Introduction to Stochastic Differential Equations and Itô Calculus

Aadam Haq

Abstract

Stochastic differential equations have many real world applications. They are used in mathematical finance, neurodynamics and electrical fields. In this essay, we will see how to solve some differential equations by looking carefully through the key components that they are made of. The aim is to be able to compute some simple stochastic integrals and stochastic differential equations by the end of the essay.

Contents

1 Markov Processes and Brownian Motion

1.1 Motivating Example - The Drunkard's Walk

To begin, we will start with a well-known motivating example. Consider a random walk, starting at point 0. There is a $\frac{1}{2}$ chance that the particle 'jumps' by 1 unit or -1 unit. What happens at the limits of the walk in this scenario, and other similar scenarios? Figure 1 simulates the particular example with over 10000 steps and Figure 2 simulates three different examples, all starting at point 0.

To investigate this question further, we must first define some properties formally and from there, some generalisations. This particular analysis is adapted from Kuo [5, pp. 4-5]. We will consider the case where the jumps are of size h and will occur at times $x, 2x, \dots$. We define the sequence of identically, independently distributed (i.i.d) random variables $\{X_n\}_{n=1}^{\infty}$ such that $\mathbb{P}[X_j = h] = \mathbb{P}[X_j = -h] = \frac{1}{2}$. We can then intuitively define the 'height variable' at time nx by

$$Y_{x,h}(xh) = X_1 + X_2 + \cdots + X_n.$$

We will focus on the limit of the random walk, $Y_{x,h}$, as x and h tend to 0. To do this, we must find the limit of the characteristic function of $Y_{x,h}$. More precisely, we are trying to find $\lim_{x,h\to 0} \mathbb{E}[i\lambda Y_{x,h}(t)]$.

$$\begin{aligned}
\mathbb{E}[i\lambda Y_{x,h}(t)] &= \prod_{j=1}^{n} \mathbb{E}[e^{i\lambda X_j}] \\
&= (\mathbb{E}[e^{i\lambda X_1}])^n \\
&= (e^{i\lambda h \frac{1}{2}} + e^{-i\lambda h \frac{1}{2}})^n \\
&= (cos(\lambda h))^n \\
&= (cos(\lambda h))^{\frac{t}{x}} \qquad (1)
\end{aligned}$$

We have used the i.i.d property of the sequence (X_n) for this calculation. λ is fixed in the real numbers and we let $t = nx$ so that $n = \frac{t}{x}$.

For a fixed t and λ, the limit does not exist when x and h tend to zero independently (we cannot have 0 in the denominator of a fraction), so we must establish a relationship[1] between these two variables. We will set $u = (\cos(\lambda h))^{\frac{1}{x}}$, and thus $\log(u) = \frac{1}{x}\log(\cos(\lambda h))$. Using small angle approximations from the Taylor series we find,

$$\cos(\lambda h) \approx 1 - \frac{\lambda^2 h^2}{2}, \qquad \log(\cos(\lambda h)) \approx \log(1 - \frac{1}{2}\lambda^2 h^2) \approx \frac{-\lambda^2 h^2}{2}.$$

[1]Depending on this relationship, we may have different answers

We reach the conclusion that $ln(u) \approx \frac{-1}{2x}\lambda^2 h^2$ by applying the laws of logarithms, and so $u \approx \exp[\frac{-1}{2x}\lambda^2 h^2]$. Substituting these into (1), we see

$$\mathbb{E}[e^{i\lambda Y_{x,h}(t)}] = u \approx e^{\frac{-1}{2x}t\lambda^2 h^2}$$

for small x and h. In particular, we can then see that if we relate the two as $h^2 = x$ we reach the result that

$$\lim_{x \to 0} \mathbb{E}[e^{i\lambda Y_{x,h}(t)}] = e^{\frac{-1}{2}t\lambda^2}$$

So for a fixed λ, we have reached our conclusion for the limit of the random walk $Y_{x,h}$ as $x, h \to 0$ in the particular case that $h^2 = x$. But how can this be interpreted? The distribution of the limit of $Y_{x,h}(t)$ satisfies a process called Brownian motion by fulfilling some key properties. Our investigation, by looking at small movements in time and space of a moving particle, in this case a random walk, is similar to the discovery of the process of Brownian motion. Over the course of this essay, Brownian motion is a key focus.

1.2 Brownian Motion

In 1825, the botanist Robert Brown was investigating the suspension of pollen particles in fluids, and noticed an irregular, random motion. A simulation of this can be seen in Figure 3. Einstein later argued this was due collisions with molecules in the fluid, leading to the atomic model we have today. After more investigation, in 1923, Wiener laid the foundations of, and formally defined Brownian motion (also known as the Wiener process) as a stochastic process.

Before we define Brownian motion, we must define a few other elements that are required.

In the following definitions, and over the course of this essay, we assume that the standard probability space is $(\Omega, \mathcal{F}, \mathbb{P})$. Ω is the sample space – all possible events that can occur, $\mathcal{F}$ is the field of events (also known as the event set) – a set of events from Ω, and $\mathbb{P}$ is the probability function – this assigns a probability between 0 and 1.

Definition 1.1 (Filtration)**.** [3, p. 17]. A filtration, $\mathbb{F}$ is the collection of fields,

$$\mathbb{F} = \{\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_t, \dots, \mathcal{F}_T\}$$

such that $\mathcal{F}_t \subset \mathcal{F}_{t+1}$.

In practice, this means that as time passes, the observer understands more detailed information. As an example, in terms of stocks, $\mathbb{F}$ describes the information about prices revealed to investors. Over time, the investors learn more and can make a more informed decision as a result.

Definition 1.2 (Stochastic Process)**.** [5, p. 17]. A stochastic process is a collection of random variables $\{X(t)\}$ for fixed $t = 0, 1, 2, \ldots, T$. $X(t)$ is a random variable over $(\Omega,\ \mathcal{F})$.

A stochastic process can be adapted to a filtration, $\mathbb{F}$ if for every $t \in \{0, 1, \ldots, T\}, X(t)$ is a random variable on $\mathcal{F}_t$. This is known as $X(t)$ being $\mathcal{F}_t$-measurable.

Remark. It is also common to see a stochastic process written as $X(t, \omega)$ defined on the product space $[0, \infty) \times \Omega$. The two will be used interchangeably during the course of this essay.

Now that we have seen filtrations and stochastic processes, we can investigate Brownian motion further.

Definition 1.3 (Brownian Motion)**.** [2, pp. 68-70]. A stochastic process $B(t, \omega)$ is called a Brownian motion if it satisfies the following properties:

i (Independence of Increments) $B(t, \omega)$ has independent increments. This means that for any $0 \leq t_1 < t_2 < \cdots < t_n$, the random variables

$$B(t_1), B(t_2) - B(t_1), \ldots, B(t_n) - B(t_{n-1})$$

are all independent.

ii (Normal Increments) For any $0 \leq s < t$, the random variable $B(t) - B(s)$ is normally distributed with mean 0, and variance $t - s$. Formally, for any real numbers a and b,

$$\mathbb{P}\{a \leq B(t) - B(s) \leq b\} = \frac{1}{2\pi(t-s)} \int_a^b e^{\frac{-x^2}{2(t-s)}} \mathrm{d}x$$

Note that if $s = 0$, we have a $N(0, t)$ distribution.

iii (Continuity of Paths) All paths $B(t)$ for $t > 0$ are continuous functions of t.

