1 Stochastic Calculus

We will extend the ideas of the integral to allow us to deal with integrands and integrators that are stochastic processes. The stochastic integrals can be used to model the value of a portfolio that trades stocks in continuous time.

1.1 Itô Integral

We first recall the notion of integration with respect to a deterministic function with bounded variation.

Definition 1.1. Let $\Pi = \{t_0, t_1 \dots, t_n\}$ be a partition of $[0, t]$. The **Riemann–Stieltjes integral** of the integrand f with respect to the integrator G is defined as

$$F(t) = \int_0^t f(s)\, \mathrm{d}G(s) = \lim_{\|\Pi\| \to 0} \sum_{i=1}^{n} f(t_i) \left(G(t_{i+1}) - G(t_i)\right),$$

where the limit is taken over all partitions as the mesh size $\|\Pi\| = \max(t_i - t_{i-1})$ tends to 0.

Remark 1.2. Recall that if f is continuous and G is a function with bounded total variation, then the sample point in $f(t_i)$ does not matter and it can be replaced with any point $c_i \in [t_i, t_{i+1}]$. Notice that $F(t)$ is also a continuous function in this case.

Our goal is to define a notion of integration against a stochastic process. Let $\{\xi_t\}_{t\geq 0}$ and $\{X_t\}_{t\geq 0}$ be two adapted stochastic processes. Typically, ξ_t measures the amount of a stock held and X_t represents the value of a stock at time t. We want to define the stochastic integral

$$I(t) = \int_0^t \xi(s)\, \mathrm{d}X(s)$$

which encodes the value of a portfolio at time t. The rules of calculus do not immediately apply since even in the simplest case when we want to integrate against a Brownian motion W_t, we have seen that W_t does not have bounded total variation, so the Riemann–Stieltjes integral does not make sense. However, we can still define the stochastic integral in the natural way, which might remind us of the betting strategies we saw in Week 4.

Definition 1.3. Let $\Pi = \{t_0, t_1 \dots, t_n\}$ be a partition of $[0, t]$ and let $\{\xi_t\}_{t\geq 0}$ and $\{X_t\}_{t\geq 0}$ be two adapted stochastic processes. The **Itô integral** of the integrand ξ with respect to the integrator X is defined as

$$I(t) = \int_0^t \xi(s)\, \mathrm{d}X(s) = \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} \xi(t_i) \left(X(t_{i+1}) - X(t_i)\right), \tag{1}$$

where the limit is taken over all partitions as the mesh size $\|\Pi\| = \max(t_{i+1} - t_i)$ tends to 0. Sometimes we write the stochastic integral in **differential form** using the notation

$$dI(t) = \xi(t)dX(t).$$

This is shorthand for precisely what is written in (1).

Interpreting what the equal sign in (1) means is a bit beyond the course. We cannot compute the interpret the stochastic integral for a realization of Brownian motion since the Riemann–Stieltjes integral is not defined so the limit for a realization of Brownain motion doesn't make much sense. Therefore, we can't make sense of $I(t)$ as an almost sure limit. However, we can make sense of $I(t)$ as a random variable and show that it makes sense if we consider a weaker notion of convergence such as convergence in probability.

Remark 1.4. In the Itô integral, it is crucial that we take the left sample point in $\xi(t_i)$. This makes sense from a modeling perspective because our strategy to invest in a stock cannot depend on future information. Different choices of sample points lead to different notions of the stochastic integral that have very different properties. Notice that $I(t)$ is a stochastic process.

1.2 The Stochastic Integral Against Brownian Motion

Throughout this section, we only consider the case of integration against Brownian motion. That is, for any adapted stochastic process $\{\xi(s)\}_{s\geq 0}$ (with possibly discontinuous sample paths), we will consider the properties of

$$I(t) = \int_0^t \xi(s)\,\mathrm{d}W(s). \tag{2}$$

We immediately see that there are some subtle differences between the stochastic integral and the Riemann integral. If $f(t)$ is a function of bounded variation, then we have by the fundamental theorem of calculus,

$$\int_0^t f(s)\,df(s) = \frac{1}{2}f(t)^2.$$

Somewhat surprisingly, from the definition of the stochastic integral, we see that (Problem 1.4)

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

so there is an extra correction term in comparison with the non-random case. This correction can be explained using new type of calculus called **Itô calculus**.

1.2.1 Basic Properties

The first three properties are direct consequences of the fact that the stochastic integral is a sum. The Riemann integral of a possibly discontinuous function gives us a continuous function. We see that continuity is preserved even for the stochastic process.

Proposition 1.5 (*Continuity*)

The stochastic integral $I(t)$ given in (2) is continuous

The Riemann integral of a function gives us another function. By the definition, we also see that the Itô integral of an adapted process is again an adapted process.

