# Stochastic Calculus and Modern Volatility Modelling

Author: Kenneth Zhang
University of Waterloo - Supervisor Dr. Alexander Schied

**Table of Contents**

# Contents

# 1 Brownian Motion

## 1.1 Symmetric Random Walks

Suppose we toss a fair coin where the probability of it landing heads is $p$ and the probability of it landing tails is $q = 1 - p$ i.e., a fair coin $p = q = \frac{1}{2}$. For outcomes $\omega = \omega_1\omega_2\omega_3\cdots$ where $\omega_n$ denotes the outcome of the $n$th toss, we define,

$$X_j = \begin{cases} 1 & \omega_j = H \\ -1 & \omega_j = T \end{cases}.$$

Then, a *Symmetric Random Walk* for $k = 0, \ldots, K$ is defined as,

$$M_k = \sum_{j=1}^{K} X_k.$$

**Theorem 1.1.** ***Properties of Symmetric Random Walks***

*(i)* ***Independent Increments.*** $M_{k_1} = (M_{k_1} - M_{k_0}), (M_{k_2} - M_{k_1}), \ldots, (M_{k_m} - M_{k_{m-1}})$ *for* $0 = k_0 < k_1 < \cdots < k_m$ *are the increments of the Symmetric Random Walk and are independent.*

*(ii)* ***Variance and Expected Value.*** $\mathbb{E}[M_{k_{i+1}} - M_{k_i}] = 0$ *and* $\operatorname{Var}(M_{k_{i+1}} - M_{k_i}) = k_{i+1} - k_i$*.*

**Proof.**

$$\mathbb{E}[M_{k_{i+1}} - M_{k_i}] = \mathbb{E}\left[\sum_{j=k_i+1}^{k_{i+1}} X_j\right] = \sum_{j=k_i+1}^{k_{i+1}} \mathbb{E}[X_j] = \sum_{j=k_i+1}^{k_{i+1}} 0 = 0,$$

$$\operatorname{Var}(M_{k_{i+1}} - M_{k_i}) = \operatorname{Var}\left(\sum_{j=k_i+1}^{k_{i+1}} X_j\right) = \sum_{j=k_i+1}^{k_{i+1}} \operatorname{Var}(X_j) = \sum_{j=k_i+1}^{k_{i+1}} 1 = k_{i+1} - k_i.$$

*(iii)* ***Martingale Property.*** $\mathbb{E}[M_\ell \mid \mathcal{F}_k] = M_k$ *for all* $k \le \ell$*.*

**Proof.**

$$M_\ell = \sum_{j=1}^{\ell} X_j = \sum_{j=1}^{k} X_j + \sum_{j=k+1}^{\ell} X_j = M_k + (M_\ell - M_k).$$

$$\begin{aligned}\mathbb{E}[M_\ell \mid \mathcal{F}_k] &= \mathbb{E}[\, M_k + (M_\ell - M_k) \,|\, \mathcal{F}_k] \\ &= \mathbb{E}[M_k \mid \mathcal{F}_k] + \mathbb{E}[M_\ell - M_k \mid \mathcal{F}_k] = M_k + \mathbb{E}\left[\sum_{j=k+1}^{\ell} X_j \,\middle|\, \mathcal{F}_k\right].\end{aligned}$$

$$\mathbb{E}\left[\sum_{j=k+1}^{\ell} X_j \,\middle|\, \mathcal{F}_k\right] = \sum_{j=k+1}^{\ell} \mathbb{E}[X_j \mid \mathcal{F}_k].$$

Since $X_j$ is independent of $\mathcal{F}_k$ for all $j > k$ and $\mathbb{E}[X_j] = 0$, we have $\mathbb{E}[X_j \mid \mathcal{F}_k] = 0$. Hence

$$\mathbb{E}[M_\ell - M_k \mid \mathcal{F}_k] = \sum_{j=k+1}^{\ell} 0 = 0.$$

Therefore,

$$\mathbb{E}[M_\ell \mid \mathcal{F}_k] = M_k.$$

*(iv)* ***Quadratic Variation.*** $[M, M]_k = \sum_{j=1}^{k} (M_j - M_{j-1})^2 = \sum_{j=1}^{k} X_j^2 = \sum_{j=1}^{k} 1 = k.$

### 1.2 Scaled Random Walks

Consider $W^{(n)}(t) = \dfrac{1}{\sqrt{n}} M_{nt}$, a continuous-time process that can converge to the Brownian motion where, at time $t$, the walk has taken $nt$ steps. We scale the random walk by $\frac{1}{\sqrt{n}}$ so that our random walk "runs" faster in time. If $t$ is odd, then for $s, u \in \mathbb{Z}^+$ we interpolate between the points $ns$ and $nu$. Then, the *Scaled Random Walk* has the following properties.

**Theorem 1.2. 1.2.1 Properties of Scaled Random Walks**

***Properties of Scaled Random Walks.***

(i) ***Scaled Random Increments.***

**Proof.**

$$W^{(n)}(t) - W^{(n)}(s) = \frac{1}{\sqrt{n}} M_{nt} - \frac{1}{\sqrt{n}} M_{ns} = \frac{1}{\sqrt{n}}(M_{nt} - M_{ns}) = \frac{1}{\sqrt{n}} \left\{ \sum_{j=1}^{nt} X_j - \sum_{j=1}^{ns} X_j \right\} = \frac{1}{\sqrt{n}} \left\{ \sum_{j=ns+1}^{nt} X_j \right\}.$$

(ii) ***Expected Value.*** $\mathbb{E}\left[W^{(n)}(t) - W^{(n)}(s)\right] = 0.$

**Proof.**

$$\mathbb{E}\left[W^{(n)}(t) - W^{(n)}(s)\right] = \mathbb{E}\left[\frac{1}{\sqrt{n}} \sum_{j=ns+1}^{nt} X_j\right] = \frac{1}{\sqrt{n}} \sum_{j=ns+1}^{nt} \mathbb{E}[X_j] = \frac{1}{\sqrt{n}} \sum_{j=ns+1}^{nt} 0 = 0.$$

(iii) ***Variance.*** $\mathrm{Var}\left(W^{(n)}(t) - W^{(n)}(s)\right) = t - s.$

**Proof.**

$$\mathrm{Var}\left(W^{(n)}(t) - W^{(n)}(s)\right) = \mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{j=ns+1}^{nt} X_j\right) = \frac{1}{n} \sum_{j=ns+1}^{nt} \mathrm{Var}(X_j) = \frac{1}{n} \sum_{j=ns+1}^{nt} 1 = \frac{1}{n}(nt - ns) = t - s.$$

*Remark* 1.1. Notice that the term $\frac{1}{\sqrt{n}} \sum X_j$ takes the form of the Central Limit Theorem (CLT). As $n \to \infty$, the distribution of $W^{(n)}(t)$ converges to a Normal distribution $N(0, t)$. This is why Brownian motion is Gaussian.

*Remark* 1.2. The convergence of the entire path of $W^{(n)}(t)$ to Standard Brownian Motion $B(t)$ as $n \to \infty$ is known as *Donsker's Invariance Principle* (or the Functional Central Limit Theorem). This theorem justifies using Brownian motion to model discrete phenomena (like stock prices) over long time horizons.

We divide by $\sqrt{n}$ (space scaling) because we accelerated time by a factor of $n$. Since $\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = n\mathrm{Var}(X_i) = n$, the standard deviation grows as $\sqrt{n}$. To keep the fluctuations finite and non-zero as $n \to \infty$, we must normalize by the standard deviation, $\sqrt{n}$.

*Remark* 1.3. While the discrete random walk $M_{nt}$ jumps at integer times, the definition of $W^{(n)}(t)$ implies linear interpolation between steps for non-integer $t$. This ensures the resulting function is continuous in $t$, a requirement for convergence to Brownian Motion (which has continuous paths almost surely).

### 1.2.2 Martingale Property of Scaled Random Walks

**Theorem 1.3.** ***Scaled Random Walk is a Martingale.***

$$\mathbb{E}[W^{(n)}(t) \mid \mathcal{F}(s)] = W^{(n)}(s).$$

**Proof.** We use the increment trick as introduced in STAT 333,

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt} = \frac{1}{\sqrt{n}}(M_{ns} + M_{nt} - M_{ns}),$$

so we have $\mathbb{E}[M_{ns} \mid \mathcal{F}(s)]$ by measurability. Then,

$$\begin{aligned}
\mathbb{E}\Big[W^{(n)}(t) \mid \mathcal{F}_{ns}\Big] &= \frac{1}{\sqrt{n}}\mathbb{E}[M_{ns} + (M_{nt} - M_{ns}) \mid \mathcal{F}_{ns}] \\
&= \frac{1}{\sqrt{n}}\left(\mathbb{E}[M_{ns} \mid \mathcal{F}_{ns}] + \mathbb{E}[M_{nt} - M_{ns} \mid \mathcal{F}_{ns}]\right), \\
\mathbb{E}[M_{nt} - M_{ns} \mid \mathcal{F}_{ns}] = \mathbb{E}\left[\sum_{j=ns+1}^{nt} X_j \,\middle|\, \mathcal{F}_{ns}\right] &= \sum_{j=ns+1}^{nt} \mathbb{E}[X_j \mid \mathcal{F}_{ns}] \\
&= \sum_{j=ns+1}^{nt} \mathbb{E}[X_j] \\
&= 0.
\end{aligned}$$

So we get that,

$$\mathbb{E}[W^{(n)}(t) \mid \mathcal{F}(s)] = \frac{1}{\sqrt{n}} M_{ns} = W^{(n)}(s),$$

as required.

**Proposition 1.1.** *Let*

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nt \rfloor} X_j,$$

*where $(X_j)_{j\geq 1}$ are i.i.d. with $\mathbb{E}[X_j] = 0$ and $\mathrm{Var}(X_j) = 1$. Then the quadratic variation of the scaled random walk is*

$$[W^{(n)}, W^{(n)}]_t = \sum_{j=1}^{\lfloor nt \rfloor} \left(W^{(n)}\left(\frac{j}{n}\right) - W^{(n)}\left(\frac{j-1}{n}\right)\right)^2 = \frac{\lfloor nt \rfloor}{n}.$$

*In particular, for each fixed $t \geq 0$,*

$$[W^{(n)}, W^{(n)}]_t \xrightarrow[n\to\infty]{} t.$$

**Proof.** Recall that

$$W^{(n)}\left(\frac{j}{n}\right) = \frac{1}{\sqrt{n}} M_j, \qquad W^{(n)}\left(\frac{j-1}{n}\right) = \frac{1}{\sqrt{n}} M_{j-1}.$$

Hence,

$$W^{(n)}\left(\frac{j}{n}\right) - W^{(n)}\left(\frac{j-1}{n}\right) = \frac{1}{\sqrt{n}}(M_j - M_{j-1}) = \frac{1}{\sqrt{n}} X_j.$$

Therefore, the quadratic variation up to time $t$ is

$$[W^{(n)}, W^{(n)}]_t = \sum_{j=1}^{\lfloor nt \rfloor} \left(W^{(n)}\left(\frac{j}{n}\right) - W^{(n)}\left(\frac{j-1}{n}\right)\right)^2 = \sum_{j=1}^{\lfloor nt \rfloor} \frac{1}{n} X_j^2.$$

Since $\mathbb{E}[X_j^2] = 1$, we have

$$[W^{(n)}, W^{(n)}]_t = \frac{\lfloor nt \rfloor}{n} \xrightarrow[n\to\infty]{} t.$$

**Theorem 1.4.** *Fix $t \geq 0$. For each $n \in \mathbb{N}$ such that $nt \in \mathbb{N}$, define*

$$W^{(n)}(t) := \frac{1}{\sqrt{n}} M_{nt}, \qquad M_k := \sum_{j=1}^{k} X_j,$$

*where $(X_j)_{j\geq 1}$ are i.i.d. random variables,*

$$\mathbb{P}(X_j = 1) = \mathbb{P}(X_j = -1) = \frac{1}{2}.$$

*Then, as $n \to \infty$ along integers with $nt \in \mathbb{N}$,*

$$W^{(n)}(t) \xrightarrow{d} \mathcal{N}(0, t).$$

**Proof.** We begin by explicitly deriving the moment generating function of the limiting normal distribution $Z \sim \mathcal{N}(0, t)$. We substitute the density, combine the exponents, and complete the square step-by-step.

$$\begin{aligned}
f_Z(z) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} \\
\phi(u) := \mathbb{E}[e^{uZ}] &= \int_{-\infty}^{\infty} e^{uz} \left( \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} \right) dz \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left( uz - \frac{z^2}{2t} \right) dz \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left( -\frac{1}{2t}(z^2 - 2utz) \right) dz \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left( -\frac{1}{2t}(z^2 - 2utz + u^2t^2 - u^2t^2) \right) dz \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left( -\frac{1}{2t}(z^2 - 2utz + (ut)^2) + \frac{u^2t^2}{2t} \right) dz \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left( -\frac{(z-ut)^2}{2t} \right) \exp\left( \frac{u^2t}{2} \right) dz \\
&= \exp\left( \frac{u^2t}{2} \right) \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left( -\frac{(z-ut)^2}{2t} \right) dz}_{=1} \\
&= \exp\left( \frac{u^2t}{2} \right).
\end{aligned}$$

Next, we derive the moment generating function for the scaled random walk $W^{(n)}(t)$, showing the expansion from the sum in the exponent to the product of expectations.

$$\begin{aligned}
\phi_n(u) := \mathbb{E}\big[e^{uW^{(n)}(t)}\big] &= \mathbb{E}\left[ \exp\left( u \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j \right) \right] \\
&= \mathbb{E}\left[ \exp\left( \sum_{j=1}^{nt} \frac{u}{\sqrt{n}} X_j \right) \right] \\
&= \mathbb{E}\left[ \prod_{j=1}^{nt} \exp\left( \frac{u}{\sqrt{n}} X_j \right) \right] \\
&= \prod_{j=1}^{nt} \mathbb{E}\left[ \exp\left( \frac{u}{\sqrt{n}} X_j \right) \right] \\
&= \prod_{j=1}^{nt} \left( P(X_j = 1) e^{\frac{u}{\sqrt{n}}(1)} + P(X_j = -1) e^{\frac{u}{\sqrt{n}}(-1)} \right) \\
&= \prod_{j=1}^{nt} \left( \frac{1}{2} e^{u/\sqrt{n}} + \frac{1}{2} e^{-u/\sqrt{n}} \right) = \left( \frac{1}{2} e^{u/\sqrt{n}} + \frac{1}{2} e^{-u/\sqrt{n}} \right)^{nt}.
\end{aligned}$$

To analyze the limit, we apply the logarithm and the variable change $x = 1/\sqrt{n}$.

$$\begin{aligned}
\log \phi_n(u) &= \log\left(\left(\frac{1}{2}e^{u/\sqrt{n}} + \frac{1}{2}e^{-u/\sqrt{n}}\right)^{nt}\right) \\
&= nt \log\left(\frac{1}{2}e^{u/\sqrt{n}} + \frac{1}{2}e^{-u/\sqrt{n}}\right) \\
x = \frac{1}{\sqrt{n}} \quad &\implies \quad n = \frac{1}{x^2} \\
\log \phi_n(u) &= \frac{t}{x^2} \log\left(\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}\right).
\end{aligned}$$

We define $g(x)$ and prepare for L'Hôpital's rule by verifying the form $0/0$ and computing the first derivative $g'(x)$.

$$\begin{aligned}
g(x) &= \log\left(\frac{1}{2}e^{ux} + \frac{1}{2}e^{-ux}\right) \implies \lim_{n\to\infty} \log \phi_n(u) = t \lim_{x\downarrow 0} \frac{g(x)}{x^2} \\
g(0) &= \log\left(\frac{1}{2}e^{0} + \frac{1}{2}e^{0}\right) = \log(1) = 0 \\
\lim_{x\downarrow 0} \frac{g(x)}{x^2} &= \lim_{x\downarrow 0} \frac{g'(x)}{2x} \\
g'(x) &= \frac{d}{dx} \log\left(\frac{1}{2}(e^{ux} + e^{-ux})\right) \\
&= \frac{1}{\frac{1}{2}(e^{ux} + e^{-ux})} \cdot \frac{d}{dx}\left(\frac{1}{2}(e^{ux} + e^{-ux})\right) \\
&= \frac{1}{\frac{1}{2}(e^{ux} + e^{-ux})} \cdot \left(\frac{1}{2}(ue^{ux} - ue^{-ux})\right) \\
&= \frac{\frac{1}{2}u(e^{ux} - e^{-ux})}{\frac{1}{2}(e^{ux} + e^{-ux})} \\
&= u\frac{e^{ux} - e^{-ux}}{e^{ux} + e^{-ux}}.
\end{aligned}$$

Since $g'(0) = u\frac{1-1}{1+1} = 0$, we apply L'Hôpital's rule again.
We expand the quotient rule terms fully to simplify the second derivative $g''(x)$.

$$\begin{aligned}
\lim_{x\downarrow 0} \frac{g'(x)}{2x} &= \lim_{x\downarrow 0} \frac{g''(x)}{2} \\
g''(x) &= u \cdot \frac{d}{dx}\left(\frac{e^{ux} - e^{-ux}}{e^{ux} + e^{-ux}}\right) \\
&= u \cdot \frac{(e^{ux} - e^{-ux})'(e^{ux} + e^{-ux}) - (e^{ux} - e^{-ux})(e^{ux} + e^{-ux})'}{(e^{ux} + e^{-ux})^2} \\
&= u \cdot \frac{(ue^{ux} + ue^{-ux})(e^{ux} + e^{-ux}) - (e^{ux} - e^{-ux})(ue^{ux} - ue^{-ux})}{(e^{ux} + e^{-ux})^2} \\
&= u^2 \cdot \frac{(e^{ux} + e^{-ux})(e^{ux} + e^{-ux}) - (e^{ux} - e^{-ux})(e^{ux} - e^{-ux})}{(e^{ux} + e^{-ux})^2} \\
&= u^2 \cdot \frac{(e^{ux} + e^{-ux})^2 - (e^{ux} - e^{-ux})^2}{(e^{ux} + e^{-ux})^2} \\
&= u^2 \cdot \frac{(e^{2ux} + 2e^{ux}e^{-ux} + e^{-2ux}) - (e^{2ux} - 2e^{ux}e^{-ux} + e^{-2ux})}{(e^{ux} + e^{-ux})^2} \\
&= u^2 \cdot \frac{(e^{2ux} + 2 + e^{-2ux}) - (e^{2ux} - 2 + e^{-2ux})}{(e^{ux} + e^{-ux})^2} \\
&= u^2 \cdot \frac{e^{2ux} - e^{2ux} + e^{-2ux} - e^{-2ux} + 2 - (-2)}{(e^{ux} + e^{-ux})^2} \\
&= u^2 \cdot \frac{4}{(e^{ux} + e^{-ux})^2}.
\end{aligned}$$

We substitute $x = 0$ into the simplified second derivative and solve the limit.

$$g''(0) = u^2 \cdot \frac{4}{(e^0 + e^0)^2} = u^2 \cdot \frac{4}{(1+1)^2} = u^2 \cdot \frac{4}{4} = u^2 \lim_{x \downarrow 0} \frac{g'(x)}{2x} \qquad = \frac{g''(0)}{2} = \frac{u^2}{2}.$$

Finally, we substitute the limit back into the expression for $\log \phi_n(u)$ and exponentiate to complete the proof.

$$\lim_{n\to\infty} \log \phi_n(u) = t\left(\frac{u^2}{2}\right) = \frac{u^2 t}{2} \quad \text{and} \quad \lim_{n\to\infty} \phi_n(u) = \exp\left(\frac{u^2 t}{2}\right) = \phi(u).$$

By uniqess of the MGF, we have that,

$$W^{(n)}(t) \xrightarrow{d} \mathcal{N}(0,t).$$

### 1.2.3 Limit of the Binomial Model

**Theorem 1.5.** *Fix $t \geq 0$. Under the binomial model with up factor*

$$u_n = 1 + \frac{\sigma}{\sqrt{n}}, \qquad d_n = 1 - \frac{\sigma}{\sqrt{n}},$$

*and risk–neutral probabilities $\tilde{p} = \tilde{q} = \frac{1}{2}$, define*

$$S_n(t) = S(0)\, u_n^{H_{nt}} d_n^{T_{nt}}, \qquad H_{nt} + T_{nt} = nt.$$

*Then, as $n \to \infty$,*

$$\log S_n(t) \xrightarrow{d} \log S(0) + \sigma W(t) - \frac{1}{2}\sigma^2 t.$$

*Consequently,*

$$S_n(t) \xrightarrow{d} S(0) \exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\},$$

*and the limiting random variable is log–normal.*

**Proof.** Let the number of heads and the number of tails be $H_{nt}$ and $T_{nt}$ respectively, such that $H_{nt} + T_{nt} = nt$ so that $M_{nt} = H_{nt} - T_{nt}$ is a symmetric random walk. Solving, we get that $H_{nt} = \frac{1}{2}(nt + M_{nt})$ and $T_{nt} = \frac{1}{2}(nt - M_{nt})$.

Then the stock price can be written as,

$$S_n(t) = S(0) u_n^{H_{nt}} d_n^{T_{nt}} = S(0)\left(1 + \frac{\sigma}{\sqrt{n}}\right)^{\frac{1}{2}(nt + M_{nt})} \left(1 - \frac{\sigma}{\sqrt{n}}\right)^{\frac{1}{2}(nt - M_{nt})}.$$

Then,

$$\log S_n(t) = \log S(0) + \frac{1}{2}(nt + M_{nt}) \log\left(1 + \frac{\sigma}{\sqrt{n}}\right) + \frac{1}{2}(nt - M_{nt}) \log\left(1 - \frac{\sigma}{\sqrt{n}}\right).$$

Recall that for $x \to 0$ we have $\log(1+x) = x - \frac{1}{2}x^2 + O(x^3)$ which means that,

$$\log\left(1 \pm \frac{\sigma}{\sqrt{n}}\right) = \pm \frac{\sigma}{\sqrt{n}} - \frac{1}{2}\frac{\sigma^2}{n} + O\left(n^{-3/2}\right).$$

Then, we can directly derive $\log S_n(t)$.

$$\begin{aligned}
\log S_n(t) &= \log S(0) + \frac{1}{2}(nt + M_{nt})\left(\frac{\sigma}{\sqrt{n}} - \frac{1}{2}\frac{\sigma^2}{n} + O\left(n^{-3/2}\right)\right) + \frac{1}{2}(nt - M_{nt})\left(-\frac{\sigma}{\sqrt{n}} - \frac{1}{2}\frac{\sigma^2}{n} + O\left(n^{-3/2}\right)\right) \\
&= \log S(0) + \frac{1}{2}\left[\frac{\sigma nt}{\sqrt{n}} + \frac{\sigma M_{nt}}{\sqrt{n}}\right] - \frac{1}{4}\frac{\sigma^2}{n}\left[(nt + M_{nt}) + (nt - M_{nt})\right] + nt\, O\left(n^{-3/2}\right) + M_{nt}\, O\left(n^{-3/2}\right) \\
&= \log S(0) + \frac{\sigma}{\sqrt{n}} M_{nt} - \frac{1}{4}\frac{\sigma^2}{n}(2nt) + nt\, O\left(n^{-3/2}\right) + M_{nt}\, O\left(n^{-3/2}\right) \\
&= \log S(0) + \frac{\sigma}{\sqrt{n}} M_{nt} - \frac{1}{2}\sigma^2 t + R_n.
\end{aligned}$$

Then, $\log S_n(t) \to \log S(0) - \dfrac{1}{2}\sigma^2 t + \sigma W(t)$ since $W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt} \xrightarrow{d} W(t) \sim N(0,t)$ where $\frac{M_{nt}}{\sqrt{n}} = W^{(n)}(t) \xrightarrow{d} W(t)$. Thus, we conclude that,

$$S_n(t) \xrightarrow{d} S(0) \exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\},$$

which matches the continuous-time Geometric Brownian Motion as required.

## 1.3 Brownian Motion

A useful way to motivate Brownian motion is via a scaled random walk. Let $W^{(n)}(t)$ denote the rescaled random walk obtained from i.i.d. mean-zero, variance-one increments. As $n \to \infty$, the process $W^{(n)}$ converges in distribution to a continuous-time stochastic process known as *Brownian motion.* Formally, a *Brownian motion* $\{W(t)\}_{t \geq 0}$ is a stochastic process with specified distributional properties, which we state below.

### 1.3.1 Properties of Brownian Motion

**Theorem 1.6.** ***Properties of Brownian Motion***

*(i)* *For* $t = 0$, *we have* $W(0) = 0$.

*(ii)* *On the grid* $0 = t_0 < t_1 < \cdots < t_m$, *Brownian Motion increments are independent,*

$$W(t_1) = W(t_2) - W(t_1) \perp W(t_2) - W(t_1) \perp \cdots \perp W(t_m) - W(t_{m-1}).$$

*(iii)* $\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0$ *and* $\operatorname{Var}[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i$.

*(iv)* *The increments* $\Delta W_i := W(t_{i+1}) - W(t_i) \sim \mathcal{N}(0, t_{i+1} - t_i)$ *i.e., normally distributed.*

*(v)* *Fix* $0 \leq s < t$, *then* $\mathbb{E}[W(s)W(t)] = s$ *and* $\operatorname{Cov}(W(s), W(t)) = s$.

**Proof.** Consider Brownian Motion at time $t$, denoted $W(t)$. Then, we split the increment,

$$\begin{aligned} W(t) &= W(s) + (W(t) - W(s)) \\ W(s) &= W(s)\left[W(s) + (W(t) - W(s))\right] \\ &= W(s)^2 + W(s)\left(W(t) - W(s)\right) \\ \implies \mathbb{E}\left[W(s)W(s)\right] &= \mathbb{E}[W(s)^2] + \mathbb{E}\left[W(s)\left(W(t) - W(s)\right)\right]. \end{aligned}$$

$W(s)$ is clearly $\mathcal{F}_s$-measurable, so we conclude that $W(s) \perp W(t) - W(s)$. Naturally,

$$\mathbb{E}\left[W(s)\left(W(t) - W(s)\right)\right] = \mathbb{E}[W(s)] \cdot \mathbb{E}[W(t) - W(s)] = 0 \cdot 0 = 0,$$

since $\mathbb{E}[W(s)] = 0$ and $\mathbb{E}[W(t) - W(s)] = 0$. Moreover, note that

$$\mathbb{E}[W(s)^2] = \operatorname{Var}(W(s)) + \left(\mathbb{E}[W(s)]\right)^2 = s + 0 = s \implies \mathbb{E}[W(s)W(t)] = \mathbb{E}[W(s)^2] + 0 = s.$$

So we conclude that,

$$\operatorname{Cov}\left(W(s)W(t)\right) = \mathbb{E}[W(s)W(t)] - \mathbb{E}[W(s)]\mathbb{E}[W(t)] = s - 0 \cdot 0 = s.$$

The increments of Brownian Motion are independently and normally distributed. Hence, we can construct a $m \times m$ covariance matrix for the $m$-dimensional random vector. We can use this to derive the moment-generating function of this random vector to further characterize the joint distribution of the random variables $W(t_1), W(t_2), \ldots, W(t_m)$.

