THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Course Reader for,

MATH7501 Mathematics for

Data Science 1

(last revision: 15/1/2021).

This course reader provides the core material for the Master of Data Science course MATH7501. You may refer to the following references for additional material, examples and exercises.

References

[1] H. Anton, *Elementary Linear Algebra*, John Wiley, 11th Ed., 2014.

[2] N. Biggs, *Discrete Mathematics*, Oxford University Press, 2nd Ed., 2002.

[3] D. Dietrich et. al., *Data Science and Big Data Analytics*, Wiley, 2015.

[4] A. J. Dobson, J. Stokoe, *Self-Paced Introductory Mathematics*, Australian National University Press. 3d Ed., 1990.

[5] S. Epp, *Discrete Mathematics with Applications*, Brooks/Cole, 4th Ed., 2011.

[6] E. Kreyszig. *Advanced Engineering Mathematics*, Wiley. 9th Ed., 2006.

[7] K. Rosen, *Discrete Mathematics and its Applications*, McGRaw-Hill, 2nd Ed., 1991.

[8] M. Spivak, *Calculus*, Cambridge University Press, 3d Ed., 1994.

[9] J. Steward. *Calculus*, Brooks/Cole, 7th Ed. 2012.

These notes contain materials and contributions from Joseph Grotowski, Phil Isaac, Michael Jennings, Birgit Loch, Barbara Maenhaut, Victor Scharaschkin, Mary Waterhouse, Poh Wah Hillock, and Sam Hambleton. Edited by Hayden Klok, and Yoni Nazarathy.

Contents

Introduction

The ability to collect, interpret, and analyse large data sets is an important skill in making informed decisions on things including efficiency of business activities, developing trends, and making predictions, among many more. In order to develop these skills, a good understanding of various mathematical techniques and a means of carrying out these methods on a computer are imperative. This course introduces students to the basic mathematical tools and methods which will provide important background for understanding the mathematics involved in data science and practice in calculating associated quantities. Through studying the material in this course, you will have a better understanding of subsequent material in the Masters of Data Science program.

In particular, this course will provide students with the mathematical foundations needed for analytical and statistical data science. The course will introduce logic, proofs and the elementary properties of functions and relations. Limits, sequences and series which form the basis of calculus are introduced. Derivatives, maxima/minima, integration, Taylor series, and some elementary numerical methods for doing these are practiced. You will study multi-dimensional problems as an extension to calculus in of one variable, including partial derivatives, and the maxima and minima of functions of two or more variables. Numerical integration and solution of non-linear equations will be studied. Linear Algebra is an important part of managing data and naturally we will study this and related topics including vectors, the scalar product, norms, simultaneous equations, determinants, eigenvalues, eigenvectors, and applications. While studying these topics you will use appropriate mathematical software for linear algebra applications and perform symbolic calculations supporting what you learn in lectures. We hope to have time this semester to introduce data fitting using least squares, however this will depend on time constraints.

1 Basic operations with matrices vectors

1.1 Matrices as a way of organizing data

There are many ways to obtain data. It may be collected from an image, a video, a music file, the Australian Bureau of Statistics website (ABS), come from an MRI image, a topographical survey, among countless many other sources. Collected data may have various degrees of structure. Data downloaded from the ABS website is called structured data since it is presented in the form of a spreadsheet with labeled columns and rows. Text documents having no such structure are known as unstructured data. Consider the following example.

1		Earnings; Males; Full Time; Adult; Ordinary time earnings ;	Earnings; Males; Full Time; Adult; Total earnings ;	Earnings; Males; Total earnings ;	Earnings; Females; Full Time; Adult; Ordinary time earnings ;	Earnings; Females; Full Time; Adult; Total earnings ;	Earnings; Females; Total earnings ;	Earnings; Persons; Full Time; Adult; Ordinary time earnings ;	Earnings; Persons; Full Time; Adult; Total earnings ;	Earnings; Persons; Total earnings ;
2	**Unit**	$	$	$	$	$	$	$	$	$
3	**Series Type**	Trend	Trend	Trend	Trend	Trend	Trend	Trend	Trend	Trend
4	**Data Type**	RATIO	RATIO	RATIO	RATIO	RATIO	RATIO	RATIO	RATIO	RATIO
5	**Frequency**	Biannual	Biannual	Biannual	Biannual	Biannual	Biannual	Biannual	Biannual	Biannual
6	**Collection Month**	2	2	2	2	2	2	2	2	2
7	**Series Start**	May-2012	May-2012	May-2012	May-2012	May-2012	May-2012	May-2012	May-2012	May-2012
8	**Series End**	May-2016	May-2016	May-2016	May-2016	May-2016	May-2016	May-2016	May-2016	May-2016
9	**No. Obs**	9	9	9	9	9	9	9	9	9
10	**Series ID**	A84990042R	A84990045W	A84990048C	A84990043T	A84990046X	A84990049F	A84990044V	A84990047A	A84990050R
11	May-2012	1447.10	1538.10	1286.00	1194.00	1211.30	822.80	1353.30	1417.40	1054.30
12	Nov-2012	1488.50	1576.60	1324.50	1226.40	1242.80	838.40	1392.80	1454.40	1081.00
13	May-2013	1515.70	1603.70	1351.30	1252.10	1268.90	852.20	1420.50	1483.10	1103.40
14	Nov-2013	1533.90	1621.80	1352.40	1267.60	1284.50	870.60	1437.20	1499.00	1114.60
15	May-2014	1560.70	1650.60	1362.90	1277.30	1295.00	881.20	1455.30	1518.70	1122.50
16	Nov-2014	1584.60	1675.10	1370.60	1291.20	1309.30	889.90	1474.50	1537.40	1129.00
17	May-2015	1593.20	1679.30	1370.20	1308.50	1326.80	905.70	1484.90	1545.60	1136.70
18	Nov-2015	1603.20	1686.30	1376.90	1327.60	1345.60	915.60	1500.00	1558.30	1146.70
19	May-2016	1613.60	1696.60	1393.50	1352.50	1370.10	925.10	1516.00	1575.40	1160.20
20										

Figure 1: An example of structured data from the ABS website on weekly earnings.

We are able to download such files as shown in Fig. 1 and save as a CSV file. It is generally simple to import a CSV file in other software and then make use of that. Doing just that to the data in Fig. 1 gives the matrix

$$A = \begin{pmatrix} 1447.1 & 1538.1 & 1286. & 1194. & 1211.3 & 822.8 & 1353.3 & 1417.4 & 1054.3 \\ 1488.5 & 1576.6 & 1324.5 & 1226.4 & 1242.8 & 838.4 & 1392.8 & 1454.4 & 1081. \\ 1515.7 & 1603.7 & 1351.3 & 1252.1 & 1268.9 & 852.2 & 1420.5 & 1483.1 & 1103.4 \\ 1533.9 & 1621.8 & 1352.4 & 1267.6 & 1284.5 & 870.6 & 1437.2 & 1499. & 1114.6 \\ 1560.7 & 1650.6 & 1362.9 & 1277.3 & 1295. & 881.2 & 1455.3 & 1518.7 & 1122.5 \\ 1584.6 & 1675.1 & 1370.6 & 1291.2 & 1309.3 & 889.9 & 1474.5 & 1537.4 & 1129. \\ 1593.2 & 1679.3 & 1370.2 & 1308.5 & 1326.8 & 905.7 & 1484.9 & 1545.6 & 1136.7 \\ 1603.2 & 1686.3 & 1376.9 & 1327.6 & 1345.6 & 915.6 & 1500. & 1558.3 & 1146.7 \\ 1613.6 & 1696.6 & 1393.5 & 1352.5 & 1370.1 & 925.1 & 1516. & 1575.4 & 1160.2 \end{pmatrix}$$

The rows and columns of the matrix A have various meanings. In this case, as we see in Fig. 1, the rows correspond to a point in time, while the columns correspond to weekly earnings in various categories. The data contains columns representing the same quantities for males and females. Anyone who wishes to examine the data further may want extract rows or columns, perform operations on specific rows and columns such as taking a difference of two columns, or even model the data in a specific column through some form of interpolation, fitting a function to it. These tasks become easier when we understand the meaning of indexing in a matrix.

Definition 1. *A rectangular array of numbers*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{pmatrix}$$

is called an $m \times n$ *matrix. It is made up of* m *rows and* n *columns.*

Entries of the jth column may be assembled into a vector

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

called the jth *column vector* of A. Similarly the ith *row vector* of A is

$$(a_{i1}\, a_{i2} \cdots\cdots a_{in}).$$

Note that an $m \times 1$ matrix is a column vector and a $1 \times n$ matrix is a row vector. The number a_{ij} in the ith row and j column of A is called the $(i,\, j)$th *entry* of A. For brevity we write $A = (a_{ij})$.

Example 1. *The* 2×3 *matrix* A *with entries* $a_{ij} = i - j$ *is*

$$A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \end{pmatrix}.$$

Example 2. *The* $m \times n$ *matrix* A *whose entries are all zero is called the* zero matrix, *denoted* 0*; e.g. the zero* 3×2 *matrix is*

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Definition 2. *Matrices* A*,* B *are said to be* **equal***, written* $A = B$*, if they have the same number of rows and columns and* $a_{ij} = b_{ij}$*, for all* $i,\ j$*.*

1.2 Matrices as images

An image file stores information on the quantity of colour or hue in each pixel. This is illustrated below in Figure 2.

Figure 2: A 16 by 16 pixel black and white image of the number 7.

The matrix corresponding to the pixel data in Figure 2 is given in the following matrix B.

$$B = \begin{pmatrix}
0.99 & 1. & 1. & 0.99 & 1. & 1. & 1. & 0.97 & 1. & 0.98 & 0.99 & 1. & 1. & 0.99 & 1. & 0.98 \\
1. & 0.96 & 1. & 1. & 0.98 & 1. & 0.96 & 1. & 1. & 0.98 & 1. & 1. & 0.97 & 0.98 & 1. & 1. \\
0.99 & 1. & 1. & 0.98 & 1. & 1. & 1. & 0.98 & 1. & 1. & 0.97 & 0.98 & 1. & 0.99 & 0.98 & 0.98 \\
0.98 & 0.95 & 0.97 & 0.019 & 0. & 0. & 0. & 0.01 & 0. & 0. & 0.01 & 0.01 & 1. & 0.99 & 1. & 1. \\
1. & 1. & 1. & 0.99 & 1. & 1. & 0.97 & 1. & 1. & 1. & 1. & 0. & 0.99 & 1. & 0.97 & 0.98 \\
0.98 & 0.98 & 1. & 0.98 & 0.99 & 1. & 0.98 & 0.99 & 1. & 1. & 0. & 0. & 1. & 0.98 & 1. & 1. \\
1. & 1. & 1. & 0.99 & 0.98 & 0.99 & 1. & 1. & 1. & 0. & 0.01 & 1. & 0.96 & 1. & 0.99 & 0.96 \\
0.99 & 1. & 0.98 & 1. & 1. & 1. & 0.99 & 0.99 & 0. & 1. & 0.99 & 0.97 & 1. & 1. & 0.96 & 1. \\
1. & 0.99 & 0.99 & 1. & 1. & 0.99 & 1. & 0.99 & 0. & 1. & 0.99 & 1. & 1. & 1. & 0.98 & 1. \\
1. & 0.99 & 0.01 & 0. & 0. & 0.01 & 0. & 0.01 & 0.01 & 0. & 0.02 & 0. & 0.98 & 0.98 & 1. & 0.99 \\
0.99 & 0.98 & 1. & 0.99 & 1. & 1. & 1. & 0. & 0.98 & 1. & 0.98 & 1. & 1. & 1. & 0.98 & 0.99 \\
1. & 0.97 & 1. & 1. & 0.98 & 1. & 0.98 & 0.01 & 0.99 & 1. & 1. & 1. & 0.99 & 0.99 & 0.99 & 1. \\
1. & 1. & 0.98 & 0.99 & 1. & 0.99 & 1. & 0. & 1. & 0.98 & 1. & 0.98 & 0.98 & 1. & 1. & 0.98 \\
0.98 & 1. & 1. & 1. & 0.97 & 1. & 0.97 & 0.01 & 1. & 1. & 0.99 & 1. & 1. & 0.97 & 1. & 0.99 \\
1. & 1. & 1. & 1. & 1. & 1. & 0.98 & 0.02 & 0.99 & 0.99 & 1. & 1. & 0.96 & 1. & 1. & 1. \\
0.99 & 1. & 1. & 0.99 & 1. & 1. & 1. & 0.98 & 1. & 1. & 0.96 & 1. & 0.99 & 0.99 & 1. & 0.98
\end{pmatrix}$$

Conversely, we can store data from a matrix as an image. This provides a means of storing a large data set a concise form. The matrix A is displayed as an image in Figure 3.

1.3 Matrix operations

Addition

The *sum* of two $m \times n$ matrices A and B is defined to be the $m \times n$ matrix $A + B$ with entries

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

We add matrices element wise, as for vectors.

Addition is only defined between matrices of the same size.

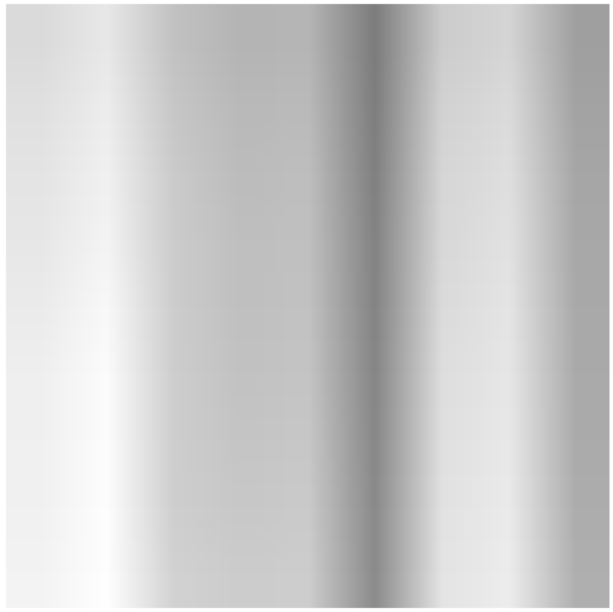


Figure 3: The data from Fig. 1 displayed as an image, after dividing throughout A by 1696.6, the maximum entry so that $0 \leq a_{ij} \leq 1$.

Properties

For matrices $A,\ B,\ C$ of the same size we have

$$\begin{aligned} A+B &= B+A, \\ (A+B)+C &= A+(B+C). \end{aligned}$$

Scalar multiplication

Let A be an $m \times n$ matrix and $\alpha \in \mathbb{R}$. We define αA to be the $m \times n$ matrix with entries

$$(\alpha A)_{ij} = \alpha \cdot a_{ij} \quad , \quad \text{for all } i,\ j.$$

Write $-A$ for $(-1) \cdot A$. Define subtraction between matrices of the same size by

$$A - B = A + (-B).$$

Example 3. *Find $A + B$ and $A - B$ if*

$$A = \begin{pmatrix} -4 & 6 & 3 \\ 0 & 1 & 2 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 5 & -1 & 0 \\ 3 & 1 & 0 \end{pmatrix}.$$

$$A + B = \begin{pmatrix} 1 & 5 & 3 \\ 3 & 2 & 2 \end{pmatrix} \quad , \quad A - B = \begin{pmatrix} -9 & 7 & 3 \\ -3 & 0 & 2 \end{pmatrix}.$$

Example 4. *If* $A = \begin{pmatrix} 6 & 0 \\ 2 & -1 \end{pmatrix}$ *then* $A + A = \begin{pmatrix} 12 & 0 \\ 4 & -2 \end{pmatrix} = 2A.$

Matrix multiplication

We generalize the multiplication of 2×2 matrices. Let $A = (a_{ij})$ be an $m \times n$ and $B = (b_{jk})$ an $n \times r$ matrix. Then AB is the $m \times r$ matrix with ik entry

$$(ab)_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots\cdots + a_{in}b_{nk} \ :$$

$$\underset{(m \times n)}{\underset{A}{\begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots\cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{pmatrix}}} \underset{(n \times r)}{\underset{B}{\begin{pmatrix} b_{11} & \cdots\cdot & b_{1k} & \cdots\cdot & b_{1r} \\ b_{21} & \cdots\cdot & b_{2k} & \cdots\cdot & b_{2r} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ b_{n1} & \cdots\cdot & b_{nk} & \cdots\cdot & b_{nr} \end{pmatrix}}} = \underset{(m \times r)}{AB}$$

where $(ab)_{ik} =$ dot product between ith row of A and kth column of B.

AB is only defined if the number of <u>columns</u> of $A =$ the number of <u>rows</u> of B.

Example 5. • *Let* $A = \underset{1 \times 3}{(-1\ 0\ 1)}$ *and* $B = \underset{3 \times 2}{\begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & 5 \end{pmatrix}}$

- *Then AB is defined, but $\underset{3\times 2}{B}\ \underset{1\times 3}{A}$ is not defined.*

Calculate AB.

$$\begin{aligned} AB &= (-1\ 0\ 1)\begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & 5 \end{pmatrix} \\ &= (-1 \cdot 0 + 0 \cdot 1 + 1 \cdot -2 \qquad -1 \cdot 1 + 0 \cdot 2 + 1 \cdot 5) \\ &= (-2 \quad 4) \end{aligned}$$

Example 6.

$$A = (1\ 3\ 9)\,, \ \ B = \begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix}.$$

Calculate AB and BA.

$$\underset{1\times 3}{(1\ \ 3\ \ 9)} \underset{3\times 1}{\begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix}} = 2 + 3 + 63 = 68$$

$$\underset{3\times 1}{\begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix}} \underset{1\times 3}{\begin{pmatrix} 1 & 3 & 9 \end{pmatrix}} = \begin{pmatrix} 2 & 6 & 18 \\ 1 & 3 & 9 \\ 7 & 21 & 63 \end{pmatrix}.$$

Properties

For matrices of appropriate size

(1) $(AB)C = A(BC)$, Associativity

(2) $(A+B)C = AC + BC$, $A(B+C) = AB + AC$ Distributive Laws

Another unusual property of matrices is that $AB = 0$ does *not* imply $A = 0$ or $B = 0$. It is possible for the product of two non-zero matrices to be zero:

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Also, if $AC = BC$, or $CA = CB$ then it is not true in general that $A = B$. (If $AC = BC$, then $AC - BC = 0$ and $(A - B)C = 0$, but this does **not** imply $A - B = 0$ or $C = 0$.)

Transposition

Definition 3. *The* transpose *of an $m \times n$ matrix $A = (a_{ij})$ is the $n \times m$ matrix A^T with entries*

$$a_{ji} = a_{ij}^T \quad , \quad \textit{for all } i,\ j$$

ie.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots\cdots & a_{m1} \\ a_{12} & a_{22} & \cdots\cdots & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \cdots\cdots & a_{mn} \end{pmatrix}$$

So the row vectors of A become column vectors of A^T and vice versa.

Example 7 (transposition)**.** *If* $A = \begin{pmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$, $C = (7 \ 5 \ -2)$ *then find A^T, B^T, C^T.*

$$A^T = \begin{pmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{pmatrix}^T = \begin{pmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{pmatrix}$$

$$B^T = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$$

$$C^T = (7 \ 5 \ -2)^T = \begin{pmatrix} 7 \\ 5 \\ -2 \end{pmatrix}.$$

Properties

For matrices of appropriate size

(1) $(\alpha A)^T = \alpha \cdot A^T \quad , \quad \alpha \in \mathbb{R}$

(2) $(A+B)^T = A^T + B^T$

(3) $(A^T)^T = A$

(4) $(AB)^T = B^T A^T \qquad$ (*not* $A^T B^T$!).

Dot product expressed as matrix multiplication

A column vector

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

may be interpreted as an $n \times 1$ matrix. Then the dot product (for two column vectors $\boldsymbol{v}$ and $\boldsymbol{w}$) may be expressed using matrix multiplication:

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{w} &= v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \underset{1 \times n}{(v_1\ v_2\ \cdots\ v_n)} \underset{n \times 1}{\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}} \\ &= \boldsymbol{v}^T \boldsymbol{w}. \end{aligned}$$

Also

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{v} &= v_1^2 + v_2^2 + \cdots + v_n^2 = \|\boldsymbol{v}\|^2 \\ &= \boldsymbol{v}^T \boldsymbol{v}. \end{aligned}$$

1.4 Identity matrix, inverses and determinants

The identity matrix

An $n \times n$ matrix ($m = n$) is called a *square matrix* of *order* n. The diagonal containing the entries

$$a_{11},\ a_{22},\ \cdots\ ,\ a_{nn}$$

is called the *main diagonal* (or *principal diagonal*) of A. If the entries above this diagonal are all zero then A is called *lower* triangular. If all the entries below the diagonal are zero, A is called *upper triangular*

$$\underset{\text{lower triangular}}{\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ \vdots & a_{22} & 0 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ a_{n1} & \cdots & & \cdots & a_{nn} \end{pmatrix}} \qquad \underset{\text{upper triangular}}{\begin{pmatrix} a_{11} & \cdots & & \cdots & a_{1n} \\ 0 & a_{22} & & & \vdots \\ 0 & 0 & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}}$$

If elements above *and* below the principal diagonal are zero, so

$$a_{ij} = 0\,, \quad i \neq j$$

then A is called a *diagonal matrix*. Note that if A is diagonal then $A = A^T$.

Example 8. $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ *is a diagonal* 3×3 *matrix.*

Definition 4 (Identity matrix)**.** *The* $n \times n$ identity matrix $I = I_n$*, is the diagonal matrix whose entries are all* 1*:*

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \end{pmatrix}.$$

Remarks on the identity matrix

Recall from linear transformations of $\mathbb{R}^2$: $A\boldsymbol{i}$ is the first column of A and $A\boldsymbol{j}$ is the second column.

For example, for the 2×2 case we have

$$AI = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A = IA.$$

In general the jth column vector of I is the coordinate vector

$$\boldsymbol{e}_j = \begin{array}{c} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \\ \\ j. \\ \\ \\ \\ \end{array}$$

Thus

$$
\begin{aligned}
A\boldsymbol{e}_j &= \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{c} \\ \\ \\ \quad j \\ \\ \\ \\ \end{array} \\
&= \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = \text{jth column vector of A.}
\end{aligned}
$$

So $AI = A$ for all A.

Replacing A by A^T, $A^T I = A^T$ also. Transpose both sides: $A = (A^T)^T = (A^T I)^T = I^T (A^T)^T = IA$, so $IA = A$ for all A.

We have proved that for all square matrices A

$$IA = AI = A.$$

Definition 5 (Inverse)**.** *Let I denote the identity matrix.*

A square matrix A is **invertible** *(or* **non-singular***) if there exists a matrix B such that*

$$AB = BA = I.$$

Then B is called the inverse *of A and is denoted A^{-1}. A matrix that is not invertible is also said to be* singular.

Inverse for the 2×2 case

Let A be a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Set $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Then

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = BA.$$

Let

$$\Delta = ad - bc.$$

If $\Delta \neq 0$, then

$$A^{-1} = \frac{1}{\Delta} B = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

We call Δ the *determinant* of A.

Example 9. *Find the inverse matrix of* $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$. *Check your answer.*

$\Delta = 1 \times 5 - 2 \times 3 = -1.$

$$\Rightarrow A^{-1} = \frac{1}{-1} \begin{pmatrix} 5 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}.$$

Check whether $AA^{-1} = A^{-1}A = I$:
Multiplying the two matrices A and A^{-1} gives

$$\begin{aligned} \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

So A is invertible with inverse

$$\begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}.$$

Example 10 (Inverse of a 3×3 matrix)**.** *Let*

$$A = \begin{pmatrix} 2 & -3 & -1 \\ 1 & -2 & -3 \\ -2 & 2 & -5 \end{pmatrix}, \qquad B = \begin{pmatrix} 16 & -17 & 7 \\ 11 & -12 & 5 \\ -2 & 2 & -1 \end{pmatrix}.$$

Show that $B = A^{-1}$.

An easy way to do this is to multiply the two matrices together.

$$AB = \begin{pmatrix} 2 & -3 & -1 \\ 1 & -2 & -3 \\ -2 & 2 & -5 \end{pmatrix} \begin{pmatrix} 16 & -17 & 7 \\ 11 & -12 & 5 \\ -2 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and, similarly

$$BA = \begin{pmatrix} 16 & -17 & 7 \\ 11 & -12 & 5 \\ -2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & -3 & -1 \\ 1 & -2 & -3 \\ -2 & 2 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Properties of inverses

(1) A has at most **one** inverse

(2) If A, B are invertible so is AB, and

$$(AB)^{-1} = B^{-1}A^{-1} \quad (\textit{not } A^{-1}B^{-1}!)$$

(3) A is invertible if and only if A^T is invertible

(4) If A is invertible then $(A^T)^{-1} = (A^{-1})^T$.

Determinants

Definition 6. *Associated to each square matrix A is a number called the* determinant *of A and denoted $|A|$ or $\det(A)$. It is defined as follows.*

- *If A is a 1×1 matrix, say $A = (a)$ then $|A|$ is defined to be a.*
- *If $A = (a_{ij})$ is 2×2 then we define*

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

- *In general if $A = (a_{ij})$ is $n \times n$, first set*

$$C_{ij} = (-1)^{i+j} \left| \begin{array}{cccc|ccc} a_{11} & a_{12} & \cdots & a_{1\ j-1} & a_{1\ j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2\ j-1} & a_{2\ j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1\ 1} & a_{i-1\ 2} & \cdots & a_{i-1\ j-1} & a_{i-1\ j+1} & \cdots & a_{i-1\ n} \\ \hline a_{i+1\ 1} & a_{i+1\ 2} & \cdots & a_{i+1\ j-1} & a_{i+1\ j+1} & \cdots & a_{i+1\ n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n\ j-1} & a_{n\ j+1} & \cdots & a_{nn} \end{array} \right|$$

  called the cofactor *of a_{ij}. The $(n-1) \times (n-1)$ determinant is obtained by* omitting *the ith row and jth column from A (indicated by the horizontal and vertical lines in the matrix).*

We then define

$$|A| = a_{11}C_{11} + a_{12}C_{12} + \cdots\cdots + a_{1n}C_{1n}$$

which gives a *recursive* definition of the determinant.

Observe that $(-1)^{i+j}$ gives the pattern $\begin{pmatrix} + & - & + & - & \cdot & \cdot & \cdot & \cdot \\ - & + & - & + & \cdot & \cdot & \cdot & \cdot \\ + & - & + & - & \cdot & \cdot & \cdot & \cdot \\ \cdot \\ \cdot\cdot \end{pmatrix}$

Thus for a 3×3 matrix A we have

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

where

$$C_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad C_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad C_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Properties of determinants

Property (1): $|A| = \left|A^T\right|$

Consider for example the 2×2 case. Then

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \Rightarrow \quad A^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$$

Thus $|A| = a_{11}a_{22} - a_{12}a_{21} = \left|A^T\right|$. It can be shown that this holds for any square matrix, not just in the 2×2 case.

This means that any results about the *rows* in a general determinant is also true about the *columns* (since the rows of A^T are the columns of A). In particular, any statement about the effect of row operations on determinants is also true for column operations.

Warning: We used row operations to simplify systems of equations, because they do not change the solution. Column operations may change the solution of linear systems, so we should not use column operations on such systems.

Property (2): The determinant may by taking cofactors along any row (not just the first) or down any column. Eg. for a 3×3 matrix A

$$\begin{aligned} |A| &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} && \text{(definition, expansion along 1st row)} \\ &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} && \text{(expansion along 2nd row)} \\ &= a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} && \text{(expansion down 3rd column) etc.} \end{aligned}$$

This is useful if one row or column contains a larger number of zeros.

Determinants of triangular matrices

Suppose $A = \begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & 0 & & \vdots \\ \vdots & & \ddots & & \vdots \\ & & & & 0 \\ a_{n1} & \cdots & \cdots & \cdots & a_{nn} \end{pmatrix}$

is lower triangular. Then by repeated expansion along $\boldsymbol{r}_1$ we obtain

$$|A| = a_{11} \begin{vmatrix} a_{22} & & & \\ \vdots & \ddots & & \\ \vdots & & \vdots & \\ a_{n2} & \cdots & \cdots & a_{nn} \end{vmatrix} = \ldots = a_{11}\ a_{22} \cdot \cdots \cdot\ a_{nn},$$

which is the product of the diagonal entries. The same result holds for upper triangular and diagonal matrices.

In particular

$$|I| = 1.$$

Connection with inverses

An important property of determinants is the following.

Fact (product of determinants)

Let $A,\ B$ be $n \times n$ matrices. Then

$$|AB| = |A| \cdot |B|\,.$$

Theorem 1 (Invertible matrices)**.**

$$A \textit{ is invertible } \iff |A| \neq 0.$$

Proof $\Longrightarrow$ If A is invertible, then

$$\begin{array}{rc} & I = AA^{-1} \\ \Rightarrow & 1 = |I| = |AA^{-1}| = |A| \cdot |A^{-1}|\,, \\ & \text{So } |A| \neq 0. \end{array}$$

$\Longleftarrow$ Follows from below.

Remark 1. *It follows immediately from the proof that if A is invertible, then*

$$|A^{-1}| = \frac{1}{|A|},$$

i.e., the inverse of the determinant is the determinant of the inverse.

1.5 Vectors operations (in 2 and n dimensions)

- A *vector* quantity has both a magnitude and a direction. Force and velocity are examples of vector quantities.
- A *scalar* quantity has only a magnitude (it has no direction). Time, area and temperature are examples of scalar quantities.

A vector is represented geometrically in the (x, y) plane (or in (x, y, z) space) by a directed line segment (arrow). The direction of the arrow is the direction of the vector, and the length of the arrow is proportional to the magnitude of the vector. Only the length and direction of the arrow are significant: it can be placed anywhere convenient in the (x, y) plane (or (x, y, z) space).

If P, Q are points, $\overrightarrow{PQ}$ denotes the vector from P to Q.

A vector $\boldsymbol{v} = \overrightarrow{PQ}$ in the (x, y) plane may be represented by a pair of numbers

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} x_Q - x_P \\ y_Q - y_P \end{pmatrix}$$

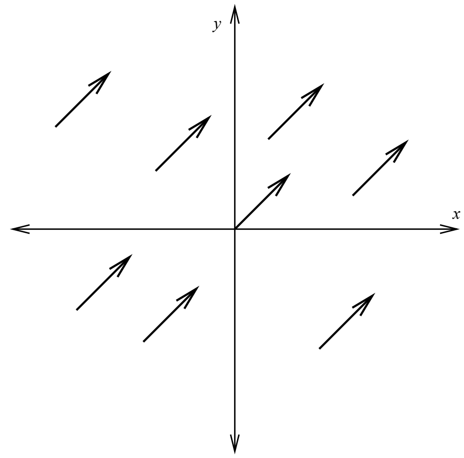


Figure 4: The vector of unit length at $45°$ to the x-axis has many representations.

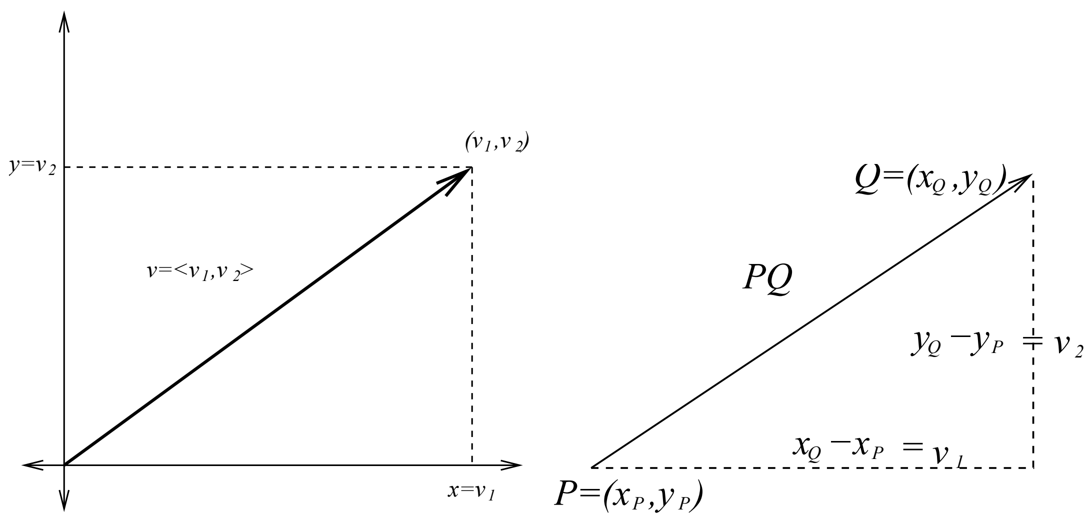


Figure 5: Geometric representation of a vector.

which is the same for all representations $\overrightarrow{PQ}$ of $\boldsymbol{v}$. We call v_1, v_2 the *components* of the vector $\boldsymbol{v}$.

We call the vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ the *zero vector*. It is denoted by $\mathbf{0}$.

Vectors will be indicated by bold lowercase letters $\boldsymbol{v}$, $\boldsymbol{w}$ etc. Writing by hand you may use $\underline{v}$ or $\vec{v}$ or $\widetilde{v}$.

Position vectors

Let $P = (x_p, y_p)$ be a point in the (x, y) plane. The vector $\overrightarrow{OP}$, where O is the origin, is called the *position vector* of P. Obviously

$$\overrightarrow{OP} = \begin{pmatrix} x_p \\ y_p \end{pmatrix}.$$

Norm

For vector $\boldsymbol{v} = \overrightarrow{PQ}$, the *norm*(or *length* or *magnitude*) of $\boldsymbol{v}$, written $\|\boldsymbol{v}\|$, is the distance between P and Q. Thus for $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ we have

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

Vector addition

We add vectors by the triangle rule.

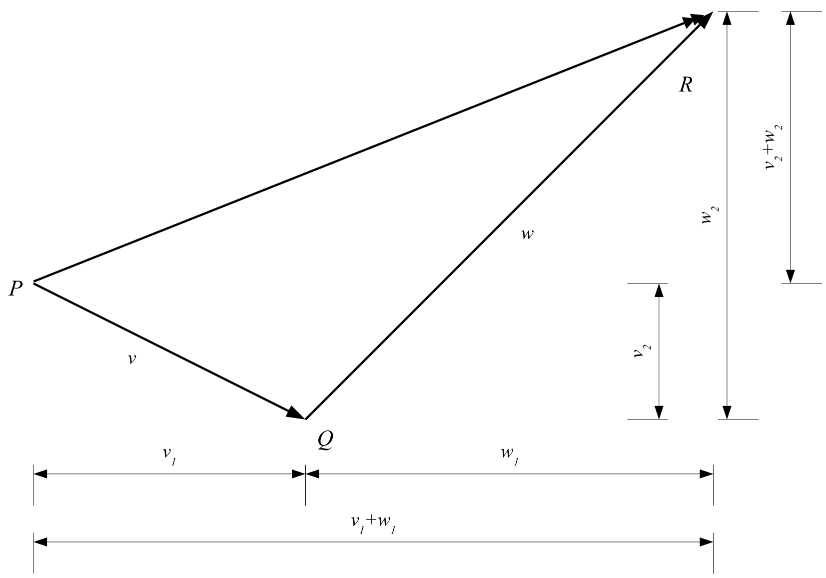


Figure 6: $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$

Consider the triangle PQR with $\boldsymbol{v} = \overrightarrow{PQ}$, $\boldsymbol{w} = \overrightarrow{QR}$. Then $\boldsymbol{v} + \boldsymbol{w} = \overrightarrow{PR}$; see Figure 6. In terms of components, if

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad \boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

then

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}.$$

It follows from the component description that vector addition satisfies the following properties:

$$\begin{aligned} \boldsymbol{v} + \boldsymbol{w} &= \boldsymbol{w} + \boldsymbol{v} && \text{(commutative law)} \\ \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) &= (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} && \text{(associative law)} \\ \boldsymbol{v} + \boldsymbol{0} &= \boldsymbol{0} + \boldsymbol{v} = \boldsymbol{v} \end{aligned}$$

Scalar multiplication

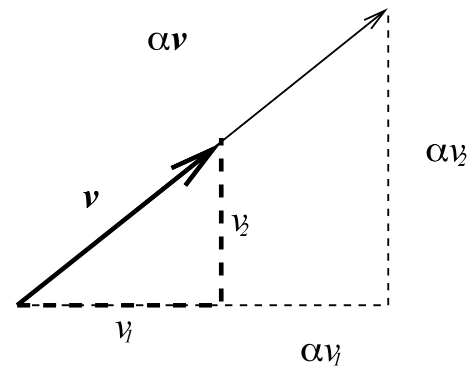


Figure 7: We can multiply the vector $\boldsymbol{v}$ by a number α (scalar).

If α is a real number (called a *scalar*), we define $\alpha\boldsymbol{v}$ to be the vector of norm

$$\|\alpha\boldsymbol{v}\| = |\alpha| \cdot \|\boldsymbol{v}\|$$

in the same direction as $\boldsymbol{v}$ if $\alpha > 0$, and opposite direction if $\alpha < 0$.

Using similar triangles it follows that if $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ then $\alpha\boldsymbol{v} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \end{pmatrix}$.

If we multiply any vector $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ by zero we obtain the zero vector:

$$0 \cdot \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \boldsymbol{0}$$

Unit vectors

A *unit vector* is a vector of norm 1. If $\boldsymbol{v} \neq \boldsymbol{0}$ is a vector, then

$$\frac{1}{\|\boldsymbol{v}\|}\,\boldsymbol{v}$$

is a unit vector in the direction of $\boldsymbol{v}$.

In particular

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

determine unit vectors along the x and y axes respectively.

For any vector $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ we have

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_2 \end{pmatrix} = v_1\boldsymbol{i} + v_2\boldsymbol{j},$$

Hence we can decompose $\boldsymbol{v}$ into a vector $v_1\boldsymbol{i}$ along the x-axis and $v_2\boldsymbol{j}$ along the y-axis. The numbers v_1 and v_2 are called the *components* of $\boldsymbol{v}$ with respect to $\boldsymbol{i}$ and $\boldsymbol{j}$.

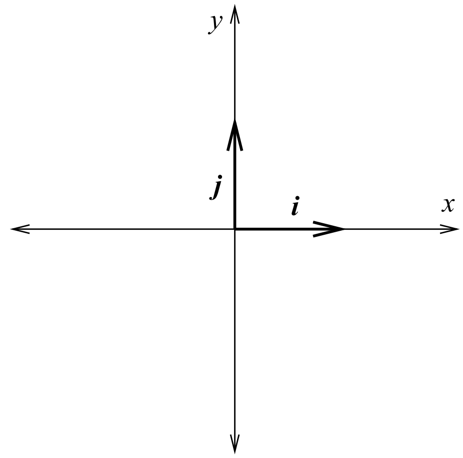


Figure 8: Unit vectors $\boldsymbol{i}$ and $\boldsymbol{j}$ in the x and y directions respectively.

Row and column vectors

We may write vectors using columns e.g. $\begin{pmatrix} a \\ b \end{pmatrix}$, or as row vectors $\begin{pmatrix} a & b \end{pmatrix}$. We usually use column vectors in this course.

Example 11. *An albatross is flying NE at 20 km/h into a 10 km/h wind in direction E60°S. Find the speed and direction of the bird relative to the ground.*

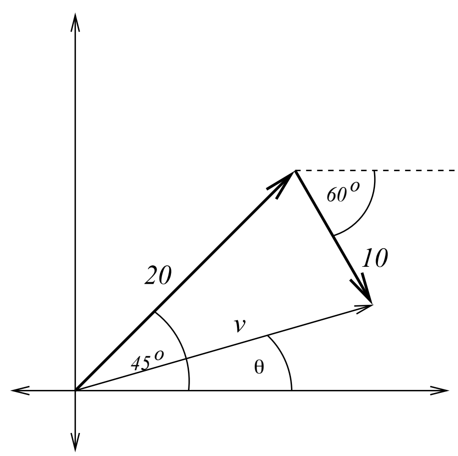


Figure 9: $\boldsymbol{v}$ gives the velocity vector of the albatross.

We have:

$$\begin{aligned} \mathbf{z} &= 20\cos 45^\circ \boldsymbol{i} + 20\sin 45^\circ \boldsymbol{j} \\ &= 10\sqrt{2}\boldsymbol{i} + 10\sqrt{2}\boldsymbol{j} \\ \mathbf{w} &= 10\cos 60^\circ \boldsymbol{i} - 10\sin 60^\circ \boldsymbol{j} \\ &= 5\boldsymbol{i} - 5\sqrt{3}\boldsymbol{j}. \end{aligned}$$

So $\mathbf{v} = \mathbf{z} + \mathbf{w} = (10\sqrt{2} + 5)\boldsymbol{i} + (10\sqrt{2} - 5\sqrt{3})\boldsymbol{j}$
Therefore the magnitude of $\mathbf{v}$ *is*

$$\begin{aligned}\|\boldsymbol{v}\| &= \sqrt{(10\sqrt{2}+5)^2+(10\sqrt{2}-5\sqrt{3})^2} \\ &= \sqrt{500+100\sqrt{2}(1-\sqrt{3})} \\ &\approx 19.91.\end{aligned}$$

To calculate the angle θ we use

$$\begin{aligned}\tan\theta &= \frac{10\sqrt{2}-5\sqrt{3}}{10\sqrt{2}+5} \\ \Rightarrow \theta &\approx 16^\circ\end{aligned}$$

So $\boldsymbol{v}$ = $19.9 kmh^{-1}$ at $E16^\circ N$.

Vectors in $\mathbb{R}^n$

We are familiar with vectors in two and three dimensional space, $\mathbb{R}^2$ and $\mathbb{R}^3$. These generalize to n dimensional space, denoted $\mathbb{R}^n$. A vector $\boldsymbol{v}$ in $\mathbb{R}^n$ is specified by n components:

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \qquad v_i \in \mathbb{R}.$$

We define addition of vectors and multiplication by a scalar component-wise. Thus if $\alpha \in \mathbb{R}$ is a scalar and $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$ is another vector then

$$\boldsymbol{v}+\boldsymbol{w} = \begin{pmatrix} v_1+w_1 \\ v_2+w_2 \\ \vdots \\ v_n+w_n \end{pmatrix} = \boldsymbol{w}+\boldsymbol{v}, \qquad \alpha\boldsymbol{v} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{pmatrix} \quad \alpha \in \mathbb{R},$$

We define the *dot* (or *scalar*) *product* between $\boldsymbol{v}$ and $\boldsymbol{w}$ as

$$\begin{aligned}\boldsymbol{v}\cdot\boldsymbol{w} &= v_1w_1+v_2w_2+\cdots\cdots+v_nw_n \\ &= \sum_{k=1}^{n} v_kw_k.\end{aligned}$$

The *length* or *norm* of the vector $\boldsymbol{v}$ is

$$\begin{aligned}\|\boldsymbol{v}\| &= \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}} \\ &= \sqrt{v_1^2+v_2^2+\cdots+v_n^2}\end{aligned}$$

Theorem 2. *Every vector $\boldsymbol{v}$ in $\mathbb{R}^n$ is expressible in terms of n special vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ called* coordinate vectors*, where*

$$\boldsymbol{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad i\textit{th position.}$$

Proof An arbitrary vector $\boldsymbol{v} \in \mathbb{R}^n$ can be written as

$$\begin{aligned} \boldsymbol{v} &= \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots\cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ v_n \end{pmatrix} \\ &= v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + \cdots\cdots + v_n\boldsymbol{e}_n. \end{aligned}$$

Notation

(1) E.g. in $\mathbb{R}^3$ we have

$$\boldsymbol{i} = \boldsymbol{e}_1\,, \quad \boldsymbol{j} = \boldsymbol{e}_2\,, \quad \boldsymbol{k} = \boldsymbol{e}_3.$$

(2) We have also the zero vector

$$\boldsymbol{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{satisfying}$$

$$\boldsymbol{0} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{0} = \boldsymbol{v}, \qquad 0 \cdot \boldsymbol{v} = \boldsymbol{0}, \qquad \text{for every vector } \boldsymbol{v}.$$

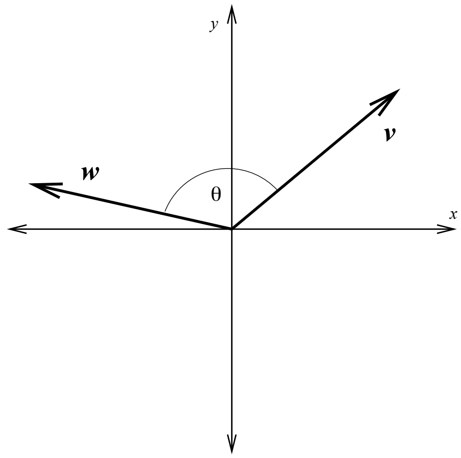
(3) If $\boldsymbol{v} \neq \boldsymbol{0}$, $\boldsymbol{w} \neq \boldsymbol{0}$, we say that vectors $\boldsymbol{v}$, $\boldsymbol{w}$ are *perpendicular* (or *orthogonal*) if

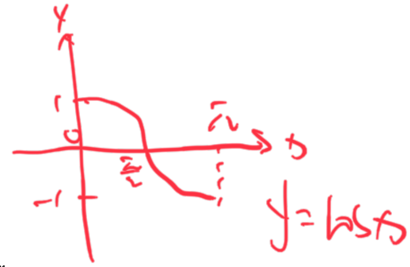
$$\boldsymbol{v} \cdot \boldsymbol{w} = 0.$$

1.6 Dot product, norm and distance

Dot product

For non zero vectors $\boldsymbol{v} = \overrightarrow{OP}$, $\boldsymbol{w} = \overrightarrow{OQ}$ the *angle between* $\boldsymbol{v}$ *and* $\boldsymbol{w}$ is the angle θ with $0 \leq \theta \leq \pi$ radians between $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ at the origin O; see Figure 10.





$y = \cos x$

$f(x) = \cos x$ is monotone on $x \in [0, \pi]$

Figure 10: θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$.

The *dot* (or *scalar* or *inner*) product of vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, denoted by $\boldsymbol{v} \cdot \boldsymbol{w}$, is the *number* given by

$$\boxed{\boldsymbol{v} \cdot \boldsymbol{w} = \begin{cases} 0\,, & \text{if } \boldsymbol{v} \text{ or } \boldsymbol{w} = \boldsymbol{0} \\ \|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\| \cos\theta\,, & \text{otherwise} \end{cases}}$$

$\cos\theta = \dfrac{v \cdot w}{\|v\| \cdot \|w\|}$

$\|v\| = \sqrt{v_1^2 + v_2^2 + v_3^2 \ldots}$

where θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$.

