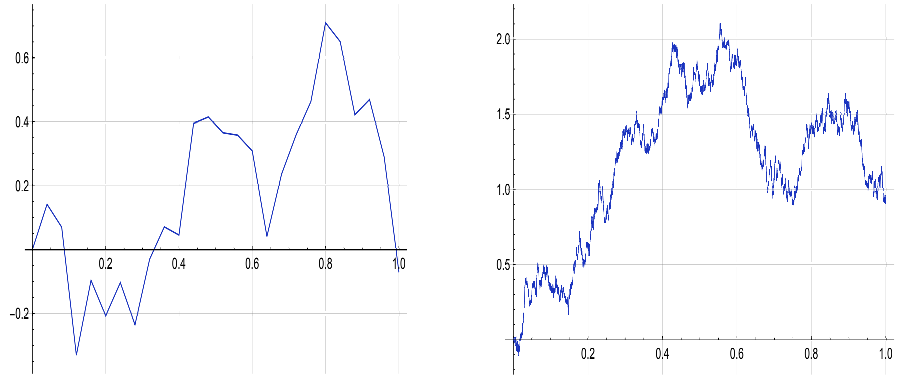

Figure 1: Euler scheme approximation for Brownian motion with 50 time steps (left) and 10000 time steps (right)

Brownian motion

A continuous time stochastic process $(W_t)_{t\geq 0}$ is said to be a standard one-dimensional **Brownian motion** if it satisfies the following four properties:

- $W_0 = 0$.
- W has **independent increments**, i.e.

$$W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2} \dots, W_{t_n} - W_{t_{n-1}}$$

  are independent for all $0 \leq t_1 < t_2 < \dots < t_n$.
- The increments are **Normally distributed**: $W_t - W_s \sim N(0, t-s)$ for all $0 \leq s \leq t$.
- W_t is continuous as a function of t almost surely (i.e. with probability one), i.e. $W_{t+h} - W_t \to 0$ as $h \to 0$.

Remark 0.1 It is not obvious that we can rigorously construct a process W which satisfies these four properties (one do this using a Fourier series representation for W known as the Karhuenen-Loève expansion, or by successively sampling W at midpoints using Hwk 1 q7, or using Haar wavelets). We will address this issue on the next page when we consider the Euler method (see also numerical simulation on next page to see what Brownian motion looks like).

Remark 0.2 For this course, the third defining property of Brownian motion is the most important to remember. Setting $s = 0$, then since $W_s = 0$ by the first property in the definition, we see that

$$W_t \quad \sim \quad N(0, t) \tag{1}$$

and thus $\mathbb{P}(W_t > x) = \mathbb{P}(\frac{W_t - 0}{\sqrt{t}} > \frac{x-0}{\sqrt{t}}) = \mathbb{P}(Z > \frac{x}{\sqrt{t}}) = \Phi^c(\frac{x}{\sqrt{t}})$ where $\Phi^c(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$, and we are using that $Z = (W_t - 0)/\sqrt{t}$ is a standard Normal random variable, so the density of W_t is $p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}x^2}$. We will use (1) repeatedly on this course.

Covariance of Brownian motion

Let $0 \leq s \leq t$. Then

$$\mathbb{E}(W_s W_t) \quad = \quad \mathbb{E}(W_s(W_s + W_t - W_s)) \quad = \quad \mathbb{E}(W_s^2) \quad = \quad \mathrm{Var}(W_s) \quad = \quad s$$

since $\mathbb{E}(W_s(W_t - W_s)) = \mathbb{E}((W_s - W_0)(W_t - W_s)) = \mathbb{E}(W_s - W_0)\mathbb{E}(W_t - W_s) = 0$, due to the independent increments property. This means that in general, for $s, t \geq 0$

$$R(s,t) \quad := \quad \mathbb{E}(W_s W_t) \quad = \quad \min(s,t)\,. \tag{2}$$

This is known as the **covariance function** of Brownian motion. For a Gaussian process i.e. a random process for which $(X_{t_1}, \dots, X_{t_n})$ has a **multivariate Normal distribution**, specifying the covariance of the process is sufficient to uniquely characterize the process.

