Contents

1 Maths Basics

1.1 Functions

Definition 1.1. A *function* is a map from one set to another. We write $f : X \to Y$ which f is the function which maps elements $x \in X$ to $y \in Y$. Another restriction we have is that for a given element $x \in X$, f maps it to at most one element $y \in Y$.

There are many different ways to write a function, let us consider the doubling function then we can write:

$$\begin{aligned} f(x) &= x^2 \\ x &\mapsto x^2 \\ f : x &\mapsto x^2 \end{aligned}$$

Definition 1.2. Let $f : X \to Y$ be a function. Then X is the *domain* of f and Y is the *image* of f.

Definition 1.3. Let $f : X \to Y$ be a function.

1. If $\forall x \in X, \exists! y \in Y : f(x) = y$ then f is a *one to one* mapping.
2. If $\exists y \in Y, x_1, x_2 \in X : f(x_1) = f x_2 = y$ then f is a *many to one* mapping.
3. Let $g : X \to Y$ be a mapping such that $\exists x \in X, y_1, y_2 \in Y : f(x) = y_1 = y_2$ then f is a *one to many* mapping and it is not a function.

Definition 1.4. Let $f : X \to Y$ be a function. We say f^{-1} is the *inverse* map of f if the following holds:

$$f^{-1}(f(x)) = f\left(f^{-1}(x)\right) = x, \quad \forall x \in X$$

Definition 1.5. Let $f : X \to Y$ be a function.

1. We say f is an *even* function if $f(-x) = f(x), \forall x \in X$.
2. We say f is an *odd* function if $f(-x) = -f(x), \forall x \in X$.
3. Let $z \in X$, we say f is *periodic* with period z if $f(x+z) = f(x), \forall x \in X$.

Definition 1.6. Let $f : X \to Y$ be a function.

1. We call f an *explicit* function if $forally \in Y$ we can write $y = f(x)$ for some $x \in X$ where all the terms of x are on the right hand side of the equation. Sometimes, we will not be able to do that so that we have to write it as $f(x, y) = 0$ which is called an *implicit* function.

The following is an example of an implicit function as we can write it with just y on the left side of the equation with all the terms of x on the right hand side:

$$4y^4 - 2y^2x^2 - yx^2 + x^2 + 3 = 0$$

Definition 1.7. Let $f : X \to Y$ be a function. We call f an *n-th degree polynomial* if $\exists a_k$ for $k = 1, \dots, n$ for some $n \in \mathbb{N}$ such that we can write f as:

$$f(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

Definition 1.8. Let $f : X \to Y$ be a polynomial. f is a *linear* polynomial if $n = 1$, i.e. the degree of the highest x term is 1. We call f a *quadratic* polynomial if $n = 2$, i.e. the highest x term has degree 2.

Lemma 1.9. Suppose we are looking at the quadratic equation $ax^2 + bx + c = 0$ then the solution to this is given by the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Proof.

$$\begin{aligned}
ax^2 + bx + c &= 0 \\
a\left(x^2 + \frac{bx}{a}\right) + c &= 0 \\
\left(x^2 + \frac{bx}{a}\right) &= -\frac{c}{a} \\
\left(x - \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} &= -\frac{c}{a} \\
\left(x - \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a^2} \\
x - \frac{b}{2a} &= \frac{\pm\sqrt{b^2 - 4ac}}{2a} \\
x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\end{aligned}$$

□

Proposition 1.10. Suppose we have the quadratic equation $ax^2 + bx + c = 0$ then we have 3 cases in terms of the roots of the equation (because of the square root):

1. $b^2 - 4ac > 0$, then we have two distinct real roots $x_1 \neq x_2 \in \mathbb{R}$ where$x_1 = \left(-b + \sqrt{b^2 - 4ac}\right) / (2a)$ and $x_2 = \left(-b - \sqrt{b^2 - 4ac}\right) / (2a)$.
2. $b^2 - 4ac = 0$, then we have one real root $x_1 = x_2 \in \mathbb{R}$ where $x_1 = x_2 = -b/(2a)$.
3. $b^2 - 4ac < 0$, then we have a complex conjugate pair roots $x_1 \neq x_2 \in \mathbb{C}$ where $x_1 = \left(-b + i\sqrt{-b^2 + 4ac}\right) / (2a)$ and $x_2 = \left(-b - i\sqrt{-b^2 + 4ac}\right) / (2a)$.

Definition 1.11. We define the *modulus* function $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ as the following:

$$f(x) = |x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$$

Definition 1.12. Let $f : X \to Y$ be a function and $x_0 \in \mathbb{R}$. We define the *limit* of f at x_0 if there exists $l \in Y$ such that:

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in R : |x - x_0| < \delta \Rightarrow |f(x) - l| < \epsilon$$

Then we say the limit of f at x_0 is l or we can write this in the following ways:

$$\begin{aligned} \lim_{x \to x_0} f(x) &= l \\ f(x) \to l \quad &\text{as} \quad x \to x_0 \end{aligned}$$

We write $x \to x_0^-$ to denote the limit from the left and $x \to x_0^+$ to denote the limit from the right.

This limit only *exists* if we have the limit from every direction exists. For us, the following has to hold:

$$\lim_{x \to x_0^-} f(x) = \lim_{x \to x_0^+} f(x)$$

To solve for limits, the easiest way is to divide by the 'strongest' term throughout the expression then work from there.

Definition 1.13. Let $f : X \to Y$ be a function. We say f is *continuous* at $x_0 \in X$ if:

$$\lim_{x \to x_0} f(x) = f(x_0)$$

Definition 1.14. Let $a, x, y \in R$, suppose we have the following equation:

$$a^x = y$$

Then we define the *logarithm* function with base a, denoted by $\log_a(\cdot)$, such that we have:

$$\log_a(y) := x$$

This can be also thought of as the inverse of the exponent function $f(x) = a^x$.

Definition 1.15. Let $e \in \mathbb{R} \backslash \mathbb{Q}$ which satisfies the following:

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

Then we define the *exponential* function, denoted by $\exp(\cdot)$ or $e^{(\cdot)}$, by:

$$\exp(x) := e^x = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n$$

The inverse of the exponential function is called the *natural logarithm* function, denoted by $\log(\cdot)$, $\log_e(\cdot)$ or $\ln(\cdot)$, which satisfies the following:

$$\log(e^x) = e^{\log(x)} = x$$

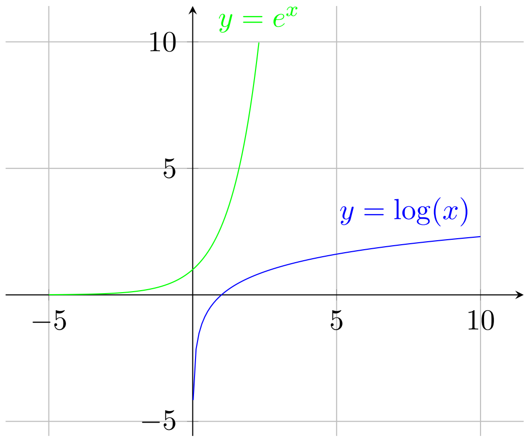

Figure 1: Exponential and Logarithm Functions

Proposition 1.16. Let $x, y \in R$, then we have the following:

- $e^{x+y} = e^x e^y$
- $(e^x)^y = e^{xy}$
- $e^{-x} = 1/e^x$
- $e^0 = 1$
- $\log(xy) = \log(x) + \log(y)$
- $\log(1/x) = -\log(x)$
- $\log(x/y) = \log(x) - \log(y)$
- $\log(x^y) = y \log(x)$
- $\log(1) = 0$

Definition 1.17. Suppose we have a right angled triangle, let $\theta \neq \pi/2$ be an angle in the right angled triangle, the length of the opposite side to *theta* be O, the length of the hypotenuse (the longest side of the triangle) be H and the length of other adjacent side be A.

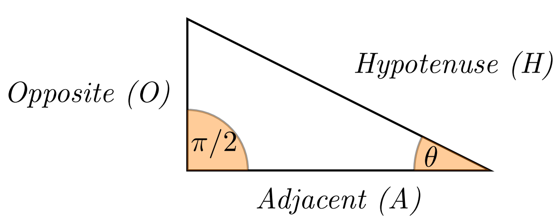

Figure 2: Right Angled Triangle

We define the *sine* function, denoted by $\sin(\cdot)$, using the following:

$$\sin(\theta) := \frac{O}{H}$$

We define the *cosine* function similarly, denoted by $\cos(\cdot)$, using the following:

$$\cos(\theta) := \frac{A}{H}$$

Finally, we define the *tangent* function, denoted by $\tan(\cdot)$, using the following:

$$\tan(\theta) := \frac{O}{A}$$

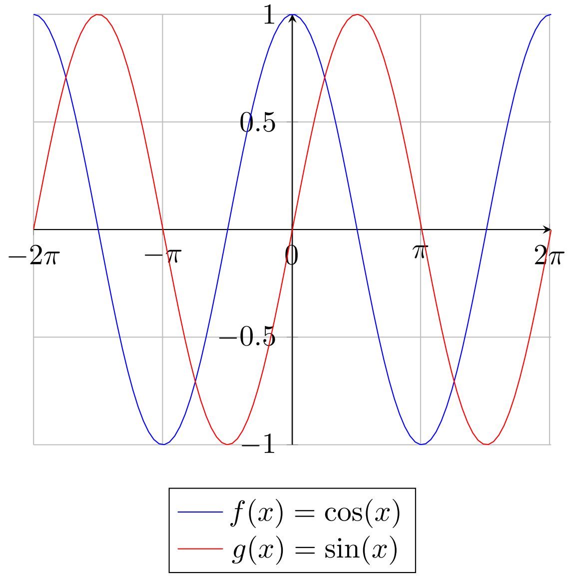

Figure 3: Sine and Cosine Functions

Proposition 1.18.

