An Introduction to Stochastic Calculus

Olivier Lefevre

May 31st 2026

Abstract

Stochastic Calculus is an important development in the space of applications of integral and differential calculus onto stochastic processes. The first main application was in finance, when Louis Bachelier first developed some theory in 1900 and used it to analyze the Parisian Bourse [Cou02]. In this report, we present a brief introduction to Stochastic Calculus, focusing on the development of Ito's Calculus. We begin with reviewing some preliminaries from Probability Theory. We then present some of the nice properties of Brownian Motion, and importantly, what makes it so difficult to perform calculus on. We then present Ito's integral, Ito's isometry, and Ito's formula, key results in stochastic calculus. To conclude the report, we present an application, where we leverage Stochastic Calculus and Reinforcement learning to solve a molecular dynamics problem.

Contents

1 Preliminaries

1.1 Martingales

Definition 1.1 ($\sigma-$algebra)**.** *We call $\mathcal{F}$, a collection of subsets of Ω, a $\sigma-$algebra if*

1. $\varnothing \in \mathcal{F}$
2. $\forall A \in \mathcal{F}, A^c \in \mathcal{F}$
3. *If* $\{A_n : n \geqslant 1\} \subseteq \mathcal{F}$ *then* $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

Definition 1.2 (Filtration)**.** *A collection $\{\mathcal{F}_n : n \geqslant 1\}$ of $\sigma-$algebras on Ω is a filtration if $\forall n \leqslant m, \mathcal{F}_n \subset \mathcal{F}_m$.*

Definition 1.3 (Martingale)**.** *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ be a filtered probability space and $\{X_n : n \geqslant 1\}$ a collection of random variables on Ω. Then we say $\{X_n\}$ is a $\{\mathcal{F}_n\}$-martingale if*

1. $\forall n \geqslant 1, \sigma(X_n) \subset \mathcal{F}_n$
2. $\forall n \geqslant 1, X_n \in L^1$
3. $\forall n \geqslant 1, \mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$

If 3. is replaced with

$$\forall n \geqslant 1, \mathbb{E}[X_{n+1}|\mathcal{F}_n] \leqslant (\geqslant) X_n$$

Then we call $\{X_n\}$ a $\{\mathcal{F}_n\}$-super-martingale (sub-martingale) respectively.

Definition 1.4. *We call $\tau : \Omega \to \mathbb{N} \cup \infty$ a $\{\mathcal{F}_n : n \geqslant 1\}-$stopping time if $\forall n \geqslant 1$*

$$\{\tau \leqslant n\} \in \mathcal{F}_n$$

Theorem 1.5 (Stopped processes)**.** *Let $\{X_n\}$ be a $\{\mathcal{F}_n\}-$ martingale, then for a $\{\mathcal{F}_n\}-$stopping time τ, $\{X_{\tau\wedge n}\}$ is also a $\{\mathcal{F}_n\}-$martingale*

1.2 Construction of Brownian motion

Definition 1.6 (Gaussian vector)**.** *For $\Sigma \in \mathbb{R}^{n\times n}$ positive definite, and $\mu \in \mathbb{R}^n$, a gaussian vector with mean μ and covariance matrix Σ is one with density function*

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)}$$

Essentially, it's simply a vector with probability density function equal to the joint probability density function of n gaussian distributed random variables.

Definition 1.7 (Standard Brownian Motion)**.** *For $T > 0$, Standard Brownian Motion (SBM) is a collection of random variables $\{X_s\}, s \in [0, T]$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that:*

1. *$X_0 = 0$ $\mathbb{P}-$a.s*
2. *$\forall\ 0 \leqslant s_1 \leqslant s_2 \leqslant t_1 \leqslant t_2 \leqslant T$ $(X_{s_2} - X_{s_1})$ is independent of $(X_{t_2} - X_{t_1})$*
3. *$\forall\ 0 \leqslant s \leqslant t \leqslant T$ $X_t - X_s \sim \mathcal{N}(0, t - s)$*
4. *$X : t \to X_t(\omega)$ is $\mathbb{P}-$almost surely continuous.*

Brownian motion is actually an example of a Gaussian process. It is one of the most important examples.

Definition 1.8 (Gaussian process)**.** *$\{X_t\}_{t\in[0,T]}$ is a gaussian process if and only if for any choice $0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant T$*

$$(X_{t_1}, X_{t_2}, \cdots, X_{t_n})^\top \textit{ is a gaussian vector.}$$

Brownian motion is a gaussian process ($X_t = X_t - X_0 \sim \mathcal{N}(0, t)$). What distinguishes it from other gaussian processes is the unique covariance structure.

Proposition 1.9. *Let X be a SBM on [0,T], then for $s, t \in [0, T]$*

$$\mathrm{Cov}(X_s, X_t) = s \wedge t$$

Definition 1.10 (Complete orthonormal basis)**.** *Let $(X, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then $\{\phi_n\}_{n\in\mathbb{N}}$ is a complete orthonormal basis if*

1. *$\forall i \neq j$, $\langle \phi_i, \phi_j \rangle = 0$*
2. *span$(\phi_i)_{i\in\mathbb{N}}$ is dense in X*

Now, to construct the Brownian motion, simply by the definition, a natural candidate would be for $\{Z_n\}_{n\in\mathbb{N}}$ a collection of mean zero independent gaussian random variables with variance 1

$$B_t = \sum_{n=0}^{\infty} a_n(t) Z_n$$

By Fubini, we know $\mathbb{E}[\sum_{n=0}^{\infty} a_n(t)Z_n] = 0$, and the covariance is given by

$$\mathbb{E}[B_t B_s] = \sum_{n=0}^{\infty} a_n(t)a_n(s)\mathbb{E}[Z_n Z_n] = \sum_{n=0}^{\infty} a_n(t)a_n(s)$$

So we are looking for $a_n : [0, T] \to \mathbb{R}$ such that $a_n(s)a_n(t) = s \wedge t$. By looking at the Hilbert space $L^2([0, T])$, with the inner product defined as

$$f, g = \int_0^T f(s)g(s)ds$$

For $f, g \in L^2([0, T])$. Now notice by Parsevla, we get that

$$f, g = \sum_{n=0}^{\infty} f, \phi_n g, \phi_n$$

So taking $f = \mathbb{1}_{[0,t]}$ and $g = \mathbb{1}_{[0,s]}$ we have that

$$f, g = \int_0^{s\wedge t} dx = \sum_{n=0}^{\infty} \left(\int_0^s \phi_n dx\right)\left(\int_0^t \phi_n dx\right)$$

Therefore a natural candidate for our a_n is $a_n(s) = (\int_0^s \phi_n dx)$ which would give the correct covariance. By taking some orthonormal basis (such as Haar wavelets) we can show that we can find a sequence of simple functions approaching a Brownian motion. For more details, refer to [Ben24]

2 Properties of Brownian Motion

Brownian motion is quite a nice object as it has some nice properties, for example stretching or shrinking a SBM is still a SBM.