Lack of differentiability

Recall that a function f is said to be differentiable at x if the following limit exists

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.$$

If $0\le s<t$, clearly the intervals $[s,s+h]$ and $[t,t+h]$ **do not intersect** for $h>0$ sufficiently small. Then using the second property of Brownian motion we see that

$$\mathbb{E}(\frac{W_{s+h}-W_s}{h}\,\frac{W_{t+h}-W_t}{h}) = \frac{1}{h^2}\mathbb{E}((W_{s+h}-W_s)(W_{t+h}-W_t)) = \frac{1}{h^2}\mathbb{E}(W_{s+h}-W_s)\mathbb{E}(W_{t+h}-W_s) = 0$$

for h sufficiently small. Conversely if $s=t$ then this expecation is $\frac{1}{h^2}\mathbb{E}((W_{t+h}-W_t)^2)=\frac{1}{h^2}h=\frac{1}{h}\to\infty$ as $h\to 0$. Thus $\Delta_t^h := \frac{1}{h}(W_{t+h}-W_t)$ is a well defined process for each $h>0$ with $\lim_{h\to 0}\mathbb{E}(\Delta_s^h\Delta_t^h)=+\infty$ if $s=t$, and zero otherwise, which cannot be the covariance of a well defined process in the limit as $h\to 0$ because we have $+\infty$ here.

Remark 0.3 We have given a sketch proof as to why W is not differentiable, but from the fourth property in the definition we know that W is **continuous** a.s. (so it does not have jumps), and (using something called the **Kolmogorov continuity theorem**) we can make a stronger statement that W is **α-Hölder continuous** for $\alpha\in(0,\frac{1}{2})$, i.e. for $0\le s\le t\le T$

$$|W_t - W_s| \le c_1|t-s|^\alpha$$

for some **random** constant c_1 which depends on $(W_t)_{0\le t\le T}$ which is finite almost surely (a.s.) (see FM04 for more details on c_1). Note this implies that W is continuous a.s. since as $t-s\to 0$, $|W_t-W_s|\to 0$ because $\alpha>0$.

Constructing and simulating Brownian motion - the Euler Monte Carlo scheme

- Let Z_i be a sequence of standard i.i.d. $N(0,1)$ random variables. Then we can approximate Brownian motion numerically as follows: fix a small step size $\Delta t>0$, set $W_0^n=0$, and then iteratively define

  $$W^n_{(i+1)\Delta t} = W^n_{i\Delta t} + \sqrt{\Delta t}\, Z_i$$

  and $\Delta t=\frac{1}{n}$, and we join $W^n_{i\Delta t}$ and $W^n_{(i+1)\Delta t}$ using linear interpolation (see figure above).

- The W_t^n process starts at zero, has independent increments (more precisely the discrete time process $X_i = W^n_{i\Delta t}$ has independent increments) and we see that $W^n_{(i+1)\Delta t}-W^n_{i\Delta t}\sim N(0,\Delta t)$, which is consistent with the third property of Brownian motion that $W_t-W_s\sim N(0,t-s)$ (recall that for *any* random variable X, we have that $\mathrm{Var}(aX)=a^2\mathrm{Var}(X)$).

- This procedure is known as the **Euler method**. If we then join the points $(W^n_{\Delta t}, W^n_{2\Delta t},...)$ with straight lines, this methods gives a *piecewise linear* approximation W_t^n to a true Brownian motion (see plots above).

- Using a deeper result called **Donsker's theorem**, it can be proved that this construction tends to a Brownian motion as the step size $\Delta t\to 0$, i.e. as $n\to\infty$ (recall that $\Delta t=\frac{1}{n}$). More precisely, for any bounded continuous function $f(x_1,...,x_n)$ and $0=t_0<t_1<t_2<...<t_k$, we have that

  $$\lim_{n\to\infty}\mathbb{E}(f(W^n_{t_1},...,W^n_{t_k})) = \mathbb{E}(f(X_1,...,X_k))$$

  where $X=(X_1,..,X_k)$ is a vector of k Normal random variables with $\mathbb{E}(X_iX_j)=R(t_i,t_j)$, where $R(.,.)$ is the covariance function computed above in (2). This is known as **weak convergence**.