If $\boldsymbol{v}, \boldsymbol{w} \neq \boldsymbol{0}$ and $\boldsymbol{v} \cdot \boldsymbol{w} = 0$ then $\boldsymbol{v}$ and $\boldsymbol{w}$ are said to be *orthogonal* or *perpendicular*.

If $\boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j} + v_3\boldsymbol{k}$ and $\boldsymbol{w} = w_1\boldsymbol{i} + w_2\boldsymbol{j} + w_3\boldsymbol{k}$ are two vectors, then $\boldsymbol{v} \cdot \boldsymbol{w}$ is given by:

$$\boldsymbol{v} \cdot \boldsymbol{w} = v_1 w_1 + v_2 w_2 + v_3 w_3,$$

In particular, for $\boldsymbol{v} \in \mathbb{R}^3$,

$$\|\boldsymbol{v}\|^2 = \boldsymbol{v} \cdot \boldsymbol{v} = v_1^2 + v_2^2 + v_3^2.$$

Example 12. *Find the angle θ between the vectors:*

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ -5 \\ 4 \end{pmatrix}, \quad \boldsymbol{w} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$$

$$\boxed{\cos\theta = \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{\|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\|}.}$$

$$\boldsymbol{v} \cdot \boldsymbol{w} = 1 \cdot 3 - 5 \cdot 3 + 4 \cdot 3 = 0.$$

So $\cos\theta = 0$, so the vectors are perpendicular, i.e. $\theta = \dfrac{\pi}{2}$.

Example 13. *If $P = (2, 4, -1)$, $Q = (1, 1, 1)$, $R = (-2, 2, 3)$, find the angle $\theta = PQR$.*

Find vectors joining Q to P, and Q to R:

$$\overrightarrow{QP} = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix},$$

$$\overrightarrow{QR} = \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}.$$

$$\therefore \quad \left\|\overrightarrow{QP}\right\| = \left\|\overrightarrow{QR}\right\| = \sqrt{1+9+4} = \sqrt{14}.$$

$$\overrightarrow{QP} \cdot \overrightarrow{QR} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} = -3 + 3 - 4 = -4$$

$$\therefore \quad \cos\theta = \frac{\overrightarrow{QP} \cdot \overrightarrow{QR}}{\left\|\overrightarrow{QP}\right\| \cdot \left\|\overrightarrow{QR}\right\|} = \frac{-4}{14} \Rightarrow \theta \simeq 107^\circ.$$

The projection formula orthogonal projection

Fix a vector $\boldsymbol{v}$. Given another vector $\boldsymbol{w}$ we can write $\boldsymbol{w}$ as

$\vec{v}$ to

$$\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2$$

where:

Parallel to

- $\boldsymbol{w}_1$ is in the direction of $\boldsymbol{v}$ → meaning there $\vec{w_1} = \alpha\vec{v}$ $\alpha \in \mathbb{R}$
- $\boldsymbol{w}_2$ is perpendicular to $\boldsymbol{v}$. → meaning $W_2 \cdot V = 0$

See Figure 11. We want to find $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ in terms of $\boldsymbol{v}$ and $\boldsymbol{w}$.

Since $\boldsymbol{w}_1$ is in the direction of $\boldsymbol{v}$, let

$$\boldsymbol{w}_1 = \alpha\boldsymbol{v} \qquad \text{for some } \alpha \in \mathbb{R}.$$

Then $\boldsymbol{w}_2 = \boldsymbol{w} - \alpha\boldsymbol{v}$.

We need to choose α to make $\boldsymbol{v}$ and $\boldsymbol{w}_2$ orthogonal.

$$0 = \boldsymbol{w}_2 \cdot \boldsymbol{v} = (\boldsymbol{w} - \alpha\boldsymbol{v}) \cdot \boldsymbol{v} = \boldsymbol{w} \cdot \boldsymbol{v} - \alpha\boldsymbol{v} \cdot \boldsymbol{v} = \boldsymbol{w} \cdot \boldsymbol{v} - \alpha\|\boldsymbol{v}\|^2.$$

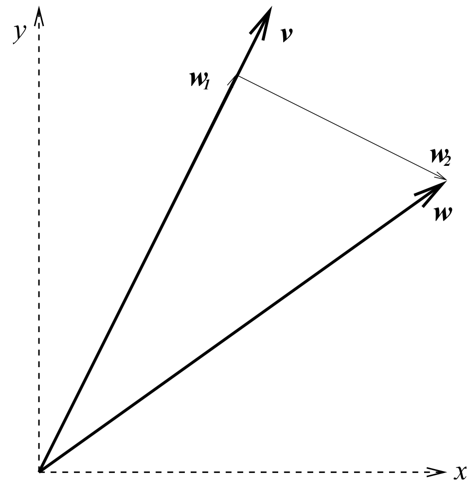


Figure 11: $\boldsymbol{w}$ can be decomposed into a component $\boldsymbol{w}_1$ in the direction of $\boldsymbol{v}$ and a component $\boldsymbol{w}_2$ perpendicular to $\boldsymbol{v}$.

So we need

$$\alpha = \frac{\boldsymbol{w} \cdot \boldsymbol{v}}{\|\boldsymbol{v}\|^2}.$$

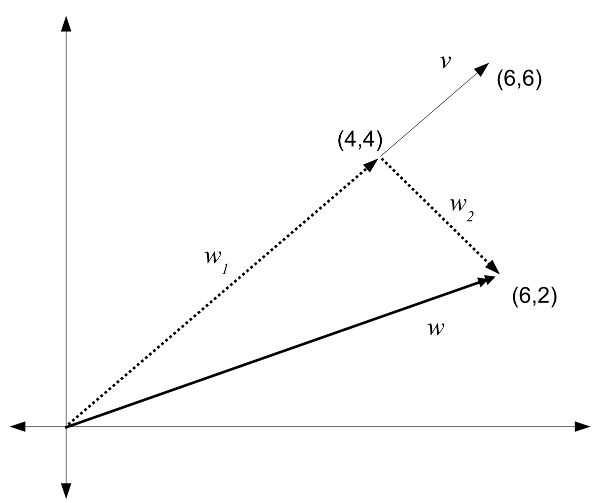
We have derived the *projection formula:*

$$\boldsymbol{w}_1 = \frac{(\boldsymbol{w} \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|^2}\boldsymbol{v}\ , \qquad \boldsymbol{w}_2 = \boldsymbol{w} - \frac{(\boldsymbol{w} \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|^2}\boldsymbol{v}. \tag{1}$$

Example 14. *Find the projection of* $\boldsymbol{w} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$ *onto* $\boldsymbol{v} = \begin{pmatrix} 6 \\ 6 \end{pmatrix}$.

$$\boldsymbol{w}_1 = \frac{\boldsymbol{w} \cdot \boldsymbol{v}}{\|\boldsymbol{v}\|^2}\boldsymbol{v} = \frac{36+12}{6^2+6^2}\begin{pmatrix} 6 \\ 6 \end{pmatrix} = \frac{2}{3}\begin{pmatrix} 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}.$$

$$\boldsymbol{w}_2 = \boldsymbol{w} - \boldsymbol{w}_1 = \begin{pmatrix} 6 \\ 2 \end{pmatrix} - \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ -2. \end{pmatrix}$$



Properties of dot product

If $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ then

(1) $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$.

(2) $\boldsymbol{v} \cdot (\boldsymbol{w} + \boldsymbol{u}) = \boldsymbol{v} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{u}$.

(3) $(\alpha \boldsymbol{v}) \cdot \boldsymbol{w} = \alpha(\boldsymbol{v} \cdot \boldsymbol{w}) = \boldsymbol{v} \cdot (\alpha \boldsymbol{w})$.

(4) $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$ and $\boldsymbol{v} \cdot \boldsymbol{v} = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}$.

1.7 Application

Concepts missing.

2 Sets, counting and cardinality

The terms *set* and *element* won't be formally defined here. Roughly speaking, a **set** will be a collection of objects called **elements**, and given any element and any set, we should be able to say whether the element belongs to the set or not.

2.1 Sets

- Recall that the notation $a \in S$ means that a is an element of the set S, or a belongs to S.
- We can list elements in a set using braces or curly brackets: $\{x_1, x_2, x_3\}$. The order in which we list the elements of a set is irrelevant, so $\{x_1, x_2, x_3\} = \{x_3, x_1, x_2\}$, etc.
  The number of times that each element is listed is also irrelevant; $\{a, b, a\} = \{a, b\}$ for example.
- The **empty** set is the set containing NO elements, denoted by $\emptyset$.
  (We'll see soon that "the" empty set is the right terminology, because we'll show that there is only one empty set.)
- For any set A, the **cardinality** of A is the number of elements in the set A; we shall denote this as $|A|$.

Note that in *Mathematica*, we can calculate with sets provided we use the Union function.

```
Union[{a,b,a}] = = Union[{a, b}]
MemberQ[{2,{2}},2]
```

Otherwise these will behave like vectors.

Example 15. **(a)** *How many elements does the set* $\{2, 2, \{2\}\}$ *have?*

(b) *Is it true that* $\{1, 1, 2\} = \{1, 2\}$ *?*

(c) *Is it true that* $1 \in \{1\}$ *?*

(d) *Is it true that* $1 \in \{\{1\}\}$ *?*

Example: *Describe the following sets in words.*

(a) $\{1, 2, \dots, 100\}$

(b) $\{x \in \mathbb{R} : x > 0\}$

(c) $\{y \in \mathbb{Z}^+ : -3 \leqslant y \leqslant 3\}$

- *If A and B are any sets, A is called a* **subset** *of B, written $A \subseteq B$, if and only if every element of A is also an element of B.*

- *For sets A and B, we say A is a* **proper subset** *of B if and only if $A \subseteq B$ and $A \neq B$. So A is a proper subset of B if and only if every element of A is also an element of B, and there is some element of B which is not in A.*

- *For sets A and B, we say sets A and B are* **equal**, *$A = B$, if and only if every element of A is in B, and also every element of B is in A.*

  So $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.

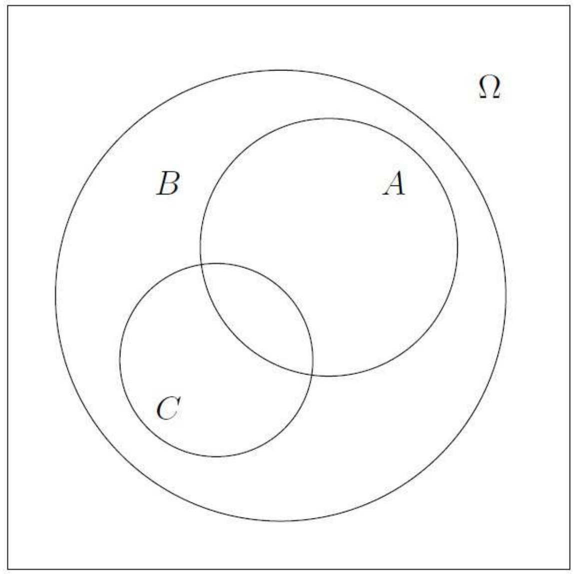


Figure 12: A Venn diagram with universe Ω.

Example 16. *Suppose $A = \{a, b, c, d\}$, $B = \{a, b, e\}$ and $C = \{a, b, c, d, e\}$. Answer the following, and give reasons.*

(a) *Is it true that $B \subseteq A$?* ✗

(b) *Is it true that $A \subseteq C$?* √

(c) *Is A a proper subset of C?*

(d) *Is it true that $B \subseteq B$?*

Example 17. *Draw a Venn diagram to represent the relationship between the following sets: $A = \{1, 2, 3\}$, $B = \{1, 4\}$, $C = \{2, 3\}$.*

Example 18. *True or false?*
(a) $\{4\} \in \{1, \{3\}, 4\}$

(b) $\{4\} \subseteq \{1, \{3\}, 4\}$

(c) $\{3\} \in \{1, \{3\}, 4\}$

(d) $1 \subseteq \{1, \{3\}, 4\}$

exam requirement : Verify or disprove $A = B$

Note the set notation used here; verbally read ":" as "such that".

Example 19. *Let* $A = \{x \in \mathbb{Z} : x = 4p - 1$ *for some* $p \in \mathbb{Z}\}$,
$B = \{y \in \mathbb{Z} : y = 4q - 5$ *for some* $q \in \mathbb{Z}\}$. *Prove that* $A = B$.

We investigate some **operations** on sets now, obtaining new sets from existing sets. Let U be some **universal set**, depending on the context.
(So U could perhaps be $\mathbb{R}$ in some contexts.)

In the following, suppose A and B are some subsets of a universal set U.

- The **union** of sets A and B, denoted $A \cup B$, is the set of all elements x in U such that $X \in A$ *or* $X \in B$ *(or both).*
- The **intersection** of sets A and B, denoted $A \cap B$, is the set of all elements x in U such that $x \in A$ *and* $x \in B$.
- The **set difference** of B minus A, denoted $B - A$, and sometimes also called the relative complement of A in B, is the set of all x in U such that $x \in B$ and $x \notin A$. Some texts write $B \setminus A$ instead of $B - A$.
- The **complement** of A, denoted A^c, is the set of all x in U such that $x \notin A$.

Summarising the above:

$$\begin{aligned} A \cup B &= \{x \in U : x \in A \text{ or } x \in B\}; \\ A \cap B &= \{x \in U : x \in A \text{ and } x \in B\}; \\ B - A &= \{x \in U : x \in B \text{ and } x \notin A\}; \\ A^c &= \{x \in U : x \notin A\} \end{aligned}$$

Example 20. *Let the universal set be* $\{1, 2, \ldots, 10\}$, *and let*
$A = \{1, 2, 3, 4\}$, $B = \{2, 4, 6, 8, 10\}$ *and* $C = \{1, 3, 5, 7, 9\}$.
Write down the following sets:

(a) $A \cap B =$

(b) $A \cup B =$

(c) $B \cup C =$

(d) $B - A =$

(e) $A - C =$

(f) $B^c =$ not in B

(g) $A^c =$ not in A

(h) $A^c \cup B =$ not in A + B

Definition 7. *Let n be a positive integer (so $n \in \mathbb{Z}^+$), and let $x_1, x_2, \ldots, x_n$ be n not necessarily distinct elements. The* **ordered n-tuple**, *denoted $(x_1, x_2, \ldots, x_n)$, consists of the n elements with their ordering: first x_1, then x_2, and so on up to x_n. (Note round brackets, not braces)*

An ordered 2-tuple is an **ordered pair**.

An ordered 3-tuple is an **ordered triple**.

Two ordered n-tuples $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ are equal if and only if $x_i = y_i$ for all i with $1 \leqslant i \leqslant n$.

Example 21. *Is it true that $(3, 1) = (1, 3)$?*

Example 22. *If $((-2)^2,\ y,\ \sqrt{9}) = (4,\ 3,\ z)$, find y and z.*

Definition 8. *Given two sets A and B, the* **Cartesian product** *of A and B, denoted $A \times B$, is the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$. (Note the word "all") So*

$$A \times B \ = \ \{(a, b) : a \in A \ \text{ and } \ b \in B\}.$$

Similarly

$$A_1 \times A_2 \times \cdots \times A_n \ = \ \{(a_1, a_2, \ldots, a_n) : a_i \in A_i,\ 1 \leqslant i \leqslant n\}.$$

A Cartesian product with which you are probably already familiar is the xy-plane. The points in the real xy-plane are the elements of the set $\mathbb{R} \times \mathbb{R}$.

If you plot $(3, 1)$ and $(1, 3)$ you'll see these ordered pairs are not equal.

Example 23. *Let $A = \{-1, 0, 1\}$ and $B = \{x, y\}$. Write out the set $A \times B$. How many elements are in $A \times B$?*

Recall that we denote a set with NO elements in it by $\emptyset$.

$\emptyset \times A = \emptyset$

- The empty set is a subset of every set. So if S is any set, we have $\emptyset \subseteq S$.
- Every set is a subset of itself. So if S is any set, we have $S \subseteq S$.

Theorem 3. *The empty set $\emptyset$ is unique.*

Proof We use a contradiction argument. So suppose that $\emptyset_1$ and $\emptyset_2$ are each sets with no elements. Since $\emptyset_1$ has no elements, it is a subset of $\emptyset_2$, that is, $\emptyset_1 \subseteq \emptyset_2$. Also since $\emptyset_2$ has no elements, we have $\emptyset_2 \subseteq \emptyset_1$. Thus $\emptyset_1 = \emptyset_2$, by definition of set equality.

Partitions of sets

- Two sets are called **disjoint** if and only if they have no elements in common.

  So A and B are disjoint if and only if $A \cap B = \emptyset$.

- Sets $A_1, A_2, \ldots, A_n$ are **mutually disjoint** (or **pairwise disjoint**) if and only if for all pairs of sets A_i and A_j with $i \neq j$, their intersection is empty; that is, if and only if $A_i \cap A_j = \emptyset$ for all $i, j = 1, 2, \ldots, n$ with $i \neq j$.

- A collection of non-empty sets $\{A_1, A_2, \ldots, A_n\}$ is a **partition** of a set A if and only if

  (1) $A = A_1 \cup A_2 \cup \cdots \cup A_n$ and

  (2) the sets $A_1, A_2, \ldots, A_n$ are mutually disjoint.

Example 24. *Determine whether the following statements are true or false.*

(a) $\emptyset = \{\emptyset\}$ → not empty

(b) $A \cup \emptyset = A$

(c) $A \cap A^c = \emptyset$

(d) $A \cup A^c = \emptyset$

(e) $A \cap \emptyset = \emptyset$

(f) $(A - B) \cap B = \emptyset$

(g) $\{a, b, c\}$ *and* $\{d, e\}$ *are disjoint sets.*

(h) $\{1, 2\}$, $\{5, 7, 9\}$ *and* $\{3, 4, 5\}$ *are mutually disjoint sets.*

Example 25. *Let*

$$A_1 = \{n \in \mathbb{Z} : n < 0\}, \qquad A_2 = \{n \in \mathbb{Z} : n > 0\}.$$

exam

Is $\{A_1, A_2\}$ a partition of $\mathbb{Z}$? If so, explain why; if not, see if you can turn it into a partition with a small change.

not cover zero

Example 26. *Find a partition of $\mathbb{Z}$ into four parts such that none of the four parts is finite in size.*

Definition 9. *Given a set X, the* **power set** *of X is the set of all subsets of X.*
It is denoted by $\mathcal{P}(X)$.

Example 27. *If $B = \{1, 2, 3\}$, write down the set $\mathcal{P}(B)$.*

Example 28. *Let $X = \emptyset$. Write down $\mathcal{P}(X)$, and $\mathcal{P}(\mathcal{P}(X))$.*

If $|S| = n$, how many elements does the power set $\mathcal{P}(S)$ have?

Subset relations

- For all sets A and B, $A \cap B \subseteq A$ and $A \cap B \subseteq B$.

- For all sets A and B, $A \subseteq A \cup B$ and $B \subseteq A \cup B$.

- For all sets A, B and C,

  if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

In the following, U denotes some universal set, and A, B and C are any subsets of U.

Set Identities

- $A \cup B = B \cup A$ and $A \cap B = B \cap A$. **(commutative)**
- $(A \cup B) \cup C = A \cup (B \cup C)$ and $(A \cap B) \cap C = A \cap (B \cap C)$. **(associative)**
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. **(distributive)** ☆ try to prove this
- $A \cup \emptyset = A$; $A \cap U = A$; $A \cup A^c = U$; $A \cap A^c = \emptyset$;
  $(A^c)^c = A$; $A \cup A = A$; $A \cap A = A$.
- $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$. **(De Morgan's laws)**

To prove $X = Y$:

First show that $X \subseteq Y$, and then that $Y \subseteq X$.

So take any x in X, and show that then $x \in Y$. This shows $X \subseteq Y$.

Next take any $y \in Y$, and show that then $y \in X$. This shows $Y \subseteq X$.

From these results we conclude that $X = Y$.

Note that: $(A \cap B)^c = A^c \cup B^c$ (one of De Morgan's laws).

Use De Morgan's law to show that: For all sets A and B, $(A \cap B)^c \ = \ A^c \cup B^c$.

First, let $x \in (A \cap B)^c$. Then $x \notin A \cap B$.

So it is *false* that "x is in A and x is in B".

Thus $x \notin A$ or $x \notin B$ (since $\sim(p \wedge q) \Leftrightarrow (\sim p) \vee (\sim q)$).

So $x \in A^c$ or $x \in B^c$.

Therefore $x \in A^c \cup B^c$. This shows that $(A \cap B)^c \ \subseteq \ A^c \cup B^c$.

To complete this example, we must show that $A^c \cup B^c \ \subseteq \ A \cap B)^c$:

- Note that $(x, y) \in A \times B$ is equivalent to "$(x \in A)$ and $(y \in B)$".

Example 29. *If $A \subseteq C$ and $B \subseteq D$, prove that $A \times B \subseteq C \times D$.*

2.2 Complex numbers

You may have heard that it is not possible to take the square root of a negative number. This is false. More precisely,

> The square root of a negative number is not real; e.g. $\sqrt{-17} \notin \mathbb{R}$.

We deal with such numbers differently and refer to them as **imaginary numbers**. We use i to denote the square root of -1. For example, $\sqrt{-17} = \sqrt{17}i$, where $i^2 + 1 = 0$. A natural extension of the set of real numbers $\mathbb{R}$ is the set of **complex numbers**, $\mathbb{C}$, given by

$$\mathbb{C} = \{a + bi \ : \ a, b \in \mathbb{R}, i^2 + 1 = 0\}. \tag{2}$$

This is an extension of degree 2 over the real numbers, essentially meaning that this is structurally equivalent to the Cartesian product $\mathbb{R} \times \mathbb{R}$. To facilitate the description of complex numbers and their properties, we often represent these numbers geometrically as points in a 2-dimensional space with the real part of $z = a + bi$, a on the horizontal axis and the imaginary part b on the vertical axis. See Figure 13.

modulus:
$|z| = \sqrt{a^2+b^2}$
conjugate
$\bar{z} = a-bi$

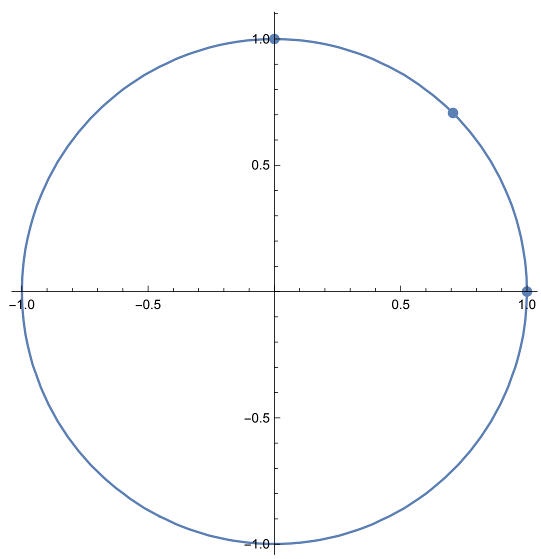


Figure 13: The complex plane where we plot $a + bi$, the circle $a^2 + b^2 = 1$, and the three complex numbers $1 + 0i$, $0 + 1i$, and $\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i$, all on the unit circle.

Definition 10. *The modulus and argument of the complex number $z = a + bi$ are respectively $|z| = \sqrt{a^2 + b^2}$, the distance from 0 also the absolute value of z, and the angle θ between the position of the point (a, b) and the real axis. We denote $\Re(z) = a$ and $\Im(z) = b$, the real and imaginary parts of z respectively. The* **conjugate** *of $z = a + bi$ is $\overline{z} = a - bi$.*

Complex numbers have the following important properties among several more:

1. $z_1 = z_1$ if and only if $\Re(z_1) = \Re(z_2)$ and $\Im(z_1) = \Im(z_2)$;

2. $z + \overline{z} = 2\Re(z)$;

3. $z_1 + z_2 = z_2 + z_1$ and $z_1 z_2 = z_2 z_1$;

4. Euler's formula: If $r = |z|$ and θ is the argument of z, then $z = re^{i\theta} = r\cos\theta + ir\sin\theta$. This is also known as polar form.

An important result in algebra states essentially that are all we need when solving equations

Theorem 4 (Fundamental theorem of algebra)**.** *Every non-constant polynomial $f(x)$ with complex coefficients has at least one complex root. Further, if the degree of $f(x)$ is equal to n, then there are exactly n such roots of $f(x)$, provided we count these as distinct when they are repeated.*

Example 30. *Find all of the roots in $\mathbb{C}$ of the polynomial $f(x) = x^5 - 3x^4 - 7x^3 - 13x^2 - 9x - 5$, determine the number of such roots according to Theorem 4. How many real roots of $f(x)$ are there? Can $f(x)$ be factorised over $\mathbb{R}$? Express each of the complex roots of $f(x)$ in polar form.*

2.3 Counting and elementary combinations

In this chapter we give an introduction to counting and probability. At first thought counting may seem a fairly simple exercise. However, when we are counting certain things it can turn out to be quite an involved process.

Counting subsets of a set: combinations

In this section we'll investigate questions of the following form.

> Given a set S with n elements, how many subsets of size r can be chosen from S?

- Let n and r be nonnegative integers with $r \leqslant n$. An r**-combination** of a set of n elements is a subset of the n elements of size r.

  The symbol $\binom{n}{r}$, which is read "n **choose** r," denotes the number of subsets of size r (so the number of r-combinations) which can be chosen from a set of n elements.

  Alternative notation for $\binom{n}{r}$ includes: $C(n, r)$, or ${}_nC_r$, or $C_{n,r}$, or ${}^nC_r$.

Theorem 5. *The number of subsets of size r (or r-combinations) which can be chosen from a set of n elements, $\binom{n}{r}$, is given by the formula*

$$\binom{n}{r} = \frac{P(n, r)}{r!}$$

or, equivalently,

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

where n and r are nonnegative integers with $r \leqslant n$.

- An **ordered selection** of r elements from a set of n elements is an r-permutation of the set (use $P(n, r)$).
- An **unordered selection** of r elements from a set of n elements is the same as a subset of size r, or an r-combination of the set (use $\binom{n}{r}$).

Example 31. *Calculate the value of:*

(a) $\binom{9}{3}$; $\qquad \frac{9!}{3!(9-3)!}$

(b) $\binom{200}{198}$;

(c) $\binom{8}{4}$.

Example 32. *A student has a maths assignment with five questions on it, but only has enough time to complete three of them. How many combinations of questions could the student complete?*

Example 33. *Imagine a word game in which a sentence has to be made using three words drawn out of a bag containing ten words.*

(a) *How many possible ways are there to choose three words from a bag of ten words?*

(b) *Suppose that the rules of the game change so that the sentence has to use the three words in the order in which they are chosen. How many possible combinations are there now?*

(c) *What is the relationship between the answers to parts (a) and (b)?*

Example 34. *In a game of straight poker, each player is dealt five cards from an ordinary deck of 52 cards, and each player is said to have a* 5*-card hand.*

(a) *How many* 5*-card poker hands contain four cards of the one denomination? (So e.g. four aces, or four threes, etc.)*

(b) *Find the error in the following calculation of the number of* 5*-card poker hands which contain at least one jack. Then calculate the true number of* 5*-card poker hands which contain at least one jack.*

Consider this in two steps:

Choose one jack from the four jacks.

Choose the other four cards in the hand.

Thus there are $\binom{4}{1}\binom{51}{4} = 999600$ *such hands.*

Watch out for the common error of counting things twice.

Wi will now work with some useful relationships involving $\binom{n}{r}$.

e x a m

Theorem 6. *Let n and r be positive integers with $r \leqslant n$. Then*

$$\binom{n}{r} = \binom{n}{n-r}.$$

$LHS = \frac{n!}{r!(n-r)!} = RHS$

Example 35. *Given that* $\binom{n}{2} = \frac{n(n-1)}{2}$, *find an expression for* $\binom{x+3}{x+1}$.

Pascal's Formula Let n and r be positive integers with $r \leqslant n$. Then

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}. \tag{3}$$

$\to$ Proof

Example 36. *Use Pascal's formula* (3) *to calculate:*

$RHS = \frac{n!}{(r-1)! \cdot (n-r+1)!} + \frac{n!}{r!(n-r)!}$

$= \frac{n! \cdot r}{r!(n-r+1)!} + \frac{n!(n-r+1)}{r!(n-r+1)!} = \ldots$

(*a*) $\binom{7}{5} + \binom{7}{6}$

(*b*) $\binom{9}{6} + \binom{9}{5}$

(*c*) $\binom{4}{2} + \binom{4}{3}$

(*d*) $\binom{6}{1} + \binom{6}{2}$

2.4 Cardinality

In this section we shall investigate the concept of the *cardinality* of a set and show that there are *infinite* sets that are larger than other infinite sets. This concept has applications in determining what can and what cannot be computed on a computer.

- A **finite** set is either one which has no elements at all, or one for which there exists a one-to-one correspondence (bijection) with a set of the form $\{1, 2, 3, \ldots, n\}$ for some fixed positive integer n.

- An **infinite** set is a nonempty set for which there does *not* exist any one-to-one correspondence (bijection) with a set of the form $\{1, 2, 3, \ldots, n\}$ for any positive integer n.

- Let $\mathcal{A}$ and $\mathcal{B}$ be any sets. Sets $\mathcal{A}$ and $\mathcal{B}$ are said to have the **same cardinality** if and only if there exists a one-to-one correspondence (bijection) from $\mathcal{A}$ to $\mathcal{B}$.

  In other words, $\mathcal{A}$ has the **same cardinality** as $\mathcal{B}$ if and only if there is a function f from $\mathcal{A}$ to $\mathcal{B}$ that is one-to-one (injective) and onto (surjective).

- A set is called **countably infinite** if and only if it has the same cardinality as the set of positive integers $\mathbb{Z}^+$.
- A set is called **countable** if and only if it is finite or it is countably infinite.

Example Consider: $f : \mathbb{Z} \mapsto \mathbb{Z}^2 \qquad f(x) = x^2$

Here we can see f is surjective but not injective, so not a bijection.

Example 37. *Show that the following is bijective:* $f : \mathbb{Z}^+ \mapsto \mathbb{Z}^2 \qquad f(x) = x^2$

Start by assuming $f(x_1) = f(x_2)$. Then $x_1^2 = x_2^2$, and taking square roots gives $x_1 = x_2$ since $\mathbb{Z}^+$ contains only positive integers. To show surjective, let $x^2 \in \mathbb{Z}^2$, then $x \in \mathbb{Z}^+$ and $f(x) = x^2$.

Example 38. *The sets* $\{1, 4, 5, 6, b\}$*,* $\mathbb{Z}^{>0}$*,* $\mathbb{Z}$*, and* $\mathbb{Q}$ *are all countable.*

- A set that is not countable is called **uncountable**.

Example 39. *The sets* $\mathbb{R}$*, and* $\mathcal{P}(\mathbb{Z}(> 0))$ *are both uncountable.*

Example 40. *Show that the set of all odd integers is countable.*

Theorem 7. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be* **finite** *sets with the same number of elements, and suppose that* f *is a function from* $\mathcal{X}$ *to* $\mathcal{Y}$*. Then* f *is one-to-one if and only if* f *is onto.*

- This theorem does *not* hold for infinite sets of the same cardinality.

  In fact if $\mathcal{A}$ and $\mathcal{B}$ are infinite sets with the same cardinality, then there exist functions from $\mathcal{A}$ to $\mathcal{B}$ that are one-to-one and not onto, and functions from $\mathcal{A}$ to $\mathcal{B}$ that are onto and not one-to-one.

Example 41. *Given that* $\mathbb{Z}$ *has the same cardinality as the set of even integers,* $\mathbb{Z}^{even}$*,*

(a) *find a map from* $\mathbb{Z}^{even}$ *to* $\mathbb{Z}$ *that is one-to-one but not onto, and*

(b) *find a map from* $\mathbb{Z}$ *to* $\mathbb{Z}^{even}$ *that is onto but not one-to-one.*

Example 42. *Verify that the set* $\mathcal{P}(\mathbb{Z}^+)$ *is uncountable.*

2.5 Application

3 Foundations in logic

3.1 Proof methods

In this section we will focus on the basic structure of simple mathematical proofs, and see how to disprove a mathematical statement using a counterexample.

To illustrate these proof techniques we will use the properties of *even* and *odd* integers, and of *prime* and *composite* integers.

- An integer n is **even** if and only if n is equal to two times some integer.
- An integer n is **odd** if and only if n is equal to two times some integer plus 1.
- An integer n is **prime** if and only if $n > 1$, and for all positive integers r and s, if we have $n = r \cdot s$, then $r = 1$ or $s = 1$.
- An integer is **composite** if and only if $n > 1$, and $n = r \cdot s$ for some positive integers r and s with $r \neq 1$ and $s \neq 1$.

Example 43. *Prove that for all* $x \in \{0, 1, 2, 3, 4, 5\}$, *the integer* $x^2 + x + 41$ *is a prime number.*

Method of Direct Proof:

To show that "$\forall x \in D$, if $P(x)$ then $Q(x)$" is **true**:

1. Suppose for a particular but *arbitrarily chosen* element x of D that the hypothesis $P(x)$ is true. (This step is often abbreviated "Suppose $x \in D$ and $P(x)$.")
2. Show that the conclusion $Q(x)$ is true using definitions, previously established results, and the rules for logical inference.

Example 44. *Prove that for all integers* a, b, c *and* m, *if*

$$a - b = rm \quad \textit{and} \quad b - c = sm, \quad \textit{then} \quad a - c = tm$$

for some integers r, s *and* t.

The following are common **mistakes** that are often made in proofs; they should be avoided.

- Arguing from examples.
- Using the same letter to mean two different things.
- Jumping to a conclusion.
- Begging the question (assuming the thing you are trying to prove).

- Misusing the word 'if'.

Example 45. *Consider the statement that the product of any two odd integers is an odd integer. The following "proof" of this statement is incorrect as it is 'begging the question'.*

NOT a Proof:

Suppose that m and n are odd integers.
If mn is odd, then $mn = 2k + 1$ for some integer k.

By the definition of odd, $m = 2a + 1$ and $n = 2b + 1$ for some integers a and b.

Thus $mn = (2a + 1)(2b + 1) = 2k + 1$, which is by definition odd. This is the statement which was to be shown.

Example 46. *Correctly prove the statement: the product of any two odd integers is an odd integer.*

Disproof by Counterexample:

To show that "$\forall x \in D$, if $P(x)$ then $Q(x)$" is **false**, find a value of $x \in D$ for which $P(x)$ is true and $Q(x)$ is false.

Example 47. Disprove the following statement:
If n is an even integer then $1 + 2 + 3 + \ldots + (n - 1) = kn$ for some integer k.
(Note that this statement is true for odd integers).

How to format a proof:

1. Write the theorem to be proved.
2. Clearly mark the beginning of the proof with the word "**Proof**".
3. Make your proof self-contained.
4. Write proofs in complete English sentences.
5. Conclude by stating what it is you have proved.

We continue our discussion of proof techniques now by considering the study of the rational numbers, that is, quotients of integers.

A real number is **rational** if and only if it can be expressed as a quotient of two integers with a *nonzero* denominator.

The set of all rational numbers is denoted by $\mathbb{Q}$.

A real number that is not rational is **irrational**.

$$r \text{ is rational} \iff \exists \text{ integers } a \text{ and } b \text{ such that } r = \frac{a}{b} \text{ and } b \neq 0.$$

Example 48. *Determine the truth values of the following statements:*

(a) (0 is rational) $\wedge$ (0.377777 . . . is rational).

(b) ($\sqrt{7}$ is rational) $\vee$ ($\sqrt{25}$ is rational).

(c) $\forall x \in \mathbb{R}$, if $3 \leqslant x \leqslant 4$ then x is rational.

Example 49. *Prove that the product of two rational numbers is a rational number.*

Example 50. *Prove that every rational number r has an additive inverse. (In other words, prove that for every rational number r, there exists another rational number s such that $r + s = 0 = s + r$.)*

Example 51. *Prove for $a \in Q$ has a unique representative with* $\gcd(p, q) = 1$.

Example 52. *Prove that every non-zero rational number r has a multiplicative inverse.*

We now describe exactly what it means to say that one integer **divides** another integer. One of the most important theorems in number theory will also be introduced, the **Unique Factorization Theorem**.

- If n and d are integers and $d \neq 0$, then n is **divisible** by d if and only if there exists some integer k such that $n = dk$.

  Alternatively, we say that:

  n is a **multiple of** d, or

  d is a **factor of** n, or

  d is a **divisor of** n, or

  d **divides** n.

- The notation $d \mid n$ is used to represent the predicate "d divides n".
- d **does not divide** n (denoted $d \nmid n$) if and only if $\frac{n}{d}$ is not an integer.

Warning: Note the difference between "$d \mid n$" and "d/n".

Example 53. Explain your answers to the questions below:

(a) *Is it true that* $4 \mid 72$*?*

(b) *Is* 24 *a multiple of* 48*?*

(c) *Is it true that* $0 \mid 5$*?*

(d) *Is* -3 *a factor of* 9*?*

Is it true that $6 \mid 2a(3b+3)$, for all $a, b \in \mathbb{Z}$? Explain.

Is $2a(4b+1)$ a multiple of 4, for all $a, b \in \mathbb{Z}$? Explain.

- An alternative definition of a prime number is:

  An integer $n > 1$ is **prime** if and only if
  its only positive integer divisors are 1 and itself.

Theorem 8 (Unique Factorization for the Integers)**.** *Given any integer $n > 1$, there exist: a positive integer k; distinct prime numbers $p_1, p_2, \ldots, p_k$; and positive integers $e_1, e_2, \ldots, e_k$, such that*

$$n = p_1^{e_1} p_2^{e_2} p_3^{e_3} \ldots p_k^{e_k},$$

and any other expression of n as a product of prime numbers is identical to this, except perhaps for the order in which the terms are written.

Example 54. *Find the unique factorization of the following integers by trial division.*

(a) 5440

(b) 43560

Suppose that k, a and b are integers.
If $k \mid a$ and $k \mid b$, prove that $k \mid (a+b)$.

Division into cases and the quotient-remainder theorem

In this section, we describe another important theorem in number theory: the **Quotient-Remainder Theorem**. We shall also encounter situations where it's easier to prove a statement by splitting the statement into cases.

Theorem 9 (Quotient-Remainder Theorem)**.** *Given any integer n and a positive integer d, there exist unique integers q and r such that*

$$n \;=\; dq + r \;\; \textit{and} \;\; 0 \leqslant r < d.$$

- Given an integer n and a positive integer d such that $n = dq + r$, where $0 \leqslant r < d$, we define

  $$n \textbf{ mod } d \;=\; r.$$

- For integers a and b, and a positive integer d, if $a \equiv r \pmod{d}$ and $b \equiv r \pmod{d}$ (so if a and b leave the same remainder upon division by d), then we say that "a is **congruent** to b **modulo** d " and write

  $$a \equiv b \pmod{d}.$$

  Note that this is the same as saying $a - b = kd$ for some integer k,
  or equivalently, $d \mid (a-b)$.

  Note: Working "$\pmod{d}$" we always assume that $d > 0$.

- If n is divisible by d, then $n \equiv 0 \pmod{d}$.

Example 55. *True or false? (Explain your answers.)*

(a) $7 \equiv 31 \pmod{6}$

(b) $-2 \equiv 8 \pmod{5}$

(c) $-27 \equiv 27 \pmod{10}$

Example 56. *Given the following values for n and d, find integers q and r such that $n = d \cdot q + r$ and $0 \leq r < d$.*

(a) $n = 102$ *and* $d = 11$.

(a) $n = -4$ *and* $d = 5$.

Example 57. *If a and b are integers such that $a = 4x + 1$ and $b = 4y + 1$ for some $x, y \in \mathbb{Z}$, then prove that the product ab is of the form $4m + 1$, for some integer m.*

Example 58. *For the positive integers u, v, w, x and d, if $u \equiv v \pmod{d}$ and $w \equiv x \pmod{d}$, prove the following two statements.*

(a) $u + w \equiv v + x \pmod{d}$.

(b) $uw \equiv vx \pmod{d}$.

Example 59. *The square of any odd integer has the form $8m + 1$ for some integer m.*

3.2 Quantified statements

We will now extend our knowledge of symbolic representation of statements to include quantified statements, that is, statements which include words such as *every*, *each*, *some*. For example:

- *every* number is either positive or negative;
- *some* integers are perfect numbers.

Many of the examples in this section refer to sets of numbers so you need to be familiar with some common notation.

A set is usually denoted by an upper case (capital) letter, and elements of the set by lower case letters. To list the elements of a set, we use curly brackets or *braces*. For example, the set of even integers between 1 and 11 can be written as $E = \{2, 4, 6, 8, 10\}$. The symbol $\in$ is used to indicate that an element belongs to a set and the symbol $\notin$ is used to indicate that an element does not belong to a set. For example, $4 \in E$ (read "four belongs to E") and $8 \in E$, but $5 \notin E$.

- $\mathbb{Z}$ denotes the set of integers. These are the positive and negative whole numbers and zero: $\{\dots, -2, -1, 0, 1, 2, \dots\}$.
- We use $\mathbb{Z}^+$ to denote the positive integers, $\{1, 2, 3, \dots\}$.
- $\mathbb{Q}$ denotes the set of rational numbers. These are the numbers that can be written as a quotient of integers, a/b, where a and b are integers and b *is nonzero*. All terminating or repeating decimals are rational numbers.
- $\mathbb{R}$ denotes the set of all real numbers. This includes all the rational numbers and all the irrational numbers (non-terminating and non-repeating decimals).

A **predicate** is a sentence that contains a finite number of variables; it becomes a statement when the variables are replaced with specific values.

Example 60. *In the following, x, a and b are integers. Which of the following are predicates?*

- *x is a positive integer.*
- *Please don't eat that.*
- *a is a factor of b.*
- *2 divides x and x divides 6.*
- *My toy elephant is grey.*
- *Paul is 20 years old.*

Predicates are often denoted by an upper case letter followed by variables listed within brackets. For example, the predicate "x is a multiple of 10" might be denoted by $P(x)$.

- The **domain** of a predicate variable is the set of all values that may be substituted in place of the variable. The set of all such elements that make the predicate true is called the **truth set** of the predicate.

Example 61. *Let $Q(n)$ be the predicate:*
n is a factor of 15.
Find the truth set of $Q(n)$ if the domain of n is the set of integers $\mathbb{Z}$.

One way to make a predicate into a statement is to substitute a value for each variable. Another way is to add quantifiers.

- The symbol $\forall$ denotes "for all" (or for each, or for every), and is called the **universal quantifier**. Let $Q(x)$ be a predicate and D be the domain of x. The **universal statement**

  $$\forall\, x \in D,\ Q(x)$$

  is true if and only if $Q(x)$ is true for every x in D.
  It is false if and only if $Q(x)$ is false for at least one x in D.

Example 62. **(a)** *Translate the English sentence "All squares are rectangles" into a universal statement.*

(b) *Translate the universal statement "$\forall\, x \in \mathbb{Z},\ x \in \mathbb{R}$" into an English sentence.*

Example 63. *Determine whether the following statements are true or false.*

(a) $\forall\, x \in \mathbb{R},\ x^2 = 2$

(b) $\forall\, x \in \{1, 2, 3\},\ x^2 < 10$

- *The symbol $\exists$ denotes "there exists" (or there is, or there are), and is called the* **existential quantifier**. *Let $Q(x)$ be a predicate and D be the domain of x. The* **existential statement**

$$\exists\, x \in D \textit{ such that } Q(x)$$

is true if and only if $Q(x)$ is true for at least one x in D. It is false if and only if $Q(x)$ is false for every x in D.

Example 64. **(a)** *Translate the English sentence "There is a real number that is also a rational number" into an existential statement.*

(b) *Let E be the set of all elephants. Translate the existential statement "$\exists\, x \in E$ such that x is white" into an English sentence.*

Example 65. *Determine whether the following statements are true or false.*

(a) $\exists\, x \in \mathbb{R}$ *such that* $x^2 = 2$

(b) $\exists\, x \in \{1, 2, 3\}$ *such that* $x > 4$

One of the most important statement forms in mathematics is:

$$\forall x,\ \text{ if } P(x) \text{ then } Q(x)$$

or equivalently,

$$\forall\, x,\ (P(x) \Rightarrow Q(x))$$

Example 66. **(a)** *Rewrite the following formal statement as an English sentence.*

$$\forall\, x \in \mathbb{R},\ \textit{ if } x^2 < 9 \textit{ then } x < 3.$$

(b) *Rewrite the following English sentence as a formal statement.*
If a number is an integer, then it is a rational number.

3.3 Indirect argument: contradiction and contraposition

In this section two powerful methods of proof are introduced: **contradiction** and **contraposition**.

These are very useful alternatives to the method of direct proof which we met earlier in this chapter.

Method of Proof by Contradiction:

- Assume that the statement to be proved is false.
  Recall that the negation of the statement $\forall\, x \in D$*, if* $P(x)$ *then* $Q(x)$ *is* $\exists\, x \in D$ *such that* $P(x)$ *and* $\sim Q(x)$*.*
- Show that this assumption leads logically to a contradiction.
- Conclude that the statement to be proved is true.

The following form of *indirect* argument is based on the logical equivalence between a statement and its contrapositive.

Method of Proof by Contraposition

- Write the statement in the form:

$$\forall\, x \in D, \text{ if } \ P(x) \ \text{ then } \ Q(x).$$

- Rewrite this statement in the contrapositive form:

$$\forall\, x \in D, \text{ if } \ \sim Q(x) \ \text{ then } \ \sim P(x).$$

- Prove the contrapositive by direct proof.
  * Suppose x is a particular (but arbitrarily chosen) element of D such that $Q(x)$ is false.
  * Show that $P(x)$ is false.

Example 67. *Prove the following statement by contradiction:*
For all integers n *and all prime numbers* p*, if* n^2 *is divisible by* p*, then* n *is divisible by* p*.*

Example 68. *Prove the following statement by contraposition:*
For all integers n*, if* n^2 *is odd, then* n *is odd.*

Example 69. *Prove the product of any nonzero rational number and any irrational number is irrational using either contradiction or contraposition.*

Example 70. *Every prime* $p > 3$ *satisfies* $p \equiv \pm 1 \pmod 6$.

