Sets

Common Sets

- $\mathbb{N}$ is the set of positive integers: $\{1, 2, 3, 4, ...\}$
- $\mathbb{Z}$ is the set of all integers: $\{..., -1, 0, 1, 2, 3, ...\}$
- $\mathbb{Q}$ is the set of rational numbers, those that can be written as fractions
- $\mathbb{R}$ is the set of real numbers.
- $\mathbb{C}$ is the set of complex numbers.

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

Some numbers cannot be written as fractions, and these we describe as irrational. These include π, e, $\sqrt{2}$, etc. These numbers are not in $\mathbb{Q}$, but they are real numbers, i.e. in $\mathbb{R}$. It can be difficult to prove numbers are irrational. An even larger set of numbers exists, that of the complex numbers $\mathbb{C}$. “Most” real numbers are inrrational.

Denoting Membership

If A is a set, saying $x \in A$ means x is in the set A. We can also use $\notin$ to denote non-membership. Sets can be expressed as 2-dimensional regions, with points within them as elements.

- $n \in \mathbb{Z}$ means n is an integer.
- $x \in \mathbb{Q}$ means x can be expressed as a fraction.

Note that n is used in the first example while x is used in the second; conventionally, n is used to represent integers.

Desciptive Notation

Descriptive notation is a way of defining a set by stating a property which all of its elements possess. For example:

$$A = \{x \in \mathbb{R} | x^2 + 1 > 37\}$$

Which can be read as the set of all real numbers x such that $x^2 + 1 > 37$. The statement must be a "predicate"; it must be true or false for all values. The vertical bar is read as "such that". A colon is sometimes used instead.

Examples

- Express the set of real numbers whose natural (base e) logarithm is positive with descriptive notation: $x \in \mathbb{R} | log(x) > 0$. Note that $log(x)$ denotes $log_e(x)$. There is however an issue with this answer; because $log(x)$ is sometimes undefined, "$log(x) > 0$" is not a predicate. The set could instead be expressed as:

$$\{x \in \mathbb{R}_{>0} | log(x) > 0\} \quad \mathbb{R}_{>0} = \{x \in \mathbb{R} | x > 0\}$$

- Express the set of integers whose cube is even in descriptive notation:

$$\{n \in \mathbb{Z} | n^3 \text{ is even}\}$$

- Describe the set $\{n \in \mathbb{N} | sin(n) > 0\}$ in words:

  The set of all natural numbers n such that sin(n) is greater than 0.

Abbreviated Notation

Set notation is often abbreviated. For example $\{x \in \mathbb{R} | sin(x) = 0\}$ could be expressed as:

$$\{k\pi | k \in \mathbb{Z}\}$$

This will be all integer multiples of pi; i.e. all values for which sin(x) = 0. This could also be expressed as:

$$\{x \in \mathbb{R} | x = k\pi \text{ for some } k \in \mathbb{Z}\}$$

This can be thought of as a kind of "generating" notation, whereas descriptive notation could be seen to excise a set from a larger set through a condition.

Examples

- Express the set of odd integers in abbreviated set notation:

$$\{2k+1|k \in \mathbb{Z}\} \text{ or } \{2k-1|k \in \mathbb{Z}\}$$

- Express the set $\{x \in \mathbb{R}|cos(x) = 0\}$ in abbreviated set notation:

$$\{\frac{\pi}{2} + k\pi|k \in \mathbb{Z}\}$$

Intervals

(a, b) means the set of real numbers between a and b (exclusive). Therefore:

$$(a, b) = \{x \in \mathbb{R}|a < x \,\text{and}\, x < b\} = \{x \in \mathbb{R}|a < x < b\}$$

This can be illustrated on a number line with an open circle at a and b. This is an open interval. To denote a closed interval (inclusive of a and b), we use a square brackets ($[a, b]$) and a filled circle.

Examples

Give definitions of $[a, b)$ and $(a, b]$ with diagrams:

- $[a, b)$ represents all numbers from a (inclusive) to a (exclusive). On a number line, it would be drawn as a filled dot at a and an open dot at b.

- $(a, b]$ represents all numbers greater than a and less than or equal to b. It could be illustrated with a number line with an empty dot at a and a filled dot at b.

Unbounded Intervals

For any $a \in \mathbb{R}$,

$$(-\infty, a) = \{x \in \mathbb{R}|x < a\}$$

Note that $-\infty \,\text{and}\, \infty$ are not elements of $\mathbb{R}$.

Small Sets

Sets with finitely many elements can be described with a list of elements.

$$\{x \in \mathbb{R} | x^2 - 1 = 0\} = \{-1, 1\}$$

$$\{x \in \mathbb{Z} | x^2 - 1 < 0\} = \{0\}$$

$$\{x \in \mathbb{Z} | x^3 - x = 0\} = \{-1, 0, 1\}$$

Example

Write the set of prime numbers less than 20 in descriptive and list of elements form:

$$\{2, 3, 5, 7, 11, 13, 17, 19\}$$

We can use a kind of "list of elements" notation to denote some infinite sets with ellipses. e.g.:

$$(2k+1)\pi | k \in \mathbb{Z} = \{..., -3\pi, -\pi, \pi, 3\pi, ...\}$$

The smallest set is taken to be the empty set, written as $\emptyset$. It can also be written as $\{\}$ in list of elements form.

Examples

$$\{x \in \mathbb{R} | x^2 + 1 = 0\} = \emptyset$$

$$\{x \in \mathbb{R} | cos(x) > 1\} = \emptyset$$

Subsets

If A and B are sets then $A \subseteq B$ means that every element of A is also an element of B. It is read as A is a subset of B.

$$x \in A \Rightarrow x \in B$$

The empty set is taken to be a subset of every set.

Proofs

A proof begins with a set of true assumptions, from which mathematical reasoning is used to prove a conclusion. To prove $A \subseteq B$, we need to show that if something meets the criteria of A, it must also meet those of B.

Example

Prove $A \subseteq B$ where:

$$A = \{n \in \mathbb{N} | sin(n) > 0\}$$
$$B = \{n \in \mathbb{N} | sin^2(n) \leq sin(n)\}$$

Let $n \in A$, thus $sin(n) > 0$.
Since $sin(n) \leq 1$ and if $x \leq y$ and $a > 0$ then $ax \leq ay$, we have $sin(n)sin(n) \leq sin(n)$. Therefore $n \in B$ so $A \subseteq B$.

To prove a statement is false, we must find a counterexample - an example for which a statement does not hold. Thus, to prove A is not a subset of B, we must find any x in A but not in B.

Example

Prove that $A \nsubseteq B$ where $A = \{3n+1 | n \in \mathbb{Z}\}$ and $B = \{6m+1 | m \in \mathbb{Z}\}$.

$$A = \{..., -2, 1, 4, 7, ...\}$$
$$B = \{..., 1, 7, ...\}$$

We guess $4 \notin B$, $4 \in A$. To prove this:

- Claim $4 \in A$; this is true since $4 = 3 * 1 + 1$.
- Claim $4 \notin B$.
- Suppose for a contradiction that $4 \in B$.
- This would mean $4 = 6m + 1$ for some $m \in \mathbb{Z}$.
- But then, $3 = 6m$

$$\therefore m = \frac{3}{6} = \frac{1}{2}$$

This is a contradiction! Thus we conclude $4 \notin B$.

Proof by Contradiction

How do we prove that $x^2 + x + 1 = 0$ has no real solutions? Proof by contradiction. Assume that there is a solution, $x \in \mathbb{R}$ and show that this leads to an absurd conclusion.

$$x^2 + x + 1 = 0 \Rightarrow (x + \frac{1}{2})^2 + \frac{3}{4} = 0 \Rightarrow (x + \frac{1}{2})^2 = \frac{-3}{4}$$

Note that $p \Rightarrow q$ means if p is true, so is q. $p \Leftrightarrow q$ means that if either is true, both are true; p is true *if and only if* q is true.

Union

The elements of $A \cup B$ are all the elements that are in either A or B.

Examples

Express the following sets using a union of intervals.

- $\{x \in \mathbb{R} | x^2 > 1\}$: $(-\infty, -1) \cup (1, \infty)$
- $\{x \in (-2\pi, 2\pi] | sin(x) \leq 0\}$:

  $\{x \in (-2\pi, 0] | sin(x) \leq 0\} \cup \{x \in (0, 2\pi] | sin(x) \leq 0\} = [-\pi, 0] \cup [\pi, 2\pi]$

- $\{x \in [-2, 2] | x \notin \mathbb{Z}\}$: $(-2, -1) \cup (-1, 0) \cup (0, 1) \cup (1, 2)$

$$(2, 8) \cup [3, 10] = (2, 10]$$

Thus, $(2, 8) \cup [3, 10]$ is an interval. However is $(0, \sqrt{2}) \cup [\frac{\pi}{2}, 3)$ an interval? It is not, because $\sqrt{2} < \frac{\pi}{2}$, thus two pieces are necessary.

Intersection

$A \cap B$ is the set of all elements that are in both A and B.

Examples

- Express $(2,8) \cap [3,10]$ as an interval: $[3,8)$
- Express $(0,\sqrt{2}) \cap [\frac{\pi}{2},3)$ in the simplest way possible: $\emptyset$
- Express $\mathbb{Z} \cap [-\pi,\pi]$ in list of elements form: $\{-3,-2,-1,0,1,2,3\}$
- Express $\mathbb{Z} \cap \{x \in \mathbb{R} | x^2 - 5 < 0\}$ in list of elements form: $\{-2,-1,0,1,2\}$
- Express the set of reals with positive sine and negative cosine with an intersection: $\{x \in \mathbb{R} | \sin(x) > 0\} \cap \{x \in \mathbb{R} | \cos(x) < 0\}$

When interpreting questions, a question involving "and" is likely to be an intersection question.

Complement

For A and B the *relative complement* of A and B is the set of elements that are in A but not in B. This is written as:

$$A \backslash B = \{x \in A | x \notin B\}$$

Examples

- Find $(0,2)\backslash(1,3)$: $(0,1]$
- Find $(1,3)\backslash(0,2)$: $[2,3)$
- Is it generally true that $A\backslash B = B\backslash A$. No. See previous questions for example.

Express each of the following as a complemet:

- $\{x \in \mathbb{R} | x^2 > 1\}$: $\mathbb{R}\backslash\{x \in \mathbb{R} | x^2 < 1\}$
- $\{x \in [-2,2] | x \notin \mathbb{Z}\}$: $[-2,2]\backslash\mathbb{Z}$
- $(-\infty,0) \cup (0,\infty)$: $\mathbb{R}\backslash\{0\}$

Cartesian Product

For sets A and B:

$$A \times B = \{(x, y) | x \in A \,\text{and}\, y \in B\}$$

a, b only equals a, b, not b, a.
e.g.:

$$\{0, 1, 5\} \times \{e, \pi\} = \{(0, e), (0, \pi), (1, e), (1, \pi), (5, e), (5, \pi)\}$$

We can also take A^2 to be $A \times A$. This is most often seen in $\mathbb{R}^2$, which is used to represent the cartesian plane. $\mathbb{Z}^2 \subseteq \mathbb{R}^2$.

Inequalities

When we add a constant to an inequality, the inequality is order preserved, it still faces the same direction. Subtraction, obviously, obeys the same principle.

$$x > y \Rightarrow x + a > y + a$$

When we multiply by a positive constant, $a > 0$, is order preserving, while multiplying by a negative constant, $a < 0$, is order reversing. As division is simply multiplication by reciprocal, it follows the same principles as multiplication; i.e. only order reversing if the divisor < 0.

Examples

Express the set $A = \{x \in \mathbb{R} | -2 - \frac{1}{2}x > -4\}$ as an interval.

- $2 - \frac{1}{2}x > -2$
- $-\frac{1}{2} > -2$
- $x < 4$
- $x \in (-\infty, 4)$

Express the set $A = \{x \in \mathbb{R} | 1 - x < 3x + 2\}$ as an interval.

- $-1 - x < 3x$
- $-1 < 4x$
- $-\frac{1}{4} < x$
- $x > -\frac{1}{4}$
- $A = (-\frac{1}{4}, \infty)$

Transitivity

$$x < y \text{ and } y < z \Rightarrow x < z$$

Example

Prove that $x < y \text{ and } a < b \Rightarrow x + a < y + b$.

- $0 < y - x$
- $a - b < 0$
- $a - b < y - x$
- $a < y - x + b$
- $a + x < y + b$

Reversing and Non-Reversing Functions

A function is order preserving if it is strictly increasing for the interval containing both sides of an inequailty. It is order-reversing if it is strictly decreasing for the same. i.e. a function is order preserving if:

$$a < b \Rightarrow f(a) < f(b) \text{ for all } a, b \in I$$

and order reversing if:

$$a < b \Rightarrow f(a) > f(b) \text{ for all } a, b \in I$$

For example, the logarithm and exponential function are order preserving. A function need not be strictly decreasing across $\mathbb{R}$ to be applied to an inequality; if both sides of an inequality are known to lie within an interval that is strictly decreasing or increasing, the function may still be applied.

