Chapter 4
Brownian Motion and Stochastic Calculus

Brownian motion is a continuous-time stochastic process having stationary and independent Gaussian distributed increments, and continuous paths. This chapter presents the constructions of Brownian motion and its associated Itô stochastic integral, which will be used for the random modeling of asset and portfolio prices in continuous time.

4.1 Brownian Motion

We start by recalling the definition of Brownian motion, which is a fundamental example of a stochastic process. The underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ of Brownian motion can be constructed on the space $\Omega = \mathcal{C}_0(\mathbb{R}_+)$ of continuous real-valued functions on $\mathbb{R}_+$ started at 0.

Definition 4.1. *The standard Brownian motion is a stochastic process* $(B_t)_{t\in\mathbb{R}_+}$ *such that*

1. $B_0 = 0$,

2. *The sample trajectories* $t \mapsto B_t$ *are continuous, with probability one.*

3. *For any finite sequence of times* $t_0 < t_1 < \cdots < t_n$, *the increments*

$$B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$$

are mutually independent random variables.

4. *For any given times* $0 \leqslant s < t$, $B_t - B_s$ *has the Gaussian distribution* $\mathcal{N}(0, t-s)$ *with mean zero and variance* $t-s$.

In particular, for $t \in \mathbb{R}_+$, the random variable $B_t \simeq \mathcal{N}(0,t)$ has a Gaussian distribution with mean zero and variance $t > 0$. Existence of a stochastic process satisfying the conditions of Definition 4.1 will be covered in Section 4.2, see also Problem 4.22.

In Figure 4.1 we draw three sample paths of a standard Brownian motion obtained by computer simulation using (4.3). Note that there is no point in "computing" the value of B_t as it is a *random variable* for all $t > 0$. However, we can generate samples of B_t, which are distributed according to the centered Gaussian distribution with variance $t > 0$ as in Figure 4.1.

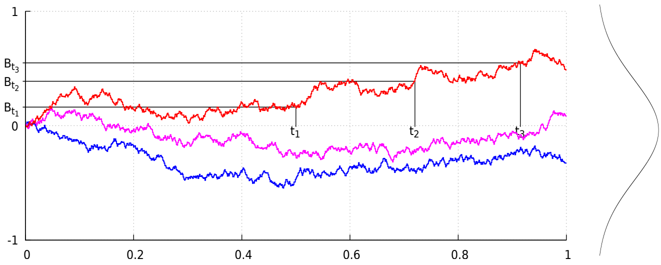


Fig. 4.1: Sample paths of a one-dimensional Brownian motion.

In particular, Property *4* in Definition 4.1 implies

$$\mathbb{E}[B_t - B_s] = 0 \quad \text{and} \quad \mathrm{Var}[B_t - B_s] = t - s, \qquad 0 \leqslant s \leqslant t,$$

and we have

$$\begin{aligned}
\mathrm{Cov}(B_s, B_t) &= \mathbb{E}[B_s B_t] \\
&= \mathbb{E}[B_s(B_t - B_s + B_s)] \\
&= \mathbb{E}\big[B_s(B_t - B_s) + (B_s)^2\big] \\
&= \mathbb{E}[B_s(B_t - B_s)] + \mathbb{E}\big[(B_s)^2\big] \\
&= \mathbb{E}[B_s]\mathbb{E}[B_t - B_s] + \mathbb{E}\big[(B_s)^2\big] \\
&= \mathrm{Var}[B_s] \\
&= s, \qquad 0 \leqslant s \leqslant t,
\end{aligned}$$

hence

$$\mathrm{Cov}(B_s, B_t) = \mathbb{E}[B_s B_t] = \min(s,t), \qquad s,t \geqslant 0, \tag{4.1}$$

cf. also Exercise 4.1. The following graphs present two examples of possible modeling of random data using Brownian motion.

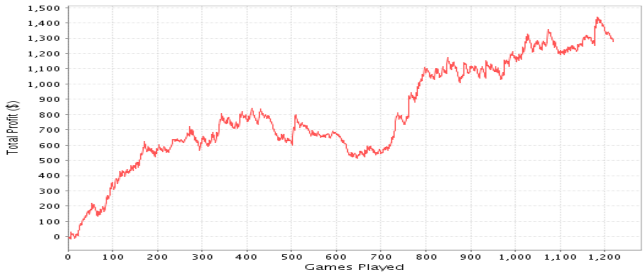


Fig. 4.2: Evolution of the fortune of a poker player *vs.* number of games played.

Fig. 4.3: Web traffic ranking.

In what follows, we denote by $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ the filtration generated by the Brownian paths up to time t, defined as

$$\mathcal{F}_t := \sigma(B_s \ : \ 0 \leqslant s \leqslant t), \qquad t \geqslant 0. \tag{4.2}$$

Property 3 in Definition 4.1 shows that $B_t - B_s$ is independent of all Brownian increments taken before time s, *i.e.*

$$(B_t - B_s) \perp\!\!\!\perp (B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}),$$

$0 \leqslant t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant s \leqslant t$, hence $B_t - B_s$ is also independent of the whole Brownian history up to time s, hence $B_t - B_s$ is in fact independent of $\mathcal{F}_s$, $s \geqslant 0$.

Definition 4.2. *A continuous-time process* $(Z_t)_{t\in\mathbb{R}_+}$ *of integrable random variables is a martingale under* $\mathbb{P}$ *and with respect to the filtration* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *if*

$$\mathbb{E}[Z_t \mid \mathcal{F}_s] = Z_s, \qquad 0 \leqslant s \leqslant t.$$

Note that when $(Z_t)_{t\in\mathbb{R}_+}$ is a martingale, Z_t is in particular $\mathcal{F}_t$-measurable at all times $t \geqslant 0$. As in Example 2 on page 272, we have the following result.

Proposition 4.3. *Brownian motion* $(B_t)_{t\in\mathbb{R}_+}$ *is a continuous-time* martingale.

Proof. We have

$$\begin{aligned}\mathbb{E}[B_t \mid \mathcal{F}_s] &= \mathbb{E}[B_t - B_s + B_s \mid \mathcal{F}_s]\\ &= \mathbb{E}[B_t - B_s \mid \mathcal{F}_s] + \mathbb{E}[B_s \mid \mathcal{F}_s]\\ &= \mathbb{E}[B_t - B_s] + B_s\\ &= B_s, \qquad 0 \leqslant s \leqslant t,\end{aligned}$$

because it has centered and independent increments, cf. Section 7.1. $\square$

The n-dimensional Brownian motion can be constructed as $\big(B_t^{(1)}, B_t^{(2)}, \ldots, B_t^{(n)}\big)_{t\in\mathbb{R}_+}$ where $\big(B_t^{(1)}\big)_{t\in\mathbb{R}_+}$, $\big(B_t^{(2)}\big)_{t\in\mathbb{R}_+}$, $\ldots$, $\big(B_t^{(n)}\big)_{t\in\mathbb{R}_+}$ are independent copies of $(B_t)_{t\in\mathbb{R}_+}$. Next, we turn to simulations of 2 dimensional and 3 dimensional Brownian motions in Figures 4.4 and 4.5. Recall that the movement of pollen particles originally observed by Brown (1828) was indeed 2-dimensional.

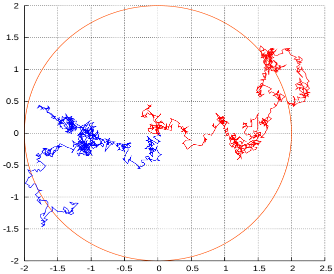


Fig. 4.4: Two sample paths of a two-dimensional Brownian motion.

This version: May 3, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

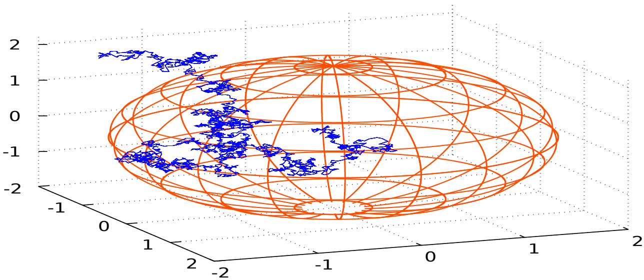


Fig. 4.5: Sample path of a three-dimensional Brownian motion.

Figure 4.6 presents an illustration of the scaling property of Brownian motion.

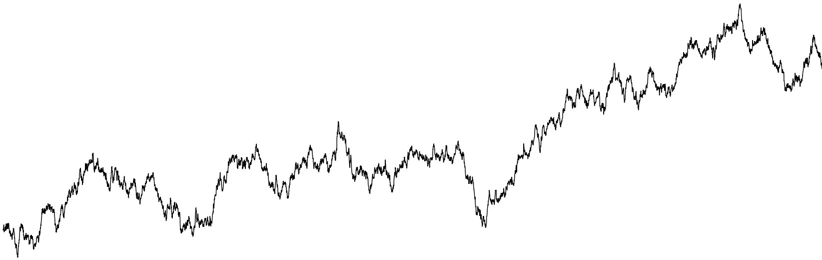


Fig. 4.6: Scaling property of Brownian motion.*

4.2 Three Constructions of Brownian Motion

We refer the reader to Chapter 1 of Revuz and Yor (1994) and to Theorem 10.28 in Folland (1999) for proofs of existence of Brownian motion as a stochastic process $(B_t)_{t\in\mathbb{R}_+}$ satisfying the Conditions *1-4* of Definition 4.1.

Brownian motion as a random walk

We start with an informal description of Brownian motion as a random walk over infinitesimal time intervals of length Δt, whose increments

* The animation works in Acrobat Reader on the entire pdf file.

$$\Delta B_t := B_{t+\Delta t} - B_t \simeq \mathcal{N}(0, \Delta t)$$

over the time interval $[t, t+\Delta t]$ will be approximated by the Bernoulli random variable

$$\Delta B_t = \pm\sqrt{\Delta t} \tag{4.3}$$

with equal probabilities $(1/2, 1/2)$. According to this representation, the paths of Brownian motion are not differentiable, although they are continuous by Property 2, as we have

$$\frac{dB_t}{dt} \simeq \pm\frac{\sqrt{dt}}{dt} = \pm\frac{1}{\sqrt{dt}} \simeq \pm\infty. \tag{4.4}$$

Figure 4.7 presents a simulation of Brownian motion as a random walk with $\Delta t = 0.1$.

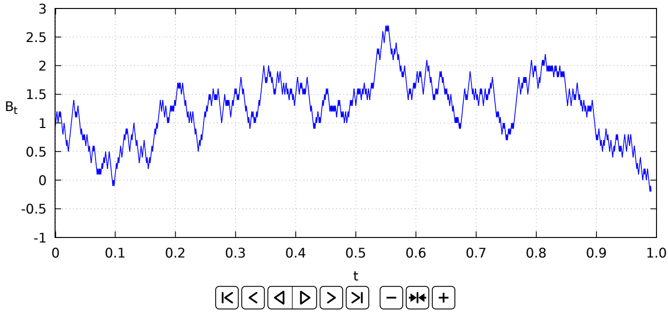


Fig. 4.7: Construction of Brownian motion as a random walk with $B_0 = 1$.*

Note that we have

$$\mathbb{E}[\Delta B_t] = \frac{1}{2}\sqrt{\Delta t} - \frac{1}{2}\sqrt{\Delta t} = 0,$$

and

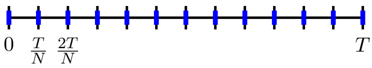
$$\mathrm{Var}[\Delta B_t] = \mathbb{E}\big[(\Delta B_t)^2\big] = \frac{1}{2}(+\sqrt{\Delta t})^2 + \frac{1}{2}(-\sqrt{\Delta t})^2 = \frac{1}{2}\Delta t + \frac{1}{2}\Delta t = \Delta t.$$

In order to recover the Gaussian distribution property of the random variable B_T, we can split the time interval $[0, T]$ into N subintervals

$$\left(\frac{k-1}{N}T, \frac{k}{N}T\right], \qquad k = 1, 2, \ldots, N,$$

of same length $\Delta t = T/N$, with N "large".

* The animation works in Acrobat Reader on the entire pdf file.



Defining the Bernoulli random variable X_k as

$$X_k := \pm\sqrt{T}$$

with equal probabilities $(1/2, 1/2)$, we have $\mathrm{Var}(X_k) = T$ and

$$\Delta B_t := \frac{X_k}{\sqrt{N}} = \pm\sqrt{\Delta t}$$

is the increment of B_t over $((k-1)\Delta t, k\Delta t]$, and we get

$$B_T \simeq \sum_{0<t<T} \Delta B_t \simeq \frac{X_1 + X_2 + \cdots + X_N}{\sqrt{N}}.$$

Hence by the central limit theorem we recover the fact that B_T has the centered Gaussian distribution $\mathcal{N}(0,T)$ with variance T, cf. point 4 of the above Definition 4.1 of Brownian motion, and the illustration given in Figure 4.8. Indeed, the central limit theorem states that given any sequence $(X_k)_{k\geqslant 1}$ of independent identically distributed centered random variables with variance $\sigma^2 = \mathrm{Var}[X_k] = T$, the normalized sum

$$\frac{X_1 + X_2 + \cdots + X_N}{\sqrt{N}}$$

converges (in distribution) to the centered Gaussian random variable $\mathcal{N}(0,\sigma^2)$ with variance σ^2 as N goes to infinity. As a consequence, ΔB_t could in fact be replaced by any centered random variable with variance Δt in the above description.

```r
N=1000; t <- 0:N; dt <- 1.0/N; dev.new(width=16,height=7); # Using Bernoulli samples
nsim=100;X <- matrix((dt)^0.5*(rbinom( nsim * N, 1, 0.5)-0.5)*2, nsim, N)
X <- cbind(rep(0, nsim), t(apply(X, 1, cumsum))); H<-hist(X[,N],plot=FALSE);
layout(matrix(c(1,2), nrow =1, byrow = TRUE));par(mar=c(2,2,2,0), oma = c(2, 2, 2, 2))
plot(t*dt, X[1, ], xlab = "", ylab = "", type = "l", ylim = c(-2, 2), col = 0,xaxs='i',las=1,
      cex.axis=1.6)
for (i in 1:nsim){lines(t*dt, X[i, ], type = "l", ylim = c(-2, 2), col = i)}
lines(t*dt,sqrt(t*dt),lty=1,col="red",lwd=3);lines(t*dt,-sqrt(t*dt), lty=1, col="red",lwd=3)
lines(t*dt,0*t, lty=1, col="black",lwd=2)
for (i in 1:nsim){points(0.999, X[i,N], pch=1, lwd = 5, col = i)}
x <- seq(-2,2, length=100); px <- dnorm(x);par(mar = c(2,2,2,2))
plot(NULL , xlab="", ylab="", xlim = c(0, max(px,H$density)), ylim = c(-2,2),axes=F)
rect(0, H$breaks[1:(length(H$breaks) - 1)], col=rainbow(20,start=0.08,end=0.6), H$density,
      H$breaks[2:length(H$breaks)]); lines(px,x, lty=1, col="black",lwd=2)
```

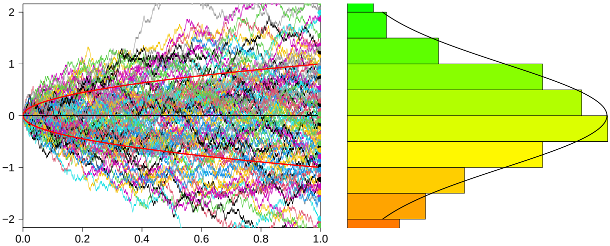


Fig. 4.8: Statistics of one-dimensional Brownian paths *vs.* Gaussian distribution.