Proposition 1.6 (*Adaptivity*)

For each t, the stochastic integral $I(t)$ given in (2) is $\mathcal{F}_t$ measurable.

A key property of the Riemann integral is the linearity, which allows us to compute the integrals of sums of functions. By the definition, the Itô integral obeys the same linearity properties.

Proposition 1.7 (*Additvity*)

If ξ and ζ are adapted processes, then

$$I(t) = \int_0^t \xi(s)\,\mathrm{d}W(s) \qquad J(t) = \int_0^t \zeta(s)\,\mathrm{d}W(s)$$

satisfies

$$\int_0^t a\xi(s) + b\zeta(s)\,\mathrm{d}W(s) = aI(t) + bJ(t).$$

We now move onto some less trivial properties. One of the main reasons why taking the sample points to be the left endpoint of the interval in the definition of the Itô integral, is that this implies that the Itô integral will be a martingale. We have already seen the intuition of this when we studied the betting strategies when we first introduced martingales in Week 4, and this computation is done for simple processes in Problem 1.1.

Proposition 1.8 (***Martingale***)

The stochastic integral $I(t)$ given in (2) is a martingale.

Since $I(t)$ is a martingale and $I(0) = 0$, we have

$$\mathbb{E}[I(t)] = 0$$

for all t. Therefore, the variance has a simpler form, $\text{Var}(I(t)) = \mathbb{E}[I^2(t)]$. The next formula allows us to compute the variance of the Itô integral. The computation is done for simple processes in Problem 1.2.

Proposition 1.9 (***Itô Isometry***)

The stochastic integral $I(t)$ given in (2) satisfies

$$\mathbb{E}[I^2(t)] = \mathbb{E}\left[\int_0^t \xi^2(s)\, ds\right]$$

We know that Brownian motion accumulates quadratic variation at a rate of 1 per unit of time. When we scale Brownian motion by $\xi(t)$, then one might expect that the quadratic variation of the Itô integral now depend on time and scale like $\Delta^2(t)$ since we are square in the definition of quadratic variation. The computation is done for simple processes in Problem 1.3.

Proposition 1.10 (***Quadratic Variation***)

The stochastic integral $I(t)$ given in (2) satisfies

$$[I, I](t) = \int_0^t \xi^2(s)\, ds.$$

Remark 1.11. The shorthand to remember this formula is using differential notation: $dI(t) = \xi(t)dW(t)$ so

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

since $dW(t)dW(t) = t$, the quadratic variation of Brownian motion.

1.3 Example Problems

1.3.1 Proofs of Main Results

Problem 1.1. Let $\Pi = \{t_0, \dots, t_n\}$ be a partition of $[0, T]$ and suppose that $\xi(t)$ is constant on each subinterval $[t_i, t_{i+1})$. Show that

$$I(t) = \int_0^t \xi(s)\, dW(s) = \sum_{i=0}^{n-1} \xi(t_i)\left(W(t_{i+1}) - W(t_i)\right)$$

is a martingale as in Proposition 1.8.

Solution 1.1. Let $0 \le s \le t \le T$. We will only show the case when s and t are in different subintervals, because the case that s and t are in the same interval is easier and considerably simpler. Suppose that $s \in [t_k, t_{k+1})$ and $t \in [t_\ell, t_{\ell+1})$. We have to show that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=0}^{\ell-1} \xi(t_i)\left(W(t_{i+1}) - W(t_i)\right) + \xi(t_\ell)(W(t) - W(t_\ell)) \mid \mathcal{F}_s\right] \\
&= \sum_{i=0}^{k-1} \xi(t_i)\left(W(t_{i+1}) - W(t_i)\right) + \xi(t_k)(W(s) - W(t_k)),
\end{aligned}$$

since we have to take the upper end of the integral to be t or s respectively instead of the end of the subinterval. Notice that

$$\sum_{i=0}^{n-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)=\sum_{i=0}^{k-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)+\xi(t_k)(W(t_{k+1})-W(t_k))$$
$$+\sum_{i=k+1}^{\ell-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)+\xi(t_\ell)(W(t)-W(t_\ell))$$

We will compute the conditional expected values of each term separately.

- If we are given information up to time s, then the first term is known. More precisely,

  $$\mathbb{E}\Big[\sum_{i=0}^{k-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)\mid\mathcal{F}_s\Big]=\sum_{i=0}^{k-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)$$

  since every random variable in the sum is $\mathcal{F}_s$ measurable.