3.4 Mathematical induction

Technique: proof by indn
→ to prove $P(n)$ for $n = 1, 2 \ldots$
→ prove $P(1)$
→ For any $k \geq 1$
assume $P(1), \ldots P(k)$
and prove $P(k+1)$

The Principle of Mathematical Induction is an extremely useful tool for proving statements about sums of sequences, about properties of integers, and about any repeated events which can be expressed in terms of consecutive integers. We'll use it to *prove* statements involving integers n, such as:

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}; \quad \text{and} \quad 2^{2n-1} + 1 \text{ is divisible by } 3.$$

- **Principle of Mathematical Induction**
  Let $P(n)$ be a statement defined for integers n, and let a be some fixed integer.
  (Often in examples a will be 0 or 1 or a small value)
  Suppose that:
  (1) $P(a)$ is a true statement; and
  (2) For all integers $k \geqslant a$, IF $P(k)$ is a true statement, then $P(k+1)$ is a true statement.
  Then the statement $P(n)$ is true for all integers n with $n \geqslant a$.

- In the two steps above, (1) is the **basis step** and (2) is the **inductive step**.

  In the inductive step you ASSUME that $P(k)$ is true, and then do some work and SHOW that $P(k+1)$ is true.

- The supposition that $P(k)$ is true is called the **inductive hypothesis**.

Consider the analogy of a ladder, with a series of rungs.

The bottom rung is rung number a; you check that rung a is really there. Then you assume rung k is there for some k where $k \geqslant a$, and do some work to show that rung $k+1$ is also there, so you can climb up one rung of your ladder from k to $k+1$.

And starting with $k = a$ you can climb to $a+1$, then on to $a+2$, and so on. So you can get to *any* rung starting from the bottom rung a.

Now check the details of the next worked example.

Prove that $1 + 2 + \cdots + n \; = \; \dfrac{n(n+1)}{2}$, for all integers $n \geqslant 1$.

Don't forget to mention the principle of induction at the end of the proof.

Let $P(n)$ denote the statement "$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$".

Now LHS of $P(1)$ is 1, while the RHS of $P(1)$ is $\frac{1(1+1)}{2} = 1$. So both sides equal 1, which means $P(1)$ is a true statement.

The statement $P(k)$ states $1+2+\cdots+k=\frac{k(k+1)}{2}$. We **assume** this is true (this is our **inductive hypothesis**).

The statement $P(k+1)$ states $1+2+\cdots+(k+1)=\frac{(k+1)(k+2)}{2}$.

We must show that IF $P(k)$ is true, then $P(k+1)$ is also true.

$$\begin{aligned}
\text{Now L.H.S. of } P(k+1) &= 1+2+\cdots+(k+1) \\
&= 1+2+\cdots+k+(k+1) \\
&= \frac{k(k+1)}{2}+(k+1)\ (P(k) \text{ is assumed true}) \\
&= \frac{k(k+1)}{2} + \frac{(k+1)\cdot 2}{2} \\
&= \frac{(k+1)(k+2)}{2} \\
&= \text{R.H.S. of } P(k+1).
\end{aligned}$$

Hence $P(k+1)$ is true.

Thus, by the principle of mathematical induction, for all integers $n \geqslant 1$, we have $1+2+\cdots+n = \frac{n(n+1)}{2}$.

Example 71. *For all integers $n \geq 1$, use mathematical induction to prove that*

$$\sum_{i=1}^{n}(2i-1)=n^2.$$

Example 72. *For all integers $n \geq 1$, prove that*

$$\sum_{j=1}^{n}\frac{1}{j(j+1)}=\frac{n}{n+1}.$$

Example 73. *For all integers $t \geqslant 1$, use induction to prove that*

$$\sum_{j=1}^{t}2^{j-1}=2^t-1.$$

3.5 Logic

In this chapter we shall consider some formal *logic*. This will enable us to determine whether the conclusion of a formal argument is true or false, given various suppositions or *premises* in the argument.

We shall see how to write truth tables for compound statements.

A **statement** or a **proposition** is a sentence that is true or false, but not both.

Which of the following are propositions?

- It is raining.
- Is it raining?
- Tom is a male and Susan is a female.
- Mary is a male.
- No smoking inside.
- The number 6 is a prime number.
- What comes next?
- That pelican is beautiful.
- Elizabeth's favourite bird is a pelican.
- Hello there.

- We often use p, q, r etc. to stand for simple statements.
  If we let p denote "it is raining", then we can denote the **negation** of this statement by $\sim p$ or $\neg p$. We read this as "not p", so the negation of the statement "it is raining" is "it is not raining".

- If statement p is FALSE, then what about the statement $\sim p$?
  Can you say whether it is true or false?

- A truth table gives the truth value of a statement for all possible instances of the truth values of its component parts. A statement p has two possible truth values: true or false.

Here is a truth table to complete; it will give the truth values for $\sim p$ in terms of the truth values for the statement p.

p	$\sim p$
T	
F	

- If we have two statements, say p and q, we can combine them in various ways.
  Suppose p denotes "it is dark" and q denotes "it is raining".
  Then the statement "it is dark and it is raining" can be written as $p \wedge q$, read "p and q". This is known as the **conjunction** of p and q.

- We can use a truth table to determine the truth value of the conjunction $p \wedge q$ in all possible cases, whatever the truth values of p and q may be. With two statements p and q we have 2^2 or 4 possible scenarios.

Here is a truth table to determine the truth value of $p \wedge q$, according to the truth values of p and q.

p	q	$p \wedge q$
T	T	
T	F	
F	T	
F	F	

- The **disjunction** of two statement forms p and q, written $p \vee q$, and read "p or q", means p or q (or possibly both). This is sometimes known as the "inclusive or".

The truth table for $p \vee q$:

p	q	$p \vee q$
T	T	
T	F	
F	T	
F	F	

- $p \wedge q$ is true when p and q are both true.
- $p \wedge q$ is false when
- $p \vee q$ is true when
- $p \vee q$ is false when

A **statement form** or **propositional form** is made up from variables such as p, q, r, and logical connectives such as $\wedge$, $\sim$, $\vee$.

Two statement forms are **logically equivalent** if and only if they have *identical* truth values for every possible combination of truth values for the variables.

Write $P \Leftrightarrow Q$, whenever P and Q are logically equivalent.

Are the statement forms $p \wedge (\sim q)$ and $(p \vee q) \wedge (\sim q)$ logically equivalent?

p	q	$(p \;\wedge\; (\sim\; q))$ $\qquad$ $((p \;\vee\; q) \;\wedge\; (\sim\; q))$
T	T	
T	F	
F	T	
F	F	

- **De Morgan's Laws** (negations of "and" and "or"):
  The statement $\sim (p \wedge q)$ is logically equivalent to the statement $(\sim p) \vee (\sim q)$.

  The statement $\sim (p \vee q)$ is logically equivalent to the statement $(\sim p) \wedge (\sim q)$.

  A **tautology** is a statement form which *always* takes the truth value TRUE, for all possible truth values of its variables.

- A **contradiction** is a statement form which *always* takes the truth value FALSE for all possible truth values of its variables.

Construct a truth table to determine the truth values for $(p \vee q) \wedge (\sim p)$.

p	q	$(p \ \vee \ q) \ \ \wedge \ \ (\sim \ p)$
T	T	
T	F	
F	T	
F	F	

If a statement form P has *three* variables, such as p, q and r, how many rows will a truth table for P need? Discuss.

Is the statement form
$(p \wedge q) \vee (\sim p \vee (p \wedge (\sim q)))$ a tautology, a contradiction, or neither?

p	q	$(p \ \wedge \ q) \ \vee \ (\sim \ p \ \vee \ (p \ \wedge \ (\sim \ q)))$
T	T	
T	F	
F	T	
F	F	

The truth table for a statement form with n statement variables will have how many rows? Discuss.

When the same connective ($\wedge$ or $\vee$) is used, the commutative and associative laws hold:

Commutativity: $p \wedge q \ \Leftrightarrow \ q \wedge p$ and $p \vee q \ \Leftrightarrow \ q \vee p$.

Associativity: $p \wedge (q \wedge r) \ \Leftrightarrow \ (p \wedge q) \wedge r$, and $p \vee (q \vee r) \ \Leftrightarrow \ (p \vee q) \vee r$.

Also note the **distributive** laws:

$p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$ and $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$.

Exercises:

1. Construct a truth table for $(p \vee q) \wedge \sim (p \vee r)$.

2. Exclusive or: We use $\underline{\vee}$ where $p \, \underline{\vee} \, q$ means p or q but NOT both. Write out the truth table for this exclusive or.

3. Verify one of the distributive laws with a truth table.

Conditional statements

You have probably heard the terms "if and only if" and "necessary and sufficient". In this subsection we'll examine these carefully. We'll see truth tables for "if p then q", and for "p if and only if q"; we shall see how to replace these with the logical connectives we've met already: $\vee$, $\wedge$ and $\sim$. We'll also see what the contrapositive of a statement is.

- **if** p **then** q is denoted $p \Rightarrow q$. You can also read this as "p implies q". Here p is the **hypothesis** and q is the **conclusion**.
- "if p then q" is *false* when p is TRUE and q is FALSE. It is true in *all* other cases. We'll complete the truth table for "implies":

p	q	$p \Rightarrow q$
T	T	
T	F	
F	T	
F	F	

Translate the following statements into symbolic form. Let p denote "I will sleep", q denote "I am worried", and r denote
"I will work hard".

(a) If I am worried, I will not sleep.

(b) I will not sleep if I am worried.

(c) If I am worried, then I will both work hard and not sleep.

- "If p then q" (denoted $p \Rightarrow q$) is logically equivalent to $(\sim p) \vee q$.

  Check with a truth table.

p	q	$p \Rightarrow q$ $\quad$ $(\sim p) \vee q$
T	T	
T	F	
F	T	
F	F	

Example 74. *Rewrite the following sentence in "if–then" form.*
Either you do not study or else you pass the test

- The **contrapositive** of $p \Rightarrow q$ is $\sim q \Rightarrow \sim p$.

  Saying "if p then q" is like saying "if not q, then not p".

  Again, a truth table shows this equivalence:

p	q	$p \Rightarrow q$ $\quad \sim q \Rightarrow \sim p$
T	T	
T	F	
F	T	
F	F	

Example 75. *Write the contrapositive of the following sentence:*
If you do not study, then you will fail the test.

Example 76. *Construct a truth table to determine the truth values for* $p \Rightarrow (q \wedge (\sim p))$.

p	q	$p \Rightarrow (q \wedge (\sim p))$

There is a quote from the book *Alice in Wonderland* by Lewis Carroll (who was in fact Charles Dodgson, an English author, mathematician, logician, Anglican deacon and photographer) which is part of a conversation between Alice and the March Hare and the mad Hatter:

"Do you mean that you think you can find out the answer to it?" said the March Hare.

"Exactly so," said Alice.

"Then you should say what you mean," the March Hare went on.

"I do," Alice hastily replied; "at least—at least I mean what I say—that's the same thing, you know."

"Not the same thing a bit!" said the Hatter. "Why, you might just as well say that 'I see what I eat' is the same thing as 'I eat what I see'!"

Rewrite the statements "*I say what I mean*" and
"*I mean what I say*" in if–then format. Use a truth table to show that the two statements are not logically equivalent.

- Given statement variables p and q, the **biconditional** of p and q is $p \Leftrightarrow q$. Read this as "p if and only if q."

  $p \Leftrightarrow q$ is true precisely when p and q take the *same* truth values. It is false when p and q take opposite truth values.

  Complete its truth table:

p	q	$p \Leftrightarrow q$
T	T	
T	F	
F	T	
F	F	

- **Necessary and sufficient:**
  If p and q are statements,
  p is a **sufficient condition** for q means that if p then q.

  p is a **necessary condition** for q means that if *not* p, then *not* q.

  So if p is a **necessary condition** for q, we have: if q, then p.

Why is necessary and sufficient the same as biconditional, according to these conditions?

3.6 Application

4 Relations and functions

4.1 Relations on sets

- If A and B are any sets, recall that their Cartesian product is

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}.$$

- A **binary relation** R from a set A to a set B is any subset of $A \times B$.

  Let $(x, y) \in A \times B$. We say x *is related to* y by R, (and write $x \, R \, y$), if and only if $(x, y) \in R$.

  The notation "$(x, y) \in R$" is equivalent to the notation "$x \, R \, y$".

  If x is *not* related to y in R, we can write $x \not R \, y$ or $(x, y) \notin R$.

- Note that we often just use the word 'relation' when we mean 'binary relation', when the context is clear.

As well as using R to denote some relation, other symbols such as the Greek letters ρ (rho), σ (sigma) and τ (tau) are also often used. For example, $a \, \rho \, b$ and $c \, \tau \, d$ (means $(a, b) \in \rho$ and $(c, d) \in \tau$, respectively).

Example 77. **(a)** *Let* $A = \{0, 2, 4\}$ *and* $B = \{1, 2, 3, 4\}$.
Then $A \times B =$

(b) *With* A *and* B *as above, suppose* $x \, R \, y$ *if and only if* $x \leqslant y$.

$0 \, R \, 2$ *because* ________________

$2 \, R \, 4$ *because* ________________

Since $2 \leqslant 3$ *(and* $2 \in A$, $3 \in B$*), we have* ________________

We have

$$R = \{ \qquad\qquad\qquad\qquad\qquad\qquad \}$$

Example 78. *Let* $A = \{1, 2, 3, 4, 5\}$ *and* $B = \{0, 2, 4, 6, 8\}$. *In each of the following cases, suppose that* ρ *is a relation from* A *to* B*, and write down the elements in* ρ.

(i) $x \, \rho \, y$ *if and only if* $x \geqslant y$.

(ii) $x \, \rho \, y$ *if and only if* $x = y$.

(iii) $x \, \rho \, y$ *if and only if* $x - y$ *is even.*

(iv) $x \, \rho \, y$ *if and only if* $x + y = 7$.

(v) $x \, \rho \, y$ *if and only if* $x + y > 9$.

Example 79. *Define (make up) three different relations from $\mathbb{Z}$ to $\mathbb{Z}^+$, the set of positive integers. Note that there are many possible answers.*

Example 80. *Consider the following relations defined on $\mathbb{Z}$.*

(i)	$R_1 = \{(a, b) : a \leqslant b\}$;	**(ii)**	$R_2 = \{(a, b) : a > b\}$;
(iii)	$R_3 = \{(a, b) : a = b$ *or* $a = -b\}$;	**(iv)**	$R_4 = \{(a, b) : a = b\}$;
(v)	$R_5 = \{(a, b) : a = b + 1\}$;	**(vi)**	$R_6 = \{(a, b) : a + b \leqslant 3\}$.

List some of the elements in each relation.

Example 81. *Consider each of the following ordered pairs in turn, and state which of the above relations the pair belongs to:* $(1, 1),\ (1, 2),\ (2, 1),\ (1, -1), (2, 2)$.

- **Arrow diagram:** A relation R from a set A to a set B can be represented by a directed bipartite graph G. The partite sets for the vertices of G are A and B, and for each $a \in A$ and $b \in B$, there is an arc (a directed edge) from a to b if and only if $(a, b) \in R$, that is,
  if and only if $a\, R\, b$.

Example 82. *Let $A = \{1, 3, 5\}$ and $B = \{1, 2, 3, 4\}$.*
Draw a directed bipartite graph to illustrate the relation σ from A to B, where $a\, \sigma\, b$ if and only if $a + 1 > b$.

- A **function** $f : A \to B$ is a relation from the set A to the set B which satisfies:

  (i) for all $x \in A$, there exists some $y \in B$ such that $(x, y) \in f$;

  (ii) for all $x \in A$ and all $y,\ z \in B$,

$$\text{if } (x, y) \in f \text{ and } (x, z) \in f, \text{ then } y = z.$$

  If f is a function from A to B we write

$$y = f(x) \text{ if and only if } (x, y) \in f.$$

Example 83. *Let $A = \{3, 6, 9\}$, $B = \{2, 4, 6, 8\}$, and let $R = \{(3, 2),\ (6, 2),\ (9, 6),\ (6, 8)\}$ be a relation. Is R a function from A to B? Explain.*

- If R is a binary relation from A to B, then the **inverse relation**, R^{-1}, is defined from B to A by:

$$R^{-1} = \{(y, x) \in B \times A : (x, y) \in R\}.$$

So for all $x \in A$ and $y \in B$, $(y, x) \in R^{-1}$ if and only if $(x, y) \in R$.

Example 84. *Let $A = \{x, y, z\}$ and $B = \{1, 4, 7, 10\}$. Suppose that*

$$\rho = \{(x, 4),\ (x, 10),\ (z, 1),\ (y, 7),\ (y, 1)\}$$

is a relation from A to B. Write down ρ^{-1}.

- A **binary relation ON a set** A is a binary relation from A to A;
so it is a subset of the Cartesian product $A \times A$.

- If a binary relation R is on a set A, then we can represent R using a directed graph G with vertex set A (instead of using a bipartite graph). The directed edge set of G satisfies:

  there is an arc (x, y) from vertex x to vertex y
  if and only if $x\,R\,y$, that is, if and only if $(x, y) \in R$.

- In a binary relation R on A, if we have $x\,R\,x$, then in the directed graph we'll have a *loop* on vertex x.

Example 85. *Let* $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. *Define a binary relation on* A *by*

$$\text{for all } x,\ y \in A, \quad x\,R\,y \ \text{ if and only if } \ 3 \mid (x - y).$$

(a) *Write down the elements of* R*.*

(b) *Write down the elements of* R^{-1}*. Is it true that* $R = R^{-1}$*?*

(c) *Draw the digraph representing* R*, with vertex set* A*.*

4.2 Reflexivity, symmetry and transitivity

In this section we'll look at three special properties which some binary relations *on a set* may have.

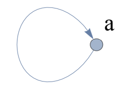
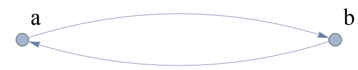
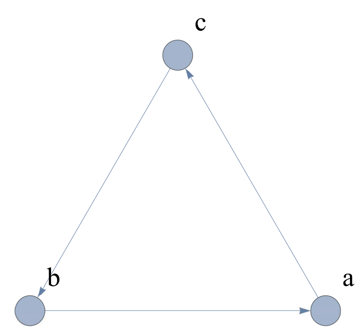
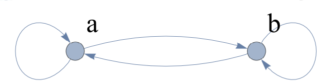
In the following, let R be a relation on a set A.

- R is **reflexive** if and only if for *all* $x \in A$, we have $x\,R\,x$.

- R is **symmetric** if and only if for all $x, y \in A$, *if* $x\,R\,y$, then $y\,R\,x$.

- R is **transitive** if and only if for all $x, y, z \in A$, *if* $x\,R\,y$ and $y\,R\,z$, then $x\,R\,z$.

  Note that there's no requirement here for x, y, z to be distinct; in particular you should consider the possibility that $x = z$ when checking transitivity of a relation.

Property of relation	Description	
	two equivalent ways	in Graphs
Reflexive	$\forall a \in A, \quad a\ R\ a.$ $\forall a \in A, \quad (a, a) \in R.$	
Symmetric	$\forall a, b \in A,$ if $a\ R\ b$ then $b\ R\ a.$ $\forall a, b \in A,$ if $(a, b) \in R$ then $(b, a) \in R.$	
Antisymmetric	$\forall a, b \in A,$ if $a\ R\ b$ and $b\ R\ a$ then $a = b.$ $\forall a, b \in A,$ if $(a, b) \in R$ and $(b, a) \in R$ then $a = b.$	
Transitive	$\forall a, b, c \in A,$ if $a\ R\ b$ and $b\ R\ c$ then $a\ R\ c.$ $\forall a, b, c \in A,$ if $(a, b) \in R$ and $(b, c) \in R$ then $(a, c) \in R.$	

In terms of the digraph G of a relation R on a set A:

- R is **reflexive** if and only if there is a loop on every vertex of G;
- R is **symmetric** if whenever there is an arc (or directed edge) (x, y) in G, then there must be an arc (y, x) in G as well.
  So whenever there is a directed edge between two distinct vertices, there is also a directed edge the other way between the same pair
- R is **transitive** if whenever the arcs (x, y) and (y, z) are in G, then the arc (x, z) is also in G.

Example 86. *Let the following relations be defined on the set* $\{1, 2, 3, 4\}$*:*

- $R_1 = \{(1, 1),\ (1, 2),\ (2, 1),\ (2, 2),\ (3, 4),\ (4, 1),\ (4, 4)\}$
- $R_2 = \{(1, 1),\ (1, 2),\ (2, 1),\ (2, 2),\ (3, 3),\ (4, 4)\}$

- $R_3 = \{(1,1),\ (1,2),\ (1,4),\ (2,1),\ (2,2),\ (3,3),\ (4,1),\ (4,4)\}$
- $R_4 = \{(2,1),\ (3,1),\ (3,2),\ (4,1),\ (4,2),\ (4,3)\}$
- $R_5 = \{(1,1),(1,2),(1,3),(1,4),(2,2),(2,3),(2,4),(3,3),(3,4),(4,4)\}$
- $R_6 = \{(1,1),\ (2,2),\ (3,3),\ (4,4)\}$

For each relation, determine whether it is reflexive, symmetric and/or transitive.

- R is the **identity relation** on A if and only if:

  for all $x,\ y \in A,\ \ x\,R\,y$ if and only if $x = y$.

  (The identity relation is also the identity function.)

Let R be a binary relation on a set A.

- R is **not reflexive** if and only if there's an element $a \in A$ with $(a,a) \notin R$. (That is, if and only if there's a vertex in the digraph of the relation with no loop.)
- R is **not symmetric** if and only if there exist elements a and b in A with precisely *one* of $(a,b),\ (b,a)$ in R.
- R is **not transitive** if there exist elements x, y and z (where possibly $x = z$) with $(x,y) \in R$ and $(y,z) \in R$ but with $(x,z) \notin R$.

So to show failure of one of the three properties, a counter-example suffices.

Example 87. *Is the relation "divides" on the set of positive integers, where we write $a\,R\,b$ if and only if $a \mid b$, (i) reflexive? (ii) symmetric? (iii) transitive?*

Example 88. *Let $A = \{1,2,3,4,5,6,7,8,9,10\}$; define ρ on A by:*

for all $x,\ y \in A,\ \ x\,\rho\,y$ if and only if $3 \mid (x - y)$.

Draw the digraph of ρ and use it to check whether ρ is reflexive, symmetric and/or transitive.

Example 89. *Let σ be defined on the real numbers $\mathbb{R}$ by:*

for all $x,\ y \in \mathbb{R},\ \ x\,\sigma\,y$ if and only if $x > y$.

(a) *Is σ reflexive?*

(b) *Is σ symmetric?*

(c) *Is σ transitive?*

Example 90. *Prove that if a relation R is reflexive, then R^{-1} is also reflexive.*

Example 91. *Define a relation ρ on $\mathbb{Z} \times (\mathbb{Z} - \{0\})$ as follows: for all (a, b), $(c, d) \in \mathbb{Z} \times (\mathbb{Z} - \{0\})$,*

$$(a, b)\, \rho\, (c, d) \text{ if and only if } \frac{a}{b} = \frac{c}{d}.$$

(i) *Is ρ reflexive?*

(ii) *Is ρ symmetric?*

(iii) *Is ρ transitive?*

Example 92. *Prove that a relation R is symmetric if and only if $R = R^{-1}$.*

Example 93. *For the binary relation R, is the statement*

"If R is transitive then R^{-1} is transitive"

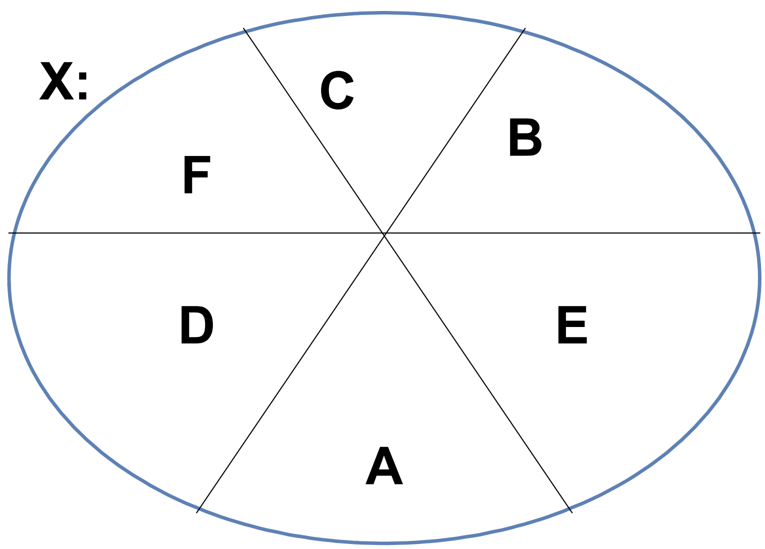
true or false? Justify your answer.

Equivalence relations

The concept of an equivalence relation is extremely important. in lots of mathematics. Before defining this:

- Recall that a **partition** of a set A is a collection of non-empty mutually disjoint subsets of A whose union is A.

  Here's an illustration of a partition of $X = A \cup B \cup C \cup D \cup E \cup F$:



- Given a partition $\{A_1, A_2, \ldots, A_n\}$ of a set A, the **binary relation** R **induced by the partition** is defined on A as follows:

  for all $x,\ y \in A,\ \ x\, R\, y$ if and only if there is a subset A_i of the partition containing both x and y.

Theorem 10. *Given a partition $\{A_1, A_2, \ldots, A_n\}$ of a set A and a binary relation R induced by the partition, it follows that R is reflexive, symmetric and transitive.*

Example 94. *Let $A = \{1, 3, 5, 7, 9, 11\}$; consider the partition of A given by*
$$\{1,\ 5,\ 9\},\ \{3,\ 7\},\ \{11\}.$$
List the elements of the relation R induced by this partition.

- Let A be a non-empty set and let R be a binary relation on A.
  R is an **equivalence relation** if and only if R is
  reflexive, *symmetric* and *transitive* .

Example 95. *Let $A = \{0, 1, 2, 3, 4\}$ and let the relation R on A be defined as follows:*

$$\begin{aligned} R \;=\; & \{(0,0),\ (2,1),\ (0,3),\ (1,1),\ (3,0),\ (1,4),\ (4,1), \\ & (2,2),\ (2,4),\ (3,3),\ (4,4),\ (1,2),\ (4,2)\}. \end{aligned}$$

Draw the digraph representing R.Then use this digraph to check whether R is an equivalence relation.

- Let A be a set and let R be an equivalence relation on A.
  For each $a \in A$, the **equivalence class** containing a, denoted $[a]$ (or $[a]_R$), is the set of all x in A such that x is related to a in R.
  That is,
  $$[a]_R \;=\; \{x \in A : \; x\,R\,a\ \}.$$

Lemma 1. *If R is an equivalence relation on a set A, and if $a,\ b \in A$ satisfy $a\,R\,b$, then $[a] = [b]$, that is, the equivalence classes in which a and b lie are identical.*

Lemma 2. *If R is an equivalence relation on a set A, and if $a,\ b \in A$, then either $[a] \cap [b] = \emptyset$ or else $[a] = [b]$.*

Lemmas 1 and 2 may be used to show that the following holds.

Theorem 11. *If R is an equivalence relation on a set A, then the distinct equivalence classes of R form a partition of A; that is, the union of all the classes is A, and the intersection of any two distinct classes is empty.*

Example 96. *Let the relation ρ be defined on the integers $\mathbb{Z}$ by:*

for all m and n in $\mathbb{Z}$,
$$m\ \rho\ n \ \text{ if and only if } \ 7 \mid (m - n).$$

(a) *Prove that ρ is an equivalence relation on $\mathbb{Z}$.*

(b) *Find the equivalence classes for ρ.*

(c) *True or false?*
$[1] = [-8]$; $[1] = [8]$; $[123] = [319]$; $[304] = [-10]$; $[-34] = [-6]$.

Example 97. *Let d be a positive integer. Define the relation ρ on the integers $\mathbb{Z}$ as follows: for all m, $n \in \mathbb{Z}$,*

$$m \,\rho\, n \text{ if and only if } m \equiv n \pmod{d}.$$

Recall that $m \equiv n \pmod{d}$ if and only if $d \mid (m - n)$. Prove that ρ is an equivalence relation. Then list its equivalence classes.

- A **total order relation** is also a partial order relation - with the extra property that any two elements in a total order relation are comparable.
- Elements of a **binary relation** on a set A are ordered pairs of elements of A, that is, elements of $A \times A$.
- But the **equivalence classes** of an equivalence relation (which is **reflexive**, **symmetric** and **transitive**) are subsets of A, not of $A \times A$.
- Also remember that the equivalence classes of an equivalence relation on a set A form a **partition** of the set A.

4.3 Functions on general sets

The concept of some dependence between two quantities is ubiquitous throughout mathematics, science and engineering. This idea was formalised in 1673 by the German mathematician Gottfried Wilhelm Leibniz who used the word **function** to indicate the dependence of one quantity on another. Leibniz was a true polymath: he was an expert in law, religion, philosophy, literature, politics, geology, metaphysics, alchemy and history, as well as mathematics.

Functions are everywhere in mathematics. In the previous chapters we have encountered many different functions, from: truth tables (which are Boolean functions); sequences (functions defined from sets of integers to some other set); mod and div (which are functions defined from Cartesian products of integers to a set of integers); and floor and ceiling (functions from $\mathbb{R}$ to $\mathbb{Z}$).

In this section we define functions and investigate some of their properties.

- A binary relation f from a set $\mathcal{X}$ to a set $\mathcal{Y}$ is said to be a **function** (or **map**, or **mapping**) if and only if it satisfies the following:

  for all $x \in \mathcal{X}$ there exists a *unique* $y \in \mathcal{Y}$ such that $(x, y) \in f$.

  That is, each and every element of $\mathcal{X}$ is associated with *precisely* one element of $\mathcal{Y}$. (However, note that two different elements of $\mathcal{X}$ may be associated with the same element of $\mathcal{Y}$.)

- A relation that satisfies the definition of a function is said to be **well-defined**.

  Note that if a relation f from a set $\mathcal{X}$ to a set $\mathcal{Y}$ is a function, then for any $x \in \mathcal{X}$, if $(x, y_1) \in f$ and $(x, y_2) \in f$ then $y_1 = y_2 \in \mathcal{Y}$.
  If $y_1 \neq y_2$ this is *not* a function.
  Furthermore, be careful: $f(x) = \sqrt{x}$ is NOT a function from $\mathbb{R}$ to $\mathbb{R}$ because $f(x)$ takes no value when $x < 0$. So this is NOT a **well-defined** function on $\mathbb{R}$!

Example 98. *Consider the binary relation ρ from the set $\mathbb{R}$ to the set $\mathbb{R}$ defined by*

$$x \rho y \text{ if and only if } y = x^2.$$

Is ρ a function?

Example 99. *Consider the binary relation S from the set $\{0, 1, 2, 3\}$ to the set $\mathbb{Z}$ defined by*

$$xSy \text{ if and only if } y \Leftrightarrow x \ (mod\ 4).$$

Is S a function?

Despite the fact that a function is a relation, it is normally written in special notation because of its distinctive properties.

- We write $f : \mathcal{X} \to \mathcal{Y}$ to denote that f is a function from the set $\mathcal{X}$ to the set $\mathcal{Y}$.

Let $f : \mathcal{X} \longrightarrow \mathcal{Y}$ be some function.

- The set $\mathcal{X}$ is called the **domain** of f.

- The set $\mathcal{Y}$ is called the **co-domain** of f.

- For each $x \in \mathcal{X}$, the single (unique) element of $\mathcal{Y}$ that f associates with x is denoted by $f(x)$. For $f(x)$ we say

  f **of** x; or
  the value of f **at** x; or
  the image of x **under** f.

  If $f(x) = y$, we sometimes write

  $$f : x \longmapsto y.$$

- The set of all values of f taken together is called the **range** of f or the **image** of $\mathcal{X}$ under f.

  $$\text{range of } f = \{y \in \mathcal{Y} : y = f(x), \text{ for some } x \text{ in } \mathcal{X}\}$$

Example 100. *Suppose we are told that a function $f : \mathbb{Q} \to \mathbb{Z}$ is to be defined by the formula $f\left(\frac{m}{n}\right) = m$ for all integers m and n when $n \neq 0$. Is f well-defined, that is, is f a function? Justify your answer.*

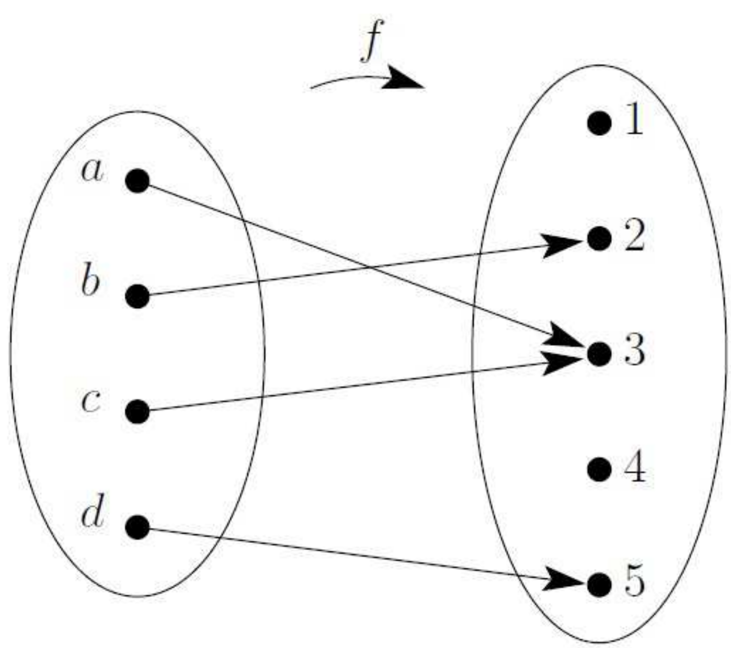
A function is a specific type of relation, so with any function between two finite sets we can associate a digraph.

(In fact if we relax our definition of a digraph so that the vertex and arc sets may be of *infinite* size, this concept can be extended to any function.)

If X and Y are two finite sets, we can illustrate a function f, mapping from X to Y, as follows:

- First, form a list of the elements of X, and a list of the elements of Y.
- Define the function by drawing an arrow from each element in X to the corresponding element in Y. Since f is a *function*, the arrows must satisfy the following property:
  - Every element of X has precisely one arrow coming out of it, pointing to one element in Y.
- Such an illustration is called an **arrow diagram**.

Example 101. *Let f be the function defined by the following arrow diagram.*



(a) Write down the domain and co-domain of f.

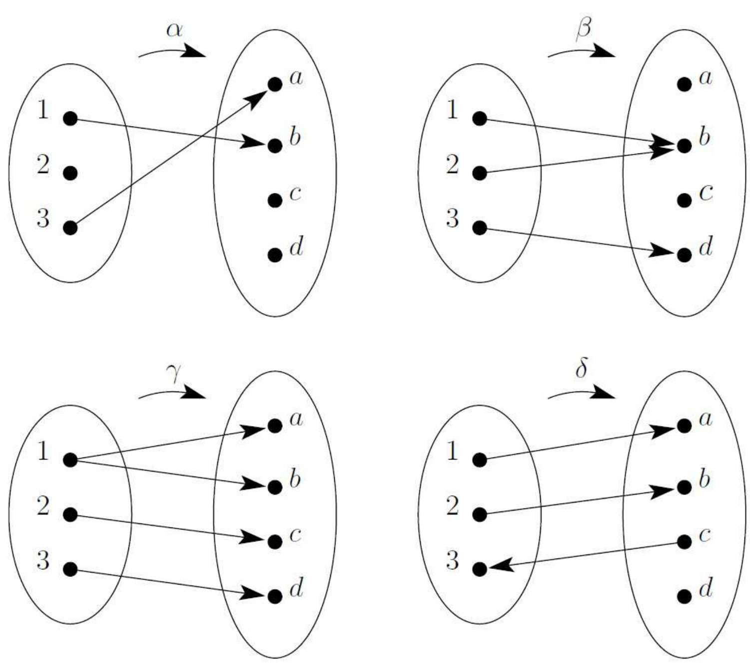
(b) Find $f(a)$, $f(b)$, $f(c)$ and $f(d)$.

(c) What is the range of f?

Note that an arrow diagram is equivalent to the graphical representation of a relation as described in the previous chapter. If f is a function from X to Y, then in the arrow diagram the following all hold:

- The arrow diagram forms a directed bipartite graph.
- The elements of X form the vertices of one partite set and the elements of Y form the vertices of the second partite set.
- Any ordered pair from the edge set (arc set) of the graph has the first coordinate in X and the second in Y.
- There is precisely one edge (arc) incident with each vertex in X.

Example 102. *Which of the following define functions from* $X = \{1, 2, 3\}$ *to* $Y = \{a, b, c, d\}$ *? Explain your answers.*



- Given a function $f : \mathcal{X} \to \mathcal{Y}$ and an element $y \in \mathcal{Y}$, the **inverse image** of y is the set of all elements $x \in \mathcal{X}$ such that $f(x) = y$.

  Symbolically, the inverse image of y is:

  $$\textbf{inverse image of } y = \{x \in \mathcal{X} : f(x) = y\}.$$

Suppose f and g are functions from $\mathcal{X}$ to $\mathcal{Y}$. Then f and g are **equal**, written $f = g$, if and only if

$$f(x) = g(x) \quad \text{for all } x \in \mathcal{X}.$$

Example 103. *Define functions* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$*, where* $f(x) = x$ *for all* $x \in \mathbb{R}$ *and* $g(x) = \sqrt[3]{x^3}$ *for all* $x \in \mathbb{R}$*. Is* $f = g$*? Explain.*

Example 104. *Define functions* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$*, where* $f(x) = x$ *for all* $x \in \mathbb{R}$ *and* $g(x) = \sqrt{x^2}$ *for all* $x \in \mathbb{R}$*. Is* $f = g$*? Explain.*

- Given a set $\mathcal{X}$, define a function $\iota_{\mathcal{X}}$ from $\mathcal{X}$ to $\mathcal{X}$ by

  $$\iota_{\mathcal{X}}(x) = x \quad \text{for all } x \text{ in } \mathcal{X}.$$

  The function $\iota_{\mathcal{X}}$ is called the **identity function on** $\mathcal{X}$.

  Here ι is the Greek letter *iota*.

Previously we said that a sequence was a list of elements. Now we have discussed functions, we can give a more formal definition of a sequence.

- A **sequence** is a function defined on the set of integers greater than or equal to some fixed integer.

Example 105. *Write the sequence*

$$-4,\ 9,\ -16,\ 25,\ \ldots,\ (-1)^n(n+1)^2,\ \ldots$$

as a function from the positive integers to the integers.

Example 106. *Now write the sequence*

$$-4,\ 9,\ -16,\ 25,\ \ldots,\ (-1)^n(n+1)^2,\ \ldots$$

as a function from the nonnegative integers to the integers.

- A **binary operation** on a set $\mathcal{X}$ is a special kind of function from the set $\mathcal{X} \times \mathcal{X}$ to the set $\mathcal{Y}$ (where usually $\mathcal{Y} \subseteq \mathcal{X}$).

Example 107. Pythagorean triples *are* (x, y, z) *in integers satisfying* $\mathcal{P} : x^2 + y^2 = z^2$*. If* $P_1 = (x_1, y_1, z_1)$*,* $P_2 = (x_2, y_2, z_2) \in \mathcal{P}(\mathbb{Z})$*, then* $P_3 = (x_3, y_3, z_3) \in \mathcal{P}(\mathbb{Z})$*, where* $P_3 = P_1 \star P_2, (x_3, y_3, z_3) = (x_1x_2 - y_1y_2, x_1y_2 + x_2 + y_1, z_1z_2)$*. Try* $(3, 4, 5) \star (5, 12, 13)$*. Incidentally, this is related to* $(\sin\theta, \cos\theta)$ *via their angle addition formulas.*

Example 108. *Let* $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ *be the function of addition on the integers.*

(a) *Evaluate* $f((3, 2))$*. (And make sure you understand why all these parentheses appear)*

(b) *Find an element in* $\mathbb{Z} \times \mathbb{Z}$ *with an image of* -1*.*

- Digital messages consist of finite sequences (strings) of zeros and ones. Ideally, we want to be able to correct any errors introduced to these messages, such as from background electronic noise or small scratches on a CD. The study of such topics is called **coding theory**.

- The **Hamming distance function** is a very important function used in coding theory. It counts the minimum number of substitutions required to change one string of zeros and ones into the other; in other words, it counts the minimum number of errors that transform one string into another. It is another example of a binary operation.

  Let S_n be the set of all strings of zeros and ones of length n.
  Define a function $H : S_n \times S_n \to \mathbb{Z}$ as follows:

$$H : (s,t) \quad \mapsto \quad \text{the number of positions in which } s \text{ and } t \text{ have different values.}$$

Example 109. *Let* $n = 8$. *Compute* $H((00111010, 00010101))$.

4.4 Examples of real function

Consider the mapping $y = x^2$, i.e. $y = f(x)$. We can see that f acts as a map which takes real numbers and maps them to the real plane. In other words: $f : \mathbb{R} \mapsto \mathbb{R}$

Exponential functions

An exponential function is one of the form $f(x) = a^x$, where the *base* a is a positive constant, and x is said to be the *exponent* or *power*. One very common exponential function which we shall see often in this course is given by $f(x) = e^x$. It has the graph shown in Figure 14. Notice that it cuts the y-axis (the line $x = 0$) at $y = 1$.

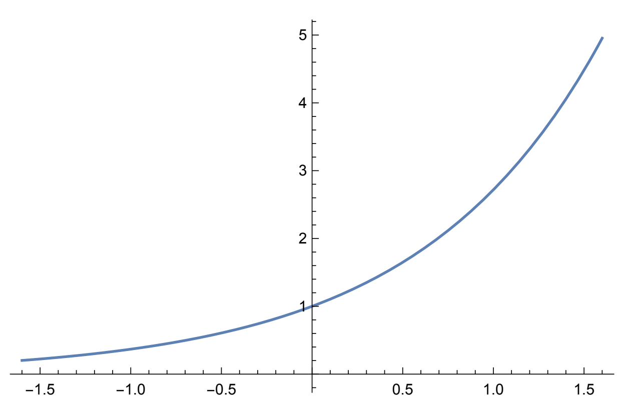


Figure 14: The function $f(x) = e^x$.

Trigonometric functions (sin, cos, tan)

In calculus, we measure angles in *radians*. One radian is defined as the angle subtended at the centre of a circle of radius 1 by a segment of arc length 1.

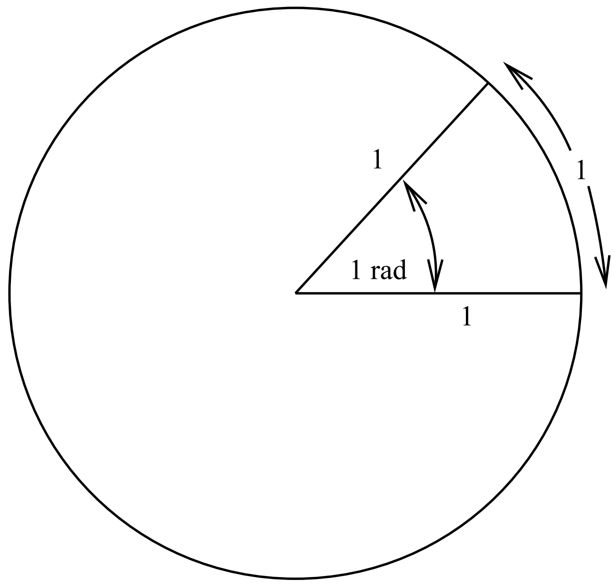


Figure 15: Angle in radians.

By definition, 2π radians = $360°$. To convert between the two measures use the following formulae:
For x angle in radians, θ angle in degrees,

$$x \times \frac{360°}{2\pi} = \theta, \quad \theta \times \frac{2\pi}{360°} = x.$$

The point on the unit circle $x^2 + y^2 = 1$ making an angle θ (in radians) anti-clockwise from the x-axis has coordinates $(\cos\theta, \sin\theta)$, so $\cos^2\theta + \sin^2\theta = 1$. This defines the cos and sin functions. This relationship is shown in Figure 16.

The graphs of the functions $\cos\theta$ and $\sin\theta$ are shown below in Figure 16. Note how the two functions have the same behaviour, but are shifted (i.e. *phase shifted)*) by $\frac{\pi}{2}$. This leads to the relationship

$$\cos\theta = \sin\left(\frac{\pi}{2} - \theta\right) = \sin\left(\theta + \frac{\pi}{2}\right).$$

In both cases in Figure 16 the range is $[-1, 1]$.

The tan function is related to the sine and cosine functions by:

$$\tan\theta = \frac{\sin\theta}{\cos\theta}.$$

It is defined for $\cos\theta \neq 0$. It is therefore not defined at values of θ which are odd integer multiples of $\frac{\pi}{2}$ (eg $\pm\frac{\pi}{2}$, $\pm\frac{3\pi}{2}$, $\pm\frac{5\pi}{2}$ etc).

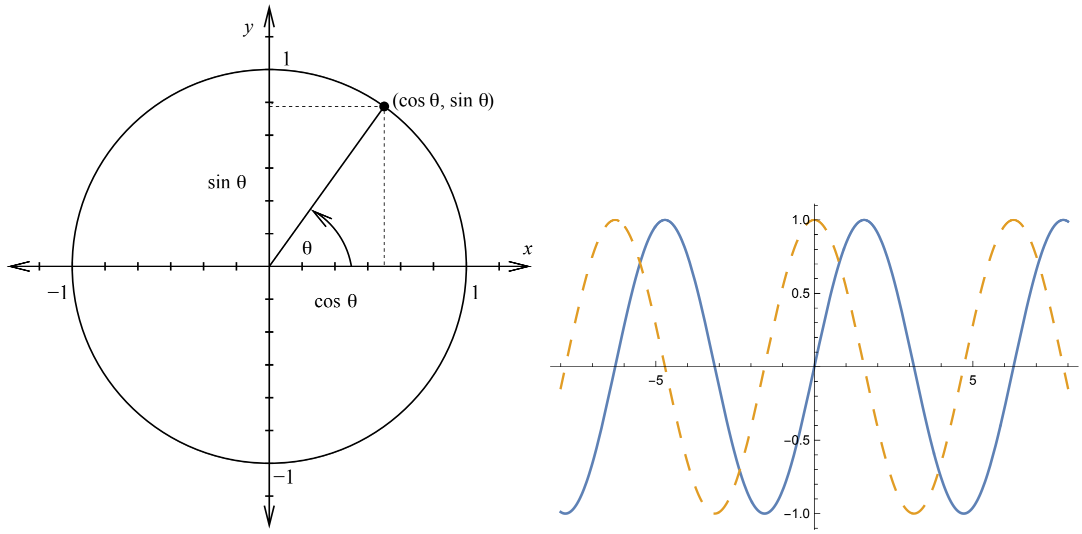


Figure 16: The relationship between the cosine and sine functions. On the right, $\cos(x)$ is dashed.

