0.1 Stochastic Differential Calculus

(Lesson 8 of 11/11/19) Compiled: December 31, 2019

0.1.1 Ito's rules of integration

We now consider a more general stochastic integral, and show that, using Ito's prescription:

$$\int_0^t H(B(\tau),\tau)(\mathrm{d}B(\tau))^k \underset{m.s.}{\overset{\text{I.p.}}{=}} \sum_{i=1}^n H(B_{i-1},\tau_{i-1})(\Delta B_i)^k =$$
$$= \begin{cases} \int_0^t H(B,\tau)\,\mathrm{d}B(\tau) & k=1 \\ \int_0^t H(B(\tau),\tau)\,\mathrm{d}\tau & k=2 \\ 0 & k>2 \end{cases}$$

This leads to the following "rules" for *Ito integrals*:

$$(\mathrm{d}B)^n = \begin{cases} \mathrm{d}B & n=1 \\ \mathrm{d}t & n=2 \\ 0 & n>2 \end{cases} \tag{1}$$

We already showed an example for $k=1$, and we now proceed with the other two cases.

Example 1 (Integral in $\mathbf{d}\boldsymbol{B}^{\mathbf{2}}$)**:**

Consider a *non-anticipating* function $G(\tau)$, and the following stochastic integral:

$$I = \int_0^t G(\tau)(\mathrm{d}B\,(\tau))^2$$

With *non-anticipating* we mean that $G(\tau)$ does not depend on $B(s)-B(\tau)\forall s > \tau$, i.e. it does not depend on the *future*. Discretizing:

$$I = \underset{n\to\infty}{\overset{\text{m.s.}}{\lim}} I_n = \underset{n\to\infty}{\overset{\text{m.s.}}{\lim}} \sum_{i=1}^n G(t_{i-1})\Delta B_i^2$$

For simplicity, denote:

$$G_i \equiv G_i \qquad \Delta B_i \equiv B_i - B_{i-1} \qquad \Delta t_i = t_i - t_{i-1}$$

We want to prove that:

$$\int_0^t G(\tau)(\mathrm{d}B(\tau))^2 \overset{?}{=} \int_0^t G(\tau)\,\mathrm{d}\tau = \lim_{n\to\infty}\sum_{i=1}^n G_{i-1}\Delta t_i$$

Applying the definition of a *mean square* limit, this is equivalent to:

$$\left\langle \left(\sum_{i=1}^n G_{i-1}\Delta B_i^2 - \sum_{i=1}^n G_{i-1}\Delta t_i \right)^2 \right\rangle \xrightarrow[n\to\infty]{?} 0$$

Expanding the square as a product of two sums over i and j, and then highlighting the case with $i = j$:

$$\langle\left[\sum_{i=1}^{n} G_{i-1}[(\Delta B_i)^2 - \Delta t_i]\right]^2\rangle = \sum_{i,j=1}^{n} \langle G_{i-1}[(\Delta B_i)^2 - \Delta t_i] G_{j-1}[(\Delta B_j)^2 - \Delta t_j]\rangle =$$

$$= \sum_{i=1}^{n} \langle G_{i-1}^2 [(\Delta B_i)^2 - \Delta t_i]^2\rangle + 2\sum_{i<j}^{n} \langle G_{i-1}[(\Delta B_i)^2 - \Delta t_i] G_{j-1} \, [(\Delta B_j)^2 - \Delta t_j]\rangle \tag{2}$$

As $i < j$, note that the yellow term *does not depend* on $\Delta B_j = B_j - B_{j-1} = B(t_j) - B(t_{j-1})$. In fact, as G is *non-anticipating*, G_{j-1} depends only on the previous steps. Thus, the yellow and blue terms are *independent* of each other, and so we can factorize the average:

$$(2) = \sum_{i=1}^{n} \langle G_{i-1}^2 [(\Delta B_i)^2 - \Delta t_i]^2\rangle + 2\sum_{i<j}^{n} \langle G_{i-1}[(\Delta B_i)^2 - \Delta t_i] G_{j-1}\rangle \langle (\Delta B_j)^2 - \Delta t_j\rangle$$