- Since Brownian motion is a martingale, $\mathbb{E}[W(t_{k+1})\mid\mathcal{F}_s]=W_s$ and $\xi(t_k)$ and $W(t_k)$ are $\mathcal{F}_s$ measurable

  $$\mathbb{E}[\xi(t_k)(W(t_{k+1})-W(t_k))\mid\mathcal{F}_s]=\xi(t_k)(\mathbb{E}[W(t_{k+1})\mid\mathcal{F}_s]-W(t_k))=\xi(t_k)(W(s)-W(t_k)).$$

- We use the tower property. Notice that for each term in the sum, we can condition on future information $\mathcal{F}_{t_j}\supseteq\mathcal{F}_s$ to see that

  $$\mathbb{E}[\xi(t_i)(W(t_{i+1})-W(t_i))\mid\mathcal{F}_s]=\mathbb{E}[\mathbb{E}[\xi(t_i)(W(t_{i+1})-W(t_i))\mid\mathcal{F}_{t_i}]\mid\mathcal{F}_s].$$

  Since Brownian motion is a martingale, $\mathbb{E}[W(t_{i+1})\mid\mathcal{F}_{t_i}]=W_{t_i}$ and $\xi(t_i)$ and $W(t_i)$ are $\mathcal{F}_{t_i}$ measurable so the same computation as the second term implies that

  $$\mathbb{E}[\xi(t_i)(W(t_{i+1})-W(t_i))\mid\mathcal{F}_{t_i}]=0.$$

  Each term in the sum is zero, so

  $$\mathbb{E}\Big[\sum_{i=k+1}^{\ell-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)\mid\mathcal{F}_s\Big]=0$$

- The same computation as the third term by conditioning on the future information $\mathcal{F}_{t_\ell}\subseteq\mathcal{F}_s$ implies that

  $$\mathbb{E}[\xi(t_\ell)(W(t)-W(t_\ell))\mid\mathcal{F}_s]=0.$$

Applying linearity of conditional expectations and using the above four simplifications completes the proof.

Problem 1.2. Let $\Pi=\{t_0,\ldots,t_n\}$ be a partition of $[0,T]$ and suppose that $\xi(t)$ is constant on each subinterval $[t_i,t_{i+1})$. Show that

$$I(t)=\int_0^t\xi(s)\,dW(s)=\sum_{i=0}^{n-1}\xi(t_i)\left(W(t_{i+1})-W(t_i)\right)$$

satisfies Itô's isometry as in Proposition 1.9.

Solution 1.2. Let $t > 0$, and suppose that $t \in [t_\ell, t_{\ell+1})$ To simplify notation, let $D_k = W(t_{k+1}) - W(t_k)$ and $D_\ell = W(t) - W(t_\ell)$. We have

$$I(t) = \sum_{i=0}^{\ell-1} \xi(t_i)(W(t_{i+1}) - W(t_i)) + \xi(t_\ell)(W(t) - W(t_\ell)) = \sum_{k=0}^{\ell} \xi(t_i)D_i.$$

Then

$$\mathbb{E}[I(t)^2] = \sum_{i=0}^{\ell} \mathbb{E}[\xi(t_i)^2 D_i^2] + 2\sum_{i<j} \mathbb{E}[\xi(t_i)\xi(t_j)D_iD_j].$$

We compute each term separately,

- Notice that for each i, we have $\xi(t_i)$ is $\mathcal{F}_{t_i}$ measurable and D_i is independent of $\mathcal{F}_{t_i}$ so
$$\mathbb{E}[\xi(t_i)^2 D_i^2] = \mathbb{E}[\xi(t_i)^2]\,\mathbb{E}[D_i^2] = \mathbb{E}[\xi(t_i)^2](t_{i+1} - t_i)$$
since $D_i \sim N(0, t_{i+1} - t_i)$. We have
$$\sum_{i=0}^{\ell} \mathbb{E}[\xi(t_i)^2 D_i^2] = \sum_{i=0}^{\ell} \mathbb{E}[\xi(t_i)^2](t_{i+1} - t_i) = \int_0^t \mathbb{E}[\xi^2(s)]\,ds = \mathbb{E}\int_0^t \xi^2(s)\,ds$$
since $\Delta(s)$ is constant on the subintervals so
$$\mathbb{E}[\xi(t_i)^2](t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} \mathbb{E}[\xi(s)^2]\,ds \quad \text{and} \quad \mathbb{E}[\xi(t_\ell)^2](t_{\ell+1} - t_\ell) = \int_{t_\ell}^{t} \mathbb{E}[\xi(s)^2]\,ds.$$
- Notice that for $i < j$, we have $\xi(t_i)\xi(t_j)D_i$ is $\mathcal{F}_{t_j}$ measurable D_j is independent of $\mathcal{F}_j$ so
$$\mathbb{E}[\xi(t_i)\xi(t_j)D_iD_j] = \mathbb{E}[\xi(t_i)\xi(t_j)D_i]\,\mathbb{E}[D_j] = 0$$
since $D_j \sim N(0, t_{j+1} - t_j)$.