2.1 Scaling properties

Proposition 2.1 (Stretching the scale)**.** *Let B be a SBM, and a>0. Then $X = (X_t)_{t\geqslant 0}$*

$$X_t = \frac{1}{\sqrt{a}} B_{at}$$

Is also a SBM with the same distribution.

Proposition 2.2 (Shrinking the scale)**.** *Let B be a SBM, then $X = (X_t)_{t\geqslant 0}$*

$$X_t = tB_{\frac{1}{t}}$$

Is a SBM of the same distribution.

2.2 Cutoff property

Proposition 2.3. *Let B be a SBM, then for $s \geqslant 0$*

$$X_t = B_{t+s} - B_s$$

Is also a SBM, in particular, X and $\sigma((B_t)_{0\leqslant t\leqslant s})$ are independent.

Notice that here, independence comes from the fact that $B_t = B_t - B_0$, and we know $B_t - B_0 \perp B_{t_1+s} - B_s$ for $t \leqslant s$.

2.3 Bounding the trajectory

Importantly, we can see that the limiting behavior of the Brownian motion is problematic, in fact it changes direction infinitely often.

Proposition 2.4 (Iterated log identities)**.** *For a SBM $B = (B_t)_{t\geqslant 0}$, we have the following properties for the limiting behavior towards ∞ and 0^+. In particular*

$$\mathbb{P}\left(\limsup_{t\to 0^+} \frac{B_t}{\sqrt{2t \ln\ln \frac{1}{t}}} = 1\right) = 1$$

And by symmetry of distribution

$$\mathbb{P}\left(\liminf_{t\to 0^+} \frac{B_t}{\sqrt{2t \ln\ln \frac{1}{t}}} = -1\right) = 1$$

And for the behavior towards ∞:

$$\mathbb{P}\left(\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\ln\ln t}} = 1\right) = \mathbb{P}\left(\liminf_{t\to\infty} \frac{B_t}{\sqrt{2t\ln\ln t}} = -1\right) = 1$$

That limiting behavior gives rise to the following corollary which ends up being very problematic. The motion crosses the x-axis finitely often within any arbitrarily small interval after a first crossing.

Corollary 2.5. *For any $\epsilon > 0$*

$$\mathbb{P}(\inf_{0\leqslant s\leqslant\epsilon} B_s < 0, \sup_{0\leqslant t\leqslant\epsilon} B_t > 0) = 1$$

Even further, we have that the set of zeroes of a SBM is closed and has no isolated points.

Proposition 2.6. *For a SBM B, let $\mathcal{Z}$ be the set of zeroes, i.e $\mathcal{Z} = \{t \in [0,1] : B_t = 0\}$. Then $\mathcal{Z}$ is almost surely closed, compact, has no isolated points and has measure 0 on* $[0,1]$.

Now this gives rise to a problem, Brownian motion is almost surely differentiable nowhere, even though it is continuous. In particular, it has unbounded total variation, which we can see quite easily.

Proposition 2.7 (Unbounded total variation)**.** *For any partition $\pi = \{t_0, \dots, t_m\}$ of $[0,t]$ such that $0 = t_0 < t_1 < \dots < t_m = M$ for some $M \in \mathbb{R}$, $m \in \mathbb{N}$ with*

$$|\pi| = \max_{0 \leqslant i \leqslant m-1} |t_{i+1} - t_i|$$

We have that

$$\lim_{|\pi| \to 0^+} \sum_{i=0}^{m-1} (B_{t_{i+1}} - B_{t_i})^2 \to^{L^2(\mathbb{P})} M$$

In particular, we have

$$\lim_{|\pi| \to 0^+} \sum_{i=0}^{m-1} |B_{t_{i+1}} - B_{t_i}| \to \infty$$

Proof. Since

$$\lim_{|\pi| \to 0^+} \sum_{i=0}^{m-1} (B_{t_{i+1}} - B_{t_i})^2 \to^{L^2(\mathbb{P})} M$$

We also know that

$$\sum_{i=0}^{m-1} (B_{t_{i+1}} - B_{t_i})^2 \leqslant \max_{0 \leqslant i \leqslant m-1} |B_{t_{i+1}} - B_{t_i}| \cdot \sum_{i=0}^{m-1} |B_{t_{i+1}} - B_{t_i}|$$

By continuity, we know that

$$\lim_{|\pi| \to 0^+} \max_{0 \leqslant i \leqslant m-1} |B_{t_{i+1}} - B_{t_i}| = 0$$

Which means that $\sum_{i=0}^{m-1} |B_{t_{i+1}} - B_{t_i}| \to \infty$. □

3 Ito's integral

3.1 The goal

The goal of this section is to define Ito's integral, i.e for a function $X : \Omega \times [0,T] \to \mathbb{R}$ and some SBM B_t, we want to find

$$I(X)(\omega) = \int_0^T X(t,\omega) dB_t(\omega)$$

Here are some examples for intuition. If we take some Brownian motion, and say that as a function of this noise, we gain 2 dollars in proporition to the noise. Then we want to find the accumulated gain, i.e

$$I(3)(\omega) = \int_0^T 3dB_t(\omega)$$

However, we have the freedom to define it as a random function. We could for example, weigh the steps of a Brownian motion, by the Brownian motion itself, i.e

$$I(B_t)(\omega) = \int_0^T B_t dB_t$$

Essentially, it's a way of integrating functions over a random path, in this case, Brownian motion. And as we saw, we'll need to define a special class of functions that this can be well defined over, especially since as we say, Brownian motion is almost surely differentiable nowhere. Let's define the class of functions we can define Brownian motion over.

Definition 3.1 (H^2)**.** *Take $\mathcal{B} := \mathbb{B}([0,T])$ to be the standard Borel $\sigma-$algebra, and $\mathcal{F}_t$ the standard brownian filtration. Then we uniquely define $\mathcal{F}_t \times \mathcal{B}$ to be the smallest $\sigma-$algebra that contains all $A \times B, A \in \mathcal{F}_t, B \in \mathcal{B}$.*

$$H^2([0,T]) := \left\{ X : X \textit{ is } \mathcal{F}_t \times \mathcal{B} \textit{ - measurable, is } \{\mathcal{F}_t\} \textit{ adapted and } \mathbb{E}\left[\int_0^T X_t(\omega)^2 dt\right] < \infty \right\}$$

Eventually, we will show that this class is also bounded under the linear functional I.

3.2 Simple functions

As with the construction of the Lebesgue integral, it's often easier to build up from simple functions. It is also known from functional analysis that simple functions are dense in $L^p(X,\mu)$, so taking $X := \mathcal{F}_t \times \mathcal{B}$ and $\mu := d\mathbb{P} \times d\lambda$, with λ the Lebesgue measure, we can extend all of our results by Monotone class arguments.

