Introducing Wiener and Itô Stochastic Processes

Definition 1.1. *A stochastic process is a collection of random variables $X(t)$ indexed by the time variable $t \in [t_0, t_n = T]$:*

$$\{X(t)\}_{t_0 \leq t \leq T} = \{X(t_1), X(t_2), X(t_3), \ldots, X(T)\}.$$

In the study of stochastic processes, we not only track how random variables evolve over time, but also how information unfolds. This is captured by a concept called a filtration, denoted by $\mathcal{F}(t)_{t\geq 0}$. Each $\mathcal{F}(t)$, known as sigma-algebra of $X(t)$, represents all the information available up to time t for the process. Think of it as the "memory" of the system at that moment: everything that has happened and can be observed by time t is contained in $\mathcal{F}(t)$. A random variable $X(t)$ is said to be measurable with respect to $\mathcal{F}(t)$ if its value is fully known given the information at that time [5, 1]. If $X(t)$ is measurable with respect to $\mathcal{F}(s)$ for $s > t$, that simply means the value of $X(t)$, which belongs to the past, is already known by time s. In contrast, $X(s)$ is not measurable with respect to $\mathcal{F}(t)$ if $s > t$, because it belongs to the future and is still uncertain from the point of view of time t [6, 5].

This evolving information structure allows us to define one of the most important tools in stochastic analysis: conditional expectation given a sigma-algebra. The conditional expectation $\mathbb{E}[X(t) \mid \mathcal{F}(t)]$ represents our best estimate of a random variable $X(t)$, based only on the information known up to time t [1]. It is itself a random variable that depends solely on the past and present — not on future events.

A key concept when dealing with stochastic processes $X(t)$ is the *martingale property*:

Definition 1.2. *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where Ω is the set of all possible outcomes, $\mathcal{F}(t)$ is the sigma-algebra and $\mathbb{Q}$ the probability measure. A stochastic process $X(t)$ for $t \in [t_0, T]$ is said to be a martingale with respect to the sigma-algebra $\mathcal{F}(t)$ under measure $\mathbb{Q}$, if for all $t < T$:*

- **(1)**: *The process $X(t)$ is adapted, i.e. $X(t)$ is $\mathcal{F}(t)$-measurable.*
- **(2)**: *The absolute expected value of the stochastic process is finite:*

$$\mathbb{E}[|X(t)|] < \infty \tag{1}$$

- **(3)**: *The conditional expectation under probability measure $\mathbb{Q}$ with filtration on time t' of $X(t)$ equals the process in time t':*

$$\mathbb{E}[X(t)|\mathcal{F}(s)] = X(s). \tag{2}$$

Property **(3)** from definition (1.2) can be stated as "the best prediction of the expectation of a martingale's future is its present value" [6]. The martingale process express a "fair game" condition to the stochastic process, meaning the expected future value equals the current value, conditional on current information [6].

The simplest of the stochastic process is the Wiener process (or Brownian Motion), a process widely used as a building block for SDEs, defined as follows:

Definition 1.3. *A Wiener process $W(t)$ can be defined by the following properties:*

- **(1)***: $W(t_0) = 0$ for t_0 being the process initial time.*
- **(2)***: $W(t)$ is a almost surely continuous process, but nowhere differentiable* [1].
- **(3)***: For any $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$ with $t_i - t_{i-1} = \Delta t$, the increments*

$$W(t_1) - W(t_0),\ W(t_2) - W(t_1),\ \ldots,\ W(t_n) - W(t_{n-1})$$

are random variables independent between themselves where

$$W(t + \Delta t) - W(t) \sim \mathcal{N}(0, \Delta t),$$

this is, the distribution of the difference of the increments depends only on Δt and not on t.

