Chapter 5

Stochastic Differential Equations

We would like to introduce stochastic ODE's without going first through the machinery of stochastic integrals.

5.1 Itô Integrals and Itô Differential Equations

Let us start with a review of the invariance principle. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables such that $\mathbf{E}\xi_n = 0$, $\mathbf{E}\xi_n^2 = 1$. Define X_t^N by

$$X_{t_n}^N = \frac{\sum_{i=1}^n \xi_i}{\sqrt{N}}, \qquad t_n = \frac{n}{N}, \qquad 0 \le n \le N \tag{5.1}$$

and piecewise linear interpolation, then the invariance principle asserts that

$$X^N \xrightarrow{d} W \tag{5.2}$$

in distribution.

Alternatively, we can define $\{X_{t_n}^N\}_{n=1}^N$ by the recursion relation

$$X_{t_{n+1}}^N = X_{t_n}^N + \sqrt{\Delta t}\,\xi_{n+1}, \qquad X_0^N = 0, \tag{5.3}$$

where $\Delta t = N^{-1}$. We can think of (5.3) as a forward Euler scheme for solving the differential equation

$$dX_t = dW_t \tag{5.4}$$

Obviously both (5.3) and (5.4) are a bit unusual. In (5.3), the multiplier in front of ξ_{n+1} is $\sqrt{\Delta t}$ instead of the usual Δt. In (5.4) we write $dX_t = dW_t$ instead of $\dot{X}_t = \dot{W}_t$. This is because $\dot{W}_t$ is not a standard stochastic process, but rather a generalized stochastic process as in generalized functions. In fact one can define $\dot{W}_t$ as a Gaussian process with mean zero and covariance

$$K(s,t) = \delta(t-s).$$

This is a very important process called the Gaussian white noise. But its sample path are not the standard functions, but rather distributions, see [5].

We can now combine (5.4) with standard ordinary differential equation and study

$$dX_t = b(X_t, t)dt + dW_t. \tag{5.5}$$

One can think of this as the distributional limit of the forward Euler scheme

$$X^N_{t_{n+1}} = X^N_{t_n} + \varDelta t\, b(X_{t_n}, t_n) + \sqrt{\varDelta t}\, \xi_{n+1} \tag{5.6}$$

where as before $\{\xi^n\}_{n=1}^{\infty}$ is a sequence of i.i.d. random variables such that $\mathbf{E}\xi_n = 0$, $\mathbf{E}\xi_n^2 = 1$. In fact if we denote by X^N_t the process obtained using (5.6) and the piecewise linear interpolation, then

Theorem 5.1.1. *There exists a stochastic process X_t such that*

$$\mathbf{E}|X^N_t - X_t| \le C\sqrt{\varDelta t} \tag{5.7}$$

for $t \in [0,1]$, where C is independent of $\varDelta t = 1/N$, and

$$\left|\mathbf{E}g[X^N_{[0,1]}] - \mathbf{E}g[X_{[0,1]}]\right| \le C\varDelta t \tag{5.8}$$

for any continuous functional g on $C[0,1]$, where C may depend on g but not on $\varDelta t$.

More generally, we can consider SDE's of the type

$$dX_t = b(X_t, W_{[0,t]}, t)dt + \sigma(X_t, W_{[0,t]}, t)dW_t \tag{5.9}$$

where B and *sigma* are functions of X_t and t, and functional of $W_{[0,t]}$. Notice that they are nonanticipative functional, i.e. the Wiener process up to time t only enters the right hand-sidde of (5.9). (5.9) is defined it as the limit of the forward Euler scheme

$$X_{n+1} = X_n + \varDelta t\, b(X_n, W_{m\le n}, t_n) + \sqrt{\varDelta t}\, \sigma(X_n, W_{m\le n}, t_n)\xi_{n+1}. \tag{5.10}$$

where, for simplicity we have used the slightly abusive notation $X^N_{t_n} = X_n$. We can also write this as

$$X_{n+1} = X_n + \varDelta t\, b(X_n, W_{m\le n}, t_n) + \sigma(X_n, W_{m\le n}, t_n)(W_{n+1} - W_n). \tag{5.11}$$

where W_t is the Wiener process and $W_n = W_{t_n}$.

