

Wiener Process

Chapters 6, 7 and 8 offer a brief introduction to stochastic differential equations (SDEs). A standard reference for the material presented hereafter is the book by R. Durrett, “Stochastic Calculus: A Practical Introduction” (CRC 1998). For a discussion of the Wiener measure and its link with path integrals see e.g. the book by M. Kac, “Probability and Related Topics in Physical Sciences” (AMS, 1991).

1. The Wiener process as a scaled random walk

Consider a simple random walk $\{X_n\}_{n \in \mathbb{N}}$ on the lattice of integers \mathbb{Z} :

$$(1) \quad X_n = \sum_{k=1}^n \xi_k,$$

where $\{\xi_k\}_{k \in \mathbb{N}}$ is a collection of independent, identically distributed (i.i.d) random variables with $\mathbb{P}(\xi_k = \pm 1) = \frac{1}{2}$. The Central Limit Theorem (see the Addendum at the end of this chapter) asserts that

$$\frac{X_N}{\sqrt{N}} \rightarrow N(0, 1) \quad (\equiv \text{Gaussian variable with mean 0 and variance 1})$$

in distribution as $N \rightarrow \infty$. This suggests to define the piecewise constant random function W_t^N on $t \in [0, \infty)$ by letting

$$(2) \quad W_t^N = \frac{X_{\lfloor Nt \rfloor}}{\sqrt{N}},$$

where $\lfloor Nt \rfloor$ denotes the largest integer less than Nt and in accordance with standard notations for stochastic processes, we have written t as a subscript, i.e. $W_t^N = W^N(t)$.

It can be shown that as $N \rightarrow \infty$, W_t^N converges in distribution to a stochastic process W_t , termed the *Wiener process* or *Brownian motion*¹, with the following properties:

- (a) *Independence.* $W_t - W_s$ is independent of $\{W_\tau\}_{\tau \leq s}$ for any $0 \leq s \leq t$.
- (b) *Stationarity.* The statistical distribution of $W_{t+s} - W_s$ is independent of s (and so identical in distribution to W_t).
- (c) *Gaussianity.* W_t is a Gaussian process with mean and covariance

$$\mathbb{E}W_t = 0, \quad \mathbb{E}W_t W_s = \min(t, s).$$

¹The Brownian motion is termed after the biologist Robert Brown who observed in 1827 the irregular motion of pollen particles floating in water. It should be noted, however, that a similar observation had been made earlier in 1765 by the physiologist Jan Ingenhousz about carbon dust in alcohol. Somehow Brown’s name became associated to the phenomenon, probably because Ingenhouszian motion does not sound that good. Some of us with complicated names are sensitive to this story.

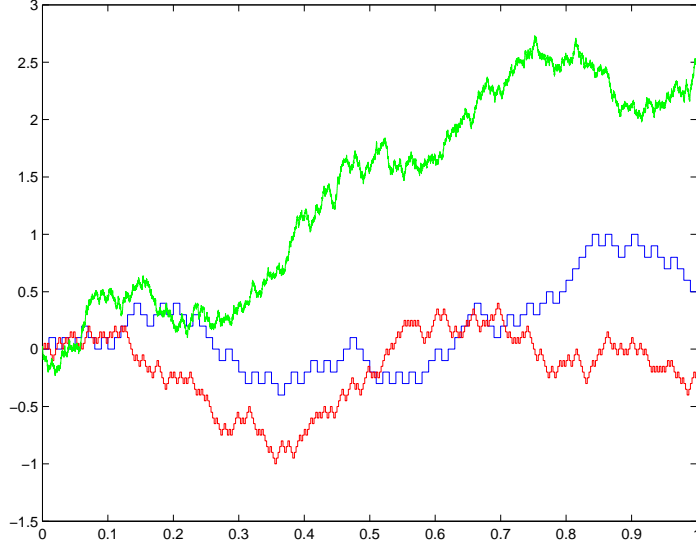


FIGURE 1. Realizations of W_t^N for $N = 100$ (blue), $N = 400$ (red), and $N = 10000$ (green).

(d) *Continuity.* With probability 1, W_t viewed as a function of t is continuous. To show independence and stationarity, notice that for $1 \leq m \leq n$

$$X_n - X_m = \sum_{k=m+1}^n \xi_k$$

is independent of X_m and is identically distributed to X_{n-m} . It follows that for any $0 \leq s \leq t$, $W_t - W_s$ is independent of W_s and satisfies

$$(3) \quad W_t - W_s \stackrel{d}{=} W_{t-s},$$

where $\stackrel{d}{=}$ means that the random process on both sides of the equality have the same distribution. To show Gaussianity, observe that at fixed time $t \geq 0$, W_t^N converges as $N \rightarrow \infty$ to Gaussian variable with mean zero and variance t since

$$W_t^N = \frac{X_{\lfloor Nt \rfloor}}{\sqrt{N}} = \frac{X_{\lfloor Nt \rfloor}}{\sqrt{\lfloor Nt \rfloor}} \frac{\sqrt{\lfloor Nt \rfloor}}{\sqrt{N}} \rightarrow N(0, 1)\sqrt{t} \stackrel{d}{=} N(0, t).$$

In other words,

$$(4) \quad \mathbb{P}(W_t \in [x_1, x_2]) = \int_{x_1}^{x_2} \rho(x, t) dx$$

where

$$(5) \quad \rho(x, t) = \frac{e^{-x^2/2t}}{\sqrt{2\pi t}}.$$

In fact, given any partition $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, the vector $(W_{t_1}^N, \dots, W_{t_n}^N)$ converges in distribution to a n -dimensional Gaussian random variable. Indeed,

using (3) recursively together with (4) and (5), it is easy to see that the probability density that $(W_{t_1}, \dots, W_{t_n}) = (x_1, \dots, x_n)$ is simply given by

$$(6) \quad \rho(x_n - x_{n-1}, t_n - t_{n-1}) \cdots \rho(x_2 - x_1, t_2 - t_1) \rho(x_1, t_1)$$

A simple calculation using

$$\mathbb{E}W_t = \int_{\mathbb{R}} x \rho(x, t) dx, \quad \mathbb{E}W_t W_s = \int_{\mathbb{R}^2} yx \rho(y - x, t - s) \rho(x, s) dx dy.$$

for $t \geq s$ and similarly for $t < s$ gives the mean and covariance specified in (b). Notice that the covariance can also be specified via

$$\mathbb{E}(W_t - W_s)^2 = |t - s|,$$

and this equation suggests that W_t is not a smooth function of t . In fact, it can be showed that even though W_t is continuous almost everywhere (in fact Hölder continuous with exponent $\gamma < 1/2$), it is differentiable almost nowhere. This is consistent with the following property of self-similarity: for $\lambda > 0$

$$W_t \stackrel{d}{=} \lambda^{-1/2} W_{\lambda t},$$

which is easily established upon verifying that both W_t and $\lambda^{-1/2} W_{\lambda t}$ are Gaussian processes with the same (zero) mean and covariance.

