Essential Engineering Mathematics

Michael Batty

bookboon
The eBook company

Michael Batty

Essential Engineering Mathematics

Essential Engineering Mathematics

ISBN 978-87-7681-735-0

Contents

Introduction

This book is partly based on lectures I gave at NUI Galway and Trinity College Dublin between 1998 and 2000. It is by no means a comprehensive guide to all the mathematics an engineer might encounter during the course of his or her degree. The aim is more to highlight and explain some areas commonly found difficult, such as calculus, and to ease the transition from school level to university level mathematics, where sometimes the subject matter is similar, but the emphasis is usually different. The early sections on functions and single variable calculus are in this spirit. The later sections on multivariate calculus, differential equations and complex functions are more typically found on a first or second year undergraduate course, depending upon the university. The necessary linear algebra for multivariate calculus is also outlined. More advanced topics which have been omitted, but which you will certainly come across, are partial differential equations, Fourier transforms and Laplace transforms.

This short text aims to be somewhere first to look to refresh your algebraic techniques and remind you of some of the principles behind them. I have had to omit many topics and it is unlikely that it will cover everything in your course. I have tried to make it as clean and uncomplicated as possible.

Hopefully there are not too many mistakes in it, but if you find any, have suggestions to improve the book or feel that I have not covered something which should be included please send an email to me at

batty.mathmo@googlemail.com

Michael Batty, Durham, 2014.

Chapter 1

Preliminaries

1.1 Number Systems: The Integers, Rationals and Reals

Calculus is a part of the mathematics of the *real* numbers. You will probably be used to the idea of real numbers, as numbers on a "line" and working with graphs of real functions in the product of two lines, i.e. a plane. To define rigorously what real numbers are is not a trivial matter. Here we will mention two important properties:

- The reals are *ordered*. That is, we can always say, for definite, whether or not one real number is greater than, smaller than, or equal to another. An example of the properties that the ordering satisfies is that if $x < y$ and $z > 0$ then $zx < zy$ but if $x < y$ and $z < 0$ then $zx > zy$. This is important for solving inequalities.

- A real number is called *rational* if it can be written as $\frac{p}{q}$ for integers p and q ($q \neq 0$). The reals, not the rationals, are the usual system in which to speak of concepts such as *limit* and *continuity* of functions, and also notions such as derivatives and integrals. This is because they have a property called *completeness* which means that if a sequence of real numbers looks like it has a limit (i.e. the distance between successive terms can always be made to be smaller than a given positive real number after a certain point) then it does have a limit. The rationals do not have this property.

The real numbers are denoted by $\mathbb{R}$ and the rational numbers are denoted by $\mathbb{Q}$. We also use the notation $\mathbb{N}$ for the set of *natural numbers* $\{1, 2, 3, \ldots\}$ and $\mathbb{Z}$ for the set of *integers* $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$. For any positive integer n, $\sqrt{n}$ is either an integer, or it is not rational. That is, it is an *irrational* number.

We will deal with two other number systems of numbers which depend on the reals: vectors of real numbers, and complex numbers.

1.2 Working With the Real Numbers

In this section we will introduce notation that is used throughout the book and explain some basics of using the real numbers.

1.2.1 Intervals

An *interval* is a "connected" subset of $\mathbb{R}$ and can be *bounded* i.e. of the form

- $[x, y] = \{r \text{ in } \mathbb{R} \mid x \leqslant r \leqslant y\}$ (closed),
- $(x, y) = \{r \text{ in } \mathbb{R} \mid x < r < y\}$ (open),
- $[x, y) = \{r \text{ in } \mathbb{R} \mid x \leqslant r < y\}$ (half-open) or
- $(x, y] = \{r \text{ in } \mathbb{R} \mid x < r \leqslant y\}$ (half-open).

or *unbounded* i.e. of the form

- $(-\infty, x] = \{r \text{ in } \mathbb{R} \mid r \leqslant x\}$,
- $(-\infty, x) = \{r \text{ in } \mathbb{R} \mid r < x\}$,
- $[x, \infty) = \{r \text{ in } \mathbb{R} \mid r \geqslant x\}$,
- $(x, \infty) = \{r \text{ in } \mathbb{R} \mid r > x\}$ or
- $(-\infty, \infty) = \mathbb{R}$.

1.2.2 Solving Inequalities

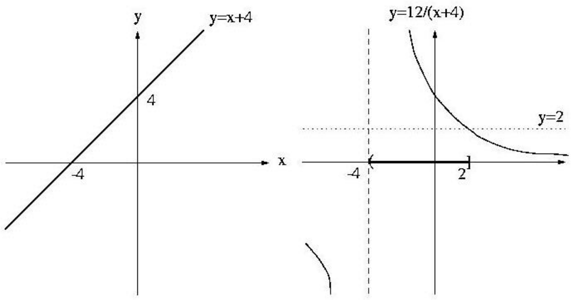
When solving inequalities, as opposed to equations, if you multiply an inequality by a negative number then it reverses the direction of the inequality.

For example $x > 1$ becomes $-x < -1$. People often remember to do this if they are multiplying by a constant but not when they multiply by a variable such as x, whose value is not known so it could be positive *or negative*. So it is very important to split your working into two cases if you have to multiply an inequality by a variable.

Example 1.2.1 Find all real numbers x such that $\frac{12}{x+4} \geqslant 2$.

If $x+4 < 0$ then $12/(x+4) < 0 < 2$. If $x+4 = 0$ then $12/(x+4)$ is not defined. Therefore $x+4 > 0$, i.e. $x > -4$. Since $x+4 > 0$ we can multiply the inequality by $x+4$ without changing the direction of inequality. This gives $12 \geqslant 2(x+4) = 2x+8$. Hence $2x \leqslant 4$ which means that $x \leqslant 2$. The solution set for the inequality is thus the interval $(-4, 2]$.

Sketching a graph is a good way to double-check your answer.



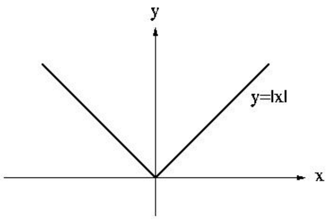
The portion of the graph above the dotted horizontal line correspons to the correct range of values.

1.2.3 Absolute Value

Let x be a real number. The *absolute value* or *modulus* of x, written $|x|$, is the distance between x and 0 on the number line. So it is always positive. It is defined by

$$|x| = \begin{cases} x & \text{if } x \geq 0; \\ -x & \text{if } x < 0. \end{cases}$$

The function $f(x) = |x|$ has the following graph:

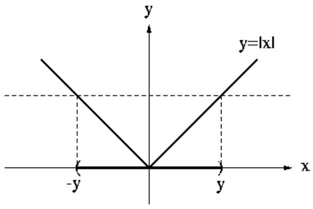


The properties of $|x|$ are as follows, for $x, y \in \mathbb{R}$:

- $|-x| = |x|$
- $|xy| = |x||y|$
- $\frac{x}{y} = \frac{|x|}{|y|}$ $(y \neq 0)$

1.2.4 Inequalities Involving Absolute Value

If $y \geqslant 0$ then the statement $|x| < y$ means $-y < x < y$, that is, $x \in (-y, y)$. Consider the following graph to see why this is true.



Similarly

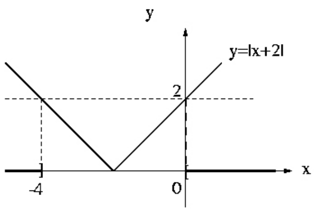
- $|x| \leqslant y$ means $-y \leqslant x \leqslant y$.
- $|x| > y$ means $x > y$ or $x < -y$.
- $|x| \geqslant y$ means $x \geqslant y$ or $x \leqslant -y$.
- $|x| = y$ means $x = y$ or $x = -y$ (it is usual to write $x = \pm y$).

Example 1.2.2 1. Find all $x \in \mathbb{R}$ for which $|x - 1| \leqslant 3$.

We have $|x-1| \leqslant 3$ if and only if $-3 \leqslant x-1 \leqslant 3$. Adding 1, $-2 \leqslant x \leqslant 4$ and the solution set is the interval $[-2, 4]$. We can check this by drawing a graph.

2. Find all real numbers $x \in \mathbb{R}$ for which $|x + 2| \geqslant 2$.

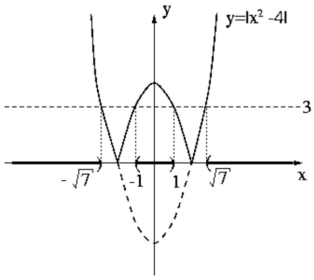
If $|x+2| \geqslant 2$ then either $x+2 \leqslant -2$ or $x+2 \geqslant 2$, i.e. $x \leqslant -4$ or $x \geqslant 0$. So x satisfies the inequality if and only if it lies in the set $(-\infty, -4] \cup [0, \infty)$.



3. Find all $x \in \mathbb{R}$ for which $|x^2 - 4| > 3$.

   Suppose that $|x^2 - 4| > 3$. Then either $x^2 - 4 > 3$ or $x^2 - 4 < -3$. If $x^2 - 4 > 3$ then either $x > \sqrt{7}$ or $x < -\sqrt{7}$, i.e. $x \in (-\infty, -\sqrt{7}) \cup (\sqrt{7}, \infty)$. If $x^2 - 4 < -3$ then $x^2 < 1$ which implies that $x \in (-1, 1)$. Thus the complete solution set to the inequality is

$$(-\infty, -\sqrt{7}) \cup (-1, 1) \cup (\sqrt{7}, \infty),$$



1.3 Complex Numbers

1.3.1 Imaginary Numbers

There is only space here to give a brief overview of the complex numbers. It is the number system formed when we want to include all square roots of the number system inside the system itself. Remarkably we can do this by adding a single extra number i (standing for *imaginary*) such that $i^2 = -1$. Of course there is no real number with this property as squaring a real number always results in something greater than or equal to zero.

Let $y < 0$. Then for some $m > 0$,

$$\sqrt{y} = \sqrt{-m} = \sqrt{-1} \cdot \sqrt{m} = i\sqrt{m}$$

A number of the form yi for $y \in \mathbb{R}$ is called a *purely imaginary number*. So square roots of real numbers are always real or purely imaginary.

1.3.2 The Complex Number System and its Arithmetic

What if we consider roots of purely imaginary numbers? Note that

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

So the expression on the left hand side, a sum of a real number and a purely imaginary number, can be thought of as the square root of a purely imaginary number.

> A *complex number* is an expression of the form $x+yi$ where x amd y are real numbers and $i^2 = -1$. The complex number system is denoted by $\mathbb{C}$.

To add, subtract and multiply complex numbers, we use ordinary rules of arithmetic and algebra and whenever it appears, we substitute i^2 with -1. For example

$$\begin{aligned}(3-2i)+(5+i) &= 8-i \\ (6+i)-(6-2i) &= 3i \\ (2-i)(3+2i) &= 6-3i+4i-2i^2 = 8+i\end{aligned}$$

To divide a complex number by another complex number first multiply the numerator and denominator by what is called the *complex conjugate* of the complex number.

If $z = a + ib$ then the *complex conjugate* of z, written $\overline{z}$, is $a - ib$.

The properties of the complex conjugate are as follows.

- $z\overline{z} = a^2 + b^2$, which is real.
- $z + \overline{z} = 2a$, also real.
- $z - \overline{z} = 2bi$, which is purely imaginary.

For the purposes of division the first of these properties is the most important. For example

$$\begin{aligned}\frac{1-i}{2+i}\cdot\frac{2-i}{2-i} &= \frac{1}{2^2+1^2}(1-i)(2-i) \\ &= \frac{1}{5}(2-2i-i-1) \\ &= \frac{1}{5}-\frac{3}{5}i\end{aligned}$$

1.3.3 Solving Polynomial Equations Using Complex Numbers

Remember that for quadratic equations

$$ax^2 + bx + c = 0 \text{ implies } x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}.$$

If the discriminant $b^2 - 4ac < 0$ then there are no real roots. But in this case we in fact have two complex roots. For example $x^2 + 4x + 5 = 0$ gives $-2 \pm \frac{1}{2}\sqrt{-4} = -2 \pm i$.

It can be shown that the roots of such a quadratic equation, with negative discriminant, always occur in conjugate pairs, i.e. they are conjugates of each other. We should also mention (but not prove, as it is difficult) the following

Fundamental Theorem of Algebra: Given a polynomial equation such as

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

the equation has at most n roots, all of which are in $\mathbb{C}$.

So by introducing a single quantity, the imaginary square root of -1, we get a lot in return - all polynomial equations can now be solved.

1.3.4 Geometry of Complex Numbers

A complex number $z = x + iy$ can be plotted as (x, y) in the plane. This plane is called the *Argand diagram* or the *complex plane* (c.f. the real line). With this geometric interpretation,

- $\overline{z}$ is the reflection of z in the x-axis.
- $|z| = \sqrt{x^2 + y^2}$ is the *modulus* of z, the length of the line segment from the origin to z.
- for $z \neq 0$, the *argument* of z, written $\arg z$, is the angle the same line segment makes with the real axis, measured anti-clockwise.

So the modulus generalises the idea of the absolute value of a real number, and argument generalises the idea of the sign of a real number, in the sense that a complex number is a real number $x > 0$ if and only if it has argument 0, and it is a real number

$x < 0$ if and only if it has argument $\frac{\pi}{2}$. For example, $|i| = 1$ and $\arg i = \frac{\pi}{2}$. We can write any complex number in *polar form*, i.e. as $z = r(\cos\theta + i\sin\theta)$ where r is $|z|$ and θ is $\arg z$. In general,

- $\arg zw = \arg z + \arg w$
- $|zw| = |z||w|$

If $x \neq 0$ then $\arg(z) = \tan^{-1}(\frac{y}{x})$. But it is important not to just put the relevant values into your calculator and work out inverse tangent because it will give you the wrong answer half of the time. You need to draw a picture and work out which quadrant of the complex plane the complex number is in. For example,

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$$

but

$$-1 - i = \sqrt{2}\left(\cos\left(\frac{-3\pi}{4}\right) + i\sin\left(\frac{-3\pi}{4}\right)\right).$$

The following is used to find roots (not just square ones) of complex numbers:

De Moivre's Theorem: $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$

For example, to calculate the fifth roots of unity (one), we write $z = r(\cos\theta + i\sin\theta)$ and seek to solve the equation

$$z^5 = r^5(\cos\theta + i\sin\theta)^5 = 1$$

So $r = 1$, because $r > 0$ is real, and by de Moivre's theorem we have

$$\cos 5\theta + i\sin 5\theta = 1.$$

Importantly, $5\theta = arg 1$ does not just mean $5\theta = 0$ but that $5\theta = 2\pi m$ for any integer m. This gives $\theta = \frac{2\pi m}{5}$ and $m = 0$, 1, 2, 3 and 4 all give unique values. So there are five fifth roots of 1:

- $z_1 = \cos 0 + i\sin 0 = 1$
- $z_2 = \cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5}$
- $z_3 = \cos\frac{4\pi}{5} + i\sin\frac{4\pi}{5}$
- $z_4 = \cos\frac{-4\pi}{5} + i\sin\frac{-4\pi}{5}$
- $z_5 = \cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5}$

We can plot them on an Argand diagram as follows: As an exercise, try to calculate the sixth and seventh roots of 1, and the fifth roots of -1.

Chapter 2

Vectors and Matrices

Vectors and matrices can be thought of as generalisations of numbers. They have their own rules of arithmetic, like numbers, and a number can be thought of as a vector of dimension 1 or a 1 by 1 matrix.

2.1 Vectors

Geometrically vectors are something with direction as well as magnitude. More abstractly they are lists of real numbers. In different situations it helps to think of them as one or the other, or both. For calculations treat them as lists of numbers. But it sometimes can help intuition to think of them as geometric objects. The vector

$$\underline{x} = \begin{bmatrix} a \\ b \end{bmatrix}$$

has magniutude, more commonly called *modulus*,

$$|\underline{x}| = \sqrt{a^2 + b^2}.$$

We will illustrate vector operations using vectors of length two or three. Vectors can have any number of entries but the operations on them are similar. Except for the cross product which is only defined for vectors of length three.

Each number in the list is called a *component*. Vectors can be added and subtracted by adding and subtracting component-wise, i.e.

$$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a+c \\ b+d \end{bmatrix}$$

Similarly vectors can be multiplied by a "scalar", i.e. a real number, as follows.

$$\lambda \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \lambda a \\ \lambda b \end{bmatrix}$$

Another useful concept is the *dot product* or *scalar product* of two vectors. This is defined as

$$\begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd$$

Note that the answer is a real number, not a vector.

The set of vectors of length n is denoted by $\mathbb{R}^n$.

If $\underline{x}$ and $\underline{y}$ are vectors in $\mathbb{R}^n$ then we have

$$\underline{x} \cdot \underline{y} = |\underline{x}||\underline{y}| \cos\theta$$

where θ is the angle between $\underline{x}$ and $\underline{y}$. In particular,

- If $\underline{x}$ and $\underline{y}$ are unit vectors (vectors whose modulus is 1) then we have $\underline{x} \cdot \underline{y} = \cos\theta$, so the dot product tells us the length of the perpendicular projection of $\underline{x}$ onto $\underline{y}$, and vice versa.
- If $\underline{x}$ and $\underline{y}$ are perpendicular, then $\cos\theta = 0$ so the dot product of the vectors is zero. Moreover, if $|\underline{x}| \neq 0$ and $|\underline{y}| \neq 0$ then $\underline{x} \cdot \underline{y} = 0$ implies that $\underline{x}$ and $\underline{y}$ are perpendicular.