A special case of (5.9) is stochastic integrals

$$dX_t = f(W_t, t)dW_t, \qquad X_0 = 0, \tag{5.12}$$

where f is continuous, whose solution can be expressed as

$$X_t = \int_0^t f(W_s, s)dW_s. \tag{5.13}$$

The meaning of X_t is define as the limit of

$$X_{n+1} = X_n + f(W_n, t_n)(W_{n+1} - W_n), \tag{5.14}$$

or in other words

$$X_n = \sum_{j=0}^{n} f(W_j, t_j)(W_{j+1} - W_j). \tag{5.15}$$

We can think of (5.15) as a "Riemann sum" in which the representative point inside each subinterval is the left-most point. This definition of the stochastic integral is called the Itô integral.

Itô integrals have several very special properties.

Theorem 5.1.2 (Itô isometry). *Itô integrals satisfy*

$$\begin{aligned} &\mathbf{E}\int_0^t f(W_s, s)dW_s = 0, \\ &\mathbf{E}\Big(\int_0^t f(W_s, s)dW_s\Big)^2 = \int_0^t \mathbf{E}f^2(W_s, s)ds. \end{aligned}$$

Proof. Let

$$I_n = \sum_{j=0}^{n} f(W_j, t_j)(W_{j+1} - W_j),$$

then

$$\begin{aligned} \mathbf{E}I_n &= \sum_{j=0}^{n} \mathbf{E}(f(W_j, t_j)(W_{j+1} - W_j)) \\ &= \sum_{j=0}^{n} \mathbf{E}f(W_j, t_j)\mathbf{E}(W_{j+1} - W_j) = 0, \end{aligned}$$

where we use the fact that W_j and $W_{j+1} - W_j$ are independent. Similarly

$$\begin{aligned} \mathbf{E}I_n^2 &= \sum_{i,j=0}^{n} \mathbf{E}(f(W_i, t_i)f(W_j, t_j)(W_{i+1} - W_i)(W_{j+1} - W_j)) \\ &= \sum_{j=0}^{n} \mathbf{E}f^2(W_j, t_j)(t_{j+1} - t_j) \end{aligned}$$

where we use the fact that $W_{i+1} - W_i$ and $W_{j+1} - W_j$ are independent unless $i = j$. Taking the limit as $\Delta t \to 0$, we obtain the equations in the theorem. $\square$

Another important property of Itô integral is that $X_t = \int_0^t f(W_s, s)dW_s$ is a martingale, i.e.

$$\mathbf{E}_{X_s}(X_t) = X_s,$$

where $\mathbf{E}_{X_s}(X_t)$ denotes the expectation of X_t conditional on X_s. However since we will not discuss Martingale theory, we will not pursue this further.

Theorem 5.1.3 (Itô formula). *Let f be a smooth function, assume that X_t is the solution of the SDE (5.9), then $g_t = f(X_t, t)$ satisfies the SDE*

$$\begin{aligned} df(X_t,t) &= \dot{f}(X_t,t)dt + f'(X_t,t)dX_t + \tfrac{1}{2}f''(X_t,t)\sigma^2(X_t,t)dt \\ &= \dot{f}(X_t,t)dt + \big(f'(X_t,t)b(X_t,t) + \tfrac{1}{2}f''(X_t,t)\sigma^2(X_t,t)\big)\,dt \\ &\quad + f'(X_t,t)\sigma(X_t,t)dW_t, \end{aligned} \tag{5.16}$$

where $\dot{f}(x,t) = \partial f/\partial t$, $f'(x,t) = \partial f/\partial x$.

Itô formula is the analog of chain rule in ordinary differential calculus. However ordinary chain rule would give

$$df(X_t) = \dot{f}(X_t,t)dt + f'(X_t,t)dX_t.$$

Here because of the non-smooth nature of X_t, we have the additional term that depends on f''.