More about the lack of regularity of the Wiener process can be understood from *first passage times*. For given $a > 0$ define the first passage time by $T_a \equiv \inf\{t : W_t = a\}$. Now, observe that

$$(7) \quad \mathbb{P}(W_t > a) = \mathbb{P}(T_a < t \ \& \ W_t > a) = \frac{1}{2} \mathbb{P}(T_a < t).$$

The first equality is obvious by continuity, the second follows from the symmetry of the Wiener process; once the system has crossed a it is equally likely to step upwards as downwards. Introducing the random variable $M_t = \sup_{0 \leq s \leq t} W_s$, we can write this identity as:

$$(8) \quad \mathbb{P}(M_t > a) = \mathbb{P}(T_a < t) = 2\mathbb{P}(W_t > a) = 2 \int_a^\infty \frac{e^{-z^2/2t}}{\sqrt{2\pi t}} dz,$$

where we have invoked the known form of the probability density function for W_t in the last equality. Similarly, if $m_t = \inf_{0 \leq s \leq t} W_s$,

$$(9) \quad \mathbb{P}(m_t < -a) = \mathbb{P}(M_t > a).$$

But this shows that the event “ W_t crosses a ” is not so tidy as it may at first appear since it follows from (8) and (9) that for all $\varepsilon > 0$:

$$(10) \quad \mathbb{P}(M_\varepsilon > 0) > 0 \quad \text{and} \quad \mathbb{P}(m_\varepsilon < 0) > 0.$$

In particular, $t = 0$ is an accumulation point of zeros: with probability 1 the first return time to 0 (and thus, in fact, to any point, once attained) is arbitrarily small.

2. Two alternative constructions of the Wiener process

Since W_t is a Gaussian process, it is completely specified by its mean and covariance,

$$(11) \quad \mathbb{E}W_t = 0 \quad \mathbb{E}W_t W_s = \min(t, s).$$

in the sense that any process with the same statistics will also form a Wiener process. This observation can be used to make other constructions of the Wiener process. In this section, we recall two of them.

The first construction is useful in simulations. Define a set of independent Gaussian random variables $\{\eta_k\}_{k \in \mathbb{N}}$, each with mean zero and variance unity, and let $\{\phi_k(t)\}_{k \in \mathbb{N}}$ be any orthonormal basis for $L^2[0, 1]$ (that is, the space of square integral functions on the unit interval). Thus any function $f(t)$ in this set can be decomposed as $f(t) = \sum_{k \in \mathbb{N}} \alpha_k \phi_k(t)$ for an appropriate choice of the coefficients α_k . Then, the stochastic process defined by:

$$(12) \quad W_t = \sum_{k \in \mathbb{N}} \eta_k \int_0^t \phi_k(t') dt',$$

is a Wiener process in the interval $[0, 1]$. To show this, it suffices to check that it has the correct pairwise covariance – since W_t is a linear combination of zero mean Gaussian random variables, it must itself be a Gaussian random variable with zero mean. Now,

$$(13) \quad \begin{aligned} \mathbb{E}B_t B_s &= \sum_{k, l \in \mathbb{N}} \mathbb{E}\eta_k \eta_l \int_0^t \phi_k(t') dt' \int_0^s \phi_l(s') ds' \\ &= \sum_{k \in \mathbb{N}} \int_0^t \phi_k(t') dt' \int_0^s \phi_k(s') ds', \end{aligned}$$

where we have invoked the independence of the random variables $\{\eta_k\}$. Now to interpret the summands, start by defining an indicator function of the interval $[0, \tau]$ and argument t

$$\chi_\tau(t) = \begin{cases} 1 & \text{if } t \in [0, \tau] \\ 0 & \text{otherwise.} \end{cases}$$

If $\tau \in [0, 1]$, then this function further admits the series expansion

$$(14) \quad \chi_\tau(t) = \sum_k \phi_k(t) \int_0^\tau \phi_k(t') dt'.$$

Using the orthogonality properties of the $\{\phi_k(t)\}$, the equation (13) may be recast as:

$$(15) \quad \begin{aligned} \mathbb{E}B_t B_s &= \sum_{k, l \in \mathbb{N}} \int_0^1 \left(\int_0^t \phi_k(t') dt' \phi_k(u) \right) \left(\int_0^s \phi_l(s') ds' \phi_l(u) \right) du \\ &= \int_0^1 \chi_t(u) \chi_s(u) du \\ &= \int_0^1 \chi_{\min(t, s)}(u) du = \min(t, s) \end{aligned}$$

as required.

One standard choice for the set of functions $\{\phi_k(t)\}$ is the *Haar basis*. The first function in this basis is equal to 1 on the half interval $0 < t < 1/2$ and to -1 on $1/2 < t < 1$, the second function is equal to 2 on $0 < t < 1/4$ and to -2 on $1/4 < t < 1/2$ and so on. The utility of these functions is that it is very easy to construct a *Brownian bridge*: that is a Wiener process for in which the initial and final values are specified: $W_0 = W_1 = 0$. This may be defined by:

$$(16) \quad \overset{\circ}{W}_t = W_t - tW_1,$$

if using the above construction then it suffices to omit the function $\phi_1(t)$ from the basis.

The second construction of the Wiener process (or, rather, of the Brownian bridge), is empirical. It comes under the name of *Kolmogorov-Smirnov statistics*. Given a random variable X uniformly distributed in the unit interval (i.e. $\mathbb{P}(0 \leq X < x) = x$), and data $\{X_1, X_2, \dots, X_n\}$, define a sample-estimate for the distribution:

$$(17) \quad \hat{F}_n(x) \equiv \frac{1}{n}(\text{number of } X_k < x, k = 1, \dots, n) = \frac{1}{n} \sum_{k=1}^n \chi_{(-\infty, x)}(X_k),$$

equal to the relative number of data points that lie in the interval $x_k < x$. Now the Law of Large numbers tells us that for fixed x , $\hat{F}_n(x) \rightarrow x$ as $n \rightarrow \infty$. In fact, the error in our determination of the distribution of the random variable can be determined from the results, valid for fixed n :

$$\mathbb{E}\hat{F}_n(x) = x, \quad \mathbb{E}(\hat{F}_n(x) - x)^2 = x(1 - x)$$

which, when combined with the Central Limit Theorem, implies that in the limit of $n \rightarrow \infty$:

$$(18) \quad \sqrt{n}(\hat{F}_n(x) - x) \xrightarrow{d} N(0, x(1 - x)).$$

This result can be generalized to x is not fixed: as $n \rightarrow \infty$

$$(19) \quad \sqrt{n}(\hat{F}_n(x) - x) \xrightarrow{d} W_x - xW_1.$$

3. The Feynman-Kac formula

Given a function $f(x)$, define

$$(20) \quad u(x, t) = \mathbb{E}f(x + W_t)$$

This is the *Feynman-Kac formula* for the solution of the diffusion equation:

$$(21) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} \quad u(x, 0) = f(x).$$

To show this note first that:

$$\begin{aligned} u(x, t + s) &= \mathbb{E}f(x + W_{t+s}) = \mathbb{E}f(x + (W_{t+s} - W_t) + W_t) \\ &= \mathbb{E}u(x + W_{t+s} - W_t, t) \equiv \mathbb{E}u(x + W_s, t) \end{aligned}$$

where we have used the independence of $W_{t+s} - W_t$ and W_t . Now, observe that

$$\begin{aligned} \frac{\partial u}{\partial t}(x, t) &= \lim_{s \rightarrow 0^+} \frac{1}{s} (u(x, t + s) - u(x, t)) \\ &= \lim_{s \rightarrow 0^+} \frac{1}{s} \mathbb{E}(u(x + W_s, t) - u(x, t)) \\ &= \lim_{s \rightarrow 0^+} \frac{1}{s} \left(\frac{\partial u}{\partial x}(x, t) \mathbb{E}W_s + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(x, t) \mathbb{E}W_s^2 + o(s) \right), \end{aligned}$$

where we have Taylor-series expanded to obtain the final equality. The result follows by noting that $\mathbb{E}W_s = 0$ and $\mathbb{E}W_s^2 = s$.

