5 The Itô formula (1 lecture)

Recall that if $x(t)$ is a differentiable function of t, we have the chain rule

$$\begin{aligned}
\frac{d}{dt}f(x(t)) &= \lim_{\delta t\to 0}\frac{f(x(t+\delta t))-f(x(t))}{\delta t} \\
&= \lim_{\delta t\to 0}\frac{f(x(t+\delta t))-f(x(t))}{\cancel{x(t+\delta t)-x(t)}}\,\frac{\cancel{x(t+\delta t)-x(t)}}{\delta t} \\
&= \frac{d}{dx}f(x(t))\frac{d}{dt}x(t) \qquad (5.1)
\end{aligned}$$

or $\frac{d}{dt}f(x(t)) = f'(x(t))\frac{d}{dt}x(t)$ or

$$df(x(t)) = f'(x(t))dx(t). \tag{5.2}$$

Here f' denotes the derivative of f, which is another function, and $f'(x(t))$ is the function obtained after replacing the independent variable in f' by $x(t)$. In other word, $f'(x(t))$ contains two operations: first f', then $(x(t))$.

If g is a function of two independent variables,

$$\begin{aligned}
\frac{d}{dt}g(t,x(t)) &= \lim_{\delta t\to 0}\frac{g(t+\delta t, x(t+\delta t))-g(t,x(t))}{\delta t} \\
&= \frac{\partial}{\partial t}g(t,x(t)) + \frac{\partial}{\partial x}g(t,x(t))\frac{d}{dt}x(t). \qquad (5.3)
\end{aligned}$$

which can be rewritten as

$$dg(t,x(t)) = \frac{\partial}{\partial t}g(t,x(t))dt + \frac{\partial}{\partial x}g(t,x(t))dx(t). \tag{5.4}$$

Here $\frac{\partial}{\partial t}g$ denotes the partial derivative of g with respect to its first independent variable, which is a function of two independent variables again. Then $\frac{\partial}{\partial t}g(t,x(t))$ asks you to take the resulting function, and replace the first independent variable by t and the second independent variable by $x(t)$. For example, $g(t,x) = t^2\sin(tx)$, $x(t) = \cos(t)$. $\frac{\partial}{\partial t}g(t,x) = 2t\sin(tx) + t^2x\cos(tx)$. Then $\frac{\partial}{\partial t}g(t,\cos(t)) = 2t\sin(t\cos(t)) + t^2x\cos(t\cos(t))$. Similarly, $\frac{\partial}{\partial x}g(t,\cos(t)) = t^2t\cos(tx)|_{t=t,x=\cos(t)} = t^3\cos(t\cos(t))$.

When studying how to price a derivative security g that depends on the price of its underlying asset X_t, conceptually, we have to deal with $g(t,X_t)$. We want to know how $g(t,X_t)$ changes with respect to t, or how the increment $g(t+\delta t, X_{t+\delta t}) - g(t,X_t)$ behaves. So, we need chain rule again. But the problem is that in general the asset price X_t is not differentiable with respect to t. There comes the Itô formula that we will study next.

Definition 5.1 *(Definition 4.4.3 of Shreve II, Defintion 4.1.1 of Oksendal) Let W_t be 1-dimensional Brownian motion on $(\Omega,\mathcal{F},\mathbb{P})$. A 1-dimensional Itô process or stochastic integral is a stochastic process X_t on $(\Omega,\mathcal{F},\mathbb{P})$ of the form*

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dW_s \tag{5.5}$$

where u, v are $\mathcal{F}_t$-adapted stochastic process and $\int_0^t |u(s,\omega)|ds < \infty$, $\int_0^t |v(s,\omega)|^2 ds < \infty$ for every t and almost all ω.

Note that $X_t - X_0 = \lim_{N\to\infty} \sum_{i=1}^N (X_{s_i} - X_{s_{i-1}})$ if $0 = s_0 < s_1 \cdots < s_N = t$ form a partition of $[0,t]$. Hence $X_t - X_0 = \int_0^t dX_s$ if we define the integral on the right hand side like in Itô integral. So, (5.5) can be written as $\int_0^t dX_s = \int_0^t (u(s,\omega)ds + v(s,\omega)dW_s)$ for every t. So, the integrand must be the same, which leads to (after changing variables from s to t),

$$dX_t = u(t,\omega)dt + v(t,\omega)dW_t. \tag{5.6}$$

One often says that the differential form (5.6) is a short notation of the integral form (5.5).

Recall what we have proved in Example 4.3 (after multiplying by 2): if $W_0 = 0$,

$$W_t^2 = \int_0^t 2W_s dW_s + \int_0^t ds. \tag{5.7}$$

With $X_t = W_t^2$, the above equation can be rewritten in differential form as

$$d\left(W_t^2\right) = 2W_t dW_t + dt. \tag{5.8}$$

But if $u(t)$ is differentiable, then $\int_0^t d(u^2(s)) = \int_0^t 2u(s)du(s)$[47], or equivalently,

$$d\left(u^2(t)\right) = 2u(t)du(t).$$

We get an extra dt term in (5.8).

In general, we have the following

Theorem 5.1 *(The 1-dimensional Ito formula[48]. Theorem 4.1.2 of Oksendal) Let X_t satisfies*

$$dX_t = udt + vdW_t.$$

Let $g(t,x) \in C^2([0,\infty)\times\mathbb{R})$ (i.e. g is twice continuously differentiable on $[0,\infty)\times\mathbb{R}$). Then

$$dg(t,X_t) = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)(dX_t)^2, \tag{5.9}$$

where $(dX_t)^2 = v^2 dt$ since $dt\cdot dt = dt\cdot dW_t = dW_t\cdot dt = 0$ and $dW_t\cdot dW_t = dt$. If g does not depend on t, then (5.9) reduces to

$$dg(X_t) = g'(X_t)dX_t + \frac{1}{2}g''(X_t)(dX_t)^2, \tag{5.10}$$

where g' and g'' are the first and second order derivatives of g.

---

[47] One directly use $\frac{d}{du}u^2 = 2u$, $du^2 = 2udu$, or one can also compute $\int_0^t d(u^2(s)) \overset{x=u^2(s)}{=} \int_{u^2(0)}^{u^2(t)} dx = u^2(t) - u^2(0)$ while $\int_0^t 2u(s)du(s) = u^2(t) - u^2(0)$ by the integration by parts formula $\int_a^b (fdg + gdf) = \int_a^b d(fg) = f(b)g(b) - f(a)g(a)$ for f and g differentiable.