### 1.3.2 Moment-Generating Function of Brownian Motion Vector

**Theorem 1.7.** *The Joint Moment-Generating Function of Brownian Motion vector* $\{W(t_1), \ldots, W(t_m)\}$, *for fixed* $0 = t_0 < t_1 < \cdots < t_m$ *and* $u_1, \ldots, u_m \in \mathbb{R}$ *is given by,*

$$\phi(u_1, \ldots, u_m) = \mathbb{E}\left[\exp\left\{\sum_{k=1}^{m} u_m W(t_k)\right\}\right] = \exp\left\{\frac{1}{2}\sum_{j=1}^{m}\left(\sum_{k=j} u_k\right)^2 (t_j - t_{j-1})\right\}.$$

**Proof.** On the same grid, for increments $\Delta W_j := W(t_j) - W(t_{j-1})$ for $j = 1, \ldots, m$, we have,

$$W(t_k) = W(t_0) + \sum_{j=1}^{k} \Delta W_j = \sum_{j=1}^{k} \Delta W_j$$

since $W(t_0) = W(0) = 0$. So, re-arranging the sum, we have that,

$$\sum_{k=1}^{m}\sum_{j=1}^{k} u_k \Delta W_j = \sum_{j=1}^{m}\left(\sum_{k=j}^{m} u_k\right) \cdot \Delta W_j.$$

Letting $a_j := \sum_{k=j}^{m} u_k$, we directly derive our MGF,

$$\begin{aligned}
\varphi(u_1, \ldots, u_m) &= \mathbb{E}\left[\exp\left\{\sum_{k=1}^{m} u_k W(t_k)\right\}\right] = \mathbb{E}\left[\exp\left\{\sum_{k=1}^{m} u_k \sum_{j=1}^{k} \Delta W_j\right\}\right] \\
&= \mathbb{E}\left[\exp\left\{\sum_{j=1}^{m}\left(\sum_{k=j}^{m} u_k\right)\Delta W_j\right\}\right] \\
&= \mathbb{E}\left[\exp\left\{\sum_{j=1}^{m} a_j \Delta W_j\right\}\right] \\
&= \prod_{j=1}^{m} \mathbb{E}\left[\exp\{a_j \Delta W_j\}\right] \qquad \text{(independent BM increments)} \\
&= \prod_{j=1}^{m} \exp\left\{\frac{1}{2} a_j^2 (t_j - t_{j-1})\right\} \qquad \text{since } \Delta W_j \sim N(0, t_j - t_{j-1}) \\
&= \exp\left\{\frac{1}{2}\sum_{j=1}^{m} a_j^2 (t_j - t_{j-1})\right\} \\
&= \exp\left\{\frac{1}{2}\sum_{j=1}^{m}\left(\sum_{k=j}^{m} u_k\right)^2 (t_j - t_{j-1})\right\}.
\end{aligned}$$

*Remark* 1.4. For $X \sim N(\mu, \sigma^2)$ we recall the mgf.

$$\begin{aligned}
\phi_X(t) &= \mathbb{E}[e^{tX}] = \exp\left\{\mu t + \tfrac{1}{2}\sigma^2 t^2\right\} \\
\phi_X'(t) &= (\mu + \sigma^2 t)\exp\left\{\mu t + \tfrac{1}{2}\sigma^2 t^2\right\} \implies \phi_X'(0) = \mu \\
\phi_X''(t) &= \sigma^2 \exp\left\{\mu t + \tfrac{1}{2}\sigma^2 t^2\right\} + (\mu + \sigma^2 t)^2 \exp\left\{\mu t + \tfrac{1}{2}\sigma^2 t^2\right\} \implies \phi_X''(0) = \mu^2 + \sigma^2
\end{aligned}$$

$$\therefore \operatorname{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (\mu^2 + \sigma^2) - \mu^2 = \sigma^2$$

### 1.3.3 Martingale Property of Brownian Motion

**Proposition 1.2.** *If $\Delta(t)$ is a $\mathcal{F}(t)$-measurable for all $t \geq 0$, we say that $\Delta(t)$ is an adapted stochastic process.*

**Proof.** Let $W(t)$ for $t \geq 0$ be Brownian Motion. Then for $0 \leq s \leq t$, with $W(t) = W(t) - W(s) + W(t)$, we can write

$$\begin{aligned}
\mathbb{E}[W(t) \mid \mathcal{F}(s)] &= \mathbb{E}[W(s) + (W(t) - W(s)) \mid \mathcal{F}(s)] \\
&= \mathbb{E}[W(t) - W(s) \mid \mathcal{F}(s)] + \mathbb{E}[W(s) \mid \mathcal{F}(s)] \\
&= \mathbb{E}[W(t) - W(s) \mid \mathcal{F}(s)] + W(s) \\
&= \mathbb{E}[W(t) - W(s)] + W(s) \\
&= 0 + W(s) \\
&= W(s).
\end{aligned}$$

### 1.3.4 First Order Variation

For a function $f(t)$ that varies with time $t$, we quantify the amount of up and down movement undergone by $f$ via *First-Order Variation* ($FV_T(f)$). Suppose we partition our time into 3 intervals (need not be equally spaced) $[0, t_1), [t_1, t_2), [t_2, T]$, then we could characterize the first-order variation of $f$ as follows,

$$\begin{aligned}
\mathrm{FV}_T(f) &= \big[f(t_1) - f(0)\big] - \big[f(t_2) - f(t_1)\big] + \big[f(T) - f(t_2)\big] \\
&= \int_0^{t_1} f'(t)\,dt + \int_{t_1}^{t_2} \big(-f'(t)\big)\,dt + \int_{t_2}^{T} f'(t)\,dt \\
&= \int_0^{T} |f'(t)|\,dt. \qquad (1.3.4.1)
\end{aligned}$$

Generally, we compute $FV_T(f)$ over a partition (set of times) $0 = t_0 < t_1 < \cdots < t_n = T$.
Formally, *First-Order Variation* can be defined as,

$$FV_T(f) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)|, \qquad (1.3.4.2)$$

i.e., the limit as the number of partition points goes to $\infty$ and the length of the longest subinterval $t_{j+1} - t_j \to 0$. It is easy to see that (1.3.4.1) and (1.3.4.2) are consistent. Take any subinterval $t_{j+1} - t_j$. Then, by MVT, we see that there exists $t_j^* \in (t_j, t_{j+1})$ such that the slope of the tangent at $t_j^*$ is parallel to the chord that connects $t_{j+1}$ to $t_j$. So,

$$f'(t_j^*) = \frac{f(t_{j+1}) - f(t_j)}{t_{j+1} - t_j} \implies f(t_{j+1}) - f(t_j) = f'(t_j^*)(t_{j+1} - t_j).$$

which is equivalent to the identity under the summation in 1.3.4.2. Thus, we can substitute,

$$FV_T(f) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)| = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} |f'(t_j^*)|(t_{j+1} - t_j) = \int_0^T |f'(t)|dt, \qquad (1.3.4.3)$$

which is exactly 1.3.4.1.

### 1.3.5 Quadratic Variation

Suppose we have a function $f(t)$ defined for $0 \leq t \leq T$. The second-order (quadratic) variation is defined as,

$$[f, f](T) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2,$$

for time partition $\Pi = \{t_0, \dots, t_n\}$ and $0 = t_0 < t_1 < \cdots < t_n = T$.

Suppose for a second that $f$ has a continuous first derivative, i.e.,

$$\sum_{j=0}^{n-1}\big[f(t_{j+1})-f(t_j)\big]^2 \;\le\; \sum_{j=0}^{n-1}|f'(t_j^*)|^2(t_{j+1}-t_j)^2 \;\le\; \|f'\|_\infty\sum_{j=0}^{n-1}|f'(t_j^*)|^2(t_{j+1}-t_j).$$

Naturally, we can substitute to get the quadratic variation of $f$ as,

$$\begin{aligned}
[f,f](T) &\le \lim_{\|\Pi\|\to 0}\left[\|\Pi\|\cdot\sum_{j=0}^{n-1}|f'(t_j^*)|^2(t_{j+1}-t_j)\right]\\
&= \lim_{\|\Pi\|\to 0}\|\Pi\|\cdot\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}|f'(t_j^*)|^2(t_{j+1}-t_j)\\
&= \lim_{\|\Pi\|\to 0}\|\Pi\|\cdot\int_0^T|f'(t)|^2\,dt = 0,
\end{aligned}$$

assuming that $\int_0^T|f'(t)|^2dt$ is finite. Otherwise, $\le \lim_{\|\Pi\|\to 0}\left[\|\Pi\|\cdot\sum_{j=0}^{n-1}|f'(t_j^*)|^2(t_{j+1}-t_j)\right]$ could be anything between $0$ and $\infty$. Quadratic variation is not used for functions with continuous derivatives as their quadratic variation is simply $0$. However, Brownian Motion is undifferentiable everywhere meaning that it cannot be differentiated wrt $t$ therefore any application of MVT would fail i.e., there is no value for $t$ for which $\frac{d}{dt}W(t)$ is defined.

We now consider the quadratic variation of the Brownian Motion which cannot be computed directly. Thus we consider the *Sampled Quadratic Variation*; a random variable, dependent on the BM path from which it is computed. Taking the sampled QV, we can show that it converges to $T$ as $\|\Pi\|\to 0$. Regardless of what path it is computed along, we must show that it converges to $T$, its expected value. We define the sampled quadratic variation as $Q_\Pi = \sum_{j=0}^{n-1}(W(t_{j+1})-W(t_j))^2$. Note because $W(t_{j+1})-W(t_j)\sim\mathcal{N}(0,\Delta t_j)$ then $W(t_{j+1})-W(t_j)=\sqrt{\Delta t_j}Z_j$ where $Z_j\sim\mathcal{N}(0,1)$. Thus, $Q_\Pi$ is the weighted average of $\chi_1^2$ variables where the weights sum to $T$ and the largest weight $\|\Pi\| := \max_j\Delta t_j$ goes to 0,

$$\begin{aligned}
\mathbb{E}[Q_\Pi] &= \sum_{j=0}^{n-1}\mathbb{E}\big[(W(t_{j+1})-W(t_j))^2\big] = \sum_{j=0}^{n-1}\Delta t_j = T,\\
\operatorname{Var}(Q_\Pi) &= \sum_{j=0}^{n-1}\operatorname{Var}\big((W(t_{j+1})-W(t_j))^2\big) = \sum_{j=0}^{n-1}\operatorname{Var}(\Delta t_j Z_j^2) = \sum_{j=0}^{n-1}(\Delta t_j)^2\operatorname{Var}(Z_j^2).
\end{aligned}$$

Since $Z\sim\mathcal{N}(0,1)$ then $\mathbb{E}[Z^2]=1,\mathbb{E}[Z^4]=3$ so $\operatorname{Var}(Z^2)=3-1=2$. Thus,

$$\operatorname{Var}(Q_\Pi) = 2\sum_{j=0}^{n-1}(\Delta t_j)^2 \le 2\Big(\max_j\Delta t_j\Big)\sum_{j=0}^{n-1}\Delta t_j = 2\|\Pi\|\,T \xrightarrow[\|\Pi\|\to 0]{} 0.$$

By Chebyshev Inequality, we have,

$$\mathbb{P}\left(|Q_\Pi - T| > \varepsilon\right) \le \frac{\operatorname{Var}(Q_\Pi)}{\varepsilon^2}\to 0,$$

so $Q_\Pi\to T$ in probability as $\|\Pi\|\to 0$. Thus $\lim_{\|\Pi\|\to 0}Q_\Pi = \mathbb{E}Q_\Pi = T$.

**Theorem 1.8.** *Let $W$ be a Brownian Motion. Then $[W,W](T)=T$ for all $T\ge 0$ almost surely.*

*Remark* 1.5. We can informally write,

$$dW(t)dW(t) = dt.$$

On an interval $[0,T]$, Brownian Motion accumulates $T$ units of quadratic variation.
Over distinct time interval $0<T_1<T_2$ we get the limit,

$$[W,W](T_2)-[W,W](T_1) = T_2 - T_1$$

so the BM accumulates $T_2-T_1$ units of quadratic variation over $[T_1,T_2]$.

### 1.3.6 Cross-Variation and Quadratic Variation with $t$

It is also easy to show the cross-variation of BM with $t$ is $[W,t]_T = 0$.

$$
\begin{aligned}
[W,t](T) &:= \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \big(W(t_{j+1}) - W(t_j)\big)\big(t_{j+1} - t_j\big) \\
&\le \lim_{\|\Pi\|\to 0} \left( \max_{0\le k\le n-1} \big|W(t_{k+1}) - W(t_k)\big| \right) \sum_{j=0}^{n-1} (t_{j+1} - t_j) \\
&= \lim_{\|\Pi\|\to 0} \left( \max_{0\le k\le n-1} \big|W(t_{k+1}) - W(t_k)\big| \right) T \\
&= 0.
\end{aligned}
$$

Concisely, we can write $dW(t)dt = 0$. Similarly, it is easy to show QV wrt $t$ is also just 0,

$$
\begin{aligned}
[t,t](T) &:= \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \big(t_{j+1} - t_j\big)^2 \\
&\le \lim_{\|\Pi\|\to 0} \left( \max_{0\le k\le n-1} (t_{k+1} - t_k) \right) \sum_{j=0}^{n-1} (t_{j+1} - t_j) \\
&= \lim_{\|\Pi\|\to 0} \|\Pi\| \cdot T \\
&= 0.
\end{aligned}
$$

We get two nice shorthand notations,

$$dW(t)dt = 0, \quad dtdt = 0.$$

## 1.4 Volatility of Geometric Brownian Motion

Under the Black-Scholes-Merton option pricing model, the asset-price model used is,

$$S(t) = S(0)\exp\left\{ \sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right) t \right\}.$$

We define the above as the *Geometric Brownian Motion.* We can derive the "realized volatility" estimator using the Quadratic Variation of Brownian Motion to identify volatility $\sigma$ from a path of this process. Suppose we let $\alpha \in \mathbb{R}$ and $\sigma > 0$ be constants. Moreover, we fix $0 \le T_1 < T_2$ and take the partition,

$$\Pi = \{T_1 = t_0 < t_1 < \cdots < t_m = T_2\}, \quad \Delta t_j := t_{j+1} - t_j, \quad \Delta W_j := W(t_{j+1}) - W(t_j),$$

and define the log-return of our stock-price process on $[t_j, t_{j+1}]$ by,

$$R_j := \log \frac{S(t_{j+1})}{S(t_j)}.$$

Starting from the explicit formula for $S(t)$,

$$\log S(t) = \log S(0) + \sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right) t.$$

Subtracting $t_{j+1}$ and $t_j$ we get,

$$
\begin{aligned}
\log \frac{S(t_{j+1})}{S(t_j)} &= \log S(t_{j+1}) - \log S(t_j) \\
&= \big(\log S(0) + \sigma W(t_{j+1}) + (\alpha - \tfrac{1}{2}\sigma^2) t_{j+1}\big) - \big(\log S(0) + \sigma W(t_j) + (\alpha - \tfrac{1}{2}\sigma^2) t_j\big) \\
&= \sigma\big(W(t_{j+1}) - W(t_j)\big) + (\alpha - \tfrac{1}{2}\sigma^2)(t_{j+1} - t_j) \\
&= \sigma\,\Delta W_j + (\alpha - \tfrac{1}{2}\sigma^2)\,\Delta t_j.
\end{aligned}
$$

So, $R_j = \sigma\Delta W_j + \mu\Delta t_j$ where $\mu := \alpha - \frac{1}{2}\sigma^2$. Note that realized volatility is mathematically just the square of log returns. Naturally, we need $R_j^2$.

$$\begin{aligned}R_j^2 &= (\sigma\Delta W_j + \mu\Delta t_j)^2 \\ &= (\sigma\Delta W_j)^2 + (\mu\Delta t_j)^2 + 2(\sigma\Delta W_j)(\mu\Delta t_j) \\ &= \sigma^2(\Delta W_j)^2 + \mu^2(\Delta t_j)^2 + 2\sigma\mu(\Delta W_j)(\Delta t_j),\end{aligned}$$

and taking the squared sum,

$$\begin{aligned}\sum_{j=0}^{m-1} R_j^2 &= \sum_{j=0}^{m-1}\Big[\sigma^2(\Delta W_j)^2 + \mu^2(\Delta t_j)^2 + 2\sigma\mu(\Delta W_j)(\Delta t_j)\Big] \\ &= \sigma^2\sum_{j=0}^{m-1}(\Delta W_j)^2 + \mu^2\sum_{j=0}^{m-1}(\Delta t_j)^2 + 2\sigma\mu\sum_{j=0}^{m-1}(\Delta W_j)(\Delta t_j).\end{aligned}$$

Lastly, taking the squared log-returns, we get our final realized volatility estimator,

$$\begin{aligned}\sum_{j=0}^{m-1}\left(\log\frac{S(t_{j+1})}{S(t_j)}\right)^2 &= \sigma^2\sum_{j=0}^{m-1}\big(W(t_{j+1}) - W(t_j)\big)^2 \\ &\quad + \left(\alpha - \tfrac{1}{2}\sigma^2\right)^2\sum_{j=0}^{m-1}(t_{j+1} - t_j)^2 \\ &\quad + 2\sigma\left(\alpha - \tfrac{1}{2}\sigma^2\right)\sum_{j=0}^{m-1}\big(W(t_{j+1}) - W(t_j)\big)(t_{j+1} - t_j).\end{aligned}$$

Suppose we have the mesh $\|\Pi\| := \max_{0\le j\le m-1}\Delta t_j$. Then we have,

(i) **Quadratic variation of Brownian motion.**

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{m-1}(\Delta W_j)^2 = T_2 - T_1.$$

(ii) **Quadratic variation of $t$ is zero.** We have the deterministic bound (pure algebra):

$$(\Delta t_j)^2 \le \|\Pi\|\,\Delta t_j \;\Rightarrow\; \sum_{j=0}^{m-1}(\Delta t_j)^2 \le \|\Pi\|\sum_{j=0}^{m-1}\Delta t_j = \|\Pi\|(T_2 - T_1),$$

so

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{m-1}(\Delta t_j)^2 = 0.$$

(iii) **Cross-variation of $W$ with $t$ is zero.** Define $M_\Pi := \max_{0\le j\le m-1}|\Delta W_j|$. Then

$$|\Delta W_j\Delta t_j| \le M_\Pi\Delta t_j,$$

and by the triangle inequality,

$$\left|\sum_{j=0}^{m-1}\Delta W_j\Delta t_j\right| \le \sum_{j=0}^{m-1}|\Delta W_j\Delta t_j| \le M_\Pi\sum_{j=0}^{m-1}\Delta t_j = M_\Pi(T_2 - T_1).$$

Because Brownian paths are continuous, $M_\Pi \to 0$ as $\|\Pi\| \to 0$, hence

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{m-1}\Delta W_j\Delta t_j = 0.$$

Using the three limits above to re-write $\sum_{j=0}^{m-1} R_j^2$. Starting from,

$$\sum_{j=0}^{m-1} R_j^2 = \sigma^2 \sum_{j=0}^{m-1} (\Delta W_j)^2 + \mu^2 \sum_{j=0}^{m-1} (\Delta t_j)^2 + 2\sigma\mu \sum_{j=0}^{m-1} \Delta W_j \Delta t_j.$$

Taking $\|\Pi\| \to 0$ we can use the three limits to obtain the following over a sufficiently fine partition,

$$\begin{aligned}\lim_{\|\Pi\|\to 0} \sum_{j=0}^{m-1} R_j^2 &= \sigma^2(T_2 - T_1) + \mu^2 \cdot 0 + 2\sigma\mu \cdot 0 \\ &= \sigma^2(T_2 - T_1)\end{aligned}$$

$$\implies \sum_{j=0}^{m-1} \left(\log \frac{S(t_{j+1})}{S(t_j)}\right)^2 \approx \sigma^2(T_2 - T_1) \implies \frac{1}{T_2 - T_1} \sum_{j=0}^{m-1} \left(\log \frac{S(t_{j+1})}{S(t_j)}\right)^2 \approx \sigma^2$$

Therefore, our realized volatility estimator is,

$$\hat{\sigma}_\Pi^2 := \frac{1}{T_2 - T_1} \sum_{j=0}^{m-1} \left(\log \frac{S(t_{j+1})}{S(t_j)}\right)^2, \quad \hat{\sigma}_\Pi := \sqrt{\hat{\sigma}_\Pi^2}.$$

### 1.5 Markov Property

**Theorem 1.9.** *Let $W(t)$ for $t \geq 0$ be Brownian Motion and let $\mathcal{F}(t)$ be the corresponding filtration for the Brownian Motion. Then, $W(t)$ for $t \geq 0$ is a Markov Process.*

**Proof.** This is a pretty standard proof. If you've taken STAT 333, we can use the the trick of splitting $W(t)$ into a past + fresh increment. Thus we get that,

$$\begin{aligned} W(t) &= W(s) + (W(t) - W(s)) \\ f(W(t)) &= f\left((W(t) - W(s)) + W(s)\right) \\ \mathbb{E}[f(W(t)) \mid \mathcal{F}(s)] &= \mathbb{E}[f\left((W(t) - W(s)) + W(s)\right) \mid \mathcal{F}(s)]. \end{aligned}$$

Now we set $Z := W(t) - W(s)$ and $X := W(s)$. Recall that $Z \sim N(0, t-s)$ and $Z \perp \mathcal{F}(s)$ with $X$ being $\mathcal{F}(s)$-measurable. Thus, the conditional expectation is of the form,

$$\mathbb{E}[\phi(Z, X) \mid \mathcal{F}(s)], \qquad Z \perp \mathcal{F}(s), \quad X \in \mathcal{F}(s)$$

For each $x \in \mathbb{R}$, $g(x) := \mathbb{E}[f(Z + x)]$. This is an ordinary (unconditional) expectation over $Z$ only. Then, the key independence lemma says that $\mathbb{E}[f(Z + X) \mid \mathcal{F}(s)] = g(X)$. Applying $X = W(s)$ gives,

$$\mathbb{E}[f(W(t)) \mid \mathcal{F}(s)] = \mathbb{E}[f(Z + W(s)) \mid \mathcal{F}(s)] = g(W(s)).$$

So, this proves the Markov property. What remains is computing $g$ explicitly. Since $Z = \mathcal{N}(0, t-s)$, then it has density,

$$\phi_{t-s}(w) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{w^2}{2(t-s)}\right).$$

Therefore,

$$\begin{aligned} g(x) = \mathbb{E}[f(Z + x)] &= \int_{-\infty}^{\infty} f(w + x)\, \varphi_{t-s}(w)\, dw \\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} f(w + x)\, \exp\left(-\frac{w^2}{2(t-s)}\right) dw. \end{aligned}$$

After a quick change-of-variables, letting $\tau := t - s > 0, y := w + x, w = y - x, dw = dy$

$$\begin{aligned} g(x) &= \frac{1}{\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} f(w+x) \exp\left(-\frac{w^2}{2\tau}\right) dw \\ &= \frac{1}{\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} f(y) \exp\left(-\frac{(y-x)^2}{2\tau}\right) dy, \end{aligned}$$

and then we get,

$$p(\tau, x, y) := \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{(y-x)^2}{2\tau}\right), \implies g(x) = \int_{-\infty}^{\infty} f(y)\, p(\tau, x, y)\, dy.$$

Subsituting, we get

$$\begin{aligned} \mathbb{E}[f(W(t)) \mid \mathcal{F}(s)] &= g(W(s)) \\ &= \int_{-\infty}^{\infty} f(y)\, p(\tau, W(s), y)\, dy \\ &= \int_{-\infty}^{\infty} f(y)\, p(t-s, W(s), y)\, dy. \end{aligned}$$

Thus, conditioning on everything you seen up to time $s$, the distribution of $W(t)$ depends on the past only through $W(s)$ and is Gaussian i.e., $W(t) \mid \mathcal{F}(s) \sim \mathcal{N}(W(s), t-s)$.

### 1.6 First Passage Time

#### 1.6.1 Exponential Martingale

The Brownian Motion is the continuous-time counterpart to the symmetric random walk thus, we introduce an *Exponential Martingale* containing Brownian Motion in the exponential function.

**Theorem 1.10.** *For $W(t)$, $t \geq 0$ Brownian Motion with associated filtration $\mathcal{F}(t)$ with constant $\sigma$, the process $Z(t)$ for $t \geq 0$ is a martingale. More specifically, we call it an Exponential Martingale,*

$$Z(t) = \exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\}.$$

**Proof.** For the exponential process $Z(t)$, in order for it to be a martingale, we know that,

$$\mathbb{E}[Z(t) \mid \mathcal{F}(s)] = Z(s)$$

must hold. This is a pretty easy proof. Let $0 \leq s \leq t$, so thus we have,

$$\begin{aligned} \mathbb{E}[Z(t) \mid \mathcal{F}(s)] &= \mathbb{E}\left[\exp\left\{\sigma W(t) - \tfrac{1}{2}\sigma^2 t\right\} \,\middle|\, \mathcal{F}(s)\right] \\ &= \mathbb{E}\left[\underbrace{\exp\left\{\sigma\left(W(t) - W(s)\right)\right\}}_{W(t)-W(s)\sim\mathcal{N}(0,t-s)} \cdot \underbrace{\exp\left\{\sigma W(s) - \tfrac{1}{2}\sigma^2 t\right\}}_{\mathcal{F}(s)\text{-measurable}} \,\middle|\, \mathcal{F}(s)\right] \\ &= \exp\left\{\sigma W(s) - \tfrac{1}{2}\sigma^2 t\right\} \cdot \mathbb{E}\left[\underbrace{\exp\left\{\sigma\left(W(t) - W(s)\right)\right\}}_{\substack{\text{independent of } \mathcal{F}(s) \\ \mathbb{E}[e^{\sigma X}] = e^{\frac{1}{2}\sigma^2(t-s)}}}\right] \\ &= \exp\left\{\sigma W(s) - \tfrac{1}{2}\sigma^2 t\right\} \cdot \exp\left\{\tfrac{1}{2}\sigma^2 (t-s)\right\} \\ &= \exp\left\{\sigma W(s) - \tfrac{1}{2}\sigma^2 s\right\} \\ &= Z(s). \end{aligned}$$

### 1.6.2 First Level $m$ Hitting Time

We wish to model the first passage time (hitting time) of level $m$, denoted $\tau_m = \inf\{t \geq 0 : W(t) = m\}$. Note if level $m$ is never reached, then $\tau_m = \infty$. We define the stopped time as $t \wedge \tau_m := \min\{t, \tau_m\}$. $Z(t \wedge \tau_m)$ is used instead of $Z(\tau_m)$ since $\tau_m$ might be $\infty$. $t \wedge \tau_m$ is a bounded stopping time, thus by *Optional Stopping Theorem* we have,

$$\mathbb{E}[Z(t \wedge \tau_m)] = \mathbb{E}[Z(0)] = 1 = \mathbb{E}\left[\exp\left\{\sigma W(t \wedge \tau_m) - \frac{1}{2}\sigma^2(t \wedge \tau_m)\right\}\right].$$

Intuitively, before Brownian motion hits level $m$, it must be $\leq m$. Thus, given the event,

(i) $\{t < \tau_m\}$, we have not yet hit $m$ by time $t$ so $W(t) \neq m$ and thus $W(t) < m$ by Continuity.

(ii) $\{t \geq \tau_m\}$ we have $t \wedge \tau_m = \tau_m$ and $W(\tau_m) = m$ by definition of $\tau_m$.