Quadratic variation of Brownian motion

A **partition** of the time interval $[0,t]$ is a set of the form $\Pi=\{t_0=0<t_1<...<t_n=t\}$, and we define the size of the partition to be

$$|\Pi| = \max_{0\le i\le n-1}(t_{i+1}-t_i)$$

i.e. equal to the largest interval of the partition. The **quadratic variation** of a random process X over a fixed time interval $[0,t]$ is then defined as

$$[X,X]_t = \langle X\rangle_t = \lim_{\|\Pi\|\to 0}\sum_{i=0}^{n-1}(X_{t_{i+1}}-X_{t_i})^2$$

if this limit exists and does not depend on the choice of the sequence of partitions Π^n. In general the quadratic variation $[X, X]_t$ of a process X is a random process, but we will see that for Brownian motion W, $[W, W]_t = t$ a.s.

We now look at the case when $X = W$ is Brownian motion. Let $\Delta W_i = W_{i\Delta t} - W_{(i-1)\Delta t}$ where $\Delta t = \frac{t}{n}$, then

$$\Pi_1^n \quad = \quad \sum_{i=0}^{n-1}(W_{\frac{(i+1)t}{n}} - W_{\frac{it}{n}})^2 \quad = \quad \sum_{i=1}^{n}(\Delta W_i)^2 \quad \sim \quad \sum_{i=1}^{n}(\sqrt{\Delta t}\, Z_i)^2 \quad = \quad \Delta t \sum_{i=1}^{n} Z_i^2 \quad = \quad \frac{t}{n}\sum_{i=1}^{n} Z_i^2$$

where $Z_1, Z_2, ...$ is an i.i.d sequence of $N(0, 1)$ random variables. Then from the **Strong Law of Large Numbers** (SLLN), $\Pi_1^n \to t$ a.s. as $n \to \infty$, as we would intuitively expect.

Morever, we also know that

$$\mathrm{Var}(\sum_{i=1}^{n} Z_i^2) \quad = \quad \sum_{i=1}^{n} \mathrm{Var}(Z_i^2) \quad = \quad 2n$$

since $\mathbb{E}((Z_i^2)^2) = \mathbb{E}(Z_i^4) = 3$ and $\mathbb{E}(Z_i^2) = 1$, so $\mathrm{Var}(Z_i^2) = \mathbb{E}((Z_i^2)^2) - \mathbb{E}((Z_i^2))^2 = 3 - 1 = 2$. Then

$$\mathrm{Var}(\sum_{i=1}^{n} (\Delta W_i)^2) \quad = \quad \mathrm{Var}(\sum_{i=1}^{n} (\sqrt{\Delta t}\, Z_i)^2) \quad = \quad n \cdot 2\,(\Delta t)^2 \quad = \quad \frac{2t^2}{n}\,.$$

Moreover

$$\sqrt{n}\,(\sum_{i=1}^{n} (\Delta W_i)^2 - t) \quad \sim \quad \sqrt{n}\,(\sum_{i=1}^{n} \Delta t Z_i^2 - t) \quad = \quad t\sqrt{n}\,(\frac{1}{n}\sum_{i=1}^{n} Z_i^2 - 1) \tag{3}$$

and from the **Central Limit Theorem** (i.e. that $\sqrt{n}(\bar{X} - \mu) \to N(0, \sigma^2)$ where $\bar{X}$ is the sample mean of n i.i.d random variables with mean μ and variance σ^2) the right hand side tends weakly to a $N(0, \sigma_1^2)$ random variable, where here $\sigma_1^2 = t^2\mathrm{Var}(Z_i^2) = 2t^2$, since we have an additional t pre-factor in (3). Hence for n large we can approximately say that

$$\sum_{i=1}^{n} (\Delta W_i)^2 - t \quad \sim \quad \frac{1}{\sqrt{n}} N(0, \sigma_1^2). \tag{4}$$

See Python code at for simulations of Brownian motion and numerical confirmation of this result `https://colab.research.google.com/drive/1hB9AOgDiw-gTTZu2Osp2in-1YU3FjuAu?usp=sharing`) The density $p(x)$ of $\chi = \sum_{i=1}^{n} Z_i^2$ is χ^2 with n degrees of freedom, and has the explicit formula

$$p(x) \quad = \quad \frac{1}{2^{n/2}\Gamma(\frac{1}{2}n)} x^{\frac{1}{2}n - 1} e^{-\frac{1}{2}x}$$

for $x > 0$.