We conclude that
$$\mathbb{E}[I(t)^2] = \int_0^t \mathbb{E}[\xi^2(s)]\,ds = \mathbb{E}\int_0^t \xi^2(s)\,ds.$$

Problem 1.3. Let $\Pi = \{t_0, \ldots, t_n\}$ be a partition of $[0, T]$ and suppose that $\xi(t)$ is constant on each subinterval $[t_i, t_{i+1})$. Show that

$$I(t) = \int_0^t \xi(s)\,dW(s) = \sum_{i=0}^{n-1} \xi(t_i)\,(W(t_{i+1}) - W(t_i))$$

satisfies the quadratic variation formula as in Proposition 1.10.

Solution 1.3. Notice that $\xi(s)$ is constant on any subinterval $[t_i, t_{i+1})$. Therefore, for any partition $\Pi = \{s_0, \ldots, s_n\}$ of this subinterval,

$$\sum_{s=0}^{n}(I(s_{i+1}) - I(s_i))^2 = \sum_{s=0}^{n}[\xi(t_i)(W(s_{i+1}) - W(s_i)]^2 = \xi^2(t_i)\sum_{s=0}^{n}(W(s_{i+1}) - W(s_i))^2 \to \xi^2(t_i)(t_{i+1} - t_i)$$

since $\sum_{s=0}^{n}(W(s_{i+1}) - W(s_i))^2 \to (t_{i+1} - t_i)$ using the quadratic variation of Brownian motion. On the other hand, since $\xi^2(t)$ is constant on $[t_i, t_{i+1})$, we have

$$\xi^2(t_i)(t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} \xi^2(s)\,ds.$$

Summing these up and using the appropriate modification for the last term proves the statement.

1.3.2 Applications

Problem 1.4. Let $W(t)$ be a Brownian motion.

1. Use the definition of the Itô Integral to show that
$$\int_0^t W(s)\,\mathrm{d}W(s) = \frac{1}{2}W_t^2 - \frac{1}{2}t.$$

2. Show that
$$\frac{1}{2}W_t^2 - \frac{1}{2}t$$
is a martingale.

Solution 1.4.

Part 1: We fix $t > 0$ and let $n \geq 1$. Let Π be a partition of $[0, t]$,

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$$

We have

$$W_{t_i}(W_{t_{i+1}} - W_{t_i}) = \frac{1}{2}W_{t_{i+1}}^2 - \frac{1}{2}W_{t_i}^2 - \frac{1}{2}(W_{t_{i+1}} - W_{t_i})^2$$

Hence, if K_n is the is the largest i such that $\frac{i}{n} \leq t$, then

$$\begin{aligned}
\int_0^t W_s\,\mathrm{d}W_s &= \lim_{n\uparrow\infty} \sum_{k=0}^{n-1} W_{t_i}(W_{t_{i+1}} - W_{t_i}) \\
&= \frac{1}{2}\lim_{\|\Pi\|\to 0} \Bigg(\underbrace{\sum_{i=0}^{n-1} \Big((W_{t_{i+1}})^2 - (W_{t_i})^2 \Big)}_{= W_{t_n}^2 - W_0^2} - \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 \Bigg) \\
&= \frac{1}{2}W_t^2 - \frac{1}{2}t
\end{aligned}$$

since

$$\lim_{\|\Pi\|\to 0} \sum_{i=0}^{n} (W_{t_{i+1}} - W_{t_i})^2 = t$$

using the quadratic variation of Brownian motion.

Part 2: The integrability conditions are clear. We have using the properties of conditional expectation and independent increments,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{2}W_t^2 - \frac{1}{2}t \mid \mathcal{F}_s\right] &= -\frac{1}{2}t + \frac{1}{2}\mathbb{E}\left[(W_t - W_s + W_s)^2 \mid \mathcal{F}_s\right] \\
&= -\frac{1}{2}t + \frac{1}{2}\mathbb{E}\left[(W_t - W_s)^2 \mid \mathcal{F}_s\right] - \mathbb{E}\left[(W_t - W_s)W_s \mid \mathcal{F}_s\right] + \frac{1}{2}\mathbb{E}\left[W_s^2 \mid \mathcal{F}_s\right] \\
&= -\frac{1}{2}t + \frac{1}{2}\mathbb{E}\left[(W_t - W_s)^2\right] - W_s\,\mathbb{E}\left[W_t - W_s\right] + \frac{1}{2}W_s^2 \\
&= -\frac{1}{2}t + \frac{1}{2}(t - s) + \frac{1}{2}W_s^2 = \frac{1}{2}W_s^2 - \frac{1}{2}s.
\end{aligned}$$

2 Itô's lemma

Itô's lemma is the counterpart of the chain rule for stochastic calculus. The main difference is the inclusion of second order correction terms.