Remark (1)*.* Usually we also have $B(0) = 0$ for a standard Brownian Motion, although this is not necessary. If it is not, we can define $W(t) = B(t + s) - B(s)$ for a fixed s. Then $W(t)$ is a Brownian Motion starting at zero.

Remark (2)*.* From property i (independence of increments), we can see that Brownian motion is actually a Markov process. This means that what has happened in the past is independent of what will happen in the future.

The following example shows how we would calculate probabilities at a point using Remark (1).

Example 1.1. Let $B(0) = 0$. What is $\mathbb{P}(B(t) \leq 0, t = 0, 1, 2)$?
It is clear that $B(t) \sim N(0, 2)$ and $B(t) \sim N(0, 1)$ from property ii. By symmetry, we note that $\mathbb{P}[B(2) \leq 0] = \mathbb{P}[B(1) \leq 0] = \frac{1}{2}$ (and it is trivial that $\mathbb{P}[B(0) = 0 \leq 0] = 1$), but we cannot simply multiply these results together to answer the question as these three are not independent. We know however, by property i, the increments are independent. If we let $B(2) = B(1) + (B(2) - B(1)) = B(1) + W(1)$, we have these two random variables as independent. Therefore,

$$\begin{aligned}\mathbb{P}[B(1) \leq 0, B(2) \leq 0] &= \mathbb{P}[(B(1) \leq 0, B(1) + W(1) \leq 0) \\ &= \mathbb{P}[B(1) \leq 0, W(1) \leq -B(1)] \\ &= \cdots = \frac{3}{8}\end{aligned}$$

The purpose of the example was to demonstrate the use of the independence of increments, but the whole solution is outlined in Klebaner [3, pp.57-58].

We will now introduce a new, related concept.

Definition 1.4 (Martingales)**.** [3, pp. 50-51] Let X_t be an adapted stochastic process to a filtration $\mathcal{F}_t$, where $\mathbb{E}[X_t]$ is finite for every $t \in T$. Then X_t is called a martingale with respect to $\mathcal{F}_t$ if for any $s \leq t$,

$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s.$$

Remark. A submartingale is where $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$, and a supermartingale is where $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$

A natural consequence of this is as follows:

Theorem 1.1. *Brownian motion, $B(t)$ is a martingale*

Proof. Let $\mathcal{F}_t = \sigma\{B(s); s \leq t\}$, where $\sigma\{B(s); s \leq t\}$ denotes all the subsets that can be made by each of the elements in $(B(s); s \leq t)$. For $s \leq t$,

$$\mathbb{E}[B(t)|\mathcal{F}_s] = \mathbb{E}[B(t) - B(s)|\mathcal{F}_s] + \mathbb{E}[B(s)|\mathcal{F}_s].$$

But as $B(t) - B(s)$ is independent of $\mathcal{F}_s$, we have $\mathbb{E}[B(t) - B(s)|\mathcal{F}_s] = \mathbb{E}[B(t) - B(s)] = 0$ since $\mathbb{E}[B(t)] = 0$, for every value of t. That leaves $\mathbb{E}[B(s)|\mathcal{F}_s] = B(s)$ since $B(s)$ is $\mathcal{F}_s$-measurable.

We have shown that $\mathbb{E}[B(t)|\mathcal{F}_s] = B(s)$, for every $s \leq t$ and so have show that Brownian motion, $B(t)$ is a martingale. $\square$

It can also be shown that $B(t)^2 - t$ and for any $u, e^{uB(t) - \frac{u^2 t}{2}}$ are also martingales Klebaner [3, pp.65-67]. These three are known as the three main martingales associated with Brownian motion.

1.3 Differentiating Brownian motion

As will be seen in the next chapter, one remarkable property of Brownian motion is very important. It is continuous everywhere, yet differentiable nowhere. This is one of the key reasons why Itô calculus was developed, which will be seen in the following sections.

One key property to prove this is the scaling invariance of Brownian motion.

Theorem 1.2 (Scaling Invariance)**.** *[6, p. 10] Let $B(t)$ be a standard Brownian motion. For any positive integer, a, $\frac{1}{a}B(a^2t)$ is also standard Brownian motion.*

Proof. Clearly, the properties of independence of increments and continuity of paths are preserved by scaling. The only thing to show is normal increments. It is clear that if $X(t) = \frac{1}{a}B(a^2t)$, we have

$$X(t) - X(s) = \frac{1}{a}\left(B(a^2t) - B(a^2s)\right),$$

and we know that the expectation of this is zero as $X(t) - X(s)$ is normally distributed. To show the variance, we have

$$\begin{aligned}
\operatorname{Var}\left[\frac{1}{a}(B(a^2t) - B(a^2s))\right] &= \frac{1}{a^2}\operatorname{Var}\left[(B(a^2t) - B(a^2s))\right] \\
&= \frac{a^2(t-s)}{a^2} \\
&= t - s
\end{aligned}$$

□

Now that we have proved the scaling invariance property of Brownian motion, we need to introduce another fact. This is a useful lemma used in measure theory.

Lemma 1.3 (Borel–Cantelli)**.** *Let E_n be the set of events with the property $\sum_{n=1}^{\infty}\mathbb{P}[E_n] < \infty$. Then, $\mathbb{P}[\limsup_{n\to\infty} E_n] (= \mathbb{P}[\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k]) = 0$.*

Proof. The proof is adapted from Leiner [6, p.3]. We know that the intersection of any sets is a subset of both the original sets ($S \cap T \subseteq S$, $S \cap T \subseteq T$). So we have

$$\mathbb{P}\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k\right) \leq \mathbb{P}\left(\cup_{k=n}^{\infty} E_k\right) \leq \sum_{k=n}^{\infty}\mathbb{P}[E_k]$$

If we take limits as $n \to \infty$, and using the null sequence rule for series, we have

$$\mathbb{P}[\limsup_{n\to\infty} E_n] \leq \lim_{n\to\infty}\sum_{k=n}^{\infty}\mathbb{P}[E_k] = 0$$

□

Remark. An easy way to understand $\mathbb{P}[\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k]$ is that an event occurs infinitely often.

With these two key elements, we can prove our key property (continuous everywhere, differentiable nowhere). From property iii, we know Brownian motion is continuous on all paths.