Complex Numbers

Complex numbers extend the real numbers by adding a new number with the property

$$i^2 = -1$$

The usual rules of algebra nonetheless still apply. Though one should note that:

$$\sqrt{-1}\sqrt{-1} \neq \sqrt{1}$$

This value can be used to solve equations such as:

$$x^2 + 1 = 0 \Rightarrow x^2 = -1 \ \Rightarrow x = \pm i$$

i can also be used to solve equations such as $x^2 + 4 = 0$:

$$x^2 + 4 = 0 \Rightarrow x^2 = \sqrt{-2} \Rightarrow x = i\sqrt{2}$$

Solutions of the form $yi|y \in \mathbb{R}$ are useful for solving many polynomials. For some others, such as $x^2 - 2x + 5 = 0$ we need solutions of the form $x + iy$:

$$(x-1)^2 + 4 = 0 \ \Rightarrow (x-1)^2 = -4 \Rightarrow x = 1 \pm 2i$$

Thus, we define complex numbers, typically denoted as z as a quantity consisting of a real number added to a real multiple of i:

$$z = x + iy$$

Thus $\mathbb{C}$ is defined as:

$$\mathbb{C} = \{x + iy|x, y \in \mathbb{R}\}$$

Examples

- Using i write down two square roots of -25: $5i$, $-5i$
- Simplify i^7: $i^0 = 1$, $i^1 = i$, $i^2 = -1$, $i^3 = ii^2 = -i$, $i^4 = (i^2)^2 = 1$, $i^5 = i$, $i^6 = -1$, $i^7 = i(i^2)^3 = -1^3 = -i$

For a complex number, the real part is x while y is known as the imaginary part.

Example

For $z = 2 - 3i$, write down:

- $\mathrm{Re}(z) = 2$
- $\mathrm{Im}(z) = -3$
- $\mathrm{Re}(z) - \mathrm{Im}(z) = 5$

The Complex Plane

A complex number z can be seen as a point on a plane, where $\mathrm{Re}(z)$ denotes the x or horizontal component while $\mathrm{Im}(z)$ denotes the y or vertical component. This is sometimes known as the Argand plane.

Addition and subtraction work as expected with complex numbers.

$$z_1 + z_2 = (a + ib) + (c + id) = (a + c) + i(b + d)$$

This can be interpreted as a vector operation, in the same way that adding two vectors entails summing their components, complex numbers seen as points are effectively vectors which add in essentially the same way. Multiplication by a real number is essentially stretching the complex number, the same as multiplying a vector by a scalar.

Multiplication is a little more complex.

$$(a + ib)(c + id) = a(c + id) + ib(c + id) =$$

$$ac + aid + ibc + ibid = ac + adi + bci - bd = (ac - bd) + i(ad + bc)$$

The complex conjugate of a number $z = a + ib$ is denoted $\bar{z}$ and is defined:

$$\bar{z} = a - ib$$

The real part remains the same, while the imaginary part has its sign reversed. In the plane, this operation represents reflection in the x or real axis.

Example

By completing the square find the solutions of $z^2 - 6z + 10 = 0$:

$$z^2 - 6z + 10 = (z-3)^2 + 1$$

$$(z-3)^2 = -1$$

$$z - 3 = \pm i$$

$$z = 3 \pm i$$

Properties of Conjugates

Let $z = x + iy$ and $w = a + ib$ be complex numbers. Then:

$$z + \bar{z} = 2x = 2\mathrm{Re}(z) \text{ (real)}$$

$$z - \bar{z} = 2yi = 2\mathrm{Im}(z)i$$

$$z\bar{z} = x^2 + y^2 \text{ (real)}$$

$$\bar{z + w} = \bar{z} + \bar{w}$$

$$\bar{zw} = \bar{z}\bar{w}$$

$\sqrt{z\bar{z}}$ can be understood as the length of a complex number by Pythagoras' theorem.

Division

$$\frac{1}{i} = -i$$

In general:

$$\frac{\bar{z}}{x^2 + y^2} = \frac{1}{z}$$

We can express the division of two complex numbers as follows:

$$\frac{a+ib}{c+id} = \frac{a+ib}{c+id} \times \frac{c-id}{c-id} = \frac{(a+ib)(c-id)}{c^2+d^2}$$

$$= \frac{1}{c^2+d^2}(ac + bd + i(-ad + bc))$$

In essence, we rationalise the denominator, much as with surds.

Examples

$$\frac{1+2i}{-1+3i} = \frac{1+2i}{-1+3i} \times \frac{-1-3i}{-1-3i} = \frac{(1+2i)(-1-3i)}{1+9}$$
$$= \frac{1-i}{2}$$

Calculate: $(1+\sqrt{3}i)^6$

$$(1+\sqrt{3}i)^6 = ((1+\sqrt{3}i)^2)^2 \times (1+\sqrt{3}i)^2$$
$$(1+\sqrt{3}i)^2 = 1 + 2\sqrt{3}i + (\sqrt{3}i)^2 = 1 + 2\sqrt{3}i - 3 = -2 + 2\sqrt{3}i$$
$$((1+\sqrt{3}i)^2)^2 = (-2+2\sqrt{3}i)^2 = (-2)^2 + 2 \times (-2) \times 2\sqrt{3}i + (2\sqrt{3}i)^2$$
$$= 4 - 8\sqrt{3}i - 12 = -8 - 8\sqrt{3}i$$
$$(1+\sqrt{3}i)^6 = (-8-8\sqrt{3}i)(-2+2\sqrt{3}i)$$
$$= 16 - 16\sqrt{3}i + 16\sqrt{3}i - 8 \times 2 \times (\sqrt{3})^2 i^2$$
$$= 16 + 8 \times 2 \times 3$$
$$= 64$$

Polar Form

To solve this problem more easily, we can instead consider complex numbers in polar form. This takes the length of the line and it's angle from the origin instead. In this case we take a distance r and an angle θ.

$$r = |z| = \sqrt{x^2 + y^2}$$

To find θ we can draw z in the plane and use triangles to find θ. $\theta \in (-\pi, \pi]$, and is called the principal argument of z and can be denoted $\mathrm{Arg}(z)$.

Examples

- $1+\sqrt{3}i$: $|z| = 2, \theta = \frac{\pi}{3}$
- $-3-3i$: $|z| = \sqrt{18} = 3\sqrt{2}, \theta = \frac{-3\pi}{4}$
- $3+4i$: $|z| = 5, \theta = \arctan(\frac{4}{3})$
- -1: $|z| = 1, \theta = \pi$

The trigonemtric polar form of a complex number is:

$$z = r(\cos(\theta) + i\sin(\theta)))$$

The Complex Exponential

For any θ, the complex exponential is given by:

$$e^{i\theta} = 1 + i\theta + \frac{1}{2!}(i\theta)^2 + \frac{1}{3!}(i\theta)^3 \ldots$$
$$e^{i\theta} = \cos(\theta) + i\sin\theta$$
$$\left|e^{i\theta}\right| = 1$$

Exponential Polar Form

Knowing this, we can write a given complex number z in exponential polar form:

$$z = re^{i\theta}$$

This happens to be an extremely useful way of expressing complex numbers.
An example:

Express $z = -2 + 2\sqrt{3}i$ in exponential polar form.

$$z = 4(-\frac{1}{2} + \frac{\sqrt{3}}{2}i)$$

$$4e^{i\frac{2\pi}{3}}$$

The complex exponential indicates a point on the unit circle, with the scalar r modifying this vector to point to the actual location of z.

The conjugate of a number in exponential form is given by:

$$e^{-i\theta} = \overline{e^{i\theta}}$$

It takes the form of inversion across the x-axis.

Properties of the Complex Exponential

$$e^{i0} = 1$$
$$e^{i\theta_1}e^{i\theta_2} = e^{i\theta_1 + i\theta_2} = e^{i(\theta_1+\theta_2)}$$
$$\frac{e^{i\theta_1}}{e^{i\theta_2}} = e^{i\theta_1 - i\theta_2} = e^{i(\theta_1-\theta_2)}$$

$$e^{i\theta_1} = e^{i\theta_2} \Leftrightarrow \theta_2 = \theta_1 + 2k\pi$$

Multiplication by $e^{i\theta}$ is equivalent to rotation byθ. For two complex numbers, z_1 and z_2, the product is given by:

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)}$$

Thus, the modulus of the result is simply the product of the two inputs modulus. The angle is increased by θ. It is intuitive that multiplication by i^2 constitutes inversion of the vector; reflection in the y-axis. The resultant angle will not necessarily be the principle argument; one can simply add or subtract 2π to find the principle argument.

Division is simply given by:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$$

Exponentiation works through:

$$z^n = r^n e^{in\theta}$$

Subsets of the Complex Plane

We can represent certain shapes in the complex plane through sets. For example: $\{z \in \mathbb{C} | \mathrm{Re}(z) = 3\}$ describes the line $x = 3$. A circle can be drawn as $\{z \in \mathbb{C} : |z - p| = r\}$, where p is the center point and r is the radius.

Roots

Suppose we are trying to find all of the n-th roots of a complex number w. This means we need to find all zs that satisfy $z^n = w$. Writing $z = re^{i\theta}$ gives us:

$$(re^{i\theta})^n = se^{i\alpha} \Rightarrow se^{i\alpha}$$

$$\Rightarrow |z| = |s| \Rightarrow r^n = s \Rightarrow r = s^{\frac{1}{n}} = \sqrt[n]{s}$$

We also have:

$$e^{in\theta} = e^{i\alpha} \Rightarrow n\theta = \alpha + 2k\pi \Rightarrow \theta = \frac{1}{n}(\alpha + 2k\pi)$$

Curiously, for power n, there will be n solutions. Using the $\alpha + 2k\pi$ equation, we can find an infinite number of solutions. However, we can stop at $n - 1$, because $n = k$ will yield the same result as $k = 0$, as it will simply add $2pi$ to the angle. Thus our solutions will be of the form:

$$\{\sqrt[n]{s} e^{i\frac{1}{n}(\alpha+2k\pi)} | k \in [0, n) \cap \mathbb{N}\}$$

Example

Find the set of cube roots of -8.

- Solve to find z such that $z^3 = -8$
- Thus $s = 8$, $\alpha = \pi$
- $z = re^{i\theta}, r^3 e^{i3\theta} = 8e^{i\pi}$
- $\therefore r^3 = 8 \Rightarrow r = 2$
- $3\theta = \pi + 2k\pi, k \in 0, 1, 2$
- $\therefore \theta = \frac{\pi}{3}, \pi, \frac{5\pi}{3}$
- $z \in \{2e^{i\frac{pi}{3}}, 2e^{i\pi}, 2e^{i\frac{5\pi}{3}}\}$
- Note we found the expected root; $2e^{i\pi} = -2$

These solutions represent 3 evenly spaced points on the circle with center $(0, 0)$ and radius 2.

Trigonometric Functions

The complex exponential can be used to express $\cos$ and $\sin$ in a useful form for many calculations.

$$
\begin{aligned}
e^{i\theta} &= \cos(\theta) + i\sin(\theta) \\
\overline{e^{i\theta}} &= \cos(\theta) - i\sin(\theta) \\
e^{i\theta} + \overline{e^{i\theta}} &= 2\cos(\theta) \\
\therefore \cos(\theta) &= \frac{1}{2}(e^{i\theta} e^{-i\theta})
\end{aligned}
$$

This is obvious when we consider that $z + \bar{z} = 2\mathrm{Re}(z)$. $\sin$ found through the same logic is:

$$\sin(\theta) = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})$$

Using the binomial formula, we can expand statements such as $\sin^4(\theta)$:

$$\sin^4(\theta) = (\frac{1}{2i})^4 (e^{i\theta} - e^{-i\theta})^4$$

$$\frac{1}{2^4}((e^{i\theta})^4 - 4(e^{i\theta})^3e^{-i\theta} + 6(e^{i\theta})^2(e^{-i\theta})^2 - 4e^{i\theta}(e^{-i\theta})^3 + (e^{-i\theta})^4)$$

$$\frac{1}{2^4}(e^{i4\theta} - 4e^{i2\theta} + 6 - 4e^{-i2\theta} + e^{-i4\theta})$$

$$6 + 2\cos(4\theta) - 4 \times 2\cos(2\theta)$$

$$\sin^4(\theta) = \frac{1}{8}(3 + cos(4\theta) - 4cos(2\theta))$$

Using this approach, we can convert expressions of the form $\sin^m(\theta)\cos^n(\theta)$ to an equivalent form that is often more useable.

Square Roots

For $w = re^{i\theta} \in \mathbb{C}\backslash\{0\}$, the root finding formula gives two solutions of $z^2 = w$:

$$z_1 = \sqrt{r}e^{i(\frac{\theta}{2})}, z_2 = \sqrt{r}e^{i(\frac{\theta}{2}+\pi)} = -z_1$$

z_1 is called the *principal square root* of w, thus the square roots are $\pm\sqrt{w}$. For a negative real number, $\sqrt{w}$ will be equal to $\sqrt{|w|}i$.

Polynomials

Real polynomials are those with coefficients exclusively in $\mathbb{R}$, while complex polynomials extend this bank of coefficients to include all of $\mathbb{C}$. All real polynomials are complex polynomials, but the inverse is not necessarily true.

Factorising Polynomials

To solve differential equations, it is important to be able to integrate polynomials of the form:

$$\frac{p(x)}{q(x)}$$

We can use the quadratic formula to solve complex polynomials just as we would for real numbers.

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

For example, factorising $p(z) = z^2 - 3iz - 2$:

$$\frac{3i \pm \sqrt{-9 - 4 \times (1) \times (-2)}}{2}$$

$$= \frac{3}{2}i \pm \frac{1}{2}\sqrt{-1} = \frac{3}{2}i \pm \frac{1}{2}i$$

$$\therefore p(z) = (z - 2i)(z - i)$$

Any complex polynomial of degree n can be factorised into n linear factors over $\mathbb{C}$.

$$p(z) = a(z - \alpha_1)(z - \alpha_2)...(z - \alpha_n)$$

Thus, there is an x intercept at each value of α. For a polynomial of degree n, there at $\leq n$ roots in $\mathbb{C}$.