The proof of Itô formula can be outlined as follows. For simplicity we assume that f does not depend explicitly on t. We Taylor expand $f(X_{n+1}) - f(X_n)$ using (5.10) and keeping terms up to $O(\Delta t)$ using $W_{j+1} - W_j = O(\sqrt{\Delta t})$:

$$\begin{aligned} &f(X_{n+1}) - f(X_n) \\ &\quad = f'(X_n)(X_{n+1} - X_n) + \tfrac{1}{2}f''(X_n)(X_{n+1} - X_n) + \cdots \\ &\quad = f'(X_n)(X_{n+1} - X_n) + \tfrac{1}{2}f''(X_n)(\Delta t b(X_n,t_n) \\ &\qquad\qquad + \sigma(X_n,t_n)(W_{n+1} - W_n))^2 + O((\Delta t)^{3/2}) \\ &\quad = f'(X_n)(X_{n+1} - X_n) \\ &\qquad + \tfrac{1}{2}f''(X_n)\sigma^2(X_n,t_n)(W_{n+1} - W_n)^2 + O((\Delta t)^{3/2}). \end{aligned}$$

The Itô formula follows in the limit as $n \to \infty$ since

$$\tfrac{1}{2}f''(X_n)\sigma^2(X_n,t_n)(W_{n+1} - W_n)^2 \to \tfrac{1}{2}f''(X_t)\sigma^2(X_t,t)dt$$

as implied by the following lemma which can be written formally as $(dW_t)^2 = dt$.

Lemma 5.1.1. *Let $W_j = W_{t_j}$ with $t_j = j/N$. Then*

$$\sum_{j=1}^{n} (W_j - W_{j-1})^2 \to t \quad a.s.$$

as $n, N \to \infty$, $n/N \to t$.

Proof. The limit of this sum can be estimated from

$$\frac{1}{N}\sum_{j=1}^{n} \xi_j^2 = \frac{n}{N}\frac{1}{n}\sum_{j=1}^{n} \xi_j^2.$$

In this last expression $n/N \to t$ by assumption; the remainder is the rescaled sum of the i.i.d random variables ξ_j^2 with mean equal to one which therefore converges to 1 as $n \to \infty$ by SLLN. □

Example 5.1.1. We compute

$$\int_0^t W_s dW_s.$$

Using the definition of Itô integral, this integral is the limit of

$$\begin{aligned}\sum_j W_j(W_{j+1} - W_j) &= \tfrac{1}{2}\sum_j (W_j - W_{j+1} + W_j + W_{j+1})(W_{j+1} - W_j) \\ &= \tfrac{1}{2}\sum_j (W_{j+1}^2 - W_j^2) - \tfrac{1}{2}\sum_j (W_{j+1} - W_j)^2.\end{aligned}$$

Therefore

$$\int_0^t W_s dW_s = \tfrac{1}{2}W_t^2 - \tfrac{1}{2}t. \tag{5.17}$$

This result can also be obtained via a straight forward application of Itô formula which gives for $f(x) = \frac{1}{2}x^2$

$$\tfrac{1}{2}dW_t^2 = W_t dW_t + \tfrac{1}{2}dt \tag{5.18}$$

Integrating (5.18) we obtain (5.17).

Compared with the standard relation $\int_0^t f(s)df(s) = \frac{1}{2}(f^2(t) - f^2(0))$ for smooth functions, we see the presence of the extra term $-\frac{1}{2}t$. It is forced to be there by the relation $\mathbf{E}\int_0^t W_s dW_s = 0$.