2.2 Matrices and Determinants

2.2.1 Arithmetic of Matrices

Matrices are things which transform vectors to other vectors. If we have a 2-dimensional vector, for example, we have to say where each component of the vector goes to, which means 2 lots of 2. So the obvious thing is to represent a matrix as a square of numbers, like this:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

This means that $\begin{bmatrix} x \\ y \end{bmatrix}$ goes to $\begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}$ under the transformation. If we follow it by another transformation

$$\begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

we get

$$\begin{bmatrix} e(ax+by)+f(cx+dy) \\ g(cx+dy)+h(ax+by) \end{bmatrix} = \begin{bmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

And this shows us how to multiply matrices. You don't need to remember this, just the pattern of how it works.

When multiplying matrices, to get the entry in row p and column q of the product of two matrices MN, form the dot product of row p of M with column q of N.

The same rule works for multiplying $n \times m$ matrices by $m \times p$ matrices. The product of two matrices is only defined when the number of columns of the first is equal to the number of rows of the second.

If a real number is not zero, we can divide another real number by it, which is the opposite of multiplication. Similarly, there is an opposite operation to multiplying by a matrix. It is just a bit more involved, as is the criterion for when and when you can't do it.

2.2.2 Inverse Matrices and Determinants

There is a special matrix called I, the *identity matrix.* For 2×2 matrices it is

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and it has the property that for every 2×2 matrix A, $IA = AI = A$. The $n \times n$ identity matrix is written I_n and is the obvious generalization of this.

A matrix B is called the *inverse* of a matrix A if $AB = BA = I$. If a matrix A has an inverse then it is unique and it is written as A^{-1}.

There is a test for whether or not a matrix A has an inverse. You calculate something called the *determinant* of the matrix, $\det A$, sometimes written as $|A|$. For the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$\det A$ is equal to $ad - bc$.

For a 3×3 matrix

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

the determinant is given by

$$\det A = a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}$$

A *square matrix* is one with the same number of rows as columns.

A square matrix A has an inverse if and only if $\det A$ is not zero.

For a 2×2 matrix, if $\det A$ is not zero then the inverse is given by

$$A^{-1} = \frac{1}{detA}\left[\begin{array}{cc} d & -b \\ -c & a \end{array}\right].$$

For larger square matrices, I find it easier to use row reduction to find the inverse.

2.2.3 The Cross Product

The *cross product* or *vector product* is a construction on vectors but it has been included here because you need to know about determinants to calculate them. You multiply two vectors and get another vector, unlike the dot product where the answer is a scalar. It is only defined for three-dimensional vectors. If $\underline{a} = (a_x, a_y, a_z)$ and $\underline{b} = (b_x, b_y, b_z)$ are vectors in $\mathbb{R}^3$ then their vector product is written $\underline{a} \times \underline{b}$ and it is calculated by evaluating the following "determinant". The top row of the matrix contains unit vectors $\underline{i}$,$\underline{j}$ and $\underline{k}$, so it isn't the determinant of a real matrix but it is calculated in the same way.

$$\underline{a} \times \underline{b} = \det \begin{bmatrix} \underline{i} & \underline{j} & \underline{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{bmatrix}$$

For example if $\underline{a} = (1, 1, 1)$ and $\underline{b} = (1, -1, 1)$ then

$$\begin{aligned} \underline{a} \times \underline{b} &= \det \begin{bmatrix} \underline{i} & \underline{j} & \underline{k} \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \\ &= \\ \underline{i}(1-(-1)) - \underline{j}(1-1) + \underline{k}(-1-1) \\ &= 2\underline{i} - 2\underline{k} \\ &= (2, 0, -2) \end{aligned}$$

Note that by calculating the dot product we can see that $a \times b$ is perpendicular to both a and b. The vector product always has these properties:

- If a and b are not parallel then $\underline{a} \times \underline{b}$ is perpendicular to $\underline{a}$ and $\underline{b}$.
- $\underline{a} \times \underline{a}$ is the zero vector.

2.3 Systems of Linear Equations and Row Reduction

2.3.1 Systems of Linear Equations

What we call "systems of linear equations" are usually called "simultaneous equations" at school. They can be written using matrices. For example,

$$\begin{aligned} x + y + z &= 2 \\ x - 2y + 2z &= 5 \\ 2x + y + z &= 3 \end{aligned}$$

can be written using a 3×3 matrix:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 2 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}$$

2.3.2 Row Reduction

There is a systematic way to solve the system of equations called *row reduction* (sometimes it is also called *Gaussian elimination*). You are allowed to use one of three moves:

- multiply a row by a nonzero number
- exchange two rows
- add a multiple of a row to another row

We aim to reduce the matrix and the right hand side to "row-echelon" form which means that all entries below the leading diagonal of the matrix (from the top left to the bottom right) should be zero. To do this

- Get zeros in the first column below the diagonal, by adding multiples of the first row to each other row in turn
- Get zeros in the second column, by adding multiples of the second row to each other row in turn (there is only one other row for a 3×3 matrix).
- And so on, if the matrix is larger than 3×3.

Once the system is row echelon form, the solution can be simply read off.

A solution to a linear system of three equations in three unknowns can be either

- a point, i.e. a unique solution
- a line of solutions, if one of the variables is unconstrained
- a plane of solutions, if two of the variables are unconstrained
- the entire three dimensional space, if all of the variables are unconstrained

For row reduction it is common to omit the x, y and z and use a bar to separate the left hand side from the right hand side. This notation is called an "augmented" matrix.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 1 & -2 & 2 & 5 \\ 2 & 1 & 1 & 3 \end{array}\right]$$

$$\underset{\longrightarrow}{R_2-R_1, R_3-2R_1} \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & 1 & 3 \\ 0 & -1 & -1 & -1 \end{array}\right]$$

$$\underset{\longrightarrow}{3R_3-R_2} \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & 1 & 3 \\ 0 & 0 & -4 & -6 \end{array}\right]$$

The third row tells us that $-4z = -6$, i.e. $z = \frac{3}{2}$. Then row two tells us that $\frac{1}{2} - y = 1$, which gives $y = -\frac{1}{2}$. Finally, row one gives $x + 1 = 2$, so $x = 1$ and we have a unique solution.

2.3.3 Finding The Inverse of a Matrix Using Row Reduction.

Row reduction can also be used to find the inverse of a matrix efficiently. To do this we make another type of augmented matrix with the identity matrix on the right hand side. When we use row reduction to reduce the left hand side to the identity, the right hand side becomes the inverse matrix. If it is not possible to reduce the left hand side to the identity then that is because the left hand side is not an invertible matrix.

Example 2.3.1 Find the inverse of

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & -1 & 1 \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2-R_1, R_3-R_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3+2R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -3 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{R_1-2R_3, R_2-R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 7 & -4 & -2 \\ 0 & 1 & 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & -3 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{R_1-R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -3 & -1 \\ 0 & 1 & 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & -3 & 2 & 1 \end{array}\right]$$

So

$$A^{-1} = \begin{bmatrix} 5 & -3 & -1 \\ 2 & -1 & -1 \\ -3 & 2 & 1 \end{bmatrix}$$

You should check that multiplying A by A^{-1} results in the 3×3 identity matrix.

2.4 Bases

A basis is like a linear co-ordinate system. Every vector in $\mathbb{R}^2$ can be written as $a(1,0) + b(0,1)$ for some real numbers a and b. The pair $(0,1)$ and $(1,0)$ are called the *standard basis* of $\mathbb{R}^2$. But there are many other pairs of vectors, like $(1,0)$ and $(0,1)$, that also have this property. For example, $(1,2)$ and $(3,4)$ will do. To write $(x,y) = a(1,2)+b(3,4)$ we solve the system of equations $x = a+3b$ and $y = 2a + 4b$ which gives

$$a = \frac{3y - 4x}{2}, b = x - \frac{y}{2}.$$

Moreover, for any (x,y) the required a and b is unique.

A *basis* of $\mathbb{R}^2$ is a pair of vectors $(\underline{v}, \underline{w})$ in $\mathbb{R}^2$ such that every vector in $\mathbb{R}^2$ can be written uniquely as $a\underline{v} + b\underline{w}$.

Of course, we can speak similarly about a basis of $\mathbb{R}^n$ for any n, the standard basis in $\mathbb{R}^n$, which has the obvious definition, and it turns out that any basis for $\mathbb{R}^n$ has to consist of precisely n vectors.

2.5 Eigenvalues and Eigenvectors

Suppose that we have a linear transformation T which sends the basis $B = \{\underline{v_1}, \underline{v_2}\}$ to $\{a\underline{v_1} + b\underline{v_2}, c\underline{v_1} + d\underline{v_2}\}$. Then the *matrix of T with respect to B* is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

In particular, if T sends B to $\{\lambda_1\underline{v_1}, \lambda_2\underline{v_2}\}$ then B is called an *eigenbasis* of T, λ_1 and λ_2 are called *eigenvalues* of T and $\underline{v_1}$ and $\underline{v_2}$ are called *eigenvectors* of T.

That is, $T\underline{v_1} = \lambda_1\underline{v_1}$ and $T\underline{v_2} = \lambda\underline{v_2}$. This is useful because we can more easily find powers and limits of a matrix representing T. With respect to an eigenbasis the matrix of a transformation is *diagonal*, meaning all entries not on the leading diagonal are 0. For example,

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

is diagonal, and A^n is just

$$\begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix}$$

There is a process for calculating eigenvalues and eigenvectors which relies on (a) calculating determinants (b) row reduction. Eigenvalues λ of an $n \times n$ matrix A, and corresponding eigenvectors $\underline{v}$ satisfy $A\underline{v} = \lambda\underline{v}$, which can be rewritten as $(A - \lambda I_n)\underline{v} = 0$. This is a system of linear equations with zeros in the right hand side, and has a solution if and only if that solution is non-unique. That is

> To calculate eigenvalues and eigenvectors of an $n \times n$ matrix A, solve
> $$\det(A - \lambda I_n) = 0$$

Note that here we are solving a polynomial equation of degree n in λ. This could potentially have repeated roots or even complex roots. A typical exam question would ask to calculate the eigenvalues and eigenvectors of a 3×3 matrix. This entails the following (usually four) computational tasks:

- Solve the above polynomial equation to find all eigenvalues.
- For each eigenvalue, substitute it into the equation and solve the resulting linear system of equations to find the eigenvectors.

Example 2.5.1 Find eigenvalues and eigenvectors of

$$\begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

We solve

$$\det\left(\begin{bmatrix} 1-\lambda & 3 & 0 \\ 3 & -2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix}\right) = 0$$

This expands to

$$\begin{aligned} (1-\lambda)\left[(-2-\lambda)(1-\lambda)-1\right] - 3\left[3(1-\lambda)-0\right] + 0 &= 0 \\ (1-\lambda)(\lambda^2+\lambda-2-1) - 9(1-\lambda) &= 0 \\ (1-\lambda)(\lambda^2+\lambda-12) &= 0 \\ (1-\lambda)(\lambda+4)(\lambda-3) &= 0 \end{aligned}$$

giving eigenvalues $\lambda = 1$, $\lambda = -4$ and $\lambda = 3$. Now for example, to find an eigenvector for $\lambda = 3$ we solve

$$\left[\begin{array}{ccc|c} 1-\lambda & 3 & 0 & 0 \\ 3 & -2-\lambda & -1 & 0 \\ 0 & -1 & 1-\lambda & 0 \end{array}\right]$$

for $\lambda = 3$. That is,

$$\left[\begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 3 & -5 & -1 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right]$$

We use row reduction to do this. We have

$$\stackrel{2R_2, R_2+3R_1}{\longrightarrow} \left[\begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right]$$

$$\stackrel{R_3-R_2}{\longrightarrow} \left[\begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Recall from the section on solving systems of equations that the solution can either be a point, a line, a plane or all of three dimensions (for a system of three equations in three variables). We saw in that section an example where the solution was a point (i.e. a unique solution). In this case the solution is a line. This is because the bottom row says $0z = 0$, which tells us nothing about z, meaning that z is *unconstrained.* When finding eigenvectors the solution sets always turn out to be at least one-dimensional. This is because any scalar multiple of an eigenvector is also an eigenvector, by the definition of an eigenvector.

Continuing with the solution, the second row tells us that $y = -2z$ and the first then tells us that $x = -3z$. The eigenvectors for the eigenvalue 3 are hence any vectors of the form

$$\alpha \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}$$

for any real value of α.

You should continue this example and show that the eigenvectors for the eigenvalue -4 are all of the form

$$\alpha \begin{bmatrix} -3 \\ 5 \\ 1 \end{bmatrix}$$

and find those for the eigenvalue 1, verifying your answer.

Chapter 3

Functions and Limits

3.1 Functions

3.1.1 Definition of a function.

Let X and Y be sets. A *function* $f : X \to Y$ assigns a *unique* value of y to every value of x. So technically, f is actually a certain type of set of pairs of points (x, y) with $x \in X$ and $y \in Y$. That is all. X is called the *domain* of f and Y is called the *range* of f.

Example 3.1.1 Let $X = Y$ be the set of all people who have ever lived. Suppose that we define $f : X \to X$ by the rule: $f(x)$ is the father of x. Then this is a well defined function (everyone has a father, and only one father). On the other hand, we may define $c : X \to X$ by defining $c(x)$ to be the child of x. This is not a function since a given individual may have (a) no children or (b) more than one child.

Example 3.1.2 Let $X = Y = \mathbb{R}$. Then $f(x) = x^2$ defines a function whereas $f(x) = \sqrt{x}$ does not, since every nonzero real number has two square roots.

We usually write f for a function and $f(x)$ for the value of f at x. We must be careful when specifying functions. Often whether or not something is defined as a function depends on the domain. For example, $f(x) = \frac{1}{x}$ is not a function from $\mathbb{R}$ to $\mathbb{R}$ because $f(0)$ is not defined. It is, however, a function from $(-\infty, 0) \cup (0, \infty)$ to $\mathbb{R}$.

3.1.2 Piping Functions Together

There are a few ways of constructing new functions from old ones. One of these is called *composition*. Given two functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ we may construct a third function $g \circ f : \mathbb{R} \to \mathbb{R}$, called g **composed with** f by the rule

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

For example, if $f(x) = x^2$ and $g(x) = x + 1$ then $g \circ f(x) = x^2 + 1$ whereas $f \circ g(x) = (x + 1)^2$.

Also, if f and g are both functions from $\mathbb{R}$ to $\mathbb{R}$ then we can define functions

1. $f + g : \mathbb{R} \to \mathbb{R}$ by $(f + g)(x) = f(x) + g(x)$,
2. $f - g : \mathbb{R} \to \mathbb{R}$ by $(f - g)(x) = f(x) - g(x)$,
3. $fg : \mathbb{R} \to \mathbb{R}$ by $(fg)(x) = f(x)g(x)$,
4. $\frac{f}{g} : \mathbb{R} \to \mathbb{R}$ by $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$ (if $g(x)$ is never equal to 0),
5. $kf : \mathbb{R} \to \mathbb{R}$ by $(kf)(x) = kf(x)$ $(k \in \mathbb{R})$ and
6. $f^k : \mathbb{R} \to \mathbb{R}$ by $(f^k)(x) = (f(x))^k$ (where defined).

Later, as we meet rules for manipulating limits, derivatives and integrals, you will see that they are usually concerned with building functions up in the ways above.

3.1.3 Inverse Functions

A function $f : \mathbb{R} \to \mathbb{R}$ is *injective* if $f(x_1) = f(x_2)$ always implies that $x_1 = x_2$. If $f(x) = 2x$ then f is injective because if $2x_1 = 2x_2$ then $x_1 = x_2$. If we draw a horizontal line across the graph of an injective function the line will never meet more than one point of the graph. The function $g : \mathbb{R} \to \mathbb{R}$, $g(x) = x^2$ is not injective. For example, $g(-1) = g(1)$.

We say that $f : \mathbb{R} \to \mathbb{R}$ is *surjective* if for every $y \in \mathbb{R}$ there is an $x \in \mathbb{R}$ with $f(x) = y$. The function $f(x) = x^3$ is surjective, since given $y \in \mathbb{R}$ the number $y^{\frac{1}{3}}$ maps onto y (every real number has a unique cube root, unlike the case for square roots). In the graph of a surjective function, every horizontal line intersects the graph in at least one point, as in the graph of x^3. The function $f(x) = \sin(x)$ is an example of a function from $\mathbb{R}$ to $\mathbb{R}$ which is not surjective. For example, there is no real number x such that $\sin(x) = 2$.

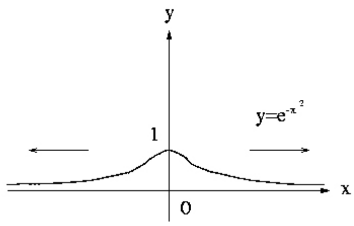
If f is both injective and surjective then we say it is *bijective*, or that it is a *bijection*. For example, $f(x) = 2x$ and $f(x) = x^3$ both have this property. In order for us to be able to define an inverse for f (i.e. a function f^{-1} such that $f^{-1} \circ f(x) = x$ for all x), f must be a bijection. For example, $f(x) = \sin(x)$ has no inverse $f^{-1} : \mathbb{R} \to \mathbb{R}$, although we can define

$$\sin^{-1} : [-1, 1] \to [-\pi/2, \pi/2].$$

3.2 Limits

The formal definition of a *limit* of a function is technical and has been omitted. We speak of the limit of a function as it *tends* to either ∞, $-\infty$ or some real number, often 0. It is perhaps easier to see the need for the first two cases. But the third is also useful when defining derivatives and similar constructions where you let Δ be a small quantity or change, then take the limit of some function as Δ tends to zero.