[48] K. Itô , "On Stochastic Differential Equations", Memoirs of the American Mathematical Society, 4 (1951) 1–51.

See (4.38) and (5.15) in the proof for the meaning of $dW_t \cdot dW_t = dt$.

Recall that in Question 6 of homework III you have proved that

$$[W,W]_{\Pi,T} = \sum_{j=0}^{n-1}(W_{t_{j+1}} - W_{t_j})^2 \to \sum_{j=0}^{n-1} \delta t = T$$

in mean square sense when $\delta t \to 0$, where $\delta t = T/n$, $t_j = j\delta t$. Hence $dW_t \cdot dW_t \approx \delta W_t \cdot \delta W_t \stackrel{\text{def}}{=} (W_{t+\delta t} - W_t)^2$ and behaves like δt in the above limit. Then, by $\sum_i |a_i b_i| \le (\sum_i a_i^2)^{1/2}(\sum_i b_i^2)^{1/2}$,

$$\sum_{j=0}^{n-1} \left|(W_{t_{j+1}} - W_{t_j})(t_{j+1} - t_j)|\right| \le \underbrace{\left(\sum_{j=0}^{n-1}(W_{t_{j+1}} - W_{t_j})^2\right)^{1/2}}_{\to T^{1/2}} \underbrace{\left(\sum_{j=0}^{n-1} \delta t^2\right)^{1/2}}_{(\delta t T)^{1/2}} \to 0$$

when $\delta t \to 0$. This indicates $dW \cdot dt = 0$. Finally, $\sum_{j=1}^{n-1} \delta t^2 = 0$ when $\delta t \to 0$ indicates $dt \cdot dt = 0$.

Example 5.1 *Let $X_t = W_t$ and $g(t,x) = g(x) = x^2$. Then $u = 0$, $v = 1$, $g' = 2x$, $g'' = 2$,*

$$d\left(W_t^2\right) = 2W_t dW_t + dt,$$

which means

$$W_t^2 = \int_0^t 2W_s dW_s + t.$$

One should compare the above result with Question 14 of Homework II.

Remark: Suppose $c(t,x)$, say, equals $e^{-t}\sin(t+2x)$, is a deterministic function. Given a stochastic process S_t, Itô formula says

$$dc(t,S_t) = \frac{\partial c}{\partial t}(t,S_t)dt + \frac{\partial c}{\partial x}(t,S_t)dS_t + \frac{1}{2}\frac{\partial^2 c}{\partial x^2}(t,S_t)(dS_t)^2.$$

In the above equation, when $c(t,x) = e^{-t}\sin(t+2x)$, $\frac{\partial c}{\partial x} = 2e^{-t}\cos(t+2x)$ which denotes the partial derivative of c with respect to its second variable. Then one replaces the first variable by t and replace the second variable by S_t to get $\frac{\partial c}{\partial x}(t,S_t) = 2e^{-t}\cos(t+2S_t)$. Similarly for $\frac{\partial c}{\partial t}(t,S_t)$ and $\frac{\partial^2 c}{\partial x^2}(t,S_t)$

Example 5.2 *Let W_t be a 1-dimensional Brownian motion. Use Ito formula to calculate $d(\sin(2t)\sin(W_t))$. You need to simplify your answer to remove the $(dW_t)^2$ term if there is any. Then rewrite the result from differential form to integral form.*
Solution: $f(t,x) = \sin(2t)\sin(x)$, $\frac{\partial f}{\partial t}(t,x) = 2\cos(2t)\sin(x)$, $\frac{\partial f}{\partial x}(t,x) = \sin(2t)\cos(x)$, $\frac{\partial^2 f}{\partial x^2}(t,x) = -\sin(2t)\sin(x)$.

In standard calculus, we keep the dt, dx terms and have

$$df(t,x) = \frac{\partial f}{\partial t}(t,x)dt + \frac{\partial f}{\partial x}(t,x)dx.$$

In stochastic calculus, we keep the dt, dx, and $(dx)^2$ terms and have

$$df(t,x) = \frac{\partial f}{\partial t}(t,x)dt + \frac{\partial f}{\partial x}(t,x)dx + \frac{1}{2}\frac{\partial f^2}{\partial x^2}(t,x)(dx)^2.$$

Then we replace (t,x) by (t,W_t) and replace dx by dW_t to get

$$\begin{aligned}
& d\left(\sin(2t)\sin(W_t)\right) \\
=& 2\cos(2t)\sin(W_t)dt + \sin(2t)\cos(W_t)dW_t - \frac{1}{2}\sin(2t)\sin(W_t)(dW_t)^2 \\
=& \left(2\cos(2t) - \frac{1}{2}\sin(2t)\right)\sin(W_t)dt + \sin(2t)\cos(W_t)dW_t.
\end{aligned}$$

To rewrite in integral form, we integrate from 0 to T, and get

$$\begin{aligned}
& \sin(2T)\sin(W_T) \\
=& \sin(0)\sin(W_0) + \int_0^T \left(2\cos(2t) - \frac{1}{2}\sin(2t)\right)\sin(W_t)dt + \int_0^T \sin(2t)\cos(W_t)dW_t \\
=& \int_0^T \left(2\cos(2t) - \frac{1}{2}\sin(2t)\right)\sin(W_t)dt + \int_0^T \sin(2t)\cos(W_t)dW_t.
\end{aligned}$$

Equivalently, we can write

$$\sin(2t)\sin(W_t) = \int_0^t \left(2\cos(2s) - \frac{1}{2}\sin(2s)\right)\sin(W_s)ds + \int_0^t \sin(2s)\cos(W_s)dW_s.$$

Example 5.3 *Let $g(t,x) = e^t x^4$. $X_t = W_t$.*

$$d\left(e^t(W_t)^4\right) = e^t(W_t)^4 dt + 4e^t(W_t)^3 dW_t + \frac{1}{2}12e^t(W_t)^2 dt.$$

while

$$d\left(e^t(\sin t)^4\right) = e^t(\sin t)^4 dt + 4e^t(\sin t)^3\ d\sin t.$$

Example 5.4 *Let $X_t = W_t$ and $g(t,x) = f(t)x$. Then*

$$d\left(f(t)W_t\right) = W_t f'(t)dt + f(t)dW_t$$

which means (recall that (5.6) is simply another way to write (5.5))

$$f(t)W_t = f(0)W_0 + \int_0^t W_s f'(s)ds + \int_0^t f(s)dW_s.$$

We have hence proved the following integration by parts formula (which still takes the same form as the standard integration by parts formula)

$$\int_0^t f(s)dW_s = f(t)W_t - f(0)W_0 - \int_0^t W_s df(s). \tag{5.11}$$

From now on, unless explicitly stated, we always assume $W_0 = 0$.

