

Brownian Motion

Brownian motion is the central object of continuous-time stochastic calculus. It arises naturally as the limit of a symmetric random walk when the step size and time increment are both shrunk to zero in a coordinated way, and it serves as the building block for stochastic differential equations and derivative pricing models. These notes follow Shreve (2004) closely.

These notes construct Brownian motion from first principles. We begin by revisiting the symmetric random walk and introducing a scaled version whose variance over any interval $[s, t]$ equals $t - s$. A characteristic function argument shows that, as the number of steps grows, the terminal value of this scaled walk converges in distribution to a normal random variable. Taking this limit for all times simultaneously yields a continuous-time process—Brownian motion—whose formal definition we then state and explore.

A key theme is that Brownian paths are far rougher than the smooth functions encountered in ordinary calculus. We make this precise by studying total variation and quadratic variation. For smooth functions both are well-behaved: total variation is finite and quadratic variation is zero. Brownian motion has the opposite character: its quadratic variation over $[0, t]$ equals t almost surely, its total variation is almost surely infinite, and its paths are nowhere differentiable. These properties are not pathologies—they are exactly what makes Brownian motion the right model for cumulative noise, and they are the source of the corrections that appear in Itô's lemma.

The Simple Random Walk Again

In this section, we revisit the symmetric random walk discussed in previous lectures. This process models the outcome of repeatedly flipping a fair coin: you gain \$1 for heads and lose \$1 for tails. Each flip is independent and has expected value zero, so the cumulative gains form a martingale.

Let $\{S_n\}$ denote the position after n steps, defined recursively by

$$S_{n+1} = S_n + X_{n+1},$$

where each X_{n+1} is an independent random variable taking values $+1$ or -1 with probability $1/2$ each. Thus, the sequence $\{X_n\}$ is iid with

$$E(X_n) = 0.5 \times 1 + 0.5 \times (-1) = 0,$$

and

$$V(X_n) = 0.5 \times (1 - 0)^2 + 0.5 \times (-1 - 0)^2 = 1.$$

For any $0 \leq m < n$, we have

$$S_n = S_m + \sum_{i=m+1}^n X_i,$$

so the increment $S_n - S_m$ depends only on the coin flips between times $m + 1$ and n . Moreover, increments over disjoint intervals are independent.

The process $\{S_n\}$ is a martingale: for $m < n$,

$$E(S_n \mid \mathcal{F}_m) = S_m + E\left(\sum_{i=m+1}^n X_i \mid \mathcal{F}_m\right) = S_m,$$

since the future coin flips are independent of the past.

The variance of the increment is

$$V(S_n - S_m) = V\left(\sum_{i=m+1}^n X_i\right) = \sum_{i=m+1}^n V(X_i) = n - m.$$

Finally, the quadratic variation of the simple symmetric random walk up to time n is

$$[S, S]_n = \sum_{i=1}^n (S_i - S_{i-1})^2 = \sum_{i=1}^n X_i^2 = n,$$

since each $X_i^2 = 1$. Unlike variance, quadratic variation is computed path-by-path, not as an average over many realizations.

A Scaled Random Walk

We now embed the symmetric random walk into a finite time interval and scale it so that the variance over any interval $[0, t]$ equals t .

Let $\Delta t = T/n$, and define discrete time points: $t_0 = 0, t_1 = \Delta t, t_2 = 2\Delta t, \dots$, so that $t_n = T$. At each step, instead of moving by $+1$ or -1 , we move by $+\sqrt{\Delta t}$ or $-\sqrt{\Delta t}$. Thus, we define the scaled random walk as

$$B_{t_m}^{(n)} = \sum_{j=1}^m \sqrt{\Delta t} X_j,$$

with $B_{t_0}^{(n)} = 0$, and each X_i is independent and takes values $+1$ or -1 with probability $1/2$.