If $p(z)$ is a real polynomial; the non-real roots of $p(z)$ occur in complex conjugate pairs z and $\bar{z}$. For example:

$$p(z) = z^3 - 3iz^2 - 2z = z(z^2 - 3iz - 2) = z(z - 2i)(z - i) = (z - 0)(z - 2i)(z - i)$$

Because in this example, our polynomial is non-real, the roots are not in conjugate pairs. How about:

$$\{z \in \mathbb{C} | z^4 + z^2 - 12 = 0\}$$

$$\text{set } w = z^2 \Rightarrow p(z) = w^2 + w - 12 = 0$$

Now, we need to factor $p(z)$ in terms of w, and solve $z^2 = w$.

$$w^2 + w - 12 = (w - 3)(w + 4)$$

$$z^2 = 3 \Rightarrow z = \sqrt{3}, -\sqrt{3}$$

$$z^2 = -4 \Rightarrow z = 2i, -2i$$

$$\{\sqrt{3}, -\sqrt{3}, 2i, -2i\}$$

In this example, because our roots are real, our solutions come in complex conjugate pairs.

If the coefficients of a polynomial $p(x)$ are all real, then $p(x)$ can be written as a product of linear and irreducible quadratic factors with exclusively real coefficients. Considering the previous example:

$$p(z) = (z - \sqrt{3})(z + \sqrt{3})(z - 2i)(z + 2i)$$
$$= (z - \sqrt{3})(z + \sqrt{3})(z^2 + 2iz - 2iz - 4i^2)$$
$$= (z - \sqrt{3})(z + \sqrt{3})(z^2 + 4)$$

Geometric Progression

For any $z \in \mathbb{C} \backslash \{1\}$:

$$z^n + z^{n-1} + ... + z + 1 = \frac{z^{n+1} - 1}{z - 1}$$

Using this, let us factor $p(z) = z^3 + z^2 + z + 1$ over $\mathbb{R}$.

$$(z - 1)(z^3 + z^2 + z + 1) = z^4 - 1$$

The roots of the RHS are the 4th roots of $w = 1$ which are:

$$1, -1, i, -i$$

We can also discover these by the general method:

$$z^4 = 1 \Rightarrow r^4 e^{i4\theta} = 1e^{i0}$$

$$\Rightarrow r = 1,\ 4\theta = 2k\pi,\ k = 0, 1, 2, 3 \ \ \theta = \frac{1}{2}k\pi,\ k = 0, 1, 2, 3$$

$$= 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2} = 1, i, -1, -i$$

Because these are the roots of $z^4 - 1$, and we know that $z - 1$ has the root of 1, we know that the roots of our polynomial are:

$$z^3 + z^2 + z + 1 = (z + 1)(z - i)(z + i)$$

Functions

Functions are a way of expressing quantities that depend on one another. For example, a function could express the temperature of a beverage placed on a counter with respect to the time passed since the object was placed there.

More technically, a function accepts an input and maps it to an output. These inputs are drawn from a nonempty set A called the *domain* of f. The outputs are drawn from a second nonempty set, B, known as the codomain of f. Finally there is a rule that associates each element of A to a unique element of B. To express a function f with domain A and codomain B we write:

$$f : A \to B$$

This can also be read as f maps A into B. For example

$$f : \mathbb{R} \to \mathbb{R}, f(x) = \frac{x+1}{x^2+4}$$

is a valid function. However,

$$h : \mathbb{C} \to \mathbb{C}, h(z) = \frac{x+1}{x^2+4}$$

is not, as the rule $h(z)$ is undefined for some $z \in \mathbb{C}$. To make it a valid function, we could change the domain to $\mathbb{C}\backslash\{2i, -2i\}$. When a function has a domain and codomain as subsets of $\mathbb{R}$, we can visualise it as a graph. Formally, a graph of $f : A \to B$ is the set defined by:

$$\{(a, f(a))|a \in A\} \subseteq A \times B$$

i.e. all points that can be generated by the form (input, output). The codomain of a function does not necessarily need to be minimal or maximal. For $f(x) = \sin(x)$, the range of the function is $[-1, 1]$, but any codomain inclusive of this range (such as $\mathbb{R}$) can be chosen.

The image of a set under a function is the outputs of each element of that set when passed through the function. So, for a function $f : A \to B$ and a set $S \subseteq A$, the image of S under f is:

$$f(S) = \{b \in B|b = f(s), s \in S\} = \{f(s)|s \in S\}$$

The range of a function is simply the image of the domain under the function. This range is the smallest valid codomain for the function, and thus $\text{range} \subseteq \text{codomain}$ in all cases.

Description of Functions

Injective functions are one-to-one functions. This means that for a given input, there will be a unique output, shared by no other inputs; for two different inputs, there will be two different outputs. To prove injectivity, we can show $f(x) = f(y) \Rightarrow x = y$.

Monotone functions are either increasing or decreasing across their entire domain. These functions are always injective.

Surjective functions map at least one input value to each legal output value. So for each element of the codomain, there is at least one element of the domain which maps to it. This can be thought of as the function having a minimal codomain; only possible outputs are included in the codomain.

Bijective functions are both injective and surjective. There is a precise correspondence between each element of A and each element of B; each element in A has a partner in B with no excess elements in B. For example a simple bijective function is

$$f : (0, \infty) \to (0, \infty) \quad f(x) = x$$

A slightly more complex example could be:

$$g : (0, \infty) \to (0, \infty) \quad g(x) = \frac{1}{x}$$

Function Equality

For functions to be equal, all of their attributes must be the same. This implies that the domain, codomain and function are the same. However, the syntax of the function could be different; $f(x) = sin(2x)$ is equivalent to $g(x) = 2sin(x)cos(x)$.

Composition

Composition of functions entails taking the output of one function and feeding it into another. i.e. taking $f(g(x))$. This is denoted by:

$$(f \circ g)(x) = f(g(x))$$

For a function composition to be valid the codomain of f must be a subset of the domain of g. The function $f \circ g$ will accept an argument from the domain of f and output from the codomain of g. If $f : A \rightarrow B$ and $g : C \rightarrow D$:

$$(f \circ g) : A \rightarrow D$$

Composition is not commutative: $g \circ f \neq f \circ g$. It is however associative: $h \circ (f \circ g) = (h \circ f) \circ g$.

When composing functions, we may know the domain and codomain of each of the parts. For example, taking the following functions, we know:

$$f : (0, \infty) \rightarrow \mathbb{R} \quad f(x) = \log(x)$$

$$g : \mathbb{R} \rightarrow (-\infty, 1] \quad g(x) = 1 - x^2$$

But what of the composition?

$$(f \circ g)(x) = \log(1 - x^2)$$

Let us consider this problem more generally. For two functions:

$$f : \mathrm{dom}(f) \rightarrow \mathrm{range}(f)$$

$$g : \mathrm{dom}(g) \rightarrow \mathrm{range}(g)$$

For composition to be possible, $\mathrm{range}(g)$ and $\mathrm{dom}(f)$ must intersect. Otherwise, the resultant function would have an empty domain. Thus, the implied domain of the resultant function is:

$$\mathrm{dom}(f \circ g) = \{x \in \mathrm{dom}(g) | g(x) \in \mathrm{dom}(f)\}$$

And the implied range is given by:

$$\mathrm{range}(f \circ g) = f(\mathrm{range}(g) \cap \mathrm{dom}(f))$$

So for our original functions:

$$\mathrm{dom}(f \circ g) = \{x \in R | g(x) \in (0, \infty)\} = (-1, 1)$$

$$\mathrm{range}(f \circ g) = f((-\infty, 1] \cap (0, \infty)) = f((0, 1]) = (-\infty, 0]$$

Inverses

If $f : A \to B$ the inverse of f will be $g : B \to A$ such that $g \circ f(x) = x$ and $f \circ g(y) = y$ i.e. f undoes g and g undoes f. Inverse functions are always unique. A function must be bijective to have an inverse function. To find an inverse function, one can observe the operations carried out by the function and then execute the inverses of these operations in the reverse order.

Inverse functions are used implicitly to solve equations such as $e^{x^3+1} = e^3$: it is our knowledge that an inverse function to e^x exists $(\log(x))$ that allows us to state with confidence that $x^3 + 1 = 3$ and solve the equation.
For example, we can ask does $f : \mathbb{C} \to \mathbb{C} \quad f(z) = \bar{z}$ have an inverse? Is it bijective?

Yes it does. Since $\bar{\bar{z}} = z$, f is its own inverse.

Sometimes, we will want to find an inverse of a function which is not bijective. To create a bijective function, we can restrict domain or codomain to ensure these characteristics. To create an injective function, we can find a range which is strictly increasing or decreasing and restrict the domain to this range. To create a surjective function, we can simply remove all elements of the codomain which the input doesn't map to i.e. we can restrict the codomain of the function to its range.

As an example, the $\sin$ function is not injective. Thus, we take the domain $[\frac{-\pi}{2}, \frac{\pi}{2}]$ to yield the inverse function arcsine. Note that the range restrictions mean that while $\theta = \arcsin(y) \Rightarrow \sin(\theta) = y$, $\sin(\theta) = x \Rightarrow \theta = \arcsin(x)$ only if $\theta \in [\frac{-\pi}{2}, \frac{\pi}{2}]$

Graphically, an inverse function can be understood stood as reflection across the line $x = y$.

Inverse Trigonometric Functions

Alongside $\arcsin$ we additionally have $\arccos$ and $\arctan$. While $\arctan$ has the same domain of $[\frac{-\pi}{2}, \frac{\pi}{2}]$ as $\arcsin$, $\arccos$ takes the domain $[0, \pi]$ as it is strictly positive in $[\frac{-\pi}{2}, \frac{\pi}{2}]$.

$$\text{cosec}(x) = \frac{1}{\sin(x)}$$

$$\text{secant}(x) = \frac{1}{\cos(x)}$$
$$\cot(x) = \frac{\cos(x)}{\sin(x)}$$

It turns out that $\cot(x)$ is not strictly the inverse function of $\tan(x)$. Because $\tan x = 0$ for multiples of π, these values would be excluded from the function of $\frac{1}{\tan(x)}$, which they are in $\cot(x)$, however, the values for which $\tan(x)$ is undefined, $(\{k\pi + \frac{\pi}{2} | k \in \mathbb{Z}\})$ are defined for $\cot(x)$ thus yielding a different domain to that of $\frac{1}{\tan(x)}$.

Vectors

Because not all quantities are measurable by a scalar value, we sometimes need to reach for a vector, which by being a representation of multiple values can offer both magnitude and direction. Formally, a vector is a member of the set $\mathbb{R}^2$ where

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \big\{(a_1, a_2) | a_1 \in \mathbb{R} \,\text{and}\, a_2 \in \mathbb{R}\big\}$$

We can consider higher dimensions through n-dimensional space with $\mathbb{R}^n$:

$$\mathbb{R}^n = \big\{(a_1, a_2, ..., a_n) | a_1, a_2, ..., a_n \in \mathbb{R}\big\}$$

Vectors have definitions of arithmetical operations *component-wise*.

- Addition: $\vec{a} + \vec{b} = (a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$
- Subtraction: $\vec{a} - \vec{b} = (a_1, a_2) - (b_1, b_2) = (a_1 - b_1, a_2 - b_2)$
- Multiplication *by scalar*: $n \times \vec{a} = (n \times a_1, n \times a_2)$

Vectors cannot be multiplied or divided by other vectors. It is important to understand what the elements of a vector stand for; in general vectors with differing numbers of elements may not be added.

For two points $A = (a_1, a_2)$ and $B = (b_1, b_2)$ the vector displacement $\vec{AB}$ is defined as the vector subtraction of A from B i.e.

$$\vec{AB} = (b_1 - a_1, b_2 - a_2)$$

If we take the origin as O, $\vec{AB} = \vec{OB} - \vec{OA}$.

Two vectors are parallel if through multiplication by some constant one can be transformed into the other. This is obvious when we consider that this is essentially the same as saying they point in the same direction, as a consequence of the fact that scalar multiplication doesn't change direction.

The length or *norm* of a vector can be determined through Pythagoras's Theorem, and is denoted $\|\vec{v}\|$:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2}$$

In higher dimensions, this extends simply through:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$$

Which essentially corresponds to the recursive application of Pythagoras in each dimension.

Some important properties of vector norms:

- For any $\vec{u}$, $\|\vec{u}\| = 0$ precisely if $\vec{u} = 0$.
- $\|\lambda\vec{u}\| = |\lambda|\|\vec{u}\|$ for any $\lambda \in \mathbb{R}$ and and vector $\vec{u}$.
- $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$ for any $\vec{u}, \vec{v}$.

Vectors of length 1 are called *unit vectors*. They are denoted by the "hat": $\hat{u}$. These are special, because they describe a direction. To find a unit vector for a given vector $\vec{u}$:

$$\hat{u} = \frac{1}{\|\vec{u}\|}\vec{u} = \frac{\vec{u}}{\|\vec{u}\|}$$

So for a piecewise example:

$$\vec{u} = (1, 2, -1)$$
$$\|\vec{u}\| = \sqrt{1^2 + 2^2 + (-1)^2} = \sqrt{6}$$
$$\hat{u} = (\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}})$$

There are standard unit vectors for the axis vectors up to $\mathbb{R}^3$:

- $\hat{i} = (1, 0, 0)$
- $\hat{j} = (0, 1, 0)$
- $\hat{k} = (0, 0, 1)$

These can be used to write any vector in $\mathbb{R}^3$.