The formula admits many generalizations. For instance: If

$$(22) \quad v(x, t) = \mathbb{E}f(x + W_t) + \int_0^t \mathbb{E}g(x + W_s) ds,$$

then the function $v(x, t)$ satisfies the diffusion equation with source-term the arbitrary function $g(x)$:

$$(23) \quad \frac{\partial v}{\partial t} = \frac{1}{2} \frac{\partial^2 v}{\partial x^2} + g(x) \quad v(x, 0) = f(x).$$

Or: If

$$(24) \quad w(x, t) = \mathbb{E} \left(f(x + W_t) \exp \left(\int_0^t c(x + W_s) ds \right) \right)$$

then $w(x, t)$ satisfies diffusive equation with an exponential growth term:

$$(25) \quad \frac{\partial w}{\partial t} = \frac{1}{2} \frac{\partial^2 w}{\partial x^2} + c(x)w \quad w(x, 0) = f(x) .$$

Addendum: The law of large numbers and the central limit theorem

Let $\{X_j\}_{j \in \mathbb{N}}$ be a sequence of i.i.d. (independent, identically distributed) random variables, let $\eta = \mathbb{E}X_1$ $\sigma^2 = \text{var}(X_1) = \mathbb{E}(Z_1 - \eta)^2$ and define

$$S_n = \sum_{j=1}^n X_j$$

The (weak) *law of large numbers* states that if $\mathbb{E}|X_j| < \infty$, then

$$\frac{S_n}{n} \rightarrow \eta \quad \text{in probability.}$$

The *central limit theorem* states that if $\mathbb{E}X_j^2 < \infty$ then

$$\frac{S_n - n\eta}{\sqrt{n\sigma^2}} \rightarrow N(0, 1) \quad \text{in distribution.}$$

We first give a proof of the law of large numbers under the stronger assumption that $\mathbb{E}|X_j|^2 < \infty$. Without loss of generality we can assume that $\eta = 0$. The proof is based the *Chebychev inequality*: Suppose X is a random variable with distribution function $F(x) = \mathbb{P}(X < x)$. Then, for any $\lambda > 0$,

$$(26) \quad \mathbb{P}(|X| \geq \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}|X|^p.$$

Indeed:

$$\lambda^p \mathbb{P}(|X| \geq \lambda) = \lambda^p \int_{|x| \geq \lambda} dF(x) \leq \int_{|x| \geq \lambda} |x|^p dF(x) \leq \int_{\mathbb{R}} |x|^p dF(x) = \mathbb{E}|X|^p.$$

Using Chebychev's inequality, we have

$$\mathbb{P} \left\{ \left| \frac{S_n}{n} \right| > \varepsilon \right\} \leq \frac{1}{\varepsilon^2} \mathbb{E} \left| \frac{S_n}{n} \right|^2$$

for any $\varepsilon > 0$. Using the i.i.d. property, this gives

$$\mathbb{E}|S_n|^2 = \mathbb{E}|X_1 + X_2 + \dots + X_n|^2 = n\mathbb{E}|X_1|^2.$$

Hence

$$\mathbb{P} \left\{ \left| \frac{S_n}{n} \right| > \varepsilon \right\} \leq \frac{1}{n\varepsilon^2} \mathbb{E}|X_1|^2 \rightarrow 0,$$

as $n \rightarrow \infty$, and this proves the law of large numbers.

Next we prove the central limit theorem. Let f be the *characteristic function* of X_1 , i.e.

$$(27) \quad f(k) \equiv \mathbb{E}e^{ikX_1}, \quad k \in \mathbb{R}.$$

and similarly let g_n be the characteristic function of $S_n/\sqrt{n\sigma^2}$. Then

$$\begin{aligned} g_n(\xi) &= \mathbb{E}e^{i\xi S_n/\sqrt{n\sigma^2}} = \prod_{j=1}^n \mathbb{E}e^{i\xi X_j/\sqrt{n\sigma^2}} = \left(\mathbb{E}e^{i\xi X_1/\sqrt{n\sigma^2}} \right)^n \\ &= \left(1 + \frac{ik}{\sqrt{n\sigma}} \mathbb{E}X_1 - \frac{k^2}{2n\sigma^2} \mathbb{E}X_1^2 + o(N^{-1}) \right)^n \\ &= \left(1 - \frac{k^2}{2n} + o(N^{-1}) \right)^n \\ &\rightarrow e^{-k^2/2} \quad \text{as } n \rightarrow \infty. \end{aligned}$$

This shows that the characteristic function of $S_n/\sqrt{n\sigma^2}$ converges to the the characteristic function of $N(0, 1)$ as $n \rightarrow \infty$ and terminates the proof.

It is instructive to note that the only property of X_1 that we have required in the central limit theorem is that $\mathbb{E}X_1^2 < \infty$. In particular, the theorem holds even if the higher moments of X_1 are infinite! For one illustration of this, consider a random variable having probability density function

$$(28) \quad \rho(x) = \frac{2}{\pi(1+x^2)^2},$$

for which no higher order moment than the variance can be computed. Nevertheless, we have:

$$\begin{aligned} f(k) &\equiv \int_{\mathbb{R}} e^{ikx} \rho(x) dx = (1 + |k|) e^{-|k|} \\ &= 1 - \frac{1}{2}k^2 + o(k^2), \end{aligned}$$

although the characteristic function is not three times differentiable at $k = 0$, so the Taylor series can not be extended beyond second order terms.

Notes by Marcus Roper and Ravi Srinivasan.

Stochastic integrals and stochastic differential equations

From (1) and (2) $W_{t_n}^N$, where $t_n = n/N$, satisfies the recurrence relation

$$(29) \quad W_{t_n}^N = W_{t_n}^N + \xi_{n+1} \sqrt{\Delta t}.$$

where $\Delta t = 1/N$ and $\{\xi_n\}_{n \in \mathbb{N}}$ are i.i.d. random variables taking values ± 1 with probability $\frac{1}{2}$ as before. A natural generalization of this relation (29) is

$$(30) \quad X_{t_{n+1}}^N = X_{t_n}^N + b(X_{t_n}^N, t_n) \Delta t + \sigma(X_{t_n}^N, t_n) \xi_{n+1} \sqrt{\Delta t}, \quad X_0 = x$$

If the last term were absent, this would be the forward Euler scheme for the ODE $\dot{X}_t = b(X_t, t)$. If $\sigma(x, t)$ meets appropriate regularity requirements, it can be shown that X_t^N converges to a stochastic process X_t as $N \rightarrow \infty$ (i.e. as $\Delta t \rightarrow 0$ with $n\Delta t \rightarrow t$). The limiting equation for X_t is denoted as the *stochastic differential equation* (SDE)

$$(31) \quad dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t, \quad X_0 = x,$$

as a reminder that the last term in (30) divided by Δt does not have a standard function as limit. The notation dW_t comes from (29) since this equation can be written as $W_{t_{n+1}}^N - W_{t_n}^N = \xi_{n+1} \sqrt{\Delta t}$. We note that the convergence of X_t^N to X_t holds provided only that the ξ_n 's are i.i.d. random variables with mean zero, $\mathbb{E}\xi_n = 0$, and variance one, $\mathbb{E}\xi_n^2 = 1$. The standard choice in numerical schemes is to take $\xi_n = N(0, 1)$, in which case

$$\sqrt{\Delta t} \xi_{n+1} \stackrel{d}{=} W_{t_{n+1}} - W_{t_n}.$$

In the discussion below, however, we will stick to the choice where $\{\xi_n\}_{n \in \mathbb{N}}$ are i.i.d. random variables taking values ± 1 with probability $\frac{1}{2}$ since it facilitates the calculations.

Next, we study the properties of X_t solution of (31) and introduce some non-standard calculus due to Itô to manipulate this solution.