Recall that:

$$\langle (\Delta B_j)^2 - \Delta t_j\rangle = \langle (\Delta B_j)^2\rangle - \Delta t_j = 0$$

and so only the first term of (2) remains. Again, noting that G_{i-1} does not depend on ΔB_i, as it is *non-anticipating*, can factorize the average:

$$(2) = \langle \sum_{i=1}^{n} G_{i-1}^2 [(\Delta B_i)^2 - \Delta t_i]^2\rangle = \sum_{i=1}^{n} \underbrace{\langle G_{i-1}^2\rangle}_{G_{i-1}^2} \langle [(\Delta B_i)^2 - \Delta t_i]^2\rangle \tag{3}$$

Expanding the stochastic term:

$$\begin{aligned}\langle [(\Delta B_i)^2 - \Delta t_i]^2\rangle &= \langle (\Delta B_i)^4 - 2\Delta t_i (\Delta B_i)^2\rangle + \Delta t_i^2 = \\ &= \underbrace{\langle (\Delta B_i)^4\rangle}_{3(\Delta t_i)^2} - 2\Delta t_i \underbrace{\langle (\Delta B_i)^2\rangle}_{\Delta t_i} + \Delta t_i^2 = 2\Delta t_i^2\end{aligned}$$

And substituting back into the sum and taking the limit completes the proof:

$$(3) = 2\sum_{i=1}^{n} G_{i-1}^2 \Delta t_i^2 \leq 2\left(\max_{i \leq j \leq n} \Delta t_j\right) \sum_{i=1}^{n} G_{i-1}^2 \Delta t_i \xrightarrow[n\to\infty]{} 2 \cdot 0 \cdot \int_0^t G^2(\tau) \, \mathrm{d}\tau = 0$$

This proves that $(\mathrm{d}B)^2 = \mathrm{d}t$.

Example 2 (The case with $\boldsymbol{n > 2}$):

We want now to show that:

$$\int_0^t G(\tau)(\mathrm{d}B(\tau))^n = \operatorname*{lim}^{\text{m.s.}}_{n\to\infty} \sum_{i=1}^{n} G_{i-1}(\Delta B_i)^n = 0$$

By definition, we want to show that:

$$\langle\left(\sum_{i=1}^{n} G_{i-1}(\Delta B_i)^n\right)^2\rangle \xrightarrow[n\to\infty]{} 0$$

Expanding the square, and factorizing the averages (as G is *non-anticipating*) leads to:

$$\begin{aligned}\langle\left(\sum_{i=1}^{n} G_{i-1}(\Delta B_i)^n\right)^2\rangle &= \sum_{i=1}^{n}\langle G_{i-1}^2(\Delta B_i)^{2n}\rangle + 2\sum_{i<j}^{n}\langle G_{i-1}G_{j-1}(\Delta B_i)^n(\Delta B_j)^n\rangle = \\ &= \sum_{i=1}^{n} G_{i-1}^2 \langle(\Delta B_i)^{2n}\rangle + 2\sum_{i<j}^{n}\langle G_{i-1}G_{j-1}(\Delta B_i)^n\rangle \langle(\Delta B_j)^n\rangle\end{aligned} \tag{4}$$

Now, recall that the p-th central moment of $X \sim \mathcal{N}(\mu,\sigma)$ can be computed with Isserlis theorem, resulting in:

$$\mathbb{E}[(X-\mu)^p] = \begin{cases} 0 & p \text{ is odd} \\ \sigma^p(p-1)!! & p \text{ is even}\end{cases}$$

where $p!! = p\cdot(p-2)\cdot\dots\cdot 1$ is a *double factorial*, that can be rewritten in terms of factorials as follows:

$$p!! = \begin{cases} 2^k k! & p = 2k \text{ even} \\ \dfrac{(2k)!}{2^k k!} & p = 2k-1 \text{ odd}\end{cases} \tag{5}$$