The expected value and variance are:

$$\begin{aligned} \mathbb{E}(B_{t_m}^{(n)}) &= \sum_{j=1}^m \sqrt{\Delta t} \mathbb{E}(X_j) = 0, \\ \mathbb{V}(B_{t_m}^{(n)}) &= \sum_{j=1}^m \Delta t \mathbb{V}(X_j) = m\Delta t = t_m. \end{aligned}$$

The quadratic variation up to time $0 \leq t_m \leq T$ is

$$[B^{(n)}, B^{(n)}]_{t_m} = \sum_{j=0}^{m-1} (B_{t_{j+1}}^{(n)} - B_{t_j}^{(n)})^2 = m\Delta t = t_m.$$

This scaling ensures that as n increases, the process has variance proportional to elapsed time, matching the behavior of a Brownian motion.

Limiting Distribution of the Scaled Random Walk

Remember that $T = t_n = n\Delta t$. To analyze the limiting distribution of $B_T^{(n)}$ as $n \rightarrow \infty$ (i.e., as $\Delta t \rightarrow 0$), we use the characteristic function, which uniquely determines the distribution of a random variable. The characteristic function of X is

$$\phi_X(u) = \mathbb{E}[e^{iuX}],$$

where i is the imaginary unit ($i^2 = -1$). For a normal random variable $X \sim \mathcal{N}(\mu, \sigma^2)$, the characteristic function is

$$\phi_X(u) = e^{iu\mu - \frac{1}{2}u^2\sigma^2}.$$

Now, consider the scaled random walk $B_T^{(n)} = \sum_{k=1}^n \sqrt{\Delta t} X_k$, where each X_k is independent and takes values ± 1 with probability $1/2$. Its characteristic function is

$$\begin{aligned} \mathbb{E}(e^{iuB_T}) &= \mathbb{E}\left(e^{iu\sum_{j=1}^n \sqrt{\Delta t} X_j}\right) = \prod_{j=1}^n \mathbb{E}\left(e^{iu\sqrt{\Delta t} X_j}\right) = \left(\frac{e^{iu\sqrt{\Delta t}} + e^{-iu\sqrt{\Delta t}}}{2}\right)^n \\ &\approx \left(\frac{1 + iu\sqrt{\Delta t} - \frac{1}{2}u^2\Delta t + 1 - iu\sqrt{\Delta t} - \frac{1}{2}u^2\Delta t}{2}\right)^n \\ &= \left(1 + \frac{-\frac{1}{2}u^2T}{n}\right)^n. \end{aligned}$$

Therefore,

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{iuB_T}] = \lim_{n \rightarrow \infty} \left(1 + \frac{-\frac{1}{2}u^2T}{n}\right)^n = e^{-\frac{1}{2}u^2T},$$

which is the characteristic function of a normal distribution with mean 0 and variance T . Thus, as $n \rightarrow \infty$, the scaled random walk $B_T^{(n)}$ converges in distribution to $\mathcal{N}(0, T)$.

Brownian Motion

We obtain Brownian motion as the limit of the scaled random walks as $n \rightarrow \infty$. The Brownian motion inherits properties from these random walks. This leads to the following definition.

Brownian Motion

A continuous-time stochastic process $\{B_t : t \geq 0\}$ is called a *Brownian motion* if it has the following four properties:

- i. $B_0 = 0$.
- ii. For any finite set of times $0 \leq t_0 < t_1 < t_2 < \dots < t_n$, the random variables $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}}$ are independent.
- iii. For any $0 \leq s \leq t$ the increment $B_t - B_s \sim \mathcal{N}(0, t - s)$.
- iv. For all ω in a set of probability one, $B_t(\omega)$ is a continuous function of t .

For any $t \geq 0$ we have $E(B_t) = E(B_t - B_0) = 0$, and for $0 \leq s \leq t$ we have

$$\begin{aligned} \text{Cov}(B_s, B_t) &= \text{Cov}(B_s, B_t - B_s + B_s) \\ &= \text{Cov}(B_s, B_t - B_s) + \text{Cov}(B_s, B_s) \\ &= 0 + s = s. \end{aligned}$$