We can express (5.17) as

$$\int_0^t W_s dW_s = \frac{1}{2!}\sqrt{\frac{t}{2}}H_2(W_t/\sqrt{2t}),$$

where $H_2(x) = 4x^2 - 2$ is the second order Hermite polynomial. In the same fashion, we have

$$\int_0^t \Big(\int_0^s W_u dW_u\Big) dW_s = \tfrac{1}{2}\int_0^t (W_s^2 - s)dW_s.$$

Using Itô formula, we have

$$\int_0^t W_s^2 dW_s = \tfrac{1}{3}W_t^3 - \int_0^t W_s ds.$$

Hence, using

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds,$$

we obtain

$$\int_0^t \Big(\int_0^s W_u dW_u) dW_s = \frac{1}{3!}\left(\frac{t}{2}\right)^{3/2} H_3(W_t/\sqrt{2t}),$$

where $H_3(x) = 8x^3 - 12x$ is the third order Hermite polynomial. As shown below, we have for $n \geq 2$

$$\int_0^t dW_{t_1} \int_0^{t_1} dW_{t_2} \dots \int_0^{t_{n-1}} dW_{t_n} = \frac{1}{n!} \left(\frac{t}{2}\right)^{n/2} H_n(W_t/\sqrt{2t}),$$

where $H_n(x)$ is the n-th order Hermite polynomial

$$H_n(x) = (-1)^n e^{x^2} \left(\frac{d}{dx}\right)^n e^{-x^2}.$$

Example 5.1.2. We solve

$$dN_t = \alpha N_t dt + \beta N_t dW_t \tag{5.19}$$

Itô's formula with $f(x) = \log x$ gives

$$d \log N_t = \frac{1}{N_t}(\alpha N_t dt + \beta N_t dW_t) - \frac{1}{2N_t^2}\beta^2 N_t^2 dt.$$

Integrating we get

$$N_t = N_0 \exp\left(\alpha t - \tfrac{1}{2}\beta^2 t + \beta W_t\right).$$

Note that if W_t were smooth, we would have obtained

$$N_t = N_0 \exp(\alpha t + \beta W_t).$$

For $N_0 = 1$ and $\alpha = 0$, N_t reduces to

$$N_t = \exp\left(-\tfrac{1}{2}\beta^2 t + \beta W_t\right)$$

Using Rodrigues' formula,

$$e^{2az - a^2} = \sum_{n=0}^{\infty} \frac{H_n(z) a^n}{n!}$$

this can be expressed as

$$N_t = \sum_{n=0}^{\infty} \frac{H_n(W_t/\sqrt{2t})}{n!} \left(\frac{t}{2}\right)^{n/2} \beta^n.$$

But iteration of the equation for N_t shows that it can also be expressed as

$$N_t = 1 + \beta W_t + \sum_{n=2}^{\infty} \beta^n \int_0^t dW_{t_1} \cdots \int_0^{t_{n-1}} dW_{t_n}.$$

We deduce that

$$\int_0^t dW_{t_1} \cdots \int_0^{t_{n-1}} dW_{t_n} = \frac{1}{n!} \left(\frac{t}{2}\right)^{n/2} H_n(W_t/\sqrt{2t})$$

as asserted before.

Example 5.1.3. We solve

$$dX_t = -\gamma X_t dt + \sigma dW_t \tag{5.20}$$

This is the Ornstein-Uhlenbeck process. Using Itô formula we have

$$d(e^{\gamma t}X_t) = \gamma e^{\gamma t}X_t dt + e^{\gamma t}dX_t = \sigma e^{\gamma t}dW_t$$

Therefore we get

$$X_t = e^{-\gamma t}X_0 + \sigma\int_0^t e^{-\gamma(t-s)}dW_s.$$

This is Duhammel principle applied to (5.20). Let

$$Q_t = \int_0^t e^{-\gamma(t-s)}dW_s.$$

Q_t is a Gaussian process with mean and variance

$$\begin{aligned}
\mathbf{E}Q_t &= 0\\
\mathbf{E}Q_t^2 &= \int_0^t \mathbf{E}(e^{-\gamma(t-s)})^2 ds\\
&= e^{-2\gamma t}\int_0^t e^{2\gamma s}ds = \frac{1}{2\gamma}(1-e^{-2\gamma t})
\end{aligned}$$

As $t \to \infty$, the variance goes to $1/2\gamma$. Thus as $t \to \infty$

$$X_t \xrightarrow{d} N\left(0, \frac{\sigma^2}{2\gamma}\right) \tag{5.21}$$

(5.20) is a special case of the so-called Langevin equations. More generally, Langevin equation with potential function $V(x)$ is given by

$$dX_t = -\nabla V(X_t)dt + \sigma dW_t \tag{5.22}$$

5.2 Numerical Solutions of SDE's

We will discuss accuracy and stability issues.