So in both cases we observe that $W(t \wedge \tau_m) \leq m$ so,

$$0 \leq \exp\{\sigma W(t \wedge \tau_m)\} \leq \exp(\sigma m).$$

So, now we have a bound for the space component of our exponential martingale. Next, we would like to analyze the limit of the time-exponential factor, specifically,

$$\exp\left\{-\frac{1}{2}\sigma^2(t \wedge \tau_m)\right\}.$$

For this, we consider when $\tau_m < \infty$ and $\tau_m = \infty$. Notice that for when $\tau_m < \infty$, for sufficiently large $t \geq \tau_m$ we have that $t \wedge \tau_m = \tau_m$ so eventually,

$$\exp\left\{-\frac{1}{2}\sigma^2(t \wedge \tau_m)\right\} = \exp\left(-\frac{1}{2}\sigma^2\tau_m\right)$$

and thus the limit exists and equals $e^{-\frac{1}{2}\sigma^2\tau_m}$. For $\tau_m = \infty$ (i.e., the Brownian motion never hits level $m$), let $t \to \infty$. Then $t \wedge \tau_m = t$ for all $t$, so,

$$\exp\left(-\frac{1}{2}\sigma^2(t \wedge \tau_m)\right) = \exp\left(-\frac{1}{2}\sigma^2 t\right) \to 0.$$

The cases for $\tau_m < \infty$ and $\tau_m = \infty$ can be written as,

$$\lim_{t\to\infty} \exp\left(-\frac{1}{2}\sigma^2(t \wedge \tau_m)\right) = \mathbf{I}\{\tau_m < \infty\} \exp\left(-\frac{1}{2}\sigma^2\tau_m\right).$$

A similar process can be applied for the limit of the whole integrand,

$$\exp\left(\sigma W(t \wedge \tau_m) - \frac{1}{2}\sigma^2(t \wedge \tau_m)\right) = \exp\left(\sigma W(t \wedge \tau_m)\right) \cdot \exp\left(-\frac{1}{2}\sigma^2(t \wedge \tau_m)\right).$$

Considering when $\tau_m < \infty$ and $\tau_m = \infty$ we achieve a well-defined bound,

$$0 \leq \exp(\sigma W(t)) \exp\left(-\frac{1}{2}\sigma^2 t\right) \leq e^{\sigma m} e^{-\frac{1}{2}\sigma^2 t} \to 0$$

as $t \to \infty$. This is obvious as for when $\tau_m = \infty$ we have that $t \wedge \tau_m = t$ so it converges to 0 almost surely. It should be pretty obvious that we can utilize the product structure of the "plus" bounds. Note that we at least know that the time term tends to 0 and the space term is bounded above $e^{\sigma m}$ since $W(t) \leq m$ for all $t$ if $\tau_m = \infty$ and $m > 0$, so naturally the path never crosses $m$. Necessarily, this implies that the product does go to 0 on that event. Combining gives us the limit,

$$\lim_{t\to\infty} \exp\left(\sigma W(t \wedge \tau_m) - \frac{1}{2}\sigma^2(t \wedge \tau_m)\right) = \mathbf{I}\{\tau_m < \infty\} \exp\left(\sigma m - \frac{1}{2}\sigma^2\tau_m\right).$$

We already know for every $t \geq 0$,

$$1 = \mathbb{E}\left[\exp\left(\sigma W(t \wedge \tau_m) - \tfrac{1}{2}\sigma^2(t \wedge \tau_m)\right)\right].$$

To let $t \to \infty$ inside the expectation, we use the Dominated Convergence Theorem. So we need an integrable dominating random variable. Using the bound (3.6.5) and the fact that the time term is always $\leq 1$ (since $-\frac{1}{2}\sigma^2(\cdot) \leq 0$),

$$0 \leq \exp\left(\sigma W(t \wedge \tau_m) - \tfrac{1}{2}\sigma^2(t \wedge \tau_m)\right) \leq \exp(\sigma m) \cdot 1 = e^{\sigma m}.$$

The constant $e^{\sigma m}$ is integrable. Hence dominated convergence applies, giving

$$1 = \lim_{t\to\infty} \mathbb{E}\big[\exp\big(\sigma W(t\wedge\tau_m) - \tfrac{1}{2}\sigma^2(t\wedge\tau_m)\big)\big] = \mathbb{E}\big[\mathbf{1}_{\{\tau_m<\infty\}} \exp\big(\sigma m - \tfrac{1}{2}\sigma^2\tau_m\big)\big] .$$

Now pull out the deterministic factor $e^{\sigma m}$:

$$1 = e^{\sigma m}\, \mathbb{E}\big[\mathbf{1}_{\{\tau_m<\infty\}} \exp\big(-\tfrac{1}{2}\sigma^2\tau_m\big)\big] .$$

Equivalently,

$$\mathbb{E}\big[\mathbf{1}_{\{\tau_m<\infty\}} \exp\big(-\tfrac{1}{2}\sigma^2\tau_m\big)\big] = e^{-\sigma m}.$$

Now send $\sigma \downarrow 0$. On the right, $e^{-\sigma m} \to 1$. On the left, for each outcome $\omega$:

- If $\tau_m(\omega) < \infty$, then $e^{-\frac{1}{2}\sigma^2\tau_m(\omega)} \to 1$.
- If $\tau_m(\omega) = \infty$, then the indicator is $0$ anyway, so the product is $0$ for all $\sigma$.

So pointwise,

$$\mathbf{1}_{\{\tau_m<\infty\}} e^{-\frac{1}{2}\sigma^2\tau_m} \xrightarrow{\sigma\downarrow 0} \mathbf{1}_{\{\tau_m<\infty\}}.$$

Also we have the uniform bound

$$0 \le \mathbf{1}_{\{\tau_m<\infty\}} e^{-\frac{1}{2}\sigma^2\tau_m} \le 1,$$

so dominated convergence applies again (now the parameter is $\sigma \downarrow 0$, but it's the same theorem). Therefore,

$$\lim_{\sigma\downarrow 0} \mathbb{E}\left[\mathbf{1}_{\{\tau_m<\infty\}} e^{-\frac{1}{2}\sigma^2\tau_m}\right] = \mathbb{E}\big[\mathbf{1}_{\{\tau_m<\infty\}}\big] = \mathbb{P}(\tau_m < \infty).$$

But the limit also equals $\lim_{\sigma\downarrow 0} e^{-\sigma m} = 1$. Hence

$$\mathbb{P}(\tau_m < \infty) = 1 \qquad (m > 0).$$

The theorem is stated in terms of $\alpha > 0$:

$$\mathbb{E}\big[e^{-\alpha\tau_m}\big] = e^{-|m|\sqrt{2\alpha}}.$$

First assume $m > 0$. Choose $\sigma = \sqrt{2\alpha}$ (positive). Then

$$\tfrac{1}{2}\sigma^2 = \tfrac{1}{2}(2\alpha) = \alpha,$$

so (3.6.8) becomes

$$\mathbb{E}\big[e^{-\alpha\tau_m}\big] = e^{-\sigma m} = e^{-m\sqrt{2\alpha}}, \qquad (m > 0,\ \alpha > 0).$$

Now handle $m < 0$. Use symmetry of Brownian motion: $(-W(t))_{t\ge 0}$ is again a standard Brownian motion. Let $a = -m > 0$. Then

$$\tau_m = \inf\{t : W(t) = m\} = \inf\{t : -W(t) = a\}.$$

But $-W$ has the same law as $W$, so the hitting time of $a > 0$ by $-W$ has the same distribution as the hitting time of $a$ by $W$. Thus $\tau_m \overset{d}{=} \tau_{|m|}$. Consequently,

$$\mathbb{E}\big[e^{-\alpha\tau_m}\big] = \mathbb{E}\big[e^{-\alpha\tau_{|m|}}\big] = e^{-|m|\sqrt{2\alpha}}.$$

From the theorem (for any $m \neq 0$),

$$\phi(\alpha) := \mathbb{E}\big[e^{-\alpha\tau_m}\big] = e^{-|m|\sqrt{2\alpha}}, \qquad \alpha > 0.$$

Differentiate both sides with respect to $\alpha$. On the right,

$$\frac{d}{d\alpha} e^{-|m|\sqrt{2\alpha}} = e^{-|m|\sqrt{2\alpha}} \cdot \frac{d}{d\alpha}\Big(-|m|\sqrt{2\alpha}\Big) .$$

Since $\sqrt{2\alpha} = (2\alpha)^{1/2}$, we have

$$\frac{d}{d\alpha}\sqrt{2\alpha} = \frac{1}{2}(2\alpha)^{-1/2} \cdot 2 = (2\alpha)^{-1/2} = \frac{1}{\sqrt{2\alpha}}.$$

Therefore,

$$\frac{d}{d\alpha} e^{-|m|\sqrt{2\alpha}} = e^{-|m|\sqrt{2\alpha}} \cdot \left(-|m| \cdot \frac{1}{\sqrt{2\alpha}}\right) = -\frac{|m|}{\sqrt{2\alpha}}\, e^{-|m|\sqrt{2\alpha}}.$$

On the left, differentiate under the expectation:

$$\phi'(\alpha) = \frac{d}{d\alpha}\mathbb{E}\left[e^{-\alpha\tau_m}\right] = \mathbb{E}\left[\frac{d}{d\alpha}e^{-\alpha\tau_m}\right] = \mathbb{E}\left[-\tau_m e^{-\alpha\tau_m}\right].$$

This interchange of differentiation and expectation is justified because for $\alpha > 0$ the random variable $\tau_m e^{-\alpha\tau_m}$ is integrable; the exponential factor kills the tail strongly enough. Hence,

$$-\mathbb{E}\left[\tau_m e^{-\alpha\tau_m}\right] = -\frac{|m|}{\sqrt{2\alpha}}\, e^{-|m|\sqrt{2\alpha}}.$$

Cancelling the minus signs yields

$$\mathbb{E}\left[\tau_m e^{-\alpha\tau_m}\right] = \frac{|m|}{\sqrt{2\alpha}}\, e^{-|m|\sqrt{2\alpha}}, \qquad \alpha > 0.$$

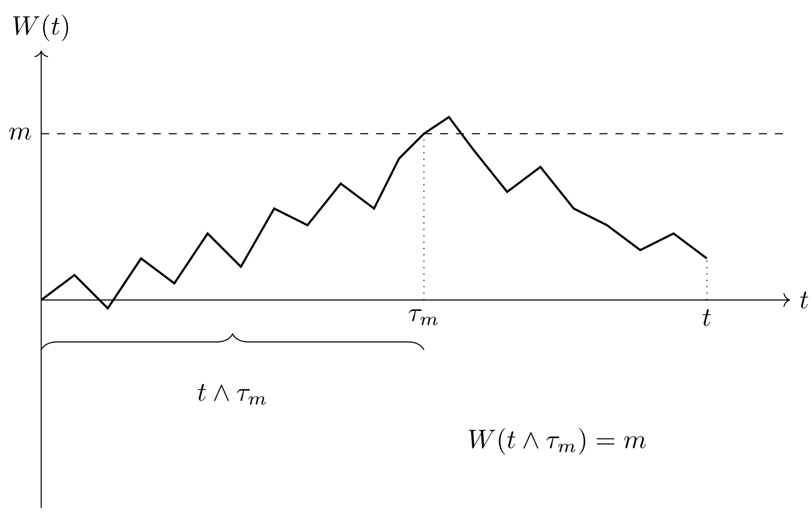
Finally, to see why $\mathbb{E}[\tau_m] = \infty$ for $m \neq 0$, note that as $\alpha \downarrow 0$,

$$\mathbb{E}\left[\tau_m e^{-\alpha\tau_m}\right] \xrightarrow[\alpha\downarrow 0]{} \mathbb{E}[\tau_m],$$

by the monotone convergence theorem, since $e^{-\alpha\tau_m} \uparrow 1$ pointwise as $\alpha \downarrow 0$ and $\tau_m e^{-\alpha\tau_m} \uparrow \tau_m$. But the explicit formula behaves like

$$\frac{|m|}{\sqrt{2\alpha}}\, e^{-|m|\sqrt{2\alpha}} \sim \frac{|m|}{\sqrt{2\alpha}} \qquad \text{as } \alpha \downarrow 0,$$

which diverges to $+\infty$. Hence $\mathbb{E}[\tau_m] = \infty$ for $m \neq 0$. (Intuitively: the hitting time is almost surely finite, but it has a heavy enough tail that its mean is infinite.)




### 1.7 Reflection Principle and First Passage Times

Let $(W(t))_{t\geq 0}$ be standard Brownian motion with $W(0) = 0$, and fix a level $m > 0$ and a time $t > 0$. Define the first passage time

$$\tau_m := \inf\{s \geq 0 : W(s) = m\}.$$

The reflection principle exploits the symmetry of Brownian motion after its first visit to the level $m$. If a path reaches level $m$ at some time $\tau_m \leq t$ and then ends at a value $w < m$ at time $t$, we may reflect the path about the horizontal line at height $m$ from time $\tau_m$ onward. The reflected path ends at $2m - w$ at time $t$. Because Brownian increments are symmetric and independent of the past, this reflection preserves probability. As a consequence, for all $w \leq m$,

$$\mathbb{P}\{\tau_m \leq t,\ W(t) \leq w\} = \mathbb{P}\{W(t) \geq 2m - w\}.$$

This identity is known as the *reflection equality*. Setting $w = m$ yields a particularly important consequence. Since $W(t) \geq m$ implies that the level $m$ was crossed prior to time $t$, we have

$$\{\tau_m \leq t,\ W(t) \geq m\} = \{W(t) \geq m\}.$$

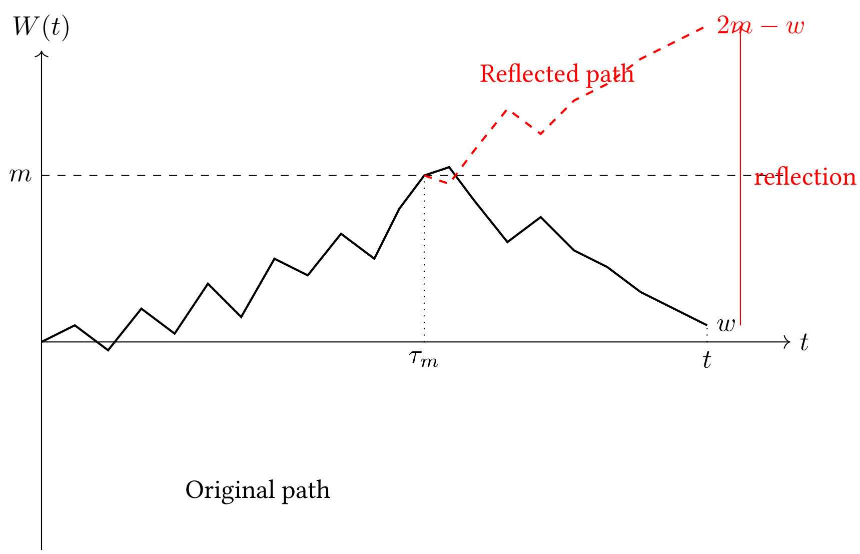
Splitting the event $\{\tau_m \le t\}$ according to whether $W(t) \le m$ or $W(t) \ge m$ therefore gives

$$\mathbb{P}\{\tau_m \le t\} = 2\,\mathbb{P}\{W(t) \ge m\}.$$

Because $W(t) \sim N(0,t)$, this leads to the explicit distribution of $\tau_m$, as shown below.

$$\begin{aligned}
\mathbb{P}\{\tau_m \le t\} &= 2\,\mathbb{P}\{W(t) \ge m\} \\
&= \frac{2}{\sqrt{2\pi t}} \int_m^\infty e^{-x^2/(2t)}\,dx \\
&= \frac{2}{\sqrt{2\pi}} \int_{m/\sqrt{t}}^\infty e^{-y^2/2}\,dy, \qquad (y = x/\sqrt{t}) \\
f_{\tau_m}(t) &= \frac{d}{dt}\mathbb{P}\{\tau_m \le t\} \\
&= \frac{m}{t\sqrt{2\pi t}} \exp\left(-\frac{m^2}{2t}\right), \qquad t > 0.
\end{aligned}$$

By symmetry of Brownian motion, the same formulas hold for $m < 0$ upon replacing $m$ by $|m|$. Thus, although the first passage time $\tau_m$ is almost surely finite for $m \neq 0$, its density exhibits a heavy tail, implying that $\mathbb{E}[\tau_m] = \infty$.



### 1.8 Exercises: Brownian Motion

#### 1.8.1 Problem 1

**Solution.** Fix times $0 \leq t < u_1 < u_2$. Observe that the increment $W(u_2) - W(u_1)$ can be written as,

$$W(u_2) - W(u_1) = [W(u_2) - W(t)] - [W(u_1) - W(t)].$$

By Definition (iii), both $W(u_2) - W(t)$ and $W(u_1) - W(t)$ are random variables independent of $\mathcal{F}(t)$. Furthermore, by Definition (iii), Brownian Motion has independent increments, so the pair,

$$[W(u_2) - W(t), W(u_1) - W(t)]$$

are jointly independent of $\mathcal{F}(t)$. Consider $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = x - y$. Then we have that,

$$W(u_2) - W(u_1) = f(W(u_2) - W(t), W(u_1) - W(t)).$$

Since, measurable functions of random variables that are independent of a $\sigma$-algebra remain independent of that $\sigma$-algebra, it follows that $W(u_2) - W(u_1)$ is independent of $\mathcal{F}(t)$, as required. ■

#### 1.8.2 Problem 2

**Solution.** We can re-write $W^2(t)$ (as mentioned in the hint) as,

$$W^2(t) = (W(t) - W(s))^2 + 2W(s)(W(t) - W(s)) + W^2(s).$$

Then, taking expectations,

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

$M(t) = W^2(t) - t$, so we get after taking expectations,

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

Thus we have shown $\mathbb{E}[M(t) \mid \mathcal{F}(s)] = M(s)$ as required. ■

#### 1.8.3 Problem 3

**Solution.** Let $Y := X - \mu$ so that $\mathbb{E}[Y] = 0$ and $\mathrm{Var}(Y) = \sigma^2$. Thus, we know that the moment-generating function of $Y$ is,

$$\phi(u) := \mathbb{E}[e^{uY}] = e^{\frac{1}{2}\sigma^2 u^2}.$$

We take the derivative of $\phi(u)$ until we get to $\phi''''(u)$.

$$\begin{aligned}\phi'(u) &= \mathbb{E}[Ye^{uY}] = \sigma^2 u e^{\frac{1}{2}\sigma^2 u^2} \implies \phi'(0) = \mathbb{E}[Y] = 0, \\ \phi''(u) &= \mathbb{E}[Y^2 e^{uY}] = (\sigma^2 + \sigma^4 u^2)e^{\frac{1}{2}\sigma^2 u^2} \implies \phi''(0) = \mathbb{E}[Y^2] = \sigma^2, \\ \phi'''(u) &= \mathbb{E}[Y^3 e^{uY}] = (3\sigma^4 u + \sigma^6 u^3)e^{\frac{1}{2}\sigma^2 u^2} \implies \phi'''(0) = \mathbb{E}[Y^3], \\ \phi''''(u) &= \mathbb{E}[Y^3 e^{uY}] = (3\sigma^4 u + 6\sigma^6 u^2 + \sigma^8 u^4)e^{\frac{1}{2}\sigma^2 u^2} \implies \phi''''(0) = \mathbb{E}[Y^4] = 3\sigma^4.\end{aligned}$$

Since $\mathrm{Var}(X) = \sigma^2$ we have that, $\mathrm{Kurtosis}(X) = \dfrac{\mathbb{E}[(X - \mu)^4]}{(\mathrm{Var}(X))^2} = \dfrac{3\sigma^4}{\sigma^4} = 3.$ ■

### 1.8.4 Problem 4

**Solution.**

(i) Define, for each $n$,

$$S_n(\omega) := \sum_{j=0}^{m_n-1} \big(\Delta_j^{(n)} W(\omega)\big)^2, \qquad V_n(\omega) := \sum_{j=0}^{m_n-1} \big|\Delta_j^{(n)} W(\omega)\big|, \qquad M_n(\omega) := \max_{0\le j\le m_n-1} \big|\Delta_j^{(n)} W(\omega)\big|.$$

The hint inequality is just the pointwise bound

$$\big(\Delta_j^{(n)} W\big)^2 = \big|\Delta_j^{(n)} W\big| \cdot \big|\Delta_j^{(n)} W\big| \le M_n \, \big|\Delta_j^{(n)} W\big|.$$

Summing over $j$ gives

$$\sum_{j=0}^{m_n-1} \big(\Delta_j^{(n)} W\big)^2 \le M_n \sum_{j=0}^{m_n-1} \big|\Delta_j^{(n)} W\big| \iff S_n \le M_n V_n. \tag{1}$$

Now isolate $V_n$ (this is the key algebra step). For $\omega \in \Omega_*$, eventually $M_n(\omega) > 0$ and in fact $M_n(\omega) \to 0$. For such $n$,

$$V_n(\omega) \ge \frac{S_n(\omega)}{M_n(\omega)}. \tag{2}$$

Take limits along $n \to \infty$. Using (QV) and (Max), for every $\omega \in \Omega_*$,

$$S_n(\omega) \to T, \qquad M_n(\omega) \to 0.$$

In particular, there exists $N(\omega)$ such that for all $n \ge N(\omega)$,

$$S_n(\omega) \ge \frac{T}{2}.$$

Plug this into (2):

$$V_n(\omega) \ge \frac{S_n(\omega)}{M_n(\omega)} \ge \frac{T/2}{M_n(\omega)}.$$

Since $M_n(\omega) \to 0$, we have $\frac{1}{M_n(\omega)} \to \infty$, hence

$$V_n(\omega) \to \infty.$$

(ii) Define the sample cubic variation along $\Pi^{(n)}$.

$$C_n(\omega) := \sum_{j=0}^{m_n-1} \big|\Delta_j^{(n)} W(\omega)\big|^3.$$

Use the elementary inequality, for each $j$,

$$\big|\Delta_j^{(n)} W\big|^3 = \big|\Delta_j^{(n)} W\big| \cdot \big|\Delta_j^{(n)} W\big|^2 \le M_n \cdot \big|\Delta_j^{(n)} W\big|^2.$$

Summing over $j$ gives

$$\sum_{j=0}^{m_n-1} \big|\Delta_j^{(n)} W\big|^3 \le M_n \sum_{j=0}^{m_n-1} \big(\Delta_j^{(n)} W\big)^2 \iff C_n \le M_n S_n. \tag{3}$$

Now take limits on $\Omega_*$. By (Max), $M_n(\omega) \to 0$. By (QV), $S_n(\omega) \to T$. Hence for every $\omega \in \Omega_*$,

$$0 \le C_n(\omega) \le M_n(\omega) S_n(\omega) \to 0 \cdot T = 0.$$

■

### 1.8.5 Problem 5

**Solution.** Note that we can write the event,

$$(x-K)^+ = x\mathbf{I}\{x>K\} - K\mathbf{I}\{x>K\}.$$

Thus, we can write $(S(T)-K)^+$ as,

$$S(T)\mathbf{I}\{S(T)>K\} - K\mathbf{I}\{S(T)>K\}.$$

Then, the fair, no-arbitrage price of a European call can be written as,

$$C_0 = \mathbb{E}\left[e^{-rT}S(T)\mathbf{I}\{S(T)>K\}\right] - Ke^{-rT}P(S(T)>K)$$

We'll let $B := e^{-rT}S(T)\mathbf{I}\{S(T)>K\}$ and $A = P(S(T)>K)$ so that,

$$C_0 = B - Ke^{-rT}A.$$

The fair value at time 0 can be interpreted as the discounted risk-neutral expectation of the payoff. We derive the explicit call-pricing formula under Black-Scholes-Merton. We start with the even $S(T)>K$.

$$\begin{aligned}
S(T)>K &\iff S_0\exp\left(\left(r-\tfrac{1}{2}\sigma^2\right)T+\sigma W(T)\right) > K \\
&\iff \exp\left(\left(r-\tfrac{1}{2}\sigma^2\right)T+\sigma W(T)\right) > \frac{K}{S_0} \\
&\iff \left(r-\tfrac{1}{2}\sigma^2\right)T+\sigma W(T) > \log\left(\frac{K}{S_0}\right) \\
&\iff \sigma W(T) > \log\left(\frac{K}{S_0}\right) - \left(r-\tfrac{1}{2}\sigma^2\right)T \\
&\iff W(T) > \frac{1}{\sigma}\left[\log\left(\frac{K}{S_0}\right) - \left(r-\tfrac{1}{2}\sigma^2\right)T\right] \\
&\iff \underbrace{\sqrt{T}\,Z}_{Z\sim\mathcal{N}(0,1)} > \frac{1}{\sigma}\left[\log\left(\frac{K}{S_0}\right) - \left(r-\tfrac{1}{2}\sigma^2\right)T\right] \\
&\iff Z > \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{K}{S_0}\right) - \left(r-\tfrac{1}{2}\sigma^2\right)T\right] \\
&\iff Z > -\underbrace{\frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_0}{K}\right) + \left(r-\tfrac{1}{2}\sigma^2\right)T\right]}_{:=d_2} \\
&\iff Z > -d_2,
\end{aligned}$$

So the even $\{S(T)>K\} = \{Z>-d_2\}$ which means that,

$$A = P(S(T)>K) = P(Z>-d_2) = 1-P(Z\le -d_2) = 1-N(-d_2),$$

since $N(-x) = 1-N(x)$ so $1-N(-d_2) = N(d_2)$. Thus, the 2nd component of $C_0$ is just $Ke^{-rT}(1-N(-d_2))$. Using similar techniques, we simplify and derive the explicit formula for $B$

$$\begin{aligned}
B &= \mathbb{E}\left[e^{-rT}S(T)\mathbf{1}_{\{S(T)>K\}}\right] \\
&= \mathbb{E}\left[e^{-rT}S_0\exp\left(\left(r-\tfrac{1}{2}\sigma^2\right)T+\sigma W(T)\right)\mathbf{1}_{\{S(T)>K\}}\right] \\
&= S_0\mathbb{E}\left[e^{-rT}\exp\left(\left(r-\tfrac{1}{2}\sigma^2\right)T+\sigma W(T)\right)\mathbf{1}_{\{S(T)>K\}}\right] \\
&= S_0\mathbb{E}\left[\exp\left(-rT+\left(r-\tfrac{1}{2}\sigma^2\right)T+\sigma W(T)\right)\mathbf{1}_{\{S(T)>K\}}\right] \\
&= S_0\mathbb{E}\left[\exp\left(-\tfrac{1}{2}\sigma^2T+\sigma W(T)\right)\mathbf{1}_{\{S(T)>K\}}\right] \\
&= S_0\mathbb{E}\left[\exp\left(-\tfrac{1}{2}\sigma^2T+\sigma\sqrt{T}\,Z\right)\mathbf{1}_{\{S(T)>K\}}\right] \\
&= S_0\mathbb{E}\left[\exp\left(-\tfrac{1}{2}\sigma^2T+\sigma\sqrt{T}\,Z\right)\mathbf{1}_{\{Z>-d_2\}}\right].
\end{aligned}$$

Now, we write the expectation as an integral and simplify.