The graph of $\tan\theta$ is given on the left of Figure 17 below. Mark the θ values in the graph where the function is not defined.

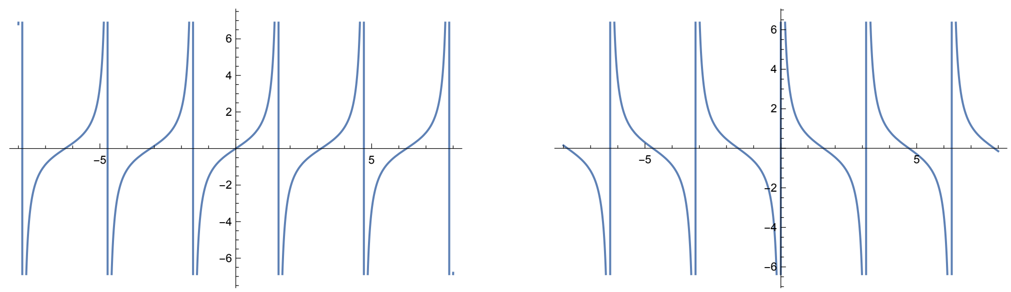


Figure 17: Left: $\tan(\theta)$, and vertical asymptote lines. Right: $\cot(\theta)$, and vertical asymptote lines.

The right of Figure 17 shows another common function,

$$\cot x \;=\; \frac{\cos x}{\sin x}.$$

The periodic nature of all the trigonometric functions we have seen makes them ideal for modelling repetitive phenomena such as tides, vibrating strings, and various types of natural wave-like behaviour.

Both the sine and cosine functions have *period* 2π. Hence

$$\begin{aligned}\sin(x + 2\pi) &= \sin x, \text{ and}\\ \cos(x + 2\pi) &= \cos x.\end{aligned}$$

Furthermore, both $\sin(x)$ and $\cos(x)$ have an *amplitude* of one.

We can “speed up" or “slow down" these functions by tinkering with their periodicity. For example, compare the functions $\sin(x)$ and $\sin\left(\frac{x}{2}\right)$ in Figure 18. Compare also the functions $\sin(x)$ and $\sin(2x)$ in Figure 18.

In addition, we can stretch or shrink trigonometric functions by multiplying these functions by constants other than one. For example, Figure 18 shows the functions $\frac{1}{2}\cos(x)$ and $3\cos(x)$. What amplitudes do these have?

Naturally, we can change both the period and amplitude of trigonometric functions simultaneously. Figure 18 show the graphs of $4\cos\left(\frac{x}{3}\right)$ and $2\sin(5x)$.

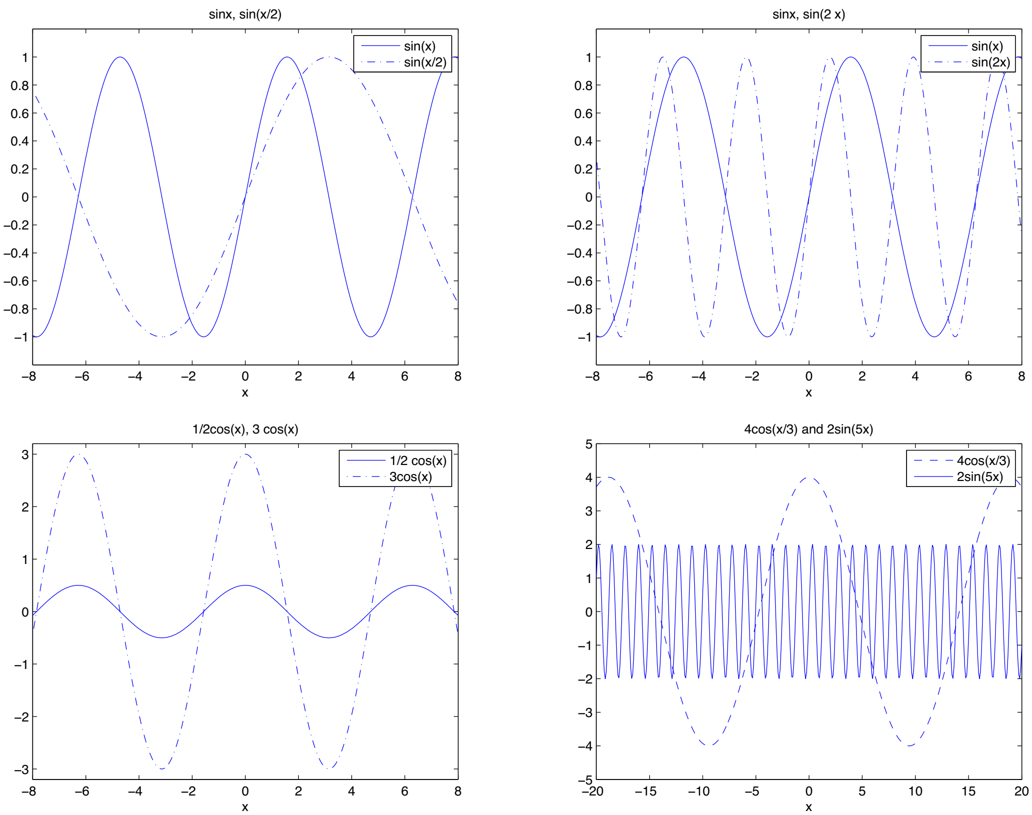


Figure 18: Graphs of $\sin(x)$ and $\sin\left(\frac{x}{2}\right)$, $\sin(x)$ and $\sin(2x)$, Graph of $\frac{1}{2}\cos x$ and $3\cos(x)$, $4\cos\left(\frac{x}{3}\right)$ and $2\sin(5x)$.

4.5 Composition of functions

Consider the functions $f : \mathbb{Z} \to \mathbb{Z}$ and $g : \mathbb{Z} \to \mathbb{Z}$ where $f(x) = x + 1$ and $g(x) = x^2$ for all $x \in \mathbb{Z}$. Then the function $h : \mathbb{Z} \to \mathbb{Z}$ where $h(x) = (x + 1)^2$ is an example of a **composition of functions**; it is in fact the **composition of** f **and** g. In this section we shall investigate this concept.

- Let $f : \mathcal{X} \to \mathcal{Y}'$ and $g : \mathcal{Y} \to \mathcal{Z}$ be functions with the property that the range of f is a subset of the domain of g. Define a new function $g \circ f : \mathcal{X} \to \mathcal{Z}$ as follows:

  $$(g \circ f)(x) = g(f(x)) \quad \text{for all } x \in \mathcal{X},$$

  where $g \circ f$ is read "g circle f" and $g(f(x))$ is read "g of f of x."

  The function $g \circ f$ is called the **composition** of f and g.

- We put the f first when we say "the composition of f and g" because an element x is acted upon first by f and then by g.
- The *domain* of $g \circ f$ is $\mathcal{X}$.
- The *co-domain* of $g \circ f$ is $\mathcal{Z}$.
- The *range* of $g \circ f$ is the image (under g) of the range of f.

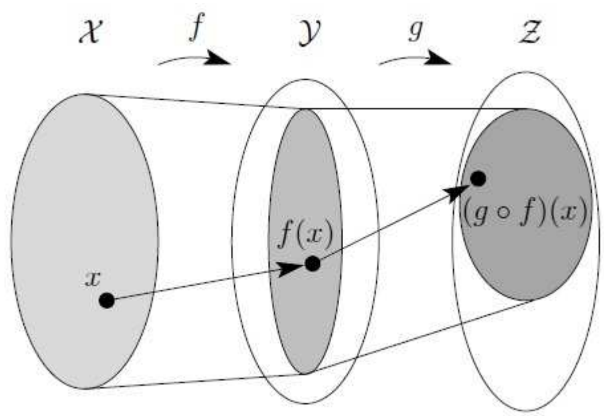


Figure 19: .

Example 110. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^3$ for all $x \in \mathbb{R}$ and let $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(x) = 2x - x^2$ for all $x \in \mathbb{R}$.*

(a) *Find $(f \circ g)(x)$ and $(g \circ f)(x)$.*

(b) *Is it true that $(g \circ f)(x) = (f \circ g)(x)$?*

Example 111. *Let $X = \{a, b, c\}$, $Y' = \{1, 2, 3\}$, $Y = \{1, 2, 3, 4\}$ and $Z = \{w, x, y, z\}$. Define the functions $f : X \to Y'$ and $g : Y \to Z$ by the arrow diagrams below.*

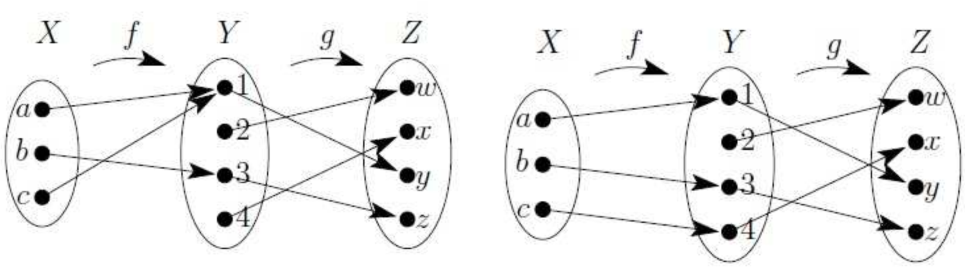


Figure 20: .

(a) *Draw the arrow diagram for $g \circ f$.*

(b) *What is the range of $g \circ f$?*

Example 112. *Suppose that $f : \mathbb{Z} \to \mathbb{Z}$ is defined by $f(x) = x + 1$ for all $x \in \mathbb{Z}$. Let $\iota_{\mathbb{Z}}$ be the identity function. Find $(f \circ \iota_{\mathbb{Z}})(x)$ and $(\iota_{\mathbb{Z}} \circ f)(x)$.*

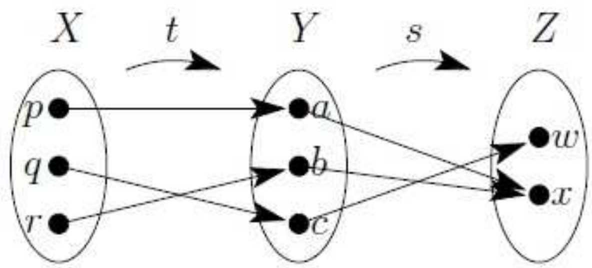


Figure 21: .

Theorem 12. *If f is a function from a set $\mathcal{X}$ to a set $\mathcal{Y}$, and $\iota_{\mathcal{X}}$ is the identity function on $\mathcal{X}$, and $\iota_{\mathcal{Y}}$ is the identity function on $\mathcal{Y}$, then*

$$f \circ \iota_{\mathcal{X}} = f \quad \text{and} \quad \iota_{\mathcal{Y}} \circ f = f.$$

Theorem 13. *Let $f : \mathcal{X} \to \mathcal{Y}$ be a one-to-one correspondence (bijection) with inverse function $f^{-1} : \mathcal{Y} \to \mathcal{X}$. Then*

$$f^{-1} \circ f = \iota_{\mathcal{X}} \quad \text{and} \quad f \circ f^{-1} = \iota_{\mathcal{Y}}.$$

Theorem 14. *If $f : \mathcal{X} \to \mathcal{Y}$ and $g : \mathcal{Y} \to \mathcal{Z}$ are both one-to-one functions, then $g \circ f$ is one-to-one.*

Example 113. *Let $X = \{p, q, r\}$, $Y = \{a, b, c\}$ and $Z = \{w, x\}$. Define the functions $t : X \to Y$ and $s : Y \to Z$ by the arrow diagrams below.*

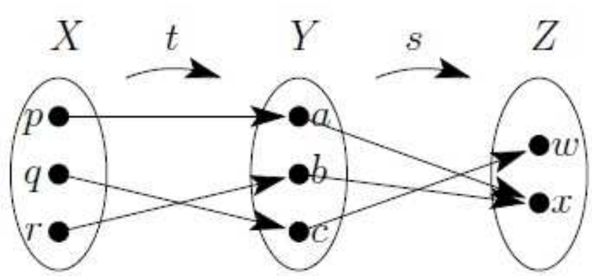


Figure 22: .

(a) Draw the arrow diagram representing $s \circ t$.

(b) Note that t and s are both onto (surjective). Is $s \circ t$ onto?

Theorem 15. *If $f : \mathcal{X} \to \mathcal{Y}$ and $g : \mathcal{Y} \to \mathcal{Z}$ are both onto functions, then $g \circ f$ is onto.*

4.6 One-to-one, onto and inverse functions

In this section we discuss two important properties which functions may or may not satisfy, namely, the property of being *one-to-one* and the property of being *onto*. When a function is one-to-one *and* onto we may define an *inverse function* from the co-domain to the domain of the original function.

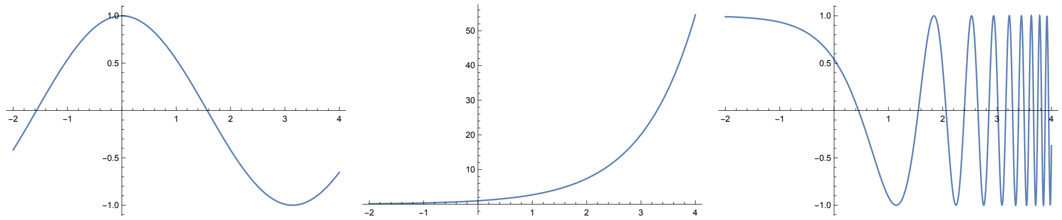


Figure 23: Plots of real functions, and their composition.

- Let f be a function from a set $\mathcal{X}$ to a set $\mathcal{Y}$.
  The function f is **one-to-one** (or **injective**) if and only if for all elements x_1 and x_2 in $\mathcal{X}$,
  $$\text{if } \ f(x_1) = f(x_2), \ \text{ then } \ x_1 = x_2.$$
  Or, equivalently,
  $$\text{if } \ x_1 \neq x_2, \ \text{ then } \ f(x_1) \neq f(x_2).$$

- A function $f : \mathcal{X} \to \mathcal{Y}$ is **not one-to-one** if and only if there exist some x_1 and x_2 in $\mathcal{X}$ with $f(x_1) = f(x_2)$ and $x_1 \neq x_2$.

Example 114. *Which of the following arrow diagrams define one-to-one (injective) functions from* $X = \{a, b, c, d\}$ *to* $Y = \{1, 2, 3, 4, 5, 6\}$*?*

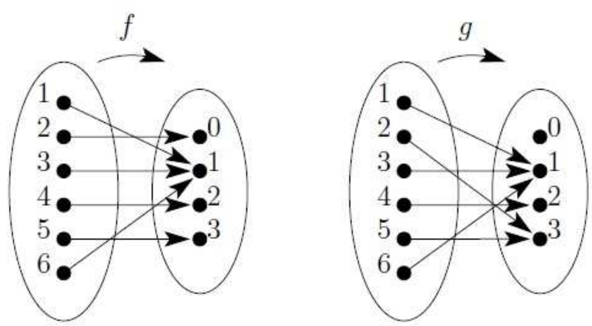


Figure 24: .

- To prove a function $f : \mathcal{X} \to \mathcal{Y}$ is one-to-one, we usually use the method of direct proof:

  suppose x_1 and x_2 are elements of $\mathcal{X}$ such that $f(x_1) = f(x_2)$, and

  show that $x_1 = x_2$.

- To prove that a function $f : \mathcal{X} \to \mathcal{Y}$ is *not* one-to-one, we usually

  find elements x_1 and x_2 in $\mathcal{X}$ such that $f(x_1) = f(x_2)$ but $x_1 \neq x_2$.

Example 115. *Define the functions* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *by* $f(x) = |x| + 1$ *and* $g(x) = 2x^3 - 1$*, for all* $x \in \mathbb{R}$.

(a) Is f one-to-one? Prove this, or else give a counterexample.

(b) Is g injective? Prove this, or else give a counterexample.

- A **hash function** is a function defined from a larger to a smaller set of integers (generally using the mod function).

  Hash functions are used to speed up searching through long lists by splitting them into a series of shorter lists.

Example 116. *Consider the student numbers of all the current students at a university (say these are 8 digit numbers). Searching through all the, say $40\,000$, student numbers would take a long time. So the university partitions the list into 100 sublists, each of approximately 400 students.*

This is done by defining the following **hash function** *which maps each 8-digit student number, say n, to an element x from the set $\{1, 2, 3, \ldots, 99\}$, such that*

$$H(n) = x, \quad \textit{where } x \Leftrightarrow n \ (\textit{mod}\ 100).$$

((**a**) *Calculate to which sublist each of the following student numbers would be allocated.*

(i) 40076832 $\quad$ (ii) 39987928 $\quad$ (iii) 41134032 $\quad$ (iv) 41004098

(**b**) *Is the function H one-to-one? Explain.*

- Let f be a function from a set $\mathcal{X}$ to a set $\mathcal{Y}$. The function f is **onto** (or **surjective**) if and only if given any element $y \in \mathcal{Y}$, it is possible to find an element $x \in \mathcal{X}$ with the property that $y = f(x)$. Or, equivalently,

  $f : \mathcal{X} \to \mathcal{Y}$ is onto if and only if $\forall y \in \mathcal{Y}, \exists x \in \mathcal{X}$ such that $f(x) = y$.

- A function $f : \mathcal{X} \to \mathcal{Y}$ is **not onto** if and only if there exists some $y \in \mathcal{Y}$ such that for all $x \in \mathcal{X}$, $f(x) \neq y$.

Example 117. *Which of the following arrow diagrams define onto (surjective) functions from $X = \{1, 2, 3, 4, 5, 6\}$ to $Y = \{0, 1, 2, 3\}$? Explain.*

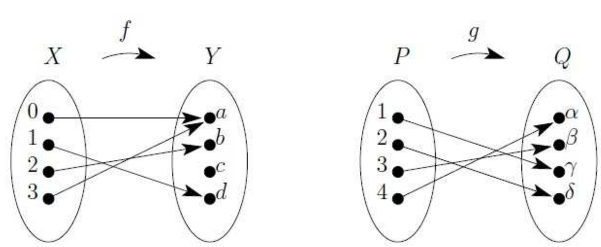


Figure 25: .

- To prove a function $f : \mathcal{X} \to \mathcal{Y}$ is onto, usually:

suppose that y is any element of $\mathcal{Y}$, and

show that there is an element of $\mathcal{X}$ with $f(x) = y$.

- To prove that a function $f : \mathcal{X} \to \mathcal{Y}$ is *not* onto, we usually

  find an element y of $\mathcal{Y}$ such that $y \neq f(x)$ for *any* $x \in \mathcal{X}$.

Example 118. *Define the functions $f : \mathbb{R} \to \mathbb{R}^+ \cup \{0\}$ and $g : \mathbb{Z} \to \mathbb{Z}$ by $f(x) = x^2$, for all $x \in \mathbb{R}$ and $g(x) = x^2$, for all $x \in \mathbb{Z}$.*

(a) Is f surjective? Prove this, or else give a counterexample.

(b) Is g onto? Prove this, or else give a counterexample.

Example 119. *Define a function $f : \mathbb{R} \to \mathbb{R}$ by $f : x \mapsto 2x + 4$ for all $x \in \mathbb{R}$. Show that f is onto.*

- A **one-to-one correspondence** (or **bijection**) from a set $\mathcal{X}$ to a set $\mathcal{Y}$ is a function $f : \mathcal{X} \to \mathcal{Y}$ that is both one-to-one and onto (so both injective and surjective).

Example 120. *Let $X = \{0, 1, 2, 3\}$ and $Y = \{0, 1, 4, 9\}$, and let the function $f : X \to Y$ satisfy $f(x) = x^2$ for all $x \in X$.*
Is f a one-to-one correspondence? Explain.

Example 121. *Is the function $f : \mathbb{R} \to \mathbb{R}$ a one-to-one correspondence, where $f(x) = x^3 - 2$ for all $x \in \mathbb{R}$? Justify your answer.*

Theorem 16. *Suppose that the function $f : \mathcal{X} \to \mathcal{Y}$ is both one-to-one and onto; so suppose f is a one-to-one correspondence (a bijection) from $\mathcal{X}$ to $\mathcal{Y}$. Then there exists a function $f^{-1} : \mathcal{Y} \to \mathcal{X}$ that is defined as follows: Given any element $y \in \mathcal{Y}$,*

$$f^{-1}(y) = \text{ the unique element } x \in \mathcal{X} \text{ such that } f(x) = y.$$

In other words,

$$f^{-1}(y) = x \text{ if and only if } y = f(x).$$

- The function f^{-1} is called the **inverse function** of f.

Example 122. *Given the functions f and g illustrated in the following arrow diagrams, find (if they exist) f^{-1} and g^{-1}. If they do exist, draw their arrow diagram.*

Example 123. *Find the inverse (if it exists) of the function $g : \mathbb{R} \to \mathbb{R}$ where $g(x) = 2x+5$ for all $x \in \mathbb{R}$.*

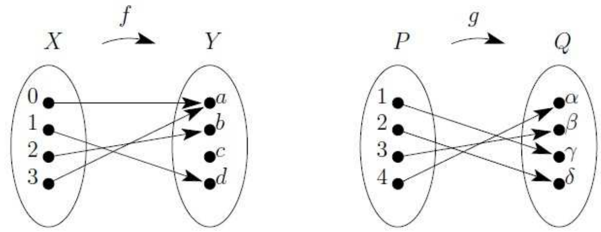


Figure 26: .

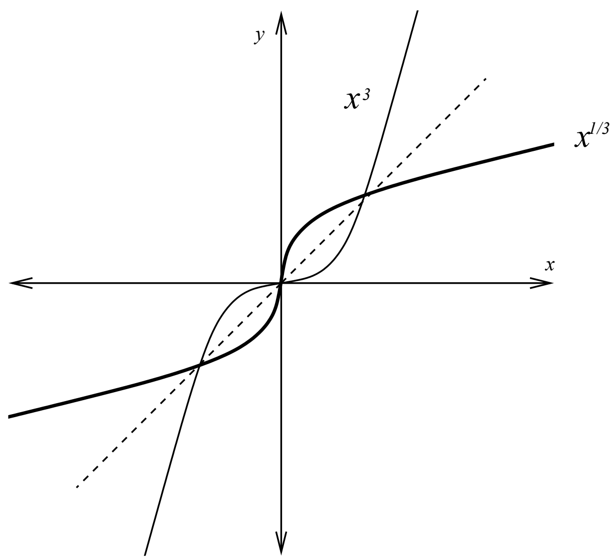


Figure 27: $f(x) = x^3$. Draw in the graph of $f^{-1}(x) = x^{1/3}$.

Inverse functions

Example 124. *The function $f(x) = x^3$ and its inverse $f^{-1}(x) = x^{1/3}$) are shown in Figure 27.*

Example 125. *$f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is not 1-1 and therefore has no inverse. However, $x \geq 0$ gives a 1-1 function $f : [0, \infty) \to \mathbb{R}$, $f(x) = x^2$ with range $[0, \infty)$. The inverse of this function is then $f^{-1} : [0, \infty) \to [0, \infty)$, $f^{-1}(x) = \sqrt{x}$. Similarly the negative half of the function $f(x) = x^2$ is 1-1, with inverse $f^{-1} : [0, \infty) \to (-\infty, 0]$, $f^{-1}(x) = -\sqrt{x}$.*

This technique is often used when the function is not 1-1 over its entire domain: just take a part where it is 1-1 and determine the inverse for that part.

Logarithms

Logarithms are the inverse functions of the exponential functions.

From the graph of $y = a^x$ ($a \neq 1$ a positive constant), we see that it is 1-1 and thus has an

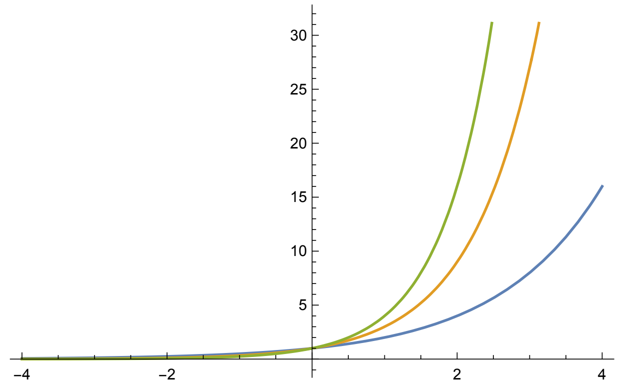


Figure 28: Three plots of $y = a^x$ with $a = 2, 3, 4$.

inverse, denoted $\log_a x$. From this definition we have the following facts:

$$\log_a(a^x) = x \quad \forall x \in \mathbb{R},$$

$$a^{\log_a x} = x \quad \forall x > 0.$$

What is the domain and range of $f(x) = \log_a(x)$?

The domain of $\log_a(x)$ is $(0, \infty)$ and its range is $\mathbb{R}$.

Inverse trigonometric functions

The function $y = \sin x$ is 1-1 if we just define it over the interval $[-\pi/2, \pi/2]$; see Figure 30. The inverse function for this part of $\sin x$ is denoted $\arcsin x$. Thus $\arcsin x$ is defined on the interval $[-1, 1]$ and takes values in the range $[-\pi/2, \pi/2]$. The graph can easily be obtained by reflecting the graph of $\sin x$ about the line $y = x$ over the appropriate interval; see Figure 29.

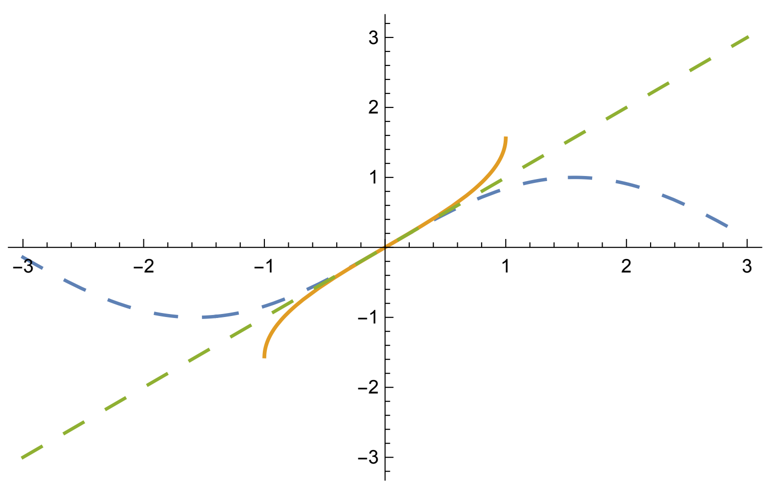


Figure 29: $f(x) = \sin x$ defined on $[-\pi/2, \pi/2]$ reflected about the line $y = x$ to give $f^{-1}(x) = \arcsin x$.

Similarly $y = \cos x$ is 1-1 on the interval $[0, \pi]$ and its inverse function is denoted $\arccos x$. The function $\arccos x$ is defined on $[-1, 1]$ and takes values in the range $[0, \pi]$. Figure 30 below shows the graphs of $f(x) = \cos x$ before and after reflection about the line $y = x$. This gives the graph of $f^{-1}(x) = \arccos x$.

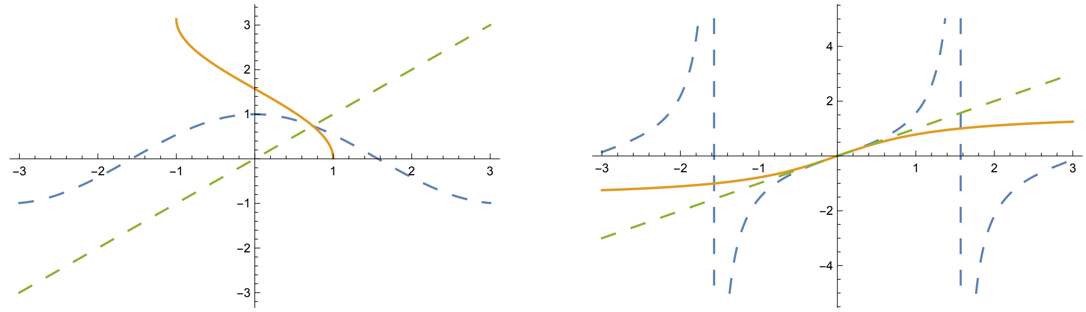


Figure 30: Left: $f(x) = \cos x$ defined on $[0, \pi]$ reflected about the line $y = x$ to give $f^{-1}(x) = \arccos x$. Right: $f(x) = \tan x$.

Also, $\tan x$ is 1-1 on the open interval $(-\pi/2, \pi/2)$ with inverse function denoted by $\arctan x$. Hence $\arctan$ has the domain $(-\infty, \infty)$ with values in the range $(-\pi/2, \pi/2)$.

4.7 Real functions of several variables

You should be familiar with functions f of one variable. The tools of calculus were useful because you could:

- Sketch the graph $y = f(x)$ of f.
- Find the minima and maxima of f.
- Analyse the slope of f by calculating f'.
- Find approximations of f using Taylor series.
- Find solutions to $f(x) = 0$.

Many familiar formulas are essentially just functions of more than one variable. For example, the volume V of a box is a function of its width, height and depth: $V(w, h, d) = whd$, and the profile of a vibrating string is a function of time and position along the string: $f(x, t) = A \sin x \cos t$.

Consider the volume of a cylinder as a function of two variables:

$$V(r, \ell) = \pi r^2 \ell.$$

We can visualise this function by graphing V in terms of ℓ and r.

The equation of a plane can be expressed using scalars or vectors. In this section we will sketch planes in $\mathbb{R}^3$ and determine their scalar equations.

In $\mathbb{R}^3$, we usually take z pointing upwards and the (x, y)-plane to be horizontal. The x and y coordinates give the position on the ground and z gives the height.

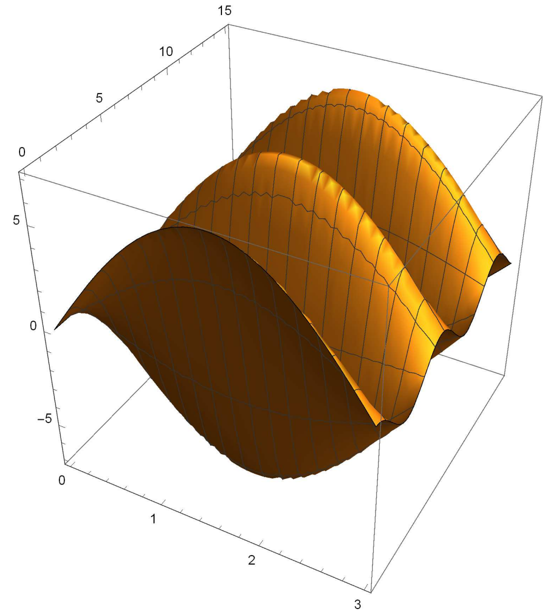


Figure 31: Vibration of a string, $f(x,t) = A\sin x\cos t$, where $t \in [0,15]$ seconds, $x \in [0,3]$ m, f is amplitude, in mm.

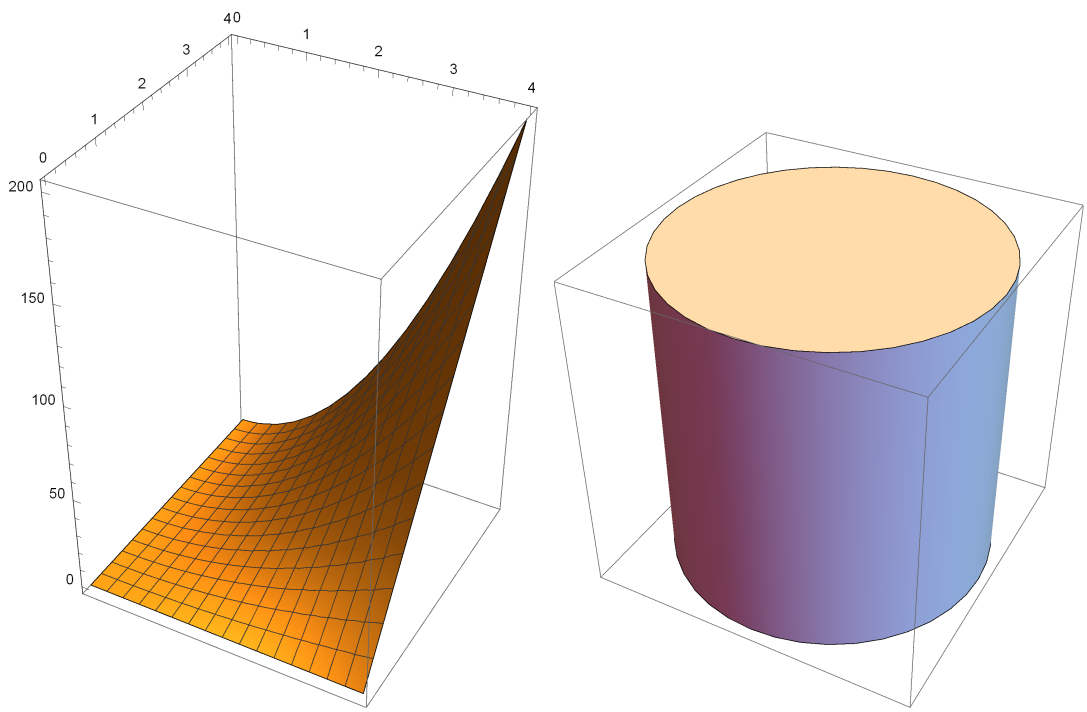


Figure 32: The volume of a cylinder, $V = (r,\ell) = \pi r^2 \ell$.

Horizontal planes

The x- and y-axes lie in the horizontal plane $z = 0$. All other horizontal planes are parallel to $z = 0$ and are given by the equation $z = c$.

Example 126. *Sketch the horizontal plane $z = 2$.*

Arbitrary planes

The general equation of a plane in $\mathbb{R}^3$ is given by

$$ax + by + cz = d$$

with a, b, c, d fixed real numbers. If the plane is not vertical, i.e., $c \neq 0$, this equation can be rearranged so that z is expressed as a function of x and y:

$$z = F(x,\, y) = -(a/c)x - (b/c)y + (d/c) = mx + ny + z_0.$$

The easiest way to sketch the plane by hand is to use the *triangle method*: If all of $a, b, c \neq 0$ the plane $ax+by+cz = d$ intercepts each axis at precisely one point. These three points make up a triangle which fixes the plane.

The plane $x + 2y + z = 4$ intersects the x-axis at $x = 4$, the y-axis at $y = 2$ and the z-axis at $z = 4$.

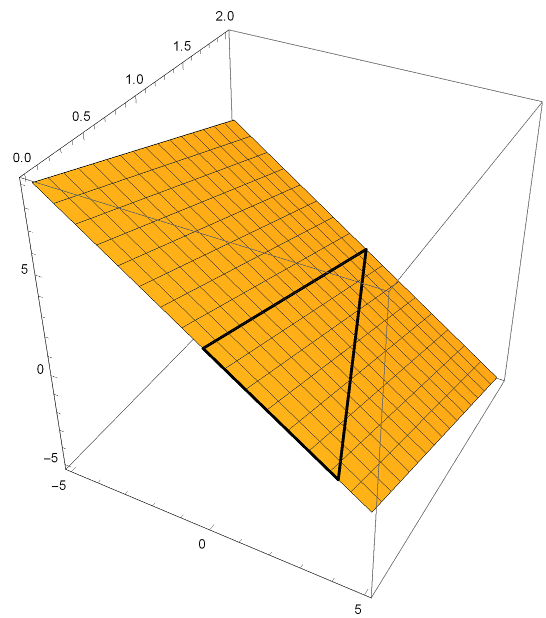


Figure 33: The plane $x + 2y + z = 4$ and lines connecting the points $(4, 0, 0)$, $(0, 2, 0)$, and $(0, 0, 4)$.

The triangle method is based on the simple fact that *any* three points that lie in a plane uniquely determine this plane *provided these three points do not lie on a single straight line*.

It is customary to say *the* equation of a plane, even though it is not unique. Multiplying the equation of a plane by a nonzero constant gives another equation for the same plane. For example, $x - 2y + 3z = 4$ and $-2x + 4y - 6z = -8$ are equations of the same plane. The same holds true for lines.

Example 127. *Find the equation of the plane through* $(0, 0, 5)$, $(1, 3, 2)$ *and* $(0, 1, 1)$.

Let $z = ax + by + d$.

The first point gives $d = 5$.

The 2nd point gives $2 = a + 3b + 5$ *and the 3rd point gives* $1 = b + 5$.

Solving simultaneously gives $a = 9$ *and* $b = -4$.

So the plane is $z = 9x - 4y + 5$.

Contour diagrams

Geographical maps have curves of constant height above sea level, or curves of constant air pressure (isobars), or curves of constant temperature (isothermals). Drawing contours is an effective method of representing a 3-dimensional surface in two dimensions. We now look at functions f of two variables. A contour is a curve corresponding to the equation $z = f(x, y) = C$.

Consider the surface $z = f(x, y) = x^2 + y^2$ sliced by horizontal planes $z = 0,\ z = 1,\ z = 2, \ldots$

Plane	Contour	Description
$z = 0$	$x^2 + y^2 = 0$	$x = 0, y = 0$
$z = 1$	$x^2 + y^2 = 1$	Circle radius 1
$z = 2$	$x^2 + y^2 = 2$	Circle radius $\sqrt{2}$
$z = 3$	$x^2 + y^2 = 3$	Circle radius $\sqrt{3}$
$z = 4$	$x^2 + y^2 = 4$	Circle radius 2

Note that as the radius increases, the contours are more closely spaced.

Example 128. *Draw a contour diagram of f given by*

$$f(x, y) = \sqrt{x^2 + y^2}.$$

If horizontal planes are equally spaced, say $z = 0,\ c,\ 2c,\ 3c, \ldots,$ it is not hard to visualise the surface from its contour diagram. Spread-out contours mean the surface is quite flat and closely spaced ones imply a steep climb.

Note that the contours of the last two functions were all circles. Such surfaces have circular symmetry. When x and y only appear as $x^2 + y^2$ in the definition of f, then the graph of f has circular symmetry about the z axis. The height z depends only on the radial distance $r = \sqrt{x^2 + y^2}$.

$z = f(x, y) = \exp^{-x^2 - y^2}$

Example 129. *Draw a contour diagram for* $z = f(x, y) = x^2 + 4y^2 - 2x + 1$.

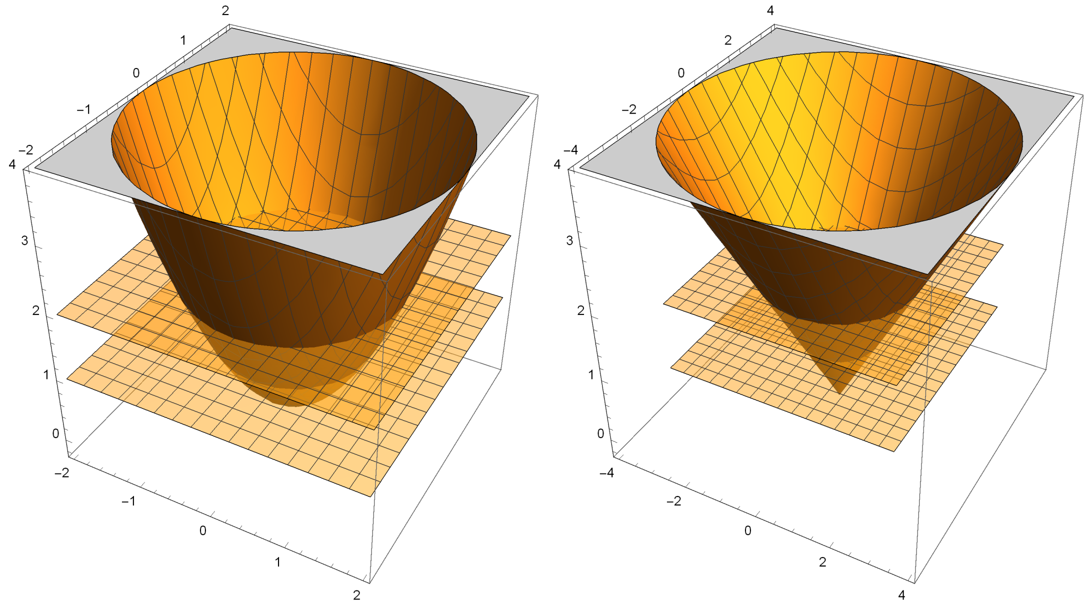


Figure 34: Left: A potential well: $z = f(x, y) = x^2 + y^2$. Right: A Cone: $z = f(x, y) = \sqrt{x^2 + y^2}$.

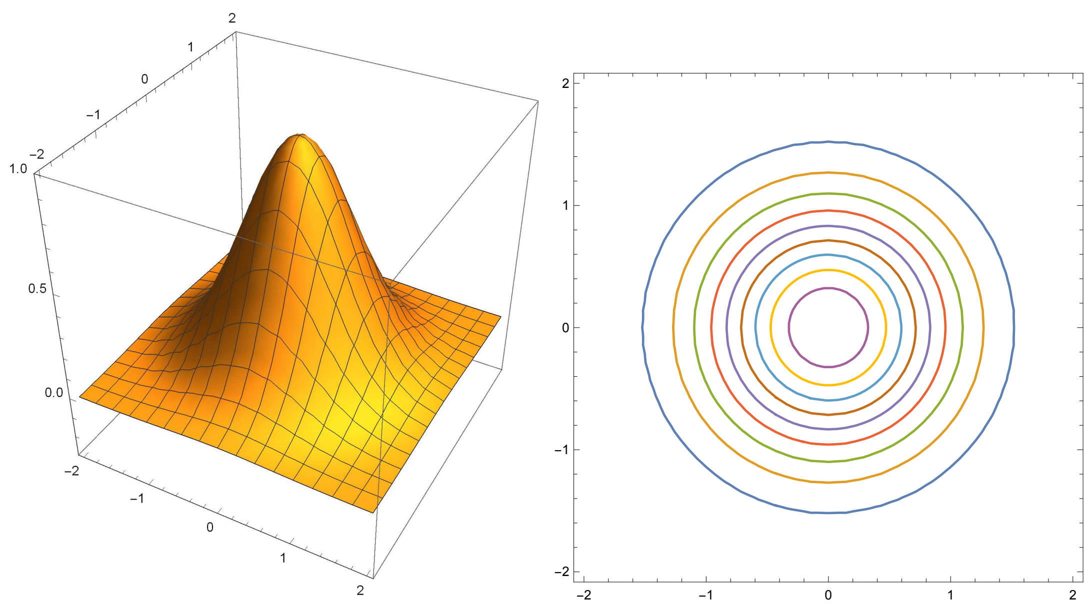


Figure 35: Left: The surface: $z = f(x, y) = e^{-x^2-y^2}$. Right: The contours $z = 0.1n$, $n = 1, 2, \ldots 10$.

Plane	*Contour*	*Description*
$z = 1$	$(x-1)^2 + 4y^2 = 1$	*ellipse centre* $(1, 0)$ x *intercepts:* $x = 0$, $x = 2$ *intersects line* $x = 1$ *at* $y = \pm\frac{1}{2}$
$z = 2$	$(x-1)^2 + 4y^2 = 2$	*ellipse centre* $(1, 0)$ x *intercepts:* $x = 1 \pm \sqrt{2}$ *intersects line* $x = 1$ *at* $y = \pm\frac{1}{\sqrt{2}}$
$z = 3$	$(x-1)^2 + 4y^2 = 3$	*ellipse centre* $(1, 0)$ x *intercepts* $x = 1 \pm \sqrt{3}$

A **saddle**, for example $z = x^2 - y^2$, has hyperbolic contours.

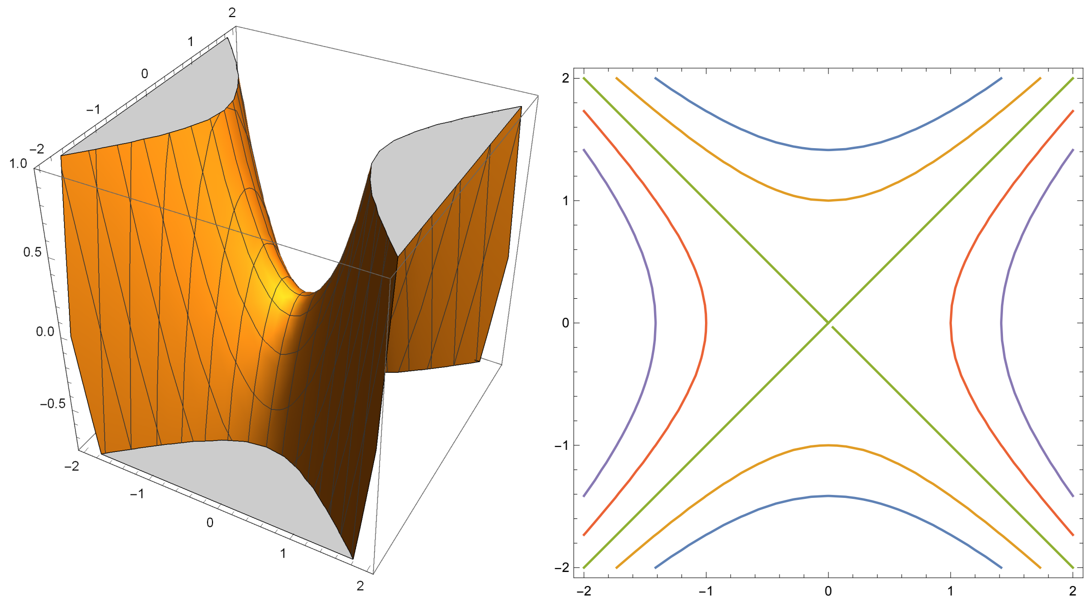


Figure 36: Left: The saddle $z = f(x, y) = x^2 - y^2$. Right: The contours $z = -2, -1, 0, 1, 2$.

$$
\begin{aligned}
z &= 0: \\
z &= 1: \\
z &= 2: \\
z &= -1: \\
z &= -2:
\end{aligned}
$$

Example 130.