Definition 3.2 (H_0^2)**.** *We denote H_0^2 the subset of H^2 consisting of simple functions. I.E $H_0^2 = \{X : X \in H^2, X \textit{ simple}\}$. Specifically, we define it as X such that $\exists n \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_n = T$ and $a_1, \ldots, a_{n-1}$ with a_i $\mathcal{F}_{t_i}-$measurable. With $\mathbb{E}[a_i^2] < \infty$ such that*

$$X(t,\omega) = \sum_{i=1}^{n-1} a_i(\omega)\mathbb{1}_{[t_i,t_{i+1}]}(t)$$

Notice we define the $a_i - s$ to be adapted to $\mathcal{F}_t$.

Now, let's define the integral for this dense subclass. However, we will notice it turns out we have already done this! It turns out by Ito's isometry that $\mathbb{E}[\int_0^T X_t(\omega)^2 dt] = \mathbb{E}[I(X)^2]$! But where does this projection come from?

Proposition 3.3 (Ito's isometry for simple functions)**.** *For $X \in H_0^2$,*

$$\mathbb{E}[I(X)^2] = ||I(X)||^2_{L^2(d\mathbb{P})} = ||X||^2_{L^2(d\mathbb{P}\times dt)} = \mathbb{E}\left[\int_0^T X_t(\omega)^2 dt\right]$$

3.3 General functions

Since we know that on simple functions, I is an isometry, we therefore get that I is Lipschitz continuous, and therefore maps Cauchy sequences to Cauchy sequences. Through this, we can get the theorem to hold on all functions in H^2.

Proposition 3.4 (Ito's isometry)**.** *For $X \in H^2$*

$$||I(X)||^2_{L^2(d\mathbb{P})} = ||X||^2_{L^2(d\mathbb{P}\times dt)}$$

Something else we might want to do is turn $I(X)(\omega) = \int_0^T X_t(\omega) dB_t(\omega)$ into a process, i.e $I(X)_{t\geqslant 0}$, with

$$I(X)_t = \int_0^t X_t(\omega) dB_t(\omega)$$

So now, each $\omega \in \Omega$ defines a determined process $I(X)_t$, i.e a new path. The idea is simple, we want to construct

$$Z_t = I(\mathbb{1}_{[0,t]}X) \in L^2(d\mathbb{P})$$

The problem with this is that we want a process that is continuous, but since $\mathbb{1}_{[0,t]}X$ is only defined as $\mathbb{P}-$a.e finite, Z_t might not be well defined. However, it turns out that we can find such a $\tilde{Z}_t$, to have our desired properties, and in fact, it is exactly what we have constructed. We also get another nice property, that Ito's isometry holds under conditional expectations.

Proposition 3.5 (Conditional Ito's)**.** *For $X \in H^2$, and $0 \leqslant s \leqslant t$, then*

$$\mathbb{E}\left[\left(\int_s^t X_u dB_u\right)^2 |\mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t X_u^2 du|\mathcal{F}_s\right]$$

3.4 Extending to almost surely finite variation

In some cases, we find that H^2 is too restrictive, so we want to extend the definition of Ito's integral to processes that have $\mathbb{P}-$almost sure finite variation.

Definition 3.6 ($\mathcal{L}^2_{loc}$)**.** *We denote*

$$\mathcal{L}^2_{loc} := \{X : X \textit{ measurable, adapted and } \mathbb{P}\left(\int_0^T X_t^2 dt < \infty\right) = 1\} \supset H^2$$

We then link $\mathcal{L}^2_{loc}$ and H^2 by noticing that for such processes that have almost surely finite variance, if we add a stopping time τ, that bounds them below n then X_τ is in H^2. Then by taking an increasing sequence, we can find a continuous function in H^2 that approximates X in $\mathcal{L}^{\in}{}_{loc}$.

Proposition 3.7 (Localizing stopping times)**.** *For $X \in \mathcal{L}^2_{loc}$, and $\{\tau_n\}_{n\geqslant 0}$ a sequence of stopping times defined as*

$$\tau_n = \inf\left\{t \in [0,T] : t = T \textit{ or } \int_0^T X_t^2 dt \geqslant n\right\}$$

is localizing, i.e we have that $\forall n \in \mathbb{N}$:

$$X_{n,t}(\omega) := \mathbb{1}_{t\leqslant\tau_n}(\omega)X_t(\omega) \in H^2$$

And that

$$\mathbb{P}(\bigcup_{n=1}^{\infty}\{\tau_n = T\}) = 1$$

Now we won't detail it here, but it seems pertinent to mention that it turns out for any localizing stopping time, we can find a unique continuous process $\mathbb{P}(Z_t = \lim_{n\to\infty} X_{n,t}) = 1$. We can therefore extend Ito's integral to all functions in $\mathcal{L}^2_{loc}$.

4 Ito's formula

4.1 For functions of Brownian motion

Proposition 4.1 (Ito's formula)**.** *For a C^2 function $f : \mathbb{R} \to \mathbb{R}$ we have*

$$f(B_t) = f(0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds$$

Which we sometimes write

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt$$

There is some nice intuition for this formula based on Taylor's expansion for C^2 functions. For some $f \in C^2(\mathbb{R})$ and two points x, y Taylor's expansion gives

$$f(y) - f(x) = (y-x)f'(x) + \frac{(y-x)^2}{2}f''(x) + r(x,y)$$

Where $r(x,y)$ is the remainder, formally:

$$r(x,y) = \int_x^y (y-u)(f''(u) - f''(x))du$$

What this gives us is a way of expressing the change of value of a function (i.e Reimann type approximation). How this differs for brownian motion, is that the $(B_{t_i} - B_{t_{i-1}})^2$ will tend to $t_i - t_{i-1}$, and so it will become an integral over dt. The other particularity, is that the remainder will tend to 0. This is because as the distance between x, y tend to 0, by Heine-Cantor we get that the modulus of continuity of a continuous function tends to 0. Recall, for some compact $A \subset \mathbb{R}$, the modulus of continuity of some continuous (so bounded on A) function f on A is

$$\sup_{x \in A} \sup_{y \in A} |f(x) - f(y)|$$

And Heine-Cantor tells us that continuous functions on compact sets are uniformly continuous (by extracting a finite subcover of open $\epsilon-$balls of A).

4.2 Proof outline

We want to show that

$$f(B_t) - f(0) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds$$

To do this, it's easier to show

$$\sum_{i=0}^{n} f(B_{i+1}) - f(B_i) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds$$

For some partition that we will refine. And by the expansion we get

$$\sum_{i=0}^{n} f(B_{i+1}) - f(B_i) = \sum_{i=0}^{n}(B_{i+1} - B_i)f'(B_i) + \frac{(B_{i+1} - B_i)^2}{2}f''(B_i) + r(B_{i+1,B_i})$$

4.2.1 Step 1

For the first step, we show

$$\sum_{i=0}^{n}(B_{i+1} - B_i)f'(B_i) \to \int_0^t f'(B_s)dB_s$$

To do this, we do 3 important things, we get a stopping time to bound B_t, $\tau_M = \inf\{t : |B_t| \geqslant M\}$. We, also consider a function f_M with $f_M(x) = f(x), \forall x \in [-M, M]$, and f_M is compactly supported and continuously twice differentiable.