Remark 4.4. *The choice of the square root in* (4.3) *is in fact not fortuitous. Indeed, any choice of* $\pm(\Delta t)^{\alpha}$ *with a power* $\alpha > 1/2$ *would lead to explosion of the process as* dt *tends to zero, whereas a power* $\alpha \in (0, 1/2)$ *would lead to a vanishing process, as can be checked from the following* R *code.*

The following R code plots a set of 72 normalized yearly return graphs of the S&P 500 index from 1950 to 2022, together with their distribution, see Figure 4.9, for comparison with the path properties and statistics of Brownian motion.

```
library(quantmod); getSymbols("^GSPC",from="1950-01-01",to="2022-12-31",src="yahoo")
stock<-Cl(`GSPC`); s=0;y=0;j=0;count=0;N=240;nsim=72; X = matrix(0, nsim, N)
for (i in 1:nrow(GSPC)){if (s==0 && grepl('-01-0',index(stock[i]))) {if (count==0 || X[y,N]>0)
       {y=y+1;j=1;s=1;count=count+1;}}
if (j<=N) {X[y,j]=as.numeric(stock[i]);};if (grepl('-02-0',index(stock[i]))) {s=0;};j=j+1;}
t <- 0:(N-1); dt <- 1.0/N; m=mean(X[,N]/X[,1]-1); sigma=sd(X[,N]/X[,1]-1);
       dev.new(width=16,height=7);
layout(matrix(c(1,2), nrow =1, byrow = TRUE));par(mar=c(2,2,2,0), oma = c(2, 2, 2, 2))
plot(t*dt, X[1,]/X[1,1]-1-m*t*dt, xlab = "", ylab = "", type = "l", ylim = c(-0.5, 0.5), col = 0,
       xaxs='i',las=1, cex.axis=1.6)
for (i in 1:nsim){lines(t*dt, X[i,]/X[i,1]-1-m*t*dt, type = "l", col = i)}
lines(t*dt,sigma*sqrt(t*dt),lty=1,col="red",lwd=3);lines(t*dt,-sigma*sqrt(t*dt), lty=1,
       col="red",lwd=3)
lines(t*dt,0*t, lty=1, col="black",lwd=2)
for (i in 1:nsim){points(0.999, X[i,N]/X[i,1]-1-m*N*dt, pch=1, lwd = 5, col = i)}
x <- seq(-0.5,0.5, length=100); px <- dnorm(x,0,sigma);
       H<-hist(X[,N]/X[,1]-1-m*N*dt,plot=FALSE);
plot(NULL , xlab="", ylab="", xlim = c(0, max(px,H$density)), ylim = c(-0.5,0.5),axes=F)
rect(0, H$breaks[1:(length(H$breaks) - 1)], col=rainbow(20,start=0.08,end=0.6), H$density,
H$breaks[2:length(H$breaks)]); lines(px,x, lty=1, col="black",lwd=2)
```

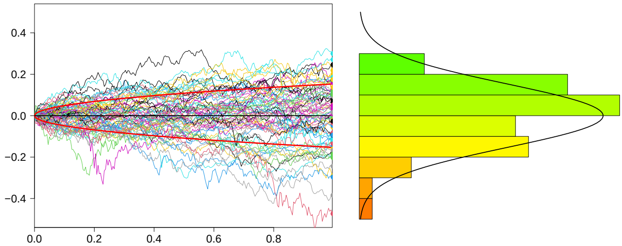


Fig. 4.9: Statistics of 72 S&P 500 yearly normalized return graphs from 1950 to 2022.

Lévy's construction of Brownian motion

Figure 4.10 represents the construction of Brownian motion by successive linear interpolations, see Problem 4.22 for a proof of existence of Brownian motion based on this construction.

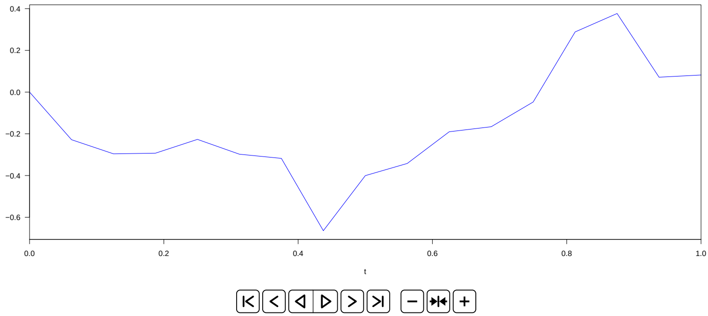


Fig. 4.10: Lévy's construction of Brownian motion.*

The following R code is used to generate Figure 4.10.†

* The animation works in Acrobat Reader on the entire pdf file.

† Download the corresponding R **code** or the **IPython notebook** that can be run **here** or **here**.

```
dev.new(width=16,height=7); alpha=1/2;t <- 0:1;dt <- 1; z=rnorm(1,mean=0,sd=dt^alpha)
plot(t*dt,c(0,z),xlab = "t",ylab = "",col = "blue",main = "",type = "l", xaxs="i", las = 1)
k=0;while (k<12) {readline("Press <return> to continue")
k=k+1;m <- (z+c(0,head(z,-1)))/2;y <- rnorm(length(t)-1,mean=0,sd=(dt/4)^alpha)
x <- m+y;x <- c(matrix(c(x,z), 2, byrow = T));n=2*length(t)-2;t <- 0:n
plot(t*dt/2, c(0, x), xlab = "t", ylab = "", col = "blue", main = "", type = "l", xaxs="i", las =
      1);z=x;dt=dt/2}
```

Construction by series expansions

Brownian motion on $[0, T]$ can also be constructed by Fourier synthesis via the Paley-Wiener series expansion

$$B_t = \sum_{n \geqslant 1} X_n f_n(t) = \frac{\sqrt{2T}}{\pi} \sum_{n \geqslant 1} X_n \frac{\sin((n-1/2)\pi t/T)}{n-1/2}, \quad 0 \leqslant t \leqslant T,$$

where $(X_n)_{n \geqslant 1}$ is a sequence of independent $\mathcal{N}(0,1)$ standard Gaussian random variables, as illustrated in Figure 4.11.*

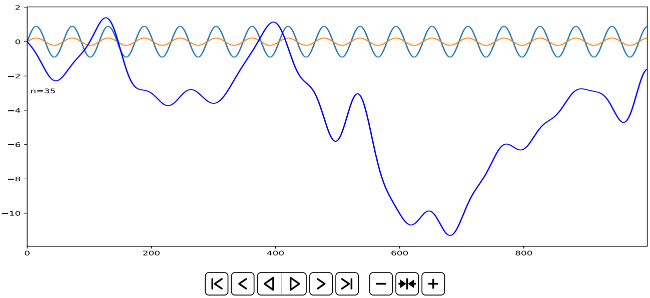


Fig. 4.11: Construction of Brownian motion by series expansions.†

4.3 Wiener Stochastic Integral

In this section, we construct the Wiener stochastic integral of square-integrable deterministic functions of time with respect to Brownian motion.

Recall that the price S_t of risky assets was originally modeled in Bachelier (1900) as $S_t := \sigma B_t$, where σ is a volatility parameter. The stochastic integral

* Download the corresponding **IPython notebook** that can be run **here** or **here**.

† The animation works in Acrobat Reader on the entire pdf file.

$$\int_0^T f(t)dS_t = \sigma \int_0^T f(t)dB_t$$

can be used to represent the value of a portfolio as a sum of profits and losses $f(t)dS_t$ where dS_t represents the stock price variation and $f(t)$ is the quantity invested in the asset S_t over the short time interval $[t, t+dt]$.

A naive definition of the stochastic integral with respect to Brownian motion would consist in letting

$$\int_0^T f(t)dB_t := \int_0^T f(t)\frac{dB_t}{dt}dt,$$

and evaluating the above integral with respect to dt. However, this definition fails because the paths of Brownian motion are not differentiable, cf. (4.4). Next we present Itô's construction of the stochastic integral with respect to Brownian motion. Stochastic integrals will be first constructed as integrals of simple step functions of the form

$$f(t) = \sum_{i=1}^n a_i \mathbb{1}_{(t_{i-1},t_i]}(t), \qquad 0 \leqslant t \leqslant T, \tag{4.5}$$

i.e. the function f takes the value a_i on the interval $(t_{i-1}, t_i]$, $i = 1, 2, \ldots, n$, with $0 \leqslant t_0 < \cdots < t_n \leqslant T$, as illustrated in Figure 4.12.

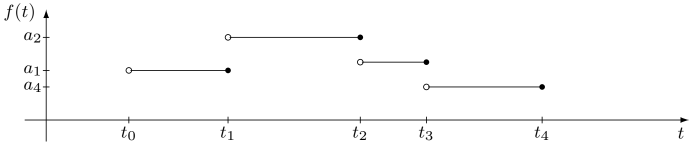


Fig. 4.12: Step function $t \mapsto f(t)$.

```
ti<-c(0,2,4.5,7,9)
ai<-c(0,3,1,2,1,0)
plot(stepfun(ti,ai),xlim = c(0,10),do.points = F,main="", col = "blue")
```

Recall that the classical integral of f given in (4.5) is interpreted as the area under the curve represented by f, and computed as

$$\int_0^T f(t)dt = \sum_{i=1}^n a_i(t_i - t_{i-1}).$$

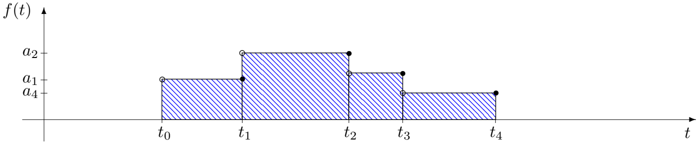


Fig. 4.13: Area under the step function $t \mapsto f(t)$.

In Definition 4.5 we use such step functions for the construction of the stochastic integral with respect to Brownian motion. The stochastic integral (4.6) for step functions will be interpreted as the sum of profits and losses $a_i(B_{t_i} - B_{t_{i-1}})$, $i = 1, 2, \ldots, n$, in a portfolio holding a quantity a_i of a risky asset whose price variation is $B_{t_i} - B_{t_{i-1}}$ at time $i = 1, 2, \ldots, n$.

Definition 4.5. *The stochastic integral with respect to Brownian motion $(B_t)_{t\in[0,T]}$ of the simple step function f of the form (4.5) is defined by*

$$\int_0^T f(t) dB_t := \sum_{i=1}^n a_i (B_{t_i} - B_{t_{i-1}}). \tag{4.6}$$

In what follows, we will make a repeated use of the space $L^2([0,T])$ of *square-integrable functions*.

Definition 4.6. *Let $L^2([0,T])$ denote the space of (measurable) functions $f : [0,T] \longrightarrow \mathbb{R}$ such that*

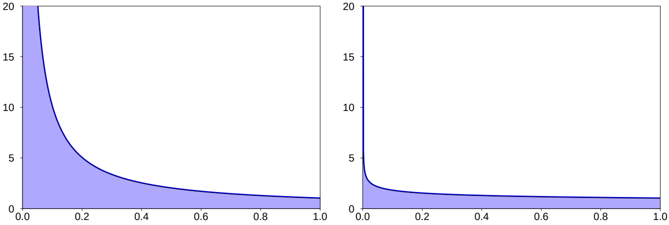
$$\|f\|_{L^2([0,T])} := \sqrt{\int_0^T |f(t)|^2 dt} < \infty, \quad f \in L^2([0,T]). \tag{4.7}$$

In the above definition, $\|f\|_{L^2([0,T])}$ represents the *norm* of the function $f \in L^2([0,T])$.

For example, the function $f(t) := t^\alpha$, $t \in (0,T]$, belongs to $L^2([0,T])$ if and only if $\alpha > -1/2$, as we have

$$\int_0^T f^2(t) dt = \int_0^T t^{2\alpha} dt = \begin{cases} +\infty & \text{if } \alpha \leqslant -1/2, \\ \left[\dfrac{t^{1+2\alpha}}{1+2\alpha} \right]_{t=0}^{t=T} = \dfrac{T^{1+2\alpha}}{1+2\alpha} < \infty & \text{if } \alpha > -1/2, \end{cases}$$

see Figure 4.14 for an illustration.



(a) Infinite area, $\alpha = -1 < -1/2$. (b) Finite area, $\alpha = -1/4 > -1/2$.

Fig. 4.14: Infinite *vs.* finite area under the curve $t \mapsto t^{2\alpha}$.

In Lemma 4.7 we determine the probability distribution of $\int_0^T f(t)dB_t$ and we show that it is independent of the particular representation (4.5) chosen for $f(t)$.

Lemma 4.7. *Let f be a simple step function f of the form (4.5). The stochastic integral $\int_0^T f(t)dB_t$ defined in (4.6) has the centered Gaussian distribution*

$$\int_0^T f(t)dB_t \simeq \mathcal{N}\Big(0, \int_0^T |f(t)|^2 dt\Big)$$

with mean $\mathbb{E}\Big[\int_0^T f(t)dB_t\Big] = 0$ and variance given by the Itô isometry

$$\mathrm{Var}\left[\int_0^T f(t)dB_t\right] = \mathbb{E}\Big[\Big(\int_0^T f(t)dB_t\Big)^2\Big] = \int_0^T |f(t)|^2 dt. \tag{4.8}$$

Proof. Recall that if $X_1, X_2, \ldots, X_n$ are independent Gaussian random variables with probability distributions $\mathcal{N}(m_1, \sigma_1^2)$,…,$\mathcal{N}(m_n, \sigma_n^2)$, then the sum $X_1 + \cdots + X_n$ is a Gaussian random variable with distribution

$$\mathcal{N}\big(m_1 + \cdots + m_n, \sigma_1^2 + \cdots + \sigma_n^2\big).$$

As a consequence, the stochastic integral

$$\int_0^T f(t)dB_t = \sum_{k=1}^n a_k (B_{t_k} - B_{t_{k-1}})$$

of the step function

$$f(t) = \sum_{k=1}^n a_k \mathbb{1}_{(t_{k-1}, t_k]}(t), \qquad 0 \leqslant t \leqslant T,$$

has the centered Gaussian distribution with mean 0 and variance

$$
\begin{aligned}
\operatorname{Var}\left[\int_0^T f(t)dB_t\right] &= \operatorname{Var}\left[\sum_{k=1}^n a_k(B_{t_k}-B_{t_{k-1}})\right] \\
&= \sum_{k=1}^n \operatorname{Var}[a_k(B_{t_k}-B_{t_{k-1}})] \\
&= \sum_{k=1}^n |a_k|^2 \operatorname{Var}[B_{t_k}-B_{t_{k-1}}] \\
&= \sum_{k=1}^n (t_k-t_{k-1})|a_k|^2 \\
&= \sum_{k=1}^n |a_k|^2 \int_{t_{k-1}}^{t_k} dt \\
&= \sum_{k=1}^n |a_k|^2 \int_0^T \mathbb{1}_{(t_{k-1},t_k]}(t)dt \\
&= \int_0^T \sum_{k=1}^n |a_k|^2 \mathbb{1}_{(t_{k-1},t_k]}(t)dt \\
&= \int_0^T |f(t)|^2 dt,
\end{aligned}
$$

since the simple function

$$
f^2(t) = \sum_{i=1}^n a_i^2 \mathbb{1}_{(t_{i-1},t_i]}(t), \qquad 0 \leqslant t \leqslant T,
$$

takes the value a_i^2 on the interval $(t_{i-1},t_i]$, $i=1,2,\ldots,n$, as can be checked from the following Figure 4.15.

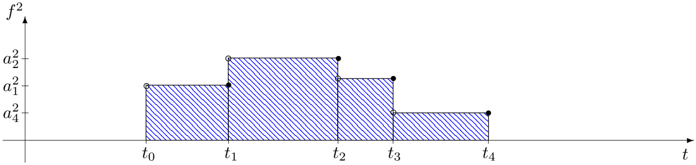


Fig. 4.15: Squared step function $t \mapsto f^2(t)$.

□

The norm $\|\cdot\|_{L^2([0,T])}$ on $L^2([0,T])$ induces a *distance* between any two functions f and g in $L^2([0,T])$, defined as

$$\|f - g\|_{L^2([0,T])} := \sqrt{\int_0^T |f(t) - g(t)|^2 dt} < \infty,$$

cf. *e.g.* Chapter 3 of Rudin (1974) for details.