- Sine function is an odd function so $\sin(-x) = -\sin(x)$ and it is 2π periodic so $\sin(x + 2\pi) = \sin(x)$ for all $x \in \mathbb{R}$.
- Cosine function is an even function so $\cos(-x) = \cos(x)$ and it is 2π periodic so $\cos(x + 2\pi) = \cos(x)$ for all $x \in \mathbb{R}$.
- We have:
$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \forall x \in \mathbb{R}$$
Tangent function is an odd function so $\tan(-x) = -\tan(x)$ and it is π periodic so $\tan(x + \pi) = \tan(x)$ for all $x \in \mathbb{R}$.
- $$\cos^2(x) + \sin^2(x) = 1$$
- $$\sin(x \pm y) = \sin(x)\cos(y) \pm \sin(y)\cos(y)$$
- $$\cos(x \pm y) = \cos(x)\cos(y) \mp \sin(x)\sin(y)$$

- $$\tan(x \pm y) = \frac{\tan(x) + \tan(y)}{1 \mp \tan(x)\tan(y)}$$

- $$\sin(x + \pi/2) = \cos(x), \quad \cos(\pi/2 - x) = \sin(x)$$

Definition 1.19. We define reciprocals of the trigonometric functions *cosecant*, *secant* and *cotangent* by the following:

$$\csc(x) := \frac{1}{\sin(x)}, \quad \sec(x) := \frac{1}{\cos(x)}, \quad \cot(x) := \frac{1}{\tan(x)}$$

Definition 1.20. We define the *hyperbolic* functions as the following:

$$\begin{aligned} \sinh(x) &:= \frac{e^x - e^{-x}}{2} \\ \cosh(x) &:= \frac{e^x + e^{-x}}{2} \\ \tanh(x) &:= \frac{\sinh(x)}{\cosh(x)} \end{aligned}$$

Proposition 1.21.

- $$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

- $$\cosh^2(x) - \sinh^2(x) = 1$$

- $$\sinh(x + y) = \sinh(x)\cosh(y) + \sinh(y)\cosh(x)$$

- $$\cosh(x + y) = \cosh(x)\cosh(y) + \sinh(x)\sinh(y)$$

Lemma 1.22. The inverse function for $\sinh(\cdot)$ is:

$$\sinh^{-1}(x) = \log\left(\left|x + \sqrt{x^2 + 1}\right|\right)$$

Proof.

$$\begin{aligned} y &= \sinh^{-1}(x) \\ \sinh(y) &= x \\ \frac{e^y - e^{-y}}{2} &= x \\ e^{2y} - 2xe^y - 1 &= 0 \\ e^y &= \frac{2x \pm \sqrt{4x^2 + 4}}{2} \\ e^y &= x \pm \sqrt{x^2 + 1} \\ & \quad (\text{As } \sqrt{x^2 + 1} > x \text{ and } e^y > 0, \forall y \in \mathbb{R}) \\ e^y &= x + \sqrt{x^2 + 1} \\ y &= \log\left(\left|x + \sqrt{x^2 + 1}\right|\right) \end{aligned}$$

$\square$

Lemma 1.23.

- $$\cosh^{-1}(x) = \log\left(\left|x + \sqrt{x^2 - 1}\right|\right)$$

- $$\tanh^{-1}(x) = \frac{1}{2}\log\left(\left|\frac{1+x}{1-x}\right|\right)$$

Lemma 1.24. The following tables describes the domains and images of the previous functions we have looked at:

Function	**Domain**	**Image**	**Odd/Even/Neither**
$\exp(\cdot)$	$\mathbb{R}$	$\mathbb{R}_{>0}$	Neither
$\log(\cdot)$	$\mathbb{R}_{>0}$	$\mathbb{R}$	Neither
$\sin(\cdot)$	$\mathbb{R}$	$[-1, 1]$	Odd
$\cos(\cdot)$	$\mathbb{R}$	$[-1, 1]$	Even
$\tan(\cdot)$	$\mathbb{R} \backslash \{(2n+1)\pi/2 : n \in \mathbb{Z}\}$	$\mathbb{R}$	Odd
$\sinh(\cdot)$	$\mathbb{R}$	$\mathbb{R}$	Odd
$\cosh(\cdot)$	$\mathbb{R}$	$[1, \infty)$	Even
$\tanh(\cdot)$	$\mathbb{R}$	$(-1, 1)$	Odd
$\sinh^{-1}(\cdot)$	$[1, \infty)$	$\mathbb{R}_{\geq 0}$	Odd
$\cosh^{-1}(\cdot)$	$\mathbb{R}$	$[1, \infty)$	Neither
$\tanh^{-1}(\cdot)$	$(-1, 1)$	$\mathbb{R}$	Odd

Table 1: Domains, images and other properties of key functions

1.2 Differentiation

We would want to know about the rate of change of a function with elements in its domain.

Definition 1.25. Let $f : X \to Y$ be a function and $x_0 \in X$, we define the *differential* of f at x_0 by the following if the limit exists and is unique:

$$f'(x_0) = \lim_{\delta \to 0} \frac{f(x_0 + \delta) - f(x_0)}{\delta}$$

Lemma 1.26. Let $f : X \to Y$ be a function and $x_0 \in X$. If f is differentiable at x_0 then f is continuous at x_0.

Proof. We want to prove the following:

$$\lim_{x \to x_0} f(x) = f(x_0)$$

We use the following which follows from being differentiable:

$$\begin{aligned}
\lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right) &= f'(x_0) \\
&\Updownarrow \\
\lim_{x \to x_0} (f(x) - (x_0)) &= \left(\lim_{x \to x_0} (x - x_0) \right) f'(x_0) \\
&\Updownarrow \\
\lim_{x \to x_0} (f(x) - (x_0)) &= 0
\end{aligned}$$

$\square$

Remark 1.27. The converse is not true and we can use $f(x) = |x|$ as the counter example as f is continuous everywhere but not differentiable at 0 as the limit is not unique:

$$\lim_{x \to 0^+} f'(x) = 1 \neq -1 = \lim_{x \to 0^-} f'(x)$$

There are different notations, we can use:

$$f'(x)\,(\text{Lagrange}) = \frac{df}{dx}\,(\text{Leibniz}) = Df\,(\text{Euler})$$

Where D can be thought of as a differential operator which maps from a function space to a function space. We will also have $\dot{y}$ (Newton) notation which will be used to for differentiation with time.

Proposition 1.28.

- $$\frac{d}{dx}(x^n) = nx^{n-1}$$
- $$\frac{d}{dx}(e^x) = e^x$$
- $$\frac{d}{dx}(e^{ax}) = ae^{ax}$$
- $$\frac{d}{dx}(\log(x)) = \frac{1}{x}$$
- $$\frac{d}{dx}(\sin(x)) = \cos(x)$$
- $$\frac{d}{dx}(\cos(x)) = -\sin(x)$$
- $$\frac{d}{dx}(\tan(x)) = \sec^2(x)$$

Definition 1.29. We say a map $T : X \to Y$ is linear if for all $\lambda, \mu \in \mathbb{R}, x_1, x_2 \in X$ we have:

$$T(\lambda x_1 + \mu x_2) = \lambda T(x_1) + \mu T(x_2)$$

Proposition 1.30. Let $f, g : X \to Y$ be differentiable functions and $\lambda, \mu \in \mathbb{R}$. The differential operator is linear which means for all $x \in X$ we have:

$$D(\lambda f(x) + \mu g(x)) = \lambda D(f(x)) + \mu D(g(x))$$

Proposition 1.31 (Product rule)**.** Let $f, g : X \to Y$ be differentiable functions then for all $x \in X$ we have:

$$D(f(x) g(x)) = f'(x) g(x) + f(x) g'(x)$$

Proposition 1.32 (Chain Rule)**.** Let $f : X \to Y, g : Y \to Z$ be differentiable functions then for all $x \in X$ we have:

$$D(f(g(x))) = g'(x) \cdot f'(g(x))$$

Proposition 1.33 (Quotient Rule)**.** Let $f, g : X \to Y$ be differentiable functions then for all $x \in X$ we have:

$$D\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x) g(x) - f(x) g'(x)}{(g(x))^2}$$

Lemma 1.34. Let $a \in \mathbb{R}$ and $y = a^x$, then we have:

$$\frac{dy}{dx} = a^x \log(a)$$

Proof. We have to use implicit differentiation here:

$$\begin{aligned} y &= a^x \\ &\Updownarrow \\ \log(y) &= x \log(a) \\ &\Downarrow \\ \frac{1}{y} \cdot \frac{dy}{dx} &= \log(a) \\ &\Updownarrow \\ \frac{dy}{dx} &= y \log(a) \\ &\Updownarrow \\ \frac{dy}{dx} &= a^x \log(a) \end{aligned}$$

□

Remark 1.35. Not all functions are differentiable everywhere so we have to be careful around singularities such as $1/x$ when $x = 0$ or $1/(x-2)$ when $x = 2$.

Proposition 1.36 (Leibniz Rule)**.** Let $f, g : X \to Y$ be differentiable functions and $n \in \mathbb{N}$ then we have:

$$D^n(fg) = \sum_{k=0}^{n} \binom{n}{k} D^k(f) D^{n-k}(g)$$

Proposition 1.37 (L'Hospital's Rule)**.** Let $f, g : X \to Y$ be differentiable functions, $x_0 \in X$ such that the following limit exists:

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

And we have the following:

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0 \text{ or } \pm\infty$$

Then we have:

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

1.3 Taylor Series

We want to find local polynomial approximations to complicated functions, we use Taylor series to do this.