The Dot Product

For two vectors $\vec{u}$ and $\vec{v}$ we take the scalar or dot product to be

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

This operation is commutative and distributive. Order does not matter, and the operation applied to a set of parenthesis will affect all operations within the parenthesis. In addition, scalar multiplication by either of the vectors before the operation is the same as scalar multiplication by the resultant vector. One other useful property is that:

$$\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$$

For $\vec{u}$ and $\vec{v}$, the *angle between* them is θ, which will always be $\leq \pi$. This angle is closely related to the dot product $\vec{u} \cdot \vec{v}$. Using the Law of Cosines ($c^2 = a^2 + b^2 - 2ab\cos(C)$), we can find that:

$$\|\vec{v} - \vec{u}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\|\vec{v}\|\cos(\theta)$$

Rearranging this, we can solve for $\cos(\theta)$

$$\cos(\theta) = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|}$$

$$\theta = \arccos\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|}\right)$$

Some information can be drawn from the dot product with using trigonemtric functions. Depending on the sign of the dot product we can determine the following:

- θ is acute if $\vec{u} \cdot \vec{v} > 0$
- θ is obtuse if $\vec{u} \cdot \vec{v} < 0$
- $\theta = \frac{\pi}{2}$ if $\vec{u} \cdot \vec{v} = 0$

Vector Projections

Vector projections describe the subsection of a vector which lies along another vector. A common example is the representation of vectors through $(a\hat{i}, b\hat{(j)}, c\hat{(k)})$. In this form, $a\hat{i}$ is the parallel projection of the vector onto $\hat{i}$.

If we have vectors $\vec{u}$ and $\vec{v}$ and we want to find the vector projection of $\vec{v}$ onto $\vec{u}$, essentially what we are trying to do is find some scalar k such that a third vector $\vec{w}$ is perpendicular (or *orthogonal*) to $\vec{u}$ when defined as:

$$\vec{w} = \vec{v} - k\vec{u}$$

Then, the parallel projection of $\vec{v}$ onto $\vec{u}$, denoted $\vec{v}_{\parallel\vec{u}}$, is $k\vec{u}$. The perpendicular projection of $\vec{v}$ onto $\vec{u}$, written as $\vec{v}_{\perp\vec{u}}$, is $\vec{w}$, or $\vec{v} - k\vec{u}$. Since vectors are perpendicular when their dot product is 0, this means we want the solution to:

$$\vec{u} \cdot (\vec{v} - k\vec{u}) = 0$$

After rearranging, this formula is:

$$k = \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$$

If the value of k is negative, the means the projetion is in the opposite direction to $\vec{u}$ and that the angle between $\vec{u}$ and $\vec{v}$ is obtuse. When k is 0, the angle is $\frac{\pi}{2}$ and If k is positive, the angle is acute.

Example

Find the closest point Q on a line L which passes through $A = (0, 1)$ and $B = (4, 3)$ to the point $P = (1, 3)$. How far away is this point from L.

- The closest point to a line will be perpendicular to the line.
- We can use a perpendicular projection to find this point.
- Taking our base vector to be $\vec{AP} = (1, 3) - (0, 1) = (1, 2)$
- And projecting on to $\vec{AB} = (4, 3) - (0, 1) = (4, 2)$
- Then $k = \vec{AB} \cdot \vec{AP} \div \vec{AB} \cdot \vec{AB} = \frac{8}{20} = \frac{2}{5}$
- Our perpendicular projection is given by $\vec{AP} - k\vec{AB}$
- $\vec{AP}_{\perp\vec{AB}} = (1, 2) - \frac{2}{5}(4, 2) = (1, 2) - (\frac{8}{5}, \frac{4}{5}) = (\frac{-3}{5}, \frac{6}{5})$
- The distance from Q to P is $\left\|\vec{AP}_{\perp\vec{AB}}\right\| = \frac{3}{\sqrt{5}}$

- To find Q, we can use $\vec{OQ} = \vec{OA} = \vec{AQ}$
- $\vec{OA} = A, \vec{AQ} = (\frac{8}{5}, \frac{4}{5}$ as calculated previously
- Thus, $\vec{OQ} = (\frac{8}{5}, \frac{9}{5})$

Parametric Curves

With the tool of vectors of the form $(x\hat{i}, y\hat{j})$, we gain access to a new construct of vectors which vary with time.

$$\vec{r}(t) = x(t)\hat{i} + y(t)\hat{j}$$

Here, $\vec{r}$ is a function which maps a real number t, usually time, to a vector. $x(t)$ and $y(t)$ are known as *parametric equations* as they depend on the parameter t. The curve which $\vec{r}$ traces out in $\mathbb{R}^2$ is known as a *parametric curve*. Functions of this form are useful for tasks such as describing the motion of a particle with time. The graph resultant from this, i.e. the set of points, is the range of the function.

Sometimes, it can be useful to find the equation of a path. For example, consider

$$\vec{r}(t) = t\hat{i} + t^2\hat{j}$$

For this equation we can solve for a function with the simultaneous equations

$$t = x$$

$$y = t^2 = x^2$$

Thus, the cartesian equal of this curve is $y = x^2$. In many cases it is difficult or impossible to find such an equation. The goal is essentially to eliminate t from the equation of one of the equations.

An application of parametric functions is in the modelling of ellipses. An ellipse of width a and height b has a cartesian equation of

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \ \ a, b > 0$$

This path can be parametrised through

$$\vec{r}(t) = a\cos(t)\hat{i} + b\sin(t)\hat{j}$$

This is proved by substituting $a\cos(t)$ for x and $b\sin(t)$ for y:

$$\left(\frac{a\cos(t)}{a}\right)^2 + \left(\frac{b\sin(t)}{b}\right)^2 = 1$$

$$\cos^2(t) + \sin^2(t) = 1$$

Which is simply a trigonemtric identity. Ellipses are fairly easy to analyse. Let us use an example to practice this:

$$\vec{r}(t) = \cos(2t)\hat{\imath} - 2\sin(2t)\hat{\jmath}$$

Here, we can immediately determine $a = 1$ and $b = 2$ from the coefficients of $\cos$ and $\sin$. We can tell that the direction is reversed from the $-$ in front of $\sin$. Finally, we can tell that the period is reduced from 2π to π by the $2t$ inner term.

A parametric equation is distinct from a path, because it contains more information. The equation can tell us where the particle is at any given time, it tells us which direction the particle is moving in at any given time, and it tells us how long the particle takes to traverse the path.

When we have a pair of parametric functions, it can be useful to determine when the two particles collide. To find these points, we can take the x and y coordinate of each function and set them equal to each other for an arbritary value of t. It is worth noting that the *paths* of two particles might intersect without the particles every colliding, as the two might pass through the same point at different times. As an example, let us consider the following functions:

$$\vec{r}_1(t) = (t+1)\hat{\imath} + (t^2 - 4t)\hat{\jmath} \text{ and } \vec{r}_2(t) = (2t)\hat{\imath} + (6t - 9)\hat{\jmath}$$

$$t + 1 = 2t \text{ and } t^2 - 4t = 6t - 9$$

$$t = 1 \text{ and } 1^2 - 4 \times 1 = 6 \times 1 - 9 = -3$$

Thus, the two particles collide at $t = 1$.

Differential Calculus

While it is difficult to model real world systems like, say, the population of rabbits in Australia, it serendipitously happens to be less difficult to write a function which is the *derivative* of such a quantity. This is the fundamental basis of differential equations, which are in turn fundamental to many modern scientific fields.

At a very basic level, a differentiable function is one for which we can find some function P' which can be used to solve the following:

$$P(t_0 + \Delta t) \approx P(t_0) + P'(t_0)\Delta t$$

i.e. we can find a value which tells us the rate of change of P across Δt. The goal of differential calculus is to find this value $P\prime$.

More precisely, we can define a derivative as a *limit*. If f is a function on an interval I, the derivative of f at any value $x \in I$ is the gradient of the tangent line to the graph of f at the point $(x, f(x))$, given by this limit

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

If and only if this limit exists is f *differentiable* at x. This in essence is stating that, for some very small value of h:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

$$f'(x)h \approx f(x+h) - f(x)$$

$$f'(x)h + f(x) \approx f(x+h)$$

This is the "fundamental formula of calculus". It tells us that we can determine the value of f at a point $x + h$ based on the derivative of f and the value of f at x. Essentially that a tangent line at this point provides a good approximation of the function for a very small interval.

Let us take a close look at the limit formula. If we notice that $f(x + h)$ is simply a point h units along from $f(x)$, then $f(x + h) - f(x)$ is the change in height of the function across the interval h, i.e. the rise. h, obviously is the distance across which this occurs, the run. Thus, this term must give us the average gradient over the interval $[x, x + h]$. When $h \approx 0$, it makes sense that this will be the instantaneous gradient at the point x. The tangent line to f at this point is the line with gradient given by this limit.

This lets us understand why a function may not be differentiable at a point. To differentiate, we need to find the gradient of a tangent line. If a tangent line cannot be drawn at a point, the function cannot be differentiated at that point. No value f' can be found. This can occur at vertical sections of a graph or at sections where a graph has a sharp point, for instance.

Standard Derivatives

The function f' is often written as

$$\frac{d}{dx}[f(x)]$$

This is perhaps more common than f' notation. For many common function structures, predetermined derivatives can be used rather than calculating on a case by case basis. Some of these common cases follow.

$$\frac{d}{dx}[x^a] = ax^{a-1}$$

$$\frac{d}{dx}[e^x] = e^x$$

$$\frac{d}{dx}[e^{f(x)}] = f'(x)e^{f(x)}$$

$$\frac{d}{dx}[\log(x)] = \frac{1}{x}$$

$$\frac{d}{dx}[\log(f(x))] = \frac{f'(x)}{f(x)}$$

$$\frac{d}{dx}[\sin(x)] = \cos(x)$$

$$\frac{d}{dx}[\cos(x)] = -\sin(x)$$

$$\frac{d}{dx}[\tan(x)] = \sec^2(x)$$

$$\frac{d}{dx}[\arcsin(x)] = \frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}[\arccos(x)] = \frac{-1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}[\arctan(x)] = \frac{1}{1+x^2}$$

Properties of Derivatives

Differentiation is a linear operator. What this means is that we can easily use derivatives for parts of a function to find a derivative for the whole of a function:

$$\frac{d}{dx}[cf(x)] = c\frac{d}{dx}[f(x)]$$

$$\frac{d}{dx}[f(x)+g(x)] = \frac{d}{dx}[f(x)] + \frac{d}{dx}[g(x)]$$

For two functions multiplied together, $f(x)g(x)$, we can find their derivative through the product rule.

$$\frac{d}{dx}[f(x)g(x)] = \frac{d}{dx}[f(x)]g(x) + f(x)\frac{d}{dx}[g(x)]$$

In the case of one function dividing another, say $\frac{f(x)}{g(x)}$, we use the quotient rule.

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{\frac{d}{dx}[f(x)]g(x) - f(x)\frac{d}{dx}[g(x)]}{(g(x))^2}$$

If a function is operating on a second function in a super-function to be differentiated, the structure is often too complex to solve with the previous rules. In these cases we use the chain rule.

$$\frac{d}{dx}[f(g(x))] = \frac{d}{dx}[f(x)](g(x))\frac{d}{dx}[g(x)]$$

These rules can be expressed simply by replacing the functions with single letter variables.

Rule	Representation
Product	$\frac{d}{dx}[uv] = vu' + uv'$
Quotient	$\frac{d}{dx}\left[\frac{u}{v}\right] = \frac{vu'-uv'}{v^2}$
Chain	$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$

Applications

If we know that a function has a positive derivative at some value, we know that the function is increasing at that point. The same is true for a negative derivative, though the function is decreasing. For an interval, if the derivative of a function is > 0 at every point in that interval, then the function is increasing on that interval. The inverse does not apply; if f is increasing on an interval, that does not imply that $f' > 0$ on that same interval.

This understanding of the implications of a positive or negative derivative suggest that points where the derivative $= 0$ have some significance. We call these points local maxima or local minima. Maxima occur where the function has a greater value at a point than at any other nearby point, minima for the inverse. These points are called *local* if the range of "nearby" is less than the domain of the function, and global if they are the highest or lowest points in the path of the function.

In addition to points where $f' = 0$, points where the gradient of a function is undefined may also be local minima or maxima. It is also worth noting that a stationary point is not necessarily a local extrema and that a local extreme is not necessarily a stationary point. To check if a local extrema is, we need to show that the points on either side have opposite sign. These points include the start and end points of a function.

It is worth noting that a derivative can change sign only when a function is undefined or when it passes through 0, which means we can find the extrema of a function by finding the sign of the function between each of the stationary points.

Higher Derivatives

We are not limited to differentiating functions once; we can take the derivative of these derivatives; what we call the *second* or *third* derivative of a function. This is very relevant in areas such as kinematics, where velocity is the first derivative of position and acceleration is the derivative of velocity, the second derivative of position. The second derivative is denoted f'' to the first derivative's f', and this trend continues to higher derivative with the alternate notation of $f^{(n)}$ for the nth derivative becoming useful for third and higher derivatives.

Just as with the first derivative, the second derivative has conditions to exist. For f'' to exist, f' must differentiable for the interval considered. If this is true, f is twice differentiable on the interval. We the second derivative as follows:

$$f''(x) = \frac{d}{dx}[f'(x)] = \frac{d^2}{dx^2}[f(x)]$$

Higher derivatives follow the same pattern. Some information can be drawn from a second derivative. The second derivative tells us the *concavity* of a function. A function is *concave up* on an interval if it displays a kind of cup shape; this occurs when f' is increasing on the interval or when f'' is positive for the interval. The function is concave down when it displays an inverted bowl shape; when f' is decreasing or f'' is less than 0.