2.1 Ito's Lemma for Brownian Motion

Recall the substitution rule,

$$f(G(t)) - f(G(0)) = \int_0^t f'(G'(s))\, dG(s)$$

or written as differentials

$$df(G(t)) = f'(G(t))G'(t)dt = f'(G(t))dG(t).$$

Itô's lemma is essentially the chain rule with a second order correction term.

Theorem 2.1 (Basic Itô Lemma)

Suppose that $\{W_t\}_{t\geq 0}$ is a Brownian motion $f : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable. Then

$$f(W_t) - f(W_0) = \int_0^t f'(W_s)\, \mathrm{d}W_s + \frac{1}{2}\int_0^t f''(W_s)\, \mathrm{d}s \tag{3}$$

or equivalently,

$$df(W(t)) = f'(W(t))\, dW(t) + \frac{1}{2} f''(W_t)\, dt.$$

Remark 2.2. The red term in (3) is a second-order correction term that appears only with functions of non-vanishing quadratic variation.

The first integral in (3) is a stochastic integral, while the second integral in (3) is the usual Riemann–Stieltjes integral with respect to time. Intuitively, the second order term is very similar to the second order correction term from Taylor's Theorem

$$f(x) = f(0) + f'(x)x + \frac{1}{2} f''(x)x^2 + o(x).$$

A heuristic derivation of Itô's lemma is a direct consequence of Taylor's Theorem (see Problem 2.1).

2.2 Time Dependent Itô's lemma

In applications, we often need to consider functions that depend on time as well. Recall the multivariate chain rule

$$df(t, G(t)) = f_x(t, G(t))G'(t)\, dt + f_t(t, G(t))\, dt = f_x(t, G(t))\, dG(t) + f_t(t, G(t))\, dt.$$

The time dependent Itô's lemma is essentially the multivariate chain rule with a second order correction term.

Theorem 2.3 (Time Dependent Itô Lemma)

Suppose that $\{W_t\}_{t\geq 0}$ is a Brownian motion $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable in space and continuously differentiable in time. Then

$$f(t, W_t) - f(0, W_0) = \int_0^t f_t(s, W_s)\,\mathrm{d}s + \int_0^t f_x(s, W_s)\,\mathrm{d}W_s + \frac{1}{2}\int_0^t f_{xx}(s, W_s)\,\mathrm{d}s \tag{4}$$

or equivalently,

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

Again, the key difference in (4) is the second order correction term highlighted in red. Notice that this formula also contains (2.1) since if f was not dependent on t, then $f_t = 0$. A heuristic derivation of Itô's lemma is a direct consequence of Taylor's Theorem (see Problem 2.1).

2.3 General Itô's lemma

We now consider integration against more complicated stochastic processes.

Definition 2.4. Let $\{W(t)\}$ be a Brownian motion with filtration $\mathcal{F}_t$. An **Itô Process** is a stochastic process of the form

$$X(t) = X(0) + \int_0^t \Delta(s)\,dW(s) + \int_0^t \Theta(s)\,ds \tag{5}$$

where $\Delta(t)$ and $\Theta(t)$ are adapted stochastic processes with respect to $\mathcal{F}_t$. Written in differentials, we have that

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

We assume that $\Delta(t)$ and $\Theta(t)$ are almost surely square integrable.

Remark 2.5. Almost all stochastic processes, except those that have jumps, are Itô processes.

The quadratic variation of the Itô process differs from Brownian motion.

Proposition 2.6

The quadratic variation of an Itô process (5) is

$$[X, X]_t = \int_0^t \Delta^2(s)\,ds.$$

This is not a surprising result since this is equal to the quadratic variation of the stochastic integral, and the Riemann integral term (although still random) in (5) is nice enough to not contribute the quadratic variation. The stochastic integral against $X(t)$ is defined in the obvious way.

Definition 2.7. Let $X(t)$ be an Itô process and let $\Gamma(t)$ be an adapted process. We define the **stochastic integral with respect to** X by

$$\int_0^t \Gamma(s)\,dX(s) = \int_0^t \Gamma(s)\Delta(s)\,dX(s) + \int_0^t \Gamma(s)\Theta(s)\,ds$$

We now state Itô's lemma for Itô processes.