An useful way of writing a Wiener process[2] is by using a standard normal random variable $z(0,1)$ (mean $= 0$ and variance $= 1$):

$$\Delta W(t) = z(0,1)\sqrt{\Delta t}. \tag{3}$$

This must be valid from property **(3)** from definition (1.3). Suppose we had $\Delta W(t) = \alpha z(0,1)$, for $\alpha \in \mathbb{R}$ constant, then evidently the expected value would be 0 for any $\alpha \in \mathbb{R}$ anyway:

$$\mathbb{E}[\alpha z(0,1)] = \alpha \mathbb{E}[z(0,1)] = 0.$$

If property **(3)** from (1.3) holds, then we must have

$$\mathrm{Var}[\Delta W(t)] = \Delta t,$$

but for that to happen, $\alpha = \sqrt{\Delta t}$, since

$$\mathrm{Var}[\alpha z(0,1)] = \alpha^2 \mathrm{Var}[z(0,1)] = \alpha^2,$$

yielding (3). From this demonstration we conclude that

$$\begin{cases} \mathbb{E}[\Delta W(t)] = 0 \\ \mathrm{Var}[\Delta W(t)] = \Delta t \end{cases} \tag{4}$$

Notice that under property **(3)** from the Wiener process definition (1.3), choosing $t + \Delta t \to t$ and $t \to t_0$, we have that

$$W(t) - W(t_0) = W(t) - 0 \sim \mathcal{N}(0, t)$$

according to property **(1)**, thus:

$$W(t) \sim \mathcal{N}(0, t), \tag{5}$$

i.e. not only the increments $W(t + \Delta t) - W(t)$ have a normal distribution as in property **(3)** from (1.3), but also the Wiener process itself $W(t)$ [7].

[1] The almost surely convergence is defined by a sequence of random variables X_n fo which $\mathbb{P}(\lim_{n \to \infty} X_n = X) = 1$

[2] This explicit way of writing the Wiener process is vastly used in numerical approaches of the GBM such as Monte Carlo simulations by sampling pseudo-random standard normal variables $z(0,1)$.

It's easy to demonstrate that a Wiener process $W(t)$ is a martingale. To show that, we prove each property from definition (1.2) for a Wiener process given definition (1.3).

By construction, $W(t)$ is $\mathcal{F}(t)$-measurable, being thus adapted [4, 6], concluding that property **(1)** is true. We can calculate the expected value of $|W(t)|$, knowing that it follows a normal distribution of mean 0 and variance t according to (5), using the symmetry of the product of functions inside the integral and a change of variables $x = w^2/2t$ (so that $t\mathrm{d}w = x\mathrm{d}x$) [1]:

$$\begin{aligned}
\mathbb{E}[|W(t)|] &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} |w| e^{-\frac{(w-0)^2}{2t}} \mathrm{d}w \\
&= \frac{2}{\sqrt{2\pi t}} \int_{0}^{\infty} w e^{-\frac{(w-0)^2}{2t}} \mathrm{d}w \\
&= \frac{2t}{(2\pi t)^{1/2}} \int_{0}^{\infty} e^{-x} \mathrm{d}x \\
&= \sqrt{\frac{2t}{\pi}} (0 + e^{-0}) \\
&= \sqrt{\frac{2t}{\pi}}.
\end{aligned}$$

thus the expected value on the absolute value of $W(t)$ is finite proving property **(2)** from (1.2):

$$\mathbb{E}[|W(t)|] = \sqrt{\frac{2t}{\pi}} < \infty. \tag{6}$$

For the third martingale property in (1.2), we use property **(3)** from the Wiener process definition: Since the increments $W(s) - W(t)$ are independent of the past, their future increments will be independent of all information available up to time s, which is exactly what $\mathcal{F}(s)$ encodes [6, 1], hence:

$$\mathbb{E}[W(s) - W(t)|\mathcal{F}(s)] = \mathbb{E}[W(s) - W(t)] = 0 \tag{7}$$

where the last equality is given by the fact that $W(s) - W(t) \sim \mathcal{N}(0, s-t)$ thus having expected value 0.