Definition 1.38. Let $f : X \to Y$ be a function and $x_0 \in X$. We define its Taylor series around x_0 by the following:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}$$

We should really check for convergence but we will not worry about that here. The following proposition can be proved by finding the coefficients of the Taylor series for each function which can be done by finding the correct differentials and evaluating them at the right point.

Proposition 1.39.

- $$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \forall x \in \mathbb{R}$$

- $$\log(1 + x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{k+1}}{k+1}, \quad \forall |x| < 1$$

- $$\log(1 - x) = -\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1}, \quad \forall |x| < 1$$

- $$\frac{1}{1 - x} = \sum_{k=0}^{\infty} x^k, \quad \forall |x| < 1$$

- $$\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}, \quad \forall x \in \mathbb{R}$$

- $$\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}, \quad \forall x \in \mathbb{R}$$

- $$\sinh(x) = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}, \quad \forall x \in \mathbb{R}$$

- $$\cosh(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, \quad \forall x \in \mathbb{R}$$

Theorem 1.40 (Binomial Theorem)**.** Let $n \in \mathbb{N}$ and $x \in \mathbb{R}$ such that $|x| < 1$ then we have:

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k$$

1.4 Integration

Definition 1.41. Let $f : X \to Y$ be a function. We definite an operator called *indefinite integral* which maps from a function space to another function space. We use the following notation:

$$f \mapsto \int f(x)\,dx$$

We define the indefinite integral by:

$$F(x) := \int f(x)\,dx, \quad \text{do that} \quad \frac{dF}{dx}(x) = f(x)$$

Remark 1.42. Since the derivative of a constant $c \in \mathbb{R}$ is 0, the indefinite integral of any function is only determined up to an arbitrary constant.

Proposition 1.43. Let $a \neq 0 \in \mathbb{R}$.

- $$\int x^n dx = \frac{x^{n+1}}{n+1} + c$$

- $$\int e^{ax} dx = \frac{e^{ax}}{a} + c$$

- $$\int \frac{1}{x} dx = \log(x) + c$$

- $$\int \sin(ax)\,dx = -\frac{\cos(ax)}{a} + c$$

- $$\int \cos(ax)\,dx = \frac{\sin(ax)}{a} + c$$

Lemma 1.44. The indefinite integral operator is linear. Let $f, g : X \to Y$ be functions and $\lambda, \mu \in \mathbb{R}$ then we have the following:

$$\int (\lambda f(x) + \mu g(x))\, dx = \lambda \int f(x)\, dx + \mu \int g(x)\, dx$$

Definition 1.45. Let $f : X \to Y$ be a function and $a, b \in X$. The *definite integral operator* is a map which maps functions into the reals. The definite integral takes in a function and two elements from the function's domain which are called *limits* and outputs a real number. The following is the notation for the definite integral for the function f between the limits a and b:

$$\int_a^b f(x)\, dx$$

Remark 1.46. The geometric understand behind the definite integral is the area under the function (between the function and the x-axis) between the to limits.

Example 1.47.

- $$\int_1^3 x^3 dx = \left[\frac{x^4}{4}\right]_1^3 = \frac{1}{4}\left(3^4 - 1^4\right) = 20$$
- $$\int_{-1}^1 e^x dx = [e^x]_{-1}^1 = e - 1/e$$

Lemma 1.48. The definite integral operator is linear, given the same limits. Let $f, g : X \to Y$ be functions, $a, b \in X$ and $\lambda, \mu \in \mathbb{R}$ then we have the following:

$$\int_a^b (\lambda f(x) + \mu g(x))\, dx = \lambda \int_a^b f(x)\, dx + \mu \int_a^b g(x)\, dx$$

Proposition 1.49. Let $f : X \to Y$ be a function and $a \geq b \geq c \in X$. then we have the following:

- $$\int_a^c f(x)\, dx = \int_a^b f(x)\, dx + \int_b^c f(x)\, dx$$
- $$\int_a^b f(x)\, dx = -\int_b^a f(x)\, dx$$

Theorem 1.50 (Integration by Substitution)**.** Let $f : X \to Y, g : Y \to Z$ be functions, f is invertible and $y_1, y_2 \in Y$. Then we have the following:

$$\int_{y_1}^{y_2} g(y)\, dy = \int_{x_1 = f^{-1}(y_1)}^{x_2 = f^{-1}(y_2)} g(f(x))\, f'(x)\, dx$$

'Proof'. We make a substitution $y = f(x)$ hence we have:

$$\frac{dy}{dx} = f'(x) \Rightarrow dy = f'(x)\, dx$$

So have have:

$$\begin{aligned}\int_{y_1}^{y_2} g(y)\, dy &= \int_{y_1}^{y_2} g(f(x))\, dy \\ &= \int_{x_1=f^{-1}(y_1)}^{x_2=f^{-1}(y_2)} g(f(x))\, f'(x)\, dx\end{aligned}$$

□

Theorem 1.51 (Integration by Parts)**.** Let $f : X \to Y$ be a function and $a \leq b \in X$. Suppose there exist functions $u, v : X \to Y$ so that:

$$f(x) = \frac{du}{dx}(x)\, v(x), \quad \forall x \in [a, b]$$

Then we have:

$$\int_a^b f(x)\, dx = \int_a^b \frac{du}{dx}(x)\, v(x)\, dx = [u(x)\, v(x)]_a^b - \int_a^b u(x) \frac{dv}{dx}(x)\, dx$$

Proof. We have the following:

$$\frac{d}{dx}(u(x)\, v(x)) = \frac{du}{dx}(x)\, v(x) + u(x) \frac{dv}{dx}(x)$$

Which is equivalent to:

$$\frac{du}{dx}(x)\, v(x) = \frac{d}{dx}(u(x)\, v(x)) - u(x) \frac{dv}{dx}(x)$$

Now integrating both side we get what we need:

$$\int_a^b \frac{du}{dx}(x)\, v(x)\, dx = [u(x)\, v(x)]_a^b - \int_a^b u(x) \frac{dv}{dx}(x)\, dx$$

□

Remark 1.52. Pick $v(x)$ so that $v'(x)$ is 'reducing' so normally this is a polynomial and you would want to pick du/dx so that it is integrable easily so something like the exponential function.

Example 1.53 (Reduction Formula)**.** Consider the following problem:

$$I_n = \int_0^\infty e^{-t} t^n dt$$

We want to use the following substitution:

$$\begin{aligned} \frac{du}{dt}(t) &= e^{-t}, & v(t) &= t^n \\ u(t) &= -e^{-t}, & \frac{dv}{dt}(t) &= nt^{n-1}\end{aligned}$$

Using the integration by parts formula we get:

$$\begin{aligned}
I_n &= \int_0^\infty e^{-t}t^n dt \\
&= \left[-e^{-t}t^n\right]_0^\infty - \int_0^\infty -e^{-t}nt^{n-1}dt \\
&= 0 + n\int_0^\infty e^{-t}t^{n-1}dt \\
&= nI_{n-1} \\
&\vdots \\
&= n!I_0
\end{aligned}$$

Where:

$$I_0 = \int_0^\infty e^{-t}dt = \left[-e^{-t}\right]_0^\infty = -(0-1) = 1$$

Therefore we have that:

$$I_n = n!$$

This is similar to the *Gamma* function.

Proposition 1.54. Let $f : X \to Y$ be a function. We have the following result:

$$\int \frac{f'(x)}{f(x)}dx = \log(|f(x)|) + c$$

Proposition 1.55 (Simple Partial Fractions)**.** Let $a_0, a_1, b_0, b_1 \in \mathbb{R}$ such that $a_0b_1 - a_1b_0 \neq 0$. We have the following:

$$\int \frac{1}{(a_1x+a_0)(b_1x+b_0)}dx = \left(\frac{1}{a_0b_1 - a_1b_0}\right)\log\left(\frac{b_1x+b_0}{a_1x+a_0}\right) + c$$

Proof. We want to equate the following and solve for A, B:

$$\frac{1}{(a_1x+a_0)(b_1x+b_0)} = \frac{A}{a_1x+a_0} + \frac{B}{b_1x+b_0}$$

We have:

$$\frac{A}{a_1x+a_0} + \frac{B}{b_1x+b_0} = \frac{A(b_1x+b_0) + B(a_1x+a_0)}{(a_1x+a_0)(b_1x+b_0)}$$

Therefore we have the following equation:

$$1 = x(Ab_1 + Ba_1) + (Ab_0 + Ba_0)$$

So equating the coefficients of $x^1 = x$ and $x^0 = 1$ we have the following simultaneous equations:

$$Ab_1 + Ba_1 = 0 \quad (1)$$
$$Ab_0 + Ba_0 = 1 \quad (2)$$

(1) implies that $A = -a_1 B/b_1$. Substituting it into (2) we get:

$$\begin{aligned}
\left(-\frac{a_1 B}{b_1}\right) b_0 + Ba_0 &= 1 \\
&\Updownarrow \\
B\left(-\frac{a_1 b_0}{b_1} + a_0\right) &= 1 \\
&\Updownarrow \\
B &= \frac{b_1}{a_0 b_1 - a_1 b_0}
\end{aligned}$$

Now solving for A we get:

$$A = -\frac{a_1}{a_0 b_1 - a_1 b_0}$$

Now we integrate:

$$\begin{aligned}
\int \frac{1}{(a_1 x + a_0)(b_1 x + b_0)} dx &= A \int \frac{1}{a_1 x + a_0} dx + B \int \frac{1}{b_1 x + b_0} dx \\
&= \left(-\frac{a_1}{a_0 b_1 - a_1 b_0}\right) \int \frac{1}{a_1 x + a_0} dx + \left(\frac{b_1}{a_0 b_1 - a_1 b_0}\right) \int \frac{1}{b_1 x + b_0} dx \\
&= \left(-\frac{a_1}{a_0 b_1 - a_1 b_0}\right) \frac{1}{a_1} \log(a_1 x + a_0) + \left(\frac{b_1}{a_0 b_1 - a_1 b_0}\right) \frac{1}{b_1} \log(b_1 x + b_0) + c \\
&= \left(\frac{1}{a_0 b_1 - a_1 b_0}\right) (\log(b_1 x + b_0) - \log(a_1 x + a_0)) + c \\
&= \left(\frac{1}{a_0 b_1 - a_1 b_0}\right) \log\left(\frac{b_1 x + b_0}{a_1 x + a_0}\right) + c
\end{aligned}$$

$\square$

1.5 Complex Numbers

Definition 1.56. We define a complex number $z \in \mathbb{C}$ where $z = x + iy$ for $x, y \in \mathbb{R}$ and $i = \sqrt{-1}$. It follows that $i^2 = -1$.