Theorem 1.4 (Differentiability of Brownian motion)**.** *Brownian motion is nowhere differentiable. Furthermore, for all t,*

$$\limsup_{h\to 0} \frac{B(t+h)-B(t)}{h} = \infty \quad \text{or} \quad \limsup_{h\to 0} \frac{B(t+h)-B(t)}{h} = -\infty \quad \text{or both.}$$

Proof. This proof is adapted from Mcknight [7, pp. 4-5]. Suppose there is a $t_0 \in [0,1]$ where $\limsup\limits_{h\to 0} \frac{|B(t+h)-B(t)|}{h} < \infty$. By the boundedness theorem, we have Brownian motion is bounded on $[0,2]$ as it is continuous. So we have

$$\sup_{h\in[0,1]} \frac{|B(t+h)-B(t)|}{h} \leq M$$

for some t_0 and for some constant, M. If this is an event, we just need to show the probability of it is zero for any general M. We can fix M and partition the interval into $[\frac{k-1}{2^n}, \frac{k}{2^n}]$, $k \in 1, 2, \ldots, 2^n$ for $n > 2$. Then there will be a sub-interval in which $t_0 \in [\frac{k-1}{2^n}, \frac{k}{2^n}]$, and we can use the triangle property on $1 \leq j \leq 2^n - k$ to obtain

$$\begin{aligned}
\left|B\left(\frac{k+j}{2^n}\right) - B\left(\frac{k+j-1}{2^n}\right)\right| &= \left|B\left(\frac{k+j}{2^n}\right) - B(t_0) - B\left(\frac{k+j-1}{2^n}\right) + B(t_0)\right| \\
&\leq \left|B\left(\frac{k+j}{2^n}\right) - B(t_0)\right| + \left|B\left(\frac{k+j-1}{2^n}\right) - B(t_0)\right| \\
&= \left|B\left(t_0 + \left(\frac{k+j}{2^n} - t_0\right)\right) - B(t_0)\right| \\
&\quad + \left|B\left(t_0 + \left(\frac{k+j-1}{2^n} - t_0\right)\right) - B(t_0)\right| \\
&\leq M\left[\frac{k+j}{2^n} - t_0 + \frac{k+j-1}{2^n} - t_0\right] \\
&= M\left[\frac{2k+2j-1}{2^n} - 2t_0\right] \\
&\leq M\left[\frac{2k+2j-1}{2^n} - 2\frac{k-1}{2^n}\right] \\
&= \frac{M(2j+1)}{2^n}
\end{aligned}$$

We then define the event

$$\Omega_{n,k} := \left\{ \left| B\left(\frac{k+j}{2^n}\right) - B\left(\frac{k+j-1}{2^n}\right) \right| \leq \frac{M(2j+1)}{2^n} \text{ for } j = 1,2,3 \right\}$$
$$\left(= \bigcap_{j=1}^{3} \left\{ \left| B\left(\frac{k+j}{2^n}\right) - B\left(\frac{k+j-1}{2^n}\right) \right| \leq \frac{M(2j+1)}{2^n} \right\} \right)$$

We must have $1 \leq k \leq 2^n - 3$ so that $\frac{k+j}{2^n} \leq 1$. We can then use independence of increments and also scaling to see that

$$\begin{aligned}
\mathbb{P}(\Omega_{n,k}) &\leq \prod_{j=1}^{3} \mathbb{P}\left(\left| B\left(\frac{k+j}{2^n}\right) - B\left(\frac{k+j-1}{2^n}\right) \right| \leq \frac{M(2j+1)}{2^n} \right) \\
&= \prod_{j=1}^{3} \mathbb{P}\left(\left| \sqrt{2^n} B\left(\frac{k+j}{2^n}\right) - \sqrt{2^n} B\left(\frac{k+j-1}{2^n}\right) \right| \leq \frac{M(2j+1)}{\sqrt{2^n}} \right) \\
&\leq \prod_{j=1}^{3} \mathbb{P}\left(\left| \sqrt{2^n} B\left(\frac{k+j}{2^n}\right) - \sqrt{2^n} B\left(\frac{k+j-1}{2^n}\right) \right| \leq \frac{7M}{\sqrt{2^n}} \right) \\
&= \mathbb{P}\left(|B(1) \leq \frac{7M}{\sqrt{2^n}} \right)^3
\end{aligned}$$

This is apparent as the scaling implies we have $B(t) - B(t-1)$ for a standard Brownian motion, which is the standard normal distribution, $B(1)$. We know the standard normal distribution has equation $\frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}}$ and we know the maximum is when $t = 0$, which is $\frac{1}{\sqrt{2\pi}}$. We can therfore bound this by $\frac{1}{2}$ since $2 < \sqrt{2\pi}$, since $2\pi > 4$, and so $\frac{1}{\sqrt{2\pi}} < \frac{1}{2}$. Using this information, we can compute the following,

$$\mathbb{P}(\Omega_{n,k}) = \left(\int_{\frac{-7M}{\sqrt{2^n}}}^{\frac{7M}{\sqrt{2^n}}} f(t)\mathrm{d}t \right)^3 \leq \left(\int_{\frac{-7M}{\sqrt{2^n}}}^{\frac{7M}{\sqrt{2^n}}} \frac{1}{2}\mathrm{d}t \right)^3 = \left(\frac{7M}{\sqrt{2^n}} \right)^3$$

where $f(t)$ is the distribution of a standard normal variable. Finally, we arrive at

$$\mathbb{P}\left(\bigcap_{k=1}^{2^n-3} \Omega_{n,k} \right) \leq (2^n - 3)(\frac{7M}{\sqrt{2^n}})^3 \leq \frac{(7M)^3}{\sqrt{2^n}}$$

and as $\sum_{n=1}^{\infty} \frac{1}{2^n} < \infty$, by Borel-Cantelli's lemma and using the fact that

$$\left\{ \sup_{h\in[0,1]} \frac{|B(t+h)-B(t)|}{h} \leq M \right\} \subseteq \Omega_{n,k},$$

$$\mathbb{P}\left\{ \text{there is a } t_0 \in [0,1] \text{ such that } \sup_{h\in[0,1]} \frac{|B(t+h)-B(t)|}{h} \leq M \right\}$$
$$\leq \mathbb{P}\left(\bigcap_{k=1}^{2^n-3} \Omega_{n,k} \text{for infinitely many n} \right) = 0$$

we have therefore shown Brownian motion is nowhere differentiable on $[0,1]$. This can then be extended onto the whole domain. This is done by considering $W_t^n := B_{t+n} - B_n$, which is a Brownian motion and noting that $(W_t^n)_{t\in[0,1]}$ is nowhere differentiable.

□

2 Itô Calculus

2.1 Introduction

We just proved that Brownian motion is differentiable nowhere, and this is key for the upcoming section. Normally, we can evaluate an integral using Riemann-Stieltjes Integration.

Definition 2.1 (Riemann-Stieltjes Integration)**.** [3, pp. 9-10]. The integral of a function, f with respect to a monotone function g over an interval $[a,b]$ is defined as

$$\int_a^b f \mathrm{d}g = \int_a^b f(t)\mathrm{d}g(t) = \lim_{\delta\to 0} \sum_{i=1}^n f(\xi_i^n)(g(t_i^n) - g(t_{i-1}^n)),$$

where t_i^n represents the partitions of the interval,

$$a = t_0^n < t_1^n < \cdots < t_n^n = b,\ \delta = \max_{1\leq i\leq n}(t_i^n - t_{i-1}^n)$$

and ξ_i^n is in the i'th sub-interval, $t_{i-1}^n \leq \xi_i^n \leq t_i^n$.

Remark. We can rewrite the integral in its more common form as,

$$\int f(t)\mathrm{d}g(t) = \int f(t)\frac{\mathrm{d}g(t)}{\mathrm{d}t}\mathrm{d}t \tag{2}$$

If we were to consider an integral such as $\int X(t)\mathrm{d}B(t)$, we would run into a problem. To understand the integral, we may try to use Riemann-Stieltjes Integration, particularly the form in (2). However we cannot do this due to the fact $\frac{\mathrm{d}B(t)}{\mathrm{d}t}$ does not exist. For this reason, we need to find an alternative approach to evaluating integrals of this form. The next section takes us through a solution to this problem by defining a new type of integral.