We first construct a simple function approximation of $f_M(B_t)$, and with Ito's isometry we change from integrating against $dt \times d\mathbb{P}$ to dB_t. Which is computationally much simpler.

4.2.2 Step 2

For the second step, we show that

$$\frac{1}{2}\sum_{i=1}^{n}(B_{t_i} - B_{t_{i-1}})^2 f''(B_{t_{i-1}}) \to \frac{1}{2}\int_0^t f''(B_t)dt$$

The strategy for this section is to say that since $s \to f''(B_s)$ is continuous almost surely, we know that

$$\frac{1}{2}\sum_{i=1}^{n} f''(B_{t_{i-1}})(t_i - t_{i-1}) \to \frac{1}{2}\int_0^t f''(B_t)dt$$

So we can make this term appear by adding zero i.e considering

$$\tilde{B}_n := \frac{1}{2}\sum_{i=1}^{n} f''(B_{t_{i-1}})\left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\right) + \frac{1}{2}\sum_{i=1}^{n} f''(B_{t_{i-1}})(t_i - t_{i-1})$$

And just need to show the first term tends to zero. We can bound the squared norm by noticing that squaring this will cancel out all the cross terms by some indepedence arguments and since we are centering the $(B_{t_i} - B_{t_{i-1}})^2$ which approach $t_i - t_{i-1}$. And by bounding the function with the sup-norm, the difference between $\mathbb{E}\left[\left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})^2\right)\right] \to$

0.

4.2.3 Step 3

Here we simply show the remainder tends to zero by showing that

$$|\epsilon_n| \leqslant \sum_{i=1}^{n} \big(B_{t_i} - B_{t_{i-1}})^2 h(t_i, t_{i-1})\big)$$

We get this tends to zero by bounding the expectation of the absolute value with Cauchy Schwartz, and Markov. Look in the textbook for the exact steps.

4.2.4 Final Step

Now that we have our convergence in probability for each of the terms, for all $t \in [0,T]$, we have some $\Omega_t : \mathbb{P}(\Omega_t) = 1$, of a sequence tending almost surely (convergence in probability implies existence of a subsequence that tends almost surely). However, we do not know that $\mathbb{P}(\bigcap_{t\in[0,T]} \Omega_t) = 1$ as it is uncountable. So a classic trick, is to take Ω' the almost sure event of continuity of $s \to f(B_s)$ and extend it with

$$\bigcap_{t\in\mathbb{Q}\cap[0,T]} \Omega_t \cap \Omega'$$

Which has probability one, and by continuity, all we need to convergence for all rational $t \in [0,T]$.

4.3 Extension to functions of Brownian motion and time

Proposition 4.2. *For $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ a function in $C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$, then we have that*

$$f(t, B_t) = f(0,0) + \int_0^t \partial_t f(s, B_s) ds + \int_0^t \partial_x f(s, B_s) dB_s + \int_0^t \partial^2_{xx} f(s, B_s) ds$$

We won't detail the proof here, as it follows very closely to the proof for functions of Ito's formula for functions of just Brownian Motion.

4.4 Extension to Ito's Processes

Definition 4.3 (Ito's process)**.** *A process $\{X_t\}_{t\in[0,T]}$ is called an Ito's process if it can be written as*

$$X_t(\omega) = X_0 + \int_0^t a_s(\omega) ds + \int_0^t b_s(\omega) dB_s$$

With a, b adapted and measurable processes with

$$\mathbb{P}\left(\int_0^T |a_s|ds < \infty\right) = 1 \qquad \textit{and} \qquad \mathbb{P}\left(\int_0^T b_s dB_s < \infty\right) = \mathbb{P}\left(\int_0^T b_s^2 ds < \infty\right) = 1$$

We also write

$$dX_t(\omega) = a_t(\omega)ds + b_t(\omega)dB_s$$

It's nice to notice that SBM is an ito's process with $a : \Omega \to \mathbb{R}$ by $a(\omega) = 0$ and $b : \Omega \to \mathbb{R}$ by $b(\omega) = 1$. A SBM with drift is such that $a = \mu$ and $b = \sigma$.

Theorem 4.4 (Ito's formula for functions of Ito's processes)**.** *For $f \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R})$ and X an Ito's process then we have*

$$f(t, X_t) = f(0,0) + \int_0^t \partial_t f(t, X_t)dt + \int_0^t \partial_x f(t, X_t)dX_t + \frac{1}{2}\langle X\rangle \int_0^t \partial_{xx}^2 f(t, X_t)dt$$

Or

$$df(t, X_t) = \partial_t f(t, X_t)dt + \partial_x f(t, X_t)dX_t + \frac{1}{2}\partial_{xx} f(t, X_t)d\langle X\rangle_t$$

Where $\langle X\rangle_t$ denotes the quadratic variation of X.

5 Connecting Stochastic Optimal Control, and Reinforcement Learning

An application of Stochastic Calculus is the field of SOC (Stochastic Optimal Control), where we are optimizing some function of a stochastic process through controls. Formally, for some stochastic process $X : [0, T] \times \Omega \to \mathbb{R}^n$, and some sufficiently smooth function $f : \mathbb{R}^n \to \mathbb{R}$, where we have that

$$dX_t = (\mu(X_t) + u(X_t))dt + \sigma(X_t)dB_t$$

For some adapted μ, σ and B_t some SBM. We would search for

$$\arg\min_u \mathbb{E}[\int_0^T f(X_t)dt]$$

For which we can use Ito's formula under some conditions. We can see how stochastic calculus could help, as we have tools for dealing with stochastic integrals.

However, once we get this problem in a nice form, we still need some tools from Optimal Control. In particular, we can use tools from Reinforcement Learning, which is a sub-field of Optimal Control, where we seek to maximize expected reward over some space of controls.