So, if ***n* is odd**, the blue term in (4) vanishes. Let's suppose, for simplicity, that G is bounded, i.e. $|G(\tau)| < K \ \forall \tau \in \mathbb{R}$. Then:

$$\begin{aligned}(4) &= \sum_{i=1}^{n} G_{i-1}^2(\Delta t_i)^n (2n-1)!! = \sum_{i=1}^{n} G_{i-1}^2(\Delta t_i)^n \frac{(2n)!}{2^n n!} \leq \frac{K^2(2n)!}{2^n n!}\sum_{i=1}^{n}(\Delta t_i)^n \\ &\leq \frac{K^2(2n)!}{2^n n!}\left(\max_{i\leq j\leq n}(\Delta t)^{n-1}\right)\underbrace{\sum_{i=1}^{n}\Delta t_i}_{t} \xrightarrow[n\to\infty]{} 0\end{aligned}$$

On the other hand, if ***n* is even**, the blue term in (4) is not null. However, the same argument for n odd can be applied to the first term, which vanishes in the limit. So we only need to study the blue term:

$$(4) = 2\sum_{i<j}^{n}\langle\underbrace{G_{i-1}G_{j-1}}_{\leq K^2}(\Delta B_i)^n\rangle\langle(\Delta B_j)^n\rangle \tag{6}$$

Here, as n is even:

$$\langle(\Delta B_i)^n\rangle = (\Delta t_i)^{n/2}(n-1)!! = (\Delta t_i)^{n/2}\left(2\frac{n}{2}-1\right)!! \underset{(5)}{=} (\Delta t_i)^{n/2}\frac{n!}{2^{n/2}(n/2)!}$$

And so:

$$\begin{aligned}(6) &\leq 2K^2\left(\frac{n!}{2^{n/2}(n/2)!}\right)^2\sum_{i<j}^{n}\Delta t_i^{n/2}\Delta t_j^{n/2}\\ &\leq 2K^2\left(\frac{n!}{2^{n/2}(n/2)!}\right)^2\left(\max_{i\leq l\leq n}\Delta t_l\right)^{2(n/2-1)}\underbrace{\sum_{i<j}^{n}\Delta t_i\Delta t_j}_{\leq t^2}\xrightarrow[n\to\infty]{}0 \qquad \square\end{aligned}$$

Example 3 (Other cases)**:**

Ito's rules allow us to consider even more general integrals. For example:

$$\int_0^t G(\tau)\,\mathrm{d}B(\tau)\,\mathrm{d}\tau = 0$$

In fact, as $(\mathrm{d}B)^2 = \mathrm{d}\tau$, $\mathrm{d}B\,\mathrm{d}\tau = 0$ because $(\mathrm{d}B)^n = 0\ \forall n > 2$.

Example 4 (Integration of polynomials)**:**

By using Ito's rules we can find a formula for integrating *powers* of the Brownian motion:

$$\int_0^t (B(\tau))^n\,\mathrm{d}B(\tau)$$

We first differentiate a polynomial, and then recover the rule for integration by performing the inverse operation.
Recall that, in general, a differential is *the increment* of a function after a small *nudge* of its argument:

$$\mathrm{d}f(t) = f(t+\mathrm{d}t) - f(t)$$

The same holds in the stochastic case. In particular:

$$\begin{aligned}\mathrm{d}(B(t))^n &= [B(t+\mathrm{d}t)]^n - (B(t))^n = [B(t)+\mathrm{d}B(t)]^n - (B(t))^n =\\ &\underset{(a)}{=}\sum_{k=0}^{n}\binom{n}{k}(\mathrm{d}B(t))^k(B(t))^{n-k} - (B(t))^n =\\ &= \cancel{(B(t))^n} + \sum_{k=1}^{n}\binom{n}{k}(\mathrm{d}B(t))^k(B(t))^{n-k} - \cancel{(B(t))^n} =\\ &\underset{(b)}{=}\underbrace{n(\mathrm{d}B(t))(B(t))^{n-1}}_{k=1} + \underbrace{\frac{n(n-1)}{2}\overbrace{(\mathrm{d}B(t))^2}^{\mathrm{d}t}(B(t))^{n-2}}_{k=2} + \underbrace{0}_{k>2}\end{aligned}$$

where in (a) we used Newton's binomial formula, and in (b) the previously found Ito's rules for integration (1). Letting $m = n-1$ and isolating $\mathrm{d}B(t)$ leads to:

$$(m+1)(B(t))^m\,\mathrm{d}B(t) = (\mathrm{d}B(t))^{m+1} - \frac{m(m+1)}{2}(B(t))^{m-1}\,\mathrm{d}t$$