- *Sketch the contour diagram for $z = f(x, y) = 9x^2 - 4y^2 + 2$ with contours at $z = -2,\ 2,\ 6$ and 10.*
- *Sketch a contour diagram for $z = f(x,\ y) = x^2$ and use this to sketch the graph of f.*
- *Sketch a contour diagram for $z = f(x, y) = x - y^2$.*
- *Sketch contour diagrams of $z = f(x, y) = x$ and $z = g(x, y) = x + y$.*

Note: Contour diagrams of functions whose graphs are planes consist of equidistant parallel lines. (Equidistant lines in $\mathbb{R}^2$are always parallel, unlike $\mathbb{R}^3$)

It is possible to construct the plane itself from its contour diagram provided the contours are labeled. Let the plane be $z = mx + ny + c$. Any point in the plane $(x_0,\ y_0,\ z_0)$ gives c:

$$c = z_0 - mx_0 - ny_0.$$

To find m consider moving along the plane in the positive x direction between two contours. Calculate the change in z, Δz, between the two contours, and the change in x, Δx, as you move from one contour line to the next. move in the x direction only, i.e. in a plane $y = C$).

Then $m = \dfrac{\Delta z}{\Delta x}$.

Similarly take a plane $x = C$ and move in the positive y direction to calculate Δy.

Then $n = \dfrac{\Delta z}{\Delta y}$.

Example 131. *Find the plane given by the following contour diagram:*

Plane	*Contour*
$z = 0$	$y = -2x + 5$
$z = 1$	$y = -2x + 2$

- *First note that $\Delta z = 1 - 0 = 1$. Moreover, if $y = 0$ then from $2x + y = 3$ and $2x + y = 2$ we obtain $x = 3/2$ and $x = 1$, respectively. So $\Delta x = 1 - 3/2 = -1/2$. It follows that $m = \dfrac{\Delta z}{\Delta x} = -2$.*
- *Similarly, if $x = 0$ then from $2x + y = 3$ and $2x + y = 2$ we obtain $y = 3$ and $y = 2$, respectively. So $\Delta y = 2 - 3 = -1$. It follows that $n = \dfrac{\Delta z}{\Delta y} = -1$.*
- *So the plane is $z = c - 2x - y$. To find c we see that the point $(1, 1, 0)$ is on the plane $z = 0$ and satisfies $2x + y = 3$. So this point must be on the plane $z = c - 2x - y$. This leads to $c = 3$.*

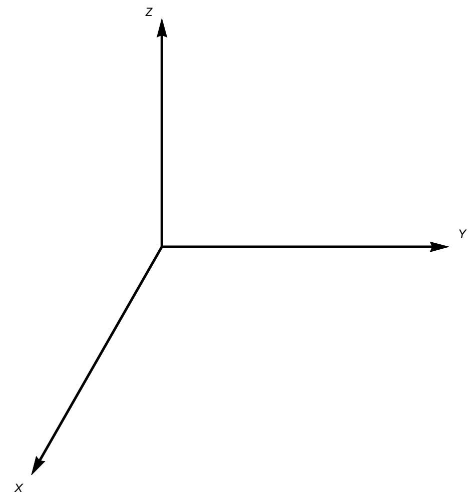
A plane $z = mx + ny + c$ does not have just one slope. It has slope m in the x direction and slope n in the y direction.

To visualise this, imagine starting at $(0,\, 0,\, c)$ and walking in the $(x,\, z)$ plane along the line $z = mx + c$ (with slope m). Or walk in the $(y,\, z)$ plane along the line $z = ny + c$ (with slope n).

Example 132. *Find the plane with slope 6 in the x direction and 4 in the y direction which passes through the point $(1,\, 5,\, 4)$.*

$z = mx + ny + c$ with $m = 6$, $n = 4$ and since

$$\begin{aligned} c &= 4 - 6.1 - 4.5, \\ &= -22, \\ \textit{we have}\quad z &= 6x + 4y - 22. \end{aligned}$$



4.8 Application

5 Sequences, their limits and series

Regular patterns and repeating processes occur all around us. We will investigate some mathematical ways to deal with these. We shall concentrate on *sequences* and *mathematical induction*. The latter is a formal method of proving the truth of some result,often involving a sequence, which you suspect is true. We shall also meet *summation notation*, and the *Well-Ordering Principle* for the integers.

5.1 Sequences of real numbers

Start counting your ancestors, remembering that each person has a mother and a father:

2 parents; they each have two parents (4 grandparents); they each have two parents ... 2, 4, 8, 16, 32, 64, 128 ...

This is an example of a *sequence* where the first term of the sequence is 2, the second term is 4 or 2^2, the third term is 8 or 2^3, ... and the kth term is 2^k. So we can denote the sequence by a_k where $a_k = 2^k$.

Such a sequence can also be written as $\{a_k\}_{k=1}$, where $a_k = 2^k$.

- In a sequence, each individual element is called a **term**.
- We can write a sequence as $a_1, a_2, \ldots, a_n$ if it has just n terms.

  If it is infinite, we can write $a_1, a_2, \ldots$.

  (The first term need not be a_1; for example $b_0, b_1, \ldots, b_m, \ldots$ is another sequence.)
- An explicit formula or **general formula** for a sequence is a rule showing how the value of the kth term a_k depends upon k.

Compute the first four terms of the sequence

$$a_i = (-1)^{i+1}\frac{n}{i+3}, \quad \text{for all integers } i \geqslant 1.$$

- Recall that a typical even integer is of the form $2k$ where k is any integer.

  And a typical odd integer is of the form $2k+1$ (or $2\ell - 1$, etc.), where k (or ℓ) is any integer.
- Note that $(-1)^{2n} = 1$ and $(-1)^{2n+1} = -1$.

Example 133. *Find an explicit formula for a sequence which has the following initial terms:* $-1,\ 4,\ -27,\ 256,\ -3125, \ldots$.

Summation notation

We use a Greek capital letter Σ (sigma) to indicate sum, as follows, where a_i is some term of a sequence and $m < n$:

$$\sum_{i=m}^{n} a_i \;=\; a_m + a_{m+1} + \cdots + a_n.$$

Note that i here is a *dummy symbol*. It does not appear in the right hand side expanded form.

Write the following summation in expanded form by writing out the first five terms and the final term.

$$\sum_{i=1}^{n}(2i-1) \;=$$

Example 134. *Express the following using summation notation.*

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{n\cdot(n+1)}$$

Given that $\displaystyle\sum_{i=1}^{n} i^2 \;=\; \frac{n(n+1)(2n+1)}{6}$*, provide a simplified expression for* $\displaystyle\sum_{i=1}^{n+1} i^2$.

- To denote a **product** of terms, we use $\prod$.

  So for instance $\displaystyle\prod_{k=m}^{n} a_k \;=\; a_m\cdot a_{m+1}\cdot\cdots\cdot a_n$. (Here $m < n$.)

- **Factorial**: For each *positive* integer n, we write $n!$ (read **n factorial**), where $n! \;=\; n(n-1)(n-2)\ldots 3\times 2\times 1$.

  So $1! = 1,\; 2! = 2\cdot 1 = 2,\; 3! = 3\cdot 2\cdot 1 = 6,\; 4! = 4\times 3\times 2\times 1 = 24$, etc.

- We define **zero factorial** or $0!$ to be 1. That is, $0! = 1$.

Simplify $\displaystyle\prod_{j=2}^{5} 2^j$.

Simplify the following: **(a)** $\dfrac{10!}{8!\,2!}$ **(b)** $\dfrac{(n+2)!}{n!}$.

Example 135. *List the first terms of the sequence* $a_n = \dfrac{n}{n+1}$.

$$0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$$

Example 136. *List the first terms of the recursive sequence* $a_{n+1} = \dfrac{1}{3 - a_n}$ *with* $a_1 = 2$*.*

$$2, 1, \frac{1}{2}, \frac{2}{5}, \ldots$$

Note that for the recursive form, you must express one or more *initial conditions* or starting values. The recursive form is actually a *difference equation*, and there are similarities with differential equations.

The Fibonacci sequence is defined by the recursive formula

$$F_1 = 1, \qquad F_2 = 1, \qquad F_n = F_{n-1} + F_{n-2}.$$

Example 137. *The first eight terms of the sequence are given by*

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Find the next three terms.

We can also write the n-th Fibonacci number in closed-form.

$$F_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n. \tag{4}$$

Example 138. *Use* (4) *to find the* F_8 *and compare your results to Example 137.*

Note that it is possible to prove the closed form of the Fibonacci sequence by using induction, which was covered previously in Section 3.4 (in particular strong induction is used for this proof).

Defining sequences recursively

A sequence is said to be defined **recursively** if (some) initial values of the sequence are given, and later terms of the sequence are defined in terms of some fixed number of earlier terms. Here's an example:

$$a_1 = 1,\ a_2 = 3 \text{ and for } n \geqslant 3,\ a_n = a_{n-1} + 2a_{n-2}.$$

We shall formally define a recurrence relation, give some examples, and explore methods for finding an explicit formula for a recursively defined sequence.

- We can define a sequence informally by giving the first few terms, until the general pattern is obvious:

  3, 5, 7, . . .

  But is the next term 9 (all odd integers from 3 on) or is it 11 (all odd primes)?

- We can define a sequence by giving its nth term, and stating where it starts; for example:

$$a_n = \frac{(-1)^n}{n+2} \text{ for all } n \geqslant 0.$$

- A third way of defining a sequence is to use **recursion**.
We give a **recurrence relation**, relating later terms in the sequence to earlier ones, and also some **initial conditions**; for example:

$$\{c_i\} \text{ for } i \geqslant 0, \text{ is given by (i) } c_0 = 1 \text{ and } c_1 = 2, \text{ and}$$
$$\text{(i) } c_s = c_{s-1} + 3c_{s-2} + 1$$

Let $c_0, c_1, c_2, \ldots$ be a sequence which satisfies the following recurrence relation.

For all integers $k \geq 2$,

$$\begin{aligned}(1)\ c_k &= (k-1)c_{k-1} + kc_{k-2} + k,\\ (2)\ c_0 &= 1 \text{ and } c_1 = 2.\end{aligned}$$

Calculate the values of c_2, c_3 and c_4.

Let $b_0, b_1, b_2, \ldots$ be a sequence which satisfies the following recurrence relation.

For all integers $i \geqslant 2$,

$$b_i = 5b_{i-1} - 6b_{i-2}.$$

Write out expressions for b_{i+1} and b_{i+2}.

Show that the sequence $a_k = 5 \cdot 2^k$, for $k \geqslant 0$, satisfies the recurrence relation $a_n = 2\,a_{n-1}$ for all integers $n \geqslant 1$.

The Catalan numbers can be defined as $C_n = \dfrac{1}{n+1}\dbinom{2n}{n}$ for all integers $n \geqslant 1$. Find C_i for $i = 1, 2, 3$. Then show that $C_k = \dfrac{4k-2}{k+1}C_{k-1}$ for all integers $k \geqslant 2$.

(a) Make a list of all bit strings of lengths 0, 1, 2, 3 and 4 which do not contain the bit pattern $1\,0$.

(b) For each integer $n \geqslant 0$, let

$$S_n = \left[\begin{array}{l}\text{the number of bit strings of length } n\\ \text{which do not contain the pattern } 1\,0\end{array}\right].$$

Find S_0, S_1, S_2, S_3 and S_4.

(c) Find the number of bit strings of length ten which do not contain the pattern $1\,0$. (Use a recurrence relation for S_n.)

Now suppose that your parents deposit \$ 1000 in a bank on your 10-th birthday, and that the interest rate was fixed for 50 years at 5% per annum, compounded annually. Set up a recurrence relation, and then calculate how much money you will have on your 30-th birthday.

How much will be in the account on your 50-th birthday, if you don't withdraw anything at any time?

5.2 Limits of sequences

Let $\{a_n\}_{n=0}^{\infty}$ be a sequence. Then

$$\lim_{n\to\infty} a_n = \ell \quad , \quad \ell \in \mathbb{R}\,, \qquad \text{means} \;:$$

a_n approaches ℓ as n gets larger and larger, ie. a_n is close to ℓ for n sufficiently large.

More formally, the definition of a limit of a sequence states that for any $\epsilon > 0$, there exists a whole number N, so that whenever the n^{th} term of our sequence is bigger than N, i.e. $n \geqslant N$, we are within epsilon of ℓ. We write this formally as:

$$\forall\, \epsilon > 0, \exists\, N \;: |a_n - \ell| < \epsilon\, \forall n > N$$

Convention

If a sequence $\{a_n\}_{n=0}^{\infty}$ has limit $\ell \in \mathbb{R}$, we say that a_n *converges* to ℓ and that the sequence $\{a_n\}_{n=0}^{\infty}$ is *convergent*. Otherwise the sequence is *divergent*.

Example 139. *Determine if the sequences $\{a_n\}$, with a_n as given below, are convergent and if so find their limits.*

1. $a_n = \dfrac{1}{n}$

$$\begin{aligned}\lim_{n\to\infty} a_n &= \lim_{n\to\infty} \frac{1}{n} \\ &= 0\end{aligned}$$

So $\{a_n\}$ converges to 0, i.e. $a_n \to 0$ as $n \to \infty$.

2. $a_n = \dfrac{1}{2} a_{n-1}$, $a_0 = -1$.

In direct form:

$$a_n = -\frac{1}{2^n},\ \textit{for}\ n = 0, 1, \ldots$$

$$\begin{aligned}\Rightarrow \quad \lim_{n\to\infty} a_n &= \lim_{n\to\infty} \left(-\frac{1}{2^n}\right) \\ &= 0\end{aligned}$$

$\Rightarrow a_n \to 0$ *as* $n \to \infty$.

3. $a_n = (-1)^n$

The first few terms are: $1, -1, 1, -1, \ldots$*. This sequence oscillates between* ± 1 *and therefore does not converge* $\Rightarrow$ *divergent sequence.*

4. $a_n = r^n$*, for* $r = \dfrac{1}{2}, r = 1, r = 2$

$a_n = r^n$*: 4 cases.*
(1) $|r| < 1$*:* $\lim\limits_{n\to\infty} a_n = 0 \;\Rightarrow\; \{a_n\}$ *converges to 0.*
(2) $|r| > 1$*:* $\lim\limits_{n\to\infty} a_n = \infty \;\Rightarrow\; \{a_n\}$ *diverges.*
(3) $r = 1$*:* $\lim\limits_{n\to\infty} a_n = 1 \;\Rightarrow\; \{a_n\}$ *converges to 1.*
(4) $r = -1$*:* $\lim\limits_{n\to\infty} a_n$*does not exist* $\Rightarrow\; \{a_n\}$ *diverges.*

So:
$r = \frac{1}{2}$*: converges to 0.*
$r = 1$*: converges to 1.*
$r = 2$*: diverges.*

The following limit laws apply provided that the separate limits exist (that is $\{a_n\}$ and $\{b_n\}$ are convergent):

Theorem 17. *Suppose that* $\{a_n\}$ *and* $\{b_n\}$ *are convergent sequences such that*

$$\lim_{n\to\infty} a_n = \ell \quad \text{and} \quad \lim_{n\to\infty} b_n = m$$

and c *is a constant. Then*

- $\lim\limits_{n\to\infty} (a_n \pm b_n) = \ell \pm m$*;*
- $\lim\limits_{n\to\infty} ca_n = c\ell$*;*
- $\lim\limits_{n\to\infty} (a_n b_n) = \ell m$*; and*
- $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{\ell}{m}$*, provided* $m \neq 0$.

Theorem 18 (Squeeze)**.** *If* $a_n \le b_n \le c_n$ *for* $n \ge n_0$ *for some* $n_0 \in \mathbb{Z}^{>0}$ *and*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = \ell,$$

then

$$\lim_{n\to\infty} b_n = \ell.$$

Example 140. *Use the squeeze theorem on* $\{a_n\}$*, where* $a_n = \dfrac{1}{n}\sin(n)$.

Since

$$-1 \le \sin(n) \le 1, \text{ for } n \ge 1$$

we have

$$-\frac{1}{n} \le a_n \le \frac{1}{n}, \text{ for } n \ge 1.$$

Now

$$\lim_{n\to\infty} -\frac{1}{n} = \lim_{n\to\infty} \frac{1}{n} = 0,$$

so the squeeze theorem gives $\lim\limits_{n\to\infty} a_n = 0$.

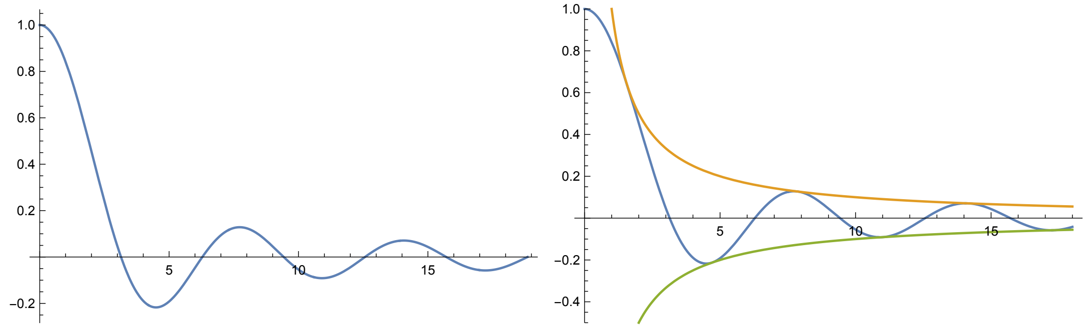


Figure 37: A plot of $f(x) = \frac{\sin(x)}{x}$ and $\pm\frac{1}{x}$ together with $f(x)$.

Useful sequences to remember

(1) For constant c, $\lim_{n\to\infty} c^n = \begin{cases} 0\,, & \text{if } |c| < 1 \\ 1\,, & \text{if } c = 1. \end{cases}$

Sequence $\{c^n\}_{n=0}^{\infty}$ is divergent if $c = -1$ or $|c| > 1$.

(2) For constant $c > 0$, $\lim_{n\to\infty} c^{1/n} = \lim_{n\to\infty} \sqrt[n]{c} = 1$.

(3) $\lim_{n\to\infty} \frac{1}{n^r} = 0$ for $r > 0$.

(4) $\lim_{n\to\infty} n^{1/n} = \lim_{n\to\infty} \sqrt[n]{n} = 1$.

(5) $\lim_{n\to\infty} \frac{1}{n!} = 0$.

(6) For constant c, $\lim_{n\to\infty} \frac{c^n}{n!} = 0$.

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

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

5.3 Series

Infinite sums (notation)

If we have an infinite sum we write

$$a_0 + a_1 + a_2 + \ldots + a_n + \ldots = \sum_{n=0}^{\infty} a_n.$$

Note that the lower bound ($n = 0$) of the sum may vary.

Motivation

Series come from many fields.

1. Approximation to problem solutions:
   - a_0 zeroth order approximation
   - $a_0 + a_1$ (a_1 small) first order approximation
   - $a_0 + a_1 + a_2$ (a_2 very small) second order approximation
   - $a_0 + a_1 + a_2 + \ldots + a_n$ nth order
   - $a_0 + a_1 + a_2 + \ldots + a_n + \ldots$ exact solution, provided the series converges
2. Current state of a process over infinite time horizon
3. Approximating functions via Taylor/Fourier Series, e.g.,

$$\begin{aligned} e^x &= 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \ldots, \\ f(t) &= c_0 + c_1 \sin t + c_2 \sin(2t) + c_3 \sin(3t) + \ldots \end{aligned}$$

4. Riemann sums

Definition: convergence

Given a series $\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \ldots$, let s_n denote its nth partial sum:

$$\begin{aligned} s_0 &= a_0 \\ s_1 &= a_0 + a_1 \\ s_2 &= a_0 + a_1 + a_2 \\ &\vdots \\ s_n &= a_0 + a_1 + \ldots + a_n = \sum_{k=0}^{n} a_k\,. \end{aligned}$$

If the sequence $\{s_n\}$ is convergent (i.e. $\lim_{n\to\infty} s_n = s$ with $s \in \mathbb{R}$), then the series $\sum_{n=0}^{\infty} a_n$ is said to be *convergent* and we write

$$\lim_{n\to\infty} s_n = \sum_{n=0}^{\infty} a_n = s.$$

The number s is called the *sum* of the series. Otherwise the series is said to be *divergent*.

Example 141. *Does the series* $\sum_{n=0}^{\infty}(-1)^n$ *converge or diverge?*

$$\begin{aligned}
s_0 &= & (-1)^0 = 1 \\
s_1 &= & (-1)^0 + (-1)^1 = 0 \\
s_2 &= & 1 \quad \textit{etc.}
\end{aligned}$$

Therefore

$$s_{2m} = 1 \quad \textit{and} \quad s_{2m+1} = 0$$

$$\Rightarrow \lim_{n\to\infty} s_n \textit{ does not exist } \Rightarrow \{s_n\} \textit{ diverges}$$

So $\sum_{n=0}^{\infty}(-1)^n$ *diverges.* $s_0 = (-1)^0 = 1,\ s_1 = (-1)^0 + (-1)^1 = 0,\ s_2 = 1,$ *etc.*
so $s_{even} = 1$ *and* $s_{odd} = 0$.
Therefore, $s_{2m} = 1,\ s_{2m+1} = 0$ *so* s_n *does not converge.*

Example 142. *Show that the series* $\sum_{n=0}^{\infty}\frac{1}{(n+1)(n+2)}$ *is convergent.*

$$\textit{First note that } \sum_{n=0}^{\infty}\frac{1}{(n+1)(n+2)} = \sum_{n=1}^{\infty}\frac{1}{n(n+1)}.$$

Then

$$\begin{aligned}
s_n &= \sum_{k=1}^{n}\frac{1}{k(k+1)} \\
&= \sum_{k=1}^{n}\left(\frac{1}{k} - \frac{1}{k+1}\right) \\
&= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\
&= 1 - \frac{1}{n+1}.
\end{aligned}$$

Therefore

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty}\left(1 - \frac{1}{n+1}\right) = 1.$$

Therefore the series is convergent and we can see that

$$\sum_{n=0}^{\infty}\frac{1}{(n+1)(n+2)} = 1.$$

This is an example of a telescoping series.

Theorem 19 (p test)**.** *For $p \in \mathbb{R}$, the p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is*

$$\textit{convergent if} \quad p > 1$$

$$\textit{and divergent if} \quad p \leq 1.$$

Note the above sum is from $n = 1$. This is just a matter of taste since

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \sum_{n=0}^{\infty} \frac{1}{(n+1)^p}$$

Theorem 20 (nth term test)**.** *If $\sum_{n=0}^{\infty} a_n$ is convergent then $\lim_{n\to\infty} a_n = 0$.*

Harmonic series

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

diverges (p-series with $p = 1$).

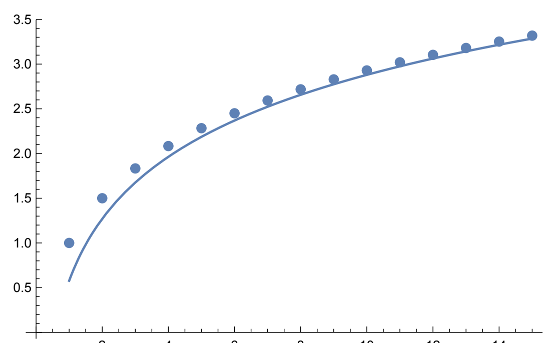


Figure 38: Euler found that $f(x) = \sum_{n=1}^{x} n^{-1}$ approaches $g(x) = \gamma + \ln(x)$ and x increases.

Note that γ is known as the Euler-Mascheroni constant. We can show that $f(x)$ approaches $g(x)$ asymptotically, however note that both f and g diverge.

Proof that the harmonic series diverges

Write the partial sums

$$
\begin{aligned}
s_1 &= 1\\
s_2 &= 1+\frac{1}{2}\\
s_3 &= 1+\frac{1}{2}+\frac{1}{3}\\
s_4 &= 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\\
&\geq 1+\frac{1}{2}+\underbrace{\frac{1}{4}+\frac{1}{4}}\\
&= 1+\frac{1}{2}+\quad\frac{1}{2}\\
\\
s_8 &= 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)\\
&\geq 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)\\
&= 1+\frac{1}{2}+\quad\frac{1}{2}+\qquad\qquad\frac{1}{2}.
\end{aligned}
$$

In general $s_{2^n} \geq 1+n/2$, so the partial sums approach infinity as $n \to \infty$.

Theorem 21 (Geometric series)**.** *The series*

$$\sum_{n=0}^{\infty} ar^n$$

is convergent if $|r| < 1$ *with* $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$ *and divergent if* $|r| \geq 1$.

Example 143. *Consider a bank located in Europe, which introduces negative real rates (charges* 5% *interest per year to store savings). If a person deposits \$* 100 *at the start of each year into their account, and this continues indefinitely, then the total savings in the account will converge to:*

$$\sum_{n=1}^{\infty} 100(0.95)^n = \frac{100}{1-0.95} = 2000$$

In the case of regular positive interest rates, the total savings will grow indefinitely (diverge).

5.4 Application

6 Real functions: Limits and continuity

Limits arise when we want to find the tangent to a curve. Once we understand limits, we can proceed to studying continuity and calculus in general. Limits are a fundamental notion to calculus, so it is important to understand them clearly.

6.1 Limits of functions

Let $f(x)$ be a function and $\ell \in \mathbb{R}$. We say $f(x)$ *approaches* the limit ℓ (or *converges* to the limit ℓ) as x approaches a if we can make the value of $f(x)$ arbitrarily close to ℓ (as close to ℓ as we like) by taking x to be sufficiently close to a (on either side of a) but not equal to a.

We write

$$\lim_{x \to a} f(x) = \ell.$$

Roughly speaking, $f(x)$ is close to ℓ for all x values sufficiently close to a, with $x \neq a$. The limit "predicts" what should happen at $x = a$ by looking at x values close to but not equal to a.

Definition 11. *Let f be a function defined on some open interval that contains the number a, except possibly a itself. Then we write*

"a" may not stay in the domain of f

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

if for every number $\varepsilon > 0$ there is a number $\delta > 0$ such that

$$|f(x) - \ell| < \varepsilon \textit{ whenever } 0 < |x - a| < \delta.$$

We can sometimes determine the limit of a function by looking at its graph. Figure 39 shows two functions with the same limit at $x = 5$.

Function 1 is described by

$$f(x) = \begin{cases} 2x & \text{for } 0 \leq x \leq 5, \\ -2x + 20 & \text{for } 5 < x \leq 10, \end{cases}$$

while function 2 is described by

$$g(x) = \begin{cases} 2x & \text{for } 0 \leq x < 5, \\ 2 & \text{for } x = 5, \\ -2x + 20 & \text{for } 5 < x \leq 10. \end{cases}$$

Each function has a limit of 10 as x approaches 5. It does not matter that the value of the second function is 2 when x equals 5 since, when dealing with limits, we are only interested in the behaviour of the function as x approaches 5.

ILO: using defination to prove lim f(x)=l is Not Required

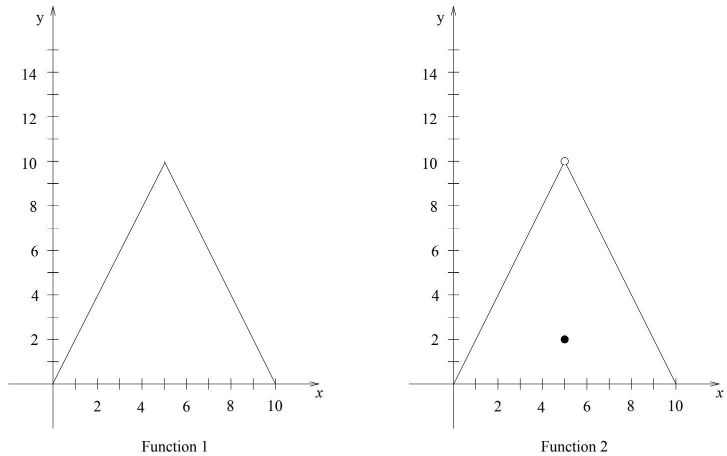


Figure 39: Two functions with limit equal to 10 as $x \to 5$.

Properties

Suppose that c is a constant and the limits $\ell = \lim_{x\to a} f(x)$ and $m = \lim_{x\to a} g(x)$ exist for some fixed $a \in \mathbb{R}$. Then

$$\lim_{x\to a}(f(x) \pm g(x)) = \lim_{x\to a} f(x) \pm \lim_{x\to a} g(x) = \ell \pm m$$

$$\lim_{x\to a} cf(x) = c \cdot \lim_{x\to a} f(x) = c \cdot \ell$$

$$\lim_{x\to a} f(x) \cdot g(x) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x) = \ell \cdot m$$

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} = \frac{\ell}{m} \text{ if } m \neq 0$$

Example 144. *Find the value of* $\lim_{x\to 1}(x^2+1)$.

$$\lim_{x\to 1}(x^2+1) = 1 + 1 = 2$$

For polynomials, $\sin x$, $\cos x$, e^x one finds limit by substitution.

Example 145. *Determine the value of* $\lim_{x\to 1} \frac{x-1}{x^2-1}$.

$f(x) = \frac{x-1}{x^2-1}$ is not defined at $x=1$

In the limit $\lim_{x\to a} f(x)$ x is always not equal to a

$$\lim_{x\to 1} \frac{x-1}{x^2-1} = \lim_{x\to 1} \frac{x-1}{(x+1)(x-1)} \qquad \textit{remember } x \neq 1$$

$$= \lim_{x\to 1} \frac{1}{x+1} = \frac{1}{2}$$

(Valid since $x \to 1$, so $x \neq -1$)

Required: Factorisation of quadratic polynomials in $\mathbb{R}$ $(a \neq 0)$

$ax^2+bx+c = a\left(x-\frac{-b+\sqrt{b^2-4ac}}{2a}\right)\left(x-\frac{-b-\sqrt{b^2-4ac}}{2a}\right)$

when $b^2-4ac \geq 0$

Theorem 22 (Squeeze principle)**.** *Suppose*

$$\lim_{x\to a} g(x) = \ell = \lim_{x\to a} h(x)$$

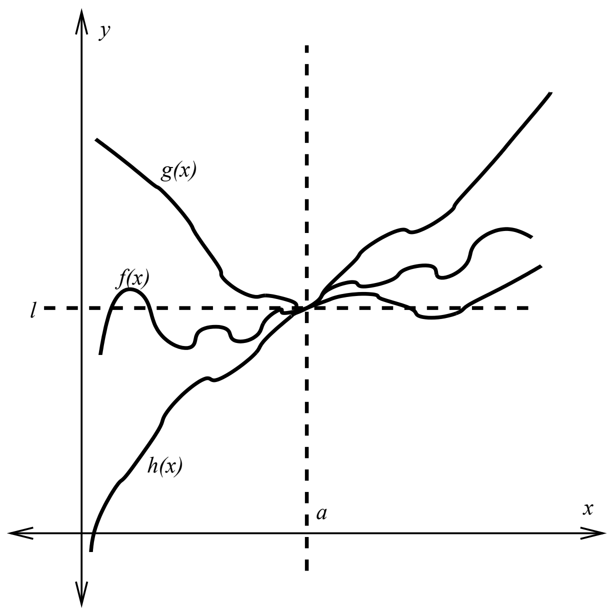
and, for x close to a ($x \neq a$)

$$h(x) \leq f(x) \leq g(x).$$

Then

$$\lim_{x\to a} f(x) = \ell.$$

See the graph in Figure 40.



area of arc = $\frac{\theta}{2}$

arc length = θ

radius = 1

θ: unit: radius

Figure 40: The squeeze principle

circumference = 2π

area of the unit disk = π

Example 146. *Prove that* $\lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right) = 0.$

For $x \neq 0$,

$$-1 \leq \sin\left(\frac{1}{x}\right) \leq 1.$$

Therefore,

$$-x^2 \leq x^2 \sin\left(\frac{1}{x}\right) \leq x^2.$$

Take $g(x) = x^2$, and $h(x) = -x^2$ in the squeeze theorem.
As $x \to 0,\ g(x) \to 0$ and $h(x) \to 0$.

So by the squeeze principle

$$\lim_{x\to 0}\ x^2 \sin\left(\frac{1}{x}\right) = 0.$$

Some important limits

The following limits are fundamental. Combined with the properties given in 6.1 and the Squeeze Principle given in 22, these will enable you to compute a range of other limits.

Here $a \in \mathbb{R}$ is arbitrary. Polynomials. finding limits by substitution

$$(1)\quad \lim_{x\to a} 1 = 1 \qquad (2)\quad \lim_{x\to a} x = a \qquad (3)\quad \lim_{x\to 0} \frac{\sin x}{x} = 1$$

$$(4)\quad \lim_{x\to 0} \frac{1-\cos(x)}{x^2} = \frac{1}{2} \qquad (5)\quad \lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

6.2 Additional real function limit concepts

Infinite Limits

Let f be a function defined on both sides of a, except possibly at a itself. Then

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

means that the values of $f(x)$ can be made arbitrarily large by taking x sufficiently close to a, but not equal to a.
Similarly,

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

means that the values of $f(x)$ can be made arbitrarily large negatively by taking x sufficiently close to a, but not equal to a. $\to +\infty, -\infty$ respectively
In these cases, we say that $f(x)$ *diverges* to $\pm\infty$. We also say that the limit does not exist in these cases. Note that the limit properties in section 6.1 do not necessarily apply if the limits diverge.

$$\lim_{x\to a} f(x) = +\infty, \ \text{ if}$$

$$\forall L > 0, \exists \delta : \forall x, | x - a | < \delta, f(x) > L$$

$0<$

One-sided limits

Consider the piecewise function

$$f(x) = \begin{cases} 1, & x \geq 0 \\ -2, & x < 0. \end{cases}$$

This function has the graph depicted in Figure 41.

Notice that $\lim\limits_{x\to 0, x>0} f(x) = 1$, but $\lim\limits_{x\to 0, x<0} f(x) = -2$. Therefore, the limit as $x \to 0$ does not exist. We can, however, talk about the *one-sided limits*.

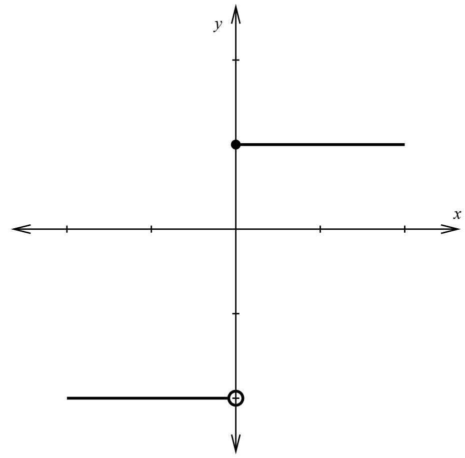


Figure 41: The limit of this function as $x \to 0$ does not exist.

In the above example, we say that the limit as $x \to 0$ *from above* (or *from the right*) equals 1 and we write

$\lim_{x \to 0, x > 0}$

$$\lim_{x \to 0^+} f(x) = 1.$$

Similarly, we say that the limit as $x \to 0$ *from below* (or *from the left*) equals -2 and we write

$\lim_{x \to 0, x < 0}$

$$\lim_{x \to 0^-} f(x) = -2.$$

In general, for $\lim_{x \to a^+} f(x) = \ell$, just consider x with $x > a$ and similarly for $\lim_{x \to a^-} f(x) = \ell$, consider only $x < a$.

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

Limits as x approaches infinity we write $\lim_{x \to \pm\infty} f(x) = l$

We say that $f(x)$ approaches ℓ as $x \to \infty$ if $f(x)$ is arbitrarily close to ℓ for all large enough values of x. Formally, we write:

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

if, for every $\varepsilon > 0$ there exists $K \in \mathbb{R}$ such that $x > K \Rightarrow |f(x) - \ell| < \varepsilon$.

That is, $f(x)$ can be made arbitrarily close to ℓ by taking x sufficiently large. Similarly, we write

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

if $f(x)$ approaches ℓ as x becomes more and more negative.

Note:

$$\lim_{x \to \infty} \frac{1}{x} = 0 \text{ and } \lim_{x \to -\infty} \frac{1}{x} = 0.$$

then $\lim_{x \to 0} \frac{1}{x}$ does not exist. Furthermore, it is irrelevant that $f(x)$ is not actually defined at the point $x = 0$. as $x \to 0$ from the left, $\frac{1}{x}$ becomes arbitrarily large in the negative

direction. As $x \to 0$ from the right, $\frac{1}{x}$ becomes arbitrarily large in the positive direction. For the one-sided limits for $x \to 0$ we thus have:
$\lim\limits_{x\to 0^+} \frac{1}{x} = \infty, \ \lim\limits_{x\to 0^-} \frac{1}{x} = -\infty.$

In this case the limit $\lim\limits_{x\to 0} \frac{1}{x}$ does not exist, since the limits from the left and the right do not agree. Note that the fact that $f(x)$ is not actually defined at the point $x = 0$ is not directly relevant to this discussion.

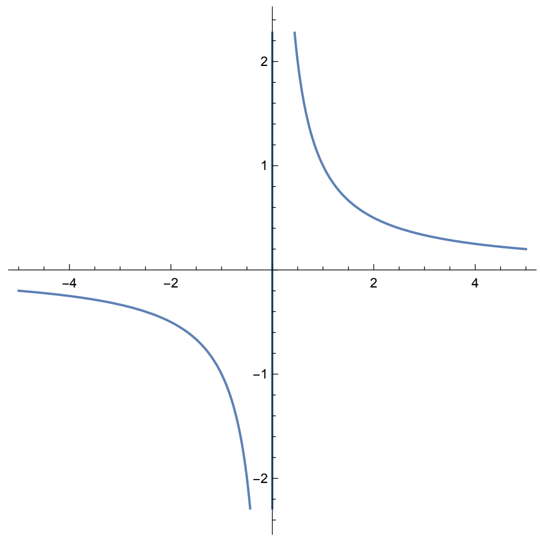


Figure 42: Graph of $f(x) = 1/x$.

Ratio of two polynomials, numerator degree = denominator degree

Example 147. *Find* $\lim\limits_{x\to\infty} \dfrac{2x^2+3}{3x^2+x}$.

Put $f(x) = 2x^2 + 3$ *and* $g(x) = 3x^2 + x$.

Then $\lim\limits_{x\to\infty} f(x) = \infty$ *and* $\lim\limits_{x\to-\infty} g(x) = \infty$
So this limit is undefined.

Technique: we divide top and bottom by the highest power of x in the denominator. The result is two new functions in the numerator and the denominator, the limits of which are both finite.

$$\begin{aligned}
\lim_{x\to\infty} \frac{2x^2+3}{3x^2+x}, &= \lim_{x\to\infty} \frac{(2x^2+3)/x^2}{(3x^2+x)/x^2}, \\
&= \lim_{x\to\infty} \frac{2+\frac{3}{x^2}}{3+\frac{1}{x}}, \\
&= \frac{2}{3}.
\end{aligned}$$

Example 148. *Find* $\lim\limits_{x\to\infty} \dfrac{8x^3+6x^2}{4x^3-x+12}$.

In this case, we divide the top and bottom by x^3, and proceed as before.

$$\begin{aligned}
\lim_{x\to\infty} \frac{8x^3+6x^2}{4x^3-x+12} &= \lim_{x\to\infty} \frac{(8x^3+6x^2)/x^3}{(4x^3-x+12)/x^3}, \\
&= \lim_{x\to\infty} \frac{8+\frac{6}{x}}{4-\frac{1}{x^2}+\frac{12}{x^3}}, \\
&= \frac{8}{4}, \\
&= 2.
\end{aligned}$$

Example 149 (ratio of two polynomials, numerator degree $>$ denominator degree)**.** *Find* $\lim_{x\to\infty} \frac{x^2+5}{x+1}$.

The highest power of x in the denominator is 1*, so we have*

$$\lim_{x\to\infty} \frac{x^2+5}{x+1} = \lim_{x\to\infty} \frac{x+\frac{5}{x}}{1+\frac{1}{x}} = \infty.$$

Example 150 (ratio of two polynomials, numerator degree $<$ denominator degree)**.** *Find* $\lim_{x\to\infty} \frac{x+1}{x^2+1}$

$$\lim_{x\to\infty} \frac{x+1}{x^2+1} = \lim_{x\to\infty} \frac{\frac{1}{x}+\frac{1}{x^2}}{1+\frac{1}{x^2}} = 0$$

The general method in the four previous examples has been to:

1. divide the numerator and denominator by the highest power of x in the denominator; and then

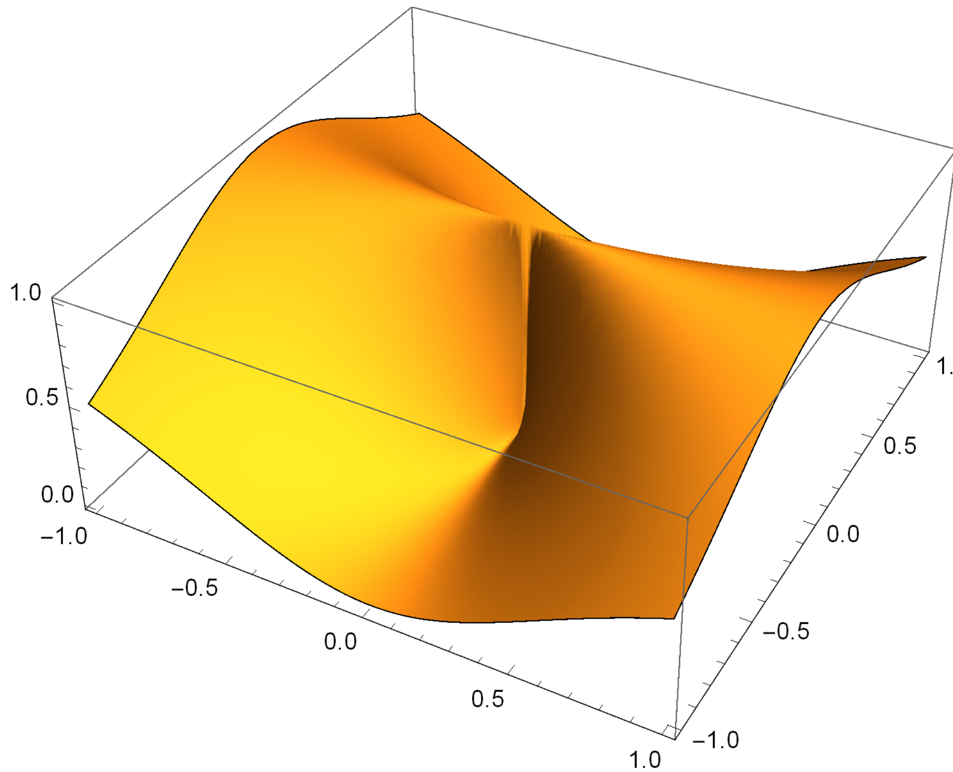
2. determine the limits of the numerator and denominator separately.

6.3 Multivariate limits

When f is a function of more than one variable, the situation is more interesting. There are more than two ways to approach a given point of interest. Consider the function

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

with domain given by $\mathbb{R}^2 \backslash \{(0,0)\}$.



Next we consider the limit as $(x, y) \to (0, 0)$.

(i) Approaching the origin along $y = 0$:

the limit of $\dfrac{x^2}{x^2 + y^2}$ as $(x, y) \to (0, 0)$ *along* $y = 0$ equals 1.

(ii) Approaching the origin along $x = 0$:

the limit of $\dfrac{x^2}{x^2 + y^2}$ as $(x,\ y) \to (0,\ 0)$ *along* $x = 0$ equals 0.

$\lim\limits_{(x,y)\to(0,0)} = 0.$

Does $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ exist?

In general, for the limit $\lim\limits_{(x,y)\to(a,b)} f(x, y)$ to exist, it is necessary that every path in D approaching (a, b) (the point (a, b) itself may or may not be in D) gives the same limiting value. This gives us the following method for finding if a limit does not exist.

The above notation is somewhat deficient and perhaps one should write

$$\lim_{(x,y)\to_D(a,b)} f(x, y)$$

to indicate that only paths in D terminating in (a, b) (which itself may or may not be in D) are considered. For example, if $f(x, y) = x^2 + y^2$ with $D = \{(x, y) :\ x^2 + y^2 < 1\}$ then $\lim\limits_{(x,y)\to(1,0)} f(x, y)$ exists and is 1. However, if

$$f(x, y) = \begin{cases} x^2 + y^2 & \text{for } D = \{(x, y) :\ x^2 + y^2 < 1\} \\ 0 & \text{for } D = \{(x, y) :\ x^2 + y^2 > 1\} \end{cases}$$

then $\lim_{(x,y)\to(1,0)} f(x,y)$ does not exist.

Let $D = \mathbb{R}^2 \backslash \{(0,0)\}$ and $f : D \to \mathbb{R}^2$ be given by $f(x,y) = \dfrac{x^2-y^2}{x^2+y^2}$. Show that $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

With the same D as above but now $f(x,y) = \dfrac{xy}{x^2+y^2}$, show that $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

In reality, there are infinitely many paths terminating at a given point, say (a,b), in $\mathbb{R}^2$, and therefore we would need to prove that $\lim_{(x,y)\to(a,b)} f(x,y)$ exists for infinitely many cases. Luckily, there are methods that can deal with infinitely many paths simultaneously, however these methods(typically ϵ-δ proofs) are outside the scope of this course.

6.4 Continuity

Definition 12. *We say that a function f is* continuous at a *if*

(i) $f(a)$ *is defined (that is, a is in the domain of $f(x)$);*

(ii) $\lim_{x\to a} f(x)$ *exists; and*

(iii) $\lim_{x\to a} f(x) = f(a)$.

If f is not continuous at a, we say that f is discontinuous at a, *or f has a* discontinuity at a.