For $t \in [t_0, T]$ and $s \in [t_0, T]$, notice that $W(t) = W(s) + [W(t) - W(s)]$, using (7) we calculate the conditional expectation of $W(t)$:

$$\begin{aligned}
\mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[W(s) + \{W(t) - W(s)\}|\mathcal{F}(s)] \\
&= \mathbb{E}[W(s)|\mathcal{F}(s)] + \mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] \\
&\overset{(7)}{=} \mathbb{E}[W(s)|\mathcal{F}(s)] + 0 \\
&= W(s).
\end{aligned}$$

Thus we conclude that the Wiener process $W(t)$ is a martingale [4].

A more general stochastic process is known as *Itô's process* and it is defined as follows [3, 1].

Definition 1.4. *The solution $X(t)$ of the following SDE is known as Itô process [3]:*

$$\mathrm{d}X(t) = a(X(t), t)\mathrm{d}t + b(X(t), t)\mathrm{d}W(t), \tag{8}$$

for $X(t_0) = X_0$ *(constant),* $\mathrm{d}W(t)$ *a differential of the Wiener process:*

$$\mathrm{d}W(t) = \lim_{\Delta t \to 0}[W(t + \Delta t) - W(t)] \tag{9}$$

and for $a = a(X(t), t)$ *and* $b = b(X(t), t)$ *(usually called drift and diffusion term) being functions that don't increase to rapidly, meaning functions* $a = a(X(t), t)$ *and* $b = b(X(t), t)$ *must satisfy the so called Lipschitz conditions [5, 4]:*

$$\begin{cases} |a(x,t) - a(y,t)|^2 + |b(x,t) - b(y,t)|^2 \leq k_1|x - y|^2 \\ |a(x,t)|^2 + |b(x,t)|^2 \leq k_2(1 + |x|^2) \end{cases}$$

for $k_1,\ k_2 \in \mathbb{R}^+$ *and* $x, y \in \mathbb{R}$.

The drift and diffusion terms receive these names due to their influence in the time evolution of $\mathrm{d}X(t)$. Imagine that for an Itô's process (8) we had the diffusion term $b(X(t), t) = 0$ and a drift term exclusively dependent on time $a = a(t)$ (not as a function of $X(t)$). In this scenario the differential equation becomes purely deterministic:

$$\mathrm{d}X = a\mathrm{d}t.$$

Integrating both sides from t_0 to t ($t_0 \leq t \leq T$) yields a straight line

$$X(t) = a(t - t_0) + X(t_0)$$

such that a is the slope. This creates some intuition about the term since it relates it to the directional movement of the stochastic process [2] (remembering that the usual process (8) will have the noisy influence of the stochastic term).

The expected value of an Itô's process (8) when $a = a(t)$ and $b = b(t)$ exclusively is given by [2]:

$$\mathbb{E}(\mathrm{d}X) = a\mathrm{d}t. \tag{10}$$

which can be quickly derived as follows (here we adopt the Δ notation, which can be made into differential through the limit $\Delta t \to 0$):

$$\mathbb{E}(\Delta X) = a\mathbb{E}(\Delta t) + b\mathbb{E}(\Delta W) \overset{(4)}{=} a\Delta t + b \cdot 0 = a\Delta t.$$

From (10) we can see that the drift term can be regarded as the expected change of the Itô's process over an infinitesimal interval $\mathrm{d}t$. If we have a pure Brownian motion $\mathrm{d}X(t) = b\mathrm{d}W(t)$, the drift will be 0, meaning this process expected value won't change in time, or according to our intuition, there will be no directional movement of the stochastic process as a whole. Summarizing, the drift term is associated with the average behavior of the process over time [2, 7].

The name drift can be associated to the nomenclature employee in physics for the average velocity of a particle perturbed by the influence of an applied field (namely electrons in a conductor): drift velocity. Among the perturbed motion (analogous to the Brownian motion although sometimes deterministic) there's also a net directional trend due to the field — this net trend is the drift.

The diffusion term $b(X(t), t)$ is evidently associated to the randomness of the process over time as it acts as the coefficient of the Wiener differential $\mathrm{d}W(t)$, or in other words as an amplitude or scale for the stochastic part of the process [2]. A large value of $b(X(t), t)$

classifies the process as *volatile*, meaning it has a higher degree of uncertainty associated with it and on the other hand, when $b(X(t), t) = 0$ the process becomes deterministic, i.e. no uncertainty. This element is essential when modeling market prices in order to replicate it's innate randomness [5].