Theorem 2.8 (***Itô's lemma***)

Suppose that $\{X_t\}_{t\geq 0}$ *is an Itô process and* $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ *is twice continuously differentiable in space and continuously differentiable in time. Then*

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t f_t(s, X_s)\,\mathrm{d}s + \int_0^t f_x(s, X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t f_{xx}(s, X_s)\,\mathrm{d}[X, X](s) \qquad (6)\\ &= \int_0^t f_t(s, X_s)\,\mathrm{d}s + \int_0^t f_x(s, X_s)\Delta(s)\,\mathrm{d}W_s + \int_0^t f_x(s, X_s)\Theta(s)\,\mathrm{d}s \\ &\quad + \frac{1}{2}\int_0^t f_{xx}(s, X_s)\Delta^2(s)\,\mathrm{d}s \end{aligned}$$

Notice that this formula contains both Theorem 2.1 and Theorem 2.3. This is because we can simply take $\Theta = 0$ and $\Delta = 1$ then $X(t)$ becomes Brownian motion and the formula simplifies.

Remark 2.9. A convenient way to remember Itô's lemma is to do a Taylor expansion (up to the second order terms) using differentials,

$$df = f_t\,dt + f_x\,dX_t + \frac{1}{2}f_{xx}dX_t dX_t.$$

Then substituting $dX = \Delta dW + \Theta dt$ gives us

$$df = f_t\,dt + f_x\Delta\,dW_t + f_x\Theta\,dt + \frac{1}{2}f_{xx}\Delta^2 dt$$

where we used the fact that

$$dX_t dX_t = \Delta^2 \underbrace{dW dW}_{=dt} + 2\Delta\Theta \underbrace{dW dt}_{=0} + \Theta^2 \underbrace{dt dt}_{=0} = \Delta^2 dt.$$

2.4 Integration by Parts

We now consider the stochastic analogue of integration by parts. Recall the rule for integration by parts,

$$\int_0^t f(s)dg(s) = f(t)g(t) - f(0)g(0) - \int_0^t f(s)dg(s)$$

which is equivalent to the product rule when written in differentials (after some rearranging)

$$d(fg) = fdg + gdf.$$

In applications, we typically take $X_t = f(W_t)$ and $Y_t = g(W_t)$ then

$$dX_t = f'(W_t)\,dW_t + \frac{1}{2}f''(W_t)\,dt \qquad \text{and} \qquad dY_t = g'(W_t)\,dW_t + \frac{1}{2}g''(W_t)\,dt$$

are Itô processes. The analogue of the integration by parts formula is

$$f(W_s)g(W_s) - f(0)g(0) = \int_0^t g(W_s)\,df(W_s) + \int_0^t f(W_s)\,dg(W_s) + \frac{1}{2}\int_0^t f'(W_s)g'(W_s)\,ds$$

The derivation of this result when $X_t = f(W_t)$ and $Y_t = g(W_t)$ is a direct application of Itô's lemma to the product $h(x) = f(x)g(x)$ and it is done in Problem 2.3. The following formula is the general formula to compute the differentials of products of Itô processes.

Theorem 2.10 (***Integration by Parts***)

Consider Itô processes

$$dX_t = \xi_t dW_t + \eta_t \, dt \qquad \text{and} \qquad dY_t = \zeta_t dW_t + \theta_t dt,$$

with respect to the same Brownian motion. Then

$$\begin{aligned} X_tY_t - X_0Y_0 &= \int_0^t X_s \, \mathrm{d}Y_s + \int_0^t Y_s \, \mathrm{d}X_s + \int_0^t \mathrm{d}[X,Y](s) \qquad (7) \\ &= \int_0^t X(s)\zeta(s) \, \mathrm{d}W(s) + \int_0^t X(s)\theta(s) \, \mathrm{d}W(s) \, ds + \int_0^t Y_s\xi(s) \, \mathrm{d}W_s + \int_0^t Y(s)\eta(s) \, \mathrm{d}s \\ &\quad + \int_0^t \xi(s)\zeta(s) \, \mathrm{d}s \end{aligned}$$

or equivalently

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

2.5 Example Problems

2.5.1 Proofs of Main Results

Problem 2.1. Derive the time dependent Itô's lemma in Proposition 2.3.