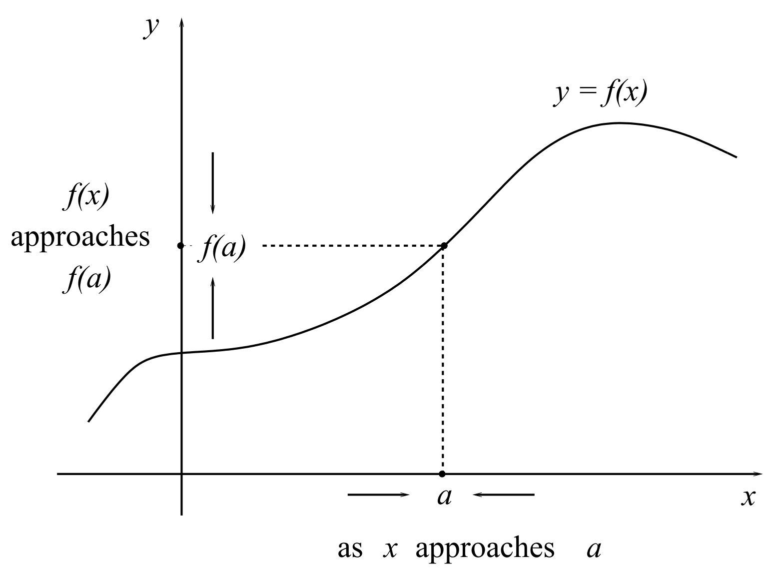


Figure 43: Graphical representation of continuity at $x = a$.

A function may not be continuous at $x = a$ for a number of reasons.

$\lim_{x\to a} f(x)$ and $f(a)$ are both defined, it is possible to have $\lim_{x\to a} f(x) \neq f(a)$.

Properties

If $f(x)$ and $g(x)$ are continuous at $x=a$ and c is a constant, then

The following functions are all continuous at $x=a$.

$$\begin{array}{ll}(1) & f(x) \pm g(x), \\ (2) & c \cdot f(x), \\ (3) & f(x) \cdot g(x), \text{ and} \\ (4) & \dfrac{f(x)}{g(x)}, \text{ provided } g(a) \neq 0,\end{array}$$

$f(x) \pm g(x), \quad f(x)g(x) \quad$ and $\quad \dfrac{f(x)}{g(x)}$
are also continuous at $x=a$ (provided $g(a) \neq 0$ in the last case).

Example 151. *Let*

$$f(x)=\begin{cases} x, & x \neq 0 \\ 1, & x=0. \end{cases}$$

Then $f(x)$ has a discontinuity at $x=0$. This is because $f(0)=1$ while $\lim_{x\to 0} f(x)=0$. Hence Condition (iii) of Definition 12 does not hold. See Figure 44.

Here $f(0)=1$, but $\lim_{x\to 0} f(x)=\lim_{x\to 0} x=0$, so Condition (ii) above does not hold. Hence, we say that f has a discontinuity at $x=0$. See Figure 44 for the graph of this function.

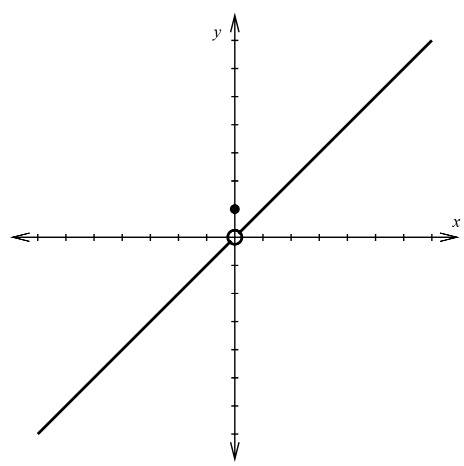


Figure 44: An example of a discontinuous function with discontinuity at $x=0$.

Example 152. *The function $f(x)=\dfrac{1}{x^2}$ (see Figure 49) is not continuous at $x=0$, since $f(0)$ is not defined, i.e. $0 \notin$ the domain of (f). Thus Condition (i) in 12 does not hold.*

Example 153.

$$f(x)=\begin{cases} x+1, & x \geq 0 \\ x^2, & x<0 \end{cases}$$

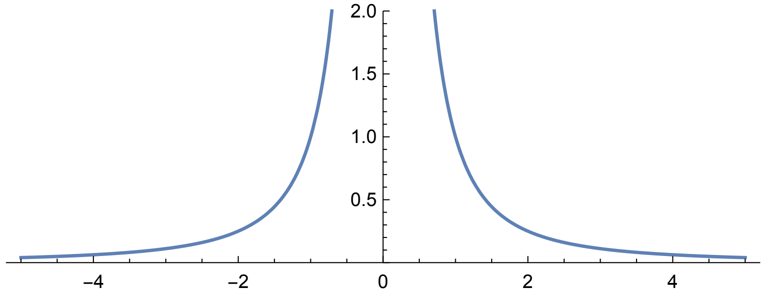


Figure 45: The function $f(x) = 1/x^2$ has a discontinuity at $x = 0$.

is not continuous at $x = 0$ *since*

$\lim_{x\to 0} f(x)$ *does not exist:*

$$\begin{aligned} \lim_{x\to 0^+} f(x) &= \lim_{x\to 0^+} (x+1) = 1 \\ \lim_{x\to 0^-} f(x) &= \lim_{x\to 0^-} x^2 = 0 \end{aligned}$$

The limit from above is different to the limit from below.

Continuity on intervals

We say that f is continuous on the *open* interval (a, b), if

f is continuous at c, for all $c \in (a, b)$.

If f is continuous on the *closed* interval $[a, b]$, then f is continuous on (a, b) and

$$\lim_{x\to a^+} f(x) = f(a), \quad \lim_{x\to b^-} f(x) = f(b).$$

You could think of a continuous function being one that on an interval can be drawn without lifting your pen.

Example 154.

- *Any polynomial in* x *is continuous on* $\mathbb{R}$. *For example,* $f(x) = ax^2 + bx + c$ *is continuous on* $\mathbb{R}$.
- e^x, $|x|$, $\sin x$, $\cos x$ *and* $\arctan x$ *are continuous on* $\mathbb{R}$.
- $f(x) = \ln\ x$ *is continuous on* $(0, \infty)$.

Theorem 23 (limit of a composite function)**.** *If* f *is continuous at* b *and* $\lim_{x\to a} g(x) = b$*, then* $\lim_{x\to a} f(g(x)) = f(b)$*. In other words,*

$$\lim_{x\to a} f(g(x)) = f(\lim_{x\to a} g(x)).$$

Theorem 23 essentially states that a limit symbol can be moved through a function symbol if the function is continuous and converges.

Similarly if $\lim_{x\to\infty} g(x) = b$ and f is continuous at $x = b$, then

$$\lim_{x\to\infty} f(g(x)) = f(b).$$

Corollary 1. *If g is continuous at a and f is continuous at $g(a)$ then $f \circ g$ is continuous at a.*

Multivariate continuity

Definition 13. *Given a function $f : D \to \mathbb{R}^2$, where D is an open subset of $\mathbb{R}^2$. Let $(a, b) \in D$. Then $f(x, y)$ is* **continuous** *at (a, b) if*

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b),$$

i.e., the limit $(x, y) \to (a, b)$ of $f(x, y)$ exists and is equal to $f(a, b)$.

If a function is continuous on all of D we say simply that it is continuous on D. Most of the functions we will consider are continuous. For example, polynomials in x and y are continuous on $\mathbb{R}^2$. As a rule of thumb, if a function with domain D is defined by a single expression it will be continuous on D.

Example 155. *Let $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$ and $D = \mathbb{R}^2 \backslash \{(0, 0)\}$. Is $f(x, y)$ a continuous function?*

Example 156. *If we edit the above example by instead defining f on all of $\mathbb{R}^2$ by taking $f(0, 0) = 0$, then is f a continuous function?*

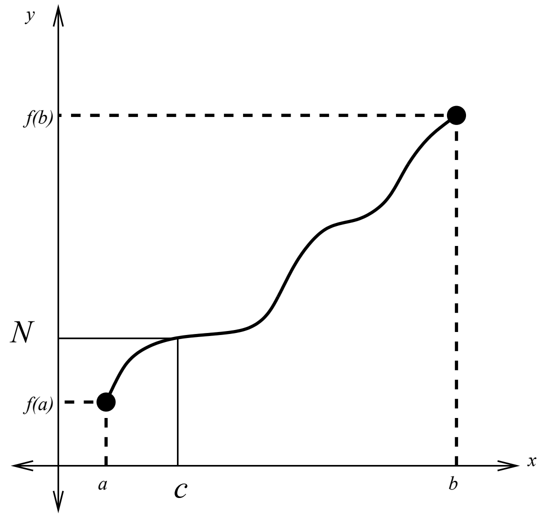
No, all of a sudden f would no longer be a continuous function: f would be continuous on all of D except for the point $(0, 0)$.

6.5 Intermediate value theorem

The following theorem is known as the intermediate value theorem (IVT).

Theorem 24 (IVT)**.** *Suppose that f is continuous on the closed interval $[a, b]$ and let N be any number between $f(a)$ and $f(b)$, where $f(a) \neq f(b)$. Then there exists a number $c \in (a, b)$ such that $f(c) = N$.*

The IVT states that a continuous function takes on *every* intermediate value between the function values $f(a)$ and $f(b)$. That is, we can choose any N between $f(a)$ and $f(b)$ which corresponds to a value c between a and b. Note that this is not necessarily true if f is discontinuous.



Example 157. *Suppose that a function f is continuous everywhere and that $f(-2) = 3$, $f(-1) = -1$, $f(0) = -4$, $f(1) = 1$, and $f(2) = 5$. Does the Intermediate-Value Theorem guarantee that f has a root on the following intervals?*

a) $[-2, -1]$ *b)* $[-1, 0]$ *c)* $[-1, 1]$ *d)* $[0, 2]$ *e)* $[1, 3]$

a) yes *b) no* *c) yes* *d) yes* *e) no*

Application of IVT (bisection method)

The *bisection method* is a procedure for approximating the zeros of a continuous function. It first cuts the interval $[a, b]$ in half (say, at a point c), and then decides in which of the smaller intervals ($[a, c]$ or $[c, b]$) the zero lies. This process is repeated until the interval is small enough to give a significant approximation to the zero itself.

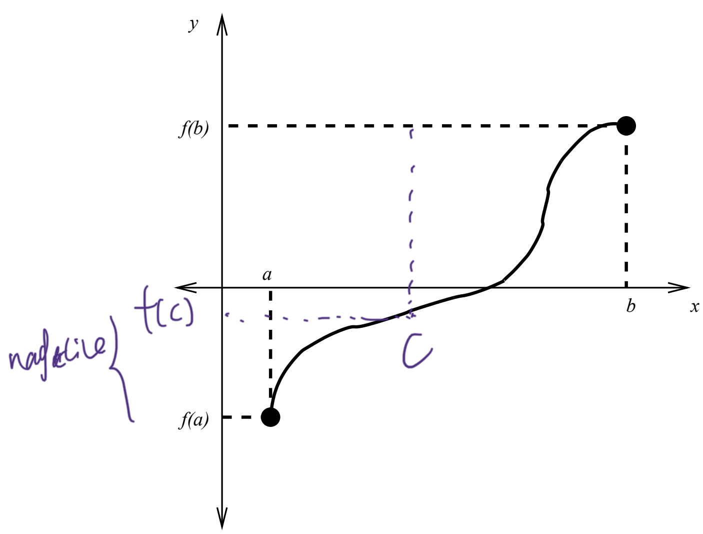
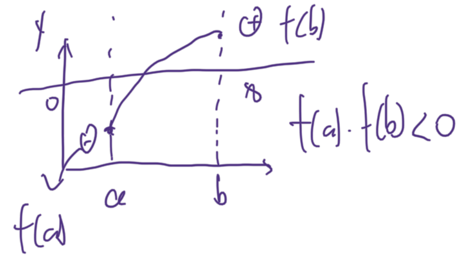


Figure 46: The bisection method is an application of the IVT.

We can present this bisection method as an *algorithm*:

(1) Given $[a, b]$ such that $f(a)f(b) < 0$, let $c = \dfrac{a+b}{2}$.



(2) If $f(c) = 0$ then quit; c is a zero of f.

(3) If $f(c) \neq 0$ then:

  (a) If $f(a)f(c) < 0$, a zero lies in the interval $[a, c)$. So replace b by c.

  (b) If $f(a)f(c) > 0$, replace a by c.

(4) If the interval $[a, b]$ is small enough to give a precise enough approximation then quit. Otherwise, go to step (1).

Note that:

$$
\begin{aligned}
f(a)f(c) < 0 \quad &\iff \quad f(a) < 0 \text{ and } f(c) > 0 \\
&\quad\text{or} \quad f(a) > 0 \text{ and } f(c) < 0 \\
f(a)f(c) > 0 \quad &\iff \quad f(a) > 0 \text{ and } f(c) > 0 \\
&\quad\text{or} \quad f(a) < 0 \text{ and } f(c) < 0
\end{aligned}
$$

6.6 Application

7 Derivatives, Optimisation and basic ODEs

Derivatives are a powerful tool when faced with situations involving *rates of change*. All rates of change can be interpreted as *slopes* of appropriate tangents. Therefore we shall consider the tangent problem and how it leads to a precise definition of the derivative.

7.1 Derivative definition and meaning

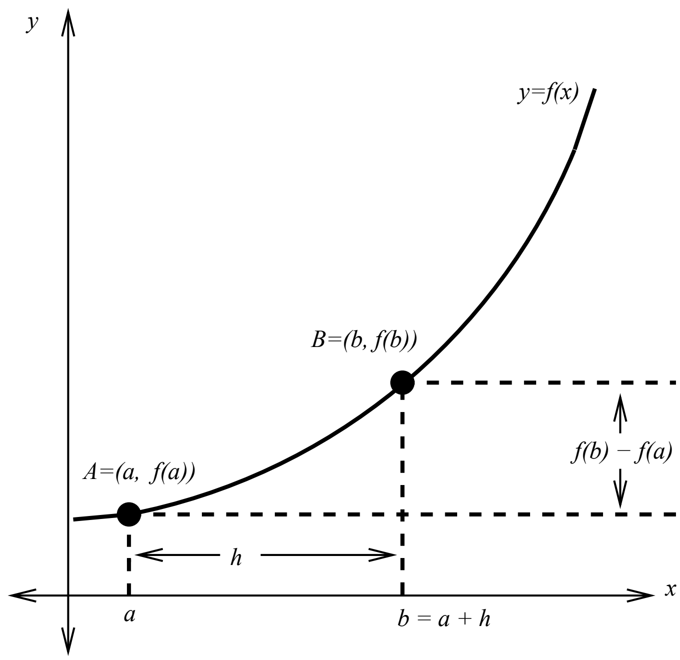


Figure 47: We want to determine the tangent line at the point A.

Consider the graph in Figure 47. We want to determine the tangent line at the point A ($x = a$) on the graph of $y = f(x)$, where $f(x)$ is 'nice enough'. We approximate the tangent at A by the chord AB where B is a point on the curve (close to A) with x-value $a+h$, where h is small. By looking at the graph, we can see that the slope m of the chord AB is given by

$$\begin{aligned} m &= \frac{\text{vertical rise}}{\text{horizontal run}} \\ &= \frac{f(b)-f(a)}{b-a} \\ &= \frac{f(a+h)-f(a)}{h}. \end{aligned}$$

As the point B gets closer to A, m will become closer to the slope at the point A. To obtain this value, we take the limit as $h \to 0$:

$$\text{The slope of the tangent to } y = f(x) \text{ at } A = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}.$$

Derivative, differentiability

The *derivative* of f at x is defined by

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

We say that f is *differentiable* at some point x if this limit exists. Further, we say that f is differentiable on an open interval if it is differentiable at every point in the interval. Note that $f'(a)$ is the slope of the tangent line to the graph of $y = f(x)$ at $x = a$.

We have thus defined a new function f', called the derivative of f. Sometimes we use the Leibniz notation $\frac{dy}{dx}$ or $\frac{df}{dx}$ in place of $f'(x)$.

Note that if f is differentiable at a, there holds:

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

Example 158. *Using the definition of the derivative (i.e., calculating "by first principles"), find the derivative of $f(x) = x^2 + x$.*

Using the definition for a derivative, we find:

$$\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} \frac{(x+h)^2 + (x+h) - (x^2 + x)}{h} \\
&= \lim_{h \to 0} \frac{x^2 + 2xh + h^2 + x + h - x^2 - x}{h} \\
&= \lim_{h \to 0} \frac{2xh + h^2 + h}{h} \\
&= \lim_{h \to 0} (2x + h + 1) \\
&= 2x + 1
\end{aligned}$$

Example 159. *Find the derivative from first principles of $f(x) = e^x$.*

We have from the definition of Euler's number e:

$$\lim_{h \to 0} \frac{e^h - 1}{h - 0} = 1, \qquad \textit{i.e.} \ \lim_{h \to 0} \frac{e^h - 1}{h} = 1.$$

Using the definition of the derivative, we then have

$$
\begin{aligned}
f'(x) &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\\
&= \lim_{h\to 0}\frac{e^{x+h}-e^x}{h}\\
&= \lim_{h\to 0}\frac{e^x e^h-e^x}{h}\\
&= \lim_{h\to 0} e^x\cdot\frac{e^h-1}{h}\\
&= e^x\lim_{h\to 0}\frac{e^h-1}{h}\\
&= e^x
\end{aligned}
$$

Rates of change

Recall that the slope of the tangent line to $y = f(x)$ at $x = a$ measures the *rate of change* of y with respect to x at the point $x = a$.

7.2 Working with derivatives

Some useful derivatives

Other useful derivatives are

$$
\begin{aligned}
\frac{d}{dx}(\text{constant}) &= 0\\
\frac{d}{dx}(x) &= 1\\
\frac{d}{dx}(x^\alpha) &= \alpha x^{\alpha-1},\ \text{for } \alpha\in\mathbb{R},\ x\geq 0 \text{ or for } x\in\mathbb{R} \text{ if } \alpha\in\mathbb{R}^+\\
\frac{d}{dx}(\sin x) &= \cos x\\
\frac{d}{dx}(\cos x) &= -\sin x
\end{aligned}
$$

Higher derivatives

Differentiating a function $y = f(x)$ n times, if possible, gives the nth derivative of f, usually denoted $f^{(n)}(x)$, $\dfrac{d^n f}{dx^n}$ or $\dfrac{d^n y}{dx^n}$. Sometimes when we deal with time derivatives we write $v(t) = \dot{d}(t)$ or $a(t) = \ddot{d}(t)$ and so on (where d, v, a are displacement, velocity and acceleration respectively.

Consider the following data set:

Rules for differentiation

Suppose that f and g are differentiable functions, and let c be a constant. Then we have the following rules:

(i) $(cf)' = cf'$;

(ii) $(f \pm g)' = f' \pm g'$;

(iii) $(fg)' = f'g + fg'$ (the product rule); and

(iv) $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$, whenever the denominator is not equal to 0 (quotient rule).

In the Leibniz notation, if f and g are both differentiable functions, and c is a constant, then the rules for differentiation can be expressed as:

(i) $\frac{d}{dx}(cf) = c\frac{df}{dx}$;

(ii) $\frac{d}{dx}(f \pm g) = \frac{df}{dx} \pm \frac{dg}{dx}$;

(iii) $\frac{d}{dx}(fg) = \frac{df}{dx}g + f\frac{dg}{dx}$ (the product rule); and

(iv) $\frac{d}{dx}\left(\frac{f}{g}\right) = \frac{\frac{df}{dx}g - f\frac{dg}{dx}}{g^2}$, for $g \neq 0$ (the quotient rule).

Example 160. *Use the rules above to find the derivative of* $f(x) = \tan x = \dfrac{\sin x}{\cos x}$.

For $x \neq (n + \frac{1}{2})\pi$, $n \in \mathbb{Z}$, *there holds* $\cos(x) \neq 0$, *so use the quotient rule:*

$$
\begin{aligned}
f'(x) &= \frac{(\sin x)' \cos x - \sin x(\cos x)'}{\cos^2 x} \\
&= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \\
&= \frac{1}{\cos^2 x} \qquad (\textit{ since } \cos^2 x + \sin^2 x = 1) \\
&= \sec^2 x
\end{aligned}
$$

The chain rule

If g and h are both differentiable and $f = g \circ h$ is the composite function defined by $f(x) = g(h(x))$, then f is differentiable and f' is given by the product

$$f'(x) = g'(h(x))h'(x).$$

In the Leibniz notation, if $y = g(u)$ and $u = h(x)$ are both differentiable functions, then

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

Example 161. *Differentiate $y = \sqrt{\sin x}$.*

Let $u = \sin x$. Then $y = u^{1/2}$. Thus, we have

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy}{du} \cdot \frac{du}{dx} \\ &= \left(\frac{1}{2}u^{-1/2}\right)(\cos x) \\ &= \frac{\cos x}{2\sqrt{u}} \\ &= \frac{\cos x}{2\sqrt{\sin x}} \end{aligned}$$

Derivative of inverse function

Suppose $y = f^{-1}(x)$, where f^{-1} is the inverse of f. To obtain $\dfrac{dy}{dx}$ we use

$$x = f(f^{-1}(x)) = f(y).$$

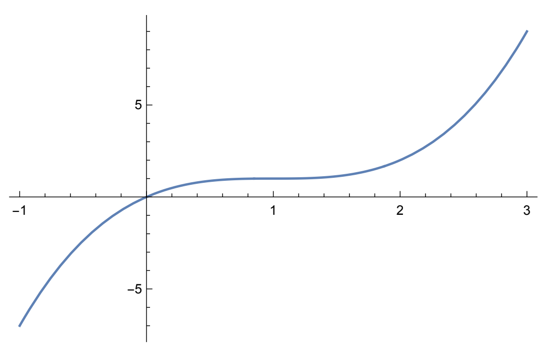


Figure 48: Plot of $f(x) = (x-1)^3 + 1$.

Differentiating both sides with respect to x using the chain rule gives

$$\begin{aligned} 1 &= \frac{d}{dx}(x) \\ &= \frac{d}{dx}f(y) \\ &= \frac{df}{dy}\cdot\frac{dy}{dx} \\ &= \frac{dx}{dy}\cdot\frac{dy}{dx} \end{aligned}$$

Hence, if $\dfrac{dx}{dy} \neq 0$:

$$\frac{dy}{dx} = \frac{1}{\left(\frac{dx}{dy}\right)}.$$

Example 162. *Find the derivative of* $y = \ln\, x$*.*

$y = \ln\, x \Rightarrow x = e^y$*. Now (noting that* e^y *is never 0)*

$$\begin{aligned} \frac{dy}{dx} &= \frac{1}{\left(\frac{dx}{dy}\right)} \\ &= \frac{1}{e^y} \\ &= \frac{1}{x}. \\ \Rightarrow \frac{d(\ln\, x)}{dx} &= \frac{1}{x}. \end{aligned}$$

Important: Our goal is to find an explicit expression in terms of x.

Example 163. *Find the derivative of* $y = \arcsin x$

First note that the domain for the $\arcsin$ *function is* $[-1,1]$ *and the range is* $[-\frac{\pi}{2},\frac{\pi}{2}]$*.*

We have $x = \sin y$*. Therefore on* $(-1,1)$

$$\begin{aligned} \frac{dy}{dx} &= \frac{1}{\left(\frac{dx}{dy}\right)} \\ &= \frac{1}{\cos y}, \left(\textit{note } \cos(y) > 0 \textit{ for } y \in (-\frac{\pi}{2},\frac{\pi}{2})\right) \\ &= \frac{1}{\sqrt{1-\sin^2 y}} \\ &= \frac{1}{\sqrt{1-x^2}}. \end{aligned}$$

Therefore,

$$\frac{d}{dx}(\arcsin x) = \frac{1}{\sqrt{1-x^2}}, \quad x \in (-1,1).$$

Other useful derivatives

$$\begin{aligned}\frac{d}{dx}(\arccos x) &= -\frac{1}{\sqrt{1-x^2}} \quad \text{for } x \in (-1,1) \\ \frac{d}{dx}(\arctan x) &= \frac{1}{1+x^2} \quad \text{for } x \in \mathbb{R}\end{aligned}$$

L'Hôpital's rule

Suppose that f and g are differentiable and $g'(x) \neq 0$ near a (except possibly at a). Suppose that

$$\lim_{x\to a} f(x) = 0 \text{ and } \lim_{x\to a} g(x) = 0$$

or

$$\lim_{x\to a} f(x) = \pm\infty \text{ and } \lim_{x\to a} g(x) = \pm\infty.$$

Then

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

if the limit on the right exists.

Example 164. *Find* $\lim_{x\to 1} \frac{\ln\ x}{x-1}$.

Note that

$$\lim_{x\to 1} \ln\ x = \lim_{x\to 1}(x-1) = 0.$$

Hence, we cannot *use*

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to 0} f(x)}{\lim_{x\to 0} g(x)}.$$

Instead, we use L'Hôpital's rule to get:

$$\begin{aligned}\lim_{x\to 1} \frac{\ln\ x}{x-1} &= \lim_{x\to 1} \frac{\frac{1}{x}}{1}, \text{ if this limit exists,} \\ &= 1.\end{aligned}$$

Example 165. *Find* $\lim_{x\to 0^+} x \ln\ x$.

Rewrite $x \ln x$ *as* $\frac{\ln\ x}{1/x}$ *and note that in this case*

$$\lim_{x\to 0^+} \ln\ x = -\infty \quad \text{and} \quad \lim_{x\to 0^+} \frac{1}{x} = \infty.$$

Hence we use L'Hôpital's rule to give:

$$\begin{aligned}\lim_{x\to 0^+} \frac{\ln\ x}{1/x} &= \lim_{x\to 0^+} \frac{1/x}{-1/x^2} \\ &= \lim_{x\to 0^+} (-x) \\ &= 0.\end{aligned}$$

The mean value theorem (MVT)

Let f be continuous on $[a, b]$ and differentiable on (a, b). Then

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

for some c, where $a < c < b$.

Note $f'(c)$ is the slope of $y = f(x)$ at $x = c$ and $\frac{f(b)-f(a)}{b-a}$ is the slope of the chord joining $A = (a, f(a))$ to $B = (b, f(b))$

Example 166. *Without actually finding them, show that the equation*

$$x^4 + 4x + 1 = 0$$

has exactly two real solutions for x.

We can use the IVT to get an idea of how many solutions there may be.
When $x = -1, y = -2$ and when $x = 1, y = 6$. This tells us that there is at least one solution between $x = -1$ and 1.
When $x = -2, y = 9$ so again, we have at least one solution between $x = -2$ and -1.
Using the IVT we can see that there are at least two solutions.

The question is asking us to show that there are exactly *two real solutions, so we can now the MVT to do this. Suppose there are at least 3 distinct zeros, say $x_1 < x_2 < x_3$. By the MVT there must be a value c with $x_1 < c < x_2$ and $f'(c) = \frac{f(x_2)-f(x_1)}{x_2-x_1} = 0$.*

Similarly, there must be a value d with $x_2 < d < x_3$ and $f'(d) = \frac{f(x_3)-f(x_2)}{x_3-x_2} = 0$. Thus if f has at least 3 distinct zeros then f' has at least 2 distinct zeros.

But $f(x) = x^4 + 4x + 1 \Rightarrow f'(x) = 4x^3 + 4$. The only time $f' = 0$ is when $4x^3 + 4 = 0 \Rightarrow x^3 = -1 \Rightarrow x = -1$ only.

So there cannot be 3 or more real zeros of f.

7.3 Smoothness

Theorem: differentiability implies continuity

If f is differentiable at a, then f is continuous at a.

The converse, however, is false. That is, if f is continuous at a, then it is not necessarily differentiable. (See the following example.)

Example 167. *The function $f(x) = |x|$ is continuous, but not differentiable in $x = 0$, since*

$$\lim_{h \to 0^+} \frac{f(0 + h) - f(0)}{h} = \lim_{h \to 0^+} \frac{h - 0}{h} = 1$$

and

$$\lim_{h \to 0^-} \frac{f(0 + h) - f(0)}{h} = \lim_{h \to 0^-} \frac{-h - 0}{h} = -1 \neq 1,$$

meaning that $\displaystyle\lim_{h \to 0} \frac{f(0 + h) - f(0)}{h}$ *does not exist.*

Can we always find partial derivatives and tangent planes?

Example 168. *Simple cusp-like functions are not smooth:*

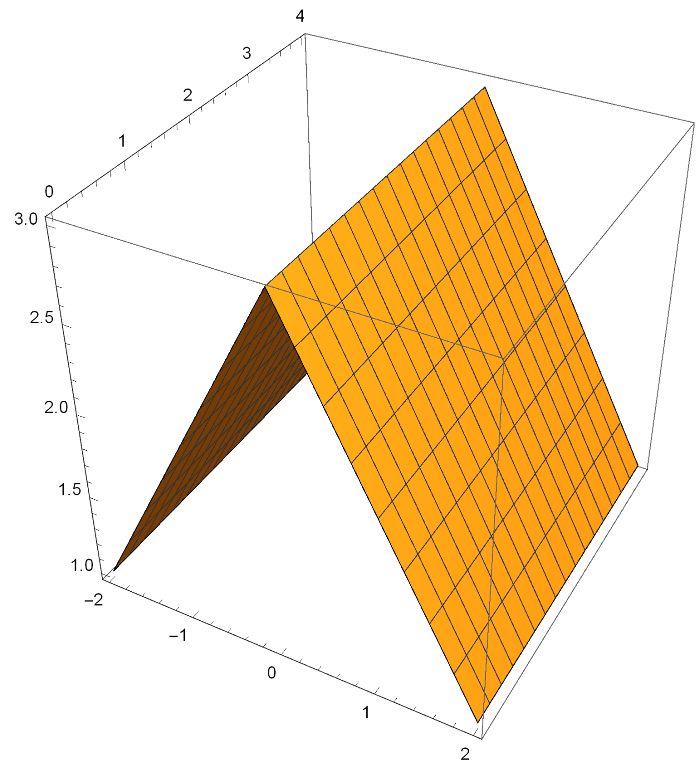


Figure 49: The function $z = f(x) = 3 - |x|$. $f_x(0)$ is undefined.

A surface $z = f(x, y)$ is smooth at (a, b) if f, f_x and f_y are all continuous at (a, b). When you zoom in close enough to a smooth surface it looks like a plane.

One way to see this is to look at the contours. The contours of a plane are straight parallel lines, the same perpendicular distance apart. As you zoom into a smooth surface the contours straighten out. This means that close to (a, b) the surface is approximated by a plane; in fact it can be approximated by the tangent plane. Do straight contours imply smoothness?

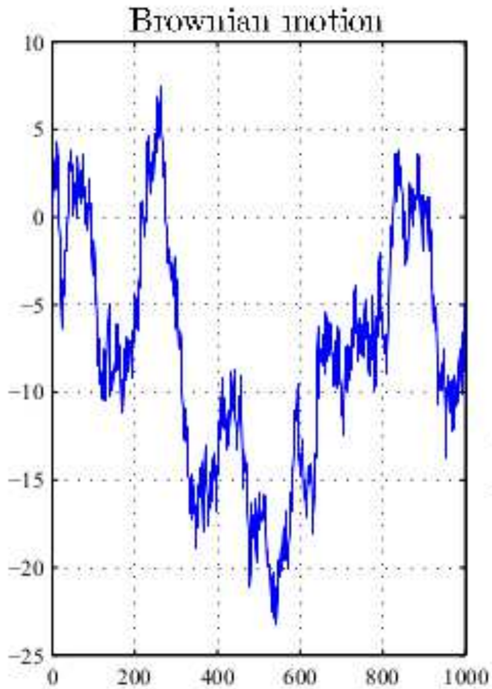
Example 169. *Brownian motion is not smooth. No matter how much you zoom in, it always looks rough, in fact, Brownian motion is a fractal. There are surface-analogues to Brownian motion, demonstrated with the fractal surface below.*

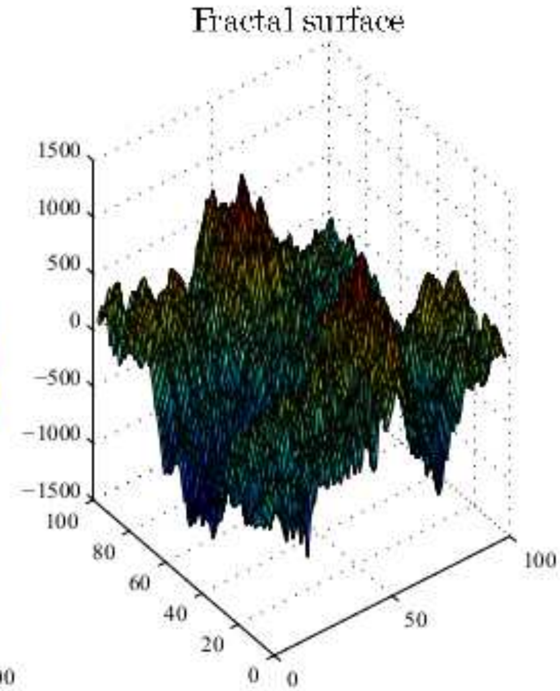
7.4 Optimization

Definition: increasing/decreasing

A function f is called *strictly increasing* on an interval I if

$f(x_1) < f(x_2)$ whenever $x_1 < x_2$





while f is called *strictly decreasing* on I if

$f(x_1) > f(x_2)$ whenever $x_1 < x_2$

This leads to the following test.

Increasing/decreasing test

Suppose that f is continuous on $[a, b]$ and differentiable on (a, b).

(a) If $f'(x) > 0$ on (a, b), then f is strictly increasing on $[a, b]$.

(b) If $f'(x) < 0$ on (a, b), then f is strictly decreasing on $[a, b]$.

(c) If $f'(x) = 0$ on (a, b), then f is constant on $[a, b]$.

Proof of (a):
Let $a \le d < e \le b$. It suffices to show $f(e) > f(d)$. Applying the MVT to f on $[d, e]$, there exists $c \in (d, e)$ with

$$\begin{aligned} \frac{f(e) - f(d)}{e - d} &= f'(c) \\ &> 0. \\ \Rightarrow f(e) - f(d) &> 0 \\ \text{i.e.} \quad f(e) &> f(d). \end{aligned}$$

The proofs of (b) and (c) are similar.

Example 170. *Test to see on what intervals the function $f(x) = x^3 + x$ is increasing or decreasing.*

$$\begin{aligned} &\textit{We have } f(x) = x^3 + x \\ &\Rightarrow f'(x) = 3x^2 + 1 > 0\,, \quad \textit{for all } x \in \mathbb{R} \\ &\textit{Therefore } f(x) \textit{ is strictly increasing on } \mathbb{R}. \end{aligned}$$

Definition 14 (local max/min)**.** *A function f has a* **local maximum** *at a if $f(a) \geq f(x)$ for all x in some open interval containing a (i.e. for x near a).*

Similarly, f has a **local minimum** *at b if $f(b) \leq f(x)$ for all x in some open interval containing b.*

Definition 15 (critical point)**.** *A function f is said to have a* **critical point** *at $x = a,\ a \in$ dom(f) if $f'(a) = 0$ or if $f'(a)$ does not exist.*

Definition 16 (global max/min)**.** *Let f be a function defined on the interval $[a, b]$, and $c \in [a, b]$.*
Then f has a global maximum *at c if*

$f(c) \geq f(x)$ *for all $x \in [a, b]$.*

Similarly f has a global minimum *at c if*

$f(c) \leq f(x)$ *for all $x \in [a, b]$.*

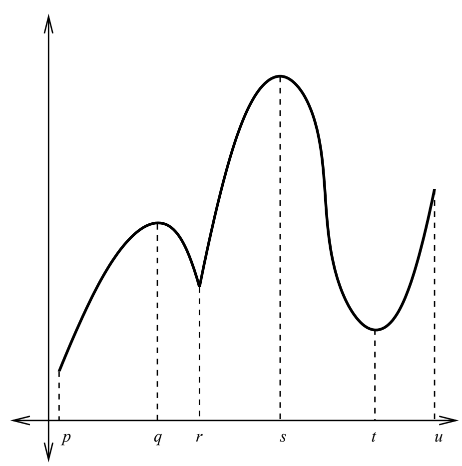


Figure 50: What do the points $x = p, q, r, s, t, u$ represent?

Example 171. *Consider the function in Figure 50. The domain of this function is the interval $[p, u]$. Which points represent the critical points? Which points represent the maximum and minimum? Which points represent local maxima and minima?*

The global maximum is at point s. The global minimum is at point p. Local maxima are at points q and s. Local minima are at points r and t.

Theorem 25. *If f has a local maximum/minimum at $x = c$ and $f'(c)$ exists then $f'(c) = 0$.*

Proof (For max.) For $x < c,\ x$ near c:

$$\begin{aligned} f(x) - f(c) &\le 0 \\ \Rightarrow \frac{f(x)-f(c)}{x-c} &\ge 0 \\ \Rightarrow f'(c) = \lim_{x\to c^-} \frac{f(x)-f(c)}{x-c} &\ge 0 \end{aligned}$$

Similarly for $x > c,\ x$ near c:

$$\begin{aligned} f(x) - f(c) &\le 0 \\ \Rightarrow \frac{f(x)-f(c)}{x-c} &\le 0 \\ \Rightarrow f'(c) = \lim_{x\to c^-} \frac{f(x)-f(c)}{x-c} &\le 0 \end{aligned}$$

Thus $f'(c) = 0$.

First derivative test

Let f be continuous on $[a, b]$ and differentiable on (a, b), and let $c \in (a, b)$.

(a) If $f'(x) > 0$ for $a < x < c$ and $f'(x) < 0$ for $c < x < b$

then f has a local maximum at c.

(b) If $f'(x) < 0$ for $a < x < c$ and $f'(x) > 0$ for $c < x < b$

then f has a local minimum at c.

Proof:

(a) Since $f'(x) > 0$ for $a < x < c$ and f continuous on $(a, c]$, f is strictly increasing on $(a, c]$ (by increasing/decreasing test). Thus $f(c) > f(x)$ for $x < c$.

Similarly (by increasing/decreasing test), $f(c) > f(x)$ for $c < x < b$.

Thus $f(c) \ge f(x)$ for all $x \in (a, b)$ and f has a maximum at $x = c$.

(b) Analogous.

Second derivative test

Suppose f'' exists and is continuous near c.

(a) If $f'(c) = 0$ and $f''(c) > 0$, then

f has a local minimum at c.

(b) If $f'(c) = 0$ and $f''(c) < 0$, then

f has a local maximum at c.

(c) If $f'(c) = 0$ and $f''(c) = 0$, then the test fails and we get no useful information. In this case the first derivative test should be used.

Example 172. *Find local minima and maxima of* $f(x) = 2x^3 + 3x^2 - 12x + 4$ *on* $\mathbb{R}$.

Using the second derivative test:

$$\begin{aligned} f'(x) &= 6x^2 + 6x - 12 \textit{ and } f''(x) = 12x + 6 \\ f'(x) &= 0 \Rightarrow 6x^2 + 6x - 12 = 0 \Rightarrow x^2 + x - 2 = 0 \Rightarrow x = -2 \textit{ or } 1 \\ f''(-2) & \textit{ is negative} \Rightarrow (-2, 24) \textit{ is a local maximum} \\ f''(1) & \textit{ is positive} \Rightarrow (1, -3) \textit{ is a local minimum} \end{aligned}$$

Theorem 26 (Extreme value theorem)**.** *If* f *is continuous on a closed interval* $[a, b]$*, then* f *attains its global maximum and a global minimum value on* $[a, b]$*. These values occur either at an end point or an interior critical point.*

Calculate the value of f *at each of these points and then work out the largest value (global maximum) and the smallest value (global minimum).*

Bonus theorem: *a) If* f *is differentiable on* (a, b) *and only has one critical point, which is a local min; then the global minimum of* f *occurs at this point.*
b) If f *is differentiable on* (a, b) *and only has one critical point, which is a local max; then the global maximum of* f *occurs at this point.*

Example 173. *Find the global maximum and minimum of the function* $f(x) = x^3 - 3x^2 + 1$ *on the interval* $[-\frac{1}{2}, 4]$.

Note that

$$f'(x) = 3x^2 - 6x = 3x(x - 2)$$

is defined everywhere and $f'(x) = 0$ *at* $x = 0$ *or 2. Both critical points therefore lie in* $(-\frac{1}{2}, 4)$.

Therefore the global maximum and minimum are either at the end points ($x = -\frac{1}{2}$ *or 4) or at the interior critical points (*$x = 0$ *or 2):*

$$\begin{aligned} f(0) &= 1 \\ f(2) &= -3 \\ f\left(-\frac{1}{2}\right) &= \frac{1}{8} \\ f(4) &= 17 \end{aligned}$$

Therefore the global maximum is 17 at $x = 4$ *and the global minimum is* -3 *at* $x = 2$.

Example 174. *Consider the following data set:*

$$\begin{aligned} S = \{ & 2, 2.24, 2.23, 1.94, 1.36, 0.56, -0.35, -1.25, -2.00, \\ & -2.49, -2.65, -2.46, -1.98, -1.30, -0.52, 0.23, 0.86, 1.31, \\ & 1.55, 1.61, 1.53, 1.38, 1.20, 1.01, 0.84, 0.64, 0.40, 0.07, \\ & -0.34, -0.84, -1.36, -1.83, -2.17, -2.27, -2.10, -1.63, -0.91, \\ & -0.03, 0.89, 1.71, 2.32, 2.62, 2.58, 2.21, 1.60, 0.84, 0.06, \\ & -0.62, -1.15, -1.48, -1.61, -1.58, -1.45, -1.27, -1.09, -0.91, \\ & -0.73, -0.51, -0.22, 0.15, 0.62, 1.15, 1.65, 2.05\}, \end{aligned}$$

obtained from the price of corn over a given period of time. One of several methods for fitting a function $f(x)$ to a data set is known as Fourier interpolation. To do this, we apply a transformation to the data to obtain a list of complex numbers $a+bi$, and then apply a formula to recover $f(x)$. In Mathematica *this is done as follows.*

```
m = Length[S];
A = ListPlot[S];
L = Sqrt[m]*Conjugate[Fourier[S]];
F = Table[
     If[(Abs[L[[j]]] < 1) || (Abs[L[[j]]] > 10000), 0, L[[j]]],
    {j, 1, m}];
g1 = ((2/m)*Cos[2*Pi*k*(x - 1)/m])*Re[F[[k + 1]]];
g2 = ((2/m)*Sin[2*Pi*k*(x - 1)/m])*Im[F[[k + 1]]];
f = Sum[g1 - g2, {k, 0, Floor[m/2] - 1}]
B = Plot[G, {x, 1, m}];
```

We obtain the function

$$
\begin{aligned}
f(x) \quad = \quad & 0.073\sin\left(\frac{1}{8}\pi(x-1)\right) - 0.003\sin\left(\frac{1}{16}\pi(x-1)\right) - 0.064\sin\left(\frac{3}{16}\pi(x-1)\right) \\
& +0.0003\sin\left(\frac{1}{32}\pi(x-1)\right) - 0.32171\sin\left(\frac{3}{32}\pi(x-1)\right) \\
& +0.667\sin\left(\frac{5}{32}\pi(x-1)\right) - 0.029\sin\left(\frac{7}{32}\pi(x-1)\right) \\
& -0.086\cos\left(\frac{1}{8}\pi(x-1)\right) + 0.128\cos\left(\frac{1}{16}\pi(x-1)\right) \\
& -0.048\cos\left(\frac{3}{16}\pi(x-1)\right) + 0.082\cos\left(\frac{1}{32}\pi(x-1)\right) \\
& +1.983\cos\left(\frac{3}{32}\pi(x-1)\right) + 0.153\cos\left(\frac{5}{32}\pi(x-1)\right) \\
& -0.029\cos\left(\frac{7}{32}\pi(x-1)\right) + 0.074.
\end{aligned}
$$

Find the

- *local maximum of the function $f(x)$,*
- *global maximum of the function $f(x)$,*
- *where $f(x)$ is increasing most rapidly,*
- *where $f(x)$ is decreasing most rapidly.*

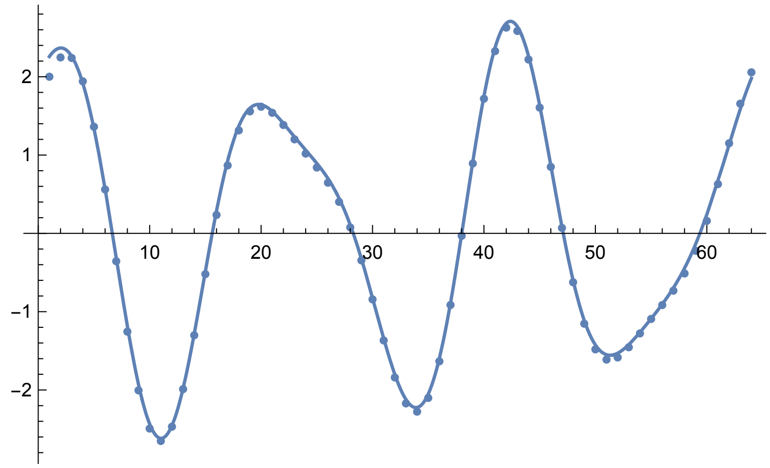


Figure 51: A Fourier fit to the data S.

7.5 Basic Ordinary Differential Equations

In almost all cases, quantifying time-dependent occurrences almost inevitably comes down to a **differential equation** (DE). A DE is an equation which contains one or more derivatives of an unknown function. There are two types of DEs:

- **ordinary** differential equations (ODEs), where the unknown function is a function of only one variable, and
- **partial** differential equations (PDEs) where the unknown function is a function of more than one variable.

How do ODEs arise?

Population modelling. Assume a population grows at a constant rate proportional to the size of the population. If $P = P(t)$ stands for the population at time t, then the model states that

$$\frac{dP}{dt} \propto P,$$

i.e.,

$$\frac{dP}{dt} = kP,$$

where k is the growth constant. If $k > 0$ the population is growing (think humans) and if $k < 0$ the population is decreasing (think tigers).

Question 1. *What happens for $k = 0$?*

Solution to an ODE

Suppose that we are given an ODE for y which is an unknown function of x. Then $y = f(x)$ is said to be a solution to the ODE if the ODE is satisfied when $y = f(x)$ and its derivatives are substituted into the equation.

Example 175. *Show that $y = y(x) = A\exp(x^2/2)$ is a* solution *to the ODE $y' = xy$.*

When asked to solve an ODE, you are expected to find *all* possible solutions. This means that you need to find the general solution to the ODE. For an ODE that involves only the unknown function y and its first derivative, the general solution will involve *one* arbitrary constant.

You should already know how to solve ODEs of the form

$$\frac{dy}{dx} = f(x) \quad \text{or} \quad \frac{d^2y}{dx^2} = g(x).$$

Example 176. *Find the general solution to the differential equation $y' = x^2$.*

The order of an ODE

The order of the highest-order derivative in an ODE defines the order of the ODE.

