Some problems on Itô calculus

Nguyen Duc Thanh

Spring 2024

1 Itô's lemma with explicit time-dependence

1.1 Introduction

Lemma: Let $B := (\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]}, (B_t)_{t\in[0,T]})$ be the standard 1-dimensional Brownian motion, and let $f : \mathbb{R}^2 \ni (t,x) \mapsto f(t,x) \in \mathbb{R}$ be continuous, with $\partial_t f, \partial_x f, \partial_x^2 f$ also continuous. Then, the following equality holds:

$$f(t,B_t) = f(0,0) + \int_0^t [\partial_x f](s,B_s)\,\mathrm{d}B_s + \int_0^t \left\{[\partial_t f](s,B_s) + \frac{1}{2}[\partial_x^2 f](s,B_s)\right\}\mathrm{d}s.$$

Let $\mathcal{P}_l$ be a partition of $[0,t]$ then by the telescopic formula, we have

$$f(t,B_t) = f(0,0) + \sum_{j=0}^{n_l-1}\Big[f\big(t_{j+1}^l, B_{t_{j+1}^l}\big) - f\big(t_j^l, B_{t_j^l}\big)\Big].$$

We get by Taylor expansion

$$\sum_{j=0}^{n_l-1}\Big[f\big(t_{j+1}^l, B_{t_{j+1}^l}\big) - f\big(t_j^l, B_{t_j^l}\big)\Big] = F_t + F_x + F_{tt} + F_{xx} + F_{xt} + R^{(3)},$$

where

$$F_t = \sum_{j=0}^{n_l-1}[\partial_t f](t_j^l, B_{t_j^l})\,(t_{j+1}^l - t_j^l),$$

$$F_x = \sum_{j=0}^{n_l-1}[\partial_x f](t_j^l, B_{t_j^l})\,\big(B_{t_{j+1}^l} - B_{t_j^l}\big),$$

$$F_{tt} = \frac{1}{2}\sum_{j=0}^{n_l-1}[\partial_t^2 f](t_j^l, B_{t_j^l})\,(t_{j+1}^l - t_j^l)^2,$$

$$F_{xx} = \frac{1}{2}\sum_{j=0}^{n_l-1}[\partial_x^2 f](t_j^l, B_{t_j^l})\,\big(B_{t_{j+1}^l} - B_{t_j^l}\big)^2,$$

$$F_{xt} = \sum_{j=0}^{n_l-1}[\partial_x\partial_t f](t_j^l, B_{t_j^l})\,\big(B_{t_{j+1}^l} - B_{t_j^l}\big)(t_{j+1}^l - t_j^l),$$

and $R^{(3)}$ being the remainder term to third order. We simplify the notation by omitting l and denote $\Delta t_j = t_{j+1} - t_j$ and $\Delta B_{t_j} = B_{t_{j+1}} - B_{t_j}$. Thus, we get

$$
\begin{aligned}
F_t &= \sum_{j=0}^{n-1} [\partial_t f](t_j, B_{t_j})\, \Delta t_j, \\
F_x &= \sum_{j=0}^{n-1} [\partial_x f](t_j, B_{t_j})\, \Delta B_{t_j}, \\
F_{tt} &= \frac{1}{2} \sum_{j=0}^{n-1} [\partial_t^2 f](t_j, B_{t_j})\, (\Delta t_j)^2, \\
F_{xx} &= \frac{1}{2} \sum_{j=0}^{n-1} [\partial_x^2 f](t_j, B_{t_j})\, \big(\Delta B_{t_j}\big)^2, \\
F_{xt} &= \sum_{j=0}^{n-1} [\partial_x \partial_t f](t_j, B_{t_j})\, \big(\Delta B_{t_j}\big)(\Delta t_j).
\end{aligned}
$$