1. Itô isometry and Itô formula

Consider the recurrence relation

$$X_{t_{n+1}}^N = X_{t_n}^N + f(W_{t_n}^N) \xi_{n+1} \sqrt{\Delta t},$$

Let us investigate the properties of the limit of $X_{n\Delta t}^N$ as $N \rightarrow \infty$, assuming that this limit exists. The limiting form of the recurrence relation above is traditionally denoted as

$$dX_t = f(W_t, t)dW_t, \quad X_0 = 0,$$

which can also be expressed as the *stochastic integral*

$$X_t = \int_0^t f(W_s, s) dW_s.$$

Stochastic integral have special properties called *Itô isometries*

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

The first Itô isometry is often written and used in differential form

$$\mathbb{E} f(W_s, s) dW_s = 0.$$

The Itô isometries are easy to demonstrate. The first is implied by

$$\begin{aligned} \mathbb{E} X_{t_n}^N &= \mathbb{E} \sum_{m=0}^{n-1} f(W_{t_m}^N, t_m) \xi_{m+1} \sqrt{\Delta t} \\ &= \sum_{m=0}^{n-1} \mathbb{E} f(W_{t_m}^N, t_m) \mathbb{E} \xi_{m+1} \sqrt{\Delta t} = 0, \end{aligned}$$

where we used the independence of the ξ_m 's and $\mathbb{E} \xi_m = 0$. The second is implied by

$$\begin{aligned} \mathbb{E} (X_{t_n}^N)^2 &= \mathbb{E} \sum_{m,p=0}^n f(W_{t_m}^N, t_m) f(\bar{W}_{t_p}^N, t_p) \xi_{m+1} \xi_{p+1} \Delta t \\ &= \sum_{m=0}^n \mathbb{E} f^2(W_{t_m}^N, t_m) \Delta t, \end{aligned}$$

where we use the fact that ξ_m and ξ_p are independent unless $m = p$, and $\xi_m^2 = 1$ by definition.

Going back to (31), a very important formula to manipulate the solution of this equation is *Itô formula* which states the following. Assume that X_t is the solution of (31) and let f be a smooth function. Then $f(X_t)$ satisfies the SDE

$$\begin{aligned} df(X_t) &= f'(X_t) dX_t + \frac{1}{2} f''(X_t) \sigma^2(X_t, t) dt \\ &= (f'(X_t) b(X_t, t) + \frac{1}{2} f''(X_t) \sigma^2(X_t, t)) dt + f'(X_t) \sigma(X_t, t) dW_t. \end{aligned}$$

If f depends explicitly on t , then an additional term $\partial f / \partial t dt$ is present at the right hand-side. Itô formula is the analog of the chain rule in ordinary differential calculus. However ordinary chain rule would give

$$df(X_t) = f'(X_t) dX_t.$$

Here because of the non-differentiability of X_t , we have the additional term that depends on $f''(x)$.

The proof of Itô formula can be outlined as follows. We Taylor expand $f(X_{T_{n+1}}^N) - f(X_{t_n}^N)$ using the recurrence relation (30) for $X_{t_n}^N$ and keep terms up to $O(\Delta t)$:

$$\begin{aligned} f(X_{t_{n+1}}^N) - f(X_{t_n}^N) &= f'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) + \frac{1}{2}f''(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N)^2 + \dots \\ &= f'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) \\ &\quad + \frac{1}{2}f''(X_{t_n}^N)\left(b(X_{t_n}^N, t_n)\Delta t + \sigma(X_{t_n}^N, t_n)\xi_{n+1}\sqrt{\Delta t}\right)^2 + O(\Delta t^{3/2}) \\ &= f'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) + \frac{1}{2}f''(X_{t_n}^N)\sigma^2(X_{t_n}^N, t_n)\xi_{n+1}^2\Delta t + O(\Delta t^{3/2}). \end{aligned}$$

The Itô formula follows in the limit as $\Delta t \rightarrow 0$ because

$$\frac{1}{2}f''(X_{t_n}^N)\sigma^2(X_{t_n}^N, t_n)\xi_{n+1}^2\Delta t \rightarrow \frac{1}{2}f''(X_t)\sigma^2(X_t, t)dt.$$

since $\xi_n^2 = 1$ by definition, and the higher order terms in the expansion gives no contribution in the limit as $\Delta t \rightarrow 0$.

2. Examples

The Itô isometries and the Itô formula are the backbone of the *Itô calculus* which we now use to compute some stochastic integrals and solve some SDEs. As an example of stochastic integral, consider

$$\int_0^t W_s dW_s.$$

Taking $f(x) = x^2$ in Itô formula gives

$$\frac{1}{2}dW_t^2 = W_t dW_t + \frac{1}{2}dt.$$

Therefore

$$\int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}t.$$

Notice that the second term at the right hand-side would be absent by the rules of standard calculus. Yet, this term must be present for consistency, since the expectation of the left hand-side is

$$\mathbb{E} \int_0^t W_s dW_s = 0,$$

using the first Itô isometry, and the expectation of the right hand-side is zero only with the term $\frac{1}{2}t$ included since $\frac{1}{2}\mathbb{E}W_t^2 = \frac{1}{2}t$.

As a first example of SDE, consider

$$dX_t = -\gamma X_t dt + \sigma dW_t, \quad X_0 = x$$

This is the *Ornstein-Uhlenbeck process*. Using Itô formula with $f(x, t) = e^{\gamma t}x$, we get (this is Duhammel principle)

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

Integrating gives

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

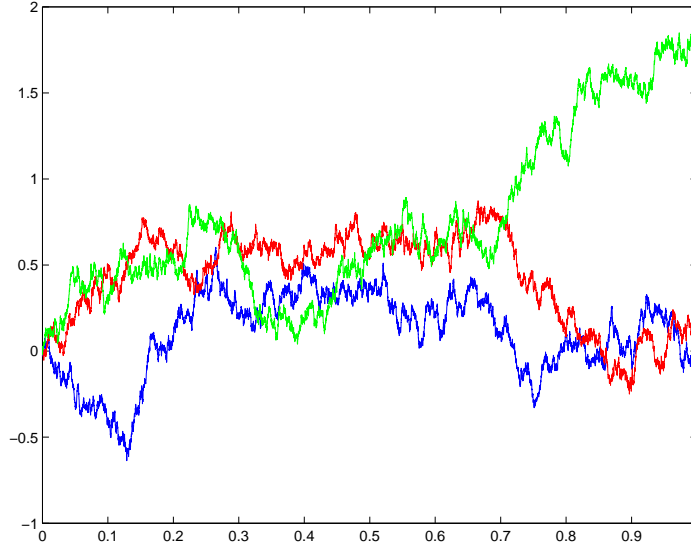


FIGURE 1. Three realizations of the Ornstein-Uhlenbeck process with $X_0 = 0$ and $\gamma = \sigma = 1$.

This process is Gaussian being a linear combination of the Gaussian process W_t . Its mean and variance are (using the Itô isometries)

$$\begin{aligned}\mathbb{E}X_t &= e^{-\gamma t}x \\ \mathbb{E}(X_t - \mathbb{E}X_t)^2 &= \sigma^2 \int_0^t (e^{-\gamma(t-s)})^2 ds = \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma t}).\end{aligned}$$

Thus when $\gamma > 0$

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

as $t \rightarrow \infty$.