Example 3.2.1 1. Let $f(x) = e^{-x^2}$, $f : \mathbb{R} \to \mathbb{R}$, which has the following graph:



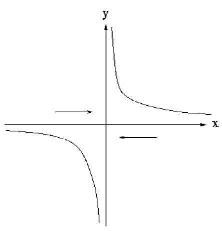
As x becomes arbitrarily large, $f(x)$ becomes arbitrarily close to 0. This also occurs as x becomes arbitrarily negative. We write

$$\lim_{x\to\infty} f(x) = 0$$

and

$$\lim_{x\to -\infty} f(x) = 0.$$

2. Let $f(x) = \frac{1}{x}$, $f : (-\infty, 0) \cup (0, \infty) \to \mathbb{R}$:

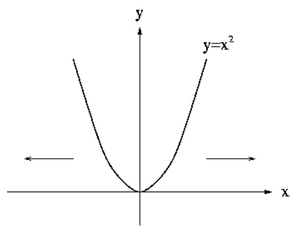


As x becomes arbitrarily close to 0, either from the left or from the right, y does not become arbitrarily close to any value. In fact it becomes arbitrarily far away from any value! In this situation we say that

$$\lim_{x \to 0} f(x)$$

does not exist.

3. Let $f(x) = x^2$.



As x becomes arbitrarily large, so does $f(x)$. Here we say that

$$\lim_{x \to \infty} f(x)$$

does not exist. Similarly,

$$\lim_{x \to -\infty} f(x)$$

does not exist.

Rules For Calculating Limits Part 1

Let k, L and M and a be real numbers and let f and g be real functions. Suppose that $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Then

- **Sum Rule:** $\lim_{x\to a}(f+g)(x) = L + M$
- **Difference Rule:** $\lim_{x\to a}(f-g)(x) = L - M$
- **Product Rule:** $\lim_{x\to a}(fg)(x) = LM$
- **Quotient Rule:** if $g(x)$ is never 0 and $M \neq 0$ then $\lim_{x\to a} \frac{f}{g}(x) = L \,/\, M$
- **Constant Rule:** $\lim_{x\to a}(kf)(x) = kL$
- **Power Rule** $\lim_{x\to a}(f^k)(x) = L^k$ (where defined)

These also hold if we replace $x \to a$ by $x \to \infty$ or $x \to -\infty$.

Example 3.2.2 1. Let

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

Then

$$f(h) = \frac{x^2 + 2hx + h^2 - x^2}{h} = \frac{h^2 + 2hx}{h}.$$

By the sum rule,

$$\begin{aligned} \lim_{h\to 0} f(h) &= \lim_{h\to 0}\frac{h^2}{h} + \lim_{h\to 0}\frac{2hx}{h} \\ &= 0 + 2x \\ &= 2x. \end{aligned}$$

2. Find

$$\lim_{x\to\infty}\left(q(x) = \frac{x^2+x}{3x^2+x+1}\right).$$

Let $x > 0$. Then we can divide $f(x) = x^2 + x$ by $g(x) = x^2$ to obtain $\frac{f}{g}(x) = 1 + 1/x \to 1$ as $x \to \infty$ by the sum rule since we know that $1/x \to 0$ as $x \to \infty$. Similarly if $h(x) = 3x^2 + x + 1$, $\frac{h}{g}(x) \to 3$ as $x \to \infty$. Thus by the quotient rule

$$\begin{aligned} \lim_{x\to\infty} q(x) &= \lim_{x\to\infty}\frac{\frac{f}{g}(x)}{\frac{h}{g}(x)} \\ &= \frac{1}{3}. \end{aligned}$$

Rules For Calculating Limits Part 2

- **Sandwich Theorems:** Suppose that $|f(x)| \leqslant |g(x)|$ for all $x \in \mathbb{R}$.
  - If $g(x) \to 0$ as $x \to 0$ then $f(x) \to 0$ as $x \to 0$.
  - If $g(x) \to 0$ as $x \to \infty$ then $f(x) \to 0$ as $x \to \infty$.
- **Exponential v. Polynomial:**
  - If $0 < a < 1$ and p is fixed then $a^x x^p \to 0$ as $x \to \infty$.
  - If $b > 1$ and p is fixed then $b^x x^p \to \infty$ as $x \to \infty$.
- **Logarithmic v. Polynomial:** Let $p < 0$ be fixed. Then $x^p \log(x) \to 0$ as $x \to \infty$.

Example 3.2.3 1. What is

$$\lim_{x\to\infty} \frac{\sin(x)}{x^2} ?$$

Since for all x we have $|\sin(x)| \leqslant 1$, it follows that $|\frac{\sin(x)}{x^2}| \leqslant |\frac{1}{x^2}|$. Since $\frac{1}{x^2} \to 0$ as $x \to \infty$, so does $\frac{\sin(x)}{x^2}$ by the second of the sandwich theorems.

2. What is

$$\lim_{x\to\infty} e^{-x} x^{2010} ?$$

In the "exponential v. polynomial" theorem, let $a = \frac{1}{e}$ and $p = 1999$. Then $e^{-x} x^{2010} = a^x x^p \to 0$ as $x \to \infty$.

3.3 Continuity

Informally, a real function is *continuous* if the graph does not suddenly jump around from one value to another. That is, a very small change in x will only produce a very small change in $f(x)$. More precisely, $f : X \to \mathbb{R}$ is said to be *continuous* if for all a in X, $\lim_{x \to a} f(x) = f(a)$.

Example 3.3.1 The function

$$f(x) = \begin{cases} x+1 & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$

is continuous everywhere except at 0.

Rules of Continuity Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be functions which are continuous at $x_0 \in \mathbb{R}$. Then

1. $f + g$ and $f - g$ are continuous at x_0
2. fg is continuous at x_0
3. kf is continuous at x_0 for any $k \in \mathbb{R}$.
4. If $g(x)$ is never 0 then $\frac{f}{g}$ is continuous at x_0.
5. f^p is continuous at x_0 for any $p \in \mathbb{R}$ where defined
6. If g is continuous at $f(x_0)$ then $g \circ f$ is continuous at x_0.

Most functions that you are familiar with are continuous. For example all polynomial functions $f(x) = a_n x^n + \cdots a_1 x + a_0$ are continuous, sin, cos, the exponential and logarithmic functions are all continuous, and rational functions

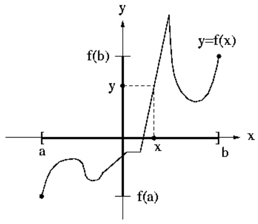
$$f(x) = \frac{a_n x^n + \cdots a_1 x + a_0}{b_m x^m + \cdots b_1 x + b_0}$$

are continuous at all points x where $b_m x^m + \cdots b_1 x + b_0 \neq 0$. They are not defined at the points x where $b_m x^m + \cdots b_1 x + b_0 = 0$, so it doesn't make sense to ask whether or not they are continuous there. For example, the function $f : (-\infty, 0) \cup (0, \infty) \to \mathbb{R}$ defined by $f(x) = \frac{1}{x}$ is continuous. However if we let L be any real value then the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x \neq 0 \\ L & \text{if } x = 0 \end{cases}$$

is not.

The Intermediate Value Theorem Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then for all $y \in \mathbb{R}$ between $f(a)$ and $f(b)$ there exists a value $x \in [a, b]$ such that $f(x) = y$.

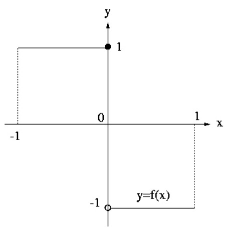


This theorem is used for showing that roots of equations exist.

Example 3.3.2 The function $f(x) = x^3 - x - 1$ is continuous because it is a polynomial function. $f(1) = -1$ and $f(2) = 5$. Since $0 \in [-1, 5]$ there exists $x_0 \in [1, 2]$ such that $f(x_0) = 0$, i.e. the equation $x^3 - x - 1 = 0$ has a root in $[1, 2]$. In fact the root is about 1.324718.

The conclusion of the intermediate value theorem is only true for continuous functions. For example let $f : [-1, 1] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in [-1, 0] \\ -1 & \text{if } x \in (0, 1]. \end{cases}$$



If we choose any value $y \in (-1, 1)$ then there is no $x \in [-1, 1]$ with $f(x) = y$.

Chapter 4

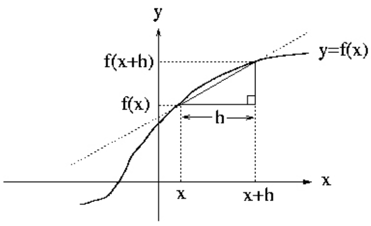
Calculus of One Variable Part 1: Differentiation

4.1 Derivatives

Some real functions have a well defined "gradient" and some do not.

For example, what is the gradient of the graph of $f(x) = |x|$ at $x = 0$? Is it 1? -1? Or should we take an average of these and say that the gradient is 0? None of those make sense, particularly, so we just say it is not defined. The mathematical name for being able to define a gradient at a point a of the graph of a function $f : \mathbb{R} \to \mathbb{R}$ is *differentiability* at a.

Recall how the gradient to a curve at a point is defined.



The tangent to the graph at x is approximated by the line which extends the sloping side of the triangle shown. The gradient of this approximation is

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

We want to take the limit of this as $h \to 0$, if the limit exists.

$f : \mathbb{R} \to \mathbb{R}$ is **differentiable** at $x \in \mathbb{R}$ if this limit exists:

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

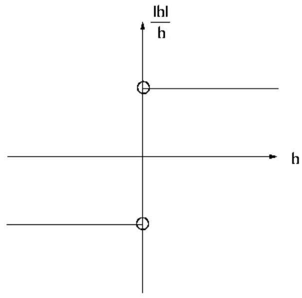
If the limit does exist, then we define the **derivative** of f at x, written $f'(x)$, to be the value of the limit. This is also written as

$$\frac{d}{dx} f(x).$$

If f is differentiable at all points $x \in \mathbb{R}$ then we say that it is **differentiable** and define the **derivative** of f to be the function f' given by

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

Example 4.1.1 1. Let $f(x) = |x|$, whose graph is as above. Then
$$\frac{f(0+h)-f(0)}{h} = \frac{|h|}{h}$$
The graph of $g : (-\infty, 0) \cup (0, \infty) \to \mathbb{R}$ given by $g(h) = |h|/h$ is as follows.



We see that $\lim_{h\to 0^+} g(h) = 1$ whereas $\lim_{h\to 0^-} g(h) = -1$. Hence $\lim_{h\to 0} g(h)$ doesn't exist, or these one sided limits would both give the same value. Thus f is not differentiable at 0.

2. Let $f(x) = x^2$. Then
$$\frac{f(x+h)-f(x)}{h} = \frac{(x+h)^2 - x^2}{h}$$
We saw in section 1.2 that this limit exists and is equal to $2x$. Thus the derivative of f is the function f' given by $f'(x) = 2x$. We can show similarly that if n is any natural number then
$$f(x) = x^n \text{ is differentiable and } f'(x) = nx^{n-1}.$$

Differentiability at a point implies continuity at that point. It follows that differentiable functions are continuous. However, continuous functions need not neccessarily be differentiable, as the example of the function f given by $f(x) = |x|$ shows.

Rules of Differentiability Let $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ be differentiable functions.

1. **Sum Rule:** $f+g$ is differentiable and $(f+g)' = f'+g'$.
2. **Difference Rule:** $f-g$ is differentiable and $(f-g)' = f'-g'$.
3. **Product Rule:** fg is differentiable and $(fg)' = fg' + gf'$.
4. **Quotient Rule:** f/g is differentiable (where defined) and
$$(f/g)' = \frac{gf' - fg'}{g^2}.$$

Exercise 4.1.2 1. Differentiate $f(x) = x^3$ from first principles (i.e use a limit argument as with x^2 above. Don't just write down "the answer is $3x^2$".)

2. Use the above rules to differentiate

$$\frac{x^{1999} + x^2 + 1}{x^{2000} + x^3 + 1}.$$

4.2 The Chain Rule

The chain rule tells us how to differentiate the composition of two functions.

The Chain Rule: If f and g are functions from $\mathbb{R}$ to $\mathbb{R}$, f is differentiable at $x \in \mathbb{R}$ and g is differentiable at $f(x)$ then $g \circ f$ is differentiable at x and

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

This theorem is also true for functions only defined on subsets of $\mathbb{R}$, provided that we can compose them.

Thus if f and g are differentiable (everywhere) then so is $g \circ f$ and we have

$$(g \circ f)' = (g' \circ f)f'.$$

We can now prove the power rule for all rational indices p.

The power rule is

$$\frac{d}{dx}(x^p) = px^{p-1} \ (x \neq 0 \text{ if } p < 1).$$

We did this for x^2 from first principles. It follows by induction for all $p \in \mathbb{N}$ by the product rule, since

$$\begin{aligned} \frac{d}{dx}(x^p) &= \frac{d}{dx}(x.x^{p-1}) \\ &= (p-1)x^{p-2}.x + 1.x^{p-1} \text{ (by the product rule)} \\ &= (p-1+1)x^{p-1} \\ &= px^{p-1}. \end{aligned}$$

If $p = 0$, $\frac{d}{dx}(1) = 0$ everywhere (this can't be incorporated into the power rule if $x = 0$).

Suppose that $p \in \mathbb{Z}$ and $p < 0$. Let $q = -p$. Then $q > 0$ and we use the quotient rule.

$$\begin{aligned}\frac{d}{dx}\left(\frac{1}{x^q}\right) &= \frac{x^q.0 - qx^{q-1}.1}{x^{2q}} \\ &= \frac{-q}{x^{q+1}} \\ &= px^{p-1}.\end{aligned}$$

Thus the power rule holds for all $p \in \mathbb{Z}$. To extend it to rational indices $p = \frac{n}{m}$ we can use the chain rule. We have

$$\begin{aligned}\frac{d}{dx}\left[(x^{\frac{n}{m}})^m\right] &= nx^{n-1} \text{ (by the power rule for } \mathbb{Z}) \\ &= \frac{d}{dx}\left(x^{\frac{n}{m}}\right).m\left(x^{\frac{n}{m}}\right)^{m-1}\end{aligned}$$

by the chain rule. Hence

$$\begin{aligned}\frac{d}{dx}\left(x^{\frac{n}{m}}\right) &= \frac{n}{m}.\frac{x^{n-1}}{(x^{\frac{n}{m}})^{m-1}} \\ &= \frac{n}{m}.x^{n-1-n+\frac{n}{m}} \\ &= \frac{n}{m}x^{\frac{n}{m}-1} \\ &= px^{p-1}.\end{aligned}$$

Thus the power rule holds for all rational indices. In fact it holds for all real indices, but we must be careful what we mean there.

4.3 Some Standard Derivatives

The following derivatives should be learnt.

$f(x)$		$f'(x)$
$\sin(x)$		$\cos(x)$
$\cos(x)$		$-\sin(x)$
$\tan(x) = \frac{\sin(x)}{\cos(x)}$	$(\cos(x) \neq 0)$	$\sec^2(x)$
$\cot(x) = \frac{\cos(x)}{\sin(x)}$	$(\sin(x) \neq 0)$	$-\mathrm{cosec}^2(x)$
$\sec(x) = \frac{1}{\cos(x)}$	$(\cos(x) \neq 0)$	$\sec(x)\tan(x)$
$\mathrm{cosec}(x) = \frac{1}{\sin(x)}$	$(\sin(x) \neq 0)$	$-\mathrm{cosec}(x)\cot(x)$
e^x		e^x
$\log(x)$	$(x > 0)$	$\frac{1}{x}$

4.4 Differentiating Inverse Functions

If f is differentiable and bijective, and hence has an inverse function f^{-1}, then we can show that f^{-1} is also differentiable at points where $f'(x) \neq 0$. In fact since $f^{-1}(f(x)) = x$ for all x we have

$$\frac{d}{dx}\left[f^{-1}(f(x))\right] = \frac{d}{dx}(x) = 1.$$

and by the chain rule, if we write $y = f(x)$ then

$$(f^{-1})'(y) = \frac{1}{f'(x)}$$

provided $f'(x) \neq 0$, i.e. we have

$$(f^{-1})' = \frac{1}{f'}$$

For example, if we let $y = \sin(x)$ then we can differentiate $\sin^{-1}(x)$ using this method. Recall from chapter 1 that

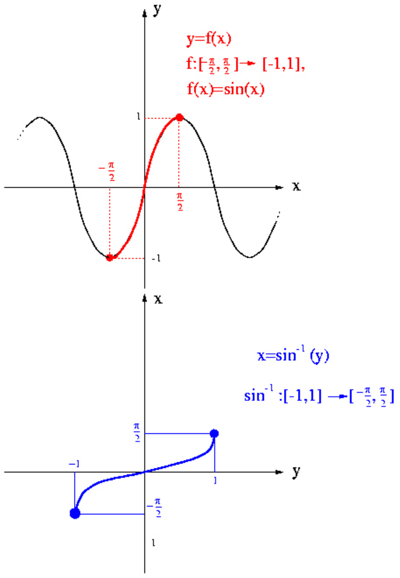
$$\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$$

is a bijection, so $\sin^{-1}$ is differentiable everywhere in its domain $[-1, 1]$ except at -1 and 1, where $\frac{dy}{dx} = 0$.