Concavity is a geometric property; it describes the direction of curvature of the graph. We can understand it by thinking about chords; if a graph is concave up, such as a normal parabola, the chords will all be above the graph, while for a concave down graph they will all be below the graph. Let us consider an example:

- Use the second derivative to determine where $f(x) = x^4$ is concave up and where it is concave down.
- $f'(x) = 4x^3$
- $f''(x) = 12x^2$
- $12x^2 > 0$ for $x \in (-\infty, 0) \cup (0, \infty)$
- Thus $f(x)$ is concave up for $(-\infty, 0))$ and $(0, \infty)$.
- It is worth noting that this implies that the first derivative is increasing on $\mathbb{R}$, so $f(x)$ is actually concave up for all $\mathbb{R}$.

A *point of inflection* is a point where a function changes between being concave up and concave down. These can be found where $f'' = 0$. In addition, they can occur at points where the function is undefined, or at start or end points. For example, in the function $f(x) = x^3$, $f''(x) = 6x$ and we can see that the only point of inflection of this function is $x = 0$.

Asymptotes

Often when a graph has a gap in its domain, an *asymptote* exists at this location. An asymptote is an area of the domain where the graph becomes increasingly (arbitrarily) closer to a value without ever reaching it. Mathematically, this can be represented as

$$\lim_{x \to a} f(x) = \pm\infty$$

Where $x = a$ is the equation of the asymptote in f. A vertical asymptote exists at $y = a$ where

$$\lim_{x \to \pm\infty} f(x) = a$$

A third case exists, that of an *oblique asymptote*, which is an asymptote defined by a linear function of the form $y = mx + b$. In this case, we can say a limit exists if

$$\lim_{x \to \pm\infty} [f(x) - (mx + b)] = 0$$

i.e. the value of the function converges with that of the line. It is worth noting that an asymptote is not impassable; it simply states that the function can be approximated by the asymptote for appropriate values of x. These asymptotes can be found through polynomial long division. As an example, let us consider the following

$$f(x) = \frac{2x^3 - 3x^2 + 2x - 2}{x^2 + 1}$$

$$2x^3 - 3x^2 + 2x - 2 = (x^2 + 1)(2x - 3) + 1$$

$$f(x) = 2x - 3 + \frac{1}{x^2 + 1}$$

Thus, $y = 2x - 3$ is an oblique asymptote to f.

Vertical asymptotes are often encountered when dividing one function by another, as a function spits out smaller and smaller values as it approaches 0, resulting in values tending towards ∞, eventually reaching the case of division by 0 and an undefined value. They can also occur in functions like the logarithm. Some functions can have surprising behaviours.

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

Looking at this function, we would assume that it was undefined for -1 and 1, as we would have a 0 denominator. However, if we factorise it we find

$$f(x) = \frac{(x + 1)(x + 1)}{(x - 1)(x + 1)} = \frac{x + 1}{x - 1}$$

So the function is indeed defined for $x = -1$. This teaches us to check for common factors before crying asymptote.

Example

Using the tools for analysis of functions discussed thus far, we can analyse

$$f(x) = \frac{x^2}{x + 1}$$

- The implied domain of f is $\mathbb{R}\setminus\{-1\}$.
- Performing long division, we can find that

$$x^2 = (x+1)(x-1)+1 \Rightarrow f(x) = x - 1 + \frac{1}{x+1}$$

  This tells us that there is a vertical asymptote at $x = -1$, approaching $-\infty$ on the negative side and ∞ on the positive side.
- In addition, we can see that the line $y = x - 1$ is an oblique asymptote.
- The y intercept of the function is 0 and the x intercept of the function is also 0.
- Using the quotient rule we find that the derivative of the function is

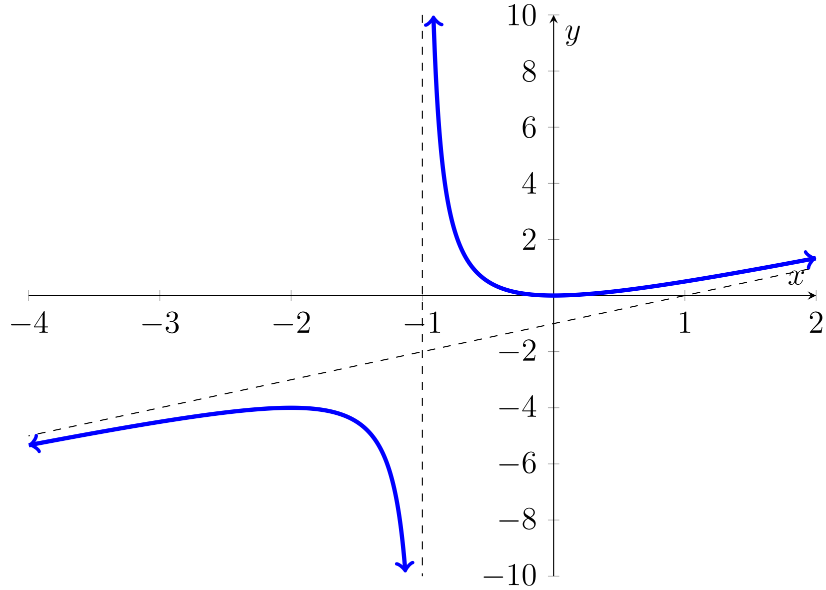
$$\frac{x(x+2)}{(x+1)^2}$$

- From this we can find the stationary points of the function at $f'(x) = 0$, $x = 0$ and $x = -2$.
- By looking at the sign of the derivative around these stationary points we can see that the function is increasing on $(-\infty, -2) \cup (0, \infty)$ and decreasing on $(-2, -1) \cup (-1, 0)$
- Through the same logic we can see that a local maximum exists at $x = -2$ and a local minimum exists at $x = 0$.
- Again using the quotient rule we can find the second derivative

$$f'' = \frac{2}{(x+1)^3}$$

- Looking at this function, we can tell there is only one stationary point at $x = -1$, and noticing that it is negative for $x < -1$ and positive for $x > -1$ we can say f is concave down on $(-\infty, 1)$ and concave up on $-1, \infty$.
- Because -1 is the only place where the sign of the second derivative changes, and this is known to be an undefined point, we can tell the function has no points of inflection.

We can use this information to plot the function.



Implicit Differentiation

In some cases, we might have a curve defined by some rule, but this rule might not express y in terms of x, it may instead be a function of both x and y or some other arcane mess of symbols. Nonetheless, it is evident that we should be able to find a rule which yields the slope of the tangent line to this curve at some point. In this case, we say there is an *implicit* dependence of y on x, and we find the derivative through *implicit differentiation*.

To do this we assume that near any given x, it is possible to write y as a function of x. For example:

$$x^2y = 1$$

$$\frac{d}{dx}[x^2y] = \frac{d}{dx}[1] = 0$$

$$\frac{d}{dx}[x^2]y + x^2\frac{d}{dx}[y] = 0$$

$$2xy + x^2\frac{d}{dx}[y] = 0$$

$$\frac{dy}{dx} = \frac{-2xy}{x^2} = \frac{-2y}{x} \ \ (x \neq 0)$$

This tells us that a tangent line to a point x, y on this curve has a slope given by this formula.

In general, implicit differentiation is performed as follows.

- For an equation involving x and y.
- Take the derivative of each side, with respect to x.
- Simplify each side as much as possible. Remember that y is assumed to be a function of x, so the product, chain and quotient rules must be appropriately applied.
- Rearrange to solve for $\frac{dy}{dx}$. Usually this solution will involve both x and y.

A more complex example:

$$x^4 + y^2 = 1 + x^2y$$

$$\frac{d}{dx}[x^4 + y^2] = \frac{d}{dx}[1 + x^2y]$$

$$4x^3 + 2y\frac{dy}{dx} = \frac{d}{dx}[x^2y]$$

$$4x^3 + 2y\frac{dy}{dx} = \frac{d}{dx}[x^2]y + x^2\frac{dy}{dx}$$

$$4x^3 + 2y\frac{dy}{dx} = 2xy + x^2\frac{dy}{dx}$$

$$(2y - x^2)\frac{dy}{dx} = 2xy - 4x^3$$

At this stage we run into a small problem; we want to divide through by $2y - x^2$, but if this is ever 0, our solution will be undefined. We can solve to see if this is ever the case:

$$2y - x^2 = 0 \Rightarrow y = \frac{1}{2}x^2$$

Subbing this into our original equation:

$$x^4 + \left(\frac{1}{2}x^2\right)^2 = 1 + x^2 \times \frac{1}{2}x^2$$

$$x^4 + \frac{1}{4}x^4 = 1 + \frac{1}{2}x^4$$

$$\frac{3}{4}x^4 = 1 \Rightarrow x^4 = \frac{4}{3} \Rightarrow x^2 = \pm\frac{2}{\sqrt{3}}$$

$$x = \pm\sqrt{\frac{2}{\sqrt{3}}}$$

So at these points, our derivative will be undefined, which tells us the curve is vertical at these points. With this caveat, we can continue solving.

$$\frac{dy}{dx} = \frac{2xy - 4x^3}{2y - x^2}$$

It is important to note that at points where both the numerator and denominator are 0, it is impossible to know what happens at this location without further analysis; the derivative could be horizontal, vertical or otherwise.

Derivatives of Inverse Functions

To find a derivative of $y = \log(x)$, we could rewrite as

$$e^y = x$$

$$\frac{d}{dx}[e^y] = \frac{d}{dx}[x]$$

$$\frac{d}{dx}[y]e^y = 1$$

$$\frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$$

Here, we have used implicit differentiation to find a derivative of $\log(x)$ based on the knowledge of the derivative of the exponential function. We can use the technique to find the derivatives of the inverse trigonemtric functions. For example, to find the derivative of $y = \arcsin(x)$

$$\sin(y) = x \quad x \in [-1, 1]$$

$$\frac{d}{dx}[sin(y)] = \frac{d}{dx}[x]$$

$$\frac{dy}{dx}\cos(y) = 1$$
$$\frac{dy}{dx} = \frac{1}{\cos(y)}$$

We can rewrite $\cos(y)$ in terms of x:

$$\sin^2(y) + \cos^2(y) = 1$$
$$\cos^2(y) = 1 - \sin^2(y)$$
$$\cos(y) = \sqrt{1 - \sin^2(y)}$$
$$\frac{d}{dx}[\arcsin(x)] = \frac{1}{\sqrt{1 - x^2}}$$

Using implicit differentiation we once again find that we can find a derivative for this function using knowledge of other derivatives.

Derivatives of Parametric Curves

Where for a function $\mathbb{R} \to \mathbb{R}$ our derivative supplies a linear approximation of the curve near some x value, for a parametric curve, we need to find a *velocity vector* which approximates the movement of a parametric function fora small time period. As opposed to

$$f(x_0 + h) \approx f(x_0) + f'(x_0) \times h$$

we have the vector equation

$$\vec{r}(t_0 + \Delta t) \approx \vec{r}(t_0) + \vec{v}\Delta t$$

Formally the definition of this vector $\vec{v}$ is

$$\vec{r}'(t) = \lim_{h \to 0} \frac{1}{h}(\vec{r}(t + h) - \vec{r}(t))$$

If and only if this limit exists is $\vec{r}$ differentiable at time t. We can understand this by examining it's components:

$$\frac{1}{h}(\vec{r}(t + h) - \vec{r}(t))$$

Here we have simply the vector displacement between the times t and $t+h$, divided by $\Delta t = h$. Thus, this is simply the instantaneous velocity at time t. This makes clear the relationship between velocity and $\vec{r}'$. To find this derivative, if the parametric function is of the form $\vec{r} = x(t)\hat{i} + y(t)\hat{j}$ then the derivative is simply

$$\vec{r}' = x'(t)\hat{i} + y'(t)\hat{j}$$

For example:

$$\vec{r}(t) = (t^2+1)\hat{i} + (\frac{1}{3}t^3+1)\hat{j}$$

$$\vec{r}'(t) = 2t\hat{i} - (3t^3+1)\hat{j}$$

Or:

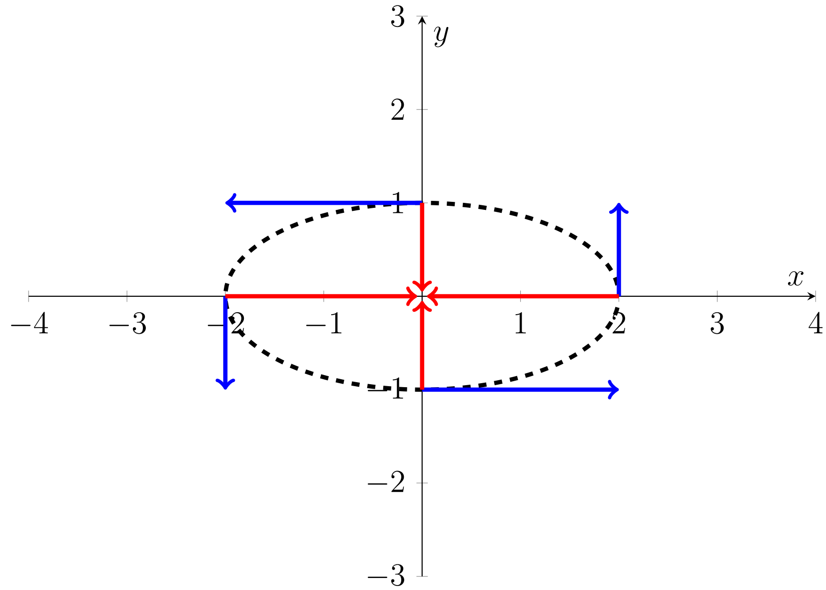
$$\vec{r}(t) = 2\cos(t)\hat{i} + \sin(t)\hat{j}$$

$$\vec{r}'(t) = -2\sin(t)\hat{i} + \cos(t)\hat{j}$$

Calculating the velocity vectors for a few values we find that:

- $\vec{r}'(0) = \hat{j}$
- $\vec{r}'(\frac{\pi}{2}) = -2\hat{i}$
- $\vec{r}'(\pi) = -\hat{j}$
- $\vec{r}'(\frac{3\pi}{2}) = 2\hat{i}$

And when plotting this graph, we find these vectors (in blue) look exactly how we would expect for the velocity of an object undergoing circular motion. The length of the resultant vectors is the instantaneous velocity of the particle at the time.