2.2 The Itô Integral

Stochastic calculus would not exist if it weren't for Professor Kiyosi Itô. Starting during the Second World War, he extended calculus to stochastic processes, which is still used heavily today in mathematical finances. For his work in the field, he won the Gauss Prize in 2006.

A stochastic integral (or Itô integral) is defined to be $\int_0^T X(t)\mathrm{d}B(t)$. Clearly we cannot use normal integration techniques to evaluate this as we are integrating with respect to Brownian motion. When defining this integral, the aim was to make sure the following properties were maintained:

1. If $X(t) = 1$, $\int_0^T X(t)\mathrm{d}B(t) = \int_0^T \mathrm{d}B(t) = B(T) - B(0)$.
2. If $X(t) = c \in \mathbb{R}$, $\int_0^T X(t)\mathrm{d}B(t) = c(B(T) - B(0))$
3. $\int_0^a X(t)\mathrm{d}B(t) + \int_a^T X(t)\mathrm{d}B(t) = \int_0^T X(t)\mathrm{d}B(t)$

If these three properties hold, the integral will be defined.

Definition 2.2 (Integration of non-random simple processes)**.** [3, pp. 91-92]. Consider a non-random simple process $X(t)$, not dependent on $B(t)$. Then there is a partition $0 = t_0 < t_1 < \cdots < t_n = T$, with real numbers $c_0, \ldots, c_{n-1}$ such that

$$X(t) = c_0 I_0(t) + \sum_{i=1}^{n-1} c_i I_{(t_i, t_{i+1})}(t),$$

and so the Itô integral is defined as

$$\int_0^T X(t)\mathrm{d}B(t) = \sum_{I=0}^{n-1} c_i(B(t_{i+1}) - B(t_i))$$

Lemma 2.1 (Distribution of Itô Integral for non-random simple processes)**.** *By independence of Brownian increments, we have the integral (and therefore summation) is equivalent to a normal distribution,* $\mu = 0$, $\sigma^2 = c_i^2$.

1. It is clear that for expectation, we have

$$\begin{aligned}
\mathbb{E}\left(\int_0^T X(t)\mathrm{d}B(t)\right) &= \mathbb{E}\left(\sum_{I=0}^{n-1} c_i(B(t_{i+1}) - B(t_i))\right) \\
&= \sum_{I=0}^{n-1} \mathbb{E}[c_i(B(t_{i+1}) - B(t_i))] \\
&= \sum_{I=0}^{n-1} c_i \mathbb{E}[(B(t_{i+1}) - B(t_i))] = 0
\end{aligned}$$

This is clear from property i of Brownian motion (the independence of increments). □

2. The proof of variance is similar. We have by following the same logic as above,

$$\begin{aligned}\mathrm{Var}\left(\int_0^T X(t)\mathrm{d}B(t)\right) &= \sum_{I=0}^{n-1} \mathrm{Var}[c_i(B(t_{i+1}) - B(t_i))] \\ &= \sum_{I=0}^{n-1} c_i^2 \mathrm{Var}((B(t_{i+1}) - B(t_i))) = c_i^2\end{aligned}$$

This is clear by property ii of Brownian motion (the normally distributed increments). □

The integral can be generalised to random variables by replacing c_i with ξ_i, where ξ_i are random variables that depend on $B(t), t \leq t_i$. Note that ξ_i does not depend on future values of $B(t)$. In order to do this, we make ξ_i $\mathcal{F}_{t_i}$-measurable.

Definition 2.3 (Integration of a simple adapted process)**.** [3, p. 93]. $X = \{X(t),\, 0 \leq t \leq T\}$ is a simple adapted process if there exists t_i such that $0 = t_0 < t_1 < \cdots < t_n = T$ and random variables $\xi_0, \ldots, c_{n-1}$ where ξ_0 is a constant. The ξ_i are $\mathcal{F}_{t_i}$-measurable. Then

$$X(t) = \xi_0 I_0(t) + \sum_{i=1}^{n-1} \xi_i I_{(t_i, t_{i+1})}(t)$$

such that the Itô Integral

$$\int_0^T X(t)\mathrm{d}B(t) = \sum_{I=0}^{n-1} \xi_i(B(t_{i+1}) - B(t_i))$$

Remark. This does not follow the normal distribution as the previous integrals did.

There are many important properties of this integral, the proofs of which are outlined in Gardiner et al. [2, §4.2] and in Durrett [1, §2].

i Linearity. If α and β are constants, and $X(t), Y(t)$ are two simple processes, then

$$\int_0^T (\alpha X(t) + \beta Y(t))\mathrm{d}B(t) = \alpha \int_0^T X(t)\mathrm{d}B(t) + \beta \int_0^T Y(t)\mathrm{d}B(t)$$

ii For the indicator function $I_{(a,b]}(t)$, we have that

$$\int_0^T I_{(a,b]}(t)\mathrm{d}B(t) = B(b) - B(a), \quad \int_0^T I_{(a,b]}(t)X(t)\mathrm{d}B(t) = \int_a^b X(t)\mathrm{d}B(t)$$

iii The Zero Mean Property: $\mathbb{E}\left[\int_0^T X(t)\mathrm{d}B(t)\right] = 0$

iv The Isometry Property:

$$\mathbb{E}\left[\int_0^T X(t)\mathrm{d}B(t)\right]^2 = \int_0^T \mathbb{E}[X^2(t)]\mathrm{d}t$$

Remark. One of the most useful properties is the isometry property. It is used in the proof of the existence and uniqueness of Itô integrals and stochastic differential equations. These are not outlined in this essay but can be found in Øksendal [4, §5.2].

2.3 Itô's Formula and Itô's Process

Itô calculus is different to ordinary calculus. When evaluating an ordinary differential equation, we have (using the chain rule) that $f(g(t)) - f(g(a)) = \int_a^t f'(g(s))g'(s)\mathrm{d}s$. However, the chain rule for Itô Calculus in the simplest form has a different form.