We have

$$\begin{aligned} \frac{d}{dy}\sin^{-1}(y) &= \frac{1}{\frac{d}{dx}(\sin(x))} \\ &= \frac{1}{\cos(x)} \\ &= \frac{1}{\sqrt{1 - sin^2(x)}} \quad (\text{since } \cos^2 x + \sin^2 x = 1) \\ &= \frac{1}{\sqrt{1 - y^2}}. \end{aligned}$$

Notice how this is undefined for $y = \pm 1$. The following graphs may help to understand this.



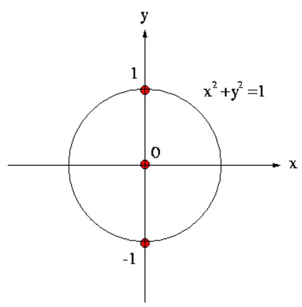
y
y=f(x)
f:[-π/2, π/2] → [-1,1],
f(x)=sin(x)
1
-π/2
x
π/2
-1
x
x=sin^-1 (y)
sin^-1 :[-1,1] → [-π/2, π/2]
π/2
-1
y
1
-π/2

4.5 Implicit Differentiation

1. Differentiate the entire equation, possibly using the chain rule.

2. Rearrange the result to get the desired derivative.

Sometimes we may have a curve defined *implicitly*. For instance in the (x, y)-plane, the circle of radius 1 is the set of all points (x, y) such that $x^2 + y^2 = 1$. Here we would say that y is given **implicitly** in terms of x as opposed to if it appears in the form $y = f(x)$ for some function x, where we would say that y is given **explicitly**. Note that the graph of the above curve is not the graph of a function. For instance $(0, 1)$ and $(0, -1)$ are both points on the graph with x-co-ordinate equal to 0. In the graph of a function there can only be one such point.



In general y appears implicitly in terms of x whenever $F(x, y) = 0$ for some function $F : \mathbb{R}^2 \to \mathbb{R}$. For instance, above we have $F(x, y) = x^2 + y^2 - 1$.

However, even though this does not always define y as a function of x we can still differentiate an implicit formula for y in terms of x, by using the chain rule.

Example 4.5.1 Differentiating the formula

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

with respect to x, we obtain

$$2x + 2y\frac{dy}{dx} = 0$$

which gives

$$\frac{dy}{dx} = -\frac{x}{y}.$$

This is only valid for $y \neq 0$. The values where $y = 0$ are when $x = 1$ and $x = -1$. In both of these places the graph of the circle has a vertical tangent and so its gradient at these points is not defined.

Notice also that if $y \neq 0$ and $|y| \leqslant 1$ there are two values of y for a given value of x and that one is $+\sqrt{1-x^2}$ and the other is $-\sqrt{1-x^2}$. This allows us to define two functions (called **branches**) $f_+ : [-1,1] \to \mathbb{R}$ defined by $f_+(x) = +\sqrt{1-x^2}$ and $f_- : [-1,1] \to \mathbb{R}$ defined by $f_-(x) = -\sqrt{1-x^2}$. The two possible values of dy/dx for each value of x are just the derivatives of each of these functions at x.

For most implicit formulae there are many different branches and many different values of the derivative for a given x. For such formulae you should *either* specify a branch *or* refer to the "derivative at (x, y)"

Example 4.5.2 Differentiate

$$(y + x)^5 = \sin(x^7)e^{3y}.$$

We have

$$5(y + x)^4(\frac{dy}{dx} + 1) = 7x^6 \cos(x^7)e^{3y} + sin(x^7).3\frac{dy}{dx}e^{3y}.$$

Therefore

$$\frac{dy}{dx}(5(y + x)^4 - 3e^{3y} \sin(x^7)) = 7x^6 \cos(x^7)e^{3y} - 5(y + x)^4$$

which gives

$$\frac{dy}{dx} = \frac{7x^6 \cos(x^7)e^{3y} - 5(y + x)^4}{5(y + x)^4 - 3e^{3y} \sin(x^7)}.$$

4.6 Logarithmic Differentiation

1. Take logs of the whole equation and apply the logarithmic manipulation rules.
2. Differentiate implicitly.
3. Rearrange to get the desired derivative.

Try to differentiate x^x. You can't treat this as a polynomial because the power is not constant. Similarly you can't treat it as an exponential function. Let $y = x^x$. Taking logs we have

$$\log(y) = x \ \log x$$

and differentiating implicitly we obtain

$$\frac{1}{y}\frac{dy}{dx} = 1 + \log(x)$$

hence

$$\frac{dy}{dx} = y(1 + \log(x)) = x^x(1 + \log(x))$$

This illustrates the principle of *logarithmic differentiation*, really just a special case of implicit differentiation. It can be used to simplify the number of operations you have to perform when differentiating a complicated product or quotient.

Example 4.6.1 Suppose that we want to differentiate

$$y = \frac{\sqrt{x^2+1}(x+5)^3}{(x+1)^2}$$

First take logs and use logarithm manipulation rules:

$$\log(y) = \frac{1}{2}\log(x^2+1) + 3\log(x+5) - 2\log(x+1)$$

Then differentiate implicitly:

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{2}.\frac{2x}{x^2+1} + \frac{3}{x+5} - \frac{2}{x+1}$$

Rearranging,

$$\begin{aligned}\frac{dy}{dx} &= y(\frac{x}{x^2+1} + \frac{3}{x+5} - \frac{2}{x+1}) \\ &= \frac{\sqrt{x^2+1}(x+5)^3}{(x+1)^2}(\frac{x}{x^2+1} + \frac{3}{x+5} - \frac{2}{x+1}).\end{aligned}$$

Of course this is not a particularly elegant expression, but using the quotient rule would require similar simplification.

Exercise 4.6.2 1. Differentiate implicitly

$$\cos(y^2) = \sin(y^3)\sqrt{1+x^2}$$

2. Use logarithmic differentiation to differentiate

(a) $f(x) = \dfrac{(x^5+1)(x^6+8)^7}{(x^4+2)^3}$

(b) $f(x) = (x^x)^x$

(c) $f(x) = x^{(x^x)}$

Note that the functions in (b) and (c) are not the same. It is a coincidence that they have the same value for $x = 2$; $2^2 = 4$ so $(2^2)^2 = 4^2 = 16$ and $2^{(2^2)} = 2^4 = 16$. However, $3^3 = 27$ and 3^{27} is *far* larger than 27^3.

4.7 Higher Derivatives

A function $f : \mathbb{R} \to \mathbb{R}$ is **twice differentiable** if f is differentiable and f' is differentiable. The derivative of f' is denoted by f'' and is called the **second derivative** of f. We may similarly define an n-**times differentiable** function and denote the n^{th} derivative of such a function by $f^{(n)}(x)$. The n should be in brackets to distinguish from the n^{th} power of f. If $y = f(x)$ then we also denote $f^{(n)}(x)$ by

$$\frac{d^n y}{dx^n}.$$

A function which is n-times differentiable for all n is called **smooth**. For example, polynomials, sin, cos and exponential functions are smooth.

4.8 L'Hôpital's Rule

To find limits of some quotients you can differentiate the numerator and denominator first, and the limit stays the same.

Quite often when we try to find a limit using the quotient rule we obtain an "answer" like

$$\frac{0}{0}, \frac{\infty}{\infty}, 0 \times \infty, \infty - \infty, 1^{\infty}, 0^0, \infty^0, \text{ or } 0^{\infty}.$$

The above pieces of nonsense are called *indeterminate forms*. If you obtain an indeterminate form you have to resort to some other means of finding the limit.

Example 4.8.1 Find

$$\lim_{x\to 0} \frac{\sin(x)}{x}.$$

We see that as $x \to 0$, both $\sin(x)$ and x tend to 0, leading to the indeterminate form $\frac{0}{0}$.

Based on differentiation, l'Hôpital's Rule is a very useful and powerful technique for evaluating these kind of limits.

When the relevant limits exist, are defined and are nonzero if in the denominator, and the relevant functions are differentiable,

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

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

3. If $\frac{f'(x)}{g'(x)} \to \infty$ as $x \to x_0$ then $\frac{f(x)}{g(x)} \to \infty$ as $x \to x_0$.

4. If $\frac{f'(x)}{g'(x)} \to -\infty$ as $x \to x_0$ then $\frac{f(x)}{g(x)} \to -\infty$ as $x \to x_0$.

We can also use the same rules to calculate left-hand and right-hand limits.

Example 4.8.2 1. Let $f(x) = \sin(x)$ and let $g(x) = x$. Both functions are differentiable. $f(0) = g(0) = 0$ and $g'(0) = 1 \neq 0$. Thus

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \frac{\cos(0)}{1} = 1.$$

2. Find

$$\lim_{x\to 0} \frac{\cos(x) - 1}{x}$$

Let $f(x) = \cos(x) - 1$ and let $g(x) = x$. $f(0) = g(0) = 0$ so by l'Hôpital's rule we have

$$\lim_{x\to 0} \frac{\cos(x) - 1}{x} = \frac{-\sin(0)}{1} = 0$$

3. Alternatively, if we assume the above limits we can differentiate $\sin(x)$ from first principles.

Let $f(x) = \sin(x)$. Then

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\sin(x+h) - \sin(x)}{h} \\ &= \lim_{h\to 0} \frac{\sin(x)\cos(h) - \cos(x)sin(h) - \sin(x)}{h} \\ &= \lim_{h\to 0} \frac{\sin(x)(\cos(h) - 1) + \cos(x)\sin(h)}{h} \\ &= \lim_{h\to 0} \left(\sin(x).\frac{\cos(h) - 1}{h}\right) + \lim_{h\to 0} \left(\cos(x).\frac{\sin(h)}{(h)}\right) \\ & \quad \text{(by the sum rule for limits)} \\ &= \sin(x).0 + \cos(x).1 \text{ (by the product rule for limits)} \\ &= \cos(x). \end{aligned}$$

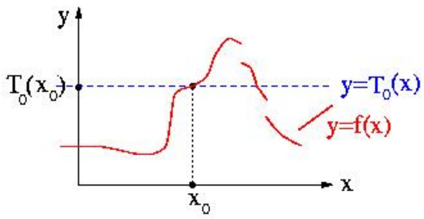
There are also forms of l'Hôpital's rule to deal with (a) limits of quotients of functions which both tend to ∞ (b) limits of quotients as $x \to \infty$.

4.9 Taylor Series

Polynomials are easy to differentiate and integrate. It is easy to find limits of rational functions (quotients of polynomials). We would therefore like to be able to approximate functions by polynomials. Recall that a polynomial is an expression of the form

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

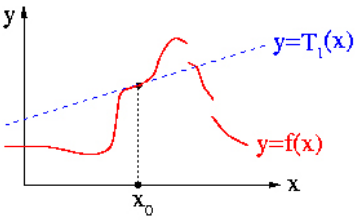
where $a_0, \ldots, a_n$ are constants. The simplest case is to take $T_0(x) = a_0$, i.e. $T_0(x)$ is a constant function with value a_0.



Clearly we have not taken much of the behaviour of f into account. If f is also **differentiable** at x_0 then we can improve our approximation by defining

$$T_1(x) = f_0 + (x - x_0)f'(x_0)$$

to be the **tangent line** to f at x_0.



The constant function T_0 is a degree 0 polynomial whose graph passes through $f(x_0)$. The tangent line is a degree 1 polynomial whose graph passes through $f(x_0)$, and whose derivative matches that of f at x_0. The idea of the n^{th} **Taylor polynomial** T_n is that it continues this sequence of approximations.

Example 4.9.1 1. Suppose that $f(x) = e^x$ and we want a degree 2 polynomial (quadratic) $T_2(x)$ satisfying

$$\begin{aligned} T_2(0) &= f(0) = 1 \\ T_2'(0) &= f'(0) = 1 \\ T_2''(0) &= f''(0) = 1 \end{aligned}$$

Let

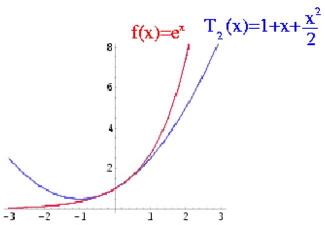
$$T_2(x) = a_2x^2 + a_1x + a_0.$$

Then we have

$$\begin{aligned} T_2'(x) &= 2a_2x + a_1 \\ T_2''(x) &= 2a_2 \end{aligned}$$

These equations hold for all x, in particular for $x = 0$. Substituting $x = 0$ in every equation gives $a_0 = 1$, $a_1 = 1$ and $a_2 = \frac{1}{2}$. We therefore have

$$T_2(x) = 1 + x + \frac{x^2}{2}$$



If you like think of this as the "tangent quadratic" to f at 0, although "tangent" now means that the second derivative of the quadratic also matches that of e^x. The name of this quadratic is **the second Taylor polynomial** of f at 0.

2. What is the second Taylor polynomial T_2 to f at 2, i.e. the quadratic analogue of the tangent line at 2? Denote this also by

$$T_2(x) = a_2x^2 + a_1x + a_0$$

We require

$$\begin{aligned} T_2(2) &= f(2) = e^2 \\ T_2'(2) &= f'(2) = e^2 \\ T_2''(2) &= f''(2) = e^2 \end{aligned}$$

The reason the previous calculation was easy was because when we substituted $x = 0$, every expression involving x vanished. Substituting $x = 2$ will not have this effect. So we "cheat" and move the y-axis! Suppose that we let $y = x - 2$.

Then we still have $T_2(y) = a_2y^2 + a_1y + a_0$. Now, as in the previous example we have

$$\begin{aligned} T_2(y) &= a_2y^2 + a_1y + a_0 \\ T_2'(y) &= 2a_2y + a_1 \\ T_2''(y) &= 2a_2 \end{aligned}$$

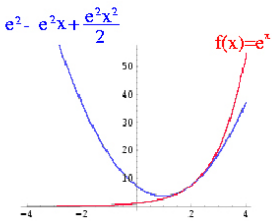
When $x = 2$, $y = 0$ so we have $2a_2 = T_2''$ when $y = 0$. This is equal to T_2'' when $x = 2$, i.e. equal to $f''(2) = e^2$. Thus $a_2 = e^2/2$. Similarly, $a_1 = e^2$ and $a_0 = e^2$. The second Taylor polynomial of f at $x = 2$ is hence given by

$$T_2(y) = e^2 + e^2y + \frac{e^2y^2}{2}$$

where $y = x - 2$.

We now have to write it in terms of x as

$$\begin{aligned} T_2(x) &= e^2 + e^2(x-2) + \frac{e^2(x-2)^2}{2} \\ &= e^2 - e^2 x + \frac{e^2x^2}{2} \end{aligned}$$



Let f be a function which is n-times differentiable. Then the n^{th} *Taylor polynomial* $T_n(x)$ at $a \in \mathbb{R}$ is the polynomial such that

$$T_n(a) = f(a), T_n'(a) = f'(a), \ldots, T_n^{(n)}(a) = f^{(n)}(a)$$

T_n is given by the following formula

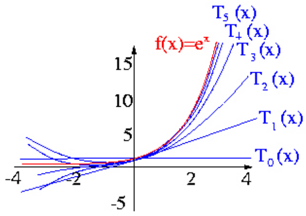
$$T_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

The proof is just an extension of the examples seen above.

Example 4.9.2 1. The n^{th} Taylor polynomial of e^x at 0 is

$$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

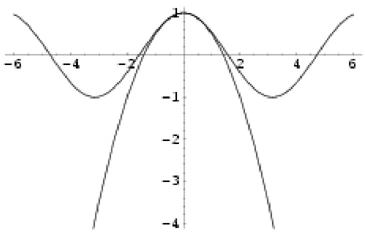
The graphs of the first few of these polynomials are as shown. Note that the higher the degree of the Taylor polynomial, the closer the approximation.