Definition 1.57. Let $z = x + iy, z_1 = x_1 + iy_1, z_2 = x_2 + iy_2 \in \mathbb{C}$ for some $x, y, x_1, y_1, x_2, y_2 \in \mathbb{R}$.

- We define the *complex conjugate* of z by:

$$\overline{z} := x - iy$$

- We define the *norm* of z by:

$$|z| := \sqrt{x^2 + y^2} = \sqrt{z \cdot \overline{z}}$$

- We 'define' the following operations:

$$z_1 \pm z_2 = (x_1 \pm x_2) + i(y_1 \pm y_2)$$

$$z_1 \times z_2 = (x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + iy_1x_2 + ix_1y_2 + i^2y_1y_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

$$\frac{z_1}{z_2} = \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{x_1 + iy_1}{x_2 + iy_2} \times \frac{x_2 - iy_2}{x_2 - iy_2} = \frac{(x_1x_2 + y_1y_2) + i(-x_1y_2 + x_2y_1)}{x_2^2 + y_2^2}$$

Proposition 1.58. Let $z, z_1, z_2 \in \mathbb{C}$.

- $\overline{(\overline{z})} = z$
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
- $\overline{(z_1 \cdot z_2)} = \overline{z_1} \cdot \overline{z_2}$
- $(z + \overline{z})/2 = \mathrm{Re}(z) =:$ Real part of z
- $(z - \overline{z})/2 = \mathrm{Im}(z) =:$ Imaginary part of z
- $(z \cdot \overline{z}) = |z|^2$
- $|\overline{z}|^2 = \overline{z} \cdot \overline{(\overline{z})} = \overline{z} \cdot z = |z|^2$
- $\frac{z_1}{z_2} = \frac{z_1}{z_2} \times \frac{\overline{z_2}}{\overline{z_2}} = \frac{z_1 z_2}{|z_2|^2}$
- $|z_1z_2|^2 = |z_1|^2 |z_2|^2$

Theorem 1.59 (Euler's Formula)**.** Let $\theta \in \mathbb{R}$, then we have:

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

Proposition 1.60. For any $z = x + iy \in \mathbb{C}$ for some $x, y \in \mathbb{R}$, we can also write it in the modulus argument form:

$$z = re^{i\theta} = r(\cos(\theta) + i\sin(\theta))$$

Where:

$$r = \sqrt{x^2 + y^2} = |z|, \quad \theta = \arctan\left(\frac{y}{x}\right)$$

Proposition 1.61. Let $z_1 = r_1e^{i\theta_1}, z_2 = r_2e^{i\theta_2} \in \mathbb{C}$ where $r_1, r_2, \theta_1, \theta_2 \in \mathbb{R}$. Then we have:

- $$z_1z_2 = r_1r_2e^{i(\theta_1+\theta_2)}$$
- Suppose $r_2 \neq 0$, then:
$$\frac{z_1}{z_2} = \frac{r_1}{r_2}e^{i(\theta_1-\theta_2)}$$

1.6 Multi-Variate Calculus

Definition 1.62. Let $f : X \times Y \to Z$ be a multivariate function. We define the *partial derivative* of f with respect to $x \in X$ as:

$$\frac{\partial f}{\partial x} := \lim_{\delta \to 0} \frac{f(x+\delta, y) - f(x,y)}{\delta} = f_x$$

Similarly the partial derivative of f with respect to $y \in Y$ is defined as:

$$\frac{\partial f}{\partial x} := \lim_{\delta \to 0} \frac{f(x, y+\delta) - f(x,y)}{\delta} = f_y$$

We also define *higher order* derivatives of f as following:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial xy} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial yx} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Theorem 1.63. Let $f : X \times Y \to Z$ be a function. Suppose the second order derivatives exist and are continuous then we have:

$$f_{xy} = \frac{\partial^2 f}{\partial xy} = \frac{\partial^2 f}{\partial yx} = f_{yx}$$

Lemma 1.64. Let $f : X \times Y \to Z$ be a function. Suppose $f(x,y) = f(x(t), y(t))$. Then we have:

$$\frac{df}{dt}(t) = \frac{\partial f}{\partial x}(x(t), y(t))\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}(x(t), y(t))\frac{dy}{dt}(t)$$

Lemma 1.65. Let $f : X \times Y \to Z$ be a function. Suppose $f(x,y) = f(x(u,v), y(u,v))$. Then we have:

$$\begin{aligned}\frac{\partial f}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \\ \frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}\end{aligned}$$

Theorem 1.66. Let $f : X \times T \to Y$ be a function and $x_0 \in X, t_0 \in T$ then we have the multivariate Taylor series as:

$$\begin{aligned} f(x,t) = \quad & f(x_0,t_0) + (x-x_0) f_x(x_0,t_0) + (t-t_0) f_t(x_0,t_0) + \dots \\ & \dots + \frac{1}{2}\Big((x-x_0)^2 f_{xx}(x_0,t_0) + 2(x-x_0)(t-t_0) f_{xt}(x_0,t_0) + (t-t_0)^2 f_{tt}(x_0,t_0)\Big)\end{aligned}$$

Example 1.67. We have:

$$f(x+\delta_x, y+\delta_y) = f(x,y) + \delta_x f_x(x,y) + \delta_y f_y(x,y) + \frac{f_{xx}(x,y)\,\delta_x^2}{2} + f_{xy}\delta_x\delta_y + \frac{f_{yy}(x,y)\,\delta_y^2}{2} + O\left(\delta_x^2, \delta_y^2\right)$$

We use the following notation:

$$df := \lim_{\delta_x \to 0, \delta_y \to 0} \left(f(x+\delta_x, y+\delta_y) - f(x,y)\right)$$

So we have the first order approximation:

$$df \approx \delta_x \frac{\partial f}{\partial x} + \delta_y \frac{\partial x}{\partial y}$$

1.7 Probability Problems

1.7.1 Birthday Problem

Suppose we have a room with $n \in \mathbb{N}$ people. We want to know the probability of at least two people in the room sharing birthdays. We need to make a few assumptions to start with:

1. There are 365 days in a year.
2. The event of a person's birth is independent of anyone else's birth.
3. The probability of a person's birthday on a given day is uniformly distributed.

Now that we have made these assumptions, let us define the event X to be the random variable which denotes the number of pairs of people that share birthdays:

$$\begin{aligned}
\mathbb{P}\left[\{X \geq 1\}\right] &= 1 - \mathbb{P}\left[\{X \geq 1\}^c\right] \\
&= 1 - \mathbb{P}\left[\{X < 1\}\right] \\
&= 1 - \mathbb{P}\left[\{X = 0\}\right]
\end{aligned}$$

So we now have to find $\mathbb{P}\left[\{X = 0\}\right]$ which is the probability of the event where no pair of people share a birthday. To find the probability of no two people sharing the same birthday, we can do this iteratively, let us define the event:

$$A_n = \{\text{No two people share a birthday in a room of } n \text{ people}\}$$

A_1 is the same is saying the one person in the room can have their birthday on any day of the year.

$$\begin{aligned}
\mathbb{P}\left[A_1\right] &= \mathbb{P}\left[\bigcup_{d_1=1}^{365} \{B_1 = d_1\}\right] \\
&= \sum_{d_1=1}^{365} \mathbb{P}\left[\{B_1 = d_1\}\right] \\
&= \sum_{d_1=1}^{365} \frac{1}{365} \\
&= \frac{365}{365} \\
&= 1
\end{aligned}$$

Where $\{B_i = d_i\}$ is the event that the i-th person's birthday happens on d_i-th day of the year. Now let us try $n = 2$, let us define the event $A_2 = \{\text{No two people share a birthday in a room with two people}\}$:

$$\begin{aligned}
\mathbb{P}\left[A_2\right] &= \mathbb{P}\left[\bigcup_{d_1=1}^{365} \bigcup_{\substack{d_2=1 \\ d_2 \neq d_1}}^{365} \left(\{B_1 = d_1\} \cap \{B_2 = d_2\}\right)\right] \\
&= \left(\sum_{j=1}^{365} \mathbb{P}\left[\{B_1 = d_j\}\right]\right) \times \left(\sum_{k=2}^{365} \mathbb{P}\left[\{B_2 = d_k\}\right] \,|\, \{B_1 = d_j\}\right)
\end{aligned}$$

2 Transition Probability

2.1 Forward Kolmogorov Equation

Proposition 2.1 (Forward Kolmogorov Equation Derivation)**.** Let $x \in \mathbb{R}$ and $s > 0$, consider a trinomial random walk which starts at point x at time s. Let $\alpha < 1/2$, the probability of an up move of δy is α, down move of $-\delta y$ is α and remaining in the same position is $1 - 2\alpha$ over a time period of δt. Let $y \in \mathbb{R}$ and $t > 0$, defining $p\,(y, t; x, s)$ to be the probability of moving from (x, s) to (y, t), p satisfies the following partial differential equation:

$$\frac{\partial p}{\partial t} = c^2 \frac{\partial^2 p}{\partial y^2} \tag{3}$$