We can also write $R^{(3)}$ explicitly

$$
\begin{aligned}
R^{(3)} = \frac{1}{6} \sum_{j=0}^{n-1} \Big(&[\partial_x^3 f](\tau_j, b_j)\, \big(\Delta B_{t_j}\big)^3 + 3[\partial_x^2 \partial_t f](\tau_j, b_j)\, \big(\Delta B_{t_j}\big)^2 (\Delta t_j) \\
&+ 3[\partial_x \partial_t^2 f](\tau_j, b_j)\, \big(\Delta B_{t_j}\big)(\Delta t_j)^2 + [\partial_t^3 f](\tau_j, b_j)\, (\Delta t_j)^3 \Big)
\end{aligned}
$$

for some $\tau_j \in (t_j, t_{j+1})$ and $b_j = B_{\tau'_j}$ where $\tau'_j \in (t_j, t_{j+1})$. By the definition of the Riemann integral, we get

$$
F_t \xrightarrow{|\mathcal{P}| \to 0} \int_0^t [\partial_t f](s, B_s)\, \mathrm{d}s,
$$

while by the definition of the Itô integral, we get

$$
F_x \xrightarrow{|\mathcal{P}| \to 0} \int_0^t [\partial_x f](s, B_s)\, \mathrm{d}B_s,
$$

For the remaining terms, we need to evaluate the sizes of the differences when we take the limit. We further assume that the map $s \mapsto f(s, B_s) \in M^2([0,T])$, meaning that the convergence is in the L^2-sense. We want to keep to first order in Δt so the only terms we need to be concerned about are $(\Delta B)^2$ and $(\Delta B)^3$ (the mixed terms can be neglected but we do not prove this). The $(\Delta B)^3$ term will also be ignored without a rigorous proof because as we shall see, $(\Delta B)^2 \sim \Delta t$.

1.2 Quadratic variation

Let us first prove some theorems in a more general setting. We will consider a martingale $M \in M^2([0,T])$ (not necessarily a Brownian motion).

Theorem: Given $t \in [0,T]$ and a partition $\mathcal{P} = (t_j)_{j=0}^n$ of $[0,t]$ with $t_0 = 0$ and $t_n = t$. We define

$$N^{\mathcal{P}} := \sum_{j=0}^{n-1} (M_{t_{j+1}} - M_{t_j})^2.$$

Then, $N^{\mathcal{P}}$ converges in L^2 to

$$[M]_t := \lim_{|\mathcal{P}|\to 0} N^{\mathcal{P}} = (M_t)^2 - (M_0)^2 - 2\int_0^t M_s \,\mathrm{d}M_s.$$

This is called the *quadratic variation* of M.

Proof. We have

$$\begin{aligned} N^{\mathcal{P}} &= \sum_{j=0}^{n-1} \left[(M_{t_{j+1}})^2 - 2M_{t_{j+1}}M_{t_j} + (M_{t_j})^2\right] \\ &= \sum_{j=0}^{n-1} \left[(M_{t_{j+1}})^2 - (M_{t_j})^2 - 2M_{t_j}(M_{t_{j+1}} - M_{t_j})\right] \\ &= (M_t)^2 - (M_0)^2 - 2\sum_{j=0}^{n-1} M_{t_j}(M_{t_{j+1}} - M_{t_j}). \end{aligned}$$

The last term converges in L^2 to

$$\sum_{j=0}^{n-1} M_{t_j}(M_{t_{j+1}} - M_{t_j}) \xrightarrow{|\mathcal{P}|\to 0} \int_0^t M_s \,\mathrm{d}M_s.$$