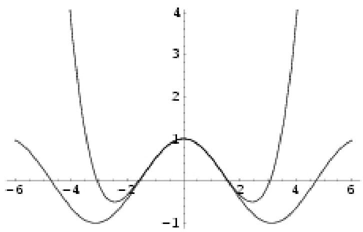


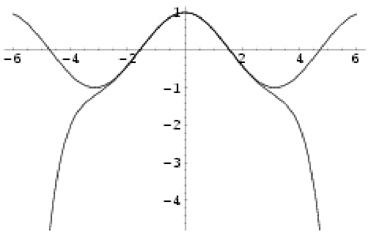
2. Suppose that we wish to calculate the n^{th} Taylor polynomial of $f(x) = \cos x$ about 0. We have $f'(x) = -\sin x$, $f''(x) = -\cos x$, $f^{(3)}(x) = \sin x$, $f^{(4)}(x) = \cos x$ etc. So if k is odd then $f^{(k)}(0) = 0$ and if k is even, say $k = 2m$, we have $f^{(k)}(0) = (-1)^m$. Thus

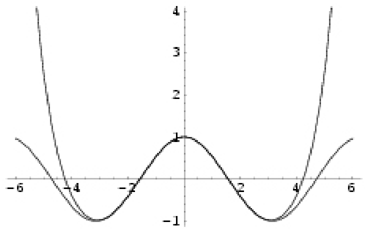
$$T_{2n}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + \frac{(-1)^n x^{2n}}{(2n)!}$$

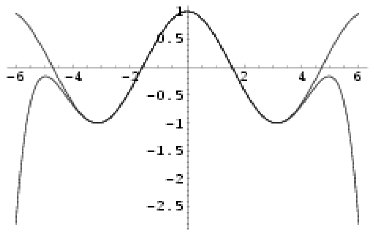
The graphs of T_{2n} are as shown for the first few values of n, along with the graph of cos:

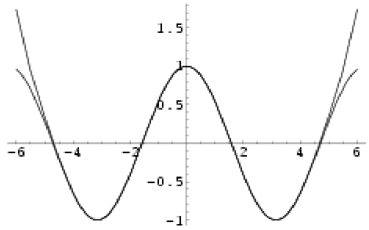





1
-6
-4
-2
2
4
6
-1
-2
-3
-4


4
3
2
1
-6
-4
-2
2
4
6
-1





Again, note that the higher the degree of Taylor polynomial, the better the approximation.

Exercise 4.9.3 For both of the following functions f find the Taylor polynomials T_0, T_1, T_2 and T_3 about the given point a. Sketch f, T_0, T_1, T_2 and T_3 on the same graph.

(a) $f(x) = \sqrt{x+4},\ a = 0$ (b) $f(x) = \sin x,\ a = \frac{\pi}{4}$

We have now seen how to construct polynomials of arbitrarily high degree n, whose m^{th} derivatives $(0 \leqslant m \leqslant n)$ all match those of f at a given point a. We use these to define the following series.

Definition 4.9.4 *Let f be a function which is smooth on an open interval I which contains the point $a \in \mathbb{R}$. Then the* **Taylor series** *of f at a is*

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k &= f(a) + f'(a)(x-a) + \\ &= \frac{f''(a)}{2!}(x-a)^2 + \cdots \\ &\quad \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + \cdots\end{aligned}$$

If $a = 0$ then the above series is called a **MacLaurin series**.

We are interested in when this series defines a function from I to $\mathbb{R}$ and when it is equal to f. It defines such a function precisely when it converges everywhere on I.

Example 4.9.5 We can't calculate a Maclaurin series for $f(x) = \frac{1}{x}$ because $f(0)$ is not defined. However, we can calculate a Taylor series for f at $a = 1$. The k^{th} derivative of f is given by

$$f^{(k)}(x) = (-1)^k.k!.x^{-(k+1)}$$

(If this general formula seems difficult, calculate the first few derivatives and the pattern will become apparent). Hence $f^{(k)}(1) = (-1)^k.k!$. The Taylor series is then given by

$$f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \cdots + \frac{f^{(k)}(1)}{k!}(x-1)^k + \cdots$$

But we have

$$\frac{f^{(k)}(1)}{k!} = (-1)^k$$

So the Taylor series is

$$1 - (x-1) + (x-1)^2 - \cdots + (-1)^k(x-1)^k + \cdots$$

Notes

- We have only covered the basics of Taylor Series here. It is also important where the series are defined (i.e. converge to a limit).
- There is a theorem called Taylor's Theorem, which we won't state here, which gives an expression for the remainder in the approximation by polynomials.

Chapter 5

Calculus of One Variable Part 2: Integration

5.1 Summing Series

Let a_n be a real number for all $n \in \{0, 1, 2, \ldots\}$. We use the **sigma notation**

$$\sum_{n=M}^{N} a_n$$

for the sum

$$a_M + \cdots + a_N,$$

which is also called a *finite series*. We have the following formal rules for manipulating finite series. These follow directly from the rules of arithmetic.

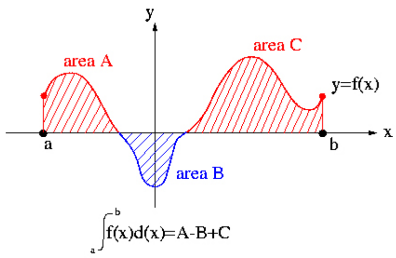
- $$\sum_{n=M}^{N}(a_n+b_n)=\sum_{n=M}^{N}a_n+\sum_{n=M}^{N}b_n$$
- $$\sum_{n=M}^{N}(a_n-b_n)=\sum_{n=M}^{N}a_n-\sum_{n=M}^{N}b_n$$
- $$\sum_{n=M}^{N}ca_n=c\sum_{n=M}^{N}a_n \ (c \text{ constant})$$
- $$\sum_{n=M}^{N}c=c(N-M) \ (c \text{ constant})$$
- $$\sum_{n=M}^{N}a_n=\sum_{n=1}^{N}a_n-\sum_{n=1}^{M-1}a_n$$

5.2 Integrals

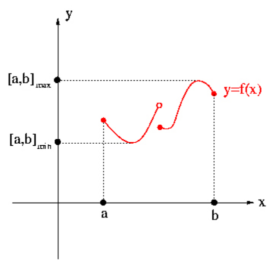
The notation

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

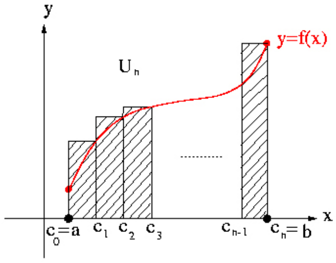
means the *signed* area under the curve $y = f(x)$ (i.e. area above the x-axis is positive and area below the x-axis is negative.)



Let f be a real function which is bounded on any closed interval, e.g. a continuous function. and let a and b be real numbers with $a \leqslant b$. Let $[a,b]_{\max}$ denote the maximum value of f on $[a,b]$ and let $[a,b]_{\min}$ denote the minimum value.



We take two approximations to the area. Let n be a positive integer and divide $[a,b]$ into n equal intervals. First take the following shaded area.

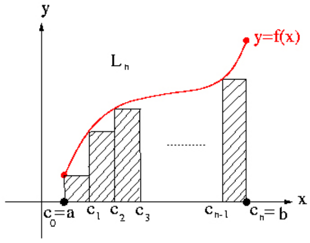


Let Δ denote the quantity $\frac{b-a}{n}$, i.e. the width of a small strip. Then $c_k = a + k\Delta$ and the shaded area, which we call U_n, is given by

$$\sum_{k=1}^{n} \Delta[c_{k-1}, c_k]_{\max} = \sum_{k=1}^{n} \Delta[a + (k-1)\Delta, a + k\Delta]_{\max}$$

This is called the n^{th} *upper Riemann sum.*

Secondly, take the following shaded area.



This is called the n^{th} *lower Riemann sum*, denoted L_n, and is given by

$$\sum_{k=1}^{n} \Delta[c_{k-1}, c_k]_{\min} = \sum_{k=1}^{n} \Delta[a + (k-1)\Delta, a + k\Delta]_{\min}$$

If for all a and b in $\mathbb{R}$, $\lim_{n\to\infty} U_n$ and $\lim_{n\to\infty} L_n$ exist and are equal then we say that f is *(Riemann) integrable* and define the *definite integral*

$$\int_a^b f(x)dx = \lim_{n\to\infty} U_n = \lim_{n\to\infty} L_n$$

The function f is called the *integrand*, x is called the *variable of integration* and a and b are called the **lower** and *upper limits of integration.*

Note that the required definition of *limit* is that of the limit of a *sequence.*

It can be shown that continuous functions are always integrable but so are many discontinuous functions.

Example 5.2.1 Evaluate, from first principles,

$$\int_0^b x^2 dx$$

$f(x) = x^2$ is continuous so it is integrable. We may choose either the upper or lower Riemann sum to evaluate it, and we choose the upper sum. We have $\Delta = b/n$ and

$$\begin{aligned} U_n &= \sum_{k=1}^{n} \frac{b}{n}.\left(\frac{kb}{n}\right)^2 \\ &= \frac{b^3}{n^3}\sum_{k=1}^{n} k^2 \text{ (since } b \text{ and } n \text{ are constants)} \\ &= \frac{b^3}{n^3}\frac{n(n+1)(2n+1)}{6} \text{ (by the last section)} \\ &= \frac{b^3}{6}\left(2+\frac{3}{n}+\frac{1}{n^2}\right) \end{aligned}$$

Hence

$$\begin{aligned} \int_0^b x^2dx &= \lim_{n\to\infty}\sum_{k=1}^{n}\frac{b}{n}\left(\frac{kb}{n}\right)^2 \\ &= \frac{b^3}{6}\lim_{n\to\infty}\left(2+\frac{3}{n}+\frac{1}{n^2}\right) \\ &= \frac{b^3}{6}.2 \\ &= \frac{b^3}{3} \end{aligned}$$

5.3 Antiderivatives

Let f be a real function. A differentiable real function F is an **antiderivative** or **indefinite integral** of f if for all x, $F'(x) = f(x)$.

Antiderivatives are unique, up to possibly adding a constant function.

We write

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

to mean that $F(x)$ is the unique derivative of x, up to possibly adding a constant C. Note that this is the same notation as used for the area under a curve. This is because of the *fundamental theorem of calculus* which states formally that integration (finding areas under curves) and differentiation (finding gradients) are inverse operations of each other.

To find antiderivatives there are various techniques. The simplest is to reverse the formulae for derivatives.

Example 5.3.1

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

since

$$\frac{d}{dx}\left(\frac{x^{n+1}}{n+1}\right) = \frac{n+1}{n+1}x^n = x^n.$$

Using this process, we obtain the following list of standard integrals:

$f(x)$	$F(x) = \int f(x)dx$
$\cos x$	$\sin x + C$
$\sin x$	$-\cos x + C$
$\sec^2 x$	$\tan x + C$
$\mathrm{cosec}^2 x$	$-\cot x + C$
$\sec x \tan$	$\sec x + C$
$\mathrm{cosec} x \cot x$	$-\cot x + C$
e^x	$e^x + C$
$\frac{1}{x}\ (x > 0)$	$\log x + C$

Certain rules hold for antiderivatives; these come from the corresponding rules for derivatives.

$$\begin{aligned} \int kf(x) &= k\int f(x)dx \text{ (for a constant } k) \\ \int (f(x)+g(x))dx &= \int f(x)dx + \int g(x)dx \end{aligned}$$

From these we can deduce that

$$\begin{aligned} \int -f(x)dx &= -\int f(x)dx \text{ and} \\ \int (f(x)-g(x))dx &= \int f(x)dx - \int g(x)dx \end{aligned}$$

The product rule for differentiation tells us that

$$\int (f'(x)g(x) + f(x)g'(x))dx = f(x)g(x) + C$$

To deal with integrals of products we require *integration by parts* (later). Similarly,

$$\int \frac{f(x)}{g(x)}dx$$

does not have a simple formula.

The technique which corresponds to operating the chain rule backwards is called *integration by substitution* in the next section, which will deal with integrals of the form

$$\int f'(x)g'(f(x))dx.$$

Examples 5.3.2 1. Using the product rule for differentiation backwards,

$$\int (x\text{sec}^2 x + \text{tan}x)dx = x\tan x + C$$

2. Using the chain rule for differentation backwards,

$$\int 2x\cos(x^2)dx = \sin(x^2) + C$$

5.4 Integration by Substitution

We saw in the last section that the chain rule may be applied in reverse to find antiderivatives. It can be difficult to do this by inspection. A good way of keeping track is to use *substitution.* For example, suppose that we wish to integrate

$$\int x^8(x^9+1)^2 dx$$

We recognise that the integrand

$$x^8(x^9+1)^2$$

looks like the derivative of

$$(x^9+1)^3.$$

So we substitute $u = x^9 + 1$ in the integral.

The important thing to notice is that we now have to integrate with respect to u and not x. So we calculate

$$\frac{du}{dx} = 9x^8 \Longrightarrow dx = \frac{du}{9x^8}$$

Which gives

$$\begin{aligned}
\int x^8(x^9+1)^2 dx &= \int x^8 u^2 \frac{du}{9x^8} \\
&= \frac{1}{9}\int u^2 du \\
&= \frac{1}{9}.\frac{u^3}{3} + C \\
&= \frac{1}{27}u^3 + C \\
\text{substituting back,} \quad &= \frac{1}{27}(x^9+1)^3 + C
\end{aligned}$$

Always check the answer by differentiating again.

$$\frac{d}{dx}\left(\frac{1}{27}(x^9+1)^3 + C\right) = \frac{1}{27}.3(x^9+1)^2.9x^8 + 0,$$

which is correct.

Formally, we are using the following rule, which is essentially the chain rule.

$$\int f(g(x))g'(x)dx = \int f(u)du \text{ where } u = g(x).$$

Substitution is a versatile technique - we now look at some more examples.

Example 5.4.1 1.

$$\int \sin x \; e^{\cos x} dx$$

Let $u = \cos x$. Then

$$\frac{du}{dx} = -\sin x \Longrightarrow dx = -\frac{du}{\sin x}$$

This gives

$$\begin{aligned}\int \sin x \; e^{\cos x} dx &= \int \sin x \; e^{u}. \left(-\frac{du}{\sin x}\right)\\ &= -\int e^{u} du\\ &= -e^{u} + C\\ &= -e^{\cos x} + C\end{aligned}$$

2. Sometimes there are different substitutions we could choose, e.g.

$$\int \sec^2 x \tan^2 x \; dx$$

Let $u = \tan x$ which gives

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

and we have

$$\int u^2 du = \frac{u^3}{3} + C = \frac{1}{3}\tan^3 x + C$$

On the other hand, for

$$\int \sec^2 x \tan x \; dx$$

Let $u = \sec x$ which gives

$$dx = \frac{du}{\sec x \tan x}$$

and the integral becomes

$$\int u \; du = \frac{u^2}{2} + C = \frac{1}{2}\sec^2 x + C$$

3. Sometimes it is easier to make more than one substitution, e.g.

$$\int \frac{dx}{x\sqrt{(x+1)}}$$

Let $u = x + 1$. Then $dx = du$ which gives

$$\int \frac{dx}{x\sqrt{(x+1)}} = \int \frac{du}{(u-1)\sqrt{u}}$$

Then let $u = v^2$ to get rid of the square root. This gives $du = 2v\ dv$ and we have

$$\begin{aligned}\int \frac{du}{(u-1)\sqrt{u}} &= \int \frac{2v\ dv}{(v^2-1).v} \\ &= \int \frac{2dv}{(v+1)(v-1)} \\ &= \int \left(\frac{-1}{v+1} + \frac{1}{v-1} \right) dv \\ &= -\log(v+1) + \log(v-1) + C \\ &= \log\left(\frac{v-1}{v+1} \right) + C.\end{aligned}$$

The identity

$$\frac{2}{(v+1)(v-1)} = \frac{-1}{v+1} + \frac{1}{v-1}$$

is obtained using *partial fractions*.

5.5 Partial Fractions

The method of *partial fractions* is an algebraic technique which can be thought of as "finding a common denominator in reverse". It allows us to express rational functions as sums of simpler rational functions.

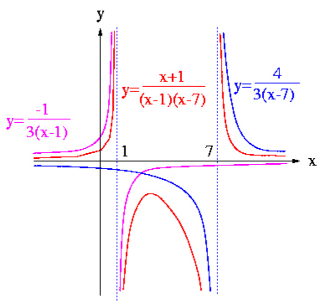
Example 5.5.1 1. Find real numbers A and B such that for all x we have

$$\frac{x+1}{(x-1)(x-7)} = \frac{A}{x-1} + \frac{B}{x-7}$$

Multiply throughout by (x-1)(x-7) to get $x+1 = A(x-7) + B(x-1)$. Let $x = 7$ to make the coefficient of A equal to 0. This gives $8 = 6B$ and so $B = \frac{4}{3}$. Let $x = 1$ to make the coefficient of B equal to 0. This gives $2 = -6A$. Thus $A = -\frac{1}{3}$. Hence

$$\frac{x+1}{(x-1)(x-7)} = \frac{-1}{3(x-1)} + \frac{4}{3(x-7)}$$

You can check that the expression is correct by finding a common denominator on the right hand side.

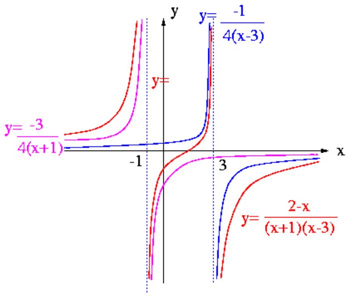


2. Find A and B such that

$$\frac{2-x}{(x+1)(x-3)} = \frac{A}{x+1} + \frac{B}{x-3}$$

We have $2 - x = A(x-3) + B(x+1)$. $x = 3$ gives $B = -\frac{1}{4}$ and $x = -1$ gives $A = -\frac{3}{4}$. Thus

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



The main application of partial fractions is to integration of rational functions.

Example 5.5.2 Find

$$\int \frac{2-x}{(x+1)(x-3)} dx$$

Using expression as partial fractions the integral becomes

$$\begin{aligned} -\int \left(\frac{3}{4(x+1)} + \frac{1}{4(x-3)} \right) dx &= -\frac{3}{4} \int \frac{dx}{x+1} - \frac{1}{4} \int \frac{dx}{x-3} \\ &= -\frac{3}{4} \log|x+1| - \frac{1}{4} \log|x-3| + C \end{aligned}$$

Note the modulus signs in the log expressions. If x can take values less than 0 then

$$\int \frac{dx}{x} = \log|x| + C.$$

There are various other possibilities when finding partial fractions.