To illustrate the influence of the drift $a(X(t), t)$ and diffusion $b(X(t), t)$ coefficients, we plot one possible stochastic process path in figure (1) using both coefficients as constants. This visualization helps highlight how the random fluctuations from the diffusion term $b\Delta W(t)$ (represented by the blue area between the path and the drift line) modify the otherwise linear deterministic trend from the drift $a\Delta t$ (represented in the figure by the dashed red line).

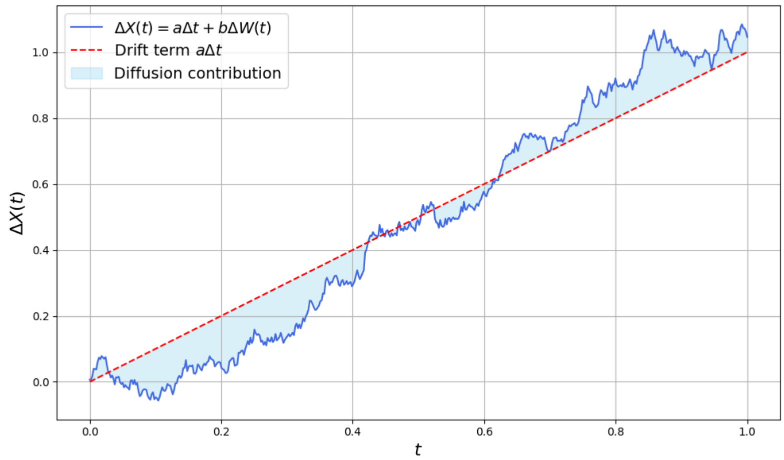

Figure 1: Sample path of Itô process $X(t) - X(0) = a(t-0) + bW(t)$ for $X(0) = 0$, time interval $t \in [0, 1]$, number of time steps $N = 500$ and drift $a = 0.3$ and diffusion $b = 0.3$ coefficients (constant). The graphic shows the drift term trend and diffusion influence to the process.

The variance of an Itô's process (8) when $a = a(t)$ and $b = b(t)$ (exclusively time dependent) is given by the following equation

$$\mathrm{Var}(\mathrm{d}X) = b^2 \mathrm{d}t, \tag{11}$$

which can be derived by using the expected value (4) [2, 1]:

$$\begin{aligned}
\mathrm{Var}(\Delta X) &= \mathrm{Var}(a\Delta t) + \mathrm{Var}(b\Delta W) + 2\mathrm{Cov}(a\Delta t, b\Delta W) \\
&= 0 + \mathrm{Var}(b\Delta W) + 0 \\
&\stackrel{(4)}{=} b^2\{\mathbb{E}[z^2\Delta t] - \mathbb{E}[z\sqrt{\Delta t}]^2\} \\
&= b^2\Delta t[\mathbb{E}(z^2) - \mathbb{E}(z)^2] \\
&= b^2\Delta t\mathrm{Var}(z) \\
&= b^2\Delta t,
\end{aligned}$$

where we've used the variance of the sum of two random variables (considered here $a\Delta t$ and $b\Delta W(t)$) and that the covariance term vanishes due to $a\Delta t$'s deterministic nature. Equation (11) shows that the squared diffusion term represents the variance of the Itô process over an infinitesimal interval $\mathrm{d}t$, connecting it directly with the randomness and spread when compared to the average behavior (diffusion) of the stochastic process[3][6, 2].

Notice that since the drift term is deterministic and $\mathrm{d}W(t) \propto z(0,1) \sim \mathcal{N}(0,1)$ the Itô's process must follow a normal distribution [3, 5], with the expected value (10) and variance (11) (for non-stochastic drift and diffusion terms):

$$\Delta X \sim \mathcal{N}(a\Delta t, b^2 \Delta t) \tag{12}$$

Given a function $f(X(t), t)$ of on an Itô process (8) we apply a powerful tool called *Itô lemma* [3] in order to derive the function dynamics, i.e. $\mathrm{d}f(X(t), t)$. It works analogously to a Taylor expansion for randomic processes when we use the $\Delta t \to 0$ limit and ignore higher order terms, namely higher than first order. The demonstration of Itô's lemma will be displayed in the Appendice section.