Theorem 2.2 (Itô's Formula)**.** *[1, p. 68]. If $B(t)$ is a Brownian motion on $[0, T]$, and $f(x)$ is a twice continuously differentiable function on $\mathbb{R}$, then for any $t \leq T$,*

$$f(B(t)) = f(0) + \int_0^t f'(B(s))\mathrm{d}B(s) + \frac{1}{2}\int_0^t f''(B(s))\mathrm{d}s \tag{3}$$

Proof. We will walk through the first part of the proof and then outline the second part. If we let $\{t_i^n\}$ be a partition of the interval $[0, t]$, then by telescoping series, we have

$$f(B(t)) - f(0) = \sum_{i=0}^{n-1} \left[f(B(t_{i+1}^n)) - f(B(t_i^n))\right]$$

$$\implies f(B(t)) = f(0) + \sum_{i=0}^{n-1} \left[f(B(t_{i+1}^n)) - f(B(t_i^n))\right] \tag{4}$$

Then, by applying Taylor's formula, we find that for some θ_i^n in the interval $(B(t_i^n), B(t_i^n))$,

$$\begin{aligned} f(B(t_{i+1}^n)) - f(B(t_i^n)) = {} & f'(B(t_i^n))\left[B(t_{i+1}^n)) - f(B(t_i^n)\right] \\ & + \frac{1}{2}f''(\theta_i^n)\left[B(t_{i+1}^n)) - f(B(t_i^n)\right]^2 \end{aligned}$$

Putting this into (4), we have

$$\begin{aligned} f(B(t)) = f(0) + {} & \sum_{i=0}^{n-1} f'(B(t_i^n))\left[B(t_{i+1}^n)) - f(B(t_i^n)\right] \\ & + \frac{1}{2}\sum_{i=0}^{n-1} f''(\theta_i^n)\left[B(t_{i+1}^n)) - f(B(t_i^n)\right]^2 \end{aligned}$$

If we take the limit as $\delta_n \to 0$, it is clear our first summation converges to $\int_0^t f'(B(s))\mathrm{d}B(s)$ from the definition of an Itô integral for a simple adapted process. This is exactly what we needed to show.

The second summation converges to $\frac{1}{2}\int_0^t f''(B(s))\mathrm{d}s$, but this is beyond the scope the essay. It uses the quadratic property of Brownian motion which we have not met. The proof shows that the summation converges in probability to $\int_0^t g(B(s))\mathrm{d}s$. The whole proof is shown in Klebaner [3, pp. 106-107]. $\square$

Itô's formula can be used to evaluate more complex integrals. We can now easily evaluate polynomials of Brownian motion and exponentials.

Example 2.1. If $f(x) = x^2$, applying Itô's formula we find that

$$B^2(t) = 2\int_0^t B(s)\mathrm{d}B(s) + t,$$

as $B^0(t) = 0$. Rearranging, we find that $\int_0^t B(s)\mathrm{d}B(s) = \frac{1}{2}(B^2(t) - t)$. This once again shows us that Itô calculus is very different to standard calculus, where we would normally see expressions such as $\int_0^t x\mathrm{d}x = \frac{t^2}{2}$.

Example 2.2. Evaluate the stochastic integral $\int_0^t \frac{\mathrm{d}B(s)}{1+B(s)^2}$.

We can rearrange Itô's formula (3) to compute this. If we let $f'(x) = \frac{1}{1+x^2}$, we see that $f(x) = \arctan(x) + C$ where C is a constant, and $f''(x) = \frac{-2x}{(1+x^2)^2}$. We then have

$$\begin{aligned} \int_0^t f'(B(s))\mathrm{d}B(s) &= f(B(t)) - f(0) - \frac{1}{2}\int_0^t f''(B(s))\mathrm{d}s \\ \implies \int_0^t \frac{\mathrm{d}B(s)}{1+B(s)^2} &= \arctan(B(t)) + \int_0^t \frac{B(s)}{(1+B(s)^2)^2}\mathrm{d}s \end{aligned}$$

Itô also proved another identity known as the Itô process. This is used to find solutions of stochastic differentials.[2]

Definition 2.4 (Itô Process)**.** [3, p. 108]. An Itô process has the form

$$Y(t) = Y(0) + \int_0^t \mu(s)\mathrm{d}s + \int_0^t \sigma(s)\mathrm{d}B(s), \quad 0 \le t \le T \tag{5}$$

where $Y(0)$ is $\mathcal{F}_0$-measurable and $\mu(t)$ and $\sigma(t)$ are $\mathcal{F}_t$-adapted so that $\int_0^T |\mu(t)|\mathrm{d}t$ and $\int_0^T |\sigma(t)|\mathrm{d}t$ are finite.

We can also say that the process $Y(t)$ has the stochastic differential on $[0, T]$

$$\mathrm{d}Y(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}B(t), \quad 0 \le t \le T \tag{6}$$

[2] Unlike normal calculus that Newton and Leibnitz invented, Itô Calculus was created for the purpose of solving stochastic differential equations, so there is no simple geometric interpretation of Itô integrals.

In this theorem, 5 and 6 are equivalent. As mentioned earlier, we can use the Itô process to solve stochastic differentials.

Example 2.3. We saw in Example 2.1 that $B^2(t) = 2\int_0^t B(s)\mathrm{d}B(s) + t$. If we let $Y(t) = B^2(t)$, we have $Y(t) = 2\int_0^t B(s)\mathrm{d}B(s) + \int_0^t \mathrm{d}s$. Clearly then by setting $\mu(s) = 1$ and $\sigma(s) = 2B(s)$, we find the stochastic differential of $B^2(s)$ is $\mathrm{d}(B^2(t)) = \mathrm{d}t + 2B(t)\mathrm{d}B(t)$.

Tying the two main theorems that we just met together, we have a final key tool for solving stochastic differential equations.

Theorem 2.3 (Itô's Formula for Itô Processes)**.** *Let $X(t)$ have a stochastic differential for $0 \le t \le T$.*

$$\mathrm{d}X(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}B(t) \tag{7}$$

If $f(x)$ is twice continuously differentiable, then the stochastic differential of the process $Y(t) = f(X(t))$ is given by

$$\mathrm{d}f(X(t)) = \left(f'(X(t))\mu(t) + \frac{1}{2}f''(X(t))\sigma^2(t)\right)\mathrm{d}t + f'(X(t))\sigma(t)\mathrm{d}B(t). \tag{8}$$

Meaning that, $f(X(t)) = f(X(0)) + \int_0^t f'(X(s))\mathrm{d}X(s) + \frac{1}{2}\int_0^t f''(X(s))\sigma^2(s)\mathrm{d}s$.

This is incredibly important and we will be using this to conclude the essay with a key example. The proof of this theorem is similar to the proof of Theorem 2.2, and is outlined in Øksendal [4, §4.1].

2.4 Stochastic Differential Equations Examples and Applications

We will finish this essay with a well-known example of stochastic differential equations. Firstly we must define a few terms.

Definition 2.5 (Strong Solutions to Stochastic Differential Equations)**.** Let $B(t)$ be a Brownian motion process. An equation of the form

$$\mathrm{d}X(t) = \mu(X(t), t)\mathrm{d}t + \sigma(X(t), t)\mathrm{d}B(t) \tag{9}$$

where $X(t)$ is unknown and $\mu(x, t)$ and $\sigma(x, t)$ are called the drift and diffusion[3] coefficients respectively.

$X(t)$ is a strong solution to equation (9) if the integral $\int_0^t \mu(X(s), s)\mathrm{d}s$ and the Itô integral $\int_0^t \sigma(X(s), s)\mathrm{d}B(s)$ exist, with

$$X(t) = X(0) + \int_0^t \mu(X(s), s)\mathrm{d}s + \int_0^t \sigma(X(s), s)\mathrm{d}B(s). \tag{10}$$

[3]Itô initially used stochastic differential equations to construct diffusion processes, therefore these names arise from there.

Our final definition is about white noise. This typically arises when forming stochastic differential equations.