Definition 4.8. Convergence in $L^2([0,T])$. *We say that a sequence $(f_n)_{n\geqslant 0}$ of functions in $L^2([0,T])$ converges in $L^2([0,T])$ to another function $f \in L^2([0,T])$ if*

$$\lim_{n\to\infty} \|f - f_n\|_{L^2([0,T])} = \lim_{n\to\infty} \sqrt{\int_0^T |f(t) - f_n(t)|^2 dt} = 0.$$

```
dev.new(width=16,height=7)
f = function(x){exp(sin(x*1.8*pi))}
for (i in 3:9){n=2^i;x<-cumsum(c(0,rep(1,n)))/n;
z<-c(NA,head(x,-1))
y<-c(f(x)-pmax(f(x)-f(z),0),f(1))
t=seq(0,1,0.01);
plot(f,from=0,to=1,ylim=c(0.3,2.9),type="l",lwd=3,col="red",main="",xaxs="i",yaxs="i",
      las=1)
lines(stepfun(x,y),do.points=F,lwd=2,col="blue",main="");
readline("Press <return> to continue");}
```

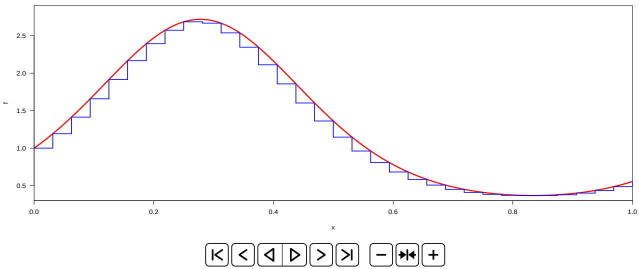


Fig. 4.16: Step function approximation.*

By *e.g.* Theorem 3.13 in Rudin (1974) or Proposition 2.4 page 63 of Hirsch and Lacombe (1999), we have the following result which states that the set of simple step functions f of the form (4.5) is a linear space which is dense in $L^2([0,T])$ for the norm (4.7), as stated in the next proposition.

Proposition 4.9. *For any function $f \in L^2([0,T])$ satisfying (4.7), there exists a sequence $(f_n)_{n\geqslant 0}$ of simple step functions of the form (4.5), converging to f in $L^2([0,T])$ in the sense that*

* The animation works in Acrobat Reader on the entire pdf file.

$$\lim_{n\to\infty} \|f - f_n\|_{L^2([0,T])} = \lim_{n\to\infty} \sqrt{\int_0^T |f(t) - f_n(t)|^2 dt} = 0.$$

In order to extend the definition (4.6) of the stochastic integral $\int_0^T f(t)dB_t$ to any function $f \in L^2([0,T])$, *i.e.* to $f : [0,T] \longrightarrow \mathbb{R}$ measurable such that

$$\int_0^T |f(t)|^2 dt < \infty, \tag{4.9}$$

we will make use of the space $L^2(\Omega)$ of *square-integrable random variables.*

Definition 4.10. *Let $L^2(\Omega)$ denote the space of random variables $F : \Omega \longrightarrow \mathbb{R}$ such that*

$$\|F\|_{L^2(\Omega)} := \sqrt{\mathbb{E}\big[F^2\big]} < \infty.$$

The norm $\|\cdot\|_{L^2(\Omega)}$ on $L^2(\Omega)$ induces the *distance*

$$\|F - G\|_{L^2(\Omega)} := \sqrt{\mathbb{E}\big[(F-G)^2\big]} < \infty,$$

between the square-integrable random variables F and G in $L^2(\Omega)$.

Definition 4.11. Convergence in $L^2(\Omega)$. *We say that a sequence $(F_n)_{n\geqslant 0}$ of random variables in $L^2(\Omega)$ converges in $L^2(\Omega)$ to another random variable $F \in L^2(\Omega)$ if*

$$\lim_{n\to\infty} \|F - F_n\|_{L^2(\Omega)} = \lim_{n\to\infty} \sqrt{\mathbb{E}\big[(F-F_n)^2\big]} = 0.$$

The next proposition allows us to extend Lemma 4.7 from simple step functions to square-integrable functions in $L^2([0,T])$.

Proposition 4.12. *The definition (4.6) of the stochastic integral $\int_0^T f(t)dB_t$ can be extended to any function $f \in L^2([0,T])$. In this case, $\int_0^T f(t)dB_t$ has the centered Gaussian distribution*

$$\int_0^T f(t)dB_t \simeq \mathcal{N}\left(0, \int_0^T |f(t)|^2 dt\right)$$

with mean $\mathbb{E}\left[\int_0^T f(t)dB_t\right] = 0$ and variance given by the Itô isometry

$$\mathrm{Var}\left[\int_0^T f(t)dB_t\right] = \mathbb{E}\left[\left(\int_0^T f(t)dB_t\right)^2\right] = \int_0^T |f(t)|^2 dt. \tag{4.10}$$

Proof. The extension of the stochastic integral to all functions satisfying (4.9) is obtained by a denseness and Cauchy* sequence argument, based on the isometry relation (4.10).

i) Given f a function satisfying (4.9), consider a sequence $(f_n)_{n\geqslant 0}$ of simple functions converging to f in $L^2([0,T])$, *i.e.*

$$\lim_{n\to\infty}\|f-f_n\|_{L^2([0,T])}=\lim_{n\to\infty}\sqrt{\int_0^T|f(t)-f_n(t)|^2dt}=0$$

as in Proposition 4.9.

ii) By the isometry relation (4.8) or (4.10) and the triangle inequality† we have

$$\begin{aligned}&\left\|\int_0^T f_k(t)dB_t-\int_0^T f_n(t)dB_t\right\|_{L^2(\Omega)}\\&\quad=\sqrt{\mathbb{E}\left[\left(\int_0^T f_k(t)dB_t-\int_0^T f_n(t)dB_t\right)^2\right]}\\&\quad=\sqrt{\mathbb{E}\left[\left(\int_0^T (f_k(t)-f_n(t))dB_t\right)^2\right]}\\&\quad=\sqrt{\int_0^T|f_k(t)-f_n(t)|^2dt}\\&\quad=\|f_k-f_n\|_{L^2([0,T])}\\&\quad\leqslant\|f_k-f\|_{L^2([0,T])}+\|f-f_n\|_{L^2([0,T])},\end{aligned}$$

which tends to 0 as k and n tend to infinity, hence $\left(\int_0^T f_n(t)dB_t\right)_{n\geqslant 0}$ is a Cauchy sequence in $L^2(\Omega)$ for the $L^2(\Omega)$-norm.

iii) Since the sequence $\left(\int_0^T f_n(t)dB_t\right)_{n\geqslant 0}$ is Cauchy and the space $L^2(\Omega)$ is *complete*, cf. *e.g.* Theorem 3.11 in Rudin (1974) or Chapter 4 of Dudley (2002), we conclude that $\left(\int_0^T f_n(t)dB_t\right)_{n\geqslant 0}$ *converges* for the L^2-norm to a limit in $L^2(\Omega)$. In this case we let

$$\int_0^T f(t)dB_t:=\lim_{n\to\infty}\int_0^T f_n(t)dB_t,$$

which also satisfies (4.10) from (4.8). From (4.10) we can check that the limit is independent of the approximating sequence $(f_n)_{n\geqslant 0}$.

---

* See MH3100 Real Analysis I.

† The triangle inequality $\|f_k-f_n\|_{L^2([0,T])}\leqslant\|f_k-f\|_{L^2([0,T])}+\|f-f_n\|_{L^2([0,T])}$ follows from the Minkowski inequality.

iv) Finally, from the convergence of Gaussian characteristic functions and a dominated convergence argument, we have

$$\begin{aligned}\mathbb{E}\left[\exp\left(i\alpha\int_0^T f(t)dB_t\right)\right] &= \mathbb{E}\left[\lim_{n\to\infty}\exp\left(i\alpha\int_0^T f_n(t)dB_t\right)\right]\\ &= \lim_{n\to\infty}\mathbb{E}\left[\exp\left(i\alpha\int_0^T f_n(t)dB_t\right)\right]\\ &= \lim_{n\to\infty}\exp\left(-\frac{\alpha^2}{2}\int_0^T |f_n(t)|^2dt\right)\\ &= \exp\left(-\frac{\alpha^2}{2}\int_0^T |f(t)|^2dt\right),\end{aligned}$$

$f\in L^2([0,T])$, $\alpha\in\mathbb{R}$, we check that $\int_0^T f(t)dB_t$ has the centered Gaussian distribution

$$\int_0^T f(t)dB_t \simeq \mathcal{N}\left(0,\int_0^T |f(t)|^2dt\right),$$

see Theorem A.13.

□

The next corollary is obtained by bilinearity from the Itô isometry (4.10).

Corollary 4.13. *The stochastic integral with respect to Brownian motion $(B_t)_{t\in\mathbb{R}_+}$ satisfies the isometry*

$$\mathbb{E}\left[\int_0^T f(t)dB_t\int_0^T g(t)dB_t\right] = \int_0^T f(t)g(t)dt,$$

for all square-integrable deterministic functions $f,g\in L^2([0,T])$.

Proof. Applying the Itô isometry (4.10) to the processes $f+g$ and $f-g$ and the relation $xy=(x+y)^2/4-(x-y)^2/4$, we have

$$\begin{aligned}&\mathbb{E}\left[\int_0^T f(t)dB_t\int_0^T g(t)dB_t\right]\\ &= \frac{1}{4}\mathbb{E}\left[\left(\int_0^T f(t)dB_t+\int_0^T g(t)dB_t\right)^2-\left(\int_0^T f(t)dB_t-\int_0^T g(t)dB_t\right)^2\right]\\ &= \frac{1}{4}\mathbb{E}\left[\left(\int_0^T (f(t)+g(t))dB_t\right)^2\right]-\frac{1}{4}\mathbb{E}\left[\left(\int_0^T (f(t)-g(t))dB_t\right)^2\right]\\ &= \frac{1}{4}\int_0^T (f(t)+g(t))^2dt-\frac{1}{4}\int_0^T (f(t)-g(t))^2dt\\ &= \frac{1}{4}\int_0^T \left((f(t)+g(t))^2-(f(t)-g(t))^2\right)dt\end{aligned}$$

$$= \int_0^T f(t)g(t)dt.$$

□

For example, the Wiener stochastic integral $\int_0^T \mathrm{e}^{-t} dB_t$ is a random variable having centered Gaussian distribution with variance

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^T \mathrm{e}^{-t} dB_t\right)^2\right] &= \int_0^T \mathrm{e}^{-2t} dt \\ &= \left[-\frac{1}{2}\mathrm{e}^{-2t}\right]_{t=0}^{t=T} \\ &= \frac{1}{2}(1-\mathrm{e}^{-2T}),\end{aligned}$$

as follows from the Itô isometry (4.8).

Remark 4.14. *The Wiener stochastic integral $\int_0^T f(s)dB_s$ is a Gaussian random variable that cannot be "computed" in the way standard integrals are computed via the use of primitives. However, when $f \in L^2([0,T])$ is in $\mathcal{C}^1([0,T])$,* we have the integration by parts relation*

$$\int_0^T f(t)dB_t = f(T)B_T - \int_0^T B_t df(t) = f(T)B_T - \int_0^T B_t f'(t)dt. \quad (4.11)$$

When $f \in L^2(\mathbb{R}_+)$ is in $\mathcal{C}^1(\mathbb{R}_+)$ we also have following formula

$$\int_0^\infty f(t)dB_t = -\int_0^\infty B_t f'(t)dt, \quad (4.12)$$

provided that $\lim_{t\to\infty} t|f(t)|^2 = 0$ and $f \in L^2(\mathbb{R}_+)$, cf. e.g. *Exercise 4.5 and Remark 2.5.9 in Privault (2009).*

For example, applying Relation (4.11) to the function $f(t) = t$ shows that

$$\int_0^T t dB_t = TB_T - \int_0^T B_t dt = T\int_0^T dB_t - \int_0^T B_t dt,$$

hence

$$\int_0^T (T-t)dB_t = \int_0^T B_t dt.$$

* This means that the function f is continuously differentiable on $[0,T]$.

4.4 Itô Stochastic Integral

In this section we extend the Wiener stochastic integral from deterministic functions in $L^2([0,T])$ to random square-integrable (random) *adapted* processes. For this, we will need the notion of *measurability*.

The extension of the stochastic integral to adapted random processes is actually necessary in order to compute a portfolio value when the portfolio process is no longer deterministic. This happens in particular when one needs to update the portfolio allocation based on random events occurring on the market.

A random variable F is said to be $\mathcal{F}_t$*-measurable* if the knowledge of F depends only on the information known up to time t. As an example, if $t =$today,

- the date of the past course exam is $\mathcal{F}_t$-measurable, because it belongs to the past.
- the date of the upcoming course exam, although it refers to a future event, is also $\mathcal{F}_t$-measurable because it is known at time t.
- the date of the next typhoon is not $\mathcal{F}_t$-measurable since it is not known at time t.
- the maturity date T of the European option is $\mathcal{F}_t$-measurable for all $t \in [0,T]$, because it has been determined at time 0.
- the exercise date τ of an American option after time t (see Section 15.1) is not $\mathcal{F}_t$-measurable because it refers to a future random event.

In Definition 4.15, $(\mathcal{F}_t)_{t\in[0,T]}$ denotes the information flow defined in (4.2), *i.e.*

$$\mathcal{F}_t := \sigma(B_s \ : \ 0 \leqslant s \leqslant t), \qquad t \geqslant 0.$$

Definition 4.15. *A stochastic process $(X_t)_{t\in[0,T]}$ is said to be $(\mathcal{F}_t)_{t\in[0,T]}$-adapted if X_t is $\mathcal{F}_t$-measurable for all $t \in [0,T]$.*

For example,

- $(B_t)_{t\in\mathbb{R}_+}$ is an $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-adapted process,
- $(B_{t+1})_{t\in\mathbb{R}_+}$ is *not* an $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-adapted process,
- $(B_{t/2})_{t\in\mathbb{R}_+}$ is an $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-adapted process,
- $(B_{\sqrt{t}})_{t\in[0,1]}$ is *not* an $(\mathcal{F}_t)_{t\in[0,1]}$-adapted process,
- $(B_{\sqrt{t}})_{t\in[1,\infty)}$ is an $(\mathcal{F}_t)_{t\in[1,\infty)}$-adapted process,

This version: May 3, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

- $\left(\operatorname{Max}_{s\in[0,t]} B_s\right)_{t\in\mathbb{R}_+}$ is an $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-adapted process,

- $\left(\int_0^t B_s ds\right)_{t\in\mathbb{R}_+}$ is an $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-adapted process,

- $\left(\int_0^t f(s) dB_s\right)_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process when $f \in L^2([0,T])$.

In other words, a stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted if the value of X_t at time t depends only on information known up to time t. Note that the value of X_t may still depend on "known" future data, for example a fixed future date in the calendar, such as a maturity time $T > t$, as long as its value is known at time t.

The next Figure 4.17 shows an adapted portfolio strategy on two assets, constructed from a sign-switching signal based on spread data, see § 2.5 in Privault (2021a) and this **code**.

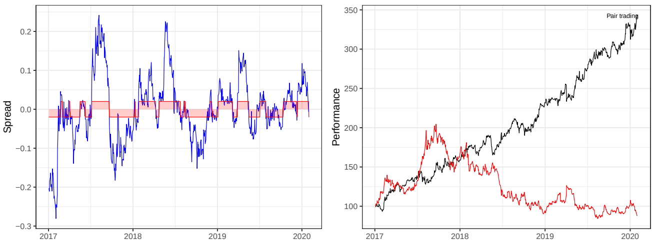


Fig. 4.17: Adapted pair trading portfolio strategy.

The stochastic integral of adapted processes is first constructed as integrals of simple predictable processes.

Definition 4.16. *A simple predictable processes is a stochastic process* $(u_t)_{t\in\mathbb{R}_+}$ *of the form*

$$u_t := \sum_{i=1}^n F_i \mathbb{1}_{(t_{i-1},t_i]}(t), \qquad t \geqslant 0, \tag{4.13}$$

where F_i *is an* $\mathcal{F}_{t_{i-1}}$*-measurable random variable for* $i = 1, 2, \ldots, n$, *and* $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$.

The notion of simple predictable process makes full sense in the context of portfolio investment, in which F_i will represent an investment allocation decided at time t_{i-1} and to remain unchanged over the time interval $(t_{i-1}, t_i]$.

By convention, $u : \Omega \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ is denoted in what follows by $u_t(\omega)$, $t \in \mathbb{R}_+$, $\omega \in \Omega$, and the random outcome ω is often dropped for convenience of notation.