As a second example of SDE, consider

$$dY_t = Y_t dt + \alpha Y_t dW_t, \quad Y_0 = y.$$

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

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

Integrating we get

$$Y_t = ye^{t - \frac{1}{2}\alpha^2 t + \alpha W_t}.$$

Note that by the rules of standard calculus, we would have obtained the wrong answer

$$Y_t = ye^{t + \alpha W_t}.$$

Indeed the term $-\frac{1}{2}\alpha^2 t$ in the exponential is important for consistency since taking the expectation of the SDE for Y_t using the first Itô isometry gives

$$d\mathbb{E}Y_t = \mathbb{E}Y_t dt,$$

and hence

$$\mathbb{E}Y_t = ye^t.$$

The solution above is consistent with this since (can you show this?)

$$\mathbb{E}e^{\alpha W_t} = e^{\frac{1}{2}\alpha^2 t}.$$

3. Generalization in multi-dimension

The definition of Itô integrals and SDE's can be extended to multi-dimension in a straightforward fashion. The SDE

$$dX_t^j = b_j(X_t, t)dt + \sum_{k=1}^K \sigma_{jk}(X_t, t)dW_t^k, \quad j = 1, \dots, J$$

where $\{W_t^k\}_{k=1}^K$ are independent Wiener processes, defines a vector-valued stochastic process $X_t = (X_t^1, \dots, X_t^J)$. The only point worth noting is the Itô formula, which in multi-dimension reads:

$$df(X_t) = \sum_{j=1}^J \frac{\partial}{\partial x_j} f(X_t) dX_t^j + \frac{1}{2} \sum_{j,j'=1}^J \frac{\partial^2 f}{\partial x_j \partial x_{j'}}(X_t) \left(\sum_{k=1}^K \sigma_{jk}(X_t, t) \sigma_{kj'}(X_t, t) \right) dt$$

4. Forward and Backward Kolmogorov equations

Consider the stochastic ODE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = y.$$

Define the transition probability density $\rho(x, t|y)$ via

$$\int_B \rho(x, t|y) dx = \mathbb{P}\{X_{t+s} \in B | X_s = y\}.$$

($\rho(x, t|y)$ does not depend on s because $b(x)$ and $\sigma(x)$ are time-independent, in which case the process X_t is stationary.) We will derive equation for ρ . Let f be an arbitrary smooth function. Using Itô formula, we have

$$f(X_t) - f(y) = \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) a(X_s) ds,$$

where $a(x) = \sigma^2(x)$. Taking expectation on both sides, we get

$$\mathbb{E}f(X_t) - f(y) = \mathbb{E} \int_0^t f'(X_s) b(X_s) ds + \frac{1}{2} \mathbb{E} \int_0^t f''(X_s) a(X_s) ds.$$

or equivalently using ρ

$$\begin{aligned} & \int_{\mathbb{R}} f(x) \rho(x, t|y) dx - f(y) \\ &= \int_0^t \int_{\mathbb{R}} f'(x) b(x) \rho(x, s|y) dx ds + \frac{1}{2} \int_0^t \int_{\mathbb{R}} f''(x) a(x) \rho(x, s|y) dx ds. \end{aligned}$$

Since this holds for all smooth f , we obtain

$$(32) \quad \frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} (b(x)\rho) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (a(x)\rho)$$

with the initial condition $\lim_{t \rightarrow 0} \rho(x, t|y) = \delta(x-y)$. This is the *forward Kolmogorov equation* for ρ in terms of the variables (x, t) . It is also called the *Fokker-Planck equation*.

Equivalently, an equation for ρ in terms of the variables (y, t) can be derived. The Markov property implies that

$$\rho(x, t + s|y) = \int_{\mathbb{R}} \rho(x, t|z)\rho(z, s|y)dz.$$

Hence

$$\begin{aligned} \rho(x, t + \Delta t|y) - \rho(x, t|y) &= \int_{\mathbb{R}} \rho(x, t|z)\rho(z, \Delta t|y)dz - \rho(x, t|y) \\ &= \int_{\mathbb{R}} \rho(x, t|z)(\rho(z, \Delta t|y) - \delta(z - y))dz. \end{aligned}$$

Dividing both side by Δt and taking the limit as $\Delta t \rightarrow 0$ using the forward Kolmogorov equation one obtains

$$\frac{\partial \rho}{\partial t} = \int_{\mathbb{R}} \rho(x, t|z) \left(-\frac{\partial}{\partial z}(b(z)\delta(z - y)) + \frac{1}{2} \frac{\partial^2}{\partial z^2}(a(z)\delta(z - y)) \right) dz,$$

which by integration by parts gives

$$(33) \quad \frac{\partial \rho}{\partial t} = b(y) \frac{\partial \rho}{\partial y} + \frac{1}{2} a(y) \frac{\partial^2 \rho}{\partial y^2}.$$

This is the *backward Kolmogorov equation* for ρ in terms of the variables (x, t) . The operator

$$L = b(y) \frac{\partial}{\partial y} + \frac{1}{2} a(y) \frac{\partial^2}{\partial y^2},$$

is called the *infinitesimal generator* of the process. The coefficient b and a can be expressed as (can you show this from the SDE?)

$$b(y) = \lim_{t \rightarrow 0} \frac{1}{t} (\mathbb{E}X_t - y), \quad a(y) = \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E}(X_t - y)^2.$$

Both the forward and the backward equations can be considered with different initial conditions. In particular, given a smooth function f , if we define

$$u(y, t) = \mathbb{E}f(X_t),$$

then $u(y, t) = \int_{\mathbb{R}} f(x)\rho(x, t|y)$ and hence it satisfies

$$\frac{\partial u}{\partial t} = b(y) \frac{\partial u}{\partial y} + \frac{1}{2} a(y) \frac{\partial^2 u}{\partial y^2},$$

with the initial condition $u(y, 0) = f(y)$. Thus, the SDE for X_t is the characteristic equation that is associated with this parabolic PDE, much in the same way as the ODE $\dot{X}_t = b(X_t)$ is the characteristic equation associated with the first order PDE $\partial u / \partial t = b(y) \partial u / \partial y$. This can be generalized in many ways. For instance, the solution of

$$\frac{\partial v}{\partial t} = c(y)v(y) + b(y) \frac{\partial v}{\partial y} + \frac{1}{2} a(y) \frac{\partial^2 v}{\partial y^2}.$$

with the initial condition $v(y, 0) = f(y)$, can be expressed as

$$v(y, t) = \mathbb{E}f(X_t) e^{\int_0^t c(X_s) ds}.$$

This is the celebrated *Feynman-Kac formula* in the context of SDEs.

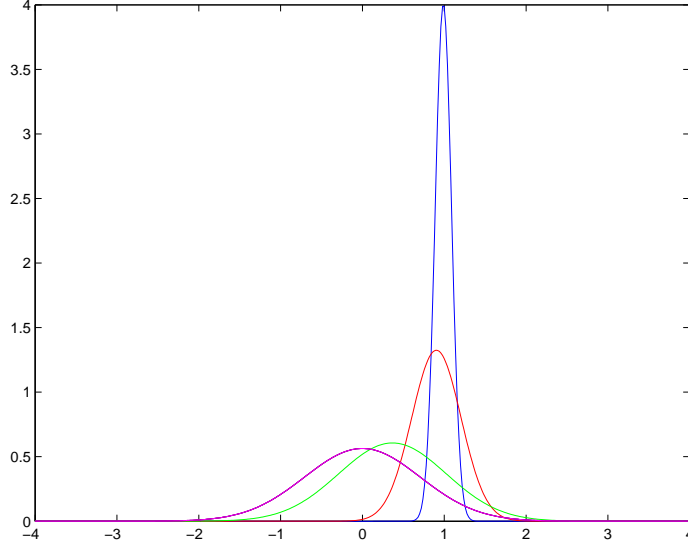


FIGURE 2. Snapshots of the density of the Ornstein-Uhlenbeck process at time $t = 0.01$ (blue), $t = 0.1$ (red), $t = 1$ (green), and $t = 10$ (magenta). Here $X_0 = y = 1$ and $\gamma = \sigma = 1$. The last snapshot at $t = 10$ is very close to the equilibrium density.