Example 5.5.3 1. (Repeated factors in the denominator)

$$\frac{6x+7}{(x+2)^2} = \frac{A}{x+2} + \frac{B}{(x+2)^2}$$

This gives $6x + 7 = A(x + 2) + B = Ax + 2A + B$.

Coefficients of x give $A = 6$. Coefficients of 1 give $B = -5$.
Thus

$$\frac{6x+7}{(x+2)^2} = \frac{6}{x+2} - \frac{5}{(x+2)^2}$$

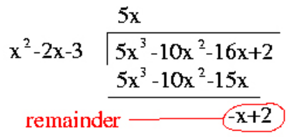
and

$$\begin{aligned}\int \frac{6x+7}{(x+2)^2}dx &= 6\int \frac{dx}{x+2}dx - 5\int \frac{dx}{(x+2)^2} \\ &= 6\log|x+2| + \frac{5}{x+2} + C\end{aligned}$$

2. (Improper fractions) How do we express

$$\frac{5x^3 - 10x^2 - 16x + 2}{x^2 - 2x - 3}$$

as partial fractions? First we use long division.



Hence we have

$$\begin{aligned}\frac{5x^3-10x^2-16x+2}{x^2-2x-3} &= 5x+\frac{2-x}{x^2-2x-3}\\ &= 5x+\frac{2-x}{(x+1)(x-3)}\\ &= 5x-\frac{3}{4(x+1)}-\frac{1}{4(x-3)}\end{aligned}$$

Using the expression from the second example.

Now we can integrate term-by-term as before.

3. (Irreducible quadratics in the denominator) Integrate

$$\int \frac{(x+3)}{(x^2+1)(x-1)}dx$$

We have to write the partial fractions expression as

$$\frac{x+3}{(x^2+1)(x-1)}=\frac{Ax+B}{x^2+1}+\frac{C}{x-1}$$

This gives $(Ax+B)(x-1)+C(x^2+1)=x+3$, i.e.

$$(A+C)x^2+(B-A)x+C-B=x+3$$

Comparing coefficients of x^2, x and 1 we obtain the following system of linear equations.

$$A+C=0 \quad B-A=1 \quad C-B=3$$

which has solution $A=-2$, $B=-1$ and $C=2$. Hence

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

and the integral becomes

$$\begin{aligned}\int\left(\frac{-2x-1}{x^2+1}+\frac{2}{x-1}\right)dx &= -\int\frac{2xdx}{x^2+1}-\int\frac{dx}{x^2+1} \\ &\quad +2\int\frac{dx}{x-1} \\ &= -\log(x^2+1)-\tan^{-1}x \\ &\quad +2\log|x-1|+C\end{aligned}$$

5.6 Integration by Parts

Integration by parts is a substitute for having no product rule for integration. Let u and v be functions of the real variable x. The product rule for differentiation is

$$(uv)' = uv' + u'v$$

Which we can integrate to get

$$uv = \int uv'dx + \int u'vdx$$

or

$$\int u'vdx = uv - \int uv'dx$$

Example 5.6.1 Find

$$\int xe^x dx$$

Let $u(x) = e^x$ and $v(x) = x$. Then

$$\int xe^x dx = xe^x - \int e^x.1dx = xe^x - e^x + C = (x-1)e^x + C$$

If we had done it the other way, i.e. $u(x) = x$ and $v(x) = e^x$ we would have got

$$\frac{x^2}{2}e^x - \frac{1}{2}\int x^2e^x dx$$

which is worse!

Sometimes we have to apply the formula more than once.

Example 5.6.2

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} x^2 \cos x dx &= \left[x^2 \sin x\right]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} 2x \sin x dx \\
&= \frac{\pi^2}{4} - 2\int_0^{\frac{\pi}{2}} x \sin x dx \\
&= \frac{\pi^2}{4} - \left(2\left[-x\cos x\right]_0^{\frac{\pi}{2}} - 2\int_0^{\frac{\pi}{2}} (-\cos x).1 dx\right) \\
&= \frac{\pi^2}{4} + 2(\frac{\pi}{2}.0 - 0.1) - 2\int_0^{\frac{\pi}{2}} \cos x dx \\
&= \frac{\pi^2}{4} - 2\left[\sin x\right]_0^{\frac{\pi}{2}} \\
&= \frac{\pi^2}{4} - 2
\end{aligned}$$

Some integrations by parts use the periodicity of functions.

Example 5.6.3

$$\begin{aligned}
\int e^x \sin x dx &= e^x \sin x - \int e^x \cos x dx \\
&= e^x \sin x - \left(e^x \cos x + \int e^x \sin x dx\right) \\
&= e^x \sin x - e^x \cos x - \int e^x \sin x dx
\end{aligned}$$

Rearranging,

$$2\int e^x \sin x dx = e^x \sin x - e^x \cos x + K$$

which gives

$$\int e^x \sin x dx = \frac{1}{2}\left(e^x \sin x - e^x \cos x\right) + C$$

Using the "trick" of writing $\log x = 1.\log x$, $\sin^{-1} x = 1.\sin^{-1} x$ etc.
we may integrate these functions.

Example 5.6.4 1.

$$\begin{aligned}\int \log x dx &= \int 1.\log x dx \\ &= x\log x - \int x.\frac{1}{x}dx \\ &= x\log x - \int dx \\ &= x\log x - x + C\end{aligned}$$

2.

$$\begin{aligned}\int \sin^{-1} x dx &= \int 1.\sin^{-1} x dx \\ &= x\sin^{-1} x - \int \frac{x}{\sqrt{(1-x^2)}}dx \\ &= x\sin^{-1} x + \sqrt{(1-x^2)} + C\end{aligned}$$

5.7 Reduction Formulae

Integration by parts becomes tedious for integrals with large powers e.g.

$$\int x^6 \cos x dx \quad \int x^{259} e^x dx$$

For such integrals it is better to use a *reduction formula*. These are best understood by example. Suppose that we want an expression for

$$I_n = \int x^n e^x dx$$

in terms of

$$I_{n-1} = \int x^{n-1} e^x dx.$$

We first use integration by parts to derive the formula, i.e.

$$\begin{aligned} I_n &= \int x^n e^x dx \\ &= x^n e^x - \int n x^{n-1} e^x dx \\ &\quad \text{(let } u = e^x \text{ and } v = x^n\text{)} \\ &= x^n e^x - n \int x^{n-1} e^x dx \\ &= x^n e^x - n I_{n-1} \end{aligned}$$

Now we can use the formula to integrate, say

$$\int x^3 e^x dx$$

We have

$$\begin{aligned} I_3 &= x^3 e^x - 3I_2 \\ &= x^3 e^x - 3(x^2 e^x - 2I_1) \\ &= x^3 e^x - 3x^2 e^x + 6(x e^x - I_0) \\ &= x^3 e^x - 3x^2 e^x + 6x e^x - 6I_0 \end{aligned}$$

But $I_0 = \int e^x dx = e^x + C$. Hence

$$I_3 = x^3 e^x - 3x^2 e^x + 6x e^x - 6e^x + C$$

We can also derive reduction formulae for definite integrals.

Example 5.7.1 Find a reduction formula for

$$I_n = \int_0^1 x^n e^{-x} dx$$

We have

$$\begin{aligned} I_n &= \left[-e^{-x}x^n\right]_0^1 - \int_0^1 nx^{n-1}.(-e^{-x})dx \\ &= -\frac{1}{e} + n\int_0^1 x^{n-1}e^{x}dx \\ &= -\frac{1}{e} + nI_{n-1} \end{aligned}$$

Thus, for example,

$$\begin{aligned} \int_0^1 x^4 e^{-x} dx &= I_4 \\ &= -\frac{1}{e} + 4I_3 \\ &= -\frac{1}{e} + 4\left(-\frac{1}{e} + 3I_2\right) \\ &= -\frac{5}{e} + 12\left(-\frac{1}{e} + 2I_1\right) \\ &= -\frac{17}{e} + 24\left(-\frac{1}{e} + I_0\right) \\ &= -\frac{41}{e} + 24I_0 \end{aligned}$$

but

$$I_0 = \int_0^1 e^{-x} dx = -\frac{1}{e} + 1$$

Hence

$$\int_0^1 x^4 e^{-x} dx = -\frac{41}{e} + 24\left(\frac{1}{e} + 1\right) = 24 - \frac{65}{e}$$

Note that this is a positive value. I_n should always be positive because x^n and e^x are both positive between 0 and 1.

Sometimes a reduction formula does not reduce to I_{n-1} but to I_{n-2}. This is what usually happens when integrate x^n multiplied by a trigonometric function.

Example 5.7.2 Find a reduction formula for

$$\int x^n \cos x \, dx$$

Let $I_n = \int x^n \cos x \, dx$. Then

$$\begin{aligned}
I_n &= x^n \sin x - \int \sin x . n x^{n-1} dx \\
&= x^n \sin x - n x^{n-1} .(-\cos x) + n \int \cos x .(n-1) x^{n-2} dx \\
&= x^n \sin x + n x^{n-1} \cos x - n(n-1) \int x^{n-2} \cos x \, dx \\
&= x^n \sin x + n x^{n-1} \cos x - n(n-1) I_{n-2}
\end{aligned}$$

Thus, for example,

$$\begin{aligned} I_4 &= x^4\sin x + 4x^3\cos x - 4.3I_2 \\ &= x^4\sin x + 4x^3\cos x - 12(x^2\sin x + 2x\cos x - 2I_0) \\ &= x^4\sin x + 4x^3\cos x - 12x^2\sin x - 24x\cos x + 24\int \cos x \ dx \\ &= x^4\sin x + 4x^3\cos x - 12x^2\sin x - 24x\cos x + 24\sin x + C \end{aligned}$$

5.8 Improper Integrals

Question

Evaluate the integral

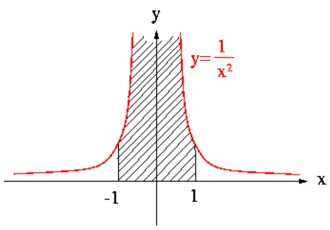
$$\int_{-1}^{1} \frac{1}{x^2} dx$$

Answer

Most of you probably did this:

$$\begin{aligned}\int_{-1}^{1} \frac{1}{x^2} dx &= \left[-\frac{1}{x}\right]_{-1}^{1} \\ &= \frac{-1}{1} - \frac{-1}{-1} \\ &= -2\end{aligned}$$

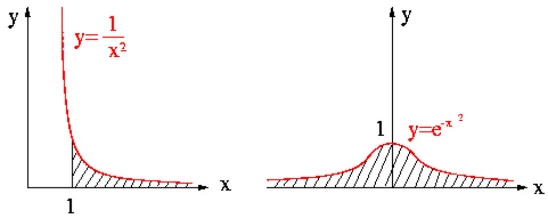
However, consider the following graph. All of the area under the curve is *above* the x-axis so we should get a positive answer.



What has gone wrong?

There are two types of integral which come under the name of *improper integrals*. One is an integral in which one or both of the limits of integration are infinite, e.g.

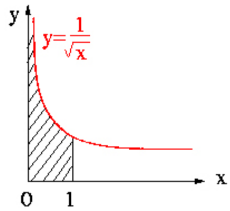
$$\int_1^\infty \frac{1}{x^2} \quad \text{or} \quad \int_{-\infty}^{\infty} e^{-x^2} dx$$



In such an integral we are integrating over an unbounded interval.

The other type is where we integrate over a bounded interval but the function which we are integrating is unbounded on this interval, e.g.

$$\int_0^1 \frac{1}{x^{1/2}} dx$$



These areas may or may not be finite.

Definition 5.8.1 Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function. We define

$$\int_a^\infty f(x)\, dx = \lim_{b \to \infty} \int_a^b f(x)\, dx$$

if this limit exists. In this case we say that the integral **converges**. If the limit doesn't exist, we say that the integral **diverges**.

Example 5.8.2 1.

$$\int_1^\infty \frac{1}{x^2} dx$$

First evaluate

$$\begin{aligned}\int_1^b \frac{1}{x^2}dx &= -\left[\frac{1}{x}\right]_1^b \\ &= -\left(\frac{1}{b} - 1\right) \\ &= 1 - \frac{1}{b}\end{aligned}$$

Now take the limit as $b \to \infty$.

$$\begin{aligned}\int_1^\infty \frac{1}{x^2}dx &= \lim_{b\to\infty}\int_1^b \frac{1}{x^2}dx \\ &= \lim_{b\to\infty}\left(1 - \frac{1}{b}\right) \\ &= 1\end{aligned}$$

2. What is

$$\int_1^\infty \frac{1}{x}dx?$$

We have

$$\begin{aligned}\int_1^b \frac{1}{x}dx &= [\log x]_1^b \\ &= \log b - \log 1 \\ &= \log b\end{aligned}$$

Since $\log b \to \infty$ as $b \to \infty$, the integral diverges.

Now we consider the second type of improper integral.

Definition 5.8.3 If f is integrable on $(a, b]$ (possibly undefined at a) then we define

$$\int_a^b f(x)\ dx = \lim_{c \to a^+} \int_c^b f(x)\ dx$$

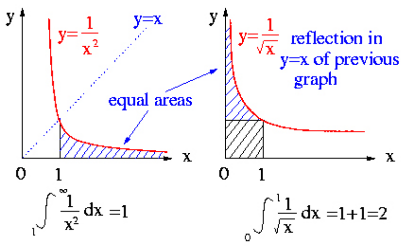
whenever it exists. If f is integrable on $[a, b)$ (possibly undefined at b) then we define

$$\int_a^b f(x)\ dx = \lim_{c \to b^-} \int_a^c f(x)\ dx$$

whenever it exists.

Example 5.8.4 1. Here is a convergent improper integral of this type.

$$\begin{aligned}\int_0^1 \frac{1}{\sqrt{x}}dx &= \lim_{c \to 0^+} \int_c^1 \frac{1}{\sqrt{x}}dx \\ &= \lim_{c \to 0^+} \left[2\sqrt{x}\right]_c^1 \\ &= \lim_{c \to 0^+} \left(2 - 2\sqrt{c}\right) \\ &= 2\end{aligned}$$



2. Here is one which is divergent.

$$\int_0^1 \frac{1}{x^2} dx$$

We have

$$\begin{aligned}\int_c^1 \frac{1}{x^2} dx &= \left[-\frac{1}{x}\right]_c^1 \\ &= -1 - \frac{1}{c} \\ &\rightarrow \ -\infty \text{ as } c \rightarrow \ 0^+\end{aligned}$$

We can also integrate from $-\infty$ to ∞. First, we define the integral

$$\int_{-\infty}^b f(x)\ dx = \lim_{a \to -\infty} \int_a^b f(x)\ dx$$

Definition 5.8.5 If f is integrable then we define

$$\int_{-\infty}^{\infty} f(x)\ dx = \int_{-\infty}^0 f(x)\ dx + \int_0^{\infty} f(x)\ dx$$

if *both* integrals on the R.H.S. exist.

We can similarly define $\int_a^b f(x)\ dx$ when there is a single point c in $[a, b]$ at which f is undefined. We define the value of the integral in this case to be

$$\int_a^c f(x)\ dx + \int_c^b f(x)\ dx$$

But both integrals must converge. The reason we got a nonsensical answer at the start of the lecture was because neither integral did converge. We were using an antiderivative which we weren't allowed to, because it wasn't defined at 0. Antiderivatives must be defined for every point in the interval we integrate over.

Chapter 6

Calculus of Many Variables

Calculus of many variables concerns functions $f : \mathbb{R}^n \to \mathbb{R}^m$ between sets of vectors. Two important types of functions which fall into this category are

- *scalar fields* which are functions $f : \mathbb{R}^n \to \mathbb{R}$
- *vector fields* which are functions $f : \mathbb{R}^n \to \mathbb{R}^n$

This section generalises the single-variable calculus of functions $f : \mathbb{R} \to \mathbb{R}$ and involves a mixture of techniques from single-variable calculus and from linear algebra, the theory of vectors and matrices. In each case when we express derivatives and partial derivatives of these functions we will assume that the functions are sufficiently differentiable.

6.1 Surfaces and Partial Derivatives

The simplest case is a function $f : \mathbb{R}^2 \to \mathbb{R}$. Geometrically, this can be viewed as a surface $z = f(x, y)$. Given such a surface we can set y to be a constant y_0, and this gives a function $f(x)$, which geometrically represents the cross section of the surface by the plane $y = y_0$. The gradient of this function f is then given by $\frac{d}{dx}f(x)$. Partial differentiation allows us to write down a function of x and y such that this gradient is given at any point (x, y). All we have to do is treat y as if it were a constant and differentiate $f(x, y)$ with respect to x to obtain a function denoted by $\frac{\partial f}{\partial x}$ or $f_x(x, y)$, the *partial derivative* of f with respect to x. Similarly we can define $\frac{\partial f}{\partial y} = f_y(x, y)$.