Definition 4.17. *The stochastic integral with respect to Brownian motion $(B_t)_{t\in\mathbb{R}_+}$ of any simple predictable process $(u_t)_{t\in\mathbb{R}_+}$ of the form (4.13) is defined by*

$$\int_0^T u_t dB_t := \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}}) F_i, \tag{4.14}$$

with $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$.

The use of predictability in the definition (4.14) is essential from a financial point of view, as F_i will represent a portfolio allocation made at time t_{i-1} and kept constant over the trading interval $[t_{i-1}, t_i]$, while $B_{t_i} - B_{t_{i-1}}$ represents a change in the underlying asset price over $[t_{i-1}, t_i]$. See also the related discussion on self-financing portfolios in Section 5.3 and Lemma 5.14 on the use of stochastic integrals to represent the value of a portfolio.

Definition 4.18. *Let $L^2(\Omega \times [0,T])$ denote the space of stochastic processes*

$$\begin{aligned} u : \Omega \times [0,T] &\longrightarrow \mathbb{R} \\ (\omega, t) &\longmapsto u_t(\omega) \end{aligned}$$

such that

$$\|u\|_{L^2(\Omega\times[0,T])} := \sqrt{\mathbb{E}\left[\int_0^T |u_t|^2 dt\right]} < \infty, \quad u \in L^2(\Omega \times [0,T]).$$

The norm $\|\cdot\|_{L^2(\Omega\times[0,T])}$ on $L^2(\Omega \times [0,T])$ induces a *distance* between two stochastic processes u and v in $L^2(\Omega \times [0,T])$, defined as

$$\|u - v\|_{L^2(\Omega\times[0,T])} = \sqrt{\mathbb{E}\left[\int_0^T |u_t - v_t|^2 dt\right]}.$$

Definition 4.19. Convergence in $L^2(\Omega \times [0,T])$. *We say that a sequence $\big(u^{(n)}\big)_{n\geqslant 0}$ of processes in $L^2(\Omega \times [0,T])$ converges in $L^2(\Omega \times [0,T])$ to another process $u \in L^2(\Omega \times [0,T])$ if*

$$\lim_{n\to\infty} \big\|u - u^{(n)}\big\|_{L^2(\Omega\times[0,T])} = \lim_{n\to\infty} \sqrt{\mathbb{E}\left[\int_0^T \big|u_t - u_t^{(n)}\big|^2 dt\right]} = 0.$$

By Lemma 1.1 of Ikeda and Watanabe (1989), pages 22 and 46, or Proposition 2.5.3 in Privault (2009), the set of simple predictable processes forms a linear space which is dense in the subspace $L^2_{\mathrm{ad}}(\Omega \times \mathbb{R}_+)$ made of square-integrable adapted processes in $L^2(\Omega \times \mathbb{R}_+)$, as stated in the next proposition.

Proposition 4.20. *Given $u \in L^2_{\rm ad}(\Omega \times \mathbb{R}_+)$ a square-integrable adapted process there exists a sequence $(u^{(n)})_{n\geqslant 0}$ of simple predictable processes converging to u in $L^2(\Omega \times \mathbb{R}_+)$, i.e.*

$$\lim_{n\to\infty} \big\| u - u^{(n)} \big\|_{L^2(\Omega\times[0,T])} = \lim_{n\to\infty} \sqrt{\mathbb{E}\left[\int_0^T \big| u_t - u_t^{(n)} \big|^2 dt \right]} = 0.$$

For example, a natural approximation of $(B_t)_{t\in\mathbb{R}_+}$ by a simple predictable process can be constructed as

$$u_t = \sum_{i=1}^n F_i \mathbb{1}_{(t_{i-1},t_i]}(t) := \sum_{i=1}^n B_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(t), \qquad t \geqslant 0, \tag{4.15}$$

where $F_i := B_{t_{i-1}}$ is $\mathcal{F}_{t_{i-1}}$-measurable for $i = 1, 2, \ldots, n$, as in Figure 4.18.

```
N=10000; t <- 0:(N-1); dt <- 1.0/N;
dB <- rnorm(N,mean=0,sd=sqrt(dt));X <- rep(0,N);X[1]=0
for (j in 2:N){X[j]=X[j-1]+dB[j]}; for (j in 1:10) {m=2**j;
plot(t/(N-1), X, xlab = "t", ylab = "", type = "l", ylim = c(1.05*min(X),1.05*max(X)),
      xaxs="i", yaxs="i", col = "blue", las = 1, cex.axis=1.6, cex.lab=1.8)
abline(h=0); t1=seq(1.0/m,1,1.0/m); Bt=c(0)
for (i in 1:m) {Bt=c(Bt,X[t1[i]*N])}
lines(stepfun(t1,Bt),xlim =c(0,T),xlab="t",ylab=expression('N'[t]),pch=1, cex=0.8, col='black',
      lwd=2, main=""); Sys.sleep(1)}
```

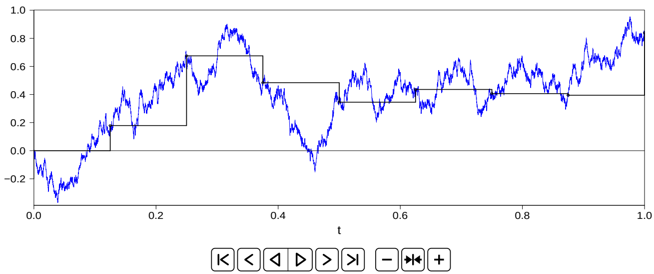


Fig. 4.18: Step function approximation of Brownian motion.*

The next Proposition 4.21 extends the construction of the stochastic integral from simple predictable processes to square-integrable $(\mathcal{F}_t)_{t\in[0,T]}$-adapted processes $(u_t)_{t\in\mathbb{R}_+}$ for which the value of u_t at time t can only depend on information contained in the Brownian path up to time t.

* The animation works in Acrobat Reader on the entire pdf file.

This restriction means that the Itô integrand u_t cannot depend on future information, for example a portfolio strategy that would allow the trader to "buy at the lowest" and "sell at the highest" is excluded as it would require knowledge of future market data. Note that the difference between Relation (4.16) below and Relation (4.10) is the presence of an expectation on the right-hand side.

Proposition 4.21. *The stochastic integral with respect to Brownian motion $(B_t)_{t\in\mathbb{R}_+}$ extends to all adapted processes $(u_t)_{t\in\mathbb{R}_+}$ such that*

$$\|u\|^2_{L^2(\Omega\times[0,T])} := \mathbb{E}\left[\int_0^T |u_t|^2 dt\right] < \infty,$$

with the Itô isometry

$$\mathrm{Var}\left[\int_0^T u_t dB_t\right] = \mathbb{E}\left[\left(\int_0^T u_t dB_t\right)^2\right] = \left\|\int_0^T u_t dB_t\right\|^2_{L^2(\Omega)} = \mathbb{E}\left[\int_0^T |u_t|^2 dt\right]. \tag{4.16}$$

In addition, the Itô integral of an adapted process $(u_t)_{t\in\mathbb{R}_+}$ is always a centered random variable:

$$\mathbb{E}\left[\int_0^T u_t dB_t\right] = 0. \tag{4.17}$$

Proof. We start by showing that the Itô isometry (4.16) holds for the simple predictable process u of the form (4.13). We have

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T u_t dB_t\right)^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})F_i\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})F_i\right)\left(\sum_{j=1}^n (B_{t_j} - B_{t_{j-1}})F_j\right)\right] \\
&= \mathbb{E}\left[\sum_{i,j=1}^n (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})F_iF_j\right] \\
&= \mathbb{E}\left[\sum_{i=1}^n |F_i|^2 (B_{t_i} - B_{t_{i-1}})^2\right] \\
&\quad +2\mathbb{E}\left[\sum_{1\leqslant i<j\leqslant n} (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})F_iF_j\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[|F_i|^2 (B_{t_i} - B_{t_{i-1}})^2\right]
\end{aligned}$$

$$
\begin{aligned}
&\quad +2 \sum_{1 \leqslant i<j \leqslant n} \mathbb{E}\left[(B_{t_i}-B_{t_{i-1}})(B_{t_j}-B_{t_{j-1}}) F_i F_j\right] \\
&= \sum_{i=1}^n \mathbb{E}[\mathbb{E}[|F_i|^2 (B_{t_i}-B_{t_{i-1}})^2 | \mathcal{F}_{t_{i-1}}]] \\
&\quad +2 \sum_{1 \leqslant i<j \leqslant n} \mathbb{E}[\mathbb{E}[(B_{t_i}-B_{t_{i-1}})(B_{t_j}-B_{t_{j-1}}) F_i F_j | \mathcal{F}_{t_{j-1}}]] \\
&= \sum_{i=1}^n \mathbb{E}[|F_i|^2 \mathbb{E}[(B_{t_i}-B_{t_{i-1}})^2 | \mathcal{F}_{t_{i-1}}]] \\
&\quad +2 \sum_{1 \leqslant i<j \leqslant n} \mathbb{E}\big[(B_{t_i}-B_{t_{i-1}}) F_i F_j \underbrace{\mathbb{E}[B_{t_j}-B_{t_{j-1}} \mid \mathcal{F}_{t_{j-1}}]}_{=0}\big] \\
&= \sum_{i=1}^n \mathbb{E}\big[|F_i|^2 \mathbb{E}\big[(B_{t_i}-B_{t_{i-1}})^2\big]\big] \\
&\quad +2 \sum_{1 \leqslant i<j \leqslant n} \mathbb{E}[(B_{t_i}-B_{t_{i-1}}) F_i F_j \underbrace{\mathbb{E}[B_{t_j}-B_{t_{j-1}}]}_{=0}] \\
&= \sum_{i=1}^n \mathbb{E}[|F_i|^2 (t_i - t_{i-1})] \\
&= \mathbb{E}\left[\sum_{i=1}^n |F_i|^2 (t_i - t_{i-1})\right] \\
&= \mathbb{E}\Big[\int_0^T |u_t|^2 dt\Big],
\end{aligned}
$$

where we applied the tower property (A.33) of conditional expectations and the facts that $B_{t_i} - B_{t_{i-1}}$ is independent of $\mathcal{F}_{t_{i-1}}$, with

$$
\mathbb{E}[B_{t_i} - B_{t_{i-1}}] = 0, \quad \mathbb{E}\left[(B_{t_i} - B_{t_{i-1}})^2\right] = t_i - t_{i-1}, \quad i = 1, 2, \ldots, n.
$$

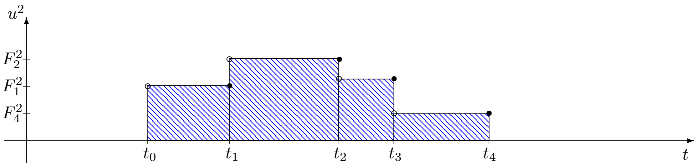


Fig. 4.19: Squared simple predictable process $t \mapsto u_t^2$.

The extension of the stochastic integral to square-integrable adapted processes $(u_t)_{t\in\mathbb{R}_+}$ is obtained by a denseness and Cauchy sequence argument using the isometry (4.16), in the same way as in the proof of Proposition 4.12.

i) By Proposition 4.20 given $u \in L^2(\Omega \times [0,T])$ a square-integrable adapted process there exists a sequence $(u^{(n)})_{n\geqslant 0}$ of simple predictable processes such that

$$\lim_{n\to\infty} \|u - u^{(n)}\|_{L^2(\Omega\times[0,T])} = \lim_{n\to\infty} \sqrt{\mathbb{E}\Big[\int_0^T \big|u_t - u_t^{(n)}\big|^2 dt\Big]} = 0.$$

ii) Since the sequence $(u^{(n)})_{n\geqslant 0}$ converges, it is a Cauchy sequence in $L^2(\Omega \times \mathbb{R}_+)$, hence by the Itô isometry (4.16), the sequence $\left(\int_0^T u_t^{(n)} dB_t\right)_{n\geqslant 0}$ is a Cauchy sequence in $L^2(\Omega)$, therefore it admits a limit in the complete space $L^2(\Omega)$. In this case we let

$$\int_0^T u_t dB_t := \lim_{n\to\infty} \int_0^T u_t^{(n)} dB_t$$

and the limit is unique from (4.16) and satisfies (4.16).

iii) The fact that the random variable $\int_0^T u_t dB_t$ is *centered* can be proved first for a simple predictable process $u^{(n)}$ of the form (4.13), as

$$\begin{aligned}
\mathbb{E}\left[\int_0^T u_t^{(n)} dB_t\right] &= \mathbb{E}\left[\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}}) F_i\right] \\
&= \sum_{i=1}^n \mathbb{E}[\mathbb{E}[(B_{t_i} - B_{t_{i-1}}) F_i \mid \mathcal{F}_{t_{i-1}}]] \\
&= \sum_{i=1}^n \mathbb{E}[F_i \mathbb{E}[B_{t_i} - B_{t_{i-1}} \mid \mathcal{F}_{t_{i-1}}]] \\
&= \sum_{i=1}^n \mathbb{E}[F_i \mathbb{E}[B_{t_i} - B_{t_{i-1}}]] \\
&= 0,
\end{aligned}$$

and this identity extends as above from simple predictable processes to adapted processes $(u_t)_{t\in\mathbb{R}_+}$ in $L^2(\Omega \times \mathbb{R}_+)$ by taking the limit as n tends to infinity in the following equality:

$$\mathbb{E}\Big[\int_0^T u_t dt\Big] = \mathbb{E}\Big[\int_0^T u_t^{(n)} dt\Big] + \mathbb{E}\Big[\int_0^T u_t - u_t^{(n)} dt\Big] = \mathbb{E}\Big[\int_0^T u_t - u_t^{(n)} dt\Big],$$

since

$$\left|\mathbb{E}\Big[\int_0^T (u_t - u_t^{(n)}) dt\Big]\right| \leqslant \mathbb{E}\Big[\int_0^T \big|u_t - u_t^{(n)}\big| dt\Big] \leqslant \sqrt{T\mathbb{E}\Big[\int_0^T \big|u_t - u_t^{(n)}\big|^2 dt\Big]}.$$

The Itô isometry (4.16) can be similarly extended from simple predictable processes to adapted processes $(u_t)_{t\in\mathbb{R}_+}$ in $L^2(\Omega \times \mathbb{R}_+)$.

This version: May 3, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

□

As an application of the Itô isometry (4.16), we note in particular the identity

$$\mathbb{E}\left[\left(\int_0^T B_t dB_t\right)^2\right] = \mathbb{E}\left[\int_0^T |B_t|^2 dt\right] = \int_0^T \mathbb{E}\left[|B_t|^2\right] dt = \int_0^T t dt = \frac{T^2}{2},$$

with

$$\int_0^T B_t dB_t \overset{L^2(\Omega)}{=} \lim_{n\to\infty} \sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})$$

from (4.15).

The next corollary is obtained by bilinearity from the Itô isometry (4.16) by the same argument as in Corollary 4.13.

Corollary 4.22. *The stochastic integral with respect to Brownian motion* $(B_t)_{t\in\mathbb{R}_+}$ *satisfies the isometry*

$$\mathbb{E}\left[\int_0^T u_t dB_t \int_0^T v_t dB_t\right] = \mathbb{E}\left[\int_0^T u_t v_t dt\right],$$

for all square-integrable adapted processes $(u_t)_{t\in\mathbb{R}_+}$, $(v_t)_{t\in\mathbb{R}_+}$.

Proof. Applying the Itô isometry (4.16) to the processes $u+v$ and $u-v$, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T u_t dB_t \int_0^T v_t dB_t\right] \\
&= \frac{1}{4}\left(\mathbb{E}\left[\left(\int_0^T u_t dB_t + \int_0^T v_t dB_t\right)^2 - \left(\int_0^T u_t dB_t - \int_0^T v_t dB_t\right)^2\right]\right) \\
&= \frac{1}{4}\left(\mathbb{E}\left[\left(\int_0^T (u_t+v_t) dB_t\right)^2\right] - \mathbb{E}\left[\left(\int_0^T (u_t-v_t) dB_t\right)^2\right]\right) \\
&= \frac{1}{4}\left(\mathbb{E}\left[\int_0^T (u_t+v_t)^2 dt\right] - \mathbb{E}\left[\int_0^T (u_t-v_t)^2 dt\right]\right) \\
&= \frac{1}{4}\mathbb{E}\left[\int_0^T \left((u_t+v_t)^2 - (u_t-v_t)^2\right) dt\right] \\
&= \mathbb{E}\left[\int_0^T u_t v_t dt\right].
\end{aligned}$$

□

In addition, when the integrand $(u_t)_{t\in\mathbb{R}_+}$ is not a deterministic function of time, the random variable $\int_0^T u_t dB_t$ no longer has a Gaussian distribution, except in some exceptional cases.