Example 177. *Now suppose you are throwing apples over your fence to your neighbour. Find an expression for the position $(x(t), y(t))$ of the apple if you assume the initial position is $x(0) = 0, y(0) = 0$ and the initial velocity is $\dot{x}(0) = u$, $\dot{y}(0) = v$.*

Let the firing device be positioned at the origin. So $x(0) = 0$ and $y(0) = 0$. We have two equations of motion;

$$\frac{d^2y}{dt^2} = -g \quad \text{and} \quad \frac{d^2x}{dt^2} = 0.$$

Suppose further that the initial horizontal velocity is $u_0 = \frac{dx}{dt}(0)$ and vertically $v_0 = \frac{dy}{dt}(0)$. Now solving $\frac{d^2y}{dt^2} = -g$ gives $y = -\dfrac{1}{2}gt^2 + v_0 t$ as before.

To solve $\dfrac{d^2x}{dt^2} = 0$ we let $u = \dfrac{dx}{dt}$

$$\Rightarrow \frac{du}{dt} = 0 \;\Rightarrow\; u = u_0 \Rightarrow \frac{dx}{dt} = u_0$$
$$\Rightarrow \int dx = \int u_0 dt \Rightarrow x = u_0 t + c \text{ but } c = 0 \text{ because } x_0 = 0.$$

Therefore $y = -\frac{1}{2}gt^2 + v_0 t$ and $x = u_0 t$.

Analytical and numerical solutions

To solve an ODE (or IVP) analytically means to give a solution curve in terms of continuous functions defined over a specified interval, where the solution is obtained exactly by analytic means (e.g., by integration). The solution satisfies the ODE (and initial conditions) on direct substitution.

To solve an ODE (or IVP) numerically means to use an algorithm to generate a sequence of points which approximates a solution curve.

As we have already seen, the ODE

$$\frac{dy}{dt} = f(t)$$

can be solved analytically. The solution simply is

$$y = \int f(t)dt.$$

Now this may seem like a cop-out because all we are saying is that the solution is given by the anti-derivative of $f(t)$, and the above solution is as informative as the actual ODE. In practice one hopes to be able perform the above integral to get a more explicit form of the solution. Depending on the form of f this may either be done exactly or by numerical means.

Example 178. *Admittedly this is not the best example of exponential growth. However it gives us an opportunity to obtain real data and practice in collecting and analysing it.*

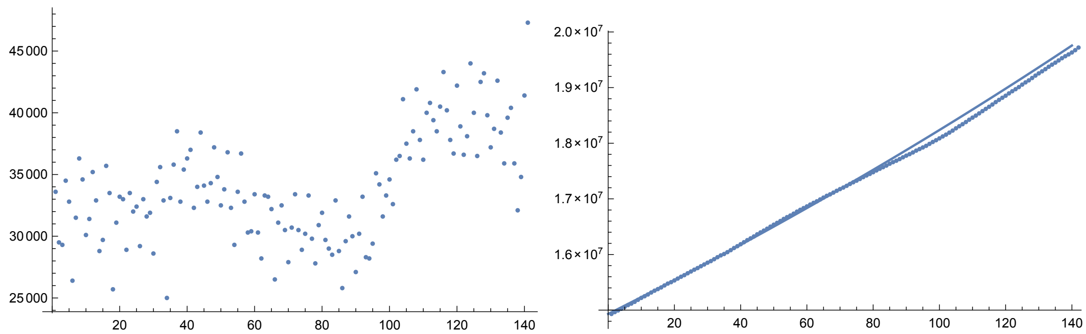


Figure 52: Left: Births minus deaths in Australia from 1981 to 2016. Right: Actual values of $P(t)$ and the model $P(t) = 14931234e^{0.002t}$.

Population is known to satisfy a differential equation. The simplest population model is given by

$$\frac{dP}{dt} = kP,$$

as we have already seen. We begin by visiting the Australian Bureau of Statistics website: `http://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/3101.0` *Next we save the desired population data as population.csv and place in the documents folder. The data displays births, deaths, immigration, and emigration in Australia from June 1981 to June 2016, among other observations. We wish to determine the difference in births and deaths at each moment in which data was collected. This can be done as follows in Mathematica. A description of these calculations are as follows.*

1. *Import the file "population.csv";*

2. *Remove the first* 10 *rows of the data since these contain descriptions of the data which we do not want;*

3. *List the rows we are interested in, births and deaths rows 2 and 3, transpose the corresponding matrix and multiply by 1000 since the original data was scaled down;*

4. *Input the initial population of Australia, from June, 1981 as* $a = 4931234$*;*

5. *Obtain a list representing births minus deaths;*

6. *Set the initial population as* $p = a$*;*

7. *Construct a list* W *consisting of the population at a given quarter plus births minus deaths;*

8. *Construct a list* Y *consisting of* W *with* a *in front as the first entry;*

9. *Plot the births minus the deaths;*

10. *Plot the population at a given quarter;*

11. *We know that* $P(t) = ae^{rt}$*. Choose an appropriate value for* $r = 0.002$*;*

12. *Plot the function* $P(t) = ae^{rt}$ *with the data* Y*.*

```
T = Import["population.csv"];
L = Drop[T, 10];
S = 1000 Table[Transpose[L][[j]], {j, 2, 3}];
a = 14931234;
V = S[[1]] - S[[2]]
p = a;
W = Table[p = p + V[[j]], {j, 1, Length[V]}];
Y = Prepend[W, a]
ListPlot[V]
M = ListPlot[Y]
r = 0.002;
H = Plot[a Exp[r x], {x, 0, 140}];
Show[H, M]
```

We see from Figure 52 that the population growth in Australia due to nature is nearly linear. This will be useful in the lesson on Taylor polynomials. Bacterial growth will be quite visibly exponential in growth.

7.6 Application

8 Linear approximations and Taylor series

8.1 Linear approximation for $f(x)$

The approximation

$$f(x) \approx f(a) + f'(a)(x-a) \tag{5}$$

can be a reasonable (linear) approximation of a function $f(x)$ provided it is not too curved.

Recall Example 178 on population growth in Australia. The growth is almost linear. We derived the population model

$$P(t) = 14931234e^{0.002t}.$$

A linear approximation of $P(t)$ at $t = 70$ quarter years, using (5), is given by

$$P(t) \approx P(70) + P'(70)(t-70).$$

To simplify this expression we must first calculate

$$P'(t) = 14931234 \times 0.002e^{0.002t} = 29862.468e^{0.002t}.$$

We also need to evaluate $P(70) = 17175007.255$ and $P'(70) = 34350.014$. This gives the linear approximation

$$\begin{aligned} P(t) &\approx 17175007.255 + 34350.014(t-70), \\ &\approx 14770506.275 + 34350.014t. \end{aligned}$$

8.2 Taylor series

Given a function $f(x)$ it is sometimes possible to expand $f(x)$ as a power series

$$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \ldots$$

To find the coefficients c_n in general, let us assume that f and all its derivatives are defined at $x = a$.

First note that

$$f(a) = c_0.$$

Now take the derivative of both sides. Assuming that we are justified in differentiating term by term on the right-hand side, this gives

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \ldots$$

so that

$$f'(a) = c_1.$$

Taking the derivative again and again gives

$$\begin{aligned} f''(a) = 2c_2 \quad &\Rightarrow \quad c_2 = f''(a)/2 \\ f'''(a) = 6c_3 \quad &\Rightarrow \quad c_3 = f'''(a)/3! \end{aligned}$$

and, in general

$$c_n = f^{(n)}(a)/n!, \text{ with } f^{(0)}(x) \Leftrightarrow f(x).$$

This leads to

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

The right-hand side is called the *Taylor series* of $f(x)$ about $x = a$. So this result says that *if* $f(x)$ has a power series representation (expansion) at $x = a$, it must be given by the formula above.

Example 179 (Maclaurin series)**.** *When $a = 0$ in the Taylor series we call the series a* Maclaurin series*.*

Note: The function f and all its derivatives must be defined at $x = a$ for the series to exist. In this case we say that f is infinitely differentiable at $x = a$.

Theorem 27 (Taylor series)**.** *Suppose $f(x)$ is infinitely differentiable at $x = a$. Then for f "sufficiently well behaved",*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \textit{ for } |x-a| < r,$$

where r is the radius of convergence.

Remark 2. *Note that if we define a sequence of numbers*

$$b_n = \max_{x \in [-r,r]} |f^{(n)}(x)|, \ n \in \mathbb{N},$$

then the "sufficiently well behaved" means, among other things,

$$\lim_{n \to \infty} \frac{b_n r^n}{n!} = 0.$$

Unless otherwise noted, all functions we consider in this section are "sufficiently well behaved".

For $|x - a| < r$, we have approximately

$$f(x) \simeq \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

which is a useful polynomial approximation. The smaller $|x - a|$, the better the approximation. This approximation can be made accurate to arbitrary order by taking k sufficiently large provided $|x - a| < r$. This begins to address the question of whether or not $f(x)$ has a power series representation.

Example 180. *Find the Taylor series of $f(x) = \ln x$ about $x = 1$. Determine its radius of convergence.*

$$\begin{aligned}
f^{(0)}(x) &= f(x) = \ln x \\
f'(x) &= \frac{1}{x} = x^{-1},\ f''(x) = -x^{-2},\ f'''(x) = 2x^{-3},\ \textit{and so on} \\
n \geq 2:\ f^{(n)}(x) &= (-1)^{n-1}(n-1)!x^{-n} \\
&\Rightarrow\ f^{(0)}(1) = 0,\ f^{(1)}(1) = 1,\ f^{(n)}(1) = (-1)^{n-1}(n-1)!
\end{aligned}$$

So the Taylor series of f about 1 is

$$\sum_{n=0}^{\infty} c_n (x-1)^n,\ \textit{where } c_n = \frac{f^n(1)}{n!}.$$

$$\begin{aligned}
\Rightarrow\ c_0 = 0,\ c_n &= \frac{(-1)^{n-1}(n-1)!}{n!} \quad n \geq 1 \\
&= \frac{(-1)^{n-1}}{n}
\end{aligned}$$

So the Taylor series of $\ln x$ *is*

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(x-1)^n \tag{6}$$

To find the radius of convergence, use the Ratio Test.

$$\begin{aligned}
L &= \lim_{n\to\infty} \left| \frac{c_{n+1}(x-1)^{n+1}}{c_n(x-1)^n} \right| = |x-1| \lim_{n\to\infty} \frac{1}{n+1} \cdot \frac{n}{1}, \\
&= |x-1| \lim_{n\to\infty} \frac{n}{n+1} = |x-1|
\end{aligned}$$

By the Ratio Test, the series converges for $|x-1| < 1$.

Remark 3.

1. *This holds for $0 < x < 2$ by what we have shown: in fact it is also okay for $x = 2$.*
2. *It can be shown that* (6) *does in fact converge to* $\ln(x)$*, i.e.*

$$\ln x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(x-1)^n \quad \textit{for } 0 < x \leq 2$$

3. *Plugging in $x = 2 \Rightarrow \ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$ (alternating harmonic series).*

Example 181. *Find the Taylor series of $f(x) = e^x$ about $x = a$. Determine its radius of convergence.*

$$\begin{aligned}
f(x) &\Rightarrow\ f^{(n)}(x) = e^x \textit{ for all } n \geq 0 \\
&\Rightarrow\ f^{(n)}(a) = e^a \textit{ for all } n \geq 0 \\
&\Rightarrow\ \textit{the Taylor series of } f \textit{ about } x = a \textit{ is}
\end{aligned}$$

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n &= \sum_{n=0}^{\infty}\frac{e^a}{n!}(x-a)^n \\ &= \sum_{n=0}^{\infty}c_n(x-a)^n\end{aligned}$$

Radius of convergence: Use the Ratio Test

$$\begin{aligned}L &= \lim_{n\to\infty}\left|\frac{c_{n+1}(x-a)^{n+1}}{c_n(x-a)^n}\right| = \lim_{n\to\infty}\left|\frac{e^a(x-a)^{n+1}}{(n+1)!}\cdot\frac{n!}{e^a(x-a)^n}\right| \\ &= |x-a|\lim_{n\to\infty}\frac{1}{n+1} = |x-a|\cdot 0 = 0\end{aligned} \tag{7}$$

Since $L<1$ (no matter what x is), by the Ratio Test the series converges for all x. The radius of convergence is ∞.

Note:

1. *Here, too, the Taylor series actually converges to the function, i.e.* $e^x = \sum_{n=0}^{\infty}\frac{e^a}{n!}(x-a)^n$.
2. *Taking $a=0$ gives the Maclaurin series:* $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$

Note: In this case coefficients in power series are given by

$$c_n = \frac{e^a}{n!}$$

so

$$\left|\frac{c_n}{c_{n+1}}\right| = \left|\frac{(n+1)!}{n!}\right| = n+1 \to \infty \textit{ as } n\to\infty$$

Example 182. *Find the Maclaurin series for $f(x)=\sin x$. Determine its radius of convergence.*

$$\begin{aligned} f(x) &= \sin x \;\Rightarrow\; f'(x)=\cos x,\; f''(x)=-\sin x,\; f'''(x)=-\cos x \\ f^{(4)}(x) &= \sin x \textit{ etc.} \\ \Rightarrow\quad f(0) &= 0,\; f'(0)=1,\; f''(0)=0,\; f'''(0)=-1,\; f^{(4)}(0)=(0) \textit{ etc.} \\ \textit{So } f^{(n)}(0) &= \begin{cases} 0 & n=2m \\ (-1)^m & n=2m+1\end{cases}\end{aligned}$$

This shows that the Maclaurin series for $\sin x$ *is:*

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n &= x-\frac{x^3}{3!}+\frac{x^5}{5!}-\dots.\\ &= \sum_{m=0}^{\infty}(-1)^m\frac{x^{2m+1}}{(2m+1)!} \qquad (8)\end{aligned}$$

To determine the radius of convergence we use the ratio test on (8)*. So*

$$\begin{aligned} L &= \lim_{m\to\infty}\left|\frac{a_{m+1}}{a_m}\right| \\ &= \lim_{m\to\infty}\left|\frac{(-1)^{m+1}x^{2m+3}}{(2m+3)!}\cdot\frac{(2m+1)!}{(-1)^m x^{2m+1}}\right| \\ &= \lim_{m\to\infty}\left|\frac{x^2}{(2m+3)(2m+2)}\right| \\ &= 0<1 \Rightarrow \text{ series converges for all } x \text{ (i.e. } r=\infty). \end{aligned}$$

Remark 4.

1. *The Maclaurin series converges to* $\sin x$.
2. *The fith-order Taylor polynomial is* $T_5(x)=x-\frac{x^3}{3!}+\frac{x^5}{5!}$*: Suppose now* $0\le x\le\frac{\pi}{4}$*: the the remainder* R_5 *satisfies*

$$\begin{aligned}|f(x)-R_5(x)|=|T_5(x)| &= \frac{|f^{(6)}(z)|}{6!}x^6, \text{ for some } z,\ 0<z<x.\\ &= \frac{|\sin(z)|}{6!}x^6 \le \frac{1}{\sqrt{2}}\cdot\frac{1}{6!}\left(\frac{\pi}{4}\right)^6 = 0.0023\end{aligned}$$

NOTE: In fact $T_5=T_6$ *for this function, so we can improve the error estimate by using* R_6 *instead of* R_5 *to* 0.000026.

8.3 Linear approximations for $f(x,y)$

The corresponding linear or first-order approximation for a function f of two variables, near a known point (a,b) is the tangent plane. The linear approximation (i.e., tangent plane) is accurate close to the known point, provided f is smooth.

The linear approximation to f at (a,b) is

$$f(x,y)\simeq f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b).$$

Example 183. *The temperature in a region is given by* $T(x,y)=100-x^2-y^2$*. Find the linear approximation to* $T(x,y)$ *near* $(0,5)$*.*

The linear approximation is the tangent plane at $(0, 5)$. *Note* $T(0, 5) = 75$*; and*

$$\frac{\partial T}{\partial x} = -2x \Rightarrow \frac{\partial T}{\partial x}(0, 5) = 0; \quad \frac{\partial T}{\partial y} = -2y \Rightarrow \frac{\partial T}{\partial y}(0, 5) = -10.$$

So $T(x,\, y) \simeq 75 - 0(x - 0) - 10(y - 5) = 75 - 10(y - 5) = 125 - 10y,$ *for* $(x, y) \approx (0, 5)$.

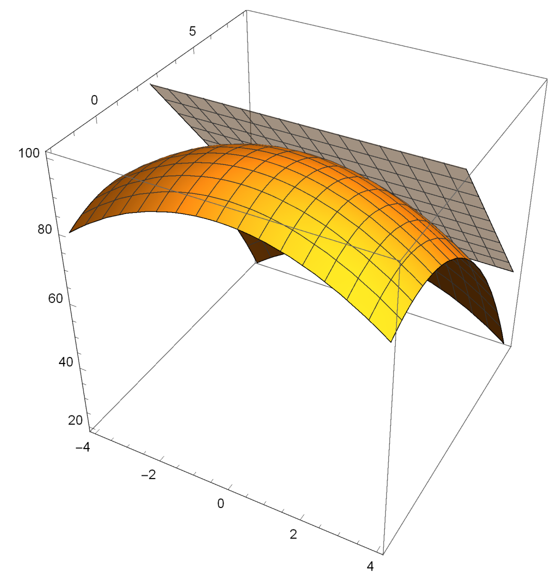


Figure 53: The function $T(x, y) = 100 - x^2 - y^2$ and the tangent plane $T(x, y) = 125 - 10y$.

Example 184. *Find the tangent plane to* $z = e^{-x^2} \sin y$ *at* $(1,\, \frac{\pi}{2})$ *and use it to find an approximate value for* $e^{-(0.9)^2} \sin(1.5)$.

Let $f(x, y) = e^{-x^2} \sin y$. *Then* $f\left(1, \frac{\pi}{2}\right) = e^{-1}.1 = \frac{1}{e}$. *So*

$$\begin{array}{l} \frac{\partial f}{\partial x} = -2x\, e^{-x^2} \sin y, \text{ which implies } \frac{\partial f}{\partial x}(1,\, \frac{\pi}{2}) = -2e^{-1}.1 = \frac{-2}{e}, \\ \text{and } \frac{\partial f}{\partial y} = e^{-x^2} \cos y, \text{ so } \frac{\partial f}{\partial y}(1,\, \frac{\pi}{2}) = 0. \end{array}$$

So the tangent plane at $(1,\, \frac{\pi}{2})$ *is*

$$z = \frac{1}{e} - \frac{2}{e}(x - 1).$$

We can use this to approximate $e^{-x^2} \sin y$ *for* $x \simeq 1$ *and* $y \simeq \frac{\pi}{2}$.

Here $x = 0.9$ *and* $y = 1.5$ *so*

$$\begin{aligned} e^{(0.9)^2} \sin(1.5) &\simeq \frac{1}{e} - \frac{2}{e}(0.9 - 1) \\ &\simeq \frac{1}{e} + \frac{0.2}{e} = \frac{1.2}{e}. \end{aligned}$$

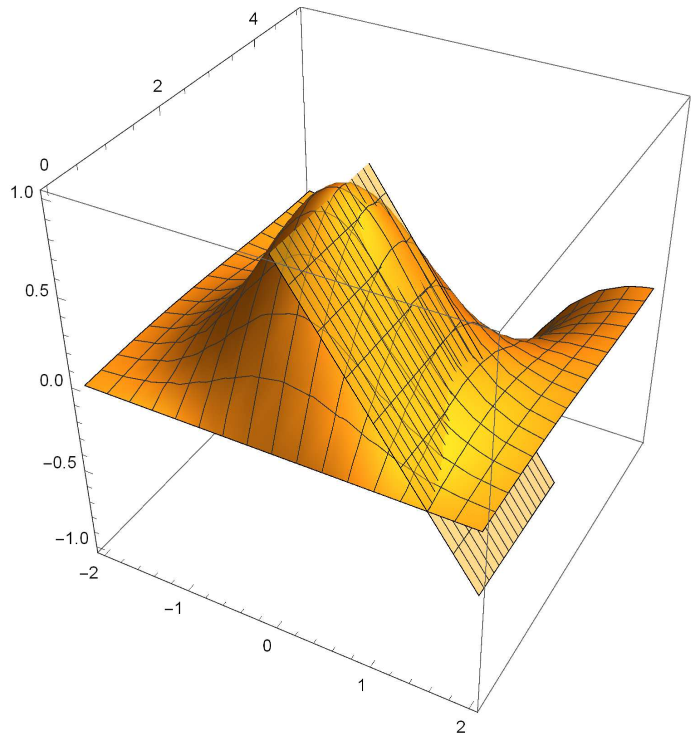


Figure 54: The function $z = e^{-x^2}\sin(y)$ and the tangent plane $z = \frac{1}{e}(3 - 2x)$.

Estimating small changes

We may use the equation for the tangent plane to infer that

$$\Delta z \approx f_x(a, b)\Delta x + f_y(a, b)\Delta y,$$

where $\Delta x = x - a$, $\Delta y = y - b$ and $\Delta z = z - f(a, b)$ represent a small change in x, y and z respectively. This is useful for estimating small changes in z arising from small changes in x and y.

Example 185. *Electric power is given by $P(E,\, R) = E^2/R$ where E is the voltage and R is the resistance. Find a linear approximation for $P(E, R)$ if $E \simeq 200$ (in Volts) and $R \simeq 400$ (in Ohms). Use this to find the effect that a change in E and R has on P.*

$$\frac{\partial P}{\partial E} = \frac{2E}{R} \Rightarrow \frac{\partial P}{\partial E}(200,\, 400) = 1; \quad \frac{\partial P}{\partial R} = -\frac{E^2}{R^2} \Rightarrow \frac{\partial P}{\partial R}(200,\, 400) = \frac{-1}{4}.$$

So $P(E,\, R) \simeq 100 + (E - 200) + \frac{-1}{4}(R - 400)$ for $(E,\, R) \approx (200,\, 400)$.
Let ΔE be the change in E; then $E = 200 + \Delta E$.

Similarly let ΔR be the change in R; then $R = 400 + \Delta R$.

So the change in P, $\Delta P = P - 100$, is given approximately by

$$\Delta P \simeq \Delta E + \frac{-1}{4}\Delta R$$

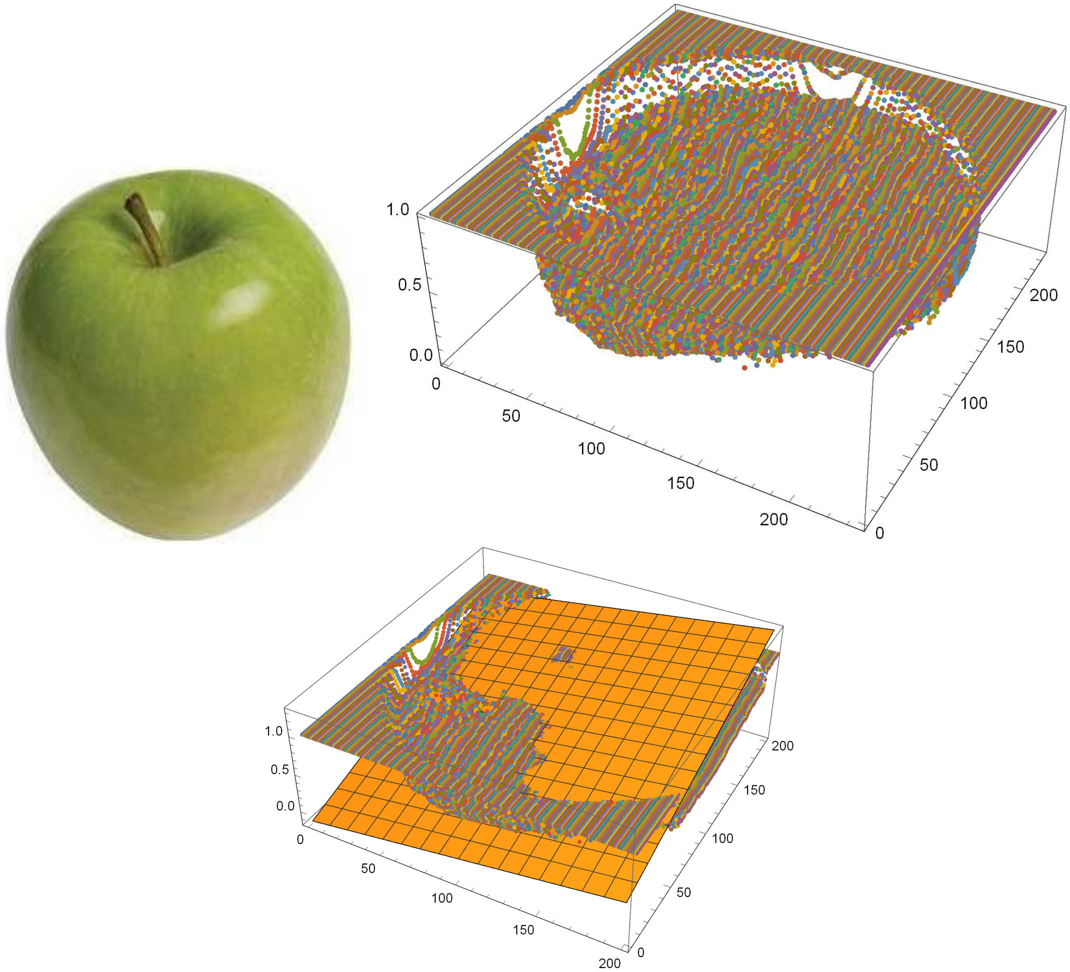


Figure 55: An image of an apple and the pixel values as a function of position in the image, $P(x, y)$, with a tangent plane.

Example 186. *We have seen how images contain data. Figure 55 shows points which represent the quantity $P(x, y)$ of RED in the RGB image of the apple shown on the left. By taking an approximation of the slopes P_x and P_y at the point $(x, y) = (85, 155)$ using a least squares method, we are able to compute an approximate tangent plane at $(85, 155)$. This was done as follows using Mathematica.*

```
L = Import["apple.jpg"]
M = ImageData[L];
n = Length[M];
T = Table[M[[i, j, 1]], {i, 1, n}, {j, 1, n}];
S = Table[{x, y, T[[x, y]]}, {x, 1, n}, {y, 1, n}];
Z = ListPointPlot3D[S]
(* Find a linear approximation of the surface at the point
  (100, 150) *)
(* Look at points near (a,b) = (75, 150) to take an average *)
{a, b} = {85, 155};
W = ListPointPlot3D[{{a, b, T[[a, b]]}},
```

```
    PlotStyle -> PointSize[0.05]];
p = 1/7^2 Sum[S[[a + i, b + j, 3]], {i, -3, 3}, {j, -3, 3}]
(* Regression line of best fit along line where y = b is
  fixed, and along line x = a is fixed, approx of f_x(a, b)
  and f_y(a, b) *)
q = Coefficient[
   Fit[Table[{S[[a, b, 2]], S[[a + j, b, 3]]}, {j, -3, 3}] ,
   {1, x}, x], x];
r = Coefficient[
   Fit[Table[{S[[a, b, 1]], S[[a, b + j, 3]]}, {j, -3, 3}] ,
   {1, y}, y], y];
Q = Plot3D[p + q (x - a) + r (y - b), {x, 0, 200},
    {y, 0, 200}];
V = Show[Q, Z]
Export["apple.eps", L]
Export["appledata.eps", V]
```

Estimating error

If the error in x is at most E_1 and in y is at most E_2, then a reasonable estimate of the worst-case error in the linear approximation of f at (a, b) is

$$|E| \approx |f_x(a,\, b)E_1| + |f_y(a,\, b)E_2|.$$

Example 187. *Suppose when making up a metal barrel of base radius 1 (in metres) and height 2 (in metres), you allow for an error of 5% in radius and height. Estimate the worst-case error in volume.*

The volume of the barrel is $V(r, h) = \pi r^2 h$. *Now* $V_r = 2\pi r h$ *and* $V_h = \pi r^2$, *and at* $(r, h) = (1, 2)$,

$$\begin{aligned} V_r(1,2) &= 4\pi \\ V_h(1,2) &= \pi \end{aligned}$$

So $\Delta V = 4\pi\Delta r + \pi\Delta h$.

Now if the error in radius is 5% then

$$\frac{\Delta r}{r} = 0.05 \Rightarrow \Delta r = 0.05, \text{ since } r = 1.$$

If the error in height is 5%,

$$\frac{\Delta h}{h} = 0.05 \Rightarrow \Delta h = 0.1, \text{ since } h = 2.$$

So $\Delta V = \pi(4.(0.05) + 0.1) = 0.3\pi$ *or the % error in volume*

$$\frac{\Delta V}{V} = \frac{0.3\pi}{(1)(2)\pi} = 0.15 \Rightarrow 15\%.$$

Extra reading: differentials

If we reduce our small changes Δx and Δy down to infinitesimal changes dx and dy, we may rewrite the linear approximation (tangent plane) as

$$dz = df = f_x(a,b)\,dx + f_y(a,b)\,dy.$$

This infinitesimal change dz is called the **total differential**. Note that this equation is not just an approximation.

8.4 Application

9 Integration

9.1 Area under a curve

Consider the problem of finding the area under a curve given by the graph of a continuous, non-negative function $y = f(x)$ between two points $x = a$ and $x = b$; see Figure 56.

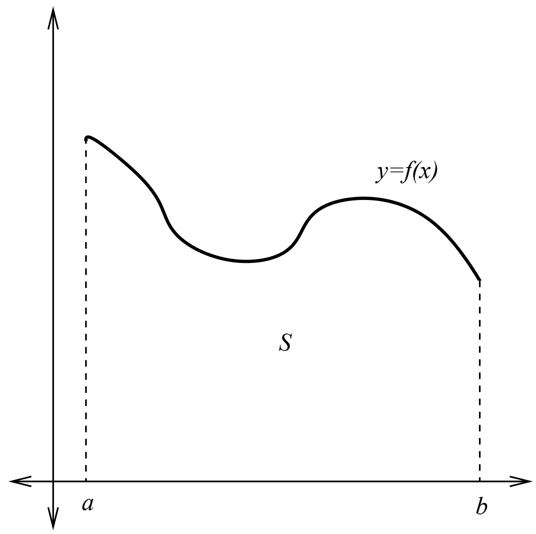


Figure 56: We need to define the area of the region $S = \{(x, y)|a \leq x \leq b, 0 \leq y \leq f(x)\}$.

Although we might have an intuitive notion of what we mean by area, how do we *define* the area under such a curve?

We start by subdividing S into n strips $S_1, S_2, \ldots, S_n$ of equal width. The width of the interval $[a, b]$ is $b - a$ so the width of each strip is

$$\Delta x = \frac{b-a}{n}.$$

These strips divide the interval $[a, b]$ into n subintervals

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n]$$

where $x_0 = a$ and $x_n = b$. If c_i is *any* point in the ith subinterval $[x_{i-1}, x_i]$, then we can approximate the ith strip S_i by a rectangle of width Δx and height $f(c_i)$, which is the value of f at the point c_i (see Figure 57).

The area of such a rectangle is $f(c_i)\Delta x$. Therefore we can approximate the area of S by taking the sum of the area of these rectangles. We call this sum R_n:

$$\begin{aligned} R_n &= f(c_1)\Delta x + f(c_2)\Delta x + \ldots + f(c_n)\Delta x \\ &= \sum_{i=1}^{n} f(c_i)\Delta x \end{aligned}$$

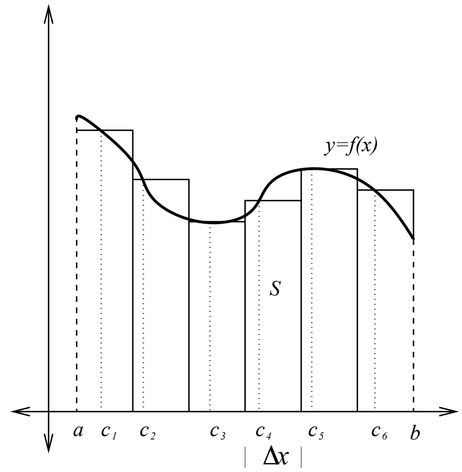


Figure 57: A method of approximating the area of the region S.

Now what happens if we increase n? Consider the graph in Figure 58. It seems that as Δx becomes smaller (when n becomes larger), the approximation gets better.

Therefore we *define* the area A of the region S as follows.

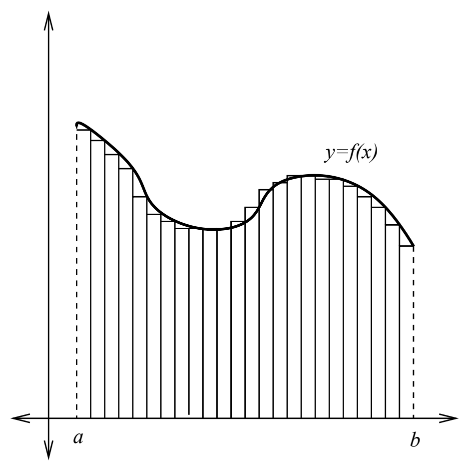


Figure 58: The smaller the width of the rectangle, the better the approximation, in general. Notice that in this diagram we chose $c_i = a + i\Delta x$.

Definition 17 (area of a region, Riemann sum)**.** *The area A of the region S that lies under the graph of the continuous, non-negative function f is*

$$\begin{aligned} A &= \lim_{n\to\infty} R_n \\ &= \lim_{n\to\infty} [f(c_1)\Delta x + f(c_2)\Delta x + \ldots + f(c_n)\Delta x] \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x. \end{aligned}$$

It can be shown that for continuous f, this limit always exists and is independent of the choice of c_i. This can be proved using the fact that f is continuous on $[a, b]$. This sum R_n is called a *Riemann sum*.

Note that $\lim_{n\to\infty}[f(c_1)\Delta x+\cdots+f(c_n)\Delta x]$ is called the *Riemann integral* of f on $[a,b]$ and is denoted $\int_a^b f(x)dx$.

IMPORTANT: this definition can be carried over to functions which are not necessarily non-negative. The intuitive interpretation of A as an area is then no longer valid. The quantity A should then be viewed as a "signed area", i.e.

$$A = (\text{area above } x\text{-axis, below graph}) - (\text{area below } x\text{-axis, above graph}).$$

The fundamental theorem of calculus

Theorem 28 (fundamental theorem of calculus)**.** *(a) If f is continuous on $[a,b]$, $a \le x \le b$, then*

$$A(x) = \int_a^x f(t)\,dt$$

gives an antiderivative of $f(x)$ such that $A(a)=0$. Since $A(x)$ is an antiderivative of f, $A' = f$. Thus

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

(b) If $F(x)$ is any antiderivative of $f(x)$, then

$$F(b) - F(a) = \int_a^b f(t)\,dt.$$

Notation

We will write

$$F(x)\Big|_a^b = \Big[F(x)\Big]_a^b = F(b) - F(a).$$

Properties of the definite integral

For some $c \in \mathbb{R}$, $f(x)$ continuous on $[a,b]$ there holds:

- $\displaystyle\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$
- $\displaystyle\int_a^b c\,dx = c(b-a)$
- $\displaystyle\int_a^b [f(x)\pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx$

- $\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$
- $\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx$
- $\int_a^a f(x)\,dx = 0$

Example 188. *Show that* $\ln x = \int\limits_1^x \frac{1}{t}\,dt,$ *for* $x > 0$*.*

$$\begin{aligned}
\textit{Set } g(x) &= \ln x - \int_1^x \frac{dt}{t} \textit{ for } x > 0.\\
\textit{We want to show: } g(x) &= 0 \textit{ for all } x > 0.\\
\textit{So } g'(x) &= \frac{d}{dx}\ln x - \frac{d}{dx}\left[\int_1^x \frac{dt}{t}\right]\\
&= \frac{1}{x} - \frac{1}{x} \textit{ by Fundamental Theorem of Calculus}\\
&= 0\\
\Rightarrow g(x) &= C, \textit{ for some constant } C.\\
\Rightarrow g(x) &= g(1) = \ln 1 - \int_1^1 \frac{dt}{t} = 0
\end{aligned}$$

Now,

$$\begin{aligned}
f'(x) &= \frac{d\ln x}{dx} - \frac{d\int\limits_1^x \frac{1}{t}dt}{dx} = \frac{1}{x} - \frac{1}{x} = 0.\\
\Rightarrow f(x) &\Leftrightarrow C, \textit{ for some constant } C\\
\Rightarrow f(x) &\Leftrightarrow 0, \textit{ since}\\
f(1) &\Leftrightarrow \ln 1 - \int\limits_1^1 \frac{1}{t}dt = 0.
\end{aligned}$$

Example 189. *Evaluate* $\int_0^{\pi/2} \sin x dx$*.*

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

Example 190. *If* $F(u) = \int\limits_2^u \frac{dt}{\sqrt{1+t^2}}$*, find* $F'(u)$.

$$\begin{aligned} \textit{Let } F(u) &= \int\limits_2^u \frac{dt}{\sqrt{1+t^2}} \\ \Rightarrow F'(u) &= \frac{1}{\sqrt{1+u^2}} \textit{ (by the Fundamental Theorem of Calculus)} \end{aligned}$$

Example 191. *If* $G(x) = \int\limits_2^{\sin x} \frac{dt}{\sqrt{1+t^2}}$*, find* $G'(x)$.

$$\begin{aligned} \textit{Let } F(u) &= \int\limits_2^u \frac{dt}{\sqrt{1+t^2}} \\ \Rightarrow F'(u) &= \frac{1}{\sqrt{1+u^2}} \textit{ (above)} \end{aligned}$$

Now, $G(x) = F(\sin x) = F(u)$*, with* $u = \sin x$*. By the chain rule,*

$$\begin{aligned} G'(x) &= \frac{d}{dx}F(u) \\ &= \frac{dF(u)}{du} \cdot \frac{du}{dx} \\ &= \frac{1}{\sqrt{1+u^2}} \cos x \\ &= \frac{\cos x}{\sqrt{1+\sin^2 x}} \end{aligned}$$

Example 192. *Find the area between the* x*-axis and the curve* $y = \sin x$ *on* $[0, \pi]$.

There holds $\sin x \geq 0$ *on* $[0, \pi]$*, so*

$$\begin{aligned} \textit{Area } = \int_0^\pi \sin x dx &= [-\cos x]_0^\pi, \\ &= -\cos\pi - (-\cos 0), \\ &= -(-1) + 1, \\ &= 2. \end{aligned}$$

$$\int_0^\pi \sin x dx = [-\cos x]_0^\pi = ((-\cos\pi) - (-\cos 0)) = 2$$

Example 193. *Find the area bounded by the curves* $y = x^2$ *and* $y = x^3$ *on the interval* $[0, 1]$.

There holds $x^2 \geq x^3$ on $[0,1]$,

$$\begin{aligned} \Rightarrow \textit{Area} &= \int_0^1 x^2 dx - \int_0^1 x^3 dx \\ &= \left[\frac{1}{3}x^3\right]_0^1 - \left[\frac{1}{4}x^4\right]_0^1 \\ &= \frac{1}{3} - \frac{1}{4} \\ &= \frac{1}{12} \end{aligned}$$

Corollary 2. *If f is non-negative and continuous on $[a,b]$ then*

$$\int_a^b f(x)dx \geq 0.$$

If f is non-positive and continuous on $[a,b]$ then

$$\int_a^b f(x)dx = -(\textit{area above graph below interval } [a,b]) \leq 0.$$

9.2 Approximate integration

We can approximate the area under a curve by taking the sum of a finite number of rectangles under the graph as outlined in the previous section. If we choose the c_i to be the left endpoints of each rectangle, we obtain the approximation

$$\ell_n = \sum_{i=1}^{n} f(x_{i-1})\Delta x.$$

If we choose the c_i to be the right endpoints of each rectangle we obtain the approximation

$$r_n = \sum_{i=1}^{n} f(x_i)\Delta x.$$

The trapezoidal rule is the approximation of the area under a curve which takes the average value of ℓ_n and r_n.

The formula for this approximation is therefore given by

$$\begin{aligned} \text{area } A &\approx \frac{\ell_n + r_n}{2} \\ &= \frac{\Delta x}{2}\left[\sum_{i=1}^{n}(f(x_{i-1}) + f(x_i))\right] \\ &= \frac{\Delta x}{2}\left[(f(x_0) + f(x_1)) + (f(x_1) + f(x_2)) + \ldots + (f(x_{n-1}) + f(x_n))\right] \\ &= \frac{\Delta x}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \ldots + 2f(x_{n-1}) + f(x_n)\right] \end{aligned}$$

where $\Delta x = (b-a)/n$ and $x_i = a + i\Delta x$.

9.3 Indefinite integrals

The *indefinite integral* $\displaystyle\int f(x)dx$, of a suitable function $f(x)$, is defined by

$$\int f(x)dx = F(x) + c,$$

where $F(x)$ is any antiderivative of $f(x)$ and c is an arbitrary constant called the *constant of integration*. Thus $\int f(x)dx$ gives all the antiderivatives of $f(x)$.

Antiderivative

A function F is called an *antiderivative* of f on an interval I if

$$F'(x) = f(x), \qquad \text{for all } x \in I.$$

See the table below for some examples of antiderivatives. We use the notation $F' = f$, $G' = g$ and c is a constant.

Function	Antiderivative
$cf(x)$	$cF(x)$
$f(x) + g(x)$	$F(x) + G(x)$
x^α, $(\alpha \neq -1)$	$\dfrac{x^{\alpha+1}}{\alpha+1}$
$\sin x$	$-\cos x$

Function	Antiderivative
$\cos x$	$\sin x$
$\sec^2 x$	$\tan x$
$\dfrac{1}{x}$	$\ln x$
e^x	e^x

The most general antiderivative can be obtained from those given in the table by adding a constant. Notice that if $\dfrac{d}{dx}F(x) = f(x)$ then $\dfrac{d}{dx}(F(x) + C) = f(x)$ is also true where C is any constant (independent of x).

Example 194. *Find the most general antiderivative of the function $f(x) = x^2 + 3x$.*

Using the table, we have

$$F(x) = \frac{1}{3}x^3 + \frac{3}{2}x^2 + C,$$

where C is a constant.

Example 195. *If $f''(x) = x - \sqrt{x}$, find $f(x)$.*

If $f''(x) = x - \sqrt{x} = x - x^{1/2}$, then

$$f'(x) = \frac{x^2}{2} - \frac{2x^{3/2}}{3} + C$$

and therefore

$$f(x) = \frac{x^3}{6} - \frac{4x^{5/2}}{15} + Cx + D$$

where C and D are constants.

9.4 Explicit integration methods

Substitution

If $u = g(x)$ is a differentiable function whose range is an interval I and f is continuous on I, then

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

If g' is continuous on $[a, b]$ and f is continuous on the range of $u = g(x)$, then

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du.$$

Example 196. *Find $\int x^3 \cos(x^4 + 2)dx$.*

Substitute $u = x^4 + 2$ so $\frac{du}{dx} = 4x^3$ and $x^3 = \frac{\frac{1}{4}du}{dx}$.

$$\begin{aligned} \Rightarrow \int x^3 \cos(x^4 + 2)dx &= \int \frac{1}{4}\frac{du}{dx} \cos u\, dx \\ &= \int \frac{\cos u}{4}\, du \\ &= \frac{1}{4}\int \cos u\, du \\ &= \frac{1}{4}\sin u + c \\ &= \frac{1}{4}\sin(x^4 + 2) + c. \end{aligned}$$

Example 197. *Evaluate $\int_1^2 \frac{dx}{(5x-3)^2}$.*

Substitute $u = 5x - 3$. Then $\frac{du}{dx} = 5$ and $dx = \frac{1}{5}du$.

Note also that $u(2) = 7$ and $u(1) = 2$. Therefore:

$$\begin{aligned}
\int_1^2 \frac{dx}{(5x-3)^2} &= \int_2^7 \frac{1}{u^2} \cdot \frac{1}{5} du \\
&= \frac{1}{5} \int_2^7 u^{-2} du \\
&= \frac{1}{5} \left[-u^{-1}\right]_2^7 \\
&= \frac{1}{5} \left(-\frac{1}{7} + \frac{1}{2}\right) = \frac{1}{14}
\end{aligned}$$

Example 198 (trigonometric substitution)**.** *Evaluate* $\displaystyle\int \frac{dx}{\sqrt{a^2 - x^2}}$.

Substitute $x = a\sin\theta$. Then $dx = a\cos\theta d\theta$. Hence:

$$\begin{aligned}
\int \frac{dx}{\sqrt{a^2 - x^2}} &= \int \frac{a\cos\theta d\theta}{\sqrt{a^2 - a^2\sin^2\theta}} \\
&= \int \frac{a\cos\theta d\theta}{a\sqrt{1 - \sin^2\theta}} \\
&= \int \frac{\cos\theta}{\cos\theta} d\theta \\
&= \int d\theta \\
&= \theta + c \\
&= \arcsin\left(\frac{x}{a}\right) + c
\end{aligned}$$

Example 199. *Evaluate* $\displaystyle\int \cos x \sin x \, dx$.

Substitute $u = \sin x$. Then $\dfrac{du}{dx} = \cos x$. Hence:

$$\begin{aligned}
\int \cos x \sin x \, dx &= \int u \, du \\
&= \frac{u^2}{2} + c \\
&= \frac{\sin^2 x}{2} + c
\end{aligned}$$

Example 200. *Evaluate* $\displaystyle\int (3x+1)(3x^2+2x)^3 \, dx$.

Substitute $u = 3x^2 + 2x$. *Then* $\dfrac{du}{dx} = 6x + 2$ *and* $3x + 1 = \dfrac{1}{2} \cdot \dfrac{du}{dx}$. *Hence:*

$$\begin{aligned}
\int (3x+1)(3x^2+2x)^3\,dx &= \int \frac{1}{2}\frac{du}{dx}u^3\,dx \\
&= \int \frac{u^3}{2}\,du \\
&= \frac{u^4}{8} + c \\
&= \frac{1}{8}(3x^2+2x)^4 + c
\end{aligned}$$

Integration by parts

Given two differentiable functions $u(x)$ and $v(x)$ we have a product rule for differentiation:

$$\frac{d}{dx}(uv) = \frac{du}{dx}v + u\frac{dv}{dx} = u'v + uv'.$$

Therefore uv is an antiderivative of $uv' + u'v$. Hence

$$uv = \int (u'v + uv')\,dx$$

$$= \int u'v\,dx + \int uv'\,dx$$

$$\Rightarrow \int uv'\,dx = uv - \int u'v\,dx.$$

The aim is to simplify the function that is integrated, i.e. $\int u'vdx$ should be easier to find than $\int uv'dx$. Also note that the antiderivative of v' needs to be found.