□

Theorem: Let X be an elementary process in $M^2([0,T])$, adapted to the stochastic process M, we have for $t \in [0,T]$

$$\mathbb{E}\left(\left(\int_0^t X_s \,\mathrm{d}M_s\right)^2\right) = \int_{[0,t]\times\Omega} X^2 \,\mathrm{d}\mu_M,$$

where

$$\mu_M\big((r,s] \times F_r\big) = \mathbb{E}\big(\mathbb{I}_{F_r}(M_s - M_r)^2\big),$$

for $r < s$, $F_r \in \mathcal{F}_r$ and is called the *Doléans measure.*

Proof. Let us denote

$$X_s(\omega) = \sum_{j=0}^{n-1} x_j \,\mathbb{I}_{(t_j,t_{j+1}](s)\times F_j(\omega)},$$

with $x_j \in \mathbb{R}$, and $F_j \in \mathcal{F}_{t_j}$. Comparing to the definition from the lecture notes [1], the above definition differs by also dividing the filtration into *predictable rectangles*, with the relation between the two definition being

$$X_j(\omega) = x_j \mathbb{I}_{F_j(\omega)}.$$

We have

$$\begin{aligned}
&\left(\int_0^t X_s \, \mathrm{d}M_s\right)^2 \\
&= \left(\sum_{j=0}^{n-1} x_j \, \mathbb{I}_{F_j} (M_{t_{j+1}} - M_{t_j})\right)^2 \\
&= \sum_{j=0}^{n-1} x_j^2 \, \mathbb{I}_{F_j} (M_{t_{j+1}} - M_{t_j})^2 + 2 \sum_{j=0}^{n-1} \sum_{k=j+1}^{n-1} x_j x_k \, \mathbb{I}_{F_j \cap F_k} (M_{t_{j+1}} - M_{t_j})(M_{t_{k+1}} - M_{t_k}).
\end{aligned}$$

When taking the expectation of the above expression, we need to evaluate two types of expression

$$\begin{aligned}
&\mathbb{E}\big(\mathbb{I}_{F_j} (M_{t_{j+1}} - M_{t_j})^2\big) =: \mu_M\big((t_j, t_{j+1}] \times F_j\big), \\
&\mathbb{E}\big(\mathbb{I}_{F_j \cap F_k} (M_{t_{j+1}} - M_{t_j})(M_{t_{k+1}} - M_{t_k})\big),
\end{aligned}$$

with $j < k$. We will use some results from conditional expectation, i.e., let W be $\mathcal{G}$-measurable ($\mathcal{G}$ a subalgebra) and bounded, then

$$\mathbb{E}(WX) = \mathbb{E}\big(W \mathbb{E}(X|\mathcal{G})\big).$$

Since $t_j < t_{j+1} \leq t_k < t_{k+1}$, we condition the second expectation with respect to $\mathcal{F}_{t_k}$. By using the fact that $\mathbb{I}_{F_j \cap F_k}(M_{t_{j+1}} - M_{t_j})$ is $\mathcal{F}_{t_k}$-measurable, we have

$$\mathbb{E}\big(\mathbb{I}_{F_j \cap F_k} (M_{t_{j+1}} - M_{t_j})(M_{t_{k+1}} - M_{t_k})\big) = \mathbb{E}\Big(\mathbb{I}_{F_j \cap F_k} (M_{t_{j+1}} - M_{t_j}) \mathbb{E}(M_{t_{k+1}} - M_{t_k} | \mathcal{F}_{t_k})\Big) = 0.$$

The second equality comes from the definition of a martingle, i.e.,

$$\mathbb{E}(M_{t_{k+1}} - M_{t_j} | \mathcal{F}_{t_k}) = \mathbb{E}(M_{t_{k+1}} | \mathcal{F}_{t_k}) - \mathbb{E}(M_{t_k} | \mathcal{F}_{t_k}) = \mathbb{E}(M_{t_k}) - \mathbb{E}(M_{t_k}) = 0.$$

Finally, we obtain

$$\mathbb{E}\left(\left(\int_0^t X_s \, \mathrm{d}M_s\right)^2\right) = \sum_{j=0}^{n-1} x_j^2 \, \mu_M\big((t_j, t_{j+1}] \times F_j\big) = \int_{[0,t] \times \Omega} X^2 \, \mathrm{d}\mu_M.$$