$$
\begin{aligned}
B &= S_0\mathbb{E}\Big[\exp\Big(-\tfrac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z\Big)\,\mathbf{1}_{\{Z>-d_2\}}\Big] \\
&= S_0\int_{-\infty}^{\infty} \exp\Big(-\tfrac{1}{2}\sigma^2 T + \sigma\sqrt{T}z\Big)\,\mathbf{1}_{\{z>-d_2\}}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz \\
&= S_0\int_{-d_2}^{\infty} \exp\Big(-\tfrac{1}{2}\sigma^2 T + \sigma\sqrt{T}z\Big)\,\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz \\
&= S_0\int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}\sigma^2 T + \sigma\sqrt{T}z - \tfrac{1}{2}z^2\Big)\,dz \\
&= S_0\int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}z^2 + \sigma\sqrt{T}z - \tfrac{1}{2}\sigma^2 T\Big)\,dz \\
&= S_0\int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}\Big[z^2 - 2\sigma\sqrt{T}z + \sigma^2 T\Big]\Big)\,dz \qquad (\text{factor } -\tfrac{1}{2}) \\
&= S_0\int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}\Big[(z-\sigma\sqrt{T})^2\Big]\Big)\,dz \qquad (\text{complete the square}) \\
&= S_0\int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}(z-\sigma\sqrt{T})^2\Big)\,dz \\
&= S_0\int_{-\infty}^{\infty} \mathbf{1}_{\{z>-d_2\}}\frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac{1}{2}(z-\sigma\sqrt{T})^2\Big)\,dz \\
&= S_0\int_{-\infty}^{\infty} \mathbf{1}_{\{y>-d_2-\sigma\sqrt{T}\}}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\,dy \qquad (y = z-\sigma\sqrt{T},\ dz = dy) \\
&= S_0\int_{-d_2-\sigma\sqrt{T}}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-y^2/2}\,dy \\
&= S_0\int_{-(d_2+\sigma\sqrt{T})}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-y^2/2}\,dy \\
&= S_0\int_{-d_1}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-y^2/2}\,dy \qquad (d_1 := d_2+\sigma\sqrt{T}) \\
&= S_0\,\mathbb{P}(Z>-d_1) \\
&= S_0\big(1-\mathbb{P}(Z\le -d_1)\big) \\
&= S_0\big(1-N(-d_1)\big) \\
&= S_0 N(d_1).
\end{aligned}
$$

Recall

$$C_0 = B - Ke^{-rT}A.$$

Insert $A = N(d_2)$ and $B = S_0 N(d_1)$:

$$C_0 = S_0 N(d_1) - Ke^{-rT}N(d_2).$$

Write $d_1 = d_+$ and $d_2 = d_-$:

$$
\begin{aligned}
d_+(T,S_0) &= \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_0}{K}\right) + \left(r+\tfrac{1}{2}\sigma^2\right)T\right], \\
d_-(T,S_0) &= \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_0}{K}\right) + \left(r-\tfrac{1}{2}\sigma^2\right)T\right].
\end{aligned}
$$

Therefore

$$\mathbb{E}\big[e^{-rT}(S(T)-K)^+\big] = S(0)\,N\big(d_+(T,S(0))\big) - Ke^{-rT}N\big(d_-(T,S(0))\big).$$

■

### 1.8.6 Problem 6

**Solution.**

(i) Let $W(t)$ be Brownian motion with filtration $\{\mathcal{F}(t)\}$ and define

$$X(t) = \mu t + W(t), \qquad \mu \in \mathbb{R}.$$

Fix $0 \le s < t$. Then

$$\begin{aligned} X(t) &= \mu t + W(t) \\ &= \mu s + W(s) + \mu(t-s) + \big(W(t) - W(s)\big) \\ &= X(s) + \mu(t-s) + \Delta W, \end{aligned}$$

where $\Delta W := W(t) - W(s)$ satisfies $\Delta W \sim N(0, t-s)$, $\Delta W \perp\!\!\!\perp \mathcal{F}(s)$, and $X(s)$ is $\mathcal{F}(s)$-measurable. Hence

$$\begin{aligned} \mathbb{E}[f(X(t)) \mid \mathcal{F}(s)] &= \mathbb{E}[f(X(s) + \mu(t-s) + \Delta W) \mid \mathcal{F}(s)] \\ &= \mathbb{E}[f(x + \mu(t-s) + \Delta W)]\Big|_{x=X(s)}. \end{aligned}$$

Writing the expectation as an integral using the density of $\Delta W$,

$$\mathbb{E}[f(x + \mu(t-s) + \Delta W)] = \int_{-\infty}^{\infty} f(x + \mu(t-s) + z)\, \frac{1}{\sqrt{2\pi(t-s)}} \exp\!\left(-\frac{z^2}{2(t-s)}\right) dz.$$

With the change of variables $y = x + \mu(t-s) + z$ (so $z = y - x - \mu(t-s)$),

$$\mathbb{E}[f(x + \mu(t-s) + \Delta W)] = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} f(y) \exp\!\left(-\frac{(y - x - \mu(t-s))^2}{2(t-s)}\right) dy.$$

Therefore

$$\mathbb{E}[f(X(t)) \mid \mathcal{F}(s)] = g(X(s)),$$

where

$$g(x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} f(y) \exp\!\left(-\frac{(y - x - \mu(t-s))^2}{2(t-s)}\right) dy.$$

Since the conditional expectation depends on $\mathcal{F}(s)$ only through $X(s)$, the process $X(t)$ is Markov with transition density

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} \exp\!\left(-\frac{(y - x - \mu\tau)^2}{2\tau}\right), \qquad \tau = t - s.$$

(ii) Let

$$S(t) = S(0)e^{\sigma W(t) + \nu t}, \qquad \sigma > 0,\ \nu \in \mathbb{R},$$

and fix $0 \le s < t$ with $\tau = t - s$. Then

$$\begin{aligned} S(t) &= S(0)e^{\sigma W(t) + \nu t} \\ &= S(0)e^{\sigma W(s) + \nu s} \exp\!\big(\sigma(W(t) - W(s)) + \nu(t-s)\big) \\ &= S(s) \exp\!\big(\sigma \Delta W + \nu\tau\big), \end{aligned}$$

where $\Delta W := W(t) - W(s) \sim N(0, \tau)$ and $\Delta W \perp\!\!\!\perp \mathcal{F}(s)$. Hence, for Borel $f$,

$$\begin{aligned} \mathbb{E}[f(S(t)) \mid \mathcal{F}(s)] &= \mathbb{E}\big[f\big(S(s)e^{\sigma\Delta W + \nu\tau}\big) \mid \mathcal{F}(s)\big] \\ &= \mathbb{E}\big[f\big(xe^{\sigma\Delta W + \nu\tau}\big)\big]\Big|_{x=S(s)}. \end{aligned}$$

Writing the expectation as an integral,

$$\mathbb{E}\big[f\big(xe^{\sigma\Delta W + \nu\tau}\big)\big] = \int_{-\infty}^{\infty} f\big(xe^{\sigma z + \nu\tau}\big)\, \frac{1}{\sqrt{2\pi\tau}} e^{-z^2/(2\tau)}\, dz.$$

With the change of variables

$$y = xe^{\sigma z + \nu\tau}, \qquad z = \frac{1}{\sigma}\left(\log\frac{y}{x} - \nu\tau\right), \qquad dz = \frac{1}{\sigma y}\,dy,$$

this becomes

$$\mathbb{E}\big[f\big(xe^{\sigma\Delta W + \nu\tau}\big)\big] = \int_0^\infty f(y)\,\frac{1}{\sigma y\sqrt{2\pi\tau}}\exp\left(-\frac{\big(\log(y/x) - \nu\tau\big)^2}{2\sigma^2\tau}\right)dy.$$

Therefore

$$\mathbb{E}[f(S(t)) \mid \mathcal{F}(s)] = g(S(s)),$$

where

$$g(x) = \int_0^\infty f(y)\,p(\tau, x, y)\,dy,$$

and the transition density is

$$p(\tau, x, y) = \frac{1}{\sigma y\sqrt{2\pi\tau}}\exp\left(-\frac{\big(\log(y/x) - \nu\tau\big)^2}{2\sigma^2\tau}\right).$$

Thus $S(t)$ is a Markov process with the above lognormal transition kernel.

■

### 1.8.7 Problem 7

**Solution.** Let

$$X(t) = \mu t + W(t), \qquad \mu \in \mathbb{R},$$

and for fixed $m > 0$ define the first hitting time

$$\tau_m = \inf\{t \geq 0 : X(t) = m\},$$

with the convention $\tau_m = \infty$ if the level is never reached. Let $\sigma > 0$ and define

$$Z(t) = \exp\big\{\sigma X(t) - \big(\sigma\mu + \tfrac{1}{2}\sigma^2\big)t\big\}.$$

(i) *Show that $Z(t)$ is a martingale.*

$$\begin{aligned} Z(t) &= \exp\big\{\sigma(\mu t + W(t)) - \big(\sigma\mu + \tfrac{1}{2}\sigma^2\big)t\big\} \\ &= \exp\big\{\sigma\mu t + \sigma W(t) - \sigma\mu t - \tfrac{1}{2}\sigma^2 t\big\} \\ &= \exp\big\{\sigma W(t) - \tfrac{1}{2}\sigma^2 t\big\}. \end{aligned}$$

For $0 \leq s < t$,

$$\begin{aligned} \frac{Z(t)}{Z(s)} &= \exp\big\{\sigma(W(t) - W(s)) - \tfrac{1}{2}\sigma^2(t-s)\big\}, \\ \mathbb{E}[Z(t) \mid \mathcal{F}(s)] &= Z(s)\mathbb{E}\big[\exp\big\{\sigma(W(t) - W(s)) - \tfrac{1}{2}\sigma^2(t-s)\big\}\big]. \end{aligned}$$

Since $W(t) - W(s) \sim N(0, t-s)$,

$$\mathbb{E}\big[e^{\sigma Y}\big] = e^{\frac{1}{2}\sigma^2(t-s)}, \qquad Y \sim N(0, t-s),$$

and therefore

$$\mathbb{E}\big[\exp\big\{\sigma(W(t) - W(s)) - \tfrac{1}{2}\sigma^2(t-s)\big\}\big] = 1,$$

so

$$\mathbb{E}[Z(t) \mid \mathcal{F}(s)] = Z(s).$$

(ii) *Apply optional stopping at $t \wedge \tau_m$.* Since $Z(t)$ is positive and a martingale,

$$\mathbb{E}[Z(t \wedge \tau_m)] = Z(0) = 1,$$
$$\mathbb{E}\big[\exp\{\sigma X(t \wedge \tau_m) - \big(\sigma\mu + \tfrac{1}{2}\sigma^2\big)(t \wedge \tau_m)\}\big] = 1.$$

(iii) *Case $\mu \geq 0$: show $\mathbb{P}(\tau_m < \infty) = 1$ and compute the Laplace transform.* On $\{\tau_m < \infty\}$,

$$t \wedge \tau_m \to \tau_m, \qquad X(t \wedge \tau_m) \to X(\tau_m) = m.$$

On $\{\tau_m = \infty\}$, since $\mu \geq 0$, $X(t) \to \infty$ a.s. By monotone convergence,

$$\mathbb{E}\left[e^{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m} \mathbf{1}_{\{\tau_m < \infty\}}\right] = 1.$$

The exponential term is strictly positive, hence

$$\mathbb{P}(\tau_m < \infty) = 1.$$

Rewriting,

$$\mathbb{E}\left[e^{-(\sigma\mu + \frac{1}{2}\sigma^2)\tau_m}\right] = e^{-\sigma m}.$$

Let $\alpha = \sigma\mu + \frac{1}{2}\sigma^2$. Solving

$$\sigma^2 + 2\mu\sigma - 2\alpha = 0 \quad \Longrightarrow \quad \sigma = \sqrt{\mu^2 + 2\alpha} - \mu,$$

we obtain

$$\begin{aligned}\mathbb{E}\big[e^{-\alpha\tau_m}\big] &= \exp\Big\{-m\big(\sqrt{\mu^2 + 2\alpha} - \mu\big)\Big\} \\ &= e^{m\mu - m\sqrt{\mu^2 + 2\alpha}}.\end{aligned}$$

(iv) *Compute $\mathbb{E}[\tau_m]$ for $\mu > 0$.* Differentiate the Laplace transform at $\alpha = 0$:

$$\begin{aligned}\mathbb{E}[\tau_m] &= -\frac{d}{d\alpha}\mathbb{E}\big[e^{-\alpha\tau_m}\big]\Big|_{\alpha=0} \\ &= -\frac{d}{d\alpha}\exp\Big\{m\mu - m\sqrt{\mu^2 + 2\alpha}\Big\}\Big|_{\alpha=0} \\ &= \frac{m}{\mu}.\end{aligned}$$

(v) *Case $\mu < 0$: hitting probability $< 1$.* Choose $\sigma > -2\mu$ so that $\sigma\mu + \frac{1}{2}\sigma^2 > 0$. Repeating the stopping argument,

$$\mathbb{E}\left[e^{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m} \mathbf{1}_{\{\tau_m < \infty\}}\right] = 1,$$

which yields

$$\mathbb{P}(\tau_m < \infty) = e^{-2m|\mu|} < 1.$$

The same quadratic calculation gives, for $\alpha > 0$,

$$\mathbb{E}\big[e^{-\alpha\tau_m}\big] = e^{m\mu - m\sqrt{\mu^2 + 2\alpha}}.$$

■

### 1.8.8 Problem 8

**Solution.** Fix $\sigma > 0$ and $r \geq 0$. For each $n$, there are $n$ steps per unit time, so over time $t$ there are $nt$ steps (assume $t$ rational and $nt \in \mathbb{N}$). The per–period interest rate is $r/n$, and $u_n = e^{\sigma/\sqrt{n}}$ and $d_n = e^{-\sigma/\sqrt{n}}$. Risk–neutral probabilities are

$$\tilde{p}_n = \frac{\frac{r}{n} + 1 - e^{-\sigma/\sqrt{n}}}{e^{\sigma/\sqrt{n}} - e^{-\sigma/\sqrt{n}}}, \qquad \tilde{q}_n = \frac{e^{\sigma/\sqrt{n}} - \frac{r}{n} - 1}{e^{\sigma/\sqrt{n}} - e^{-\sigma/\sqrt{n}}}.$$

Let $X_{1,n}, \ldots, X_{nt,n}$ be i.i.d. with

$$\tilde{\mathbb{P}}(X_{k,n} = 1) = \tilde{p}_n, \qquad \tilde{\mathbb{P}}(X_{k,n} = -1) = \tilde{q}_n,$$

and define

$$M_{nt,n} = \sum_{k=1}^{nt} X_{k,n}.$$

Then

$$S_n(t) = S(0) u_n^{\frac{1}{2}(nt + M_{nt,n})} d_n^{\frac{1}{2}(nt - M_{nt,n})} = S(0) \exp\left(\frac{\sigma}{\sqrt{n}} M_{nt,n}\right).$$

The goal is to show that $\frac{\sigma}{\sqrt{n}} M_{nt,n} \Rightarrow N\big((r - \frac{1}{2}\sigma^2)t, \sigma^2 t\big)$, so $S_n(t) \Rightarrow S(0) e^{\sigma W(t) + (r - \frac{1}{2}\sigma^2)t}$.

(i) *Compute the mgf* $\varphi_n(u)$ *of* $\frac{1}{\sqrt{n}} M_{nt,n}$.

$$\begin{aligned}
\varphi_n(u) &= \tilde{\mathbb{E}} \exp\left(u \frac{1}{\sqrt{n}} M_{nt,n}\right) = \tilde{\mathbb{E}} \exp\left(\frac{u}{\sqrt{n}} \sum_{k=1}^{nt} X_{k,n}\right) \\
&= \prod_{k=1}^{nt} \tilde{\mathbb{E}} \exp\left(\frac{u}{\sqrt{n}} X_{k,n}\right) = \left(\tilde{\mathbb{E}} e^{\frac{u}{\sqrt{n}} X_{1,n}}\right)^{nt} \\
&= \left(e^{u/\sqrt{n}} \tilde{p}_n + e^{-u/\sqrt{n}} \tilde{q}_n\right)^{nt}.
\end{aligned}$$

Let $x = 1/\sqrt{n}$. Then $e^{\pm\sigma/\sqrt{n}} = e^{\pm\sigma x}$ and

$$\tilde{p}_n = \frac{rx^2 + 1 - e^{-\sigma x}}{e^{\sigma x} - e^{-\sigma x}}, \qquad \tilde{q}_n = \frac{e^{\sigma x} - rx^2 - 1}{e^{\sigma x} - e^{-\sigma x}}.$$

Hence

$$e^{ux} \tilde{p}_n + e^{-ux} \tilde{q}_n = \frac{e^{ux}(rx^2 + 1 - e^{-\sigma x}) + e^{-ux}(e^{\sigma x} - rx^2 - 1)}{e^{\sigma x} - e^{-\sigma x}}.$$

(ii) *Compute* $\log \varphi_{1/x^2}(u)$ *and rewrite via hyperbolic functions.* Since $n = 1/x^2$ and $nt = t/x^2$,

$$\varphi_{1/x^2}(u) = \left(e^{ux}\tilde{p} + e^{-ux}\tilde{q}\right)^{t/x^2}.$$

Writing the numerator explicitly and grouping,

$$\begin{aligned}
e^{ux}\tilde{p} + e^{-ux}\tilde{q} &= \frac{(rx^2 + 1)(e^{ux} - e^{-ux}) + (e^{(\sigma - u)x} - e^{-(\sigma - u)x})}{e^{\sigma x} - e^{-\sigma x}} \\
&= \frac{(rx^2 + 1)\, 2\sinh(ux) + 2\sinh((\sigma - u)x)}{2\sinh(\sigma x)} \\
&= \frac{(rx^2 + 1)\sinh(ux) + \sinh((\sigma - u)x)}{\sinh(\sigma x)}.
\end{aligned}$$

Therefore

$$\log \varphi_{1/x^2}(u) = \frac{t}{x^2} \log\left(\frac{(rx^2 + 1)\sinh(ux) + \sinh((\sigma - u)x)}{\sinh(\sigma x)}\right).$$

Using $\sinh(\sigma x - ux) = \sinh(\sigma x)\cosh(ux) - \cosh(\sigma x)\sinh(ux)$,

$$\frac{(rx^2 + 1)\sinh(ux) + \sinh((\sigma - u)x)}{\sinh(\sigma x)} = \cosh(ux) + \frac{(rx^2 + 1 - \cosh(\sigma x))\sinh(ux)}{\sinh(\sigma x)}.$$

(iii) *Taylor expansion as* $x \to 0$. Using $\cosh z = 1 + \frac{z^2}{2} + O(z^4)$ and $\sinh z = z + O(z^3)$,

$$\begin{aligned}\cosh(ux) &= 1 + \frac{u^2x^2}{2} + O(x^4),\\ \sinh(ux) &= ux + O(x^3),\\ \cosh(\sigma x) &= 1 + \frac{\sigma^2x^2}{2} + O(x^4),\\ \sinh(\sigma x) &= \sigma x + O(x^3).\end{aligned}$$

Hence

$$\begin{aligned}\frac{(rx^2+1-\cosh(\sigma x))\sinh(ux)}{\sinh(\sigma x)} &= \frac{x^2(r-\frac{1}{2}\sigma^2)+O(x^4)}{\sigma x}\,(ux+O(x^3))\\ &= \frac{u}{\sigma}\Big(r-\tfrac{1}{2}\sigma^2\Big)x^2 + O(x^4).\end{aligned}$$

Therefore

$$\cosh(ux) + \frac{(rx^2+1-\cosh(\sigma x))\sinh(ux)}{\sinh(\sigma x)} = 1 + \frac{1}{2}u^2x^2 + \frac{u}{\sigma}\Big(r-\tfrac{1}{2}\sigma^2\Big)x^2 + O(x^4).$$

Taking logs,

$$\log\varphi_{1/x^2}(u) = \frac{t}{x^2}\left[\frac{1}{2}u^2x^2 + \frac{u}{\sigma}\Big(r-\tfrac{1}{2}\sigma^2\Big)x^2 + O(x^4)\right],$$

so

$$\log\varphi_{1/x^2}(u) \;\longrightarrow\; t\left(\frac{1}{2}u^2 + \frac{u}{\sigma}\Big(r-\tfrac{1}{2}\sigma^2\Big)\right).$$

This is the log–mgf of $N\big((r-\frac{1}{2}\sigma^2)t/\sigma,\ t\big)$ for $\frac{1}{\sqrt{n}}M_{nt,n}$, hence of $N\big((r-\frac{1}{2}\sigma^2)t,\ \sigma^2 t\big)$ for $\frac{\sigma}{\sqrt{n}}M_{nt,n}$.

(iv) *Use* $\log(1+z) = z + O(z^2)$ *to compute the limit mgf.*

$$\begin{aligned}\log\varphi_{1/x^2}(u) &= \frac{t}{x^2}\log\left[\cosh(ux) + \frac{(rx^2+1-\cosh(\sigma x))\sinh(ux)}{\sinh(\sigma x)}\right]\\ &= \frac{t}{x^2}\log\left[1 + \Big(\tfrac{1}{2}u^2 + \frac{ur}{\sigma} - \tfrac{1}{2}u\sigma\Big)x^2 + O(x^4)\right].\end{aligned}$$

Define

$$A(u) := \tfrac{1}{2}u^2 + \frac{ur}{\sigma} - \tfrac{1}{2}u\sigma,$$

so the bracket equals $1 + A(u)x^2 + O(x^4)$. Using $\log(1+z) = z + O(z^2)$ with $z = A(u)x^2 + O(x^4)$ and noting $z^2 = O(x^4)$,

$$\log\big(1 + A(u)x^2 + O(x^4)\big) = A(u)x^2 + O(x^4),$$

and hence

$$\log\varphi_{1/x^2}(u) = \frac{t}{x^2}\big(A(u)x^2 + O(x^4)\big) = tA(u) + O(x^2).$$

Letting $x \downarrow 0$ (equivalently $n \to \infty$),

$$\lim_{x\downarrow 0}\log\varphi_{1/x^2}(u) = t\Big(\tfrac{1}{2}u^2 + \frac{ur}{\sigma} - \tfrac{1}{2}u\sigma\Big),$$

so

$$\lim_{n\to\infty}\varphi_n(u) = \exp\Big\{t\Big(\tfrac{1}{2}u^2 + \frac{ur}{\sigma} - \tfrac{1}{2}u\sigma\Big)\Big\}.$$

Comparing with the mgf of $Y \sim N(m, v)$,

$$\mathbb{E}[e^{uY}] = \exp\big(mu + \tfrac{1}{2}vu^2\big),$$

we identify $v = t$ and $m = t\big(\frac{r}{\sigma} - \frac{\sigma}{2}\big)$, hence

$$\frac{1}{\sqrt{n}} M_{nt,n} \;\Rightarrow\; N\Big(t\Big(\frac{r}{\sigma} - \frac{\sigma}{2}\Big)\,,\, t\Big)\,.$$

Multiplying by $\sigma$,

$$\frac{\sigma}{\sqrt{n}} M_{nt,n} \;\Rightarrow\; N\big(\big(r - \tfrac{1}{2}\sigma^2\big)\,t,\, \sigma^2 t\big)\,.$$

Finally, since

$$S_n(t) = S(0)\exp\Big\{\frac{\sigma}{\sqrt{n}} M_{nt,n}\Big\}\,,$$

the limiting distribution is

$$S(0)\exp\big\{\sigma W(t) + \big(r - \tfrac{1}{2}\sigma^2\big)\,t\big\}\,.$$

■

# 2 Stochastic Calculus

## 2.1 Ito Integral for Simple Integrands

Suppose we have Brownian Motion, $W(t)$ for $t \geq 0$ together with a filtration $\mathcal{F}(t)$ (information up to time $t$). Let $\Delta(t)$ be an adapted stochastic process (e.g., the position we take in an asset at time $t$) such that it is required to be a $\mathcal{F}(t)$-measurable for all $t \geq 0$. Thus, information up to time $t$ is sufficient to evaluate $\Delta(t)$ and its randomness has been resolved. Intuitively, we cannot make a trading decision now based off of something that has not occurred yet, i.e., and event that does not occur in $[0, t]$. However, once we get to $\Delta(t)$, we know what position we must take in the underlying to hedge our position in an option, thus the random nature of $\Delta(t)$ has been resolved. Since $\Delta(t)$ is $\mathcal{F}(t)$-measurable, it is also independent of future increments since positions we take in assets may depend on the price history of those assets but must be independent of the future increments of the Brownian motion that drives the price process.

### 2.1.1 Integral Construction

Brownian motion paths are differentiable nowhere with respect to time. Thus,

$$\int_0^T \Delta(t) dg(t) = \int_0^T \Delta(t) g'(t) dt,$$

for some differentiable function $g(t)$, is not valid for Brownian motion. Similar to our approaches for other undifferentiable processes wrt $t$, Ito constructed the integral for simple $\Delta(t)$ and extended it as a limit of integration of simple integrands for general integrands. Suppose on $[t_j, t_{j+1})$, $\Delta(t)$ is constant (a simple process) where holding a single value at $t_j$ up to $t_{j+1}$ depending on the same $\omega$ on which $W(t)$ depends. Sampling different $\omega_i$'s from $\Omega$ would result in different realized Brownian Motion paths. By definition, $\Delta(0)$

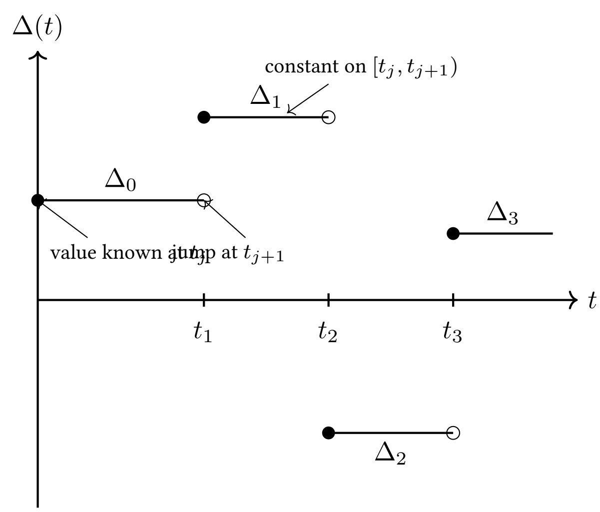



Figure 1: A path of a simple adapted process $\Delta(t)$ used in the construction of the Itô integral.

must be the same for all paths since there is no information at time 0.

Suppose $t_0, \dots, t_{n-1}$ are trading dates in an underlying so $\Delta(t_0), \dots, \Delta(t_{n-1})$ are positions taken in the underlying at each trading date (held to the next trading date). Intuitively, the gain from trading at $t$ is,

$$\begin{aligned}
I(t) &= \Delta(t_0)\big[W(t) - W(t_0)\big] = \Delta(0)W(t), && 0 < t < t_1, \\
I(t) &= \Delta(0)W(t_1) + \Delta(t_1)\big[W(t) - W(t_1)\big], && t_1 \leq t \leq t_2, \\
I(t) &= \Delta(0)W(t_1) + \Delta(t_1)\big[W(t_2) - W(t_1)\big] + \Delta(t_2)\big[W(t) - W(t_2)\big], && t_2 \leq t \leq t_3.
\end{aligned}$$

Or generally,

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j)\big[W(t_{j+1}) - W(t_j)\big] + \Delta(t_k)\big[W(t) - W(t_k)\big] \iff I(t) = \int_0^t \Delta(u) dW(u).$$

At time $t_j$, we lock in our position $\Delta(t_j)$ then the asset undergoes the random price change $W(t_{j+1}) - W(t_j)$. The Ito integral $I(t)$ sums the resulting gains over completed integrals plus currently partially realized gain on $[t_k, t]$.