Example 5.5 *(Another proof of Question 3 of Homework V.) If $S(t) = e^{rt+\sigma W_t}$, we can use (5.9) with $X_t = rt + \sigma W_t$ ($dX_t = rdt + \sigma dW_t$) to get*

$$\begin{aligned} dS_t =& e^{X_t}dX_t + \frac{1}{2}e^{X_t}(dX_t)^2 \\ =& e^{X_t}(rdt + \sigma dW_t) + \frac{1}{2}e^{X_t}\sigma^2 dt \\ =& \left(r + \frac{1}{2}\sigma^2\right) S_t dt + \sigma S_t dW_t. \end{aligned}$$

In particular, if $r = -\frac{1}{2}\sigma^2$, $dS_t = \sigma S_t dW_t$, or

$$S_t = S_0 + \int_0^t \sigma S_\tau dW_\tau.$$

Since $\int_0^t \sigma S_\tau dW_\tau$ is a martingale, so is $S_t - S_0$. (or simply S_t since $S_0 = 1$ is a constant, assuming $W_0 = 0$.) Question 11 of Homework IV asks you to prove the same conclusion by direct calculation.

Example 5.6 *(Exercise 4.11 of Oksendal) Use Itô formula to prove that*

$$Y_t = e^{\frac{1}{2}t}\cos(W_t)$$

is a $\mathcal{F}_t$-martingale, i.e., $\mathbb{E}[Y_t|\mathcal{F}_\tau] = Y_\tau$ if $\tau \le t$.
Proof*: We take $f(t,x) = e^{\frac{1}{2}t}\cos(x)$. Then*

$$df(t,W_t) = \frac{1}{2}f(t,W_t)dt - e^{\frac{1}{2}t}\sin(W_t)dW_t - \frac{1}{2}e^{\frac{1}{2}t}\cos(W_t)dt = -e^{\frac{1}{2}t}\sin(W_t)dW_t.$$

Hence $e^{\frac{1}{2}t}\cos(W_t) = e^{\frac{1}{2}\times 0}\cos(W_0) + \int_0^t$ something dW_s which is a constant plus a martingale. So $e^{\frac{1}{2}t}\cos(W_t)$ is a martingale.

Intuitive explanation of why (5.9) is true: Consider the increment $\delta W_t \stackrel{\text{def}}{=} W_{t+\delta t} - W_t = \sqrt{\delta t}\theta$ with $\theta \sim N(0,1)$. Hence roughly speaking, δW_t is a term of size $\sqrt{\delta t}$. Note that when δt is small,

$$\delta t^2 < \delta t^{3/2} < \delta t < \sqrt{\delta t} < 1.$$

Since $X_t = X_0 + \int_0^t u ds + \int_0^t v dW_s$,

$$\delta X_t \stackrel{\text{def}}{=} X_{t+\delta t} - X_t \approx u\delta t + v\delta W_t,$$
$$(\delta X_t)^2 = v^2(\delta W_t)^2 + \text{h.o.t of } \delta t.$$

Here h.o.t of δt denotes "higher order terms" of δt which include terms containing $\delta t \delta W_t$, $(\delta W_t)^3$, δt^2, $\cdots$.

$(\delta X_t)^2 = v^2(\delta W_t)^2 +$ h.o.t. $= v^2\theta^2\delta t +$ h.o.t.. The variance of $\text{Var}[(\delta X_t)^2] \approx \text{Var}[v^2\theta^2\delta t] = C\delta t^2$ which is a small term comparing with the variance of δX_t since $\text{Var}[\delta X_t] = \mathbb{E}[(\delta X_t - \mathbb{E}[\delta X_t])^2] \approx \mathbb{E}[(v\delta W_t)^2] \approx \tilde{C}\delta t$. When δX_t is present, the variance of $(\delta X_t)^2$ is therefore too small for $(\delta X_t)^2$ to have a stochastic component. Hence we can ignore the randomness of $(\delta X_t)^2$ and replace it with its mean, which is $v^2\delta t$ (recall $\mathbb{E}\theta^2 = 1$).

$$\begin{aligned}
\delta g \stackrel{\text{def}}{=}& g(t+\delta t, X_{t+\delta t}) - g(t, X_t) \\
\stackrel{\text{Taylor expansion}}{=}& \frac{\partial g}{\partial t}\delta t + \frac{\partial g}{\partial x}(\delta X_t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(\delta X_t)^2 + \text{h.o.t. of } \delta t
\end{aligned}$$

In the last step, we have used Taylor expansion for multivariable functions

$$\begin{aligned}
f(t+\delta t, x+\delta x) = f(t,x) &+ \begin{pmatrix} \frac{\partial}{\partial t} f(t,x) \\ \frac{\partial}{\partial x} f(t,x) \end{pmatrix} \cdot \begin{pmatrix} \delta t \\ \delta x \end{pmatrix} \\
&+ \frac{1}{2}(\delta t, \delta x) \begin{pmatrix} \frac{\partial^2}{\partial t^2} f(t,x) & \frac{\partial^2}{\partial t \partial x} f(t,x) \\ \frac{\partial^2}{\partial t \partial x} f(t,x) & \frac{\partial^2}{\partial x^2} f(t,x) \end{pmatrix} \begin{pmatrix} \delta t \\ \delta x \end{pmatrix} \\
&+ \text{h.o.t of } (\delta t)^2 \text{ and } (\delta x)^2.
\end{aligned}$$

If you do not like the mutivariate Taylor expansion, you can use the following trick

$$\begin{aligned}
\delta g =& g(t+\delta t, X_{t+\delta t}) - g(t, X_{t+\delta t}) + g(t, X_{t+\delta t}) - g(t, X_t) \\
=& \frac{\partial g}{\partial t}\delta t + \frac{\partial g}{\partial x}(\delta X_t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(\delta X_t)^2 + \text{h.o.t. of } \delta t.
\end{aligned}$$