Example 6.1.1 If $f(x, y) = x^2y + 2y^2x^3$ then

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy + 6y^2x^2 \\ \frac{\partial f}{\partial y} &= x^2 + 4yx^3 \end{aligned}$$

6.2 Scalar Fields

A *scalar field*, i.e. a function $f : \mathbb{R}^n \to \mathbb{R}$, is something more general than the surfaces from the previous section, although if we could visualize more than three dimensions it would be a direct analogy of a surface. Suppose that a scalar field be given by $f(x_1, \ldots, x_n)$. Then we can define the *gradient* of f, which is denoted by $\underline{\nabla} f$ and is defined by

$$\underline{\nabla} f = (\frac{\partial f}{x_1}, \ldots, \frac{\partial f}{x_n}).$$

Or it may just be written as ∇f, but it is underlined here to remind us that it is a vector operator. The gradient is a *vector field*, that is a function $\mathbb{R}^n \to \mathbb{R}^n$. It has a geometric interpretation for $n = 2$:

> the value of $\underline{\nabla} f(x, y)$ at a point (x, y) is a vector pointing in the direction where the surface slopes the most steeply upwards

Example 6.2.1 If we take the function $f(x, y) = x^2y + 2y^2x^3$ as in the last section then

$$\underline{\nabla} f(x, y) = (2xy + 6y^2x^2, x^2 + 4yx^3)$$

You might like to try to visualise this surface. For example, at $(x, y) = (1, 1)$, $\underline{\nabla} f(x, y) = (8, 5)$. You might also like to consifer what happens at $(0, 0)$.

It is also possible to define the slope of the surface at (x, y) in any direction. Suppose that $\underline{\hat{u}}$ is a unit vector (i.e. $|\underline{\hat{u}}| = 1$). Then the slope of the surface at (x, y) in the direction of $\underline{\hat{u}}$ is called the *directional derivative* of f in the direction $\underline{\hat{u}}$. We write it as $\underline{\nabla}_{\underline{\hat{u}}} f(x, y)$ and it is given by

$$\underline{\nabla}_{\underline{\hat{u}}} f(x, y) = \underline{\nabla} f(x, y) \cdot \underline{\hat{u}}.$$

Example 6.2.2 if we take $\underline{\hat{u}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ then

$$\begin{aligned}\underline{\nabla}_{\hat{\mathbf{u}}} f(x,y) &= (2xy + 6y^2x^2, x^2 + 4yx^3) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \\ &= \frac{1}{\sqrt{2}}(2xy + 6y^2x^2 + x^2 + 4yx^3)\end{aligned}$$

6.3 Vector Fields

Recall that a vector field is a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$. It is more commonly denoted by

$$\underline{v} = (v_1(x_1, \dots, x_n), \dots, v_n(x_1, \dots x_n)).$$

We have already seen one example, the gradient of a scalar field. If a vector field $\underline{v}$ can be obtained as $\underline{v} = \underline{\nabla f}$ for some scalar field f then it is called a *conservative field* and f is called a *(scalar) potential* for f. It will be apparent later why such fields are called conservative, when we study many-variable integration.

We can view vector fields by showing the vector values on a grid

There are two commonly used differential operators on vector fields, the *divergence* and the *curl* of a vector field. The latter is defined in terms of the cross product and hence is only defined on vector fields $\underline{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

The *divergence* of a vector field $\underline{v}$ is a scalar field given by

div

$$\text{div } \underline{v} = \underline{\nabla} \cdot \underline{v} = \frac{\partial v_1}{x_1} + \cdots + \frac{\partial v_n}{x_n}$$

The *curl* of a vector field $\underline{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, which will be written as

$$(v_x(x, y, z), v_y(x, y, z), v_z(x, y, z))$$

is a vector field given by

$$\text{curl } \underline{v} = \underline{\nabla} \times \underline{v} = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}, \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}, \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right)$$

Curl has a nice interpretation. If you imagine the vector field as a fluid, then the curl of the vector field at a point is the axis a ball would spin around if it were inside the fluid at that point.

Example 6.3.1 Let $\underline{v} = (xy, xyz, y^2z)$. Then we have

$$\begin{aligned}\text{div } \underline{v} &= \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(xyz) + \frac{\partial}{\partial z}(y^2z) \\ &= y + xz + y^2\end{aligned}$$

and

$$\begin{aligned}\text{curl } \underline{v} &= \left(\frac{\partial}{\partial y}(y^2z) - \frac{\partial}{\partial z}(xyz), \frac{\partial}{\partial z}(xy) - \frac{\partial}{\partial x}(y^2z),\right. \\ &\left.\frac{\partial}{\partial x}(xyz) - \frac{\partial}{\partial y}(xy)\right) \\ &= (2yz - xy, 0, yz - x)\end{aligned}$$

6.4 Jacobians and The Chain Rule

This section concerns functions $\underline{f} : \mathbb{R}^n \to \mathbb{R}^m$ for arbitrary m and n. Such a function can be written as

$$\underline{f}(x_1, \cdots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$$

Given such a (differentiable) function we can write down an $m \times n$ matrix of partial derivatives called the *Jacobian matrix* of $\underline{f}$:

$$J\underline{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Recall the chain rule for functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, that

$$\frac{d}{dx}(g(f(x)) = \frac{dg}{dx}(f(x)) \cdot \frac{df}{dx}(x)$$

In a very aesthetically pleasing way, this translates into something about Jacobian matrices for functions $\underline{f} : \mathbb{R}^n \to \mathbb{R}^m$ where the product in the ordinary chain rule is replaced by a matrix product. Note that if $\underline{g} : \mathbb{R}^m \to \mathbb{R}^p$ then $\underline{g} \circ \underline{f} : \mathbb{R}^n \to \mathbb{R}^p$. Let $\underline{a}$ be a point in $\mathbb{R}^n$. Then

Theorem 6.4.1 (Multivariate Chain Rule)

$$J[\underline{g} \circ \underline{f}](\underline{a}) = J\underline{g}(\underline{f}(\underline{a}))J\underline{f}(\underline{a}).$$

Example 6.4.2 Let $\underline{f}(x, y) = (x + y, xy)$, $\underline{f} : \mathbb{R}^2 \to \mathbb{R}^2$, and let $g(x, y) = (x + y)^2$, $g : \mathbb{R}^2 \to \mathbb{R}$. Then

$$g \circ \underline{f}(x, y) = (x + y + xy)^2$$

We have

$$\begin{aligned} J[g \circ \underline{f}](x, y) &= \underline{\nabla}^T [g \circ \underline{f}](x, y) \\ &= (2x + 2y + 4xy + 2y^2 + 2xy^2, \\ &\quad 2x + 2y + 4xy + 2x^2 + 2yx^2) \end{aligned}$$

Note here that A^T is the *transpose* of a matrix, where the $(i, j)^{\text{th}}$ entry of A^T is the $(j, i)^{\text{th}}$ entry of A. We also have

$$J\underline{f} = \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix}$$

and

$$Jg(\underline{f}(x, y)) = (2(x + y + xy), 2(x + y + xy))$$

And multiplying the matrix $Jg(\underline{f}(x, y))$ on the right by $J(\underline{f})$ does indeed result in $J[g \circ \underline{f}](x, y)$, illustrating the chain rule.

6.5 Line Integrals

A *line integral* is an integral done on a vector field. It has a physical interpretation in the case where the field represents a force; the value of the line integral is the work against the vector field in moving from a point a to another point b. Line integrals are of the form

$$\int_a^b \underline{v} \cdot d\underline{r}$$

where $\underline{v}$ is a vector field, and $\underline{r}$ is a curve from a to b. It is not easy to work with in this form; usually the curve is parameterized as $\underline{r}(\lambda) = (x(\lambda), y(\lambda), z(\lambda))$ for $\lambda_a \leqslant \lambda \leqslant \lambda_b$. We can then write down a function $f : \mathbb{R} \to \mathbb{R}$

$$f(\lambda) = \underline{v}(x(\lambda), y(\lambda), z(\lambda)) \cdot (x(\lambda), y(\lambda), z(\lambda))$$

and treat it as a normal definite integral

$$\int_{\lambda_a}^{\lambda_b} f(\lambda) d\lambda$$

Example 6.5.1 Let $\underline{v}(x, y) = (y, x)$ and the curve C be parameterized by $\underline{r}(\lambda) = (\cos\lambda, \sin\lambda)$ for $0 \leqslant \lambda \leqslant \frac{\pi}{4}$. Then $\underline{r}'(\lambda) = (-\sin\lambda, \cos\lambda)$ and

$$\begin{aligned}
\int_C \underline{v} \cdot d\underline{r} &= \int_0^{\frac{\pi}{4}} (y, x) \cdot (-\sin\lambda, \cos\lambda) d\lambda \\
&= \int_0^{\frac{\pi}{4}} (\sin\lambda, \cos\lambda) \cdot (-\sin\lambda, \cos\lambda) d\lambda \\
&= \int 0^{\frac{\pi}{4}} (-\sin^2\lambda + \cos^2\lambda) d\lambda
\end{aligned}$$

Since $\cos^2\lambda - \sin^2\lambda = \cos 2\lambda$ we have

$$\begin{aligned}
\int_0^{\frac{\pi}{4}} \cos 2\lambda d\lambda &= \left[\frac{1}{2}\sin 2\lambda\right]_0^{\frac{\pi}{4}} \\
&= \frac{1}{2}(\sin\frac{\pi}{2} - \sin 0) \\
&= \frac{1}{2}
\end{aligned}$$

A special case is where $a = b$ and the integral is over a closed curve C it is called a *circulation integral*, and written as

$$\oint_C \underline{v} \cdot d\underline{r}$$

In the case where v is a conservative field, the value of the line integral only depends on the endpoints, not the curve $\underline{r}$. It follows that a circulation integral over a conservative field is always zero. More generally we have:

Green's Theorem: If C is a smooth closed curve in $\mathbb{R}^2$ and $v : \mathbb{R}^2 \to \mathbb{R}^2$ is a smooth vector field ($\underline{v} = (v_x, v_y)$) then

$$\oint_C \underline{v} \cdot \mathrm{d}\underline{r} = \iint\limits_R \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y$$

where R is the region enclosed by C.

6.6 Surface and Volume Integrals

Analogous to line integrals, we can also define *surface integrals* and *volume integrals* over a vector field. A surface integral looks like

$$\iint_S \underline{v} \cdot d\underline{S}$$

where v is a vector field and S is a smooth surface. d$\underline{\mathrm{S}}$ means the rate of change of the normal vector to the surface S. A normal vector to S at a point p on S is a vector perpendicular to the tangent plane to S at p. Like with line integrals, it helps if we have a of S parameterisation of S. A parameterisation of S takes the form

$$\underline{s}(\lambda, \mu) = (x(\lambda, \mu), y(\lambda, \mu), z(\lambda, \mu))$$

where λ and μ are real numbers. The normal is given by the cross product

$$\underline{n}(\lambda, \mu) = \underline{s}_\lambda(\lambda, \mu) \times \underline{s}_\mu(\lambda, \mu),$$

where

$$\begin{aligned} \underline{s}_\lambda &= \left(\frac{\partial x}{\partial \lambda}, \frac{\partial y}{\partial \lambda}, \frac{\partial z}{\partial \lambda}\right) \\ \underline{s}_\mu &= \left(\frac{\partial x}{\partial \mu}, \frac{\partial y}{\partial \mu}, \frac{\partial z}{\partial \mu}\right) \end{aligned}$$

We typically integrate over an area of the surface S demarcated by allowing λ and μ to vary within intervals $\lambda \in [\lambda_a, \lambda_b]$ and $\mu \in [\mu_a, \mu_b]$. Thus, as with a line integral we can rephrase the surface integral as some more easily understood real integrals

$$\int_{\lambda_a}^{\lambda_b} \int_{\mu_a}^{\mu_b} f(\lambda, \mu) d\mu d\lambda$$

where

$$f(\lambda, \mu) = \underline{v}(s(\lambda, \mu)) \cdot \underline{n}(\lambda, \mu)$$

The following theorem, Stokes' theorem, allows us to evaluate a surface integral of the curl of a vector field in terms of a line integral over the boundary of the surface. One word of warning. There are in fact two possible normal vectors. A choice which varies smoothly over the surface S is called an *orientation of S*. Not all surfaces possess an orientation. If you are interested then google “mobius band”. In the following theorem C must be oriented in the same way as S (otherwise there will be a sign error).

Theorem 6.6.1 (Stokes' Theorem) *If S is a smooth orientable surface with boundary C and $\underline{v}$ is a smooth vector field in $\mathbb{R}^3$, then*

$$\oint_C \underline{v} \cdot d\underline{r} = \iint_S (\nabla \times \underline{v}) \cdot d\underline{S}$$

A *volume integral* is simply a triple integral of some function $f : \mathbb{R}^3 \to \mathbb{R}^3$

$$\iiint_V f(x, y, z) dV = \int_{x_a}^{x_b} \int_{y_a}^{y_b} \int_{z_a}^{z_b} f(x, y, z) dz dy dx$$

The following theorem is usually used to evaluate a surface integral of a vector field $\underline{v}$ in terms of a volume integral of the divergence of $\underline{v}$. Again the orientations should be compatible.

Theorem 6.6.2 (The Divergence Theorem) *If S is a smooth closed surface which encloses a solid region R, and $\underline{v}$ is a smooth vector field in $\mathbb{R}^3$ then*

$$\iint_S \underline{v} \cdot d\underline{S} = \iiint_R (\nabla \cdot \underline{v})\, dV$$

Chapter 7

Ordinary Differential Equations

Differential equations involve a function and its derivatives. The solution to such an equation is not a number but a function. Differential equations are important in engineering. There are many kinds which can be solved algebraically but it is often not possible. Linear equations are often more easily solved than nonlinear equations. There are also ordinary differential equations (ODE) and partial differential equations (PDE). The solution to the former is a function in one variable. The latter contain partial derivatives and their solution is a function of more than one variable. We will not cover partial differential equations in this text as they are a topic in their own right.

7.1 First Order Differential Equations Solvable By Integrating Factor

A differential equation of the form

$$\frac{dy}{dx} + f(x)y = g(x)$$

123

can be solved for $y = f(x)$ by using what is called an *integrating factor*. This is given by

$$I(x) = e^{\int f(x)dx}$$

The required steps are as follows.

1. Calculate the integrating factor.
2. Multiply the differential equation by the integrating factor.
3. The left hand side will now be of the form $\frac{d}{dx}\left(I(x)y(x)\right)$
4. Rearrange to get $y(x) = \frac{1}{I(x)} \int g(x)I(x)dx$ and integrate.

This is quite a general method but of course it depends upon being able to algebraically integrate the functions $f(x)$ and $g(x)I(x)$, which is not always possible.

Example 7.1.1 Solve

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

We have $f(x) = 1$ and $g(x) = x$. The integrating factor is

$$e^{\int 1dx} = e^x.$$

(Ignore the constant of integration when calculating the integrating factor.) Multiplying the differential equation by the integrating factor gives

$$\begin{aligned} e^x \frac{dy}{dx} + ye^x &= e^{2x} \\ \frac{d}{dy}\left(ye^x\right) &= e^{2x} \\ ye^x &= \int e^{2x} dx \\ ye^x &= \frac{1}{2}e^{2x} + c \\ y &= e^{-x}\left(\frac{1}{2}e^{2x} + c\right) \\ &= \frac{1}{2}e^x + ce^{-x} \end{aligned}$$

Which we can check by substituting into the original differential equation.

7.2 First Order Separable Differential Equations

A first order *separable* differential equation is one of the form

$$\frac{dy}{dx} = f(x)g(y)$$

To solve we simply "rearrange" it and integrate both sides.

Example 7.2.1

$$\begin{aligned}\frac{dy}{dx} &= xy \\ \int \frac{dy}{y} &= \int x dx \\ \ln y &= \frac{x^2}{2} + c \\ y &= e^c e^{\frac{x^2}{2}} \\ &= K e^{\frac{x^2}{2}}\end{aligned}$$

Again you can check by substituting into the differential equation.

7.3 Second Order Linear Differential Equations with Constant Coefficients: The Homogenous Case

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0$$

These can be compared with quadratic equations in a sense. Their solutions are an interesting combination of what happens with first order differential equations with constant coefficients (where the solutions are typically exponential functions) and what happens with quadratic equations. As a quadratic equation has two roots, the equations in this subsection have two solutions, called *particular* solutions. The general solution is then a superposition of the particular solutions.

First we form what is called the *characteristic equation* which is the quadratic equation

$$p^2 + ap + b = 0$$

We solve this to get the roots p_1 and p_2. There are now three cases.

1. *Two Real Roots* p_1 *and* p_2

   The particular solutions are

$$\begin{aligned} y_1(x) &= e^{p_1 x} \\ y_2(x) &= e^{p_2 x} \end{aligned}$$

and the general solution is

$$Ay_1(x) + By_2(x) = Ae^{p_1 x} + Be^{p_2 x}.$$

2. *One Real Root* p

   The particular solutions are

$$\begin{aligned} y_1(x) &= e^{px} \\ y_2(x) &= pe^{px} \end{aligned}$$

   and the general solution is

$$Ay_1(x) + By_2(x) = Ae^{px} + Bpe^{px}.$$

3. *Complex Roots* $r \pm is$

   Note that because the equation has real coefficients, the complex roots are conjugates of each other. The particular solutions are

$$\begin{aligned} y_1(x) &= e^{rx}\cos(sx) \\ y_2(x) &= e^{rx}\sin(sx) \end{aligned}$$

   (these follow from Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$) and the general solution is

$$\begin{aligned} Ay_1(x) + By_2(x) &= Ae^{rx}\cos sx + Be^{rx}\sin sx \\ &= e^{rx}(A\cos sx + B\sin sx). \end{aligned}$$

Example 7.3.1 1.