Definite stochastic integral

The definite stochastic integral of an adapted process $u \in L^2_{ad}(\Omega \times \mathbb{R}_+)$ over an interval $[a,b] \subset [0,T]$ is defined as

$$\int_a^b u_t dB_t := \int_0^T \mathbb{1}_{[a,b]}(t) u_t dB_t,$$

with in particular

$$\int_a^b dB_t = \int_0^T \mathbb{1}_{[a,b]}(t) dB_t = B_b - B_a, \quad 0 \leqslant a \leqslant b,$$

We also have the Chasles relation

$$\int_a^c u_t dB_t = \int_a^b u_t dB_t + \int_b^c u_t dB_t, \qquad 0 \leqslant a \leqslant b \leqslant c,$$

and the stochastic integral has the following linearity property:

$$\int_0^T (u_t + v_t) dB_t = \int_0^T u_t dB_t + \int_0^T v_t dB_t, \qquad u, v \in L^2(\mathbb{R}_+).$$

4.5 Stochastic Calculus

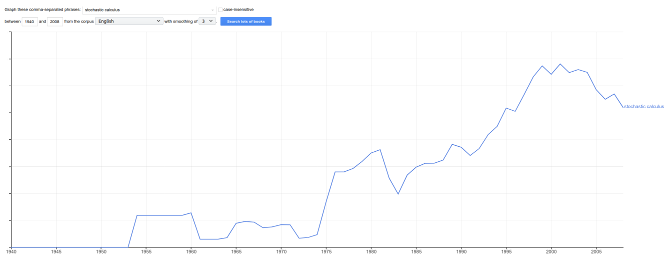


Fig. 4.20: NGram Viewer output for the term "stochastic calculus".

Stochastic modeling of asset returns

In the sequel, we consider the return at time $t \in \mathbb{R}_+$ of the risky asset price process $(S_t)_{t \in \mathbb{R}_+}$, defined as

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t, \quad \text{or} \quad dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{4.18}$$

with $\mu \in \mathbb{R}$ and $\sigma > 0$. Using the relation

$$X_T = X_0 + \int_0^T dX_t, \qquad T > 0,$$

which holds for any process $(X_t)_{t\in\mathbb{R}_+}$, Equation (4.18) can be rewritten in integral form as

$$S_T = S_0 + \int_0^T dS_t = S_0 + \mu \int_0^T S_t dt + \sigma \int_0^T S_t dB_t, \tag{4.19}$$

hence the need to define an integral with respect to dB_t, in addition to the usual integral with respect to dt. Note that in view of the definition (4.14), this is a continuous-time extension of the notion portfolio value based on a predictable portfolio strategy.

In Proposition 4.21 we have defined the stochastic integral of square-integrable processes with respect to Brownian motion, thus we have made sense of the equation (4.19), where $(S_t)_{t\in\mathbb{R}_+}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process, which can be rewritten in differential notation as in (4.18).

This model will be used to represent the random price S_t of a risky asset at time t. Here the return dS_t/S_t of the asset is made of two components: a constant return μdt and a random return σdB_t parametrized by the coefficient σ, called the volatility.

Our goal is now to solve Equation (4.18), and for this we will need to introduce Itô's calculus in Section 4.5 after a review of classical deterministic calculus.

Deterministic calculus

The *fundamental theorem of calculus* states that for any continuously differentiable (deterministic) function f we have the integral relation

$$f(x) = f(0) + \int_0^x f'(y)dy.$$

In differential notation this relation is written as the first-order expansion

$$df(x) = f'(x)dx, \tag{4.20}$$

where dx is "infinitesimally small". Higher-order expansions can be obtained from *Taylor's formula*, which, letting

$$\Delta f(x) := f(x + \Delta x) - f(x),$$

states that

$$\Delta f(x) = f'(x)\Delta x + \frac{1}{2} f''(x)(\Delta x)^2 + \frac{1}{3!} f'''(x)(\Delta x)^3 + \frac{1}{4!} f^{(4)}(x)(\Delta x)^4 + \cdots .$$

Note that Relation (4.20), *i.e.* $df(x) = f'(x)dx$, can be obtained by neglecting all terms of order higher than one in Taylor's formula, since $(\Delta x)^n << \Delta x$, $n \geqslant 2$, as Δx becomes "infinitesimally small".

Stochastic calculus

Let us now apply Taylor's formula to Brownian motion, taking

$$\Delta B_t = B_{t+\Delta t} - B_t \simeq \pm\sqrt{\Delta t},$$

and letting

$$\Delta f(B_t) := f(B_{t+\Delta t}) - f(B_t),$$

we have

$$\begin{aligned}
&\Delta f(B_t) \\
&= f'(B_t)\Delta B_t + \frac{1}{2}f''(B_t)(\Delta B_t)^2 + \frac{1}{3!}f'''(B_t)(\Delta B_t)^3 + \frac{1}{4!}f^{(4)}(B_t)(\Delta B_t)^4 + \cdots .
\end{aligned}$$

From the construction of Brownian motion by its small increments $\Delta B_t = \pm\sqrt{\Delta t}$, it turns out that the terms in $(\Delta t)^2$ and $\Delta t \Delta B_t \simeq \pm(\Delta t)^{3/2}$ can be neglected in Taylor's formula at the first order of approximation in Δt. However, the term of order two

$$(\Delta B_t)^2 = (\pm\sqrt{\Delta t})^2 = \Delta t$$

can no longer be neglected in front of Δt itself.

Basic Itô formula

For $f \in \mathcal{C}^2(\mathbb{R})$,* Taylor's formula written at the second order for Brownian motion reads

$$\boxed{df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt,} \qquad (4.21)$$

for "infinitesimally small" dt. Note that writing this formula as

$$\frac{df(B_t)}{dt} = f'(B_t)\frac{dB_t}{dt} + \frac{1}{2}f''(B_t)$$

does not make sense because the pathwise derivative

$$\frac{dB_t}{dt} \simeq \pm\frac{\sqrt{dt}}{dt} \simeq \pm\frac{1}{\sqrt{dt}} \simeq \pm\infty$$

* This means that f is twice continuously differentiable on $[0, T]$.

of B_t with respect to t does not exist. Integrating (4.21) on both sides and using the relation

$$f(B_t) - f(B_0) = \int_0^t df(B_s)$$

together with (4.21), we get the integral form of Itô's formula for Brownian motion, *i.e.*

$$\boxed{f(B_t) = f(B_0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds.}$$

Itô processes

We now turn to the general expression of Itô's formula, which is stated for Itô processes.

Definition 4.23. *An Itô process is a stochastic process $(X_t)_{t\in\mathbb{R}_+}$ that can be written as*

$$X_t = X_0 + \int_0^t v_s ds + \int_0^t u_s dB_s, \qquad t \geqslant 0, \tag{4.22}$$

or in differential notation

$$dX_t = v_t dt + u_t dB_t,$$

where $(u_t)_{t\in\mathbb{R}_+}$ and $(v_t)_{t\in\mathbb{R}_+}$ are square-integrable adapted processes.

In what follows, we let $\mathcal{C}_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ denote the set of functions $f(t,x)$ of two variables which are continuously differentiable on $t \in \mathbb{R}_+$ and twice differentiable in $x \in \mathbb{R}$, with bounded derivatives. Given $f \in \mathcal{C}_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$, we let $\dfrac{\partial f}{\partial t}$ denote partial differentiation with respect to the *first* (time) variable in $f(t,x)$, while $\dfrac{\partial f}{\partial x}$ denotes partial differentiation with respect to the *second* (price) variable in $f(t,x)$.

Theorem 4.24. *(Itô formula for Itô processes). For any Itô process $(X_t)_{t\in\mathbb{R}_+}$ of the form (4.22) and any $f \in \mathcal{C}_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$, we have*

$$\begin{aligned} &f(t,X_t) \\ &= f(0,X_0) + \int_0^t \frac{\partial f}{\partial s}(s,X_s)ds + \int_0^t v_s \frac{\partial f}{\partial x}(s,X_s)ds + \int_0^t u_s \frac{\partial f}{\partial x}(s,X_s)dB_s \\ &\quad + \frac{1}{2}\int_0^t |u_s|^2 \frac{\partial^2 f}{\partial x^2}(s,X_s)ds. \end{aligned} \tag{4.23}$$

Proof. The proof of the Itô formula can be outlined as follows in the case where $(X_t)_{t\in\mathbb{R}_+} = (B_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion and $f(x)$ does not depend on time t. We refer to Theorem II-32, page 79 of Protter (2004) for the general case.

Let $\{0 = t_0^n \leqslant t_1^n \leqslant \cdots \leqslant t_n^n = t\}$, $n \geqslant 1$, be a refining sequence of partitions of $[0,t]$ tending to the identity. We have the telescoping identity

$$f(B_t) - f(B_0) = \sum_{k=1}^n \big(f(B_{t_i^n}) - f(B_{t_{i-1}^n}) \big),$$

and from Taylor's formula

$$f(y) - f(x) = (y-x)\frac{\partial f}{\partial x}(x) + \frac{1}{2}(y-x)^2\frac{\partial^2 f}{\partial x^2}(x) + R(x,y),$$

where the remainder $R(x,y)$ satisfies $R(x,y) \leqslant o(|y-x|^2)$, we get

$$\begin{aligned} f(B_t) - f(B_0) = & \sum_{k=1}^n (B_{t_i^n} - B_{t_{i-1}^n})\frac{\partial f}{\partial x}(B_{t_{i-1}^n}) + \frac{1}{2}\sum_{k=1}^n |B_{t_i^n} - B_{t_{i-1}^n}|^2\frac{\partial^2 f}{\partial x^2}(B_{t_{i-1}^n}) \\ & + \sum_{k=1}^n R(B_{t_i^n}, B_{t_{i-1}^n}). \end{aligned}$$

It remains to show that as n tends to infinity the above converges to

$$f(B_t) - f(B_0) = \int_0^t \frac{\partial f}{\partial x}(B_s)dB_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(B_s)ds.$$

□

From the relation

$$\int_0^t df(s, X_s) = f(t, X_t) - f(0, X_0),$$

we can rewrite (4.23) as

$$\begin{aligned} \int_0^t df(s, X_s) = & \int_0^t \frac{\partial f}{\partial s}(s, X_s)ds + \int_0^t v_s\frac{\partial f}{\partial x}(s, X_s)ds + \int_0^t u_s\frac{\partial f}{\partial x}(s, X_s)dB_s \\ & + \frac{1}{2}\int_0^t |u_s|^2\frac{\partial^2 f}{\partial x^2}(s, X_s)ds, \end{aligned}$$

which allows us to rewrite (4.23) in differential notation, as

$$
\begin{aligned}
&df(t, X_t) \qquad\qquad (4.24)\\
&= \frac{\partial f}{\partial t}(t, X_t)dt + v_t \frac{\partial f}{\partial x}(t, X_t)dt + u_t \frac{\partial f}{\partial x}(t, X_t)dB_t + \frac{1}{2}|u_t|^2 \frac{\partial^2 f}{\partial x^2}(t, X_t)dt.
\end{aligned}
$$

In case the function $x \mapsto f(x)$ does not depend on the time variable t we get

$$
df(X_t) = v_t \frac{\partial f}{\partial x}(X_t)dt + u_t \frac{\partial f}{\partial x}(X_t)dB_t + \frac{1}{2}|u_t|^2 \frac{\partial^2 f}{\partial x^2}(X_t)dt.
$$

Taking $u_t = 1$, $v_t = 0$ and $X_0 = 0$ in (4.22) yields $X_t = B_t$, in which case the Itô formula (4.23)-(4.24) reads

$$
f(t, B_t) = f(0, B_0) + \int_0^t \frac{\partial f}{\partial s}(s, B_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s)dB_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s)ds,
$$

i.e. in differential notation:

$$
df(t, B_t) = \frac{\partial f}{\partial t}(t, B_t)dt + \frac{\partial f}{\partial x}(t, B_t)dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, B_t)dt. \qquad (4.25)
$$

Bivariate Itô formula

Next, consider two Itô processes $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ written in *integral form* as

$$
X_t = X_0 + \int_0^t v_s ds + \int_0^t u_s dB_s, \qquad t \geqslant 0,
$$

and

$$
Y_t = Y_0 + \int_0^t b_s ds + \int_0^t a_s dB_s, \qquad t \geqslant 0,
$$

or in *differential notation* as

$$
dX_t = v_t dt + u_t dB_t, \quad \text{and} \quad dY_t = b_t dt + a_t dB_t, \qquad t \geqslant 0.
$$

The Itô formula can also be written for functions $f \in \mathcal{C}^{1,2,2}(\mathbb{R}_+ \times \mathbb{R}^2)$ of two state variables as

$$
\begin{aligned}
df(t, X_t, Y_t) &= \frac{\partial f}{\partial t}(t, X_t, Y_t)dt + \frac{\partial f}{\partial x}(t, X_t, Y_t)dX_t + \frac{1}{2}|u_t|^2 \frac{\partial^2 f}{\partial x^2}(t, X_t, Y_t)dt\\
&+ \frac{\partial f}{\partial y}(t, X_t, Y_t)dY_t + \frac{1}{2}|a_t|^2 \frac{\partial^2 f}{\partial y^2}(t, X_t, Y_t)dt + u_t a_t \frac{\partial^2 f}{\partial x \partial y}(t, X_t, Y_t)dt.
\end{aligned}
\qquad (4.26)
$$

Itô multiplication table

Applying the bivariate Itô formula (4.26) to the function $f(x,y) := xy$ shows that

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + a_t u_t dt = X_t dY_t + Y_t dX_t + dX_t \bullet dY_t \qquad (4.27)$$

where the product

$$\begin{aligned} dX_t \bullet dY_t &= (v_t dt + u_t dB_t) \bullet (b_t dt + a_t dB_t) \\ &= b_t v_t \, dt \bullet dt + b_t u_t \, dt \bullet dB_t + a_t v_t \, dt \bullet dB_t + a_t u_t \, dB_t \bullet dB_t \\ &= a_t u_t dt \end{aligned}$$

can be computed according to the *Itô rule*

$$dt \bullet dt = 0, \qquad dt \bullet dB_t = 0, \qquad dB_t \bullet dB_t = dt, \qquad (4.28)$$

which can be encoded in the following Itô multiplication table:

$\bullet$	dt	dB_t
dt	0	0
dB_t	0	dt

Table 4.1: Itô multiplication table.