Example 201. *Evaluate* $\displaystyle\int x^3 \ln(x)dx$.

Set $u = \ln x$ *and* $v' = x^3$. *Then* $u' = \dfrac{1}{x}$ *and* $v = \dfrac{x^4}{4}$. *Hence:*

$$\begin{aligned}
\int x^3 \ln x dx &= \frac{1}{4}x^4 \ln x - \int \frac{1}{4}x^4 \cdot \frac{1}{x}dx \\
&= \frac{1}{4}x^4 \ln x - \frac{1}{4}\int x^3 dx \\
&= \frac{1}{4}x^4 \ln x - \frac{1}{16}x^4 + c
\end{aligned}$$

Note that if $u = x^3$ *and* $v' = \ln x$ *is chosen, then* $v = \int \ln x dx$ *needs to be found.*

Example 202. *Evaluate* $\int xe^x dx$.

Set $u = x$ *and* $v' = e^x$. *Then* $u' = 1$ *and* $v = e^x$.

$$\begin{aligned}\Rightarrow \int xe^x dx &= xe^x - \int e^x dx \\ &= (x-1)e^x + c\end{aligned}$$

Note that for the choice $u = e^x$ *and* $v' = x$*, the integral becomes more complicated than before.*

Integrals involving ln function

If $y = \ln x$ for $x > 0$, then

$$\frac{dy}{dx} = \frac{1}{x} \Rightarrow \int \frac{dx}{x} = \ln x + c, \text{ for } x > 0.$$

For $x < 0$, if $y = \ln(-x)$ we have

$$\frac{dy}{dx} = \frac{-1}{-x} = \frac{1}{x},$$

so we have

$$\int \frac{dx}{x} = \ln(-x) + c, \text{ for } x < 0.$$

We often combine these cases and write

$$\int \frac{dx}{x} = \ln|x| + c, \quad x \neq 0.$$

Integration of rational functions by partial fractions

A ratio of two polynomials,

$$f(x) = \frac{P(x)}{Q(x)}$$

is called a *rational function*. Integration of a rational function is made possible by expressing it as the sum of simpler fractions, called *partial fractions*.

(In the three examples which follow, we will only consider rational functions for which the degree of $P(x)$ is less than the degree of $Q(x)$.)

The general process for obtaining the sum of partial fractions is as follows:

- Factor the denominator $Q(x)$ as much as possible;
- Express $f(x)$ as a sum of fractions each of which take either the form $\dfrac{A}{(ax+b)^i}$ or $\dfrac{Ax+B}{(ax^2+bx+c)^j}$, where the denominators of these fractions come from the factorisation of $Q(x)$.

Consider the following examples.

Example 203. *Evaluate* $\displaystyle\int \frac{x+2}{x^2+x}dx$.

The general technique for solving integrals of the form $\displaystyle\int \frac{cx+d}{(x-a)(x-b)}$ *is to write* $\dfrac{cx+d}{(x-a)(x-b)} = \dfrac{A}{x-1} + \dfrac{B}{x-b}$.

In this case: $\dfrac{x+2}{x^2+x} = \dfrac{x+2}{x(x+1)}$.

$$
\begin{aligned}
\textit{We want: } \frac{x+2}{x(x+1)} &= \frac{A}{x} + \frac{B}{x+1} \\
&= \frac{A(x+1)+Bx}{x(x+1)}. \\
\textit{equate numerators } \Rightarrow x+2 &= A(x+1)+Bx \\
&= (A+B)x + A \qquad \circledast \\
\textit{equate coefficients } \Rightarrow (A+B) &= 1, \textit{ and } A = 2. \\
&\Rightarrow \quad B = -1. \\
\textit{Hence } \Rightarrow \int \frac{x+2}{x(x+1)}dx &= \int \frac{2}{x}dx - \int \frac{dx}{x+1} \\
&= 2\ln|x| - \ln|x+1| + C.
\end{aligned}
$$

An alternate method at $\circledast$ *would be to pick convenient values of* x.

$$\textit{i.e.} \qquad x=-1 \Rightarrow B=-1 \qquad x=0 \;\Rightarrow A=2$$

Example 204. *Evaluate* $\displaystyle\int \frac{dx}{x^2-a^2},\ a \neq 0$.

We write

$$
\begin{aligned}
\frac{1}{x^2-a^2} &= \frac{1}{(x+a)(x-a)} = \frac{A}{x+a} + \frac{B}{x-a} \\
&= \frac{A(x-a)+B(x+b)}{(x+a)(x-a)} \\
&= \frac{(A+B)x - a(A-B)}{(x+a)(x-a)}
\end{aligned}
$$

Equating numerators and equating coefficients:

$$\textit{i.e.} \qquad \begin{cases} A+B=0 \\ A-B=-\frac{1}{a} \end{cases}$$

Hence

$$\begin{aligned}
\Rightarrow A &= -\frac{1}{2a} \textit{ and } B = \frac{1}{2a} \\
\Rightarrow \int \frac{dx}{x^2-a^2} &= \frac{1}{2a}\int \left(-\frac{1}{x+a}+\frac{1}{x-a}\right)dx \\
&= \frac{1}{2a}(-\ln|x+a|+\ln|x-a|)+C.
\end{aligned}$$

9.5 Volume integrals

Recall that if $y = f(x)$, the area under the curve over the interval $I = [a, b]$ is

$$\int_I f(x)dx = \lim \sum_{i=1}^{n} f(x_i^*)(x_i - x_{i-1})$$

where $x_i^* \in [x_i, x_{i-1}]$.

Double integrals

Suppose we have a surface $z = f(x, y)$ above a planar region R in the x-y plane.

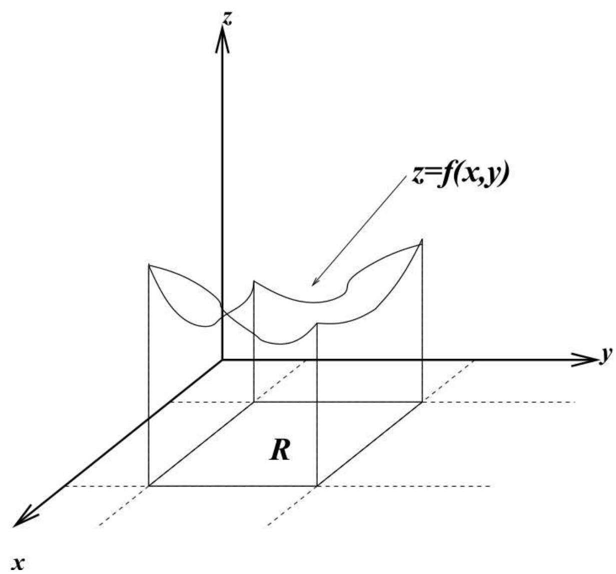
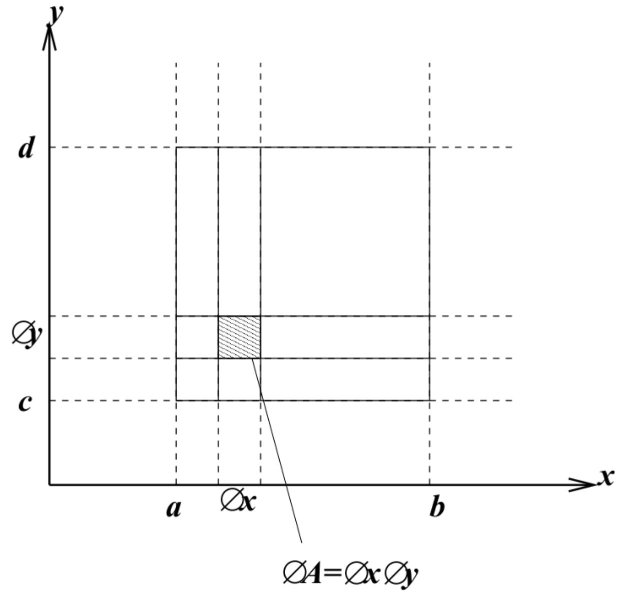


Figure 59: What is the volume V under the surface?

Before moving onto general regions, we start by considering the case where R is a rectangle. That is,

$$R = \{(x, y) \in \mathbb{R}^2 \mid a \le x \le b,\ c \le y \le d\}.$$

Start by dividing R into subrectangles by dividing the interval $[a, b]$ into m subintervals $[x_{i-1}, x_i]$, each of width $\Delta x = \dfrac{b-a}{m}$ and $[c, d]$ into n subintervals $[y_{j-1}, y_j]$ of equal width $\Delta y = \dfrac{d-c}{n}$.



Combining these gives a rectangular grid R_{ij} with subrectangles each of area $\Delta A = \Delta x \Delta y$.

In each subrectangle take any point P_{ij} with co-ordinates (x_{ij}^*, y_{ij}^*).

The volume of the box with base the rectangle ΔA and height the value of the function $f(x, y)$ at the point P_{ij} (so the box touches the surface at a point directly above P_{ij} - see figure 60) is

$$V_{ij} = f(x_{ij}^*, y_{ij}^*)\Delta A.$$

Then for all the subrectangles we have an approximation to the required volume V:

$$V \approx \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)\Delta A,$$

the double Riemann sum.

Let $\Delta x \to 0$ and $\Delta y \to 0$, ie $m \to \infty$ and $n \to \infty$, then we *define* the volume to be

$$V = \lim_{m\to\infty} \lim_{n\to\infty} \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)\Delta A,$$

if the limits exist and we write this as

$$\iint\limits_R f(x, y) dA.$$

We call f integrable if the limits exist. Note that every continuous function is integrable.

Properties of the double integral

(i) $\iint\limits_R (f \pm g) dA = \iint\limits_R f dA \pm \iint\limits_R g dA$

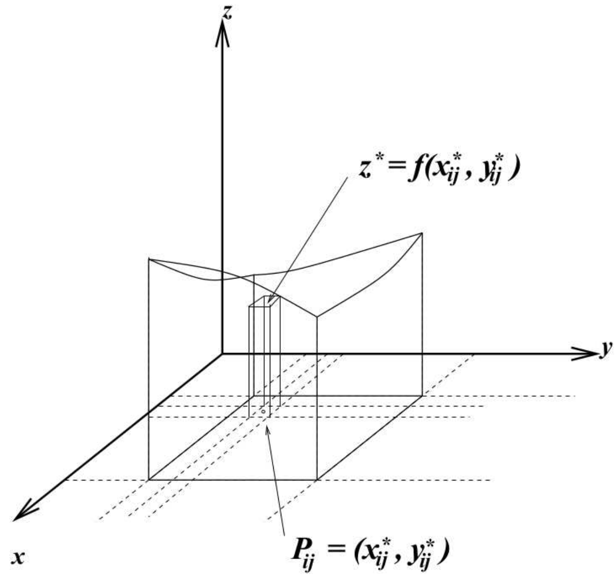


Figure 60: The rectangular box whose volume is $z^* \Delta A$.

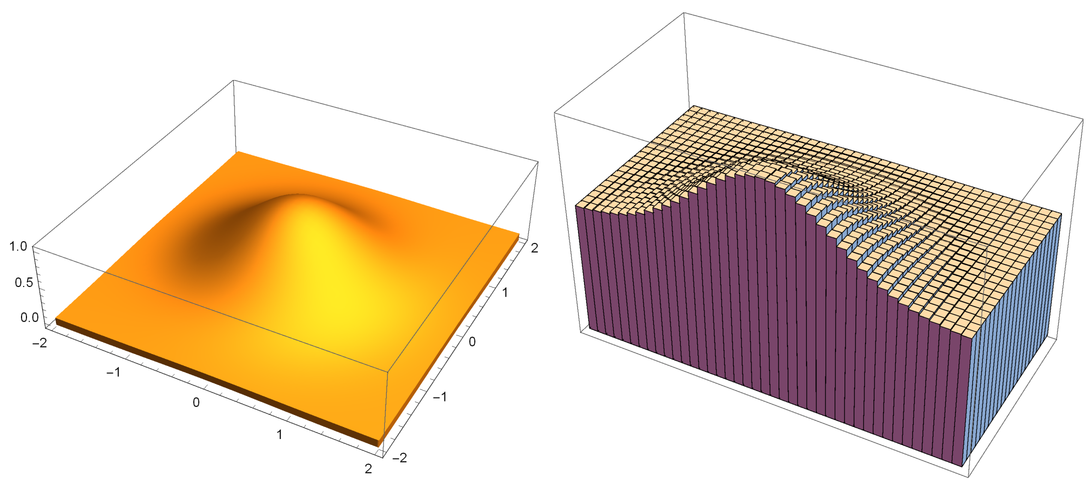


Figure 61: An illustration of the volume calculation of $f(x, y) - e^{-x^2-y^2}$ via the sum of volumes in rectangular coordinates.

(ii) $\displaystyle\iint\limits_R cf dA = c \iint\limits_R f dA$

(iii) $\displaystyle\iint\limits_R f dA = \iint\limits_{R_1} f dA + \iint\limits_{R_2} f dA$

(iv) If $f(x, y) \geq g(x, y)$ for all $(x, y) \in R$ then

$$\iint\limits_R f dA \geq \iint\limits_R g dA$$

Iterated integrals

We define $\int_c^d f(x,y)dy$ to mean that x is fixed and $f(x,y)$ is integrated with respect to y from $y=c$ to $y=d$. So

$$A(x) = \int_c^d f(x,y)dy$$

is a function of x only.

If we now integrate $A(x)$ with respect to x from $x=a$ to $x=b$ we have

$$\begin{aligned}\int_a^b A(x)dx &= \int_a^b \left[\int_c^d f(x,y)dy\right] dx\\ &= \int_a^b \int_c^d f(x,y)\, dy\, dx\end{aligned}$$

This is called an iterated integral.

Example 205. *Evaluate* $\int_0^2 \int_1^3 x^2 y\, dy\, dx$

$$\begin{aligned}\int_0^2 \left(\int_1^3 x^2 y\, dy\right) dx &= \int_0^2 \left[\frac{x^2y^2}{2}\right]_1^3 dx\\ &= \int_0^2 (\frac{9}{2}x^2 - \frac{1}{2}x^2)\, dx\\ &= \int_0^2 4x^2\, dx\\ &= \left[\frac{4}{3}x^3\right]_0^2 = \frac{32}{3}\end{aligned}$$

Now try integrating the other way around:

Example 206. *Evaluate* $\int_1^3 \int_0^2 x^2 y\, dx\, dy$

$$\begin{aligned}\int_1^3 \left(\int_0^2 x^2 y\, dx\right)\, dy &= \int_1^3 \left[\frac{x^3y}{3}\right]_0^2 dy\\ &= \int_1^3 \frac{8}{3}y\, dy\\ &= \left[\frac{8}{6}y^2\right]_1^3\\ &= \frac{8}{6}\times 9 - \frac{8}{6}\times 1 = \frac{64}{6} = \frac{32}{3}\end{aligned}$$

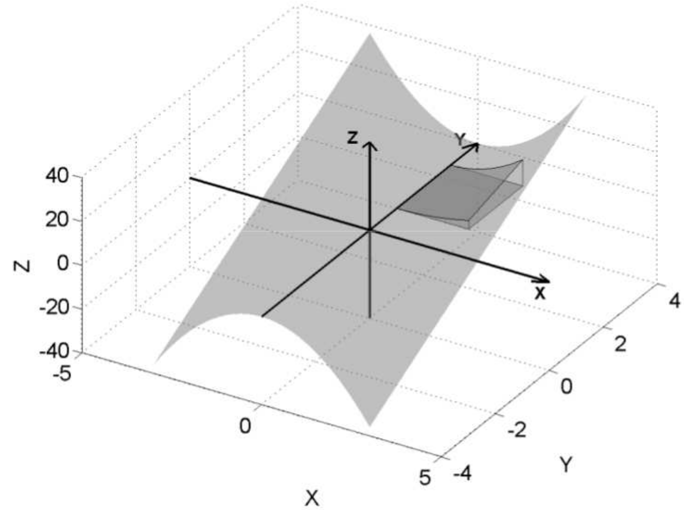


Figure 62: We have just calculated the volume of the solid outlined above.

Integrals over general regions

To find the double integral over a general region D instead of just a rectangle we consider a rectangle which encloses D and define

$$F(x,y) = \begin{cases} f(x,y), & \text{if } (x,y) \in D \\ 0, & \text{if } (x,y) \in R \text{ but } \notin D \end{cases}$$

then

$$\iint_D f(x,y)dA = \iint_R F(x,y)dA$$

and we can proceed as before. It is possible to show that F is integrable if the boundary of D is bounded by a finite number of smooth curves of finite length. Note that F may still be discontinuous at the boundary of D.

Double integrals in polar coordinates

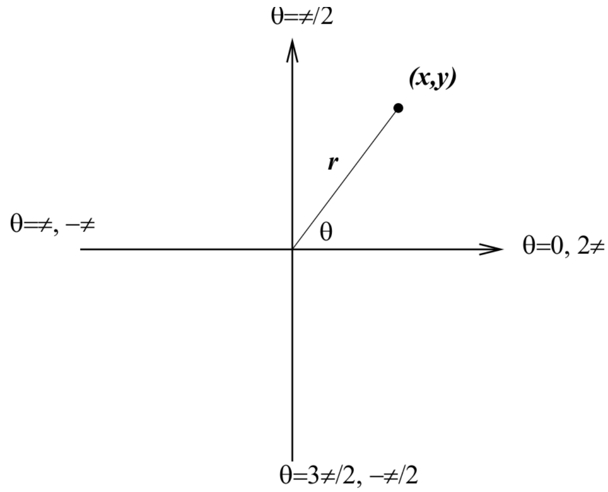
For annular regions with circular symmetry, rectangular coordinates are difficult. It can be more convenient to use *polar coordinates*.

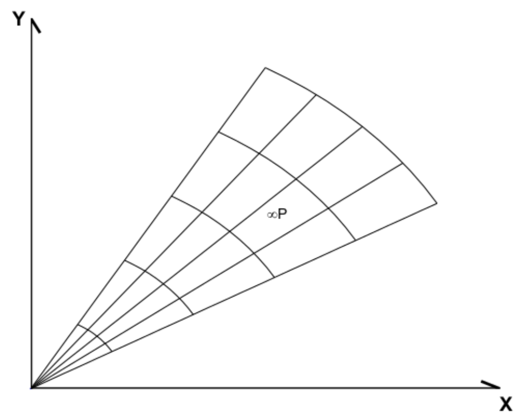
The following diagram explains the relationship between the polar variables r, θ and the usual rectangular ones x, y.

For polar coordinates, we have

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Consider the volume of a solid beneath a surface $z = f(x,y)$ and above a circular region in the x-y plane.





We divide the region into a polar grid as in the following diagram:

We first approximate the area of each polar rectangle as a regular rectangle. We do this as follows. Choose a point P inside each polar rectangle in the polar grid. Let $P = (x^*, y^*)$ or in polar coordinates $P = (r^*, \theta^*)$, where

$$x^* = r^* \cos\theta^*, \quad y^* = r^* \sin\theta^*.$$

The area of the polar rectangle containing P can be approximated as $r^* \Delta\theta \Delta r$. Therefore the volume under the surface and above each polar rectangle can be approximated as

$$\text{vol. one box } \approx r^* \Delta\theta \Delta r f(r^* \cos\theta^*, r^* \sin\theta^*).$$

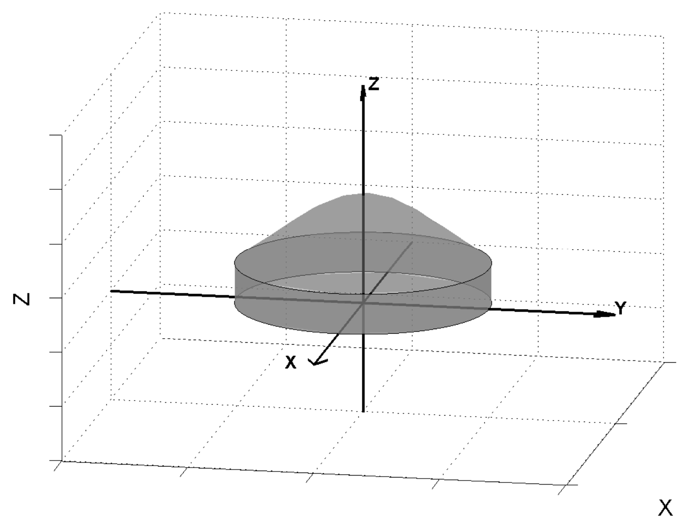
Here $f(r^* \cos\theta^*, r^* \sin\theta^*)$ is the value of the function at the point P, which is also the height of the box used in the approximation. To obtain an approximation for the entire volume below the surface, we sum over the entire polar grid:

$$
\begin{aligned}
\text{vol.} &\approx \sum_{\text{(polar grid)}} r^*\Delta\theta\Delta r f(r^*\cos\theta^*, r^*\sin\theta^*) \\
\Rightarrow \quad \text{vol.} &= \lim_{\Delta r, \Delta\theta \to 0} \sum_{\text{(polar grid)}} r^*\Delta\theta\Delta r f(r^*\cos\theta^*, r^*\sin\theta^*) \\
&= \iint_D f(r\cos\theta, r\sin\theta) r \, d\theta \, dr
\end{aligned}
$$

The double integral in rectangular coordinates is then transformed as follows:

$$
\iint_R f(x,y)\, dx\, dy = \iint_S f(r\cos\theta, r\sin\theta) r \, dr \, d\theta.
$$

Example 207. *Find* $\iint_D e^{-(x^2+y^2)}\, dx\, dy$ *where* D *is the region bounded by the circle* $x^2 + y^2 = R^2$



Use polar coordinates. Set $x = r\cos\theta$*,* $y = r\sin\theta$*, so that the region can be expressed in our familiar notation*

$$
D = \{(r,\theta) | 0 \le r \le R, \ 0 \le \theta \le 2\pi\}.
$$

In the integral, we also have $dx\, dy = r\, dr\, d\theta$*.*

$$
\begin{aligned}
\Rightarrow \quad I &= \iint_D e^{-(x^2+y^2)} dx\, dy \\
&= \int_0^{2\pi} \left(\int_0^R e^{-r^2} r\, dr \right) d\theta \\
&= \left(\int_0^{2\pi} d\theta \right) \left(\int_0^R e^{-r^2} r\, dr \right)
\end{aligned}
$$

Set $u = r^2 \Rightarrow du = 2r\ dr$

$$\Rightarrow \quad I = 2\pi \times \int_{u=0}^{u=R^2} e^{-u} \left(\frac{1}{2} du\right)$$
$$= -\pi \times \left[e^{-u}\right]_0^{R^2}$$
$$= \pi \left(1 - e^{-R^2}\right)$$

What happens when $R \to \infty$*? We have*

$$\int_0^{2\pi} \int_0^{\infty} e^{-r^2} r\ dr\ d\theta = \lim_{R\to\infty} \left(\pi(1 - e^{-R^2})\right) = \pi$$

Now consider the same integral, but this time over the square of side a *in the positive quadrant of the* x*-*y *plane:*

$$\int_0^a \int_0^a e^{-(x^2+y^2)} dx\ dy = \left(\int_0^a e^{-x^2} dx\right) \left(\int_0^a e^{-y^2} dy\right)$$
$$= \left(\int_0^a e^{-x^2} dx\right)^2$$

Now if we let $a \to \infty$ *and multiply by 4 to take into account the entire* x*-*y *plane,*

$$\Rightarrow \quad \left(\int_0^{\infty} e^{-x^2} dx\right)^2 = \frac{\pi}{4}$$
$$\Rightarrow \quad \int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

As a point of interest, at some stage in your future studies, you may come across the "error function"

$$\textit{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{9}$$

which arises in various applications. See Figure 63.

9.6 Improper integrals

Definition 18. *If* $\int_a^t f(x)dx$ *exists for every* $t \geq a$*, then*

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

provided the limit exists.

If $\int_t^b f(x)dx$ *exists for every* $t \leq b$*, then*

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

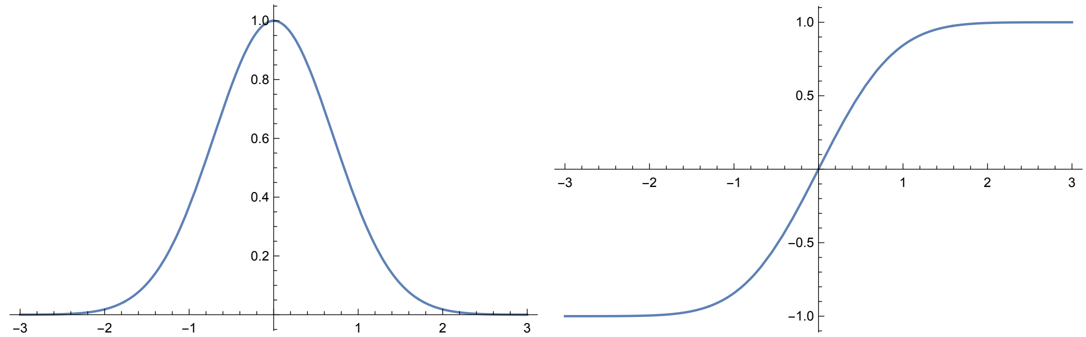


Figure 63: Left: $f(x) = e^{-x^2}$. Right: The error function erf(x), (9).

The integrals $\displaystyle\int_a^\infty f(x)dx$ and $\displaystyle\int_{-\infty}^b f(x)dx$ are called *convergent* if the limit exists and divergent otherwise.

If both $\displaystyle\int_a^\infty f(x)dx$ and $\displaystyle\int_{-\infty}^a f(x)dx$ are convergent, then we define

$$\int_{-\infty}^\infty f(x)dx = \int_{-\infty}^a f(x)dx + \int_a^\infty f(x)dx \quad \text{for any real } a.$$

Note that this must be independent of a. Note also that the usual properties of integrals apply provided all integrals are convergent.

Example 208. *Find* $\displaystyle\int_0^\infty e^{-x}dx.$

For $a > 0$,

$$\int_0^a e^{-x}dx = [-e^{-x}]_0^a = 1 - e^{-a}.$$

As $a \to \infty$, $1 - e^{-a} \to 1$, *so*

$$\int_0^\infty e^{-x}dx = \lim_{a\to\infty}(1 - e^{-a}) = 1.$$

9.7 Application

10 Partial derivatives and gradient descent

10.1 Partial derivatives and tangent planes

Slope in the x-direction

Consider the surface $z = f(x, y) = 1 - x^2 - y^2$ and the point $P = (1, -1, -1)$ on the surface. Use the "y-is-constant" cross-section through P to find the slope in the x-direction at P. The plane meets the surface $z = f(x, y)$ in the curve $z = f(x, -1) = g(x)$,

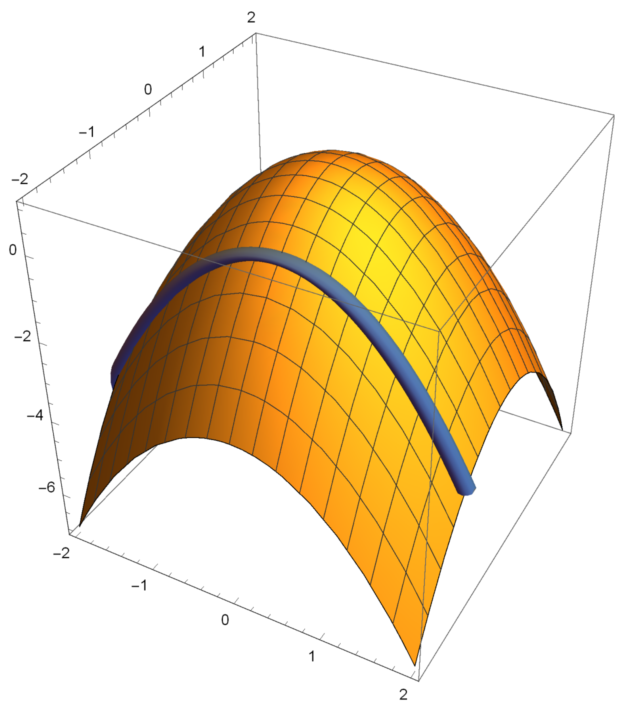


Figure 64: The surface $z = f(x, y) = 1 - x^2 - y^2$ and the intersection of the plane $y = -1$ with the surface.

say.

Here $g(x) = f(x, -1) = 1 - x^2 - 1$ and $g'(x) = -2x = -2$.

The slope in the x-direction, with y held fixed, is called the partial derivative of f with respect to x at the point (a, b)

$$\frac{\partial f}{\partial x}(a, b) = f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}.$$

Slope in the y-direction

Use the "x-is-constant" cross-section to find the slope at $P = (1, -1, -1)$ in the y direction, i.e., where $x = 1$.

When $x = 1$ we have $z = f(1, y) = h(y)$ say, where $h(y) = -y^2$. So $h'(y) = -2y$ and $h'(-1) = -2$.

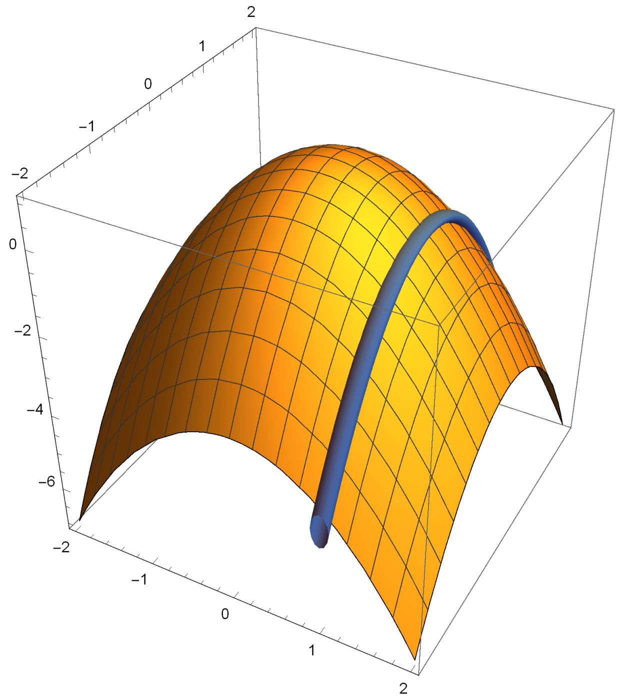


Figure 65: The surface $z = f(x, y) = 1 - x^2 - y^2$ and the intersection of the plane $x = 1$ with the surface.

Similarly, the slope in the y-direction, with x held fixed, is called the partial derivative of f with respect to y at the point (a, b)

$$\frac{\partial f}{\partial y}(a, b) = f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h}.$$

Normal rules of differentiation apply, we simply think of the variables being held fixed as constants when doing the differentiation.

Find the partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ of $f(x, y) = x \sin y + y \cos x$.

Given $f(x, y) = xy^3 + x^2$, find $f_x(1, 2)$ and $f_y(1, 2)$.

Partial derivatives for $f(x, y, z)$

Example 209. *The volume of a box $V(x, y, z) = xyz$.*

If x changes by a small amount, say Δx, denote the corresponding change in V by ΔV. We can easily visualise that $\Delta V = yz\Delta x$.

Therefore,

$$\frac{\Delta V}{\Delta x} = yz.$$

Letting $\Delta x \to 0$ we have $\dfrac{\partial V}{\partial x} = yz$.

For partial derivatives only one independent variable changes and all other independent variables remain fixed.

Higher order derivatives

The second order partial derivatives of f, if they exist, are written as

$$f_{xx} = \frac{\partial^2 f}{dx^2}, \qquad f_{yx} = \frac{d^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\Big(\frac{\partial f}{\partial y}\Big),$$

$$f_{yy} = \frac{\partial^2 f}{\partial y^2}, \qquad f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\Big(\frac{\partial f}{\partial x}\Big).$$

If f_{xy} and f_{yx} are both continuous, then $f_{xy} = f_{yx}$.

Example 210. *Returning to the example where $f(x, y) = x \sin y + y \cos x$, calculate all of the second order partial derivatives of f and show that $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$.*

Multivariate chain rule

Example 211. *Suppose the radius of a cylinder decreases at a rate of $r'(t) = -2$ cm/s. How fast is the volume decreasing when $r = 1$ cm and $h = 2$ cm?*

Volume of a cylinder is given by $V = \pi r^2 h$. Since h is constant, it follows that

$$\frac{dV}{dt} = \frac{dV}{dr}\frac{dr}{dt} = 2\pi r h \frac{dr}{dt} = 2\pi 100.200.(-2) = -80,000\pi \textit{ cm}^3\textit{/sec}.$$

The chain rule for $f(x, y)$

Given $f(x, y)$ with x and y functions of t. Then

$$\frac{df}{dt} = \lim_{\Delta t \to 0} \frac{\Delta f}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(x(t + \Delta t),\, y(t + \Delta t)) - f(x(t),\, y(t))}{\Delta t}.$$

Now if f is smooth and Δf is small, we can relate it to Δx and Δy through the linear approximation

$$\Delta f \simeq f_x \Delta x + f_y \Delta y.$$

Hence

$$\frac{\Delta f}{\Delta t} \simeq f_x \frac{\Delta x}{\Delta t} + f_y \frac{\Delta y}{\Delta t}.$$

Now we let $\Delta t \to 0$, and provided $x(t)$ and $y(t)$ are smooth,

$$\frac{df}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt},$$

which is the chain rule for $f(x(t), y(t))$.

Example 212. *Continuing the previous example, suppose that not just the radius but also the height h is decreasing: $\frac{dh}{dt} = -1$ cm/s. What is the rate of change in volume?*

In this case we have

$$\begin{aligned}\frac{dV}{dt} &= \frac{\partial V}{dr}\frac{dr}{dt} + \frac{\partial V}{\partial h}\frac{dh}{dt} = 2\pi r h\frac{dr}{dt} + \pi r^2\frac{dh}{dt} \\ &= 2\pi.100.200.(-2) + \pi(100)^2(-1) = -90,000\pi \, cm^3/sec.\end{aligned}$$

The chain rule can be extended to any number of dimensions.

Example 213. *If $V(a(t)), b(t), c(t)) = abc$ is the volume of a box then find $\dfrac{dV}{dt}$.*

$$\frac{dV}{dt} = V_a\frac{da}{dt} + V_b\frac{db}{dt} + V_c\frac{dc}{dt} = bc\frac{da}{dt} + ac\frac{db}{dt} + ab\frac{dc}{dt}$$

10.2 Gradients

Directional derivatives

The partial derivative f_x (or f_y) corresponds to the slope of $f(x, y)$ in the x-direction (or y-direction). We now turn our attention to the question of slopes in arbitrary directions, such as $\boldsymbol{i} + 2\boldsymbol{j}$ or $-\boldsymbol{j}$.

Let $\boldsymbol{u} = (u_1, u_2)$ be an arbitrary unit vector in $\mathbb{R}^2$, i.e., $\|\boldsymbol{u}\| = 1$. We have already seen that the equation of the tangent plane at (a, b) of a function f may be written as

$$(\Delta x, \Delta y, \Delta z) = \lambda\big(1, 0, f_x(a, b)\big) + \mu\big(0, 1, f_y(a, b)\big), \quad \lambda, \mu \in \mathbb{R},$$

where $\Delta x = x - a$, $\Delta y = y - a$ and $\Delta z = z - f(a, b)$. This allows us to easily describe the slope of f at the point (a, b) in the direction of $\boldsymbol{u}$: we must "measure" Δz when $\Delta x = u_1$ and $\Delta y = u_2$. Hence we take $\lambda = u_1$ and $\mu = u_2$ in order to find that

$$\begin{aligned}\Delta z &= u_1 f_x(a, b) + u_2 f_y(a, b) \\ &= \big(f_x(a, b), f_y(a, b)\big) \cdot \boldsymbol{u}.\end{aligned}$$

This is usually denoted as $f_{\boldsymbol{u}}(a, b)$ and known as the slope of f at the point (a, b) in the direction of $\boldsymbol{u}$, or the directional derivative of f at (a, b) in the direction of $\boldsymbol{u}$.

Because it is not always convenient to work with unit vectors, we more generally have that the directional derivative of f at (a, b) in the direction of an arbitrary nonzero vector $\boldsymbol{u}$ is given by

$$f_{\boldsymbol{u}}(a, b) = \big(f_x(a, b), f_y(a, b)\big) \cdot \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}.$$

Example 214. *Find the directional derivative of $f(x, y) = 4 - x^2 - 4y^2$ at (a, b) in the $(1, 1)$ direction.*

Here, we evaluate the partial derivatives at $(a, b) = (1, 1)$, so

$$\frac{\partial f}{\partial x}(1,\ 1) = -2x\Big|_{(1,1)} = -2$$

and

$$\frac{\partial f}{\partial y}(1,\ 1) = -8y\Big|_{(1,1)} = -8.$$

The slope in the $\boldsymbol{u} = \dfrac{1}{\sqrt{2}}\boldsymbol{i} + \dfrac{1}{\sqrt{2}}\boldsymbol{j}$ direction is

$$= \frac{1}{\sqrt{2}}(-2) + \frac{1}{\sqrt{2}}(-8) \quad = \quad \frac{-10}{\sqrt{2}} \ = \ -5\sqrt{2}.$$

Example 215. *If $f(x, y) = x^2 - 3y^2 + 6y$, find the slope at $(1, 0)$ in the direction $\boldsymbol{i} - 4\boldsymbol{j}$.*

Here $\|\boldsymbol{i} - 4\boldsymbol{j}\| = \sqrt{1 + 16} = \sqrt{17}$.

So $\boldsymbol{u} = \dfrac{1}{\sqrt{17}}(\boldsymbol{i} - 4\boldsymbol{j})$ is a unit vector in the direction $\boldsymbol{i} - 4\boldsymbol{j}$.

Now slope $= f_{\boldsymbol{u}}(1, 0) = \dfrac{\partial f}{\partial x}\Big|_{(1,0)} \dfrac{1}{\sqrt{17}} + \dfrac{\partial f}{\partial y}\Big|_{(1,0)} \left(\dfrac{-4}{\sqrt{17}}\right)$,

$$\frac{\partial f}{\partial x}(1, 0) = 2x\Big|_{(1,0)} = 2$$

and

$$\frac{\partial f}{\partial y}(1,\ 0) = (-6y + 6)\Big|_{(1,0)} = 6.$$

So slope $= f_{\boldsymbol{u}}(1,\ 0) = \dfrac{2}{\sqrt{17}} - \dfrac{6.4}{\sqrt{17}} = -\dfrac{22}{\sqrt{17}}$.

The gradient vector ∇f

The gradient vector or simply gradient of f is a vector with the partial derivatives as components

$$\text{grad}\, f = \nabla f = (f_x,\ f_y) = f_x\boldsymbol{i} + f_y\boldsymbol{j}.$$

Example 216. *Find the gradient of $f(x, y) = x^2 - 3(y - 1)^2 + 3$.*

$\nabla f = 2x\boldsymbol{i} - 6(y - 1)\boldsymbol{j}$.

Similarly, for a function $f(x, y, z, w)$ of four variables,

$$\textit{grad}\, f = \nabla f = (f_x, f_y, f_z, f_w).$$

Note that directional derivative can be conveniently expressed in terms of the gradient as

$$f_{\boldsymbol{u}} = \nabla f \cdot \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}.$$

Example 217. *Find $f_{(1,-1)}(0,1)$ for $f(x,y) = x - x^2y^2 + y$.*

Example 218. *Find the directional derivative of $g(x,y) = x^{x^2} \cos y$ at $(1,\pi)$ in the direction $-3\boldsymbol{i} + 4\boldsymbol{j}$.*

First we need to find the unit vector in the given direction. Now $\|-3\boldsymbol{i}+4\boldsymbol{j}\| = \sqrt{9+16} = 5$. So $\boldsymbol{u} = \dfrac{-3}{5}\boldsymbol{i} + \dfrac{4}{5}\boldsymbol{j}$.

Next find the gradient vector at $(1,\pi)$:

$$\begin{aligned}
\nabla g &= \frac{\partial g}{\partial x}\boldsymbol{i} + \frac{\partial g}{\partial y}\boldsymbol{j} \\
&= 2xe^{x^2} \cos y\boldsymbol{i} - e^{x^2} \sin y\boldsymbol{j} \\
\nabla g(1,\,\pi) &= 2e^1(-1)\boldsymbol{i} - 0\boldsymbol{j} = -2e\boldsymbol{i} \\
g_{\boldsymbol{u}}(1,\,\pi) &= \nabla g(1,\,\pi).\left(\frac{-3}{5}\boldsymbol{i} + \frac{4}{5}\boldsymbol{j}\right) \\
&= \frac{6e}{5}.
\end{aligned}$$

Properties of the gradient vector

Example 219. *Consider the contour diagram of a plane $z = f(x,y) = mx + ny + c$. For $n \neq 0$ the contours have slope $-m/n$ in the xy-plane.*

The gradient vector, $m\boldsymbol{i} + n\boldsymbol{j}$ ($m \neq 0$), has slope n/m and so is perpendicular to the contours. It points in the direction of increasing f. In fact, the direction in which it points is the direction of greatest slope.

Example 220. *Consider $f(x,y) = x^2 + y^2$. The contours are circles centered at the origin, and $\nabla f = 2x\boldsymbol{i} + 2y\boldsymbol{j}$ points radially out. Once again, ∇f points in the direction of greatest slope, perpendicular to the contour lines.*

In general, two important properties of the gradient of a function are:

> The gradient $\nabla f(a,b)$ is **perpendicular to the contour line through** (a,b) and points in the direction of increasing f. In fact, the direction and magnitude of **steepest slope** at (a,b) are given by $\nabla f(a,b)$ and $\|\nabla f(a,b)\|$.

We can understand these two facts by considering the value of $\cos\theta$ in

$$f_{\boldsymbol{u}} = \nabla f \cdot \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} = \|\nabla f\| \cos\theta.$$

Example 221. *$T(x, y) = 20 - 4x^2 - y^2$ describes the temperature on the surface of a metal plate.*

In which direction away from the point $(2, -3)$ does the temperature increase most rapidly?

In which directions away from the point $(2, -3)$ is the temperature not changing?

The direction is $\nabla T = \frac{\partial T}{\partial x}\boldsymbol{i} + \frac{\partial T}{\partial y}\boldsymbol{j} = -8x\boldsymbol{i} - 2y\boldsymbol{j}$. *So*

$$\nabla T(2, -3) = -16\boldsymbol{i} + 6\boldsymbol{j}.$$

The direction **in terms of angles** *is* $\pi - \arctan\left(\frac{6}{16}\right)$.

Example 222. *A team from a large British oil company is mapping the ocean floor to assist in the plugging of a leaking oil well in the Gulf of Mexico.*

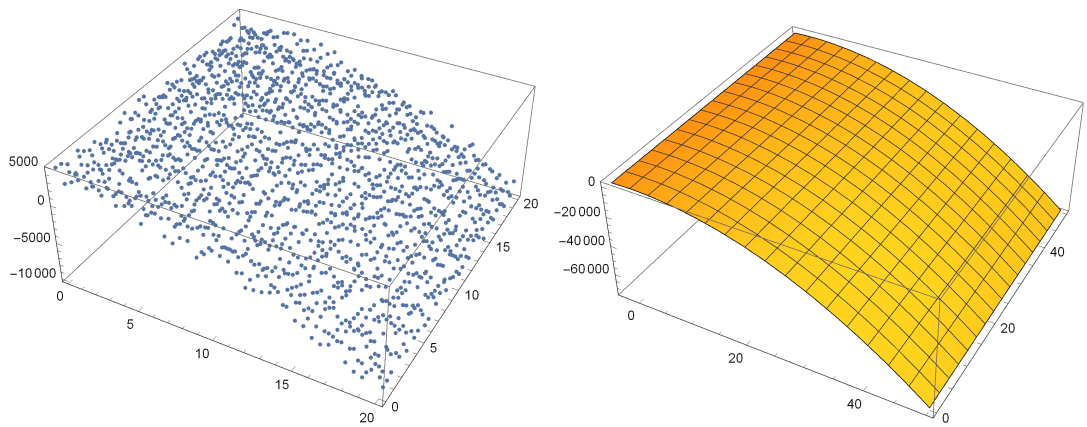


Figure 66: Data collected on points of the ocean floor and the corresponding model $D(x, y)$.

Using sonar, they collect data illustrated in Fig. 66 and develop the model

$$D(x, y) = 1700 - 30x^2 - 50\sin\left(\frac{\pi y}{2}\right),$$

where x and y are distance in kilometres, D is depth in metres, and $-2 \leq x \leq 2$ and $-2 \leq y \leq 2$.

(a) The well is located at $(1, 0.5)$. What is its depth?

$$D(1, 0.5) = 250 - 30 - 50\sin\frac{\pi}{4} \simeq 184.6m.$$

(b) Determine the slope of the ocean floor in the positive x-direction and in the positive y-direction in the area considered by the clean-up team.

Slope in the x direction is

$$\frac{\partial h}{\partial x}(1, 0.5) = 60x\Big|_{(1,0.5)} = 60.$$

But we must be careful here because h is in meters while x and y are in kilometers. So in fact the slope is $\dfrac{60}{1000} = 0.06$.

Slope in the y direction is

$$\begin{aligned}\frac{\partial h}{\partial y}(1,\, 0.5) &= \frac{50\pi}{2}\cos\frac{\pi y}{2}\Big|_{(1,0.5)} = 25\pi\cos\frac{\pi}{4} \\ &= \frac{25}{\sqrt{2}}\pi.\end{aligned}$$

So slope is $\dfrac{25}{\sqrt{2}.1000}\pi = \dfrac{\pi}{40\sqrt{2}}$.

(c) Determine the magnitude and direction of greatest rate of change of depth at the position of the well.

The direction of greatest rate of change is given by

$$\nabla h = \frac{60\boldsymbol{i} + \frac{25\pi}{\sqrt{2}}\boldsymbol{j}}{1000}$$

or $\arctan\left(\frac{25\pi}{60\sqrt{2}}\right) = \arctan\left(\frac{5\pi}{12\sqrt{2}}\right)$ *and the magnitude is* $\dfrac{\sqrt{3600 + \frac{625\pi^2}{2}}}{1000}$.

10.3 Gradient descent

10.4 Application