5.2.1 Accuracy

Let $\{X_t^{\Delta t}\}$ be the numerical solution of a SDE with time step size Δt, and $\{X_t\}$ be the exact solution. There are two different notions of measuring the error $X_t - X_t^{\Delta t}$. We say that $\{X_t^{\Delta t}\}$ converges to $\{X_t\}$ with *strong order* α if

$$\max_{0\le t\le T}\mathbf{E}|X_t^{\Delta t} - X_t| \le C(\Delta t)^\alpha \tag{5.23}$$

holds with a constant C independent of Δt. We say that $\{X_t^{\Delta t}\}$ converges to $\{X_t\}$ with *weak order* β if

$$\max_{0\le t\le T} |\mathbf{E}g(X_t^{\Delta t}) - \mathbf{E}g(X_t)| \le C(\Delta t)^\beta \tag{5.24}$$

holds with a constant C independent of Δt but which may depend on the smooth function g. It can be shown that $\alpha \le \beta$, since

$$\begin{aligned} |\mathbf{E}g(X_t) - \mathbf{E}g(X_t^{\Delta t})| &\le \mathbf{E}\Big(|X_t - X_t^{\Delta t}| \int_0^1 |g'(X_t + \theta(X_t^{\Delta t} - X_t)|)d\theta\Big) \\ &\le M\mathbf{E}|X_t - X_t^{\Delta t}|, \end{aligned}$$

where $M = \max|g'|$.

Exercise 5.2.1. Show that $\alpha = \frac{1}{2}$, $\beta = 1$ for the forward Euler scheme discussed earlier.

In order to find numerical schemes that improve the order of accuracy, we proceed as follows. Consider the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \tag{5.25}$$

From Itô formula, we have

$$\begin{aligned} df(X_t) &= f'(X_t)dX_t + \tfrac{1}{2}f''(X_t)\sigma^2(X_t)dt \\ &= (L_1 f)(X_t)dt + (L_2 f)(X_t)dW_t \end{aligned}$$

where

$$\begin{aligned} (L_1 f)(x) &= f'(x)b(x) + \tfrac{1}{2}f''(x)\sigma^2(x) \\ (L_2 f)(x) &= f'(x)\sigma(x) \end{aligned}$$

Integrating, we get

$$f(X_{t+\Delta t}0 - f(X_t) = \int_t^{t+\Delta t} (L_1 f)(X_s)ds + \int_t^{t+\Delta t} (L_2 f)(X_s)dW_s. \tag{5.26}$$

Apply (5.26) with $f(x) = b(x)$, and $f(x) = \sigma(x)$ respectively, and use the result in

$$X_{t+\Delta t} - X_t = \int_t^{t+\Delta t} b(X_s)ds + \int_t^{t+\Delta t} \sigma(X_s)dW_s,$$

we get

$$\begin{aligned} X_{t+\Delta t} = X_t &+ \Delta t b(X_t) + \sigma(X_t)(W_{t+\Delta t} - W_t) \\ &+ \int_t^{t+\Delta t}\int_t^s (L_1 b)(X_\tau)d\tau ds + \int_t^{t+\Delta t}\int_t^s (L_2 b)(X_\tau)dW_\tau ds \\ &+ \int_t^{t+\Delta t}\int_t^s (L_1 \sigma)(X_\tau)d\tau dW_s + \int_t^{t+\Delta t}\int_t^s (L_2 \sigma)(X_\tau)dW_\tau dW_s. \end{aligned} \tag{5.27}$$