We can differentiate again to find

$$\vec{r}''(t) = -2\cos(t)\hat{i} - \sin(t)\hat{j}$$

Solving for the same values as earlier:

- $\vec{r}''(0) = -2\hat{i}$
- $\vec{r}''(\frac{\pi}{2}) = -\hat{j}$
- $\vec{r}''(\pi) = 2\hat{i}$
- $\vec{r}''(\frac{3\pi}{2}) = \hat{j}$

These vectors are marked in red on the graph; the acceleration is of course toward the center.

We can study the motions of objects through their derivatives. In the case that the derivative of a parametric function is 0, the velocity of the particle is 0. These points usually correspond to a sharp point on a curve, known as a *cusp*. To check, one can check the signs of the two components either side of the point; at least one must change sign for a cusp to exist at the point. There is no tangent line at the curve at this point, however it is still differentiable. If a curve has no cusps, it is smooth.

The angle between the acceleration and velocity vectors contains some information. When the angle is acute, the speed is increasing. When it is obtuse, the speed is decreasing. If the two are perpendicular, the speed is at a turning point. This is sensible; when the angle is acute, the acceleration is adding in the direction of the velocity, while if it is obtuse, it is working against the velocity. This can be stated through the dot product of the two vectors:

- If $\vec{r}'(t) \cdot \vec{r}''(t) < 0$, the speed is decreasing.
- If $\vec{r}'(t) \cdot \vec{r}''(t) = 0$, t is at a stationary point.
- If $\vec{r}'(t) \cdot \vec{r}''(t) > 0$, the speed is increasing.

The reason we have sharp points in differentiable parametric functions is because the plot of the curve is not a plot of the graph; the plot of the graph would be in $\mathbb{R}^3$, as the input is one dimensional and the output is two dimensional, much as the input to a single dimensional function needs a two dimensional graph in a plane.

An application of vector valued functions is projectile motion. For a projectile with initial velocity $a\hat{i}, b\hat{j}$ and launch height y_0 the particles position at a given time t is given by

$$\vec{r}(t) = at\hat{i} + (y_0 + bt - \frac{1}{2}gt^2)\hat{j}$$

Integral Calculus

The definite integral of a function f in the interval $[a, b]$ is denoted

$$\int_a^b f(x)\,\mathrm{d}x$$

It is defined to be the sum of the *signed areas* bounded by the curve $y = f(x)$, the x axis and the lines $x = a$ and $x = b$. Signed area here refers to the fact that areas which are under the graph and above the x axis are added to the area while those above the graph and below the axis are subtracted. a and b are the *terminals* of the integral, and together for the interval of integration.

Definite integrals share some of the properties of derivatives

$$\int_a^b f(x) + g(x)\,\mathrm{d}x = \int_a^b f(x)dx + \int_a^b g(x)\,\mathrm{d}x$$

$$\int_a^b kf(x) + g(x)\,\mathrm{d}x = k\int_a^b f(x)\,\mathrm{d}x$$

i.e. linearity. Integrals also have the property that for an interval $[a, a]$ the result will be 0. An integral of f on $[a, b]$ and an integral on $[b, c]$ will add to the same value as an integral on $[a, c]$.

Conceptually, integrals can be understood as being an approximation of the area under a graph through a method using vertical rectangles. These rectangles have width h and height given by $f(x)$ at their boundary (upper or lower depending on implementation). By taking a limit as h approaches 0, we can find the exact area under the curve. However, rather than doing this manually each time we can use antidifferentiation. For a function f, we call F an antiderivative of f if

$$F'(x) = f(x)$$

A property of this gives rise to the *Fundamental Theorem of Calculus*:

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) = [F(x)]_a^b$$

This tells us that the area under $f(x)$ is given by the value of an antiderivative of f at b minus the value of that antiderivative at a. We can use this to easily find the area under a curve.

$$\int_0^1 x^2\,\mathrm{d}x$$

$$\frac{d}{dx}[\frac{1}{3}x^3] = x^2$$

$$\int_0^1 x^2\,\mathrm{d}x = F(1) - F(0) = \frac{1}{3}1^3 - \frac{1}{3}0^3 = \frac{1}{3}$$

It is important to note that antiderivatives are not unique. For any F that is an antiderivative of f, there is an arbitrary number of possible variations. We thus introduce the concept of the indefinite integral.

$$\int f(x) = F(x) + C$$

Because the constant term C dissapears when differentiating, all possible values of C are valid.

If we have two curves and we want to find the area enclosed horizontally between them, all we need to do is to rearrange to find x as the dependent variable, and we can then integrate dy to find the area.

Standard Antiderivatives

$$\boxed{\begin{gathered}
\int x^a \,\mathrm{d}x = \tfrac{x^{a+1}}{a+1} + C \; (a \neq -1) \\
\int \tfrac{1}{x} \,\mathrm{d}x = \log(\,|x|\,) + C \\
\int e^x \,\mathrm{d}x = e^x + C \\
\int e^{ax} \,\mathrm{d}x = \tfrac{1}{a} e^{ax} + C \\
\int \sin(x) \,\mathrm{d}x = -\cos(x) + C \\
\int \cos(x) \,\mathrm{d}x = \sin(x) + C \\
\int \sec^2(x) \,\mathrm{d}x = \tan(x) + C \\
\int \tfrac{1}{\sqrt{s^2-x^2}} \,\mathrm{d}x = \arcsin(\tfrac{x}{s}) + C \\
\int \tfrac{-1}{\sqrt{s^2-x^2}} \,\mathrm{d}x = \arccos(\tfrac{x}{s}) + C \\
\int \tfrac{1}{s^2+x^2} \,\mathrm{d}x = \tfrac{1}{s} \arctan(\tfrac{x}{s}) + C
\end{gathered}}$$

For the inverse trigonometric functions, $s > 0$. To find the area between two curves, we can simply substitute $f(x) - g(x)$ into our integral.

Integration by Substitution

Integration has no general rules like the product or quotient rules. We can however use the process of integration by substitution in some situations. If we have a function $f(x)$ that we can't differentiate, but we can replace some sections of this function with a second function $h(x)$, then we can use integration by substitution. For example, we might replace x^4 with u. The *substitution rule* which makes this work is

$$\int g(u) \frac{du}{dx} \,\mathrm{d}x = \int g(u) du$$

So we then find an antiderivative in terms of u and then substitute back. For example

$$\int 2x(x^2-5)^4\,\mathrm{d}x$$

$$u = x^2 - 5 \Rightarrow g(u) = u^4$$

$$\frac{du}{dx} = 2x$$

$$\int 2x(x^2-5)^4\,\mathrm{d}x = \int u^4 \frac{du}{dx}\,\mathrm{d}x = \int u^4 du = \frac{1}{5}u^5$$

$$\frac{1}{5}(x^2-5)^5 + C$$

Note that here dx cancels with the denominator of $\frac{du}{dx}$. Another example:

$$\int cos(3x)\sqrt{\sin(3x+4)}\,\mathrm{d}x$$

$$u = \sin(3x) + 4 \Rightarrow g(u) = \sqrt{u}$$

$$\frac{du}{dx} = 3\cos(3x) \Rightarrow \frac{1}{3}\frac{du}{dx} = \cos(3x)$$

$$\int cos(3x)\sqrt{u}\,\mathrm{d}x = \int \sqrt{u}\frac{1}{3}\frac{du}{dx}\,\mathrm{d}x = \frac{1}{3}\int \sqrt{u}\frac{du}{dx}\,\mathrm{d}x = \frac{1}{3}\int u^{\frac{1}{2}}\,du$$

$$= \frac{1}{3}\frac{2}{3}u^{\frac{3}{2}} = \frac{2}{9}u^{\frac{3}{2}}$$

$$\Rightarrow \int cos(3x)\sqrt{\sin(3x+4)}\,\mathrm{d}x = \frac{2}{9}(\sin(3x)+4)^{\frac{3}{2}}$$

Integration by substitution is essentially a backwards application of the chain rule. To demonstrate, the chain rule gives us

$$\frac{d}{dx}\left[\sin^5(x)\right] = 5\sin^4(x)\cos(x)$$

Using the Fundamental Theorem

$$\int_a^b 5\sin^4(x)\cos(x)\,\mathrm{d}x = \int_a^b \frac{d}{dx}\left[\sin^5(x)\right]\,\mathrm{d}x = \sin^5(b) - \sin^5(a)$$

The key here is noticing that the integrand $5\sin^4(x)\cos(x)$ can be written as a product of a function on a function multiplied by a derivative, the form of output of the chain rule.

$$\frac{d}{dx}\big[G(h(x))\big] = g(h(x))h'(x)$$

Where $g(x) = 5x^4$ and $h(x) = \sin(x)$. We can integrate any function in this form, as long as we know an antiderivative of g (that is G above).

$$\int_a^b g(h(x))h'(x)\,\mathrm{d}x = \int_a^b G'(h(x))h'(x)\,dx = \int_a^b \frac{d}{dx}\big[G(h(x))\big]\,dx =$$

$$\big[G(h(x))\big]_a^b = G(h(b)) - G(h(a))$$

Here, an inverse chain rule is applied to transform the second integral into the third, by recognising that $G'(h(x))h'(x)$ is the output of a chain rule on $G(h(x))$. One interesting factor of this form is that

$$\big[G(h(x))\big]_a^b = \big[G(x)\big]_{h(a)}^{h(b)}$$

This knowledge lets us skip a few steps. Backtracking up the process we find

$$\big[G(x)\big]_{h(a)}^{h(b)} = \int_{h(a)}^{h(b)} G'(x)\,\mathrm{d}x = \int_{h(a)}^{h(b)} g(x)\,\mathrm{d}x$$

This gives us the final rule for integration by substitution for definite integrals.

$$\int_a^b g(h(x))h'(x)\,\mathrm{d}x = \int_{h(a)}^{h(b)} g(x)\,\mathrm{d}x$$

As an example:

$$\int_1^{e^{\frac{\pi}{2}}} \frac{\sin(\log(x))}{x}\,\mathrm{d}x$$

$$u = \log(x) \Rightarrow \frac{du}{dx} = \frac{1}{x}$$

$$\int_1^{e^{\frac{\pi}{2}}} \sin(u)\frac{du}{dx}\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \sin(u)\,du$$

Note that at this stage the bounds of integration have changed; we have moved from $[a, b]$ to $[\log(a), \log(b)]$, because in this case $\log$ was our h.

$$\int_0^{\frac{\pi}{2}} \sin(u)\,du = \big[-\cos(u)\big]_0^{\frac{\pi}{2}}$$

Linear Substitution

Not all functions are obviously transformed to use integration by substitution. However, sometimes we can use the trick of *linear substitution* to find a way to do so. The key to using integration by substitution is to rearrange the function into the form

$$\int g(u)\frac{du}{dx}\,dx$$

So one way we can do this is by finding a linear function $u(x)$ and rearranging to find x in terms of u. Example:

$$\int (2x+1)\sqrt{x-3}\,\mathrm{d}x$$

$$u = x-3 \Rightarrow x = u+3$$

$$\int (2(u+3)+1)\sqrt{u}\,\mathrm{d}x = \int (2u+7)\sqrt{u}\frac{du}{dx}\,\mathrm{d}x \left(\frac{du}{dx}=1\right)$$

$$= \int 2u^{\frac{3}{2}} + 7u^{\frac{1}{2}}\,du = 2\frac{1}{\frac{3}{2}+1}u^{\frac{3}{2}+1} + 7\frac{1}{\frac{1}{2}+1}u^{\frac{1}{2}+1} + C$$

$$= 2\frac{2}{5}u^{\frac{5}{2}} + 7\frac{2}{3}u^{\frac{3}{2}} + C$$

$$= \frac{4}{5}(x-3)^{\frac{5}{2}} + \frac{14}{3}(x-3)^{\frac{3}{2}} + C$$

Here, we used the fact that we could easily find x in terms of u to simplify the integral. This trick can also be used for high powers of $\sin$ and $\cos$.

$$\int \sin^4\cos^3\,\mathrm{d}x = \int u^4\cos^2(x)\cos(x)\,\mathrm{d}x = \int u^4\cos^2(x)\frac{du}{dx}\,\mathrm{d}x$$

$$\cos^2(x) + sin^2(x) = 1 \Rightarrow cos^2(x) = 1-\sin^2(x) = 1-u^2$$

$$\int u^4(1-u^2)\,du = \int u^4 - u^6 = \frac{1}{5}u^5 - \frac{1}{7}u^7 + C$$

$$\int \sin^4\cos^3\,\mathrm{d}x = \frac{1}{5}\sin^5(x) - \frac{1}{7}\sin^7(x) + C$$