This can be extended to general X in $M^2([0,T])$ since elementary processes are dense in this space, so they can be approximated by elementary processes. □

The above theorem implies that there is an isometry between convergence in the L^2-space containing X and that L^2-space containing $\int X_s \, \mathrm{d}M_s$. In relation to the quadratic variation, we state without proof the following result (Theorem 4.2, [3])

$$\mathbb{E}\left(\left(\int_0^t X_s \, \mathrm{d}M_s\right)^2\right) = \int_{[0,t] \times \Omega} X^2 \, \mathrm{d}\mu_M = \mathbb{E}\left(\int_0^t \big(X_s^2\big) \, \mathrm{d}[M]_s\right).$$

Theorem: With $M, \mathcal{P}, [M]$ defined as before, let P be a bounded continuous adapted process on $[0,T]$. Similar to the quadratic variation case, we define

$$Q^{\mathcal{P}} := \sum_{j=0}^{n-1} P_{t_j}(M_{t_{j+1}} - M_{t_j})^2,$$

$$q^{\mathcal{P}} := \sum_{j=0}^{n-1} P_{t_j}([M]_{t_{j+1}} - [M]_{t_j}).$$

Then, $Q^{\mathcal{P}} - q^{\mathcal{P}}$ converges to 0 in L^2, i.e.

$$\lim_{|\mathcal{P}|\to 0} \mathbb{E}\big((Q^{\mathcal{P}} - q^{\mathcal{P}})^2\big) = 0.$$

In other words, $Q^{\mathcal{P}}$ converges in L^2 to

$$\lim_{|\mathcal{P}|\to 0} Q^{\mathcal{P}} = \int_0^t P_s \, \mathrm{d}[M]_s.$$

Proof. We want to evaluate

$$Q^{\mathcal{P}} - q^{\mathcal{P}} = \sum_{j=0}^{n-1} P_{t_j}\Big[(M_{t_{j+1}} - M_{t_j})^2 - ([M]_{t_{j+1}} - [M]_{t_j})\Big].$$

From the above theorem

$$\begin{aligned}
[M]_{t_{j+1}} - [M]_{t_j} &= \Big[(M_{t_{j+1}})^2 - (M_0)^2 - 2\int_0^{t_{j+1}} M_s \, \mathrm{d}M_s\Big] - \Big[(M_{t_j})^2 - (M_0)^2 - 2\int_0^{t_j} M_s \, \mathrm{d}M_s\Big] \\
&= (M_{t_{j+1}})^2 - (M_{t_j})^2 - 2\int_{t_j}^{t_{j+1}} M_s \, \mathrm{d}M_s,
\end{aligned}$$

so the term in the bracket is simplified to

$$\begin{aligned}
(M_{t_{j+1}} - M_{t_j})^2 - ([M]_{t_{j+1}} - [M]_{t_j}) &= 2(M_{t_j})^2 - 2M_{t_{j+1}}M_{t_j} + 2\int_{t_j}^{t_{j+1}} M_s \, \mathrm{d}M_s \\
&= 2\Big[\int_{t_j}^{t_{j+1}} M_s \, \mathrm{d}M_s - M_{t_j}(M_{t_{j+1}} - M_{t_j})\Big].
\end{aligned}$$

Then, we get

$$\begin{aligned}
Q^{\mathcal{P}} - q^{\mathcal{P}} &= 2\sum_{j=0}^{n-1} P_{t_j}\Big[\int_{t_j}^{t_{j+1}} M_s \, \mathrm{d}M_s - M_{t_j}(M_{t_{j+1}} - M_{t_j})\Big] \\
&= 2\sum_{j=0}^{n-1}\Big[\int_0^t \mathbb{I}_{(t_j,t_{j+1}]} P_{t_j} M_s \, \mathrm{d}M_s - \int_0^t \mathbb{I}_{(t_j,t_{j+1}]} P_{t_j} M_{t_j} \mathrm{d}M_s\Big] \\
&= 2\int_0^t P_s^{\mathcal{P}}(M_s - M_s^{\mathcal{P}}) \, \mathrm{d}M_s,
\end{aligned}$$

where

$$P_s^{\mathcal{P}} = \sum_{j=0}^{n-1} \mathbb{I}_{(t_j,t_{j+1}](s)} P_{t_j} \quad , \quad M_s^{\mathcal{P}} = \sum_{j=0}^{n-1} \mathbb{I}_{(t_j,t_{j+1}](s)} M_{t_j}$$