Theorem 1.1. *Let $X(t)$ be an Itô process as defined in equation (8). Consider a function $f = f(X(t), t)$, depending on $X(t)$ and time t, with continuous first and second partial derivatives with respect to X, and a continuous first partial derivative with respect to t. Then, the stochastic process $f(X(t), t)$ follows the dynamics given by [3]:*

$$\mathrm{d}f(X(t), t) = \left[\frac{\partial f}{\partial t} + a(X,t)\frac{\partial f}{\partial X} + \frac{1}{2} b(X,t)\frac{\partial^2 f}{\partial X^2}\right]\mathrm{d}t + b(X,t)\frac{\partial f}{\partial X}\mathrm{d}W(t). \tag{13}$$

Notice that the lemma equation (13) in (1.1) is itself an Itô process with

$$\text{drift} \to \frac{\partial f}{\partial t} + a(X,t)\frac{\partial f}{\partial X} + \frac{1}{2} b(X,t)\frac{\partial^2 f}{\partial X^2}, \quad \text{diffusion} \to b(X,t)\frac{\partial f}{\partial X}.$$

Appendices

Derivation of Itô's Lemma

The Taylor series expansion for a function $f = f(X(t), t)$ (of two variables) around (X_0, t_0) (with $X(t_0) = X_0$) with $\Delta X = X - X_0$ and $\Delta t = t - t_0$ is the given by

$$\begin{aligned} f(X,t) = f(X_0, t_0) &+ \left.\frac{\partial f}{\partial t}\right|_{t=t_0} \Delta t + \left.\frac{\partial f}{\partial X}\right|_{X=X_0} \Delta X + \\ &+ \frac{1}{2}\left.\frac{\partial^2 f}{\partial t^2}\right|_{t=t_0} (\Delta t)^2 + \frac{1}{2}\left.\frac{\partial^2 f}{\partial X^2}\right|_{X=X_0} (\Delta X)^2 \\ &+ \left.\frac{\partial^2 f}{\partial X \partial t}\right|_{(X,t)=(X_0,t_0)} \Delta X \Delta t + \cdots, \end{aligned}$$

where we take the limits $\lim_{t\to t_0} \Delta t = \mathrm{d}t$ and $\lim_{X \to X_0} \Delta X = \mathrm{d}X$. Neglecting higher order terms of $\mathrm{d}t$ (this is, up to $\mathrm{d}t$ and higher order terms), we arrive at

$$\mathrm{d}f(X(t), t) = \frac{\partial f}{\partial t}\mathrm{d}t + \frac{\partial f}{\partial X}\mathrm{d}X + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(\mathrm{d}X)^2. \tag{14}$$

[3] Here we relate the variance with the uncertainty of the process given its definition: $\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2]$ for a random variable X with a distribution of mean μ. This expression captures the average squared deviation of X from its mean, quantifying how much the values of X typically fluctuate around μ.

Notice that a differential Wiener process can be written as in (3), thus $\mathrm{d}X(t)\mathrm{d}t$ according to (8) depends on $\mathrm{d}t^2$ (for the drift term) and $\mathrm{d}t^{3/2}$ (for the diffusion term), both higher than $\mathrm{d}t$'s order. The term $(\mathrm{d}X)^2$ must be developed carefully:

$$\begin{aligned}(\mathrm{d}X)^2 &\overset{(8)}{=} (a\mathrm{d}t + b\mathrm{d}W)(a\mathrm{d}t + b\mathrm{d}W)\\ &= a^2(\mathrm{d}t)^2 + 2ab\mathrm{d}t\mathrm{d}W + b^2(\mathrm{d}W)^2\\ &\overset{(3)}{=} a^2(\mathrm{d}t)^2 + 2abz(\mathrm{d}t)^{3/2} + b^2z^2\mathrm{d}t\\ &\approx b^2z^2\mathrm{d}t\end{aligned}$$