### 2.1.2 $I(t)$ is a martingale.

$W(t)$ is a martingale so $I(t)$ (an upper limit of integration $t$) is also a martingale.

**Theorem 2.1.** *The Ito integral defined by $I(t)$ below,*

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j)\big[W(t_{j+1}) - W(t_j)\big] + \Delta(t_k)\big[W(t) - W(t_k)\big] \iff I(t) = \int_0^t \Delta(u)dW(u),$$

*is a martingale.*

**Proof.** We fix $0 \le s \le t \le T$ and define the partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$ of $\Delta$. WLOG, we choose indices $t_l \le s < t_{l+1}$ and $t_k \le t < t_{k+1}$ so $s \in [t_l, t_{l+1})$ and $t \in [t_k, t_{k+1})$. For the simple Ito integral,

$$I(u) = \int_0^u \Delta(r)dW(r),$$

we have the explicit sum representation for $u \in [t_m, t_{m+1})$,

$$I(u) = \sum_{j=0}^{m-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)] + \Delta(t_m)[W(u) - W(t_m)].$$

Using $l, k$ we write $I(t)$ as a sum of positions up to $t_l$ plus from $t_l$ to $s$ and from $t_{l+1}$ onward.

$$\begin{aligned} I(t) &= \sum_{j=0}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_k)\big(W(t) - W(t_k)\big) \\ &= \sum_{j=0}^{\ell-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_\ell)\big(W(t_{\ell+1}) - W(t_\ell)\big) \\ &\quad + \sum_{j=\ell+1}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_k)\big(W(t) - W(t_k)\big). \end{aligned}$$

Next, we isolate for $I(s)$,

$$I(s) = \sum_{j=0}^{\ell-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_\ell)\big(W(s) - W(t_\ell)\big).$$

And, subtract $I(s)$ from $I(t)$ to expose the future increment,

$$\begin{aligned} I(t) - I(s) &= \Delta(t_\ell)\Big(\big(W(t_{\ell+1}) - W(t_\ell)\big) - \big(W(s) - W(t_\ell)\big)\Big) \\ &\quad + \sum_{j=\ell+1}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_k)\big(W(t) - W(t_k)\big) \\ &= \Delta(t_\ell)\big(W(t_{\ell+1}) - W(s)\big) + \sum_{j=\ell+1}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_k)\big(W(t) - W(t_k)\big). \end{aligned}$$

Thus the resulting is an $I(s)$ + future decomposition of $I(t)$,

$$I(t) = I(s) + \Delta(t_\ell)\big(W(t_{\ell+1}) - W(s)\big) + \sum_{j=\ell+1}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) + \Delta(t_k)\big(W(t) - W(t_k)\big).$$

We know that the first term is $\mathcal{F}(s)$-measurable,

$$\mathbb{E}[I(s) \mid \mathcal{F}(s)] = I(s).$$

And, for the straddling term, we can take out what is known and apply the martingale property of Brownian Motion. Since $t_l \le w$ then $\Delta(t_l)$ is $\mathcal{F}(t_l) \subseteq \mathcal{F}(s)$, so we have that,

$$\begin{aligned}
\mathbb{E}\big[\Delta(t_\ell)\big(W(t_{\ell+1}) - W(s)\big) \mid \mathcal{F}(s)\big] &= \Delta(t_\ell)\,\mathbb{E}[W(t_{\ell+1}) - W(s) \mid \mathcal{F}(s)] \\
&= \Delta(t_\ell)\Big(\mathbb{E}[W(t_{\ell+1}) \mid \mathcal{F}(s)] - W(s)\Big) \\
&= \Delta(t_\ell)\big(W(s) - W(s)\big) \\
&= 0.
\end{aligned}$$

Then, we apply iterated conditioning for a generic future summand with $j \ge l+1$ for $t_j \ge t_{l+1} > s$,

$$\begin{aligned}
&\mathbb{E}\big[\Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) \mid \mathcal{F}(s)\big] \\
&\quad= \mathbb{E}\Big[\mathbb{E}\big[\Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) \mid \mathcal{F}(t_j)\big] \Bigm| \mathcal{F}(s)\Big] \\
&\quad= \mathbb{E}\Big[\Delta(t_j)\,\mathbb{E}[W(t_{j+1}) - W(t_j) \mid \mathcal{F}(t_j)] \Bigm| \mathcal{F}(s)\Big] \\
&\quad= \mathbb{E}\Big[\Delta(t_j)\Big(\mathbb{E}[W(t_{j+1}) \mid \mathcal{F}(t_j)] - W(t_j)\Big) \Bigm| \mathcal{F}(s)\Big] \\
&\quad= \mathbb{E}\big[\Delta(t_j)\big(W(t_j) - W(t_j)\big) \mid \mathcal{F}(s)\big] \\
&\quad= 0.
\end{aligned}$$

Summing the zeroes gives,

$$\mathbb{E}\left[\sum_{j=\ell+1}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) \Bigm| \mathcal{F}(s)\right] = 0.$$

The last partial interval term can be handled using the same iterated-conditioning trick since $t_k \ge t_{l+1} > s$,

$$\begin{aligned}
\mathbb{E}\big[\Delta(t_k)\big(W(t) - W(t_k)\big) \mid \mathcal{F}(s)\big] &= \mathbb{E}\Big[\mathbb{E}\big[\Delta(t_k)\big(W(t) - W(t_k)\big) \mid \mathcal{F}(t_k)\big] \Bigm| \mathcal{F}(s)\Big] \\
&= \mathbb{E}\Big[\Delta(t_k)\,\mathbb{E}[W(t) - W(t_k) \mid \mathcal{F}(t_k)] \Bigm| \mathcal{F}(s)\Big] \\
&= \mathbb{E}\Big[\Delta(t_k)\Big(\mathbb{E}[W(t) \mid \mathcal{F}(t_k)] - W(t_k)\Big) \Bigm| \mathcal{F}(s)\Big] \\
&= \mathbb{E}\big[\Delta(t_k)\big(W(t_k) - W(t_k)\big) \mid \mathcal{F}(s)\big] \\
&= 0.
\end{aligned}$$

Putting the conditioned terms together yields,

$$\begin{aligned}
\mathbb{E}[I(t) \mid \mathcal{F}(s)] &= \mathbb{E}[I(s) \mid \mathcal{F}(s)] + \mathbb{E}\big[\Delta(t_\ell)\big(W(t_{\ell+1}) - W(s)\big) \mid \mathcal{F}(s)\big] \\
&\quad + \mathbb{E}\left[\sum_{j=\ell+1}^{k-1} \Delta(t_j)\big(W(t_{j+1}) - W(t_j)\big) \Bigm| \mathcal{F}(s)\right] + \mathbb{E}\big[\Delta(t_k)\big(W(t) - W(t_k)\big) \mid \mathcal{F}(s)\big] \\
&= I(s) + 0 + 0 + 0 \\
&= I(s).
\end{aligned}$$

Hence, $I(t)$ for $t \ge 0$ is a martingale.

*Remark* 2.1. Since $I(t)$ is a martingale and $I(0) = 0$ and $\mathbb{E}I(t) = 0$ for all $t \ge 0$, we have that, $\mathrm{Var}I(t) = \mathbb{E}I^2(t)$.

### 2.1.3 Ito Isometry

**Theorem 2.2.** *The Ito integral defined in **Theorem 2.1** satisfies,*

$$\mathbb{E}I^2(t) = \mathbb{E}\int_0^t \Delta^2(u)du.$$

**Proof.** We fix $t \in [t_k, t_{k+1})$ and define $D_j := W(t_{j+1}) - W(t_j)$ for $j = 0, \dots, k-1$ and $D_k = W(t) - W(t_k)$. Then the Ito integral admits the representation,

$$I(t) = \sum_{j=0}^{k} \Delta(t_j)\, D_j.$$

We expand the square,

$$\begin{aligned} I(t)^2 &= \left(\sum_{j=0}^{k} \Delta(t_j) D_j\right)^2 \\ &= \sum_{j=0}^{k} \Delta^2(t_j) D_j^2 + 2\sum_{0\le i<j\le k} \Delta(t_i)\Delta(t_j) D_i D_j. \end{aligned}$$

Then, taking expectations of the above yields,

$$\mathbb{E}I(t)^2 = \sum_{j=0}^{k} \mathbb{E}\big[\Delta^2(t_j) D_j^2\big] + 2\sum_{0\le i<j\le k} \mathbb{E}[\Delta(t_i)\Delta(t_j) D_i D_j]\,.$$

We wish to show that the cross-terms vanish. Fix some $i < j$. Then we have that $\Delta(t_i)\Delta(t_j)D_i$ is $\mathcal{F}(t_j)$-measurable while $D_j$ is independent of $\mathcal{F}(t_j)$ and has mean zero. Thus we get that,

$$\begin{aligned} \mathbb{E}[\Delta(t_i)\Delta(t_j)D_iD_j] &= \mathbb{E}[\mathbb{E}[\Delta(t_i)\Delta(t_j)D_iD_j \mid \mathcal{F}(t_j)]] \\ &= \mathbb{E}[\Delta(t_i)\Delta(t_j)D_i\,\mathbb{E}[D_j \mid \mathcal{F}(t_j)]] \\ &= \mathbb{E}[\Delta(t_i)\Delta(t_j)D_i \cdot 0] \\ &= 0 \\ \implies 2\sum_{0\le i<j\le k} \mathbb{E}[\Delta(t_i)\Delta(t_j)D_iD_j] &= 0. \end{aligned}$$

Each $\Delta^2(t_j)$ is $\mathcal{F}(t_j)$-measurable and independent of $D_j^2$ so we know $\mathbb{E}\big[\Delta^2(t_j)D_j^2\big] = \mathbb{E}\big[\Delta^2(t_j)\big]\,\mathbb{E}[D_j^2]$.
For Brownian Motion,

$$\mathbb{E}[D_j^2] = t_{j+1} - t_j, \quad \mathbb{E}[D_k^2] = t - t_k. \qquad j = 0, \dots, k-1,$$

Thus yields $\mathbb{E}I^2(t)$ as,

$$\mathbb{E}I(t)^2 = \sum_{j=0}^{k-1} \mathbb{E}\big[\Delta^2(t_j)\big]\,(t_{j+1} - t_j) + \mathbb{E}\big[\Delta^2(t_k)\big]\,(t - t_k).$$

Since $\Delta(u)$ is constant over $[t_j, t_{j+1})$,

$$\begin{aligned} \Delta^2(t_j)(t_{j+1} - t_j) &= \int_{t_j}^{t_{j+1}} \Delta^2(u)\,du, \qquad j = 0, \dots, k-1, \\ \Delta^2(t_k)(t - t_k) &= \int_{t_k}^{t} \Delta^2(u)\,du. \end{aligned}$$

Substituting the expectation we get,

$$\begin{aligned}
\mathbb{E}I(t)^2 &= \sum_{j=0}^{k-1} \mathbb{E}\left[\int_{t_j}^{t_{j+1}} \Delta^2(u)\, du\right] + \mathbb{E}\left[\int_{t_k}^{t} \Delta^2(u)\, du\right] \\
&= \mathbb{E}\left[\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \Delta^2(u)\, du + \int_{t_k}^{t} \Delta^2(u)\, du\right] \\
&= \mathbb{E}\left[\int_{0}^{t} \Delta^2(u)\, du\right].
\end{aligned}$$

### 2.1.4 Quadratic Variation of Ito Integrals

**Theorem 2.3.** *The quadratic variation accumulated up to time $t$ by the Ito Integral is,*

$$[I,I](t) = \int_0^t \Delta^2(u)du.$$

**Proof.** Fix an interval $[t_j, t_{j+1})$ on which $\Delta(u) = \Delta(t_j)$ is consistent. We can choose a refinement $t_j = s_0 < s_1 < \cdots < s_m = t_{j+1}$. On this interval, the increment of the Ito integral satisfies,

$$I(s_{i+1}) - I(s_i) = \Delta(t_j)[W(s_{i+1}) - W(s_i)].$$

Hence, we compute the quadratic variation sum over $[t_j, t_{j+1})$ as,

$$\begin{aligned}
\sum_{i=0}^{m-1} \big[I(s_{i+1}) - I(s_i)\big]^2 &= \sum_{i=0}^{m-1} \Delta^2(t_j)\big(W(s_{i+1}) - W(s_i)\big)^2 \\
&= \Delta^2(t_j) \sum_{i=0}^{m-1} \big(W(s_{i+1}) - W(s_i)\big)^2.
\end{aligned}$$

As our mesh $\max_{0\le i\le m-1}(s_{i+1} - s_i) \to 0$, we know that the Brownian Motion quadratic variation converges pathwise,

$$\sum_{i=0}^{m-1} \big(W(s_{i+1}) - W(s_i)\big)^2 \;\longrightarrow\; t_{j+1} - t_j.$$

Therefore the quadratic variation accumulated by the Ito integral on $[t_j, t_{j+1})$ is,

$$\Delta^2(t_j)(t_{j+1} - t_j).$$

We consider only simple integrands, so $\Delta(u)$ is constant on $[t_j, t_{j+1})$, so it can be written as,

$$\Delta^2(t_j)(t_{j+1} - t_j) = \int_{t_j}^{t_{j+1}} \Delta^2(u)\, du.$$

Performing the same computation on the final partial interval $[t_j, t]$ gives us,

$$\Delta^2(t_k)(t - t_k) = \int_{t_k}^{t} \Delta^2(u)\, du.$$

Thus, summing quadratic variation contributions over all subintervals yields us,

$$\begin{aligned}
[I,I](t) &= \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \Delta^2(u)\, du + \int_{t_k}^{t} \Delta^2(u)\, du \\
&= \int_0^t \Delta^2(u)\, du.
\end{aligned}$$

*Remark* 2.2 (Differential vs. integral notation for the Itô integral). Recall that Brownian motion has quadratic variation $dW(t)\,dW(t) = dt$ which is a concise way of stating that

$$[W, W](t) = t.$$

If the Itô integral is written in integral form as

$$I(t) = \int_0^t \Delta(u)\,dW(u),$$

then its differential form is

$$dI(t) = \Delta(t)\,dW(t).$$

This notation expresses that an infinitesimal change in the Itô integral is obtained by multiplying the current integrand value $\Delta(t)$ by the corresponding Brownian increment. Using the quadratic variation rule $dW(t)\,dW(t) = dt$, we compute

$$dI(t)\,dI(t) = \Delta^2(t)\,dW(t)\,dW(t) = \Delta^2(t)\,dt.$$

Thus the Itô integral accumulates quadratic variation at instantaneous rate $\Delta^2(t)$, which is another way of expressing the result

$$[I, I](t) = \int_0^t \Delta^2(u)\,du.$$

The notations

$$I(t) = \int_0^t \Delta(u)\,dW(u) \qquad \text{and} \qquad dI(t) = \Delta(t)\,dW(t)$$

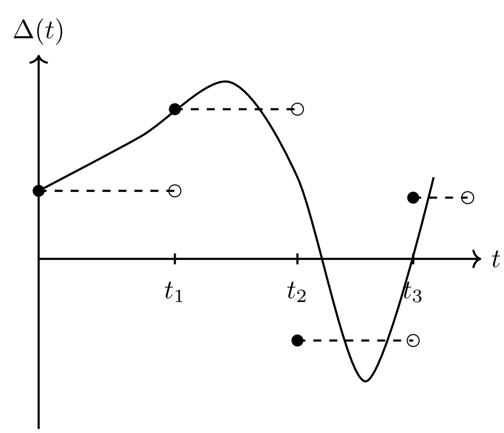
describe the same object. The integral form has a precise definition via limits of simple integrands, while the differential form is shorthand that encodes how increments behave. Integrating the differential form yields

$$I(t) = I(0) + \int_0^t \Delta(u)\,dW(u).$$

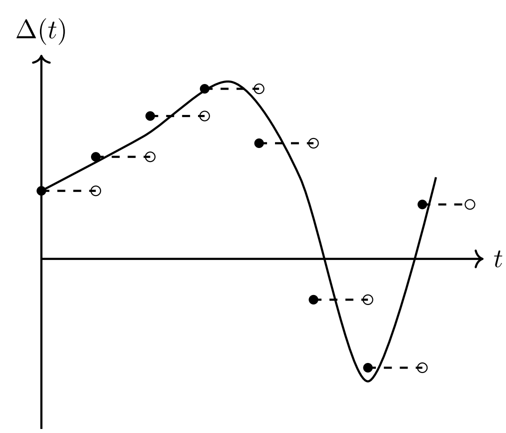
When one writes $I(t) = \int_0^t \Delta(u)\,dW(u)$, it is implicitly assumed that $I(0) = 0$. In contrast, the differential and integral forms above allow $I(0)$ to be an arbitrary constant.

### 2.2 Itô Integral for General Integrands

We again assume that $\Delta(t), t \geq 0$, is an adapted stochastic process with respect to a filtration $\{\mathcal{F}(t)\}_{t \geq 0}$. In contrast to the previous subsection, we now allow $\Delta(t)$ to be a general process that may be continuous or possess jumps. We assume square integrability, namely $\mathbb{E}\displaystyle\int_0^T \Delta^2(t)\,dt < \infty$ which ensures that the Itô integral is well defined via the Itô isometry. For such general integrands, the Itô integral $\int_0^T \Delta(t)\,dW(t)$ is constructed by approximating $\Delta(t)$ with simple adapted processes. Specifically, for a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, we replace $\Delta(t)$ by a step process that is constant on each interval $[t_j, t_{j+1})$ and equal to $\Delta(t_j)$. As the mesh of the partition, $\|\Pi\| = \max_j(t_{j+1} - t_j)$, tends to zero, these simple processes converge to $\Delta(t)$ in $L^2$, and the corresponding Itô integrals converge in $L^2$. The limit defines the Itô integral for general square-integrable adapted integrands.




(coarse partition)




$(\max |t_{j+1} - t_j| \to 0)$

### 2.2.1 Approximating Sequence $\Delta_n(t)$

In the $L^2$ sense, we choose a sequence $\Delta_n(t)$ of simple processes that converge to $\Delta(t)$ as $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{E} \int_0^T |\Delta_n(t) - \Delta(t)|^2 dt = 0.$$

Thus, we can write the continuously varying integrand $\Delta(t)$ using the Ito integral of $\Delta_n(t)$,

$$\int_0^t \Delta(u) dW(u) = \lim_{n\to\infty} \int_0^t \Delta_n(u) dW(u), \quad 0 \le t \le T.$$

The integral of the continuously varying adapted process, $\Delta(t)$ inherits properties of the Ito integrals of simple processes.

**Theorem 2.4** (Properties of the Itô Integral)**.** *Let $T > 0$ and let $\Delta(t)$, $0 \le t \le T$, be an adapted stochastic process satisfying*

$$\mathbb{E} \int_0^T \Delta^2(t)\, dt < \infty.$$

*Define the Itô integral by*

$$I(t) = \int_0^t \Delta(u)\, dW(u), \qquad 0 \le t \le T.$$

*Then the following properties hold.*

*(i)* ***(Continuity)*** *The paths of $I(t)$ are continuous.*

**Proof.** The Itô integral is the $L^2$-limit of integrals of simple processes, each of which has continuous paths.

$$I(t) = \lim_{n\to\infty} \int_0^t \Delta_n(u)\, dW(u),$$

where each $\int_0^t \Delta_n(u)\, dW(u)$ is continuous in $t$.

*(ii)* ***(Adaptivity)*** *For each $t$, $I(t)$ is $\mathcal{F}(t)$-measurable.*

**Proof.** Each simple-process integral is $\mathcal{F}(t)$-measurable, and measurability is preserved under $L^2$ limits.

$$I(t) = \lim_{n\to\infty} \sum_j \Delta_n(t_j)\big(W(t_{j+1} \wedge t) - W(t_j \wedge t)\big).$$

*(iii)* ***(Linearity)*** *If $I(t) = \int_0^t \Delta(u)\, dW(u)$ and $J(t) = \int_0^t \Gamma(u)\, dW(u)$, then*

$$I(t) + J(t) = \int_0^t (\Delta(u) + \Gamma(u))\, dW(u), \quad cI(t) = \int_0^t c\,\Delta(u)\, dW(u).$$

**Proof.** Linearity holds for simple integrands and passes to limits.

$$\begin{aligned}\int_0^t (\Delta + \Gamma)\, dW &= \lim_{n\to\infty} \int_0^t (\Delta_n + \Gamma_n)\, dW \\ &= \lim_{n\to\infty} \left( \int_0^t \Delta_n\, dW + \int_0^t \Gamma_n\, dW \right).\end{aligned}$$

*(iv)* ***(Martingale property)*** *$\{I(t)\}_{t\ge 0}$ is a martingale.*

**Proof.** For $0 \le s \le t$, future Brownian increments have zero conditional expectation.

$$\mathbb{E}[I(t) \mid \mathcal{F}(s)] = I(s) + \mathbb{E}\left[ \int_s^t \Delta(u)\, dW(u) \,\middle|\, \mathcal{F}(s) \right] = I(s).$$

(v) ***(Itô isometry)***

$$\mathbb{E}I(t)^2 = \mathbb{E}\int_0^t \Delta^2(u)\,du.$$

**Proof.** Square, expand, and use independence and zero mean of Brownian increments.

$$\mathbb{E}I(t)^2 = \mathbb{E}\left[\left(\sum_j \Delta(t_j)\Delta W_j\right)^2\right] = \sum_j \mathbb{E}[\Delta^2(t_j)]\,\mathbb{E}[(\Delta W_j)^2].$$

(vi) ***(Quadratic variation)***

$$[I,I](t) = \int_0^t \Delta^2(u)\,du.$$

**Proof.** Quadratic variation accumulates pathwise from squared increments.

$$\sum_i (I(t_{i+1}) - I(t_i))^2 = \sum_i \Delta^2(t_i)\big(W(t_{i+1}) - W(t_i)\big)^2 \longrightarrow \int_0^t \Delta^2(u)\,du.$$

### 2.2.2 Brownian Integral

*Example* 2.1. We wish to compute the integral,

$$\int_0^T W(t)dW(t),$$

by defining an approximating sequence of simple adapted processes for $W(t)$.

**Solution.** Fix $n \in \mathbb{N}$ and define the partition,

$$0 = 0\cdot\frac{T}{n} < \frac{T}{n} < \frac{2T}{n} < \cdots < \frac{(n-1)T}{n} < T.$$

We construct the simple process, $\Delta_n(t)$ as,

$$\Delta_n(t) = W\left(\frac{jT}{n}\right), \quad \frac{jT}{n} \le t < \frac{(j+1)T}{n}, \quad j = 0,\dots,n-1.$$

Then, in the $L^2$ sense, we have that,

$$\lim_{n\to\infty} \mathbb{E}\int_0^T [\Delta_n(t) - W(t)]^2 dt = 0.$$

By the definition of the Ito integral, we have,

$$\int_0^T W(t)dW(t) = \lim_{n\to\infty}\int_0^T \Delta_n(t)dW(t).$$

For each $n$, we can write the Ito sum explicitly,

$$\int_0^T \Delta_n(t)\,dW(t) = \sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right)\left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right].$$

Let $W_j := W\left(\frac{jT}{n}\right)$ for $j = 0,\dots,n$ with $W_0 = 0$. Then our sum becomes $\sum_{j=0}^{n-1} W_j\,(W_{j+1} - W_j)$. Starting from the elementary identity, then summing from $j = 0$ to $n-1$ we get,

$$(W_{j+1} - W_j)^2 = W_{j+1}^2 - 2W_jW_{j+1} + W_j^2,$$
$$\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2 = \sum_{j=0}^{n-1} W_{j+1}^2 - 2\sum_{j=0}^{n-1} W_jW_{j+1} + \sum_{j=0}^{n-1} W_j^2.$$

We can re-indent the first sum and use $W_0 = 0$ then re-arranging, we get

$$\sum_{j=0}^{n-1} W_{j+1}^2 = \sum_{k=1}^{n} W_k^2 = W_n^2 + \sum_{j=1}^{n-1} W_j^2.$$

Substituting this back, we get,

$$\begin{aligned}\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2 &= W_n^2 + \sum_{j=1}^{n-1} W_j^2 - 2\sum_{j=0}^{n-1} W_j W_{j+1} + \sum_{j=0}^{n-1} W_j^2 \\ &= W_n^2 + 2\sum_{j=0}^{n-1} W_j^2 - 2\sum_{j=0}^{n-1} W_j W_{j+1}, \\ \sum_{j=0}^{n-1} W_j(W_{j+1} - W_j) &= \frac{1}{2}W_n^2 - \frac{1}{2}\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2.\end{aligned}$$

Next, we take the limit as $n \to \infty$ to get,

$$\begin{aligned}\int_0^T W(t)\,dW(t) &= \lim_{n\to\infty}\sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right)\left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right] \\ &= \frac{1}{2}W^2(T) - \frac{1}{2}\lim_{n\to\infty}\sum_{j=0}^{n-1}\left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right]^2.\end{aligned}$$

Using the fact of quadratic variation of Brownian motion,

$$\sum_{j=0}^{n-1}\left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right]^2 \longrightarrow T.$$

Thus, we get that,

$$\int_0^T W(t)\,dW(t) = \frac{1}{2}W^2(T) - \frac{1}{2}T.$$

The additional $-\frac{1}{2}T$ term comes from non-zero quadratic variation of Brownian motion and left-endpoint evaluation. Replacing left endpoints by midpoints yields the Stratonovich sum,

$$\lim_{n\to\infty}\sum_{j=0}^{n-1} W\left(\frac{(j+\frac{1}{2})T}{n}\right)\left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right],$$

for which the quadratic variation term disappears. Thus,

$$\int_0^t W(u)\,dW(u) = \frac{1}{2}W^2(t) - \frac{1}{2}t.$$

This process is a martingale with expectation zero, as required. ■

### 2.3 Ito-Doeblin

Given $f(W(t))$ where $f$ is differentiable and $W(t)$ is Brownian motion, calculus yields $\frac{d}{dt}f(W(t)) = f'(W(t))W'(t)$. Alternatively, we could alos write $df(W(t)) = f'(W(t))W'(t)dt = f'(W(t))dW(t)$. However, $W(\cdot)$ has non-zero quadratic variation, so we have an extra $+\frac{1}{2}f''(W(t))dt$ term. Integrating yields the *Ito-Doeblin Formulation*,

$$f(W(t)) - f(W(0)) = \int_0^t f'(W(u))dW(u) + \frac{1}{2}\int_0^t f''(W(u))du.$$

$df(W(t))$ is the change in $f(W(t))$ when $t$ makes an infinitesimally small change and $dW(t)$ is the change in Brownian motion when $t$ makes an infinitesimally small change. We encounter similar things in calculus when trying to compute $\int f(u)f'(u)du$ and do a change-of-variables to get $\frac{1}{2}f^2(u) + C$.