The integrand in the expression for $Q^{\mathcal{P}} - q^{\mathcal{P}}$ converges to 0 as $|\mathcal{P}| \to 0$. Then, from the previous theorem on isometry, convergence of the integrand to 0 implies the statement. □

1.3 Back to Itô's lemma

We want to apply this to the Brownian motion $(\Delta B)^2$. We will show that

$$[B]_t = t,$$

meaning that

$$\mathbb{E}\bigg(\Big[\sum_{j=0}^{n-1} (\Delta B_{t_j})^2 - t\Big]^2\bigg) \to 0.$$

Since $t = \sum \Delta t_j$, the above equation is equivalent to

$$0 = \mathbb{E}\bigg(\Big[\sum_{j=0}^{n-1} \big[(\Delta B_{t_j})^2 - \Delta t_j\big]\Big]^2\bigg) = \mathbb{E}\bigg(\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} X_j X_k\bigg) = \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} \mathbb{E}(X_j X_k),$$

where we have denoted $X_j = (\Delta B_{t_j})^2 - \Delta t_j$. Evaluating the r.h.s.

$$\begin{aligned}\mathbb{E}(X_j X_k) &= \mathbb{E}\bigg(\big[(\Delta B_{t_j})^2 - \Delta t_j\big]\big[(\Delta B_{t_k})^2 - \Delta t_k\big]\bigg) \\ &= \mathbb{E}\big((\Delta B_{t_j})^2 (\Delta B_{t_k})^2\big) - \Delta t_k\, \mathbb{E}\big((\Delta B_{t_j})^2\big) - \Delta t_j\, \mathbb{E}\big((\Delta B_{t_k})^2\big) + \Delta t_j \Delta t_k.\end{aligned}$$

Since we are working with the Brownian motion, we have by definition

$$\mathbb{E}\big((\Delta B_{t_j})^2\big) = \Delta t_j, \quad \mathbb{E}\big((\Delta B_{t_k})^2\big) = \Delta t_k.$$

The increments at different times are independent, so we can write

$$\begin{aligned}\mathbb{E}\big((\Delta B_{t_j})^2 (\Delta B_{t_k})^2\big) &= \begin{cases} \mathbb{E}\big((\Delta B_{t_j})^4\big) & \text{if } j = k, \\ \mathbb{E}\big((\Delta B_{t_j})^2\big)\, \mathbb{E}\big((\Delta B_{t_k})^2\big) & \text{if } j \neq k, \end{cases} \\ &= \begin{cases} \mathbb{E}\big((\Delta B_{t_j})^4\big) & \text{if } j = k, \\ \Delta t_j\, \Delta t_k & \text{if } j \neq k. \end{cases}\end{aligned}$$

For the fourth moment, let us use the moment generating function for some Gaussian distribution X with $\mathbb{E}(X) = m$ and $\text{Var}(X) = \sigma^2$

$$\mathbb{E}(e^{\lambda X}) = e^{\lambda m + \lambda^2 \sigma^2/2}.$$

The proof of this equality is given in the Appendix. For $X = \Delta B_{t_j}$, $m = 0$, $\sigma^2 = \Delta t_j$. We can get the fourth moment by simply reading off the coefficient of the λ^4 term (and multiply by 4!) in the expansion

$$\mathbb{E}(e^{\lambda\, \Delta B_{t_j}}) = e^{\lambda^2\, \Delta t_j/2} = 1 + \frac{\lambda^2\, \Delta t_j}{2} + \frac{\lambda^4\, (\Delta t_j)^2}{8} + \ldots,$$