Either way, we get

$$\begin{aligned}
\delta g =& \frac{\partial g}{\partial t}\delta t + \frac{\partial g}{\partial x}(u\delta t + v\delta W_t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}v^2\delta t + \text{h.o.t. of } \delta t \\
=& \sqrt{\delta t} \text{ term} \quad + \delta t \text{ term} \quad + \text{h.o.t. of } \delta t.
\end{aligned}$$

Proof of Theorem 5.1: We need to prove that

$$\begin{aligned}
dg(t, X_t) =& \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t)(dX_t)^2 \\
=& \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)(u_t dt + v_t dW_t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t)v_t^2 dt,
\end{aligned}$$

which means that

$$g(t, X_t) - g(0, X_0) = \int_0^t \left(\frac{\partial g}{\partial t}(s, X_s) + u_s \frac{\partial g}{\partial x}(s, X_s) + \frac{1}{2} v_s^2 \frac{\partial^2 g}{\partial x^2}(s, X_s) \right) ds$$
$$+ \int_0^t v_s \frac{\partial g}{\partial x}(s, X_s) dW_s. \tag{5.12}$$

To explain the idea, let us look at the special case when $X_t = W_t$ and $g(t,x)$ does not depend on t, i.e., $g(t,x) = g(x)$, which means $\frac{\partial g}{\partial t} = 0$, $\frac{\partial}{\partial x} g = g'$, $\frac{\partial^2}{\partial x^2} g = g''$.

$$g(W_t) - g(W_0) = \int_0^t \frac{1}{2} g''(W_s) ds + \int_0^t g'(W_s) dW_s. \tag{5.13}$$

Let $0 = t_0 < t_1 \cdots < t_n = t$ be a partition of $[0,t)$. Introduce the notation $\Delta g(W_{t_j}) = g(W_{t_{j+1}}) - g(W_{t_j})$, $\delta t_j = t_{j+1} - t_j$, $\Delta W_{t_j} = W_{t_{j+1}} - W_{t_j}$. Using Taylor expansion $g(a+b) - g(a) = g'(a)b + \frac{1}{2} g''(a) b^2 + \text{h.o.t. of } b^2$

$$g(W_t) = g(W_0) + \sum_{j=0}^{n-1} \Delta g(W_{t_j}) \approx g(W_0) + \sum_j \frac{\partial g}{\partial x}(W_{t_j}) \Delta W_{t_j} + \frac{1}{2} \sum_j \frac{\partial^2 g}{\partial x^2}(W_{t_j})(\Delta W_{t_j})^2.$$

If $\delta t_j \to 0$, then

$$\sum_j \frac{\partial g}{\partial x}(W_{t_j}) \Delta W_{t_j} \to \int_0^t \frac{\partial g}{\partial x}(W_s) dW_s.$$

We are left to show

$$\frac{1}{2} \sum_j \frac{\partial^2 g}{\partial x^2}(W_{t_j})(\Delta W_{t_j})^2 \to \frac{1}{2} \int_0^t \frac{\partial^2 g}{\partial x^2}(W_s) ds. \tag{5.14}$$

To prove (5.14), let $a(t) = \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(W_t)$, $a_j = a(t_j)$. We want to show

$$\begin{aligned} &\mathbb{E}\left[\left(\sum_j a_j (\Delta W_{t_j})^2 - \sum_j a_j \delta t_j \right)^2 \right] \\ =& \sum_{i,j} \mathbb{E}\left[a_i a_j ((\Delta W_{t_i})^2 - \delta t_i)((\Delta W_{t_j})^2 - \delta t_j) \right] \\ \to& 0. \end{aligned}$$

If $i < j$, $a_i a_j ((\Delta W_{t_i})^2 - \delta t_i)$ and $((\Delta W_{t_j})^2 - \delta t_j)$ are independent so

$$\begin{aligned} &\mathbb{E}\left[a_i a_j ((\Delta W_{t_i})^2 - \delta t_i)((\Delta W_{t_j})^2 - \delta t_j) \right] \\ =& \mathbb{E}\left[a_i a_j ((\Delta W_{t_i})^2 - \delta t_i) \right] \mathbb{E}\left[(\Delta W_{t_j})^2 - \delta t_j \right] \\ =& \mathbb{E}\left[a_i a_j ((\Delta W_{t_i})^2 - \delta t_i) \right] \times 0 = 0. \end{aligned}$$

The same is true when $i > j$. When $i = j$,

$$\sum_j \mathbb{E}\left[\underbrace{a_j^2}_{\text{depends on } \mathcal{F}_{t_j}} \underbrace{((\Delta W_{t_j})^2 - \delta t_j)^2)}_{\text{independent of } \mathcal{F}_{t_j}}\right] = \sum_j \mathbb{E}a_j^2\, \mathbb{E}((\Delta W_{t_j})^2 - \delta t_j)^2)$$
$$= 2\sum_j \mathbb{E}a_j^2 \delta t_j^2 \to 0.$$

In the second to last step, we have used $\mathbb{E}[(W_{t_{j+1}} - W_{t_j})^2] = \delta t_j$ and $\mathbb{E}[(W_{t_{j+1}} - W_{t_j})^4] = 3\delta t_j^2$ (Question 2 of Homework III). Hence we have proved

$$\mathbb{E}\left[\left(\sum_j a_j(\Delta W_{t_j})^2 - \sum_j a_j \delta t_j\right)^2\right] \to 0.$$

This finishes the proof of (5.14). Indeed, we have established that

$$\sum_j a_j(\Delta W_{t_j})^2 \to \int_0^t a(s)ds \qquad \text{in the mean square sense as } \delta t_j \to 0 \tag{5.15}$$

and this is often expressed shortly by the striking formula

$$(dW_t)^2 = dt. \tag{5.16}$$

This completes the proof (5.10) for $X_t = W_t$. For the proof of (5.9), see Theorem 4.1.2 of Oksendal or Theorem 4.4.1 of Shreve II. □