Definition 2.6 (White Noise)**.** [3, p. 124] White noise, $\xi(t)$ is defined as the derivative of Brownian motion.

$$\xi(t) = \frac{\mathrm{d}B(t)}{\mathrm{d}t}$$

We have already shown this does not exist in Theorem 1.4, however its use is as follows. If $\sigma(x,t)$ is the intensity of noise at point x and time t, then $\int_0^T \sigma(X(t),t)\xi(t)\mathrm{d}t = \int_0^T \sigma(X(t),t))\mathrm{d}B(t)$.

Finally using these two definitions, we can meet an important example with real world applications.

Example 2.4 (Black-Scholes-Merton Model for growth at an uncertain rate of return)**.** Let $x(t)$ be the value of \$1 after time t invested into an account. By compound interest, it satisfies the differential equation $\frac{\mathrm{d}x(t)}{x(t)} = r\mathrm{d}t$, where r is the interest rate. If r is uncertain, we can say there is some noise $r+\sigma\xi(t)$ and we reach the following stochastic differential equation.

$$\begin{aligned} \frac{\mathrm{d}X(t)}{X(t)} &= (r+\sigma\xi(t))\mathrm{d}t, \quad X(0)=1 \\ \implies \frac{\mathrm{d}X(t)}{X(t)} &= r\mathrm{d}t + \sigma\mathrm{d}B(t) \end{aligned}$$

Noticing that this the left hand side looks similar to the derivative of $\log(X(t))$, we will multiply by $X(t)$ and apply Theorem 2.3 to $Y(t) = \log(X(t))$ and we let $r = \mu$ for consistency. This gives,

$$\begin{aligned} \mathrm{d}(\log(X(t))) &= \left(\frac{1}{X(t)}\mu X(t)\mathrm{d}t + \frac{1}{2}\left(\frac{-1}{X^2(t)}\right)\sigma^2X^2(t)\right)\mathrm{d}t + \frac{1}{X(t)}\sigma X(t)\mathrm{d}B(t) \\ &= \left(\mu - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}B(t) \\ \implies \mathrm{d}Y(t) &= \left(\mu - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}B(t) \\ \implies Y(t) &= Y(0) + \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t) \\ \implies X(t) &= X(0)e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma B(t)} \qquad (11) \\ \therefore X(t) &= e^{(r-\frac{1}{2}\sigma^2)t+\sigma B(t)} \end{aligned}$$

Equation (11) is extremely important in financial mathematics. It is known as Geometric Brownian Motion and is used to model financial stock prices. A graph of it is shown in Figure 4 and 5.

3 Figures and Python Code

3.1 Figure 1

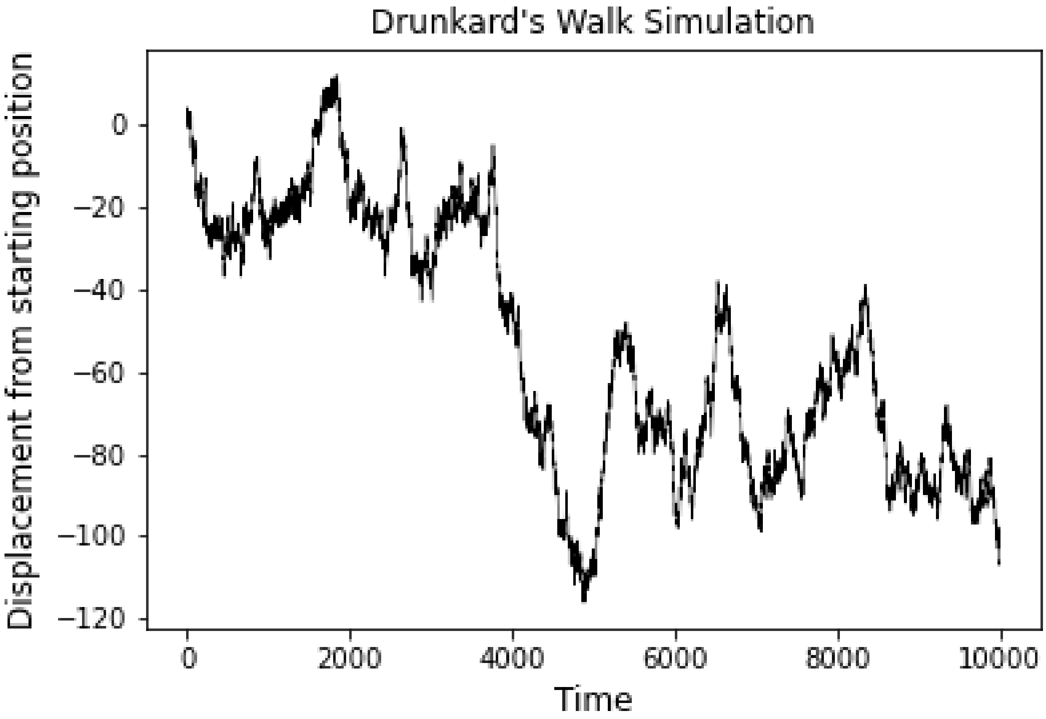

Figure 1: Simulation of the drunkard's walk over 10,000 steps

```
import numpy as np
import matplotlib.pyplot as plt
import random

def randomwalk1D(n):
    x = 0
    y = 0

    ts = np.arange(n + 1) #timesteps
    height = [y]
    move = [-1, 1]
    for i in range(1, n + 1):
        # Randomly select either UP or DOWN
        step = random.choice(move)
```

```
        # Move the object up or down
        if step == 1:
            y += 1
        elif step == -1:
            y -= 1
        # Keep track of the positions
        height.append(y)
    return ts, height

walk = randomwalk1D(10000)
plt.plot(walk[0], walk[1], color = 'black', lw = 0.5)
plt.title("Drunkard's Walk Simulation")
plt.xlabel('Time', fontsize = 12)
plt.ylabel('Displacement from starting position', fontsize = 12)
plt.savefig('fig1.png')
plt.show()
```

Listing 1: Python Code for Figure 1 – Simulation of Drunkard's Walk over 10,000 steps

3.2 Figure 2

```
walk1 = randomwalk1D(10000)
walk2 = randomwalk1D(10000)
walk3 = randomwalk1D(10000)
plt.plot(walk1[0], walk1[1], color = 'blue', lw = 0.5)
plt.plot(walk2[0], walk2[1], color = 'red', lw = 0.5)
plt.plot(walk3[0], walk3[1], color = 'green', lw = 0.5)
plt.title("3 Drunkard Walk Simulations (1D)")
plt.xlabel('Time', fontsize = 12)
plt.ylabel('Displacement from starting position', fontsize = 12)
plt.savefig('fig2.png')
plt.show()
```

Listing 2: Python Code for Figure 2 – Simulation of 3 Drunkard's Walks over 10,000 steps

3.3 Figure 3

```
def randomwalk2D(n):
    x = np.zeros(n)
    y = np.zeros(n)

    move = ['l','r','u','d','ul','ur','dl','dr']
```

