

# Lecture 6: Wiener Process

Eric Vanden-Eijnden

Chapters 6, 7 and 8 offer a (very) brief introduction to stochastic analysis. These lectures are based in part on a book project with Weinan E. 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  $\mathbb{Z}$ :

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

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

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

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*<sup>1</sup>, with the following properties:

- (a) *Independence.*  $W_t - W_s$  is independent of  $\{W_\tau\}_{\tau \leq s}$  for any  $0 \leq s \leq t$ .

---

<sup>1</sup>The 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 very good. Some of us with complicated names are moved by this story.

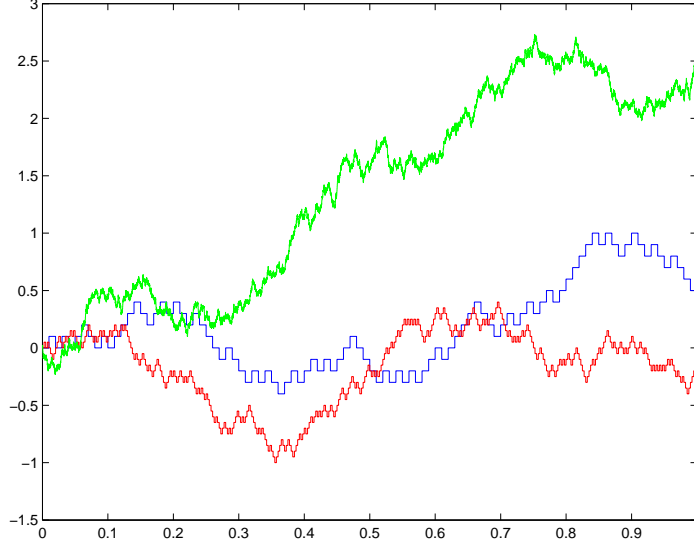


Figure 1: Realizations of  $W_t^N$  for  $N = 100$  (blue),  $N = 400$  (red), and  $N = 10000$  (green).

- (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

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

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

$$W_t - W_s \stackrel{d}{=} W_{t-s}, \tag{3}$$

where  $\stackrel{d}{=}$  means that the random processes 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,

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

where

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

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),(5) and the independence property (a), 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

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

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

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

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:

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

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

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

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  $\varepsilon > 0$ :

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

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,

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

in the sense that any process with the same statistics is also 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  $f(t) = \sum_{k \in \mathbb{N}} \alpha_k \phi_k(t)$  where (assuming that the  $\phi_k$ 's are real)  $\alpha_k = \int_0^1 f(t) \phi_k(t) dt$ . Then, the stochastic process defined by:

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

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,

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

where we have invoked the independence of the random variables  $\{\eta_k\}$ . 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

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

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

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

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 on  $[0, 1]$  for which the initial and final values are specified:  $W_0 = W_1 = 0$ . This may be defined by:

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

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 probability distribution of  $X$ :

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

equal to the relative number of data points that lie in the interval  $x_k < x$ . For fixed  $x$   $F_n(x) \rightarrow x$  as  $n \rightarrow \infty$  by the Law of Large Numbers tells us that, whereas

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

by the Central Limit Theorem. This result can be generalized to the function  $\hat{F}_n : [0, 1] \mapsto [0, 1]$  (i.e. when  $x$  is not fixed): as  $n \rightarrow \infty$

$$\sqrt{n}(F_n(x) - x) \xrightarrow{d} W_x - xW_1 = \hat{W}_x. \quad (19)$$

### 3 The Feynman-Kac formula

Given a function  $f(x)$ , define

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

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

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

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

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

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

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

Or: If

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

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

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

## 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 *Chebyshev inequality*: Suppose  $X$  is a random variable with distribution function  $F(x) = \mathbb{P}(X < x)$ . Then, for any  $\lambda > 0$ ,

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

provided only that  $\mathbb{E}|X|^p < \infty$ . 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.

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

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

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

for which all moments of order higher than 2 are infinite. Nevertheless, we have:

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

and hence the Central Limit Theorem applies. Intuitively, the reason is that the fat tails of the density  $\rho(x)$  disappear in the limit owing to the rescaling of the partial sum by  $1/\sqrt{nrtn}$ .

*Notes by Marcus Roper and Ravi Srinivasan.*