Let us consider an example. The forward differential equation associated with the Ornstein-Uhlenbeck process introduced in the last section is

$$\frac{\partial \rho}{\partial t} = \gamma \frac{\partial}{\partial x}(x\rho) + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}$$

The solution of this equation is

$$\rho(x, t|y) = \frac{1}{\sqrt{\pi\sigma^2(1 - e^{-2\gamma t})/\gamma}} \exp\left(-\frac{\gamma(x - ye^{-\gamma t})^2}{\sigma^2(1 - e^{-2\gamma t})}\right).$$

This shows that the Ornstein-Uhlenbeck process is a Gaussian process with mean $ye^{-\gamma t}$ and variance $\sigma^2(1 - e^{-2\gamma t})/2\gamma$. It also confirms that this process tends to $N(0, \sigma^2/2\gamma)$ as $t \rightarrow \infty$ since

$$\rho(x) = \lim_{t \rightarrow \infty} \rho(x, t|y) = \frac{e^{-\gamma x^2/\sigma^2}}{\sqrt{\pi\sigma^2/\gamma}}.$$

Generally, the limit of $\rho(x, t|y)$ as $t \rightarrow \infty$, when it exists, gives the equilibrium density ρ of the process. It satisfies

$$0 = -\frac{\partial}{\partial x}(b(x)\rho) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(a(x)\rho).$$

Forward and backward Kolmogorov equations can also be derived for multi-dimensional processes. They read respectively

$$\frac{\partial \rho}{\partial t} = -\sum_{j=1}^J \frac{\partial}{\partial x_j}(b_j(x)\rho) + \frac{1}{2} \sum_{j,j'=1}^J \frac{\partial^2}{\partial x_i \partial x_j}(a_{jj'}(x)\rho)$$

and

$$\frac{\partial \rho}{\partial t} = \sum_{j=1}^J b_j(x) \frac{\partial \rho}{\partial x_j} + \frac{1}{2} \sum_{j,j'=1}^J a_{jj'}(x) \frac{\partial^2 \rho}{\partial x_j \partial x_{j'}},$$

where $a_{jj'}(x) = \sum_{k=1}^K \sigma_{jk}(x) \sigma_{j'k}(x)$.

Notes by Walter Pauls and Arghir Dani Zarnescu.

Asymptotic techniques for SDEs

Here we discuss techniques by which one can study SDEs evolving on very different time-scales and derive closed equations for the slow variables.

1. The case of stiff ordinary differential equations

We start with an ODE example. Consider

$$(34) \quad \begin{cases} \dot{X}_t = -Y_t^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t) & X_0 = x \\ \dot{Y}_t = -\frac{1}{\varepsilon}(Y_t - X_t) & Y_0 = y. \end{cases}$$

If ε is very small, Y_t is very fast and one expects that it will adjust rapidly to the current value of X_t , i.e. $Y_t = X_t + O(\varepsilon)$ at all times. Then the equation for X_t reduces to

$$(35) \quad \dot{X}_t = -X_t^3.$$

The solutions of (34) and (35) are compared in figure 1.

Here is a formal derivation of the limiting equation (35) which uses the backward Kolmogorov equation. For simplicity we drop the term $\sin(\pi t) + \cos(\sqrt{2}\pi t)$. Generalizing the derivation below with this term included is easy but requires a slightly different backward equation because (35) is non-autonomous. Let f be a smooth function and consider

$$u(x, y, t) = f(X_t).$$

(This function depends on both x and y since X_t depends on both these variable because X_t and Y_t are coupled in (34), and there is no expectation since (34) is deterministic.) The backward equation is

$$\frac{\partial u}{\partial t} = L_x u + \frac{1}{\varepsilon} L_y u,$$

where

$$L_x = -y \frac{\partial}{\partial x}, \quad L_y = -(y - x) \frac{\partial}{\partial y}.$$

Look for a solution of the form $u = u_0 + \varepsilon u_1 + O(\varepsilon^2)$, so that $u \rightarrow u_0$ as $\varepsilon \rightarrow 0$. Inserting this expansion into the backward equation, and grouping terms of same order in ε , one obtains

$$(36) \quad \begin{aligned} L_y u_0 &= 0, \\ L_y u_1 &= \frac{\partial u_0}{\partial t} - L_x u_0, \end{aligned}$$

and so on. The first equation tells that u_0 belong to the null-space of L_y , i.e. $u_0 = u_0(x, t)$. The second equation requires as a solvability condition that the right hand-side belongs to the range of L_y . To see what this condition actually is,

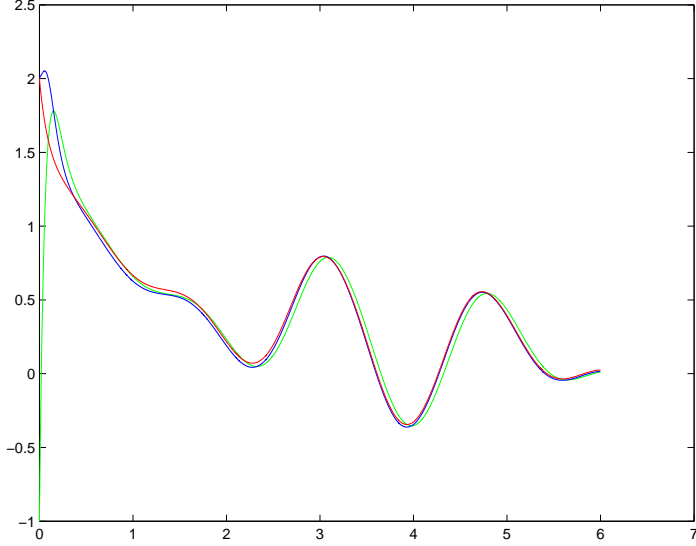


FIGURE 1. The solution of (34) when $\varepsilon = 0.05$ and we took $X_0 = 2$, $Y_0 = -1$. X_t is shown in blue, and Y_t in green. Also shown in red is the solution of the limiting equation (35).

multiply the second equation in (36) by a test function $\rho(y)$, and integrate both sides over \mathbb{R} . After integration by part at the left hand-side, this gives

$$\int_{\mathbb{R}} L_y^* \rho(y) u_1 dy = \int_{\mathbb{R}} \rho(y) \left(\frac{\partial u_0}{\partial t} - L_x u_0 \right) dy.$$

where L_y^* is the adjoint of L_y viewed as an operator in y at fixed x , i.e.

$$L_y^* \rho(y) = \frac{\partial}{\partial y} ((y-x)\rho(y)).$$

Choosing $\rho(y)$ such that

$$(37) \quad 0 = L_y^* \rho(y),$$

one concludes that the solvability of (36) requires that

$$(38) \quad 0 = \int_{\mathbb{R}} \rho(y) \left(\frac{\partial u_0}{\partial t} - L_x u_0 \right) dy.$$

It can be shown that this equation is also sufficient for the solvability of (36) – the calculation above actually tells the range of L_y is the space perpendicular to the null-space of the adjoint of L_y . Now, (37) is simply the forward Kolmogorov equation for the equilibrium density of the process Y_t at fixed $X_t = x$. Here the equilibrium density is a generalized function

$$\rho(y|x) = \delta(y-x).$$

Using this $\rho(y|x)$, the solvability condition (38) becomes

$$0 = \frac{\partial u_0}{\partial t} + x \frac{\partial u_0}{\partial x},$$

which is the backward equation for

$$\dot{X}_t = -X_t^3, \quad X_0 = x.$$

A similar argument with the term $\sin(\pi t) + \cos(\sqrt{2}\pi t)$ included gives the backward equation for (35).