Finally, dividing by $m+1$ and integrating leads to the desired formula:

$$\begin{aligned}\int_0^\tau (B(t))^m \,\mathrm{d}B(t) &= \frac{1}{m+1}\int_0^\tau \mathrm{d}(B(t))^{m+1} - \frac{m}{2}\int_0^\tau (B(t))^{m-1}\,\mathrm{d}t = \\ &= \frac{1}{m+1}(B(t))^{m+1}\Big|_0^\tau - \frac{m}{2}\int_0^\tau (B(t))^{m-1}\,\mathrm{d}t = \\ &= \frac{(B(\tau))^{m+1} - (B(0))^{m+1}}{m+1} - \frac{m}{2}\int_0^\tau (B(t))^{m-1}\,\mathrm{d}t\end{aligned}$$

And in the case $m=1$ we retrieve the previously obtained result:

$$\int_0^\tau B(t)\,\mathrm{d}B(t) = \frac{B^2(\tau) - B^2(0)}{2} - \frac{\tau}{2}$$

Example 5 (General differentiation rule)**:**

Because $(\mathrm{d}B)^2 = \mathrm{d}t$, when computing differentials from a Taylor expansion up to $O(\mathrm{d}t^2)$ one must compute even the terms of order $\mathrm{d}B^2$. For example, consider a generic function $f(B(t),t)$:

$$\begin{aligned}\mathrm{d}f(B(t),t) = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial B}\,\mathrm{d}B(t) + \underbrace{\frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\mathrm{d}t)^2}_{O([\mathrm{d}t]^2)} + \frac{1}{2}\frac{\partial^2 f}{\partial B^2}\underbrace{[\mathrm{d}B(t)]^2}_{\mathrm{d}t} + \\ + \frac{\partial^2 f}{\partial B(t)\partial t}\underbrace{\mathrm{d}t\,\mathrm{d}B(t)}_{0} + O([\mathrm{d}t]^2) = \\ = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial B}\,\mathrm{d}B(t) + \frac{1}{2}\frac{\partial^2 f}{\partial B^2}\,\mathrm{d}t + O([\mathrm{d}t]^2)\end{aligned}$$

0.2 Derivation of the Fokker-Planck equation

Starting from the Master Equation and taking the continuum limit we arrived at the Fokker-Planck equation:

$$\dot{W}(x,t) = -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}W(x,t)D(x,t)\right] \tag{7}$$

At the same time, if we consider the dynamics of a *single path*, adding a *stochastic term* to the second law of motion, we arrive at the Langevin equation (in the overdamped limit):

$$\mathrm{d}x(t) = f(x(t),t)\,\mathrm{d}t + \sqrt{2D(x(t),t)}\,\mathrm{d}B(t) \tag{8}$$

We want now to show that these two formulations are equivalent, by deriving (7) from (8). The main idea is to introduce a *test function* $h(x(t))$, and compute its

expected value at the instant t over *all possible points* that can be reached by the trajectory $x(t)$, thus obtaining a value that will depend on the *global* probability distribution $W(x,t)$. Then, we can use Langevin equation to describe the dynamics of each *single path*. In this way, we will obtain a relation between a quantity involving $W(x,t)$ and the parameters $f(x,t)$ and $D(x,t)$ appearing in (8), which will hopefully be (7).
So, let's start by computing the average of $h(x(t))$ at a fixed time:

$$\langle h(x(t))\rangle = \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)h(x)$$