where we neglected the second and 3/2-order $\mathrm{d}t$ terms. We now study the nature of the $z^2\mathrm{d}t$ term by calculating its expected value through

$$\mathbb{E}(z^n) = \begin{cases}\frac{(n)!}{2^{n/2}(n/2)!}, & \text{if n is even,}\\ 0, & \text{if n is odd}\end{cases} \tag{15}$$

for $n \in \mathbb{Z}_{\geq 0}$ and $z \sim \mathcal{N}(0,1)$ a standard normal random variable [4]:

- **Expected Value**:

$$\mathbb{E}[(\mathrm{d}W)^2] = \mathbb{E}[z^2\mathrm{d}t] = \mathrm{d}t\ \mathbb{E}[z^2] \overset{(15)}{=} \frac{2!}{2^1 \cdot 1!}\mathrm{d}t = \mathrm{d}t$$

- **Variance**:

$$\begin{aligned}\mathrm{Var}[(\mathrm{d}W)^2] &= \mathbb{E}[(\mathrm{d}W)^4] - \mathbb{E}[(\mathrm{d}W)^2]^2\\ &= (\mathrm{d}t)^2\mathbb{E}(z^4) - (\mathrm{d}t)^2\mathbb{E}(z^2)^2\\ &\overset{(15)}{=} (\mathrm{d}t)^2\left[\frac{4!}{2^2 2!} - 1^2\right]\\ &= (\mathrm{d}t)^2(3-1)\\ &= 2(\mathrm{d}t)^2\\ &\approx 0.\end{aligned}$$

This implies that the variance of $(\mathrm{d}W)^2$ converges to 0 quadratically while the expected value converges to 0 linearly in $\mathrm{d}t$ when $\mathrm{d}t \to 0$, vanishing the variance according to our first order approximation in $\mathrm{d}t$. Since the variance is 0 for $\mathrm{d}t \to 0$, then $(\mathrm{d}W)^2$ is deterministic and equal to its expected value:

$$(\mathrm{d}W)^2 \approx \mathrm{d}t, \ \text{when } \mathrm{d}t \to 0. \tag{16}$$

This statement can be intuitively derived graphically. For the purpose of the visualization and practicability, lets consider a normal distributed variable $X \sim \mathcal{N}(\mu, \sigma^2)$, presented in figure (2).

[4] This derivation can be made by using the moment generating function of the standard normal distribution $M_X(\tau) = e^{\tau^2/2}$ and knowing that $\mathbb{E}(X^n) = \frac{\mathrm{d}^n M_X}{\mathrm{d}\tau^n}(\tau)\Big|_{\tau=0}$, noticing that $\frac{\mathrm{d}^n M_X}{\mathrm{d}\tau^n}(\tau) = H_n(\tau)M_X(\tau)$ where $H_n(\tau)$ is the nth Hermite polynomial.