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 2y = 0$$

The characteristic equation is

$$p^2 + p - 2 = 0$$

which has solutions $p = -2$ and $p = 1$ so the differential equation has general solution

$$y(x) = Ae^{-2x} + Be^{x}.$$

2.

$$\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9 = 0$$

The characteristic equation is $p^2 + 6p + 9 = 0$

which is $(p+3)^2 = 0$ giving $p = -3$ as a repeated root. So the differential equation has general solution

$$y(x) = Ae^{-3x} + Bxe^{-3x}.$$

3.

$$\frac{d^2y}{dx^2} + 4y = 0$$

(Undamped simple harmonic motion). The characteristic equation is

$$p^2 + 4 = 0$$

which has purely imaginary roots $p = 2i$ and $p = -2i$. So the differential equation has the general solution

$$y(x) = A\cos 2x + B\sin 2x.$$

4.
$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + 2 = 0$$

(Damped simple harmonic motion). The characteristic equation is

$$p^2 + 2p + 2 = 0$$

which has complex roots $-1 + i$ and $-1 - i$. This gives the general solution

$$y(x) = e^{-x}(A\cos x + B\sin x)$$

Notice that if such an equation has roots with positive real part, then the motion is forced, not damped.

7.4 Second Order Linear Differential Equations with Constant Coefficients: The Inhomogenous Case

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = f(x)$$

To solve this type of equation we

1. Solve the corresponding homogenous equation, that is when $f(x) = 0$, as in the previous section. This solution is called the *complementary function*
2. Find a *particular integral*. We will shortly see how to do this.
3. Write down the general solution as *complementary function* + *particular integral*

The second step in this process involves using a *trial solution*. Depending on the function f there are different trial functions you can use. The most common are as follows

$f(x)$	trial solution
polynomial of degree n	polynomial of degree n
$k \sin x$ or $k \cos x$	$K \cos x + L \sin x$
e^{kx}	Ke^{kx}

Note that the first case includes

- constants, for which the trial solution is a constant.
- kx, for which the trial solution is $Kx + L$
- kx^2, for which the trial solution is $Kx^2 + Lx + M$

Example 7.4.1

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

We found the complementary function in the last section, that is

$$y(x) = Ae^{-2x} + Be^{x}.$$

A trial solution should be $Kx^2 + Lx + M$. Substituting into the differential equation, we get

$$2K + (2Kx + L) - 2(Kx^2 + Lx + M) = x^2$$

Comparing coefficients of 1, x, x^2 we get

$$\begin{aligned} 2K + L - 2M &= 0 \\ 2K - 2L &= 0 \\ -2K &= 1 \end{aligned}$$

which has solution $K = L = -\frac{1}{2}$, $M = -\frac{3}{4}$ and the particular integral is

$$y(x) = -\frac{1}{2}x^2 - \frac{1}{2}x - \frac{3}{4}.$$

giving the solution to the differential equation as

$$y(x) = Ae^{-2x} + Be^{x} - \frac{1}{2}x^2 - \frac{1}{2}x - \frac{3}{4}.$$

7.5 Initial Value Problems

An *initial value problem* or IVP for short, is where we are also supplied with enough information to determine the constants in the solution (e.g. A and B in the previous section). Suppose that for the homogeneous ODE

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 2y = 0$$

we are also told that $y(0) = 2$ and $\frac{dy}{dx}(0) = 5$. Then

$$\begin{aligned} y(x) &= Ae^{-2x} + Be^{x} \\ y'(x) &= -2Ae^{-2x} + Be^{x} \end{aligned}$$

evaluated at 0 tells us

$$\begin{aligned} 2 &= A + B \\ 5 &= -2A + B \end{aligned}$$

which has the solution $A = -1$ and $B = 3$. So the solution to the IVP is

$$y(x) = -e^{-2x} + 3e^{x}.$$

Chapter 8

Complex Function Theory

The subject usually referred as "complex function theory" concerns the calculus of the complex numbers. This calculus has more structure than real calculus and as a result has some surprising and strong properties.

8.1 Standard Complex Functions

Recall that complex numbers are written $z = x+iy$ where $i^2 = -1$. Suppose that we are given a complex function $f : \mathbb{C} \to \mathbb{C}$, so that $w = f(z)$. Then in general we can write $w = u(x, y) + iv(x, y)$. u is the *real part* of f and v is the *imaginary part* of f. For example,

$$z^2 = (x + iy)^2 = x^2 + 2xyi + i^2y^2 = (x^2 - y^2) + i(2xy)$$

So in this case, $u(x, y) = x^2 - y^2$ and $v(x, y) = 2xy$. We use power series to define some complex functions, for example

$$
\begin{aligned}
e^z &= 1 + z + \frac{z^2}{2!} + \cdots \\
\sin z &= z - \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots \\
\cos z &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \cdots
\end{aligned}
$$

By manipulating power series we can prove that certain identities from the real case carry through to the complex case, such as:

$$
\begin{aligned}
e^z e^w &= e^{z+w} \\
\sin(-z) &= -\sin z \\
\cos(-z) &= \cos z \\
\sin(z+w) &= \sin z \cos w + \cos w \sin z \\
\cos(x+w) &= \cos z \cos w - \sin z \sin w
\end{aligned}
$$

Most importantly we have *Euler's formula*

$$e^{iz} = \cos z + i \sin z$$

In particular for $\theta \in \mathbb{R}$, $e^{i\theta} = \cos\theta + i\sin\theta$ giving the famous identity $e^{i\pi} = -1$, and *de Moivre's theorem*

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i \sin n\theta$$

Also, $e^z = e^{x+iy} = e^x e^{iy}$ gives

$$e^z = e^x(\cos y + i \sin y)$$

Thus $|e^z| = e^x$ and $Arg(e^z) = y$. It follows that e^z, like its real counterpart, is never equal to zero. Recall that the real function sin is *periodic*. This means that there is a number $0 \neq a \in \mathbb{R}$ with $\sin(x+a) = \sin x$. The least such a is 2π and is called the *period* of sin. Since

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z$$

we have the following fact

e^z is periodic, with period $2\pi i$.

So the complex function e^z displays behaviour like both the real exponential function and the real trigonometric functions.

We could also define the complex logarithm using a power series. But it is more revealing to define it as the inverse of the exponential function. Suppose that $w \in \mathbb{C}$ satisfies $e^w = z$. Let $z = re^{i\theta}$ and let $w = u + iv$. Then

$$e^w = e^{u+iv} = e^u e^{iv} = re^{i\theta}$$

So $r = e^u$ which gives $u = \log r$, and $e^{iv} = e^{i\theta}$ implies that $v = \theta + 2n\pi$ for $n \in \mathbb{Z}$. That is,

$$\log z = \log r + i(\theta + 2n\pi),\ n \in \mathbb{Z}$$

Note that the complex logarithm is many-valued. If we restrict to imaginary parts within the interval $(-\pi, \pi]$ we obtain what is called the *principal logarithm*.

The complex logarithm allows us to define complex powers of complex numbers. Since $z^w = e^{logz^w} = e^{w \log z}$, this serves as a definition:

$$z^w = e^{w \log z}$$

For example, $i^i = e^{i \log i} = e^{i(i\frac{\pi}{2})} = e^{-\frac{\pi}{2}}$.

There is also a simple ezpression for the complex sin and cos functions. Recall the definitions of the *hyperbolic functions*

$$\sinh x = \frac{1}{2}(e^x - e^{-x}),\ \ \cosh x = \frac{1}{2}(e^x + e^{-x})$$

from

$$e^{iz} = \cos z + i \sin z,\ \ e^{-iz} = \cos z - i \sin z$$

we obtain

$$\cos z = \frac{1}{2}(e^{iz} + e^{-iz}),\ \ \sin z = \frac{1}{2i}(e^{iz} - e^{-iz})$$

giving

$$\cos(iy) = \cosh y,\ \ \sin(iy) = i \sinh y$$

Since

$$\sin z = \sin(x + iy) = \sin x \cos(iy) + \cos x \sin(iy)$$

we have

$$
\begin{aligned}
\sin z &= \sin x \cosh y + i \cos x \sinh y \\
\cos z &= \cos x \cosh y - i \sin x \sinh y
\end{aligned}
$$

8.2 The Cauchy-Riemann Equations

There is a similar notion of differentiability for complex functions to the one for real functions. $f : \mathbb{C} \to \mathbb{C}$ is *differentiable* at $z \in \mathbb{C}$ if the limit

$$\lim_{h \to 0} \left\{ \frac{f(z+h) - f(z)}{h} \right\}$$

exists; if it does them the limit is called the *derivative* $f'(z) = \frac{df}{dz}$ of f at z. The only difference from the real case is that h can tend to 0 *in any manner*, i.e. along any curve. By taking two different limits, in the direction of the real axis, and in the direction of the imaginary axis, we obtain two different expressions for the derivative:

> Let $f(z) = u(x,y) + iv(x,y)$. If f is differentiable at $z = (x,y)$ then
>
> $$\begin{aligned} f'(z) &= u_x(x,y) + iv_x(x,y) \\ &= v_y(x,y) - iu_y(x,y) \end{aligned}$$

It follows that if $f : \mathbb{C} \to \mathbb{C}$ is differentiable then the following partial differential equations hold for u and v.

> **The Cauchy-Riemann Equations**
>
> $$u_x = v_y, \; v_x = -u_y$$

Examples 8.2.1 1. Let $f(z) = \bar{z}$, the complex conjugate of z. Then $f(z) = x - iy$ which means that $u(x,y) = x$ and $v(x,y) = -y$. Since $u_x = 1$ and $v_y = -1$, the Cauchy-Riemann equations are not satisfied, so f is not differentiable.

2. Let $f(z) = \sin z = \sin z \cosh y + i \cos x \sinh y$. Then we can differentiate f to get

$$\begin{aligned} f'(z) &= u_x(x,y) + iv_x(x,y) \\ &= \cos x \cosh y - \sin x \sinh y \\ &= \cos z \end{aligned}$$

Check that you get the same answer using

$$f'(z) = v_y(x, y) - iu_y(x, y).$$

As an exercise, try to show that $\frac{d}{dz}(e^z) = e^z$, $\frac{d}{dz}(\cos z) = -\sin z$ and $\frac{d}{dz}(\log z) = \frac{1}{z}$.

Recall that the partial differential equation $\nabla^2 u = u_{xx} + u_{yy} = 0$ is called *Laplace's equation*. If f is a differentiable complex function (whose components are sufficently differentiable) then by the Cauchy-Riemann equations,

$$u_{xx} + u_{yy} = v_{yx} - v_{xy} = 0$$

Because by the theorem of mixed partial derivatives $v_{yx} = v_{xy}$. Thus u (and similarly v) is a solution to Laplace's equation, what is called a *harmonic* function. If two harmonic functions are the real and imaginary parts of a complex function then they are called *harmonic conjugates.*

Example 8.2.2 Let $u(x, y) = x^3 - 3xy^2$. Then $u_x x = 6x$ and $u_y y = -6x$ so u is harmonic. To find a harmonic conjugate v for u we solve the Cauchy-Riemann equations. We have $v_y = u_x = 3x^2 - 3y^2$ and $v_x = -u_y = 6xy$. From the second of these,

$$v(x, y) = \int v_x dx = 3x^2 y + g(y).$$

Differentiating with respect to y, we obtain $v_y = 3x^2 + g'(y)$, giving $g'(y) = -3y^2$. So $g(y) = -y^3 + c$ for any constant c. Taking $c = 0$ gives $v(x, y) = 3x^2 y - y^3$. Hence

$$\begin{aligned} f(z) &= x^3 - 3xy^2 + i(3x^2 y - y^3) \\ &= (x + iy)^3 \\ &= z^3 \end{aligned}$$

8.3 Complex Integrals

In this final section we will only touch on the ideas of complex integration. Complex integrals are similar to line integrals, that is they are performed on a curve in the complex plane. In this context the curve is usually called a *contour*. It is often convenient to parameterize using the complex exponential.

Example 8.3.1 1. Let C be the circle $|z| = r$, with centre $a \in \mathbb{C}$ parameterized in an anticlockwise direction. Then $z = a + re^{i\theta}$ on C, for $0 \leqslant \theta \leqslant 2\pi$. We wish to evaluate $\oint_C \frac{1}{z-a} dz$. Now, $dz = ire^{i\theta} d\theta$ and we have $z - a = re^{i\theta}$. Hence

$$\begin{aligned} \oint_C \frac{1}{z-a} dz &= \int_0^{2\pi} \frac{ire^{i\theta} d\theta}{re^{i\theta}} \\ &= i \int_0^{2\pi} d\theta = 2\pi i \end{aligned}$$

2. On the other hand suppose that we integrate, on the same contour, $(z-a)^n$ for any integer $n \neq -1$. Then we have

$$\begin{aligned}\oint_C (z-a)^n dz &= \int_0^{2\pi} (re^{i\theta})^n ire^{i\theta} d\theta \\ &= ir^{n+1} \int_0^{2\pi} e^{i(n+1)\theta} d\theta \\ &= \frac{ir^{n+1}}{i(n+1)} \left[e^{i(n+1)\theta} \right]_0^{2\pi} \\ &= \frac{r^{n+1}}{n+1} \left(e^{2\pi(n+1)i} - 1 \right) \\ &= 0\end{aligned}$$

Since $e^{2\pi i m} = 1$ for any integer m.

These two important examples illustrate, in a sense, much of the important behaviour about complex functions. Note

- The value of the integral did not depend on a or r
- There is something special about the power -1.

In fact given any complex function that is sufficiently smooth to have an expansion as a Taylor series, such as

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

we have

$$\frac{e^z}{z} = \frac{1}{z} + 1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots$$

and for C as in the previous example, with $a = 0$

$$\oint_C \frac{e^z}{z} = \oint_C \frac{1}{z} dz + \oint_C dz + \oint_C \frac{z}{2!} dz + \cdots$$

Here only the first term in the right hand side, which equals $2\pi i$, is nonzero. Hence this is the value of the integral. Generalizing these ideas we obtain

Theorem 8.3.2 (Cauchy's Integral Formula) *Let* $f : \mathbb{C} \to \mathbb{C}$ *be any function which is differentiable on and inside* any *simple closed contour* C *which encloses (but doesn't pass through) the point* $a \in \mathbb{C}$*. Then*

$$\oint_C \frac{f(z)}{z-a} dz = 2\pi i f(a)$$

Cauchy's integral formula is very powerful - it allows us to integrate most contour integrals in terms of the value of the function we are integrating at a singularity (given that we can express it as a function with a Taylor series (called an *analytic function* divided by z. But you might ask

- What happens for a function which can be expressed by a Taylor series (without having to divide by z)?
- What about functions divided by other powers of z?

In the first case we simply get 0. You can look at this two ways. Firstly the computation above with, say, e^z instead of $\frac{e^z}{z}$ will clearly give 0. On the other hand complex functions can be thought of as vector fields $\mathbb{R}^2 \to \mathbb{R}^2$. In the analytic case, they correspond to conservative fields. We don't have room here for the details, but to address the second point you can do a similar calculation which results in the following generalized formula.

Theorem 8.3.3 *Let* C *and* f *be as in the statement of Cauchy's integral formula. Then*

$$\oint_C \frac{f(z)}{(z-a)^k} dz = \frac{2\pi i f^{(k-1)}(a)}{(k-1)!}$$

Here, $f^{(k-1)}$ denotes the $(k-1)^{\text{th}}$ derivative of f.

All of these ideas lead up to the most powerful theorem for computing complex integrals, called the *Theorem of Residues.* If you learn any theorem about complex integrals then learn this one! It basically says that we can evaluate a complex integral by computing properties at only finitely many points, the singularities or *poles* of the function we are integrating.

Suppose that

$$f(z) = \frac{g(z)}{(z-a)^k}$$

where g is differentiable on and inside a circle of radius $r > 0$ about a. Then

- We say that f has a *pole of order* k at a.
- We call the expression

$$\operatorname{res}(f, a) = \frac{g^{(k-1)}(a)}{(k-1)!}$$

  the *residue* at the pole a.

Theorem 8.3.4 *If $f : \mathbb{C} \to \mathbb{C}$ is differentiable on and inside a simple closed contour C except at finitely many poles $a_1, a_2, \ldots, a_n$ inside C. Then*

$$\oint_C f(z)dz = 2\pi i \sum_{m=1}^{n} res(f, a_m)$$

Example 8.3.5 Calculate

$$\oint_C \frac{dz}{(z-1)(z+1)(z-2)}$$

where C is a circle of radius 2 about $z = 2$. Let

$$f(z) = \frac{1}{(z-1)(z+1)(z-2)}.$$

Then to find $\text{res}(f, 1)$ we evaluate $(z-1)f(z)$ at $z = 1$ to get $-\frac{1}{2}$. Similarly, $\text{res}(f, -1) = \frac{1}{6}$ and $\text{res}(f, 2) = \frac{1}{3}$. Only the poles at 1 and 2 lie inside C so we have

$$\begin{aligned} \oint_C f(z)dz &= 2\pi i(\text{res}(f, 1) + \text{res}(f, 2)) \\ &= 2\pi i(-\frac{1}{2} + \frac{1}{3}) \\ &= -\frac{\pi i}{3} \end{aligned}$$

Index