**Theorem 2.5.** *Let $f(t,x)$ be a function for which the partial derivatives $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$ are defined and continuous, and let $W(t)$ be a Brownian motion. Then for every $T \geq 0$,*

$$f(T,W(T)) = f(0,W(0)) + \int_0^T f_t(t,W(t))dt + \int_0^T f_x(t,W(t))dW(t) + \frac{1}{2}\int_0^T f_{xx}(t,W(t))dt.$$

**Proof.** We first apply Taylor's formula on increments $W_{j+1} - W_j$ with $W_0 = 0$,

$$\begin{aligned} f(x_{j+1}) - f(x_j) &= f'(x_j)(x_{j+1} - x_j) + \frac{1}{2}f''(x_j)(x_{j+1} - x_j)^2, \\ &= x_j(W_{j+1} - W_j) + \frac{1}{2}(W_{j+1} - W_j)^2. \end{aligned}$$

Because all higher-order derivatives vanish, the expansion above is exact.
Then, we can split the change in $f(W(t))$ over an arbitrary partition,

$$\begin{aligned} f(W(T)) - f(W(0)) &= \sum_{j=0}^{n-1}\left[f(W_{j+1}) - f(W_j)\right], \\ &= \sum_{j=0}^{n-1} W_j(W_{j+1} - W_j) + \frac{1}{2}\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2. \end{aligned}$$

Starting from $(W_{j+1} - W_j)^2 = W_{j+1}^2 - 2W_jW_{j+1} + W_j^2$, we can sum from $j = 0$ to $n-1$, and re-arranging we get,

$$\begin{aligned} \sum_{j=0}^{n-1}(W_{j+1} - W_j)^2 &= W_n^2 + \sum_{j=0}^{n-1} W_j^2 - 2\sum_{j=0}^{n-1} W_jW_{j+1} \\ \sum_{j=0}^{n-1} W_j(W_{j+1} - W_j) &= \frac{1}{2}W_n^2 - \frac{1}{2}\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2. \end{aligned}$$

Recall that Brownian motion quadratic variation converges pathwise, so we have that,

$$\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2 \longrightarrow T \implies \int_0^T W(t)\,dW(t) = \frac{1}{2}W^2(T) - \frac{1}{2}T.$$

For $i < j$ use independence of our increment $D_j$ from $\mathcal{F}(t_j)$ and $\mathbb{E}[D_j] = 0$,

$$\mathbb{E}[\Delta(t_i)\Delta(t_j)D_iD_j] = \mathbb{E}[\Delta(t_i)\Delta(t_j)D_i]\mathbb{E}[D_j] = 0.$$

Recall the Ito isometry property of Ito integrals,

$$\mathbb{E}I(t)^2 = \sum_{j=0}^{k-1}\mathbb{E}[\Delta^2(t_j)](t_{j+1} - t_j) + \mathbb{E}[\Delta^2(t_k)](t - t_k) = \mathbb{E}\left[\int_0^t \Delta^2(u)\,du\right].$$

So, we have our pathwise quadratic variation is,

$$[I,I](t) = \sum_{j=0}^{k-1}\Delta^2(t_j)(t_{j+1} - t_j) + \Delta^2(t_k)(t - t_k) = \int_0^t \Delta^2(u)\,du.$$

*Remark* 2.3. A similar approach to writing the Ito-Doeblin formula is by writing in its full differential form,

$$df(t,W(t)) = f_t(t,W(t))dt + f_x(t,W(t))dW(t) + \frac{1}{2}f_{xx}(t,W(t))dW(t)dW(t) + f_{tx}(t,W(t))dtdW(t) + \frac{1}{2}f_{tt}(t,W(t))dtdt.$$

However, note that $dW(t)dW(t) = dt$, $dtdW(t) = dW(t)dt = 0$ and $dtdt = 0$, so we get that,

$$df(t,W(t)) = f_t(t,W(t))dt + f_x(t,W(t))dW(t) + \frac{1}{2}f_{xx}(t,W(t))dt.$$

*Remark* 2.4. We can formulate a Taylor series approximation of $f(W(t_{j+1})) - f(W(t_j))$ for a function $f(x)$ that does not depend on $t$. The first order approximation has error resulting from the convexity of $f(\cdot)$. Adding a second order term captures more of the curvature of the function $f(x)$ for $x = W(t_j)$.

$$f(W(t_{j+1})) - f(W(t_j)) = f'(W(t_j))(W(t_{j+1}) - W(t_j)) + \text{small error.}$$
$$f(W(t_{j+1})) - f(W(t_j)) = f'(W(t_j))(W(t_{j+1}) - W(t_j)) + \frac{1}{2}f''(W(t_j))(W(t_{j+1}) - W(t_j))^2 + \text{smaller error.}$$

In both equations above, as the max step size of both partitions approaches 0, their errors approach zero as well. However, in the first one, when we sum both sides, the errors accumulate although the error in each summand approaches 0, the errors do not. The extra 2nd order term provides more accuracy since the paths of Brownian motion are so volatile.

*Example* 2.2. Let $f(x) = \frac{1}{2}x^2$ for $x := W(t)$, then Ito-Doeblin yields,

$$\begin{aligned}
\frac{1}{2}W^2(T) &= f(W(T)) - f(W(0)) \\
&= \int_0^T f'(W(t))dW(t) + \frac{1}{2}\int_0^t f''(W(t))dt \\
&= \int_0^T W(t)dW(t) + \frac{1}{2}T.
\end{aligned}$$

### 2.4 General Ito Stochastic Processes

**Definition 2.1.** For Brownian motion $W(t)$ for $t \geq 0$ and associated filtration $\mathcal{F}(t)$ containing all information up to time $t$ for $t \geq 0$. Then an Ito Process is a stochastic process of the form,

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du,$$

where $X(0)$ is deterministic and $\Delta(u), \Theta(u)$ are adapted stochastic processes.

*Remark* 2.5. The volatility associated with Ito processes is determined by the rate at which they accumulate quadratic variation.

**Lemma 2.1.** *The Quadratic Variation of an Ito Process is,*

$$[X, X](t) = \int_0^t \Delta^2(u)du.$$

**Proof.** Write

$$I(t) = \int_0^t \Delta(u)\,dW(u), \qquad R(t) = \int_0^t \Theta(u)\,du,$$

so that

$$X(t) = X(0) + I(t) + R(t).$$

Fix $t > 0$ and let $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ be a partition of $[0, t]$. The sampled quadratic variation of $X$ along $\Pi$ is

$$\sum_{j=0}^{n-1} \big[X(t_{j+1}) - X(t_j)\big]^2.$$

Since $X(0)$ is constant, increments satisfy

$$X(t_{j+1}) - X(t_j) = \big(I(t_{j+1}) - I(t_j)\big) + \big(R(t_{j+1}) - R(t_j)\big).$$

Expanding the square and summing gives

$$\begin{aligned}
\sum_{j=0}^{n-1} \big[X(t_{j+1}) - X(t_j)\big]^2 &= \sum_{j=0}^{n-1} \big[I(t_{j+1}) - I(t_j)\big]^2 + \sum_{j=0}^{n-1} \big[R(t_{j+1}) - R(t_j)\big]^2 \\
&\quad + 2\sum_{j=0}^{n-1} \big[I(t_{j+1}) - I(t_j)\big]\big[R(t_{j+1}) - R(t_j)\big].
\end{aligned}$$

As $\|\Pi\| \to 0$, the first term converges to the quadratic variation of the Itô integral:

$$\sum_{j=0}^{n-1} \big[I(t_{j+1}) - I(t_j)\big]^2 \longrightarrow [I, I](t) = \int_0^t \Delta^2(u)\, du.$$

For the second term, note that

$$R(t_{j+1}) - R(t_j) = \int_{t_j}^{t_{j+1}} \Theta(u)\, du.$$

Hence

$$\begin{aligned}
\sum_{j=0}^{n-1} \big[R(t_{j+1}) - R(t_j)\big]^2 &\le \max_{0\le k\le n-1} \big|R(t_{k+1}) - R(t_k)\big| \sum_{j=0}^{n-1} \big|R(t_{j+1}) - R(t_j)\big| \\
&= \max_{0\le k\le n-1} \left|\int_{t_k}^{t_{k+1}} \Theta(u)\, du\right| \sum_{j=0}^{n-1} \left|\int_{t_j}^{t_{j+1}} \Theta(u)\, du\right| \\
&\le \max_{0\le k\le n-1} \left|\int_{t_k}^{t_{k+1}} \Theta(u)\, du\right| \int_0^t |\Theta(u)|\, du.
\end{aligned}$$

Since $R(t)$ is continuous, the maximum term tends to 0 as $\|\Pi\| \to 0$, and therefore

$$\sum_{j=0}^{n-1} \big[R(t_{j+1}) - R(t_j)\big]^2 \longrightarrow 0.$$

For the cross term, we estimate

$$\begin{aligned}
&\left| 2\sum_{j=0}^{n-1} \big[I(t_{j+1}) - I(t_j)\big]\big[R(t_{j+1}) - R(t_j)\big] \right| \\
&\qquad \le 2 \max_{0\le k\le n-1} \big|I(t_{k+1}) - I(t_k)\big| \sum_{j=0}^{n-1} \big|R(t_{j+1}) - R(t_j)\big| \\
&\qquad = 2 \max_{0\le k\le n-1} \big|I(t_{k+1}) - I(t_k)\big| \int_0^t |\Theta(u)|\, du.
\end{aligned}$$

Since $I(t)$ is continuous, the maximum increment $\max_k |I(t_{k+1}) - I(t_k)|$ converges to 0 as $\|\Pi\| \to 0$, and hence

$$\sum_{j=0}^{n-1} \big[I(t_{j+1}) - I(t_j)\big]\big[R(t_{j+1}) - R(t_j)\big] \longrightarrow 0.$$

Combining all three limits, we conclude that

$$[X, X](t) = \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \big[X(t_{j+1}) - X(t_j)\big]^2 = [I, I](t) = \int_0^t \Delta^2(u)\, du.$$

*Remark* 2.6. You may notice, many of our conclusions can be simplified by Ito-Doeblin by first converting definition 2.1 into differential form, given by,

$$dX(t) = \Delta(t)dW(t) + \Theta(t)dt,$$

which can then give us,

$$dX(t)dX(t) = \Delta^2(t)dW(t)dW(t) + 2\Delta(t)\Theta(t)dW(t)dt + \Theta^2(t)dtdt = \Delta^2(t)dt.$$

The formulations above essentially tell us the Ito process $X$ accumulates quadratic variation at rate $\Delta^2(t)$ per unit time, which can be attributed purely to Ito integral $I(t) = \int_0^t \Delta(u)dW(u)$ since $R(t)$ does not contribute quadratic variation. This does not imply that $R(t)$ is non-random since $\Theta(u)$ is allowed to be random. Similar to investing in the money market, we have a good estimate of the return over the near future knowing today's rate. However, interest rates are subject to change, thus are random.

### 2.5 Ito Processes with respect to Adapted Processes

We also must define Ito Integrals with respect to Ito Processes. Thus we have the following definition.

**Definition 2.2.** Suppose we have an Ito process $X(t)$ for $t \geq 0$ and an adapted process $\Gamma(t)$ for $t \geq 0$. Then the integral wrt an Ito process is written,

$$\int_0^t \Gamma(u)dX(u) = \int_0^t \Gamma(u)\Delta(u)dW(u) + \int_0^t \Gamma(u)\Theta(u)du.$$

**Proof.** We consider an Itô process

$$X(t) = X(0) + \int_0^t \Delta(u)\,dW(u) + \int_0^t \Theta(u)\,du,$$

and fix $T > 0$. Let $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$ be a partition of $[0, T]$. We study the increment

$$f(T, X(T)) - f(0, X(0)).$$

By telescoping over the partition,

$$f(T, X(T)) - f(0, X(0)) = \sum_{j=0}^{n-1} \big[f(t_{j+1}, X(t_{j+1})) - f(t_j, X(t_j))\big].$$

Applying Taylor's formula in both variables $(t, x)$ at $(t_j, X(t_j))$, we obtain

$$\begin{aligned} f(t_{j+1}, X(t_{j+1})) - f(t_j, X(t_j)) &= f_t(t_j, X(t_j))(t_{j+1} - t_j) \\ &+ f_x(t_j, X(t_j))(X(t_{j+1}) - X(t_j)) \\ &+ \frac{1}{2} f_{xx}(t_j, X(t_j))(X(t_{j+1}) - X(t_j))^2 \\ &+ f_{tx}(t_j, X(t_j))(t_{j+1} - t_j)(X(t_{j+1}) - X(t_j)) \\ &+ \frac{1}{2} f_{tt}(t_j, X(t_j))(t_{j+1} - t_j)^2 \\ &+ \text{higher-order terms.} \end{aligned}$$

Summing from $j = 0$ to $n - 1$ yields

$$\begin{aligned} f(T, X(T)) - f(0, X(0)) &= \sum_{j=0}^{n-1} f_t(t_j, X(t_j))(t_{j+1} - t_j) \\ &+ \sum_{j=0}^{n-1} f_x(t_j, X(t_j))(X(t_{j+1}) - X(t_j)) \\ &+ \frac{1}{2} \sum_{j=0}^{n-1} f_{xx}(t_j, X(t_j))(X(t_{j+1}) - X(t_j))^2 \\ &+ \sum_{j=0}^{n-1} f_{tx}(t_j, X(t_j))(t_{j+1} - t_j)(X(t_{j+1}) - X(t_j)) \\ &+ \frac{1}{2} \sum_{j=0}^{n-1} f_{tt}(t_j, X(t_j))(t_{j+1} - t_j)^2 \\ &+ \text{higher-order terms.} \end{aligned}$$

We now pass to the limit as $\|\Pi\| \to 0$. For the first term,

$$\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} f_t(t_j, X(t_j))(t_{j+1} - t_j) = \int_0^T f_t(t, X(t))\,dt.$$

For the second term, using the definition of the stochastic integral with respect to $X$,

$$\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} f_x(t_j, X(t_j))(X(t_{j+1}) - X(t_j)) = \int_0^T f_x(t, X(t))\, dX(t) = \int_0^T f_x(t, X(t))\Delta(t)\, dW(t) + \int_0^T f_x(t, X(t))\Theta(t)\, dt.$$

For the third term, using Lemma 4.4.4 on quadratic variation,

$$\lim_{\|\Pi\|\to 0} \frac{1}{2} \sum_{j=0}^{n-1} f_{xx}(t_j, X(t_j))(X(t_{j+1}) - X(t_j))^2 = \frac{1}{2}\int_0^T f_{xx}(t, X(t))\, d[X,X](t) = \frac{1}{2}\int_0^T f_{xx}(t, X(t))\Delta^2(t)\, dt.$$

For the mixed term, we estimate

$$\begin{aligned}&\left|\sum_{j=0}^{n-1} f_{tx}(t_j, X(t_j))(t_{j+1} - t_j)(X(t_{j+1}) - X(t_j))\right| \\ &\qquad \le \max_{0\le k\le n-1} |X(t_{k+1}) - X(t_k)| \sum_{j=0}^{n-1} |f_{tx}(t_j, X(t_j))|(t_{j+1} - t_j),\end{aligned}$$

which converges to 0 as $\|\Pi\| \to 0$ since $X$ has continuous paths. Similarly, for the time–time term,

$$\left|\frac{1}{2}\sum_{j=0}^{n-1} f_{tt}(t_j, X(t_j))(t_{j+1} - t_j)^2\right| \le \frac{1}{2} \max_{0\le k\le n-1} (t_{k+1} - t_k) \sum_{j=0}^{n-1} |f_{tt}(t_j, X(t_j))|(t_{j+1} - t_j),$$

which also converges to 0. The higher-order terms vanish as $\|\Pi\| \to 0$. Collecting all limits, we obtain

$$f(T, X(T)) - f(0, X(0)) = \int_0^T f_t(t, X(t))\, dt + \int_0^T f_x(t, X(t))\, dX(t) + \frac{1}{2}\int_0^T f_{xx}(t, X(t))\Delta^2(t)\, dt.$$

### 2.6 Ito-Doeblin for Ito processes

**Theorem 2.6.** *Suppose we have an Ito process $X(t)$ for $t \ge 0$ and a differentiable function $f(t,x)$ with partials $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$ which are defined and continuous. Then for all $T \ge 0$, by Ito-Doeblin we have that,*

$$\begin{aligned} f(T, X(T)) &= f(0, X(0)) + \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))dX(t) + \frac{1}{2}\int_0^T f_{xx}(t, X(t))d[X,X](t) \\ &= f(0, X(0)) + \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))\Delta(t)dW(t) + \int_0^T f_x(t, X(t))\Theta(t)dt + \frac{1}{2}\int_0^T f_{xx}(t, X(t))\Delta^2(t)dt.\end{aligned}$$

*Remark* 2.7. We can simplify Theorem 2.6 by applying differential notation,

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2}f_{xx}(t, X(t))dX(t)dX(t).$$

You may notice that we often write Taylor series expansions of $f(t, X(t))$ where $X(t)$ is an Ito process, with respect to all arguments i.e., $t$ and $X(t)$. The formulation above can be further simplified such that it only involves change wrt time and wrt Brownian motion, since we know that rate at which $X(t)$ (an Ito process) accumulates quadratic variation. Thus,

$$df(t, X(t)) = f_t(t, X(t)) + f_x(t, X(t))\Delta(t)dW(t) + f_x(t, X(t))\Theta(t)dt + \frac{1}{2}f_{xx}(t, X(t))\Delta^2(t)dt.$$

### 2.7 Applications of Ito-Doeblin

*Example* 2.3. Suppose we have Brownian motion $W(t)$ associated with filtration $\mathcal{F}(t)$ for $t \geq 0$. Let $\alpha(t)$ and $\sigma(t)$ be adapted processes. Suppose we have an Ito process defined by,

$$X(t) = \int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right) ds \iff dX(t) = \sigma(t)dW(t) + \left(\alpha(t) - \frac{1}{2}\sigma^2(t)\right) dt$$
$$\iff dX(t)dX(t) = \sigma^2(t)dW(t)dW(t) = \sigma^2(t)dt.$$

Suppose our asset price process is written as, $S(t) = S(0)e^{X(t)}$ where $S(0)$ is deterministic. Thus, $S(t)$ is a function of $X(t)$ so, $f(x) = S(0)e^x, \quad f'(x) = S(0)e^x, \quad f''(x) = S(0)e^x$. Applying Ito-Doeblin gives,

$$\begin{aligned} dS(t) &= df(X(t)) \\ &= f'(X(t))dX(t) + \frac{1}{2}f''(X(t))dX(t)dX(t) \\ &= S(0)e^{X(t)}dX(t) + \frac{1}{2}S(0)e^{X(t)}dX(t)dX(t) \\ &= S(t)dX(t) + \frac{1}{2}S(t)dX(t)dX(t) \\ &= \alpha(t)S(t)dt + \sigma(t)S(t)dW(t). \end{aligned}$$

The mean rate of return $\alpha(t)$ and volatility $\sigma(t)$ are allowed to vary with time and are stochastic in nature. Note, however, in this example we assume that the asset price process is driven by a single Brownian motion. In the case that $\alpha, \sigma$ are constant, $S(t)$ follows a log-normal distribution. Otherwise, $S(t)$ need not be log-normally distributed. In the case of constant mean rate of return and volatility, we yield the Geometric Brownian motion (solution to Black-Scholes-Merton equation).

$$S(t) = S(0)\exp\left\{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right)t\right\}.$$

In order for $S(0)e^{\sigma W(t)}$ to also be a martingale since $e^{\sigma x}$ is convex, we subtract $\frac{1}{2}\sigma^2 t$ in the exponential. Adding $\alpha t$ to the exponential now yields a price process with mean rate of return $\alpha$. In the case that $\alpha = 0$, the $\sigma(t)S(t)dW(t)$ term contributes no drift, just pure vol., to the asset price. Otherwise, the instantaneous mean rate of return would depend on the time and sample path from which it is sampled.

### 2.8 Ito Integral of Deterministic Integrands

**Theorem 2.7.** *Suppose we have Brownian motion $W(s)$ for $s \geq 0$, with a deterministic function of time $\Delta(s)$. Then, the Ito integral $I(t) = \int_0^t \Delta(s)dW(s)$ is a random variable normally distributed with $\mathbb{E}[I(t)] = 0$ and $\mathrm{Var}[I(T)] = \int_0^t \Delta^2(s)ds$.*

**Proof.** $I(t)$ is a martingale so we have that $I(0) = 0$ and thus $\mathbb{E}I(t) = I(0) = 0$. By Ito isometry, we have that,

$$\mathrm{Var}I(t) = \mathbb{E}I^2(t) = \int_0^t \Delta^2(s)ds.$$

$\Delta(s)$ is deterministic so we need not take expectations. Showing $I(t)$ may seem intimidating at first, but recall the *Uniqueness of Moment-generating functions*. We can show that $I(t)$ has the mgf of a normal random variable with mean 0 and variance $\int_0^t \Delta^2(s)ds$. We start with the definition of an mgf,

$$\begin{aligned} \mathbb{E}e^{uI(t)} &= \exp\left\{\frac{1}{2}u^2\int_0^t \Delta^2(s)ds\right\} \\ 1 &= \mathbb{E}\exp\left\{uI(t) - \frac{1}{2}u^2\int_0^t \Delta^2(s)ds\right\} \\ &= \mathbb{E}\exp\left\{\int_0^t u\Delta(s)dW(s) - \frac{1}{2}\int_0^t (u\Delta(s))^2 ds\right\}. \end{aligned}$$

However, notice that $\exp\left\{\int_0^t u\Delta(s)dW(s) - \frac{1}{2}\int_0^t (u\Delta(s))^2ds\right\}$ is a martingale, so we have a generalized geometric Brownian motion with mean rate of return $\alpha = 0$ and $\sigma = u\Delta(s)$. Note that the expectation above in the last line of our derivation will always hold. However, we assumed that $\Delta(s)$ was deterministic in order to derive the mgf.

### 2.9 Vasicek Interest Rate Model

Suppose we have Brownian motion $W(t)$ for $t \geq 0$. The *Vasicek* model for interest rate process $R(t)$ is,

$$dR(t) = (\alpha - \beta R(t))dt + \sigma dW(t).$$

We wish to show that that closed-form solution to the model above is,

$$R(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s).$$

**Solution.** If you are familiar with differential equations then it is trivial to simply write,

$$\begin{aligned} dR(t) &= (\alpha - \beta R(t))dt + \sigma dW(t) \\ &= \alpha dt - \beta R(t)dt + \sigma dW(t) \end{aligned}$$

Re-arranging the formulation above we get that,

$$\begin{aligned} dR(t) + \beta R(t)dt &= \alpha dt + \sigma dW(t) \\ e^{\beta t}dR(t) + \beta e^{\beta t}R(t)dt &= \alpha e^{\beta t}dt + \sigma e^{\beta t}dW(t) \\ d(e^{\beta t}R(t)) &= \alpha e^{\beta t}dt + \sigma e^{\beta t}dW(t) \end{aligned}$$

Taking the integral of the differential we get,

$$\begin{aligned} \int_0^t d(e^{\beta t}R(s))ds &= \int_0^t \alpha e^{\beta s}ds + \int_0^t \sigma e^{\beta s}dW(s) \\ &= e^{\beta t}R(t) - e^0R(0) = e^{\beta t}R(t) - R(0) \end{aligned}$$

Thus this yields,

$$e^{\beta t}R(t) - R(0) = \alpha\int_0^t e^{\beta s}ds + \sigma\int_0^t e^{\beta s}dW(s)$$

Re-arranging we get,

$$\begin{aligned} e^{\beta t}R(t) &= R(0) + \alpha\int_0^t e^{\beta s}ds + \sigma\int_0^t e^{\beta s}dW(s) \\ &= R(0) + \alpha\cdot\frac{1}{\beta}(e^{\beta t} - 1) + \sigma\int_0^t e^{\beta s}dW(s) \\ &= R(0) + \frac{\alpha}{\beta}(e^{\beta t} - 1) + \sigma\int_0^t e^{\beta s}dW(s) \end{aligned}$$

Dividing through by $e^{\beta t}$ we get,

$$\begin{aligned} R(t) &= e^{-\beta t}R(0) + \frac{\alpha}{\beta}e^{-\beta t}(e^{\beta t} - 1) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s) \\ &= e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s), \end{aligned}$$

which is exactly the closed-form solution to the Vasicek model. However, given that we have just learned that Ito-Doeblin yields a very elegant method of deriving a solution to a given SDE, could we somehow apply it to deriving the closed-form solution of the Vasicek model? Let's re-write the original SDE in a standard $dR + \beta R dt = \cdots$ form.

$$dR(t) + \beta R(t)dt = \alpha dt + \sigma dW(t).$$

This is the stochastic analogue of the linear ODE $R'(t) + \beta R(t) = \alpha$. For a deterministic ODE, we multiple by $e^{\beta t}$ because,

$$\frac{d}{dt}\left(e^{\beta t}R(t)\right) = e^{\beta t}R'(t) + \beta e^{\beta t}R(t).$$

Since we want the same cancellation in the SDE, we let $Y(t) := e^{\beta t}R(t)$ and compute $dY(t)$ carefully using the product rule for Ito processes. Moreover, let $A(t) = e^{\beta t}$ (deterministic, finite variation) and let $B(t) = R(t)$. Then we have that,

$$d(AB) = AdB + BdA + dAdB.$$

Clearly, $dA(t) = d(e^{\beta t}) = \beta e^{\beta t}dt$, $dB(t) = dR(t)$ given by the original equation for $dR(t)$. The cross term $dAdB$ is zero because $dA$ is proportional to $dt$ and recall from Ito calculus that $dt \cdot dt = 0$ and $dt \cdot dW(t) = 0$. Since $dA = \beta e^{\beta}dt$, multiplying by anything gives no quadratic variation term so $dAdB = (\beta e^{\beta t}dt) \cdot dR(t) = 0$. Then we have that,

$$d(e^{\beta t}R(t)) = e^{\beta t}dR(t) + R(t)\beta e^{\beta t}dt.$$

From our original equation, after substituting, we get that,

$$\begin{aligned} d(e^{\beta t}R(t)) &= e^{\beta t}\left[(\alpha - \beta R(t))dt + \sigma dW(t)\right] + \beta e^{\beta t}R(t)dt \\ &= e^{\beta t}(\alpha - \beta R(t))dt + \sigma e^{\beta t}dW(t) + \beta e^{\beta t}R(t)dt \\ &= \alpha e^{\beta t}dt - \cancel{\beta e^{\beta t}R(t)dt} + \sigma e^{\beta t}dW(t) + \cancel{\beta e^{\beta t}R(t)dt} \\ &= \alpha e^{\beta t}dt + \sigma e^{\beta t}dW(t). \end{aligned}$$

So we get the simpler SDE,

$$d(e^{\beta t}R(t)) = \alpha e^{\beta t}dt + \sigma e^{\beta t}dW(t).$$

Now, we integrate both sides from 0 to $t$,

$$\begin{aligned} \int_0^t d(e^{\beta s}R(s)) &= \int_0^t \alpha e^{\beta s}ds + \int_0^t \sigma e^{\beta s}dW(s) \\ &= e^{\beta t}R(t) - e^{\beta \cdot 0}R(0) = e^{\beta t}R(t) - R(0). \end{aligned}$$

The LHS is just the increment of the process $e^{\beta s}R(s)$ so we get that,

$$e^{\beta t}R(t) - R(0) = \alpha \int_0^t e^{\beta s}ds + \sigma \cdot \int_0^t e^{\beta s}dW(s)$$

So,

$$e^{\beta t}R(t) = R(0) + \alpha \int_0^t e^{\beta s}ds + \sigma \int_0^t e^{\beta s}dW(s).$$

Note that $\int_0^t e^{\beta s}ds$ is a deterministic integral, so we can simply evaluate. After some basic algebra, we get that $\int_0^t e^{\beta s}ds = \frac{1}{\beta}(e^{\beta t} - 1)$. Plugging this into the equation we derived for $e^{\beta t}R(t)$ we get that,

$$e^{\beta t}R(t) = R(0) + \alpha \cdot \frac{1}{\beta}(e^{\beta t} - 1) + \sigma \int_0^t e^{\beta s}dW(s).$$

Dividing both sides by $e^{\beta t}$ we get that,

$$\begin{aligned} R(t) &= e^{-\beta t}R(0) + \frac{\alpha}{\beta}e^{-\beta t}(e^{\beta t} - 1) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s) \\ &= e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s). \end{aligned}$$

Given we now have a closed-form solution for $R(t)$, define the stochastic integral part as a separate process i.e., $X(t) := \int_0^t e^{\beta s}dW(s)$. Then, $R(t)$ is literally just,

$$R(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma e^{-\beta t}X(t).$$

At this point, if we wanted to verify by Ito, it's natural to view the RHS as a deterministic function of $(t, X(t))$.

$$f(t, x) := e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma e^{-\beta t}x \implies R(t) = f(t, X(t)).$$

Therefore, $f(t, X(t))$ is simply just the explicit solution formula with the random part renamed as $x$.