If we just keep the first two terms at the right hand side, we get Euler scheme. Out of the four double integrals, the last one is the biggest. Approximate the last integral by

$$\begin{aligned}
&\int_t^{t+\Delta t}\int_t^s (L_2\sigma)(X_\tau)dW_\tau dW_s \\
&\quad\approx (L_2\sigma)(X_t)\int_t^{t+\Delta t}\int_t^s dW_\tau dW_s \\
&\quad= \tfrac{1}{2}(L_2\sigma)(X_t)\left((W_{t+\Delta t}-W_t)^2-\Delta t\right).
\end{aligned} \tag{5.28}$$

Since

$$\begin{aligned}
\mathbf{E}(W_{\Delta t}^2-\Delta t)^2 &= \mathbf{E}W_{\Delta t}^4 - 2\Delta t\mathbf{E}W_{\Delta t}^2 + (\Delta t)^2 \\
&= 3(\Delta t)^2 - 2(\Delta t)^2 + (\Delta t)^2 = 2(\Delta t)^2,
\end{aligned}$$

the term in (5.28) is of order Δt; we also have

$$\mathbf{E}(W_{\Delta t}^2-\Delta t) = 0.$$

So the total error from the final term in (5.27) can be estimated as

$$\mathbf{E}(\text{total error})^2 \approx \frac{1}{\Delta t}O(\Delta t^2) = O(\Delta t). \tag{5.29}$$

This suggests that the error for the forward Euler scheme scales as $O(\sqrt{\Delta t})$.

This analysis also suggests an obvious extension of the forward Euler scheme that promises to be more accurate.

$$\begin{aligned}
X_{t+\Delta t} &= X_t + b(X_t)\Delta t + \sigma(X_t)(W_{t+\Delta t}-W_t) \\
&\quad + \tfrac{1}{2}(\sigma\sigma')(X_t)((W_{t+\Delta t}-W_t)^2-\Delta t).
\end{aligned}$$

This is the well-known Milstein scheme.

Despite the fact that Milstein's scheme is formally more accurate, it also uses the derivative of σ. This is unpleasant for many applications.

5.2.2 Stability

5.3 Path Integral Representation

We have seen in Chapter 4 that the Wiener measure over $[0,T]$ can be formally expressed as

$$d\mu_W = Z^{-1}\exp\left(-\tfrac{1}{2}\int_0^T \dot{h}^2(t)dt\right)Dh(\cdot)$$

The solution of the SDE

$$dX_t = b(X_t,t) + \sigma(X_t,t)dW_t.$$

can be viewed as a map between the Wiener path $W_{[0,T]}$ and $X_{[0,T]}$:

$$W_{[0,T]} \xrightarrow{\Phi} X_{[0,T]}.$$

Consequently, Φ induces another measures on $C[0,T]$, which is nothing but the distribution of $X_{[0,T]}$.

We now ask the question how the measure $d\mu_W$ change under the mapping Φ? Let us first consider the case when $\sigma = I$. In this case, we claim the measure $d\mu_X$ can be formally expressed as

$$d\mu_X = Z^{-1} \exp\left(-\tfrac{1}{2} \int_0^T (\dot{h}(t) - b(h(t),t))^2 dt \right) Dh(\cdot)$$

We proceed at the discrete level. The recurrence relation

$$X_{n+1} = X_n + b(X_n, t_n)\Delta t + W_{n+1} - W_n, \qquad X_0 = x,$$

in principle allows to express X_n as a function of $W_{m\le n}$ and x, i.e.