Solution 2.1. We do a heuristic derivation of Itô's lemma. By Taylor's theorem up to the second order, we have

$$df(t,W_t) = f_t(t,W_t) \, dt + f_x(t,W-t) \, dW_t + \frac{1}{2} f_{tt}(t,W_t) \, dtdt + f_{tx}(t,W_t) \, dtdW_t + \frac{1}{2} f_{xx}(t,W_t) \, dW_tdW_t$$

where we ignored the third order terms since they are all of a smaller scale. We have the following limits of the quadratic variation and cross variation,

- For a partition Π of $[0,t]$ we have

$$dW_tdW_t = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t$$

- For a partition Π of $[0,t]$ we have

$$dtdW_t = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \underbrace{(W_{t_{i+1}} - W_{t_i})}_{\le \max_i (W_{t_{i+1}} - W_{t_i})} (t_{i+1} - t_i) \le \lim_{\|\Pi\|\to 0} \max_{i \le n-1} (W_{t_{i+1}} - W_{t_i}) \cdot t = 0.$$

  by continuity of Brownian motion and the fact that $\sum_{i=0}^{n-1} (t_{i+1} - t_i) = t$.

- For a partition Π of $[0,t]$ we have

$$dtdt = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \underbrace{(t_{i+1} - t_i)}_{\le \max_i (t_{i+1} - t_i)} (t_{i+1} - t_i) = \lim_{\|\Pi\|\to 0} \max_{i \le n} (t_{i+1} - t_i) \cdot t = 0$$

  by continuity and the fact that $\sum_{i=0}^{n-1} (t_{i+1} - t_i) = t$.

We conclude that

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

as required.

Remark 2.11. A complete rigorous proof of this fact follows the same strategy, but we start from the definition and use much more care to control the error terms as we take the mesh size to 0.

Problem 2.2. Derive Itô's isometry,

$$\mathbb{E}\left[\left(\int_0^t \xi(s)\,dW(s)\right)^2\right] = \mathbb{E}\left[\int_0^t \xi^2(s)\,ds\right]$$

Solution 2.2. We prove this fact using Itô's lemma. We have that

$$dX(s) = \xi(s)dW_s$$

is an Itô process. Applying Itô's lemma to $f(x) = x^2$ implies that

$$\left(\int_0^t \xi(s)\,dW(s)\right)^2 = X^2(s) = 2\int_0^t W_s\,dX_s + \int_0^t d[X,X](s) = 2\int_0^t W(s)\xi(s)\,dW(s) + \int_0^t \xi^2(s)\,ds.$$

Taking expected values implies

$$\mathbb{E}\left(\int_0^t \xi(s)\,dW(s)\right)^2 = \underbrace{\mathbb{E}\left[2\int_0^t W(s)\xi(s)\,dW(s)\right]}_{=0} + \mathbb{E}\left[\int_0^t \xi^2(s)\,ds\right] = \mathbb{E}\left[\int_0^t \xi^2(s)\,ds\right]$$

where we used the fact that

$$\mathbb{E}\left[2\int_0^t W(s)\xi(s)\,dW(s)\right] = 0$$

since $I(t) = 2\int_0^t W(s)\xi(s)\,dW(s)$ is a martingale and $I(0) = 0$.

Problem 2.3. Let $X_t = f(W_t)$ and $Y_t = g(W_t)$. Prove the integration by parts formula in Proposition 2.10 in this setting.

Solution 2.3. If we let $h(x) = f(x)g(x)$ then

$$h' = f'g + fg' \qquad \text{and} \qquad h'' = f''g + g''f + 2f'g'.$$

Itô's lemma implies that

$$\begin{aligned}
X_tY_t - X_0Y_0 &= f(W_t)g(W_t) - f(0)g(0)\\
&= \int_0^t f'(W_s)g(W_s)\,dW_s + \int_0^t f(W_s)g'(W_s)\,dW_s\\
&\quad + \frac{1}{2}\int_0^t f''(W_s)g(W_s) + g''(W_s)f(W_s) + 2f'(W_s)g'(W_s)\,ds\\
&= \int_0^t g(W_s)\,dX_s + \int_0^t f(W_s)\,dY_s + \frac{1}{2}\int_0^t f'(W_s)g'(W_s)\,ds
\end{aligned}$$

where we used the fact that Itô's lemma implies that

$$dX_t = f'(W_t)\,dW_t + \frac{1}{2}f''(W_t)\,dt \qquad \text{and} \qquad dY_t = g'(W_t)\,dW_t + \frac{1}{2}g''(W_t)\,dt.$$

2.5.2 Applications

Problem 2.4. Compute the differentials of the following

1. $f(t, W_t) = W_t^2 - t$.
2. $f(t, W_t) = W_t^3 - 3tW_t$

Solution 2.4. All of these computations follow from a direct application of Itô's lemma.