It follows similarly from the Itô Table 4.1 that

$$\begin{aligned} (dX_t)^2 &= (v_t dt + u_t dB_t) \bullet (v_t dt + u_t dB_t) \\ &= (v_t)^2 dt \bullet dt + (u_t)^2 \, dB_t \bullet dB_t + 2 u_t v_t \, dt \bullet dB_t \\ &= (u_t)^2 dt. \end{aligned}$$

Consequently, (4.24) can be rewritten as

$$df(t,X_t) = \frac{\partial f}{\partial t}(t,X_t)dt + \frac{\partial f}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t)dX_t \bullet dX_t, \qquad (4.29)$$

and the Itô formula for functions $f \in \mathcal{C}^{1,2,2}(\mathbb{R}_+ \times \mathbb{R}^2)$ of two state variables can be similarly rewritten as

$$\boxed{\begin{aligned} df(t,X_t,Y_t) &= \frac{\partial f}{\partial t}(t,X_t,Y_t)dt + \frac{\partial f}{\partial x}(t,X_t,Y_t)dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t,Y_t)(dX_t)^2 \\ &+ \frac{\partial f}{\partial y}(t,X_t,Y_t)dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(t,X_t,Y_t)(dY_t)^2 + \frac{\partial^2 f}{\partial x\partial y}(t,X_t,Y_t)(dX_t \bullet dY_t).\end{aligned}}$$

Examples

Applying Itô's formula (4.25) to $(B_t)^2$ with

$$(B_t)^2 = f(t,B_t) \quad \text{and} \quad f(t,x) = x^2,$$

and

$$\frac{\partial f}{\partial t}(t,x) = 0, \qquad \frac{\partial f}{\partial x}(t,x) = 2x, \qquad \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,x) = 1,$$

we find

$$\begin{aligned} d((B_t)^2) &= df(B_t) \\ &= \frac{\partial f}{\partial t}(t,B_t)dt + \frac{\partial f}{\partial x}(t,B_t)dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,B_t)dt \\ &= 2B_t dB_t + dt. \end{aligned}$$

Note that from the Itô Table 4.1 we could also write directly

$$d((B_t)^2) = B_t dB_t + B_t dB_t + (dB_t)^2 = 2B_t dB_t + dt.$$

Next, by integration in $t \in [0,T]$ we find

$$B_T^2 = B_0^2 + 2\int_0^T B_s dB_s + \int_0^T dt = 2\int_0^T B_s dB_s + T, \tag{4.30}$$

hence the relation

$$\boxed{\int_0^T B_s dB_s = \frac{1}{2}\left(B_T^2 - T\right), \tag{4.31}}$$

see Exercises 4.7 and 4.15 for the probability distribution of $\int_0^T B_s dB_s$.

Similarly, we have

i) $d\big((B_t)^3\big) = 3(B_t)^2 dB_t + 3B_t dt$.

Letting $f(x) := x^3$ with $f'(x) = 3x^2$ and $f''(x) = 6x$, we have

$$d((B_t)^3) = df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt = 3(B_t)^2 dB_t + 3B_t dt.$$

ii) $d(\sin B_t) = \cos(B_t)dB_t - \frac{1}{2}\sin(B_t)dt.$

Letting $f(x) := \sin(x)$ with $f'(x) = \cos(x)$, $f''(x) = -\sin(x)$, we have

$$\begin{aligned} d\sin(B_t) &= df(B_t) \\ &= f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt \\ &= \cos(B_t)dB_t - \frac{1}{2}\sin(B_t)dt. \end{aligned}$$

iii) $d\,\mathrm{e}^{B_t} = \mathrm{e}^{B_t}dB_t + \frac{1}{2}\mathrm{e}^{B_t}dt.$

Letting $f(x) := \mathrm{e}^x$ with $f'(x) = \mathrm{e}^x$, $f''(x) = \mathrm{e}^x$, we have

$$\begin{aligned} d\,\mathrm{e}^{B_t} &= df(B_t) \\ &= f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt \\ &= \mathrm{e}^{B_t}dB_t + \frac{1}{2}\mathrm{e}^{B_t}dt. \end{aligned}$$

iv) $d\log B_t = \frac{1}{B_t}dB_t - \frac{1}{2(B_t)^2}dt.$

Letting $f(x) := \log x$ with $f'(x) = 1/x$ and $f''(x) = -1/x^2$, we have

$$d\log B_t = df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt = \frac{dB_t}{B_t} - \frac{dt}{2(B_t)^2}.$$

v) $d\,\mathrm{e}^{tB_t} = B_t\,\mathrm{e}^{tB_t}dt + \frac{t^2}{2}\mathrm{e}^{tB_t}dt + t\,\mathrm{e}^{tB_t}dB_t.$

Letting $f(t,x) := \mathrm{e}^{xt}$ with

$$\frac{\partial f}{\partial t}(t,x) = x\,\mathrm{e}^{xt}, \quad \frac{\partial f}{\partial x}(t,x) = t\,\mathrm{e}^{xt}, \quad \frac{\partial^2 f}{\partial x^2}(t,x) = t^2\,\mathrm{e}^{xt},$$

we have

$$\begin{aligned} d\,\mathrm{e}^{tB_t} &= df(t,B_t) \\ &= \frac{\partial f}{\partial t}(t,B_t)dt + \frac{\partial f}{\partial x}(t,B_t)dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,B_t)dt \\ &= B_t\,\mathrm{e}^{tB_t}dt + t\,\mathrm{e}^{tB_t}dB_t + \frac{t^2}{2}\mathrm{e}^{tB_t}dt. \end{aligned}$$

vi) $d\cos(2t + B_t) = -2\sin(2t+B_t)dt - \sin(2t+B_t)dB_t - \frac{1}{2}\cos(2t+B_t)dt.$

Letting $f(t,x) := \cos(2t+x)$ with

$$\frac{\partial f}{\partial t}(t,x) = -2\sin(2t+x), \quad \frac{\partial f}{\partial x}(t,x) = -\sin(2t+x), \quad \frac{\partial^2 f}{\partial x^2}(t,x) = -\cos(2t+x),$$

we have

$$\begin{aligned} d\cos(2t+B_t) &= df(t,B_t) \\ &= \frac{\partial f}{\partial t}(t,B_t)dt + \frac{\partial f}{\partial x}(t,B_t)dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,B_t)dt \\ &= -2\sin(2t+B_t)dt - \sin(2t+B_t)dB_t - \frac{1}{2}\cos(2t+B_t)dt. \end{aligned}$$

Notation

We close this section with some comments on the practice of Itô's calculus. In certain finance textbooks, Itô's formula for *e.g.* geometric Brownian motion $(S_t)_{t\in\mathbb{R}_+}$ given by

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

can be found written in the notation

$$\begin{aligned} f(T,S_T) = f(0,S_0) &+ \sigma\int_0^T S_t\frac{\partial f}{\partial S_t}(t,S_t)dB_t + \mu\int_0^T S_t\frac{\partial f}{\partial S_t}(t,S_t)dt \\ &+ \int_0^T \frac{\partial f}{\partial t}(t,S_t)dt + \frac{1}{2}\sigma^2\int_0^T S_t^2\frac{\partial^2 f}{\partial S_t^2}(t,S_t)dt, \end{aligned}$$

or

$$df(S_t) = \sigma S_t\frac{\partial f}{\partial S_t}(S_t)dB_t + \mu S_t\frac{\partial f}{\partial S_t}(S_t)dt + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 f}{\partial S_t^2}(S_t)dt.$$

The notation $\frac{\partial f}{\partial S_t}(S_t)$ can in fact be easily misused in combination with the fundamental theorem of classical calculus, and potentially leads to the *wrong* identity

$$\cancel{df(S_t) = \frac{\partial f}{\partial S_t}(S_t)dS_t},$$

as in the following actual example:

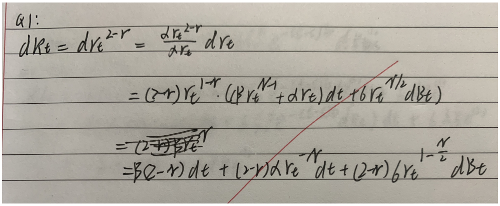


Fig. 4.21: Wrong application of Itô's formula (sample).

Similarly, writing

$$df(B_t) = \frac{\partial f}{\partial x}(B_t)dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(B_t)dt$$

is consistent, while writing

$$df(B_t) = \frac{\partial f(B_t)}{\partial B_t}dB_t + \frac{1}{2}\frac{\partial^2 f(B_t)}{\partial B_t^2}dt$$

is a potential *source of confusion*. Note also that the right-hand side of the Itô formula uses *partial derivatives* while its left-hand side is a *total derivative*.

Stochastic differential equations

In addition to geometric Brownian motion there exists a large family of stochastic differential equations that can be studied, although most of the time they cannot be explicitly solved. Let now

$$\sigma : \mathbb{R}_+ \times \mathbb{R}^n \longrightarrow \mathbb{R}^d \otimes \mathbb{R}^n$$

where $\mathbb{R}^d \otimes \mathbb{R}^n$ denotes the space of $d \times n$ matrices, and

$$b : \mathbb{R}_+ \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

satisfy the global Lipschitz condition

$$\|\sigma(t,x) - \sigma(t,y)\|^2 + \|b(t,x) - b(t,y)\|^2 \leqslant K^2\|x-y\|^2,$$

$t \in \mathbb{R}_+$, $x, y \in \mathbb{R}^n$. Then there exists a unique "strong" solution to the stochastic differential equation

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s, \quad t \geqslant 0, \tag{4.32}$$

i.e., in differential notation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad t \geqslant 0,$$

where $(B_t)_{t\in\mathbb{R}_+}$ is a d-dimensional Brownian motion, see *e.g.* Theorem V-7 in Protter (2004). In addition, the solution process $(X_t)_{t\in\mathbb{R}_+}$ of (4.32) has the *Markov property*, see § V-6 of Protter (2004).

The term $\sigma(s, X_s)$ in (4.32) will be interpreted later on in Chapters 8-9 as a *local volatility* component.

Stochastic differential equations can be used to model the behaviour of a variety of quantities, such as

- stock prices,

- interest rates,
- exchange rates,
- weather factors,
- electricity/energy demand,
- commodity (*e.g.* oil) prices, etc.

Next, we consider several examples of stochastic differential equations that can be solved explicitly using Itô's calculus, in addition to geometric Brownian motion. See *e.g.* § II-4.4 of Kloeden and Platen (1999) for more examples of explicitly solvable stochastic differential equations.

Examples of stochastic differential equations

1. Consider the mean-reverting stochastic differential equation

$$dX_t = -\alpha X_t dt + \sigma dB_t, \quad X_0 = x_0, \tag{4.33}$$

with $\alpha > 0$ and $\sigma > 0$.

```
N=10000; t <- 0:(N-1); dt <- 1.0/N;alpha=5; sigma=0.4;
dB <- rnorm(N,mean=0,sd=sqrt(dt));X <- rep(0,N);X[1]=0.5
for (j in 2:N){X[j]=X[j-1]-alpha*X[j-1]*dt+sigma*dB[j]}
plot(t*dt, X, xlab = "t", ylab = "", type = "l", ylim = c(-0.5,1), col = "blue")
abline(h=0)
```

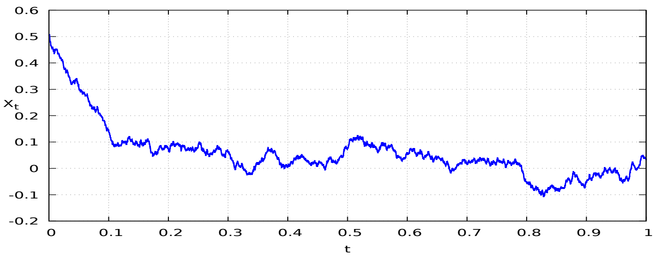


Fig. 4.22: Simulated path of (4.33) with $\alpha = 10$, $\sigma = 0.2$ and $X_0 = 0.5$.

We look for a solution of the form

$$X_t = a(t)Y_t = a(t)\left(x_0 + \int_0^t b(s)dB_s\right),$$

where

$$Y_t := x_0 + \int_0^t b(s)dB_s,$$

and $a(\cdot)$, $b(\cdot)$ are deterministic functions of time. After applying Theorem 4.24 to the Itô process $x_0 + \int_0^t b(s)dB_s$ of the form (4.22) with

$u_t = b(t)$ and $v(t) = 0$, and to the function $f(t, x) = a(t)x$, we find

$$\begin{aligned} dX_t &= d(a(t)Y_t) \\ &= Y_t a'(t)dt + a(t)dY_t \\ &= Y_t a'(t)dt + a(t)b(t)dB_t. \end{aligned} \tag{4.34}$$

By identification of (4.33) with (4.34), we get

$$\begin{cases} a'(t) = -\alpha a(t) \\ \\ a(t)b(t) = \sigma, \end{cases}$$

hence $a(t) = a(0)\,\mathrm{e}^{-\alpha t} = \mathrm{e}^{-\alpha t}$ and $b(t) = \sigma/a(t) = \sigma\,\mathrm{e}^{\alpha t}$, which shows that

$$X_t = x_0\,\mathrm{e}^{-\alpha t} + \sigma \int_0^t \mathrm{e}^{-(t-s)\alpha} dB_s, \qquad t \geqslant 0, \tag{4.35}$$

Using integration by parts, we can also write

$$X_t = x_0\,\mathrm{e}^{-\alpha t} + \sigma B_t - \sigma\alpha \int_0^t \mathrm{e}^{-(t-s)\alpha} B_s ds, \qquad t \geqslant 0, \tag{4.36}$$

Remark: the solution of the equation (4.33) *cannot* be written as a function $f(t, B_t)$ of t and B_t as in the proof of Proposition 5.15.

2. Consider the stochastic differential equation

$$dX_t = tX_t dt + \mathrm{e}^{t^2/2} dB_t, \qquad X_0 = x_0. \tag{4.37}$$

```
N=10000; T<-2.0; t <- 0:(N-1); dt <- T/N;
dB <- rnorm(N,mean=0,sd= sqrt(dt));X <- rep(0,N);X[1]=0.5
for (j in 2:N){X[j]=X[j-1]+j*X[j-1]*dt*dt+exp(j*dt*j*dt/2)*dB[j]}
plot(t*dt, X, xlab = "t", ylab = "", type = "l", ylim = c(-0.5,10), col = "blue")
abline(h=0)
```

Looking for a solution of the form $X_t = a(t)\left(X_0 + \int_0^t b(s)dB_s\right)$, where $a(\cdot)$ and $b(\cdot)$ are deterministic functions of time, we get $a'(t)/a(t) = t$ and $a(t)b(t) = \mathrm{e}^{t^2/2}$, hence $a(t) = \mathrm{e}^{t^2/2}$ and $b(t) = 1$, which yields $X_t = \mathrm{e}^{t^2/2}(X_0 + B_t)$, $t \geqslant 0$.

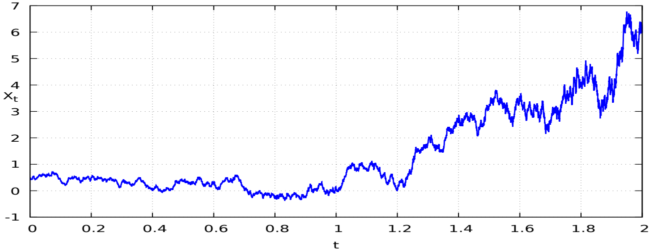


Fig. 4.23: Simulated path of (4.37).

3. Consider the stochastic differential equation

$$dY_t = (\sigma^2 - 2\alpha Y_t)dt + 2\sigma\sqrt{Y_t}dB_t, \qquad (4.38)$$

where $Y_0 > 0$, $\alpha \in \mathbb{R}$, and $\sigma > 0$.

```
N=10000; t <- 0:(N-1); dt <- 1.0/N;mu=-5;sigma=1;
dB <- rnorm(N,mean=0,sd=sqrt(dt));Y <- rep(0,N);Y[1]=0.5
for (j in 2:N){Y[j]=max(0,Y[j-1] +(2*mu*Y[j-1]+sigma*sigma)*dt
        +2*sigma*sqrt(Y[j-1])*dB[j])}
plot(t*dt, Y, xlab = "t", ylab = "", type = "l", ylim = c(-0.1,1), col = "blue")
abline(h=0)
```

Letting

$$X_t := \mathrm{e}^{-\alpha t}\sqrt{Y_0} + \sigma \int_0^t \mathrm{e}^{-(t-s)\alpha} dB_s, \qquad t \geqslant 0,$$

denote the solution of $dX_t = -\alpha X_t dt + \sigma dB_t$, see (4.35), by the Itô formula the process $Y_t := (X_t)^2$ satisfies the stochastic differential equation

$$\begin{aligned}
dY_t &= 2X_t dX_t + \sigma^2 dt \\
&= -2\alpha X_t^2 dt + 2\sigma X_t dB_t + \sigma^2 dt \\
&= (\sigma^2 - 2\alpha X_t^2)dt + 2\sigma |X_t| \text{sign}\,(X_t) dB_t \\
&= (\sigma^2 - 2\alpha Y_t)dt + 2\sigma\sqrt{Y_t}dW_t,
\end{aligned}$$

where the process

$$W_t := \int_0^t \text{sign}\,(X_\tau) dB_\tau, \qquad t \geqslant 0,$$

is a standard Brownian motion by the Lévy characterization theorem, see *e.g.* Theorem IV.3.6 in Revuz and Yor (1994). In this case, $Y_t = (X_t)^2$ is called a weak solution of (4.38).