$$X_n = \Phi_n(W_{m\le n}, x).$$

Therefore the expectation of any function $F(X_{n\le N})$ can be expressed as

$$\mathbf{E}F(X_{n\le N}) = Z_N^{-1} \int_{\mathbb{R}^N} F(\Phi_{n\le N}(y_{m\le n}, x)) \exp\left(-I_{t_{n\le N}}(y_{n\le N})\right) Dy_{n\le N}$$

where $Dy_{n\le N} = dy_1 \cdots dy_N$, Z_N is a normalization factor, and $I_{t_{n\le N}}(y_{n\le N})$ is the kernel defined in the last chapter

$$I_{t_{n\le N}}(y_{n\le N}) = \tfrac{1}{2} \sum_{i=1}^{N} \frac{(y_i - y_{i-1})^2}{t_i - t_{i-1}}, \qquad y_0 = t_0 = 0$$

In the expression above, we wish to integrate over

$$h_n = \Phi_n(y_{m\le n}, x),$$

instead of y_n. Since the recurrence relation for X_n implies that

$$W_{n+1} - W_n = X_{n+1} - X_n - b(X_n, t_n)\Delta t,$$

it follows that the kernel $I_{t_{n\le N}}(y_{n\le N})$ can be expressed in terms of $h_{n\le N}$ simply as

$$\bar{I}_{t_{n\le N}}(h_{n\le N}) = \tfrac{1}{2} \sum_{i=1}^{N} \frac{(h_i - h_{i-1} - b(h_{i-1}, t_{i-1})\Delta t)^2}{t_i - t_{i-1}}, \qquad y_0 = t_0 = 0$$

On the other hand, the Jocobian of the transformation for $y_{n\le N}$ to $h_{n\le N}$ is one because the matrix

$$\frac{\partial h_n}{\partial y_m}$$

is a lower triangular matrix (i.e. it has zero element above the diagonal) with ones on the diagonal and hence has determinant one. Therefore we have

$$\mathbf{E}F(X_{n\le N}) = Z_N^{-1}\int_{\mathbb{R}^N} F(h_{n\le N})\exp\left(-\bar{I}_{t_{n\le N}}(h_{n\le N})\right) Dh_{n\le N}$$

In the limit as $\Delta t\to 0$, this expression formally gives

$$\mathbf{E}F[X_{[0,T]}] = \int F[h_{[0,T]}]d\mu_X[h_{[0,T]}],$$

with the measure $d\mu_X$ given before.

Quite remmarkably, we can give a rigorous meaning to these formal manipulations. To see how, notice that the kernel $\bar{I}_{t_{n\le N}}(h_{n\le N})$ can be expanded as (using $t_i - t_{i-1} = \Delta t$)

$$\begin{aligned}\bar{I}_{t_{n\le N}}(h_{n\le N}) &= \tfrac{1}{2}\sum_{i=1}^N \frac{(h_i-h_{i-1})^2}{\Delta t} - \sum_{i=1}^N b(h_{i-1},t_{i-1})(h_i-h_{i-1}) + \tfrac{1}{2}\sum_{i=1}^N b^2(h_{i-1},t_{i-1})\Delta t\\ &= I_{t_{n\le N}}(h_{n\le N}) - \sum_{i=1}^N b(h_{i-1},t_{i-1})(h_i-h_{i-1}) + \tfrac{1}{2}\sum_{i=1}^N b^2(h_{i-1},t_{i-1})\Delta t.\end{aligned}$$

Therefore we can also express the expectation of $\mathbf{E}F(X_{n\le N})$ as the ratio

$$\mathbf{E}F(X_{n\le N}) = A/B.$$

Here

$$\begin{aligned}A &= Z_N^{-1}\int_{\mathbb{R}^N} F(h_{n\le N})e^{Q(h_{n\le N})}\exp\left(-I_{t_{n\le N}}(h_{n\le N})\right) Dh_{n\le N}\\ &= \mathbf{E}\left(F(W^x_{n\le N})e^{Q(W^x_{n\le N})}\right),\end{aligned}$$

and

$$\begin{aligned}B &= Z_N^{-1}\int_{\mathbb{R}^N} e^{Q(h_{n\le N})}\exp\left(-I_{t_{n\le N}}(h_{n\le N})\right) Dh_{n\le N}\\ &= \mathbf{E}e^{Q(W^x_{n\le N})}\end{aligned}$$

where $W^x_n = x + W_n$ and

$$Q(h_{n\le N}) = \sum_{i=1}^N b(h_{i-1},t_{i-1})(h_i-h_{i-1}) - \tfrac{1}{2}\sum_{i=1}^N b^2(h_{i-1},t_{i-1})\Delta t$$