5.1 Importance Sampling Problem in Stochastic Optimal Control

In molecular dynamics, we often want to calculate the expected value of some observable of a molecular system [Que+18]. To do this we collect ergodic trajectories. Ergodicity means that the distribution of states of a system over some infinite trajectory is the same as the distribution of initial states. In our case, we can find the expected value of some observable of our molecular system by simulating the trajectory of the system for a long enough time. Formally, for some system $X_t \in \mathbb{R}^n$, governed by the SDE

$$dX_t = -\nabla V(X_t)dt + \sigma(X_t)dB_t$$

For some $X_0 = x_0$, where X_t is the state of the system at time $t \geqslant 0$, B_t a n-dimensional standard Brownian motion and $V : \mathbb{R}^{\ltimes} \to \mathbb{R}$ a sufficiently smooth potential energy function. We would want to study the behavior of this system up to some stopping time. We can define such a stopping time τ as

$$\tau = \inf\{t > 0 : X_t \in \mathcal{T}\}$$

For some closed boundary $\mathcal{T} \subseteq \mathbb{R}^n$. For this type of system, we want to measure the expected value of some function of the path. In our case, for some $g : \mathbb{R}^n \to \mathbb{R}$, and $f : \mathbb{R}^n \to \mathbb{R}$, where f,g are sufficiently smooth and bounded, we want to measure

$$\Psi(x) := \mathbb{E}\left[\exp\left(-g(X_\tau) - \int_0^\tau f(X_t)dt\right)\middle| X_0 = x\right]$$

To find $\Psi(x)$, we could use a Monte Carlo estimator, i.e collect sample trajectories, and we have that

$$\sum_{n=1}^{N} \frac{\exp\left(-g(X_\tau) - \int_0^\tau f(X_t)dt\right)}{N} \to_{N\to\infty} = \Psi(x)$$

Since the average of iid samples of some distribution D tends to the average of the samples by the Strong Law of Large Numbers. However this estimator has extremely high variance, as molecular systems exhibit metastability.

Metastability intuitively means that the system is stuck in a small stable subregion of $\mathcal{T}^c \subseteq \mathbb{R}^n$. We can see this in the dynamics of the system, which are composed of a step in the direction of greatest descent of the potential energy function, as well as some Gaussian noise. We can therefore imagine the state reaches some local minimum $X \in \mathbb{R}^n$, i.e for some $r > 0$

$$\forall x \in B_r(X), V(x) \geqslant V(X) \iff V(x) - V(X) \geqslant 0$$

Therefore to escape this local state, we need the Brownian motion to ”push us out”, which rarely happens.

Due to this problem, sampled trajectories have high variance as the systems can get stuck in local minimums, making the Monte Carlo estimator a poor estimate of the true $\Psi(x)$. This is where the importance sampling problem comes from, we want to sample these trajectories in a better way, such that they are more representative of our desired value. In particular, we will see that we can find an unbiased estimator of $\Psi(x)$ with no variance.

In particular, when looking for

$$
\begin{aligned}
-\log \Psi(x) &= -\log \mathbb{E}_P[I(X_{0:\tau})] \\
&= -\log \mathbb{E}_Q[I(X_{0:\tau})m_{0:\tau}] \\
&= -\log \mathbb{E}_P\left[\exp\left(\int_0^\tau u_t dB_t - \frac{1}{2}\int_0^\tau |u_t|^2 dt - g(Y_t^u) - \int_0^\tau f(Y_t^u)dt\right)\right] \\
&\leqslant \mathbb{E}\left[\frac{1}{2}\int_0^\tau |u_t|^2 dt + g(Y_t^u) + \int_0^\tau f(Y_t^u)dt\right]
\end{aligned}
$$

Where we use Jensen's inequality [Øks03] . We also know we achieve inequality when the elements are constant. I.E we get equality here for some optimal control that gives zero variance for the stuff inside. So the control that minimizes the variance is the control that will give a zero variance estimator of $-\log \Psi(x)$, also called the free-path energy [Har+17]. We can see that we can find such an optimal control by seeing the above as

$$
\begin{aligned}
-\log \mathbb{E}_\pi[\int \exp(-W)dx] &= -\log \mathbb{E}_\rho[\int \exp(-W)\frac{\pi}{\rho}dx] \\
&= -\log \mathbb{E}_\rho[\int \exp(-W - \log \frac{\rho}{\pi})dx] \\
&\leqslant \mathbb{E}_\rho[\int W + \log \frac{\rho}{\pi}dx]
\end{aligned}
$$

Where we get equality for some optimal ρ^* such that

$$
-W - \log \frac{\rho^*}{\pi} = C
$$

And we can solve for $\frac{\rho^*}{\pi}$.

The approach to the sampling problem we will see here involves adding a drift term to our SDE.

Let's prove that a control actually induces a change of measure. In particular, we want to consider $Y_t \in \mathbb{R}^n$ for $t > 0$ governed by

$$
dY_t = (-\nabla V(Y_t) + \sigma(Y_t)u(Y_t))dt + \sigma(Y_t)dB_t^u
$$

where here u : $\mathbb{R}^n \to \mathbb{R}^n$ belongs to the space of time-independent feedback controls $\mathcal{U}$ (this makes u a $\mathbb{R}^n$ adapted stochastic process, i.e not forward looking). Here we are also consider the auxiliary process induced by u B_t^u defined by

$$
B_t^u = B_t - \int_0^t u_s ds
$$

However, under the probability measure $\mathbb{P}$, this is not a Brownian motion. Luckily for us, Girsanov's theorem [Øks03] tell us that $(B_t^u)_{\tau \geqslant t \geqslant 0}$ is a Brownian motion under a measure $\mathcal{Q}$ with Radon-Nikodym derivative

$$\left.\frac{d\mathcal{Q}}{d\mathbb{P}}\right|_{[0,\tau]} = \exp\left(\int_0^\tau u_t dB_t^u - \frac{1}{2}\int_0^\tau |u_t|^2 dt\right)$$

So we get that

$$\gamma = \mathbb{E}_{\mathcal{Q}}\left[g(Y_\tau) + \int_0^\tau f(Y_t) + |u_t|^2 dt\right]$$

Since the distribution of B_t^u is the same under $\mathcal{Q}$ as the distribution of B_t under $\mathbb{P}$, we have that the above is equal to

$$\gamma = \mathbb{E}_{\mathbb{P}}\left[g(Y_\tau^u) + \int_0^\tau f(Y_\tau^u) + |u_t|^2 dt\right]$$

Where Y_t^u is the process that satisfies

$$dY_t^u = (-\nabla V(Y_t^u) + \sigma(Y_t^u)u(Y_t^u))dt + \sigma(Y_t^u)dB_t$$

This shows that a control u induces a change of measure by Girsanov. All of this just tells us that we can add a drift, and still recover the original $\Psi(x)$ by adding an additional term, (the Girsanov change of measure term). Notice as well that Y_t is the same process as X_t since B_t^u absorbs the drift term from the control.

So, we want to find an optimal control u^* which minimizes the variance of our estimator, i.e we want

$$u^* = \arg\min_{u \in \mathcal{U}}\{\mathrm{Var}_{\mathcal{Q}}\left(\exp\left(\int_0^\tau u_t dB_t - \frac{1}{2}\int_0^\tau |u_t|^2 dt - g(Y_t^u) - \int_0^\tau f(Y_t^u)dt\right)\right)\}$$

In particular, we will see that we can find a control with zero variance.