As we seek to construct a *time derivative*, we start by differentiating:

$$\mathrm{d}\langle h(x(t))\rangle = \left(\frac{\partial}{\partial t}\int_{\mathbb{R}} \mathrm{d}x\, W(x,t)h(x)\right)\mathrm{d}t = \mathrm{d}t\int_{\mathbb{R}} \mathrm{d}x\, \dot{W}(x,t)h(x) \tag{9}$$

And then dividing by $\mathrm{d}t$ leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle h(x(t))\rangle = \int_{\mathbb{R}} \mathrm{d}x\, \dot{W}(x,t)h(x) \tag{10}$$

However, we could also start by differentiating $h(x(t))$:

$$\mathrm{d}h(x(t)) = h(x(t)+\mathrm{d}x\,(t)) - h(x(t)) = \tag{11}$$

$$\underset{(a)}{=} h'(x(t))\,\mathrm{d}x\,(t) + \frac{1}{2}h''(x(t))[\mathrm{d}x\,(t)]^2 + O([\mathrm{d}x\,(t)]^2) \tag{12}$$

where in (a) we used a Taylor expansion for the first term. From (8), and applying Ito's rules, we can obtain explicit expressions for the $[\mathrm{d}x(t)]^n$:

$$\begin{aligned}
&[\mathrm{d}x(t)]^2 = f^2[\mathrm{d}t]^2 + 2D\overbrace{[\mathrm{d}B(t)]^2}^{\mathrm{d}t} + f\sqrt{2D}\,\overbrace{\mathrm{d}B(t)\,\mathrm{d}t}^{0}\\
&[\mathrm{d}x(t)]^3 = O([\mathrm{d}t]^2)
\end{aligned}$$

And substituting in (12) leads to:

$$\begin{aligned}
\mathrm{d}h(x(t)) &= h'[f\,\mathrm{d}t + \sqrt{2D}\,\mathrm{d}B] + \frac{1}{2}h''2D\,\mathrm{d}t + O([\mathrm{d}t]^2) =\\
&= \mathrm{d}t\,[h'f + h''D] + h'\sqrt{2D}\,\mathrm{d}B
\end{aligned}$$

Taking the expected value:

$$\begin{aligned}
\mathrm{d}\langle h(x(t))\rangle &= \langle \mathrm{d}t\,[h'f + h''D]\rangle + \langle h'\sqrt{2D}\,\mathrm{d}B\rangle =\\
&\underset{(a)}{=} \langle \mathrm{d}t\,[h'f + h''D]\rangle + \langle\sqrt{2D}h'\rangle\underbrace{\langle \mathrm{d}B\rangle}_{0} =\\
&= \langle \mathrm{d}t\,[h'f + h''D]\rangle
\end{aligned}$$

where in (a) we used the fact that $D(x(t),t)$ is *non-anticipating*, allowing to factor the average.

Dividing by $\mathrm{d}t$ and expanding the average leads to:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle h(x(t))\rangle &= \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)[h'(x)f(x,t) + h''(x)D(x,t)] = \\
&= \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)f(x,t)h'(x) + \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)D(x,t)h''(x) = \\
&\underset{(a)}{=} \cancel{Whf\Big|_{-\infty}^{+\infty}} - \int_{\mathbb{R}} \mathrm{d}x\, h\frac{\partial}{\partial x}(Wf) + \\
&+ \cancel{WDh'\Big|_{-\infty}^{+\infty}} - \cancel{h\frac{\partial}{\partial x}(DW)\Big|_{-\infty}^{+\infty}} + \int_{\mathbb{R}} \mathrm{d}x\, h\frac{\partial^2}{\partial x^2}(WD) = \\
&= \int_{\mathbb{R}} \mathrm{d}x\, h(x)\left[\frac{\partial^2}{\partial x^2}(W(x,t)D(x,t)) - \frac{\partial}{\partial x}(W(x,t)f(x,t))\right] \qquad (13)
\end{aligned}$$

where in (a) we integrated by parts the first integral once, and the second one twice.
Finally, equating (10) and (**??**) leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle h(x(t))\rangle = \int_{\mathbb{R}} \mathrm{d}x\, \frac{\partial}{\partial t}W(x,t)h(x) = \int_{\mathbb{R}} \mathrm{d}x\, h(x)\left[\frac{\partial^2}{\partial x^2}(W(x,t)D(x,t)) - \frac{\partial}{\partial x}(W(x,t)f(x,t))\right]$$