Taking the limit as $\Delta t\to 0$, we deduce that

$$\mathbf{E}F(X_{[0,T]}) = \frac{\mathbf{E}\left(e^{Q[W^x_{[0,T]}]}F[W^x_{[0,T]}]\right)}{\mathbf{E}e^{Q[W^x_{[0,T]}]}},$$

where $W_t^x = x + W_t$ is the shifted Wiener process and

$$Q[W^x_{[0,T]}] = \int_0^T b(W_t^x, t)dW_t - \tfrac{1}{2}\int_0^T b^2(W_t^x, t)dt.$$

Quite remarkably, this expression simplify even more because one can show that $\mathbf{E}e^{Q[W_{[0,T]}]} = 1$ and hence

$$\mathbf{E}F(X_{[0,T]}) = \mathbf{E}\left(e^{Q[W^x_{[0,T]}]}F[W^x_{[0,T]}]\right)$$

This formula is known as the *Girsanov formula*. To see that $\mathbf{E}e^{Q[W^x_{[0,T]}]} = 1$, let

$$B_t = \int_0^t b(W_s^x, s)dW_s - \tfrac{1}{2}\int_0^t b^2(W_s^x, s)ds$$

so that

$$dB_t = b(W_t^x, t)dW_t - \tfrac{1}{2}b^2(W_t^x, t)dt, \qquad B_0 = 0,$$

Therefore

$$de^{B_t} = e^{B_t}dB_t + \tfrac{1}{2}e^{B_t}b^2(W_t^x, t)dt = e^{B_t}b(W_t^x, t)dW_t$$

and we obtain

$$e^{B_t} = 1 + \int_0^t e^{B_s}b(W_s^x, s)dW_s$$

The first Itô isometry then implies that

$$\mathbf{E}e^{B_t} = 1 \qquad \forall t$$

and, in particular, $\mathbf{E}e^B_T = \mathbf{E}e^{Q[W_{[0,T]}]} = 1$.

The Girsanov formula asserts that $d\mu_X$ and $d\mu_W$ are absolutely continuous with respect to each other with *Radon-Nikodym derivative* given by

$$\frac{d\mu_X}{d\mu_W} = \exp\left(\int_0^1 \left(b(W_s^x, s)dX_s - \tfrac{1}{2}b^2(W_s^x, s)ds\right)\right).$$

A special case of this representation is the *Cameron-Martin formula*, for the transformation

$$X_t = W_t + \phi(t) \tag{5.30}$$

where ϕ is a smooth function. This can be obtained from SDE with $b(X, t) = \dot{\phi}(t)$. In this case, we get

$$\frac{d\mu_X}{d\mu_W} = \exp\left(\int_0^1 \left(\dot{\phi}(s)dW_s - \tfrac{1}{2}\dot{\phi}^2(s)ds\right)\right).$$

A slight generalization is the *Girsanov formula*. Consider two SDE's:

$$\begin{cases} dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t, & X_0 = x \\ dY_t = (b(Y_t, t) + c(Y_t, t))dt + \sigma(Y_t, t)dW_t, & Y_0 = x \end{cases}$$

(Note that the initial conditions are the same.) Then the distributions of $X_{[0,T]}$ and $Y_{[0,T]}$ are absolutely continuous with respect to each other. Moreover the Radon-Nikodym derivative is given by

$$\frac{d\mu_Y}{d\mu_X}[X.] = \exp\left(\int_0^T \phi(X_t,t)dW_t - \tfrac{1}{2}\int_0^T |\phi(X_t,t)|^2 dt\right),$$

where ϕ is the solution of

$$\sigma(x,t)\phi(x,t) = c(x,t).$$