[X|\mathcal{G}](\omega)1_A(\omega)d\mathbb{P}(\omega) = \int_\Omega X(\omega)1_A(\omega)d\mathbb{P}(\omega)$$

or

$$\int_A \mathbb{E}[X|\mathcal{G}](\omega)d\mathbb{P}(\omega) = \int_A X(\omega)d\mathbb{P}(\omega)$$

which is precisely (4.20).

"(4.20)⇒(4.50)": (4.20) implies (4.22) $\mathbb{E}[ZX|\mathcal{G}] = Z\mathbb{E}[X|\mathcal{G}]$. Taking $\mathbb{E}$ on both sides, we obtain

$$\mathbb{E}[\mathbb{E}[ZX|\mathcal{G}]] = \mathbb{E}[Z\mathbb{E}[X|\mathcal{G}]].$$

Since the left hand side is $\mathbb{E}[ZX]$ by (4.24), we get (4.50).

13. (Continue with Question 12. Page 324 of "Dynamic Asset Pricing Theory", 3rd edition, by Duffie. This is for your information only. It won't be tested.) If Y is a nonnegative random variable with $\mathbb{E}Y = 1$, then we can create a new probability measure $\tilde{\mathbb{P}}$ from the old probability measure $\mathbb{P}$ by defining

$$\tilde{\mathbb{P}}(A) = \mathbb{E}(1_A Y) \tag{4.51}$$

for any event A, where $1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise.} \end{cases}$ Note that (4.51) can also be written as $\int_A d\tilde{\mathbb{P}}(\omega) = \int_\Omega Y(\omega)1_A(\omega)d\mathbb{P}(\omega) = \int_A Y(\omega)d\mathbb{P}(\omega)$. So, we write $Y = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}$ and call Y the Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$. (4.51) can also be written as $\tilde{\mathbb{E}}(1_A) = \mathbb{E}(Y1_A)$ for any set $A \in \Omega$, where $\tilde{\mathbb{E}}$ denotes the expectation of under $\tilde{\mathbb{P}}$ and $\mathbb{E}$ denotes the expectation of under $\mathbb{P}$. With some standard mathematics/probability

[46] meaning $\mathcal{G}$ is a σ-algebra and is also a subset of $\mathcal{F}$.

technics which you do not need to know the details, the last equation implies that for any random variable X,

$$\tilde{\mathbb{E}}(X) = \mathbb{E}(YX). \tag{4.52}$$

Definition 4.7 *If $\tilde{\mathbb{P}}(A) > 0$ whenever $\mathbb{P}(A) > 0$, and vice verse, then $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are said to be equivalent measures; they have the same events of probability zero.*

Prove that if $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$ and $\tilde{\mathbb{P}}$ is equivalent to $\mathbb{P}$, then

$$\tilde{\mathbb{E}}(Z|\mathcal{G}) = \frac{1}{\mathbb{E}(\xi|\mathcal{G})}\mathbb{E}(\xi Z|\mathcal{G}), \tag{4.53}$$

where $\xi = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}$.

Proof: By definition (4.50), we need to prove that for any random variable Y being $\mathcal{G}$-measurable,

$$\tilde{\mathbb{E}}[ZY] = \tilde{\mathbb{E}}[\frac{1}{\mathbb{E}(\xi|\mathcal{G})}\mathbb{E}(\xi Z|\mathcal{G})Y], \tag{4.54}$$

which, by the definition of $\tilde{\mathbb{E}}$, is equivalent to

$$\mathbb{E}[\xi ZY] = \mathbb{E}[\xi\frac{1}{\mathbb{E}(\xi|\mathcal{G})}\mathbb{E}(\xi Z|\mathcal{G})Y]. \tag{4.55}$$

But by (4.24) $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]]$, the right hand side of (4.55) is

$$\mathbb{E}\left[\mathbb{E}\left[\xi\frac{1}{\mathbb{E}(\xi|\mathcal{G})}\mathbb{E}(\xi Z|\mathcal{G})Y\middle|\mathcal{G}\right]\right] \overset{(4.22)}{=} \mathbb{E}\left[\mathbb{E}[\xi|\mathcal{G}]\frac{1}{\mathbb{E}(\xi|\mathcal{G})}\mathbb{E}(\xi Z|\mathcal{G})Y\right]$$

since Y and $\frac{1}{\mathbb{E}(\xi|\mathcal{G})}\mathbb{E}(\xi Z|\mathcal{G})$ are already $\mathcal{G}$-measurable. Hence the right hand side of (4.55) becomes $\mathbb{E}[\mathbb{E}(\xi Z|\mathcal{G})Y] \overset{(4.22)}{=} \mathbb{E}[\mathbb{E}[\xi ZY|\mathcal{G}]] \overset{(4.24)}{=} \mathbb{E}[\xi ZY]$. This proves (4.55).