Finally, we present (whose proof is similar to that of Theorem 5.1 and is skipped) the multi-dimensional Itô formula (Section 4.2 of Oksendal). Let $W(t) = (W_1(t), W_2(t), \cdots, W_m(t))^\top \in \mathbb{R}^{m\times 1}$ denote m-dimensional Brownian motion. $W_1(t), \cdots, W_m(t)$ are independent [49]. Like in Definition 5.1, we can introduce the following n-dimensional Itô process

$$\begin{cases} dX_1 = & u_1 dt + v_{11} dW_1 + \cdots + v_{1m} dW_m \\ \vdots & \vdots \\ dX_n = & u_n dt + v_{n1} dW_1 + \cdots + v_{nm} dW_m \end{cases} \tag{5.17}$$

where u_i, v_{ij} are $\mathcal{F}_t$-adapted stochastic process and $\int_0^t |u_i(s)|ds < \infty$, $\int_0^t |v_{ij}(s)|^2 ds < \infty$ for every t and almost all ω. We can also rewrite (5.17) using matrix notation

$$dX(t) = udt + vdW(t) \tag{5.18}$$

where

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}, u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix}, dW(t) = \begin{pmatrix} dW_1(t) \\ \vdots \\ dW_m(t) \end{pmatrix} \tag{5.19}$$

[49] Recall that since the covariance matrix is a diagonal matrix tI_m in (3.28), our construction ensures that the components of m-dimensional Brownian motion are independent.

Theorem 5.2 *(Theorem 4.2.1 of Oksendal) Let $dX(t) = udt + vdW(t)$ be the n-dimensional Itô process (5.17). Let $g(t,x) = (g_1(t,x), \cdots, g_p(t,x))^\top \in \mathbb{R}^{p\times 1}$ be twice continuously differentiable with respect to its variables t and $x = (x_1, \cdots, x_n)$. Then for each $k \in \{1, 2, ..., p\}$,*

$$dg_k(t, X(t,\omega)) = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j \tag{5.20}$$

where $dW_i(t)dW_j(t) = \begin{cases} dt & \textit{if } i = j \\ 0 & \textit{otherwise} \end{cases}$, $dtdW_i(t) = dW_i(t)dt = 0.$

Remark: : If X_t is a stochastic process, then

$$d(e^{-rt}X_t) = X_t de^{-rt} + e^{-rt}dX_t. \tag{5.21}$$

To prove (5.21), one can use the 1-dimensional Itô formula on $g(t, X_t)$ with $g(t,x) = e^{-rt}x$ (hence $\frac{\partial^2 g}{\partial x^2} = 0$)

$$dg(t, X_t) = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\cancel{\frac{\partial^2 g}{\partial x^2}}(t, X_t)(dX_t)^2.$$

Alternatively, one can use the product rule in Q16 of Homework IV: Because $de^{-rt} = -re^{-rt}dt$ does not contain a dW_t term, $(de^{-rt})(dX_t) = 0$. Hence

$$d(e^{-rt}X_t) = X_t de^{-rt} + e^{-rt}dX_t + \cancel{(de^{-rt})(dX_t)}.$$

5.1 Homework V

(Only submit solutions to Questions 5,17 of Homework IV and Questions 5,6,8 of this section.)

1. (An application of conditional expectation) Show that if M_t is a martingale and $\theta \in \mathcal{V}(0,T)$, then
$$\mathbb{E}\left[\int_0^T \theta_t dM_t\right] = 0.$$
Here, you can assume $\int_0^T \theta_t dM_t = \lim_{n\to\infty} \sum_{i=0}^{n-1} \theta_{t_i}(M_{t_{i+1}} - M_{t_i})$. (This problem just want to show you a heuristic argument which uses conditional expectation. It shows again how conditional expectation helps in the calculation. If you want to see a fully rigorous proof, see Page 109 of Duffie's Dynamic Asset Pricing Theory, 3rd edition.)

   Proof: Using iterated conditioning and martingale property i.e., $\mathbb{E}\left[M_{t_{i+1}} \middle| \mathcal{F}_{t_i}\right] = M_{t_i}$, we get
$$\begin{aligned}
\mathbb{E}\left[\int_0^T \theta_t dM_t\right] &= \mathbb{E}\left[\lim_{n\to\infty} \sum_{i=0}^{n-1} \theta_{t_i}(M_{t_{i+1}} - M_{t_i})\right] = \lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[\theta_{t_i}(M_{t_{i+1}} - M_{t_i})\right] \\
&= \lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[\mathbb{E}\left[\theta_{t_i}(M_{t_{i+1}} - M_{t_i}) \middle| \mathcal{F}_{t_i}\right]\right] \\
&= \lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[\theta_{t_i}\mathbb{E}\left[(M_{t_{i+1}} - M_{t_i}) \middle| \mathcal{F}_{t_i}\right]\right] = \lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[\theta_{t_i}(M_{t_i} - M_{t_i})\right] = 0.
\end{aligned}$$

2. (Precise meaning of the coefficients. But omit the proof on first reading) Suppose $S_t = S_0 + \int_0^t \mu(s,\omega)ds + \int_0^t \sigma(s,\omega)dW_s$ where S_0 is a given number, μ is an adapted process, $\sigma \in \mathcal{V}(0,T)$, $\mu(t,\omega)$ and $\sigma(t,\omega)$ are continuous functions of t for every ω. This integral form is equivalent to the differential form
$$dS_t = \mu(t,\omega)dt + \sigma(t,\omega)dW_t.$$
Show that
$$\frac{d}{d\tau}\,\mathbb{E}[S_\tau|\mathcal{F}_t]|_{\tau\downarrow t} = \mu(t,\omega) \tag{5.22}$$
and
$$\frac{d}{d\tau}\,\mathrm{Var}[S_\tau|\mathcal{F}_t]|_{\tau\downarrow t} = \sigma^2(t,\omega) \tag{5.23}$$
where the conditional variance is defined to be $\mathrm{Var}[X|\mathcal{G}] = \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2 \middle| \mathcal{G}\right]$. (These are the statements on Page 86 of Duffie's Dynamic Asset Pricing Theory, 3rd edition. The following proof, which requires conditional variance and Ito isometry for $\int_t^\tau \sigma dW_s$ conditioning on $\mathcal{F}_t$, is very technical and definitely won't be tested. It is presented merely because it is an application of conditional expectation and gives a precise meaning of the coefficients in the differential form.)