Proof. To find the probability of going from (x, s) to (y, t), we would like to go a timestep back to $t - \delta t$ and look at the relationship from there. At timestep $t - \delta t$ we could have been at $y + \delta y$, y or $y - \delta y$. Therefore leaving the dependency on (x, s), we have the following equation:

$$p\,(y, t) = \alpha p\,(y - \delta y, t - \delta t) + \alpha p\,(y + \delta y, t - \delta t) + (1 - 2\alpha)\, p\,(y, t - \delta t)$$

Therefore, we can expand all the terms on the right using Taylor expansion for the first order terms for t and second order terms for y:

$$\begin{aligned}
p\,(y, t) \quad &= \quad \alpha \left(p\,(y, t) - \delta t \frac{\partial p}{\partial t} - \delta y \frac{\partial p}{\partial y} + \frac{1}{2} \delta y^2 \frac{\partial^2 p}{\partial y^2} \right) + \dots \\
\dots + \quad & \alpha \left(p\,(y, t) - \delta t \frac{\partial p}{\partial t} + \delta y \frac{\partial p}{\partial y} + \frac{1}{2} \delta y^2 \frac{\partial^2 p}{\partial y^2} \right) + \dots \\
\dots + \quad & (1 - 2\alpha) \left(p\,(y, t) - \delta t \frac{\partial p}{\partial t} \right)
\end{aligned}$$

Cancelling out terms, we get the following:

$$\begin{aligned}
0 \quad &= \quad \alpha \delta y^2 \frac{\partial^2 p}{\partial y^2} - \delta t \frac{\partial p}{\partial t} \\
&\Updownarrow \\
\frac{\partial p}{\partial t} \quad &= \quad \alpha \frac{\delta y^2}{\delta t} \frac{\partial^2 p}{\partial y^2}
\end{aligned}$$

Assuming that δy^2 is proportional to δt and letting $c^2 = \alpha \delta y^2 / \delta t$, we get the following partial differential equation:

$$\frac{\partial p}{\partial t} = c^2 \frac{\partial^2 p}{\partial y^2}$$

$\square$

Proposition 2.2. Let $x, y \in \mathbb{R}$ and $s, t > 0$, suppose we want the transition probability of going from (x, s) to (y, t), denoted by $p(y, t; x, s)$, and we know that it satisfies the Forward Kolmogorov Equation:

$$\frac{\partial p}{\partial t} = c^2 \frac{\partial^2 p}{\partial y^2}$$

Then we have the following result:

$$p(y, t; x, s) = \frac{1}{2c\sqrt{\pi(t-s)}} \exp\left(\frac{(y-x)^2}{4c^2(t-s)}\right) \tag{4}$$

Proof. Suppose $p(y, t)$ satisfies the Forward Kolmogorov Equation, then we use the method of similarity to obtain a solution. Suppose the solution is of the following form:

$$p(y, t) = t^{\alpha} f\left(\frac{y}{t^{\beta}}\right) = t^{\alpha} f(\xi), \quad \xi = \frac{y}{t^{\beta}}$$

Then we have the following differentials:

$$\begin{aligned}
\frac{\partial p}{\partial t} &= \alpha t^{\alpha-1} f(\xi) + t^{\alpha} f'(\xi) \frac{\partial \xi}{\partial t} \\
&= \alpha t^{\alpha-1} f(\xi) + t^{\alpha} f'(\xi) \left(-bt^{-b-1} y\right) \\
&= t^{\alpha-1} \left(\alpha f(\xi) - b\xi f'(\xi)\right) \\
\frac{\partial p}{\partial y} &= t^{\alpha} f'(\xi) \frac{\partial \xi}{\partial y} = t^{\alpha-\beta} f'(\xi) \\
\frac{\partial^2 p}{\partial y^2} &= t^{\alpha-\beta} f''(\xi) \frac{\partial \xi}{\partial y} = t^{\alpha-2\beta} f''(\xi)
\end{aligned}$$

Therefore, we have the following relation:

$$\begin{aligned}
\frac{\partial p}{\partial t} &= c^2 \frac{\partial^2 p}{\partial y^2} \\
&\Updownarrow \\
t^{\alpha-1} \left(\alpha f(\xi) - b\xi f'(\xi)\right) &= c^2 t^{\alpha-2\beta} f''(\xi)
\end{aligned}$$

For the above equation to have a similarity solution, both sides of the equation have to be independent of t hence we need:

$$\alpha - 1 = \alpha - 2\beta \Leftrightarrow \beta = \frac{1}{2}$$

Therefore we have:

$$p(y, t) = t^{\alpha} f(\xi) = t^{\alpha} f\left(\frac{y}{\sqrt{t}}\right)$$

As $p(y, t)$ is a transition probability, we must have that the integral over $y \in \mathbb{R}$ should add up to one:

$$\begin{aligned}
\int_{y\in\mathbb{R}} p\left(y,t\right)dy &= 1\\
&\Updownarrow\\
t^{\alpha}\int_{y\in\mathbb{R}} f\left(\frac{y}{\sqrt{t}}\right)dy &= 1\\
\left(\text{Subsititute } u=y/\sqrt{t},\ du=dy/\sqrt{t}\right) &\Updownarrow\\
t^{\alpha+1/2}\int_{u\in\mathbb{R}} f\left(u\right)du &= 1
\end{aligned}$$

Again we need the above relation ship to be true for all $t>0$, therefore, we need:

$$\alpha+\frac{1}{2}=0\Leftrightarrow\alpha=-\frac{1}{2}$$

Now we have the following differential equation:

$$\begin{aligned}
t^{\alpha-1}\left(\alpha f\left(\xi\right)-b\xi f'\left(\xi\right)\right) &= c^2t^{\alpha-2\beta}f''\left(\xi\right)\\
&\Updownarrow\\
-t^{-1/2}\frac{1}{2}\left(f\left(\xi\right)+\xi f'\left(\xi\right)\right) &= c^2t^{-1/2}f''\left(\xi\right)\\
&\Updownarrow\\
-\frac{1}{2c^2}d\left(\xi f\left(\xi\right)\right) &= f''\left(\xi\right)\\
&\Updownarrow\\
-\frac{1}{2c^2}\xi f\left(\xi\right) &= f'\left(\xi\right)+K
\end{aligned}$$

For a fixed $t>0$, we have that $f\left(\xi\right)\to 0$ and $f'\left(\xi\right)\to 0$ as $\xi\to\infty$ due to the fact that its part of a probability density function. This will mean that $K=0$. Therefore, we now have the following differential equation:

$$\begin{aligned}
-\frac{1}{2c^2}\xi f\left(\xi\right) &= f'\left(\xi\right)\\
&\Updownarrow\\
\frac{f'\left(\xi\right)}{f\left(\xi\right)} &= -\frac{1}{2c^2}\xi\\
&\Updownarrow\\
f\left(\xi\right) &= A\exp\left(-\frac{1}{4c^2}\xi^2\right)
\end{aligned}$$

Where A is a normalising constant which calculated the following way:

$$\begin{aligned}
A\int_{\xi\in\mathbb{R}}\exp\left(-\frac{1}{4c^2}\xi^2\right) &= 1 \\
\text{(Subsititute } u=\xi/(2c)\text{, } du=d\xi/(2c)\text{)} &\Updownarrow \\
2cA\underbrace{\int_{u\in\mathbb{R}}\exp\left(-u^2\right)}_{=\sqrt{\pi}} &= 1 \\
&\Updownarrow \\
A &= \frac{1}{2c\sqrt{\pi}}
\end{aligned}$$

Therefore we have the fundamental solution:

$$\begin{aligned}
p(y,t) &= t^{-1/2}f(\xi) \\
&= \frac{1}{2c\sqrt{\pi t}}f\left(\frac{y}{\sqrt{t}}\right) \\
&= \frac{1}{2c\sqrt{\pi t}}\exp\left(-\frac{1}{4c^2t}y^2\right)
\end{aligned}$$

As you can see this is the Gaussian probability density function with $\mu=0$ and $\sigma=c\sqrt{2t}$.

We need that $p(x,s)=p(x,s;x,s)=1$. To achieve this we need to adjust our fundamental solution to get the right answer. We need to have a Gaussian probability density function with $\mu=x$ and $\sigma=t-s$ as we should have that $\sigma\downarrow 0$ as $t\downarrow s$. Therefore, we have the following final solutions:

$$p(y,t;x,s)=\frac{1}{2c\sqrt{\pi(t-s)}}\exp\left(-\frac{(y-x)^2}{4c^2(t-s)}\right)$$

$\square$

3 $\sigma_t \propto \sigma_1\sqrt{t}$

We have that the volatility of the returns of an asset scales with square root of time. Assume X_m denotes the discrete asset price at period m. Then we define the n-period return from X_m to X_{m+n}, $r_{m+n,n}$, as:

$$X_{n+m}=X_m\exp\left(r_{m+n,n}\right)\Leftrightarrow r_{m+n,n}=\log\left(\frac{X_{n+m}}{X_m}\right)$$

Suppose we know the distribution of $r_{i,1}:=r_i$ which is i.i.d. with variance σ^2. Then we know that $\mathbb{V}\left[\log\left(\frac{X_{m+1}}{X_m}\right)\right]=\mathbb{V}\left[r_{m+1}\right]=\sigma^2$ which means that the volatility which is the standard deviation of the returns for one period is σ. From here, we can show that the n-period volatility scales with