Thus, for any $s, t \geq 0$ it must be the case that

$$\text{Cov}(B_s, B_t) = E(B_s B_t) = \min(s, t).$$

This covariance structure is enough to characterize the Brownian motion as well. It is in fact one way to rigorously construct the Brownian motion process by noting that for $0 \leq s \leq 1$ and $0 \leq t \leq 1$ we have

$$E(B_s B_t) = \int_0^1 \mathbf{1}_{\{[0,s]\}}(u) \mathbf{1}_{\{[0,t]\}}(u) du = \min(s, t).$$

The previous relationship establishes an isometry between the Hilbert space of normal random

variables $\mathcal{H} \in \mathcal{L}^2(\mathbb{P})$ with inner product $\langle X, Y \rangle = E[XY]$, and the Hilbert space $\mathcal{L}^2[0, 1]$ with inner product

$$\langle f, g \rangle = \int_0^1 f(u)g(u)du.$$

We can then show that for any $0 \leq t \leq 1$ we have

$$B_t = \sum_{k=0}^{\infty} \langle \phi_k, \mathbf{1}_{[0,t]} \rangle Z_k,$$

where $\{Z_k : 0 \leq k < \infty\}$ is a sequence of independent $\mathcal{N}(0, 1)$ random variables, and $\phi_k : 0 \leq k < \infty$ is an orthonormal basis of $\mathcal{L}^2[0, 1]$. By using the Haar basis in the previous expression, it is possible to show that the series representation on the right generates a Brownian motion.

Total and Quadratic Variation

Total Variation of a Function

To study the variation of a function over an interval, we begin by dividing it into smaller pieces. Consider a partition of the time interval $[0, T]$ given by

$$\Pi = \{t_0, t_1, \dots, t_n\},$$

where the partition points satisfy

$$0 = t_0 < t_1 < \dots < t_n = T.$$

The *mesh* (or *norm*) of the partition Π is defined as the length of the longest subinterval:

$$\|\Pi\| = \max_{j=0,1,\dots,n-1} (t_{j+1} - t_j).$$

As we refine the partition by adding more points, the mesh $\|\Pi\|$ decreases and approaches zero.

The *total variation* of a function f over the interval $[0, T]$ measures the total “distance traveled” by the function. It is defined as

$$V_T(f) = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)|.$$

In words, we partition the interval into smaller pieces, sum the absolute changes in f across each piece, and then take the limit as the partition becomes arbitrarily fine.

For *differentiable* functions, the total variation has a simple integral representation. By the mean-value theorem, for each subinterval $[t_j, t_{j+1}]$ there exists a point $t_j^* \in [t_j, t_{j+1}]$ where the derivative equals the average rate of change:

$$\frac{f(t_{j+1}) - f(t_j)}{t_{j+1} - t_j} = f'(t_j^*).$$

Rearranging and taking absolute values gives

$$\sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)| = \sum_{j=0}^{n-1} |f'(t_j^*)|(t_{j+1} - t_j).$$

The right-hand side is a Riemann sum for the integral of $|f'(t)|$. Taking the limit as the mesh goes to zero yields

$$V_T(f) = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |f'(t_j^*)|(t_{j+1} - t_j) = \int_0^T |f'(t)| dt.$$

Thus, for differentiable functions, total variation equals the integral of the absolute value of the derivative—a quantity that is always finite when f' is integrable.

Quadratic Variation of a Function

While total variation measures the absolute distance traveled by a function, *quadratic variation* measures the sum of squared changes. For a function f over the interval $[0, T]$, the quadratic

variation is defined as

$$[f, f]_T = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} (f(t_{j+1}) - f(t_j))^2.$$

For continuously differentiable functions, the quadratic variation vanishes. To see why, we again apply the mean-value theorem: for each subinterval there exists a point $t_j^* \in [t_j, t_{j+1}]$ such that

$$f(t_{j+1}) - f(t_j) = f'(t_j^*)(t_{j+1} - t_j).$$

Squaring both sides and summing over all subintervals gives

$$\sum_{j=0}^{n-1} (f(t_{j+1}) - f(t_j))^2 = \sum_{j=0}^{n-1} (f'(t_j^*))^2 (t_{j+1} - t_j)^2.$$