which gives
$$\mathbb{E}\big((\Delta B_{t_j})^4\big) = 3(\Delta t_j)^2.$$
Therefore, we get
$$\mathbb{E}(X_j X_k) = \begin{cases} 2(\Delta t_j)^2 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$
This means that
$$\mathbb{E}\left(\left[\sum_{j=0}^{n-1} X_j\right]^2\right) = 2\sum_{j=0}^{n-1}(\Delta t_j)^2 \leq 2|\mathcal{P}|\sum_{j=0}^{n-1}\Delta t_j = 2|\mathcal{P}|t \xrightarrow{|\mathcal{P}|\to 0} 0,$$
so the quadratic variation of the Brownian motion indeed converges to t.

Using the result from the previous section
$$F_{xx} \xrightarrow{|\mathcal{P}|\to 0} \frac{1}{2}\int_0^t [\partial_x^2 f](s, B_s)\,\mathrm{d}[B]_s = \frac{1}{2}\int_0^t [\partial_x^2 f](s, B_s)\,\mathrm{d}s,$$
so finally, we have
$$f(t, B_t) = f(0,0) + \int_0^t [\partial_x f](s, B_s)\,\mathrm{d}B_s + \int_0^t \left\{[\partial_t f](s, B_s) + \frac{1}{2}[\partial_x^2 f](s, B_s)\right\}\mathrm{d}s.$$

2 Examples

2.1 Itô exponential

We consider the function
$$f(t, x) = e^{x - t/2},$$
called the Itô exponential. The derivatives are given by
$$\begin{aligned} [\partial_x f](t, x) &= e^{x-t/2} = f(t, x), \\ [\partial_t f](t, x) &= -\frac{1}{2}e^{x-t/2}, \\ [\partial_x^2 f](t, x) &= e^{x-t/2}. \end{aligned}$$
This satisfies
$$[\partial_t f](t, x) + \frac{1}{2}[\partial_x^2 f](t, x) = 0,$$
so Itô's lemma reduces to
$$f(t, B_t) = f(0,0) + \int_0^t f(s, B_s)\,\mathrm{d}B_s.$$
It can be seen that this equation resembles that of the exponential function from ordinary calculus, hence the name.

2.2 Gambler's ruin for Brownian motion with drift

Let $\sigma, \mu \neq 0$ be some constants, we want to study the probability of the random variable

$$X_t = \sigma B_t + \mu t$$

reaching $\mathcal{W}$ before $\mathcal{L}$ for $\mathcal{L} < 0 < \mathcal{W}$. We define the stopping time

$$\tau = \min\{t \geq 0 | X_t = \mathcal{W} \text{ or } X_t = \mathcal{L}\}.$$

Even if $\mu = 0$, since $\mathbb{E}(B_t^2) = t$, we expect the motion of X_t to become more spread out as time goes on, meaning that X_t will eventually reach either one of the boundaries and $\tau < \infty$. With $\mu \neq 0$, there is a bias towards one boundary so it is even more likely that $\tau < \infty$ a.s.

We want to apply the optional stopping theorem so we need to find a bounded martingale of X_t adapted to the Brownian filtration

$$M_t = g(X_t).$$

The theorem states that

$$\mathbb{E}[g(X_\tau)] = \mathbb{E}[g(X_0)] = g(X_0),$$

since we assume the initial condition at $t = 0$. By definition X_τ can only take the values $X_\tau = \mathcal{L}$ or $X_\tau = \mathcal{W}$, so we get

$$\mathbb{P}(X_\tau = \mathcal{L})\, g(X_\tau = \mathcal{L}) + \mathbb{P}(X_\tau = \mathcal{W})\, g(X_\tau = \mathcal{W}) = g(X_0).$$