This trick will work as long as the power of either $\sin$ or $\cos$ is odd for any similar function. If both are odd, either can be used. If both are even, the complex exponential can be used. Linear substitution happens to make one particular common form of substitution very simple. For a function of the form $f(kx)$ where $\int f(x)$ is already known, linear substitution tells us that $\int f(kx)$ is simply

$$\int f(kx)\,\mathrm{d}x = \frac{1}{k}F(kx) + C$$

A couple of examples of this

$$\int \sin(kx)\,\mathrm{d}x = -\frac{1}{k}\cos(kx) + C$$

$$\int e^{kx}\,dx = \frac{1}{k}e^{kx} + C$$

Integration of Rational Functions

Functions of the form

$$\frac{p(x)}{q(x)}$$

Turn out to be a very important structure to integrate, and a general process for this exists. We already know that

$$\int \frac{1}{x}\,dx = \log(|x|\,) + C$$

And now we extend this to examine integrals of the form

$$\int \frac{1}{ax+b}\,\mathrm{d}x$$

Here, we use the substitution $u = ax + b$.

$$u = ax + b \Rightarrow \frac{du}{dx} = a$$

So for an integral

$$\int \frac{1}{2x-3}\,\mathrm{d}x$$

$$u = 2x - 3 \Rightarrow \frac{du}{dx} = 2$$

We can multiply $\frac{du}{dx}$ by $\frac{1}{2}$ to simplify this process somewhat.

$$\int \frac{1}{u}\frac{1}{2}\frac{du}{dx}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{u}\,du = \frac{1}{2}\log(|u|\,) + C$$

$$\int \frac{1}{2x-3}\,\mathrm{d}x = \frac{1}{2}\log(\,|2x-3|\,) + C$$

For functions with an *irreducible quadratic denominator* of the form $q(x) = x^2 + bx + c$, we have a special technique. First, we complete the square of the denominator, to yield an equation of form

$$q(x) = u^2 + s^2$$

with $s > 0$. We can then perform a linear u substitution. This will generally split the integrand into a term which can be directly integrated via substitution, yield a $\log(q(x))$ term, and a second which can be integrated with the special antiderivative

$$\int \frac{1}{s^2+u^2}\,du = \frac{1}{s}\arctan(\frac{u}{s}) + C$$

In certain cases, one of these terms dissapears. To complete the square we use the following method.

$$r = c - \frac{b^2}{4a}$$

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + r = \left(\sqrt{a}x + \frac{b}{2\sqrt{a}}\right)^2 + r$$

Examples

Here we complete the square and then use u substitution to rearrange into the derivative of $\arctan$.

$$\int \frac{1}{x^2+2x+2}\,\mathrm{d}x \Rightarrow p(x) = x^2 + 2x + 2$$

$$x^2 + 2x + 2 = (x+1)^2 + 1$$

$$\int \frac{1}{1+(x+1)^2}\,\mathrm{d}x = \int \frac{1}{1+u^2}\frac{du}{dx}dx$$

$$= \int \frac{1}{1+u^2}\,du = \arctan(u) + C = \arctan(x+1) + C$$

In this second example, substitution is used twice to rearrange into a simpler form.

$$\int \frac{x+1}{x^2+2x+2}\,\mathrm{d}x = \int \frac{x+1}{(x+1)^2+1}\,\mathrm{d}x = \int \frac{u}{u^2+1}\,\frac{du}{dx}dx$$

$$u = (x+1) \Rightarrow \frac{du}{dx} = 1$$

$$v = u^2 + 1 \Rightarrow \frac{dv}{du} = 2u \Rightarrow \frac{1}{2}\frac{dv}{du} = u$$

$$\int \frac{u}{u^2+1}\, du = \frac{1}{2}\int \frac{1}{v}\,\frac{dv}{du} du = \frac{1}{2}\int \frac{1}{v}\, dv$$

$$\Rightarrow \int \frac{x+1}{x^2+2x+2}\, \mathrm{d}x = \int \frac{u}{u^2+1}\, du = \frac{1}{2}\int \frac{1}{v}\, dv$$

$$= \frac{1}{2}\log(|v|) + C = \frac{1}{2}\log(u^2+1) + C = \frac{1}{2}\log((x+1)^2+1) + C$$

Note that the $|v|$ becomes u^2+1 because this term is always positive, so the absolute value is the same. Finally, a less ideal example; here, the linear equation on top doesn't appear in the quadratic factorisation.

$$\int \frac{x-2}{x^2+2x+2}\, \mathrm{d}x = \int \frac{x-2}{(x+1)^2+1}\, \mathrm{d}x$$

$$u = x+1 \Rightarrow x = u-1 \Rightarrow \frac{du}{dx} = 1$$

$$\int \frac{(u-1)-1}{u^2+1}\frac{du}{dx} dx = \int \frac{u-3}{u^2+1}\, du = \int \left(\frac{u}{u^2+1} - \frac{3}{u^2+1}\right)\, du$$

$$v = u^2+1 \Rightarrow \frac{1}{2}\frac{dv}{du} = u$$

$$\int \frac{1}{v}\frac{1}{2}\frac{dv}{du} du - 3\arctan(u) + C = \frac{1}{2}\log(|v|) - 3\arctan(u) + C$$

$$= \frac{1}{2}\log(u^2+1) - 3\arctan(u) + C = \frac{1}{2}\log((x+1)^2+1) - 3\arctan(x+1) + C$$

The key to this approach is the ability to express a more complex fraction as an equation of order 1 or less over an equation of higher order. To do this, we can use a variety of tools such as long division or partial fractions.

Partial Fractions

To break a rational function down into partial fractions, one must first use polynomial long division to break a function down into terms where the order of the numerator is less than that of the denominator. Then, the denominator must be factorised into linear and irreducible quadratic factors. Only then can it be broken into partial fractions, and only if its factors are distinct. A simple example

$$f(x) = \frac{9x+1}{(x-3)(x+1)} = \frac{A_i}{x-3} + \frac{A_2}{x+1}$$

$$9x+1 = \frac{A_1}{x-3}(x-3)(x+1) + \frac{A_2}{x+1}(x-3)(x+1)$$

$$9x+1 = A_1(x+1) + A_2(x-3) = (A_1+A_2)x + (A_1 - 3A_2)$$

$$9 = A_1 + A_2 \quad 1 = A_1 - 3A_2$$

$$A_2 = 2, A_1 = 7$$

This is a general method for finding values of A, one can also choose values of x and substitute them in to find the appropriate values.

$$\int \frac{9x+1}{(x-3)(x+1)}\,\mathrm{d}x = \int \frac{7}{x-3}\,\mathrm{d}x + \int \frac{2}{x+1}\,\mathrm{d}x$$

$$= 7\log(|x-3|) + 2\log(|x+1|) + C$$

Note that the reason constants A are found is because the factors of the denominator are linear. For quadratic factors, the relevant numerators will be of the form $Bx+C$. In the case of non-distinct factors of the denominator, a special case applies. If $(x-a)^n$ appears, the related terms in the result will be of the form

$$\frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \ldots + \frac{A_n}{(x-a)^n}$$

The general algorithm for integration by this method is as follows.

- Perform long division to find a numerator of lower degree than the denominator.
- Factorise the denominator into linear and irreducible quadratic factors.
- Use partial fractions to break the more complex terms into simple terms.

- Integrate the linear terms as logarithms and the irreducible quadratic terms as $\arctan$ derivatives.

Yet another example.

$$\int \frac{2x^3-3x^2-8x+24}{x^2-4}\,\mathrm{d}x = \int \frac{(2x-3)(x^2-4)+12}{x^2-4}\,\mathrm{d}x$$

$$= \int 2x-3\,\mathrm{d}x + \int \frac{12}{x^2-4}\,\mathrm{d}x = x^2-3x+12\int \frac{1}{x^2-4}\,\mathrm{d}x$$

$$x^2-4=(x-2)(x+2) \Rightarrow \frac{1}{(x-2)(x+2)} = \frac{A_1}{x-2}+\frac{A_2}{x+2}$$

$$\Rightarrow 1 = A_1(x+2)+A_2(x-2)$$

$$x=2 \Rightarrow 1 = 4A_1 \Rightarrow A_1 = \frac{1}{4} \Rightarrow A_2 = -\frac{1}{4}$$

$$12\int \frac{1}{x^2-4} = 12\int \frac{1}{4(x-2)} - \frac{1}{4(x+2)} = 3\int \frac{1}{x-2}\,\mathrm{d}x - 3\int \frac{1}{x+2}\,\mathrm{d}x$$

$$\int \frac{2x^3-3x^2-8x+24}{x^2-4}\,\mathrm{d}x = x^2-3x+3\log(|x-2|)-3\log(|x+2|)+C$$

Differential Equations

Differential equations are a key tool in many fields of science for mathematically modelling phenomena. The reason for this is that it happens to be much easier to describe phenomena in terms of their second derivative, such as acceleration, than it is to directly describe their state. It is therefore very useful to be able to move from a second derivative to an equation describing the state. In a way, differential equations are to algebra as functions are to numbers; where we aim to find a constant in algebra, we aim to find a function with a differential equation.

A differential equation, in a mathematical sense, is an equation involving a variable x, and unknown function y and the derivatives of y with respect to x. This is an *ordinary* differential equation. The order of the highest derivative of in the equation is called the order of the differential equation. Some examples of differential equations:

- $\frac{dy}{dx} = -5y$

- $\frac{d^3y}{dt^3} + t\frac{dy}{dt} + (t^4 - 1)y = \sin(t)$
- $e^x f'(x) + f''(x) = f^3(x) + x$

For an algrebraic equation, we can check if a value is a solution by substitution, simply checking if a value of x satisfies the equality. Likewise to check a solution for a differential equation, we can substitute in a function and see if the statement satisfies function equality. For example

$$y + \cos^2(x) = x^2 + 1$$

- Is $y = \sin^2(x) + x^2$ a solution?

$$\sin^2(x) + cos^2(x) = 1 \Rightarrow \sin^2(x) + x^2 + \cos^2(x) = x^2 + 1$$

- Is $y = x^2$ a solution?

$$x^2 + \cos^2(x) = x^2 + 1 \Rightarrow \cos^2(x) = 1$$

$$\cos^2(\frac{\pi}{2}) = 0 \neq 1$$

Let us now try this in the case of a differential equation, by checking whether $y = e^{3x}$ is a solution to

$$\frac{d^2y}{dx^2} = 15y - 2\frac{dy}{dx}$$

$$\frac{d^2}{dx^2}\left[e^{3x}\right] = \frac{d}{dx}\left[3e^{3x}\right] = 9e^{3x}$$

$$15e^{3x} - 2(3e^{3x}) = 9e^{3x}$$

Thus, both sides are equivalent functions and y is a solution to the differential equation. Note that while in this case, both sides were identical, that need not always be the case. As long as the functions are equivalent, a solution has been found.

When solving a differential equation, there are usually infinitely many solutions. The reason for this is much the same as the reason for the $+C$ when solving an integral, and indeed an integral is simply a special case of a differential equation, as can been seen through the following differential equation.

$$\frac{dy}{dx} = \cos(x)$$

$$\int \frac{dy}{dx}\,\mathrm{d}x = \int \cos(x)\,\mathrm{d}x \Rightarrow y = sin(x) + C$$

The general solution for a differential equation is a formula for all of the possible solutions to that differential equation. For an equation involving an order of one, a single integration can yield a solution, with a single constant of integration in the solution. For an order two equation, there will be more constants of integration and so on. If we fix the values of these constants of integration, we have a particular solution of the differential equation.

$$\frac{d^2y}{dx^2} = \cos(x) \Rightarrow \int \frac{d}{dx}\left[\frac{dy}{dx}\right]\,\mathrm{d}x = \int \cos(x)\,\mathrm{d}x = \sin(x) + C$$

$$\frac{dy}{dx} = \sin(x) + C \Rightarrow \int \frac{dy}{dx}\,\mathrm{d}x = \int \sin(x) + C\,\mathrm{d}x$$

$$y = \int \sin(x)\,\mathrm{d}x + \int C\,\mathrm{d}x = -\cos(x) + Cx + D$$

This tells us that all of the equations which satisfy the differential equation are of the form $-\cos(x) + Cx + D$ with varying values of C and D. To find a particular solution, we usually need additional information, such as the solution of y at a particular x, which can be used to solve for constants. Solving for these constants is known as solving an *initial value problem*. More values are required for higher order differential equations. An example:

$$f''(x) = x, f(1) = 2, f'(0) = 1$$

$$\int f''(x)\,\mathrm{d}x = \int x\,\mathrm{d}x$$

$$\int f'(x)\,\mathrm{d}x = \int \frac{1}{2}x^2 + C\,\mathrm{d}x$$

$$f(x) = \frac{1}{6}x^3 + Cx + D$$

$$f(1) = 2 \Rightarrow 2 = \frac{1}{6} \times 1^3 + C \times 1 + D \Rightarrow 11 = 6C + 6D$$

$$f'(0) = 1 \Rightarrow 1 = \frac{1}{2} \times 0^2 + C \Rightarrow C = 1$$

$$11 = 6 \times 1 + 6D \Rightarrow D = \frac{5}{6}$$

$$f(x) = \frac{1}{6}x^3 + x + \frac{5}{6}$$

Seperable Differential Equations

A special case of differential equations is *seperable differential equations*, which are of the form

$$\frac{dy}{dx} = F(x)G(y)$$

Where F and G are known functions. An example of this form is found in some of the simpler equations solved in previous examples

$$\frac{dy}{dx} = F(x)$$

Another important form, known as *autonomous* differential equations is

$$\frac{dy}{dx} = G(y)$$

A simple example:

$$\frac{dy}{dx} = y, y > 0$$

$$F(x) = 1, G(y) = y \Rightarrow \frac{1}{y}\frac{dy}{dx} = 1$$

$$\int \frac{1}{y}\frac{dy}{dx}\,\mathrm{d}x = \int 1\,\mathrm{d}x = x + C \Rightarrow \int \frac{1}{y}\,\mathrm{d}y = x + C$$

$$\log(y) = x + C \Rightarrow e^{\log(y)} = e^{x+C} \Rightarrow y = e^C e^x \Rightarrow$$

$$y = Ae^x, A > 0$$

Here, integration by substitution was used to deal with the $\frac{dy}{dx}$ term. The fact that e^C could take any positive value was used to replace e^C with A. The general method for solving seperable differential equations is

- First divide both sides by $G(y)$, on the proviso that $G(y) \neq 0$
- Then, integrate both sides and evaluate the integrals.
- If possible, solve the resultant function to obtain a formula for y.
- The case $G(y) = 0$ must be handled seperately.