Since $(t_{j+1} - t_j) \leq \|\Pi\|$ for all j , we can factor out one power of the mesh:

$$\sum_{j=0}^{n-1} (f(t_{j+1}) - f(t_j))^2 \leq \|\Pi\| \sum_{j=0}^{n-1} (f'(t_j^*))^2 (t_{j+1} - t_j).$$

The sum on the right is a Riemann sum for $\int_0^T (f'(t))^2 dt$, which converges to this integral as $\|\Pi\| \rightarrow 0$. Therefore,

$$[f, f]_T \leq \lim_{\|\Pi\| \rightarrow 0} \|\Pi\| \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} (f'(t_j^*))^2 (t_{j+1} - t_j) = \lim_{\|\Pi\| \rightarrow 0} \|\Pi\| \int_0^T (f'(t))^2 dt = 0.$$

In other words, for smooth functions the quadratic variation is always zero because the mesh shrinks faster than the Riemann sum can accumulate.

Quadratic Variation of Brownian Motion

Unlike smooth functions whose quadratic variation is zero, Brownian motion has non-trivial quadratic variation that accumulates linearly over time. We now prove that the quadratic varia-

tion of Brownian motion over $[0, T]$ converges to T in mean square.

Analysis

Consider a partition Π of $[0, T]$ and define the quadratic variation sum:

$$Q_{\Pi} = \sum_{j=0}^{n-1} (B_{t_{j+1}} - B_{t_j})^2.$$

Our strategy is to show two things: first, that the expected value of Q_{Π} equals T for any partition; second, that the variance of Q_{Π} vanishes as the mesh size goes to zero. Together, these facts imply convergence in $L^2(\mathbb{P})$.

Step 1: Expected value of Q_{Π} .

Since each increment $B_{t_{j+1}} - B_{t_j}$ is normally distributed with mean zero and variance $t_{j+1} - t_j$, we have

$$\mathbb{E}(Q_{\Pi}) = \sum_{j=0}^{n-1} \mathbb{E}(B_{t_{j+1}} - B_{t_j})^2 = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = t_n - t_0 = T.$$

Thus, regardless of how we partition the interval, the expected quadratic variation is always T .

Step 2: Variance of Q_{Π} .

To show that Q_{Π} concentrates around its mean as $\|\Pi\| \rightarrow 0$, we compute its variance. For a standard normal random variable $Z \sim \mathcal{N}(0, 1)$, we have $\mathbb{E}(Z^4) = 3$. Since $(B_{t_{j+1}} - B_{t_j})/\sqrt{t_{j+1} - t_j} \sim \mathcal{N}(0, 1)$, it follows that

$$\begin{aligned} \mathbb{V}(B_{t_{j+1}} - B_{t_j})^2 &= \mathbb{E}(B_{t_{j+1}} - B_{t_j})^4 - \left(\mathbb{E}(B_{t_{j+1}} - B_{t_j})^2\right)^2 \\ &= 3(t_{j+1} - t_j)^2 - (t_{j+1} - t_j)^2 = 2(t_{j+1} - t_j)^2. \end{aligned}$$

Because Brownian increments over disjoint intervals are independent, the variance of their sum

equals the sum of their variances:

$$V(Q_\Pi) = \sum_{j=0}^{n-1} V(B_{t_{j+1}} - B_{t_j})^2 = 2 \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2.$$

Each difference $(t_{j+1} - t_j)$ is bounded by the mesh $\|\Pi\|$, so we can write

$$V(Q_\Pi) = 2 \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \leq 2\|\Pi\| \sum_{j=0}^{n-1} (t_{j+1} - t_j) = 2\|\Pi\| T.$$

As the partition is refined and $\|\Pi\| \rightarrow 0$, the variance $V(Q_\Pi)$ also goes to zero:

$$\lim_{\|\Pi\| \rightarrow 0} V(Q_\Pi) \leq 2 \lim_{\|\Pi\| \rightarrow 0} \|\Pi\| T = 0.$$

Step 3: Convergence in $L^2(\mathbb{P})$ and in probability.