As this relation holds for *any* test function $h(x)$, it means that the *integrands* are equal. So, by collecting a derivative, we retrieve the the Fokker-Planck equation (7):

$$\frac{\partial}{\partial t}W(x,t) = -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}(W(x,t)D(x,t))\right]$$

0.3 The role of temperature

From physical observations, we expect the amplitude of stochastic oscillations in Brownian motion to be dependent on temperature - as it is a direct effect of collisions with molecules in thermal equilibrium. So, we want to derive an explicit relation between the diffusion parameter D and T.
We start by assuming that, for $t \to \infty$, the particle will be *at equilibrium*, meaning that its distribution will be given by the Maxwell-Boltzmann:

$$W(x,t) \xrightarrow[t\to\infty]{} P_{\mathrm{eq}}(x) = \frac{e^{-\beta V(x)}}{Z} \qquad Z = \int_{\mathbb{R}} \mathrm{d}x\, e^{-\beta V(x)}; \qquad \beta = \frac{1}{k_B T}$$

Recall the Fokker-Planck equation:

$$\frac{\partial}{\partial t}W(x,t) = -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}(D(x,t)W(x,t))\right]$$

From the Langevin equivalence, and some physical reasoning, we found that:

$$f(x,t) = \frac{F_{\mathrm{ext}}}{\gamma} = -\frac{1}{\gamma}\frac{\partial V(x)}{\partial x} \qquad \gamma = 6\pi\eta a$$

Where F_{ext} is an external conservative force with potential $V(x)$ acting on the Brownian particle, assumed to be a sphere of radius a moving through a medium of viscosity η. Assuming $D(x,t) \equiv D$ for simplicity, the Fokker-Planck equation becomes:

$$\frac{\partial W^*}{\partial t} = \frac{\partial}{\partial x}\left[\frac{W^*}{\gamma}\frac{\partial V}{\partial x} + D\frac{\partial W^*}{\partial x}\right]$$

Here we are interested in the *particular solution* $W^*(x)$ that will be reached at the equilibrium, as it *does not* depend on time. So:

$$\frac{\partial W^*}{\partial t} \stackrel{!}{=} 0$$

Meaning that:

$$\left[\frac{W^*(x)}{\gamma}\frac{\partial V}{\partial x}(x) + D\frac{\partial W^*}{\partial x}(x)\right] = \text{constant} \qquad \forall x \tag{14}$$

As this relation holds for *any* x, we can examine it in the limit $x \to \infty$ to find the value of the constant. In fact, as $W^*(x)$ is a normalized pdf, we expect:

$$W^*, \frac{\partial W^*}{\partial x} \xrightarrow[x\to\infty]{} 0$$

And so the constant in (14) must be 0, leading to:

$$\frac{\partial W^*}{\partial x} = -\frac{1}{\gamma D}W^*\frac{\partial V}{\partial x} \Rightarrow \frac{1}{W^*}\frac{\partial W^*}{\partial x} = \ln\left(W^*\right) = -\frac{1}{\gamma D}\frac{\partial V}{\partial x}$$

Integrating, we find:

$$\ln W^*(x) = -\frac{1}{\gamma D}V(x) + c \Rightarrow W^*(x) = K\exp\left(-\frac{1}{\gamma D}V(x)\right) \stackrel{!}{=} \frac{1}{Z}\exp\left(-\beta V(x)\right)$$

And by comparing the two functions we obtain the desired relation:

$$\beta = \frac{1}{\gamma D} = \frac{1}{k_B T} \Rightarrow D = \frac{k_B T}{\gamma} = \frac{k_B T}{6\pi\eta a}$$

This is indeed the same relation that Einstein found when examining Brownian motion (*fluctuation-dissipation relationship*, 1905). As $D(x,t) \propto T$, the amplitude of stochastic oscillations (from Langevin equation) is proportional $\sqrt{2D} \propto \sqrt{T}$.