Example. Let us prove this result with our bare hands without using the CLT. The mgf of $\chi = \sum_{i=1}^{n} Z_i^2$ is known to be $M(p) := (1 - 2p)^{-\frac{1}{2}n}$ for $p < \frac{1}{2}$, and $+\infty$ otherwise. Hence the mgf of $\chi/n - 1$ is $f(p) = (1 - \frac{2p}{n})^{-\frac{1}{2}n} e^{-p}$, and the mgf of $\sqrt{n}(\chi/n - 1)$ is

$$f(p\sqrt{n}) \quad = \quad (1 - \frac{2p}{\sqrt{n}})^{-\frac{1}{2}n} e^{-p\sqrt{n}}\,.$$

The log of this expression is $\log f(p\sqrt{n}) = -p\sqrt{n} - \frac{1}{2}n\log(1 - \frac{2p}{\sqrt{n}})$. Then using that $\log(1 - x) = -x - \frac{x^2}{2} + O(x^3)$, we find that

$$\log f(p\sqrt{n}) \quad = \quad -p\sqrt{n} - \frac{1}{2}n(-2p/\sqrt{n}) + \frac{1}{2}n(-2p/\sqrt{n})^2/2 \,+\, O(\frac{1}{\sqrt{n}}) \quad = \quad p^2 \,+\, O(\frac{1}{\sqrt{n}})$$

as $n \to \infty$. Exponentiating this, we recover e^{p^2} as $n \to \infty$, i.e. the mgf of an $N(0, 2)$ random variable as above.

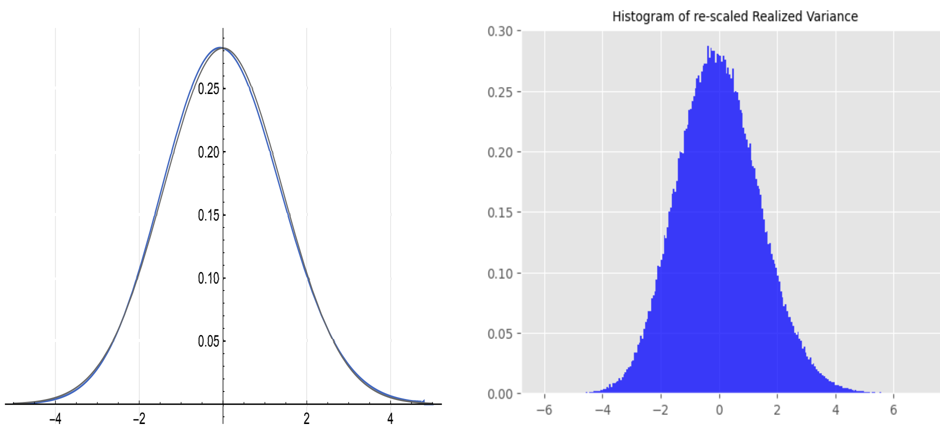

Figure 2: On the left we have plotted the density of the re-scaled Realized Variance: $\sqrt{n}\left(\sum_{i=1}^{n}(\Delta W_i)^2 - t\right)$ (blue) with $t = 1$ and $n = 800$, verses the density of an $N(0, 2)$ random variable (grey), and we see they are almost identical, as the CLT shows above. On the right we have plotted a histogram of the re-scaled Realized Variance with 250,000 Brownian paths and $n = 200$ time steps which again tends to an $N(0, 2)$ distribution as $n \to \infty$, and the sample variance here came out to 2.0111, i.e. close to 2.