2. Generalization to stochastic differential equation

The derivation that lead to (35) can be generalized to SDEs. Consider

$$(39) \quad \begin{cases} dX_t = f(X_t, Y_t)dt, & X_0 = x \\ dY_t = \frac{1}{\varepsilon}b(X_t, Y_t)dt + \frac{1}{\sqrt{\varepsilon}}\sigma(X_t, Y_t)dt, & Y_0 = y, \end{cases}$$

and assume that the equation for Y_t at $X_t = x$ fixed has an equilibrium density $\rho(y|x)$ for every x . Then going through a derivation as above with

$$u(x, y, t) = \mathbb{E}f(X_t),$$

one concludes that the backward equation associated with this SDE also reduces to (38) as $\varepsilon \rightarrow 0$, i.e.

$$\frac{\partial u_0}{\partial t} = F(x) \frac{\partial u_0}{\partial x},$$

where

$$F(x) = \int_{\mathbb{R}} f(x, y) \rho(y|x) dy.$$

Thus the limiting equation for X_t is

$$\dot{X}_t = F(X_t), \quad X_0 = 0.$$

The main difference with the deterministic example treated before is that the fast process Y_t does not rapidly settle to an equilibrium point depending on the current value of X_t – only its density does.

Here is an example generalizing (34). Consider

$$(40) \quad \begin{cases} dX_t = -Y_t^3 dt + \sin(\pi t) + \cos(\sqrt{2}\pi t), & X_0 = x \\ dY_t = -\frac{1}{\varepsilon}(Y_t - X_t)dt + \frac{\alpha}{\sqrt{\varepsilon}}dW_t, & Y_0 = y. \end{cases}$$

The equation for Y_t at fixed $X_t = x$ defines an Ornstein-Uhlenbeck process whose equilibrium density is

$$\rho(y|x) = \frac{e^{-(y-x)^2/\alpha^2}}{\sqrt{\pi}\alpha}.$$

Therefore

$$F(x) = - \int_{\mathbb{R}} y^3 \frac{e^{-(y-x)^2/\alpha^2}}{\sqrt{\pi}\alpha} dy = -x^3 - \frac{3}{2}\alpha^2 x,$$

and the limiting equation is

$$(41) \quad \dot{X}_t = -X_t^3 - \frac{3}{2}\alpha^2 X_t + \sin(\pi t) + \cos(\sqrt{2}\pi t), \quad X_0 = x.$$

Note the new term $-\frac{3}{2}\alpha^2 X_t$, due to the noise in (40). The solution of (40) and (41) are shown in figure 2.

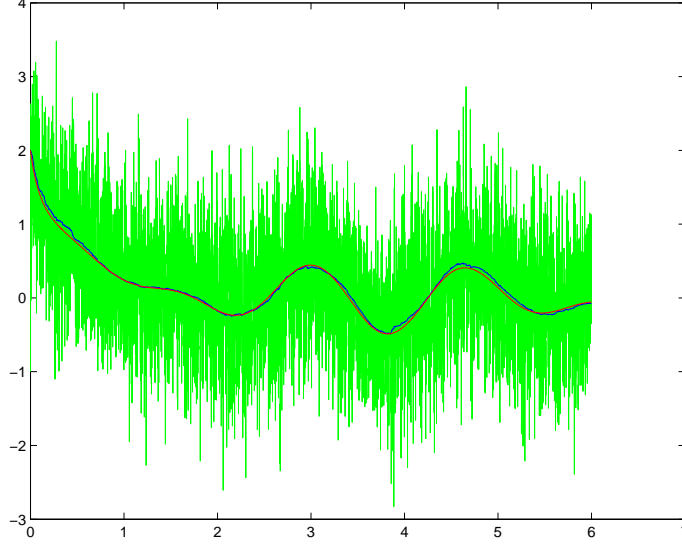


FIGURE 2. The solution of (40) with $X_0 = 2$, $Y_0 = -1$ when $\varepsilon = 10^{-3}$ and $\alpha = 1$. X_t is shown in blue, and Y_t in green. Also shown in red is the solution of the limiting equation (41). Notice how noisy Y_t is.

3. Strong convergence and the property of self-averaging

The derivation in section 2 only give weak convergence, or convergence in distribution. But stronger results can be obtained. Consider a system of the form

$$(42) \quad \dot{X}_t^\varepsilon = f(X_t^\varepsilon, Y_{t/\varepsilon}),$$

where Y_t is a given stochastic process. Assume that Y_t is ergodic, in the sense that for any fixed x ,

$$(43) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x, Y_s) ds = \bar{f}(x).$$

Then we can show that, as $\varepsilon \rightarrow 0$, X_t^ε converges strongly to the solution of

$$(44) \quad \dot{\bar{X}}_t = \bar{f}(\bar{X}_t)$$

To see this, consider the integral form of (42):

$$(45) \quad X_{t+\Delta t}^\varepsilon - X_t^\varepsilon = \int_t^{t+\Delta t} f(X_s, Y_{s/\varepsilon}) ds.$$

We rewrite this equation in a way that allows us to exploit the self-averaging property (43).

$$(46) \quad X_{t+\Delta t}^\varepsilon - X_t^\varepsilon = \int_t^{t+\Delta t} f(X_t, Y_{s/\varepsilon}) ds + \int_t^{t+\Delta t} (f(X_s, Y_{s/\varepsilon}) - f(X_t, Y_{s/\varepsilon})) ds.$$

We will consider the behavior of these two integrals as $\varepsilon \rightarrow 0$ separately.

Using (43), the first integral

$$(47) \quad \int_t^{t+\Delta t} f(X_t, Y_{s/\varepsilon}) ds = \varepsilon \int_{t/\varepsilon}^{(t+\Delta t)/\varepsilon} f(X_t, Y_s) ds \rightarrow \Delta t \bar{f}(X_t),$$

as $\varepsilon \rightarrow 0$. To investigate the contribution of the second integral, let

$$(48) \quad A(t, \Delta t, \varepsilon) = \int_t^{t+\Delta t} (f(X_s, Y_{s/\varepsilon}) - f(X_t, Y_{s/\varepsilon})) ds.$$

We then have

$$(49) \quad |A(t, \Delta t, \varepsilon)| \leq \int_t^{t+\Delta t} |f(X_s, Y_{s/\varepsilon}) - f(X_t, Y_{s/\varepsilon})| ds.$$

Assuming f is uniformly Lipschitz in Y_t with constant K , we then write

$$\begin{aligned} |A(t, \Delta t, \varepsilon)| &\leq \int_t^{t+\Delta t} K |X_s - X_t| ds \\ &\leq \int_t^{t+\Delta t} K \left| X_s - X_t - \int_t^s f(X_t, Y_{s'/\varepsilon}) ds' \right| \\ &\quad + \int_t^{t+\Delta t} K \left| \int_t^s f(X_t, Y_{s'/\varepsilon}) ds' \right| ds \end{aligned}$$

It is straightforward to show using (47) that, for sufficiently small ε ,

$$(50) \quad \int_t^{t+\Delta t} K \left| \int_t^s f(X_t, Y_{s'/\varepsilon}) ds' \right| ds < C \Delta t^2$$

for some constant $C < \infty$. Gronwall's lemma then implies that

$$(51) \quad \left| X_{t+\Delta t} - X_t - \int_t^{t+\Delta t} f(X_t, Y_{s/\varepsilon}) ds \right| = |A(t, \Delta t, \varepsilon)| \leq C \Delta t^2 \exp K \Delta t = o(\Delta t).$$

This shown that

$$(52) \quad \lim_{\varepsilon \rightarrow 0} (X_{t+\Delta t}^\varepsilon - X_t^\varepsilon) = \Delta t \bar{f}(X_t^\varepsilon) + o(\Delta t).$$

which is sufficient to demonstrate that X_t^ε converges strongly to \bar{X}_t .