For a given differential equation with a known (x, y) point there will be a single unique solution which passes through that point. This extends generally to tell us that no two different solutions may meet at any point.

Constant Solutions

A differential equation might have a constant solution, where a single value satisfies the equation. As an example we can consider the equation

$$\frac{dy}{dx} = -6xy^2, y(0) = 0$$

$$G(y) = y^2 \Rightarrow y = 0 \text{ is a constant solution}$$

As we discovered that a solution can never cross another solution, this constant solution tells us that another solution can never touch the line $y = 0$. For the differential equation

$$\frac{dy}{dx} = y - \frac{y^2}{4}$$

Has constant solutions of $y = 0$ and $y = 4$, which tells us that all solutions must be one of $y > 4$, $4 > y > 0$ or $y < 0$. This tool can allow us to use the technique for seperable differential equations if a constant solution of $y = 0$ exists, because we know y can never be 0 in any other solution.

Example

$$\frac{dy}{dx} = y - \frac{y^2}{3}, y(0) = 1$$

$$y \equiv 0, y \equiv 4$$

$$\frac{dy}{dx} = y(1 - \frac{y}{4}) \Rightarrow \frac{1}{y(1-\frac{y}{4})}\frac{dy}{dx} = 1$$

$$\int \frac{1}{y(1-\frac{y}{4})}\frac{dy}{dx}\,\mathrm{d}x = x + C \Rightarrow \int \frac{1}{y(1-\frac{y}{4})}\,\mathrm{d}y = x + C$$

We first use integration by substitution to rearrange into a more workable form, through the knowledge that this is a seperable differential equation.

$$\int \frac{1}{y(1-\frac{y}{4})}\,\mathrm{d}y = \int \frac{4}{y(4-y)}\,\mathrm{d}y = 4\int \frac{1}{y(4-y)}\,\mathrm{d}y = -4\int \frac{1}{y(y-4)}\,\mathrm{d}y$$

After rearranging and simplifying we can easily use partial fractions to integrate this integral.

$$\frac{1}{y(y-4)} = \frac{A_1}{y} + \frac{A_2}{y-4} \Rightarrow 1 = A_1(y-4) + A_2 y$$

$$y = 0 \Rightarrow 1 = -4A_1 \Rightarrow -\frac{1}{4} = A_1$$
$$y = 4 \Rightarrow 1 = 4A_2 \Rightarrow \frac{1}{4} = A_2$$
$$\frac{1}{y(y-4)} = \frac{-1\frac{1}{4}}{y} + \frac{\frac{1}{4}}{y-4}$$
$$-4\int \frac{-1\frac{1}{4}}{y} + \frac{\frac{1}{4}}{y-4} = \int \frac{1}{y}\,\mathrm{d}y - \int \frac{1}{y-4}\,\mathrm{d}y = \log(|y|) - \log(|y-4|) =$$

The knowledge that the difference of logarithms is a ratio allows us to write this as

$$\log\left(\left|\frac{y}{y-4}\right|\right) = x + C \Rightarrow \left|\frac{y}{y-4}\right| = Ae^x$$

Because we know that for this function as y is 1 at $x = 0$, the function must always be between the two constant functions, $y \equiv 0$ and $y \equiv 4$, thus we can assume $0 < y < 4$.

$$\frac{y}{y-4} < 0 \Rightarrow \left|\frac{y}{y-4}\right| = \frac{-y}{y-4}$$
$$\frac{-y}{y-4} = Ae^x \Rightarrow -y = Ae^x(y-4) = Ae^x y - 4Ae^x$$
$$y = -Ae^x y + 4Ae^x \Rightarrow y + Ae^x y = 4Ae^x \Rightarrow y(1 + Ae^x) = 4Ae^x \Rightarrow$$
$$y = \frac{4Ae^x}{1 + Ae^x}$$

This form has a known simplification. By dividing through by Ae^x we find

$$y = \frac{4}{\frac{1}{A}e^{-x} + 1}$$

Solving our initial value problem of $y(0) = 1$

$$1 = \frac{4}{\frac{1}{A} + 1} \Rightarrow \frac{1}{A} + 1 = 4 \Rightarrow \frac{1}{A} = 3 \Rightarrow A = \frac{1}{3}$$

$$y = \frac{4}{3e^{-x} + 1}$$

Population Model

A simple application of differential equations is an exponential population growth model. This model begins with an assumption that population varies due to births and deaths. This is the classic process of constructing differential equation; begin with some reasonable assumptions, describe them with equations and solve these equations to understand the model better. For our population growth model, we might say that for a change in time Δt a change in population is given by

$$P(t_0 + \Delta t) = P(t_0) + \text{births} - \text{deaths}$$

$$\text{births} = (\text{probability of giving birth})P(t_0) = bP(t_0)$$

$$\text{deaths} = (\text{probability of dying})P(t_0)) = dP(t_0)$$

$$P(t_0 + \Delta t) = P(t_0) + bP(t_0) - dP(t_0) = P(t_0)\Delta t(b - d) \Rightarrow$$

$$\frac{P(t_0 + \Delta t) - P(t_0)}{\Delta t} = (b - d)P(t_0)$$

$$\lim_{\Delta t \to 0} \frac{P(t_0 + \Delta t) - P(t_0)}{\Delta t} = P'(t_0) \Rightarrow$$

What we have found is that the rate of change of the population is independent of the time. As we used a general t_0, if our assumptions were appropriate, the population must be modelled by the differential equation

$$\frac{dP}{dt} = (b - d)P(t_0)$$

As an example, let us consider a population of mice where $b - d = 0.2$, with an initial population of $M = 50$ mice.

$$\frac{dM}{dt} = 0.2M$$

$$\int \frac{1}{M}\frac{dM}{dt}\,\mathrm{dt} = \int 0.2\,\mathrm{dt}$$

$$\log(|M|) + = 0.2t + C \Rightarrow M = Ae^{0.2t}$$

$$50 = Ae^{0.2\times 0} \Rightarrow A = 50$$

$$M = 50e^{0.2t}$$

If we wanted to find at what time the population reaches 500, we can simply perform a substitution.

$$500 = 50e^{0.2t} \Rightarrow 10 = e^{0.2t} \Rightarrow \log(10) = 0.2t \Rightarrow t = \frac{\log(10)}{0.2}$$

It is worth noting that this model indicates that the population of mice would outweigh the earth in around six years. This is an important point to observe; a model is only valid as long as it properly represents all of the relevant variables. This model might work for small populations and short times, but it has the issue of ignoring competition and other relevant variables.

To rectify this, we can make some modifications to our model. Let us replace $b - d$ with b, which we will change from a simple probability of giving birth per unit time to a more complex factor, dependent on P. The simplest function we can use for this is a linear function of P, with a maximum birth rate at $P = 0$ and a x-intercept at C, the birth rate capacity of the environment. This gives us a linear function of

$$b = b_0 - \frac{b_0}{C}P$$

$$P(t_0 + \Delta t) = P(t_0) + (b_0 - \frac{b_0}{C}P(t_0))\Delta t P(T_0))$$

$$\frac{P(t_0 + \Delta t) - P(t_0)}{\Delta t} = (b_0 - \frac{b_0}{C}P(t_0))$$

$$\lim_{\Delta t \to 0} \frac{P(t_0 + \Delta t) - P(t_0)}{\Delta t} = \frac{dP}{dt}$$

$$\frac{dP}{dt} = (b_0 - \frac{b_0}{C}P)P = b_0 P(1 - \frac{1}{C}P)$$

This is known as the logistic differential equation, and refines the exponential growth model to incorporate constraints in a minimal way. As long as $0 < P < C$, $\frac{dP}{dt} > 0$. As $P \to C$, the growth rate drops off to 0. For a P small relative to C, this equation is roughly equivalent to the exponential example. Let us take a b_0 of 1 and attempt to solve, with an initial population of 1 and carrying capacity of 4.

$$\frac{dP}{dt} = P(1 - \frac{P}{4}) \Rightarrow \frac{1}{P(1 - \frac{P}{4})}\frac{dP}{dt} = 1$$

$$1 - \frac{P}{4} = -\frac{1}{4}(P-4) \Rightarrow \frac{1}{P(1-\frac{P}{4})} = \frac{1}{P(-\frac{1}{4}(P-4))}$$
$$\int \frac{1}{P(-\frac{1}{4}(P-4))}\,\mathrm{d}p = \int \frac{-4}{P(P-4)}\,\mathrm{d}p = t + C$$
Using partial fractions we can rearrange to find that
$$\frac{-4}{P(P-4)} = \frac{1}{P} - \frac{1}{P-4}$$
$$\int \frac{1}{P} - \int \frac{1}{P-4}\,\mathrm{d}p = t + C$$
$$\log(|P|) - \log(|P-4|) = t + C = \log\left(\left|\frac{P}{P-4}\right|\right)$$
$$\left|\frac{P}{P-4}\right| = e^C e^t = Ae^t$$
Because the solutions to this problem must be between 0 and 4, as these are constant solutions to this equation (our minimum and maximum population), we know that P is always positive and $P-4$ is always negative, meaning
$$\left|\frac{P}{P-4}\right| = -\frac{P}{P-4} = Ae^t$$
$$-P = (P-4)Ae^t = PAe^t - 4Ae^t \Rightarrow 4Ae^t = PAe^t + P = P(Ae^t + 1)$$
$$P = \frac{4Ae^t}{1 + Ae^t}$$
$$P(0) = 1 \Rightarrow 1 = \frac{4Ae^0}{1 + Ae^0} = \frac{4A}{1 + A}$$
$$1 + A = 4A \Rightarrow A = \frac{1}{3}$$
We have found more or less the general form for a logistic differential equation.
$$\frac{dP}{dt} = b_0 P(1 - \frac{1}{C}P)$$
$$P = \frac{C}{1 + Ae^{-b_0 t}}$$
With carrying capactiy C, birth rate b_0 and a constant A yielded by the initial condition problem. For this model, there is a point of inflection at $P = \frac{C}{2}$. If $P(0) > C$, there will be no point of inflection and the population will asymptotically approach C from above.

Newton's Law of Cooling

The rate at which an object cools (or heats) is proportional to the difference between its temperature T and the (taken to be constant) temperature T_s of its surroundings. For a time t minutes:

$$\frac{dT}{dt} = -k(T - T_s)$$

As an example, let us consider a loaf of bread in a freezer of $-15°\mathrm{C}$. The bread is initially $20°\mathrm{C}$ and takes 20 minutes to reach $10°\mathrm{C}$. How long will the bread take to reach $0°\mathrm{C}$?

$$\frac{dT}{dt} = -k(T - (-15)) = -k(T + 15)$$

$$\frac{1}{T+15}\frac{dT}{dt} = -k \Rightarrow \int \frac{1}{T+15}\frac{dT}{dt}\,\mathrm{d}t = \int -k\,\mathrm{d}t$$

$$\log(|T+15|) = -kt + C \Rightarrow |T+15| = e^{-kt}e^{D}$$

Assuming $T > -15$, as we are looking at the case of a cooling object,

$$T + 15 > 0 \Rightarrow |T+15| = T + 15$$

$$T = Ae^{-kt} - 15$$

$$T(0) = 20 \Rightarrow 20 = Ae^{-0k} - 15 \Rightarrow 35 = A$$

$$T(20) = 10 \Rightarrow 10 = 35e^{-20k} - 15 \Rightarrow 25 = 35e^{-20k}$$

$$\frac{25}{35} = e^{-20k} \Rightarrow \log\left(\frac{25}{35}\right) = -20k$$

$$\frac{\log\left(\frac{25}{35}\right)}{-20} = k \approx 0.0168 \Rightarrow T = 35e^{-0.0168t} - 15$$

$$0 = 35e^{-0.0168t} - 15 \Rightarrow \frac{15}{35} = e^{-0.0168t} \Rightarrow \log\left(\frac{15}{35}\right) = -0.0168t$$

$$\frac{\log\left(\frac{15}{35}\right)}{-0.0168} = t \approx 50.434$$

The general form of Newton's law of cooling is therefore

$$T = T_s + Ae^{-kt}$$

General Process

The general process of generating and using differential equations is as follows.

- First, one makes observations of a phenomena, using these observations to generate a hypothesis of some behaviour of the system.
- This hypothesis is then encoded in a differential equation, which is analysed to find either a solution, or to identify some behaviours of the system.
- If a solution can't be found, one can instead find an approximation of the equation to approximate the state of the system.