Since $E(Q_\Pi) = T$ and $V(Q_\Pi) \rightarrow 0$, the mean-squared error converges to zero:

$$\lim_{\|\Pi\| \rightarrow 0} E(Q_\Pi - T)^2 = \lim_{\|\Pi\| \rightarrow 0} V(Q_\Pi) = 0.$$

This is L^2 convergence: $Q_\Pi \rightarrow T$ in $L^2(\mathbb{P})$.

Moreover, by Chebyshev's inequality, for any $\varepsilon > 0$,

$$P(|Q_\Pi - T| > \varepsilon) \leq \frac{V(Q_\Pi)}{\varepsilon^2} \rightarrow 0 \quad \text{as } \|\Pi\| \rightarrow 0.$$

This shows that Q_Π converges to T in probability: $Q_\Pi \xrightarrow{\mathbb{P}} T$.

The Main Results

We see that the quadratic variation of Brownian motion over $[0, t]$ for any $0 \leq t \leq T$ equals t :

$$[B, B]_t = t.$$

The result $[B, B]_t = t$ is summarized by the shorthand differential notation $(dB_t)^2 = dt$. This is the first entry of the Itô multiplication table:

$$(dB)(dB) = dt, \quad (dB)(dt) = 0, \quad (dt)(dt) = 0.$$

This notation is heuristic and will be made precise when we study Itô's lemma. These equalities hold in the mean-square sense, not path by path: each squared increment $(B_{t_{j+1}} - B_{t_j})^2$ is a random variable with mean $t_{j+1} - t_j$, not a deterministic quantity equal to it.

To justify the remaining two entries, let $M_\Pi = \max_{0 \leq j \leq n-1} |B_{t_{j+1}} - B_{t_j}|$ denote the maximum absolute increment over the partition. Since Brownian motion is uniformly continuous on $[0, T]$, we have $M_\Pi \rightarrow 0$ as $\|\Pi\| \rightarrow 0$. It follows that

$$\left| \sum_{j=0}^{n-1} (B_{t_{j+1}} - B_{t_j})(t_{j+1} - t_j) \right| \leq M_\Pi \sum_{j=0}^{n-1} (t_{j+1} - t_j) = M_\Pi T \xrightarrow{\|\Pi\| \rightarrow 0} 0,$$

and

$$\sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \leq \|\Pi\| \sum_{j=0}^{n-1} (t_{j+1} - t_j) = \|\Pi\| T \xrightarrow{\|\Pi\| \rightarrow 0} 0.$$

These limits confirm $(dB)(dt) = 0$ and $(dt)(dt) = 0$.

Brownian Motion is Nowhere Differentiable

Recall that $M_\Pi = \max_{0 \leq j \leq n-1} |B_{t_{j+1}} - B_{t_j}| \rightarrow 0$ as $\|\Pi\| \rightarrow 0$, and let $Q_\Pi = \sum_{j=0}^{n-1} (B_{t_{j+1}} - B_{t_j})^2$. For each increment we have the inequality

$$(B_{t_{j+1}} - B_{t_j})^2 \leq M_\Pi |B_{t_{j+1}} - B_{t_j}|.$$

Summing over all subintervals gives $Q_\Pi \leq M_\Pi \sum_{j=0}^{n-1} |B_{t_{j+1}} - B_{t_j}|$, or equivalently,

$$\sum_{j=0}^{n-1} |B_{t_{j+1}} - B_{t_j}| \geq \frac{Q_\Pi}{M_\Pi}.$$

As $\|\Pi\| \rightarrow 0$, the right-hand side diverges since $Q_\Pi \rightarrow T > 0$ while $M_\Pi \rightarrow 0$. The total variation of Brownian motion over $[0, T]$ is therefore almost surely infinite.

Since the total variation is infinite on every interval, Brownian paths cannot be continuously differentiable on any subinterval of $[0, T]$: if a path were C^1 on $[a, b]$, its total variation would equal $\int_a^b |B'(t)| dt$, which is finite. In fact, a stronger result holds: Brownian paths are almost surely nowhere differentiable. Establishing this in full rigor requires a separate argument, but the infinite total variation already reveals how irregular these trajectories are.

Shreve, Steven E. 2004. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance. Springer.