A convenient choice is for $g(X_\tau = \mathcal{L}) = 0$ and $g(X_\tau = \mathcal{W}) = 1$, so that

$$\mathbb{P}(X_\tau = \mathcal{W}) = g(X_0).$$

For the condition of $t \mapsto g(X_t)$ being a bounded martingale, let us rewrite to Itô's lemma in a different way, given that it is well-defined

$$f(0,0) + \int_0^t [\partial_x f](s, B_s)\, \mathrm{d}B_s = f(t, B_t) - \int_0^t \left\{[\partial_t f](s, B_s) + \frac{1}{2}[\partial_x^2 f](s, B_s)\right\} \mathrm{d}s.$$

Notice that $f(0,0)$ is just a constant and $\int [\partial_x f](s, B_s)\, \mathrm{d}B_s$ is a martingale transform so the l.h.s. is a martingale adapted to the Brownian filtration. This implies that the r.h.s. is also a martingale. Furthermore, if we impose the condition that

$$\partial_t f + \frac{1}{2}\partial_x^2 f = 0,$$

then $f(t, B_t)$ itself is a martingale. Notice that the Itô exponential is one example that satisfies this equation so we can expect a similar form of the equation for our problem.

Since X_t is a function of t and B_t itself, we can write

$$g(X_t) = g(\sigma B_t + \mu t).$$

The problem becomes finding a function $f(t, x) = g(\sigma x + \mu t)$ such that

$$\partial_t f + \frac{1}{2}\partial_x^2 f = 0,$$

with the boundary conditions

$$\begin{cases} g(\mathcal{L}) = 0, \\ g(\mathcal{W}) = 1. \end{cases}$$

Using the chain rule

$$\partial_t f = \frac{\partial y}{\partial t}\frac{\partial g}{\partial y} = \mu g',$$
$$\partial_x^2 f = \frac{\partial y}{\partial x}\frac{\partial}{\partial y}\left(\frac{\partial y}{\partial x}\frac{\partial g}{\partial y}\right) = \sigma^2 g'',$$

so we have

$$\sigma^2 g'' + 2\mu g' = 0.$$

Solving for g', we find

$$g'(y) = Ae^{-2\mu y/\sigma^2}$$

for some constant A. Integrating g' then gives

$$g(y) = A'e^{-2\mu y/\sigma^2} + B,$$

where $A' = -\sigma^2 A/2\mu$ and B some constant. Using the boundary conditions,

$$\begin{cases} g(\mathcal{L}) = A'e^{-2\mu\mathcal{L}/\sigma^2} + B = 0, \\ g(\mathcal{W}) = A'e^{-2\mu\mathcal{W}/\sigma^2} + B = 1, \end{cases}$$

we can get A' and B

$$\begin{cases} A' = \frac{1}{e^{-2\mu\mathcal{W}/\sigma^2} - e^{-2\mu\mathcal{L}/\sigma^2}}, \\ B = \frac{-e^{-2\mu\mathcal{L}/\sigma^2}}{e^{-2\mu\mathcal{W}/\sigma^2} - e^{-2\mu\mathcal{L}/\sigma^2}}. \end{cases}$$

Thus, we get

$$g(y) = \frac{1 - e^{-2\mu(y-\mathcal{L})/\sigma^2}}{1 - e^{-2\mu(\mathcal{W}-\mathcal{L})/\sigma^2}} = \frac{1 - e^{-2\mu(y+|\mathcal{L}|)/\sigma^2}}{1 - e^{-2\mu(\mathcal{W}+|\mathcal{L}|)/\sigma^2}},$$

(recall that $\mathcal{L} < 0 < \mathcal{W}$) and the probability of reaching $\mathcal{W}$ before $\mathcal{L}$ is

$$\mathbb{P}(X_\tau = \mathcal{W}) = g(0) = \frac{1 - e^{-2\mu|\mathcal{L}|/\sigma^2}}{1 - e^{-2\mu(\mathcal{W}+|\mathcal{L}|)/\sigma^2}}.$$