Theorem 5.1 (Existence of a solution)**.** *For a process $X_t \in \mathbb{R}^m$ for $0 \leqslant t \leqslant T$ for some $T \in \mathbb{R}$, governed by the SDE*

$$dX_t = -\nabla V(X_t)dt + \sigma(X_t)dB_t$$

For $x \in \mathbb{R}^m$, $0 \leqslant t \leqslant \tau \leqslant T$, and

$$\Psi(x,t) = \mathbb{E}_{\mathbb{P}}\left[\exp\left(-\int_t^\tau f(X_s, s)ds - g(X_\tau)\right)\middle| X_t = x\right].$$

Then the measure Q induced by

$$u_s^* = \sigma(X_s^{u^*})^T \nabla_x \log \Psi(X_s^{u^*}, s)$$

is a zero variance estimator, i.e

$$\exp(-Z_\tau^{u^*} - W_\tau^{u^*}) = \Psi(x,0) \ - Q \text{ a.s.}$$

Where $X_s^{u^*}$ is the solution to the controlled SDE

$$dX_s^{u^*} = (-\nabla V(X_s^{u^*}) + \sigma(X_s^{u^*})u^*(X_s^{u^*}))dt + \sigma(X_s^{u^*})dB_t^{u^*}$$

And with $W_s^{u^*} = \exp(-g(X_\tau^{u^*}) - \int_0^\tau f(X_\tau^{u^*})dt)$ and $Z_\tau^{u^*}$ is the change of measure term from Girsanov on $X_s^{u^*}$. (For more details, see [Har+17])

Proof. By the Feynmac-Kac formula [Øks03], we have that Ψ solves the following elliptic boundary problem

$$\begin{aligned} (\mathcal{A} - f)\Psi = 0 \qquad & \text{for } (x,t) \in D \\ \Psi = \exp(-g) \quad & \text{for } (x,t) \in D^C \end{aligned}$$

Where $\mathcal{A}$ is the infintesimal generator of, i.e

$$\mathcal{A} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma\sigma^T : \nabla_x^2 + -\nabla V \cdot \nabla_x$$

In particular, we get this formula via Ito's formula. For our function $\zeta_s^{u^*} = -\log\Psi(X_s^{u^*}, s)$ and using the shorthand $\Psi = \Psi(X_s^{u^*}, s), \sigma = \sigma(X_s^{u^*}), V = \nabla V(X_s^{u^*}), u^* = u^*(X_s^{u^*})$ We get

$$\begin{aligned} d\zeta_s^{u^*} &= -\frac{\partial \log \Psi}{\partial t} \cdot ds - \nabla_x \log\Psi(-V + \sigma u^*) \cdot ds \\ &- \frac{1}{2}\sigma\sigma^T : \nabla_x^2(\log\Psi) \cdot ds - \sigma^T \nabla_x \log\Psi \cdot dB_s^{u^*} \\ &= -\left[\frac{\mathcal{A}\Psi}{\Psi} + \left((\sigma)^T \frac{\nabla_x \Psi}{\Psi}\right) \cdot u^* - \frac{1}{2}\frac{|\sigma^T \nabla_x \Psi|^2}{\Psi^2}\right] \cdot ds - \sigma^T \frac{\nabla_x \Psi}{\Psi} \cdot dB_s^{u^*} \\ &= -\left[\frac{\Psi f(X_s^{u^*}, s)}{\Psi} + \left((\sigma)^T \frac{\nabla_x \Psi}{\Psi}\right) \cdot u^* - \frac{1}{2}\frac{|\sigma^T \nabla_x \Psi|^2}{\Psi^2}\right] \cdot ds - \sigma^T \frac{\nabla_x \Psi}{\Psi} \cdot dB_s^{u^*} \end{aligned}$$

Where we used Feynman-Kac for the last equality. If we substitute $u_s^* = \sigma^T \nabla_x \log\Psi$, we get

$$\begin{aligned} d\zeta_s^{u^*} &= -\left[f(X_s^{u^*}, s) + |u^*|^2 - \frac{1}{2}|u^*|^2\right] \cdot ds - u^* \cdot dB_s^{u^*} \\ &= \left[-f(X_s^{u^*}, s) - \frac{|u^*|^2}{2}\right] \cdot ds - u^* \cdot dB_s^{u^*} \end{aligned}$$

And we can recognize the terms in u^* as the change of measure term we obtained before,

$$Z_{s,\tau}^{u^*} = \int_s^\tau u_t^* dB_t^{u^*} + \frac{1}{2}\int_s^\tau |u_t^*|^2 dt$$

So we get

$$d\zeta_s^{u^*} = -f(X_s^{u^*}, s) \cdot ds - dZ_{s,\tau}^{u^*}$$

By continuity of $\zeta_s^{u^*}$ as $s \downarrow 0$, we get

$$\begin{aligned}\zeta_\tau^{u^*} - \zeta_0^{u^*} &= \int_0^\tau d\zeta_s^{u^*} \\ &= -\int_0^\tau f(X_s^{u^*}, s)ds - Z_{0,\tau}^{u^*}\end{aligned}$$

By definition, we have $\zeta_0^{u^*} = -\log \Psi(x, 0)$, we also by Feynman-Kac that $\zeta_\tau^{u^*} = -\log \Psi(X_\tau^{u^*}, \tau) = g(X_\tau^{u^*})$, so we get

$$\Psi(x, 0) = \exp\left(-g(X_\tau^{u^*}) - \int_0^\tau f(X_s^{u^*})ds - Z_{0,\tau}^{u^*}\right) \text{ Q-a.s}$$

This shows that we get a zero variance estimator of $\Psi(x, 0)$, showing that the optimal u^* can be reached. □

5.2 Reinforcement Learning

Reinforcement learning is a sub-field of optimal control that deals with agents, that can take actions in an environment. The goal is to maximize a reward signal over sequences of decision epochs. We will first introduce the theoretical framework most RL problems are set in. (For a more robust introduction, see [SB+98]).

5.2.1 Markov Decision Process

The theoretical framework for solving most RL problems is a Markov Decision Process. Most examples consist of

1. S, the state space. The space of all possible states the environment can be in.
2. A, an action space. The set of actions an agent can choose at each state
3. R: $S \times A \to \mathbb{R}$, a function that maps state action pairs to a reward, quantified by a real number.
4. P: $S \times S \times A \to [0, 1]$ a transition probability function, which takes a state $s_1 \in S$, $s_2 \in S$ and $a \in A$, to the probability that taking action a in s_1 will lead to s_2. In deterministic environments, this is a simple function.
5. $T_r : S \to \{0, 1\}$. Where it maps states whether they are terminal or not, i.e whether there are any actions that can send the state to another state.
6. $\mathcal{T}$ a set of decision epochs on which the action can choose actions. In most cases, we would have $t \in \mathcal{T}$, $s_t \in S$ and $a_t \in A$, a sequence of states and actions.

5.2.2 Value, Return and Trajectories

We call τ a trajectory if it is an ordered sequence of states and actions. I.E $s_0, a_0, s_1, a_1, \ldots, s_T$ for some $T \in \mathcal{T} = \mathbb{N}$.