Proof: (a) Let $\tau > t$. $\int_0^t \sigma(s,\omega)dW_s$ is a martingale implies $\mathbb{E}[\int_0^\tau \sigma(s,\omega)dW_s|\mathcal{F}_t] = \int_0^t \sigma(s,\omega)dW_s$.

$$\mathbb{E}[S_\tau|\mathcal{F}_t] = S_0 + \int_0^\tau \mathbb{E}[\mu(s,\omega)|\mathcal{F}_t]ds + \int_0^t \sigma(s,\omega)dW_s.$$

$$\frac{d}{d\tau}\mathbb{E}[S_\tau|\mathcal{F}_t] = \mathbb{E}[\mu(\tau,\omega)|\mathcal{F}_t].$$

$$\frac{d}{d\tau}\,\mathbb{E}[S_\tau|\mathcal{F}_t]|_{\tau\downarrow t} = \mathbb{E}[\mu(\tau,\omega)|\mathcal{F}_t]|_{\tau\downarrow t} = \mathbb{E}[\mu(t,\omega)|\mathcal{F}_t] = \mu(t,\omega).$$

In the last step, we have used the fact that $\mu(t,\omega)$ is $\mathcal{F}_t$-measurable.

(b) Let $\tau > t$. $\int_0^\tau \mathbb{E}[\mu(s,\omega)|\mathcal{F}_t]ds = \int_0^t \mathbb{E}[\mu(s,\omega)|\mathcal{F}_t]ds + \int_t^\tau \mathbb{E}[\mu(s,\omega)|\mathcal{F}_t]ds = \int_0^t \mu(s,\omega)ds + \int_t^\tau \mathbb{E}[\mu(s,\omega)|\mathcal{F}_t]ds$. So

$$S_\tau - \mathbb{E}[S_\tau|\mathcal{F}_t] = \int_t^\tau (\mu(s,\omega) - \mathbb{E}[\mu(s,\omega)|\mathcal{F}_t])ds + \int_t^\tau \sigma(s,\omega)dW_s \stackrel{\text{def}}{=} g + \int_t^\tau \sigma(s,\omega)dW_s.$$

Note that

$$\frac{d}{d\tau}g = \mu(\tau,\omega) - \mathbb{E}[\mu(\tau,\omega)|\mathcal{F}_t]$$

which vanishes after setting $\tau\downarrow t$. Hence when computing $\frac{d}{d\tau}\,\mathbb{E}[(S_\tau - \mathbb{E}[S_\tau|\mathcal{F}_t])^2|\mathcal{F}_t]|_{\tau\downarrow t}$, both g^2 and $g\int_t^\tau \sigma(s,\omega)dW_s$ vanish after taking $\frac{d}{d\tau}\big|_{\tau\downarrow t}$. (Basically, it says that g is a small term and can be ignored or will disappear in the end.) One should also notice that since we are conditioning on $\mathcal{F}_t$, and since we are integrating from t to a later time, the whole proof of Itô isometry (e.g. Lemma 4.1) remains valid with $\mathbb{E}$ replaced by $\mathbb{E}[\cdot|\mathcal{F}_t]$, and we still have $\mathbb{E}[\left(\int_t^\tau \sigma(s,\omega)dW_s\right)^2|\mathcal{F}_t] = \int_t^\tau \mathbb{E}[\sigma^2(s,\omega)|\mathcal{F}_t]ds$. So, using the aforementioned Itô isometry with conditional expectation to $\int_t^\tau \sigma(s,\omega)dW_s$, we get

$$\begin{aligned}\frac{d}{d\tau}\,\mathbb{E}[(S_\tau - \mathbb{E}[S_\tau|\mathcal{F}_t])^2|\mathcal{F}_t]\big|_{\tau\downarrow t} &= \frac{d}{d\tau}\left(\int_t^\tau \mathbb{E}[\sigma^2(s,\omega)|\mathcal{F}_t]ds\right)\Big|_{\tau\downarrow t}\\ &= \mathbb{E}[\sigma^2(\tau,\omega)|\mathcal{F}_t]\big|_{\tau\downarrow t}\\ &= \mathbb{E}[\sigma^2(t,\omega)|\mathcal{F}_t] = \sigma^2(t,\omega).\end{aligned}$$

In the last step, we have used again the fact that $\sigma(t,\omega)$ is $\mathcal{F}_t$-measurable.

3. a) Assume r and σ are constants. Show that if f is continuously differentiable, then $S_t = e^{rt+\sigma f(t)}$ satisfies

$$dS_t = rS_tdt + \sigma S_t f'(t)dt = rS_tdt + \sigma S_t df(t). \tag{5.24}$$

b) Suppose we want to solve for S_t that satisfies

$$dS_t = rS_tdt + \sigma S_t dW_t. \tag{5.25}$$

Motivated by (1), we may try $S_t = e^{rt+\sigma W_t}$. Use Itô formula with $g(t,x) = e^{rt+\sigma x}$ to show that

$$dS_t = \left(r + \frac{1}{2}\sigma^2\right) S_t dt + \sigma S_t dW_t.$$

c) Show that if $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}$, then it satisfies (5.25).

d) Show that if S_t is the solution of (5.25), then

$$\mathbb{E}[\log(S_{t+\delta t}/S_t)] = (r - \frac{\sigma^2}{2})\delta t, \qquad \mathrm{Var}[\log(S_{t+\delta t}/S_t)] = \sigma^2 \delta t. \tag{5.26}$$

Solution: (a) If $S_t = e^{rt+\sigma f(t)}$, $\frac{dS_t}{dt} = (r + \sigma f'(t))e^{rt+\sigma f(t)}$. Hence S_t satisfies (5.24).