$\sqrt{n}$ by the following:

$$\begin{aligned}
\mathbb{V}\left[\log\left(\frac{X_{m+n}}{X_m}\right)\right] &= \mathbb{V}\left[\sum_{i=m+1}^{m+n} r_i\right] \\
&= \sum_{i=m+1}^{m+n}\sum_{j=m+1}^{m+n} \operatorname{cov}(r_i, r_j) \\
&= \sum_{i=m+1}^{m+n} \mathbb{V}[r_i] + 2\sum_{i=m+1}^{m+n}\sum_{j>i}^{m+n} \underbrace{\operatorname{cov}(r_i, r_j)}_{=0 \text{ as returns are i.i.d.}} \\
&= n\sigma^2
\end{aligned}$$

4 Stochastic Calculus

Theorem 4.1 (Ito's Lemma)**.** Let $\left(X_t^i : t > 0\right)$ be stochastic processes for $i = 1, \ldots, n$ where $n \in \mathbb{N}$. Let $f_t = f(t, X_t^1, \ldots, X_t^n)$ be a function of time and the stochastic processes. Then we have the following:

$$f_t = f_s + \int_s^t \frac{\partial f_r}{\partial r} dr + \sum_{i=1}^{n} \int_s^t \frac{\partial f_r}{\partial X_r^i} dX_r^i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_s^t \frac{\partial^2 f_r}{\partial X_r^i X_r^j} d\left\langle X^i, X^j\right\rangle_r \tag{5}$$

Which is equivalent to the following:

$$df_t = \frac{\partial f_t}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial f_t}{\partial X_t^i} dX_t^i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f_t}{\partial X_t^i X_t^j} d\left\langle X^i, X^j\right\rangle_t \tag{6}$$

Where $\left\langle X^i, X^j\right\rangle_t$ is the quadratic variation process of X^i and X^j.

Corollary 4.2. Let $\left(B_t^i : t > 0\right)$ for Brownian motions for $i = 1, \ldots, n$ where $n \in \mathbb{N}$. Let $f_t = f(t, B_t^1, \ldots, B_t^n)$ be a function of time and the stochastic processes. Let the quadratic variation process between B_t^i and B_t^j for $i, j = 1, \ldots, n$ be $\left\langle B^i, B^j\right\rangle_t = \rho_{i,j} dt$. Then we have the following:

$$df_t = \left(\frac{\partial f_t}{\partial t} + \frac{1}{2}\sum_{i=1}^{n} \frac{\partial^2 f_t}{\partial B_t^i} + \rho_{i,j}\sum_{i=1}^{n}\sum_{j>i}^{n} \frac{\partial^2 f_t}{\partial B_t^i B_j^i}\right) dt + \sum_{i=1}^{n} \frac{\partial f_t}{\partial B_t^i} dB_t^i \tag{7}$$

4.1 Geometric Brownian Motion

Proposition 4.3. Let X_t follow a geometric Brownian motion such that:

$$dX_t = \mu X_t dt + \sigma X_t dB_t \tag{8}$$

Where $\mu, \sigma \in \mathbb{R}$. Then we have the following formula for X_t where $s < t$:

$$X_t = X_s \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t - s) + \sigma\left(B_t - B_s\right)\right) \tag{9}$$

Proof. Let us look at the function $\log(X_t)$. We can use Ito's Lemma to get the following:

$$\begin{aligned}
d(\log(X_t)) &= \frac{\partial \log(X_t)}{\partial X_t} dX_t + \frac{1}{2}\frac{\partial^2 \log(X_t)}{\partial X_t^2} d\langle X, X\rangle_t \\
&= \frac{1}{X_t}(\mu X_t dt + \sigma X_t dB_t) + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)\sigma^2 X_t^2 d\langle B, B\rangle_t \\
&= \mu dt + \sigma dB_t - \frac{1}{2}\sigma^2 dt \\
&= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t
\end{aligned}$$

Now integrating both sides between times s and t, we get the following:

$$\begin{aligned}
\int_s^t d(\log(X_t)) &= \left(\mu - \frac{1}{2}\sigma^2\right)\int_s^t dr + \sigma \int_s^t dB_r \\
&\Updownarrow \\
\log(X_t) - \log(X_s) &= \left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(B_t - B_s) \\
&\Updownarrow \\
X_t &= X_s \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(B_t - B_s)\right)
\end{aligned}$$

□

4.2 Vasicek Model

Proposition 4.4. Let X_t satisfy the following stochastic differential equation:

$$dX_t = \gamma(\theta - X_t)\,dt + \sigma dB_t \tag{10}$$

Where $\gamma, \theta, \sigma \in \mathbb{R}$ and B_t is a Brownian motion. Then we have the following formula for X_t where $s < t$:

$$X_t = X_s e^{-\gamma(t-s)} + \theta\left(1 - e^{-\gamma(t-s)}\right) + \sigma\int_s^t e^{-\gamma(t-r)} dB_r \tag{11}$$

Proof. Let X_t satisfy the following stochastic differential equation shown in Equation 10 then we have the following:

$$\begin{aligned}
dX_t &= \gamma(\theta - X_t)\,dt + \sigma dB_t \\
&\Updownarrow \\
dX_t + \gamma X_t dt &= \gamma\theta dt + \sigma dB_t \\
&\Updownarrow \\
d\left(X_t e^{\gamma t}\right) &= \gamma\theta e^{\gamma t} dt + \sigma e^{\gamma t} dB_t \\
&\Updownarrow \\
X_t e^{\gamma t} - X_s e^{\gamma s} &= \theta\int_s^t \gamma e^{\gamma r} dr + \sigma\int_s^t e^{\gamma r} dB_r \\
&\Updownarrow \\
X_t &= X_s e^{-\gamma(t-s)} + \theta\left(1 - e^{-\gamma(t-s)}\right) + \sigma\int_s^t e^{-\gamma(t-r)} dB_r
\end{aligned}$$

□

5 Black-Scholes Equation

Theorem 5.1 (Black-Scholes Equation)**.** Assumptions:

- The underlying asset follows a geometric Brownian motion as follows:

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

  Where $\mu, \sigma \in \mathbb{R}$ are constants and B_t is a Brownian motion.

- Riskless borrowing and lending rate is constant for any tenor.
- Buying and selling of underlying asset and borrowing and lending of any amount is possible even fractional amounts.
- Possible to short the underlying asset.
- No taxes and transaction costs.
- There is no arbitrage opportunities.

Let $V_t = V(t, S_t)$ be a derivative whose value depends on the underlying S_t and time. Then we have the following partial differential equation known as the Black-Scholes Equation:

$$\frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + rS_t \frac{\partial V_t}{\partial S_t} - rV_t = 0 \tag{12}$$

Proof. We would like to create a hedging portfolio Π_t such that:

$$\Pi_t = V_t + \Delta S_t$$

Where Δ is constant over a small time increment, it is chosen at the start of the time increment then held constant then changed at the end of the time increment. Hence we have the following differential equation:

$$d\Pi_t = dV_t + \Delta dS_t$$

Using Ito's Lemma on $V_t = V(t, S_t)$, we get the following:

$$\begin{aligned} d\Pi_t &= \frac{\partial V_t}{\partial t} dt + \frac{1}{2}\frac{\partial^2 V_t}{\partial S_t^2} d\left\langle S, S \right\rangle_t + \frac{\partial V_t}{\partial S_t} dS_t - \Delta dS_t \\ &= \frac{\partial V_t}{\partial t} dt + \frac{1}{2}\frac{\partial^2 V_t}{\partial S_t^2} \sigma^2 S_t^2 d\left\langle B, B \right\rangle_t + \left(\frac{\partial V_t}{\partial S_t} - \Delta \right) dS_t \\ &= \left(\frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} \right) dt + \left(\frac{\partial V_t}{\partial S_t} - \Delta \right) dS_t \end{aligned}$$

Choosing $\Delta = \frac{\partial V_t}{\partial S_t}$, we can get rid of the dS_t term. Therefore no randomness is left. This means our portfolio should grow at the riskfree rate, hence $d\Pi_t = r\Pi dt$:

$$\begin{aligned} d\Pi_t &= \left(\frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2}\right) dt + \left(\frac{\partial V_t}{\partial S_t} - \frac{\partial V_t}{\partial S_t}\right) dS_t && (13)\\ &= \left(\frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2}\right) dt && (14)\\ &= r\Pi dt && (15)\\ &= r\left(V_t - \frac{\partial V_t}{\partial S_t} S_t\right) dt && (16)\end{aligned}$$

Equating Equations 14 and 16, we get the Black-Scholes Equation:

$$\begin{aligned} \frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} &= r\left(V_t - \frac{\partial V_t}{\partial S_t} S_t\right) && (17)\\ &\Updownarrow && (18)\\ \frac{\partial V_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + rS_t\frac{\partial V_t}{\partial S_t} - rV_t &= 0 && (19)\end{aligned}$$

□

Proposition 5.2. We can transform the Black-Scholes Equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

Into the following:

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial x^2}$$

Where:

$$V(t, S) = e^{-r(T-t)} U(t, S) = e^{-r\tau} U\left(T - \tau, e^{x - \left(r - \sigma^2/2\right)\tau}\right)$$

Proof. Our first transformation is $V = Ue^{-r(T-t)}$, therefore we have the following differentials:

$$\frac{\partial V}{\partial t} = e^{-r(T-t)}\left(\frac{\partial U}{\partial t} + rU\right), \quad \frac{\partial V}{\partial S} = e^{-r(T-t)}\frac{\partial U}{\partial S}, \quad \frac{\partial^2 V}{\partial S^2} = e^{-r(T-t)}\frac{\partial^2 U}{\partial S^2}$$

Hence we get the following equation:

$$e^{-r(T-t)}\left(\frac{\partial U}{\partial t} + rU + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S} - rU\right) = 0$$