4. Diffusive time-scale

An interesting generalization of the situation presented in section 2 arises when

$$(53) \quad \int_{\mathbb{R}} f(x, y) \rho(y|x) dy = 0.$$

In this case the limiting equation reduces to the trivial ODE, $\dot{X}_t = 0$, i.e. no evolution at all. In fact, the interesting evolution then occurs on a longer time-scale of order ε^{-1} , and the right scaling to study (39) is

$$(54) \quad \begin{cases} dX_t = \frac{1}{\varepsilon} f(X_t, Y_t) dt, & X_0 = x \\ dY_t = \frac{1}{\varepsilon^2} b(X_t, Y_t) dt + \frac{1}{\varepsilon} \sigma(X_t, Y_t) dt, & Y_0 = y, \end{cases}$$

To obtain the limiting equation for X_t as $\varepsilon \rightarrow 0$, we proceed as above and consider the backward equation for $u(x, y, t) = \mathbb{E}f(X_t)$, which is now rescaled as

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon}L_x u + \frac{1}{\varepsilon^2}L_y u.$$

Inserting the expansion $u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + O(\varepsilon^2)$ (we will have to go one order in ε higher than before) in this equation now gives

$$(55) \quad \begin{aligned} L_y u_0 &= 0, \\ L_y u_1 &= -L_x u_0, \\ L_y u_2 &= \frac{\partial u_0}{\partial t} - L_x u_1, \end{aligned}$$

and so on. The first equation tells that $u_0(x, y, t) = u_0(x, t)$. The solvability condition for the second equation is satisfied by assumption because of (53) and therefore this equation can be formally solved as

$$u_1 = -L_y^{-1}L_x u_0.$$

Inserting this expression in the third equation in (55) and considering the solvability condition for this equation, we obtain the limiting equation for u_0 :

$$\frac{\partial u_0}{\partial t} = \bar{L}_x u_0,$$

where

$$\bar{L}_x = \int_{\mathbb{R}} dy \rho(y|x) L_x L_y^{-1} L_x.$$

To see what this equation is explicitly, notice that $-L_y^{-1}g(y)$ is the steady state solution of

$$\frac{\partial v}{\partial t} = L_y v + g(y).$$

The solution of this equation with the initial condition $v(y, 0) = 0$ can be represented by Feynman-Kac formula as

$$v(y, t) = \mathbb{E} \int_0^t g(Y_s^x) ds,$$

where Y_t^x denotes the solution of the second SDE in (54) at $X_t = x$ fixed and $\varepsilon = 1$, i.e.

$$dY_t^x = b(x, Y_t^x)dt + \sigma(x, Y_t^x)dW_t, \quad Y_0^x = y.$$

Therefore

$$-L_y^{-1}g(y) = \mathbb{E} \int_0^\infty g(Y_t^x) dt,$$

and the limiting backward equation above can be written as

$$\frac{\partial u_0}{\partial t} = \mathbb{E} \int_0^\infty dt \int_{\mathbb{R}} dy \rho(y|x) f(x, y) \frac{\partial}{\partial x} \left(f(x, Y_t^x) \frac{\partial u_0}{\partial x} \right),$$

This is the backward equation of the SDE

$$dX_t = \bar{b}(X_t)dt + \bar{\sigma}(X_t)dW_t, \quad X_0 = x,$$

where

$$\begin{aligned}\bar{b}(x) &= \mathbb{E} \int_0^\infty \int_{\mathbb{R}} \rho(y|x) f(x, y) \frac{\partial}{\partial x} f(x, Y_t^x) dy dt, \\ \bar{\sigma}^2(x) &= 2\mathbb{E} \int_0^\infty \int_{\mathbb{R}} \rho(y|x) f(x, y) f(x, Y_t^x) dy dt.\end{aligned}$$

The interesting new phenomena is that the limiting equation for X_t has become an SDE. This means that fluctuations are important on the long-time scale and give rise to stochastic effects in the evolution of X_t that were absent on the shorter time-scale.

The calculation above is easy to generalize if there is a slow term in the original equation for X_t , i.e. if instead of (54) one considers

$$\begin{cases} dX_t = g(X_t, Y_t)dt + \frac{1}{\varepsilon} f(X_t, Y_t)dt, & X_0 = x \\ dY_t = \frac{1}{\varepsilon^2} b(X_t, Y_t)dt + \frac{1}{\varepsilon} \sigma(X_t, Y_t)dt, & Y_0 = y, \end{cases}$$

The limiting equation for X_t is then

$$dX_t = G(X_t)dt + \bar{b}(X_t)dt + \bar{\sigma}(X_t)dW_t, \quad X_0 = x,$$

with $\bar{b}(x)$ and $\bar{\sigma}(x)$ as above, and

$$G(x) = \int_{\mathbb{R}} \rho(y|x) g(x, y) dy.$$

It is also straightforward to generalize to higher dimensions.

Here is an example.

$$\begin{cases} dX_t = \frac{2\alpha}{\varepsilon} Y_t Z_t dt - (X_t + X_t^3)dt, \\ dY_t = \frac{3\alpha}{\varepsilon} Z_t X_t dt - \frac{1}{\varepsilon^2} Y_t dt + \frac{1}{\varepsilon} dW_t^y, \\ dZ_t = -\frac{\alpha}{\varepsilon} b_3 Y_t X_t dt - \frac{1}{\varepsilon^2} Z_t dt + \frac{1}{\varepsilon} dW_t^z. \end{cases}$$

where W_t^y, W_t^z are independent Wiener processes and α is a parameter. There are two fast variables, Y_t and Z_t , in this example. There is also a slow term, $-(X_t + X_t^3)dt$, in the equation for X_t which, in the absence of coupling with Y_t and Z_t , would drive X_t to the position $x = 0$. We ask to what extent this equilibrium of the uncoupled dynamics is relevant with coupling with Y_t and Z_t .

The limiting equation for X_t is

$$dX_t = ((\alpha^2 - 1)X_t - X_t^3)dt + \alpha dW_t.$$

The equilibrium density for this equation is

$$\rho(x) = Z^{-1} e^{\frac{1}{2}(\alpha^2 - 1)x^2 - \frac{1}{4}x^4}.$$

This density is shown in figure 3. For $|\alpha| \leq 1$, $\rho(x)$ is mono-modal and centered around $x = 0$, the stable equilibrium of the uncoupled dynamics. However, for $|\alpha| > 1$, $\rho(x)$ becomes bi-modal, with two maxima at $x = \pm\sqrt{\alpha^2 - 1}$ and a minimum at $x = 0$. Thus coupling with the fast modes may destroy the structures apparent in the uncoupled dynamics and induce bifurcations.

Notes by Inga Koszalka and Alex Hasha.

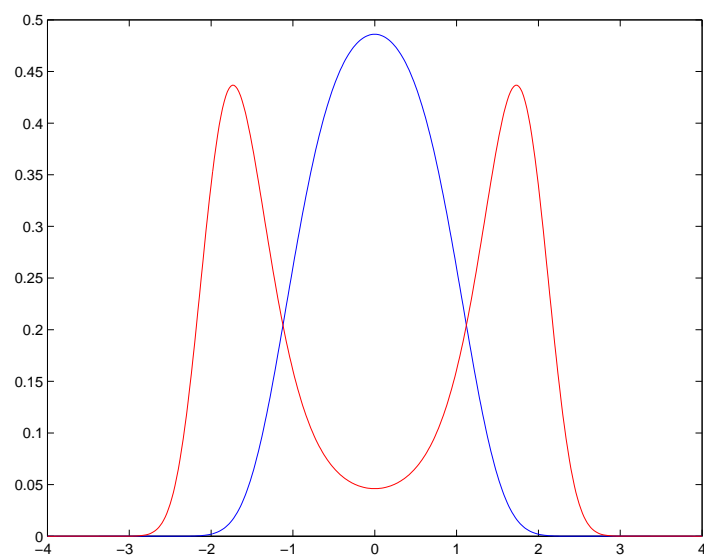


FIGURE 3. The equilibrium density $\rho(x) = Z^{-1}e^{\frac{1}{2}(\alpha^2-1)x^2 - \frac{1}{4}x^4}$ for $\alpha = \frac{1}{2}$ (blue) and $\alpha = 2$ (red).