Part 1: Let $f(t, x) = x^2 - t$. Then

$$\frac{\partial f}{\partial t}(t, x) = -1, \quad \frac{\partial f}{\partial x}(t, x) = 2x, \quad \frac{\partial^2 f}{\partial x^2}(t, x) = 2$$

So Itô's lemma implies that

$$W_t^2 - t = f(t, W_t) - f(0, W_0) = \int_0^T 2W_t \, \mathrm{d}W_t - \int_0^t 1 \mathrm{d}s + \frac{1}{2} \int_0^t 2 \, \mathrm{d}s = 2 \int_0^t W_t \, \mathrm{d}W_t.$$

That is,

$$df = 2W_t dW_t$$

Part 2: Let $f(t, x) = x^3 - 3tx$. Then

$$\frac{\partial g}{\partial t}(t, x) = -3x, \quad \frac{\partial g}{\partial x}(t, x) = 3x^2 - 3t, \quad \frac{\partial^2 g}{\partial x^2}(t, x) = 6x$$

So Itô's lemma implies that

$$\begin{aligned} W_t^3 - 3tW_t &= f(t, W_t) - f(0, W_0) \\ &= \int_0^t (3W_s^2 - 3s) \, \mathrm{d}W_s - \int_0^t 3W_s \, \mathrm{d}s + \frac{1}{2} \int_0^t 6W_s \, \mathrm{d}s \\ &= 3 \int_0^t (W_s^2 - s) \, \mathrm{d}W_s. \end{aligned}$$

That is,

$$df = 3(W_t^2 - t) dW_t$$

Problem 2.5. Simplify

$$\int_0^t s dW(s).$$

Solution 2.5. We pick a function so that its derivative in x gives us the integrand. In particular, if we let $f(t, x) = tx$. Then

$$\frac{\partial f}{\partial t}(t, x) = x, \quad \frac{\partial f}{\partial x}(t, x) = t, \quad \frac{\partial^2 f}{\partial x^2}(t, x) = 0.$$

Itô's lemma implies that

$$\begin{aligned} tW_t &= f(t, W_t) - f(0, W_0) \\ &= \int_0^t s \, \mathrm{d}W_s + \int_0^t W_s \, \mathrm{d}s + \frac{1}{2} \int_0^t 0 \, \mathrm{d}s. \end{aligned}$$

Rearranging implies that

$$\int_0^t s \, \mathrm{d}W_s = tW_t - \int_0^t W_s \, \mathrm{d}s.$$

Remark 2.12. This computation should remind us of the integration by parts formula. Indeed, we have that by integration by parts,

$$\int_0^t s dW_s = tW_t - 0W_0 - \int_0^t W_s ds - \int_0^t d[W,t](s) = tW_t - \int_0^t W_s\, ds$$

Problem 2.6. For some constant $\sigma > 0$, define

$$Z_t = e^{\sigma W_t - \frac{1}{2}\sigma^2 t}.$$

Show that Z_t is a martingale.

Solution 2.6. We verified this directly by hand in Week 11. We now use stochastic calculus to prove this fact. In particular, if we let $f(t,x) = e^{\sigma x - \frac{1}{2}\sigma^2 t}$. Then

$$\frac{\partial f}{\partial t}(t,x) = -\frac{\sigma^2}{2} e^{\sigma x - \frac{1}{2}\sigma^2 t}, \quad \frac{\partial f}{\partial x}(t,x) = \sigma e^{\sigma x - \frac{1}{2}\sigma^2 t}, \quad \frac{\partial^2 f}{\partial x^2}(t,x) = \sigma^2 e^{\sigma x - \frac{1}{2}\sigma^2 t}.$$

So

$$\begin{aligned} e^{\sigma W_t - \frac{1}{2}\sigma^2 t} - 1 &= f(t, W_t) - f(0, W_0) \\ &= \int_0^t \sigma e^{\sigma x - \frac{1}{2}\sigma^2 t}\, \mathrm{d}W_s - \int_0^t \frac{\sigma^2}{2} e^{\sigma x - \frac{1}{2}\sigma^2 t}\, \mathrm{d}s + \frac{1}{2}\int_0^t \sigma^2 e^{\sigma x - \frac{1}{2}\sigma^2 t}\, \mathrm{d}s \\ &= \int_0^t \sigma e^{\sigma x - \frac{1}{2}\sigma^2 t}\, \mathrm{d}W_s. \end{aligned}$$

Rearranging implies that

$$Z_t = e^{\sigma W_t - \frac{1}{2}\sigma^2 t} = \int_0^t \sigma e^{\sigma x - \frac{1}{2}\sigma^2 t}\, \mathrm{d}W_s + 1.$$

We have that $\int_0^t \sigma e^{\sigma x - \frac{1}{2}\sigma^2 t}\, \mathrm{d}W_s$ is a martingale since $\sigma e^{\sigma x - \frac{1}{2}\sigma^2 t}$ is an adapted (integrable) process and 1 is clearly a martingale. The sum of martingales is a martingale, so Z_t is a martingale.