As $e^{-r(T-t)} > 0$, we can divide both sides by it and the rU cancels giving us the following:

$$\frac{\partial U}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S} = 0$$

Now we substitute $\tau' = T - t$, giving us the following differential:

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau'} \cdot \frac{\partial \tau'}{\partial t} = -\frac{\partial}{\partial \tau}$$

Therefore now we have the following equation:

$$\frac{\partial U}{\partial \tau'} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S}$$

We will now make the following substitution:

$$\xi = \log(S)$$

So, we have the following differentials:

$$\frac{\partial}{\partial S} = \frac{\partial}{\partial \xi} \cdot \frac{\partial \xi}{\partial S} = \frac{1}{S}\frac{\partial}{\partial \xi}$$

$$\begin{aligned}
\frac{\partial^2}{\partial S^2} &= \frac{\partial}{\partial S}\left(\frac{1}{S}\frac{\partial}{\partial \xi}\right) && (20)\\
&= -\frac{1}{S^2}\frac{\partial}{\partial \xi} + \frac{1}{S}\frac{\partial}{\partial S}\left(\frac{\partial}{\partial \xi}\right) && (21)\\
&= -\frac{1}{S^2}\frac{\partial}{\partial \xi} + \frac{1}{S}\frac{\partial \xi}{\partial S}\frac{\partial}{\partial \xi}\left(\frac{\partial}{\partial \xi}\right) && (22)\\
&= -\frac{1}{S^2}\frac{\partial}{\partial \xi} + \frac{1}{S^2}\frac{\partial^2}{\partial \xi^2} && (23)\\
&= \frac{1}{S^2}\left(\frac{\partial^2}{\partial \xi^2} - \frac{\partial}{\partial \xi}\right) && (24)
\end{aligned}$$

Therefore we have the following differential equation:

$$\begin{aligned}
\frac{\partial U}{\partial \tau'} &= \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S}\\
&= \frac{1}{2}\sigma^2 S^2 \frac{1}{S^2}\left(\frac{\partial^2 U}{\partial \xi^2} - \frac{\partial U}{\partial \xi}\right) + rS\frac{1}{S}\frac{\partial U}{\partial \xi}\\
&= \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial \xi} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial U}{\partial \xi}
\end{aligned}$$

Now we make two final substitutions:

$$\begin{aligned}
\tau &= \tau'\\
x &= \xi + \left(r - \frac{1}{2}\sigma^2\right)\tau'
\end{aligned}$$

We get the following differentials:

$$\begin{aligned}
\frac{\partial}{\partial \tau'} &= \frac{\partial}{\partial \tau}\frac{\partial \tau}{\partial \tau'} + \frac{\partial}{\partial x}\frac{\partial x}{\partial \tau'} = \frac{\partial}{\partial \tau} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial}{\partial x}\\
\frac{\partial}{\partial \xi} &= \frac{\partial}{\partial \tau}\underbrace{\frac{\partial \tau}{\partial \xi}}_{=0} + \frac{\partial}{\partial x}\frac{\partial x}{\partial \xi} = \frac{\partial}{\partial x}\\
\frac{\partial^2}{\partial \xi^2} &= \underbrace{\frac{\partial}{\partial \xi}}_{\partial/\partial x}\left(\frac{\partial}{\partial x}\right) = \frac{\partial^2}{\partial x^2}
\end{aligned}$$

Therefore, we get the final differential equation:

$$\begin{aligned}
\frac{\partial U}{\partial \tau'} &= \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial \xi} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial U}{\partial \xi} \\
&\Updownarrow \\
\frac{\partial U}{\partial \tau} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial U}{\partial x} &= \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial U}{\partial x} \\
&\Updownarrow \\
\frac{\partial U}{\partial \tau} &= \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial x^2}
\end{aligned}$$

Where we have:

$$\begin{aligned}
V(t,S) &= e^{-r(T-t)}U(t,S) \\
&= e^{-r\tau'}U\left(T-\tau', S\right) \\
&= e^{-r\tau'}U\left(T-\tau', e^{\xi}\right) \\
&= e^{-r\tau}U\left(T-\tau, e^{x-\left(r-\sigma^2/2\right)\tau}\right)
\end{aligned}$$

□

6 Greeks

We will start with the closed form formulas for vanilla European calls and puts. We use the following notation:

- t - current time
- T - time at expiry
- r - interest rates, assumed to be constant
- σ - volatility, assumed to be constant
- K - strike price
- S_t - price of the underlying at time t
- C_t - price of the call option at time t
- P_t - price of the put option at time t
- $\mathrm{N}(z)$ is the standard normal distribution where:

$$\mathrm{N}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

We have the following formulas for calls and puts assuming the Black-Scholes-Merton model assumptions:

$$C_t = S_t \mathrm{N}(d_1) - Ke^{-r(T-t)}\mathrm{N}(d_2)$$
$$P_t = Ke^{-r(T-t)}\mathrm{N}(-d_2) - S_t \mathrm{N}(-d_1)$$

Where:

$$d_1 = \frac{\log(S_t/K) + \left(r + \sigma^2/2\right)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t} = \frac{\log(S_t/K) + \left(r - \sigma^2/2\right)(T-t)}{\sigma\sqrt{T-t}}$$

Note that we will drop the subscripts when taking derivatives.

Proposition 6.1. We have the following relationship:

$$S_t \mathrm{N}'(d_1) = Ke^{-r(T-t)}\mathrm{N}'(d_2)$$

Where:

$$\mathrm{N}'(z) = \frac{\partial \mathrm{N}(z)}{\partial z} = \frac{e^{-z^2/2}}{\sqrt{2\pi}}$$

Proof.

$$\begin{aligned}
\mathrm{N}'(d_1) &= \frac{e^{-d_1^2/2}}{\sqrt{2\pi}} \\
&= \frac{\exp\left(-\frac{\left(d_2+\sigma\sqrt{T-t}\right)^2}{2}\right)}{\sqrt{2\pi}} \\
&= \frac{e^{-d_2^2/2}}{\sqrt{2\pi}} \exp\left(-d_2\sigma\sqrt{T-t} - \frac{\sigma^2(T-t)}{2}\right) \\
&= \mathrm{N}'(d_2)\exp\left(-\log(S_t/K) - \left(r - \frac{\sigma^2}{2}\right)(T-t) - \frac{\sigma^2(T-t)}{2}\right) \\
&= \mathrm{N}'(d_2)\frac{K}{S_t}e^{-r(T-t)}
\end{aligned}$$

□

Proposition 6.2. The following is the call-put parity:

$$S_t + P_t = C_t + Ke^{-r(T-t)}$$

Proof. Prove using the absence of arbitrage argument. □

Proposition 6.3.

$$\Delta_C := \frac{\partial C_t}{\partial S_t} = \mathrm{N}(d_1)$$
$$\Delta_P := \frac{\partial P_t}{\partial S_t} = \mathrm{N}(-d_1)$$

Proof. Starting from the pricing formula of the call option $C_t = S_t \mathrm{N}(d_1) - Ke^{-r(T-t)}\mathrm{N}(d_2)$ and taking the derivative with respect to S_t, we get the following:

$$
\begin{aligned}
\Delta_C &= \frac{\partial C_t}{\partial S_t} \\
&= \mathrm{N}(d_1) + S_t \frac{\partial \mathrm{N}(d_1)}{\partial S_t} - Ke^{-r(T-t)} \frac{\partial \mathrm{N}(d_2)}{\partial S_t} \\
&= \mathrm{N}(d_1) + S_t \mathrm{N}'(d_1) \frac{\partial d_1}{\partial S_t} - Ke^{-r(T-t)} \underbrace{\frac{\partial d_2}{\partial S_t}}_{=\partial d_1/\partial S_t} \\
&= \mathrm{N}(d_1) + \underbrace{\left(S_t \mathrm{N}'(d_1) - Ke^{-r(T-t)}\mathrm{N}'(d_2)\right)}_{=0 \text{ by Proposition 6.1}} \frac{\partial d_1}{\partial S_t} \\
&= \mathrm{N}(d_1)
\end{aligned}
$$

Now using the call-put parity $S_t + P_t = C_t + Ke^{-r(T-t)}$ and taking the derivative with respect to S_t we get:

$$1 + \frac{\partial P_t}{\partial S_t} = 1 + \Delta_P = \Delta_C$$

Therefore:

$$\Delta_P = 1 - \Delta_C = 1 - \mathrm{N}(d_1) = \mathrm{N}(-d_1)$$

□

Proposition 6.4.

$$\rho_C := \frac{\partial C_t}{\partial r} = Ke^{-r(T-t)}(T-t)\mathrm{N}(d_2)$$

$$\rho_P := -\frac{\partial C_t}{\partial r} = Ke^{-r(T-t)}(T-t)\mathrm{N}(-d_2)$$

Proof. Starting with the pricing formula for a cal option and taking the derivative with respect to r we get:

$$
\begin{aligned}
\rho_C &= \frac{\partial C_t}{\partial r} \\
&= S_t \mathrm{N}'(d_1) \frac{\partial d_1}{\partial r} - Ke^{-r(T-t)}\mathrm{N}'(d_2) \underbrace{\frac{\partial d_2}{\partial r}}_{=\partial d_1/\partial r} + (T-t)Ke^{-r(T-t)}\mathrm{N}(d_2) \\
&= \underbrace{\left(S_t \mathrm{N}'(d_1) - Ke^{-r(T-t)}\mathrm{N}'(d_2)\right)}_{=0 \text{ by Proposition 6.1}} \frac{\partial d_1}{\partial r} + (T-t)Ke^{-r(T-t)}\mathrm{N}(d_2) \\
&= (T-t)Ke^{-r(T-t)}\mathrm{N}(d_2)
\end{aligned}
$$