14. (Generalization of Question 9 of Homework III) Let X and Y be a pair of jointly normal random variables with joint density

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right) \\ &= \frac{1}{2\pi\det(\Sigma)^{1/2}}\exp\left(-\frac{1}{2}(x-\mu_1, y-\mu_2)\Sigma^{-1}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}\right) \end{aligned} \tag{4.56}$$

where σ_1, $\sigma_2 > 0$, $|\rho| < 1$, μ_1, μ_2 are real numbers, and $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. Define $W = Y - \frac{\rho\sigma_2}{\sigma_1}X$. Prove W and X are independent and show that

$$\mathbb{E}[Y|X] = \frac{\rho\sigma_2}{\sigma_1}(X-\mu_1) + \mu_2. \tag{4.57}$$

Proof: Since we know Σ is the covariance matrix, we know $\mathrm{Var}(X) = \sigma_1^2$, $\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mu_1)(Y - \mu_2)] = \rho\sigma_1\sigma_2$. Then

$$\mathrm{Cov}(X, W) = \mathrm{Cov}(X, Y) - \frac{\rho\sigma_2}{\sigma_1}\mathrm{Cov}(X, X) = \rho\sigma_1\sigma_2 - \rho\sigma_1\sigma_2 = 0.$$

Since X and W are jointly normal distributed and are un-correlated, X and W are independent.

Because $Y = \frac{\rho\sigma_2}{\sigma_1}X + W$ and X and W are indenpendent,

$$\mathbb{E}[Y|X] = \mathbb{E}[\frac{\rho\sigma_2}{\sigma_1}X + W|X] \overset{(4.21)}{=} \frac{\rho\sigma_2}{\sigma_1}X + \mathbb{E}[W|X] = \frac{\rho\sigma_2}{\sigma_1}(X - \mu_1) + \mu_2.$$

In the last step, we have used $\mathbb{E}[W|X] \overset{(4.25)}{=} \mathbb{E}[W] = \mu_2 - \frac{\rho\sigma_2}{\sigma_1}\mu_1$.

15. Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$, be a filtration for this Brownian motion. Show that $W^2(t) - t$ is a martingale, i.e., $\mathbb{E}[W^2(t) - t|\mathcal{F}(s)] = W^2(s) - s$ or equivalently $\mathbb{E}[W^2(t) - W^2(s)|\mathcal{F}(s)] = t - s$ for $0 \leq s \leq t$. (Hint: Write $W^2(t)$ as $(W(t) - W(s))^2 + 2W(t)W(s) - W^2(s)$.)

    Proof:

$$\begin{aligned}\mathbb{E}[W^2(t) - W^2(s)|\mathcal{F}(s)] =& \mathbb{E}[(W(t) - W(s))^2 + 2W(t)W(s) - 2W^2(s)|\mathcal{F}(s)] \\ =& \mathbb{E}[(W(t) - W(s))^2|\mathcal{F}(s)] + 2W(s)\mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] = t - s.\end{aligned}$$

16. Let $\{W_s : s \geq 0\}$ be a 1-dimensional Brownian motion with $W_0 = 0$. Show that the covariance of $\int_0^s W_u du$ and $\int_0^t W_v dv$ is

$$\mathrm{Cov}\left(\int_0^s W_u du, \int_0^t W_v dv\right) = \frac{1}{3}\min\{s^3, t^3\} + \frac{1}{2}|t - s|\min\{s^2, t^2\}. \tag{4.58}$$

Proof: By definition, when $t \geq s$,

$$
\begin{aligned}
&\mathrm{Cov}\left(\int_0^s W_u du, \int_0^t W_v dv\right)\\
=&\mathbb{E}\left[\left(\int_0^s W_u du\right)\left(\int_0^t W_v dv)\right)\right] - \mathbb{E}\left[\left(\int_0^s W_u du\right)\right]\mathbb{E}\left[\left(\int_0^t W_v dv)\right)\right]\\
=&\mathbb{E}\left[\int_0^s W_u\left(\int_0^t W_v dv\right)du\right] = \mathbb{E}\left[\int_0^s\left(\int_0^t W_u W_v dv\right)du\right]\\
=&\left[\int_0^s\left(\int_0^t \mathbb{E}[W_u W_v]dv\right)du\right] = \left[\int_0^s\left(\int_0^t \min(u,v)dv\right)du\right]\\
=&\left[\int_0^s\left(\int_u^t \min(u,v)dv\right)du\right] + \left[\int_0^s\left(\int_0^u \min(u,v)dv\right)du\right] \qquad \text{use } t \geq s\\
=&\left[\int_0^s\left(\int_u^t u dv\right)du\right] + \left[\int_0^s\left(\int_0^u v dv\right)du\right]\\
=&\left[\int_0^s u(t-u)du\right] + \left[\int_0^s u^2/2 du\right]\\
=&ts^2/2 - s^3/6 = \frac{1}{3}s^3 + \frac{1}{2}(t-s)s^2.
\end{aligned}
$$

When $t \leq s$, we switch s and t in the above computation and get

$$
\mathrm{Cov}\left(\int_0^s W_u du, \int_0^t W_v dv\right) = \mathrm{Cov}\left(\int_0^t W_v dv, \int_0^s W_u du\right) = st^2/2 - t^3/6 = \frac{1}{3}t^3 + \frac{1}{2}(s-t)t^2.
$$

Combining them together, we have proved the desired result.

17. Let $\{W_s : s \geq 0\}$ be a 1-dimensional Brownian motion with $W_0 = 0$. Show that $X_t = W_t^3 - 3tW_t$ is a martingale, i.e., show that for $s \leq t$

$$
\mathbb{E}[W_t^3 - 3tW_t|\mathcal{F}_s] = W_s^3 - 3sW_s.
$$

[Hint: Rewrite $W_t^3 - 3tW_t$ in terms of the increment $W_t - W_s$ and derive $W_t^3 - 3tW_t = (W_t - W_s)^3 + 3(W_t - W_s)^2 W_s + 3(W_t - W_s)W_s^2 + W_s^3 - 3t(W_t - W_s) - 3tW_s$.]