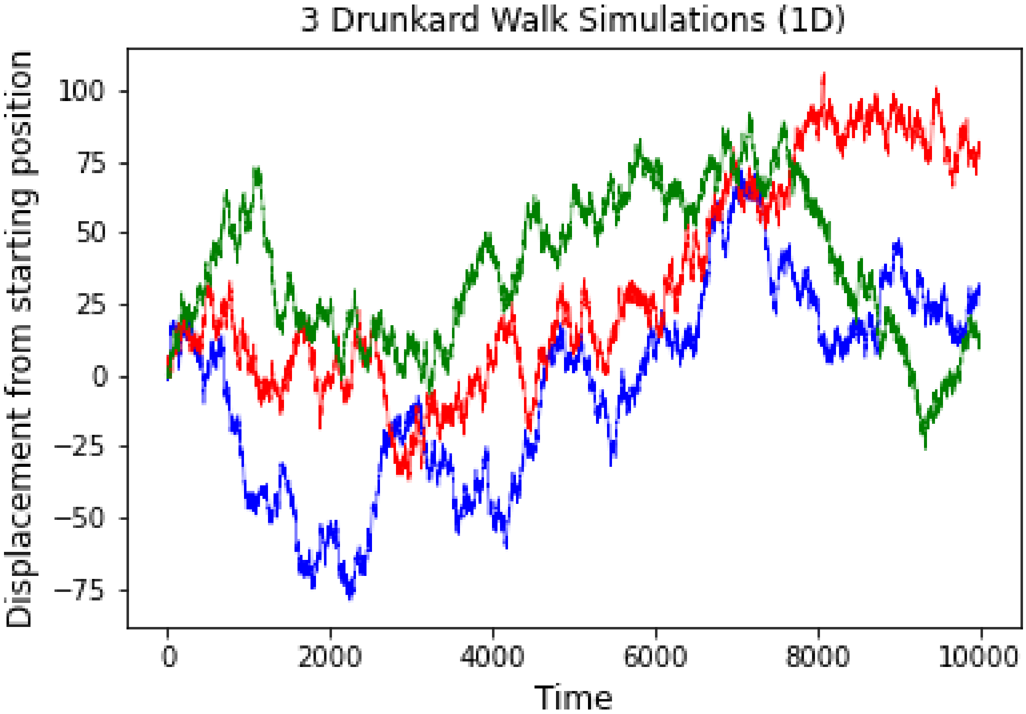

Figure 2: Simulation of 3 drunkards' walks over 10,000 steps

```
for i in range(1, n):
    # Randomly select either UP or DOWN
    step = random.choice(move)

    # Move the object up or down
    if step == 'r':
        x[i] = x[i-1]+1
        y[i] = y[i-1]
    elif step == 'l':
        x[i] = x[i-1]-1
        y[i] = y[i-1]
    elif step == 'u':
        y[i] = y[i-1]+1
        x[i] = x[i-1]
    elif step == 'd':
        y[i] = y[i-1]-1
        x[i] = x[i-1]
    elif step == 'ul':
        x[i] = x[i-1]-1
```

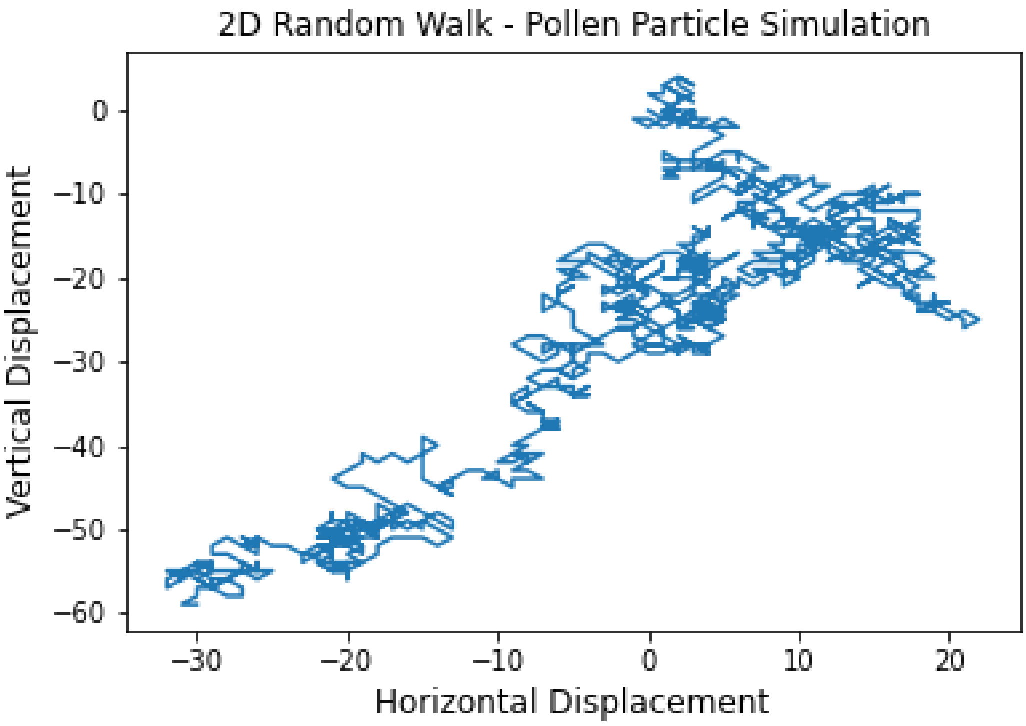

Figure 3: Simulation of the random motion of a pollen particle , starting from the origin

```
            y[i] = y[i-1]+1
        elif step == 'ur':
            y[i] = y[i-1]+1
            x[i] = x[i-1]+1
        elif step == 'dl':
            y[i] = y[i-1]-1
            x[i] = x[i-1]-1
        elif step == 'dr':
            y[i] = y[i-1]-1
            x[i] = x[i-1]+1
        # Keep track of the positions
    return x, y
x_data, y_data = randomwalk2D(1000)
plt.title("2D Random Walk - Pollen Particle Simulation")
plt.xlabel("Horizontal Displacement", fontsize = 12)
plt.ylabel("Vertical Displacement", fontsize = 12)
plt.plot(x_data, y_data)
plt.savefig('fig3.png')
```

```
plt.show()
```