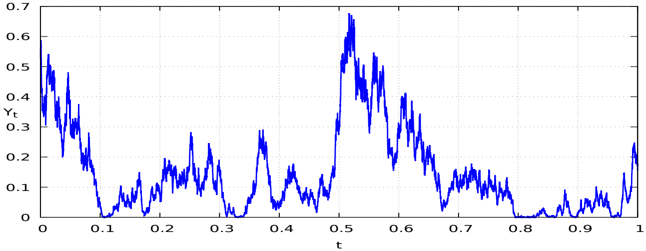


Fig. 4.24: Simulated path of (4.38) with $\alpha = -5$ and $\sigma = 1$.

See Proposition 2.1 in Hefter and Herzwurm (2017) for the representation of the strong solution of (4.38).

Exercises

Exercise 4.1 Compute $\mathbb{E}[B_s B_t]$ in terms of $s, t \geqslant 0$.

Exercise 4.2 Let $(B_t)_{t\in\mathbb{R}_+}$ denote a standard Brownian motion. Let $c > 0$. Among the following processes, tell which is a standard Brownian motion and which is not. Justify your answer.

a) $(X_t)_{t\in\mathbb{R}_+} := \big(B_{c+t} - B_c\big)_{t\in\mathbb{R}_+}$,
b) $(X_t)_{t\in\mathbb{R}_+} := \big(B_{ct^2}\big)_{t\in\mathbb{R}_+}$,
c) $(X_t)_{t\in\mathbb{R}_+} := \big(cB_{t/c^2}\big)_{t\in\mathbb{R}_+}$,
d) $(X_t)_{t\in\mathbb{R}_+} := \big(B_t + B_{t/2}\big)_{t\in\mathbb{R}_+}$.

Exercise 4.3 Let $(B_t)_{t\in\mathbb{R}_+}$ denote a standard Brownian motion. Compute the stochastic integrals

$$\int_0^T 2dB_t \qquad \text{and} \qquad \int_0^T \big(2\times \mathbb{1}_{[0,T/2]}(t) + \mathbb{1}_{(T/2,T]}(t)\big)dB_t$$

and determine their probability distributions (including mean and variance).

Exercise 4.4 Determine the probability distribution (including mean and variance) of the stochastic integral $\displaystyle\int_0^{2\pi} \sin(t)\,dB_t$.

Exercise 4.5 Let $T > 0$. Show that for $f : [0,T] \mapsto \mathbb{R}$ a differentiable function such that $f(T) = 0$, we have

$$\int_0^T f(t) dB_t = -\int_0^T f'(t) B_t dt.$$

Hint: Apply Itô's calculus to $t \mapsto f(t) B_t$.

Exercise 4.6 Let $(B_t)_{t \in \mathbb{R}_+}$ denote a standard Brownian motion.

a) Find the probability distribution of the stochastic integral $\int_0^1 t^2 dB_t$.
b) Find the probability distribution of the stochastic integral $\int_0^1 t^{-1/2} dB_t$.

Exercise 4.7 Find the mean, variance and probability distribution of the stochastic integral $\displaystyle\int_0^T B_t \, dB_t$.

Exercise 4.8 Given $(B_t)_{t \in \mathbb{R}_+}$ a standard Brownian motion and $n \geqslant 1$, let the random variable X_n be defined as

$$X_n := \int_0^{2\pi} \sin(nt) dB_t, \qquad n \geqslant 1.$$

a) Give the probability distribution of X_n for all $n \geqslant 1$.
b) Show that $(X_n)_{n \geqslant 1}$ is a sequence of identically distributed and pairwise independent random variables.

Hint: We have $\sin a \sin b = \dfrac{1}{2}\big(\cos(a-b) - \cos(a+b)\big)$, $a, b \in \mathbb{R}$.

Exercise 4.9 Apply the Itô formula to the process $X_t := \sin^2(B_t)$, $t \geqslant 0$.

Exercise 4.10 Let $(B_t)_{t \in \mathbb{R}_+}$ denote a standard Brownian motion.

a) Using the Itô isometry and the known relations

$$B_T = \int_0^T dB_t \quad \text{and} \quad B_T^2 = T + 2\int_0^T B_t dB_t,$$

compute the third and fourth moments $\mathbb{E}[B_T^3]$ and $\mathbb{E}[B_T^4]$.
b) Give the third and fourth moments of the centered normal distribution with variance σ^2.

Exercise 4.11 Let $(B_t)_{t \in \mathbb{R}_+}$ be a standard Brownian motion.

a) Show that

$$\mathbb{E}\left[\int_0^t \frac{|B_s|}{s} ds\right] < \infty, \qquad t > 0.$$

Hint: The Gaussian distribution $\mathcal{N}(0, s)$ has the probability density function $x \mapsto \mathrm{e}^{-x^2/(2s)}/\sqrt{2\pi s}$.

b) We let

$$\widehat{B}_t := B_t - \int_0^t \frac{B_s}{s} ds, \qquad t > 0.$$

Compute the mean and variance of $\widehat{B}_t$.

c) Show that $\widehat{B}_t$ is independent of B_t for all $t > 0$.
Hint: As the random vector $(B_t, \widehat{B}_t)$ has a bivariate Gaussian distribution, the random variables B_t and $\widehat{B}_t$ are independent if and only if they are uncorrelated.

Exercise 4.12 Let $(B_t)_{t\in\mathbb{R}_+}$ denote a standard Brownian motion. Given $T > 0$, find the stochastic integral decomposition of $(B_T)^3$ as

$$(B_T)^3 = C + \int_0^T \zeta_{t,T} dB_t \tag{4.39}$$

where $C \in \mathbb{R}$ is a constant and $(\zeta_{t,T})_{t\in[0,T]}$ is an adapted process to be determined.

Exercise 4.13 Let $f \in L^2([0,T])$, and consider a standard Brownian motion $(B_t)_{t\in[0,T]}$.

a) Compute the conditional expectation

$$\mathbb{E}\left[\mathrm{e}^{\int_0^T f(s)dB_s} \Big| \mathcal{F}_t \right], \qquad 0 \leqslant t \leqslant T,$$

where $(\mathcal{F}_t)_{t\in[0,T]}$ denotes the filtration generated by $(B_t)_{t\in[0,T]}$.

b) Using the result of Question (a), show that the process

$$t \longmapsto \exp\left(\int_0^t f(s)dB_s - \frac{1}{2}\int_0^t f^2(s)ds \right), \qquad 0 \leqslant t \leqslant T,$$

is an $(\mathcal{F}_t)_{t\in[0,T]}$-martingale, where $(\mathcal{F}_t)_{t\in[0,T]}$ denotes the filtration generated by $(B_t)_{t\in[0,T]}$.

c) By applying the result of Question (b) to the function $f(t) := \sigma \mathbb{1}_{[0,T]}(t)$, show that the geometric Brownian motion process $\left(\mathrm{e}^{\sigma B_t - \sigma^2 t/2}\right)_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-martingale for any $\sigma \in \mathbb{R}$.

Exercise 4.14 Consider two assets whose prices $S_t^{(1)}$, $S_t^{(2)}$ follow the Bachelier dynamics

$$dS_t^{(1)} = \mu S_t^{(1)} dt + \sigma_1 dW_t^{(1)}, \quad dS_t^{(2)} = \mu S_t^{(2)} dt + \sigma_2 dW_t^{(2)}, \quad t \in [0,T],$$

where $\left(W_t^{(1)}\right)_{t\in[0,T]}$, $\left(W_t^{(2)}\right)_{t\in[0,T]}$ are two Brownian motions with correlation $\rho \in [-1,1]$, *i.e.* we have $dW_t^{(1)} \bullet dW_t^{(2)} = \rho dt$. Show that the spread $S_t :=$

$S_t^{(2)} - S_t^{(1)}$ also satisfies an equation of the form

$$dS_t = \mu S_t dt + \sigma dW_t,$$

where $\mu \in \mathbb{R}$, $(W_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion, and $\sigma > 0$ should be given in terms of σ_1, σ_2 and ρ.

Hint: By the Lévy characterization theorem, see *e.g.* Theorem IV.3.6 in Revuz and Yor (1994), Brownian motion $(W_t)_{t\in\mathbb{R}_+}$ is the only continuous martingale such that $dW_t \bullet dW_t = dt$.

Exercise 4.15

a) Compute the moment generating function

$$\mathbb{E}\left[\exp\left(\beta\int_0^T B_t dB_t\right)\right]$$

for all $\beta < 1/T$.

Hint: Expand $(B_T)^2$ using the Itô formula as in (4.30).

b) Find the probability distribution of the stochastic integral $\int_0^T B_t dB_t$.

Exercise 4.16

a) Solve the stochastic differential equation

$$dX_t = -bX_t dt + \sigma\,\mathrm{e}^{-bt} dB_t, \qquad t \geqslant 0, \tag{4.40}$$

where $(B_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion and $\sigma, b \in \mathbb{R}$.

b) Solve the stochastic differential equation

$$dX_t = -bX_t dt + \sigma\,\mathrm{e}^{-at} dB_t, \qquad t \geqslant 0, \tag{4.41}$$

where $(B_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion and $a, b, \sigma > 0$ are positive constants.

c) Find the probability distribution of X_t, $t > 0$.

Exercise 4.17 Given $T > 0$, let $(X_t)_{t\in[0,T)}$ denote the solution of the stochastic differential equation

$$dX_t = \sigma dB_t - \frac{X_t}{T-t}dt, \qquad t \in [0, T), \tag{4.42}$$

under the initial condition $X_0 = 0$ and $\sigma > 0$.

a) Show that
$$X_t = (T-t)\int_0^t \frac{\sigma}{T-s} dB_s, \qquad 0 \leqslant t < T.$$
Hint: Start by computing $d(X_t/(T-t))$ using the Itô formula.
b) Show that $\mathbb{E}[X_t] = 0$ for all $t \in [0,T)$.
c) Show that $\mathrm{Var}[X_t] = \sigma^2 t(T-t)/T$ for all $t \in [0,T)$.
d) Show that $\lim_{t\to T} X_t = 0$ in $L^2(\Omega)$. The process $(X_t)_{t\in[0,T]}$ is called a *Brownian bridge*.

Exercise 4.18 Exponential Vašíček (1977) model (1). Consider a Vasicek process $(r_t)_{t\in\mathbb{R}_+}$ solving of the stochastic differential equation

$$dr_t = (a - br_t)dt + \sigma dB_t, \qquad t \geqslant 0,$$

where $(B_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion and $\sigma, a, b > 0$ are positive constants. Show that the exponential $X_t := \mathrm{e}^{r_t}$ satisfies a stochastic differential equation of the form

$$dX_t = X_t\big(\widetilde{a} - \widetilde{b}f(X_t)\big)dt + \sigma g(X_t)dB_t,$$

where the coefficients $\widetilde{a}$ and $\widetilde{b}$ and the functions $f(x)$ and $g(x)$ are to be determined.

Exercise 4.19 Exponential Vasicek model (2). Consider a short-term rate interest rate process $(r_t)_{t\in\mathbb{R}_+}$ in the exponential Vasicek model:

$$dr_t = (\eta - a\log r_t)r_t dt + \sigma r_t dB_t, \tag{4.43}$$

where η, a, σ are positive parameters and $(B_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion.

a) Find the solution $(Z_t)_{t\in\mathbb{R}_+}$ of the stochastic differential equation
$$dZ_t = -aZ_t dt + \sigma dB_t$$
as a function of the initial condition Z_0, where a and σ are positive parameters.
b) Find the solution $(Y_t)_{t\in\mathbb{R}_+}$ of the stochastic differential equation
$$dY_t = (\theta - aY_t)dt + \sigma dB_t \tag{4.44}$$
as a function of the initial condition Y_0. *Hint*: Let $Z_t := Y_t - \theta/a$.
c) Let $X_t = \mathrm{e}^{Y_t}$, $t \in \mathbb{R}_+$. Determine the stochastic differential equation satisfied by $(X_t)_{t\in\mathbb{R}_+}$.
d) Find the solution $(r_t)_{t\in\mathbb{R}_+}$ of (4.43) in terms of the initial condition r_0.

e) Compute the conditional mean* $\mathbb{E}[r_t|\mathcal{F}_u]$.
f) Compute the conditional variance

$$\mathrm{Var}[r_t|\mathcal{F}_u] := \mathbb{E}[r_t^2|\mathcal{F}_u] - (\mathbb{E}[r_t|\mathcal{F}_u])^2$$

of r_t, $0 \leqslant u \leqslant t$, where $(\mathcal{F}_u)_{u\in\mathbb{R}_+}$ denotes the filtration generated by the Brownian motion $(B_t)_{t\in\mathbb{R}_+}$.
g) Compute the asymptotic mean and variance $\lim_{t\to\infty} \mathbb{E}[r_t]$ and $\lim_{t\to\infty} \mathrm{Var}[r_t]$.

Exercise 4.20 Cox-Ingersoll-Ross (CIR) model. Consider the equation

$$dr_t = (\alpha - \beta r_t)dt + \sigma\sqrt{r_t}dB_t \tag{4.45}$$

modeling the variations of a short-term interest rate process r_t, where α, β, σ and r_0 are positive parameters and $(B_t)_{t\in\mathbb{R}_+}$ is a standard Brownian motion.

a) Write down the equation (4.45) in integral form.
b) Let $u(t) = \mathbb{E}[r_t]$. Show, using the integral form of (4.45), that $u(t)$ satisfies the differential equation

$$u'(t) = \alpha - \beta u(t),$$

and compute $\mathbb{E}[r_t]$ for all $t \geqslant 0$.
c) By an application of Itô's formula to r_t^2, show that

$$dr_t^2 = r_t(2\alpha + \sigma^2 - 2\beta r_t)dt + 2\sigma r_t^{3/2}dB_t. \tag{4.46}$$

d) Using the integral form of (4.46), find a differential equation satisfied by $v(t) := \mathbb{E}[r_t^2]$ and compute $\mathbb{E}[r_t^2]$ for all $t \geqslant 0$.
e) Show that

$$\mathrm{Var}[r_t] = r_0\frac{\sigma^2}{\beta}\left(\mathrm{e}^{-\beta t} - \mathrm{e}^{-2\beta t}\right) + \frac{\alpha\sigma^2}{2\beta^2}\left(1 - \mathrm{e}^{-\beta t}\right)^2, \quad t \geqslant 0.$$

Problem 4.21 Itô-Tanaka formula. Let $(B_t)_{t\in\mathbb{R}_+}$ be a standard Brownian motion started at $B_0 \in \mathbb{R}$.

a) Does the Itô formula apply to the European call option payoff function $f(x) := (x - K)^+$? Why?
b) For every $\varepsilon > 0$, consider the approximation $f_\varepsilon(x)$ of $f(x) := (x - K)^+$ defined by

* One may use the Gaussian moment generating function $\mathbb{E}[\mathrm{e}^X] = \mathrm{e}^{\alpha^2/2}$ for $X \simeq \mathcal{N}(0, \alpha^2)$.

$$f_\varepsilon(x) := \begin{cases} x-K & \text{if } x > K+\varepsilon, \\ \\ \dfrac{1}{4\varepsilon}(x-K+\varepsilon)^2 & \text{if } K-\varepsilon < x < K+\varepsilon, \\ \\ 0 & \text{if } x < K-\varepsilon. \end{cases}$$

Plot the graph of the function $x \mapsto f_\varepsilon(x)$ for $\varepsilon = 1$ and $K = 10$.

c) Using the Itô formula, show that

$$\begin{aligned} f_\varepsilon(B_T) = f_\varepsilon(B_0) + \int_0^T f_\varepsilon'(B_t)dB_t \\ + \frac{1}{4\varepsilon}\ell\left(\left\{t \in [0,T] \ : \ K-\varepsilon < B_t < K+\varepsilon\right\}\right), \end{aligned} \tag{4.47}$$

where ℓ denotes the measure of time length (Lebesgue measure) in $\mathbb{R}$.

d) Show that $\lim_{\varepsilon\to 0} \|\mathbb{1}_{[K,\infty)}(\cdot) - f_\varepsilon'(\cdot)\|_{L^2(\mathbb{R}_+)} = 0$.

e) Show, using the Itô isometry,* that the limit

$$\mathcal{L}_{[0,T]}^K := \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\ell(\{t \in [0,T] \ : \ K-\varepsilon < B_t < K+\varepsilon\})$$

exists in $L^2(\Omega)$, and prove the Itô-Tanaka formula

$$(B_T - K)^+ = (B_0 - K)^+ + \int_0^T \mathbb{1}_{[K,\infty)}(B_t)dB_t + \frac{1}{2}\mathcal{L}_{[0,T]}^K. \tag{4.48}$$

The quantity $\mathcal{L}_{[0,T]}^K$ is called the *local time* spent by Brownian motion at the level K.