If $\sigma = \mu = 1$ and $\mathcal{W} = |\mathcal{L}| = 1$, then

$$\mathbb{P}(X_\tau = \mathcal{W}) = \frac{1 - e^{-2}}{1 - e^{-4}} \approx 0.88 > 0.5,$$

as expected. For the case $\mu = 0$, the previous derivation does not work because the differential equation becomes

$$g_0'' = 0,$$

and the general solution is

$$g_0(y) = Cy + D.$$

With the boundary conditions, we get

$$\begin{cases} C = \frac{1}{\mathcal{W}-\mathcal{L}}, \\ D = \frac{-\mathcal{L}}{\mathcal{W}-\mathcal{L}}, \end{cases}$$

so

$$\mathbb{P}_0(X_\tau = \mathcal{W}) = g_0(0) = \frac{-\mathcal{L}}{\mathcal{W} - \mathcal{L}} = \frac{|\mathcal{L}|}{\mathcal{W} + |\mathcal{L}|}.$$

If $\mathcal{W} = |\mathcal{L}|$ then we get the familiar result of $\mathbb{P}_0(X_\tau = \mathcal{W}) = 0.5$. We can also arrive at this equation by approximation of e^α at small α

$$e^\alpha = 1 + \alpha + O(\alpha^2),$$

to get

$$\mathbb{P}(X_\tau = \mathcal{W}) = \frac{1 - e^{-2\mu|\mathcal{L}|/\sigma^2}}{1 - e^{-2\mu(\mathcal{W}+|\mathcal{L}|)/\sigma^2}} \approx \frac{1 - (1 - 2\mu|\mathcal{L}|/\sigma^2)}{1 - [1 - 2\mu(\mathcal{W} + |\mathcal{L}|)/\sigma^2]} = \frac{|\mathcal{L}|}{\mathcal{W} + |\mathcal{L}|} = \mathbb{P}_0(X_\tau = \mathcal{W}).$$

3 Appendix

Given a random variable X, the moment generating function is defined (if it exists) by

$$\mathbb{E}(e^{\lambda X}) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \mathbb{E}(X^n).$$

The n^{th} moment can be determined either from the coefficient of λ^n or by differentiation

$$\mathbb{E}(X^n) = \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} \mathbb{E}(e^{\lambda X}) \bigg|_{\lambda=0}.$$

For our case of a 1D Gaussian distribution $X = N(m, \sigma^2)$, we have the probability density function

$$\Pi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right].$$

The moment generating function is then

$$\begin{aligned}
\mathbb{E}(e^{\lambda X}) &= \int_{\mathbb{R}} e^{\lambda x}\, \Pi(x)\, \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left[\lambda x - \frac{(x-m)^2}{2\sigma^2}\right] \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left[-\frac{x^2 - 2mx + m^2 - 2\lambda\sigma^2 x}{2\sigma^2}\right] \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left(-\frac{[x - (m + \lambda\sigma^2)]^2 - (m + \lambda\sigma^2)^2 + m^2}{2\sigma^2}\right) \mathrm{d}x \\
&= \left[\frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \exp\left(-\frac{[x - (m + \lambda\sigma^2)]^2}{2\sigma^2}\right) \mathrm{d}x\right] e^{\lambda m + \lambda^2\sigma^2/2}.
\end{aligned}$$

It can be seen that the term in the bracket is just the normalization for a Gaussian distribution $N(m + \lambda\sigma^2, \sigma^2)$ so it just gives 1. Thus, we find

$$\mathbb{E}(e^{\lambda X}) = e^{\lambda m + \lambda^2\sigma^2/2},$$

as required.

4 Reference

[1] Introduction to Stochastic Calculus, lecture notes by S. Richard.
[2] A first course in stochastic calculus, book by J.-L. Arguin.
[3] Introduction to Stochastic Integration, book by K. L. Chung and R. J. Williams.