(b) We use (5.9) with $X_t = W_t$ to get

$$\begin{aligned} dS_t =& re^{rt+\sigma W_t} dt + \sigma e^{rt+\sigma W_t} dW_t + \frac{1}{2}\sigma^2 e^{rt+\sigma W_t} dt \\ =& \left(r + \frac{1}{2}\sigma^2\right) S_t dt + \sigma S_t dW_t. \end{aligned}$$

(c) Let $\tilde{r} = r - \frac{1}{2}\sigma^2$. If $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t} = S_0 e^{\tilde{r}t+\sigma W_t}$,

$$dS_t = \left(\tilde{r} + \frac{1}{2}\sigma^2\right) S_t dt + \sigma S_t dW_t = rS_t dt + \sigma S_t dW_t.$$

(d) Since $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}$, $S_{t+\delta t}/S_t = e^{(r-\frac{\sigma^2}{2})\delta t+\sigma(W_{t+\delta t}-W_t)}$. Hence

$$\log(S_{t+\delta t}/S_t) = \left(r - \frac{1}{2}\sigma^2\right)\delta t + \sigma(W_{t+\delta t} - W_t).$$

This leads to $\mathbb{E}[\log(S_{t+\delta t}/S_t)] = \left(r - \frac{1}{2}\sigma^2\right)\delta t$ and $\mathrm{Var}[\log(S_{t+\delta t}/S_t)] = \sigma^2\delta t$.

4. Find a stochastic process $X_t \geq 0$ and a Brownian motion B_t so that

$$dX_t = dt + 2\sqrt{X_t} dB_t, \qquad X_0 = x_0 \geq 0.$$

Proof: Let W_t be any Brownian motion with $W_0 = \sqrt{x_0}$. Let $X_t = (W_t)^2$. By (5.9),

$$dX_t = 2W_t dW_t + dt = 2|W_t|\mathrm{sign}(W_t)dW_t + dt = 2\sqrt{X_t}\mathrm{sign}(W_t)dW_t + dt$$

where $\mathrm{sign}(a) = \begin{cases} 1 & \text{if } a > 0 \\ 0 & \text{if } a = 0 \\ -1 & \text{if } a < 0 \end{cases}$. Define $B_t = \int_0^t \mathrm{sign}(W_s)dW_s$. Then B_t is continuous (Theorem 4.2) and is a martingale (Theorem 4.4). Moreover, by definition (equivalence between integral form and differential form) $dB_t = \mathrm{sign}(W_t)dW_t$. So

$$dB_t \cdot dB_t = \mathrm{sign}(W_t)dW_t \cdot \mathrm{sign}(W_t)dW_t = dW_t \cdot dW_t = dt.$$

By Theorem 3.5 of Levy, we know B_t is a Brownian motion. Hence we have find $X_t \geq 0$ and a Brownian motion B_t so that

$$dX_t = dt + 2\sqrt{X_t}dB_t, \qquad X_0 = x_0 \geq 0.$$

5. Let α and σ be constants. We want to find X_t that satisfies

$$dX_t = -\alpha X_t dt + \sigma dW_t, \qquad X_0 = x_0. \tag{5.27}$$

(The solution is the so-called Ornstein-Uhlenbeck process.) Show that by Itô formula with $g(t,x) = e^{\alpha t}x$, we have

$$d(e^{\alpha t}X_t) = \sigma e^{\alpha t}dW_t. \tag{5.28}$$

By the equivalence between (5.6) and (5.5), prove that

$$X_t = e^{-\alpha t}x_0 + e^{-\alpha t}\sigma \int_0^t e^{\alpha s}dW_s.$$

6. (Continuation of Questions 4 and 5) Suppose that $X_1(t)$ and $X_2(t)$ are the Ornstein-Uhlenbeck processes with $dX_i(t) = -\alpha X_i(t)dt + \sigma dW_i(t)$ for $i = 1, 2$, where $(W_1(t), W_2(t))$ is a 2-dimensional Brownian motion. Let $\mathbf{X} = (X_1, X_2)$. Recall the 2-dimensional Itô formula (Theorem 5.2)

$$dg(t, \mathbf{X}(t,\omega)) = \frac{\partial g}{\partial t}(t,\mathbf{X})dt + \sum_{i=1}^{2} \frac{\partial g}{\partial x_i}(t,\mathbf{X})dX_i + \frac{1}{2}\sum_{i,j=1}^{2} \frac{\partial^2 g}{\partial x_i \partial x_j}(t,\mathbf{X})dX_i dX_j$$

where $dW_i(t)dW_j(t) = \begin{cases} dt & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$, $dtdW_i(t) = dW_i(t)dt = 0$. **Find** the constant β so that $Y(t) = (X_1(t))^2 + (X_2(t))^2$ satisfies the Cox-Ingersoll-Ross (CIR) model

$$dY_t = (\beta - 2\alpha Y_t)dt + 2\sigma\sqrt{Y_t}d\tilde{W}_t,$$

where $\tilde{W}_t = \int_0^t \frac{X_1(s)}{\sqrt{X_1^2(s) + X_2^2(s)}}dW_1(s) + \frac{X_2(s)}{\sqrt{X_1^2(s) + X_2^2(s)}}dW_2(s)$ is a continuous martingale. For your information, $d\tilde{W}_t d\tilde{W}_t = dt$[50]. Hence by Levy's criteria of Brownian motion (Theorem 3.5), $\tilde{W}_t$ is a Brownian motion. [Hint: Take $g(t, x_1, x_2) = x_1^2 + x_2^2$. Then $Y_t = g(t, X_1(t), X_2(t))$]

---

[50] $d\tilde{W}_t = \frac{X_1(t)}{\sqrt{(X_1(t))^2+(X_2(t))^2}}dW_1(t) + \frac{X_2(t)}{\sqrt{(X_1(t))^2+(X_2(t))^2}}dW_2(t)$. Hence (there is no cross tem because $dW_1(t)dW_2(t) = 0$)

$$\begin{aligned} d\tilde{W}_t d\tilde{W}_t &= \left(\frac{X_1(t)}{\sqrt{(X_1(t))^2 + (X_2(t))^2}}\right)^2 (dW_1)^2 + \left(\frac{X_2(t)}{\sqrt{(X_1(t))^2 + (X_2(t))^2}}\right)^2 (dW_2)^2 \\ &= \left(\frac{X_1(t)}{\sqrt{(X_1(t))^2 + (X_2(t))^2}}\right)^2 dt + \left(\frac{X_2(t)}{\sqrt{(X_1(t))^2 + (X_2(t))^2}}\right)^2 dt = dt. \end{aligned}$$

7. Consider the following CIR model for the evolution of interest rates

$$dr_t = \alpha(\gamma - r_t)dt + \rho\sqrt{r_t}dW_t. \tag{5.29}$$

Let $L_t = \log(r_t)$. Use Itô formula to prove that

$$dL_t = \left((\alpha\gamma - \frac{\rho^2}{2})e^{-L_t} - \alpha\right)dt + \rho e^{-L_t/2}dW_t. \tag{5.30}$$

(For your information, one can prove that if $\alpha\gamma - \frac{\rho^2}{2} \geq 0$, r_t remains strictly positive if $r_0 > 0$. See for example Cairns "Interest rate models" for a rigorous proof.)