The return of a trajectory is for some discount rate $\gamma \in [0, 1]$,

$$G(\tau) = \sum_{i=0}^{T-1} \gamma^i R(s_i, a_i)$$

Therefore, since transitions are stochastic, we can define the value of a state for some policy π to be

$$V^\pi(s) = \mathbb{E}[G(\tau)|\pi, s_0 = s]$$

Where the trajectories are generated by the policy π.

Therefore, we would want optimize the quality of our policy, what we call the objective function

$$J(\pi) = \mathbb{E}[G(\tau)|\pi, s_0 \sim p_0]$$

Where p_0 is the distribution over initial states, and the expectation is over the path-measure space.

We search for an optimal policy π^* such that

$$\pi^* = \arg\max_{\pi \in \mathcal{P}} \{\mathbb{E}_{x_0 \sim p_{x_0}, \pi}[G(\tau)|s_0 = x_0]\}$$

5.2.3 RL Algorithms

Two of the main model-free and model-based. Model-free methods do not rely on an accurate model of the environment. Model-based methods use a given P for the environment and can use it for their policies. (For a more general overview of Reinforcement Learning algorithms, see [SB+98])

Two of the main approaches to improving a policy are value based methods and policy gradient methods.

Value based methods, such as Q-learning aim to estimate the value of $(s, a) \in S \times A$ state action pairs, and at each decision epoch $t \in \mathcal{T}$ choose the action $\max_{a \in A} V^\pi(s_t, a)$, where V^π is some parametrized method of estimating the expected return of taking action a in state s_t.

The Bellman optimality equation

$$Q^*(s, a) = \mathbb{E}[R(s, a) + \max_{a' \in A} Q^*(s_{t+1}, a')|s_t = s, a_t = a, \pi^*]$$

is an algorithm, that under some conditions provide a guarantee of converging to the optimal Q value. If this function is obtained, the optimal policy is then given by

$$\pi^*(s) := \arg\max_{a \in A} Q^*(s,a)$$

Policy based methods aim to parametrize the policy and optimize the expected discounted return with respect to the parameters. One can find a closed form for the gradient, called the policy gradient theorem, was first shown in [Sut+99] given by

$$\begin{aligned}
\nabla_\theta J(\pi_\theta) &= \nabla_\theta \mathbb{E}[G(\tau)|\pi_\theta] \\
&= \nabla_\theta \int_\Omega G(\tau) p(\tau|\pi_\theta) d\tau \\
&= \int_\Omega G(\tau) \nabla_\theta \left(p_0(s_0) \prod_{i=1}^{T-1} p(s_{i+1}|s_i, a_i) \pi_\theta(a_i|s_i) \right) d\tau \\
&= \int_\Omega G(\tau) p(\tau|\pi_\theta) \nabla_\theta \log p(\tau|\pi_\theta) d\tau \\
&= \mathbb{E}[G(\tau) \sum_{i=1}^{T-1} \nabla_\theta \log \pi_\theta(a_i|s_i)|\pi_\theta]
\end{aligned}$$

Where we use the log-trick, i.e taking that

$$\nabla_\theta \log \pi_\theta(s,a) = \frac{\nabla_\theta \pi_\theta(s,a)}{\pi_\theta(s,a)}$$

which implies that

$$\pi_\theta(s,a) \nabla_\theta \log \pi_\theta(s,a) = \nabla_\theta \pi_\theta(s,a)$$

Which allows us to take the log, take a sum and remove all of the $p_i(s_i)$ terms.

5.3 Solving the Sampling Problem with Reinforcement Learning

The following example of using RL to solve the sampling problem, and forming a link between RL and Stochastic Calculus comes from [QB24]. In our case, we want to minimize the variance of our estimator, so we set our reward is $-1/2 \int_0^T |a_t|^2 dt - g(s_T) - \int_0^T f(s_t) dt$ for each trajectory. Therefore, by using RL methods, we can iteratively find an optimal control that gives us a zero variance estimator of the desired quantity.

6 Exercises

Exercise 6.1 (Doob's optional stopping theorem)**.** Let $\{X_n\}$ be a $\{\mathcal{F}_n\}-$sub-martingale. And let $\nu \leqslant \tau, \quad \mathbb{P} - a.s$ be bounded $\{\mathcal{F}_n\}-$stopping times. Then

$$\mathbb{E}[X_\nu] \leqslant \mathbb{E}[X_\tau]$$

Proof. Notice that $\nu \leqslant \tau$-$\mathbb{P}$ almost surely implies $\forall n \geqslant 1, \nu \wedge n \leqslant \tau \wedge n - \mathbb{P}$ almost surely. By induction we have shown that for $m > n$ $\mathbb{E}[X_m] \geqslant \mathbb{E}[X_n]$, so we have

$$\begin{aligned}
\mathbb{E}[X_{\tau\wedge n}] &= \mathbb{E}[X_{\tau\wedge n}\mathbb{1}_{\tau\wedge n\geqslant\nu\wedge n}] + \mathbb{E}[X_{\tau\wedge n}\mathbb{1}_{\tau\wedge n<\nu\wedge n}] \\
&\geqslant \mathbb{E}[X_{\nu\wedge n}\mathbb{1}_{\tau\wedge n\geqslant\nu\wedge n}] + \mathbb{E}[X_{\nu\wedge n}\mathbb{1}_{\tau\wedge n<\nu\wedge n}] \\
&= \mathbb{E}[X_{\nu\wedge n}]
\end{aligned}$$

Since we have τ, ν bounded stopping times, we have $C, D \in \mathbb{N} : \tau \leqslant C$ and $\nu \leqslant D$. So we get $\mathbb{E}[X_\tau] \leqslant \mathbb{E}[X_C] < \infty$, and $\mathbb{E}[X_\nu] \leqslant \mathbb{E}[X_D] < \infty$. So by dominated convergence theorem:

$$\mathbb{E}[X_{\tau\wedge n}] \geqslant \mathbb{E}[X_{\nu\wedge n}]$$

$$\lim_{n\to\infty} \mathbb{E}[X_{\tau\wedge n}] = \mathbb{E}[\lim_{n\to\infty} X_{\tau\wedge n}] = \mathbb{E}[X_\tau]$$

$$\lim_{n\to\infty} \mathbb{E}[X_{\nu\wedge n}] = \mathbb{E}[\lim_{n\to\infty} X_{\nu\wedge n}] = \mathbb{E}[X_\nu]$$

And finally

$$\mathbb{E}[X_\tau] \geqslant \mathbb{E}[X_\nu]$$

□

Exercise 6.2. Let $(B_t)_{t\geqslant 0}$ be a SBM, and let $s < t$.