Problem 4.22 Lévy's construction of Brownian motion. The goal of this problem is to prove the existence of standard Brownian motion $(B_t)_{t\in[0,1]}$ as a stochastic process satisfying the four properties of Definition 4.1, *i.e.*:

1. $B_0 = 0$ almost surely,

2. The sample trajectories $t \mapsto B_t$ are continuous, with probability 1.

3. For any finite sequence of times $t_0 < t_1 < \cdots < t_n$, the increments

$$B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$$

are independent.

4. For any given times $0 \leqslant s < t$, $B_t - B_s$ has the Gaussian distribution $\mathcal{N}(0, t-s)$ with mean zero and variance $t-s$.

* Hint: Show that $\displaystyle\lim_{\varepsilon\to 0} \mathbb{E}\left[\int_0^T \left(\mathbb{1}_{[K,\infty)}(B_t) - f_\varepsilon'(B_t)\right)^2 dt\right] = 0$.

The construction will proceed by the linear interpolation scheme illustrated in Figure 4.10. We work on the space $\mathcal{C}_0([0,1])$ of continuous functions on $[0,1]$ started at 0, with the norm

$$\|f\|_\infty := \operatorname*{Max}_{t\in[0,1]} |f(t)|$$

and the distance

$$\|f-g\|_\infty := \operatorname*{Max}_{t\in[0,1]} |f(t)-g(t)|.$$

The following ten questions are interdependent.

a) Show that for any Gaussian random variable $X \simeq \mathcal{N}(0,\sigma^2)$, we have

$$\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\sigma}{\varepsilon\sqrt{\pi/2}}\,\mathrm{e}^{-\varepsilon^2/(2\sigma^2)}, \qquad \varepsilon > 0.$$

Hint: Start from the inequality $\mathbb{E}[(X-\varepsilon)^+] \geqslant 0$ and compute the left-hand side.

b) Let X and Y be two independent centered Gaussian random variables with variances α^2 and β^2. Show that the conditional distribution

$$\mathbb{P}(X \in dx \mid X+Y=z)$$

of X given $X+Y=z$ is Gaussian with mean $\alpha^2 z/(\alpha^2+\beta^2)$ and variance $\alpha^2\beta^2/(\alpha^2+\beta^2)$.

Hint: Use the definition

$$\mathbb{P}(X \in dx \mid X+Y=z) := \frac{\mathbb{P}(X\in dx \text{ and } X+Y\in dz)}{\mathbb{P}(X+Y\in dz)}$$

and the formulas

$$d\mathbb{P}(X\leqslant x) := \frac{1}{\sqrt{2\pi\alpha^2}}\,\mathrm{e}^{-x^2/(2\alpha^2)}dx, \quad d\mathbb{P}(Y\leqslant x) := \frac{1}{\sqrt{2\pi\beta^2}}\,\mathrm{e}^{-x^2/(2\beta^2)}dx,$$

where dx (resp. dy) represents a "small" interval $[x,x+dx]$ (resp. $[y,y+dy]$).

c) Let $(B_t)_{t\in\mathbb{R}_+}$ denote a standard Brownian motion and let $0<u<v$. Give the distribution of $B_{(u+v)/2}$ given that $B_u=x$ and $B_v=y$.

Hint: Note that given that $B_u=x$, the random variable B_v can be written as

$$B_v = (B_v - B_{(u+v)/2}) + (B_{(u+v)/2} - B_u) + x, \tag{4.49}$$

and apply the result of Question (b) after identifying X and Y in the above decomposition (4.49).

d) Consider the random sequences

$$
\begin{cases}
Z^{(0)} = \big(0, Z^{(0)}_1\big) \\
Z^{(1)} = \big(0, Z^{(1)}_{1/2}, Z^{(0)}_1\big) \\
Z^{(2)} = \big(0, Z^{(2)}_{1/4}, Z^{(1)}_{1/2}, Z^{(2)}_{3/4}, Z^{(0)}_1\big) \\
Z^{(3)} = \big(0, Z^{(3)}_{1/8}, Z^{(2)}_{1/4}, Z^{(3)}_{3/8}, Z^{(1)}_{1/2}, Z^{(3)}_{5/8}, Z^{(2)}_{3/4}, Z^{(3)}_{7/8}, Z^{(0)}_1\big) \\
\quad \vdots \qquad \vdots \\
Z^{(n)} = \big(0, Z^{(n)}_{1/2^n}, Z^{(n)}_{2/2^n}, Z^{(n)}_{3/2^n}, Z^{(n)}_{4/2^n}, \ldots, Z^{(n)}_1\big) \\
Z^{(n+1)} = \big(0, Z^{(n+1)}_{1/2^{n+1}}, Z^{(n)}_{1/2^n}, Z^{(n+1)}_{3/2^{n+1}}, Z^{(n+1)}_{2/2^n}, Z^{(n+1)}_{5/2^{n+1}}, Z^{(n+1)}_{3/2^n}, \ldots, Z^{(n+1)}_1\big)
\end{cases}
$$

with $Z^{(n)}_0 = 0$, $n \geqslant 0$, defined recursively as

i) $Z^{(0)}_1 \simeq \mathcal{N}(0,1)$,

ii) $Z^{(1)}_{1/2} \simeq \dfrac{Z^{(0)}_0 + Z^{(0)}_1}{2} + \mathcal{N}(0,1/4)$,

iii) $Z^{(2)}_{1/4} \simeq \dfrac{Z^{(1)}_0 + Z^{(1)}_{1/2}}{2} + \mathcal{N}(0,1/8)$, $\quad Z^{(2)}_{3/4} \simeq \dfrac{Z^{(1)}_{1/2} + Z^{(0)}_1}{2} + \mathcal{N}(0,1/8)$,

and more generally

$$
Z^{(n+1)}_{(2k+1)/2^{n+1}} = \frac{Z^{(n)}_{k/2^n} + Z^{(n)}_{(k+1)/2^n}}{2} + \mathcal{N}(0,1/2^{n+2}), \qquad k = 0,1,\ldots,2^n-1,
$$

where $\mathcal{N}(0,1/2^{n+2})$ is an independent centered Gaussian sample with variance $1/2^{n+2}$, and $Z^{(n+1)}_{k/2^n} := Z^{(n)}_{k/2^n}$, $k = 0,1,\ldots,2^n$.

In what follows we denote by $\big(Z^{(n)}_t\big)_{t\in[0,1]}$ the continuous-time random path obtained by linear interpolation of the sequence points in $\big(Z^{(n)}_{k/2^n}\big)_{k=0,1,\ldots,2^n}$.

Draw a sample of the first four linear interpolations $\big(Z^{(0)}_t\big)_{t\in[0,1]}$, $\big(Z^{(1)}_t\big)_{t\in[0,1]}$, $\big(Z^{(2)}_t\big)_{t\in[0,1]}$, $\big(Z^{(3)}_t\big)_{t\in[0,1]}$, and label the values of $Z^{(n)}_{k/2^n}$ on the graphs for $k = 0,1,\ldots,2^n$ and $n = 0,1,2,3$.

e) Using an induction argument, explain why for all $n \geqslant 0$ the sequence

$$Z^{(n)} = \big(0, Z^{(n)}_{1/2^n}, Z^{(n)}_{2/2^n}, Z^{(n)}_{3/2^n}, Z^{(n)}_{4/2^n}, \ldots, Z^{(n)}_1\big)$$

has same distribution as the sequence

$$B^{(n)} := \big(B_0, B_{1/2^n}, B_{2/2^n}, B_{3/2^n}, B_{4/2^n}, \ldots, B_1\big).$$

Hint: Compare the constructions of Questions (c) and (d) and note that under the above linear interpolation, we have

$$Z^{(n)}_{(2k+1)/2^{n+1}} = \frac{Z^{(n)}_{k/2^n} + Z^{(n)}_{(k+1)/2^n}}{2}, \qquad k = 0, 1, \ldots, 2^n - 1.$$

f) Show that for any $\varepsilon_n > 0$ we have

$$\mathbb{P}\big(\|Z^{(n+1)} - Z^{(n)}\|_\infty \geqslant \varepsilon_n\big) \leqslant 2^n \mathbb{P}\big(|Z^{(n+1)}_{1/2^{n+1}} - Z^{(n)}_{1/2^{n+1}}| \geqslant \varepsilon_n\big).$$

Hint: Use the inequality

$$\mathbb{P}\left(\bigcup_{k=0}^{2^n-1} A_k\right) \leqslant \sum_{k=0}^{2^n-1} \mathbb{P}(A_k)$$

for a suitable choice of events $(A_k)_{k=0,1,\ldots,2^n-1}$.

g) Use the results of Questions (a) and (f) to show that for any $\varepsilon_n > 0$ we have

$$\mathbb{P}\left(\|Z^{(n+1)} - Z^{(n)}\|_\infty \geqslant \varepsilon_n\right) \leqslant \frac{2^{n/2}}{\varepsilon_n \sqrt{2\pi}} \mathrm{e}^{-\varepsilon_n^2 2^{n+1}}.$$

h) Taking $\varepsilon_n = 2^{-n/4}$, show that

$$\mathbb{P}\left(\sum_{n \geqslant 0} \|Z^{(n+1)} - Z^{(n)}\|_\infty < \infty\right) = 1.$$

Hint: Show first that

$$\sum_{n \geqslant 0} \mathbb{P}\left(\|Z^{(n+1)} - Z^{(n)}\|_\infty \geqslant 2^{-n/4}\right) < \infty,$$

and apply the Borel-Cantelli lemma.

i) Show that with probability one, the sequence $\left\{\big(Z^{(n)}_t\big)_{t \in [0,1]},\ n \geqslant 1\right\}$ converges uniformly on $[0,1]$ to a continuous (random) function $(Z_t)_{t \in [0,1]}$.

Hint: Use the fact that $\mathcal{C}_0([0,1])$ is a complete space for the $\|\cdot\|_\infty$ norm.

j) Argue that the limit $(Z_t)_{t \in [0,1]}$ is a standard Brownian motion on $[0,1]$ by checking the four relevant properties.

Problem 4.23 Consider $(B_t)_{t\in\mathbb{R}_+}$ a standard Brownian motion, and for any $n \geqslant 1$ and $T > 0$, define the discretized quadratic variation

$$Q_T^{(n)} := \sum_{k=1}^{n} (B_{kT/n} - B_{(k-1)T/n})^2, \qquad n \geqslant 1.$$

a) Compute $\mathbb{E}\left[Q_T^{(n)}\right]$, $n \geqslant 1$.

b) Compute $\mathrm{Var}[Q_T^{(n)}]$, $n \geqslant 1$.

c) Show that

$$\lim_{n\to\infty} Q_T^{(n)} = T,$$

where the limit is taken in $L^2(\Omega)$, that is, show that

$$\lim_{n\to\infty} \|Q_T^{(n)} - T\|_{L^2(\Omega)} = 0,$$

where

$$\big\|Q_T^{(n)} - T\big\|_{L^2(\Omega)} := \sqrt{\mathbb{E}\big[\big(Q_T^{(n)} - T\big)^2\big]}, \qquad n \geqslant 1.$$

d) By the result of Question (c), show that the limit

$$\int_0^T B_t dB_t := \lim_{n\to\infty} \sum_{k=1}^{n} (B_{kT/n} - B_{(k-1)T/n}) B_{(k-1)T/n}$$

exists in $L^2(\Omega)$, and compute it.

Hint: Use the identity

$$(x-y)y = \frac{1}{2}(x^2 - y^2 - (x-y)^2), \qquad x, y \in \mathbb{R}.$$

e) Consider the modified quadratic variation defined by

$$\widetilde{Q}_T^{(n)} := \sum_{k=1}^{n} (B_{(k-1/2)T/n} - B_{(k-1)T/n})^2, \qquad n \geqslant 1.$$

Compute the limit $\lim_{n\to\infty} \widetilde{Q}_T^{(n)}$ in $L^2(\Omega)$ by repeating the steps of Questions (a)-(c).

f) By the result of Question (e), show that the limit

$$\int_0^T B_t \circ dB_t := \lim_{n\to\infty} \sum_{k=1}^{n} (B_{kT/n} - B_{(k-1)T/n}) B_{(k-1/2)T/n}$$

This version: May 3, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

exists in $L^2(\Omega)$, and compute it.

Hint: Use the identities

$$(x-y)y = \frac{1}{2}(x^2 - y^2 - (x-y)^2),$$

and

$$(x-y)x = \frac{1}{2}(x^2 - y^2 + (x-y)^2), \qquad x, y \in \mathbb{R}.$$

g) More generally, by repeating the steps of Questions (e) and (f), show that for any $\alpha \in [0,1]$ the limit

$$\int_0^T B_t \circ d^\alpha B_t := \lim_{n\to\infty} \sum_{k=1}^n (B_{kT/n} - B_{(k-1)T/n}) B_{(k-\alpha)T/n}$$

exists in $L^2(\Omega)$, and compute it.

h) Comparison with deterministic calculus. Compute the limit

$$\lim_{n\to\infty} \sum_{k=1}^n (k-\alpha)\frac{T}{n}\left(k\frac{T}{n} - (k-1)\frac{T}{n}\right)$$

for all values of α in $[0,1]$.

Exercise 4.24 Let $(B_t)_{t\in\mathbb{R}_+}$ be a standard Brownian motion generating the information flow $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

a) Let $0 \leqslant t \leqslant T$. What is the probability distribution of $B_T - B_t$?

b) From the answer to Exercise A.4-(c), show that

$$\mathbb{E}[(B_T)^+ \mid \mathcal{F}_t] = \sqrt{\frac{T-t}{2\pi}}\,\mathrm{e}^{-(B_t)^2/(2(T-t))} + B_t \Phi\left(\frac{B_t}{\sqrt{T-t}}\right),$$

$0 \leqslant t \leqslant T$, where Φ denotes the standard Gaussian cumulative distribution function. *Hint*: Use the time splitting decomposition $B_T = B_T - B_t + B_t$.

c) Let $\sigma > 0$, $\nu \in \mathbb{R}$, and $X_t := \sigma B_t + \nu t$, $t \geqslant 0$. Compute e^{X_t} by applying the Itô formula

$$f(X_t) = f(X_0) + \int_0^t u_s \frac{\partial f}{\partial x}(X_s) dB_s + \int_0^t v_s \frac{\partial f}{\partial x}(X_s) ds + \frac{1}{2}\int_0^t u_s^2 \frac{\partial^2 f}{\partial x^2}(X_s) ds$$

to $f(x) = \mathrm{e}^x$, where X_t is written as $X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds$, $t \geqslant 0$.

d) Let $S_t = \mathrm{e}^{X_t}$, $t \geqslant 0$, and $r > 0$. For which value of ν does $(S_t)_{t\in\mathbb{R}_+}$ satisfy the stochastic differential equation

$$dS_t = rS_t dt + \sigma S_t dB_t \quad ?$$

Exercise 4.25 From the answer to Exercise A.4-(c), show that for any $\beta \in \mathbb{R}$ we have

$$\mathbb{E}[(\beta - B_T)^+ \mid \mathcal{F}_t] = \sqrt{\frac{T-t}{2\pi}}\, \mathrm{e}^{-(\beta - B_t)^2/(2(T-t))} + (\beta - B_t)\Phi\left(\frac{\beta - B_t}{\sqrt{T-t}}\right),$$

$0 \leqslant t \leqslant T$.

Hint: Use the time splitting decomposition $B_T = B_T - B_t + B_t$.