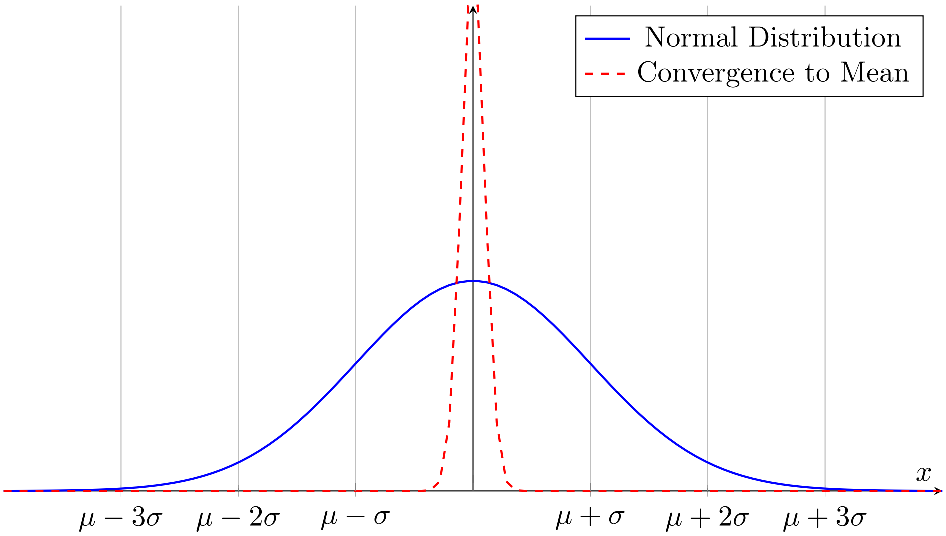

Figure 2: Illustration of a standard normal distribution (blue curve) with marked intervals at $\mu \pm \sigma$, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$. The red dashed curve approximates a Dirac delta function, representing convergence of probability mass toward the mean.

Considering figure (2), in the limit where $\sigma \to 0$ (intervals becoming smaller: $\mu \pm \sigma$, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$), we have that the blue curve becomes more centered sharped, or equivalently there's a higher change of sampling a value close to μ. As we take the limit $\sigma \to 0$ further, the distribution becomes a spike around μ (the red dashed line in the illustration) called Dirac delta function [5], sampling exclusively the mean value μ. At this point our random variable becomes deterministic, with value $X = \mu$.

We summarize the derived results regarding the products of $\mathrm{d}t$ with $\mathrm{d}W(t)$ with the so called Itô's multiplication table:

	$\mathrm{d}t$	$\mathrm{d}W(t)$
$\mathrm{d}t$	0	0
$\mathrm{d}W(t)$	0	$\mathrm{d}t$

Table 1: Itô's multiplication table for a Wiener process $\mathrm{d}W(t)$.

With the derivation of $(\mathrm{d}W)^2 \approx \mathrm{d}t$ for $\mathrm{d}t \to 0$, we figure out that

$$(\mathrm{d}X)^2 \approx b^2 \mathrm{d}t \tag{17}$$

from (11). With this result in mind together and the definition of Itô's process (8), we

[5] Formally the Dirac delta $\delta(x - x_0)$ is not a function but a distribution, defined by

$$\delta(x - x_0) = \begin{cases} \infty, & \text{if } x = x_0, \\ 0, & \text{otherwise.} \end{cases}$$

continue from (14) by replacing $\mathrm{d}X$ and $(\mathrm{d}X)^2$:

$$\begin{aligned}
\mathrm{d}f(X,t) &= \frac{\partial f}{\partial t}\mathrm{d}t + \frac{\partial f}{\partial X}\left(a\mathrm{d}t + b\mathrm{d}W\right) + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\left(b^2\mathrm{d}t\right) \\
\mathrm{d}f(X,t) &= \left(\frac{\partial f}{\partial t}\mathrm{d}t + a\frac{\partial f}{\partial X} + \frac{b^2}{2}\frac{\partial^2 f}{\partial X^2}\right)\mathrm{d}t + b\frac{\partial f}{\partial X}\mathrm{d}W,
\end{aligned}$$

matching (13) and concluding our demonstration.

References

[1] Lech A. Grzelak and Cornelis W. Oosterlee. *Mathematical Modeling and Computation in Finance: With Exercises and Python and MATLAB Computer Codes*. World Scientific Publishing Europe Ltd, 2019. ISBN: 978-1786348050.

[2] John C. Hull. *Options, Futures, and Other Derivatives*. 10th ed. Pearson, 2022. ISBN: 978-1-292-40856-2.

[3] Kiyosi Ito et al. *On stochastic differential equations*. Vol. 4. American Mathematical Society New York, 1951.

[4] Ioannis Karatzas and Steven E Shreve. *Brownian motion*. Springer, 2021, pp. 47–127.

[5] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[6] Steven E Shreve et al. *Stochastic calculus for finance II: Continuous-time models*. Vol. 11. Springer, 2004.

[7] Paul Wilmott, Jeff Dewynne, and Sam Howison. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, 1994. ISBN: 978-0952208204.