1. Find $\mathbb{E}[B_s B_t^2]$
2. Find $\mathbb{E}[B_s^2 B_t^2]$
3. Show $\mathbb{E}[B_s e^{B_s}] = s e^{s/2}$
4. Find $\mathbb{E}[B_s e^{B_t}]$

Proof. First, notice B_s and B_t are not independent. In particular $B_t = B_s + Z, Z \sim \mathcal{N}(0, t-s)$. Therefore

$$\begin{aligned}
\mathbb{E}[B_s B_t^2] &= \mathbb{E}[\mathbb{E}[B_s B_t^2 | B_s = b_s]] \\
&= \mathbb{E}[\mathbb{E}[b_s(b_s + Z)^2]] \\
&= \mathbb{E}[\mathbb{E}[b_s^3] + \mathbb{E}[2b_s^2 Z] + \mathbb{E}[b_s Z^2]] \\
&= \mathbb{E}[B_s^3 + 0 + B_s(t-s)] \\
&= \mathbb{E}[B_s^3] \\
&= 0
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{E}[B_s^2 B_t^2] &= \mathbb{E}[\mathbb{E}[B_s^2 B_t^2 | B_s = b_s]] \\
&= \mathbb{E}[\mathbb{E}[b_s^2(b_s + Z)^2]] \\
&= \mathbb{E}[\mathbb{E}[b_s^4] + \mathbb{E}[2b_s^3 Z] + \mathbb{E}[b_s^2 Z^2]] \\
&= \mathbb{E}[B_s^4 + 0 + B_s^2(t-s)] \\
&= \mathbb{E}[B_s^4] + (t-s)(s) \\
&= 3s^2
\end{aligned}$$

We get the $4th$ moment by doing repeated integration by parts. Now,

$$\begin{aligned}
\mathbb{E}[B_s e^{B_s}] &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi s}} e^x e^{-\frac{x^2}{2s}} dx \\
&= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi s}} e^{-\frac{x^2 - 2xs}{2s}} dx \\
&= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi s}} e^{-\frac{(x-s)^2}{2s}} e^{\frac{s}{2}} dx \\
&= e^{\frac{s}{2}} \mathbb{E}[B_s + s] \\
&= s e^{\frac{s}{2}}
\end{aligned}$$

By identifying the expectation. Now for the final exercise:

$$\begin{aligned}
\mathbb{E}[B_S e^{B_t}] &= \mathbb{E}[\mathbb{E}[b_s e^{b_s + Z} | B_s = b_s]] \\
&= \mathbb{E}[B_s e^{B_s} \mathbb{E}[e^Z]] \\
&= \mathbb{E}[B_s e^{B_s} e^{\frac{t-s}{2}}] \\
&= e^{\frac{t-s}{2}} s e^{\frac{s}{2}} \\
&= s e^{\frac{t}{2}}
\end{aligned}$$

□

Exercise 6.3. Let $(B_t)_{t\geqslant 0}$ be a SBM, and let $a \in \mathbb{R}$. Find

1. $\lim_{t\to\infty} \mathbb{E}[\mathbb{1}_{B_t\leqslant a}]$
2. $\lim_{t\to\infty} \mathbb{E}[B_t \mathbb{1}_{B_t\leqslant a}]$

Proof. First notice $\mathbb{E}[\mathbb{1}_{B_t\leqslant a}] = \mathbb{P}(B_t \leqslant a)$. And since $B_t \sim \mathcal{N}(0,t)$ has a symmetric distribution,

$$x := \mathbb{P}(B_t \leqslant a) = 1 - F_{B_t}(-a) = 1 - \mathbb{P}(B_t \leqslant -a) = \mathbb{P}(B_t \geqslant -a)$$

We can also simply notice $\forall t \geqslant 0$

$$1 = \mathbb{P}(B_t \leqslant a) + \mathbb{P}(B_t \geqslant -a) - \mathbb{P}(|B_t| \leqslant a) = 2x - \mathbb{P}(|B_t| \leqslant a)$$

Now switching back, $\mathbb{P}(|B_t| \leqslant a) = \mathbb{E}[\mathbb{1}_{|B_t|\leqslant a}] \leqslant 1$. So by the dominated convergence theorem,

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}[\mathbb{1}_{|B_t|\leqslant a}] &= \lim_{t\to\infty} \int_{-a}^{a} \frac{1}{\sqrt{t2\pi}} e^{-\frac{x^2}{2t}} dx \\
&= \int_{-a}^{a} \lim_{t\to\infty} \frac{1}{\sqrt{t2\pi}} e^{-\frac{x^2}{2t}} dx \\
&= \int_{-a}^{a} 0 dx \\
&= 0
\end{aligned}$$

Therefore we have $\lim_{t\to\infty} \mathbb{P}(B_t \leqslant a) = \frac{1}{2}$.
Now notice by calculation, $\forall a \in \mathbb{R}$:

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}[B_t \mathbb{1}_{B_t\leqslant a}] &= \lim_{t\to\infty} \int_{-\infty}^{a} \frac{x}{\sqrt{2t\pi}} e^{-\frac{x^2}{2t}} dx \\
&= \lim_{t\to\infty} \sqrt{\frac{t}{2\pi}} \int_{\infty}^{\frac{a^2}{2t}} e^{-u} du \\
&= \lim_{t\to\infty} -\sqrt{\frac{t}{2\pi}} e^{-\frac{a^2}{2t}} \\
&= -\infty
\end{aligned}$$

□

References

[Ben24] Lucas Benigni. *Calcul Stochastique*. 2024.

[Cou02] Jean-Michel Courtault. "LOUIS BACHELIER ON THE CENTENARY OF THÉORIE DE LA SPÉCULATION[1]". In: *Louis Bachelier: aux origines de la finance mathématique* 838 (2002), p. 5.

[Har+17] Carsten Hartmann et al. "Variational Characterization of Free Energy: Theory and Algorithms". In: *Entropy* 19.11 (2017). ISSN: 1099-4300. DOI: 10.3390/e19110626. URL: https://www.mdpi.com/1099-4300/19/11/626.

[Øks03] Bernt Øksendal. "Stochastic differential equations". In: *Stochastic differential equations: an introduction with applications*. Springer, 2003, pp. 38–50.

[QB24] Jannes Quer and Enric Ribera Borrell. *Connecting Stochastic Optimal Control and Reinforcement Learning*. 2024. arXiv: 2211.02474 [math.OC]. URL: https://arxiv.org/abs/2211.02474.

[Que+18] J. Quer et al. "An Automatic Adaptive Importance Sampling Algorithm for Molecular Dynamics in Reaction Coordinates". In: *SIAM Journal on Scientific Computing* 40.2 (2018), A653–A670. DOI: 10.1137/17M1124772. eprint: https://doi.org/10.1137/17M1124772. URL: https://doi.org/10.1137/17M1124772.

[SB+98] Richard S Sutton, Andrew G Barto, et al. *Reinforcement learning: An introduction*. Vol. 1. 1. MIT press Cambridge, 1998.

[Sut+99] Richard S Sutton et al. "Policy gradient methods for reinforcement learning with function approximation". In: *Advances in neural information processing systems* 12 (1999).