Listing 3: Python Code for Figure 3 – Simulation of pollen particle

3.4 Figure 4

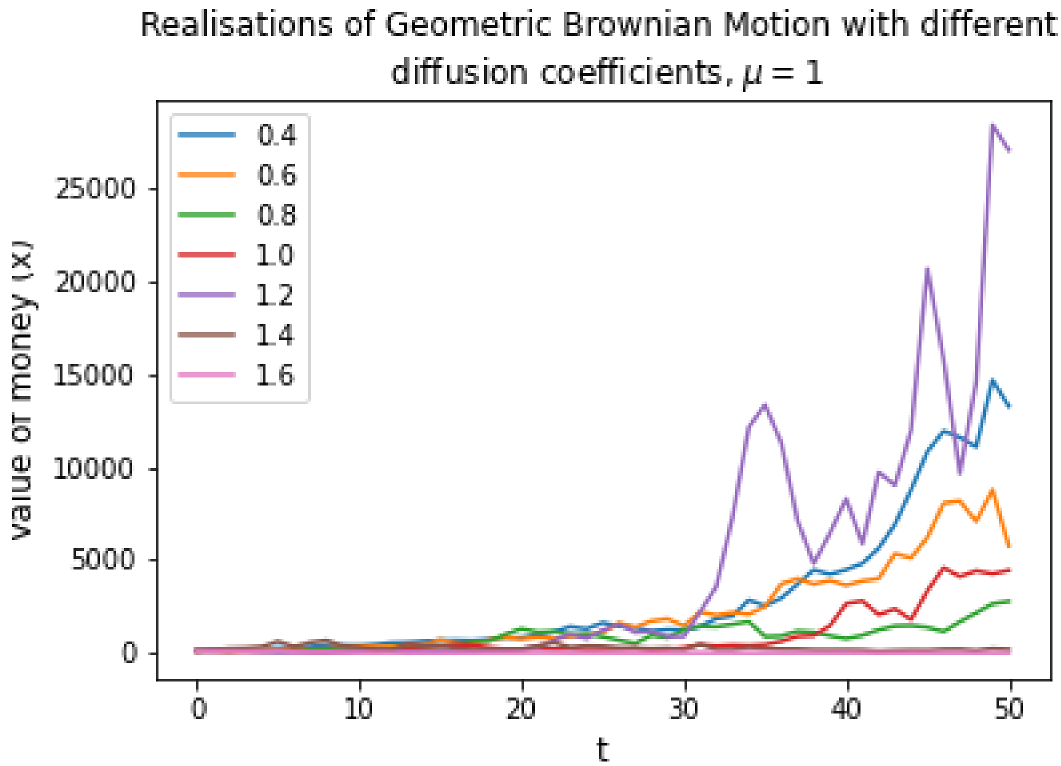

Figure 4: Realisations of geometric Brownian motion for different values of σ

```
mu = 1
n = 50
dt = 0.1
x0 = 100 #initial value (must have some money initially)

sigma = np.arange(0.4, 1.6, 0.2)

#Model Brownian motion term as normal distribution, mean=0, var = sqrt(t)
#due to quadratic property of Brownian motion
x = np.exp(
    (mu - sigma ** 2 / 2) * dt
    + sigma * np.random.normal(0, np.sqrt(dt), size=(len(sigma), n)).T
```

```
)

x = np.vstack([np.ones(len(sigma)), x])
x = x0 * x.cumprod(axis=0) #'compound interest'

plt.plot(x)
plt.legend(np.round(sigma, 2))
plt.xlabel("t")
plt.ylabel("Value of money (x)")
plt.title(
    "Realisations of Geometric Brownian Motion with different \n diffusio
)
plt.savefig('fig4.png')
plt.show()
```

Listing 4: Python Code for Figure 4 – Realisations of Geometric Brownian Motion

3.5 Figure 5

```
X0 = 100
tf = 5
mu = 0.02
sigma = 0.2
Npaths = 10
Nsteps = 365 * tf

def SDE_GBM(X0, tf, mu, sigma, Npaths): # Define as asked
    t, dt = np.linspace(0 ,tf , Nsteps + 1, retstep = True)
    X = np.zeros((Nsteps + 1, Npaths))
    root_dt = np.sqrt(dt)

    X[0, :] = X0
    for n in range (Nsteps):
        X_of_t = mu * X[n, :]
        X[n + 1, :] = X[n, :] + dt * (X_of_t) + (sigma * X[n, :] * root_d
        #SDE Equation. Express dB(t) as sqrt(dt)*Random Normal Distributi

    return t, X

# Plotting
t , X = SDE_GBM(X0, tf, mu, sigma, Npaths)
plt.plot(t, X)
plt.title('Simulation of Stock Prices, \n $X_0=100$, $\mu=0.02$, $\sigma=
```

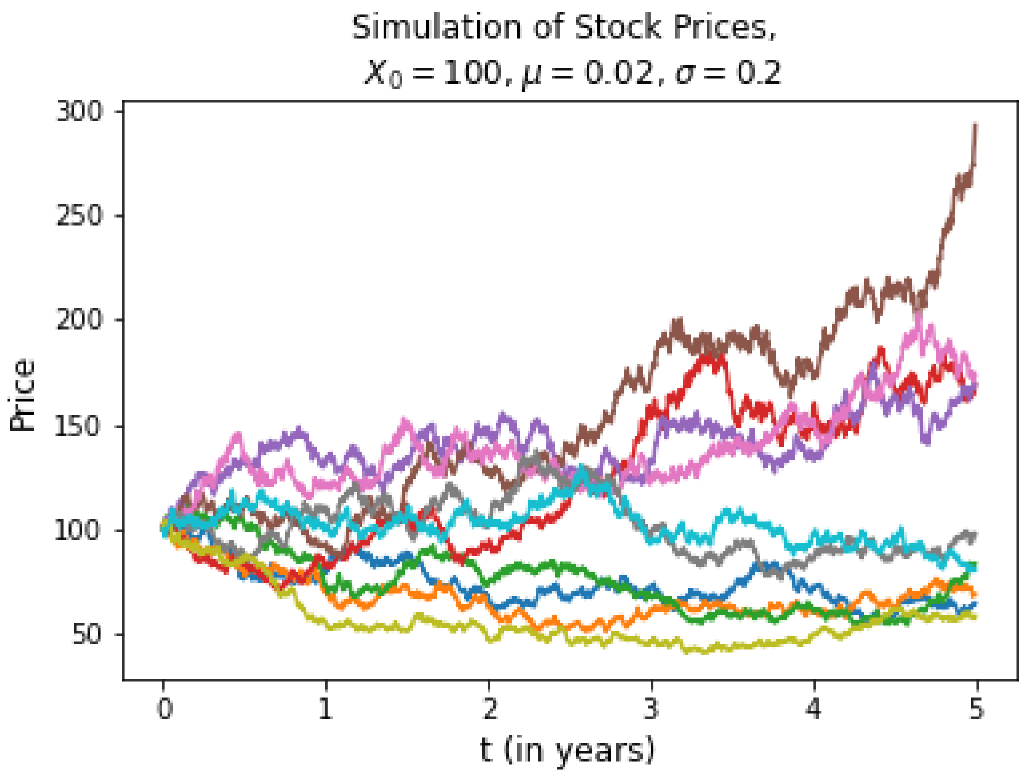

Figure 5: Simulations of stock prices to demonstrate use of geometric Brownian motion SDE

```
plt.xlabel('t (in years)', fontsize = 12)
plt.ylabel('Price', fontsize = 12)
plt.savefig('fig5.png')
plt.show()
```

Listing 5: Python Code for Figure 5 – Simulation of Stock Price Graphs

References

[1] Richard Durrett. *Stochastic calculus: a practical introduction.* CRC press, 2018.

[2] Crispin W Gardiner et al. *Handbook of stochastic methods for physics, chemistry and the natural sciences.* 3rd edition, 2004.

[3] Fima C Klebaner. *Introduction to stochastic calculus with applications.* World Scientific Publishing Company, 3rd edition, 2012.

[4] Bernt Øksendal. *Stochastic differential equations: an introduction with applications.* Springer Science & Business Media, 6th edition, 2013.

[5] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. Springer, 2006.

[6] James Leiner. Brownian motion and the strong markov property. `https://math.uchicago.edu/~may/REU2012/REUPapers/Leiner.pdf`, 2012. Accessed on 17th March 2022.

[7] Aaron Mcknight. Some basic properties of brownian motion. `https://www.math.uchicago.edu/~may/VIGRE/VIGRE2009/REUPapers/McKnight.pdf`, 2009. Accessed on 19th March 2022.