From our equation for $R(t)$, we can re-write (decompose) it as,

$$R(t) = \underbrace{e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t})}_{\text{deterministic}} + \underbrace{\sigma e^{-\beta t} \int_0^t e^{\beta s} dW(s)}_{\text{random}}.$$

In the "random" component, notice that $\int_0^t e^{\beta s} dW(s)$ is a Gaussian random variable (an Ito integral with deterministic integrand) with mean 0 and variance $\int_0^t (e^{\beta s})^2 ds$, so we can write that,

$$\mathbb{E}\left[\int_0^t e^{\beta s} dW(s)\right] = 0, \quad \text{Var}\left(\int_0^t e^{\beta s} dW(s)\right) = \int_0^t e^{2\beta s} ds.$$

Computing the variance integral explicitly gives us,

$$\int_0^t e^{2\beta s} ds = \left[\frac{1}{2\beta} e^{2\beta s}\right]_{s=0}^{s=t} = \frac{1}{2\beta}(e^{2\beta t} - e^0) = \frac{1}{2\beta}(e^{2\beta t} - 1).$$

Thus, the integral,

$$\int_0^t e^{\beta s} dW(s) \sim N\left(0, \frac{1}{2\beta}(e^{2\beta t} - 1)\right).$$

Taking expectations makes the stochastic integral term drop out, so we can obtain the mean of $R(t)$,

$$\mathbb{E}[R(t)] = e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}).$$

To compute the variance of $R(t)$, the deterministic part has zero variance,

$$\begin{aligned}
\text{Var}(R(t)) &= \text{Var}\left(\sigma e^{-\beta t} \int_0^t e^{\beta s} dW(s)\right) \\
&= \sigma^2 e^{-2\beta t} \text{Var}\left(\int_0^t e^{\beta s} dW(s)\right) \\
&= \sigma^2 e^{-2\beta t} \cdot \frac{1}{2\beta}(e^{2\beta t} - 1) \\
&= \frac{\sigma^2}{2\beta} e^{-2\beta t}(e^{2\beta t} - 1) \\
&= \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}).
\end{aligned}$$

Since the interest rate process $R(t)$ is deterministic + Gaussian, we conclude that it is Gaussian i.e.,

$$R(t) \sim \mathcal{N}\left(e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}), \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})\right).$$

Recall from the equation we derived for $\mathbb{E}[R(t)]$,

$$\mathbb{E}[R(t)] = \frac{\alpha}{\beta} + e^{-\beta t}\left(R(0) - \frac{\alpha}{\beta}\right).$$

So as $\mathbb{E}[R(t)] \to \alpha/\beta$ and $t \to \infty$, the drift in the SDE,

$$\alpha - \beta R(t) = \beta\left(\frac{\alpha}{\beta} - R(t)\right),$$

is positive when $R(t) < \alpha/\beta$ and negative when $R(t) > \alpha/\beta$ pushing the process back toward $\alpha/\beta$. ■

### 2.10 Cox-Ingersoll Interest Rate Model

*Example* 2.4. Consider the Cox-Ingersoll SDE,

$$dR(t) = (\alpha - \beta R(t))dt + \sigma\sqrt{R(t)}dW(t), \quad \alpha, \beta, \sigma > 0.$$

As in the linear Vasicek case, the drift contains a $-\beta R(t)$ term, so we will try the same integrating factor $e^{\beta t}$. Define $X(t) := e^{\beta t}R(t)$ and we will compute $dX(t)$ directly from this definition. Using the product rule for Ito processes with $A(t) = e^{\beta t}$ and $B(t) = R(t)$ we have that,

$$dX(t) = d(A(t)B(t)) = A(t)dB(t) + B(t)dA(t) + dA(t)dB(t).$$

Here we see that $dA(t) = d(e^{\beta t}) = \beta e^{\beta t}dt$ and $dB(t) = dR(t)$. Also, note that $dA(t)dB(t) = (\beta e^{\beta t}dt)dR(t) = 0$ since $dtdt = 0$ and $dtdW(t) = 0$. Therefore, we have now,

$$dX(t) = e^{\beta t}dR(t) + \beta e^{\beta t}R(t)dt.$$

Now, substituting $dR(t) = (\alpha - \beta R(t))dt + \sigma\sqrt{R(t)}dW(t)$,

$$\begin{aligned} dX(t) &= e^{\beta t}\left[(\alpha - \beta R(t))dt + \sigma\sqrt{R(t)}dW(t)\right] + \beta e^{\beta t}R(t)dt \\ &= \alpha e^{\beta t}dt - \beta e^{\beta t}R(t)dt + \sigma e^{\beta t}\sqrt{R(t)}dW(t) + \beta e^{\beta t}R(t)dt \\ &= \alpha e^{\beta t}dt + \sigma e^{\beta t}\sqrt{R(t)}dW(t). \end{aligned}$$

Notice that the $-\beta$ terms cancel by construction, so now we integrate from $0$ to $t$,

$$\int_0^t dX(u) = \int_0^t \alpha e^{\beta u}du + \int_0^t \sigma e^{\beta u}\sqrt{R(u)}dW(u).$$

The LHS is $X(t) - X(0) = e^{\beta t}R(t) - R(0)$ and the deterministic integral is given by,

$$\int_0^t \alpha e^{\beta u}du = \alpha\left[\frac{1}{\beta}e^{\beta u}\right]_0^t = \frac{\alpha}{\beta}\left(e^{\beta t} - 1\right).$$

So now we have the identity,

$$e^{\beta t}R(t) = R(0) + \frac{\alpha}{\beta}(e^{\beta t} - 1) + \sigma\int_0^t e^{\beta u}\sqrt{R(u)}dW(u).$$

Taking expectations gives $e^{\beta t}\mathbb{E}[R(t)] = R(0) + \frac{\alpha}{\beta}(e^{\beta t} - 1)$. Dividing through by $e^{\beta t}$ gives us the expectation,

$$\mathbb{E}[R(t)] = e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t}).$$

Now, the variance is pretty easy to get. Since we know that $X(t) = e^{\beta t}R(t)$ and we have that $R(t) = e^{-\beta t}X(t)$ then we know that $\mathbb{E}[R(t)^2] = e^{-2\beta t}[X(t)^2]$ and $\text{Var}(R(t)) = \mathbb{E}[R(t)^2] - (\mathbb{E}[R(t)])^2$. Given that we already know $dX(t) = \alpha e^{\beta t}dt + \sigma e^{\beta t}\sqrt{R(t)}dW(t)$, we can re-write $\sqrt{R(t)}$ in terms of $X(t)$ i.e.,

$$R(t) = e^{-\beta t}X(t) \implies \sqrt{R(t)} = e^{-\beta t/2}\sqrt{X(t)}.$$

Thus $e^{\beta t}\sqrt{R(t)} = e^{\beta t}e^{-\beta t/2}\sqrt{X(t)} = e^{\beta t/2}\sqrt{X(t)}$ and the SDE for $X$ therefore becomes $dX(t) = \alpha e^{\beta t}dt + \sigma e^{\beta t/2}\sqrt{X(t)}dW(t)$. Next, let's use Ito's Formula to compute $d(X(t)^2)$ i.e., using $g(x) = x^2$. We have that $g'(x) = 2x$, $g''(x) = 2$ so $dg(X) = g'(X)dX + \frac{1}{2}g''(X)(dX)^2$ gives us $d(X(t)^2) = 2X(t)dX(t) + (dX(t))^2$. Now we compute $d(X(t))^2$ given the equation $dX(t)$. Notice that all terms involving $dt$ vanish in quadratic variation leaving only the $dW$ term,

$$(dX(t))^2 = \left(\sigma e^{\beta t/2}\sqrt{X(t)}dW(t)\right)^2 = \sigma^2 e^{\beta t}X(t)(dW(t))^2 = \sigma^2 e^{\beta t}X(t)dt.$$

Plugging in $dX(t)$ and $(dX(t))^2$ into $d(X^2) = 2XdX + (dX)^2$ we get that,

$$\begin{aligned} d(X(t)^2) &= 2X(t)\left(\alpha e^{\beta t}dt + \sigma e^{\beta t/2}\sqrt{X(t)}dW(t)\right) + \sigma^2 e^{\beta t}X(t)dt \\ &= \left(2\alpha e^{\beta t}X(t) + \sigma^2 e^{\beta t}X(t)\right)dt + 2\sigma e^{\beta t/2}X(t)\sqrt{X(t)}dW(t) \\ &= (2\alpha + \sigma^2)e^{\beta t}X(t)dt + 2\sigma e^{\beta t/2}X(t)^{3/2}dW(t). \end{aligned}$$

Now we integrate from 0 to $t$ and since $X(0) = e^0 R(0) = R(0)$ we get that,

$$\begin{aligned} X(t)^2 - X(0)^2 &= (2\alpha + \sigma^2)\int_0^t e^{\beta u} X(u)du + 2\sigma \int_0^t e^{\beta u/2} X(u)^{3/2} dW(u) \\ X(t)^2 &= R(0)^2 + (2\alpha + \sigma^2)\int_0^t e^{\beta u} X(u)du + 2\sigma \int_0^t e^{\beta u/2} X(u)^{3/2} dW(u) \\ \mathbb{E}[X(t)^2] &= R(0)^2 + (2\alpha + \sigma^2)\int_0^t e^{\beta u}\mathbb{E}[X(u)]du. \end{aligned}$$

Since the Ito integral has mean 0 we get the expectation above. We also need $\mathbb{E}[X(u)]$ but $X(u) = e^{\beta u}R(u)$ so $\mathbb{E}[X(u)] = e^{\beta u}\mathbb{E}[R(u)]$. We must use the formula for $\mathbb{E}[R(u)]$ to compute $\mathbb{E}[X(u)]$, which gives us,

$$\mathbb{E}[X(u)] = e^{\beta u}\mathbb{E}[R(u)] = e^{\beta u}\left(e^{-\beta u}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta u})\right) = R(0) + \frac{\alpha}{\beta}(e^{\beta u} - 1).$$

Substituting $\mathbb{E}[X(u)]$ into the integral for $\mathbb{E}[X(t)^2]$ and expanding we get,

$$\begin{aligned} \mathbb{E}[X(t)^2] &= R(0)^2 + (2\alpha + \sigma^2)\int_0^t e^{\beta u}\left(R(0) + \frac{\alpha}{\beta}(e^{\beta u} - 1)\right) du \\ &= R(0)^2 + (2\alpha + \sigma^2)\left[\int_0^t R(0)e^{\beta u}du + \frac{\alpha}{\beta}\int_0^t (e^{2\beta u} - e^{\beta u})du\right] \\ &= R(0)^2 + (2\alpha + \sigma^2)\left[\frac{R(0)}{\beta}(e^{\beta t} - 1) + \frac{\alpha}{\beta}\left(\frac{1}{2\beta}(e^{2\beta t} - 1) - \frac{1}{\beta}(e^{\beta t} - 1)\right)\right] \\ &= R(0)^2 + (2\alpha + \sigma^2)\left[\frac{R(0)}{\beta}(e^{\beta t} - 1) - \frac{\alpha}{\beta^2}(e^{\beta t} - 1) + \frac{\alpha}{2\beta^2}(e^{2\beta t} - 1)\right] \\ &= R(0)^2 + (2\alpha + \sigma^2)\left[\frac{1}{\beta}\left(R(0) - \frac{\alpha}{\beta}\right)(e^{\beta t} - 1) + \frac{\alpha}{2\beta^2}(e^{2\beta t} - 1)\right]. \end{aligned}$$

Converting this equation back to $\mathbb{E}[R(t)^2]$, since we know that $R(t) = e^{-\beta t}X(t)$ and $R(t)^2 = e^{-2\beta t}X(t)^2$, we then have that $\mathbb{E}[R(t)^2] = e^{-2\beta t}\mathbb{E}[X(t)^2]$. So, multiplying the expression we just derived for $\mathbb{E}[X(t)^2]$ above by $e^{-2\beta t}$ we get that,

$$\mathbb{E}[R(t)^2] = e^{-2\beta t}R(0)^2 + (2\alpha + \sigma^2)e^{-2\beta t}\cdot\frac{1}{\beta}\left(R(0) - \frac{\alpha}{\beta}\right)(e^{\beta t} - 1) + (2\alpha + \sigma^2)e^{-2\beta t}\cdot\frac{\alpha}{2\beta^2}(e^{2\beta t} - 1).$$

Given that we have $\mathbb{E}[R(t)^2]$ now, solving for $\mathrm{Var}(R(t))$ is purely algebraic now,

$$\begin{aligned} \mathrm{Var}(R(t)) &= \mathbb{E}[R(t)^2] - (\mathbb{E}[R(t)])^2 \\ &= e^{-2\beta t}\mathbb{E}[X(t)^2] - \left(e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t})\right)^2 \\ &= e^{-2\beta t}\left(R(0)^2 + \frac{2\alpha + \sigma^2}{\beta}\left(R(0) - \frac{\alpha}{\beta}\right)(e^{\beta t} - 1) + \frac{\alpha(2\alpha + \sigma^2)}{2\beta^2}(e^{2\beta t} - 1)\right) \\ &\quad - \left(e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t})\right)^2 \\ &= e^{-2\beta t}R(0)^2 + \frac{2\alpha + \sigma^2}{\beta}\left(R(0) - \frac{\alpha}{\beta}\right)(e^{-\beta t} - e^{-2\beta t}) \\ &\quad + \frac{\alpha(2\alpha + \sigma^2)}{2\beta^2}(1 - e^{-2\beta t}) - \left(e^{-\beta t}R(0) + \frac{\alpha}{\beta}(1 - e^{-\beta t})\right)^2 \\ &= e^{-2\beta t}R(0)^2 + \frac{2\alpha + \sigma^2}{\beta}\left(R(0) - \frac{\alpha}{\beta}\right)(e^{-\beta t} - e^{-2\beta t}) \\ &\quad + \frac{\alpha(2\alpha + \sigma^2)}{2\beta^2}(1 - e^{-2\beta t}) \\ &\quad - \left(e^{-2\beta t}R(0)^2 + 2\frac{\alpha}{\beta}e^{-\beta t}R(0)(1 - e^{-\beta t}) + \frac{\alpha^2}{\beta^2}(1 - e^{-\beta t})^2\right) \\ &= \frac{2\alpha + \sigma^2}{\beta}\left(R(0) - \frac{\alpha}{\beta}\right)(e^{-\beta t} - e^{-2\beta t}) + \frac{\alpha(2\alpha + \sigma^2)}{2\beta^2}(1 - e^{-2\beta t}) - 2\frac{\alpha}{\beta}R(0)(e^{-\beta t} - e^{-2\beta t}) - \frac{\alpha^2}{\beta^2}(1 - 2e^{-\beta t} + e^{-2\beta t}) \end{aligned}$$

$$
\begin{aligned}
&= \frac{2\alpha+\sigma^2}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) - \frac{\alpha(2\alpha+\sigma^2)}{\beta^2}(e^{-\beta t}-e^{-2\beta t}) \\
&\quad + \frac{\alpha(2\alpha+\sigma^2)}{2\beta^2}(1-e^{-2\beta t}) - 2\frac{\alpha}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) \\
&\quad - \frac{\alpha^2}{\beta^2}(1-2e^{-\beta t}+e^{-2\beta t}) \\
&= \left(\frac{2\alpha+\sigma^2}{\beta} - \frac{2\alpha}{\beta}\right)R(0)(e^{-\beta t}-e^{-2\beta t}) \\
&\quad - \frac{\alpha(2\alpha+\sigma^2)}{\beta^2}(e^{-\beta t}-e^{-2\beta t}) + \frac{\alpha(2\alpha+\sigma^2)}{2\beta^2}(1-e^{-2\beta t}) \\
&\quad - \frac{\alpha^2}{\beta^2}(1-2e^{-\beta t}+e^{-2\beta t}) \\
&= \frac{\sigma^2}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) \\
&\quad + \frac{\alpha}{2\beta^2}\left[-2(2\alpha+\sigma^2)(e^{-\beta t}-e^{-2\beta t}) + (2\alpha+\sigma^2)(1-e^{-2\beta t}) - 2\alpha(1-2e^{-\beta t}+e^{-2\beta t})\right] \\
&= \frac{\sigma^2}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) \\
&\quad + \frac{\alpha}{2\beta^2}\left[(2\alpha+\sigma^2) - (2\alpha+\sigma^2)e^{-2\beta t} - 2(2\alpha+\sigma^2)e^{-\beta t} + 2(2\alpha+\sigma^2)e^{-2\beta t}\right. \\
&\qquad\qquad \left. -2\alpha + 4\alpha e^{-\beta t} - 2\alpha e^{-2\beta t}\right] \\
&= \frac{\sigma^2}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) \\
&\quad + \frac{\alpha}{2\beta^2}\left[\sigma^2 - 2\sigma^2 e^{-\beta t} + \sigma^2 e^{-2\beta t}\right] \\
&= \frac{\sigma^2}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) + \frac{\alpha\sigma^2}{2\beta^2}(1-2e^{-\beta t}+e^{-2\beta t}) \\
&= \frac{\sigma^2}{\beta}R(0)(e^{-\beta t}-e^{-2\beta t}) + \frac{\alpha\sigma^2}{2\beta^2}(1-e^{-\beta t})^2.
\end{aligned}
$$

Thus, we see that,

$$\lim_{t\to\infty} \operatorname{Var}(R(t)) = \frac{\alpha\sigma^2}{2\beta^2}.$$

### 2.11 Black-Scholes-Merton Equation

#### 2.11.1 Evolution of Portfolio Value

Suppose we have the stock process,

$$dS(t) = \alpha S(t)dt + \sigma S(t)dW(t).$$

At time $t$ suppose we hold $\Delta(t)$ shares of stock with each share being worth $S(t)$ making our total value invested in the stock $\Delta(t)S(t)$ at time $t$. Assume that $X(t)$ represents our portfolio at time $t$ with any remaining value $X(t)-\Delta(t)S(t)$ being invested into a money-market account. Suppose the money market account grows at constant rate $r$ (interest rate). Then, over a small time interval $dt$, the interest earned on our cash position in the money market is $r(X(t)-\Delta(t)S(t))dt$.

Simultaneously, our position in the stock gains/loses an amount equal to $\Delta(t)dS(t)$ i.e., the number of shares we hold at time $d$ multiplied by the change in the underlying stock price. Thus, the total change in our portfolio value is,

$$dX(t) = \Delta(t)dS(t) + r(X(t)-\Delta(t)S(t))dt.$$

Note that in our stock price process, $\alpha$ represents the instantaneous expected rate of return of the stock and $\sigma$ is our volatility i.e., if $\alpha = 0.08$ then the stock's expected annual growth rate is $8\%$. The randomness around that growht is comes from $\sigma S(t)dW(t)$ i.e., the term driven by Brownian motion.

To simplify this formula, we can substitute the stock SDE into the formula for $dX(t)$ and simplify.

$$\begin{aligned} dX(t) &= \Delta(t)dX(t) + r(X(t) - \Delta S(t))dt \\ &= \Delta(t)(\alpha S(t)dt + \sigma S(t)dW(t)) + r(X(t) - \Delta(t)S(t))dt \\ &= \Delta(t)\alpha S(t)dt + \Delta(t)\sigma S(t)dW(t) + r(X(t) - \Delta(t)S(t))dt \\ &= \Delta(t)\alpha S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt - r\Delta(t)S(t)dt \\ &= \Delta(t)\alpha S(t)dt - r\Delta(t)S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt \\ &= (\alpha - r)\Delta(t)S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt \\ &= rX(t)dt + \Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t). \end{aligned}$$

*Remark* 2.8. $rX(t)dt$ represents the earnings of the portfolio if all wealth simply grew at the risk-free rate. $\Delta(t)(\alpha - r)S(t)dt$ is the excess expected return from holding stock instead of cash. $\alpha - r$ is the stock's drift above the money-market rate, so multiplying by the dollar amount invested in stock, $\Delta(t)S(t)$ gives the extra drift contribution from the risky investment. $\Delta(t)\sigma S(t)dW(t)$ is the random fluctuation coming from stock-price uncertainty and is proportional to the size of the stock position $\Delta(t)S(t)$.

### 2.11.2 Discounting

We discount to remove the baseline growth rate $r$ and start by studying the discounted stock price $e^{-rt}S(t)$. Computing this differential is precise under Ito-Doeblin so we let $f(t,x) = e^{-rt}x$ which means that $e^{rt}S(t) = f(t, S(t))$.

We'll use Ito's formula and start by computing the partials,

$$f_t(t,x) = \frac{\partial}{\partial x}(e^{-rt}x) = -re^{-rt}x, \quad f_x(t,x) = \frac{\partial}{\partial x}(e^{-rt}x) = e^{-rt}, \quad f_{xx}(t,x) = 0.$$

Applying Ito's formula gives us,

$$d(f(t,S(t))) = f_t(t,S(t))dt + f_x(t,S(t))dS(t) + \frac{1}{2}f_{xx}(t,S(t))dS(t)dS(t).$$

Since $f_{xx} = 0$ the quadratic variation term vanishes, thus we are left with,

$$d(e^{-rt}S(t)) = -re^{-rt}S(t)dt + e^{-rt}dS(t).$$

Now, we substituting and simplify,

$$\begin{aligned} d(e^{-rt}S(t)) &= -re^{-rt}S(t)dt + e^{-rt}(\alpha S(t)dt + \sigma S(t)dW(t)) \\ &= -re^{-rt}S(t)dt + \alpha e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t) \\ &= (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t). \end{aligned}$$

We have derived the discounted stock-price dynamic. Equivalently, we can do the same for the discounted portfolio value $e^{-rt}X(t)$. Again, let's define $g(t,x) = e^{-rt}x$ which means that $e^{-rt}X(t) = g(t, X(t))$. Taking partial derivatives again we get,

$$g_t(t,x) = -re^{-rt}x, \quad g_x(t,x) = e^{-rt}, \quad g_{xx}(t,x) = 0.$$

Substituting into our formula for $dX(t)$ we get that,

$$\begin{aligned} d(e^{-rt}X(t)) &= -re^{-rt}X(t)dt + e^{-rt}\left(rX(t)dt + \Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t)\right) \\ &= -re^{-rt}X(t)dt + re^{-rt}X(t)dt + \Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t) \\ &= \Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t). \end{aligned}$$

Notice that if we $\Delta(t)d(e^{-rt}S(t))$, we get exactly $d(e^{-rt}X(t))$, thus we can write that,

$$d(e^{-rt}X(t)) = \Delta(t)d(e^{-rt}S(t)).$$

This means that once everything is measured in units of the money market account, the only source of change in portfolio value is the change in the discounted stock price. The cash account no longer explicit appears because discounting has already removed its deterministic accumulation.

### 2.11.3 Evolution of Option Value

Recall our stock-price model $dS(t) = \alpha S(t)dt + \sigma S(t)dW(t)$. Moreover, suppose that at time $t$, an option's price is a deterministic function of time $t$ and $x$ (stock price) i.e., $c(t, x)$. If at time $t$ the stock price happens to be equal to $x$, then the option value is $c(t, x)$. Since the actual stock price at time $t$ is $S(t)$ which is a random variable, then the option value process can be defined as $c(t, S(t))$. While $c(t, x)$ is not random, $c(t, S(t))$ is random because $S(t)$ is random.

### 2.11.4 Differential of Option-Value Process

Because $c$ depends on $t, x$ and because $x := S(t)$, we will apply Ito's formula.

For a sufficiently smooth function $c(t, x)$ Ito says,

$$\begin{aligned}
dc(t, S(t)) &= c_t(t, S(t))dt + c_x(t, S(t))dS(t) + \frac{1}{2}c_{xx}(t, S(t))dS(t)dS(t) \\
&= c_t(t, S(t))dt + c_x(t, S(t))(\alpha S(t)dt + \sigma S(t)dW(t)) + \frac{1}{2}c_{xx}(t, S(t))dS(t)dS(t).
\end{aligned}$$

As we stated before,

$$dS(t)dS(t) = (\alpha S(t)dt + \sigma S(t)dW(t))^2 = \alpha^2 S(t)^2 dt^2 + 2\alpha\sigma S(t)^2 dt dW(t) + \sigma^2 S(t)^2 dW(t)^2.$$

Because $dt^2 = dtdW(t) = 0$ and $dW(t)^2 = dt$, we see that the first two terms vanish and only,

$$dS(t)dS(t) = \sigma^2 S(t)^2 dt,$$

remains. Thus, substituting this into our expanded formulation for $dc(t, S(t))$ we get that,

$$\begin{aligned}
dc(t, S(t)) &= c_t(t, S(t))dt + c_x(t, S(t))\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S(t)^2 dt \\
&= c_t(t, S(t))\, dt + \alpha S(t)c_x(t, S(t))\, dt + \sigma S(t)c_x(t, S(t))\, dW(t) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t))\, dt \\
&= \left[c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t))\right] dt + \sigma S(t)c_x(t, S(t))dW(t).
\end{aligned}$$

Now, let's derive the discounted option-price dynamics i.e., compute $d(e^{-rt}c(t, S(t))$.

We will apply Ito to the function $f(t, x) = e^{-rt}x$ starting with computing the partial derivative of $f$,

$$f_t(t, x) = \frac{\partial}{\partial t}(e^{-rt}x) = -re^{-rt}x, \quad f_x(t, x) = \frac{\partial}{\partial x}(e^{-rt}x) = e^{-rt}, \quad f_{xx}(t, x) = 0.$$

Applying Ito's formula to $f(t, c(t, S(t)))$ and substituting the partials gives us,

$$\begin{aligned}
d\big(e^{-rt}c(t, S(t))\big) &= f_t(t, c(t, S(t)))\ dt + f_x(t, c(t, S(t)))\ dc(t, S(t)) + \frac{1}{2}f_{xx}(t, c(t, S(t)))\ dc(t, S(t))\, dc(t, S(t)) \\
&= f_t(t, c(t, S(t)))\ dt + f_x(t, c(t, S(t)))\ dc(t, S(t)) \\
&= -re^{-rt}c(t, S(t))\, dt + e^{-rt}\, dc(t, S(t)) \\
&= -re^{-rt}c(t, S(t))\, dt \\
&\quad + e^{-rt}\left(\left[c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t))\right] dt + \sigma S(t)c_x(t, S(t))\, dW(t)\right) \\
&= -re^{-rt}c(t, S(t))\, dt + e^{-rt}\left[c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t))\right] dt \\
&\quad + e^{-rt}\sigma S(t)c_x(t, S(t))\, dW(t) \\
&= e^{-rt}\left[-r\, c(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t))\right] dt \\
&\quad + e^{-rt}\sigma S(t)c_x(t, S(t))\, dW(t).
\end{aligned}$$

In the 2nd line, $f_{xx} = 0$ so the last term disappears.