Now to use the call-put parity $S_t + P_t = C_t + Ke^{-r(T-t)}$ and take the derivative with respect to r:

$$
\begin{aligned}
\rho_P &= \rho_C - (T-t)Ke^{-r(T-t)} \\
&= (T-t)Ke^{-r(T-t)}(\mathrm{N}(d_2) - 1) \\
&= -(T-t)Ke^{-r(T-t)}\mathrm{N}(-d_2)
\end{aligned}
$$

□

Proposition 6.5.

$$\nu = \nu_C = \frac{\partial C_t}{\partial \sigma} = \nu_P = \frac{\partial P_t}{\partial \sigma} = Ke^{-r(T-t)}\mathrm{N}'(d_2)\sqrt{T-t}$$

Proof. First let us note that $\frac{\partial C_t}{\partial \sigma} = \frac{\partial P_t}{\partial \sigma}$ by using the call-put parity. $S_t + P_t = C_t + Ke^{-r(T-t)}$ and taking the derivative with respect to σ to get:

$$\frac{\partial P_t}{\partial \sigma} = \frac{\partial C_t}{\partial \sigma}$$

Now let us start with the option pricing formula for a call option and take the derivative with respect to σ:

$$\begin{aligned}
\nu &= \nu_C \\
&= \frac{\partial C_t}{\partial \sigma} \\
&= S_t\mathrm{N}'(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{-r(T-t)}\mathrm{N}'(d_2)\frac{\partial d_2}{\partial \sigma} \\
&= S_t\mathrm{N}'(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{-r(T-t)}\mathrm{N}'(d_2)\left(\frac{\partial d_1}{\partial \sigma} - \sqrt{T-t}\right) \\
&= \underbrace{\left(S_t\mathrm{N}'(d_1) - Ke^{-r(T-t)}\mathrm{N}'(d_2)\right)}_{=0 \text{ by Proposition 6.1}}\frac{\partial d_1}{\partial \sigma} + Ke^{-r(T-t)}\mathrm{N}'(d_2)\sqrt{T-t} \\
&= Ke^{-r(T-t)}\mathrm{N}'(d_2)\sqrt{T-t}
\end{aligned}$$

□

7 Bond Pricing Equation

Proposition 7.1 (Bond Pricing Equation Derivation)**.** Let $Z_t = Z(t, r_t; T)$ be a security whose value depends on current time t, time to maturity T and the spot interest rate r_t which satisfies the following stochastic differential equation:

$$dr_t = u(t, r_t)\,dt + w(t, r_t)\,dB_t$$

Where $u(t, r_t)$ and $w(t, r_t)$ are functions and B_t is a Brownian motion. Then there exists a function $\lambda(t, r_t)$ such that Z_t satisfies the following partial differential equation known as the Bond Pricing Equation:

$$\frac{\partial Z_t}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t}{\partial r_t^2} + (u(t, r_t) + \lambda(t, r_t)\,w(t, r_t))\frac{\partial Z_t}{\partial r_t} - rZ_t = 0 \tag{25}$$

Proof. Let $Z_t^1 = Z^1(t, r_t; T_1)$ and $Z_t^2 = Z^2(t, r_t; T_2)$ be two different securities. We will create a hedging portfolio Π_t as we did in Theorem 5.1:

$$\Pi_t = Z_t^1 - \Delta Z_t^2$$

Where δ does not change over a small time increment, it is decided before the time increment then chosen again at the end of the time increment, therefore using Ito's Lemma we have the following:

$$\begin{aligned} d\Pi_t &= dZ_t^1 - \Delta dZ_t^2 \\ &= \left(\frac{\partial Z_t^1}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t^1}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t^1}{\partial r_t}\right)dt + \cdots \\ \cdots &\quad -\Delta\left(\frac{\partial Z_t^2}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t^2}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t^2}{\partial r_t}\right)dt + \cdots \\ \cdots &\quad +\left(\frac{\partial Z_t^1}{\partial r_t} - \Delta\frac{\partial Z_t^2}{\partial r_t}\right)dB_t \end{aligned}$$

Therefore, we can cancel out the dB_t term by choosing:

$$\Delta = \frac{\partial Z_t^1/\partial r_t}{\partial Z_t^2 \partial r_t}$$

Now as Π_t is not dependent on any random terms, it has to grow with the time value of money:

$$d\Pi_t = r\Pi_t dt = r\left(Z_t^1 - \frac{\partial Z_t^1/\partial r_t}{\partial Z_t^2 \partial r_t}Z_t^2\right)dt$$

Therefore, we have the following relationship:

$$\begin{aligned} d\Pi_t &= dZ_t^1 - \Delta dZ_t^2 \\ &= \left(\frac{\partial Z_t^1}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t^1}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t^1}{\partial r_t}\right)dt + \cdots \\ \cdots &\quad -\Delta\left(\frac{\partial Z_t^2}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t^2}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t^2}{\partial r_t}\right)dt \\ &= r\left(Z_t^1 - \frac{\partial Z_t^1/\partial r_t}{\partial Z_t^2 \partial r_t}Z_t^2\right)dt \end{aligned}$$

Equating the dt terms from the above relation and rearranging the terms we get the following:

$$\frac{\frac{\partial Z_t^1}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t^1}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t^1}{\partial r_t} - rZ_t^1}{\frac{\partial Z_t^1}{\partial r_t}} = \frac{\frac{\partial Z_t^2}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t^2}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t^2}{\partial r_t} - rZ_t^2}{\frac{\partial Z_t^2}{\partial r_t}} \tag{26}$$

Each side of Equation 26 are dependent on different T_i for $i = 1, 2$. This means that exists a function $\alpha(t, r_t)$ independent of T_i for $i = 1, 2$ such that each side of Equation 26 is equal to that. We can also drop the distinction of $i = 1, 2$ for Z_t^i as both sides of the equation are of the same form and generalise to a security $Z_t = Z(t, r_t; T)$ and we get:

$$\frac{\frac{\partial Z_t}{\partial t} + \frac{1}{2}w^2(t, r_t)\frac{\partial^2 Z_t}{\partial r_t^2} + u(t, r_t)\frac{\partial Z_t}{\partial r_t} - rZ_t}{\frac{\partial Z_t}{\partial r_t}} = \alpha(t, r_t)$$

Without loss of generality, for some function $\lambda(t, r_t)$ we set:

$$\alpha(t, r_t) = \lambda(t, r_t)\, w(t, r_t)$$

Finally rearranging, to get the Bond Pricing Equation:

$$\frac{\partial Z_t}{\partial t} + \frac{1}{2} w^2 (t, r_t) \frac{\partial^2 Z_t}{\partial r_t^2} + (u (t, r_t) - \lambda (t, r_t) w (t, r_t)) \frac{\partial Z_t}{\partial r_t} - r Z_t = 0$$

$\square$

8 Other

$$\begin{aligned}
\mathbb{E}[N_1] &= \mathbb{E}[N_1, T_1 = T] + \mathbb{E}[N_1, T_1 = H] && (27)\\
&= \mathbb{E}[N_1 | T_1 = T] \mathbb{P}[T_1 = T] + \mathbb{E}[N_1 | T_1 = H] \mathbb{P}[T_1 = H] && (28)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[N_1 | T_1 = T] &= \sum_{n=1}^{\infty} n \mathbb{P}[N_1 = n | T_1 = T] && (29)\\
&= \sum_{n=2}^{\infty} n \mathbb{P}[N_1 = n | T_1 = T] && (30)\\
&= \sum_{k=1}^{\infty} (k+1) \mathbb{P}[N_1 = k+1 | T_1 = T] && (31)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}[N_1 = k+1 | T_1 = T] + \sum_{k=1}^{\infty} \mathbb{P}[N_1 = k+1 | T_1 = T] && (32)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}[N_1 = k+1 | T_1 = T] + 1 && (33)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{T_i = T, i < n\}} = k+1 | T_1 = T\right] + 1 && (34)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}\left[\sum_{i=1}^{k} \mathbb{1}_{\{T_i = T\}} + \mathbb{1}_{\{T_{k+1} = H\}} = k+1 | T_1 = T\right] + 1 && (35)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}\left[\sum_{i=2}^{k} \mathbb{1}_{\{T_i = T\}} + \mathbb{1}_{\{T_{k+1} = H\}} = k\right] + 1 && (36)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}\left[\sum_{i=1}^{k} \mathbb{1}_{\{T_i = T\}} + \mathbb{1}_{\{T_k = H\}} = k\right] + 1 && (37)\\
&= \sum_{k=1}^{\infty} k \mathbb{P}[N_1 = k] + 1 && (38)\\
& && (39)
\end{aligned}$$

$$\mathbb{E}[N_1 | T_1 = T] = \mathbb{E}[N_1 - 1 | T_1 = T] + \mathbb{E}[1 | T_1 = T] \quad (40)$$

$$
\begin{aligned}
\mathbb{E}\left[N_2\right] &= \mathbb{E}\left[N_2|TT\right]\mathbb{P}\left[TT\right] + \mathbb{E}\left[N_2|HT\right]\mathbb{P}\left[HT\right] + \mathbb{E}\left[N_2|TH\right]\mathbb{P}\left[TH\right] + \mathbb{E}\left[N_2|HH\right]\mathbb{P}\left[HH\right] \\
&= 0.25\left(2 + \mathbb{E}\left[N_2\right]\right) + 0.25\left(2 + \mathbb{E}\left[N_2\right]\right) + 0.25\left(0.5 \times 3 + 0.5\left(3 + \mathbb{E}\left[N_2\right]\right)\right) + 0.25 \times 2 \\
&= 2.25 + 5/8\mathbb{E}\left[N_2\right] +
\end{aligned}
$$