Proof: Let $g(t,x) = g(x) = \log x$. Then $g'(x) = \frac{1}{x}$, $g''(x) = -\frac{1}{x^2}$. Note that $r_t^\beta = e^{\beta \log r_t} = e^{\beta L_t}$. By Itô formula

$$\begin{aligned}
d(\log r_t) &= \frac{1}{r_t}dr_t - \frac{1}{2}\frac{1}{r_t^2}(dr_t)^2 \\
&= \alpha(\gamma r_t^{-1} - 1)dt + \rho r_t^{-1/2}dW_t - \frac{1}{2}\frac{1}{r_t^2}\rho^2 r_t dt \\
&= \left((\alpha\gamma - \frac{\rho^2}{2})r_t^{-1} - \alpha\right)dt + \rho r_t^{-1/2}dW_t \\
&= \left((\alpha\gamma - \frac{\rho^2}{2})e^{-L_t} - \alpha\right)dt + \rho e^{-L_t/2}dW_t.
\end{aligned}$$

8. Use Itô formula to prove that

$$\int_0^t W_s^2 dW_s = \frac{1}{3}W_t^3 - \int_0^t W_s ds.$$

Here we assume $W_0 = 0$ as usual.

9. (Shreve II, Corollary 4.6.3) We want to generalize Example 5.4 to the product of two stochastic processes. Let X and Y be 1-dimensional Itô processes. Prove that

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t). \tag{5.31}$$

Proof: Let $g(t,x_1,x_2) = x_1x_2$. Note that $\frac{\partial^2 g}{\partial x_1^2} = 0 = \frac{\partial^2 g}{\partial x_2^2}$. By (5.20)

$$\begin{aligned}
&dg(t, X(t), Y(t)) \\
=&\frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x_1}(t,X,Y)dX + \frac{\partial g}{\partial x_2}(t,X,Y)dY + 2 \times \frac{1}{2}\frac{\partial^2 g}{\partial x_1 \partial x_2}(t,X,Y)dXdY \\
=&YdX + XdY + dXdY.
\end{aligned}$$

10. Let $\{W_t : t \geq 0\}$ be a 1-dimensional Brownian motion with $W_0 = 0$. Show that

$$\int_0^t W_s ds = \int_0^t (t-s) dW_s \tag{5.32}$$

and prove that

$$\int_0^t W_s ds \sim N(0, t^3/3).$$

You can use the fact that $\int_0^t f(s) dW_s$ is normally distributed if $f(s)$ is a deterministic function of s. By the way, in the midterm, you have learned how to compute the variance of $\int_0^t W_s ds$ directly without this integration by parts trick.

Proof: By the integration by parts formula 5.11 with $f(s) = t - s$ where t is fixed. Hence $f(t) = 0$, $df(s) = -ds$.

$$\int_0^t (t-s) dW_s = \int_0^t W_s ds.$$

By (4.15), $\mathbb{E}[\int_0^t (t-s) dW_s] = 0$.

$$\text{Var}\left[\int_0^t (t-s) dW_s\right] = \mathbb{E}\left[\left(\int_0^t (t-s) dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t (t-s)^2 ds\right] = \left.\frac{-(t-s)^3}{3}\right|_{s=0}^{s=t} = \frac{t^3}{3}.$$

11. Let $\{W_t : t \geq 0\}$ be a 1-dimensional Brownian motion with $W_0 = 0$. Show that

$$\mathbb{E}(W_t^n) = \frac{1}{2} n(n-1) \int_0^t \mathbb{E}\left(W_s^{n-2}\right) ds \tag{5.33}$$

and then use induction to prove that

$$\mathbb{E}(W_t^n) = \begin{cases} \frac{n! t^{n/2}}{2^{n/2}(n/2)!} & n = 2, 4, 6, \cdots \\ 0 & n = 1, 3, 5, \cdots \end{cases}. \tag{5.34}$$

Proof: Let $g(x) = x^n$. Apply Itô formula to $g(W_t)$, we get

$$dW_t^n = n W_t^{n-1} dW_t + \frac{1}{2} n(n-1) W_t^{n-2} (dW_t)^2$$

which can be rewritten into integral form

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

We obtain (5.33) after taking expectation on both sides of the above equation and using (4.15).

We prove (5.34) by induction. It is true when $n = 1$ and 2. Assume it is true for $n = 2k-1$ and $2k$. It is certainly true for $n = 2k+1$ by (5.33). For $n = 2k+2$,

$$\begin{aligned}\mathbb{E}(W_t^{2k+2}) &= \frac{1}{2}(2k+2)(2k+1)\int_0^t \mathbb{E}\left(W_s^{2k}\right) ds = \frac{1}{2}(2k+2)(2k+1)\int_0^t \frac{(2k)!s^k}{2^k k!} ds \\ &= \frac{1}{2}(2k+2)(2k+1)\frac{(2k)!t^{k+1}}{2^k k!(k+1)} = \frac{(2k+2)!t^{k+1}}{2^{k+1}(k+1)!}.\end{aligned}$$

This finishes the proof.

12. Let $M_t = \int_0^t f(s)dW_s$ and

$$Y_t = e^{\theta M_t - \frac{1}{2}\theta^2 \int_0^t f^2(s)ds}.$$

Show that Y_t satisfies

$$dY_t = \theta f(t) Y_t dW_t.$$

Proof: Let $X_t = \theta M_t - \frac{1}{2}\theta^2 \int_0^t f^2(s)ds$ which means $dX_t = -\frac{1}{2}\theta^2 f^2(t)dt + \theta f(t)dW_t$.

$$\begin{aligned}de^{X_t} &= e^{X_t}dX_t + \frac{1}{2}e^{X_t}(dX_t)^2 \\ &= e^{X_t}(-\frac{1}{2}\theta^2 f^2(t)dt + \theta f(t)dW_t) + \frac{1}{2}e^{X_t}\theta^2 f^2(t)dt \\ &= e^{X_t}\theta f(t)dW_t.\end{aligned}$$