*Remark* 2.9. The option value is a function of $t$ and $S(t)$ so it changes because of the explicit passage of time, represented by $c_t dt$ and movement in the stock price, represented by $c_x dS$ and $\frac{1}{2}c_{xx}(dS)^2$ terms. After discounting, an extra $-rc$ term appears because multiplying by $e^{-rt}$ removes the risk-free growth rate from the value process.

### 2.11.5 Black-Scholes-Merton PDE

Recall that we have the following three discounted equations,

$$\begin{aligned} d(e^{-rt}X(t)) &= \Delta(t)d(e^{-rt}S(t)), \\ d(e^{-rt}S(t)) &= (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t), \\ d(e^{-rt}X(t)) &= \Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t). \end{aligned}$$

Moreover, we have computed the discounted option-value as,

$$d(e^{-rt}c(t,S(t))) = e^{-rt}\left[-rc(t,S(t)) + c_t(t,S(t)) + \alpha S(t)c_x(t,S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t,S(t))\right]dt + e^{-rt}\sigma S(t)c_x(t,S(t))dW(t).$$

We aim to devise a trading strategy $\Delta(t)$ such that the portfolio replicates the option i.e., $X(t)$ equals the option-value process $c(t,S(t))$ at every timestep $t \in [0,T]$. Recall that we get this if $e^{-rt}X(t) = e^{-rt}c(t,S(t))$ for all $t$ which is equivalent to saying $X(t) = c(t,S(t))$. The rationale for comparing discounted processes rather than original ones is because discounting removes the trivial risk-free accumulation. After discounting, the only thing left is the true risky evolution.

To ensure equality, we require that,

$$d(e^{-rt}X(t)) = d(e^{-rt}c(t,S(t))), \quad \text{for all } t \in [0,T].$$

We also trivially impose that $X(0) = c(0,S(0))$. If two continuous semi-martingales have the same initial value and the same differential at every time $t$, then their increments over every time interval are the same. Thus, integrating from 0 to $t$ gives us,

$$\int_0^t d(e^{-ru}X(u)) = \int_0^t d(e^{-ru}c(u,S(u))) \iff e^{-rt}X(t) - X(0) = e^{-rt}c(t,S(t))) - c(0,S(0)).$$

Given that $X(0) = c(0,S(0))$ we can cancel these terms and get $e^{-rt}X(t) = e^{-rt}c(t,S(t))$ which means that $X(t) = c(t,S(t))$. Matching the discounted differentials and the initial value is exactly what makes the portfolio replicate the option.

Equating the discounted differential formulas gives us,

$$\begin{aligned} &\Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t) \\ &\quad = e^{-rt}\left[-rc(t,S(t)) + c_t(t,S(t)) + \alpha S(t)c_x(t,S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t,S(t))\right]dt + e^{-rt}\sigma S(t)c_x(t,S(t))dW(t). \end{aligned}$$

Since $e^{-rt} > 0$ dividing both sides by $e^{-rt}$ gives us,

$$\begin{aligned} &\Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) \\ &\quad = \left[-rc(t,S(t)) + c_t(t,S(t)) + \alpha S(t)c_x(t,S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t,S(t))\right]dt + \sigma S(t)c_x(t,S(t))dW(t). \end{aligned}$$

Notice that both the right and left sides of the equality have one $dt$ term and one $dW(t)$ term. For the differentials to be equal for all $t$, the $dW(t)$ and $dt$ coefficients must match. Fundamentally, $dW(t)$ and $dt$ are different objects. $dt$ represents predictable finite-variation drift whilst $dW(t)$ represents martingale noise. A non-zero Brownian shock cannot be canceled by a drift term and vice versa. On the LHS, the $dW(t)$ coefficient is,

$$\Delta(t)\sigma S(t),$$

and on the RHS, the $dW(t)$ coefficient is,

$$\sigma S(t)c_x(tS(t)).$$

Equating the two sides gives us,

$$\Delta(t)\sigma S(t) = \sigma S(t)c_x(t,S(t)).$$

Assuming the $\sigma, S(t) > 0$, dividing both sides by $\sigma S(t)$ tells us that $\Delta(t) = c_x(t,S(t))$. $c_x(t,S(t))$ is also known as the delta of an option, so we have effectively identified the **delta-hedging rule**. The quantity $c_x(t,S(t))$ is the derivative of the option value wrt the stock price evaluated at the current stock price measuring local sensitivity of the option price to a small change in the stock i.e., $c(t, S(t)+h) - c(t, S(t)) \approx c_x(t, S(t)) \cdot h$. Thus, if the stock moves by a small amount, $dS(t)$, then the option value changes approximately by $c_x(t, S(t))dS(t)$. If our portfolio holds $\Delta(t)$ shares of stock, then its stock component changes by $\Delta(t)S(t)$. In order to make the random Brownian part of the portfolio move exactly like the option, we therefore choose $\Delta(t) = c_x(t, S(t))$.

Moreover, $\Delta(t)(\alpha - r)S(t)$ is the $dt$ coefficient on the LHS and the RHS yields,

$$-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t)).$$

Equating the $dt$ drift terms gives,

$$\Delta(t)(\alpha - r)S(t) = -rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t)).$$

Substituting $\Delta(t) = c_x(t, S(t))$ we get the LHS to be $c_x(t, S(t))(\alpha - r)S(t)$ so our equality above becomes,

$$\begin{aligned}(\alpha - r)S(t)c_x(t, S(t)) &= -rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t)) \\ \alpha S(t)c_x(t, S(t)) - rS(t)c_x(t, S(t)) &= -rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t)).\end{aligned}$$

Adding $rc(t, S(t))$ to both sides we get,

$$\begin{aligned}-rS(t)c_x(t, S(t)) &= -rc(t, S(t)) + c_t(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t)) \\ rc(t, S(t)) &= c_t(t, S(t)) + rS(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S(t)^2 c_{xx}(t, S(t)).\end{aligned}$$

Since this must hold for every $t$ and along every sampled stock path, the natural conclusion is that the deterministic function $c(t, x)$ must satisfy for all $t \in [0, T]$ and $x \geq 0$,

$$c_t(t, x) + rxc_x(t, x) + \frac{1}{2}\sigma^2 x^2 c_{xx}(t, x) = rc(t, x).$$

*Remark* 2.10. At a first glance, since the stock itself evolves under physical drift $\alpha$ we may be led to believe that the option price should depend on it as well. However, after constructing a self-financing portfolio that replicates the option's randomness, the random risk is hedged away. After matching the $dW(t)$ terms with $\Delta(t) = c_x(t, S(t))$, the only remaining condition comes from the deterministic drift of a locally riskless position which must earn risk-free rate $r$ and not the stock's expected return $\alpha$.

*The option price is determined by replication and absence of arbitrage, not by the stock's physical expected return.*

### 2.11.6 Terminal Condition of BS-Merton PDE

A European call has payoff $(S(T) - K)^+$ so the option value at expiration must equal that payoff $c(T, x) = (x - K)^+$. Equivalently, along the realized stock path, $c(T, S(T)) = (S(T) - K)^+$. So, our option pricing problem is to find a smooth function $c(t, x)$ satisfying,

$$c_t(t, x) + rxc_x(t, x) + \frac{1}{2}\sigma^2 x^2 c_{xx}(t, x) = rc(t, x),$$

for $0 \leq t < T$ and $x \geq 0$ together with $c(T, x) = (x - K)^+$.

Assume that we found such $c(t, x)$. Consider a portfolio that starts with initial capital $X(0) = c(0, S(0))$ and at each time holds $\Delta(t) = c_x(t, S(t))$ shares of stock. By construction, the $dW(t)$ terms (as we saw before) match. Moreover, since $c$ satisfies the BS PDE and the $dt$ terms match, we have that,

$$d(e^{-rt}X(t)) = d(e^{-rt}c(t, S(t))).$$

Integrating from 0 to $t$,

$$e^{-rt}X(t) - X(0) = e^{-rt}c(t, S(t)) - c(0, S(0)).$$

Multiplying both sides by $e^{rt}$ gives $X(t) = c(t, S(t))$ for all $t \in [0, T]$. Specifically, if we take $t = T$ (terminal time $T$), then $X(T) = c(T, S(T)) = (S(T) - K)^+$ i.e., a short position in the option can be perfectly hedged.

If you sell the option at time 0 and use the proceeds to set up the replicating strategy, then no matter what stock path occurs, the portfolio value at maturity will equal the amount owed on the option.

*Remark* 2.11. My solution to **Exercise 1.8.5** provides a probabilistic solution to the Black-Scholes-Merton PDE.

**Theorem 2.8** (Black–Scholes–Merton Pricing Formula)**.** *Let $(S_t)_{t\geq 0}$ denote the price of a non–dividend paying asset following the geometric Brownian motion*

$$dS_t = rS_t\,dt + \sigma S_t\,dW_t$$

*under the risk–neutral measure $\mathbb{Q}$, where $r$ is the constant risk–free interest rate, $\sigma > 0$ is the constant volatility, and $(W_t)$ is a standard Brownian motion. Consider a European call option with strike $K$ and maturity $T$, whose payoff is*

$$(S_T - K)^+.$$

*Then the arbitrage–free price of the option at time $t$ is given by*

$$C(t, S_t) = S_t N(d_1) - Ke^{-r(T-t)}N(d_2),$$

*where*

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}.$$

*Equivalently, this price satisfies the risk–neutral valuation identity*

$$C(t, S_t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S_T - K)^+ \,\middle|\, S_t\right].$$

*Moreover, $C(t, S)$ is the unique classical solution to the Black–Scholes partial differential equation*

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0,$$

*subject to the terminal condition*

$$C(T, S) = (S - K)^+.$$

### 2.11.7 The Greeks

For a European Call option, the Black-Scholes price can be written as $c(t, x) = xN(d_+) - Ke^{-r(T-t)}N(d_-)$ where,

$$d_\pm = d_\pm(T-t, x) = \frac{1}{\sigma\sqrt{T-t}}\left[\ln(x/K) - \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)\right].$$

Moreover, we can write $d_- = d_+ - \sigma\sqrt{T-t}$. Note that here $N$ is the standard normal cdf and $N'$ is it's respective density function $N'(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$. The greeks are simply differentiating the formulation above with respect to $x, t$, and then interpreting the result geometrically and financially.

**Delta.** Since $c(t, x) = xN(d_+) - Ke^{-r(T-t)}N(d_-)$, to get the delta we differentiate wrt $x$.

$$c_x(t, x) = \frac{\partial}{\partial x}(xN(d_+)) - \frac{\partial}{\partial x}(Ke^{-r(T-t)}N(d_-)).$$

Since $Ke^{-r(T-t)}$ does not depend on $x$, we have that,

$$c_x(t, x) = N(d_+) + xN'(d_+)\frac{\partial d_+}{\partial x} - Ke^{-r(T-t)}N'(d_-)\frac{\partial d_-}{\partial x}.$$

We now only need the derivatives of $d_\pm$ with respect to $x$. Notice that the only $x$-dependence is $\ln(x/K)$ whose derivative is $1/x$. Thus we have that,

$$\frac{\partial d_+}{\partial x} = \frac{1/x}{\sigma\sqrt{T-t}} - \frac{1}{x\sigma\sqrt{T-t}}, \qquad \frac{\partial d_-}{\partial x} = \frac{1}{x\sigma\sqrt{T-t}}.$$

We can substitute these into the formula for $c_x$ which gives us,

$$c_x(t, x) = N(d_+) + xN'(d_+)\frac{1}{x\sigma\sqrt{T-t}} - Ke^{-r(T-t)}N'(d_-)\frac{1}{x\sigma\sqrt{T-t}} = N(d_+) + \frac{N'(d_+)}{\sigma\sqrt{T-t}} - \frac{Ke^{-r(T-t)}N'(d_-)}{x\sigma\sqrt{T-t}}.$$

To get that $c_x(t,x) = N(d_+)$ we need two extra terms to cancel, that is we need,

$$\frac{N'(d_+)}{\sigma\sqrt{T-t}} = \frac{Ke^{-r(T-t)}N'(d_-)}{x\sigma\sqrt{T-t}} \iff xN'(d_+) = Ke^{-r(T-t)}N'(d_-).$$

We know that $N'(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$ so logically we have that,

$$\frac{N'(d_+)}{N'(d_-)} = \exp\left\{-\frac{1}{2}d_+^2 + \frac{1}{2}d_-^2\right\} = \exp\left\{\frac{1}{2}(d_-^2 - d_+^2)\right\}.$$

We can use the fact that $d_- = d_+ - \sigma\sqrt{T-t}$ so then we have that $d_-^2 - d_+^2 = (d_- - d_+)(d_- + d_+)$ and since $d_- - d_+ = -\sigma\sqrt{T-t}$ and $d_- + d_+ = \frac{1}{\sigma\sqrt{T-t}}2\ln(x/K) + 2r(T-t)$, note that the $+\frac{1}{2}\sigma^2(T-t)$ and $-\frac{1}{2}\sigma^2(T-t)$ terms cancel, we therefore obtain,

$$d_-^2 - d_+^2 = -\sigma\sqrt{T-t}\cdot\frac{2\ln(x/K) + 2r(T-t)}{\sigma\sqrt{T-t}} = -2\ln(x/K) - 2r(T-t).$$

Thus we have that $\frac{1}{2}(d_-^2 - d_+^2) = -\ln(x/K) - r(T-t)$ so $N'(d_+)/N'(d_-) = \frac{K}{x}e^{-r(T-t)}$.

If we simply multiply both sides by $xN'(d_-)$ we get,

$$xN'(d_+) = Ke^{-r(T-t)}N'(d_-).$$

Hence the extra terms cancel which gives us $c_x(t,x) = N(d_+)$. This is the delta and is always positive since $N(d_+) \in (0,1)$.

For a call option, this makes sense since as the stock price rises, the call becomes more valuable so it's derivative wrt stock price should be positive. How about the 2nd derivative?

**Gamma.** The Gamma is the second derivative wrt stock price. Since we saw that $c_x(t,x) = N(d_+)$, we can differentiate once more wrt $x$ to get $c_{xx}(t,x) = N'(d_+)\frac{\partial d_+}{\partial x}$. We had already computed $\frac{\partial d_+}{\partial x} = \frac{1}{x\sigma\sqrt{T-t}}$. Thus, differentiating once more wrt $x$ gives,

$$c_{xx}(t,x) = N'(d_+)\frac{\partial d_+}{\partial x}.$$

Since we already have what $\frac{\partial d_+}{\partial x}$ is, then we have that,

$$c_{xx}(t,x) = \frac{1}{x\sigma\sqrt{T-t}}N'(d_+).$$

Note that because of this, Gamma is always positive because $x > 0, \sigma > 0$ and $T - t > 0$ and $N'(d_+) > 0$. Thus the call price is convex as a function of stock price. A positive 2nd derivative means that the graph of $x \mapsto c(t,x)$ bends upwards. Thus, the tangent line at a point lies below the graph away from the tangency point. Thus, if we match the first order stock sensitivity by shorting delta shares, the residual position benefits from movement in either direction i.e., hence being long gamma.

**Theta.** I have reduced the derivation for theta but it follows the same principles.

$$\begin{aligned}-xN'(d_+)\frac{\partial d_+}{\partial\tau} + Ke^{-r\tau}N'(d_-)\frac{\partial d_-}{\partial\tau} &= -x\left(\frac{1}{\sqrt{2\pi}}e^{-d_+^2/2}\right)\frac{\partial d_+}{\partial\tau} + Ke^{-r\tau}\left(\frac{1}{\sqrt{2\pi}}e^{-d_-^2/2}\right)\frac{\partial d_-}{\partial\tau} \\ &= -\frac{x}{\sqrt{2\pi}}e^{-d_+^2/2}\frac{\partial d_+}{\partial\tau} + \frac{Ke^{-r\tau}}{\sqrt{2\pi}}e^{-d_-^2/2}\frac{\partial d_-}{\partial\tau}.\end{aligned}$$

And we separately simplify the second coefficient,

$$\begin{aligned}Ke^{-r\tau}e^{-d_-^2/2} &= K\exp\left(-r\tau - \frac{d_-^2}{2}\right) = K\exp\left(-r\tau - \frac{(d_+ - \sigma\sqrt{\tau})^2}{2}\right) \\ &= K\exp\left(-r\tau - \frac{1}{2}\left(d_+^2 - 2\sigma\sqrt{\tau}\,d_+ + \sigma^2\tau\right)\right) \\ &= K\exp\left(-r\tau - \frac{d_+^2}{2} + \sigma\sqrt{\tau}\,d_+ - \frac{\sigma^2\tau}{2}\right).\end{aligned}$$

Computing the exponent term $\sigma\sqrt{\tau}d_+$ explicitly gives us,

$$\begin{aligned}\sigma\sqrt{\tau}\,d_+ &= \sigma\sqrt{\tau}\left(\frac{\ln(x/K)+(r+\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\right)\\ &= \ln(x/K)+(r+\tfrac{1}{2}\sigma^2)\tau.\end{aligned}$$

Substituting this exponent term back gives us,

$$\begin{aligned}-r\tau-\frac{d_+^2}{2}+\sigma\sqrt{\tau}\,d_+-\frac{\sigma^2\tau}{2} &= -r\tau-\frac{d_+^2}{2}+\ln(x/K)+(r+\tfrac{1}{2}\sigma^2)\tau-\frac{\sigma^2\tau}{2}\\ &= -r\tau-\frac{d_+^2}{2}+\ln(x/K)+r\tau+\frac{\sigma^2\tau}{2}-\frac{\sigma^2\tau}{2}\\ &= \ln(x/K)-\frac{d_+^2}{2}.\end{aligned}$$

$$\begin{aligned}Ke^{-r\tau}e^{-d_-^2/2} &= K\exp\left(\ln(x/K)-\frac{d_+^2}{2}\right)\\ &= Ke^{\ln(x/K)}e^{-d_+^2/2}\\ &= K\left(\frac{x}{K}\right)e^{-d_+^2/2}\\ &= xe^{-d_+^2/2}.\end{aligned}$$

We have that,

$$\begin{aligned}\frac{Ke^{-r\tau}}{\sqrt{2\pi}}e^{-d_-^2/2} &= \frac{x}{\sqrt{2\pi}}e^{-d_+^2/2},\\ Ke^{-r\tau}N'(d_-) &= xN'(d_+).\end{aligned}$$

Now, returning to the original equation,

$$\begin{aligned}-xN'(d_+)\frac{\partial d_+}{\partial\tau}+Ke^{-r\tau}N'(d_-)\frac{\partial d_-}{\partial\tau} &= -xN'(d_+)\frac{\partial d_+}{\partial\tau}+xN'(d_+)\frac{\partial d_-}{\partial\tau}\\ &= xN'(d_+)\left(-\frac{\partial d_+}{\partial\tau}+\frac{\partial d_-}{\partial\tau}\right)\\ &= xN'(d_+)\left(\frac{\partial d_-}{\partial\tau}-\frac{\partial d_+}{\partial\tau}\right)\\ &= xN'(d_+)\left(\frac{\partial}{\partial\tau}(d_--d_+)\right)\\ &= xN'(d_+)\left(\frac{\partial}{\partial\tau}(-\sigma\sqrt{\tau})\right)\\ &= xN'(d_+)\left(\frac{\partial}{\partial\tau}(-\sigma\tau^{1/2})\right)\\ &= xN'(d_+)\left(-\sigma\frac{\partial}{\partial\tau}\tau^{1/2}\right)\\ &= xN'(d_+)\left(-\sigma\cdot\frac{1}{2}\tau^{-1/2}\right)\\ &= -\frac{\sigma}{2}\tau^{-1/2}xN'(d_+)\\ &= -\frac{\sigma x}{2\sqrt{\tau}}N'(d_+).\end{aligned}$$

Clearly the theta is negative. This makes sense since for a plain call without dividends, as time passes, there is less opportunity for the stock to move favorably so time decay hurts the option holder.

### 2.11.8 Hedge Portfolio

At time $t$, if the stock price is $x$, the short-option hedge from the replication argument holds $c_x(t,x)$ shares of stock. Since $c_x(t,x) = N(d_+)$, then the dollar amount invested in the stock is $xc_x(t,x) = xN(d_+)$. The option itself is worth,

$$c(t,x) = xN(d_+) - Ke^{-r(T-t)}N(d_-).$$

Therefore, the amount invested in the money-market account is,

$$c(t,x) - xc_x(t,x) = \left[xN(d_+) - Ke^{-r(T-t)}N(d_-)\right] - xN(d_+) = -Ke^{-r(T-t)}N(d_-).$$

So naturally we have that,

$$c(t,x) - xc_x(t,x) = -Ke^{-r(T-t)}N(d_-).$$

This is negative so the hedging portfolio borrows cash. To hedge a short call, you buy some stock and you may need to finance that purchase by borrowing. Furthmore, suppose we fix a time $t$ and current stock price $x$ and consider the portfolio of one long call and short $c_x(t,x)$ shares of stock and put the residual amount in the money market account. The money market amount is chosen to make the initial portfolio value zero. The call costs $c(t,x_1)$ so shorting $c_x(t,x_1)$ shares of stock brings in $x_1c_x(t,x_1)$. Thus, the leftover cash invested in the money market is $M = x_1c_x(t,x_1) - c(t,x_1)$.

Therefore the initial portfolio value is,

$$c(t,x_1) - x_1c_x(t,x_1) + M = c(t,x_1) - x_1c_x(t,x_1) + x_1c_x(t,x_1) - c(t,x_1) = 0.$$

Suppose the stock price jumps instantaneously from $x_1$ to some other value $x$ and suppose that we do not rebalance immediately. The money market amount $M$ does not change instantaneously and the call value becomes $c(t,x)$. The short stock position is now worth $-xc_x(t,x_1)$ because we are still short the older number of shares $c(t,x_1)$. Thus the portfolio value becomes,

$$\Pi(x) = c(t,x) - xc_x(t,x_1) + M.$$

Substituting and re-arranging we can solve to get,

$$\begin{aligned}
M &= x_1c_x(t,x_1) - c(t,x_1) \\
\Pi(x) &= c(t,x) - xc_x(t,x_1) + x_1c_x(t,x_1) - c(t,x_1) \\
&= c(t,x) - c_x(t,x_1)(x - x_1) - c(t,x_1) \\
&= c(t,x) - \left(c(t,x_1) + c_x(t,x_1)(x - x_1)\right).
\end{aligned}$$

The expression in parentheses is the tangent line to the curve $y = c(t,x)$ at $x_1$ so $\Pi(x)$ is the vertical distance between the option-price curve and its tangent line at $x_1$. Because gamma is positive, the function $x \mapsto c(t,x)$ is convex. For a convex function the graph always lies above its tangent line,

$$c(t,x) \geq c(t,x_1) + c_x(t,x_1)(x - x_1) \implies \Pi(x) \geq 0.$$

Note we have equality when $x = x_1$ meaning the portfolio gains value if the stock moves either up or down.

Now we have two other cases to consider.

(i) If the stock falls from $x_1$ to $x_0$ then the portfolio values becomes,

$$\Pi(x_0) = c(t,x_0) - c_x(t,x_1)(x_0 - x_1) - c(t,x_1).$$

Since $x_0 - x_1 < 0$ then the term $-c_x(t,x_1)(x_0 - x_1)$ is positive. Note that convexity also guarantees that the whole quantity is positive so the portfolio benefits from an instantaneous drop.

(ii) if the stock rises from $x_1$ to $x_2$ then the portfolio values becomes,

$$\Pi(x_2) = c(t,x_2) - c_x(t,x_1)(x_2 - x_1) - c(t,x_1).$$

The distance between the convex curve and its tangent line and positive so the portfolio also benefits from an instantaneous rise. The instant it is set up, the first-order sensitivity to the stock is zero.

The call has delta $c_x(t, x_1)$ and the short stock position has delta $-c_x(t, x_1)$ so the total delta is $c_x(t, x_1) - c_x(t, x_1) = 0$. Thus, very small stock moves have no first-order effect on the portfolio.

However, what remains is the 2nd-order effect which is govered by Gamma. Since Gamma is positive, the second-order effect is favorable. You can also see this from a Taylor expansion. For $x$ near $x_1$,

$$c(t, x) = c(t, x_1) + c_x(t, x_1)(x - x_1) + \frac{1}{2} c_{xx}(t, \zeta)(x - x_1)^2$$

for some intermediate point $\zeta$. Substituting into the equation for $\Pi(x)$ since $c_{xx}(t, \zeta) > 0$ we get that $\Pi(x) > 0$ for $x \neq x_1$. After neutralizing the linear term, the remaining quadratic term is positive.

### 2.11.9 Toward Volatility

A long-gamma portfolio profits from movement regardless of direction so its value depends not on whether the stock drifts upward or downward on average, but on how much it moves around. This is why volatility matters so much. Greater volatility means more opportunity to realize gains from convexity.

Moreover, this explains why the stock's drift $\alpha$ does not appearin the BS-PDE. A delta-neutral portfolio eliminates directional exposure to first order. Once direction is hedged away, what matters for pricing is the magnitude of fluctuations captured by $\sigma$ not the physical expected return $\alpha$. Long gamma typically comes with negative theta as the option holder gains from convexity but pays for the passage of time. Mathematically, we had derived,

$$c_t(t, x) = -rKe^{-r(T-t)} N(d_-) - \frac{\sigma x}{2\sqrt{T - t}} N'(d_+) < 0.$$

As $t \to \infty$ with fixed $x$, the option value tends to fall. A long-gamma position therefore benefits from realized movement but suffers from time decay if the stock does not move enough.

**Vega.** Vega is the derivative of the option wrt volatility. Differentiating gives,

$$c_\sigma = xN'(d_+) \frac{\partial d_+}{\partial \sigma} - Ke^{-r(T-t)} N'(d_-) \frac{\partial d_-}{\partial \sigma}.$$

Recall the identity $xN'(d_+) = Ke^{-r(T-t)} N'(d_-)$ which becomes,

$$c_\sigma x N'(d_+) \left( \frac{\partial d_+}{\partial \sigma} - \frac{\partial d_-}{\partial \sigma} \right).$$

Since we know that $d_- - d_+ = -\sigma\sqrt{T - t}$ then we have that $\frac{\partial d_-}{\partial \sigma} - \frac{\partial d_+}{\partial \sigma} = -\sqrt{T - t}$ so $\frac{\partial d_+}{\partial \sigma} - \frac{\partial d_-}{\partial \sigma} = \sqrt{T - t}$. Thus,

$$c_\sigma(t, x) = x\sqrt{T - t} N'(d_+) > 0.$$

Higher volatility raises call prices which matches our intuition from long gamma that more volatility means more chance to benefit from convexity. So the full picture is this. Delta measures first-order sensitivity to stock price. Gamma measures curvature and tells you whether a delta-neutral position benefits from stock movement. Theta measures time decay and is typically negative for a long call. Vega measures sensitivity to volatility and is positive for a long call.

The economic reason these fit together is that an option is a convex claim: convexity is valuable when the underlying moves a lot, but that convexity costs time value, which decays as expiration approaches.